Contents

Problem: Consider a random, infinite string of 0's and 1's

$$(l_1, l_2, \ldots, l_m, \ldots), l_m \in \{0,1\} \forall m \in \mathbb{N}$$

What is the probability that somewhere along the string, I find 100 consecutive 0's?

Comment: The statement of the problem is unclear. Lesbegue's machinery will clarify this.

$X := \{x = (l_1, l_2, \ldots, l_m, \ldots) \mid l_m \in \{0,1\} \forall m \in \mathbb{N}\}$ This is, in a natural way, a compact metric space. Use the "Hamming-style" distance. For

$$x = (l_1, l_2, \ldots, l_m, \ldots), x' = (l'_1, l'_2, \ldots, l'_m, \ldots)$$

in X,

$$d(x, x') = \frac{|l_1 - l'_1|}{2} + \frac{|l_2 - l'_2|}{4} + \ldots + \frac{|l_n - l'_n|}{2^n}$$

We can consider

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

We can consider $\mathcal{B} = \sigma\text{-Alg}(\mathcal{T})$ (Borel σ-algebra)
Probability is $P : \mathcal{B} \to [0,1]$. Have $P(B) \in [0,1] \forall B \in \mathcal{B}$.
"Problem" says: let
$S = \{x = (l_1, l_2, \ldots, l_m, \ldots) \in X \mid \exists m \in \mathbb{N} \text{ such that } l_m = l_{m+1} = \ldots = l_{m+99} = 0\}$.
Check that $S \in \mathcal{B}$ and compute $P(S) = ?$

1 Additive set-function on an algebra of sets

Definition 1.1

What it means for $\mathscr{A} \subseteq 2^X$ to be an **algebra of sets**:

1. (AS1) $\varnothing \in \mathscr{A}$
2. (AS2) If $A \in \mathscr{A}$, then $X \setminus A \in \mathscr{A}$
3. (AS3) If $A, B \in \mathscr{A}$, then $A \cap B \in \mathscr{A}$

Proposition 1.2

$\mathscr{A} \subseteq 2^X$ algebra of sets. Then

1. $X \in \mathscr{A}$
2. $A, B \in \mathscr{A} \Rightarrow A \cup B \in \mathscr{A}$
3. $A, B \in \mathscr{A} \Rightarrow A \cap B \in \mathscr{A}$
4. For every $n \geq 2$ and every $A_1, \ldots, A_n \in \mathscr{A}$, one has that $A_1 \cap \ldots \cap A_n \in \mathscr{A}$
5. For every $n \geq 2$ and every $A_1, \ldots, A_n \in \mathscr{A}$, one has that $A_1 \cup \ldots \cup A_n \in \mathscr{A}$

Example

$\mathscr{A} = 2^X$ and $\mathscr{A} = \{\varnothing, X\}$

Lemma 1.3 **Intersection Property of algebras of sets**

Let $\mathscr{C}$ be a collection of algebras of sets of a set X. Then $\cap\mathscr{C}$ is an algebra of sets.

Proof:

1. For each $\mathscr{A} \in \mathscr{C}$, $\varnothing \in \mathscr{A}$ since $\mathscr{A}$ is an algebra of sets, so $\varnothing \in \cap\mathscr{C}$.
2. For $\mathcal{A} \in \cap\mathscr{C}$, we have that for all $\mathscr{A} \in \mathscr{C}$, $\mathcal{A} \in \mathscr{A}$. But $\mathscr{A}$ is an algebra of sets, so $X \setminus \mathcal{A} \in \mathscr{A}$. But this holds for all $\mathscr{A} \in \mathscr{C}$, so $X \setminus A \in \cap\mathscr{C}$.
3. For $A, B \in \cap\mathscr{C}$, we have that $A, B \in \mathscr{A}$ for all $\mathscr{A} \in \mathscr{C}$, but $\mathscr{A}$ is an algebra of sets, so $A \cap B \in \mathscr{A}$ for all $\mathscr{A} \in \mathscr{C}$, hence $A \cap B \in \cap\mathscr{C}$.

This suffices to show that $\cap\mathscr{C}$ is an algebra of sets. □

Proposition 1.4

Let $\mathscr{C} \subseteq 2^X$. Then there exists a unique algebra of sets $\mathscr{C}_0 \subseteq 2^X$ such that $\mathscr{C}_0 \supseteq \mathscr{C}$ and $\mathscr{C}_0$ is minimal in the sense of inclusion (or equivalently, for any algebra of sets $\mathscr{D} \subseteq 2^X$ with $\mathscr{D} \supseteq \mathscr{C}$, we have $\mathscr{D} \supseteq \mathscr{C}_0$).

Proof: Let $\mathscr{C}_0$ be the intersection of the collection of algebras of sets that contain $\mathscr{C}$. We have that $\mathscr{C}_0 \supseteq \mathscr{C}$ by construction. Furthermore, for any algebra of sets $\mathscr{D} \supseteq \mathscr{C}$, we have that $\mathscr{D}$ is one of the sets in the collection, so $\mathscr{C}_0 \subseteq \mathscr{D}$. For uniqueness, let A be a minimal sigma-algebra containing $\mathscr{C}$. Then $A \subseteq \mathscr{C}_0$ (since A is minimal) and $\mathscr{C}_0 \subseteq A$ (since $\mathscr{C}_0$ is minimal), so $\mathscr{C}_0$ is unique. □

Definition 1.5

Let $\mathcal{S}$ be a collection of subsets of X. Then, the algebra of sets generated by $\mathcal{S}$, denoted by $\operatorname{Alg}(\mathcal{S})$, is $\cap\mathscr{C}$, where $\mathscr{C}$ is the collection of algebras of sets of X which contain $\mathcal{S}$.

Definition 1.6 **(Finite) Additive set function**

$\mathscr{C} \subseteq 2^X$ algebra of sets. Define $\mu : \mathscr{C} \to [0, \infty]$ such that

1. $\mu(A \cup B) = \mu(A) \cup \mu(B)$, where $A \cap B = \varnothing$
2. $\mu(\varnothing) = 0$

Remark 1.7

We use the arithmetic of $[0, \infty]$. Convention specific to measure theory: $0 \cdot \infty = 0 = \infty \cdot 0$ (“zero is absorbant”).

Remark 1.8

Why do we require $\mu(\varnothing) = 0$? It is to avoid the case where $\mu(A) = \infty \forall A \in \mathscr{C}$.

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

1. (Add) For $A \subseteq C$ in $\mathscr{C}$, we have $\mu(A) \leq \mu(C)$.
2. (Subadd) For A, B in $\mathscr{C}$, we have $\mu(A \cup B) \leq \mu(A) + \mu(B)$.

2 Positive measure on a sigma-algebra

Definition 2.1 **Sigma-Algebra**

What it means for $\mathfrak{M} \subseteq 2^X$ to be a **sigma-algebra**:

1. (Sigma AS3) $\varnothing \in \mathfrak{M}$
2. (Sigma AS2) If $A \in \mathfrak{M}$, then $X \setminus A \in \mathfrak{M}$
3. (Sigma AS3) If $\mathscr{C}$ is a countably infinite collection of sets in $\mathfrak{M}$, then $\cap \mathscr{C} \in \mathfrak{M}$

Remark 2.2

1. $\mathfrak{M}$ is an algebra of sets.
2. $\mathfrak{M}$ is stable under unions and under set difference.

Lemma 2.3

Let $\mathscr{C}$ be a collection of sigma-algebras of a set X. Then $\cap \mathscr{C}$ is a sigma-algebra.

Proposition 2.4

Let $\mathscr{C} \subseteq 2^X$. Then there exists a unique sigma-algebra $\mathscr{C}_0 \subseteq 2^X$ containing $\mathscr{C}$ such that $\mathscr{C}_0$ is minimal in the sense of inclusion (or equivalently, for any algebra of sets $\mathscr{D} \subseteq 2^X$ with $\mathscr{D} \supseteq \mathscr{C}$, we have $\mathscr{D} \supseteq \mathscr{C}_0$).

Definition 2.5

The sigma-algebra generated by a collection of sets $\mathscr{C} \subseteq 2^X$ is the sigma-algebra obtained in Proposition 2.4.

Definition 2.6 **Borel sigma-algebra**

The Borel sigma-algebra of a metric space (X, d), given that $\mathcal{T}$ is the collection of open subsets of X, is $\sigma\text{-Alg}(\mathcal{T})$.

Remark 2.7

1. Let $(s_n)_{n=1}^{\infty}$ be a sequence of elements in $[0, \infty]$ which is increasing, in the sense that
$$0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq \ldots$$
Then the sequence approaches a limit $L \in [0, \infty]$ and $L = \sup\{s_n \mid n \in \mathbb{N}\}$.
2. Let $(t_n)_{n=1}^{\infty}$ be any sequence in $[0, \infty]$. We can consider the finite sums
$$s_1 = t_1, s_2 = t_1 + t_2, \ldots, s_n = t_1 + \ldots + t_n, \ldots$$
which form an increasing sequence in $[0, \infty]$. We can consider the limit
$$L = \lim_{n \to \infty} s_n \in [0, \infty]$$
which we typically denote as
$$\sum_{n=1}^{\infty} t_n$$

Definition 2.8 **Positive measure**

$\mathfrak{M} \subseteq 2^X$ sigma-algebra of subsets of X. A positive measure on $\mathfrak{M}$ is a function $\mu : \mathfrak{M} \to [0, \infty]$ which has $\mu(\varnothing) = 0$ and satisfies: (Sigma-Add)
$$\begin{cases} \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n), \text{ whenever } A_1, A_2, \ldots, A_n, \ldots \in \mathfrak{M} \\ \text{are such that } A_i \cap A_j = \varnothing \text{ for } i \neq j \end{cases}$$

Remark 2.9

A positive measure is, in particular, a (finite) additive set function, so we can use all the properties found in Proposition 1.9.

Definition 2.10

1. If we are given a sigma-algebra $\mathfrak{M} \subseteq 2^X$, then we say that the pair $(X, \mathfrak{M})$ is a *measureable space.*
2. If we are given a sigma-algebra $\mathfrak{M} \subseteq 2^X$ and positive measure $\mu : \mathfrak{M} \to [0, \infty]$, then the triple $(X, \mathfrak{M}, \mu)$ is said to be a *measure space.*
3. Let $(X, \mathfrak{M}, \mu)$ be as above. If $\mu(X) < \infty$, then we say that μ is a *finite measure.* In the special case when $\mu(X) = 1$, we call $(X, \mathfrak{M}, \mu)$ a *probability space).*

Proposition 2.11 **Sigma-Sub-Add**

Let $\mathfrak{M}$ be a sigma-algebra. Let $\mu : \mathfrak{M} \to [0, \infty]$ be a positive measure. Then

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

for $A_1, A_2, ..., A_n, ... \in \mathfrak{M}$.

Proposition 2.12 **Continuity along increasing and decreasing chains**

Let $\mathfrak{M} \subseteq 2^X$ be a sigma-algebra and let $\mu : \mathfrak{M} \to [0, \infty]$ be a postive measure.

1. Let $(B_n)_{n=1}^{\infty}$ be a sequence (or an *increasing chain*) of sets in $\mathfrak{M}$ such that

$$B_1 \subseteq B_2 \subseteq ... \subseteq B_n \subseteq ...$$

Then

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n) \in [0, \infty]$$

2. Let $(C_n)_{n=1}^{\infty}$ be a sequence (or a *decreasing chain*) of sets in $\mathfrak{M}$ such that

$$C_1 \supseteq C_2 ... \supseteq C_n \supseteq ...$$

and $\mu(C_1) < \infty$. Then

$$\mu\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} \mu(C_n) \in [0, \infty)$$

Proof of 2.11: Define $A := \cup_{n=1}^{\infty} A_n$ for $A_n \in \mathfrak{M}$ and $S := \sum_{n=1}^{\infty} \mu(A_n) \in [0, \infty]$. For every $m \in \mathbb{N}$, define $B_m = A_1 \cup A_2 \cup ... \cup A_m$. Then $(B_m)_{m=1}^{\infty}$ is an increasing chain with

$$\begin{aligned} \bigcup_{n=1}^{\infty} B_n &= \bigcup_{n=1}^{\infty} (A_1 \cup ... \cup A_n) \\ &= \bigcup_{n=1}^{\infty} A_n \\ &= A \end{aligned}$$

So we get

$$\begin{aligned} \mu(A) &= \lim_{n\to\infty} \mu(B_n) \text{ (by Prop 2.12.1)} \\ &= \sup\{\mu(B_n) \mid n \in \mathbb{N}\} \end{aligned}$$

Want to show that the sup is upper bounded by S. But for every $n \in \mathbb{N}$, we indeed have

$$\begin{aligned} \mu(B_n) &= \mu(A_1 \cup ... \cup A_n) \\ &\leq \mu(A_1) + ... + \mu(A_n) \text{ (by Sub-Add)}_M \\ &\leq S \end{aligned}$$

□

Proof of 2.12.2: Define $D_n = C_1 \setminus C_n$ for $n \in \mathbb{N}$, with $D_n \in \mathfrak{M}$. Since $(C_n)_{n=1}^{\infty}$ is a decreasing chain, then $(D_n)_{n=1}^{\infty}$ is an increasing chain. But note that

$$\bigcup_{n=1}^{\infty} D_n = \bigcup_{n=1}^{\infty} (C_1 \setminus C_n) = C_1 \setminus \bigcap_{n=1}^{\infty} C_n$$

So, we have

$$\begin{aligned} \mu\left(C_1 \setminus \bigcap_{n=1}^{\infty} C_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} D_n\right) \\ &= \lim_{n\to\infty} \mu(D_n) \text{ (by Prop 2.12.1)} \\ &= \lim_{n\to\infty} \mu(C_1 \setminus C_n) \quad \mathbf{(1)} \end{aligned}$$

Since μ is an additive set function, then

$$\begin{aligned} \mu(C_1) &= \mu\left(C_1 \setminus \bigcap_{n=1}^{\infty} C_n\right) + \mu\left(\bigcap_{n=1}^{\infty} C_n\right) \\ &= \mu(C_1 \setminus C_n) + \mu(C_n) \end{aligned}$$

But $\mu(C_1) < \infty$, so

$$\begin{aligned} \mu\left(C_1 \setminus \bigcap_{n=1}^{\infty} C_n\right) &= \mu(C_1) - \mu\left(\bigcap_{n=1}^{\infty} C_n\right) < \infty \\ \mu(C_1 \setminus C_n) &= \mu(C_1) - \mu(C_n) < \infty \end{aligned}$$

Therefore,

$$\begin{aligned} \mu(C_1) - \mu\left(\bigcap_{n=1}^{\infty} C_n\right) &= \mu\left(C_1 \setminus \bigcap_{n=1}^{\infty} C_n\right) \\ &= \lim_{n\to\infty} \mu(C_1 \setminus C_n) \text{ (by } \mathbf{1}) \\ &= \mu(C_1) - \lim_{n\to\infty} \mu(C_n) \end{aligned}$$

Which implies that

$$\mu\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} \mu(C_n)$$

□

Example **Borel-Cantelli Lemma**

Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $(A_n)_{n=1}^{\infty}$ be a family of sets from $\mathfrak{M}$. Let $T := \{x \in X \mid x \text{ belongs to infinitely many of the } A_n\text{'s}\}$. If $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then $\mu(T) = 0$.
Proof: Let $B_n = \cup_{k=n}^{\infty} A_k$.
`claim` $T = \cap_{n=1}^{\infty} B_n$
Proof:

$$\begin{aligned}
\bigcap_{n=1}^{\infty} B_n &= \{x \in X \mid x \in B_n \forall n \in \mathbb{N}\} \\
&= \{x \in X \mid x \in A_k \text{ for some } k \geq n, \forall n \in \mathbb{N}\} \\
&= \{x \in X \mid x \text{ belongs to infinitely many } A_n\text{'s}\} \\
&= T
\end{aligned}$$

□

Then $(B_n)_{n=1}^{\infty}$ is a decreasing chain. Note that

$$\begin{aligned}
\mu(B_1) &= \mu\left(\bigcup_{k=1}^{\infty} A_k\right) \\
&\leq \sum_{n=1}^{\infty} \mu(A_n) \text{ (by Sigma-Sub-Add)} \\
&< \infty \text{ (by assumption)}
\end{aligned}$$

So, we have

$$\begin{aligned}
\mu(T) &= \mu\left(\bigcap_{n=1}^{\infty} B_n\right) \\
&= \lim_{n\to\infty} \mu(B_n) \text{ (by Prop 2.12.2)} \\
&= \lim_{n\to\infty} \mu\left(\bigcup_{k=n}^{\infty} A_k\right) \\
&\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(A_k) \text{ (by Sigma-Sub-Add)} \\
&= 0
\end{aligned}$$

since the tail sums of a convergent series converge to zero. Therefore, $\mu(T) = 0$. □

Example **Atomic Measure**

An atomic measure μ is based on weights of individual points $x \in X$. Suppose you have $w : X \to [0, \infty]$. Take $\mathfrak{M} = 2^X$. Define $\mu : \mathfrak{M} \to [0, \infty]$ by

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

Then $(X, 2^X, \mu)$ is a measure space. We can show this by verifying that Sigma-Add is satisfied by the measure.
Special cases:

1. $w(x) = 1$ for all $x \in X$, to get the *counting measure* on X.
2. Dirac measure concentrated at $x_o \in X$. $w(x_0) = 1, w(x) = 0$ for all $x \neq x_0$.

Example

$X = \mathbb{R}, \mathfrak{M} = \mathcal{B}_{\mathbb{R}}$ (Borel σ-algebra).

Note **Fact 2D-1 to D2-3**

1. $\mathcal{B}_{\mathbb{R}}$ can be generated by the collection of half-open intervals in $\mathbb{R}$.
2. If $\mu, v : \mathcal{B}_{\mathbb{R}} \to [0, \infty]$ are positive measures such that $\mu((a, b]) = v((a, b]) < \infty$ for all $a < b$ in $\mathbb{R}$, then $\mu = v$.
3. The Borel sigma-algebra is stable under translation.

Note **Lesbegue Measure (Fact 2D-4)**

The Lesbegue measure $\lambda : \mathcal{B}_{\mathbb{R}} \to [0, \infty]$ is determined uniquely via the requirement that $\lambda((a, b]) = b - a$ for all $a < b$ in $\mathbb{R}$.
Another approach: λ is uniquely determined by the requirement that $\lambda(B + t) = \lambda(B)$ for all $B \in \mathcal{B}_{\mathbb{R}}$ for all $t \in \mathbb{R}$, with the normalization that $\lambda((0, 1]) = 1$.
These conditions are equivalent.
For existence of λ, we will involve an extension theorem of Caratheodory (to be done later). For now, we accept that λ exists.

Note **Fact 2D-5**

There is no positive measure $\mu : 2^{\mathbb{R}} \to [0, \infty]$, which is invariant under translations and with $\mu((0, 1]) = 1$. This shows in particular that $\mathcal{B}_{\mathbb{R}} \neq 2^{\mathbb{R}}$ Construction of a set in $2^{\mathbb{R}} \setminus \mathcal{B}_{\mathbb{R}}$ uses a Hamel basis (basis of $\mathbb{R}$ over $\mathbb{Q}$).

3 Measurable functions, the space of functions $\mathrm{Bor}(X, \mathbb{R})$

Definition 3.1

$(X, \mathfrak{M}), (Y, \mathfrak{N})$. $f : X \to Y$ is said to be $\mathfrak{M}/\mathfrak{N}$-*measurable* when it has the property that $f^{-1}(N) \in \mathfrak{M}$ for every $N \in \mathfrak{N}$.

Remark 3.2 **Pre-images are nice**

Pre-images of unions/intersections of (measurable) sets are unions/intersections of the pre-images.

Proposition 3.3 **Compositions of Measurable Functions are Measurable**

$(X, \mathfrak{M}), (Y, \mathfrak{N}), (Z, \mathfrak{P})$. $f : \mathfrak{M} \to \mathfrak{N}$ and $g : \mathfrak{N} \to \mathfrak{P}$ measurable. Then $h = g \circ f$ is measurable.

Proposition 3.4

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

Proof: Use the method of civilized sets. Let $\mathfrak{Q} = \{Q \subseteq Y \mid f^{-1}(Q) \in \mathfrak{M}\} \subseteq 2^Y$ be the collection of "civilized sets". We want to show that $\mathfrak{Q} = \mathfrak{N}$ (ie: all sets in sigma algebra $\mathfrak{N}$ are civilized).
Claim 1: $\mathfrak{Q}$ is a sigma-algebra.
Claim 2: $\mathfrak{Q}$ contains $\mathfrak{N}$.
Hence, $\mathfrak{N} = \sigma\text{-Alg}(\mathfrak{C}) \subseteq \mathfrak{Q}$, so $\mathfrak{N} = \mathfrak{Q}$. Therefore, f is measurable. □

Corollary 3.5

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

Proof: Note that $\mathcal{T}_Y$ is a generator for $\mathfrak{B}_Y$. For $G \in \mathcal{T}_Y$, $f^{-1}(G)$ is open (since f is continuous), hence $f^{-1}(G) \in \mathfrak{B}_X$. By Prop 3.4, since $\sigma\text{-Alg}(\mathcal{T}_Y) = \mathcal{B}_Y$, we have that f is measurable. □

Definition 3.6

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

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

Proposition 3.7 **Stability properties of $\mathrm{Bor}(X, \mathbb{R})$**

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

1. $f + g \in \mathrm{Bor}(X, \mathbb{R})$, where $(f + g)(x) := f(x) + g(x)$.
2. $\alpha f + \beta g \in \mathrm{Bor}(X, \mathbb{R})$ for $\alpha, \beta \in \mathbb{R}$, where $(\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x)$.
3. $f \cdot g \in \mathrm{Bor}(X, \mathbb{R})$, where $(f \cdot g) = f(x)g(x)$.
4. $f \vee g, f \wedge g \in \mathrm{Bor}(X, \mathbb{R})$, where $(f \vee g)(x) := \max(f(x), g(x))$ and $(f \wedge g)(x) := \min(f(x), g(x))$.

Proof of 3: Let $h : X \to \mathbb{R}^2$ be defined by $h(x) = (f(x), g(x))$ for all $x \in X$. Let $p : \mathbb{R}^2 \to \mathbb{R}$ be defined by $p(s, t) = s \cdot t$ for all $s, t \in \mathbb{R}$.
Claim 1: h is $\mathfrak{M}/\mathcal{B}_{\mathbb{R}^2}$ measurable.
Proof: We claim that $\mathcal{B}_{\mathbb{R}^2} = \sigma\text{-Alg}(\mathcal{T})$, where $\mathcal{T} = \{\text{all open squares in } \mathbb{R}^2\}$. Once this claim is established, it suffices to show that $h^{-1}(S) \in \mathfrak{M}$ for all

$$S = (\alpha_1, \alpha_2) \times (\beta_1, \beta_2) \in \mathcal{T}$$

□

Claim 2: p is $\mathcal{B}_{\mathbb{R}^2} / \mathcal{B}_{\mathbb{R}}$-measurable.

<u>*Claim 3:*</u> $f \cdot g = p \circ h$.

Proof: By direct verification. □

□

Exercise 3.8

The Borel sigma-algebra on $\mathbb{R}^2$ is equal to the sigma-algebra generated by the open squares in $\mathbb{R}^2$, ie: $\sigma\text{-Alg}(\mathfrak{S}) = \mathcal{B}_{\mathbb{R}^2}$.

Remark 3.9

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

Remark 3.10

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

Example **Problem 3.1**

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

Proof: Denote $\sigma\text{-Alg}(\mathscr{C}) = \mathscr{P}$. Clearly, $\mathscr{P} \subseteq \mathcal{B}_{\mathbb{R}}$. If we can show that $G \in \mathscr{P}$ for all $G \subseteq \mathbb{R}$ open, then we will be done because we will have that $\mathscr{P}$ is a sigma-algebra and that $\mathscr{P} \supseteq \mathcal{T}_{\mathbb{R}}$.

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 $\sigma\text{-Alg}(\mathscr{P})$ contains all of the open intervals of $\mathbb{R}$. This is left as an exercise. □

By the claim, we have that $f^{-1}((-\infty, t]) \in \mathfrak{M}$ for all $t \in \mathbb{R}$. That is, $f^{-1}(C) \in \mathfrak{M}$ for all $C \in \mathscr{C}$. Then by Prop 3.4, $f^{-1}(B) \in \mathfrak{M}$ for all $B \in \mathcal{B}_{\mathbb{R}}$, hence $f \in \mathrm{Bor}(X, \mathbb{R})$.

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

Proposition 4.1 **Closure under Countable Supremum and Infimum**

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 $t \in \mathbb{R}$. Fix $t \in \mathbb{R}$. For $x \in X$, we have
$$\begin{aligned} x \in f^{-1}((-\infty, t]) &\iff f(x) \le t \\ &\iff \sup\{f_n(x) \mid n \in \mathbb{N}\} \le t \\ &\iff f_n(x) \le t \quad \text{for all } n \in \mathbb{N} \\ &\iff x \in f_n^{-1}((-\infty, t]) \quad \text{for all } n \in \mathbb{N} \\ &\iff x \in \bigcap_{n=1}^{\infty} f_n^{-1}((-\infty, t]). \end{aligned}$$
Hence
$$f^{-1}((-\infty, t]) = \bigcap_{n=1}^{\infty} f_n^{-1}((-\infty, t]).$$
Since each $f_n \in \mathrm{Bor}(X, \mathbb{R})$, we have $f_n^{-1}((-\infty, t]) \in \mathfrak{M}$ for every $n \in \mathbb{N}$. Since $\mathfrak{M}$ is a sigma-algebra, it is closed under countable intersections, so $f^{-1}((-\infty, t]) \in \mathfrak{M}$. Therefore $f \in \mathrm{Bor}(X, \mathbb{R})$.
2. Define $f : X \to \mathbb{R}$ by $f(x) = -g(x)$ for all $x \in X$, and define $f_n : X \to \mathbb{R}$ by $f_n(x) = -g_n(x)$ for all $n \in \mathbb{N}$. Since $g_n \in \mathrm{Bor}(X, \mathbb{R})$ for every $n \in \mathbb{N}$, we have $f_n \in \mathrm{Bor}(X, \mathbb{R})$ for every $n \in \mathbb{N}$.

   For every $x \in X$,
$$\begin{aligned} f(x) &= -g(x) \\ &= -\inf\{g_n(x) \mid n \in \mathbb{N}\} \\ &= \sup\{-g_n(x) \mid n \in \mathbb{N}\} \\ &= \sup\{f_n(x) \mid n \in \mathbb{N}\}. \end{aligned}$$

By part 1, $f \in \mathrm{Bor}(X, \mathbb{R})$. Since $g = -f$, it follows that $g \in \mathrm{Bor}(X, \mathbb{R})$.

$\square$

Remark 4.2 **Characterization of $\limsup$ and $\liminf$.**

Let $(t_n)_{n=1}^{\infty}$ be a bounded sequence in $\mathbb{R}$. Define $\ell_+ = \limsup_{n\to\infty} t_n$ to be the unique number satisfying:

1. (LIM-SUP-1) Some subsequence of $(t_n)_{n=1}^{\infty}$ converges to ℓ_+, and no subsequence converges to a limit $\ell > \ell_+$.
2. (LIM-SUP-2)

$$\ell_+ = \lim_{k\to\infty} \sup_{n\geq k} t_n = \inf_{k\geq 1} \sup_{n\geq k} t_n.$$

Similarly, $\ell_- = \liminf_{n\to\infty} t_n$ is the smallest possible subsequential limit, and

$$\liminf_{n\to\infty} t_n = \lim_{k\to\infty} \inf_{n\geq k} t_n = \sup_{k\geq 1} \inf_{n\geq k} t_n.$$

Proposition 4.3 **Measurability of Pointwise Limsup and Liminf**

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

1. Let $(f_n)_{n=1}^{\infty}$ be a sequence in $\mathrm{Bor}(X, \mathbb{R})$ such that $(f_n(x))_{n=1}^{\infty}$ is bounded for every $x \in X$. If $f : X \to \mathbb{R}$ is given by
$$f(x) = \limsup_{n\to\infty} f_n(x),$$
then $f \in \mathrm{Bor}(X, \mathbb{R})$.
2. Let $(g_n)_{n=1}^{\infty}$ be a sequence in $\mathrm{Bor}(X, \mathbb{R})$ such that $(g_n(x))_{n=1}^{\infty}$ is bounded for every $x \in X$. If $g : X \to \mathbb{R}$ is given by
$$g(x) = \liminf_{n\to\infty} g_n(x),$$
then $g \in \mathrm{Bor}(X, \mathbb{R})$.

Proof:

1. For every $k \in \mathbb{N}$, define $h_k : X \to \mathbb{R}$ by
$$h_k(x) = \sup\{f_n(x) \mid n \geq k\}.$$
By Proposition 4.1.1, $h_k \in \mathrm{Bor}(X, \mathbb{R})$ for every $k \in \mathbb{N}$. By LIM-SUP-2,
$$f(x) = \limsup_{n\to\infty} f_n(x) = \inf\{h_k(x) \mid k \in \mathbb{N}\}.$$
Hence $f \in \mathrm{Bor}(X, \mathbb{R})$ by Proposition 4.1.2.
2. Apply part 1 to the sequence $(-g_n)_{n=1}^{\infty}$:
$$\limsup_{n\to\infty}(-g_n(x)) = -\liminf_{n\to\infty} g_n(x) = -g(x).$$

Thus $-g \in \mathrm{Bor}(X, \mathbb{R})$, and hence $g \in \mathrm{Bor}(X, \mathbb{R})$.

$\square$

Corollary 4.4 **Pointwise Limits of Borel Functions are Borel**

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

Proof: Fix $x \in X$. Since $f(x) = \lim_{n\to\infty} f_n(x)$, the sequence $(f_n(x))_{n=1}^{\infty}$ is convergent. Hence every subsequence of $(f_n(x))_{n=1}^{\infty}$ also converges to $f(x)$. By LIM-SUP-2,

$$\limsup_{n\to\infty} f_n(x) = \lim_{k\to\infty} \sup_{n\geq k} f_n(x) = f(x).$$

Thus f is the pointwise lim sup of the sequence $(f_n)_{n=1}^{\infty}$. Therefore $f \in \mathrm{Bor}(X, \mathbb{R})$ by Proposition 4.3.1. $\square$

Example

Suppose $(X, \mathfrak{M}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable (which implies continuous, which implies f Borel measurable) and consider the function $f' : \mathbb{R} \to \mathbb{R}$.

We claim that $f' \in \mathrm{Bor}(\mathbb{R}, \mathbb{R})$.

Solution: the idea is to produce $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}(\mathbb{R}, \mathbb{R})$ such that

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

is in fact continuous. If we find $f_n{}'s$, then we get $f' \in \mathrm{Bor}(\mathbb{R}, \mathbb{R})$. due to Corollary 4.4. For $x \in \mathbb{R}$, we write

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

with $u_n(x) := f\left(x + \frac{1}{n}\right)$ for $x \in X, n \in \mathbb{N}$. So then, for every $n \in N$, define $f_n = n(u_n - f)$. Every f_n is continuous, hence in $\mathrm{Bor}(\mathbb{R}, \mathbb{R})$ and

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

as required.

Definition 4.5 **Simple functions**

We say that $f \in \mathrm{Bor}(X, \mathbb{R})$ is *simple* when $f(X)$ is finite. We denote

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

$\mathrm{Bor}_s(X, \mathbb{R})$ is stable under the algebraic operations from Proposition 3.7: linear combinations, multiplication, and min/max.

Also,

$$\mathrm{Bor}_s(X, \mathbb{R}) = \mathrm{span}\{\mathcal{X}_A \mid A \in \mathfrak{M}\},$$

where $\mathcal{X}_A$ denotes the indicator function of A. The equality is left as part of A2.

We also use

1. $$\mathrm{Bor}^+(X, \mathbb{R}) := \{f \in \mathrm{Bor}(X, \mathbb{R}) \mid f(x) \geq 0 \text{ for all } x \in X\}$$
2. $$\begin{aligned} \mathrm{Bor}_s^+(X, \mathbb{R}) &:= \mathrm{Bor}_s(X, \mathbb{R}) \cap \mathrm{Bor}^+(X, \mathbb{R}) \\ &= \{f \in \mathrm{Bor}_s(X, \mathbb{R}) \mid f(x) \geq 0 \text{ for all } x \in X\} \end{aligned}$$

Remark 4.6 **Binning of values of $f \in \mathrm{Bor}^+(X, \mathbb{R})$**

Let $f \in \mathrm{Bor}^+(X, \mathbb{R})$. For every $n \in \mathbb{N}$, the *n'th binning of the values of f* is the function $f_n : X \to \mathbb{R}$ defined by

$$f_n(x) := \begin{cases} n, \text{if } f(x) \geq n \\ \frac{k-1}{2^n}, \text{if } f(x) < n \text{ and } k \in \{1, \ldots, n2^n\} \text{ is such that } \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \end{cases}$$

For this fixed n, let

$$B = f^{-1}([n, \infty)) \in \mathfrak{M}$$

and

$$A_k = f^{-1}\left(\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right) \in \mathfrak{M}, \quad 1 \leq k \leq n2^n$$

Then

$$f_n = n\mathcal{X}_B + \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \mathcal{X}_{A_k} \in \mathrm{span}\{\mathcal{X}_A \mid A \in \mathfrak{M}\} = \mathrm{Bor}_s(X, \mathbb{R})$$

Since $f_n \geq 0$, it follows that $f_n \in \mathrm{Bor}_s^+(X, \mathbb{R})$.

Proposition 4.7

Let $f \in \mathrm{Bor}^+(X,\mathbb{R})$. Then there exists a sequence $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}_s^+(X,\mathbb{R})$ such that for every $x \in X$,

$$0 \le f_1(x) \le f_2(x) \le \dots \le f_n(x) \le \dots$$

and

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

Proof: For every $n \in \mathbb{N}$, let f_n be the n'th binning of the values of f from Remark 4.6. By Remark 4.6, $f_n \in \mathrm{Bor}_s^+(X,\mathbb{R})$ for every $n \in \mathbb{N}$.

Fix $x \in X$. By Problem P4.4, $f_n(x) \le f_{n+1}(x)$ for all $n \in \mathbb{N}$, and there exists $n_0 \in \mathbb{N}$ such that

$$f_n(x) \le f(x) < f_n(x) + \frac{1}{2^n}, \quad \text{for all } n \ge n_0.$$

Hence $(f_n(x))_{n=1}^{\infty}$ is increasing, and the inequalities above imply $\lim_{n\to\infty} f_n(x) = f(x)$. $\square$

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

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

Recall that

$$\mathrm{Bor}^+(X,\mathbb{R}) = \{f \in \mathrm{Bor}(X,\mathbb{R}) \mid f(x) \ge 0 \ \text{ for all } \ x \in X\}.$$

The goal is to define a map

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

where $L^+(f)$ is the integral of f against the measure μ. The value is defined for every $f \in \mathrm{Bor}^+(X,\mathbb{R})$, but it may be ∞.

We construct L^+ in two steps:

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

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

Remark 5.1 **Basic stability 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. For every $A \in \mathfrak{M}$, the indicator function $\mathcal{X}_A$ is in $\mathrm{Bor}_s^+(X,\mathbb{R})$. In particular, $\mathcal{X}_\varnothing = \underline{0}$ and $\mathcal{X}_X = \underline{1}$ are in $\mathrm{Bor}_s^+(X,\mathbb{R})$.

Theorem 5.2

There exists a unique map

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

such that the following conditions are satisfied:

1. Additivity: $L_s^+(f+g) = L_s^+(f) + L_s^+(g)$ for all $f, g \in \mathrm{Bor}_s^+(X, \mathbb{R})$.
2. Positive homogeneity: $L_s^+(\alpha f) = \alpha \cdot L_s^+(f)$ for all $\alpha \in [0, \infty)$ and $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$.
3. Relation to the measure: $L_s^+(\mathcal{X}_A) = \mu(A)$ for all $A \in \mathfrak{M}$.

Definition 5.3 **Measurable divisions**

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

1. A *measurable division of* X is a set $\Delta = \{A_1, ..., A_n\}$ of pairwise disjoint non-empty sets in $\mathfrak{M}$ which cover X.
2. If $\Delta = \{A_1, ..., A_n\}$ and $\Gamma = \{B_1, ..., B_m\}$ are measurable divisions of X, then their intersection is the measurable division

$$\Delta \wedge \Gamma := \{A \cap B \mid A \cap B \neq \varnothing, A \in \Delta, B \in \Gamma\}.$$

Lemma 5.4

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

$$f = \alpha_1 \mathcal{X}_{A_1} + ... + \alpha_n \mathcal{X}_{A_n}.$$

2. Suppose $\Delta = \{A_1, ..., A_n\}$ and $\alpha_1, ..., \alpha_n$ are as in part 1, and let $\Gamma = \{B_1, ..., B_m\}$ be a measurable division of X. Then

$$f = \sum_{\substack{1 \le i \le n, \\ 1 \le j \le m, \\ A_i \cap B_j \neq \varnothing}} \alpha_i \mathcal{X}_{A_i \cap B_j}.$$

Proof:

1. Exercise: See P4.3 (very similar to it).
2. For every $1 \le i \le n$, write

$$\begin{aligned}\chi_{A_i} &= \chi_{A_i} \cdot 1 \\ &= \chi_{A_i}\left(\chi_{B_1} + \dots + \chi_{B_m}\right) \\ &= \sum_{j=1}^{m} \chi_{A_i \cap B_j} \\ &= \sum_{\substack{1 \le j \le m, \\ A_i \cap B_j \ne \varnothing}} \chi_{A_i \cap B_j}\end{aligned}$$

Now, we substitute the expression for χ_{A_i} into Simple Span, as follows:

$$\begin{aligned}f = \sum_{i=1}^{n} \alpha_i \chi_{A_i} &= \sum_{i=1}^{n} \alpha_i \left(\sum_{\substack{1 \le j \le m, \\ A_i \cap B_j \ne \varnothing}} \chi_{A_i \cap B_j} \right) \\ &= \end{aligned}$$

□

Lemma 5.5

Let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$. Suppose we have two ways of writing f as a linear combination:

$$f = \alpha_1 \mathcal{X}_{A_1} + \dots + \alpha_n \mathcal{X}_{A_n}$$

and

$$f = \beta_1 \mathcal{X}_{B_1} + \dots + \beta_m \mathcal{X}_{B_m},$$

where $\Delta = \{A_1, \dots, A_n\}$ and $\Gamma = \{B_1, \dots, B_m\}$ are measurable divisions of X, and $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in [0, \infty)$. 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 $1 \le i \le n$, let us write

$$\begin{aligned}A_i &= A_i \cap X \\ &= A_i \cap (B_1 \cup \dots \cup B_m) \\ &= \bigcup_{j=1}^{m} (A_i \cap B_j) \\ &= \bigcup_{\substack{1 \le j \le m, \\ A_i \cap B_j \ne \varnothing}} (A_i \cap B_j)\end{aligned}$$

which is a disjoint union. So, we have

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

Hence, we have that

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

Similarly, we obtain that

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

The following claim suffices to complete the proof.

Claim : If $A_i \cap B_j \ne \varnothing$, then $\alpha_i = \beta_j$.
Proof: Suppose $A_i \cap B_j \ne \varnothing$. Pick $x \in A_i \cap B_j$, and evaluate $f(x)$ in two ways. From 1, we have $f(x) = \alpha_1 \chi_{A_1}(x) + ... \alpha_n \chi_{A_n}(x) = \alpha_i$.
From $f(x) = \beta_1 \chi_{\beta_1}(x) + ... + \beta_m \chi_{B_m}(x)$, we get $f(x) = \beta_j$. Hence, $\alpha_i = f(x) = B_j$. □

□

Proof of existence (Thm 5.2): Let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$. By Lemma 5.4, choose a measurable division

$$\Delta = \{A_1, ..., A_n\}$$

of X and numbers $\alpha_1, ..., \alpha_n \in [0, \infty)$ such that

$$f = \sum_{i=1}^{n} \alpha_i \chi_{A_i}.$$

Define

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

Lemma 5.5 says exactly that this value is independent of the chosen writing of f as a linear combination of indicators coming from a measurable division. Hence L_s^+ is well-defined as a map

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

We check the three required properties from Theorem 5.2:

1. Let $f, g \in \mathrm{Bor}_s^+(X, \mathbb{R})$. Choose measurable divisions

$$\Delta = \{A_1, ..., A_n\}, \quad \Gamma = \{B_1, ..., B_m\},$$

and coefficients $\alpha_i, \beta_j \in [0, \infty)$ such that

$$f = \sum_{i=1}^{n} \alpha_i \chi_{A_i}, \quad g = \sum_{j=1}^{m} \beta_j \chi_{B_j}.$$

Refine both writings to the measurable division

$$\Delta \wedge \Gamma = \left\{A_i \cap B_j \mid A_i \cap B_j \neq \varnothing\right\}.$$

By Lemma 5.4,

$$f = \sum_{A_i \cap B_j \neq \varnothing} \alpha_i \chi_{A_i \cap B_j}, \quad g = \sum_{A_i \cap B_j \neq \varnothing} \beta_j \chi_{A_i \cap B_j}.$$

Thus

$$f + g = \sum_{A_i \cap B_j \neq \varnothing} (\alpha_i + \beta_j) \chi_{A_i \cap B_j}.$$

Therefore

$$\begin{aligned} L_s^+(f+g) &= \sum_{A_i \cap B_j \neq \varnothing} (\alpha_i + \beta_j) \mu(A_i \cap B_j) \\ &= \sum_{A_i \cap B_j \neq \varnothing} \alpha_i \mu(A_i \cap B_j) + \sum_{A_i \cap B_j \neq \varnothing} \beta_j \mu(A_i \cap B_j) \\ &= \sum_{i=1}^{n} \alpha_i \mu(A_i) + \sum_{j=1}^{m} \beta_j \mu(B_j) \\ &= L_s^+(f) + L_s^+(g). \end{aligned}$$

2. Let $\alpha \in [0, \infty)$ and let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$. Write

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

using a measurable division $\Delta = \{A_1, ..., A_n\}$. Then

$$\alpha f = \sum_{i=1}^{n} (\alpha \alpha_i) \chi_{A_i},$$

so

$$L_s^+(\alpha f) = \sum_{i=1}^{n} (\alpha \alpha_i) \mu(A_i) = \alpha \sum_{i=1}^{n} \alpha_i \mu(A_i) = \alpha L_s^+(f).$$

3. Let $A \in \mathfrak{M}$. If $\varnothing \neq A \neq X$, then $\{A, X \setminus A\}$ is a measurable division of X and

$$\chi_A = 1 \cdot \chi_A + 0 \cdot \chi_{X \setminus A}.$$

Hence

$$L_s^+(\chi_A) = 1 \cdot \mu(A) + 0 \cdot \mu(X \setminus A) = \mu(A).$$

If $A = \varnothing$ or $A = X$, the same conclusion follows from the one-piece measurable division $\{X\}$.

□

Proof of uniqueness: Exercise. □

Definition 5.6 **Integral in $\mathrm{Bor}^+(X, \mathbb{R})$**

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

1. We use $u, v, \ldots$ for general functions in $\mathrm{Bor}^+(X, \mathbb{R})$ and $f, g, \ldots$ for functions in $\mathrm{Bor}_s^+(X, \mathbb{R})$.
2. The inequality $f \leq u$ means $f(x) \leq u(x)$ for every $x \in X$.
3. The set whose supremum is taken is non-empty, since $0 \in \mathrm{Bor}_s^+(X, \mathbb{R})$, $0 \leq u$, and $L_s^+(0) = 0$.
4. The supremum is taken in $[0, \infty]$: if the set is bounded above by a finite number, this is the usual least upper bound; otherwise its supremum is defined to be ∞.

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 $L^+(u) = L_s^+(u)$.

Proof: We must show that

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

$\geq$ holds because $L_s^+(u)$ is an element of $\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \leq u\}$.
$\leq$ holds because if $f \in \mathrm{Bor}(X, \mathbb{R})$ with $f \leq u$, then
$L_s^+(u) = L_s^+(f + (u - f)) = L_s^+(f) + L_s^+(u - f)$, with $L_s^+(u - f) \geq 0$. □

Proposition 5.9

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

$$L^+(\alpha u) = \alpha L^+(u).$$

Proof:

1. Suppose $u \le v$. Then
$$\{f \in \mathrm{Bor}_s^+(X,\mathbb{R}) \mid f \le u\} \subseteq \{f \in \mathrm{Bor}_s^+(X,\mathbb{R}) \mid f \le v\}.$$
Therefore,
$$\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X,\mathbb{R}), f \le u\} \subseteq \{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X,\mathbb{R}), f \le v\}.$$
Taking suprema gives $L^+(u) \le L^+(v)$.
2. If $\alpha = 0$, then $\alpha u = 0$, so $L^+(\alpha u) = L^+(0) = 0$. Also $\alpha L^+(u) = 0 \cdot L^+(u) = 0$, using the convention $0 \cdot \infty = 0$ in $[0,\infty]$.

   Now suppose $\alpha > 0$. We claim that
$$\{g \in \mathrm{Bor}_s^+(X,\mathbb{R}) \mid g \le \alpha u\} = \{\alpha f \mid f \in \mathrm{Bor}_s^+(X,\mathbb{R}), f \le u\}.$$
   Indeed, if $f \in \mathrm{Bor}_s^+(X,\mathbb{R})$ and $f \le u$, then $\alpha f \in \mathrm{Bor}_s^+(X,\mathbb{R})$ and $\alpha f \le \alpha u$. Conversely, if $g \in \mathrm{Bor}_s^+(X,\mathbb{R})$ and $g \le \alpha u$, then $f := \frac{g}{\alpha}$ is in $\mathrm{Bor}_s^+(X,\mathbb{R})$ and $f \le u$, so $g = \alpha f$.

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

□

Remark 5.10

The superadditivity of the following desirable equality can be easily proved, but not the subadditivity part:
$$L^+(u+v) = L^+(u) + L^+(v)$$
for all $u, v \in \mathrm{Bor}^+(X,\mathbb{R})$. So we pivot to Monotone Convergence Theorem to find a way to prove it.

6 The Monotone Convergence Theorem (MCT)

Notation 6.1

Let u and $u_1, u_2, \dots, u_n, \dots$ be functions in $\mathrm{Bor}^+(X, \mathbb{R})$. We write

$$u_n \nearrow u$$

to mean that, for every $x \in X$,

$$0 \leq u_1(x) \leq u_2(x) \leq \dots \leq u_{n(x)} \leq \dots$$

and

$$\lim_{n \to \infty} u_n(x) = u(x).$$

Theorem 6.2 **Monotone Convergence Theorem**

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

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

where the limit is an increasing limit in $[0, \infty]$.

Proposition 6.3 **Subadditivity of L^+**

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

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

Proof: By Proposition 4.7, there are sequences $(f_n)_{n=1}^{\infty}$ and $(g_n)_{n=1}^{\infty}$ in $\mathrm{Bor}_s^+(X, \mathbb{R})$ such that

$$f_n \nearrow u \quad \text{and} \quad g_n \nearrow v.$$

Then $(f_n + g_n)_{n=1}^{\infty}$ is a sequence in $\mathrm{Bor}_s^+(X, \mathbb{R})$ and $f_n + g_n \nearrow u + v$. Hence

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

The first equality follows from the Monotone Convergence Theorem. The second and fourth use that L^+ extends L_s^+, and the third uses the additivity of L_s^+.

Since $f_n \leq u$ and $g_n \leq v$, the monotonicity of L^+ gives

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

for every $n \in \mathbb{N}$. Taking the increasing limit yields

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

as required. □

Remark 6.4

To prove the Monotone Convergence Theorem, it remains to establish the following inequality. Suppose $u_n \nearrow u$ in $\mathrm{Bor}^+(X, \mathbb{R})$, and set

$$\Lambda := \lim_{n\to\infty} L^+(u_n) = \sup\{L^+(u_n) \mid n \in \mathbb{N}\},$$

where the limit is increasing in $[0, \infty]$. We need to prove that

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

This inequality is sufficient because $u_n \leq u$ for every n. By the monotonicity of L^+,

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

for every $n \in \mathbb{N}$. Taking the supremum over n gives

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

Thus, once the reverse inequality $\Lambda \geq L^+(u)$ is proved, we obtain $\Lambda = L^+(u)$, which is precisely the Monotone Convergence Theorem.

We will prove the reverse inequality in Proposition 6.7. First, we establish two special cases in Lemmas 6.5 and 6.6.

Lemma 6.5 **Essence of the Monotone Convergence Theorem**

Let $u_1 \leq u_2 \leq \ldots \leq u_n \leq \ldots$ be functions in $\mathrm{Bor}^+(X, \mathbb{R})$, and let $A \in \mathcal{M}$ satisfy $u_n \nearrow \chi_A$. Define

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

Then

$$\Lambda \geq \mu(A) = L^+(\chi_A).$$

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

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

Observe:

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 $\bigcup_{n=1}^{\infty} A_n = A$. By continuity of μ along increasing chains,

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

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 the monotonicity and homogeneity of L^+, the fact that L^+ extends L_s^+, and the formula for L_s^+ on indicator functions,

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

So we have

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

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

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

where we use Equation ($\star$) for the right-hand side. Finally, unfix θ. 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 $u_1 \leq u_2 \leq \ldots \leq u_n \leq \ldots$ be functions in $\mathrm{Bor}^+(X, \mathbb{R})$, and let $f_0 \in \mathrm{Bor}_s^+(X, \mathbb{R})$ satisfy $u_n \nearrow f_0$. Define

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

Then

$$\Lambda \geq L_s^+(f_0) = L^+(f_0).$$

Proof: Write

$$f_0 = \sum_{i=1}^{m} \alpha_i \chi_{A_i},$$

where $\{A_1, \ldots, A_m\}$ is a measurable division of X and $\alpha_1, \ldots, \alpha_m \in [0, \infty)$. Fix $c \in (0, 1)$ and, for each i, define

$$A_{i,n} := \{x \in A_i \mid u_n(x) \geq c\alpha_i\}.$$

For each i, the sets $A_{i,n}$ increase to A_i: indeed, if $x \in A_i$, then $u_n(x) \to f_0(x) = \alpha_i$, so eventually $u_n(x) \geq c\alpha_i$. Define the simple function

$$h_n := \sum_{i=1}^{m} c\alpha_i \chi_{A_{i,n}}.$$

Because the sets $A_{i,n}$ lie in the disjoint sets A_i, we have $h_n \leq u_n$. Therefore,

$$\begin{aligned}\Lambda &\geq L^+(u_n)\\ &\geq L^+(h_n)\\ &= L_s^+(h_n)\\ &= c\sum_{i=1}^{m} \alpha_i \mu(A_{i,n}).\end{aligned}$$

Letting $n \to \infty$ and using continuity of μ along increasing chains gives

$$\Lambda \geq c\sum_{i=1}^{m} \alpha_i \mu(A_i) = cL_s^+(f_0).$$

If $L_s^+(f_0) = \infty$, this implies $\Lambda = \infty$. Otherwise, let $c \to 1$ to obtain $\Lambda \geq L_s^+(f_0)$. Since L^+ extends L_s^+, $L_s^+(f_0) = L^+(f_0)$. $\square$

Proposition 6.7

Let $(u_n)_{n=1}^{\infty}$ be a sequence in $\mathrm{Bor}^+(X, \mathbb{R})$ that increases pointwise to $u \in \mathrm{Bor}^+(X, \mathbb{R})$, and 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) $L^+(\mathcal{X}_A) = \mu(A)$ 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 $(u_n)_{n=1}^{\infty}$ increases pointwise to u.

Note

For the existence part of the theorem, property (iv) is the Monotone Convergence Theorem (Theorem 6.2). Property (i) follows from the formula for L_s^+ in Theorem 5.2 and the fact that L^+ extends L_s^+ (Proposition 5.8). Property (ii) is Proposition 6.3, and property (iii) is the homogeneity statement in Proposition 5.9.

For the uniqueness part, it is convenient to record the following lemma, which will also be useful in subsequent lectures.

Lemma 6.9 **Method of the Friendly Functions**

Let $(X, \mathfrak{M})$ be a measurable space, and let $\mathcal{G} \subseteq \mathrm{Bor}^+(X, \mathbb{R})$ satisfy the following properties:

(i) $\mathcal{X}_A \in \mathcal{G}$ for every $A \in \mathfrak{M}$.
(ii) If $g_1, g_2 \in \mathcal{G}$, then $g_1 + g_2 \in \mathcal{G}$.
(iii) If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then $\alpha g \in \mathcal{G}$.
(iv) If $g \in \mathrm{Bor}^+(X, \mathbb{R})$ and $(g_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{G}$ such that $g_n \nearrow g$, then $g \in \mathcal{G}$.

Then $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$.

Proof: Since $\mathcal{G} \subseteq \mathrm{Bor}^+(X, \mathbb{R})$ by assumption, it remains to prove the reverse inclusion.

First, we show that $\mathcal{G}$ contains every nonnegative simple function. Let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$. By Lemma 5.4, there exist measurable sets $A_1, \ldots, A_m \in \mathfrak{M}$ and constants $\alpha_1, \ldots, \alpha_m \in [0, \infty)$ such that

$$f = \sum_{j=1}^{m} \alpha_j \chi_{A_j}.$$

By property (i), $\chi_{A_j} \in \mathcal{G}$ for each j. Property (iii) then gives $\alpha_j \chi_{A_j} \in \mathcal{G}$, and repeated application of property (ii) gives $f \in \mathcal{G}$. Hence

$$\mathrm{Bor}_s^+(X, \mathbb{R}) \subseteq \mathcal{G}.$$

Now fix $u \in \mathrm{Bor}^+(X, \mathbb{R})$. By the binning approximation of Proposition 4.7, there exists a sequence $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}_s^+(X, \mathbb{R})$ such that

$$f_n \nearrow u.$$

The first part shows that $f_n \in \mathcal{G}$ for every $n \in \mathbb{N}$. Therefore, property (iv) gives $u \in \mathcal{G}$. Since $u \in \mathrm{Bor}^+(X, \mathbb{R})$ was arbitrary, $\mathrm{Bor}^+(X, \mathbb{R}) \subseteq \mathcal{G}$, and consequently

$$\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R}).$$

□

Proof of Uniqueness for Theorem 6.8:
Suppose $M^+ : \mathrm{Bor}^+(X, \mathbb{R}) \to [0, \infty]$ satisfies the conditions $(i) - (iv)$ from Theorem 6.8. Let $\mathcal{G} = \{g \in \mathrm{Bor}^+(X, \mathbb{R}) \mid L^+(g) = M^+(g)\}$. We want to show that $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$ (by method of friendly functions). We do this by showing that $\mathcal{G}$ satisfies $(i) - (iv)$ from Lemma 6.9.

(i) $\chi_A \in \mathcal{G}$ because $L^+(\chi_A) = \mu(A) = M^+(\chi_A)$

(ii) If $g_1, g_2 \in \mathcal{G}$, then $g_1 + g_2 \in \mathcal{G}$ because

$$\begin{aligned} L^+(g_1 + g_2) &= L^+(g_1) + L^+(g_2) \\ &= M^+(g_1) + M^+(g_2) \\ &= M^+(g_1 + g_2) \end{aligned}$$

(iii) If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then

$$L^+(\alpha g) = \alpha L^+(g) = \alpha M^+(g) = M^+(\alpha g).$$

Hence $\alpha g \in \mathcal{G}$.

(iv) Suppose $g \in \mathrm{Bor}^+(X, \mathbb{R})$ and $(g_n)_{n=1}^\infty$ is a sequence in $\mathcal{G}$ such that $g_n \nearrow g$. Since both L^+ and M^+ satisfy property (iv) of Theorem 6.8,

$$L^+(g) = \lim_{n \to \infty} L^+(g_n) = \lim_{n \to \infty} M^+(g_n) = M^+(g).$$

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

Therefore, $\mathcal{G}$ satisfies all the hypotheses of Lemma 6.9, so $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$. Hence $L^+(g) = M^+(g)$ for every $g \in \mathrm{Bor}^+(X, \mathbb{R})$, and therefore $M^+ = L^+$. $\square$

7 The Space $\mathcal{L}^1(\mu)$

Notation 7.1

Fix $(X, \mathfrak{M}, \mu)$ as before and have $L^+ : \mathrm{Bor}^+(X, \mathbb{R}) \to [0, \infty]$, which is trying to become the integral with respect to μ. We will see a map $L : \mathcal{L}^1(\mu) \to \mathbb{R}$ which also tries to become the integral with respect to μ.

To define L, we reduce each $f \in \mathrm{Bor}(X, \mathbb{R})$ to two non-negative Borel functions. The *positive part* and *negative part* of f are respectively defined by

$$f_+ := f \vee 0 \quad \text{and} \quad f_- := (-f) \vee 0.$$

Since $\mathrm{Bor}(X, \mathbb{R})$ is stable under $\vee$, we have $f_+, f_- \in \mathrm{Bor}^+(X, \mathbb{R})$. Explicitly, for every $x \in X$,

$$f_+(x) = \begin{cases} f(x), & \text{if } \ f(x) \geq 0 \\ 0, & \text{if } \ f(x) < 0 \end{cases} \quad \text{and} \quad f_-(x) = \begin{cases} 0, & \text{if } \ f(x) \geq 0 \\ -f(x), & \text{if } \ f(x) < 0. \end{cases}$$

These functions satisfy

$$f_+ + f_- = |f| \quad \text{and} \quad f_+ - f_- = f.$$

Equivalently (by adding or subtrating the above two equations), we obtain

$$f_+ = \frac{f + |f|}{2} \quad \text{and} \quad f_- = \frac{|f| - f}{2}.$$

Lemma 7.2

For $f \in \mathrm{Bor}(X, \mathbb{R})$, we have the equivalence

$$L^+(|f|) < \infty \Leftrightarrow (L^+(f_+) < \infty \quad \text{and} \quad L^+(f_-) < \infty).$$

Proof:

($\Longrightarrow$) Since $|f| = f_+ + f_-$, then $|f| \geq f_+$ and $|f| \geq f_-$, so by the increasing property of L^+, we get the RHS.

($\Longleftarrow$)

$$\begin{aligned} L^+(|f|) &= L^+(f_+ + f_-) \text{ (by additivity of } L^+) \\ &= L^+(f_+) + L^+(f_-) < \infty \end{aligned}$$

$\square$

Definition 7.3

A function $f : X \to \mathbb{R}$ is *integrable with respect to* μ when $f \in \mathrm{Bor}(X, \mathbb{R})$ and the equivalent conditions of the preceding lemma hold. We denote

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

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

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

This is the difference of two finite numbers in $[0, \infty)$.

We will prove that L is linear. First, we establish a lemma that gives us some flexibility in calculating $L(f)$: instead of using only $f = f_+ - f_-$, it allows us to use other representations $f = h_1 - h_2$.

Lemma 7.4

Let $f \in \mathcal{L}^1(\mu)$ and suppose that

$$f = h_1 - h_2$$

for some $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ such that $L^+(h_1), L^+(h_2) < \infty$. Then

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

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

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

Applying L^+ to both sides and using its additivity gives

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

All four quantities are finite: this holds for f_+ and f_- because $f \in \mathcal{L}^1(\mu)$, and for h_1 and h_2 by assumption. We may therefore rearrange the equality to obtain

$$L(f) = L^+(f_+) - L^+(f_-) = L^+(h_1) - L^+(h_2),$$

as required. □

Theorem 7.5

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: **1.** Pick $f, g \in \mathcal{L}^1(\mu)$ 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

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

which implies, by the monotonicity, additivity, and homogeneity of L^+,

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

The last inequality holds because $f, g \in \mathcal{L}^1(\mu)$. It follows that $h \in \mathcal{L}^1(\mu)$, proving 1.

2. We verify separately the two properties 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 2(a). Let $f, g \in \mathcal{L}^1(\mu)$. We may write

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

Set $h_1 := f_+ + g_+$ and $h_2 := f_- + g_-$. Then $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$, and additivity of L^+ gives

$$L^+(h_1) = L^+(f_+) + L^+(g_+) < \infty$$

and, similarly, $L^+(h_2) < \infty$. By Lemma 7.4,

$$\begin{aligned} L(f+g) &= L^+(h_1) - L^+(h_2) \\ &= L^+(f_+) + L^+(g_+) - L^+(f_-) - L^+(g_-) \\ &= (L^+(f_+) - L^+(f_-)) + (L^+(g_+) - L^+(g_-)) \\ &= L(f) + L(g). \end{aligned}$$

Verification of 2(b). Distinguish the cases $\alpha > 0, \alpha = 0, \alpha < 0$. □

Proposition 7.6

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

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

Proof:

1. Since $f \geq 0$, we have $f_+ = f$ and $f_- = 0$. Therefore

$$L(f) = L^+(f_+) - L^+(f_-) = L^+(f) - L^+(0) = L^+(f).$$

This shows that the newly defined integral L agrees with the old non-negative integral L^+ whenever both are defined.

2. Since $f \leq g$, the function $g - f$ is non-negative. By Theorem 7.5, $g - f \in \mathcal{L}^1(\mu)$, so part 1 gives

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

Using linearity of L,

$$L(g) - L(f) = L(g - f) \geq 0,$$

and hence $L(f) \leq L(g)$.

3. By Theorem 7.5, the function $|f| = f_+ + f_-$ belongs to $\mathcal{L}^1(\mu)$, since $f_+, f_- \in \mathcal{L}^1(\mu)$. Also,

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

Applying part 2 to these two inequalities gives

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

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

$\square$

Notation 7.7

Switching to integral notation. For the measure space $(X, \mathfrak{M}, \mu)$, we have constructed the two maps

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

On the intersection of their domains, the two maps agree: if $f \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu)$, then Proposition 7.6 gives $L(f) = L^+(f)$. It is therefore safe to use the same integral notation for both maps.

For $f \in \mathrm{Bor}^+(X, \mathbb{R}) \cup \mathcal{L}^1(\mu)$, we denote the corresponding value of L^+ or L by

$$\int_X f(x)\,\mathrm{d}\mu(x),$$

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

Although the notation is uniform, it can still represent two meanings:
- the L^+ meaning, with values in $[0, \infty]$, for non-negative Borel functions;
- the L meaning, with values in $\mathbb{R}$, for integrable functions.

These meanings agree whenever both apply. In particular, the inequality in Proposition 7.6 becomes

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

The integral on the right may be interpreted in either the L^+ sense or the L sense.

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

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

(integral in the L^+ sense), where, as usual, χ_A is the indicator function of A.

2. For any $f \in \mathcal{L}^1(\mu)$ and $A \in \mathfrak{M}$, we observe that $f \cdot \chi_A \in \mathcal{L}^1(\mu)$ (please fill in why), and we denote

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

(integral in the L sense).

Proposition 7.9

Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $h \in \mathrm{Bor}^+(X, \mathbb{R})$. 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 ν is a positive measure on $\mathfrak{M}$.

Proof:

1. First, we see that $\mu(\varnothing) = \int_X h \cdot \chi_\varnothing d\mu = \int_X 0 d\mu = 0$.
2. For the rest of the proof, we fix a sequence $(A_n)_{n=1}^{\infty}$ in $\mathfrak{M}$ such that $A_i \cap A_j = \varnothing$ for $i \neq j$. Let $U = \cup_{n=1}^{\infty} A_n$. We want to verify that

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

$\square$

Proposition 7.10

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

$$\int_X f \, \mathrm{d}\nu = \int_X fh \, \mathrm{d}\mu \in [0, \infty] \quad (\text{insert leaf})$$

(equality of integrals considered in the L^+ sense).

Proof: We use the method of the friendly functions (see Lemma 6.9). Let

$$\mathcal{G} = \left\{ g \in \mathrm{Bor}^+(X, \mathbb{R}) \mid \int_X g \, \mathrm{d}\nu = \int_X gh \, \mathrm{d}\mu \right\}$$

We verify that $\mathcal{G}$ satisfies conditions j to jv from Lemma 6.9.

The rest of the proof is on A4.

Then Lemma 6.9 gives $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$. This means that “leaf” holds for every $g \in \mathrm{Bor}^+(X, \mathbb{R})$. $\square$

8 Dominated Convergence Theorem

8.1 LDCT

Theorem 8.1 **LDCT**

Let f and $(f_n)_{n=1}^{\infty}$ be functions from $\mathrm{Bor}(X, \mathbb{R})$ 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 $f_1, f_2, \ldots, f_n, \ldots$ are all in $\mathcal{L}^1(\mu)$, and

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

Note

Note that the last equation can be written as

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

and thus really is a formula where one interchanges integration with a limit. In principle, doing such an interchange can have issues. The insight of Lebesgue was to spot a sufficient condition, the existence of the dominating function h, which makes the issues go away. The popularity of LDCT comes from the fact that this sufficient condition is easy to verify, and turns out to hold in many applications.

The proof of LDCT will be organized as follows:

$$\text{MCT} \implies \text{Reverse-MCT} \implies \text{LDCT}_0 \implies \text{LDCT}.$$

Proposition 8.2 **Reverse-MCT**

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

$$u_1 \geq u_2 \geq \ldots \geq u_n \geq \ldots$$

and such that

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

Assume, in addition, that

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

Then

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

***Proof* 8.3:**

TODO: See connection between this and the continuity of decreasing chains proof.

Define

$$w_n := u_1 - u_n, \quad n \in \mathbb{N}.$$

Since $u_1 \geq u_n$, each w_n is non-negative and Borel. Also, (w_n) is increasing: indeed, $u_n \geq u_{n+1}$ implies $u_1 - u_n \leq u_1 - u_{n+1}$. Finally, since $u_n \nearrow u$, we have

$$w_n \nearrow u_1 - u.$$

By MCT,

$$\lim_{n\to\infty} \int_X w_n \, \mathrm{d}\mu = \int_X (u_1 - u) \, \mathrm{d}\mu.$$

Since $\int_X u_1 \, \mathrm{d}\mu < \infty$, all quantities involved are finite, and linearity gives

$$\int_X w_n \, \mathrm{d}\mu = \int_X u_1 \, \mathrm{d}\mu - \int_X u_n \, \mathrm{d}\mu$$

and

$$\int_X (u_1 - u) \, \mathrm{d}\mu = \int_X u_1 \, \mathrm{d}\mu - \int_X u \, \mathrm{d}\mu.$$

Therefore

$$\lim_{n\to\infty}\left(\int_X u_1 \,\mathrm{d}\mu - \int_X u_n \,\mathrm{d}\mu\right) = \int_X u_1 \,\mathrm{d}\mu - \int_X u \,\mathrm{d}\mu.$$

Cancelling the finite number $\int_X u_1 \,\mathrm{d}\mu$ gives the desired conclusion. □

Proposition 8.4 LDCT$_0$

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

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

Proof 8.5:

Step 1: We define a decreasing sequence so that we can apply Reverse-MCT.

For each $n \in \mathbb{N}$, define using point-wise supremum the following functions:

$$u_n := \sup_{k\ge n} g_k.$$

Note that $u_n(x) \in [0,\infty)$ for all $x \in X$ because $(g_n(x))_{n=1}^{\infty}$ is convergent for all $x \in X$. Then $u_n \in \mathrm{Bor}^+(X,\mathbb{R})$ (by Proposition 4.1), and

$$u_1 \ge u_2 \ge \ldots \ge u_n \ge \ldots.$$

Step 2: We show that the u_n's converge to the zero function.

Moreover, since $g_n(x) \to 0$ for every $x \in X$ and all g_n are non-negative, we have

$$u_n(x) = \sup_{k\ge n} g_k(x) \to 0, \quad \forall x \in X.$$

Step 3: We use the reverse-MCT for (u_n).

Also $0 \le u_n \le h$ for every n, so

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

By Reverse-MCT applied to (u_n),

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

Since $0 \le g_n \le u_n$, monotonicity gives

$$0 \le \int_X g_n \, \mathrm{d}\mu \le \int_X u_n \, \mathrm{d}\mu.$$

Hence, by squeeze theorem, we get

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

as required. □

> **Remark 8.6**
>
> Fatou's lemma is about constructing a decreasing sequence using lim sups.
>
> $$\lim_{n\to\infty} u_n(x) = \lim_{n\to\infty} \sup\{g_k(x) \mid k \ge n\}$$

Proof of LDCT 8.7: Steps of proof:

1. For every $n \in \mathbb{N}$, we have $|f_n| \le h$ implies that $\int_X |f_n| \, \mathrm{d}\mu \le \int_X h \, \mathrm{d}\mu < \infty$. Therefore, $f_n \in \mathcal{L}^1(\mu)$. Observe that $|f| \le h$, which holds because $|f(x)| = \lim_{n\to\infty} |f_n(x)| \le h(x)$ for all $x \in X$ (since $|f_n| \le h$ for all $n \in \mathbb{N}$). Thus, $\int_X |f| d\mu \le \int_X h d\mu < \infty$, hence $f \in \mathcal{L}^1(\mu)$ as well.
2. It remains to check that $\int_X f_n \, \mathrm{d}\mu \to \int_X f \, \mathrm{d}\mu$. Towards that end, consider $g_n := |f_n - f| \in \mathrm{Bor}^+(X, \mathbb{R})$ for all $n \in \mathbb{N}$. We claim that the g_n's satisfy the hypothesis of LDCT_0.
   (i) $g_n(x) \underset{n\to\infty}{\to} 0$ for all $x \in \mathbb{X}$.
   (ii) There exists $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R}) \cap L^1(\mu)$ such that $g_n \le \tilde{h}$ for all $n \in \mathbb{N}$.

   By the claim we can apply LDCT_0 to $(g_n)_{n=1}^\infty$ to get that

   $$\int_X g_n d\mu \underset{n\to\infty}{\longrightarrow} 0$$

   This implies the fact that

   $$\int_X f_n \, \mathrm{d}\mu \underset{n\to\infty}{\to} \int_X f \, \mathrm{d}\mu$$

   because we can write

$$\begin{aligned}
0 &\leq \left| \int_X f_n \, \mathrm{d}\mu - \int_X f \, \mathrm{d}\mu \right| \\
&\leq \left| \int_X (f_n - f) \, \mathrm{d}\mu \right| \\
&\leq \int_X |f_n - f| \, \mathrm{d}\mu \text{ (by Prop 7.6)} \\
&= \int_X g_n \, \mathrm{d}\mu \underset{n\to\infty}{\to} 0
\end{aligned}$$

By Squeeze Theorem, we are done.

□

8.2 a.e.-μ and the a.e.-μ version of LDCT

Note

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* μ, denoted by f=g a.e.-μ, to mean that there exists a set $Z \in \mathfrak{M}$ such that:

$$\mu(Z) = 0 \quad \text{and} \quad f(x) = g(x) \ \text{ for all } \ x \in X \setminus Z.$$

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

Proposition 8.9

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

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

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

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

Lemma 8.10

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

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

and in particular we can be sure that $f \in \mathcal{L}^1(\mu)$.

Proof: For every $n \in \mathbb{N}$, let $f_n = f \wedge n$, where $n : X \to \mathbb{R}$ is constantly 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 that $f_n \in \mathrm{Bor}(X, \mathbb{R})$ because $\mathrm{Bor}(X, \mathbb{R})$ is stable under $\wedge$. It is trivial to check that $f_n \nearrow f$ (fill in proof). By MCT,

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

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 all $n \in \mathbb{N}$. But now observe: $f_n \leq n\chi_Z$ for all $n \in \mathbb{N}$, where $Z = \{x \in X \mid f(x) \neq 0\}$. This is true because for $x \in Z$, use that $f_n(x) \leq n = n\chi_Z(x)$, and for $x \in X \setminus Z = \{x \in X \mid f(x) = 0\}$, we have $f_n(x) = 0 = n\chi_Z(x)$. So then indeed, we have

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

Hence, $\int_X f_n \, \mathrm{d}\mu = 0$, as desired. We further have that $\int_X f \, \mathrm{d}\mu = L^+(|f|) = L^+(f) < \infty$, so $f \in \mathcal{L}^1(\mu)$. □

Proposition 8.11

Let f, g be two functions in $\mathrm{Bor}(X, \mathbb{R})$ such that $f = g$ a.e.-μ. Suppose that one of these functions, say f, is in $\mathcal{L}^1(\mu)$. Then it follows that the other function is in $\mathcal{L}^1(\mu)$ as well, and we have the equality of integrals

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

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

$$|f - g| \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu), \quad \int_X |f - g| \, \mathrm{d}\mu = 0.$$

But then $f - g$ must be 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, because $\mathcal{L}^1(\mu)$ is stable under linear combinations.

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

$$\left|\int_X f \,\mathrm{d}\mu - \int_X g \,\mathrm{d}\mu\right| = \left|\int_X (f - g) \,\mathrm{d}\mu\right| \leq \int_X |f - g| \,\mathrm{d}\mu = 0.$$

Hence

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

as required. □

Remark 8.12

The converse of Lemma 8.10 is true as well: if $f \in \mathrm{Bor}^+(X, \mathbb{R})$ is such that

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

then it follows that $f = 0$ a.e.-μ.

Proof: Let $Z = \{x \in X \mid f(x) \neq 0\} = \{x \in X \mid x > 0\}$ has $\mu(Z) = 0$. For every $n \in \mathbb{N}$, let $Z_n = \left\{x \in X \mid f(x) \geq \frac{1}{n}\right\}$. Then $Z_1 \subseteq Z_2 \subseteq \ldots \subseteq Z_n \subseteq \ldots$ with $\cup_{n=1}^{\infty} Z_n = Z$.

Why is it that $Z_n \subseteq Z_{n+1}$? Because for $x \in X$, $f(x) \geq \frac{1}{n} \Rightarrow f(x) \geq \frac{1}{n+1}$.

Why is it that $\cup_{n=1}^{\infty} Z_n \supseteq Z$? Because if $x \in Z$, then $f(x) > 0$, hence there exists $n \in N$ such that $f(x) \geq \frac{1}{n}$.

By continuity of μ along increasing chains, we have

$$\mu(Z) = \lim_{n \to \infty} \mu(Z_n)$$

In order to prove that $\mu(Z) = 0$, it suffices to prove that $\mu(Z_n) = 0$ for all $n \in \mathbb{N}$. And indeed, for every $n \in \mathbb{N}$, we observe $f \geq \frac{1}{n}\chi_{Z_n}$. Check: $x \in Z_n \Longrightarrow \frac{1}{n}\chi_{Z_n}(x) = \frac{1}{n} \leq f(x)$ and $x \in X \setminus Z_n \Rightarrow \frac{1}{n}\chi_{Z_n}(x) = 0 \leq f(x)$. Hence,

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

This implies that $\mu(Z_n) = 0$. Therefore, $\mu(Z) = 0$. Therefore, $f = 0$ a.e-μ. □

Theorem 8.13 a.e.-μ **version of the LDCT**

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

1. 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 $x \in X \setminus Z_1$;
2. there exist 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 $Z = Z_1 \cup Z_2$, where $Z \in \mathfrak{M}$, $\mu(Z) = 0$. Note that the hypotheses 1 and 2 still hold in connection to Z.

Idea: Let $\tilde{f} = f\chi_{X\setminus Z}$ and for every $n \in \mathbb{N}$, let $\tilde{f}_n = f_n \chi_{X\setminus Z}$. Observe that $\tilde{f}$ and $\left(\tilde{f}_n\right)_{n=1}^{\infty}$ satisfy the hypotheses of Theorem 8.1 (LDCT). Hence, Theorem 8.1 can be applied to conclude that $\tilde{f}$ and $\tilde{f}_1, \tilde{f}_2, \ldots, \tilde{f}_n$ are in $\mathcal{L}^1(\mu)$ and
$$\int_X \tilde{f}_n \underset{n\to\infty}{\to} \int_X \tilde{f} \,\mathrm{d}\mu$$

On the other hand, observe that $\tilde{f} = f$ a.e-μ and $\tilde{f}_n = f_n$ a.e-μ for all $n \in \mathbb{N}$ (this is because $\mu(Z) = 0$). Then Prop 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 $\int_X f_n \,\mathrm{d}\mu = \int_X \tilde{f}_n \,\mathrm{d}\mu$. Finally,
$$\int_X \tilde{f}_n \,\mathrm{d}\mu \underset{n\to\infty}{\longrightarrow} \int_X \tilde{f} \,\mathrm{d}\mu$$

becomes
$$\int_X f_n \,\mathrm{d}\mu \underset{n\to\infty}{\to} \int_X f \,\mathrm{d}\mu$$

$\square$

9 $\mathcal{L}^1(\mu)$ spaces

9.1 Basic stuff about $\mathcal{L}^1(\mu)$ spaces

Notation 9.1

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 $\|\cdot\|_1$, where the dot is a placeholder for a function $f \in \mathcal{L}^1(\mu)$.

Claim : $\|\cdot\|_1$ is a seminorm on $\mathcal{L}^1(\mu)$. That is:

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

Proof: Non-negativity and homogeneity are immediate from the definition of $\|\cdot\|_1$ and the corresponding properties of the integral.

For subadditivity, let $f, g \in \mathcal{L}^1(\mu)$. Since

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

as functions in $\mathrm{Bor}^+(X, \mathbb{R})$, monotonicity and linearity of the integral give

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

□

Remark 9.2

The only norm axiom that can fail for $\|\cdot\|_1$ is positive definiteness. To be a norm, we would need

$$\|f\|_1 = 0 \Longrightarrow f = 0$$

as functions on X, where $0 : X \to \mathbb{R}$ denotes the identically zero function.

However, Lemma 8.10 and Remark 8.12 give only the weaker statement

$$\|f\|_1 = 0 \Longleftrightarrow f = 0 \quad \text{a.e.- } \mu.$$

This does not imply pointwise equality to 0. For example, if there is a nonempty measurable set A with $\mu(A) = 0$, then $\chi_A \neq 0$ as a function, but

$$\|\chi_A\|_1 = \int_X \chi_A \, \mathrm{d}\mu = \mu(A) = 0.$$

Thus $\|\cdot\|_1$ is generally only a seminorm on $\mathcal{L}^1(\mu)$.

Note

We do not use equivalence classes to get a normed space and instead use semi-norms.

Definition 9.3

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

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

where the integral is considered in the L^+ sense. For $f \in \mathcal{L}^p(\mu)$, we put

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

Note

1. Observe that $\|\alpha f\|_p = |\alpha| \|f\|_p$ for all $\alpha \in \mathbb{R}$, $f \in \mathcal{L}^1(\mu)$. This is why we define $\|f\|_p$ by exponentiating the integral with $\frac{1}{p}$ (to guarantee homogeneity). Indeed, we have that

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

2. In the special case $p = 1$, this recovers the definition of $\mathcal{L}^1(\mu)$ and the seminorm $\|\cdot\|_1$.

Proposition 9.4

Let $p \in [1, \infty)$. One has that $\mathcal{L}^p(\mu)$ is a linear subspace of $\operatorname{Bor}(X, \mathbb{R})$.

Proof: We need to check that $\mathcal{L}^p(\mu)$ is stable under addition and scalar multiplication. It is non-empty since it contains the zero function.

Stability under addition. First observe that for every $a, b \in \mathbb{R}$,

$$|a + b|^p \leq (2 \max\{|a|, |b|\})^p \leq 2^p (|a|^p + |b|^p).$$

Hence, for $f, g \in \operatorname{Bor}(X, \mathbb{R})$,

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

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

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

If $f, g \in \mathcal{L}^p(\mu)$, then the right-hand side is finite. Therefore

$$\int_X |f + g|^p \, \mathrm{d}\mu < \infty,$$

so $f + g \in \mathcal{L}^p(\mu)$.

Stability under scalar multiplication. Let $\alpha \in \mathbb{R}$ and $f \in \mathcal{L}^p(\mu)$. Then $\alpha f \in \operatorname{Bor}(X, \mathbb{R})$, and

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

Hence $\alpha f \in \mathcal{L}^p(\mu)$. $\square$

Remark 9.5

1. Pick $p \in (1, \infty)$ and consider the function $\|\cdot\|_p$ on the vector space $\mathcal{L}^p(\mu)$. Is $\|\cdot\|_p$ a semi-norm (similar to how $\|\cdot\|_1$ was found to be a semi-norm on $\mathcal{L}^1(\mu)$)? It is easy to see that $\|\cdot\|_p$ is non-negative and homogeneous, but sub-additivity is not obvious. So, we need
$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$
for all $f, g \in \mathcal{L}^p(\mu)$. This is called *Minkowski's Inequality.*
2. We also record the analogue of the positive-definiteness issue from $\mathcal{L}^1(\mu)$. For $f \in \mathcal{L}^p(\mu)$,
$$\|f\|_p = 0 \iff f = 0 \;\text{ a.e.- } \mu.$$
Indeed,
$$\|f\|_p = 0 \iff \int_X |f|^p \,\mathrm{d}\mu = 0 \iff |f|^p = 0 \;\text{ a.e.- } \mu \iff f = 0 \;\text{ a.e.- } \mu,$$
where the middle equivalence uses Lemma 8.10 and Remark 8.12. Thus, just as for $\|\cdot\|_1$, the map $\|\cdot\|_p$ is generally not a norm on $\mathcal{L}^p(\mu)$ unless we identify functions that agree a.e.

Note

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 $\mathcal{L}^q(\mu)$.

Definition 9.6 **Conjugate Exponents**

Two numbers $p, q \in (1, \infty)$ are said to be *conjugate exponents* when they satisfy
$$\frac{1}{p} + \frac{1}{q} = 1.$$
Thus, given $p \in (1, \infty)$, there is precisely one conjugate exponent for it, namely
$$q = \frac{p}{p-1}.$$

Note

The point is that even if we want to prove a result only about $\mathcal{L}^p(\mu)$, such as Minkowski's inequality, it may still be useful to bring in the sibling space $\mathcal{L}^q(\mu)$ associated to the conjugate exponent.

Proposition 9.7 **Holder's Inequality**

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

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q \text{ (Holder-1)}$$

Consequently,

$$\left| \int_X f \cdot g \, \mathrm{d}\mu \right| \leq \|f\|_p \cdot \|g\|_q. \text{ (Holder-2)}$$

Note

Holder's Inequality is a generalization of Cauchy-Schwarz.

It also introduces a *duality* between $\mathcal{L}^p(\mu)$ and $\mathcal{L}^q(\mu)$. To every function $f \in \mathcal{L}^p(\mu)$, we associate a linear functional $\varphi_f : \mathcal{L}^q(\mu) \to \mathbb{R}$,

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

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

Note

$p = 1, 2, \infty$ are the most important $\mathcal{L}^p(\mu)$ spaces.

Remark 9.8

Before proving Holder's inequality, let us first look at the important special case $p = q = 2$. In this case, Holder's inequality says that

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

This lets us introduce an inner-product notation by putting

$$\langle f, g \rangle := \int_X f \cdot g \, \mathrm{d}\mu \in \mathbb{R}$$

for $f, g \in \mathcal{L}^2(\mu)$.

The corresponding form of Holder's inequality is the Cauchy-Schwarz inequality for $\mathcal{L}^2$-functions:

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

We now prove Holder's inequality. The first step is to record the elementary inequality used in the proof.

Lemma 9.9 **Young's Inequality**

Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. We have $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ for all $a, b \in [0, \infty)$.

Proof: Based on the shape of the graph of the exponential function (since it is concave up).
□

Lemma 9.10

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, and let $f \in \mathcal{L}^p(\mu)$ 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 $\|f\|_p = 1$ implies that $\left(\int_X |f|^p \, \mathrm{d}\mu\right)^{\frac{1}{p}} = 1$, which implies that $\int_X |f|^p \, \mathrm{d}\mu = 1$. Likewise, $\int_X |f|^q \, \mathrm{d}\mu = 1$. Applying lemma 9.9 to $a = |f(x)|$ and $b = |g(x)|$ with $x \in X$, we obtain

$$|f(x)||g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}$$

for all $x \in X$. Hence, $|f \cdot g| \leq \frac{1}{p}|f|^p + \frac{1}{q}|g|^q$, so

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

as required. □

Proof of Holder 1: We have $(p, q) \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $f \in \mathcal{L}^p(\mu), g \in \mathcal{L}^q(\mu)$. We want $f \cdot g \in \mathcal{L}^1(\mu)$ and $\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q$.

Denote $\|f\|_p = \alpha \in [0, \infty)$ and $\|g\|_q = \beta \in [0, \infty)$.

1. Case 1: $\alpha = 0$. Then $\|f\|_p = 0$ implies that $f = 0$ a.e-μ. Hence $f \cdot g = 0$ a.e-μ, so $\|f \cdot g\|_1 = 0 = \alpha \cdot \beta = \|f\|_p \|g\|_q$.
2. Case 2: $\alpha \neq 0 \neq \beta$. Define $\tilde{f} = \frac{1}{\alpha}f, \tilde{g} = \frac{1}{\beta}g$. We have $\tilde{f} \in \mathcal{L}^p(\mu)$ with

$$\left\|\tilde{f}\right\|_p = \left\|\frac{1}{\alpha}f\right\|_p = \frac{1}{\alpha}\|f\|_p = \frac{1}{\alpha}\alpha = 1$$

Likewise, $\tilde{g} \in \mathcal{L}^q(\mu)$ with $\|\tilde{g}\|_q = 1$. Applying Lemma 9.10, we get that $\tilde{f} \cdot \tilde{g} \in \mathcal{L}^1(\mu)$ with $\left\|\tilde{f} \cdot \tilde{g}\right\|_1 \leq 1$. But then $f = \alpha\tilde{f}, g = \beta\tilde{g}$. So $f \cdot g = (\alpha\beta) \cdot \left(\tilde{f} \cdot \tilde{g}\right)$. Therefore, $f \cdot g \in \mathcal{L}^1(\mu)$ with $\|f \cdot g\|_1 = \left\|(\alpha\beta) \cdot \tilde{f} \cdot \tilde{g}\right\|_1 = \alpha\beta\left\|\tilde{f} \cdot \tilde{g}\right\|_1 \leq \alpha\beta = \|f\|_p \cdot \|g\|_q$.

□

Notation 9.11

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

1. Define $\varphi_f : \mathcal{L}^q(\mu) \to \mathbb{R}$ by

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

   This definition makes sense by Holder's inequality, since $f \cdot g \in \mathcal{L}^1(\mu)$.
2. The map φ_f is linear, and Holder's inequality (Holder-2) gives

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

3. Hence the set

$$\left\{\left|\varphi_f(g)\right| \mid g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1\right\}$$

   is bounded above by $\|f\|_p$. We may therefore define

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

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

Lemma 9.12

Let f be a function in $\mathcal{L}^p(\mu)$, and consider the number $\lambda(f) \in [0, \infty)$ defined above. Then

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

Proof: $\lambda(f) \leq \|f\|_p$ was seen in Remark 9.11. For $\lambda(f) \geq \|f\|_p$, we will define a function $g_0 \in \mathcal{L}^q(\mu)$ such that

$$\|g_0\|_q \leq 1 \text{ and } \varphi_f(g_0) = \|f\|_p. \quad \blacklozenge$$

Using this g_0, we will infer that

$$\begin{aligned}\lambda(f) &\geq \left|\varphi_f(g_0)\right| \\ &= \|f\|_p\end{aligned}$$

which will give $\lambda(f) \geq \|f\|_p$.

Question: How do we find g_0? Think of the case $p = q = 2$. There, we have $\varphi_f(g) = \int_X f \cdot g \,\mathrm{d}\mu = \langle f, g\rangle$ for $f, g \in \mathcal{L}^2(\mu)$. Given $f \in \mathcal{L}^2(\mu)$, want to find $g_0 \in \mathcal{L}^2(\mu)$ with $\|g_0\|_2 = 1$ and $|\langle f, g_0\rangle| = \|f\|_2 \cdot \|g_0\|_2$ (equality in Cauchy-Schwarz). So try $g_0 = \alpha f$ for suitable $\alpha \in \mathbb{R}$. For general $p \in (1, \infty)$, we have $q \neq p$. So we cannot do $g_0 = \alpha f$, but observe that $h = |f|^{\frac{p}{q}}$ is in $\mathcal{L}^q(\mu)$ because $h^q = |f|^p$ hence $\int_X h^q \,\mathrm{d}\mu = \int_X |f|^p \,\mathrm{d}\mu < \infty$.

Claim 1: One can find $u \in \mathrm{Bor}(X, \mathbb{R})$ such that $|u(x)| = 1$ for all $x \in X$ and such that $f = u \cdot |f|$.

Verification of Claim 1: Let $u : X \to \mathbb{R}$ be defined by

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

This does the job (please fill in why: in particular, explain $u \in \mathrm{Bor}(X, \mathbb{R})$).

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

$$(*) \quad \|h\|_q = \|f\|_p^{p-1}$$

and

$$(**) \quad \int_X fh \,\mathrm{d}\mu = \|f\|_p^p$$

Verification of Claim 2: Straightforward calculations. For $(*)$, we have

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

For $(**)$, we have

$$\begin{aligned} fh &= (u|f|)\left(u \cdot |f|^{\frac{p}{q}}\right) \\ &= u^2|f|^{1+\frac{p}{q}} = |f|^p \end{aligned}$$

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

Conclusion of the proof: recall that we need $g_0 \in \mathcal{L}^q(\mu)$ with $\|g_0\|_q \leq 1$ and

$$\int_X f g_0 \, \mathrm{d}\mu = \|f\|_p$$

If $\|f\|_p = 0$, then we can use $g_0 = 0$. So suppose $\|f\|_p \neq 0$. Let h be as in Claim 2, and put $g_0 = \frac{1}{\|f\|_p^{p-1}} h \in \mathcal{L}^q(\mu)$. Then g_0 will work in ◆. Claim 2 gives $\|g_0\|_q = 1$, and Claim 2 gives $\int_X f g_0 \, \mathrm{d}\mu = \|f\|_p$. □

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: Due to Lemma 9.12, this proposition is equivalent to proving

$$\lambda(f_1 + f_2) \leq \lambda(f_1) + \lambda(f_2) \quad ◆$$

Since

$$\begin{aligned} \lambda(f_1 + f_2) &= \sup\left\{\left|\varphi_{f_1+f_2}(g)\right| g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1\right\} \\ &= \sup\left\{\left|\int_X (f_1 + f_2) g \, \mathrm{d}\mu\right| \mid g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1\right\} \end{aligned}$$

It suffices to check that

$$\left|\int_X (f_1 + f_2) g \, \mathrm{d}\mu\right| \leq \lambda(f_1) + \lambda(f_2)$$

for every $g \in \mathcal{L}^q(\mu)$ with $\|g\|_q \leq 1$. And indeed, for such a g, we write

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

Taking the supremum over all $g \in \mathcal{L}^q(\mu)$ with $\|g\|_q \leq 1$ gives $(*)$, and therefore

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

□

Corollary 9.14

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

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

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

In this part of the lecture, we fix a couple $(\mathcal{V}, \|\cdot\|)$ where $\mathcal{V}$ is a vector space over $\mathbb{R}$, and where the map $\|\cdot\| : \mathcal{V} \to [0, \infty)$ is a *seminorm*. A possible name for the couple $(\mathcal{V}, \|\cdot\|)$ is thus a *semi-normed vector space*.

We have to keep in mind that $\|\cdot\|$ is not assumed to be a norm, so the null-space

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

may include non-zero vectors. But even without assuming that $\|\cdot\|$ is a norm, we can still define (with some care) the notions of convergent sequence and of Cauchy sequence in $\mathcal{V}$.

Definition 10.1

Let $(v_n)_{n=1}^{\infty}$ be a sequence of vectors in $\mathcal{V}$, and let v be a vector in $\mathcal{V}$. We say that $(v_n)_{n=1}^{\infty}$ *converges* to v to mean that one has

$$\lim_{n\to\infty} \|v_n - v\| = 0$$

Remark 10.2

In the framework we use here, the limit of a convergent sequence is not uniquely determined. Indeed, suppose that $(v_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{V}$ for which we proved that $(v_n)_{n=1}^{\infty}$ converges to v, and someone else has proved that $(v_n)_{n=1}^{\infty}$ converges to v'. The familiar trick of writing

$$\|v - v'\| \leq \|v - v_n\| + \|v_n - v'\|, \quad \forall n \in \mathbb{N}$$

and then making $n \to \infty$ will lead us to the conclusion that $\|v - v'\| = 0$. But this will not necessarily imply that $v = v'$.

Definition 10.3

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

1. We say that $(v_n)_{n=1}^{\infty}$ is a *Cauchy sequence* to mean that it satisfies the following condition:

$$\text{(Cauchy)} \quad \begin{cases} \text{For every } \varepsilon > 0 \text{ one can find } n_o \in \mathbb{N} \\ \text{such that: whenever } m, n \geq n_o, \text{it follows that } \|v_m - v_n\| < \varepsilon \end{cases}$$

2. We say that $(v_n)_{n=1}^{\infty}$ is a $\frac{1}{2}$*-geometric sequence* to mean that one has

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

Remark 10.4

The familiar arguments you know from studying normed vector spaces in PMath 351 work to prove the following implications about a sequence $(v_n)_{n=1}^{\infty}$ of vectors in $\mathcal{V}$.

1. If $(v_n)_{n=1}^{\infty}$ is convergent to some $v \in \mathcal{V}$, 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 it is possible to select indices $1 \leq n(1) < n(2) < \ldots < n(k) < \ldots$ such that the subsequence $\left(v_{n(k)}\right)_{k=1}^{\infty}$ is $\frac{1}{2}$-geometric.

Definition 10.5

A semi-normed space $(\mathcal{V}, \|\cdot\|)$ is said to be *complete* when for every Cauchy sequence $(v_n)_{n=1}^{\infty}$ in $\mathcal{V}$, there exists $v \in \mathcal{V}$ (not necessarily unique) such that $(v_n)_{n=1}^{\infty}$ converges to v.

Remark 10.6 **Completeness Criterion**

Fact: Let $(\mathcal{V}, \|\cdot\|)$ be a semi-normed space such that every $\frac{1}{2}$-geometric sequence in $\mathcal{V}$ converges to some vector in $\mathcal{V}$. Then $(\mathcal{V}, \|\cdot\|)$ is complete.

Fact$_0$: Let $(\mathcal{V}, \|\cdot\|)$ be a semi-normed space for which we know that whenever $(w_n)_{n=1}^{\infty}$ is a $\frac{1}{2}$-geometric sequence in $\mathcal{V}$ with $w_1 = 0_{\mathcal{V}}$, there exists $w \in \mathcal{V}$ such that $(w_n)_{n=1}^{\infty}$ converges to w. Then $(\mathcal{V}, \|\cdot\|)$ is complete.

Proof of Fact: Let $(v_n)_{n=1}^{\infty}$ be a Cauchy sequence in $\mathcal{V}$. By the preceding remark, choose indices

$$1 \leq n(1) < n(2) < \ldots < n(k) < \ldots$$

such that $\left(v_{n(k)}\right)_{k=1}^{\infty}$ is $\frac{1}{2}$-geometric. By hypothesis, there is $v \in \mathcal{V}$ such that

$$\lim_{k\to\infty}\left\|v_{n(k)} - v\right\| = 0.$$

We claim that the whole sequence converges to v. Fix $\varepsilon > 0$. Since $(v_n)_{n=1}^{\infty}$ is Cauchy, choose $N \in \mathbb{N}$ such that

$$\|v_m - v_n\| < \frac{\varepsilon}{2}$$

whenever $m, n \geq N$. Since $v_{n(k)} \to v$, choose $k \in \mathbb{N}$ with $n(k) \geq N$ and

$$\left\|v_{n(k)} - v\right\| < \frac{\varepsilon}{2}.$$

Then, for every $n \geq N$,

$$\|v_n - v\| \leq \left\|v_n - v_{n(k)}\right\| + \left\|v_{n(k)} - v\right\| < \varepsilon.$$

Hence $v_n \to v$, so $(\mathcal{V}, \|\cdot\|)$ is complete. □

***Proof of* Fact$_0$:** It is enough to check the hypothesis of Fact above. Let $(v_n)_{n=1}^{\infty}$ be any $\frac{1}{2}$-geometric sequence in $\mathcal{V}$, and set $w_n := v_n - v_1$. Then $w_1 = 0_{\mathcal{V}}$ and

$$\|w_n - w_{n+1}\| = \|v_n - v_{n+1}\| \leq \frac{1}{2^n}.$$

By the hypothesis of Fact$_0$, there exists $w \in \mathcal{V}$ such that $w_n \to w$. Hence

$$\|v_n - (w + v_1)\| = \|w_n - w\| \to 0.$$

Thus every $\frac{1}{2}$-geometric sequence in $\mathcal{V}$ converges. By Fact, $(\mathcal{V}, \|\cdot\|)$ is complete. □

10.2 Completeness of $\left(\mathcal{L}^p(\mu), \|\cdot\|_p\right)$

Lemma 10.7 Convergence mechanism – how to find a candidate for limit

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

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

and, in general,

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

We observe that $(h_n)_{n=1}^{\infty}$ is increasing and we denote

$$Z := \left\{x \in X \mid \lim_{n\to\infty} h_n(x) = \infty\right\}$$

Then, for every $x \in X \setminus Z$, the sequence $(f_n(x))_{n=1}^{\infty}$ is convergent in $\mathbb{R}$.

Proof: $x \in X \setminus Z$. Since $x \notin Z$ and $(h_n(x))_{n=1}^{\infty}$ is increasing, then there exists $\ell \in [0, \infty)$ such that $h_n(x) \to \ell$. For $m < n$ in $\mathbb{N}$, the telescoping triangle inequality gives

$$|f_m(x) - f_n(x)| \leq \sum_{k=m+1}^{n} |f_k(x) - f_{k-1}(x)| = h_n(x) - h_m(x).$$

Since $(h_n(x))_{n=1}^{\infty}$ converges, it is Cauchy. Hence, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n > m \geq N$, then $h_n(x) - h_m(x) < \varepsilon$. The displayed inequality gives $|f_m(x) - f_n(x)| < \varepsilon$, so $(f_n(x))_{n=1}^{\infty}$ is Cauchy. Since $\mathbb{R}$ is complete, $(f_n(x))_{n=1}^{\infty}$ converges. $\square$

Note **Addendum to MCT**

Let $(X, \mathfrak{M}, \mu)$ be a measure space. Let $u_1 \leq u_2 \leq \ldots \leq u_n \leq \ldots$ be functions in $\mathrm{Bor}^+(X, \mathbb{R})$, and suppose there exists $Z \in \mathfrak{M}$ with $\mu(Z) > 0$ such that

$$\lim_{n\to\infty} u_n(x) = \infty \quad \text{for every } x \in Z.$$

Then

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

Proposition 10.8

Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $p \in [1, \infty)$, and 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 Lemma 10.7, we get a function $f : X \to \mathbb{R}$, as described there.

Then $f \in \mathcal{L}^p(\mu)$, and

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Proof: We use the help functions $(h_n)_{n=1}^{\infty}$ and the set Z defined in Lemma 10.7. Since $f_1 = \underline{0}$, we have

$$h_1 = 0, \quad h_n = \sum_{m=2}^{n} |f_m - f_{m-1}| \quad (n \geq 2).$$

In addition, define $h : X \to [0, \infty)$ by

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

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

Proof of Claim 1: The case $n = 1$ is clear. Now suppose $n \geq 2$. For $2 \leq m \leq n$, put $u_m := |f_m - f_{m-1}|$. Since $\mathcal{L}^p(\mu)$ is stable under linear combinations and lattice operations, $u_m \in \mathcal{L}^p(\mu)$. Also,

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

Since $h_n = u_2 + \ldots + u_n$, Minkowski's inequality gives

$$\|h_n\|_p \leq \|u_2\|_p + \ldots + \|u_n\|_p \leq \frac{1}{2} + \ldots + \frac{1}{2^{n-1}} \leq 1.$$

Hence $h_n \in \mathcal{L}^p(\mu)$ and $\|h_n\|_p \leq 1$.

Claim 2:
1. Let $Z = \{x \in X \mid \lim_{n\to\infty} h_n(x) = \infty\}$. Then $Z \in \mathfrak{M}$ and $\mu(Z) = 0$.
2. Moreover, let $h : X \to [0, \infty)$ be defined by

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

   Then $h \in \mathcal{L}^p(\mu)$ with $\|h\|_p \leq 1$.

Proof of Claim 2: We consider the increasing sequence

$$0 = h_1^p \leq h_2^p \leq \ldots \leq h_n^p \leq \ldots$$

in $\mathrm{Bor}^+(X, \mathbb{R})$. By Claim 1,

$$\int_X h_n^p \,\mathrm{d}\mu = \|h_n\|_p^p \leq 1, \quad \forall n \in \mathbb{N},$$

and hence

$$\lim_{n\to\infty} \int_X h_n^p \,\mathrm{d}\mu \leq 1 < \infty.$$

The addendum to MCT implies that

$$\left\{x \in X \mid \lim_{n\to\infty} h_n^p(x) = \infty\right\} \in \mathfrak{M}$$

and has measure zero. This set is precisely Z, since $\lim_{n\to\infty} h_n^p(x) = \infty$ if and only if $\lim_{n\to\infty} h_n(x) = \infty$. Thus $Z \in \mathfrak{M}$ and $\mu(Z) = 0$.

Moreover, applying the same addendum to the present situation gives

$$\lim_{n\to\infty} \int_X h_n^p \,\mathrm{d}\mu = \int_X h^p \,\mathrm{d}\mu.$$

Combining this with the preceding bound gives

$$\int_X h^p \,\mathrm{d}\mu \leq 1.$$

Hence $h \in \mathcal{L}^p(\mu)$ and $\|h\|_p \leq 1$.

Claim 3: $f \in \mathrm{Bor}(X, \mathbb{R})$.

Proof of Claim 3: The function f can be written as the pointwise limit of the sequence $\left(f_n \chi_{X \setminus Z}\right)_{n=1}^{\infty}$. Each function in this sequence is in $\mathrm{Bor}(X, \mathbb{R})$, and $\mathrm{Bor}(X, \mathbb{R})$ is closed under pointwise limits. Hence $f \in \mathrm{Bor}(X, \mathbb{R})$.

Claim 4: Fix an $n_0 \in \mathbb{N}$ and consider $g := f_{n_0} \chi_{X \setminus Z} - f \in \mathrm{Bor}(X, \mathbb{R})$. Then $g \in \mathcal{L}^p(\mu)$ with $\|g\|_p \leq \frac{1}{2^{n_0 - 1}}$.

Proof of Claim 4: For every $n \in \mathbb{N}$, let $g_n = f_{n_0} \chi_{X \setminus Z} - f_n \chi_{X \setminus Z}$. Then $g(x) = \lim_{n \to \infty} g_n(x)$ for all $x \in X$. Observe that for $n > n_0$,

$$\begin{aligned}
\|g_n\|_p &= \left\| \left(f_{n_0} - f_n\right) \chi_{X \setminus Z} \right\|_p \\
&\leq \left\| f_{n_0} - f_n \right\|_p \\
&= \left\| \sum_{k=n_0}^{n-1} (f_k - f_{k+1}) \right\|_p \\
&\leq \sum_{k=n_0}^{n-1} \left\| f_k - f_{k+1} \right\|_p \\
&\leq \sum_{k=n_0}^{n-1} \frac{1}{2^k} \\
&\leq \sum_{k=n_0}^{\infty} \frac{1}{2^k} \\
&= \frac{1}{2^{n_0 - 1}}.
\end{aligned}$$

Also, $g_{n(x)} \to g(x)$ for every $x \in X$, and $|g_n| \leq h$ for all $n > n_0$. By LDCT applied to $|g_n|^p$, we get

$$\int_X |g|^p \,\mathrm{d}\mu = \lim_{n \to \infty} \int_X |g_n|^p \,\mathrm{d}\mu \leq \left(\frac{1}{2^{n_0 - 1}}\right)^p.$$

Therefore $g \in \mathcal{L}^p(\mu)$ and $\|g\|_p \leq \frac{1}{2^{n_0 - 1}}$.

Claim 5: $f \in \mathcal{L}^p(\mu)$.

Proof of Claim 5: Fix $n_0 \in \mathbb{N}$. We can write

$$f = f_{n_0} \chi_{X \setminus Z} - \left(f_{n_0} \chi_{X \setminus Z} - f\right).$$

The first term belongs to $\mathcal{L}^p(\mu)$ because $f_{n_0} \in \mathcal{L}^p(\mu)$ and multiplication by $\chi_{X\setminus Z}$ cannot increase the p-seminorm. The second term belongs to $\mathcal{L}^p(\mu)$ by Claim 4. Since $\mathcal{L}^p(\mu)$ is stable under linear combinations, $f \in \mathcal{L}^p(\mu)$.

Claim 6: We have $\lim_{n\to\infty} \|f_n - f\|_p = 0$.

Proof of Claim 6: Apply Claim 4 with $n_0 = n$. Since $\mu(Z) = 0$, the functions $f_n - f$ and $f_n\chi_{X\setminus Z} - f$ agree a.e.-μ. Therefore

$$\|f_n - f\|_p = \left\|f_n\chi_{X\setminus Z} - f\right\|_p \leq \frac{1}{2^{n-1}}.$$

Letting $n \to \infty$ gives $\|f_n - f\|_p \to 0$. □

Corollary 10.9

$\left(\mathcal{L}^p(\mu), \|\cdot\|_p\right)$ is complete.

Corollary 10.10

$\left(L^p(\mu), \|\cdot\|_p\right)$ is complete (hence a Banach space).

Definition 10.11

Let $(X, \mathfrak{M}, \mu)$ be a measure space, and let f and $(f_n)_{n=1}^\infty$ be functions in $\operatorname{Bor}(X, \mathbb{R})$. We say that $(f_n)_{n=1}^\infty$ *converges to f almost everywhere with respect to μ* (or a.e.-μ, for short) to mean that there exists a set $Z \in \mathfrak{M}$ with $\mu(Z) = 0$ and such that

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{for every } x \in X \setminus Z.$$

Note

In the proof of Prop 10.8, at the same time with $\|f_n - f\|_p \underset{n\to\infty}{\to} 0$, we also obtained convergence μ-almost everywhere.

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 of functions in $\mathcal{L}^p(\mu)$ such that

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

Then there exists a function $f \in \mathcal{L}^p(\mu)$ such that

$$\lim_{n\to\infty} f_n = f \quad \text{a.e.-}\mu.$$

11 Bounded Linear Functionals on $\mathcal{L}^2(\mu)$.

For this lecture, we fix a measure space $(X, \mathfrak{M}, \mu)$.

Definition 11.1

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a linear functional. We say that φ is *bounded* to mean that there exists a constant $c \geq 0$ such that

$$\text{(Bdd)} \quad |\varphi(f)| \leq c \cdot \|f\|_2, \quad \forall f \in \mathcal{L}^2(\mu).$$

Note

The above definition is equivalent to saying that

$$\sup\{|\varphi(f)| \mid f \in \mathcal{L}^2(\mu), \|f\|_2 \leq 1\}$$

is finite. In that case, we define

$$\|\varphi\| := \sup\{|\varphi(f)| \mid f \in \mathcal{L}^2(\mu), \|f\|_2 \leq 1\}.$$

This is the operator norm of φ.

Remark 11.2

If $f, g \in \mathcal{L}^2(\mu)$, then $f \cdot g \in \mathcal{L}^1(\mu)$ by Holder's Inequality. So, it makes sense to define $\varphi_f : \mathcal{L}^2(\mu) \to \mathbb{R}$ by

$$g \mapsto \langle f, g \rangle$$

where $\langle f, g \rangle := \int_X f \cdot g \, d\mu$. Note that φ_f is a linear functional since integration is linear and φ_f is bounded because for any $g \in \mathcal{L}^2(\mu)$, $\int_X f \cdot g \, d\mu \leq c\|g\|_2$, where $c := \|f\|_2$.

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 an $f_o \in \mathcal{L}^2(\mu)$ such that

$$\varphi(g) = \langle f_o, g \rangle, \quad \forall g \in \mathcal{L}^2(\mu).$$

Lemma 11.4 **Parallelogram Law**

For every $f, g \in \mathcal{L}^2(\mu)$ we have

$$\text{(Parllg-Law)} \quad \|f+g\|_2^2 + \|f-g\|_2^2 = 2(\|f\|_2^2 + \|g\|_2^2).$$

Proof: Regard f and g as vectors in the (almost) Hilbert space $\mathcal{L}^2(\mu)$ and consider the subspace

$$\operatorname{span}\{f, g\} := \{af + bg \mid a, b \in \mathbb{R}\}.$$

This subspace contains f and g and is generated by two vectors, so its dimension is at most 2. Thus all four vertices $0, f, g, f+g$ lie in a Euclidean plane (or in a line in the degenerate case).

These vertices form a parallelogram whose four side lengths are $\|f\|_2, \|g\|_2, \|f\|_2, \|g\|_2$. Its diagonals have vectors $f+g$ and $f-g$, and hence lengths $\|f+g\|_2$ and $\|f-g\|_2$.

If either f or g is zero, the result is immediate. Otherwise, let θ be the angle between them. Applying the cosine law to the two diagonals gives

$$\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2 \cos\theta$$

and

$$\|f-g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 - 2\|f\|_2\|g\|_2 \cos\theta.$$

Adding these identities cancels the angle-dependent terms and yields

$$\|f+g\|_2^2 + \|f-g\|_2^2 = 2\|f\|_2^2 + 2\|g\|_2^2,$$

as required. □

Lemma 11.5

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a bounded linear functional which is not identically equal to 0. Consider the set

$$\mathcal{U} := \{u \in \mathcal{L}^2(\mu) \mid \varphi(u) = 1\}$$

Then:

1. The set $\mathcal{U}$ is not empty.
2. Denoting

$$\alpha_o := \inf\{\|u\|_2 \mid u \in \mathcal{U}\},$$

   we have that $\alpha_o > 0$. View α_0 as the distance between $\mathcal{U}$ and 0 in $\mathcal{L}^2(\mu)$.
3. There exists a $u_o \in \mathcal{U}$ such that $\|u_o\|_2 = \alpha_o$.

Proof:

1. Since φ is not identically equal to 0, we can choose $f \in \mathcal{L}^2(\mu)$ such that $\varphi(f) \neq 0$. Put

$$u := \frac{1}{\varphi(f)} \cdot f.$$

   Then

$$\varphi(u) = \frac{1}{\varphi(f)} \cdot \varphi(f) = 1,$$

so $u \in \mathcal{U}$. Thus $\mathcal{U}$ is not empty.

2. Since φ is bounded, there exists $c \geq 0$ such that

$$|\varphi(f)| \leq c \cdot \|f\|_2, \quad \forall f \in \mathcal{L}^2(\mu).$$

In fact $c > 0$, since $c = 0$ would imply $\varphi = 0$. If $u \in \mathcal{U}$, then

$$1 = |\varphi(u)| \leq c \cdot \|u\|_2,$$

hence $\|u\|_2 \geq \frac{1}{c}$. Taking the infimum over $u \in \mathcal{U}$ gives

$$\alpha_o \geq \frac{1}{c} > 0.$$

3. By the definition of α_o, choose a sequence $(u_n)_{n=1}^{\infty}$ in $\mathcal{U}$ such that $\|u_n\|_2 \to \alpha_o$. We claim that (u_n) is Cauchy. For $m, n \in \mathbb{N}$, the midpoint $\frac{u_m + u_n}{2}$ belongs to $\mathcal{U}$, since

$$\varphi\left(\frac{u_m + u_n}{2}\right) = 1.$$

Hence

$$\left\|\frac{u_m + u_n}{2}\right\|_2 \geq \alpha_o,$$

or equivalently $\|u_m + u_n\|_2^2 \geq 4\alpha_o^2$. By the parallelogram law,

$$\begin{aligned} \|u_m - u_n\|_2^2 &= 2\|u_m\|_2^2 + 2\|u_n\|_2^2 - \|u_m + u_n\|_2^2 \\ &\leq 2\|u_m\|_2^2 + 2\|u_n\|_2^2 - 4\alpha_o^2. \end{aligned}$$

The right-hand side tends to 0 as $m, n \to \infty$, so (u_n) is Cauchy in $\mathcal{L}^2(\mu)$. By completeness of $\mathcal{L}^2(\mu)$, there exists $u_o \in \mathcal{L}^2(\mu)$ such that $\|u_n - u_o\|_2 \to 0$.

Since φ is bounded, $\varphi(u_n - u_o) \to 0$, and therefore

$$\varphi(u_o) = \lim_{n \to \infty} \varphi(u_n) = 1.$$

Thus $u_o \in \mathcal{U}$. Finally,

$$\left|\|u_o\|_2 - \|u_n\|_2\right| \leq \|u_o - u_n\|_2,$$

so $\|u_o\|_2 = \lim_{n \to \infty} \|u_n\|_2 = \alpha_o$.

□

Lemma 11.6

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a bounded linear functional which is not identically equal to 0. Let the set

$$\mathcal{U} := \{u \in \mathcal{L}^2(\mu) \mid \varphi(u) = 1\}$$

and the special element $u_o \in \mathcal{U}$ be as in Lemma 11.5. On the other hand, consider the set

$$\mathcal{V} := \{v \in \mathcal{L}^2(\mu) \mid \varphi(v) = 0\}$$

(the kernel of φ). Then $\langle u_o, v\rangle = 0$ for all $v \in \mathcal{V}$.

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 definition of α_o and the choice of u_o,

$$\|u_o + tv\|_2 \geq \alpha_o = \|u_o\|_2.$$

Squaring and expanding gives

$$\alpha_o^2 + 2t\langle u_o, v\rangle + t^2\|v\|_2^2 \geq \alpha_o^2,$$

and hence

$$t^2\|v\|_2^2 + 2t\langle u_o, v\rangle \geq 0, \quad \forall t \in \mathbb{R}.$$

If $v = 0$, the conclusion is immediate. Otherwise, the discriminant of this nonnegative quadratic is nonpositive, so

$$4\langle u_o, v\rangle^2 \leq 0.$$

Therefore $\langle u_o, v\rangle = 0$. □

Proof of Theorem 11.3: If $\varphi = 0$, take $f_o = 0$. Then

$$\varphi(g) = 0 = \langle f_o, g\rangle, \quad \forall g \in \mathcal{L}^2(\mu).$$

Suppose now that φ is not identically zero, and let u_o be the element given by Lemma 11.5. For any $g \in \mathcal{L}^2(\mu)$, put

$$v := g - \varphi(g)u_o.$$

Since $\varphi(u_o) = 1$, we have

$$\varphi(v) = \varphi(g) - \varphi(g)\varphi(u_o) = 0,$$

so $v \in \mathcal{V}$. Lemma 11.6 therefore gives

$$
\begin{aligned}
0 &= \langle u_o, v \rangle \\
&= \langle u_o, g \rangle - \varphi(g) \langle u_o, u_o \rangle \\
&= \langle u_o, g \rangle - \varphi(g) \|u_o\|_2^2.
\end{aligned}
$$

Since $\|u_o\|_2 = \alpha_o > 0$, we may define

$$f_o := \frac{u_o}{\|u_o\|_2^2}.$$

It follows that, for every $g \in \mathcal{L}^2(\mu)$,

$$\varphi(g) = \frac{\langle u_o, g \rangle}{\|u_o\|_2^2} = \langle f_o, g \rangle.$$

This proves the theorem. □

12 Radon-Nikodym Theorem

Fix a measureable space $(X, \mathfrak{M})$ for this lecture.

Notation 12.1

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. We write $\nu \leq \mu$ to mean that

$$\nu(A) \leq \mu(A), \quad \forall A \in \mathfrak{M}.$$

Proposition 12.2

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 friendly functions. Let

$$\mathcal{G} := \left\{ g \in \mathrm{Bor}^+(X, \mathbb{R}) \mid \int_X g \, \mathrm{d}\nu \leq \int_X g \, \mathrm{d}\mu \right\}.$$

We verify the four properties in Lemma 6.9.

(i) If $A \in \mathfrak{M}$, then

$$\int_X \chi_A \, \mathrm{d}\nu = \nu(A) \leq \mu(A) = \int_X \chi_A \, \mathrm{d}\mu \quad (\text{since} \quad \nu \leq \mu \ ).$$

Hence $\chi_A \in \mathcal{G}$.

(ii) If $g_1, g_2 \in \mathcal{G}$, then additivity of the integral gives

$$\begin{aligned}\int_X (g_1 + g_2)\,\mathrm{d}\nu &= \int_X g_1\,\mathrm{d}\nu + \int_X g_2\,\mathrm{d}\nu \\ &\leq \int_X g_1\,\mathrm{d}\mu + \int_X g_2\,\mathrm{d}\mu \quad (\text{since } g_1, g_2 \in \mathcal{G}\ ) \\ &= \int_X (g_1 + g_2)\,\mathrm{d}\mu.\end{aligned}$$

Thus $g_1 + g_2 \in \mathcal{G}$.

(iii) If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then homogeneity gives

$$\begin{aligned}\int_X \alpha g\,\mathrm{d}\nu &= \alpha \int_X g\,\mathrm{d}\nu \\ &\leq \alpha \int_X g\,\mathrm{d}\mu \quad (\text{since } g \in \mathcal{G} \text{ and } \alpha \geq 0\ ) \\ &= \int_X \alpha g\,\mathrm{d}\mu.\end{aligned}$$

Therefore $\alpha g \in \mathcal{G}$.

(iv) Suppose $g \in \mathrm{Bor}^+(X, \mathbb{R})$ and $(g_n)_{n=1}^\infty$ is a sequence in $\mathcal{G}$ such that $g_n \nearrow g$. By the Monotone Convergence Theorem,

$$\begin{aligned}\int_X g\,\mathrm{d}\nu &= \lim_{n\to\infty} \int_X g_n\,\mathrm{d}\nu \\ &\leq \lim_{n\to\infty} \int_X g_n\,\mathrm{d}\mu \quad (\text{since } g_n \in \mathcal{G} \text{ for every } n \in \mathbb{N}\ ) \\ &= \int_X g\,\mathrm{d}\mu.\end{aligned}$$

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

Lemma 6.9 now gives $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$, which is precisely the desired conclusion. □

Remark 12.3 **The Converse of Special case of RN Theorem**

Suppose μ is a positive measure and $h \in \mathrm{Bor}^+(X, \mathbb{R})$. We can define a positive measure ν by

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

We denote this relationship by $\mathrm{d}\nu = h \, \mathrm{d}\mu$.

Suppose, in addition, that $h(x) \leq 1$ for every $x \in X$. Then $h\chi_A \leq \chi_A$ for every $A \in \mathfrak{M}$, so monotonicity of the integral gives

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

Therefore $\nu \leq \mu$.

Theorem 12.4 **A special case of the Radon–Nikodym Theorem**

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 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 uniquely determined in the a.e.-μ sense. That is, if $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R})$ also satisfies $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$, then

$$\tilde{h} = h \quad \text{a.e.-}\mu.$$

Remark 12.5 **Comments about the Theorem**

1. It is easy to see that the converse of Radon-Nikodym (Remark 12.3).
2. The function h obtained by the theorem is called the *R-N Derivative* of ν with respect to μ. We denote it by

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

3. The R-N derivative is unique modulo almost everywhere equivalence. This follows from Q2 of A3.
4. The difficult part of Theorem 12.4 is the existence of h. We will prove this using an $\mathcal{L}^2$-space.

Some lemmas toward the proof:

Lemma 12.6

Let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures such that $\nu \leq \mu$. Then

$$\mathcal{L}^2(\mu) \subseteq \mathcal{L}^2(\nu),$$

and, for every $f \in \mathcal{L}^2(\mu)$,

$$\|f\|_{2,\nu} \leq \|f\|_{2,\mu}.$$

Proof: Fix $f \in \mathcal{L}^2(\mu)$. Since $|f|^2 \in \mathrm{Bor}^+(X, \mathbb{R})$, Proposition 12.2 gives

$$\int_X |f|^2 \, \mathrm{d}\nu \leq \int_X |f|^2 \, \mathrm{d}\mu < \infty.$$

Hence $f \in \mathcal{L}^2(\nu)$, proving $\mathcal{L}^2(\mu) \subseteq \mathcal{L}^2(\nu)$. Moreover, the same inequality says

$$\|f\|_{2,\nu}^2 \leq \|f\|_{2,\mu}^2.$$

Taking square roots yields $\|f\|_{2,\nu} \leq \|f\|_{2,\mu}$. □

Lemma 12.7

Let $\nu : \mathfrak{M} \to [0, \infty)$ be a finite positive measure. Then

$$\mathcal{L}^2(\nu) \subseteq \mathcal{L}^1(\nu).$$

Moreover, for every $f \in \mathcal{L}^2(\nu)$,

$$\left| \int_X f \, \mathrm{d}\nu \right| \leq \sqrt{\nu(X)} \cdot \|f\|_{2,\nu}.$$

Proof: Since ν is finite, the constant function 1 belongs to $\mathcal{L}^2(\nu)$ and

$$\|1\|_{2,\nu} = \sqrt{\int_X 1 \, \mathrm{d}\nu} = \sqrt{\nu(X)}.$$

Fix $f \in \mathcal{L}^2(\nu)$. By the Cauchy–Schwarz inequality (or Holder-2 with $p = q = 2$),

$$\int_X |f| \, \mathrm{d}\nu = \int_X |f| \cdot 1 \, \mathrm{d}\nu \leq \|f\|_{2,\nu} \|1\|_{2,\nu} = \sqrt{\nu(X)} \cdot \|f\|_{2,\nu} < \infty.$$

Thus $f \in \mathcal{L}^1(\nu)$, proving $\mathcal{L}^2(\nu) \subseteq \mathcal{L}^1(\nu)$. Finally,

$$\left| \int_X f \, \mathrm{d}\nu \right| \leq \int_X |f| \, \mathrm{d}\nu \leq \sqrt{\nu(X)} \cdot \|f\|_{2,\nu},$$

as required. □

Proof of Special case of Radon-Nikodym: Suppose $\mu, \nu : \mathfrak{M} \to [0, \infty)$ are finite positive measures such that $\nu \leq \mu$. We want a Radon-Nikodym derivative $h \in \mathrm{Bor}^+(X, \mathbb{R})$.

Scratchwork: the desired h will have

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

for all $A \in \mathfrak{M}$ which implies that

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

for all $g \in \mathrm{Bor}^+(X, \mathbb{R})$ (by Prop 7.10 / Q1 on A3). Note that $\int_X gh \, \mathrm{d}\mu$ looks like an inner product, which reminds us of Riesz Representation.

Claim 1: It makes sense to define $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ by

$$\varphi(g) = \int_X g \, \mathrm{d}\nu$$

for all $g \in \mathcal{L}^2(\mu)$. Moreover, φ is a bounded linear functional on $\mathcal{L}^2(\mu)$.

Verification of Claim 1: We have that $\mathcal{L}^2(\mu) \subseteq \mathcal{L}^2(\nu) \subseteq \mathcal{L}^1(\nu)$ (by Lemma 12.6 and Lemma 12.7, respectively). So for $g \in \mathcal{L}^2(\mu)$, it makes sense to define $\varphi(g) = \int_X g \, \mathrm{d}\nu \in \mathbb{R}$. φ is linear because of the linearity of the integral with respect to ν. Moreover, for $g \in \mathcal{L}^2(\mu)$, we get

$$|\varphi(g)| = \left| \int_X g \, \mathrm{d}\nu \right| \leq \sqrt{\nu(X)} \cdot \|g\|_{2,\nu} \leq \sqrt{\nu(X)} \cdot \|g\|_{2,\mu}$$

by Lemma 12.7 and 12.6, respectively, with g viewed in $\mathcal{L}^2(\nu)$.

Claim 2: There exists $f_0 \in \mathcal{L}^2(\mu)$ such that

$$\heartsuit \int_X g \, \mathrm{d}\nu = \int_X f_0 g \, \mathrm{d}\mu$$

for all $g \in \mathcal{L}^2(\mu)$.

Verification of Claim 2: By Claim 1, φ is a bounded linear functional on $\mathcal{L}^2(\mu)$. Therefore, the Riesz Representation Theorem gives an $f_0 \in \mathcal{L}^2(\mu)$ such that

$$\varphi(g) = \langle f_0, g \rangle, \quad \forall g \in \mathcal{L}^2(\mu).$$

By the definitions of φ and the inner product on $\mathcal{L}^2(\mu)$, this identity becomes

$$\int_X g \, \mathrm{d}\nu = \int_X f_0 g \, \mathrm{d}\mu, \quad \forall g \in \mathcal{L}^2(\mu),$$

which is $\heartsuit$.

Claim 3: Let $f_0 \in \mathcal{L}^2(\mu)$ be as in Claim 2. Let $Z_- = \{x \in X \mid f_0(x) < 0\}$ and $Z_+ = \{x \in X \mid f_0(x) > 1\}$. Then $Z_-, Z_+ \in \mathfrak{M}$ and $\mu(Z_-) = \mu(Z_+) = 0$.

Verification of Claim 3: Since $f_0 \in \mathcal{L}^2(\mu) \subseteq \mathrm{Bor}(X, \mathbb{R})$,

$$Z_- = f_0^{-1}((-\infty, 0)) \quad \text{and} \quad Z_+ = f_0^{-1}((1, \infty))$$

belong to $\mathfrak{M}$. Since μ is finite, both χ_{Z_-} and χ_{Z_+} belong to $\mathcal{L}^2(\mu)$. We may therefore substitute either indicator function for g in ♡.

First, using $g = \chi_{Z_-}$ gives

$$0 \leq \nu(Z_-) = \int_X f_0 \chi_{Z_-} \, \mathrm{d}\mu \leq 0,$$

because $f_0 \chi_{Z_-} \leq 0$. Hence

$$\int_X (-f_0) \chi_{Z_-} \, \mathrm{d}\mu = 0.$$

The integrand is nonnegative, so Remark 8.12 implies that $(-f_0)\chi_{Z_-} = 0$ a.e.-μ. Its set of positive values is precisely Z_-, and therefore $\mu(Z_-) = 0$.

Next, using $g = \chi_{Z_+}$ and the assumption $\nu \leq \mu$, we obtain

$$\begin{aligned}\int_X (f_0 - 1)\chi_{Z_+} \, \mathrm{d}\mu &= \int_X f_0 \chi_{Z_+} \, \mathrm{d}\mu - \int_X \chi_{Z_+} \, \mathrm{d}\mu \\ &= \nu(Z_+) - \mu(Z_+) \\ &\leq 0.\end{aligned}$$

On the other hand, $(f_0 - 1)\chi_{Z_+} \geq 0$, so its integral is also nonnegative and must therefore equal 0. Remark 8.12 now gives $(f_0 - 1)\chi_{Z_+} = 0$ a.e.-μ. Since its set of positive values is exactly Z_+, it follows that $\mu(Z_+) = 0$.

Conclusion of the Proof: Consider the notation as in Claims 2 and 3. Let

$$h = f_0 \left(1 - \chi_{Z_+ \cup Z_-}\right)$$

That is,

$$h(x) = \begin{cases} f_0(x), \text{if } 0 \leq f_0(x) \leq 1 \\ 0, \text{otherwise} \end{cases}$$

Then $h \in \mathrm{Bor}^+(X, \mathbb{R})$ with $h(x) \leq 1$ for all $x \in X$. Moreover, note that $h = f_0$ ae-μ. Indeed, $h - f_0 = -f_0 \chi_{Z_- \cup Z_+} = 0$ ae-μ because $\mu(Z_- \cup Z_+) = 0$. For every $A \in \mathfrak{M}$, define $g = \chi_A$ in ♡. So we get $\nu(A) = \int_X \chi_A \, \mathrm{d}\nu = \int_X f_0 \chi_A \, \mathrm{d}\mu = \int_X h \chi_A \, \mathrm{d}\mu$, showing that h is a R-N derivative. □

13 Radon-Nikodym for $\nu \ll \mu$

13.1 Trick of the Connecting Equation

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

Definition 13.1 **Sum of Measures**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. Define $\mu + \nu : \mathfrak{M} \to [0, \infty]$ by

$$(\mu + \nu)(A) := \mu(A) + \nu(A), \quad \forall A \in \mathfrak{M} \quad (13.1)$$

Then $\mu + \nu$ is itself a positive measure. Moreover,

$$\int_X f \, \mathrm{d}(\mu + \nu) = \int_X f \, \mathrm{d}\mu + \int_X f \, \mathrm{d}\nu, \quad \forall f \in \mathrm{Bor}^+(X, \mathbb{R}) \quad (13.2)$$

The integral identity follows from the method of friendly functions.

Moreover, if μ and ν are finite, then so is $\mu + \nu$.

Proposition 13.2 **Trick of the connecting equation**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures. Then there exist functions $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ such that

$$h_1(x) + h_2(x) = 1, \quad \forall x \in X,$$

and such that

$$\heartsuit \quad \int_X g h_1 \, \mathrm{d}\mu = \int_X g h_2 \, \mathrm{d}\nu, \quad \forall g \in \mathrm{Bor}^+(X, \mathbb{R}).$$

Proof: Put $\rho := \mu + \nu$. Then ρ is a finite positive measure and $\nu \leq \rho$. By Theorem 12.4, there exists $h \in \mathrm{Bor}^+(X, \mathbb{R})$ such that $h(x) \leq 1$ for every $x \in X$ and

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

Define

$$h_1 := h \quad \text{and} \quad h_2 := 1 - h.$$

Since $0 \leq h \leq 1$, we have $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$, and clearly

$$h_1(x) + h_2(x) = 1, \quad \forall x \in X.$$

We prove $\heartsuit$ by the method of friendly functions. Let

$$\mathcal{G} := \left\{ g \in \mathrm{Bor}^+(X, \mathbb{R}) \mid \int_X g h_1 \, \mathrm{d}\mu = \int_X g h_2 \, \mathrm{d}\nu \right\}.$$

We verify the four hypotheses of Lemma 6.9.

(i) Fix $A \in \mathfrak{M}$. By the choice of h_1 and the integration identity in Definition 13.1,

$$\begin{aligned}\nu(A) &= \int_X \chi_A h_1 \,\mathrm{d}\rho \\ &= \int_X \chi_A h_1 \,\mathrm{d}\mu + \int_X \chi_A h_1 \,\mathrm{d}\nu.\end{aligned}$$

On the other hand, since $h_1 + h_2 = 1$,

$$\begin{aligned}\nu(A) &= \int_X \chi_A \,\mathrm{d}\nu \\ &= \int_X \chi_A (h_1 + h_2) \,\mathrm{d}\nu \\ &= \int_X \chi_A h_1 \,\mathrm{d}\nu + \int_X \chi_A h_2 \,\mathrm{d}\nu.\end{aligned}$$

All the integrals above are finite because $0 \le h_1, h_2 \le 1$ and μ, ν are finite. We may therefore cancel $\int_X \chi_A h_1 \,\mathrm{d}\nu$ from the two identities, obtaining

$$\int_X \chi_A h_1 \,\mathrm{d}\mu = \int_X \chi_A h_2 \,\mathrm{d}\nu.$$

Hence $\chi_A \in \mathcal{G}$.

(ii) If $g_1, g_2 \in \mathcal{G}$, then additivity of the integral gives

$$\begin{aligned}\int_X (g_1 + g_2) h_1 \,\mathrm{d}\mu &= \int_X g_1 h_1 \,\mathrm{d}\mu + \int_X g_2 h_1 \,\mathrm{d}\mu \\ &= \int_X g_1 h_2 \,\mathrm{d}\nu + \int_X g_2 h_2 \,\mathrm{d}\nu \\ &= \int_X (g_1 + g_2) h_2 \,\mathrm{d}\nu.\end{aligned}$$

Thus $g_1 + g_2 \in \mathcal{G}$.

(iii) If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then homogeneity gives

$$\begin{aligned}\int_X (\alpha g) h_1 \,\mathrm{d}\mu &= \alpha \int_X g h_1 \,\mathrm{d}\mu \\ &= \alpha \int_X g h_2 \,\mathrm{d}\nu \\ &= \int_X (\alpha g) h_2 \,\mathrm{d}\nu.\end{aligned}$$

Therefore $\alpha g \in \mathcal{G}$.

(iv) Suppose $g \in \mathrm{Bor}^+(X, \mathbb{R})$ and $(g_n)_{n=1}^{\infty}$ is a sequence in $\mathcal{G}$ such that $g_n \nearrow g$. Since $h_1, h_2 \geq 0$, we have $g_n h_1 \nearrow g h_1$ and $g_n h_2 \nearrow g h_2$. By the Monotone Convergence Theorem,

$$\begin{aligned}\int_X g h_1 \,\mathrm{d}\mu &= \lim_{n\to\infty} \int_X g_n h_1 \,\mathrm{d}\mu \\ &= \lim_{n\to\infty} \int_X g_n h_2 \,\mathrm{d}\nu \\ &= \int_X g h_2 \,\mathrm{d}\nu.\end{aligned}$$

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

Lemma 6.9 now gives $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$, which is exactly $\heartsuit$. □

Lemma 13.3

Consider the framework and notation of Proposition 13.2, and let

$$S := \{x \in X \mid h_2(x) = 0\}$$

Then $S \in \mathfrak{M}$ and $\mu(S) = 0$.

Proof: Note that $S \in \mathfrak{M}$ because h_2 is Borel measurable.

Apply the identity $\heartsuit$ from Proposition 13.2 with $g = \chi_S$. Since $h_1 = 1$ and $h_2 = 0$ on S, we obtain

$$\begin{aligned}\mu(S) &= \int_X \chi_S \,\mathrm{d}\mu \\ &= \int_X \chi_S h_1 \,\mathrm{d}\mu \\ &= \int_X \chi_S h_2 \,\mathrm{d}\nu \\ &= 0.\end{aligned}$$

Therefore $\mu(S) = 0$. □

13.2 Radon-Nikodym with absolute continuity

Definition 13.4 **Absolute continuity of measures**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. We say that ν is *absolutely continuous with respect to* μ, denoted by $\nu \ll \mu$, if

$$[A \in \mathfrak{M} \quad \text{and} \quad \mu(A) = 0] \Longrightarrow [\nu(A) = 0].$$

Remark 13.5

Comparing the definitions of $\nu \le \mu$ and $\nu \ll \mu$, we have

$$[\nu \le \mu] \implies [\nu \ll \mu].$$

The converse is not generally true.

Remark 13.6

Recall from Remark 12.3 how a positive measure $\nu : \mathfrak{M} \to [0,\infty]$ can be constructed from a positive measure $\mu : \mathfrak{M} \to [0,\infty]$ and a density $h \in \mathrm{Bor}^+(X,\mathbb{R})$ by setting

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

Unlike in Remark 12.3, we no longer assume that $h(x) \le 1$ for every $x \in X$. In fact, h may be unbounded. Consequently, the resulting measure ν need not satisfy $\nu \le \mu$. For example, suppose that

$$A := \{x \in X \mid h(x) \ge 2\}$$

satisfies $0 < \mu(A) < \infty$. Then

$$\nu(A) \ge 2\mu(A) > \mu(A),$$

so $\nu \le \mu$ does not hold.

Nevertheless, even when h is unbounded, we still have $\nu \ll \mu$. Why is that?

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 uniquely determined in the a.e.-μ sense. That is, if $\tilde{h} \in \mathrm{Bor}^+(X,\mathbb{R})$ also satisfies $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$, then

$$\tilde{h} = h \quad \text{a.e.-}\mu.$$

Proof: Let h_1, h_2 be the functions satisfying Prop 13.2. Let $A = \{x \in X \mid h_1(x) < 1\}$ and $S = \{x \in X \mid h_1(x) = 1\}$. By lemma 13.3, $\mu(S) = 0$. Also, $X = A \cup S$ (disjoint union). Notice that for every $x \in A$, $h_2(x) > 0$. So, given $f \in \mathrm{Bor}^+(X,\mathbb{R})$, set

$$g(x) = \begin{cases} \frac{f(x)}{h_2(x)}, x \in A \\ 0, x \in S \end{cases}$$

We can easily verify that $g \in \mathrm{Bor}^+(X, \mathbb{R})$. Now, applying 13.2 to g gives

$$\begin{aligned} \int_X g h_1 \, \mathrm{d}\mu &= \int_X f \cdot \frac{h_1}{h_2} \, \mathrm{d}\mu \\ &= \int_X g h_2 \, \mathrm{d}\nu \\ &= \int_A f \, \mathrm{d}\nu \end{aligned}$$

Let

$$h(x) = \begin{cases} \frac{h_1(x)}{h_2(x)}, x \in A \\ 0, x \in X \setminus A \end{cases}$$

Then $\int_A f h \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\nu$ Therefore, $\int_S f h \, \mathrm{d}\mu = 0$ and $\int_S f \, \mathrm{d}\nu = 0$ since $\mu(S) = 0$ and $\nu \ll \mu$, so $\nu(S) = 0$. Altogether, we obtain

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

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

For uniqueness, suppose $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R})$ such that $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$. Notice that $\int_X \tilde{h} \, \mathrm{d}\mu = \nu(X) < \infty$. Therefore, $\int_X f h \, \mathrm{d}\mu = \int_X f \tilde{h} \, \mathrm{d}\mu$. Therefore, $\int_X f\big(h - \tilde{h}\big) \, \mathrm{d}\mu = 0$ for all $f \in \mathrm{Bor}^+(X, \mathbb{R})$. Define

$$P = \left\{x \in X \mid \big(h - \tilde{h}\big)(x) \geq 0\right\}$$

and

$$N = \left\{x \in X \mid \big(h - \tilde{h}\big)(x) < 0\right\}$$

Then $\int_X \chi_P\big(h - \tilde{h}\big) \, \mathrm{d}\mu = \int_P \big(h - \tilde{h}\big) \, \mathrm{d}\mu = \int_X \big(h - \tilde{h}\big)^+ \, \mathrm{d}\mu = 0$. So $\big(h - \tilde{h}\big)^+ = 0$ a.e-μ. Similarly, with N shows that $\big(h - \tilde{h}\big)^- = 0$ a.e-μ. So it holds that $h = \tilde{h}$ a.e-μ. □

14 Radon-Nikodym Applications

14.1 Lesbegue Decomposition Theorem

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 μ 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 μ and ν are *mutually singular*, denoted by $\mu \perp \nu$, if there exist sets $P, Q \in \mathfrak{M}$ such that

$$P \cap Q = \varnothing,$$

μ is concentrated on P, and ν is concentrated on Q.

Theorem 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) $\nu_1 + \nu_2 = \nu$, where $\nu = \nu_1 + \nu_2$ is called the *Lesbegue decomposition* of ν with respect to μ

(ii) $\nu_1 \ll \mu$, where ν_1 is the *absolutely continuous part* of ν with respect to μ

(iii) $\nu_2 \perp \mu$, where ν_2 is the *singular part* of ν with respect to μ

Remark 14.3

The uniqueness part of Theorem 14.2 is an elementary verification, left for you to figure out: Problem P14.1.

For the existence part of Theorem 14.2, we will rely on the connecting equation $\heartsuit$ for the measures μ and ν established in Proposition 13.2 of the preceding lecture. This is discussed below.

Proof of Existence in Theorem 14.2: Consider the connecting equation $\heartsuit$ provided by Proposition 13.2: 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

$$\heartsuit \quad \int_X g h_1 \, \mathrm{d}\mu = \int_X g h_2 \, \mathrm{d}\nu, \quad \forall g \in \mathrm{Bor}^+(X, \mathbb{R}).$$

In connection with $\heartsuit$, Lemma 13.3 also tells us that, upon defining

$$S := \{x \in X \mid h_1(x) = 1\} = \{x \in X \mid h_2(x) = 0\} \in \mathfrak{M}, \quad (14.1)$$

we have $\mu(S) = 0$.

The key idea for finding the decomposition $\nu = \nu_1 + \nu_2$ is to use the set S from (14.1) to decompose ν. Namely, define

$$\nu_1(A) := \nu(A \cap (X \setminus S)) \quad \text{and} \quad \nu_2(A) := \nu(A \cap S), \quad \forall A \in \mathfrak{M}. \quad (14.2)$$

It is easy to check that the formulas in (14.2) define two finite positive measures $\nu_1, \nu_2 : \mathfrak{M} \to [0, \infty)$. Moreover, it is immediate that $\nu_1 + \nu_2 = \nu$. Indeed, 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 we used the additivity of ν and the fact that

$$(A \cap (X \setminus S)) \cup (A \cap S) = A$$

is a disjoint union.

It remains to verify that ν_1 and ν_2 satisfy conditions (ii) and (iii) in the statement of the theorem. That is, we must verify the following two claims.

Claim 1. $\nu_2 \perp \mu$.

Verification of Claim 1. From its definition, ν_2 is concentrated on S. Equivalently, $\nu_2(X \setminus S) = \nu((X \setminus S) \cap S) = \nu(\varnothing) = 0$. But μ is concentrated on $Q := X \setminus S$ because $\mu(X \setminus Q) = \mu(S) = 0$ (by lemma 13.3).

Claim 2. $\nu_1 \ll \mu$.

Verification of Claim 2. Let $A \in \mathfrak{M}$ with $\mu(A) = 0$. We want to show that $\nu_1(A) = 0$. But $\nu_1(A) = \nu(A \cap (X \setminus S))$ For $\heartsuit$, let $g = \chi_A$. So

$$\int_X \chi_A h_1 \,\mathrm{d}\mu = \int_X \chi_A h_2 \,\mathrm{d}\nu$$

But $\int_X \chi_A h_1 \,\mathrm{d}\mu = 0$ because $\chi_A h_1 = 0$ a.e-μ (due to $\mu(A) = 0$). Hence, we have $\chi_A h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ which has $\int_X \chi_A h_2 \,\mathrm{d}\nu = 0$. It follows that $\chi_A h_2 = 0$ a.e-μ. Hence,

$$Z = \{x \in X \mid \chi_A(x) h_2(x) \neq 0\}$$

has $\nu(Z) = 0$. But

$$\begin{aligned} Z &= \{x \in X \mid \chi_A(x) \neq 0\} \cap \{h_2(x) \neq 0\} \\ &= A \cap (X \setminus S) \end{aligned}$$

Hence, $\nu_1(A) = \nu(A \cap (X \setminus S)) = 0$, as required. $\square$

14.2 Conditional Expectations

Notation 14.4

Suppose $(X, \mathfrak{M})$ is a measurable space and $\mathfrak{N}$ is a sigma-algebra of subsets of X such that $\mathfrak{N} \subseteq \mathfrak{M}$. In this part of Lecture 14, we use the more elaborate notation

$$\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) := \{f : X \to \mathbb{R} \mid f^{-1}(B) \in \mathfrak{N} \text{ for every } B \in \mathcal{B}_{\mathbb{R}}\}, \tag{14.3}$$

in parallel with

$$\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R}) := \{f : X \to \mathbb{R} \mid f^{-1}(B) \in \mathfrak{M} \text{ for every } B \in \mathcal{B}_{\mathbb{R}}\}. \tag{14.4}$$

The reason for deviating from the familiar notation $\operatorname{Bor}(X, \mathbb{R})$ is that it would now be unclear whether we mean the space in (14.3) or the space in (14.4).

Since $\mathfrak{N} \subseteq \mathfrak{M}$, it follows immediately from the definitions that

$$\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) \subseteq \operatorname{Bor}((X, \mathfrak{M}), \mathbb{R}).$$

This inclusion can be strict. In fact, $\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R})$ can be much smaller than $\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R})$, as shown by the next example.

Example 14.5

Consider the measurable space $(X, \mathfrak{M})$, where

$$X = [0,1] \quad \text{and} \quad \mathfrak{M} = \mathcal{B}_{[0,1]},$$

the Borel sigma-algebra of $[0, 1]$. Partition $[0,1] = J_1 \cup J_2 \cup J_3$, where

$$J_1 = \left[0, \frac{1}{3}\right], \quad J_2 = \left(\frac{1}{3}, \frac{1}{2}\right], \quad J_3 = \left(\frac{1}{2}, 1\right].$$

It is immediate to check that

$$\mathfrak{N} := \{\varnothing, J_1, J_2, J_3, J_1 \cup J_2, J_1 \cup J_3, J_2 \cup J_3, X\} \tag{14.5}$$

is a sigma-algebra of subsets of X satisfying $\mathfrak{N} \subseteq \mathfrak{M}$.

Let us compare $\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R})$ and $\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R})$. First, $\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R})$ is a large infinite-dimensional vector space; for instance, it contains every continuous function $f : [0,1] \to \mathbb{R}$. On the other hand, for a function $g : X \to \mathbb{R}$,

$$g \in \operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) \iff g \text{ is constant on each of } J_1, J_2, J_3. \tag{14.6}$$

Consequently, $\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R})$ is a three-dimensional space:

$$\begin{aligned} \operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) &= \operatorname{span}\left\{\chi_{J_1}, \chi_{J_2}, \chi_{J_3}\right\} \quad (14.7) \\ &= \left\{\alpha_1 \chi_{J_1} + \alpha_2 \chi_{J_2} + \alpha_3 \chi_{J_3} \mid \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\right\}. \end{aligned}$$

The verification of (14.6) and (14.7) is left as an exercise–Problem P14.2 in the problem set for this lecture.

We now move to the situation in which the framework also has a finite positive measure $\mu : \mathfrak{M} \to [0, \infty)$.

Lemma 14.6

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, and let $\mathfrak{N}$ be a sigma-algebra of subsets of X such that $\mathfrak{N} \subseteq \mathfrak{M}$. Let $\mu_0 : \mathfrak{N} \to [0, \infty)$ be the restriction of μ to $\mathfrak{N}$; that is,

$$\mu_0(N) = \mu(N), \quad \forall N \in \mathfrak{N}.$$

Then μ_0 is a finite positive measure on $\mathfrak{N}$, and

$$\int_X g \, \mathrm{d}\mu_0 = \int_X g \, \mathrm{d}\mu, \quad \forall g \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R}). \quad (14.8)$$

The proof of Lemma 14.6 is left as an exercise. One possible approach is the method of friendly functions: define

$$\mathcal{G} := \{g \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R}) \mid \text{Equation (14.8) holds for } g\},$$

and verify that $\mathcal{G}$ satisfies conditions (i)–(iv) of Lemma 6.9. That lemma then gives

$$\mathcal{G} = \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R}),$$

which says precisely that every $g \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R})$ satisfies (14.8).

Definition 14.7 **Conditional expectation**

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, let $\mathfrak{N}$ be a sigma-algebra of subsets of X such that $\mathfrak{N} \subseteq \mathfrak{M}$, and let $\mu_0 : \mathfrak{N} \to [0, \infty)$ be the finite positive measure obtained by restricting μ to $\mathfrak{N}$. Consider the measure space $(X, \mathfrak{N}, \mu_0)$.

Let $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$. A function $g : X \to \mathbb{R}$ is called a *(version of the) conditional expectation of f onto $\mathfrak{N}$* if $g \in \mathcal{L}^1(X, \mathfrak{N}, \mu_0)$ and

$$\int_X \chi_N g \, \mathrm{d}\mu_0 = \int_X \chi_N f \, \mathrm{d}\mu, \quad \forall N \in \mathfrak{N}. \quad (14.9)$$

Comment related to Definition 14.7. The issue with the function f in (14.9) is that it may not be $\mathfrak{N}/\mathcal{B}_{\mathbb{R}}$-measurable. Consequently, integrals of functions $\chi_N f$ may not make sense in the measure space $(X, \mathfrak{N}, \mu_0)$. The conditional expectation g is introduced precisely to fix this issue. Also, g is only determined in an a.e-μ sense, which is why we call it a "version".

Proposition 14.8

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, let $\mathfrak{N}$ be a sigma-algebra of subsets of X such that $\mathfrak{N} \subseteq \mathfrak{M}$, 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 $\mathfrak{N}$ in the sense of Definition 14.7, and g may be chosen so that

$$g(x) \geq 0, \quad \forall x \in X.$$

Proof: We look at $\mu : \mathfrak{M} \to [0, \infty)$ and also at $\nu : \mathfrak{M} \to [0, \infty)$ defined by $\mathrm{d}\nu = f \, \mathrm{d}\mu$. That is,

$$\nu(A) = \int_A f \, \mathrm{d}\mu$$

for all $A \in \mathfrak{M}$. Then ν is a finite positive measure on $\mathfrak{M}$ (by Prop 7.9) with $\nu \ll \mu$. Now, let us restrict to $\mathfrak{N}$. Let $\mu_0, \nu_0 : \mathfrak{N} \to [0, \infty)$ be defined by $\mu_0(N) = \mu(N)$ and $\nu_0(N) = \nu(N)$ for all $N \in \mathfrak{N}$. μ_0, ν_0 are finite positive measures on $\mathfrak{N}$ with $\nu_0 \ll \mu_0$ (since $\nu \ll \mu$ on $\mathfrak{M}$). Apply Radon-Nikodym to get a R-N derivative $g \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R})$ such that $\mathrm{d}\nu_0 = g \, \mathrm{d}\mu_0$. That is, $\nu_0(N) = \int_N g \, \mathrm{d}\mu_0$ for all $N \in \mathfrak{N}$. So, for every $N \in \mathfrak{N}$, we have

$$\begin{aligned} \int_N g \, \mathrm{d}\mu_0 &= \nu_0(N) \\ &= \nu(N) \\ &= \int_N f \, \mathrm{d}\mu \end{aligned}$$

Wait, is this g in $\mathcal{L}^1(X, \mathfrak{N}, \mu_0)$?

Yes, because for $N = X$, we get $\int_X g \, \mathrm{d}\mu_0 = \int_X f \, \mathrm{d}\mu < \infty$ since $f \in \mathcal{L}^1(\mu, \mathfrak{M}, \mu)$. □

Corollary 14.9

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, let $\mathfrak{N}$ be a sigma-algebra of subsets of X such that $\mathfrak{N} \subseteq \mathfrak{M}$, and let $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$. Then:

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 $Z \in \mathfrak{N}$ such that $\mu(Z) = 0$ and

$$g(x) = g'(x), \quad \forall x \in X \setminus Z.$$

The proof of Corollary 14.9 is left as an exercise:

- To prove 1, consider the positive and negative parts f_+ and f_- of $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$ and apply Proposition 14.8 to each of them.
- To prove 2, use Q2 of A3 in connection with the function

$$|g - g'| \in \text{Bor}^+((X, \mathfrak{N}), \mathbb{R}).$$

Notation 14.10

Consider the same framework and notation as in Corollary 14.9: a measure space $(X, \mathfrak{M}, \mu)$, a sigma-algebra $\mathfrak{N} \subseteq \mathfrak{M}$, and a function $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$. A conditional expectation g of f onto $\mathfrak{N}$ is often denoted by

$$g = E[f \mid \mathfrak{N}].$$

Thus $E[f \mid \mathfrak{N}]$ is a function in $\mathcal{L}^1(X, \mathfrak{N}, \mu_0)$, where μ_0 is the restriction of μ to $\mathfrak{N}$. Keep in mind that $E[f \mid \mathfrak{N}]$ is not uniquely determined: it is defined only up to equality a.e.-μ_0. This is why we speak about *versions* of the conditional expectation of f onto $\mathfrak{N}$.

Example

Continuing on Example 14.5.

Let $f(x) = x^2$ for all $x \in [0, 1]$. We accept that there exists probability measure $\mu : \mathcal{B}_{[0,1]} \to [0, 1]$, uniquely determined, such that $\mu([a, b]) = b - a$ for all $0 \leq a \leq b \leq 1$. In particular, we have $\mu(\{a\}) = 0$ for all $a \in [0, 1]$. This implies that $\mu(J) = \sup(J) - \inf(J)$ for every (open, half-open, closed, ...) interval $J \subseteq [0, 1]$.

Important property of μ: for $f : [0, 1] \to \mathbb{R}$ continuous, we have

$$\int_X f \, \mathrm{d}\mu = \int_0^1 f(x) \, \mathrm{d}x$$

where the right hand side denotes a Riemann integral, and moreover,

$$\int_J f \, \mathrm{d}\mu = \int_{\inf(J)}^{\sup(J)} f(x) \, \mathrm{d}x$$

for every interval $J \subseteq [0, 1]$.

We want to find $E[f \mid \mathfrak{N}]$ for $f(x) = x^2$ and $\mathfrak{N}$ as above.

Have: $E[f \mid \mathfrak{N}] = g \in \text{Bor}((X, \mathfrak{N}), \mathbb{R})$ such that

$$\int_N g \, \mathrm{d}\mu = \int_N x^2 \, \mathrm{d}\mu(x)$$

for all $N \in \mathfrak{N}$.

But $\operatorname{Bor}((X,\mathfrak{N}),\mathbb{R}) = \operatorname{span}\{\chi_{J_1},\chi_{J_2},\chi_{J_3}\}$. Hence, g must be of the form

$$g = \alpha_1\chi_{J_1} + \alpha_2\chi_{J_2} + \alpha_3\chi_{J_3}$$

for some $\alpha_1,\alpha_2,\alpha_3 \in \mathbb{R}$.

How do we compute the α_i's?

Start with α_1. We have $\int_{J_1} g\,\mathrm{d}\mu = \int_{J_1} x^2\,\mathrm{d}\mu(x) = \int_0^{\frac{1}{3}} x^2\,\mathrm{d}x = \left[\frac{1}{3}x^3\right]_0^{\frac{1}{3}} = \frac{1}{81}$. On the other hand,

$$\begin{aligned}\int_{J_1} g\,\mathrm{d}\mu &= \int_X g\chi_{J_1}\,\mathrm{d}\mu \\ &= \int_X \left(\alpha_1\chi_{J_1} + \alpha_2\chi_{J_2} + \alpha_3\chi_{J_3}\right)\chi_{J_1}\,\mathrm{d}\mu \\ &= \int_X \alpha_1\chi_{J_1}\,\mathrm{d}\mu \\ &= \alpha_1 \cdot \mu(J_1) \\ &= \alpha_1\frac{1}{3}\end{aligned}$$

Therefore, $\frac{\alpha_1}{3} = \frac{1}{81}$, so $\alpha_1 = \frac{1}{27}$. Use the same procedure to compute α_2, α_3.

15 How to Construct Measures via the Caratheodory Extension Theorem

15.1 Motivation: Why an Extension Theorem Is Useful

Remark 15.1

For a metric space X, there is usually no intrinsic description of all sets in the Borel sigma-algebra $\mathcal{B}_X$, nor a direct way to define a measure on every such set. What we often do have is:

- an algebra $\mathfrak{A} \subseteq \mathcal{B}_X$ with $\sigma\text{-Alg}(\mathfrak{A}) = \mathcal{B}_X$;
- an explicit additive set-function $\mu_0 : \mathfrak{A} \to [0,\infty]$.

The extension problem is to find conditions ensuring that μ_0 extends to a positive measure $\mu : \mathcal{B}_X \to [0,\infty]$. The relevant condition will be that μ_0 is a pre-measure.

Example 15.2

Let $X = \mathbb{R}$ with its usual metric and consider the semi-algebra

$$\mathcal{J} := \{\varnothing\} \cup \{(a,b] \mid a < b \text{ in } \mathbb{R}\} \cup \{(-\infty,b] \mid b \in \mathbb{R}\} \cup \{(a,\infty) \mid a \in \mathbb{R}\} \cup \{\mathbb{R}\}.$$

Let $\mathcal{E} := \mathrm{Alg}(\mathcal{J})$. Equivalently, $\mathcal{E}$ consists of all finite disjoint unions of sets in $\mathcal{J}$. Moreover,

$$\sigma\text{-Alg}(\mathcal{E}) = \mathcal{B}_{\mathbb{R}}.$$

Now let $G : \mathbb{R} \to \mathbb{R}$ be increasing and put

$$\ell_+ := \lim_{t\to\infty} G(t) \quad \text{and} \quad \ell_- := \lim_{s\to-\infty} G(s).$$

Define μ_0 on $\mathcal{J}$ by $\mu_0(\varnothing) = 0$ and

$$\begin{aligned} \mu_0((a,b]) &:= G(b) - G(a), \quad a < b, \quad (15.1) \\ \mu_0((-\infty,b]) &:= G(b) - \ell_-, \\ \mu_0((a,\infty)) &:= \ell_+ - G(a), \\ \mu_0(\mathbb{R}) &:= \ell_+ - \ell_-. \quad (15.2) \end{aligned}$$

Finite additivity extends μ_0 to $\mathcal{E}$.

The question is whether $\mu_0 : \mathcal{E} \to [0,\infty]$ extends to a positive measure on $\mathcal{B}_{\mathbb{R}}$. It does when G is cadlag: in that case μ_0 is a pre-measure, so the Caratheodory extension theorem applies.

15.2 Pre-measures and Caratheodory's Extension Theorem

Definition 15.3 **Pre-measure**

Let X be nonempty, let $\mathfrak{A}$ be an algebra of subsets of X, and let $\mu_0 : \mathfrak{A} \to [0,\infty]$ be additive. We call μ_0 a *pre-measure* if it satisfies the following *pre-sigma-additivity* condition:

$$\text{(Pre-Sigma-Add)} \quad \begin{cases} \text{If } (A_n)_{n=1}^{\infty} \text{ are sets in } \mathfrak{A} \text{ with } A_i \cap A_j = \varnothing \text{ for } i \neq j \\ \text{and if } \cup_{n=1}^{\infty} A_n \text{ still belongs to } \mathfrak{A} \\ \text{then } \mu_0(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu_0(A_n). \end{cases}$$

Comment: Since $\mathfrak{A}$ need not be a sigma-algebra, a countable union of sets in $\mathfrak{A}$ may leave $\mathfrak{A}$. Thus (Pre-Sigma-Add) applies only when the countable union happens to remain in $\mathfrak{A}$.

The condition is *necessary* for an extension: if a positive measure μ on σ-Alg$(\mathfrak{A})$ extends μ_0, then the sigma-additivity of μ forces (Pre-Sigma-Add) for μ_0. Caratheodory's theorem says that the condition is also *sufficient* for such an extension to exist.

Theorem 15.4 **Caratheodory extension theorem**

Let X be nonempty, let $\mathfrak{A}$ be an algebra of subsets of X, and put

$$\mathfrak{M} := \sigma\text{-Alg}(\mathfrak{A}).$$

If $\mu_0 : \mathfrak{A} \to [0, \infty]$ is an additive pre-measure, then there exists a positive measure $\mu : \mathfrak{M} \to [0, \infty]$ extending μ_0:

$$\mu(A) = \mu_0(A), \quad \forall A \in \mathfrak{A}.$$

Remark 15.5 **Caratheodory's blueprint**

The construction has two stages:

1. *Overshoot.* Extend μ_0 from $\mathfrak{A}$ to an outer measure $\mu^* : 2^X \to [0, \infty]$.
2. *Trim down.* Find a sigma-algebra $\mathcal{G} \subseteq 2^X$ containing $\mathfrak{A}$ such that $\tilde{\mu} := \mu^*|_{\mathcal{G}}$ is a positive measure.

This suffices because:

- any sigma-algebra $\mathcal{G}$ containing $\mathfrak{A}$ also contains $\mathfrak{M} = \sigma\text{-Alg}(\mathfrak{A})$;
- the restriction $\mu := \tilde{\mu}|_{\mathfrak{M}}$ is therefore a positive measure;
- since μ^* extends μ_0, so does μ.

Definition 15.6 **Outer measure**

Let X be nonempty. A function $\mu^* : 2^X \to [0, \infty]$ is an *outer measure* if it satisfies:

(OM1) $\mu^*(\varnothing) = 0$.
(OM2) If $E \subseteq F \subseteq X$, then $\mu^*(E) \leq \mu^*(F)$.
(OM3) For every sequence $(E_n)_{n=1}^{\infty}$ of subsets of X,

$$\mu^*(\cup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu^*(E_n).$$

Proposition 15.7 **Overshoot step**

Let X be nonempty, let $\mathfrak{A}$ be an algebra of subsets of X, and let $\mu_0 : \mathfrak{A} \to [0, \infty]$ be a pre-measure. For $E \subseteq X$, define

$$\mu^*(E) := \inf\left\{ \sum_{n=1}^{\infty} \mu_0(A_n) \mid A_n \in \mathfrak{A} \text{ for every } n \in \mathbb{N}, E \subseteq \cup_{n=1}^{\infty} A_n \right\}. \quad (15.3)$$

Then:

1. μ^* is an outer measure.
2. $\mu^*(A) = \mu_0(A)$ for every $A \in \mathfrak{A}$.

A sequence $(A_n)_{n=1}^{\infty}$ occurring in (15.3) is called a *countable cover* of E.

Proposition 15.8 **Trim-down step, Part 1**

Let X be nonempty and let $\mu^* : 2^X \to [0, \infty]$ be an outer measure. Call a set $G \subseteq X$ *good for* μ^* if

$$\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$$

whenever $E \subseteq G$ and $F \subseteq X \setminus G$. Then:

1. The collection

   $$\mathcal{G} := \{G \subseteq X \mid G \text{ is good for } \mu^*\}$$

   is a sigma-algebra.
2. The restriction $\tilde{\mu} := \mu^*|_{\mathcal{G}}$ is a positive measure on $\mathcal{G}$.

Proposition 15.9 **Trim-down step, Part 2**

In the framework of Proposition 15.7, every $A \in \mathfrak{A}$ is good for the outer measure μ^* defined there. Consequently,

$$\mathfrak{A} \subseteq \mathcal{G}.$$

15.3 Lebesgue–Stieltjes Measures on the Real Line

Return to Example 15.2 and its increasing function $G : \mathbb{R} \to \mathbb{R}$.

Definition 15.10 **Cadlag function**

The function G is *cadlag* if it is right-continuous at every $b \in \mathbb{R}$:

$$\lim_{s \to b^+} G(s) = G(b), \quad \forall b \in \mathbb{R}. \quad (15.4)$$

The name comes from the French *continue à droite, limite à gauche*: it is continuous from the right and has limits from the left. Since G is increasing, its left-hand limit exists at every $a \in \mathbb{R}$ and satisfies

$$\lim_{t \to a^-} G(t) \leq G(a),$$

although strict inequality may occur. (15.5)

Proposition and Definition 15.11

Suppose the increasing function $G : \mathbb{R} \to \mathbb{R}$ from Example 15.2 is cadlag. Then the additive set-function $\mu_0 : \mathcal{E} \to [0, \infty]$ constructed there is a pre-measure, and hence extends to a positive measure

$$\mu : \mathcal{B}_{\mathbb{R}} \to [0, \infty].$$

The resulting measure μ is called the *Lebesgue–Stieltjes measure defined by* G. It is uniquely determined by

$$\mu((a, b]) = G(b) - G(a), \quad \forall a < b \ \text{ in } \ \mathbb{R}.$$

Proof of Uniqueness of μ: Let $\mu, \tilde{\mu} : \mathcal{B}_{\mathbb{R}} \to [0, \infty]$ are such that $\mu((a, b]) = G(b) - G(a) = \tilde{\mu}((a, b])$. Then for $a < b$ in $\mathbb{R}$, we also have

$$\mu((a, b)) = \lim_{\substack{n \to \infty \\ n \geq n_0}} \mu\left(\left(a, b - \frac{1}{n}\right]\right) = \ldots = \tilde{\mu}((a, b))$$

From $\mu((a, b)) = \tilde{\mu}((a, b))$ for all $a < b$ in $\mathbb{R}$, we further concldue that $\mu(U) = \tilde{\mu}(U)$ for all $U \subseteq \mathbb{R}$ open. Write U as a countable disjoint union of open intervals. Finally, get $\mu(B) = \tilde{\mu}(B)$ for all $B \in \mathcal{B}_{\mathbb{R}}$ by using regularity considerations from P2.5-Q3 on A1. □

Remark 15.12

How does the cadlag property arise in Example 15.2? In particular, where does the asymmetry between the left and the right come from?

Think in reverse, starting from μ. Suppose $(s_n)_{n=1}^{\infty}$ decreases to b, with $s_n > b$, and fix $a < b$. Then

$$(a, s_n] \downarrow (a, b].$$

Since $\mu((a, s_1]) < \infty$, continuity from above gives

$$\begin{aligned} G(b) - G(a) &= \mu((a, b]) \\ &= \lim_{n\to\infty} \mu((a, s_n]) \\ &= \lim_{n\to\infty} (G(s_n) - G(a)). \end{aligned}$$

Therefore $\lim_{n\to\infty} G(s_n) = G(b)$, which is right-continuity.

Now suppose $(t_n)_{n=1}^{\infty}$ increases to a, with $t_n < a$, and fix $b > a$. Then $(t_n, b] \downarrow [a, b]$, so

$$\begin{aligned} \lim_{n\to\infty} (G(b) - G(t_n)) &= \lim_{n\to\infty} \mu((t_n, b]) \\ &= \mu([a, b]) \\ &= \mu(\{a\}) + G(b) - G(a). \end{aligned}$$

Hence

$$\lim_{n\to\infty} G(t_n) = G(a) - \mu(\{a\}) \leq G(a).$$

The inequality is strict exactly when a is an atom of μ, meaning that $\mu(\{a\}) > 0$. This explains why right-continuity is required while a jump from the left is allowed.

Example — **Lebesgue measure on an interval**

Fix $a < b$ in $\mathbb{R}$ and define

$$G(t) := \begin{cases} 0 & \text{if } \ t \leq a \\ t - a & \text{if } \ a < t \leq b \\ b - a & \text{if } \ t > b. \end{cases}$$

This function is increasing and continuous, hence cadlag. Its Lebesgue–Stieltjes measure is Lebesgue measure restricted to $[a, b]$:

$$\mu(E) = \lambda(E \cap [a, b]), \quad \forall E \in \mathcal{B}_{\mathbb{R}}.$$

In particular, $\mu(\mathbb{R}) = b - a$.