Measure Theory

R.Rusev

1   σ-Algebras

Definition 1.1. A σ-algebra $\mathcal{A}$ on a set X is a family of subsets of X such that:

- $X \in \mathcal{A}$ ($\mathbf{\Sigma}_1$)

- If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ ($\mathbf{\Sigma}_2$)

- If $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$, then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ ($\mathbf{\Sigma}_3$)

A set $A \in \mathcal{A}$ is said to be *measurable* or $\mathcal{A}$-*measurable*.

Example 1.1.

1. $\mathcal{P}(X)$ is a σ-algebra (the maximal σ-algebra on X).

2. $\{\emptyset, X\}$ is a σ-algebra (the minimal σ-algebra on X).

3. $A := \{A \subseteq X : \#A < \infty \text{ or } \#A^c < \infty\}$ is a σ-algebra.

4. (Trace σ-algebra) Let $E \subseteq X$ be any set and let $\mathcal{A}$ be a σ-algebra on X. Then

$$\mathcal{A}_E := \{E \cap A : A \in \mathcal{A}\}$$

   is a σ-algebra on E.

   Proof. We verify the three defining properties of a σ-algebra on E:

   - Since $X \in \mathcal{A}$, we have $E = E \cap X \in \mathcal{A}_E$.
   - If $E \cap A \in \mathcal{A}_E$, then $E \setminus (E \cap A) = E \cap A^c$, and since $A^c \in \mathcal{A}$, it follows that $E \cap A^c \in \mathcal{A}_E$.
   - If $(E \cap A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}_E$, then $\bigcup_n (E \cap A_n) = E \cap \bigcup_n A_n$, and since $\bigcup_n A_n \in \mathcal{A}$, we conclude that $\bigcup_n (E \cap A_n) \in \mathcal{A}_E$.

   Hence, $\mathcal{A}_E$ is a σ-algebra on E. $\square$

5. (Pre-image σ-algebra) Let $f : X \to X'$ be a function and let $\mathcal{A}'$ be a σ-algebra on X'. Then
$$\mathcal{A} := \{f^{-1}(A') : A' \in \mathcal{A}'\}$$
   is a σ-algebra on X.

Theorem 1.1. Let X be a set and let $\{\mathcal{A}_i : i \in I\}$ be a family of σ-algebras on X. Define

$$\mathcal{A} := \bigcap_{i \in I} \mathcal{A}_i = \{A \subseteq X : A \in \mathcal{A}_i \text{ for all } i \in I\}.$$

Then, $\mathcal{A}$ is a σ-algebra on X.

Proof. We verify the σ-algebra properties for $\mathcal{A}$:

- Since $X \in \mathcal{A}_i$ for all $i \in I$, we have $X \in \mathcal{A}$.

- If $A \in \mathcal{A}$, then $A \in \mathcal{A}_i$ for all $i \in I$, so $A^c = X \setminus A \in \mathcal{A}_i$ for all $i \in I$, hence $A^c \in \mathcal{A}$.

- If $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$, then $A_n \in \mathcal{A}_i$ for all n and i, so $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}_i$ for all $i \in I$, and thus $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Therefore, $\mathcal{A}$ is a σ-algebra on X. $\qquad\square$

Definition 1.2. Let X be a set and let $\mathcal{E} \subseteq \mathcal{P}(X)$ be a collection of subsets of X. The *σ-algebra generated by $\mathcal{E}$*, denoted by $\sigma(\mathcal{E})$, is the smallest σ-algebra on X containing all sets in $\mathcal{E}$. That is,

$$\sigma(\mathcal{E}) := \bigcap \Big\{ \mathcal{A} \subseteq \mathcal{P}(X) : \mathcal{A} \text{ is a } \sigma\text{-algebra on } X, \ \mathcal{E} \subseteq \mathcal{A} \Big\}.$$

Remark 1.1 (Generated σ-algebras)**.**

- If $\mathcal{A}$ is a σ-algebra, then $\sigma(\mathcal{A}) = \mathcal{A}$.

- For $A \subseteq X$, we have $\sigma(\{A\}) = \{\emptyset, A, A^c, X\}$.

- If $\mathcal{F} \subseteq \mathcal{G} \subseteq \mathcal{A}$, then $\sigma(\mathcal{F}) \subseteq \sigma(\mathcal{G}) \subseteq \sigma(\mathcal{A})$.

Definition 1.3 (Topological Space)**.** A *topological space* is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T} \subseteq \mathcal{P}(X)$ is a collection of subsets of X, called *open sets*, satisfying the following properties:

- $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,

- If $\{U_\alpha \in \mathcal{T} : \alpha \in I\}$ is an arbitrary collection of open sets, then the union $\bigcup_{\alpha \in I} U_\alpha \in \mathcal{T}$,

- If $\{U_i \in \mathcal{T} : i = 1, \ldots, n\}$ is a finite collection of open sets, then the intersection $\bigcap_{i=1}^{n} U_i \in \mathcal{T}$.

The collection $\mathcal{T}$ is called a *topology* on X. The complement of an open set is called a *closed set*.

Remark 1.2 (Standard Topology on $\mathbb{R}^n$)**.** A subset $U \subseteq \mathbb{R}^n$ is called *open* if for every point $x \in U$, there exists an $\varepsilon > 0$ such that the open ball

$$B_\varepsilon(x) := \{y \in \mathbb{R}^n : \|x - y\| < \varepsilon\},$$

where $\|\cdot\|$ denotes the Euclidean norm, is contained in U; that is, $B_\varepsilon(x) \subseteq U$.

The collection of all such open sets is denoted by $\mathcal{O} = \mathcal{O}_{\mathbb{R}^n}$ and forms the *standard topology* on $\mathbb{R}^n$.

Definition 1.4 (Borel σ-algebra)**.** The σ-algebra $\sigma(\mathcal{O})$ generated by the collection of open sets $\mathcal{O} = \mathcal{O}_{\mathbb{R}^n}$ of $\mathbb{R}^n$ is called the *Borel σ-algebra* on $\mathbb{R}^n$.

Its elements are called *Borel sets* or *Borel measurable sets*. We denote the Borel σ-algebra by $\mathcal{B}(\mathbb{R}^n)$.

Definition 1.5. Let X be a topological space and let $A \subseteq X$. A collection $\{U_\alpha\}_{\alpha \in A} \subseteq \mathcal{T}$ of open sets is called an *open cover* of A if

$$A \subseteq \bigcup_{\alpha \in A} U_\alpha.$$

A *subcover* is a subcollection that still covers A. The set A is called *compact* if every open cover of A admits a finite subcover.

Remark 1.3. In $\mathbb{R}^n$, a set is compact if and only if it is closed and bounded (Heine–Borel Theorem).

Theorem 1.2 (Borel σ-algebra from Different Generators)**.** Let $\mathcal{O}, \mathcal{C}, \mathcal{K} \subseteq \mathcal{P}(\mathbb{R}^n)$ denote the collections of open, closed, and compact subsets of $\mathbb{R}^n$, respectively. Then,

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{O}) = \sigma(\mathcal{C}) = \sigma(\mathcal{K}).$$

Proof. Since compact sets are closed, we have $\mathcal{K} \subseteq \mathcal{C}$, and by Remark 1.1, $\sigma(\mathcal{K}) \subseteq \sigma(\mathcal{C})$. Conversely, for any $C \in \mathcal{C}$, define $C_k := C \cap B_k(0)$, where $B_k(0)$ is the closed ball of radius k centered at the origin. Each C_k is closed and bounded, hence compact, so $C_k \in \mathcal{K}$. Since $C = \bigcup_{k \in \mathbb{N}} C_k$, it follows that $C \in \sigma(\mathcal{K})$, and thus $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{K})$.

Next, since $\mathcal{C} = \mathcal{O}^c := \{U^c : U \in \mathcal{O}\}$, and complements of sets in a σ-algebra are again in the σ-algebra, it follows that $\mathcal{C} \subseteq \sigma(\mathcal{O})$, hence $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{O})$. The reverse inclusion follows similarly from $\mathcal{O} = \mathcal{C}^c$. Therefore, we have:

$$\sigma(\mathcal{K}) = \sigma(\mathcal{C}) = \sigma(\mathcal{O}).$$

$\square$

Generating Sets of the Borel Algebra. The Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ can be generated by various systems of sets. Of particular importance are:

- The family of open rectangles:

$$\mathcal{J}_{o,n} := \{(a_1, b_1) \times \cdots \times (a_n, b_n) : a_i, b_i \in \mathbb{R}\},$$

- The family of half-open rectangles:

$$\mathcal{J}_n := \{[a_1, b_1) \times \cdots \times [a_n, b_n) : a_i, b_i \in \mathbb{R}\}.$$

We denote by $\mathcal{J}_n^{\mathrm{rat}}, \mathcal{J}_{o,n}^{\mathrm{rat}}$ the subsets with rational endpoints. These sets represent intervals in $\mathbb{R}$, rectangles in $\mathbb{R}^2$, cuboids in $\mathbb{R}^3$, and hypercubes in higher dimensions.

Theorem 1.3. We have the following equality of Borel σ-algebras on $\mathbb{R}^n$:

$$\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{J}_n^{\mathrm{rat}}) = \sigma(\mathcal{J}_{o,n}^{\mathrm{rat}}) = \sigma(\mathcal{J}_n) = \sigma(\mathcal{J}_{o,n}),$$

Remark 1.4. Let $D \subseteq \mathbb{R}$ be a dense subset, for example $D = \mathbb{Q}$ or $D = \mathbb{R}$. Then the Borel sets on $\mathbb{R}$ can also be generated by any of the following families of intervals:

$$\{(-\infty, a) : a \in D\}, \quad \{(-\infty, a] : a \in D\}, \quad \{(a, \infty) : a \in D\}, \quad \{[a, \infty) : a \in D\}.$$

2 Measure Spaces

Definition 2.1. A *(positive) measure* μ on X is a map $\mu : \mathcal{A} \to [0, \infty]$, where $\mathcal{A}$ is a σ-algebra on X, satisfying:

$$\mu(\emptyset) = 0, \tag{M1}$$

and for any pairwise disjoint sequence $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$,

$$\mu\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n). \tag{M2}$$

Property (M2) is also called *countable additivity*.

If μ satisfies (M1), (M2), but $\mathcal{A}$ is not a σ-algebra, then μ is called a *pre-measure*.

Remark 2.1. (M2) requires implicitly that $\bigsqcup_n A_n$ is again in $\mathcal{A}$ this is clearly the case for σ-algebras, but needs special attention when dealing with pre-measures.

Definition 2.2 (Monotone sequences of sets). Let $(A_n)_{n\in\mathbb{N}}$ and $(B_n)_{n\in\mathbb{N}}$ be sequences of subsets of X.

We say (A_n) is *increasing* if

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$$

and write $A_n \uparrow A$ where

$$A := \bigcup_{n\in\mathbb{N}} A_n$$

Similarly, (B_n) is *decreasing* if

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$$

and write $B_n \downarrow B$ where

$$B := \bigcap_{n\in\mathbb{N}} B_n$$

Definition 2.3. Let X be a set and $\mathcal{A}$ a σ-algebra on X. The pair $(X, \mathcal{A})$ is called a *measurable space*. If μ is a measure on $(X, \mathcal{A})$, then $(X, \mathcal{A}, \mu)$ is called a *measure space*.

A measure μ is called:

- *finite* if $\mu(X) < \infty$,

- a *probability measure* if $\mu(X) = 1$.

Accordingly, we speak of a *finite measure space* and a *probability space*.

Definition 2.4. A measure μ on a measurable space $(X, \mathcal{A})$ is called σ-*finite* if there exists a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ such that:

$$A_n \uparrow X \quad \text{and} \quad \mu(A_n) < \infty \quad \text{for all } n \in \mathbb{N}.$$

In this case, the measure space $(X, \mathcal{A}, \mu)$ is called σ-*finite*.

Lemma 2.1 (Basic properties of measures). Let $(X, \mathcal{A}, \mu)$ be a measure space. Then:

(i) If $A_0, \ldots, A_k \in \mathcal{A}$ are pairwise disjoint, then $\mu\left(\bigsqcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} \mu(A_i)$.

(ii) If $A, B \in \mathcal{A}$ with $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

(iii) If $A, B \in \mathcal{A}$, $A \subseteq B$, and $\mu(A) < \infty$, then $\mu(B \setminus A) = \mu(B) - \mu(A)$.

Proof. (i) Extend (A_n) by $A_n = \varnothing$ for $n > k$. Then by countable additivity,

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

(ii) Since $B = A \sqcup (B \setminus A)$ we have

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

(iii) Rearranging gives

$$\mu(B \setminus A) = \mu(B) - \mu(A),$$

which is well-defined if $\mu(A) < \infty$. $\qquad\square$

Lemma 2.2 (Main properties of measures)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. Then:

(i) **Countable subadditivity:** For any countable family $\{A_i\}_{i \in \mathbb{N}} \subseteq \mathcal{A}$,

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) \leq \sum_{i \in \mathbb{N}} \mu(A_i).$$

(ii) **Continuity from below (increasing sequence):** If $A_1 \subseteq A_2 \subseteq \cdots$ (i.e., $A_n \uparrow A$), then

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

(iii) **Continuity from above (decreasing sequence):** If $B_1 \supseteq B_2 \supseteq \cdots$ (i.e., $B_n \downarrow B$), then

$$\mu\left(\bigcap_{n \in \mathbb{N}} B_n\right) = \lim_{n \to \infty} \mu(B_n).$$

Proof. (i) For countable subadditivity, set $B_k := A_k \setminus \bigcup_{i=1}^{k-1} A_i$, so that (B_k) are disjoint with $B_k \subseteq A_k$. Then,

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

(ii) Let $A_n \uparrow A$, i.e., $A_n \subseteq A_{n+1}$ and $A = \bigcup_{n \in \mathbb{N}} A_n$. Define $B_n := A_n \setminus A_{n-1}$ with $A_0 := \emptyset$. Then (B_n) is disjoint and $\bigsqcup_n B_n = A$. By countable additivity,

$$\mu(A) = \sum_{n=1}^{\infty} \mu(B_n) = \lim_{n \to \infty} \sum_{k=1}^{n} \mu(B_k) = \lim_{n \to \infty} \mu(A_n)$$

(iii) Assume $B_n \downarrow B$, i.e., $B_n \supseteq B_{n+1}$ and $B = \bigcap_n B_n$, with $\mu(B_1) < \infty$. Set $A_n := B_1 \setminus B_n$, so $A_n \uparrow A := B_1 \setminus B$. Then

$$\mu(B) = \mu(B_1) - \mu(A) = \mu(B_1) - \lim_{n \to \infty} \mu(A_n) = \lim_{n \to \infty} \mu(B_n)$$

$\square$

Remark 2.2. With appropriate modifications, these properties also hold for pre-measures, i.e., when $\mathcal{A}$ is not necessarily a σ-algebra.

Example 2.1 (Dirac measure)**.** Let $(X, \mathcal{A})$ be a measurable space and let $x \in X$. Define $\delta_x : \mathcal{A} \to \{0, 1\}$ by

$$\delta_x(A) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Then δ_x is a measure on $(X, \mathcal{A})$, called the *Dirac measure* (or unit mass) at the point x.

Example 2.2 (Counting measure)**.** Let $(X, \mathcal{A})$ be a measurable space. Define $\#A : \mathcal{A} \to [0, \infty]$ by

$$\#A := \begin{cases} \text{number of elements in } A & \text{if } A \text{ is finite}, \\ \infty & \text{if } A \text{ is infinite}. \end{cases}$$

Then $\#$ is a measure on $(X, \mathcal{A})$, called the *counting measure.*

Example 2.3 (Discrete probability measure)**.** Let $\Omega = \{\omega_1, \omega_2, \dots\}$ be a countable set, and let $(p_n)_{n \in \mathbb{N}} \subseteq [0, 1]$ be a sequence such that $\sum_{n \in \mathbb{N}} p_n = 1$. Define the set function $P : \mathcal{P}(\Omega) \to [0, 1]$ by

$$P(A) := \sum_{\{n \in \mathbb{N} : \omega_n \in A\}} p_n = \sum_{n \in \mathbb{N}} p_n \, \delta_{\omega_n}(A), \quad A \subseteq \Omega,$$

where δ_{ω_n} denotes the Dirac measure at ω_n. Then P is a probability measure on $(\Omega, \mathcal{P}(\Omega))$, and the triplet $(\Omega, \mathcal{P}(\Omega), P)$ is called a *discrete probability space.*

Example 2.4 (Linear combination of measures)**.** Let $(X, \mathcal{A})$ be a measurable space, and let $(\mu_n)_{n \in \mathbb{N}}$ be a sequence of measures on $(X, \mathcal{A})$. Let $(x_n)_{n \in \mathbb{N}} \subseteq [0, \infty]$. Then the set function

$$\mu := \sum_{n \in \mathbb{N}} x_n \mu_n$$

defined by

$$\mu(A) := \sum_{n \in \mathbb{N}} x_n \mu_n(A), \quad \text{for all } A \in \mathcal{A},$$

is a measure on $(X, \mathcal{A})$

Proof. We verify the axioms of a measure:

(M1) *(Null empty set):* For all $n \in \mathbb{N}$, $\mu_n(\emptyset) = 0$, so

$$\mu(\emptyset) = \sum_{n \in \mathbb{N}} x_n \mu_n(\emptyset) = \sum_{n \in \mathbb{N}} x_n \cdot 0 = 0.$$

(M2) *(Countable additivity):* Let $(A_k)_{k \in \mathbb{N}} \subseteq \mathcal{A}$ be pairwise disjoint. Since each μ_n is a measure, we have

$$\mu_n \left(\bigsqcup_{k \in \mathbb{N}} A_k \right) = \sum_{k \in \mathbb{N}} \mu_n(A_k), \quad \text{for all } n \in \mathbb{N}.$$

Then,

$$\mu \left(\bigsqcup_{k \in \mathbb{N}} A_k \right) = \sum_{n \in \mathbb{N}} x_n \mu_n \left(\bigsqcup_{k \in \mathbb{N}} A_k \right) = \sum_{n \in \mathbb{N}} x_n \sum_{k \in \mathbb{N}} \mu_n(A_k).$$

Since all terms are non-negative, we may exchange the order of summation:

$$\sum_{n \in \mathbb{N}} x_n \sum_{k \in \mathbb{N}} \mu_n(A_k) = \sum_{k \in \mathbb{N}} \sum_{n \in \mathbb{N}} x_n \mu_n(A_k) = \sum_{k \in \mathbb{N}} \mu(A_k).$$

Therefore, μ is countably additive. $\square$

Example 2.5 (Restriction of a measure). Let $(X, \mathcal{A}, \mu)$ be a measure space and let $A \in \mathcal{A}$. Define the set function $\mu_A : \mathcal{A} \to [0, \infty]$ by

$$\mu_A(B) := \mu(A \cap B), \quad \text{for all } B \in \mathcal{A}.$$

Then μ_A is a measure on $(X, \mathcal{A})$, called the *restriction of μ to A*.

Proof. We verify the two defining properties of a measure:
  (M1): $\mu_A(\emptyset) = \mu(A \cap \emptyset) = \mu(\emptyset) = 0$.

  (M2): Let $(B_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$ be pairwise disjoint. Then $(A \cap B_n)_{n \in \mathbb{N}}$ are also pairwise disjoint, and

$$\mu_A\left(\bigsqcup_{n \in \mathbb{N}} B_n\right) = \mu\left(A \cap \bigsqcup_{n \in \mathbb{N}} B_n\right) = \mu\left(\bigsqcup_{n \in \mathbb{N}} (A \cap B_n)\right) = \sum_{n \in \mathbb{N}} \mu(A \cap B_n) = \sum_{n \in \mathbb{N}} \mu_A(B_n).$$

Hence, μ_A is a measure. $\qquad\square$

Definition 2.5 (Lebesgue measure on $\mathbb{R}^n$). Define the set function λ_n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ by

$$\lambda_n\left([\![a, b[\![\right) := \prod_{i=1}^{n} (b_i - a_i),$$

for all $[\![a, b[\![:= [a_1, b_1) \times \cdots \times [a_n, b_n) \in \mathcal{J}_n$. This is called the n-dimensional Lebesgue measure.

Remark 2.3. The set function λ_n is defined only on the family $\mathcal{J}_n$ of half-open rectangles and hence is not yet a measure. Extending λ_n to a measure on $\mathcal{B}(\mathbb{R}^n)$ requires the Carathéodory extension theorem, which will be developed later.

Lemma 2.3. Let $(X, \mathcal{A})$ be a measure space, and let $\mu : \mathcal{A} \to [0, \infty]$ be an additive set function with $\mu(\emptyset) = 0$. Then μ is a measure if and only if it is **continuous from below**, i.e., for every increasing sequence $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$ with $A_n \uparrow A$, we have

$$\mu(A) = \lim_{n \to \infty} \mu(A_n) = \sup_{n \in \mathbb{N}} \mu(A_n).$$

Proof. Any measure μ is continuous from below.
  Conversely, suppose μ is finitely additive, $\mu(\emptyset) = 0$, and μ is continuous from below. Let $(B_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$ be disjoint, and define $A_n := \bigcup_{i=1}^{n} B_i$. Then (A_n) is increasing with $\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} B_n$. By finite additivity,

$$\mu(A_n) = \sum_{i=1}^{n} \mu(B_i),$$

and by continuity from below,

$$\mu\left(\bigsqcup_{n=1}^{\infty} B_n\right) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu(B_n).$$

Hence μ is countably additive, i.e., a measure. $\qquad\square$

Lemma 2.4. Let $(X, \mathcal{A})$ be a measurable space and $\mu : \mathcal{A} \to [0, \infty)$ an additive set function with $\mu(\emptyset) = 0$ and $\mu(A) < \infty$ for all $A \in \mathcal{A}$. Then μ is a measure if and only if it satisfies one of the following continuity properties:

(i) μ is continuous from below;

(ii) μ is continuous from above;

(iii) μ is continuous at $\emptyset$, i.e., for every decreasing sequence $(B_n)_{n \in \mathbb{N}}$ in $\mathcal{A}$ with $\bigcap_{n=1}^{\infty} B_n = \emptyset$, we have
$$\lim_{n \to \infty} \mu(B_n) = 0.$$

Proof. Clearly, every measure satisfies properties (i)–(iii), so we only need to show that (iii) implies countable additivity.

Assume μ is additive, $\mu(\emptyset) = 0$, and satisfies continuity at $\emptyset$. Let $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$ be pairwise disjoint and define $A := \bigsqcup_{n \in \mathbb{N}} A_n$. For each n, let
$$B_n := A \setminus \bigcup_{i=1}^{n} A_i.$$

Then (B_n) is a decreasing sequence in $\mathcal{A}$ with $\bigcap_{n \in \mathbb{N}} B_n = \emptyset$, so by continuity at $\emptyset$, we have $\mu(B_n) \to 0$.

Using additivity, we compute
$$\mu(A) = \mu\left(B_n \sqcup \bigcup_{i=1}^{n} A_i\right) = \mu(B_n) + \sum_{i=1}^{n} \mu(A_i).$$

Taking the limit as $n \to \infty$, we get
$$\mu(A) = \lim_{n \to \infty} \left(\mu(B_n) + \sum_{i=1}^{n} \mu(A_i)\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Thus, μ is countably additive, hence a measure. $\qquad\square$

3 Uniqueness of Measures

Definition 3.1. A *Dynkin system* (or λ-system) $\mathcal{D} \subseteq \mathcal{P}(X)$ is a collection of subsets of X such that:

- $X \in \mathcal{D}$ **(D1)**

- If $D \in \mathcal{D}$, then $D^c \in \mathcal{D}$ **(D2)**

- If $(D_n)_{n \in \mathbb{N}} \subseteq \mathcal{D}$ are pairwise disjoint, then $\bigsqcup_{n \in \mathbb{N}} D_n \in \mathcal{D}$ **(D3)**

Remark 3.1. As with σ-algebras one easily checks that $\emptyset \in \mathcal{D}$ and that finite disjoint unions are in $\mathcal{D}$: if $D, E \in \mathcal{D}$ with $D \cap E = \emptyset$, then $D \sqcup E \in \mathcal{D}$. Every σ-algebra is a Dynkin system, but the converse is not true in general.

Lemma 3.1. Let $\mathcal{E} \subseteq \mathcal{P}(X)$. Then there exists a smallest Dynkin system $\mathcal{D}(\mathcal{E})$ containing $\mathcal{E}$, called the *Dynkin system generated by $\mathcal{E}$*. Moreover,

$$\mathcal{E} \subseteq \mathcal{D}(\mathcal{E}) \subseteq \sigma(\mathcal{E}),$$

where $\sigma(\mathcal{E})$ denotes the σ-algebra generated by $\mathcal{E}$.

Proof. The proof is analogous to that of Theorem 1.1 for σ-algebras. Let $\mathcal{F}$ be the family of all Dynkin systems on X that contain $\mathcal{E}$. Then $\mathcal{F}$ is nonempty, since $\mathcal{P}(X)$ is a Dynkin system containing $\mathcal{E}$. Define

$$\mathcal{D}(\mathcal{E}) := \bigcap_{\mathcal{D} \in \mathcal{F}} \mathcal{D}.$$

Then $\mathcal{D}(\mathcal{E})$ is a Dynkin system, being the intersection of Dynkin systems (which are closed under complements, disjoint unions, and contain X). Moreover, it is the smallest such system containing $\mathcal{E}$ by construction. Since every σ-algebra is in particular a Dynkin system, we also have

$$\mathcal{D}(\mathcal{E}) \subseteq \sigma(\mathcal{E}).$$

$\square$

Lemma 3.2. A Dynkin system $\mathcal{D}$ is a σ-algebra if and only if it is closed under finite intersections; that is,

$$D, E \in \mathcal{D} \quad \Rightarrow \quad D \cap E \in \mathcal{D}.$$

Proof. The "only if" direction follows immediately from Remark 3.1 and the fact that every σ-algebra is closed under finite intersections.

For the converse, assume $\mathcal{D}$ is a Dynkin system closed under finite intersections. Let $(D_n)_{n \in \mathbb{N}} \subseteq \mathcal{D}$, and define

$$E_1 := D_1 \in \mathcal{D}, \quad E_{n+1} := D_{n+1} \setminus \bigcup_{k=1}^{n} D_k = D_{n+1} \cap \bigcap_{k=1}^{n} D_k^c.$$

Each $E_n \in \mathcal{D}$ by the Dynkin properties and the assumed stability under finite intersections. The sets (E_n) are disjoint and satisfy

$$\bigcup_{n=1}^{\infty} D_n = \bigsqcup_{n=1}^{\infty} E_n \in \mathcal{D},$$

so $\mathcal{D}$ is closed under countable unions. Hence, $\mathcal{D}$ is a σ-algebra. $\square$

While Lemma 3.2 characterizes when a Dynkin system is a σ-algebra, it is not directly applicable when the Dynkin system $\mathcal{D}$ is defined via a generator $\mathcal{E} \subseteq \mathcal{P}(X)$, as is often the case in practice. The following theorem overcomes this limitation and plays a central role in many applications.

Theorem 3.3 (Dynkin's π-λ Theorem). Let $\mathcal{E} \subseteq \mathcal{P}(X)$ be a collection of sets that is closed under finite intersections. Then,

$$\mathcal{D}(\mathcal{E}) = \sigma(\mathcal{E}).$$

Proof. By Lemma 3.1, we have $\mathcal{D}(\mathcal{E}) \subseteq \sigma(\mathcal{E})$. To show equality, it suffices to prove that $\mathcal{D}(\mathcal{E})$ is a σ-algebra. Since it contains $\mathcal{E}$, it would then contain $\sigma(\mathcal{E})$ by minimality.

By Lemma 3.2, it is enough to show that $\mathcal{D}(\mathcal{E})$ is closed under finite intersections.

Fix $D \in \mathcal{D}(\mathcal{E})$, and define

$$\mathcal{D}_D := \{A \subseteq X : A \cap D \in \mathcal{D}(\mathcal{E})\}.$$

We claim that $\mathcal{D}_D$ is a Dynkin system:

(D1): Since $D = X \cap D \in \mathcal{D}(\mathcal{E})$, we have $X \in \mathcal{D}_D$.

(D2): If $A \in \mathcal{D}_D$, then

$$A^c \cap D = ((A \cap D) \sqcup D^c)^c \cap D \in \mathcal{D}(\mathcal{E}),$$

using that $A \cap D \in \mathcal{D}(\mathcal{E})$, $D^c \in \mathcal{D}(\mathcal{E})$, and that Dynkin systems are closed under disjoint unions and complements.

(D3): Let $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{D}_D$ be disjoint. Then the sets $A_n \cap D \in \mathcal{D}(\mathcal{E})$ are disjoint, and

$$\left(\bigsqcup_{n=1}^{\infty} A_n \right) \cap D = \bigsqcup_{n=1}^{\infty} (A_n \cap D) \in \mathcal{D}(\mathcal{E}).$$

Thus, $\mathcal{D}_D$ is a Dynkin system. Since $\mathcal{E} \subseteq \mathcal{D}_G$ for all $G \in \mathcal{E}$ by the assumed $\cap$-stability of $\mathcal{E}$, and each $\mathcal{D}_G$ is a Dynkin system, it follows that

$$\mathcal{D}(\mathcal{E}) \subseteq \mathcal{D}_G \quad \text{for all } G \in \mathcal{E}.$$

Hence, for all $D \in \mathcal{D}(\mathcal{E})$ and $G \in \mathcal{E}$, we have $D \cap G \in \mathcal{D}(\mathcal{E})$, i.e., $\mathcal{D}(\mathcal{E})$ is closed under finite intersections.

By Lemma 3.2, we conclude that $\mathcal{D}(\mathcal{E})$ is a σ-algebra. Since $\mathcal{D}(\mathcal{E}) \subseteq \sigma(\mathcal{E})$ and both are σ-algebras containing $\mathcal{E}$, we have

$$\mathcal{D}(\mathcal{E}) = \sigma(\mathcal{E}).$$

$\square$

Theorem 3.4 (Uniqueness of Measures). Let $(X, \mathcal{A})$ be a measurable space with $\mathcal{A} = \sigma(\mathcal{E})$, where $\mathcal{E} \subseteq \mathcal{P}(X)$ satisfies:

- $\mathcal{E}$ is closed under finite intersections;

- there exists an increasing sequence $(E_n)_{n \in \mathbb{N}} \subseteq \mathcal{E}$ with $E_n \uparrow X$.

Suppose μ and ν are measures on $\mathcal{A}$ such that $\mu(E) = \nu(E)$ for all $E \in \mathcal{E}$, and $\mu(E_n) = \nu(E_n) < \infty$ for all $n \in \mathbb{N}$. Then $\mu = \nu$ on $\mathcal{A}$; that is,

$$\mu(A) = \nu(A) \quad \text{for all } A \in \mathcal{A}.$$

Proof. Fix $n \in \mathbb{N}$, and define

$$\mathcal{D}_n := \{A \in \mathcal{A} : \mu(E_n \cap A) = \nu(E_n \cap A)\}.$$

We claim that $\mathcal{D}_n$ is a Dynkin system:

(D1): Since $E_n \in \mathcal{E} \subseteq \mathcal{A}$, and $\mu(E_n) = \nu(E_n)$, it follows that $X \in \mathcal{D}_n$.

(D2): If $A \in \mathcal{D}_n$, then

$$\mu(E_n \cap A^c) = \mu(E_n) - \mu(E_n \cap A) = \nu(E_n) - \nu(E_n \cap A) = \nu(E_n \cap A^c),$$

so $A^c \in \mathcal{D}_n$.

(D3): Let $(A_k)_{k \in \mathbb{N}} \subseteq \mathcal{D}_n$ be disjoint. Then:

$$\mu \left(E_n \cap \bigsqcup_{k=1}^{\infty} A_k \right) = \sum_{k=1}^{\infty} \mu(E_n \cap A_k) = \sum_{k=1}^{\infty} \nu(E_n \cap A_k) = \nu \left(E_n \cap \bigsqcup_{k=1}^{\infty} A_k \right),$$

so $\bigsqcup_{k=1}^{\infty} A_k \in \mathcal{D}_n$.

Thus, $\mathcal{D}_n$ is a Dynkin system. Since $\mathcal{E} \subseteq \mathcal{D}_n$ (as $\mu(E_n \cap E) = \nu(E_n \cap E)$ for all $E \in \mathcal{E}$, by the $\cap$-stability of $\mathcal{E}$), and since $\sigma(\mathcal{E}) = \mathcal{A}$, Theorem 3.3 yields

$$\mathcal{A} = \sigma(\mathcal{E}) = \mathcal{D}(\mathcal{E}) \subseteq \mathcal{D}_n.$$

Hence,
$$\mu(E_n \cap A) = \nu(E_n \cap A) \quad \text{for all } A \in \mathcal{A}, \ n \in \mathbb{N}.$$

Now fix $A \in \mathcal{A}$. Since $E_n \uparrow X$, we have $E_n \cap A \uparrow A$, and by continuity from below,

$$\mu(A) = \lim_{n \to \infty} \mu(E_n \cap A) = \lim_{n \to \infty} \nu(E_n \cap A) = \nu(A).$$

Therefore, $\mu = \nu$ on $\mathcal{A}$. $\qquad \square$

Theorem 3.5 (Translation Invariance and Uniqueness of Lebesgue Measure)**.** Let λ^n denote the n-dimensional Lebesgue measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Then:

(i) **(Translation invariance)** For all $x \in \mathbb{R}^n$ and all $B \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\lambda^n(x + B) = \lambda^n(B),$$

where $x + B := \{x + y : y \in B\}$ is the translation of B by x.

(ii) **(Uniqueness up to scalar)** Let μ be a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ that is translation invariant and finite on the unit cube:

$$\mu(x + B) = \mu(B) \quad \text{for all } x \in \mathbb{R}^n, \ B \in \mathcal{B}(\mathbb{R}^n), \quad \text{and} \quad \mu([0,1)^n) < \infty.$$

Then μ is a scalar multiple of Lebesgue measure:

$$\mu = \mu([0,1)^n) \cdot \lambda^n.$$

4 Existence of Measures

Definition 4.1 (Semi-ring)**.** Let X be a set. A family $\mathcal{S} \subseteq \mathcal{P}(X)$ is called a *semi-ring* if:

- $\emptyset \in \mathcal{S}$ (S1)

- For all $S, T \in \mathcal{S}$, we have $S \cap T \in \mathcal{S}$ (S2)

- For all $S, T \in \mathcal{S}$, there exist disjoint sets $S_1, \ldots, S_M \in \mathcal{S}$ such that

$$S \setminus T = \bigsqcup_{i=1}^{M} S_i \qquad (S3)$$

Theorem 4.1 (Carathéodory Extension Theorem)**.** Let $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-ring and let $\mu : \mathcal{S} \to [0, \infty]$ be a pre-measure, i.e.,

- $\mu(\emptyset) = 0$,

- For every sequence $(S_n)_{n \in \mathbb{N}} \subseteq \mathcal{S}$ of disjoint sets with $\bigsqcup_{n \in \mathbb{N}} S_n \in \mathcal{S}$, we have

$$\mu \left(\bigsqcup_{n \in \mathbb{N}} S_n \right) = \sum_{n \in \mathbb{N}} \mu(S_n).$$

Then μ has an extension to a measure $\bar{\mu}$ on $\sigma(\mathcal{S})$.

Moreover, if $\mathcal{S}$ contains an increasing sequence $(S_n)_{n \in \mathbb{N}}$ with $S_n \uparrow X$ and $\mu(S_n) < \infty$ for all n, then the extension is unique.

Idea of the proof. The fundamental problem is how to extend the pre-measure μ. The following auxiliary set function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ will play a central role. For any $A \subseteq X$, define the family of countable $\mathcal{S}$-coverings

$$\mathcal{C}(A) := \left\{ (S_n)_{n \in \mathbb{N}} \subseteq \mathcal{S} : A \subseteq \bigcup_{n \in \mathbb{N}} S_n \right\},$$

and the set function

$$\mu^*(A) := \inf \left\{ \sum_{n \in \mathbb{N}} \mu(S_n) : (S_n)_{n \in \mathbb{N}} \in \mathcal{C}(A) \right\}.$$

If A cannot be covered by sets from $\mathcal{S}$, we define $\mathcal{C}(A) = \emptyset$ and hence $\mu^*(A) := \inf \emptyset = \infty$.

The proof proceeds in four main steps:

1. **(Outer measure)** Show that μ^* is an outer measure, i.e., it satisfies:

$$\begin{aligned}
&\text{(OM1)} \quad \mu^*(\emptyset) = 0, \\
&\text{(OM2)} \quad A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B), \\
&\text{(OM3)} \quad \mu^* \left(\bigcup_{n \in \mathbb{N}} A_n \right) \leq \sum_{n \in \mathbb{N}} \mu^*(A_n).
\end{aligned}$$

2. **(Extension)** Show that μ^* extends μ, i.e., $\mu^*(S) = \mu(S)$ for all $S \in \mathcal{S}$.

3. **(μ^*-measurable sets)** Define the collection of μ^*-measurable sets by

$$\mathcal{A}_{\mu^*} := \{A \subseteq X : \mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \setminus A) \quad \text{for all } Q \subseteq X\}.$$

Then $\mathcal{A}_{\mu^*}$ is a σ-algebra with $\mathcal{S} \subseteq \mathcal{A}_{\mu^*}$ and $\sigma(\mathcal{S}) \subseteq \mathcal{A}_{\mu^*}$.

4. **(Measure on σ-algebra)** The restriction of μ^* to $\mathcal{A}_{\mu^*}$ is a measure. In particular, $\mu^*|_{\sigma(\mathcal{S})}$ is a measure extending μ.

If $\mathcal{S}$ contains an increasing sequence $(S_n)_{n \in \mathbb{N}}$ with $S_n \uparrow X$ and $\mu(S_n) < \infty$ for all n, then the extension is unique. $\qquad \square$

Existence of Lebesgue Measure on $\mathbb{R}$

Lemma 4.2. Let $\mathcal{J}_1 := \{[a, b) \subseteq \mathbb{R} : a < b\}$ be the family of half-open intervals. Define the set function

$$\lambda_1([a, b)) := b - a \quad \text{for all } [a, b) \in \mathcal{J}_1.$$

Then $\lambda_1 : \mathcal{J}_1 \to [0, \infty)$ is a pre-measure.

Proof. Let $[a, b) \in \mathcal{J}_1$, and suppose it can be written as a disjoint union of intervals:

$$[a, b) = \bigsqcup_{n \in \mathbb{N}} I_n, \quad \text{with } I_n \in \mathcal{J}_1 \text{ for all } n.$$

Our goal is to show that

$$\lambda_1([a, b)) = \sum_{n=1}^{\infty} \lambda_1(I_n).$$

Fix $\varepsilon > 0$. For each $n \in \mathbb{N}$, choose a closed interval $I_n^{(\varepsilon)}$ such that

$$I_n \subseteq I_n^{(\varepsilon)} \quad \text{and} \quad \lambda_1(I_n^{(\varepsilon)}) \leq \lambda_1(I_n) + \frac{\varepsilon}{2^n}.$$

These intervals slightly extend each I_n, allowing us to approximate the union $\bigsqcup I_n$ from above.

Since the I_n cover $[a, b)$ disjointly, the union of the extended intervals will eventually cover most of $[a, b)$. More precisely, for sufficiently large N, we have

$$[a, b - \varepsilon) \subseteq \bigcup_{n=1}^{N} I_n^{(\varepsilon)}.$$

Now we estimate the difference:

$$\lambda_1([a, b)) - \sum_{n=1}^{N} \lambda_1(I_n) = (\lambda_1([a, b)) - \lambda_1([a, b - \varepsilon)))$$

$$+ \left(\lambda_1([a, b - \varepsilon)) - \sum_{n=1}^{N} \lambda_1(I_n^{(\varepsilon)}) \right)$$

$$+ \sum_{n=1}^{N} \left(\lambda_1(I_n^{(\varepsilon)}) - \lambda_1(I_n) \right)$$

$$\leq \varepsilon + 0 + \sum_{n=1}^{N} \frac{\varepsilon}{2^n} \leq 2\varepsilon.$$

14

On the other hand, since $\bigsqcup_{n=1}^{N} I_n \subseteq [a, b)$ and the intervals I_n are disjoint, finite additivity and monotonicity of λ_1 imply:

$$\sum_{n=1}^{N} \lambda_1(I_n) = \lambda_1\left(\bigsqcup_{n=1}^{N} I_n\right) \leq \lambda_1([a, b)).$$

Therefore,

$$0 \leq \lambda_1([a, b)) - \sum_{n=1}^{N} \lambda_1(I_n),$$

which justifies the lower bound in the previous inequality.

Combining both sides, we have

$$0 \leq \lambda_1([a, b)) - \sum_{n=1}^{N} \lambda_1(I_n) \leq 2\varepsilon.$$

Letting $N \to \infty$ and then $\varepsilon \to 0$, we conclude:

$$\lambda_1([a, b)) = \sum_{n=1}^{\infty} \lambda_1(I_n).$$

Thus, λ_1 is countably additive on $\mathcal{J}_1$, and hence a pre-measure. $\qquad\square$

Lemma 4.3 (Lebesgue measure on $\mathbb{R}$)**.** The set function λ_1, defined on $\mathcal{J}_1$ by $\lambda_1([a, b)) = b - a$ for $a < b$, extends to a measure on $\mathcal{B}(\mathbb{R})$. This extension is the unique measure μ on $\mathcal{B}(\mathbb{R})$ such that

$$\mu([a, b)) = b - a \quad \text{for all } a < b.$$

Proof. We have already shown that λ_1 is a pre-measure on $\mathcal{J}_1$. By Theorem 1.3, $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{J}_1)$, i.e., the Borel σ-algebra is generated by $\mathcal{J}_1$.

Consider the sequence of half-open intervals $[-k, k) \subseteq \mathbb{R}$ for $k \in \mathbb{N}$. This forms an increasing sequence for $\mathbb{R}$, and we have

$$\lambda_1([-k, k)) = 2k < \infty \quad \text{for all } k \in \mathbb{N}.$$

Thus, all the conditions of Theorem 4.1 (Carathéodory's extension theorem) are satisfied. It follows that λ_1 extends uniquely to a measure on $\mathcal{B}(\mathbb{R})$, yielding the one-dimensional Lebesgue measure on $\mathbb{R}$. $\qquad\square$

Existence of Lebesgue Measure on $\mathbb{R}^n$

Lemma 4.4. Let $\mathcal{J}_n$ denote the collection of half-open rectangles in $\mathbb{R}^n$ of the form

$$[\![a, b)\!] = \prod_{i=1}^{n} [a_i, b_i), \quad \text{where } a = (a_1, \dots, a_n), \ b = (b_1, \dots, b_n), \ a_i < b_i.$$

Then $\mathcal{J}_n$ is a semi-ring.

Proof. We prove the statement by induction on n. Assume $\mathcal{J}_n \subset \mathbb{R}^n$ is a semi-ring. Define

$$\mathcal{J}_{n+1} := \mathcal{J}_n \times \mathcal{J}_1,$$

i.e., the collection of rectangles of the form $R = R_n \times R_1$, where $R_n \in \mathcal{J}_n$ and $R_1 \in \mathcal{J}_1$.

We verify the properties of a semi-ring:

(S1) *Closure under the empty set:* Since $\emptyset \in \mathcal{J}_n$ and $\mathcal{J}_1$, we have

$$\emptyset = \emptyset \times [a, b) \in \mathcal{J}_{n+1}.$$

(S2) *Closure under intersection:* Let $R = R_n \times R_1$ and $S = S_n \times S_1$ be in $\mathcal{J}_{n+1}$. Then

$$R \cap S = (R_n \cap S_n) \times (R_1 \cap S_1),$$

which belongs to $\mathcal{J}_{n+1}$, since both $R_n \cap S_n \in \mathcal{J}_n$ and $R_1 \cap S_1 \in \mathcal{J}_1$, by the inductive hypothesis.

(S3) *Closure under set difference (finite disjoint union):* Consider

$$R \setminus S = (R_n \times R_1) \setminus (S_n \times S_1).$$

This set can be decomposed as

$$(R_n \setminus S_n) \times (R_1 \setminus S_1) \sqcup (R_n \cap S_n) \times (R_1 \setminus S_1) \sqcup (R_n \setminus S_n) \times (R_1 \cap S_1).$$

Each of the components $R_n \setminus S_n$, $R_n \cap S_n$, $R_1 \setminus S_1$, and $R_1 \cap S_1$ can be written as finite disjoint unions of sets in $\mathcal{J}_n$ and $\mathcal{J}_1$, respectively. Therefore, their Cartesian products yield finite disjoint unions of elements in $\mathcal{J}_{n+1}$.

Hence, $\mathcal{J}_{n+1}$ is a semi-ring. By induction, it follows that $\mathcal{J}_n$ is a semi-ring for all $n \in \mathbb{N}$. $\qquad\square$

Lemma 4.5. The function $\lambda_n \colon \mathcal{J}_n \to [0, \infty)$, defined by

$$\lambda_n\Big([a_1, b_1) \times \cdots \times [a_n, b_n)\Big) = \prod_{i=1}^{n}(b_i - a_i),$$

is a pre-measure on the semi-ring $\mathcal{J}_n$.

Corollary 4.5.1 (Lebesgue measure on $\mathbb{R}^n$)**.** The set function λ_n extends to a measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$, called the *Lebesgue measure.* It is the unique measure satisfying

$$\lambda_n\left([a_1, b_1) \times \cdots \times [a_n, b_n)\right) = \prod_{i=1}^{n}(b_i - a_i), \quad \text{for all } a_i < b_i.$$

Remark 4.1 (Relation to Elementary Volume)**.** The uniqueness of Lebesgue measure and its properties imply that Lebesgue measure coincides with the familiar volume function $\mathrm{vol}^{(n)}(\,\cdot\,)$ from elementary geometry. More precisely, $\mathrm{vol}^{(n)}$ can be extended to a measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ in only one way namely, as the Lebesgue measure λ^n.

5 Measurable Mappings

Definition 5.1 (Measurable Map)**.** Let $(X, \mathcal{A})$, $(X', \mathcal{A}')$ be measurable spaces. A map $T : X \to X'$ is called $\mathcal{A}/\mathcal{A}'$-measurable (or simply measurable) if the pre-image of every measurable set is measurable:

$$T^{-1}(A') \in \mathcal{A} \quad \text{for all } A' \in \mathcal{A}'.$$

Remark 5.1.

- Probabilists often refer to a measurable map defined on a probability space as a *random variable.*

- The symbolic notation $T^{-1}(\mathcal{A}') := \{T^{-1}(A') : A' \in \mathcal{A}'\}$ is often used. We also write $T^{-1}(\mathcal{A}') \subset \mathcal{A}$ as shorthand for measurability.

- It is common to write $T : (X, \mathcal{A}) \to (X', \mathcal{A}')$ to indicate that T is measurable.

- A measurable map between $\mathcal{B}(\mathbb{R}^n)$ and $\mathcal{B}(\mathbb{R}^m)$ is often called a *Borel measurable map.*

Example 5.1. Let $(X, \mathcal{A})$ be a measurable space and let $A \in \mathcal{A}$. We show that the indicator function

$$\mathbf{1}_A : X \to \{0, 1\}, \quad \mathbf{1}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

is $\mathcal{A}/\mathcal{P}(\{0, 1\})$-measurable.

We check that the preimage of each subset of $\{0, 1\}$ lies in $\mathcal{A}$:

- $\mathbf{1}_A^{-1}(\emptyset) = \emptyset \in \mathcal{A}$,

- $\mathbf{1}_A^{-1}(\{0\}) = A^c \in \mathcal{A}$,

- $\mathbf{1}_A^{-1}(\{1\}) = A \in \mathcal{A}$,

- $\mathbf{1}_A^{-1}(\{0, 1\}) = X \in \mathcal{A}$.

Therefore, $\mathbf{1}_A$ is measurable.

Lemma 5.1. Let $(X, \mathcal{A})$, $(X', \mathcal{A}')$ be measurable spaces, and suppose $\mathcal{A}' = \sigma(\mathcal{E}')$. Then a map $T : X \to X'$ is $\mathcal{A}/\mathcal{A}'$-measurable if and only if $T^{-1}(\mathcal{E}') \subseteq \mathcal{A}$, i.e. if

$$T^{-1}(E') \in \mathcal{A} \quad \text{for all } E' \in \mathcal{E}'.$$

Proof. If T is measurable, then by definition $T^{-1}(A') \in \mathcal{A}$ for all $A' \in \mathcal{A}'$. Since $\mathcal{E}' \subset \mathcal{A}'$, it follows immediately that $T^{-1}(E') \in \mathcal{A}$ for all $E' \in \mathcal{E}'$.

Conversely, suppose $T^{-1}(E') \in \mathcal{A}$ for every $E' \in \mathcal{E}'$. Define

$$\mathcal{D}' := \{A' \subseteq X' : T^{-1}(A') \in \mathcal{A}\}.$$

By assumption, $\mathcal{E}' \subseteq \mathcal{D}'$. We now show that $\mathcal{D}'$ is a σ-algebra:

- Since $T^{-1}(X') = X \in \mathcal{A}$, we have $X' \in \mathcal{D}'$.

- If $A' \in \mathcal{D}'$, then $T^{-1}(A'^c) = T^{-1}(A')^c \in \mathcal{A}$, so $A'^c \in \mathcal{D}'$.

- If $A_1', A_2', \cdots \in \mathcal{D}'$, then

$$T^{-1}\left(\bigcup_{i=1}^{\infty} A_i'\right) = \bigcup_{i=1}^{\infty} T^{-1}(A_i') \in \mathcal{A},$$

hence $\bigcup_{i=1}^{\infty} A_i' \in \mathcal{D}'$.

Thus, $\mathcal{D}'$ is a σ-algebra containing $\mathcal{E}'$, so it contains $\sigma(\mathcal{E}') = \mathcal{A}'$. Therefore, $T^{-1}(A') \in \mathcal{A}$ for all $A' \in \mathcal{A}'$, i.e., T is measurable. $\qquad\square$

Example 5.2. Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a continuous function. Then T is $\mathcal{B}(\mathbb{R}^m)/\mathcal{B}(\mathbb{R}^n)$-measurable.

Indeed, from elementary analysis, we know that T is continuous if and only if

$$T^{-1}(A') \subset \mathbb{R}^m \text{ is open for every open set } A' \subset \mathbb{R}^n.$$

Since the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ is generated by the open sets $\mathcal{O}_{\mathbb{R}^n}$, it follows that

$$T^{-1}(\mathcal{O}_{\mathbb{R}^n}) \subset \mathcal{O}_{\mathbb{R}^m} \subset \sigma(\mathcal{O}_{\mathbb{R}^m}) = \mathcal{B}(\mathbb{R}^m).$$

Hence, by Lemma 5.1, T is $\mathcal{B}(\mathbb{R}^m)/\mathcal{B}(\mathbb{R}^n)$-measurable.

Theorem 5.2. Let $(X_i, \mathcal{A}_i)$, $i = 1, 2, 3$, be measurable spaces, and let

$$T : X_1 \to X_2, \quad S : X_2 \to X_3$$

be $\mathcal{A}_1/\mathcal{A}_2$- and $\mathcal{A}_2/\mathcal{A}_3$-measurable maps, respectively. Then the composition

$$S \circ T : X_1 \to X_3$$

is $\mathcal{A}_1/\mathcal{A}_3$-measurable.

Proof. Let $A_3 \in \mathcal{A}_3$. Then

$$(S \circ T)^{-1}(A_3) = T^{-1}\left(S^{-1}(A_3)\right).$$

Since S is $\mathcal{A}_2/\mathcal{A}_3$-measurable, we have $S^{-1}(A_3) \in \mathcal{A}_2$. Since T is $\mathcal{A}_1/\mathcal{A}_2$-measurable, it follows that $T^{-1}(S^{-1}(A_3)) \in \mathcal{A}_1$. Therefore, $S \circ T$ is $\mathcal{A}_1/\mathcal{A}_3$-measurable. $\qquad\square$

Remark 5.2. Given a map $T : X \to X'$, where X' carries a natural σ-algebra $\mathcal{A}'$ (e.g., $\mathcal{B}(\mathbb{R})$), but no σ-algebra is specified on X, one may ask: is there a smallest σ-algebra on X that makes T measurable?

While $\mathcal{P}(X)$ trivially works, it is too large to be useful. On the other hand, $T^{-1}(\mathcal{A}')$ is a σ-algebra, and removing any set from it may break measurability. This leads to the following definition.

Definition 5.2. Let $(T_i)_{i \in I}$, with $T_i : X \to X_i$, be an arbitrary family of mappings from the same space X into measurable spaces $(X_i, \mathcal{A}_i)$. The smallest σ-algebra on X that makes all T_i simultaneously measurable is given by

$$\sigma(T_i : i \in I) := \sigma\left(\bigcup_{i \in I} T_i^{-1}(\mathcal{A}_i)\right).$$

We say that $\sigma(T_i : i \in I)$ is the σ-algebra generated by the family $(T_i)_{i \in I}$.

Although each $T_i^{-1}(\mathcal{A}_i)$ is a σ-algebra, the union $\bigcup_{i \in I} T_i^{-1}(\mathcal{A}_i)$ is, in general, not a σ-algebra if $\#I > 1$; this is why we must take the σ-hull in the definition above.

Theorem 5.3. Let $(X, \mathcal{A})$, $(X', \mathcal{A}')$ be measurable spaces and let $T : X \to X'$ be an $\mathcal{A}/\mathcal{A}'$-measurable map. For every measure μ on $(X, \mathcal{A})$,

$$\mu'(A') := \mu(T^{-1}(A')), \quad A' \in \mathcal{A}'$$

defines a measure μ' on $(X', \mathcal{A}')$.

Proof. If $A' = \emptyset$, then $T^{-1}(\emptyset) = \emptyset$ and $\mu'(\emptyset) = \mu(\emptyset) = 0$.

Let $(A'_n)_{n \in \mathbb{N}} \subset \mathcal{A}'$ be a sequence of pairwise disjoint sets. Then

$$\mu'\left(\bigsqcup_{n \in \mathbb{N}} A'_n\right) = \mu\left(T^{-1}\left(\bigsqcup_{n \in \mathbb{N}} A'_n\right)\right) = \mu\left(\bigsqcup_{n \in \mathbb{N}} T^{-1}(A'_n)\right) = \sum_{n \in \mathbb{N}} \mu\left(T^{-1}(A'_n)\right) = \sum_{n \in \mathbb{N}} \mu'(A'_n).$$

Hence, μ' is a measure on $(X', \mathcal{A}')$. $\qquad\square$

Definition 5.3. The measure $\mu'(\cdot)$ from Theorem 7.6 is called the *image measure* or *pushforward* of μ under T. It is commonly denoted by one of the following:

- $T(\mu)(\cdot)$,

- $T_*\mu(\cdot)$,

- $\mu \circ T^{-1}(\cdot)$.

Example 5.3. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\xi : \Omega \to \mathbb{R}$ be a random variable, i.e., an $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable map. Then the pushforward measure

$$\xi(\mathbb{P})(A') = \mathbb{P}(\xi^{-1}(A')) = \mathbb{P}(\{\omega \in \Omega : \xi(\omega) \in A'\}) = \mathbb{P}(\xi \in A')$$

defines a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which is called the *law* or *distribution* of the random variable ξ.

Example 5.4. Suppose we model the experiment of rolling two fair six-sided dice. The underlying probability space is given by

$$\Omega := \{(i, k) : 1 \le i, k \le 6\}, \quad \mathcal{A} := \mathcal{P}(\Omega), \quad \mathbb{P}(\{(i, k)\}) := \frac{1}{36}.$$

Each outcome (i, k) represents the result of the first and second die, respectively. Define the map

$$\xi : \Omega \to \{2, 3, \ldots, 12\}, \quad \xi(i, k) := i + k,$$

which assigns to each outcome the total number of points rolled. This function ξ is measurable and thus a random variable.

The pushforward measure $\xi(\mathbb{P})$, also called the *law* or *distribution* of ξ, gives the probabilities of the possible total number of points. For instance,

$$\xi(\mathbb{P})(7) = \mathbb{P}(\xi^{-1}(\{7\})) = \mathbb{P}(\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}) = \frac{6}{36} = \frac{1}{6}.$$

Remark 5.3. A matrix $T \in \mathbb{R}^{n \times n}$ is called orthogonal if and only if

$$T^\top T = I,$$

i.e., the transpose of T is equal to its inverse.

Orthogonal matrices preserve lengths and angles. That is, for all $x, y \in \mathbb{R}^n$, we have

$$\langle x, y \rangle = \langle Tx, Ty \rangle,$$
$$\|x\| = \|Tx\|,$$

where the standard Euclidean inner product and norm are defined by

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i, \qquad \|x\|^2 := \langle x, x \rangle.$$

Theorem 5.4. Let $T \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Then the Lebesgue measure λ^n is invariant under T, i.e.,

$$T(\lambda^n) = \lambda^n.$$

Proof. The matrix $T \in \mathbb{R}^{n \times n}$ defines a linear map $x \mapsto Tx$, i.e.,

$$T(ax + by) = aTx + bTy \quad \text{for all } a, b \in \mathbb{R}, \ x, y \in \mathbb{R}^n.$$

From the orthogonality condition, it follows that T is an isometry:

$$\|Tx - Ty\| = \|T(x - y)\| = \|x - y\|.$$

Thus, T is continuous and therefore $\mathcal{B}(\mathbb{R}^n)/\mathcal{B}(\mathbb{R}^n)$-measurable by Example 5.2. Furthermore by Theorem 5.3, the pushforward measure

$$\nu(B) := \lambda^n(T^{-1}(B))$$

is well-defined on $\mathcal{B}(\mathbb{R}^n)$.

We now show that ν is translation invariant. For any $x \in \mathbb{R}^n$ and $B \in \mathcal{B}(\mathbb{R}^n)$, we compute

$$\begin{aligned}
\nu(x + B) &= \lambda^n \left(T^{-1}(x + B) \right) \\
&= \lambda^n \left(T^{-1}x + T^{-1}B \right) \\
&= \lambda^n \left(T^{-1}B \right) = \nu(B),
\end{aligned}$$

where we used linearity of T and translation invariance of Lebesgue measure (Theorem 3.5(i)).

Hence ν is a translation-invariant measure on $\mathbb{R}^n$. By Theorem 3.5(ii), since ν is also finite on bounded sets (e.g., the unit ball), it must be a scalar multiple of Lebesgue measure:
$$\nu = c\lambda^n \quad \text{for some } c > 0.$$

To determine the constant c, consider the unit ball $B_1(0) := \{x \in \mathbb{R}^n : \|x\| < 1\}$. Since T is orthogonal, we have
$$x \in B_1(0) \iff \|x\| < 1 \iff \|Tx\| < 1 \iff x \in T^{-1}(B_1(0)),$$

so $T^{-1}(B_1(0)) = B_1(0)$. Therefore,
$$\lambda^n(B_1(0)) = \lambda^n(T^{-1}(B_1(0))) = \nu(B_1(0)) = c\lambda^n(B_1(0)),$$

which implies $c = 1$, since $0 < \lambda^n(B_1(0)) < \infty$. Thus, $\nu = \lambda^n$, and the theorem follows. $\square$

Remark 5.4. Theorem 5.4 is a special case of the following general change-of-variable formula for Lebesgue measure. Recall that a matrix $S \in \mathbb{R}^{n \times n}$ is invertible if and only if $\det S \neq 0$.

Theorem 5.5 (Change of Variables)**.** Let $S \in \mathbb{R}^{n \times n}$ be an invertible matrix. Then
$$S(\lambda^n) = |\det S^{-1}| \, \lambda^n = |\det S|^{-1} \, \lambda^n. \tag{7.7}$$

Proof. Since S is invertible, both S and S^{-1} are linear maps on $\mathbb{R}^n$, and hence continuous and measurable (by Example 5.2). Define a measure ν on $\mathbb{R}^n$ by
$$\nu(B) := \lambda^n(S^{-1}(B)), \quad B \in \mathcal{B}(\mathbb{R}^n).$$

For any $x \in \mathbb{R}^n$, we have
$$\nu(x + B) = \lambda^n(S^{-1}(x + B)) = \lambda^n(S^{-1}x + S^{-1}B) = \lambda^n(S^{-1}B) = \nu(B),$$

so ν is translation invariant.

By Theorem 3.5(ii), any translation-invariant measure finite on the unit cube must be a scalar multiple of Lebesgue measure, so there exists a constant $c > 0$ such that
$$\nu = c\lambda^n.$$

To determine c, we evaluate both sides on the unit cube $[0,1)^n$, which satisfies $\lambda^n([0,1)^n) = 1$:
$$\nu([0,1)^n) = \lambda^n\big(S^{-1}([0,1)^n)\big).$$

The set $S^{-1}([0,1)^n)$ is a parallelepiped spanned by the vectors $S^{-1}e_i$, where $(e_i)_{i=1}^n$ is the standard basis of $\mathbb{R}^n$. Its volume is given by the absolute value of the determinant:
$$\operatorname{vol}^{(n)}\big(S^{-1}([0,1)^n)\big) = |\det S^{-1}| = |\det S|^{-1}.$$

By Remark 4.1, Lebesgue measure coincides with this volume on Borel sets, so
$$\nu([0,1)^n) = \lambda^n\big(S^{-1}([0,1)^n)\big) = |\det S|^{-1}.$$

Hence,
$$\nu = |\det S|^{-1}\lambda^n,$$

which completes the proof. $\square$

Definition 5.4. A *motion* in $\mathbb{R}^n$ is a linear transformation of the form

$$Mx = \tau_x \circ T(x),$$

where $\tau_x(y) = y + x$ denotes translation by $x \in \mathbb{R}^n$, and $T \in \mathbb{R}^{n \times n}$ is an orthogonal matrix (i.e., $T^\top T = \mathrm{id}_n$). In particular, two sets are said to be *congruent* if one can be obtained from the other by a motion.

Theorem 5.6 (Invariance under Motions)**.** Lebesgue measure is invariant under motions: for any motion M in $\mathbb{R}^n$, we have

$$\lambda^n = M(\lambda^n).$$

In particular, congruent sets have the same Lebesgue measure.

Proof. By definition, any motion M can be written as $M = \tau_x \circ T$, where T is orthogonal (so $|\det T| = 1$). By Theorem 5.4,

$$T(\lambda^n) = \lambda^n.$$

Translation invariance of Lebesgue measure (Theorem 3.5(i)) gives

$$\tau_x(\lambda^n) = \lambda^n.$$

Hence,

$$M(\lambda^n) = \tau_x(T(\lambda^n)) = \tau_x(\lambda^n) = \lambda^n. \qquad \square$$

6   Measurable Functions

Definition 6.1 (Measurable Function)**.** Let $(X, \mathcal{A})$ be a measurable space. A function $u : X \to \mathbb{R}$ is called *measurable* if it is $\mathcal{A}$-$\mathcal{B}(\mathbb{R})$-measurable; that is,

$$u^{-1}(B) \in \mathcal{A} \quad \text{for all } B \in \mathcal{B}(\mathbb{R}).$$

Remark 6.1. By Lemma 5.1, a function $u : X \to \mathbb{R}$ is $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable if and only if

$$u^{-1}(G) \in \mathcal{A} \quad \text{for all } G \in \mathcal{E},$$

where $\mathcal{E}$ is a generator of $\mathcal{B}(\mathbb{R})$.

Definition 6.2 (Level Set)**.** Let $(X, \mathcal{A})$ be a measurable space, and let $u : X \to \mathbb{R}$ be a function. For any $y \in \mathbb{R}$, the *level set* of u at the value y is defined as

$$\{u = y\} := \{x \in X : u(x) = y\}.$$

More generally, we define:

- Strict upper level set: $\quad \{u > y\} := \{x \in X : u(x) > y\}$

- Strict lower level set: $\quad \{u < y\} := \{x \in X : u(x) < y\}$

- Upper level set: $\quad \{u \geq y\} := \{x \in X : u(x) \geq y\}$

- Lower level set: $\quad \{u \leq y\} := \{x \in X : u(x) \leq y\}$

Remark 6.2. As noted in Remark 1.4, the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is generated by intervals of the form $[a, \infty)$, (a, ∞), $(-\infty, a)$, or $(-\infty, a]$, with $a \in \mathbb{R}$ (or $\mathbb{Q}$). To verify that a function $u : X \to \mathbb{R}$ is measurable, it suffices to check that

$$u^{-1}([a, \infty)) = \{x \in X : u(x) \in [a, \infty)\} = \{x \in X : u(x) \geq a\} \in \mathcal{A}$$

for all such a, and likewise for the other interval types.

We write
$$\{u > v\} := \{x \in X : u(x) > v(x)\},$$
and similarly $\{u < v\}$, $\{u \leq v\}$, $\{u = v\}$, $\{u \neq v\}$, $\{u \in B\}$, etc., for measurable $u, v : X \to \mathbb{R}$ and Borel sets $B \subseteq \mathbb{R}$.

Lemma 6.1. Let $(X, \mathcal{A})$ be a measurable space. A function $u : X \to \mathbb{R}$ is $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable if and only if any one (and hence all) of the following equivalent conditions hold:

(i) $\{u > a\} \in \mathcal{A} \quad \forall a \in \mathbb{R}$ or $a \in \mathbb{Q}$ $\qquad$ (iii) $\{u < a\} \in \mathcal{A} \quad \forall a \in \mathbb{R}$ or $a \in \mathbb{Q}$

(ii) $\{u \geq a\} \in \mathcal{A} \quad \forall a \in \mathbb{R}$ or $a \in \mathbb{Q}$ $\qquad$ (iv) $\{u \leq a\} \in \mathcal{A} \quad \forall a \in \mathbb{R}$ or $a \in \mathbb{Q}$

Remark 6.3. It is often helpful to use the values $+\infty$ and $-\infty$ in calculations. To do this properly, we consider the extended real line $\overline{\mathbb{R}} := [-\infty, +\infty]$. If we agree that $-\infty < x$ and $y < +\infty$ for all $x, y \in \mathbb{R}$, then $\overline{\mathbb{R}}$ inherits the usual ordering from $\mathbb{R}$, as well as the standard rules of addition and multiplication for real numbers. The latter, however, must be augmented as shown below.

Addition in $\overline{\mathbb{R}}$

$+$	0	x	$+\infty$	$-\infty$
0	0	x	$+\infty$	$-\infty$
y	y	$x+y$	$+\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$	$+\infty$	undef.
$-\infty$	$-\infty$	$-\infty$	undef.	$-\infty$

Multiplication in $\overline{\mathbb{R}}$

$\cdot$	0	a	$+\infty$	$-\infty$
0	0	0	0	0
b	0	ab	$+\infty$	$-\infty$
$+\infty$	0	$+\infty$	$+\infty$	$-\infty$
$-\infty$	0	$-\infty$	$-\infty$	$+\infty$

Remark 6.4. Caution: The extended real line $\overline{\mathbb{R}} = [-\infty, +\infty]$ is *not* a field. Expressions such as $\infty - \infty$ or $\frac{\infty}{\infty}$ are undefined and must be avoided.

The Borel σ-algebra on $\overline{\mathbb{R}}$, denoted $\mathcal{B}(\overline{\mathbb{R}})$, is defined by

$$B^* \in \mathcal{B}(\overline{\mathbb{R}}) \quad \Longleftrightarrow \quad B^* = B \cup S,$$

for some $B \in \mathcal{B}(\mathbb{R})$ and $S \in \{\emptyset, \{-\infty\}, \{+\infty\}, \{-\infty, +\infty\}\}$.

It is straightforward to verify that $\mathcal{B}(\overline{\mathbb{R}})$ is a σ-algebra, and its trace on $\mathbb{R}$ coincides with the usual Borel σ-algebra $\mathcal{B}(\mathbb{R})$.

Lemma 6.2. The Borel σ-algebra on the extended real line $\overline{\mathbb{R}} = [-\infty, +\infty]$ satisfies

$$\mathcal{B}(\mathbb{R}) = \mathbb{R} \cap \mathcal{B}(\overline{\mathbb{R}}) \quad \text{or equivalently,} \quad \mathcal{B}(\mathbb{R}) = \{A \cap \mathbb{R} : A \in \mathcal{B}(\overline{\mathbb{R}})\}.$$

Lemma 6.3. The Borel σ-algebra $\mathcal{B}(\overline{\mathbb{R}})$ is generated by any one of the following families of sets:

$$[a, +\infty], \quad (a, +\infty], \quad [-\infty, a), \quad \text{or} \quad [-\infty, a],$$

with $a \in \mathbb{R}$ or $\mathbb{Q}$.

Proof. Let $\mathcal{E} := \sigma\big(\{[a, +\infty] : a \in \mathbb{R}\}\big)$. Since

$$[a, +\infty] = [a, +\infty) \cup \{+\infty\} \quad \text{with} \quad [a, +\infty) \in \mathcal{B}(\mathbb{R}),$$

we see that $[a, +\infty] \in \mathcal{B}(\overline{\mathbb{R}})$, hence $\mathcal{E} \subseteq \mathcal{B}(\overline{\mathbb{R}})$.

Conversely, for $-\infty < a \le b < +\infty$,

$$[a, b) = [a, +\infty] \setminus [b, +\infty] \in \mathcal{E},$$

so $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{E}$. Moreover,

$$\{+\infty\} = \bigcap_{j \in \mathbb{N}} [j, +\infty], \qquad \{-\infty\} = \bigcap_{j \in \mathbb{N}} [-\infty, -j) = \bigcap_{j \in \mathbb{N}} [-j, +\infty]^c,$$

so $\{-\infty\}, \{+\infty\} \in \mathcal{E}$. Hence for any $B \in \mathcal{B}(\mathbb{R})$,

$$B, \quad B \cup \{+\infty\}, \quad B \cup \{-\infty\}, \quad B \cup \{-\infty, +\infty\} \in \mathcal{E},$$

implying $\mathcal{B}(\overline{\mathbb{R}}) \subseteq \mathcal{E}$. Therefore, $\mathcal{B}(\overline{\mathbb{R}}) = \mathcal{E}$.

The same argument applies if the generating system uses $a \in \mathbb{Q}$, or other families like $(a, +\infty], [-\infty, a)$, or $[-\infty, a]$. $\qquad\square$

Definition 6.3. Let $(X, \mathcal{A})$ be a measurable space. We define

$$\mathcal{M} := \mathcal{M}(\mathcal{A}) \quad \text{and} \quad \mathcal{M}_{\overline{\mathbb{R}}} := \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A})$$

as the collections of real-valued and extended real-valued measurable functions, respectively:

$$\mathcal{M} = \{u : X \to \mathbb{R} \mid u \text{ is } \mathcal{A}/\mathcal{B}(\mathbb{R})\text{-measurable}\},$$
$$\mathcal{M}_{\overline{\mathbb{R}}} = \{u : X \to \overline{\mathbb{R}} \mid u \text{ is } \mathcal{A}/\mathcal{B}(\overline{\mathbb{R}})\text{-measurable}\}.$$

Example 6.1. Let $(X, \mathcal{A})$ be a measurable space. The indicator function $f(x) := \mathbf{1}_A(x)$ is measurable if and only if $A \in \mathcal{A}$.

Proof. Recall that a function $f : X \to \mathbb{R}$ is measurable if for all $\alpha \in \mathbb{R}$, the set $\{f > \alpha\} \in \mathcal{A}$. Now observe:

$$\{\mathbf{1}_A > \alpha\} = \begin{cases} \varnothing & \text{if } \alpha \geq 1, \\ A & \text{if } 0 < \alpha < 1, \\ X & \text{if } \alpha \leq 0. \end{cases}$$

Thus, $\{\mathbf{1}_A > \alpha\} \in \mathcal{A}$ for all $\alpha \in \mathbb{R}$ if and only if $A \in \mathcal{A}$, proving the claim. $\qquad\square$

Definition 6.4. Let $(X, \mathcal{A})$ be a measurable space.
   A *simple function* is a function $f : X \to \mathbb{R}$ of the form

$$f(x) = \sum_{m=1}^{M} y_m \mathbf{1}_{A_m}(x),$$

where $M \in \mathbb{N}$, $y_m \in \mathbb{R}$, and $A_1, \ldots, A_M \in \mathcal{A}$ are pairwise disjoint.
   A representation of the form

$$f(x) = \sum_{n=1}^{N} z_n \mathbf{1}_{B_n}(x),$$

with $N \in \mathbb{N}$, $z_n \in \mathbb{R}$, $B_n \in \mathcal{A}$, and $\bigsqcup_{n=1}^{N} B_n = X$, is called a *standard representation* of f.
   The set of all simple functions on $(X, \mathcal{A})$ is denoted by $\mathcal{E}$ or $\mathcal{E}(\mathcal{A})$.

Remark 6.5. Simple functions may have multiple representations; in particular, standard representations are not unique.

Example 6.2. A measurable function $h : X \to \mathbb{R}$ that attains only finitely many values is a simple function.
   Indeed, let $h(X) = \{y_0, \ldots, y_M\}$. The sets $\{h = \beta\}$ with $\beta \in h(X)$ are mutually disjoint and satisfy

$$\{h = \beta\} = \{h \leq \beta\} \setminus \{h < \beta\} \in \mathcal{A},$$

and

$$\bigcup_{\beta \in h(X)} \{h = \beta\} = X.$$

Thus, h admits the standard representation

$$h(x) = \sum_{\beta \in h(X)} \beta \, \mathbf{1}_{\{h=\beta\}}(x).$$

This shows that every measurable function with finitely many values is a simple function. Conversely, since every simple function only takes finitely many values, it always admits at least one standard representation.
   In particular, $\mathcal{E}(\mathcal{A}) \subset \mathcal{M}(\mathcal{A})$, where $\mathcal{M}(\mathcal{A})$ is the space of measurable functions.

Example 6.3. If $f, g \in \mathcal{E}(\mathcal{A})$, then

$$f \pm g \in \mathcal{E}(\mathcal{A}) \quad \text{and} \quad f \cdot g \in \mathcal{E}(\mathcal{A}).$$

Proof. Let

$$f = \sum_{m=1}^{M} y_m \, \mathbf{1}_{A_m}, \qquad g = \sum_{n=1}^{N} z_n \, \mathbf{1}_{B_n}$$

be standard representations, where $A_m, B_n \in \mathcal{A}$ are disjoint families and $y_m, z_n \in \mathbb{R}$.
Consider the intersections $A_m \cap B_n$. Then:

- Each $A_m \cap B_n \in \mathcal{A}$, since $\mathcal{A}$ is closed under intersections.

- The sets $A_m \cap B_n$ are pairwise disjoint for distinct pairs $(m, n) \neq (m', n')$.

- Their union covers X, because

$$X = \bigcup_{m=1}^{M} A_m = \bigcup_{n=1}^{N} B_n \quad \Rightarrow \quad X = \bigcup_{m=1}^{M} \bigcup_{n=1}^{N} (A_m \cap B_n).$$

On each set $A_m \cap B_n$, the functions f and g are constant with values y_m and z_n, respectively. Thus,

$$f(x) \pm g(x) = y_m \pm z_n, \qquad f(x) \cdot g(x) = y_m z_n \quad \text{for } x \in A_m \cap B_n.$$

Therefore, we obtain:

$$f \pm g = \sum_{m=1}^{M} \sum_{n=1}^{N} (y_m \pm z_n) \, \mathbf{1}_{A_m \cap B_n}, \qquad f \cdot g = \sum_{m=1}^{M} \sum_{n=1}^{N} (y_m z_n) \, \mathbf{1}_{A_m \cap B_n}.$$

These are finite sums over pairwise disjoint measurable sets, so $f \pm g$ and $f \cdot g$ are both simple functions.
Hence, $f \pm g, \ f \cdot g \in \mathcal{E}(\mathcal{A})$. $\qquad\square$

Definition 6.5. Let $f \in \mathcal{E}(\mathcal{A})$. Define its *positive part*, *negative part*, and *absolute value* by

$$f^+(x) := \max(f(x), 0), \qquad f^-(x) := \max(-f(x), 0), \qquad |f(x)| := f^+(x) + f^-(x).$$

Then $f^+, f^-, |f| \in \mathcal{E}(\mathcal{A})$, and

$$f = f^+ - f^-, \qquad |f| = f^+ + f^-.$$

Theorem 6.4 (Sombrero Lemma)**.** Let $(X, \mathcal{A})$ be a measurable space. Every positive $\mathcal{A}/\mathcal{B}(\overline{\mathbb{R}})$-measurable function $u : X \to [0, \infty]$ is the pointwise limit of an increasing sequence of simple functions $f_n \in \mathcal{E}(\mathcal{A})$, with $f_n \geq 0$, that is,

$$u(x) = \sup_{n \in \mathbb{N}} f_n(x) = \lim_{n \to \infty} f_n(x), \qquad \text{with} \quad f_1 \leq f_2 \leq f_3 \leq \dots$$

Corollary 6.4.1. Let $(X, \mathcal{A})$ be a measurable space. Every $\mathcal{A}/\mathcal{B}(\overline{\mathbb{R}})$-measurable function $u : X \to \overline{\mathbb{R}}$ is the pointwise limit of simple functions $f_n \in \mathcal{E}(\mathcal{A})$ such that

$$|f_n(x)| \leq |u(x)| \quad \text{for all } x \in X.$$

If u is bounded, the convergence is uniform.

Corollary 6.4.2. Let $(X, \mathcal{A})$ be a measurable space. If $u_n : X \to \overline{\mathbb{R}}$ are measurable functions for all $n \in \mathbb{N}$, then the following functions are measurable:

$$\sup_{n \in \mathbb{N}} u_n, \quad \inf_{n \in \mathbb{N}} u_n, \quad \limsup_{n \to \infty} u_n, \quad \liminf_{n \to \infty} u_n,$$

and, whenever it exists in $\overline{\mathbb{R}}$, also the pointwise limit

$$\lim_{n \to \infty} u_n.$$

Corollary 6.4.3. Let $u, v : X \to \overline{\mathbb{R}}$ be $\mathcal{A}/\mathcal{B}(\overline{\mathbb{R}})$-measurable functions. Then the following functions are also measurable.

$$u + v, \quad u \cdot v, \quad u \vee v := \max\{u, v\}, \quad u \wedge v := \min\{u, v\}.$$

Proof. If $u, v \in \mathcal{E}(\mathcal{A})$ are simple functions, then each of the functions

$$u + v, \quad u \cdot v, \quad u \vee v = \max\{u, v\}, \quad u \wedge v = \min\{u, v\}$$

is again a simple function (and hence measurable).

For general $u, v \in \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A})$, let $(f_n)_{n \in \mathbb{N}}, (g_n)_{n \in \mathbb{N}} \subset \mathcal{E}(\mathcal{A})$ be sequences of simple functions such that

$$f_n \to u \quad \text{and} \quad g_n \to v \quad \text{pointwise as } n \to \infty.$$

By the algebraic rules for limits and the fact that the pointwise limit of measurable functions is measurable (Corollary 6.4.2), the operations

$$f_n + g_n \to u + v, \quad f_n \cdot g_n \to u \cdot v, \quad f_n \vee g_n \to u \vee v, \quad f_n \wedge g_n \to u \wedge v$$

preserve measurability.

Hence, $u + v, \ u \cdot v, \ u \vee v, \ u \wedge v \in \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A})$. $\qquad \square$

Corollary 6.4.4. A function $u : X \to \mathbb{R}$ is $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable if and only if both its positive and negative parts,

$$u^+(x) := \max\{u(x), 0\}, \qquad u^-(x) := \max\{-u(x), 0\},$$

are $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable.

Corollary 6.4.5. Let $u, v : X \to \mathbb{R}$ be $\mathcal{A}/\mathcal{B}(\mathbb{R})$-measurable functions. Then the sets

$$\{u < v\}, \quad \{u \leq v\}, \quad \{u = v\}, \quad \{u \neq v\}$$

are all elements of $\mathcal{A}$.

7 Integration of Positive Functions

Throughout this chapter, let $(X, \mathcal{A}, \mu)$ be a measure space.

We recall the following spaces of measurable functions:

- $\mathcal{M}_+(\mathcal{A})$: the set of all **non-negative measurable functions**, that is,

$$\mathcal{M}_+(\mathcal{A}) := \{f : X \to [0, \infty] \mid f \text{ is } \mathcal{A}/\mathcal{B}(\mathbb{R})\text{-measurable}\}.$$

- $\mathcal{E}_+(\mathcal{A})$: the set of all **non-negative simple functions**, i.e., functions of the form

$$f(x) = \sum_{i=1}^{M} y_i \, \mathbf{1}_{A_i}(x), \quad y_i \geq 0, \ A_i \in \mathcal{A}.$$

- $\mathcal{E}(\mathcal{A})$: the set of all **(real-valued) simple functions**, i.e., functions of the form

$$f(x) = \sum_{i=1}^{M} y_i \, \mathbf{1}_{A_i}(x), \quad y_i \in \mathbb{R}, \ A_i \in \mathcal{A}.$$

Lemma 7.1. Let

$$f = \sum_{i=1}^{M} y_i \, \mathbf{1}_{A_i} = \sum_{k=1}^{N} z_k \, \mathbf{1}_{B_k}$$

be two standard representations of the same function $f \in \mathcal{E}_+(\mathcal{A})$. Then,

$$\sum_{i=1}^{M} y_i \, \mu(A_i) = \sum_{k=1}^{N} z_k \, \mu(B_k).$$

Proof. Since both representations of f are standard, the sets $A_1, \ldots, A_M$ and $B_1, \ldots, B_N$ form partitions of X. That is,

$$X = \bigsqcup_{i=1}^{M} A_i = \bigsqcup_{k=1}^{N} B_k.$$

We can express each A_i as a disjoint union of the intersections $A_i \cap B_k$, and similarly for each B_k, i.e.,

$$A_i = \bigsqcup_{k=1}^{N} (A_i \cap B_k), \qquad B_k = \bigsqcup_{i=1}^{M} (B_k \cap A_i).$$

By finite additivity of the measure μ, we have:

$$\sum_{i=1}^{M} y_i \mu(A_i) = \sum_{i=1}^{M} y_i \sum_{k=1}^{N} \mu(A_i \cap B_k) = \sum_{i=1}^{M} \sum_{k=1}^{N} y_i \mu(A_i \cap B_k).$$

Similarly,

$$\sum_{k=1}^{N} z_k \mu(B_k) = \sum_{k=1}^{N} z_k \sum_{i=1}^{M} \mu(B_k \cap A_i) = \sum_{k=1}^{N} \sum_{i=1}^{M} z_k \mu(A_i \cap B_k).$$

Now, since the function values agree on all of X, we must have $y_i = z_k$ whenever $A_i \cap B_k \neq \emptyset$. For disjoint sets where $A_i \cap B_k = \emptyset$, we have $\mu(A_i \cap B_k) = 0$. Therefore,

$$y_i \mu(A_i \cap B_k) = z_k \mu(A_i \cap B_k) \quad \text{for all } i, k.$$

Thus, the double sums are equal, and we conclude:

$$\sum_{i=1}^{M} y_i \mu(A_i) = \sum_{k=1}^{N} z_k \mu(B_k).$$

$\square$

Definition 7.1 (Integral of a Simple Function)**.** Let

$$f = \sum_{i=1}^{M} y_i \, \mathbf{1}_{A_i} \in \mathcal{E}_+(\mathcal{A})$$

be a simple function in standard representation. Then the number

$$I_\mu(f) := \sum_{i=1}^{M} y_i \, \mu(A_i) \in [0, \infty]$$

is called the μ-*integral* of f. This value is independent of the particular standard representation of f.

Lemma 7.2 (Properties of Integral of Simple Functions)**.** Let $f, g \in \mathcal{E}_+(\mathcal{A})$. Then the integral $I_\mu : \mathcal{E}_+(\mathcal{A}) \to [0, \infty]$ satisfies:

(i) **Normalization:** $I_\mu(\mathbf{1}_A) = \mu(A)$ for all $A \in \mathcal{A}$.

(ii) **Positive Homogeneity:** For all $\lambda > 0$, we have $\mathcal{I}_\mu(\lambda f) = \lambda \, I_\mu(f)$.

(iii) **Additivity:** $I_\mu(f + g) = \mathcal{I}_\mu(f) + \mathcal{I}_\mu(g)$.

(iv) **Monotonicity:** If $f \leq g$, then $I_\mu(f) \leq I_\mu(g)$.

Remark 7.1. By Theorem 6.4, every function $u \in \mathcal{M}_+(\mathcal{A})$ can be written as the pointwise limit of an increasing sequence of simple functions. Corollary 6.4.2 further ensures that the supremum of measurable functions is measurable. Hence, we have the characterization:

$$u \in \mathcal{M}_+(\mathcal{A}) \iff u = \sup_{n \in \mathbb{N}} f_n, \quad \text{where } f_n \in \mathcal{E}_+(\mathcal{A}), \ f_n \leq f_{n+1}.$$

This observation allows us to *approximate from below* any positive measurable function u by simple functions whose integrals are well-defined. By doing so, we can systematically exhaust the area under the graph of u using integrals of simple functions.

Definition 7.2 (Integral of a Positive Measurable Function)**.** Let $(X, \mathcal{A}, \mu)$ be a measure space. The μ-integral of a positive measurable function $u \in \mathcal{M}_+(\mathcal{A})$ is defined as

$$\int u \, d\mu := \sup \left\{ I_\mu(g) : 0 \leq g \leq u, \ g \in \mathcal{E}_+(\mathcal{A}) \right\} \in [0, \infty].$$

If we wish to emphasize the integration variable, we may write

$$\int u(x) \, \mu(dx).$$

Lemma 7.3. For all $f \in \mathcal{E}_+(\mathcal{A})$, we have

$$\int f \, d\mu = I_\mu(f).$$

Proof. Let $f \in \mathcal{E}_+(\mathcal{A})$. By definition of the integral for positive measurable functions (Definition 7.2), we take the supremum over all simple functions $g \in \mathcal{E}_+(\mathcal{A})$ such that $g \leq f$.

But since f is itself a simple function and clearly satisfies $f \leq f$, it is one of the admissible candidates in this supremum. This gives

$$\int f \, d\mu = \sup \left\{ I_\mu(g) : g \in \mathcal{E}_+(\mathcal{A}), \ g \leq f \right\} \geq I_\mu(f).$$

On the other hand, any simple function $g \in \mathcal{E}_+(\mathcal{A})$ with $g \leq f$ satisfies $I_\mu(g) \leq I_\mu(f)$, by the monotonicity of I_μ (Lemma 7.2). Hence, the supremum over such $I_\mu(g)$ can never exceed $I_\mu(f)$, so

$$\int f \, d\mu \leq I_\mu(f).$$

Putting both inequalities together, we conclude

$$\int f \, d\mu = I_\mu(f). \qquad \square$$