Measure Theory
Jan-May 2026 Anoop V. P.

Introduction: From Riemann to Lebesgue

Measure theory provides a systematic way to assign numerical values to sets and functions. Specifically, we assign a **measure** to a measurable set and an **integral value** to an integrable function. While Riemann integration is a staple of introductory calculus, Lebesgue integration theory represents a significant generalization and completion of those ideas. Lebesgue's framework allows us to assign numbers to a much broader collection of sets and functions than is possible within the Riemann system.

The Shift from Finite to Countable Operations

The fundamental difference between these two theories can be summarized by their relationship with infinity:

- **Riemann Integration** is centered on **finiteness**. It relies on approximations of a finite nature, such as summing the areas of finitely many rectangles to find the area of a bounded subset of $\mathbb{R}^2$. This theory is robust when dealing with finite operations on sets $A_1, \ldots, A_n$ or functions $f_1, \ldots, f_n$.
- **Lebesgue Integration** is centered on **countable infiniteness**. Riemann integration often fails to behave consistently when faced with operations of a countably infinite nature, such as the union $\bigcup_{n=1}^{\infty} A_n$, the sum $\sum_{n=1}^{\infty} f_n$, or the limit $\lim_{n\to\infty} f_n$. Lebesgue integration was developed to rectify these disadvantages.

Methodological and Structural Differences

Beyond their handling of infinity, the two theories differ in their execution and scope:

- **Partitioning:** In Riemann's theory, we partition the **domain** of the function, whereas in Lebesgue's theory, we partition the **range**.
- **Generalization:** Riemann's theory is largely restricted to Euclidean space. Conversely, the principles of Lebesgue's theory are applicable to more general spaces, leading to the development of **abstract measure theory**. This abstract approach is highly versatile and intersects with numerous branches of mathematics.

I suggest all students read these two series of notes by **Prof. A. K. Nandakumaran**, Department of Mathematics, IISc Bangalore. These are fantastic articles that provide excellent motivation for studying Measure Theory:

- Meaning of Integration 1
- Meaning of Integration 2

The Vitali Construction: Limits of Generalizing Volume

In this section, we attempt to generalize the concept of volume (specifically length in $\mathbb{R}$) by searching for a set function μ that assigns a "size" to subsets of $\mathbb{R}$. We want this function μ to satisfy certain natural properties:

1. **Normalization:** The length of the unit interval is one, i.e., $\mu([0, 1]) = 1$.
2. **Countable Additivity:** If $E_1, E_2, \ldots$ are disjoint subsets, then the measure of their union is the sum of their individual measures: $\mu(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu(E_j)$.
3. **Translation Invariance:** If a set E is shifted by a distance r to form a new set $F = E + r$, then $\mu(F) = \mu(E)$.

The following construction, known as the **Vitali Set**, demonstrates that these three assumptions are mutually inconsistent. It proves that there exist sets for which a consistent "length" cannot be defined.

Step-by-Step Construction of the Vitali Set

We work within the interval $[0, 1]$ in $\mathbb{R}$.

Step 1: Define an Equivalence Relation We define a relation $\sim$ on $[0, 1]$. For any $a, b \in [0, 1]$, we say $a \sim b$ if and only if their difference is a rational number:

$$a \sim b \iff a - b \in \mathbb{Q}$$

This partitions $[0, 1]$ into disjoint equivalence classes, where each class contains all points that differ from each other by only a rational value.

Step 2: Form the Vitali Set Y Using the **Axiom of Choice**, we construct a set $Y \subset [0, 1]$ by picking exactly one element from each equivalence class. This set Y is our "Vitali Set."

Step 3: Analyze Rational Translations Let $Q = \mathbb{Q} \cap [-1, 1]$ be the set of rational numbers between -1 and 1. Since $\mathbb{Q}$ is countable, we can list these rationals as $\{r_1, r_2, r_3, \dots\}$. Consider the translated sets:

$$Y_i = Y + r_i = \{y + r_i : y \in Y\}$$

Step 4: Verify Disjointness We claim that if $r_i \neq r_j$, then $Y_i \cap Y_j = \emptyset$. Suppose there was an element in both. Then $y_1 + r_i = y_2 + r_j$ for some $y_1, y_2 \in Y$. This implies $y_1 - y_2 = r_j - r_i$. Since r_j and r_i are rational, their difference is rational, meaning $y_1 \sim y_2$. However, since Y contains only one element from each class, we must have $y_1 = y_2$, which forces $r_i = r_j$. Thus, the translates are disjoint.

Step 5: Containment and Contradiction Note that since $Y \subset [0, 1]$ and each $r_i \in [-1, 1]$, every translated set Y_i is contained within the interval $[-1, 2]$. Therefore:

$$\bigcup_i Y_i \subset [-1, 2]$$

By translation invariance, all Y_i must have the same measure: $\mu(Y_i) = \mu(Y) = \delta$. By countable additivity:

$$\mu\left(\bigcup_i Y_i\right) = \sum_{i=1}^{\infty} \mu(Y_i) = \sum_{i=1}^{\infty} \delta$$

- If $\delta = 0$, the total measure is 0. But the union of these translates actually covers the interval $[0, 1]$, so the measure must be at least 1. (Contradiction: $0 \geq 1$)
- If $\delta > 0$, the infinite sum $\sum \delta$ diverges to ∞. But the union is contained in $[-1, 2]$, so the measure cannot exceed 3. (Contradiction: $\infty \leq 3$)

Thus, no such function μ can exist for all subsets of $\mathbb{R}$.

Step 6: Finite Additivity Failure Suppose $\mu(Y) = \delta$. Since $Y \subset [0, 1]$, δ must be some non-negative value.

- If $\delta = 0$, then any countable union of its translates would also have measure 0, which cannot cover $[0, 1]$ (as $\mu([0, 1]) = 1$).
- If $\delta > 0$, we can find a natural number k such that $k\delta > 2$.

Now, choose k distinct rational numbers $q_1, q_2, \dots, q_k$ from $\mathbb{Q} \cap [0, 1]$. Define the translated sets $Y_i = Y + q_i$.

1. **Disjointness:** As shown previously, these translates are disjoint because Y picks only one element per class.
2. **Containment:** Since $Y \subset [0, 1]$ and $q_i \in [0, 1]$, it follows that each $Y_i \subset [0, 2]$. Thus:

$$\bigcup_{i=1}^{k} Y_i \subset [0, 2]$$

3. **The Contradiction:** By monotonicity and normalization, $\mu(\bigcup_{i=1}^{k} Y_i) \leq \mu([0, 2]) = 2$. However, by finite additivity and translation invariance:

$$\mu\left(\bigcup_{i=1}^{k} Y_i\right) = \sum_{i=1}^{k} \mu(Y_i) = \sum_{i=1}^{k} \delta = k\delta$$

Since we chose k such that $k\delta > 2$, we get $k\delta \leq 2$, which is a contradiction.

Thus, even finite additivity is inconsistent with translation invariance and normalization.

Remark. *The set Y constructed above is an example of a **non-measurable set**. It shows that we cannot assign a meaningful "length" to every possible collection of real numbers if we want to keep the properties of additivity and translation invariance.*

Exercise 1. *Find two disjoint subsets $A, B \subset [0, 1]$ such that $\mu(A \cup B) \neq \mu(A) + \mu(B)$.*

Solution. Consider the translates $Y_1, Y_2, \dots, Y_k$ from the construction above. We know that:

$$\mu(Y_1 \cup Y_2 \cup \cdots \cup Y_k) \neq \sum_{j=1}^{k} \mu(Y_j)$$

We can find the specific pair by induction:

1. Check Y_1 and Y_2. If $\mu(Y_1 \cup Y_2) \neq \mu(Y_1) + \mu(Y_2)$, we are done. Set $A = Y_1, B = Y_2$.
2. If they are equal, consider the set $(Y_1 \cup Y_2)$ and Y_3. If $\mu((Y_1 \cup Y_2) \cup Y_3) \neq \mu(Y_1 \cup Y_2) + \mu(Y_3)$, then set $A = Y_1 \cup Y_2$ and $B = Y_3$.
3. Continue this process. We must reach a first natural number $r \leq k$ such that:

$$\mu\left(\left(\bigcup_{j=1}^{r-1} Y_j\right) \cup Y_r\right) \neq \mu\left(\bigcup_{j=1}^{r-1} Y_j\right) + \mu(Y_r)$$

   This stage r is guaranteed to exist because if additivity held for every step up to k, the total sum would equal the measure of the union, which we already proved is impossible.

Taking $A = \bigcup_{j=1}^{r-1} Y_j$ and $B = Y_r$ provides the required disjoint sets. □

σ-Algebras and Measurable Spaces

Definition 1. *Let X be a nonempty set. An **algebra** of sets on X (also called a **field**) is a nonempty collection $\mathcal{A}$ of subsets of X that satisfies the following closure properties:*

1. ***Closed under finite unions:*** *If $E_1, \dots, E_n \in \mathcal{A}$, then $\bigcup_{j=1}^{n} E_j \in \mathcal{A}$.*
2. ***Closed under complements:*** *If $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$.*

Definition 2. *A **σ-algebra** (or **σ-field**) is an algebra that is also closed under countable unions. That is, if $\{E_j\}_{j=1}^{\infty} \subset \mathcal{A}$, then $\bigcup_{j=1}^{\infty} E_j \in \mathcal{A}$.*

Remark. *By applying De Morgan's laws, $\bigcap E_j = (\bigcup E_j^c)^c$, we observe that algebras and σ-algebras are also closed under finite and countable intersections, respectively. Furthermore, any algebra $\mathcal{A}$ necessarily contains the empty set $\emptyset$ and the whole set X, since for any $E \in \mathcal{A}$, we have $\emptyset = E \cap E^c$ and $X = E \cup E^c$.*

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

Proof. Suppose $\{E_j\}_1^{\infty} \subset \mathcal{A}$. To show the countable union is in $\mathcal{A}$, we can define a sequence of disjoint sets $\{F_k\}$ as follows:

$$F_1 = E_1, \quad F_k = E_k \setminus \left[\bigcup_{j=1}^{k-1} E_j\right] = E_k \cap \left[\bigcup_{j=1}^{k-1} E_j\right]^c \quad \text{for } k > 1$$

These F_k belong to $\mathcal{A}$ because $\mathcal{A}$ is an algebra (closed under finite unions and complements). By construction, the F_k are disjoint and satisfy $\bigcup_1^{\infty} E_j = \bigcup_1^{\infty} F_k$. Thus, if $\mathcal{A}$ is closed under countable disjoint unions, it must also be closed under all countable unions. □

Remark. *The technique of replacing an arbitrary sequence of sets with a disjoint sequence such that their union remains unchanged is a fundamental tool in measure theory.*

Example 1. *Let X be any non-empty set. The following collections of subsets are σ-algebras on X:*

1. ***The Power Set:*** *$\mathcal{P}(X)$, the collection of all subsets of X.*
2. ***The Trivial σ-algebra:*** *$\{\emptyset, X\}$, the smallest possible σ-algebra.*
3. ***The Countable or Co-countable σ-algebra:*** *If X is an uncountable set, then the collection*

$$\mathcal{A} = \{E \subset X : E \text{ is countable or } E^c \text{ is countable}\}$$

is a σ-algebra.

Remark. *The verification of Example 3 relies on the following logic regarding countable unions:*

- ***Case 1 (All sets are countable):*** *If we have a sequence $\{E_j\}_{j=1}^{\infty} \subset \mathcal{A}$ where every E_j is countable, then their union $\bigcup_{j=1}^{\infty} E_j$ is a countable union of countable sets, which is itself countable. Thus, the union belongs to $\mathcal{A}$.*
- ***Case 2 (At least one set is co-countable):*** *If at least one E_{j_0} is co-countable (meaning $E_{j_0}^c$ is countable), then the complement of the union $(\bigcup E_j)^c = \bigcap E_j^c$ is a subset of $E_{j_0}^c$. Since $E_{j_0}^c$ is countable and any subset of a countable set is countable, the complement of the union is countable. Thus, the union is co-countable and belongs to $\mathcal{A}$.*

Exercise 2. *Prove that the intersection of any family of σ-algebras on a set X is again a σ-algebra on X.*

Solution. Let $\{\mathcal{M}_\alpha\}_{\alpha \in I}$ be a family of σ-algebras on X. Let $\mathcal{M} = \bigcap_{\alpha \in I} \mathcal{M}_\alpha$.

1. Since $X \in \mathcal{M}_\alpha$ for every α, $X \in \mathcal{M}$.
2. If $E \in \mathcal{M}$, then $E \in \mathcal{M}_\alpha$ for all α. Since each $\mathcal{M}_\alpha$ is a σ-algebra, $E^c \in \mathcal{M}_\alpha$ for all α, so $E^c \in \mathcal{M}$.
3. If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$, then $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}_\alpha$ for all α. Thus $\bigcup E_j \in \mathcal{M}_\alpha$ for all α, implying $\bigcup E_j \in \mathcal{M}$.

□

Definition 3. *Let $\mathcal{E}$ be any subset of the power set $\mathcal{P}(X)$. There exists a unique smallest σ-algebra containing $\mathcal{E}$, denoted by $\mathcal{M}(\mathcal{E})$. It is defined as the intersection of all σ-algebras on X that contain $\mathcal{E}$:*

$$\mathcal{M}(\mathcal{E}) = \bigcap \{\mathcal{M} : \mathcal{M} \text{ is a } \sigma\text{-algebra and } \mathcal{E} \subset \mathcal{M}\}$$

*This intersection is well-defined because there is always at least one such σ-algebra, namely $\mathcal{P}(X)$. $\mathcal{M}(\mathcal{E})$ is called the **σ-algebra generated by** $\mathcal{E}$.*

Lemma. *If $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$, then $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}(\mathcal{F})$.*

Proof. By definition, $\mathcal{M}(\mathcal{E})$ is the *smallest* σ-algebra containing the collection of sets $\mathcal{E}$. This means that if $\mathcal{A}$ is any σ-algebra such that $\mathcal{E} \subset \mathcal{A}$, then it must follow that $\mathcal{M}(\mathcal{E}) \subset \mathcal{A}$.

In this case, we are given that $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$. We already know by definition that $\mathcal{M}(\mathcal{F})$ is a σ-algebra. Therefore, $\mathcal{M}(\mathcal{F})$ is a σ-algebra that contains $\mathcal{E}$. By the minimality of $\mathcal{M}(\mathcal{E})$ discussed above, we conclude that $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}(\mathcal{F})$. □

Further Definitions in σ-Algebras

Definition 4. *Let $\mathcal{E}$ be a collection of subsets of X. The **σ-algebra generated by** $\mathcal{E}$, denoted by $\mathcal{M}(\mathcal{E})$, is the unique smallest σ-algebra containing $\mathcal{E}$. It is defined as the intersection of all σ-algebras on X that contain $\mathcal{E}$. Since the power set $\mathcal{P}(X)$ is always one such σ-algebra, this intersection is well-defined.*

Lemma. *If $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$, then $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}(\mathcal{F})$.*

Proof. By definition, $\mathcal{M}(\mathcal{F})$ is a σ-algebra. Since it is given that $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$, then $\mathcal{M}(\mathcal{F})$ is one of the σ-algebras included in the intersection that defines $\mathcal{M}(\mathcal{E})$. Because $\mathcal{M}(\mathcal{E})$ is the intersection of *all* such σ-algebras, it must be a subset of any individual σ-algebra in that collection, specifically $\mathcal{M}(\mathcal{F})$. □

Definition 5. *Let X be a topological space (such as a metric space). The σ-algebra generated by the family of all open sets in X is called the **Borel σ-algebra** on X, denoted by $\mathcal{B}_X$. The members of $\mathcal{B}_X$ are called **Borel sets**.*

Remark. *Since σ-algebras are closed under complements, the Borel σ-algebra is equivalently generated by the family of all closed sets in X. $\mathcal{B}_X$ includes all open sets, closed sets, countable intersections of open sets, and countable unions of closed sets.*

Definition 6. *Within the hierarchy of Borel sets, there is standard terminology for specific levels:*

- *A G_δ **set** is a countable intersection of open sets.*
- *An F_σ **set** is a countable union of closed sets.*
- *A $G_{\delta\sigma}$ **set** is a countable union of G_δ sets.*
- *An $F_{\sigma\delta}$ **set** is a countable intersection of F_σ sets.*

The symbols δ and σ originate from the German words Durchschnitt (intersection) and Summe (union).

Generators of the Borel σ-algebra on $\mathbb{R}$

Proposition. *The Borel σ-algebra $\mathcal{B}_\mathbb{R}$ is generated by each of the following collections of intervals:*

a. *The open intervals: $\mathcal{E}_1 = \{(a, b) : a < b\}$*

b. *The closed intervals: $\mathcal{E}_2 = \{[a, b] : a < b\}$*

c. *The half-open intervals: $\mathcal{E}_3 = \{(a, b] : a < b\}$ or $\mathcal{E}_4 = \{[a, b) : a < b\}$*

d. *The open rays: $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$ or $\mathcal{E}_6 = \{(-\infty, a) : a \in \mathbb{R}\}$*

e. *The closed rays: $\mathcal{E}_7 = \{[a, \infty) : a \in \mathbb{R}\}$ or $\mathcal{E}_8 = \{(-\infty, a] : a \in \mathbb{R}\}$*

Proof. Recall that $\mathcal{B}_\mathbb{R}$ is generated by the collection of all open sets in $\mathbb{R}$. We will use Lemma 1 to show that $\mathcal{M}(\mathcal{E}_j) = \mathcal{B}_\mathbb{R}$ for each j.

Step 1: Establishing $\mathcal{M}(\mathcal{E}_1) = \mathcal{B}_\mathbb{R}$ Since every open interval (a, b) is an open set, we have $\mathcal{E}_1 \subset \mathcal{B}_\mathbb{R}$, which implies $\mathcal{M}(\mathcal{E}_1) \subset \mathcal{B}_\mathbb{R}$. Conversely, every open set in $\mathbb{R}$ can be written as a countable union of open intervals. Since σ-algebras are closed under countable unions, any open set must belong to $\mathcal{M}(\mathcal{E}_1)$. Therefore, $\mathcal{B}_\mathbb{R} \subset \mathcal{M}(\mathcal{E}_1)$, and we conclude $\mathcal{M}(\mathcal{E}_1) = \mathcal{B}_\mathbb{R}$.

Step 2: Showing $\mathcal{M}(\mathcal{E}_j) \subset \mathcal{B}_\mathbb{R}$ for all j The elements of $\mathcal{E}_j$ for $j \neq 3, 4$ are either open or closed sets, which are inherently Borel sets. For $\mathcal{E}_3$ and $\mathcal{E}_4$, notice that:

$$(a, b] = \bigcap_{n=1}^{\infty} \left(a, b + \frac{1}{n}\right) \quad \text{and} \quad [a, b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b\right)$$

Since these are countable intersections of open sets (G_δ sets), they are Borel sets. Thus, $\mathcal{E}_j \subset \mathcal{B}_\mathbb{R}$ for all j, which implies $\mathcal{M}(\mathcal{E}_j) \subset \mathcal{B}_\mathbb{R}$.

Step 3: Showing $\mathcal{B}_\mathbb{R} \subset \mathcal{M}(\mathcal{E}_j)$ for $j \geq 2$ To prove this, it suffices to show that any open interval (a, b) can be constructed using elements from each $\mathcal{E}_j$.

- **For $\mathcal{E}_2$ (Closed intervals):** We can write $(a, b) = \bigcup_{n=1}^{\infty} [a + \frac{1}{n}, b - \frac{1}{n}]$. Since each $[a + \frac{1}{n}, b - \frac{1}{n}] \in \mathcal{E}_2$, the union is in $\mathcal{M}(\mathcal{E}_2)$.

- **For $\mathcal{E}_3$ (Half-open $(a, b]$):** We can write $(a, b) = \bigcup_{n=1}^{\infty}(a, b - \frac{1}{n}]$. Thus $\mathcal{E}_1 \subset \mathcal{M}(\mathcal{E}_3)$.
- **For $\mathcal{E}_5$ (Open rays (a, ∞)):** Note that $(-\infty, b] = (b, \infty)^c$, so $(-\infty, b] \in \mathcal{M}(\mathcal{E}_5)$. Then $(a, b] = (a, \infty) \cap (-\infty, b] \in \mathcal{M}(\mathcal{E}_5)$. Using the result from $\mathcal{E}_3$, we get $\mathcal{B}_{\mathbb{R}} \subset \mathcal{M}(\mathcal{E}_5)$.
- **For $\mathcal{E}_7$ (Closed rays $[a, \infty)$):** Note that $(a, \infty) = \bigcup_{n=1}^{\infty}[a + \frac{1}{n}, \infty)$. Thus $\mathcal{E}_5 \subset \mathcal{M}(\mathcal{E}_7)$, which implies $\mathcal{B}_{\mathbb{R}} \subset \mathcal{M}(\mathcal{E}_7)$.

Similar constructions apply to $\mathcal{E}_4, \mathcal{E}_6$, and $\mathcal{E}_8$ by using symmetry and complements. Since $\mathcal{E}_1 \subset \mathcal{M}(\mathcal{E}_j)$ in all cases, we have $\mathcal{B}_{\mathbb{R}} \subset \mathcal{M}(\mathcal{E}_j)$. □

Product σ-Algebras

Definition 7. *Let $\{X_\alpha\}_{\alpha \in A}$ be an indexed collection of nonempty sets, and let $X = \prod_{\alpha \in A} X_\alpha$ denote their Cartesian product. For each $\alpha \in A$, let $\pi_\alpha : X \to X_\alpha$ be the coordinate maps (or projection maps). If $\mathcal{M}_\alpha$ is a σ-algebra on X_α for each α, the **product σ-algebra** on X is defined as the σ-algebra generated by the following collection:*

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

This σ-algebra is denoted by $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$.

Remark. *In the specific case where the index set is finite, such as $A = \{1, \dots, n\}$, we use the notation $\bigotimes_1^n \mathcal{M}_j$ or $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$. The importance of this definition is primarily established when discussing the measurability of functions on product spaces.*

Proposition. *If the index set A is countable, then the product σ-algebra $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$ is generated by the collection of measurable rectangles:*

$$\mathcal{F} = \left\{ \prod_{\alpha \in A} E_\alpha : E_\alpha \in \mathcal{M}_\alpha \right\}$$

Proof. To prove that $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha = \mathcal{M}(\mathcal{F})$, we show containment in both directions using Lemma 1.

Step 1: Showing $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha \subset \mathcal{M}(\mathcal{F})$ Let $\pi_\alpha^{-1}(E_\alpha)$ be a generator of the product σ-algebra, where $E_\alpha \in \mathcal{M}_\alpha$. This inverse image can be written as a special type of product:

$$\pi_\alpha^{-1}(E_\alpha) = \prod_{\beta \in A} E_\beta, \quad \text{where } E_\beta = \begin{cases} E_\alpha & \text{if } \beta = \alpha \\ X_\beta & \text{if } \beta \neq \alpha \end{cases}$$

Since X_β is always in $\mathcal{M}_\beta$, this set is a member of the collection $\mathcal{F}$. Thus, every generator of the product σ-algebra is contained in $\mathcal{F}$. By Lemma 1, it follows that $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha \subset \mathcal{M}(\mathcal{F})$.

Step 2: Showing $\mathcal{M}(\mathcal{F}) \subset \bigotimes_{\alpha \in A} \mathcal{M}_\alpha$ Now consider an arbitrary element of $\mathcal{F}$, which is a product of the form $\prod_{\alpha \in A} E_\alpha$. We can express this product as an intersection of inverse images:

$$\prod_{\alpha \in A} E_\alpha = \bigcap_{\alpha \in A} \pi_\alpha^{-1}(E_\alpha)$$

By the definition of the product σ-algebra, each set $\pi_\alpha^{-1}(E_\alpha)$ is a generator and thus belongs to $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$. Since the index set A is countable, this is a **countable intersection** of measurable sets. Since σ-algebras are closed under countable intersections, the resulting product belongs to $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$.

Therefore, $\mathcal{F} \subset \bigotimes_{\alpha \in A} \mathcal{M}_\alpha$, which by Lemma 1 implies $\mathcal{M}(\mathcal{F}) \subset \bigotimes_{\alpha \in A} \mathcal{M}_\alpha$. We conclude that the two σ-algebras are identical. □

Proposition. *Suppose that for each $\alpha \in A$, the σ-algebra $\mathcal{M}_\alpha$ is generated by a collection $\mathcal{E}_\alpha$ (i.e., $\mathcal{M}_\alpha = \mathcal{M}(\mathcal{E}_\alpha)$). Then the product σ-algebra $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$ is generated by the collection:*

$$\mathcal{F}_1 = \{\pi_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{E}_\alpha, \alpha \in A\}$$

Furthermore, if the index set A is countable and $X_\alpha \in \mathcal{E}_\alpha$ for all α, then $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$ is also generated by the collection of "simple" rectangles:

$$\mathcal{F}_2 = \left\{ \prod_{\alpha \in A} E_\alpha : E_\alpha \in \mathcal{E}_\alpha \right\}$$

Proof. We begin by proving the first assertion regarding $\mathcal{F}_1$.

Step 1: Establishing $\mathcal{M}(\mathcal{F}_1) \subset \bigotimes_{\alpha\in A} \mathcal{M}_\alpha$ By definition, each $\mathcal{E}_\alpha \subset \mathcal{M}_\alpha$. Consequently, every set in $\mathcal{F}_1$ is of the form $\pi_\alpha^{-1}(E_\alpha)$ where $E_\alpha \in \mathcal{M}_\alpha$. These are precisely the generators of the product σ-algebra. Since $\mathcal{F}_1 \subset \bigotimes \mathcal{M}_\alpha$, it follows by Lemma 1 that $\mathcal{M}(\mathcal{F}_1) \subset \bigotimes \mathcal{M}_\alpha$.

Step 2: Establishing $\bigotimes_{\alpha\in A} \mathcal{M}_\alpha \subset \mathcal{M}(\mathcal{F}_1)$ For a fixed $\alpha \in A$, consider the collection of subsets of X_α whose inverse images under the coordinate map fall into $\mathcal{M}(\mathcal{F}_1)$:

$$\mathcal{G}_\alpha = \{E \subset X_\alpha : \pi_\alpha^{-1}(E) \in \mathcal{M}(\mathcal{F}_1)\}$$

We can verify that $\mathcal{G}_\alpha$ is a σ-algebra on X_α:

- $X_\alpha \in \mathcal{G}_\alpha$ because $\pi_\alpha^{-1}(X_\alpha) = X$, which is in any σ-algebra.
- If $E \in \mathcal{G}_\alpha$, then $\pi_\alpha^{-1}(E^c) = (\pi_\alpha^{-1}(E))^c \in \mathcal{M}(\mathcal{F}_1)$, so $E^c \in \mathcal{G}_\alpha$.
- If $\{E_j\} \subset \mathcal{G}_\alpha$, then $\pi_\alpha^{-1}(\cup E_j) = \cup \pi_\alpha^{-1}(E_j) \in \mathcal{M}(\mathcal{F}_1)$, so $\cup E_j \in \mathcal{G}_\alpha$.

By the construction of $\mathcal{F}_1$, we know that $\mathcal{E}_\alpha \subset \mathcal{G}_\alpha$. Since $\mathcal{M}_\alpha$ is the *smallest* σ-algebra containing $\mathcal{E}_\alpha$, we must have $\mathcal{M}_\alpha \subset \mathcal{G}_\alpha$. This implies that for *any* $E \in \mathcal{M}_\alpha$, the set $\pi_\alpha^{-1}(E)$ belongs to $\mathcal{M}(\mathcal{F}_1)$. Since these inverse images generate the product σ-algebra, we conclude $\bigotimes \mathcal{M}_\alpha \subset \mathcal{M}(\mathcal{F}_1)$.

Step 3: Proving the second assertion for $\mathcal{F}_2$ If A is countable and $X_\alpha \in \mathcal{E}_\alpha$, we can use the same logic as the previous Proposition. Every element $\pi_\alpha^{-1}(E_\alpha)$ in $\mathcal{F}_1$ can be written as a product $\prod E_\beta$ where $E_\beta = X_\beta$ for $\beta \neq \alpha$. Since $X_\beta \in \mathcal{E}_\beta$, this product is in $\mathcal{F}_2$. Thus $\mathcal{F}_1 \subset \mathcal{F}_2$, and conversely, every element of $\mathcal{F}_2$ is a countable intersection of elements from $\mathcal{F}_1$:

$$\prod_{\alpha\in A} E_\alpha = \bigcap_{\alpha\in A} \pi_\alpha^{-1}(E_\alpha)$$

The result then follows directly from the closure of σ-algebras under countable intersections. □

Borel Sets on Product Spaces

Preliminary: Separability and Product Metrics

Before relating product σ-algebras to Borel σ-algebras, we recall some essential concepts from the theory of metric spaces.

Definition 8. *A metric space (X, ρ) is called* ***separable*** *if it contains a countable dense subset. That is, there exists a set $C = \{c_1, c_2, \dots\}$ such that for every $x \in X$ and every $\epsilon > 0$, the open ball $B(x, \epsilon)$ contains at least one point of C.*

Example 2. *The space $\mathbb{R}^n$ with the usual Euclidean metric is separable. The set $\mathbb{Q}^n = \{(q_1, \dots, q_n) : q_i \in \mathbb{Q}\}$ is a countable dense subset of $\mathbb{R}^n$.*

Remark. *An important property of separable metric spaces is that they possess a* ***countable base****. Let C be a countable dense subset of X. Let $\mathcal{E}$ be the collection of all open balls with centers in C and rational radii:*

$$\mathcal{E} = \{B(c, r) : c \in C, r \in \mathbb{Q}^+\}$$

Since C and $\mathbb{Q}^+$ are countable, $\mathcal{E}$ is a countable collection. Every open set $V \subset X$ can be written as a countable union of members of $\mathcal{E}$. Proof idea: For any $x \in V$, there is a ball $B(x, \epsilon) \subseteq V$. By density, we can find $c_i \in C$ very close to x and a rational radius r such that $x \in B(c_i, r) \subseteq B(x, \epsilon) \subseteq V$. Thus, V is the union of all such balls from $\mathcal{E}$ contained within it.

Definition 9. *Let $(X_1, \rho_1), \dots, (X_n, \rho_n)$ be metric spaces. The* ***product metric*** *ρ on $X = \prod_{j=1}^n X_j$ is typically defined as:*

$$\rho((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max_{1\le j\le n} \{\rho_j(x_j, y_j)\}$$

Under this metric, an open ball in X is simply the product of open balls in the component spaces:

$$B_X(x, r) = B_{X_1}(x_1, r) \times B_{X_2}(x_2, r) \times \cdots \times B_{X_n}(x_n, r)$$

Furthermore, if each (X_j, ρ_j) is separable with countable dense subset C_j, then X is separable with countable dense subset $C = \prod_{j=1}^n C_j$.

The Relationship between $\mathcal{B}_X$ and $\bigotimes \mathcal{B}_{X_j}$

Proposition. *Let $X_1, \ldots, X_n$ be metric spaces and let $X = \prod_1^n X_j$ be equipped with the product metric. Then $\bigotimes_1^n \mathcal{B}_{X_j} \subset \mathcal{B}_X$. If each X_j is separable, then $\bigotimes_1^n \mathcal{B}_{X_j} = \mathcal{B}_X$.*

Proof. We approach the proof in steps to distinguish between the general case and the case requiring separability.

Step 1: Understanding the Two σ-Algebras We are comparing two distinct ways of constructing a σ-algebra on the product space X:

- **The Borel σ-algebra $\mathcal{B}_X$:** This is generated by the collection of all open sets in the product metric space X. This topology is equivalent to the product topology.
- **The Product σ-algebra $\bigotimes \mathcal{B}_{X_j}$:** This is generated by "measurable cylinders" of the form $\pi_j^{-1}(E_j)$, where $E_j \in \mathcal{B}_{X_j}$.

Step 2: Proving the General Inclusion $\bigotimes_1^n \mathcal{B}_{X_j} \subset \mathcal{B}_X$ By our previous results, the product σ-algebra $\bigotimes \mathcal{B}_{X_j}$ is generated by the collection:

$$\mathcal{F} = \{\pi_j^{-1}(U_j) : U_j \text{ is open in } X_j, 1 \leq j \leq n\}$$

For each j, the projection map π_j is continuous. Therefore, if U_j is an open set in X_j, its inverse image $\pi_j^{-1}(U_j)$ is an open set in the product space X. Since $\mathcal{B}_X$ is the σ-algebra generated by *all* open sets in X, it must contain this sub-collection of open sets. By the minimality of generated σ-algebras (Lemma 1), we conclude that $\bigotimes \mathcal{B}_{X_j} \subset \mathcal{B}_X$.

Step 3: Leveraging Separability in X_j Suppose each X_j is separable. Let C_j be a countable dense subset of X_j. Let $\mathcal{E}_j$ be the countable collection of open balls in X_j with centers in C_j and rational radii. As noted in our preliminary remarks, every open set in X_j is a countable union of balls from $\mathcal{E}_j$. This implies that the Borel σ-algebra $\mathcal{B}_{X_j}$ is generated by the countable collection $\mathcal{E}_j$.

Step 4: Completion of the Equality via Separability Now consider the product space X. Since X is a product of finitely many separable spaces, it is also separable with a countable dense subset $C = \prod C_j$.

1. **Generation of $\mathcal{B}_X$:** Because X is separable, $\mathcal{B}_X$ is generated by the collection of all "product balls":
$$\mathcal{F}_2 = \left\{ \prod_{j=1}^n E_j : E_j \in \mathcal{E}_j \right\}$$
This is because any open set in X is a countable union of such product balls.
2. **Generation of the Product σ-algebra:** From our previous Proposition regarding generators, since $X_j \in \mathcal{E}_j$ (by taking a sufficiently large rational radius) and the index set is finite (and thus countable), the product σ-algebra $\bigotimes \mathcal{B}_{X_j}$ is also generated by the same collection of "simple" rectangles $\mathcal{F}_2$.

Since both $\mathcal{B}_X$ and $\bigotimes \mathcal{B}_{X_j}$ are generated by the exact same collection of sets $\mathcal{F}_2$, they must be identical. □

Remark. *The assumption of separability is crucial. In non-separable spaces, the Borel σ-algebra $\mathcal{B}_X$ can be strictly larger than the product σ-algebra because there may be open sets in X that cannot be represented as countable unions of measurable rectangles.*

Corollary 1. $\mathcal{B}_{\mathbb{R}^n} = \bigotimes_1^n \mathcal{B}_{\mathbb{R}}$.

Proof. This follows directly from the previous Proposition. Since $\mathbb{R}$ is a separable metric space (containing the countable dense subset $\mathbb{Q}$), its n-fold Cartesian product $\mathbb{R}^n$ satisfies the condition for the Borel σ-algebra to coincide with the product σ-algebra. □

Eelementary Family

Definition 10. *A collection $\mathcal{E}$ of subsets of X is called an* ***elementary family*** *if it satisfies the following three properties:*

1. $\emptyset \in \mathcal{E}$.
2. *If $E, F \in \mathcal{E}$, then $E \cap F \in \mathcal{E}$ (closed under finite intersections).*
3. *If $E \in \mathcal{E}$, then its complement E^c is a finite disjoint union of members of $\mathcal{E}$. That is, $E^c = \bigsqcup_{j=1}^n E_j$ for some $E_1, \ldots, E_n \in \mathcal{E}$.*

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

Proof. To prove that $\mathcal{A}$ is an algebra, we must show that it is closed under finite unions and complements. Recall that an element of $\mathcal{A}$ is of the form $\bigsqcup_{i=1}^n E_i$ where each $E_i \in \mathcal{E}$.

Step 1: Closed under Finite Unions First, we observe that if we can show the union of two elements of $\mathcal{E}$ belongs to $\mathcal{A}$, the general case for elements of $\mathcal{A}$ follows.

Let $A, B \in \mathcal{E}$. We want to show $A \cup B \in \mathcal{A}$. Since $\mathcal{E}$ is an elementary family, the complement B^c can be written as a finite disjoint union of members of $\mathcal{E}$, say $B^c = \bigsqcup_{j=1}^J C_j$ where $C_j \in \mathcal{E}$.

- We can write the set difference as $A \setminus B = A \cap B^c = A \cap (\bigcup_{j=1}^J C_j) = \bigcup_{j=1}^J (A \cap C_j)$.
- Since $\mathcal{E}$ is closed under finite intersections, each $(A \cap C_j) \in \mathcal{E}$. Because the C_j are disjoint, these intersections are also disjoint. Thus, $A \setminus B \in \mathcal{A}$.
- Now, we can write the union as $A \cup B = (A \setminus B) \sqcup B$. Since both $(A \setminus B)$ and B are finite disjoint unions of elements of $\mathcal{E}$, their union is also a finite disjoint union of elements of $\mathcal{E}$. Hence, $A \cup B \in \mathcal{A}$.

By induction, it follows that if $A_1, \ldots, A_n \in \mathcal{E}$, then $\bigcup_1^n A_j \in \mathcal{A}$. Specifically, for the inductive step, we assume the union of $n-1$ sets is in $\mathcal{A}$ and write:

$$\bigcup_{j=1}^n A_j = A_n \cup \bigcup_{j=1}^{n-1} (A_j \setminus A_n)$$

This expresses the union as a disjoint union of elements in $\mathcal{A}$, confirming that $\mathcal{A}$ is closed under finite unions.

Step 2: Closed under Complements Let $A \in \mathcal{A}$. By definition, $A = \bigcup_{m=1}^n A_m$ where $A_m \in \mathcal{E}$ and the A_m are disjoint. We need to show $A^c \in \mathcal{A}$. Using De Morgan's Laws:

$$A^c = \left(\bigcup_{m=1}^n A_m \right)^c = \bigcap_{m=1}^n A_m^c$$

Since $\mathcal{E}$ is an elementary family, each A_m^c is a finite disjoint union of members of $\mathcal{E}$. Let $A_m^c = \bigsqcup_{j=1}^{J_m} B_m^j$, where $B_m^j \in \mathcal{E}$. Substituting this into the intersection gives:

$$A^c = \bigcap_{m=1}^n \left(\bigcup_{j=1}^{J_m} B_m^j \right)$$

By the distributive law for sets, this intersection of unions can be rewritten as a union of intersections:

$$A^c = \bigcup \{ B_1^{j_1} \cap B_2^{j_2} \cap \cdots \cap B_n^{j_n} : 1 \leq j_m \leq J_m \text{ for each } m \}$$

Each term $B_1^{j_1} \cap \cdots \cap B_n^{j_n}$ is an intersection of elements of $\mathcal{E}$, so it belongs to $\mathcal{E}$. Furthermore, these terms are disjoint because if we change any index j_m, we are intersecting with a disjoint set from the m-th family. Thus, A^c is a finite disjoint union of members of $\mathcal{E}$, meaning $A^c \in \mathcal{A}$. $\square$

Example 3. *Let $\mathcal{E}$ be the collection of all half-open intervals of the form $(a, b]$ in $\mathbb{R}$, including $(a, a] = \emptyset$ and infinite intervals like (a, ∞) and $(-\infty, b]$. This collection forms an* ***elementary family*** *because:*

1. *$\emptyset \in \mathcal{E}$ (by setting $a = b$).*
2. *The intersection of any two such intervals $(a, b] \cap (c, d] = (\max(a, c), \min(b, d)]$ is either another half-open interval or empty.*
3. *The complement of any interval $(a, b]^c = (-\infty, a] \cup (b, \infty)$ is a finite disjoint union of two elements from $\mathcal{E}$.*

Measures and Measure Spaces

Definition 11. *Let X be a set equipped with a σ-algebra $\mathcal{M}$. A* ***measure*** *on $\mathcal{M}$ (or on $(X, \mathcal{M})$, or simply on X if $\mathcal{M}$ is understood) is a function $\mu : \mathcal{M} \to [0, \infty]$ that satisfies the following two properties:*

i. ***Null Empty Set:*** *$\mu(\emptyset) = 0$.*

ii. ***Countable Additivity:*** *If $\{E_j\}_{j=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{M}$, then:*

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j)$$

Remark. *Property (ii) implies* ***finite additivity****. If $E_1, \dots, E_n$ are disjoint sets in $\mathcal{M}$, then:*

$$\mu\left(\bigcup_{j=1}^{n} E_j\right) = \sum_{j=1}^{n} \mu(E_j)$$

This is easily seen by taking $E_j = \emptyset$ for all $j > n$. A function μ that satisfies property (i) and finite additivity, but not necessarily countable additivity, is called a ***finitely additive measure****.*

Definition 12. *We use the following standard terminology for the structures involving measures:*

- *The pair $(X, \mathcal{M})$ is called a* ***measurable space****.*
- *The elements of the σ-algebra $\mathcal{M}$ are called* ***measurable sets****.*
- *The triple $(X, \mathcal{M}, \mu)$ is called a* ***measure space****.*

Classification of Measures by Size

Let $(X, \mathcal{M}, \mu)$ be a measure space. We classify the measure μ (and sets within $\mathcal{M}$) based on their "size" as follows:

Definition 13. *A measure μ is called* ***finite*** *if $\mu(X) < \infty$. Note: Since $\mu(X) = \mu(E) + \mu(E^c)$ for any $E \in \mathcal{M}$, the finiteness of $\mu(X)$ implies that $\mu(E) < \infty$ for all measurable sets E.*

Definition 14. *A measure μ is called* ***σ-finite*** *if X can be written as a countable union of measurable sets with finite measure. That is, $X = \bigcup_1^{\infty} E_j$ where $E_j \in \mathcal{M}$ and $\mu(E_j) < \infty$ for all j.*

Definition 15. *A set $E \in \mathcal{M}$ is said to be* ***σ-finite for*** *μ if $E = \bigcup_1^{\infty} E_j$ where $E_j \in \mathcal{M}$ and $\mu(E_j) < \infty$ for all j.*

Definition 16. *A measure μ is called* ***semifinite*** *if for every $E \in \mathcal{M}$ with $\mu(E) = \infty$, there exists a measurable subset $F \subset E$ such that $0 < \mu(F) < \infty$.*

Theorem 2. *Let $(X, \mathcal{M}, \mu)$ be a measure space.*

a. ***(Monotonicity)*** *If $E, F \in \mathcal{M}$ and $E \subset F$, then $\mu(E) \leq \mu(F)$.*

b. ***(Subadditivity)*** *If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$, then $\mu\left(\bigcup_1^{\infty} E_j\right) \leq \sum_1^{\infty} \mu(E_j)$.*

c. ***(Continuity from below)*** *If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$ and $E_1 \subset E_2 \subset \dots$, then:*

$$\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \lim_{j \to \infty} \mu(E_j)$$

d. ***(Continuity from above)*** *If* $\{E_j\}_{j=1}^{\infty} \subset \mathcal{M}$, $E_1 \supset E_2 \supset \dots$, *and* $\mu(E_1) < \infty$, *then:*

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

Proof. **(a) Proof of Monotonicity:** Suppose $E, F \in \mathcal{M}$ with $E \subset F$. We can decompose F into two disjoint sets: E and the set difference $F \setminus E$. Thus, $F = E \cup (F \setminus E)$. By the finite additivity of measures:

$$\mu(F) = \mu(E) + \mu(F \setminus E)$$

Since the range of μ is $[0, \infty]$, we know $\mu(F \setminus E) \geq 0$. Therefore, $\mu(F) \geq \mu(E)$.

(b) Proof of Subadditivity: To use the additivity property of measures on a sequence that might not be disjoint, we must first "disjointify" it. Let $\{E_j\}_1^{\infty}$ be any sequence in $\mathcal{M}$. Define:

$$F_1 = E_1, \quad F_k = E_k \setminus \left(\bigcup_{j=1}^{k-1} E_j\right) \text{ for } k > 1$$

These F_k are disjoint by construction, and their union is the same as the union of the E_j's (i.e., $\bigcup_1^{\infty} F_j = \bigcup_1^{\infty} E_j$). Furthermore, for each j, $F_j \subset E_j$. By countable additivity and the monotonicity proved in (a):

$$\mu\left(\bigcup_1^{\infty} E_j\right) = \mu\left(\bigcup_1^{\infty} F_j\right) = \sum_1^{\infty} \mu(F_j) \leq \sum_1^{\infty} \mu(E_j)$$

(c) Proof of Continuity from Below: Let $E_1 \subset E_2 \subset \dots$ be an increasing sequence of sets. We define $E_0 = \emptyset$. We can represent the union as a disjoint union of "rings":

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{j=1}^{\infty} (E_j \setminus E_{j-1})$$

Applying countable additivity:

$$\mu\left(\bigcup_1^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j \setminus E_{j-1})$$

By the definition of an infinite series, this is the limit of partial sums:

$$\mu\left(\bigcup_1^{\infty} E_j\right) = \lim_{n\to\infty} \sum_{j=1}^{n} \mu(E_j \setminus E_{j-1})$$

Since the sets are nested, the finite sum $\sum_{j=1}^{n} \mu(E_j \setminus E_{j-1})$ is a telescoping sum that simplifies to $\mu(E_n)$. Thus:

$$\mu\left(\bigcup_1^{\infty} E_j\right) = \lim_{n\to\infty} \mu(E_n)$$

(d) Proof of Continuity from Above: Let $E_1 \supset E_2 \supset \dots$ with $\mu(E_1) < \infty$. We transform this into an increasing sequence to apply part (c). Let $F_j = E_1 \setminus E_j$. Since the E_j are decreasing, the F_j are increasing ($F_1 \subset F_2 \subset \dots$). Note that $\bigcup F_j = E_1 \setminus (\bigcap E_j)$. By part (c):

$$\mu(E_1 \setminus \bigcap E_j) = \lim_{j\to\infty} \mu(F_j)$$

Using the property $\mu(A \setminus B) = \mu(A) - \mu(B)$ (valid here since $\mu(E_1) < \infty$):

$$\mu(E_1) - \mu\left(\bigcap_1^{\infty} E_j\right) = \lim_{j\to\infty} [\mu(E_1) - \mu(E_j)]$$

Because $\mu(E_1)$ is finite, we can subtract it from both sides of the equation:

$$-\mu\left(\bigcap_1^{\infty} E_j\right) = -\lim_{j\to\infty} \mu(E_j)$$

Multiplying by -1 gives the desired result: $\mu\left(\bigcap_1^{\infty} E_j\right) = \lim_{j\to\infty} \mu(E_j)$. ∎

Remark. *The condition $\mu(E_1) < \infty$ in part (d) of Theorem 1.8 is a sufficient but not strictly necessary condition for the first term of the sequence; it could be replaced by $\mu(E_n) < \infty$ for any $n > 1$. This is because the first $n-1$ sets in the sequence do not affect the ultimate intersection $\bigcap_1^\infty E_j$.*

*However, **some** finiteness assumption is absolutely necessary. It is possible for $\mu(E_j) = \infty$ for all j, yet $\mu(\bigcap_1^\infty E_j) < \infty$. In such cases, the limit of the measures may not equal the measure of the intersection.*

Example 4 (Failure of Continuity from Above without Finiteness)**.** *Let μ be the **counting measure** on the measurable space $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Define a decreasing sequence of sets as follows:*

$$E_j = \{n \in \mathbb{N} : n \geq j\} = \{j, j+1, j+2, \dots\}$$

- *For every $j \in \mathbb{N}$, the set E_j contains infinitely many elements, so $\mu(E_j) = \infty$. This implies:*

$$\lim_{j\to\infty} \mu(E_j) = \infty$$

- *However, the intersection of all these sets is empty: $\bigcap_{j=1}^\infty E_j = \emptyset$.*
- *Therefore, $\mu(\bigcap_1^\infty E_j) = \mu(\emptyset) = 0$.*

In this case, $0 \neq \infty$, demonstrating that the property fails when the finiteness condition is violated.

Remark. *Every σ-finite measure is semifinite (Exercise 1.13), but not conversely. Most measures that arise in practice are σ-finite, which is fortunate since non-σ-finite measures tend to exhibit pathological behavior.*

Exercise 3 (Exercise 1.13)**.** *Prove that every σ-finite measure is semifinite.*

Proof. Let $(X, \mathcal{M}, \mu)$ be a σ-finite measure space. By definition, $X = \bigcup_{n=1}^\infty E_n$ where $E_n \in \mathcal{M}$ and $\mu(E_n) < \infty$ for all n. To show μ is semifinite, let $E \in \mathcal{M}$ be such that $\mu(E) = \infty$. We must find a measurable subset $F \subset E$ with $0 < \mu(F) < \infty$.

- We can decompose E using the sequence $\{E_n\}$:

$$E = E \cap X = E \cap \left(\bigcup_{n=1}^\infty E_n \right) = \bigcup_{n=1}^\infty (E \cap E_n)$$

- By subadditivity, $\mu(E) \leq \sum_{n=1}^\infty \mu(E \cap E_n)$. Since $\mu(E) = \infty$, the sum $\sum \mu(E \cap E_n)$ must also be infinite.
- This implies that there must exist at least one index k for which $\mu(E \cap E_k) > 0$.
- Furthermore, since $E \cap E_k \subseteq E_k$, by monotonicity we have $\mu(E \cap E_k) \leq \mu(E_k) < \infty$.
- Thus, $F = E \cap E_k$ is a measurable subset of E that satisfies $0 < \mu(F) < \infty$.

Therefore, μ is semifinite. □

Measures Defined by Functions

A common way to construct measures on a set X is by using a non-negative function. Let X be any nonempty set, and let $\mathcal{M} = \mathcal{P}(X)$ be its power set. Let $f : X \to [0, \infty]$ be any function.

Definition 17. *The function f determines a measure μ on $\mathcal{M}$ defined by the formula:*

$$\mu(E) = \sum_{x \in E} f(x)$$

For the definition of such possibly uncountable sums, one typically defines $\sum_{x\in E} f(x) = \sup\{\sum_{x\in F} f(x) : F \subseteq E, F \text{ is finite}\}$.

Remark. *The properties of the measure μ are directly related to the behavior of the function f:*

- *μ is **semifinite** if and only if $f(x) < \infty$ for every $x \in X$.*
- *μ is **σ-finite** if and only if μ is semifinite and the set $\{x : f(x) > 0\}$ is countable.*

Special Cases of Significance

Two specific instances of this construction are used throughout measure theory and probability:

1. **Counting Measure:** If $f(x) = 1$ for all $x \in X$, then μ is called the **counting measure**.
$$\mu(E) = \begin{cases} |E| & \text{if } E \text{ is finite} \\ \infty & \text{if } E \text{ is infinite} \end{cases}$$
Note that the counting measure is σ-finite if and only if X is a countable set.

2. **Dirac Measure (Point Mass):** If, for some fixed $x_0 \in X$, the function f is defined by $f(x_0) = 1$ and $f(x) = 0$ for $x \neq x_0$, then μ is called the **point mass** or **Dirac measure** at x_0, often denoted by δ_{x_0}.
$$\delta_{x_0}(E) = \begin{cases} 1 & \text{if } x_0 \in E \\ 0 & \text{if } x_0 \notin E \end{cases}$$
The Dirac measure is always a finite measure (since $\mu(X) = 1$), and thus it is always σ-finite and semifinite.

Note: These same names are also applied to the restrictions of these measures to smaller σ-algebras on X.

Semifinite measure which is not σ-finite : To show that a semifinite measure is not necessarily σ-finite, consider the **counting measure** μ on an uncountable set X (for example, $X = \mathbb{R}$) with the σ-algebra $\mathcal{P}(X)$.

1. μ **is semifinite:** Let $E \subset X$ with $\mu(E) = \infty$. This means E is an infinite set. Since X consists of points, we can pick any single point $x \in E$. The set $F = \{x\}$ is a measurable subset of E, and $\mu(F) = 1$. Since $0 < 1 < \infty$, the measure is semifinite.

2. μ **is not σ-finite:** Suppose $X = \bigcup_{n=1}^{\infty} E_n$ where $\mu(E_n) < \infty$ for all n. Under the counting measure, $\mu(E_n) < \infty$ means that each E_n must be a finite set. However, a countable union of finite sets is at most countable. Since we assumed X is uncountable, no such sequence $\{E_n\}$ can exist.

Thus, μ is semifinite but not σ-finite.

More Examples

Example 5 (The Countable/Co-countable Measure)**.** *Let X be an uncountable set, and let $\mathcal{M}$ be the σ-algebra of countable or co-countable sets. Define μ on $\mathcal{M}$ by:*
$$\mu(E) = \begin{cases} 0 & \text{if } E \text{ is countable} \\ 1 & \text{if } E \text{ is co-countable} \end{cases}$$
This is a valid measure because in any sequence of disjoint sets $\{E_j\} \subset \mathcal{M}$, at most one E_j can be co-countable. This ensures that countable additivity holds.

Example 6 (A Finitely Additive Function that is Not a Measure)**.** *Let X be an infinite set and $\mathcal{M} = \mathcal{P}(X)$. Define $\mu(E) = 0$ if E is finite and $\mu(E) = \infty$ if E is infinite.*

- ***Finite Additivity:*** *If A, B are disjoint and finite, $A \cup B$ is finite ($0 + 0 = 0$). If one is infinite, $A \cup B$ is infinite ($\infty + 0 = \infty$).*
- ***Failure of Countable Additivity:*** *Let $X = \mathbb{N}$ and $E_j = \{j\}$. Then $\sum \mu(E_j) = \sum 0 = 0$, but $\mu(\bigcup E_j) = \mu(\mathbb{N}) = \infty$.*

*Thus, μ is finitely additive but is **not** a measure.*

Definition 18 (Null Set)**.** *If $(X, \mathcal{M}, \mu)$ is a measure space, a set $E \in \mathcal{M}$ such that $\mu(E) = 0$ is called a **null set**. By subadditivity, any countable union of null sets is a null set.*

Definition 19 (Almost Everywhere)**.** *If a statement about points $x \in X$ is true except for x in some null set, we say that it is true **almost everywhere** (abbreviated **a.e.**), or for **almost every** x. If more precision is needed, we shall speak of a μ-**null set**, or μ-**almost everywhere**.*

Definition 20 (σ-finite Set)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. A set $E \in \mathcal{M}$ is said to be σ-**finite for** μ if $E = \bigcup_1^\infty E_j$ where $E_j \in \mathcal{M}$ and $\mu(E_j) < \infty$ for all j.*

Definition 21 (Semifinite Measure)**.** *A measure μ is called **semifinite** if for each $E \in \mathcal{M}$ with $\mu(E) = \infty$, there exists some $F \in \mathcal{M}$ with $F \subset E$ and $0 < \mu(F) < \infty$.*

Definition 22 (Finitely Additive Measure)**.** *A function μ that satisfies $\mu(\emptyset) = 0$ and finite additivity ($\mu(\bigcup_1^n E_j) = \sum_1^n \mu(E_j)$ for disjoint sets), but not necessarily countable additivity, is called a **finitely additive measure**.*

Complete Measures

In a general measure space, if E is a **null set** ($\mu(E) = 0$) and $F \subset E$, it follows from monotonicity that $\mu(F) = 0$ provided that $F \in \mathcal{M}$. However, it is not always true that every subset of a null set belongs to the σ-algebra $\mathcal{M}$.

Definition 23. *A measure whose domain includes all subsets of null sets is called **complete**.*

Completeness can often simplify technical arguments and can always be achieved by enlarging the domain of the measure μ.

—

Theorem 3 (Completion of a Measure)**.** *Suppose that $(X, \mathcal{M}, \mu)$ is a measure space. Let $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$ be the collection of null sets, and define:*

$$\overline{\mathcal{M}} = \{E \cup F : E \in \mathcal{M} \text{ and } F \subset N \text{ for some } N \in \mathcal{N}\}$$

Then $\overline{\mathcal{M}}$ is a σ-algebra, and there is a unique extension $\overline{\mu}$ of μ to a complete measure on $\overline{\mathcal{M}}$.

Proof. Let $(X, \mathcal{M}, \mu)$ be a measure space and $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$. We define:

$$\overline{\mathcal{M}} = \{E \cup F : E \in \mathcal{M} \text{ and } F \subset N \text{ for some } N \in \mathcal{N}\}$$

Step 1: Closed under Countable Unions
Let $\{A_j\}_{j=1}^\infty$ be a sequence in $\overline{\mathcal{M}}$. By definition, $A_j = E_j \cup F_j$ where $E_j \in \mathcal{M}$ and $F_j \subset N_j \in \mathcal{N}$. We look at the union $A = \bigcup A_j$:

- Let $E = \bigcup E_j$. Since $\mathcal{M}$ is a σ-algebra, $E \in \mathcal{M}$.
- Let $F = \bigcup F_j$ and $N = \bigcup N_j$. Since $\mathcal{M}$ is a σ-algebra, $N \in \mathcal{M}$.
- By subadditivity, $\mu(N) \leq \sum \mu(N_j) = 0$, so $N \in \mathcal{N}$.
- Since $F_j \subset N_j$, it follows that $F \subset N$.

Thus $A = E \cup F \in \overline{\mathcal{M}}$.
Step 2: Closed under Complements
Let $A = E \cup F \in \overline{\mathcal{M}}$ with $F \subset N \in \mathcal{N}$. We can assume $E \cap N = \emptyset$ by replacing F and N with $F \setminus E$ and $N \setminus E$ if necessary. Observe that $E \cup F = (E \cup N) \cap (N^c \cup F)$. Taking the complement:

$$(E \cup F)^c = (E \cup N)^c \cup (N \setminus F)$$

- $(E \cup N)^c \in \mathcal{M}$ because $E, N \in \mathcal{M}$.
- $N \setminus F \subset N$, and since N is a null set, $N \setminus F$ is a subset of a null set.

Thus $A^c \in \overline{\mathcal{M}}$, and $\overline{\mathcal{M}}$ is a σ-algebra.
Step 3: Well-definedness of $\overline{\mu}$
Define $\overline{\mu}(E \cup F) = \mu(E)$. Suppose $E_1 \cup F_1 = E_2 \cup F_2$ with $F_j \subset N_j \in \mathcal{N}$. Then:

$$E_1 \subset E_1 \cup F_1 = E_2 \cup F_2 \subset E_2 \cup N_2$$

By monotonicity: $\mu(E_1) \leq \mu(E_2 \cup N_2) \leq \mu(E_2) + \mu(N_2) = \mu(E_2)$. By symmetry, $\mu(E_2) \leq \mu(E_1)$. Hence $\mu(E_1) = \mu(E_2)$, so $\overline{\mu}$ is well-defined.

Step 4: Completeness of $\overline{\mu}$
Suppose $\overline{\mu}(A) = 0$ for some $A \in \overline{\mathcal{M}}$, and let $B \subset A$. $A = E \cup F$ where $\mu(E) = 0$. Since $F \subset N \in \mathcal{N}$, $A \subset E \cup N$. Let $N' = E \cup N \in \mathcal{N}$. Since $B \subset A \subset N'$, B is a subset of a null set. We can write $B = \emptyset \cup B$, so $B \in \overline{\mathcal{M}}$. Thus $\overline{\mu}$ is complete.

Step 5: Uniqueness
If ν is another measure on $\overline{\mathcal{M}}$ extending μ, then for $A = E \cup F$:

$$\mu(E) = \nu(E) \leq \nu(E \cup F) \leq \nu(E \cup N) \leq \nu(E) + \nu(N) = \mu(E) + 0$$

This forces $\nu(A) = \mu(E) = \overline{\mu}(A)$. □

The measure $\overline{\mu}$ in the above theorem is called the **completion of** μ, and $\overline{\mathcal{M}}$ is called the **completion of** $\mathcal{M}$ **with respect to** μ.

Outer Measures

In this section, we develop the tools used to construct measures. To motivate the ideas, it is useful to recall the procedure used in calculus to define the area of a bounded region E in the plane $\mathbb{R}^2$.

Motivation: Area in the Plane

One typically draws a grid of rectangles in the plane and approximates the area of E from two sides:

- **Inner Area:** Approximated from below by the sum of the areas of the rectangles in the grid that are *subsets* of E.
- **Outer Area:** Approximated from above by the sum of the areas of the rectangles in the grid that *intersect* E.

The limits of these approximations as the grid is taken finer and finer give the "inner area" and "outer area" of E. If they are equal, their common value is the "area" of E.

The key idea here is that of **outer area**. If R is a large rectangle containing E, the inner area of E is simply the area of R minus the outer area of $R \setminus E$. Thus, the notion of outer area is the fundamental concept we generalize to abstract spaces.

Definition of Outer Measure

The abstract generalization of the notion of outer area is as follows:

Definition 24. *An **outer measure** on a nonempty set X is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ that satisfies the following three properties:*

1. *$\mu^*(\emptyset) = 0$.*
2. ***Monotonicity:** $\mu^*(A) \leq \mu^*(B)$ if $A \subset B$.*
3. ***Countable Subadditivity:** $\mu^*(\bigcup_1^\infty A_j) \leq \sum_1^\infty \mu^*(A_j)$ for any sequence $\{A_j\} \subset \mathcal{P}(X)$.*

Remark. *Note that the domain of an outer measure is the entire power set $\mathcal{P}(X)$, but it does not require countable additivity-only subadditivity.*

Proposition (Proposition 1.10)**.** *Let $\mathcal{E} \subset \mathcal{P}(X)$ and $\rho : \mathcal{E} \to [0, \infty]$ be such that $\emptyset \in \mathcal{E}$, $X \in \mathcal{E}$, and $\rho(\emptyset) = 0$. For any $A \subset X$, define*

$$\mu^*(A) = \inf \left\{ \sum_{j=1}^{\infty} \rho(E_j) : E_j \in \mathcal{E} \text{ and } A \subset \bigcup_{j=1}^{\infty} E_j \right\}$$

Then μ^* is an outer measure.

Proof. We verify the three axioms of an outer measure step-by-step.

Step 1: The definition of μ^* makes sense.
For any $A \subset X$, we must ensure that the set over which we take the infimum is non-empty. Since $X \in \mathcal{E}$ and $A \subset X$, we can choose the sequence $E_1 = X$ and $E_j = \emptyset$ for all $j > 1$. This sequence covers A, so the infimum is well-defined.

Step 2: $\mu^*(\emptyset) = 0$.
Since $\emptyset \in \mathcal{E}$ and $\rho(\emptyset) = 0$, we can cover the empty set using $E_j = \emptyset$ for all j. Then $\sum \rho(E_j) = 0$. Since μ^* is non-negative and the infimum must be ≤ 0, we have $\mu^*(\emptyset) = 0$.

Step 3: Monotonicity.
Suppose $A \subset B$.

- Let $\mathcal{C}_A$ be the collection of all sequences $\{E_j\} \subset \mathcal{E}$ that cover A, and $\mathcal{C}_B$ be those that cover B.
- If a sequence $\{E_j\}$ covers B, it automatically covers A because $A \subset B$. Therefore, $\mathcal{C}_B \subset \mathcal{C}_A$.
- Taking the infimum over a larger set ($\mathcal{C}_A$) results in a smaller or equal value. Thus, $\inf(\mathcal{C}_A) \leq \inf(\mathcal{C}_B)$, which means $\mu^*(A) \leq \mu^*(B)$.

Step 4: Countable Subadditivity.
Suppose $\{A_j\}_1^\infty \subset \mathcal{P}(X)$ and let $\epsilon > 0$.

1. By the property of the infimum, for each j, $\mu^*(A_j) + \epsilon 2^{-j}$ is not a lower bound for the sums. Thus, there exists a sequence $\{E_j^k\}_{k=1}^\infty \subset \mathcal{E}$ such that $A_j \subset \bigcup_{k=1}^\infty E_j^k$ and:
$$\sum_{k=1}^\infty \rho(E_j^k) \leq \mu^*(A_j) + \epsilon 2^{-j}$$
2. Let $A = \bigcup_{j=1}^\infty A_j$. Then $A \subset \bigcup_{j,k=1}^\infty E_j^k$. This double union is a countable union of elements in $\mathcal{E}$, so it is a valid candidate for the infimum defining $\mu^*(A)$.
3. By the definition of μ^* as the *infimum*, we have:
$$\mu^*(A) \leq \sum_{j,k} \rho(E_j^k) = \sum_j \left(\sum_k \rho(E_j^k) \right)$$
4. Substituting the inequality from step (1):
$$\mu^*(A) \leq \sum_j (\mu^*(A_j) + \epsilon 2^{-j}) = \sum_j \mu^*(A_j) + \epsilon \sum_j 2^{-j} = \sum_j \mu^*(A_j) + \epsilon$$

Since ϵ is arbitrary, we conclude $\mu^*(A) \leq \sum_j \mu^*(A_j)$. □

Remark (Observation: Construction of Lebesgue Outer Measure)**.** *We can apply the general result of above Proposition to $\mathbb{R}^n$ to define a measure for arbitrary sets. Let $\mathcal{E}$ be the collection of all n-**boxes** in $\mathbb{R}^n$, where an n-box is a product of n intervals. For any box B, let $Vol_n(B)$ denote its standard volume.*

Definition 25. *The n-**dimensional Lebesgue outer measure** $\mu^*_{L,n}$ of an arbitrary set $Y \subset \mathbb{R}^n$ is defined as:*
$$\mu^*_{L,n}(Y) = \inf \left\{ \sum_{j=1}^\infty Vol_n(A_j) : A_j \text{ are } n\text{-boxes with } Y \subset \bigcup_{j=1}^\infty A_j \right\}$$

By Proposition above, since $\emptyset$ is a box with volume 0 and $\mathbb{R}^n$ can be covered by boxes, $\mu^*_{L,n}$ is a valid outer measure on $\mathcal{P}(\mathbb{R}^n)$. This gives us a concrete way to measure the "outer size" of any set in Euclidean space using simple geometric building blocks.

The Carathéodory Condition

The fundamental step that leads from outer measures to measures is as follows. If μ^* is an outer measure on X, a set $A \subset X$ is called **μ^*-measurable** if it satisfies the **Carathéodory condition**:

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

Key Observations for Students:

- **Subadditivity is Given:** The inequality $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c)$ holds for any A and E by the definition of an outer measure.
- **The Proof Strategy:** To prove that A is μ^*-measurable, it is sufficient to prove only the reverse inequality:
$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$
- **Handling Infinite Measure:** The reverse inequality is trivial if $\mu^*(E) = \infty$.
- **Simplified Criterion:** Therefore, A is μ^*-measurable if and only if:
$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subset X \text{ such that } \mu^*(E) < \infty.$$

Theorem 4 (Carathéodory's Theorem)**.** *If μ^* is an outer measure on X, the collection $\mathcal{M}$ of μ^*-measurable sets is a σ-algebra, and the restriction of μ^* to $\mathcal{M}$ is a complete measure.*

Proof. Let μ^* be an outer measure on X and $\mathcal{M}$ be the collection of sets $A \subset X$ satisfying the Carathéodory condition:

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

Step 1: $\mathcal{M}$ is closed under complements.
The definition of the Carathéodory condition is perfectly symmetric with respect to A and A^c. Since $(A^c)^c = A$, the equation for A^c is identical to the equation for A. Thus, $A \in \mathcal{M} \implies A^c \in \mathcal{M}$.

Step 2: $\mathcal{M}$ is an algebra and μ^* is finitely additive.
Suppose $A, B \in \mathcal{M}$. To show $A \cup B \in \mathcal{M}$, let $E \subset X$ be any test set.

1. Since $A \in \mathcal{M}$, we split E using A:
$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$
2. Since $B \in \mathcal{M}$, we split the pieces $(E \cap A)$ and $(E \cap A^c)$ using B:
$$\mu^*(E \cap A) = \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c)$$
$$\mu^*(E \cap A^c) = \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)$$
3. Substitute these back into the original equation for $\mu^*(E)$:
$$\mu^*(E) = \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)$$
4. Note that $A \cup B = (A \cap B) \cup (A \cap B^c) \cup (A^c \cap B)$. By the subadditivity of μ^*:
$$\mu^*(E \cap (A \cup B)) \leq \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B)$$
5. Since $A^c \cap B^c = (A \cup B)^c$, substituting the inequality from (4) into (3) yields:
$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$

This proves $A \cup B \in \mathcal{M}$. If A, B are disjoint, setting $E = A \cup B$ in the split $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$ gives $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$, showing finite additivity.

Step 3: $\mathcal{M}$ is a σ-algebra.
Let $\{A_j\}_1^\infty$ be disjoint sets in $\mathcal{M}$. Let $B_n = \bigcup_1^n A_j$ and $B = \bigcup_1^\infty A_j$.

1. By finite additivity and induction: $\mu^*(E \cap B_n) = \sum_1^n \mu^*(E \cap A_j)$.
2. Since $B_n \in \mathcal{M}$, we have $\mu^*(E) = \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c)$.
3. By monotonicity, $B^c \subset B_n^c$, so $\mu^*(E \cap B_n^c) \geq \mu^*(E \cap B^c)$. Thus:

$$\mu^*(E) \geq \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*(E \cap B^c)$$

4. Taking $n \to \infty$: $\mu^*(E) \geq \sum_1^\infty \mu^*(E \cap A_j) + \mu^*(E \cap B^c)$.
5. By countable subadditivity: $\sum_1^\infty \mu^*(E \cap A_j) \geq \mu^*(\bigcup (E \cap A_j)) = \mu^*(E \cap B)$.
6. Therefore: $\mu^*(E) \geq \mu^*(E \cap B) + \mu^*(E \cap B^c)$, which proves $B \in \mathcal{M}$.

Step 4: μ^* is countably additive on $\mathcal{M}$.
Setting $E = B$ in the inequality $\mu^*(E) \geq \sum \mu^*(E \cap A_j) + \mu^*(E \cap B^c)$ gives:

$$\mu^*(B) \geq \sum_{j=1}^{\infty} \mu^*(A_j) + 0$$

Since $\mu^*(B) \leq \sum \mu^*(A_j)$ by subadditivity, we have equality.
Step 5: The measure is complete.
If $\mu^*(A) = 0$, then for any $E \subset X$, $\mu^*(E \cap A) = 0$ by monotonicity. Since $\mu^*(E) \geq \mu^*(E \cap A^c)$ also by monotonicity, we have:

$$\mu^*(E) \geq 0 + \mu^*(E \cap A^c) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

This satisfies the Carathéodory condition, so $A \in \mathcal{M}$. $\square$

Premeasures

Before extending a measure to a fully developed σ-algebra, we frequently begin by defining a well-behaved set function on a simpler algebraic structure known as an algebra. This foundational set function is called a premeasure.

Definition 26. *Let X be a non-empty set and let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra of subsets of X. A function $\mu_0 : \mathcal{A} \to [0, \infty]$ is called a **premeasure** if it satisfies the following two conditions:*

1. *The empty set has a measure of zero:*
$$\mu_0(\emptyset) = 0$$
2. *It is countably additive on elements within the algebra. Specifically, if $\{A_j\}_{j=1}^\infty$ is a sequence of mutually disjoint sets in $\mathcal{A}$ such that their countable union $\bigcup_{j=1}^\infty A_j$ also happens to belong to $\mathcal{A}$, then:*
$$\mu_0\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu_0(A_j)$$

Remark. *The definition of a premeasure naturally inherits the standard behaviors associated with classic measures, which can be summarized in two key structural properties:*

1. ***Finite Additivity:** Every premeasure is automatically finitely additive. If $A_1, A_2, \dots, A_n$ is a finite collection of mutually disjoint sets belonging to the algebra $\mathcal{A}$, we can convert this into an infinite sequence by defining $A_j = \emptyset$ for all index values $j > n$. Because algebras are closed under finite unions, the resulting countable union matches the finite union $\bigcup_{j=1}^n A_j \in \mathcal{A}$. Applying the countable additivity property and utilizing $\mu_0(\emptyset) = 0$ collapses the tail of the summation:*

$$\mu_0\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \mu_0(A_j)$$

2. ***Finite and σ-Finite Character:*** *The foundational concepts of finite and σ-finite properties for a premeasure are defined in the exact same manner as they are for standard measures:*

   - *A premeasure μ_0 is called* ***finite*** *if the total measure of the entire underlying space is bounded, meaning $\mu_0(X) < \infty$.*
   - *A premeasure μ_0 is called* ***σ-finite*** *if the full space X can be decomposed into a countable union of sets from the algebra, $X = \bigcup_{j=1}^{\infty} B_j$, such that each individual component set satisfies $B_j \in \mathcal{A}$ and possesses a finite premeasure value, $\mu_0(B_j) < \infty$.*

Inducing an Outer Measure from a Premeasure

We recall from general measure theory that an outer measure can be constructed on a set X starting from a highly minimal amount of structure. Specifically, if $\mathcal{E} \subset \mathcal{P}(X)$ is any subcollection of sets that contains both the empty set $\emptyset$ and the whole space X, and $\rho : \mathcal{E} \to [0, \infty]$ is a set function satisfying $\rho(\emptyset) = 0$, then ρ induces a valid outer measure μ^* on $\mathcal{P}(X)$ by looking at the infimum of the total weights of countable coverings from $\mathcal{E}$.

Because an algebra $\mathcal{A}$ automatically contains $\emptyset$ and X, and a premeasure $\mu_0 : \mathcal{A} \to [0, \infty]$ inherently satisfies $\mu_0(\emptyset) = 0$, a premeasure fulfills these minimal structural requirements perfectly. We can therefore systematically define the outer measure induced by μ_0 on any arbitrary subset $E \subset X$ as follows:

Definition 27. *Let μ_0 be a premeasure on an algebra $\mathcal{A} \subset \mathcal{P}(X)$. The* ***induced outer measure*** *$\mu^* : \mathcal{P}(X) \to [0, \infty]$ is defined for any set $E \subset X$ by:*

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu_0(A_j) : A_j \in \mathcal{A} \text{ for all } j \geq 1, \text{ and } E \subset \bigcup_{j=1}^{\infty} A_j \right\}$$

Remark. *This definition explicitly means that to calculate the outer measure of an arbitrary set E, we look at all possible ways to cover E using a countable family of sets $\{A_j\}_{j=1}^{\infty}$ chosen exclusively from our algebra $\mathcal{A}$. We compute the sum of the premeasures of those covering elements, $\sum \mu_0(A_j)$, and then take the greatest lower bound (infimum) over all such valid coverings.*

Theorem 5. *Let μ_0 be a premeasure on an algebra $\mathcal{A}$, and let μ^* be the induced outer measure defined by:*

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu_0(A_j) : A_j \in \mathcal{A} \text{ for all } j \geq 1, \text{ and } E \subset \bigcup_{j=1}^{\infty} A_j \right\}$$

Then the following properties hold:

a. *The restriction of the induced outer measure μ^* to the algebra $\mathcal{A}$ matches the original premeasure μ_0; that is, $\mu^*|_{\mathcal{A}} = \mu_0$.*

b. *Every set belonging to the algebra $\mathcal{A}$ is μ^*-measurable in the sense of Carathéodory.*

Proof. **Part (a): Consistency on the Algebra ($\mu^*|_{\mathcal{A}} = \mu_0$)**

Let E be an arbitrary set belonging to the algebra $\mathcal{A}$. To establish equality, we prove both directional inequalities step-by-step:

- **Step 1: Establishing $\mu^*(E) \leq \mu_0(E)$.** By the definition of the induced outer measure, $\mu^*(E)$ is computed as the infimum over all valid countable coverings from $\mathcal{A}$. We can construct a highly trivial countable covering for E using elements of $\mathcal{A}$ by choosing:

  $$A_1 = E, \quad \text{and} \quad A_j = \emptyset \quad \text{for all } j > 1$$

  Since $E \in \mathcal{A}$ and $\emptyset \in \mathcal{A}$, this family forms a valid countable covering such that $E \subset \bigcup_{j=1}^{\infty} A_j$. Summing their premeasures yields:

  $$\sum_{j=1}^{\infty} \mu_0(A_j) = \mu_0(E) + \mu_0(\emptyset) + \mu_0(\emptyset) + \cdots = \mu_0(E)$$

  Because the infimum must be less than or equal to the total sum of this specific valid covering, we obtain:

  $$\mu^*(E) \leq \mu_0(E)$$

- **Step 2: Establishing $\mu_0(E) \le \mu^*(E)$.** Let $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ be any arbitrary countable sequence of sets in the algebra that covers E, meaning $E \subset \bigcup_{j=1}^{\infty} A_j$. To leverage the countable additivity of the premeasure μ_0, we must convert this potentially overlapping family into a mutually disjoint sequence of sets. We define the disjoint sequence $\{B_n\}_{n=1}^{\infty}$ by intersectively trimming away prior overlaps:

$$B_n = E \cap \left(A_n \setminus \bigcup_{i=1}^{n-1} A_i \right)$$

By this construction, the sets B_n satisfy the following properties:

1. They are mutually disjoint elements of the algebra $\mathcal{A}$ because $\mathcal{A}$ is closed under finite unions, relative complements, and intersections.
2. Their countable union satisfies $\bigcup_{n=1}^{\infty} B_n = E \cap (\bigcup_{n=1}^{\infty} A_n) = E$ since $E \subset \bigcup_{n=1}^{\infty} A_n$.
3. Each individual set satisfies $B_n \subset A_n$, which by the monotonicity of the premeasure implies $\mu_0(B_n) \le \mu_0(A_n)$.

Because $\bigcup_{n=1}^{\infty} B_n = E \in \mathcal{A}$, we apply the defining property of countable additivity for premeasures to obtain:

$$\mu_0(E) = \sum_{n=1}^{\infty} \mu_0(B_n) \le \sum_{n=1}^{\infty} \mu_0(A_n)$$

Since this inequality holds firmly for every conceivable countable covering $\{A_n\}$ of E within $\mathcal{A}$, it must also hold boundedly for their greatest lower bound (infimum):

$$\mu_0(E) \le \mu^*(E)$$

Combining the conclusions of Step 1 and Step 2 proves that $\mu^*(E) = \mu_0(E)$ for all $E \in \mathcal{A}$.

—

Part (b): Carathéodory Measurability of the Algebra

To prove that an arbitrary set $A \in \mathcal{A}$ is μ^*-measurable, we must show that it satisfies the Carathéodory splitting criterion for any test set $E \subset X$:

$$\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

(Note: The reverse inequality $\mu^*(E) \le \mu^*(E \cap A) + \mu^*(E \cap A^c)$ always holds automatically due to the subadditivity of outer measures).

- **Step 1: Setting up the covering approach via ϵ.** Fix any test subset $E \subset X$ and choose an arbitrary tolerance parameter $\epsilon > 0$. By the infimum-based definition of $\mu^*(E)$, there must exist a countable covering sequence $\{B_j\}_{j=1}^{\infty} \subset \mathcal{A}$ such that $E \subset \bigcup_{j=1}^{\infty} B_j$, which satisfies the approximating upper bound:

$$\sum_{j=1}^{\infty} \mu_0(B_j) \le \mu^*(E) + \epsilon$$

- **Step 2: Splitting the covering elements using additivity.** For every set B_j in the covering sequence, we decompose it into two disjoint components using our target algebra element A:

$$B_j = (B_j \cap A) \cup (B_j \cap A^c)$$

Since $B_j \in \mathcal{A}$ and $A \in \mathcal{A}$, both components $(B_j \cap A)$ and $(B_j \cap A^c)$ are disjoint elements belonging to the algebra $\mathcal{A}$. Utilizing the finite additivity property inherent to premeasures on an algebra, we can separate the value of $\mu_0(B_j)$ precisely as:

$$\mu_0(B_j) = \mu_0(B_j \cap A) + \mu_0(B_j \cap A^c)$$

- **Step 3: Distributing the full summation and applying definitions.** Substituting this exact identity into our epsilon covering inequality from Step 1 yields:

$$\mu^*(E) + \epsilon \ge \sum_{j=1}^{\infty} [\mu_0(B_j \cap A) + \mu_0(B_j \cap A^c)] = \sum_{j=1}^{\infty} \mu_0(B_j \cap A) + \sum_{j=1}^{\infty} \mu_0(B_j \cap A^c)$$

We now look at the two separate sum channels on the right side:

1. The sequence $\{B_j \cap A\}_{j=1}^{\infty} \subset \mathcal{A}$ forms a valid countable covering for the portion $E \cap A$ because $E \subset \bigcup B_j \implies E \cap A \subset \bigcup (B_j \cap A)$. By definition of the outer measure, we have $\sum_{j=1}^{\infty} \mu_0(B_j \cap A) \geq \mu^*(E \cap A)$.
2. The sequence $\{B_j \cap A^c\}_{j=1}^{\infty} \subset \mathcal{A}$ forms a valid countable covering for the remaining portion $E \cap A^c$ because $E \cap A^c \subset \bigcup (B_j \cap A^c)$. By definition of the outer measure, we have $\sum_{j=1}^{\infty} \mu_0(B_j \cap A^c) \geq \mu^*(E \cap A^c)$.

Substituting these outer measure bounds directly into our inequality chain gives:

$$\mu^*(E) + \epsilon \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

- **Step 4: Vanishing epsilon limit.** Since this inequality is preserved consistently for any arbitrarily small value of $\epsilon > 0$, we take the limit as $\epsilon \to 0$, which yields the required Carathéodory splitting relation:
$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

Thus, every set $A \in \mathcal{A}$ is verified to be μ^*-measurable. □

Theorem 6 (The Carathéodory Extension Theorem)**.** *Let X be a non-empty set, let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra of subsets, and let $\mu_0 : \mathcal{A} \to [0, \infty]$ be a premeasure. Let $\mathcal{M} = \sigma(\mathcal{A})$ be the σ-algebra generated by $\mathcal{A}$. Then:*

a. Existence and Minimality: *There exists a measure μ on $\mathcal{M}$ whose restriction to $\mathcal{A}$ is exactly μ_0. Specifically, this extension is given by $\mu = \mu^*|_{\mathcal{M}}$, where μ^* is the outer measure induced by μ_0. Furthermore, if ν is any other measure on $\mathcal{M}$ that extends μ_0, then ν satisfies:*

$$\nu(E) \leq \mu(E) \quad \text{for all } E \in \mathcal{M}$$

with guaranteed equality whenever $\mu(E) < \infty$.

b. Uniqueness: *If the initial premeasure μ_0 is σ-finite, then μ is the unique measure on $\mathcal{M}$ that extends μ_0.*

Proof. **Part 1: Existence of the Extension Measure μ**

We define μ^* on $\mathcal{P}(X)$ as the outer measure induced by μ_0. Let $\mathcal{M}^*$ denote the collection of all μ^*-measurable sets in the sense of Carathéodory.

- By Carathéodory's Theorem, $\mathcal{M}^*$ is a fully developed σ-algebra, and the restriction of μ^* to $\mathcal{M}^*$ is a complete measure.
- We proved previously that every set in the algebra $\mathcal{A}$ is μ^*-measurable. Therefore, $\mathcal{A} \subset \mathcal{M}^*$.
- Since $\mathcal{M}^*$ is a σ-algebra containing $\mathcal{A}$, it must contain the smallest σ-algebra containing $\mathcal{A}$, which is $\mathcal{M} = \sigma(\mathcal{A})$. Thus, $\mathcal{M} \subset \mathcal{M}^*$.
- Restricting μ^* to this sub-σ-algebra $\mathcal{M}$ yields a valid measure, which we call $\mu = \mu^*|_{\mathcal{M}}$. Furthermore, for any $A \in \mathcal{A}$, we have $\mu(A) = \mu^*(A) = \mu_0(A)$, proving that μ truly extends μ_0.

—

Part 2: The Minimality of μ Compared to Other Extensions

Let ν be any arbitrary measure on $\mathcal{M}$ that also satisfies $\nu|_{\mathcal{A}} = \mu_0$. Fix any set $E \in \mathcal{M}$.

- **Step 1: Establishing $\nu(E) \leq \mu(E)$ via coverings.** Let $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ be any countable collection of sets in the algebra that covers E, meaning $E \subset \bigcup_{j=1}^{\infty} A_j$. Using the countable subadditivity of the measure ν, and the fact that ν matches μ_0 on $\mathcal{A}$, we write:
$$\nu(E) \leq \nu\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \sum_{j=1}^{\infty} \nu(A_j) = \sum_{j=1}^{\infty} \mu_0(A_j)$$
Since this bound holds for every conceivable countable covering $\{A_j\} \subset \mathcal{A}$ of E, it must remain valid when taking the infimum over all such coverings. By the definition of the induced outer measure, this infimum is exactly $\mu^*(E) = \mu(E)$. Thus:
$$\nu(E) \leq \mu(E) \quad \text{for all } E \in \mathcal{M}$$

- **Step 2: Proving equality when $\mu(E) < \infty$.** Suppose that $\mu(E) < \infty$. Fix an arbitrary tolerance parameter $\epsilon > 0$. By the definition of $\mu(E)$ as an infimum, we can find a countable covering sequence $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ for E such that:

$$\sum_{j=1}^{\infty} \mu_0(A_j) < \mu(E) + \epsilon$$

Let $A = \bigcup_{j=1}^{\infty} A_j$. Since $\mathcal{M}$ is a σ-algebra, $A \in \mathcal{M}$. By using the continuity from below for the measures ν and μ, we can look at the finite union limits:

$$\nu(A) = \lim_{n\to\infty} \nu\left(\bigcup_{j=1}^{n} A_j\right) \quad \text{and} \quad \mu(A) = \lim_{n\to\infty} \mu\left(\bigcup_{j=1}^{n} A_j\right)$$

Because $\bigcup_{j=1}^{n} A_j$ is a finite union of sets in $\mathcal{A}$, it belongs to the algebra $\mathcal{A}$. Since ν and μ are completely identical on the algebra $\mathcal{A}$, their limits must also be equal:

$$\nu(A) = \mu(A)$$

Now, since A covers E, we have $A = E \cup (A \setminus E)$ as a disjoint union. Using the additivity of measures, we can write:

$$\mu(A) = \mu(E) + \mu(A \setminus E) \implies \mu(A \setminus E) = \mu(A) - \mu(E)$$

Since $\mu(A) \le \sum_{j=1}^{\infty} \mu_0(A_j) < \mu(E) + \epsilon$, we substitute this bound to find:

$$\mu(A \setminus E) < (\mu(E) + \epsilon) - \mu(E) = \epsilon$$

We now bound $\mu(E)$ using ν by inserting our known relation $\mu(A) = \nu(A)$:

$$\mu(E) \le \mu(A) = \nu(A) = \nu(E) + \nu(A \setminus E)$$

Since we already proved in Step 1 that $\nu \le \mu$ across all sets in $\mathcal{M}$, it follows that $\nu(A \setminus E) \le \mu(A \setminus E) < \epsilon$. Substituting this back into our equation yields:

$$\mu(E) \le \nu(E) + \epsilon$$

Because $\epsilon > 0$ can be made arbitrarily small, taking the limit as $\epsilon \to 0$ gives $\mu(E) \le \nu(E)$. Combined with our Step 1 result ($\nu(E) \le \mu(E)$), this yields $\nu(E) = \mu(E)$.

—

Part 3: Complete Uniqueness Under σ-Finiteness
Now suppose that μ_0 is σ-finite. By definition, the entire space X can be decomposed into a countable union $X = \bigcup_{j=1}^{\infty} A_j$, where each $A_j \in \mathcal{A}$ and $\mu_0(A_j) < \infty$.

Without loss of generality, we can convert this sequence into a disjoint sequence of algebra elements by setting $A_1' = A_1$ and $A_n' = A_n \setminus \bigcup_{i=1}^{n-1} A_j$. Thus, we assume the elements $\{A_j\}_{j=1}^{\infty} \subset \mathcal{A}$ are mutually disjoint, cover X, and satisfy $\mu_0(A_j) < \infty$.

Let $E \in \mathcal{M}$ be an arbitrary set. Because the sets A_j are mutually disjoint and cover X, we can slice E into a countable union of disjoint pieces:

$$E = E \cap X = E \cap \left(\bigcup_{j=1}^{\infty} A_j\right) = \bigcup_{j=1}^{\infty} (E \cap A_j)$$

Since μ and ν are both valid measures on $\mathcal{M}$, they are countably additive. We apply this property to expand both measures:

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E \cap A_j) \quad \text{and} \quad \nu(E) = \sum_{j=1}^{\infty} \nu(E \cap A_j)$$

For every index j, the set $E \cap A_j$ is a subset of A_j. By the monotonicity of measures, we have:

$$\mu(E \cap A_j) \le \mu(A_j) = \mu_0(A_j) < \infty$$

Because $\mu(E \cap A_j)$ is finite for every single term in the sum, we can apply our conclusion from Part 2 (which guarantees that ν and μ match exactly on any set where μ is finite):

$$\mu(E \cap A_j) = \nu(E \cap A_j) \quad \text{for all } j \geq 1$$

Substituting this term-by-term equivalence into our summation yields:

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E \cap A_j) = \sum_{j=1}^{\infty} \nu(E \cap A_j) = \nu(E)$$

Since $E \in \mathcal{M}$ was chosen arbitrarily, $\nu = \mu$, proving that the extension is completely unique. □

—

Remark (Failure of Uniqueness for Non-σ-Finite Premeasures)**.** *If the premeasure μ_0 is not σ-finite, the extension to $\mathcal{M}$ does not have to be unique. Here is a classic, straightforward counterexample that demonstrates this failure and confirms the minimality property of the induced measure μ.*

Let $X = \mathbb{R}$ be the set of real numbers, and let $\mathcal{A}$ be the standard algebra consisting of all finite unions of intervals of the form $(a, b]$ (where $-\infty \leq a \leq b \leq \infty$). The σ-algebra generated by this collection is the standard Borel σ-algebra, $\mathcal{M} = \mathcal{B}_{\mathbb{R}}$.

We define a set function $\mu_0 : \mathcal{A} \to [0, \infty]$ on this algebra as follows:

$$\mu_0(A) = \begin{cases} 0 & \text{if } A = \emptyset \\ \infty & \text{if } A \neq \emptyset \end{cases}$$

Let us verify the properties of this system:

1. ***It is a valid premeasure:*** *If $\{A_j\}_{j=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{A}$ such that $\bigcup A_j \in \mathcal{A}$, then:*
   - *If all $A_j = \emptyset$, then their union is empty, and $\mu_0(\emptyset) = 0 = \sum 0$.*
   - *If at least one set $A_k \neq \emptyset$, then the union is non-empty. This gives $\mu_0(\bigcup A_j) = \infty$, and the summation contains at least one ∞ term, which evaluates to ∞. Thus, countable additivity holds.*
2. ***It is not σ-finite:*** *Any non-empty set in the algebra has an infinite premeasure. Therefore, it is impossible to decompose $\mathbb{R}$ into a countable union of sets from $\mathcal{A}$ that have finite premeasure.*

Now, let us examine two distinct, valid extensions of this premeasure to the Borel σ-algebra $\mathcal{M} = \mathcal{B}_{\mathbb{R}}$:

- ***Extension 1 (The Minimal Induced Measure μ):*** *If we construct the outer measure induced by μ_0 using our standard formula, any non-empty set E covered by elements of $\mathcal{A}$ will result in a covering sum of ∞. However, if $E = \emptyset$, it can be covered by the empty set, which gives a measure of 0. Thus, the minimal extension $\mu = \mu^*|_{\mathcal{M}}$ evaluates to:*

  $$\mu(E) = \begin{cases} 0 & \text{if } E = \emptyset \\ \infty & \text{if } E \neq \emptyset \end{cases}$$

  This is known as the counting-type envelope measure, and it clearly matches μ_0 on $\mathcal{A}$.

- ***Extension 2 (The Counting Measure ν):*** *Let ν be the standard counting measure on $\mathbb{R}$, defined for any set $E \in \mathcal{B}_{\mathbb{R}}$ by:*

  $$\nu(E) = \text{number of elements in } E$$

  Let us check if ν matches μ_0 on the algebra $\mathcal{A}$:
  - *If $A = \emptyset$, then $\nu(\emptyset) = 0 = \mu_0(\emptyset)$.*
  - *If $A \in \mathcal{A}$ is non-empty, it must contain at least one non-empty interval $(a, b]$ with $a < b$. Because intervals contain uncountably many points, the counting measure evaluates to $\nu(A) = \infty$, which matches $\mu_0(A) = \infty$.*

  Therefore, ν is a completely valid alternative measure extension of μ_0 onto $\mathcal{B}_{\mathbb{R}}$.

Verifying Minimality

We can compare these two extensions to see that they behave exactly as predicted by our theorem:

- *Consider a single-point set, such as $E = \{5\}$. This set is Borel measurable ($E \in \mathcal{B}_{\mathbb{R}}$).*
- *Our minimal extension gives $\mu(\{5\}) = \infty$.*
- *The alternative counting measure extension gives $\nu(\{5\}) = 1$.*

This shows that $\nu(E) \leq \mu(E)$, which confirms that the induced extension μ is indeed the maximal possible baseline value among extensions here. Because $\nu(\{5\}) = 1 < \infty = \mu(\{5\})$, this simple example proves that uniqueness fails completely without the σ-finiteness condition.

Remark (The Relationship Between Induced Outer Measures, Completions, and Saturation)**.** *Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose we treat the measure μ as a premeasure on the algebra $\mathcal{M}$. This premeasure induces an outer measure μ^* on X defined for any set $E \subset X$ by:*

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(A_j) : A_j \in \mathcal{M} \text{ for all } j \geq 1, \text{ and } E \subset \bigcup_{j=1}^{\infty} A_j \right\}$$

Let $\mathcal{M}^*$ denote the σ-algebra of all μ^*-measurable sets in the sense of Carathéodory, and let $\overline{\mu} = \mu^*|_{\mathcal{M}^*}$ be the restriction of the outer measure to these measurable sets. The structural behavior of this extension $(\mathcal{M}^*, \overline{\mu})$ depends directly on whether the original measure space is σ-finite:

a. ***The σ-Finite Case (Completion):*** *If the original measure μ is σ-finite, then $(\mathcal{M}^*, \overline{\mu})$ is exactly the standard **completion** of the measure space $(X, \mathcal{M}, \mu)$.*

To understand why this happens, recall that the completion of a measure space is constructed by adding all subsets of sets of measure zero to the original σ-algebra. When μ is σ-finite, any set E with outer measure $\mu^*(E) = 0$ can be enclosed inside an original measurable set $N \in \mathcal{M}$ such that $\mu(N) = 0$. Under these conditions, the Carathéodory condition does not capture any sets beyond the standard null-set completion. Thus, every set in $\mathcal{M}^*$ can be written uniquely in the form $A \cup N$, where $A \in \mathcal{M}$ and N is a subset of a μ-null set.

b. ***The General Case (Saturation of the Completion):*** *If the original measure μ is not σ-finite, the Carathéodory construction $(\mathcal{M}^*, \overline{\mu})$ can produce a space strictly larger than the standard completion. Instead, it yields the **saturation** of the completion of μ.*

A measure space is said to be saturated (or maximally extended) if any set $E \subset X$ that intersects every set of finite measure in a measurable way is itself automatically declared to be measurable. When σ-finiteness fails, the induced outer measure $\mu^*(E)$ evaluates to ∞ for any set that cannot be covered by a countable collection of finite-measure sets.

As a result, the Carathéodory condition allows large, unwieldy sets to become elements of $\mathcal{M}^*$ as long as their intersections with all finite-measure chunks of the space look like standard measurable sets. This means the outer measure extension automatically forces the resulting space to be closed under these locally measurable structures, which matches the definition of a saturated measure space.

Borel Measures on the Real Line

Let $X = \mathbb{R}$ be the set of real numbers. A measure whose domain is the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$ is called a **Borel measure** on $\mathbb{R}$.

To build a definitive theory for measuring subsets of $\mathbb{R}$ based on the core geometric idea that the measure of an interval corresponds to its length, we look at the properties of a bounding function F. Let μ be a finite Borel measure on $\mathbb{R}$, and let $F : \mathbb{R} \to \mathbb{R}$ be defined by:

$$F(x) = \mu((-\infty, x])$$

Exercises on the Properties of F

Exercise 4. *Prove that the function $F(x) = \mu((-\infty, x])$ is monotonically increasing and right-continuous.*

Proof. We establish the two structural properties using the foundational axioms of a measure:

1. **Monotonically Increasing:** Let $x, y \in \mathbb{R}$ such that $x \leq y$. The interval $(-\infty, y]$ can be decomposed into the disjoint union of two measurable subsets:

$$(-\infty, y] = (-\infty, x] \cup (x, y]$$

By the finite additivity of the measure μ, we have:

$$\mu((-\infty, y]) = \mu((-\infty, x]) + \mu((x, y])$$

Substituting our definition of F yields $F(y) = F(x) + \mu((x, y])$. Since a measure is non-negative ($\mu((x, y]) \geq 0$), it follows that $F(x) \leq F(y)$, proving that F is monotonically increasing.

2. **Right-Continuous:** Let $x \in \mathbb{R}$ and consider any decreasing sequence $\{x_n\}_{n=1}^{\infty}$ that converges to x from the right ($x_n \searrow x$). We can construct a sequence of nested intervals $I_n = (-\infty, x_n]$. Since x_n is decreasing, the sets are nested downwards ($I_1 \supset I_2 \supset I_3 \supset \dots$). The countable intersection of these sets satisfies:

$$\bigcap_{n=1}^{\infty} (-\infty, x_n] = (-\infty, x]$$

Since μ is a finite measure ($\mu(X) < \infty$), we apply the continuity property of measures from above for nested sequences:

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mu((-\infty, x_n]) = \mu\left(\bigcap_{n=1}^{\infty} (-\infty, x_n]\right) = \mu((-\infty, x]) = F(x)$$

This demonstrates that $\lim_{x_n \searrow x} F(x_n) = F(x)$, establishing that F is right-continuous.

□

Remark (The Identity Function Case)**.** *The special canonical choice where $F(x) = x$ serves as the core foundation for geometric integration. Turning our construction around using this identity function yields the standard, classic **Lebesgue measure** on $\mathbb{R}$, where the measure of any half-open interval matches its geometric length precisely:*

$$\mu((a, b]) = F(b) - F(a) = b - a$$

—

Definition of h-intervals

The primary building blocks for constructing Borel measures are left-open, right-closed intervals in $\mathbb{R}$. We define an **h-interval** (short for "half-open interval") to be any subset of $\mathbb{R}$ belonging to the following collection:

$$\{(a, b] : -\infty \leq a < b < \infty\} \cup \{(a, \infty) : -\infty \leq a < \infty\} \cup \{\emptyset\}$$

—

Exercises on the Geometry of h-intervals

Exercise 5. *Let $\mathcal{A}$ be the collection of all subsets of $\mathbb{R}$ that can be expressed as a finite disjoint union of h-intervals. Prove that $\mathcal{A}$ forms an algebra of sets on $\mathbb{R}$.*

Proof. To prove that $\mathcal{A}$ is an algebra, we verify that it contains the empty set and is closed under both complements and finite unions:

1. **Contains the Empty Set:** By definition, $\emptyset$ is an h-interval. Thus, $\emptyset \in \mathcal{A}$ as a trivial single-element union.

2. **Closed Under Complements:** Let I be an h-interval. We observe its complement behavior:

- If $I = \emptyset$, then $I^c = (-\infty, \infty)$, which is the h-interval $(-\infty, \infty)$.
- If $I = (a, b]$, its complement is $(-\infty, a] \cup (b, \infty)$, which is a disjoint union of two valid h-intervals.
- If $I = (a, \infty)$, its complement is $(-\infty, a]$, which is a single valid h-interval.

Since the complement of any single h-interval is a finite disjoint union of h-intervals, it belongs to $\mathcal{A}$. By extension, if $A = \bigcup_{i=1}^{n} I_i \in \mathcal{A}$ is a disjoint union, its complement is the intersection of these individual complements:

$$A^c = \bigcap_{i=1}^{n} I_i^c$$

Because the intersection of any two h-intervals is clearly another h-interval, distributing this finite intersection across the finite unions of h-intervals demonstrates that A^c can be simplified back into a finite disjoint union of h-intervals. Thus, $A^c \in \mathcal{A}$.

3. **Closed Under Finite Unions:** Let $A, B \in \mathcal{A}$. We can express B as a finite disjoint union of h-intervals. The union $A \cup B$ can be rewritten by parsing out any overlapping regions:

$$A \cup B = A \cup (B \cap A^c)$$

Since $\mathcal{A}$ is closed under complements and intersections, $B \cap A^c \in \mathcal{A}$. Expressing this piece as a collection of disjoint h-intervals that do not overlap with A allows us to append them directly, showing that $A \cup B$ remains a finite disjoint union of h-intervals. Hence, $A \cup B \in \mathcal{A}$.

Therefore, $\mathcal{A}$ forms a valid algebra of sets on $\mathbb{R}$. □

Exercise 6. *Prove that the σ-algebra generated by the algebra $\mathcal{A}$ of finite disjoint unions of h-intervals is exactly the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$.*

Proof. To demonstrate that $\sigma(\mathcal{A}) = \mathcal{B}_{\mathbb{R}}$, we prove containment in both directions:

1. **Proving $\sigma(\mathcal{A}) \subset \mathcal{B}_{\mathbb{R}}$:** By definition, the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$ is generated by all open sets in $\mathbb{R}$. Any open interval (a, b) is a Borel set. A standard bounded h-interval $(a, b]$ can be viewed as the intersection of an open interval and a closed set:

$$(a, b] = (a, b + 1) \cap (-\infty, b]$$

Since open and closed sets are elements of $\mathcal{B}_{\mathbb{R}}$, every h-interval is a Borel set. Because $\mathcal{B}_{\mathbb{R}}$ is a σ-algebra containing all individual h-intervals, it must contain their finite unions. Thus, $\mathcal{A} \subset \mathcal{B}_{\mathbb{R}}$, which implies that the generated structure satisfies $\sigma(\mathcal{A}) \subset \mathcal{B}_{\mathbb{R}}$.

2. **Proving $\mathcal{B}_{\mathbb{R}} \subset \sigma(\mathcal{A})$:** It suffices to show that every open interval (a, b) belongs to $\sigma(\mathcal{A})$, since open intervals generate all open sets in $\mathbb{R}$. Let (a, b) be an arbitrary open interval. We can express it as a countable union of increasing h-intervals by pulling the right boundary slightly inward:

$$(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \frac{1}{n}\right]$$

For each n, the set $\left(a, b - \frac{1}{n}\right]$ is a valid h-interval, meaning it belongs to the algebra $\mathcal{A}$. Since a σ-algebra is closed under countable unions, it follows that the open interval (a, b) must belong to $\sigma(\mathcal{A})$. Therefore, $\mathcal{B}_{\mathbb{R}} \subset \sigma(\mathcal{A})$.

Combining both containment directions yields $\sigma(\mathcal{A}) = \mathcal{B}_{\mathbb{R}}$. □

Construction of Borel Measures via Cumulative Functions

Let $\mathcal{A}$ be the algebra on $\mathbb{R}$ consisting of all subsets that can be expressed as a finite disjoint union of h-intervals (half-open intervals of the form $(a, b]$, (a, ∞), or $\emptyset$). To construct a unique Borel measure on $\mathbb{R}$, we begin by defining a localized set function on this base algebra using a right-continuous, monotonically increasing function.

Proposition (Proposition 1.15)**.** *Let $F : \mathbb{R} \to \mathbb{R}$ be a monotonically increasing and right-continuous function. If $I_j = (a_j, b_j]$ ($j = 1, \dots, n$) are mutually disjoint bounded h-intervals, we define:*

$$\mu_0\left(\bigcup_{j=1}^{n}(a_j, b_j]\right) = \sum_{j=1}^{n}[F(b_j) - F(a_j)]$$

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

—

Proof. The proof requires establishing two main properties: first, that μ_0 is completely independent of the choice of interval representation (well-defined and finitely additive); second, that it satisfies countable additivity for sequences of disjoint sets whose union remains within the algebra $\mathcal{A}$.

Part 1: Well-Definedness and Finite Additivity

Because an element $A \in \mathcal{A}$ can be decomposed into disjoint h-intervals in more than one way, we must prove that its assigned value $\mu_0(A)$ is unique.

- **Single Interval Case:** Suppose a single h-interval $I = (a, b]$ is partitioned into a finite collection of disjoint sub-h-intervals $\{I_j\}_{j=1}^n = \{(a_j, b_j]\}_{j=1}^n$. After sorting and relabeling the indices, the intervals must align end-to-end across the main span:

  $$a = a_1 < b_1 = a_2 < b_2 = \cdots < b_n = b$$

  Evaluating the sum of the increments of F results in a telescoping series:

  $$\sum_{j=1}^{n}[F(b_j) - F(a_j)] = [F(b_1) - F(a)] + [F(b_2) - F(b_1)] + \cdots + [F(b) - F(a_n)]$$
  $$= F(b) - F(a)$$

  Thus, the sum matches the single interval definition perfectly.

- **General Set Case:** Let $A \in \mathcal{A}$ be represented by two different finite collections of disjoint h-intervals, say $\{I_i\}_{i=1}^n$ and $\{J_k\}_{k=1}^m$, such that $\bigcup_i I_i = \bigcup_k J_k$. Because the intersections $I_i \cap J_k$ are themselves h-intervals, we can employ our single interval telescoping rule twice:

  $$\sum_i \mu_0(I_i) = \sum_i \left(\sum_k \mu_0(I_i \cap J_k)\right) = \sum_{i,k} \mu_0(I_i \cap J_k)$$

  Reversing the order of summation yields:

  $$\sum_{i,k} \mu_0(I_i \cap J_k) = \sum_k \left(\sum_i \mu_0(I_i \cap J_k)\right) = \sum_k \mu_0(J_k)$$

  This guarantees that μ_0 is well-defined and automatically finitely additive by construction.

—

Part 2: Proving Countable Additivity

Let $\{I_j\}_{j=1}^\infty$ be a sequence of mutually disjoint h-intervals in $\mathcal{A}$ such that their infinite union forms a single cohesive element $I \in \mathcal{A}$. That is, $I = \bigcup_{j=1}^\infty I_j$. We assume without loss of generality that I is a bounded interval of the form $(a, b]$. We must establish that $\mu_0(I) = \sum_{j=1}^\infty \mu_0(I_j)$ by proving inequalities in both directions.

Step 2a: Establishing the Lower Bound ($\mu_0(I) \geq \sum_{j=1}^{\infty} \mu_0(I_j)$)

For any finite integer $n \geq 1$, the partial union $\bigcup_{j=1}^{n} I_j$ forms a subset of the total interval I. Since the remainder $I \setminus \bigcup_{j=1}^{n} I_j$ can be expressed as a finite disjoint union of h-intervals, the finite additivity of μ_0 implies:

$$\mu_0(I) = \mu_0\left(\bigcup_{j=1}^{n} I_j\right) + \mu_0\left(I \setminus \bigcup_{j=1}^{n} I_j\right) \geq \mu_0\left(\bigcup_{j=1}^{n} I_j\right) = \sum_{j=1}^{n} \mu_0(I_j)$$

Taking the limit as $n \to \infty$ on the right side preserves the inequality:

$$\mu_0(I) \geq \sum_{j=1}^{\infty} \mu_0(I_j)$$

Step 2b: Establishing the Upper Bound ($\mu_0(I) \leq \sum_{j=1}^{\infty} \mu_0(I_j)$)

To prove the reverse direction, we construct a compact topological space to force a finite cover from the infinite collection. Fix an arbitrary tolerance parameter $\epsilon > 0$.

1. **Shrinking the Domain Interval:** Since F is right-continuous, there exists an associated parameter $\delta > 0$ small enough such that:

$$F(a + \delta) - F(a) < \epsilon$$

   This allows us to look at the strictly smaller, closed interval $[a + \delta, b]$.

2. **Expanding the Covering Intervals:** Let each disjoint piece in the sequence be denoted by $I_j = (a_j, b_j]$. Using the right-continuity of F on each individual sub-component, for every index $j \geq 1$, there exists a chosen $\delta_j > 0$ such that:

$$F(b_j + \delta_j) - F(b_j) < \epsilon 2^{-j}$$

   We can convert each original h-interval into a larger open interval:

$$I'_j = (a_j, b_j + \delta_j)$$

Notice that the collection of open intervals $\{I'_j\}_{j=1}^{\infty}$ covers our constructed compact set:

$$[a + \delta, b] \subset (a, b] = \bigcup_{j=1}^{\infty} (a_j, b_j] \subset \bigcup_{j=1}^{\infty} (a_j, b_j + \delta_j)$$

By the Heine-Borel Theorem, every open cover of a compact set contains a finite subcover. Thus, there exists a large finite integer N such that:

$$[a + \delta, b] \subset \bigcup_{j=1}^{N} (a_j, b_j + \delta_j)$$

To systematically evaluate this finite cover, we can eliminate redundant intervals that are entirely contained within larger ones. After sorting and relabeling, we assume the elements satisfy the interlocking condition:

- The intervals $(a_1, b_1 + \delta_1), \ldots, (a_N, b_N + \delta_N)$ cover the set $[a + \delta, b]$.
- To ensure continuous coverage across the boundary points, we must have $a_1 < a + \delta$ and $b_N + \delta_N > b$, with the intermediate spans interlocking such that $b_j + \delta_j \in (a_{j+1}, b_{j+1} + \delta_{j+1})$ for all $j = 1, \ldots, N - 1$.

We now calculate an upper bound for the total measure $\mu_0(I)$ by utilizing the monotonic behavior of F and tracking these overlapping finite increments:

$$\mu_0(I) = F(b) - F(a) < F(b) - F(a + \delta) + \epsilon$$

Since $b < b_N + \delta_N$ and $a + \delta > a_1$, this can be bounded by:

$$\leq F(b_N + \delta_N) - F(a_1) + \epsilon$$

We insert the intermediate boundary values into the expression to form a series:

$$= F(b_N + \delta_N) - F(a_N) + \sum_{j=1}^{N-1} [F(a_{j+1}) - F(a_j)] + \epsilon$$

Because the intervals interlock ($a_{j+1} < b_j + \delta_j$), the monotonicity of F gives $F(a_{j+1}) \leq F(b_j + \delta_j)$. Substituting this bound yields:

$$\leq F(b_N + \delta_N) - F(a_N) + \sum_{j=1}^{N-1} [F(b_j + \delta_j) - F(a_j)] + \epsilon$$

$$= \sum_{j=1}^{N} [F(b_j + \delta_j) - F(a_j)] + \epsilon$$

We isolate our engineered delta parameters from the main interval evaluations:

$$< \sum_{j=1}^{N} [F(b_j) + \epsilon 2^{-j} - F(a_j)] + \epsilon$$

$$= \sum_{j=1}^{N} \mu_0(I_j) + \epsilon \sum_{j=1}^{N} 2^{-j} + \epsilon$$

Extending the finite sum to an infinite series reveals our final structural bound:

$$\mu_0(I) < \sum_{j=1}^{\infty} \mu_0(I_j) + 2\epsilon$$

Since equation (3) holds true for any arbitrarily chosen parameter $\epsilon > 0$, taking the limit as $\epsilon \to 0$ yields:

$$\mu_0(I) \leq \sum_{j=1}^{\infty} \mu_0(I_j)$$

Combining the inequalities from Equation (1) and Equation (4) proves complete countable additivity:

$$\mu_0(I) = \sum_{j=1}^{\infty} \mu_0(I_j)$$

Handling Unbounded Limits ($a = -\infty$ or $b = \infty$)

If the domain interval extends to infinity, we can verify the limit using a sequence of expanding compact sets:

- If $a = -\infty$, for an arbitrarily large cutoff parameter $M < \infty$, the open intervals cover the truncated set $[-M, b]$. The same interlocking calculation yields $F(b) - F(-M) \leq \sum_1^\infty \mu_0(I_j) + 2\epsilon$. Letting $M \to \infty$ and $\epsilon \to 0$ recovers the full expression.
- If $b = \infty$, evaluating the cover across the truncated segment $[a + \delta, M]$ yields the inequality $F(M) - F(a) \leq \sum_1^\infty \mu_0(I_j) + 2\epsilon$. Letting $M \to \infty$ ensures the theorem holds across all boundary types.

Thus, μ_0 is a valid premeasure on $\mathcal{A}$. $\square$

The Correspondence Between Distribution Functions and Borel Measures

Theorem 7. *If $F : \mathbb{R} \to \mathbb{R}$ is any increasing, right-continuous function, there is a unique Borel measure μ_F on $\mathbb{R}$ such that $\mu_F((a, b]) = F(b) - F(a)$ for all $a, b \in \mathbb{R}$ with $a < b$. If G is another such function, then $\mu_F = \mu_G$ if and only if $F - G$ is constant.*

Conversely, if μ is a Borel measure on $\mathbb{R}$ that is finite on all bounded Borel sets and we define:

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

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

Proof. The verification is split into three main steps: establishing the existence and uniqueness of μ_F from F, characterizing when two functions yield the same measure, and constructing F from a given locally finite measure μ.

Step 1: Existence and Uniqueness of μ_F

Let $\mathcal{A}$ be the algebra on $\mathbb{R}$ consisting of all finite disjoint unions of half-open intervals of the form $(a, b]$, (a, ∞), or $\emptyset$.

- **Premeasure Assignment:** We define an assignment μ_0 on the intervals of $\mathcal{A}$ by setting $\mu_0((a, b]) = F(b) - F(a)$ and $\mu_0(\emptyset) = 0$. Because F is increasing and right-continuous, this assignment defines a well-behaved, countably additive premeasure on the algebra $\mathcal{A}$.
- **Establishing σ-Finiteness:** To guarantee that this premeasure can be uniquely extended to the entire Borel σ-algebra, we must show that the space is σ-finite. We can decompose the entire real line into a countable union of bounded, adjacent half-open intervals:

  $$\mathbb{R} = \bigcup_{j=-\infty}^{\infty} (j, j + 1]$$

  For each integer j, the interval $(j, j + 1]$ belongs to $\mathcal{A}$, and its premeasure value is $\mu_0((j, j + 1]) = F(j + 1) - F(j)$. Since F maps strictly to real numbers, this difference is finite for every $j \in \mathbb{Z}$. Thus, μ_0 is a σ-finite premeasure.
- **Carathéodory Extension:** By the Carathéodory Extension Theorem, any σ-finite premeasure on an algebra $\mathcal{A}$ possesses a unique extension to a full measure on the generated σ-algebra $\sigma(\mathcal{A})$. Since the collection of half-open intervals generates the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$, there exists a unique Borel measure μ_F extending μ_0, which satisfies $\mu_F((a, b]) = F(b) - F(a)$.

Step 2: Equivalence of Functions ($\mu_F = \mu_G \iff F - G = C$)

- **Direction ($\Longleftarrow$):** Suppose $F(x) - G(x) = C$ for some constant $C \in \mathbb{R}$, which implies $G(x) = F(x) - C$. Evaluating the measure of any half-open interval under G gives:

  $$\mu_G((a, b]) = G(b) - G(a) = [F(b) - C] - [F(a) - C] = F(b) - F(a) = \mu_F((a, b])$$

  Since their induced premeasures match identically on the generating algebra $\mathcal{A}$, their unique extensions to $\mathcal{B}_{\mathbb{R}}$ must also be identical ($\mu_F = \mu_G$).
- **Direction ($\Longrightarrow$):** Conversely, suppose $\mu_F = \mu_G$. For any bounded interval $(a, b]$, this implies:

  $$F(b) - F(a) = G(b) - G(a) \implies F(b) - G(b) = F(a) - G(a)$$

  Fixing the reference point $a = 0$ and letting $C = F(0) - G(0)$, substituting any arbitrary point $b = x$ into the equation yields:

  $$F(x) - G(x) = C$$

  Thus, $F - G$ is a constant function over the entire real line.

Step 3: Deriving the Converse

Let μ be a Borel measure on $\mathbb{R}$ that is finite on all bounded sets. We analyze the function F defined sectionally in the theorem statement.

3.1 Monotonicity of F We must verify that $F(x) \leq F(y)$ whenever $x \leq y$. We consider three cases based on the positions of x and y relative to 0:

1. **Case** $0 \leq x \leq y$**:** Here, $F(x) = \mu((0, x])$ and $F(y) = \mu((0, y])$. We can split the larger interval into two disjoint components: $(0, y] = (0, x] \cup (x, y]$. By the finite additivity of measures:
$$\mu((0, y]) = \mu((0, x]) + \mu((x, y]) \implies F(y) = F(x) + \mu((x, y])$$
Since measures are non-negative ($\mu((x, y]) \geq 0$), it follows directly that $F(x) \leq F(y)$.

2. **Case** $x < 0 < y$**:** By definition, $F(x) = -\mu((-x, 0])$, which is less than or equal to 0. Conversely, $F(y) = \mu((0, y]) \geq 0$. Thus, $F(x) \leq 0 \leq F(y)$ holds trivially.

3. **Case** $x \leq y \leq 0$**:** Here, $F(x) = -\mu((-x, 0])$ and $F(y) = -\mu((-y, 0])$. Multiplying the inequality $x \leq y \leq 0$ by -1 reverses the order: $0 \leq -y \leq -x$. This sets up a nested containment of intervals: $(-y, 0] \subset (-x, 0]$. By the monotonicity of measures:
$$\mu((-y, 0]) \leq \mu((-x, 0]) \implies -\mu((-x, 0]) \leq -\mu((-y, 0]) \implies F(x) \leq F(y)$$

Thus, F is a monotonically increasing function.

3.2 Right-Continuity of F To prove right-continuity, we must establish that $\lim_{n\to\infty} F(x_n) = F(x)$ for any decreasing sequence $x_n \searrow x$.

1. **Subcase** $x \geq 0$**:** For sufficiently large n, we have $x_n > 0$. Consider the sequence of shrinking, nested intervals $I_n = (0, x_n]$. As $x_n \searrow x$, their intersection yields $\bigcap_{n=1}^{\infty}(0, x_n] = (0, x]$. Because the initial interval $(0, x_1]$ is bounded, its measure is finite. Applying the continuity of measures from above for nested sequences:
$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mu((0, x_n]) = \mu\left(\bigcap_{n=1}^{\infty}(0, x_n]\right) = \mu((0, x]) = F(x)$$

2. **Subcase** $x < 0$**:** For sufficiently large n, x_n is negative. Here, $F(x_n) = -\mu((-x_n, 0])$. As $x_n \searrow x$, their negations increase to $-x$ (i.e., $-x_n \nearrow -x$). This forms an expanding sequence of nested sets $J_n = (-x_n, 0]$ whose countable union is $\bigcup_{n=1}^{\infty}(-x_n, 0] = (-x, 0]$. Applying the continuity of measures from below for expanding sequences:
$$\lim_{n\to\infty} F(x_n) = -\lim_{n\to\infty} \mu((-x_n, 0]) = -\mu\left(\bigcup_{n=1}^{\infty}(-x_n, 0]\right) = -\mu((-x, 0]) = F(x)$$

Thus, F is right-continuous everywhere.

3.3 Verifying $\mu = \mu_F$ Finally, we show that $\mu_F((a, b]) = \mu((a, b])$ for any bounded interval. By the construction of μ_F, we know $\mu_F((a, b]) = F(b) - F(a)$. We test the interval positions relative to the origin:

- **If** $0 < a < b$**:**
$$F(b) - F(a) = \mu((0, b]) - \mu((0, a]) = \mu((0, b] \setminus (0, a]) = \mu((a, b])$$

- **If** $a < 0 < b$**:**
$$F(b) - F(a) = \mu((0, b]) - [-\mu((-a, 0])] = \mu((0, b]) + \mu((a, 0])$$
Since $(a, 0]$ and $(0, b]$ are disjoint and their union is $(a, b]$, additivity yields $\mu((a, b])$.

- **If** $a < b < 0$**:**

$$F(b) - F(a) = -\mu((-b, 0]) - [-\mu((-a, 0])] = \mu((-a, 0]) - \mu((-b, 0])$$

  Since $a < b < 0$, we have $0 < -b < -a$, meaning $(-a, 0] \setminus (-b, 0] = (-a, -b]$. The measure of this reflected interval is equivalent to $\mu((a, b])$.

Because μ and μ_F match perfectly on all half-open bounded intervals, they agree across the entire generating algebra $\mathcal{A}$. By the uniqueness of the Carathéodory extension, they must be globally identical on the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$, meaning $\mu = \mu_F$. □

Remarks on the Integration and Measure Correspondence

Remark. *The theory could equally well be developed by using intervals of the form $[a, b)$ and left-continuous functions F. In that framework, the left-continuity of F guarantees that the set function matches the geometric requirements of the intervals properly.*

Remark. *If μ is a finite Borel measure on $\mathbb{R}$, then $\mu = \mu_F$ where $F(x) = \mu((-\infty, x])$ is the standard cumulative distribution function of μ. This classical choice differs from the F specified in the main theorem by the constant $\mu((-\infty, 0])$. Because adding or subtracting a constant shift does not change the differences $F(b) - F(a)$, both choices generate the exact same Borel measure.*

Remark. *The outer measure construction gives, for each increasing and right-continuous function F, not only the Borel measure μ_F but a complete measure $\overline{\mu}_F$ whose domain includes the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$. In fact, $\overline{\mu}_F$ is just the completion of μ_F, and one can show that its domain is always strictly larger than $\mathcal{B}_{\mathbb{R}}$.*

Remark. *We shall usually denote this complete measure also by μ_F. It is called the* ***Lebesgue-Stieltjes measure*** *associated to F.*

Regularity Properties of Lebesgue-Stieltjes Measures

Fix a complete Lebesgue-Stieltjes measure μ on $\mathbb{R}$ associated to an increasing, right-continuous function F. We denote the domain of this complete measure by $\mathcal{M}_\mu$.

For any set $E \in \mathcal{M}_\mu$, the measure $\mu(E)$ can be characterized from the outer measure construction by covering E with a countable sequence of half-open intervals $(a_j, b_j]$. This yields the following equivalent infimum expressions:

$$\begin{aligned}\mu(E) &= \inf \left\{ \sum_{j=1}^{\infty} [F(b_j) - F(a_j)] : E \subset \bigcup_{j=1}^{\infty} (a_j, b_j] \right\} \\ &= \inf \left\{ \sum_{j=1}^{\infty} \mu((a_j, b_j]) : E \subset \bigcup_{j=1}^{\infty} (a_j, b_j] \right\}\end{aligned}$$

In the second formulation, the half-open intervals $(a_j, b_j]$ can also be replaced by open intervals without changing the value of the infimum.

Lemma. *For any set $E \in \mathcal{M}_\mu$, the Lebesgue-Stieltjes measure can be equivalently calculated by covering E with open intervals instead of half-open h-intervals:*

$$\mu(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu((a_j, b_j)) : E \subset \bigcup_{j=1}^{\infty} (a_j, b_j) \right\}$$

Proof. Let us denote the quantity on the right side of the equation (the infimum over open interval covers) by $\nu(E)$. We must show that $\mu(E) = \nu(E)$ by proving both inequalities $\nu(E) \geq \mu(E)$ and $\nu(E) \leq \mu(E)$.

Part 1: Proving $\nu(E) \geq \mu(E)$

Suppose $\{(a_j, b_j)\}_{j=1}^{\infty}$ is an arbitrary countable collection of open intervals that covers the set E, so that:

$$E \subset \bigcup_{j=1}^{\infty} (a_j, b_j)$$

Each open interval (a_j, b_j) can be decomposed into a countable disjoint union of half-open h-intervals, say $I_j^k = (c_j^k, c_j^{k+1}]$ for $k = 1, 2, \ldots$. Specifically, for each interval j, we can choose a strictly increasing sequence $\{c_j^k\}_{k=1}^{\infty}$ such that the initial point starts at the left boundary, $c_j^1 = a_j$, and the sequence limits directly to the right boundary, $c_j^k \to b_j$ as $k \to \infty$.

By constructing these sequences for every j, we can express the total cover as a countable collection over both indices:

$$E \subset \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} I_j^k$$

Since μ is a countably additive measure on its σ-algebra $\mathcal{M}_\mu$, the measure of the open interval is exactly equal to the sum of the measures of its disjoint h-interval components:

$$\mu((a_j, b_j)) = \sum_{k=1}^{\infty} \mu(I_j^k)$$

Summing over all open intervals j, and using the countable subadditivity property of the outer measure μ, we obtain:

$$\sum_{j=1}^{\infty} \mu((a_j, b_j)) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \mu(I_j^k) \geq \mu(E)$$

Because this inequality holds true for every possible open interval cover of E, taking the infimum over all such open covers yields:

$$\nu(E) \geq \mu(E)$$

Part 2: Proving $\nu(E) \leq \mu(E)$

To prove the reverse inequality, we use an ϵ-argument combined with the right-continuity property of the distribution function F. Let $\epsilon > 0$ be given.

By the original definition of the measure $\mu(E)$ as an infimum over half-open h-intervals, there must exist a half-open covering collection $\{(a_j, b_j]\}_{j=1}^{\infty}$ such that:

$$E \subset \bigcup_{j=1}^{\infty} (a_j, b_j] \quad \text{and} \quad \sum_{j=1}^{\infty} \mu((a_j, b_j]) \leq \mu(E) + \epsilon$$

Because F is right-continuous everywhere, the measure μ is right-continuous on interval endpoints. For each interval index j, we can find a small expansion factor $\delta_j > 0$ such that expanding the right endpoint slightly increases the interval's measure by less than $\epsilon 2^{-j}$:

$$F(b_j + \delta_j) - F(b_j) < \epsilon 2^{-j} \implies \mu((a_j, b_j + \delta_j]) - \mu((a_j, b_j]) < \epsilon 2^{-j}$$

Now, consider the slightly enlarged open intervals $(a_j, b_j + \delta_j)$. Since $(a_j, b_j] \subset (a_j, b_j + \delta_j)$ for every j, these open intervals form a valid open cover for our original set E:

$$E \subset \bigcup_{j=1}^{\infty} (a_j, b_j + \delta_j)$$

By the definition of $\nu(E)$ as the absolute infimum over all open covers, the sum of the measures of this specific open cover must be greater than or equal to $\nu(E)$:

$$\nu(E) \leq \sum_{j=1}^{\infty} \mu((a_j, b_j + \delta_j))$$

Using the monotonicity of measures, we know $\mu((a_j, b_j + \delta_j)) \leq \mu((a_j, b_j + \delta_j])$. Substituting this in and expanding the terms gives:

$$\nu(E) \leq \sum_{j=1}^{\infty} \mu((a_j, b_j + \delta_j])$$

$$\nu(E) \leq \sum_{j=1}^{\infty} \left[\mu((a_j, b_j]) + \epsilon 2^{-j}\right]$$

$$\nu(E) \leq \sum_{j=1}^{\infty} \mu((a_j, b_j]) + \epsilon \sum_{j=1}^{\infty} 2^{-j}$$

Since the geometric series $\sum_{j=1}^{\infty} 2^{-j} = 1$, and our chosen half-open cover was bounded by $\mu(E) + \epsilon$, we can substitute it into the expression:

$$\nu(E) \leq \mu(E) + \epsilon + \epsilon = \mu(E) + 2\epsilon$$

Because $\epsilon > 0$ can be chosen to be arbitrarily small, this forces:

$$\nu(E) \leq \mu(E)$$

Combining Parts 1 and 2, we conclude that $\nu(E) = \mu(E)$, completing the proof. □

Theorem 8. *If $E \in \mathcal{M}_\mu$, then the Lebesgue-Stieltjes measure μ is both outer regular and inner regular:*

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

Proof. We split the proof into two independent sections: first establishing the outer regularity via open sets, and second proving the inner regularity via compact approximations.

Part 1: Verification of Outer Regularity

We wish to show that $\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ open}\}$. Let us denote this infimum by $\nu_o(E)$.

- **Direction $\nu_o(E) \geq \mu(E)$:** Let U be any arbitrary open set such that $U \supset E$. By the monotonicity property of measures, since $E \subset U$, it follows directly that $\mu(E) \leq \mu(U)$. Taking the infimum over all such open supersets U preserves the inequality, yielding:

$$\mu(E) \leq \nu_o(E)$$

- **Direction $\nu_o(E) \leq \mu(E)$:** Let $\epsilon > 0$ be given. By our previous lemma, we can approximate the measure of any set E arbitrarily closely using a countable cover of open intervals. Thus, there exists a sequence of open intervals $\{(a_j, b_j)\}_{j=1}^{\infty}$ such that:

$$E \subset \bigcup_{j=1}^{\infty} (a_j, b_j) \quad \text{and} \quad \sum_{j=1}^{\infty} \mu((a_j, b_j)) \leq \mu(E) + \epsilon$$

  Let us define the set $U = \bigcup_{j=1}^{\infty} (a_j, b_j)$. Because U is a countable union of open intervals, U is inherently an open set. Furthermore, by construction, $U \supset E$.

  Applying the countable subadditivity of measures to the union defining U, we see that:

$$\mu(U) \leq \sum_{j=1}^{\infty} \mu((a_j, b_j)) \leq \mu(E) + \epsilon$$

  Since U is a specific open set containing E, the infimum over all such open sets must be less than or equal to $\mu(U)$:

$$\nu_o(E) \leq \mu(U) \leq \mu(E) + \epsilon$$

  Because $\epsilon > 0$ is completely arbitrary, letting $\epsilon \to 0$ forces $\nu_o(E) \leq \mu(E)$.

Combining both directions yields the first equality: $\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ open}\}$.

Part 2: Verification of Inner Regularity

We now wish to show that $\mu(E) = \sup\{\mu(K) : K \subset E, K \text{ compact}\}$. We break this analysis down into two scenarios based on whether E is bounded or unbounded.

Case 2.1: Suppose E is bounded Let $\overline{E}$ denote the topological closure of E. Since E is bounded, its closure $\overline{E}$ is both closed and bounded, making $\overline{E}$ a compact set under the Heine-Borel theorem.

- **Subcase if E is closed:** If E is already closed, then since it is bounded, E is itself compact. We can simply choose $K = E$. Then $K \subset E$ is compact and $\mu(K) = \mu(E)$, making the supremum equality trivially true.
- **Subcase if E is not closed:** Let $\epsilon > 0$ be given. Consider the set difference $\overline{E} \setminus E$. Because E is bounded, $\overline{E} \setminus E$ is a bounded set. We apply the outer regularity property established in Part 1 to this set. Thus, there exists an open set U such that:

  $$U \supset (\overline{E} \setminus E) \quad \text{and} \quad \mu(U) \leq \mu(\overline{E} \setminus E) + \epsilon$$

  Let us construct our compact inner approximation candidate by setting:

  $$K = \overline{E} \setminus U$$

  Because $\overline{E}$ is closed and U is open, $K = \overline{E} \cap U^c$ is a intersection of two closed sets, hence K is closed. Since $K \subset \overline{E}$ and $\overline{E}$ is bounded, K is a closed and bounded set, making K compact.

  Furthermore, we check that $K \subset E$. If $x \in K$, then $x \in \overline{E}$ and $x \notin U$. Since U contains all points of $\overline{E}$ that are *not* in E, any point in $\overline{E}$ that avoids U must necessarily belong to E. Thus, $K \subset E$.

  Now we compute the measure of K. Since $K = \overline{E} \setminus U$ and $K \cup (\overline{E} \cap U) = \overline{E}$ forms a disjoint partitioning of $\overline{E}$, additivity gives:

  $$\mu(K) = \mu(\overline{E}) - \mu(\overline{E} \cap U)$$

  Using the identity $\overline{E} \cap U = (E \cap U) \cup ((\overline{E} \setminus E) \cap U)$, and noting that $\overline{E} \setminus E \subset U$ simplifies the second component to just $\overline{E} \setminus E$, we can expand the measure expression:

  $$\mu(\overline{E} \cap U) = \mu(E \cap U) + \mu(\overline{E} \setminus E)$$

  Substituting this back into our calculation for $\mu(K)$ yields:

  $$\mu(K) = \mu(\overline{E}) - \mu(\overline{E} \setminus E) - \mu(E \cap U)$$

  Since $\mu(\overline{E}) - \mu(\overline{E} \setminus E) = \mu(E)$, this simplifies directly to:

  $$\mu(K) = \mu(E) - \mu(E \cap U)$$

  From our original open set selection, we know that $\mu(U) \leq \mu(\overline{E} \setminus E) + \epsilon$. Rewriting this via subtraction rules, we have $\mu(U) - \mu(U \setminus \overline{E}) \leq \mu(\overline{E} \setminus E) + \epsilon$, which implies:

  $$\mu(E \cap U) \leq \epsilon$$

  Applying this upper bound to our subtraction formula for K yields:

  $$\mu(K) = \mu(E) - \mu(E \cap U) \geq \mu(E) - \epsilon$$

  Since we found a compact set $K \subset E$ whose measure is within ϵ of $\mu(E)$, taking the supremum over all compact subsets confirms that the supremum equals $\mu(E)$.

Case 2.2: Suppose E is unbounded If E spans infinitely, we slice the real line into a countable collection of bounded, disjoint blocks. For each integer j, let $E_j = E \cap (j, j+1]$. Each slice E_j is a bounded set.

Let $\epsilon > 0$. Since each E_j is bounded, we can apply the result from Case 2.1 to each slice. For every index j, there exists a compact subset $K_j \subset E_j$ such that:

$$\mu(K_j) \geq \mu(E_j) - \epsilon 2^{-|j|-1}$$

For a fixed natural number n, let us assemble a finite union of these compact components:

$$H_n = \bigcup_{j=-n}^{n} K_j$$

Because H_n is a finite union of compact sets, H_n is itself a compact set. Furthermore, since each $K_j \subset E_j \subset E$, it is clear that $H_n \subset E$.

Since the slices $(j, j+1]$ are pairwise disjoint, the sets K_j are mutually disjoint. Therefore, the measure of their union is the sum of their individual measures:

$$\mu(H_n) = \sum_{j=-n}^{n} \mu(K_j) \geq \sum_{j=-n}^{n} \left(\mu(E_j) - \epsilon 2^{-|j|-1}\right) \geq \mu\left(\bigcup_{j=-n}^{n} E_j\right) - \epsilon$$

Now, consider the behavior as $n \to \infty$. The expanding finite unions grow monotonically to cover the entire set:

$$\lim_{n\to\infty} \bigcup_{j=-n}^{n} E_j = E$$

By the continuity of measures from below for increasing sequences of sets, taking the limit as $n \to \infty$ on both sides gives:

$$\sup_n \mu(H_n) \geq \mu(E) - \epsilon$$

Since $\epsilon > 0$ was arbitrary, we can find compact sets $H_n \subset E$ with measures arbitrarily close to $\mu(E)$ if $\mu(E) < \infty$, or limiting to ∞ if $\mu(E) = \infty$. Thus, the supremum over all compact subsets is precisely equal to $\mu(E)$. □

Theorem 9. *Let $E \subset \mathbb{R}$. The following topological and measure-theoretic conditions are completely equivalent:*

1. *$E \in \mathcal{M}_\mu$ (i.e., E is Lebesgue-Stieltjes measurable).*
2. *$E = V \setminus N_1$, where V is a G_δ set and $\mu(N_1) = 0$.*
3. *$E = H \cup N_2$, where H is an F_σ set and $\mu(N_2) = 0$.*

Proof. We will construct the complete proof by establishing the cycle of implications: (2) $\implies$ (1), (3) $\implies$ (1), and finally (1) $\implies$ (2) and (3).

Part 1: Structural Sufficiency — proving (2) $\implies$ (1) and (3) $\implies$ (1)

- **Proving (2) $\implies$ (1):** Suppose $E = V \setminus N_1$, where V is a G_δ set and $\mu(N_1) = 0$. By definition, a G_δ set is a countable intersection of open sets. Because open sets belong to the Borel σ-algebra $\mathcal{B}_\mathbb{R}$, and σ-algebras are stable under countable intersections, V must be a Borel set ($V \in \mathcal{B}_\mathbb{R} \subset \mathcal{M}_\mu$).

  Furthermore, because μ is a complete measure space on $\mathcal{M}_\mu$, any subset of a set with measure zero is automatically measurable and has measure zero. Thus, the null set N_1 satisfies $N_1 \in \mathcal{M}_\mu$. Since $\mathcal{M}_\mu$ is a σ-algebra, it is closed under set differences. Therefore, $E = V \setminus N_1 \in \mathcal{M}_\mu$.

- **Proving (3) $\implies$ (1):** Suppose $E = H \cup N_2$, where H is an F_σ set and $\mu(N_2) = 0$. An F_σ set is defined as a countable union of closed sets. Since closed sets are Borel sets, their countable union H is also a Borel set ($H \in \mathcal{B}_\mathbb{R} \subset \mathcal{M}_\mu$).

  By the same completeness property of μ, the null set N_2 belongs to $\mathcal{M}_\mu$. Since a σ-algebra is closed under countable unions, it must be closed under finite unions as well. Thus, $E = H \cup N_2 \in \mathcal{M}_\mu$.

Part 2: Necessity under Finite Measure — proving (1) $\implies$ (2) and (3) assuming $\mu(E) < \infty$

Assume $E \in \mathcal{M}_\mu$ and that its measure is finite ($\mu(E) < \infty$). Let $j \in \mathbb{N}$ represent an arbitrary indexing parameter.

- **Constructing the G_δ Envelope (V):** By the outer regularity of Lebesgue-Stieltjes measures, we can approximate E from the outside using open supersets. For each natural number j, we can choose an open set $U_j \supset E$ such that:

$$\mu(U_j) \le \mu(E) + 2^{-j} \implies \mu(U_j) - 2^{-j} \le \mu(E)$$

Let us define V as the countable intersection of these approximating open sets:

$$V = \bigcap_{j=1}^{\infty} U_j$$

By definition, V is a G_δ set. Since $E \subset U_j$ for every single j, it follows directly that $E \subset V$.

- **Constructing the F_σ Core (H):** By the inner regularity of Lebesgue-Stieltjes measures, we can approximate E from the inside using compact subsets. For each natural number j, there exists a compact set $K_j \subset E$ such that:

$$\mu(K_j) \ge \mu(E) - 2^{-j} \implies \mu(E) \le \mu(K_j) + 2^{-j}$$

Let us define H as the countable union of these compact sets:

$$H = \bigcup_{j=1}^{\infty} K_j$$

Since every compact set in $\mathbb{R}$ is closed, H is a countable union of closed sets, making it an F_σ set. Furthermore, since $K_j \subset E$ for all j, we have $H \subset E$.

Combining our constructed sets gives the chain of inclusions $H \subset E \subset V$. We now evaluate their measures. Because $H = \bigcup_{j=1}^{\infty} K_j$, for any specific index j we know $K_j \subset H \subset E \subset V \subset U_j$. By monotonicity:

$$\mu(K_j) \le \mu(H) \le \mu(E) \le \mu(V) \le \mu(U_j)$$

Substituting our boundary limits $\mu(K_j) \ge \mu(E) - 2^{-j}$ and $\mu(U_j) \le \mu(E) + 2^{-j}$ into this chain yields:

$$\mu(E) - 2^{-j} \le \mu(H) \le \mu(E) \le \mu(V) \le \mu(E) + 2^{-j}$$

Taking the limit as $j \to \infty$, the error factor $2^{-j} \to 0$. By the squeeze theorem, we deduce:

$$\mu(V) = \mu(H) = \mu(E)$$

Now we look at the remainder sets. Let $N_1 = V \setminus E$ and $N_2 = E \setminus H$. Because $\mu(E) < \infty$, we can expand the subtraction rules for nested sets:

$$\mu(N_1) = \mu(V \setminus E) = \mu(V) - \mu(E) = \mu(E) - \mu(E) = 0$$

$$\mu(N_2) = \mu(E \setminus H) = \mu(E) - \mu(H) = \mu(E) - \mu(E) = 0$$

Rewriting these structural relations, we obtain $E = V \setminus N_1$ and $E = H \cup N_2$, completing the proof for the finite measure case.

Part 3: Extension to General Unbounded Sets — proving (1) $\implies$ (2) and (3) when $\mu(E) = \infty$

If E is an unbounded set with infinite measure, the direct subtraction $\mu(V) - \mu(E) = \infty - \infty$ is undefined. To circumvent this, we leverage the σ-finiteness of the real line. We decompose $\mathbb{R}$ into a countable union of bounded intervals: $\mathbb{R} = \bigcup_{n=-\infty}^{\infty} (n, n+1]$.

Let $E_n = E \cap (n, n+1]$. Because $(n, n+1]$ is a bounded set, each slice E_n is a bounded measurable set with a finite measure ($\mu(E_n) < \infty$).

- **Global G_δ Construction:** Applying the finite-measure result from Part 2 to each bounded slice E_n, there exists a G_δ set $V_n \subset \mathbb{R}$ and a null set $N_{1,n}$ such that:

$$E_n = V_n \setminus N_{1,n} \implies E_n \subset V_n \quad \text{and} \quad \mu(V_n \setminus E_n) = 0$$

Let us define the global envelope V as the union of these components: $V = \bigcup_{n=-\infty}^{\infty} V_n$. Because a countable union of G_δ sets is not guaranteed to remain a G_δ set automatically, we rewrite the presentation explicitly. Each V_n is an intersection of open sets, say $V_n = \bigcap_{j=1}^{\infty} U_{n,j}$. We can define an adjusted open sequence $W_j = \bigcup_{n=-\infty}^{\infty} U_{n,j}$, which means $V = \bigcap_{j=1}^{\infty} W_j$ can be structured as a valid G_δ set.

Evaluating the global error set $N_1 = V \setminus E$, we observe:

$$V \setminus E = \left(\bigcup_{n=-\infty}^{\infty} V_n \right) \setminus \left(\bigcup_{n=-\infty}^{\infty} E_n \right) \subset \bigcup_{n=-\infty}^{\infty} (V_n \setminus E_n)$$

By countable subadditivity, the measure of this global error is bounded by the individual slice errors:

$$\mu(N_1) \leq \sum_{n=-\infty}^{\infty} \mu(V_n \setminus E_n) = \sum_{n=-\infty}^{\infty} 0 = 0$$

Thus, $\mu(N_1) = 0$, confirming $E = V \setminus N_1$.

- **Global F_σ Construction:** Similarly, applying Part 2 to each slice E_n from the inside, there exists an F_σ set H_n and a null set $N_{2,n}$ such that:

$$E_n = H_n \cup N_{2,n} \implies H_n \subset E_n \quad \text{and} \quad \mu(E_n \setminus H_n) = 0$$

Let us define the global core H as the countable union of these localized cores:

$$H = \bigcup_{n=-\infty}^{\infty} H_n$$

Because each H_n is a countable union of closed sets, the global union H is a countable union of countable unions of closed sets, which remains a countable union of closed sets. Thus, H is a valid F_σ set.

We collect the leftover global error set $N_2 = E \setminus H$:

$$E \setminus H = \left(\bigcup_{n=-\infty}^{\infty} E_n \right) \setminus \left(\bigcup_{n=-\infty}^{\infty} H_n \right) \subset \bigcup_{n=-\infty}^{\infty} (E_n \setminus H_n)$$

Applying countable subadditivity to the error components:

$$\mu(N_2) \leq \sum_{n=-\infty}^{\infty} \mu(E_n \setminus H_n) = \sum_{n=-\infty}^{\infty} 0 = 0$$

Thus, $\mu(N_2) = 0$, establishing that $E = H \cup N_2$ holds universally across the entire real line.

□

Proposition. *If $E \in \mathcal{M}_\mu$ and $\mu(E) < \infty$, then for every $\epsilon > 0$ there exists a set A which is a finite union of open intervals such that:*

$$\mu(E \Delta A) < \epsilon$$

Proof. Let $\epsilon > 0$ be given. By our previous lemma, the measure of E can be computed as an infimum over coverings consisting of open intervals. Therefore, there exists a countable sequence of open intervals $\{(a_j, b_j)\}_{j=1}^{\infty}$ such that:

$$E \subset \bigcup_{j=1}^{\infty} (a_j, b_j) \quad \text{and} \quad \sum_{j=1}^{\infty} \mu((a_j, b_j)) \leq \mu(E) + \frac{\epsilon}{2}$$

Since the series $\sum_{j=1}^{\infty} \mu((a_j, b_j))$ converges to a finite value (bounded by $\mu(E) + \frac{\epsilon}{2} < \infty$), its tail must shrink to zero. Thus, we can find a sufficiently large index $N \in \mathbb{N}$ such that the sum of the remaining terms satisfies:

$$\sum_{j=N+1}^{\infty} \mu((a_j, b_j)) < \frac{\epsilon}{2}$$

Let us define our candidate set A as the finite union of the first N open intervals:

$$A = \bigcup_{j=1}^{N} (a_j, b_j)$$

We now estimate the measure of the symmetric difference $E\Delta A = (E \setminus A) \cup (A \setminus E)$.

1. **Estimating $E \setminus A$:** Since $E \subset \bigcup_{j=1}^{\infty}(a_j, b_j)$, removing the first N intervals leaves behind at most the infinite tail of the collection:
$$E \setminus A \subset \bigcup_{j=N+1}^{\infty} (a_j, b_j)$$
By countable subadditivity, the measure of this component is bounded by:
$$\mu(E \setminus A) \leq \sum_{j=N+1}^{\infty} \mu((a_j, b_j)) < \frac{\epsilon}{2}$$

2. **Estimating $A \setminus E$:** Let $U = \bigcup_{j=1}^{\infty}(a_j, b_j)$ denote the full open cover. Since $A \subset U$, we have:
$$\mu(A \setminus E) \leq \mu(U \setminus E) = \mu(U) - \mu(E)$$
Using countable subadditivity on U, we know $\mu(U) \leq \sum_{j=1}^{\infty} \mu((a_j, b_j))$. Substituting our original coverage bound gives:
$$\mu(A \setminus E) \leq \sum_{j=1}^{\infty} \mu((a_j, b_j)) - \mu(E) \leq \left(\mu(E) + \frac{\epsilon}{2}\right) - \mu(E) = \frac{\epsilon}{2}$$

Combining these two pieces via finite subadditivity yields:

$$\mu(E\Delta A) \leq \mu(E \setminus A) + \mu(A \setminus E) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

This completes the proof. □

Lebesgue Measure

We now examine the most fundamentally important example of a Lebesgue-Stieltjes measure on $\mathbb{R}$, which occurs when the underlying distribution function is the identity mapping.

Definition 28. *The **Lebesgue measure** on $\mathbb{R}$, denoted by m, is defined as the complete Lebesgue-Stieltjes measure μ_F associated with the linear identity function:*

$$F(x) = x$$

Under this specific configuration, the measure of any basic half-open interval simplifies directly to its geometric length:

$$m((a, b]) = b - a$$

Domains and Notation Restrictions

- **Lebesgue Measurable Sets:** The complete domain of m is denoted by $\mathcal{L}$. This collection forms the class of all **Lebesgue measurable sets** on the real line.
- **Restriction to Borel Sets:** In many analytical contexts, the domain is restricted exclusively to the Borel σ-algebra. By standard convention, the restriction of the complete measure m to $\mathcal{B}_{\mathbb{R}}$ (written as $m|_{\mathcal{B}_{\mathbb{R}}}$) is also referred to simply as the **Lebesgue measure**.

Linear Shifts and Dilations

To analyze structural properties such as translation invariance and scale scaling, we establish algebraic transformations of sets. For any subset $E \subset \mathbb{R}$ and structural real constants $s, r \in \mathbb{R}$, we define:

- **Translation (Shift):**
$$E + s = \{x + s : x \in E\}$$
- **Dilation (Scaling):**
$$rE = \{rx : x \in E\}$$

Theorem 10 (Translation Invariance and Dilation of Lebesgue Measure)**.** *If $E \in \mathcal{L}$, then $E + s \in \mathcal{L}$ and $rE \in \mathcal{L}$ for all $s, r \in \mathbb{R}$. Moreover, the measures satisfy:*

$$m(E + s) = m(E) \quad \textit{and} \quad m(rE) = |r|m(E)$$

Proof. The proof is structured by first verifying the preservation and scaling rules on the underlying Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$, analyzing the structural stability of null sets, and finally extending the properties globally to the complete Lebesgue σ-algebra $\mathcal{L}$ via structural decompositions.

Part 1: Verification on the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$

Let us consider the basic geometric behavior of open intervals under shifts and scaling transformations:

$$(a, b) + s = (a + s, b + s)$$
$$r(a, b) = \begin{cases} (ra, rb) & \text{if } r > 0 \\ \emptyset & \text{if } r = 0 \\ (rb, ra) & \text{if } r < 0 \end{cases}$$

In all instances, shifting or scaling an open interval produces another open interval. Because the collection of open intervals forms a topology that generates $\mathcal{B}_{\mathbb{R}}$, these algebraic operations map open generators to open generators. Consequently, the Borel σ-algebra is structurally invariant under translations and dilations; that is, if $E \in \mathcal{B}_{\mathbb{R}}$, then $E + s \in \mathcal{B}_{\mathbb{R}}$ and $rE \in \mathcal{B}_{\mathbb{R}}$.

Now, we introduce two auxiliary set functions for a fixed translation factor s and dilation factor r:

$$m_s(E) = m(E + s) \quad \text{and} \quad m^r(E) = m(rE)$$

We test these functions on an elementary bounded half-open interval $I = (a, b] \in \mathcal{B}_{\mathbb{R}}$:

- **Translation Measure:** Shifting the interval yields $I + s = (a + s, b + s]$. Evaluating its Lebesgue measure gives:
$$m_s((a, b]) = m((a + s, b + s]) = (b + s) - (a + s) = b - a = m((a, b])$$
Thus, m_s matches m identically on all basic intervals.
- **Dilation Measure:** If $r \geq 0$, scaling yields $rI = (ra, rb]$. Its measure is:
$$m^r((a, b]) = m((ra, rb]) = rb - ra = r(b - a) = |r|m((a, b])$$
If $r < 0$, the orientation of the endpoints reverses, yielding $rI = [rb, ra)$, whose Lebesgue measure is identical to the half-open alternative $(rb, ra]$:
$$m^r((a, b]) = m((rb, ra]) = ra - rb = -r(b - a) = |r|m((a, b])$$
Thus, m^r matches the scaled measure $|r|m$ on all basic intervals.

Because m_s matches m, and m^r matches $|r|m$ on the generating algebra of intervals, they represent valid σ-finite premeasures that agree perfectly on the generators. By the uniqueness of the Carathéodory Extension Theorem, their unique extensions to the generated σ-algebra must match everywhere. Hence, for any Borel set $E \in \mathcal{B}_{\mathbb{R}}$:

$$m(E + s) = m(E) \quad \text{and} \quad m(rE) = |r|m(E)$$

Part 2: Preservation of Lebesgue Null Sets

Suppose N is a set of Lebesgue measure zero ($m(N) = 0$). By the outer regularity of the measure, N can be embedded inside a Borel set $N_0 \in \mathcal{B}_{\mathbb{R}}$ that also possesses a measure of zero ($N \subset N_0$ and $m(N_0) = 0$).

Applying the Borel transformation rules established in Part 1 to the envelope N_0, we obtain:

$$m(N_0 + s) = m(N_0) = 0 \quad \text{and} \quad m(rN_0) = |r|m(N_0) = |r| \cdot 0 = 0$$

Because the containment relations are preserved under these point-wise transformations, we have:

$$(N + s) \subset (N_0 + s) \quad \text{and} \quad (rN) \subset (rN_0)$$

Because the Lebesgue measure space $(\mathbb{R}, \mathcal{L}, m)$ is complete, any subset of a set with measure zero is automatically Lebesgue measurable and has measure zero. Since $N_0 + s$ and rN_0 are Borel sets with a measure of zero, their subsets $N + s$ and rN are automatically pushed into $\mathcal{L}$ with measures:

$$m(N + s) = 0 \quad \text{and} \quad m(rN) = 0$$

Thus, the class of sets of Lebesgue measure zero is strictly preserved under both translations and dilations.

Part 3: Extension to the Complete Lebesgue σ-algebra $\mathcal{L}$

Recall from our structural characterization theorem that any Lebesgue measurable set $E \in \mathcal{L}$ can be uniquely split into the union of a Borel set $H \in \mathcal{B}_{\mathbb{R}}$ (specifically an F_σ set) and a Lebesgue null set N ($m(N) = 0$):

$$E = H \cup N$$

We apply the linear transformations directly to this decomposed representation:

- **For Translations:** Distributing the translation shift across the union gives:

  $$E + s = (H + s) \cup (N + s)$$

  From Part 1, since H is a Borel set, $H + s$ is a Borel set and belongs to $\mathcal{L}$. From Part 2, since N is a null set, $N + s \in \mathcal{L}$ and has a measure of zero. Because $\mathcal{L}$ is a σ-algebra, it is closed under finite unions. Therefore, $E + s \in \mathcal{L}$.

  To evaluate its measure, note that since H and N can be chosen to be disjoint, their shifts $H + s$ and $N + s$ remain strictly disjoint. By the additivity of measures:

  $$m(E + s) = m(H + s) + m(N + s) = m(H) + 0 = m(H)$$

  Since $m(E) = m(H) + m(N) = m(H) + 0 = m(H)$, we substitute this back into the equation to conclude:

  $$m(E + s) = m(E)$$

- **For Dilations:** Distributing the scaling factor across the union yields:

  $$rE = (rH) \cup (rN)$$

  By identical reasoning, since H is Borel, rH is Borel and belongs to $\mathcal{L}$. Since N is a null set, $rN \in \mathcal{L}$ and has a measure of zero. The union of these two components yields $rE \in \mathcal{L}$.

  Evaluating the scaled measure via additivity on the disjoint components gives:

  $$m(rE) = m(rH) + m(rN) = |r|m(H) + 0 = |r|m(H)$$

  Since $m(E) = m(H)$, we substitute this into the expression to obtain our final scaling relation:

  $$m(rE) = |r|m(E)$$

This completes the verification across the entire Lebesgue σ-algebra. $\square$

The Lebesgue Measure of the Cantor Set

The Cantor ternary set provides a foundational example in real analysis of a set that is compact, uncountable, and nowhere dense, yet carries a Lebesgue measure of exactly zero.

The Construction of the Cantor Set

The Cantor set, denoted by C, is constructed through an iterative geometric deletion process starting from the closed unit interval $I_0 = [0, 1]$:

- **Step 1 (I_1):** Remove the open middle third interval $\left(\frac{1}{3}, \frac{2}{3}\right)$ from $[0, 1]$. This leaves behind two disjoint closed intervals:
$$I_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$
- **Step 2 (I_2):** Remove the open middle third from each of the two remaining intervals in I_1 (deleting $\left(\frac{1}{9}, \frac{2}{9}\right)$ and $\left(\frac{7}{9}, \frac{8}{9}\right)$). This leaves $2^2 = 4$ closed intervals:
$$I_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right]$$
- **Step j (I_j):** Continuing indefinitely, at the j-th step, remove the open middle third of length 3^{-j} from each of the 2^{j-1} closed intervals remaining from the previous stage.

The **Cantor set** C is defined as the intersection of all remaining sets across this infinite sequence of operations:
$$C = \bigcap_{j=0}^{\infty} I_j$$
Alternatively, C can be characterized arithmetically as the set of all numbers $x \in [0, 1]$ that possess a base-3 (ternary) expansion containing no 1s in its digits:
$$x = \sum_{j=1}^{\infty} \frac{a_j}{3^j} \quad \text{where } a_j \in \{0, 2\} \text{ for all } j \in \mathbb{N}$$

Proposition. *Let C be the Cantor set. Then:*

1. *C is compact, nowhere dense, and totally disconnected (i.e., the only connected subsets of C are single points). Moreover, C has no isolated points.*
2. *$m(C) = 0$.*
3. $\operatorname{card}(C) = \mathfrak{c}$ *(i.e., C is uncountable and has the cardinality of the continuum).*

Proof. We break down the verification of these distinct properties section by section.

Part 1: Topological Properties (Item a)

- **Compactness:** Each intermediate set I_j is a finite union of closed intervals, making it a closed subset of $\mathbb{R}$. Because the arbitrary intersection of closed sets is closed, $C = \bigcap_{j=0}^{\infty} I_j$ is a closed set. Since $C \subset [0, 1]$, it is also bounded. By the Heine-Borel theorem, since C is closed and bounded, it is compact.
- **Nowhere Dense:** At each step j, open intervals are removed from the domain. The union of all removed open intervals forms a dense open subset of $[0, 1]$. Because C is the complement of this open set relative to $[0, 1]$, its interior is completely empty ($\operatorname{int}(C) = \emptyset$). A closed set with an empty interior is, by definition, nowhere dense.
- **Totally Disconnected:** Suppose C contained a connected component consisting of more than a single point. Such a component would have to be a closed interval $[a, b]$ with $a < b$. The length of this interval would be $b - a > 0$. However, at any stage j where $3^{-j} < b - a$, the remaining intervals in I_j have a length of exactly 3^{-j}. Therefore, the interval $[a, b]$ cannot fit inside any single component interval of I_j, which means it must contain a deleted middle-third open interval. This breaks the connectivity, proving that the only connected components of C are singletons.

Part 2: Lebesgue Measure Evaluation (Item b)

To compute the Lebesgue measure of C, we analyze the measure of its open complement within $[0,1]$, which is the union of all open middle-third intervals removed during the construction:

- At stage $j = 0$, we remove $2^0 = 1$ interval of length $\frac{1}{3}$.
- At stage $j = 1$, we remove $2^1 = 2$ intervals, each of length $\frac{1}{9} = \frac{1}{3^2}$.
- Globally, at stage j, we remove 2^j disjoint open intervals, each having a length of $\frac{1}{3^{j+1}}$.

Because all these removed intervals are pairwise disjoint, we can sum their measures directly using the countable additivity of the Lebesgue measure. The total measure of the removed segments is given by a convergent geometric series:

$$m([0,1] \setminus C) = \sum_{j=0}^{\infty} \frac{2^j}{3^{j+1}} = \frac{1}{3}\sum_{j=0}^{\infty} \left(\frac{2}{3}\right)^j$$

Applying the standard summation formula for an infinite geometric series $\sum_{j=0}^{\infty} q^j = \frac{1}{1-q}$ with ratio $q = \frac{2}{3}$:

$$m([0,1] \setminus C) = \frac{1}{3} \cdot \frac{1}{1 - \frac{2}{3}} = \frac{1}{3} \cdot \frac{1}{\frac{1}{3}} = 1$$

Since the total measure of the initial interval is $m([0,1]) = 1$, we calculate the measure of the Cantor set by subtraction:

$$m(C) = m([0,1]) - m([0,1] \setminus C) = 1 - 1 = 0$$

Part 3: Uncountability and Cardinality (Item c)

To establish that $\operatorname{card}(C) = \mathfrak{c}$, we construct an explicit surjective mapping from the Cantor set onto the entire closed unit interval $[0,1]$.

Let $x \in C$. We write x using its unique base-3 expansion consisting only of the digits 0 and 2:

$$x = \sum_{j=1}^{\infty} a_j 3^{-j} \quad \text{where } a_j \in \{0,2\}$$

We define the mapping $f : C \to [0,1]$ by dividing each ternary digit a_j by 2 to get $b_j = \frac{a_j}{2} \in \{0,1\}$, and interpreting the resulting digits as a base-2 (binary) expansion:

$$f(x) = \sum_{j=1}^{\infty} b_j 2^{-j} = \sum_{j=1}^{\infty} \frac{a_j}{2} 2^{-j}$$

Every real number $y \in [0,1]$ can be written in a binary expansion form $\sum_{j=1}^{\infty} b_j 2^{-j}$ where $b_j \in \{0,1\}$. By setting $a_j = 2b_j$, we obtain a sequence of ternary digits $a_j \in \{0,2\}$ that corresponds exactly to a point $x \in C$ such that $f(x) = y$.

This demonstrates that f maps C **onto** $[0,1]$. Consequently, the cardinality of C must be greater than or equal to the cardinality of $[0,1]$:

$$\operatorname{card}(C) \geq \operatorname{card}([0,1]) = \mathfrak{c}$$

Since C is a subset of $[0,1]$, its cardinality cannot exceed $\mathfrak{c}$. By the Cantor-Schröder-Bernstein theorem, we conclude that $\operatorname{card}(C) = \mathfrak{c}$, meaning it is uncountably infinite despite having a measure of zero. $\square$

Remarks on the Cantor Set, Cantor Function, and Non-Borel Measurable Sets

The geometric and arithmetic construction of the Cantor ternary set, alongside its associated distribution function, serves as one of the most critical foundational testing grounds in modern measure theory. Below is a systematic, itemized breakdown highlighting exactly why these structures were developed, what classical mathematical intuitions they disrupt, and how they function as essential counterexamples.

Remarks

1. **Disrupting Intuitions on Size and Dimension (The Cantor Set as a Counterexample)** Historically, a set containing "as many points" as a solid geometric continuous block was intuitively expected to occupy a non-zero portion of space. The Cantor set explicitly shatters this intuition:

   - **The Paradox of Uncountable Null Sets:** The Cantor set C is uncountably infinite with the cardinality of the continuum ($\text{card}(C) = \mathfrak{c}$), yet it possesses a total Lebesgue measure of exactly zero ($m(C) = 0$). This proves that measure-theoretic size (length) and set-theoretic size (cardinality) are completely independent concepts.
   - **Topological vs. Measure Isolation:** Topologically, C has no isolated points (it is a perfect set), meaning every point is tightly surrounded by infinitely many other members of the set. Yet, measure-theoretically, it is entirely hollow, being nowhere dense with an empty interior.

2. **Pathological Continuity without Growth (The Cantor-Lebesgue Function)** By extending the surjective mapping $f : C \to [0, 1]$ to the open intervals removed during the construction of the Cantor set, we obtain the famous **Cantor-Lebesgue Function** (or "Devil's Staircase"). This function serves as a profound counterexample to fundamental assumptions regarding calculus and differentiation:

   - **The Construction Purpose:** The function is defined to be flat (constant) on every single open interval omitted from the Cantor set. Since the total length of these omitted intervals sums to exactly 1, the function is constant almost everywhere on $[0, 1]$.
   - **The Pathological Counterexample:**
     - The function is continuous on $[0, 1]$ and increases monotonically from $f(0) = 0$ to $f(1) = 1$.
     - Because it is constant on the complements, its derivative exists and equals zero almost everywhere ($f'(x) = 0$ a.e.).
     - This strongly demonstrates that a function can continuously climb from 0 to 1 even though its rate of growth is zero almost everywhere. It shows that the classical Fundamental Theorem of Calculus ($\int_a^b f'(x)dx = f(b) - f(a)$) fails to hold for arbitrary continuous functions without enforcing a stricter condition known as *absolute continuity*.

3. **Decoupling Open Invariance from Metric Mass (Generalized Cantor Sets)** One might hypothesize that the empty interior and nowhere dense nature of the standard Cantor set are direct consequences of its measure being zero. Generalized Cantor sets are constructed to disprove this link:

   - **The Construction Purpose:** Instead of always removing a fixed middle third, at each stage j, an arbitrary open middle proportion $\alpha_j \in (0, 1)$ is deleted from the remaining closed intervals.
   - **Nowhere Dense Sets with Positive Measure:** If the sequence $\alpha_j \to 0$ sufficiently fast as $j \to \infty$, the resulting limiting set $K = \bigcap_1^\infty K_j$ retains all the topological properties of the original Cantor set (it remains compact, totally disconnected, and nowhere dense), but its Lebesgue measure is strictly positive ($m(K) > 0$). In fact, for any target value $\beta \in (0, 1)$, one can calibrate α_j such that $m(K) = \beta$. This proves that topological "hollowness" does not force a set to have zero mass.

4. **Proving that $\mathcal{B}_{\mathbb{R}}$ is a Strict Subset of $\mathcal{L}$ (The Existential Counterexample)** A crucial structural question in measure theory is whether completing a σ-algebra actually adds new sets, or if every Lebesgue measurable set ($\mathcal{L}$) is automatically a Borel set ($\mathcal{B}_{\mathbb{R}}$). The Cantor set resolves this explicitly by acting as a factory for non-Borel sets:

   - **The Cardinality Discrepancy Framework:**
     - The total number of Borel sets on the real line is bounded by the cardinality of the continuum: $\text{card}(\mathcal{B}_{\mathbb{R}}) = \mathfrak{c}$.
     - Because the Cantor set has Lebesgue measure zero ($m(C) = 0$), the completeness of the Lebesgue measure space demands that *every single subset* of C must also be Lebesgue measurable.

  - The power set of the Cantor set, $\mathcal{P}(C)$, is therefore completely contained within the Lebesgue σ-algebra ($\mathcal{P}(C) \subset \mathcal{L}$).
  - The cardinality of this power set is strictly larger than the continuum: $\text{card}(\mathcal{P}(C)) = 2^{\mathfrak{c}} > \mathfrak{c}$.
- **The Conclusion:** Because $\mathcal{L}$ contains at least $2^{\mathfrak{c}}$ elements while $\mathcal{B}_{\mathbb{R}}$ only contains $\mathfrak{c}$ elements, it is mathematically impossible for them to be equal. Thus, $\mathcal{B}_{\mathbb{R}} \subsetneq \mathcal{L}$ is a strict inclusion. There are vastly more Lebesgue measurable sets than Borel sets, proving that the completion process adds an uncountably massive family of sets to the domain.

Integration

Example 7. *We recall that any mapping $f : X \to Y$ between two sets induces a mapping $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$, defined by $f^{-1}(E) = \{x \in X : f(x) \in E\}$, which preserves unions, intersections, and complements. Thus, if $\mathcal{N}$ is a σ-algebra on Y, the collection $\{f^{-1}(E) : E \in \mathcal{N}\}$ is a σ-algebra on X.*

Verification:

(i) ***Empty set:*** *Since $\mathcal{N}$ is a σ-algebra, $\varnothing \in \mathcal{N}$. Thus, $f^{-1}(\varnothing) = \varnothing$ belongs to the collection.*

(ii) ***Complements:*** *If $f^{-1}(E)$ is in the collection for some $E \in \mathcal{N}$, then its complement is given by $(f^{-1}(E))^c = f^{-1}(E^c)$. Since $\mathcal{N}$ is closed under complements, $E^c \in \mathcal{N}$, so $f^{-1}(E^c)$ belongs to the collection.*

(iii) ***Countable unions:*** *If $\{f^{-1}(E_j)\}_{j=1}^{\infty}$ is a sequence in the collection with $E_j \in \mathcal{N}$, then $\bigcup_{j=1}^{\infty} f^{-1}(E_j) = f^{-1}\left(\bigcup_{j=1}^{\infty} E_j\right)$. Since $\mathcal{N}$ is closed under countable unions, $\bigcup_{j=1}^{\infty} E_j \in \mathcal{N}$, so the union belongs to the collection.*

Definition 29. *If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable spaces, a mapping $f : X \to Y$ is called $(\mathcal{M}, \mathcal{N})$-**measurable**, or just **measurable** when $\mathcal{M}$ and $\mathcal{N}$ are understood, if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.*

This proposition provides a highly practical shortcut for checking measurability. Instead of verifying that the preimage of *every* set in the σ-algebra $\mathcal{N}$ is measurable, we only need to check the preimages of the smaller collection of sets that generate $\mathcal{N}$.

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

Proof. **Forward Direction ($\implies$):**
Assume f is $(\mathcal{M}, \mathcal{N})$-measurable. By definition, $f^{-1}(E) \in \mathcal{M}$ for every $E \in \mathcal{N}$. Since $\mathcal{E} \subset \mathcal{N}$, it follows immediately that $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$.

Backward Direction ($\impliedby$):
Assume that $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$. Let us define the collection:

$$\mathcal{A} = \{E \subset Y : f^{-1}(E) \in \mathcal{M}\}$$

Our goal is to show that $\mathcal{N} \subset \mathcal{A}$. We proceed by verifying two facts:

(i) $\mathcal{A}$ **contains the generator:** By our initial assumption, $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$, which means $\mathcal{E} \subset \mathcal{A}$.

(ii) $\mathcal{A}$ **is a σ-algebra on** Y**:**

- $f^{-1}(\varnothing) = \varnothing \in \mathcal{M} \implies \varnothing \in \mathcal{A}$.
- If $E \in \mathcal{A}$, then $f^{-1}(E) \in \mathcal{M}$. Since $\mathcal{M}$ is a σ-algebra, $(f^{-1}(E))^c = f^{-1}(E^c) \in \mathcal{M}$, which implies $E^c \in \mathcal{A}$.
- If $\{E_j\}_{j=1}^{\infty} \subset \mathcal{A}$, then $f^{-1}(E_j) \in \mathcal{M}$ for each j. Since $\mathcal{M}$ is closed under countable unions, $\bigcup_{j=1}^{\infty} f^{-1}(E_j) = f^{-1}\left(\bigcup_{j=1}^{\infty} E_j\right) \in \mathcal{M}$, which implies $\bigcup_{j=1}^{\infty} E_j \in \mathcal{A}$.

Since $\mathcal{A}$ is a σ-algebra containing $\mathcal{E}$, and $\mathcal{N}$ is the *smallest* σ-algebra containing $\mathcal{E}$ (as $\mathcal{N} = \sigma(\mathcal{E})$), it must be that $\mathcal{N} \subset \mathcal{A}$.

Therefore, for every $E \in \mathcal{N}$, we have $E \in \mathcal{A}$, meaning $f^{-1}(E) \in \mathcal{M}$. This proves that f is $(\mathcal{M}, \mathcal{N})$-measurable. $\square$

Corollary 11. *If X and Y are metric (or topological) spaces, every continuous $f : X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.*

Proof. Recall that by the topological definition of continuity, a mapping $f : X \to Y$ is continuous if and only if $f^{-1}(U)$ is open in X for every open set $U \subset Y$.

Let $\mathcal{E}$ be the collection of all open sets in Y. By definition, the Borel σ-algebra $\mathcal{B}_Y$ is generated by $\mathcal{E}$ (i.e., $\mathcal{B}_Y = \sigma(\mathcal{E})$). Since f is continuous, for every open set $U \in \mathcal{E}$, its preimage $f^{-1}(U)$ is open in X, and therefore $f^{-1}(U) \in \mathcal{B}_X$.

Applying the previous Proposition, since the preimages of all generating sets $U \in \mathcal{E}$ are measurable in X, it follows that f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable. □

Example 8. *The composition of measurable mappings is measurable; that is, if $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable and $g : Y \to Z$ is $(\mathcal{N}, \mathcal{O})$-measurable, then $g \circ f : X \to Z$ is $(\mathcal{M}, \mathcal{O})$-measurable.*

Proof: *To show that $g \circ f$ is $(\mathcal{M}, \mathcal{O})$-measurable, we must verify that $(g \circ f)^{-1}(E) \in \mathcal{M}$ for any measurable set $E \in \mathcal{O}$.*

Let $E \in \mathcal{O}$. Using the properties of preimages for composite functions, we have:

$$(g \circ f)^{-1}(E) = f^{-1}\big(g^{-1}(E)\big)$$

Since $g : Y \to Z$ is $(\mathcal{N}, \mathcal{O})$-measurable, the preimage $g^{-1}(E)$ belongs to $\mathcal{N}$. Furthermore, since $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable, the preimage under f of any set in $\mathcal{N}$ belongs to $\mathcal{M}$. Therefore, $f^{-1}\big(g^{-1}(E)\big) \in \mathcal{M}$, which proves that $g \circ f$ is $(\mathcal{M}, \mathcal{O})$-measurable.

Remark. *If $(X, \mathcal{M})$ is a measurable space, a real- or complex-valued function f on X will be called **$\mathcal{M}$-measurable**, or just **measurable**, if it is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$ or $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$ measurable. $\mathcal{B}_{\mathbb{R}}$ or $\mathcal{B}_{\mathbb{C}}$ is always understood as the σ-algebra on the range space unless otherwise specified.*

*In particular, $f : \mathbb{R} \to \mathbb{C}$ is **Lebesgue** (resp. **Borel**) **measurable** if it is $(\mathcal{L}, \mathcal{B}_{\mathbb{C}})$ (resp. $(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{C}})$) measurable; likewise for $f : \mathbb{R} \to \mathbb{R}$.*

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

a. f is $\mathcal{M}$-measurable.

b. $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

c. $f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

d. $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

e. $f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

Proof. We will prove the equivalence by showing the implications a $\implies$ b $\implies$ c $\implies$ e $\implies$ d $\implies$ b and finally showing how any of these implies a.

Step 1: (a) $\implies$ (b)
Assume f is $\mathcal{M}$-measurable. By definition, f is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable, meaning $f^{-1}(E) \in \mathcal{M}$ for every Borel set $E \in \mathcal{B}_{\mathbb{R}}$. Since the open interval (a, ∞) is an open set in $\mathbb{R}$, it is a Borel set. Thus, $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

Step 2: (b) $\implies$ (c)
Assume $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. We can express the closed-open interval $[a, \infty)$ as a countable intersection of open intervals:

$$[a, \infty) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, \infty\right)$$

Taking the preimage on both sides gives:

$$f^{-1}([a, \infty)) = f^{-1}\left(\bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, \infty\right)\right) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(a - \frac{1}{n}, \infty\right)\right)$$

By assumption, each $f^{-1}((a - 1/n, \infty)) \in \mathcal{M}$. Since $\mathcal{M}$ is a σ-algebra, it is closed under countable intersections. Therefore, $f^{-1}([a, \infty)) \in \mathcal{M}$.

Step 3: (c) $\implies$ (e)
Assume $f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. The interval $(-\infty, a]$ is the complement of the interval (a, ∞), but it can also be linked directly via complements to $[a, \infty)$ by noting that $(-\infty, a] = \mathbb{R} \setminus (a, \infty)$. Alternatively, using standard complements:

$$(-\infty, a) = \mathbb{R} \setminus [a, \infty)$$

Let's establish (c) $\implies$ (d) first for elegance:

$$f^{-1}((-\infty, a)) = f^{-1}(\mathbb{R} \setminus [a, \infty)) = \left(f^{-1}([a, \infty))\right)^c$$

Since $f^{-1}([a, \infty)) \in \mathcal{M}$ and $\mathcal{M}$ is closed under complements, $f^{-1}((-\infty, a)) \in \mathcal{M}$, which proves (d).

To get (e) from here, observe that:

$$(-\infty, a] = \bigcap_{n=1}^{\infty} \left(-\infty, a + \frac{1}{n}\right)$$

Taking preimages:

$$f^{-1}((-\infty, a]) = \bigcap_{n=1}^{\infty} f^{-1}\left(\left(-\infty, a + \frac{1}{n}\right)\right)$$

Since each term in the intersection belongs to $\mathcal{M}$ by statement (d), their countable intersection belongs to $\mathcal{M}$. Thus, $f^{-1}((-\infty, a]) \in \mathcal{M}$.

Step 4: (e) $\implies$ (d)
Assume $f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$. We can write the open interval $(-\infty, a)$ as a countable union of closed-bounded-above intervals:

$$(-\infty, a) = \bigcup_{n=1}^{\infty} \left(-\infty, a - \frac{1}{n}\right]$$

Taking preimages:

$$f^{-1}((-\infty, a)) = \bigcup_{n=1}^{\infty} f^{-1}\left(\left(-\infty, a - \frac{1}{n}\right]\right)$$

Since $\mathcal{M}$ is closed under countable unions and each component belongs to $\mathcal{M}$ by statement (e), $f^{-1}((-\infty, a)) \in \mathcal{M}$.

Step 5: (d) $\implies$ (b)
Assume $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. Since (a, ∞) is the complement of $(-\infty, a]$, we write:

$$(a, \infty) = \mathbb{R} \setminus (-\infty, a]$$

From Step 4, we know statement (e) holds if (d) holds because they are duals. Thus, taking the complement:

$$f^{-1}((a, \infty)) = \left(f^{-1}((-\infty, a])\right)^c$$

Since $f^{-1}((-\infty, a]) \in \mathcal{M}$, its complement $f^{-1}((a, \infty)) \in \mathcal{M}$. Thus, statements (b), (c), (d), and (e) are completely equivalent to one another.

Step 6: Equivalence with (a)
To complete the proof, we must show that any of these conditions implies (a). Let us take condition (b).

Assume $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. Let $\mathcal{E} = \{(a, \infty) : a \in \mathbb{R}\}$. It is a standard result in real analysis that the collection of open rays $\mathcal{E}$ generates the Borel σ-algebra on $\mathbb{R}$ (i.e., $\sigma(\mathcal{E}) = \mathcal{B}_{\mathbb{R}}$).

By our earlier Proposition, if the preimage of every generating set in $\mathcal{E}$ belongs to $\mathcal{M}$, then the mapping f is $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable. Therefore, f is $\mathcal{M}$-measurable. $\square$

Definition 30. *If $(X, \mathcal{M})$ is a measurable space, f is a function on X, and $E \in \mathcal{M}$, we say that f is **measurable on** E if:*

$$f^{-1}(B) \cap E \in \mathcal{M} \quad \text{for all Borel sets } B.$$

Equivalently, the restriction $f|_E$ is $\mathcal{M}_E$-measurable, where $\mathcal{M}_E = \{F \cap E : F \in \mathcal{M}\}$ is the subspace σ-algebra on E.

Remark. ***Explanation:*** *This definition allows us to discuss the measurability of a function restricted to a specific domain patch E without requiring f to be well-behaved or even measurable across the entire space X. Notice that $f^{-1}(B) \cap E$ is precisely $(f|_E)^{-1}(B)$, making it a natural internalization of subspace measurability.*

Definition 31. *Given a set X and a family of measurable spaces $\{(Y_\alpha, \mathcal{N}_\alpha)\}_{\alpha \in A}$ with a map $f_\alpha : X \to Y_\alpha$ for each $\alpha \in A$, the **σ-algebra generated by** $\{f_\alpha\}_{\alpha \in A}$ is the unique smallest σ-algebra on X with respect to which all the f_α's are measurable. It is explicitly generated by the collection of sets:*

$$\left\{ f_\alpha^{-1}(E_\alpha) : E_\alpha \in \mathcal{N}_\alpha \text{ and } \alpha \in A \right\}$$

Definition 32. *In particular, if $X = \prod_{\alpha \in A} Y_\alpha$ is a cartesian product space, the **product σ-algebra** on X is defined as the σ-algebra generated by the coordinate projection maps $\pi_\alpha : X \to Y_\alpha$.*

Remark. ***Explanation:*** *Think of this as pulling back information. If X doesn't have an inherent structure yet, we can build a tailor-made σ-algebra on it by taking all necessary preimages from the target spaces Y_α so that every function f_α is forced to satisfy the definition of measurability. In the case of product spaces, this construction guarantees that working with a multi-dimensional structure scales down predictably to checking properties on individual target component paths via coordinates π_α.*

Proposition. *Let $(X, \mathcal{M})$ and $(Y_\alpha, \mathcal{N}_\alpha)$ $(\alpha \in A)$ be measurable spaces, $Y = \prod_{\alpha \in A} Y_\alpha$, $\mathcal{N} = \bigotimes_{\alpha \in A} \mathcal{N}_\alpha$, and $\pi_\alpha : Y \to Y_\alpha$ the coordinate projection maps. Then a mapping $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable iff $f_\alpha = \pi_\alpha \circ f$ is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable for all $\alpha \in A$.*

Proof. **Forward Direction ($\Longrightarrow$):**
Assume $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable. By definition, the coordinate projection map $\pi_\alpha : Y \to Y_\alpha$ is $(\mathcal{N}, \mathcal{N}_\alpha)$-measurable because the product σ-algebra $\mathcal{N}$ is constructed precisely to make these projection maps measurable.

Since $f_\alpha = \pi_\alpha \circ f$ is the composition of two measurable mappings (f and π_α), it follows naturally that f_α is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable for each $\alpha \in A$.

Backward Direction ($\Longleftarrow$):
Assume that each component function $f_\alpha = \pi_\alpha \circ f$ is $(\mathcal{M}, \mathcal{N}_\alpha)$-measurable for all $\alpha \in A$.

To show that f is $(\mathcal{M}, \mathcal{N})$-measurable, we must prove that $f^{-1}(B) \in \mathcal{M}$ for every set $B \in \mathcal{N}$. One of our previous propositions states that instead of checking the inverse image of every single measurable set, it suffices to prove this condition just for the collection of sets that generates the σ-algebra. We will perform that verification directly here.

Let us define the collection of all subsets in Y whose preimages under f are measurable in X:

$$\mathcal{A} = \{B \subset Y : f^{-1}(B) \in \mathcal{M}\}$$

We first check that $\mathcal{A}$ forms a valid σ-algebra on Y:

- $f^{-1}(\varnothing) = \varnothing \in \mathcal{M} \implies \varnothing \in \mathcal{A}$.
- If $B \in \mathcal{A}$, then $f^{-1}(B) \in \mathcal{M}$. Since $\mathcal{M}$ is a σ-algebra, its complement $(f^{-1}(B))^c = f^{-1}(B^c) \in \mathcal{M}$, which implies $B^c \in \mathcal{A}$.
- If $\{B_j\}_{j=1}^\infty \subset \mathcal{A}$, then $f^{-1}(B_j) \in \mathcal{M}$ for each j. Since $\mathcal{M}$ is closed under countable unions, $\bigcup_{j=1}^\infty f^{-1}(B_j) = f^{-1}\left(\bigcup_{j=1}^\infty B_j\right) \in \mathcal{M}$, which implies $\bigcup_{j=1}^\infty B_j \in \mathcal{A}$.

Thus, $\mathcal{A}$ is a σ-algebra on Y.

Now, recall that the product σ-algebra $\mathcal{N}$ is defined as the smallest σ-algebra generated by the elementary cylinder sets of the form $\pi_\alpha^{-1}(E_\alpha)$ where $E_\alpha \in \mathcal{N}_\alpha$ and $\alpha \in A$. Let us test if these generating sets belong to $\mathcal{A}$. Taking the preimage under f:

$$f^{-1}\big(\pi_\alpha^{-1}(E_\alpha)\big) = (\pi_\alpha \circ f)^{-1}(E_\alpha) = f_\alpha^{-1}(E_\alpha)$$

By our initial assumption, each component function f_α is measurable, meaning $f_\alpha^{-1}(E_\alpha) \in \mathcal{M}$. Therefore, every generating set $\pi_\alpha^{-1}(E_\alpha)$ belongs to $\mathcal{A}$.

Since $\mathcal{A}$ is a σ-algebra containing all the generating sets of Y, and $\mathcal{N}$ is the unique smallest σ-algebra containing those same generating sets, it must hold that $\mathcal{N} \subset \mathcal{A}$. Consequently, for every $B \in \mathcal{N}$, we have $B \in \mathcal{A}$, meaning $f^{-1}(B) \in \mathcal{M}$. This proves that f is $(\mathcal{M}, \mathcal{N})$-measurable. □

Corollary 12. *A function $f : X \to \mathbb{C}$ is $\mathcal{M}$-measurable iff* $\operatorname{Re} f$ *and* $\operatorname{Im} f$ *are $\mathcal{M}$-measurable.*

Proof. We can identify the complex plane $\mathbb{C}$ topologically with $\mathbb{R}^2$ via the standard mapping $z \mapsto (\operatorname{Re} z, \operatorname{Im} z)$. Under this identification, the Borel σ-algebra on $\mathbb{C}$ satisfies:

$$\mathcal{B}_{\mathbb{C}} = \mathcal{B}_{\mathbb{R}^2} = \mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$$

where $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$ denotes the product σ-algebra on $\mathbb{R} \times \mathbb{R}$.

Consider the coordinate projection maps $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ and $\pi_2 : \mathbb{R}^2 \to \mathbb{R}$ defined by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. Under our identification, these projections correspond exactly to taking the real and imaginary parts of a complex number:

$$\pi_1 \circ f = \operatorname{Re} f \quad \text{and} \quad \pi_2 \circ f = \operatorname{Im} f$$

By the above Proposition , a mapping into a product space equipped with its product σ-algebra is measurable if and only if its composition with each coordinate projection map is measurable. Applying this directly to $f : X \to \mathbb{R}^2$, it follows that f is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$-measurable if and only if $\pi_1 \circ f = \operatorname{Re} f$ and $\pi_2 \circ f = \operatorname{Im} f$ are both $(\mathcal{M}, \mathcal{B}_{\mathbb{R}})$-measurable. □

Remark (Functions with Values in the Extended Real Number System)**.** *In measure theory and integration, we frequently encounter functions whose values can become infinitely large (such as the limit of an increasing sequence of positive functions, or the measure of an unbounded set). To handle these situations rigorously without breaking our mathematical framework, we extend the real line by adding positive and negative infinity.*

1. Defining the Extended Real Number System ($\overline{\mathbb{R}}$)

*The **extended real number system**, denoted by $\overline{\mathbb{R}}$, is defined as the standard real line $\mathbb{R}$ unioned with the two symbolic endpoints $\{-\infty, +\infty\}$:*

$$\overline{\mathbb{R}} = [-\infty, \infty] = \mathbb{R} \cup \{-\infty, +\infty\}$$

2. The Extended Borel σ-Algebra ($\mathcal{B}_{\overline{\mathbb{R}}}$)

To define measurability for functions mapping into $\overline{\mathbb{R}}$, we must establish what it means for a subset of $\overline{\mathbb{R}}$ to be a Borel set. We define the extended Borel σ-algebra, $\mathcal{B}_{\overline{\mathbb{R}}}$, by tracking how sets intersect with the standard real numbers:

$$\mathcal{B}_{\overline{\mathbb{R}}} = \{E \subset \overline{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\}$$

What this means for students: *A set E is an extended Borel set if its "finite part" ($E \cap \mathbb{R}$) is a standard real Borel set. The set E itself may or may not contain the isolated points $\{+\infty\}$ and $\{-\infty\}$.*

3. Topological View: The Arc-Tangent Distance

An alternative and elegant way to visualize $\overline{\mathbb{R}}$ is to treat it as a legitimate metric space. We can "squish" the infinite line into a finite interval using the homeomorphism $A(x) = \arctan x$.

- *The function* $\arctan x$ *maps the infinite line $\mathbb{R}$ bijectively onto the open interval $(-\pi/2, \pi/2)$.*
- *By assigning* $\arctan(\infty) = \pi/2$ *and* $\arctan(-\infty) = -\pi/2$*, the extended system $\overline{\mathbb{R}}$ maps perfectly onto the closed, bounded interval $[-\pi/2, \pi/2]$.*

We can then define a rigorous metric (distance function) ρ on $\overline{\mathbb{R}}$ by:

$$\rho(x, y) = |\arctan x - \arctan y|$$

The open sets generated by this metric give rise to a topology whose Borel sets are exactly the collection $\mathcal{B}_{\overline{\mathbb{R}}}$ defined above.

4. Checking Measurability via One-Sided Rays

By applying our earlier proposition (Proposition 2.1), we do not need to check the preimage of every single complex set in $\mathcal{B}_{\overline{\mathbb{R}}}$ to prove a function is measurable. The σ-algebra $\mathcal{B}_{\overline{\mathbb{R}}}$ can be completely generated by simple, one-sided extended intervals (rays).

Therefore, a function $f : X \to \overline{\mathbb{R}}$ is defined to be $\mathcal{M}$***-measurable*** *if it is $(\mathcal{M}, \mathcal{B}_{\overline{\mathbb{R}}})$-measurable, which is equivalent to verifying just one of the following straightforward conditions for all $a \in \mathbb{R}$:*

(i) $f^{-1}((a, \infty]) \in \mathcal{M}$ *(the preimage of every ray extending up to positive infinity is measurable)*

(ii) $f^{-1}([-\infty, a)) \in \mathcal{M}$ *(the preimage of every ray extending down to negative infinity is measurable)*

Proposition. *If $f, g : X \to \mathbb{C}$ are $\mathcal{M}$-measurable, then so are $f + g$ and fg.*

Proof. To make this proof perfectly intuitive for students, we break it down into three logical phases: constructing a joint mapping into a product space, verifying the continuity of arithmetic operations, and analyzing their compositions.

Phase 1: Defining a Joint Function into a Product Space

Let us combine our two individual functions f and g into a single, joint mapping $F : X \to \mathbb{C} \times \mathbb{C}$ defined by:

$$F(x) = \big(f(x), g(x)\big)$$

We need to determine if F is a measurable function. By identifying $\mathbb{C}$ with $\mathbb{R}^2$, the Borel σ-algebra on the product space satisfies $\mathcal{B}_{\mathbb{C}\times\mathbb{C}} = \mathcal{B}_{\mathbb{C}} \otimes \mathcal{B}_{\mathbb{C}}$.

Recall from Proposition 2.4 that a multi-dimensional function mapping into a product space is measurable if and only if each of its component projection coordinate paths is individually measurable. The coordinate projections of F are exactly our starting functions:

$$\pi_1 \circ F = f \quad \text{and} \quad \pi_2 \circ F = g$$

Since we are given that f and g are $\mathcal{M}$-measurable, both component projections are measurable. Therefore, by Proposition 2.4, the joint map F is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}\times\mathbb{C}})$-measurable.

Phase 2: The Measurability of Arithmetic Operations

Now let us look at the target operations themselves. Define the addition map $\phi : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ and the multiplication map $\psi : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ as follows:

$$\phi(z, w) = z + w \quad \text{and} \quad \psi(z, w) = zw$$

From basic topology, we know that standard algebraic operations—addition and multiplication of complex numbers—are continuous functions.

By Corollary 2.2, any continuous function between topological spaces is automatically measurable with respect to their corresponding Borel σ-algebras. Therefore, ϕ and ψ are $(\mathcal{B}_{\mathbb{C}\times\mathbb{C}}, \mathcal{B}_{\mathbb{C}})$-measurable mappings.

Phase 3: Composition and Conclusion

Finally, notice that the functions $f+g$ and fg can be written perfectly as compositions of our operational maps with our joint map:

$$(f + g)(x) = f(x) + g(x) = \phi\big(f(x), g(x)\big) = (\phi \circ F)(x)$$

$$(fg)(x) = f(x)g(x) = \psi\big(f(x), g(x)\big) = (\psi \circ F)(x)$$

We have already established that:

(i) F is a measurable mapping from X to $\mathbb{C} \times \mathbb{C}$.

(ii) ϕ and ψ are measurable mappings from $\mathbb{C} \times \mathbb{C}$ to $\mathbb{C}$.

Since the composition of two measurable mappings is always measurable, it follows immediately that $f + g = \phi \circ F$ and $fg = \psi \circ F$ are $\mathcal{M}$-measurable functions. □

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

$$g_1(x) = \sup_j f_j(x), \qquad g_3(x) = \limsup_{j\to\infty} f_j(x),$$

$$g_2(x) = \inf_j f_j(x), \qquad g_4(x) = \liminf_{j\to\infty} f_j(x)$$

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

Proof. To make this proof accessible to students, we will break it down into clear stages. We leverage the fact that for extended real-valued functions ($\overline{\mathbb{R}}$), we can establish full measurability by checking only one-sided infinite rays like $(a, \infty]$ or $[-\infty, a)$.

Step 1: Proving the Measurability of the Supremum (g_1)

Let $g_1(x) = \sup_j f_j(x)$. We choose to evaluate the preimage of the open ray $(a, \infty]$ for an arbitrary $a \in \mathbb{R}$.

By definition of the supremum, the value $g_1(x) = \sup_j f_j(x)$ is strictly greater than a if and only if **at least one** individual function $f_j(x)$ in the sequence exceeds a. Translating this logical "at least one" statement into set theory gives us a countable union of preimages:

$$g_1^{-1}((a, \infty]) = \bigcup_{j=1}^{\infty} f_j^{-1}((a, \infty])$$

Since every function f_j in the sequence is given to be $\mathcal{M}$-measurable, each individual set $f_j^{-1}((a, \infty])$ belongs to the σ-algebra $\mathcal{M}$. Because a σ-algebra is closed under countable unions, their union must also belong to $\mathcal{M}$. Therefore, $g_1 = \sup_j f_j$ is $\mathcal{M}$-measurable.

Step 2: Proving the Measurability of the Infimum (g_2)

Let $g_2(x) = \inf_j f_j(x)$. This time, it is contextually simpler to check the downward ray $[-\infty, a)$ for an arbitrary $a \in \mathbb{R}$.

By definition of the infimum, the greatest lower bound $g_2(x) = \inf_j f_j(x)$ is strictly less than a if and only if **at least one** individual function $f_j(x)$ drops strictly below a. Once again, this logical "at least one" corresponds to a countable union:

$$g_2^{-1}([-\infty, a)) = \bigcup_{j=1}^{\infty} f_j^{-1}([-\infty, a))$$

Because each f_j is $\mathcal{M}$-measurable, each set $f_j^{-1}([-\infty, a)) \in \mathcal{M}$. Since $\mathcal{M}$ is closed under countable unions, the resulting union belongs to $\mathcal{M}$. This proves $g_2 = \inf_j f_j$ is $\mathcal{M}$-measurable.

Step 3: Proving the Measurability of the Limit Superior (g_3)

Recall the standard analytical definition of the limit superior as a sequential optimization:

$$\limsup_{j\to\infty} f_j(x) = \inf_{k\geq 1} \left(\sup_{j>k} f_j(x) \right)$$

Let us analyze this from the inside out:

1. For a fixed integer k, define the "tail supremum" function:

$$h_k(x) = \sup_{j>k} f_j(x)$$

   Since h_k is the supremum of a countable sub-sequence of our measurable functions, it is measurable by the exact logic we established in **Step 1**.

2. Now, notice that the total limit superior function $g_3(x)$ can be rewritten as:

$$g_3(x) = \inf_{k \geq 1} h_k(x)$$

Since g_3 is the infimum of a countable collection of functions $\{h_k\}_{k=1}^{\infty}$ which we just proved are measurable, it must be measurable by the exact logic established in **Step 2**.

Thus, $g_3 = \limsup f_j$ is $\mathcal{M}$-measurable.

Step 4: Proving the Measurability of the Limit Inferior (g_4)

Similarly, the limit inferior is defined as:

$$\liminf_{j \to \infty} f_j(x) = \sup_{k \geq 1} \left(\inf_{j > k} f_j(x) \right)$$

Applying the identical two-step composition logic:

1. For a fixed k, the tail infimum $m_k(x) = \inf_{j>k} f_j(x)$ is a countable infimum of measurable functions, making it measurable per **Step 2**.
2. The final operation, $g_4(x) = \sup_{k \geq 1} m_k(x)$, represents a countable supremum of these measurable tail functions, making it fully measurable per **Step 1**.

Thus, $g_4 = \liminf f_j$ is $\mathcal{M}$-measurable.

Step 5: Analysis of the Pointwise Limit (f)

Finally, suppose the pointwise limit $f(x) = \lim_{j \to \infty} f_j(x)$ exists for every $x \in X$.

From real analysis, if a sequence of numbers converges, its limit superior and its limit inferior must be exactly equal to each other, and equal to the limit itself:

$$f(x) = \limsup_{j \to \infty} f_j(x) = \liminf_{j \to \infty} f_j(x)$$

Because we have already proven that $g_3 = \limsup f_j$ is a perfectly valid $\mathcal{M}$-measurable function, and $f(x) = g_3(x)$ everywhere, it follows immediately that f is $\mathcal{M}$-measurable. □

Corollary 13 (Corollary 2.8)**.** *If $f, g : X \to \overline{\mathbb{R}}$ are $\mathcal{M}$-measurable, then so are* $\max(f, g)$ *and* $\min(f, g)$.

Proof. To make this intuitive for students, we can connect the maximum and minimum of two functions directly to the supremum and infimum operations we established in Proposition 2.7.

Consider the maximum function, $h_1(x) = \max\big(f(x), g(x)\big)$. For any given point $x \in X$, the maximum of these two values is equivalent to taking the supremum over a finite sequence containing just these two functions:

$$\max\big(f(x), g(x)\big) = \sup_{j \in \{1,2\}} \phi_j(x) \quad \text{where } \phi_1 = f \text{ and } \phi_2 = g$$

According to Proposition 2.7, the supremum of any countable family of extended real-valued measurable functions is automatically measurable. Since a finite set is a trivial case of a countable set, $\max(f, g)$ is $\mathcal{M}$-measurable.

Similarly, consider the minimum function, $h_2(x) = \min\big(f(x), g(x)\big)$. We can reframe this as an infimum operation over the same pair of functions:

$$\min\big(f(x), g(x)\big) = \inf_{j \in \{1,2\}} \phi_j(x)$$

By the second part of Proposition 2.7, the infimum of a countable family of measurable functions is measurable. Therefore, $\min(f, g)$ is $\mathcal{M}$-measurable. □

Corollary 14. *If $\{f_j\}$ is a sequence of complex-valued measurable functions and $f(x) = \lim_{j \to \infty} f_j(x)$ exists for all $x \in X$, then f is $\mathcal{M}$-measurable.*

Proof. This proof reduces the complex convergence problem down to separate real convergence problems by leveraging real and imaginary parts.

Let each function f_j in the sequence be decomposed into its real component functions:

$$f_j(x) = u_j(x) + iv_j(x) \quad \text{where } u_j = \operatorname{Re} f_j \text{ and } v_j = \operatorname{Im} f_j$$

By Corollary 2.5, since each f_j is given to be complex-measurable, its real part sequence $\{u_j\}_{j=1}^{\infty}$ and its imaginary part sequence $\{v_j\}_{j=1}^{\infty}$ are sequences of real-valued $\mathcal{M}$-measurable functions.

We are told that the pointwise limit function $f(x)$ exists everywhere. Let us decompose this limit function as well:

$$f(x) = u(x) + iv(x) \quad \text{where } u = \operatorname{Re} f \text{ and } v = \operatorname{Im} f$$

From the properties of complex limits, a sequence of complex numbers converges if and only if their real and imaginary components converge independently. Thus, for every $x \in X$, we have:

$$u(x) = \lim_{j\to\infty} u_j(x) \qquad \text{and} \qquad v(x) = \lim_{j\to\infty} v_j(x)$$

Now we apply our results from the above Proposition:

(i) Since u is the pointwise limit of a sequence of real-valued measurable functions $\{u_j\}$, $u = \operatorname{Re} f$ is $\mathcal{M}$-measurable.

(ii) Since v is the pointwise limit of a sequence of real-valued measurable functions $\{v_j\}$, $v = \operatorname{Im} f$ is $\mathcal{M}$-measurable.

Finally, by applying Corollary 8 in the forward direction, since both $\operatorname{Re} f$ and $\operatorname{Im} f$ are proven to be $\mathcal{M}$-measurable, the joint complex limit function f is $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$-measurable. □

Definition 33 (Positive and Negative Parts of a Function)**.** *Let $(X, \mathcal{M})$ be a measurable space and let $f : X \to \overline{\mathbb{R}}$ be an extended real-valued function. The **positive part** of f, denoted by f^+, and the **negative part** of f, denoted by f^-, are non-negative functions on X defined pointwise by:*

$$f^+(x) = \max\big(f(x), 0\big)$$

$$f^-(x) = \max\big(-f(x), 0\big)$$

Remark. *This decomposition is one of the most powerful tools in integration theory. It allows us to split any real-valued function—which might oscillate between positive and negative values-into two purely non-negative pieces.*

Here are the key insights your students should keep in mind:

- ***Behavior at a point** x:*
  - *If $f(x) \geq 0$, then $f^+(x) = f(x)$ and $f^-(x) = 0$.*
  - *If $f(x) < 0$, then $f^+(x) = 0$ and $f^-(x) = -f(x)$ (which is a positive number).*
- ***Reconstructing the Function:** The original function and its absolute value can be easily recovered by combining these two parts:*

$$f(x) = f^+(x) - f^-(x)$$

$$|f(x)| = f^+(x) + f^-(x)$$

- ***Measurability Inheritence:** Since f is measurable, the constant function 0 is measurable, and the scalar multiplication $-f$ is measurable. By Corollary 9, the maximum of two measurable functions is measurable. Therefore, if f is a measurable function, its positive part f^+ and negative part f^- are automatically measurable functions.*

Example 9. *Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to \mathbb{C}$ is an $\mathcal{M}$-measurable function, then its absolute value function $|f| : X \to \mathbb{R}$, defined pointwise by $|f|(x) = |f(x)|$, is also $\mathcal{M}$-measurable.*

Proof. To make this clear for students, we express the function $|f|$ as a composition of two separate mathematical mappings.

Consider the absolute value function on the complex plane, $g : \mathbb{C} \to \mathbb{R}$, defined by:

$$g(z) = |z|$$

From classical topology, the map g is continuous everywhere on $\mathbb{C}$ because it satisfies the standard triangle inequality $\big||z_1| - |z_2|\big| \le |z_1 - z_2|$.

By Corollary 2.2, since g is a continuous function between topological spaces, it is automatically a Borel measurable mapping. This means g is $(\mathcal{B}_{\mathbb{C}}, \mathcal{B}_{\mathbb{R}})$-measurable.

Now, observe that our function of interest, $|f|$, can be written perfectly as the composition of g and f:

$$|f|(x) = |f(x)| = g\big(f(x)\big) = (g \circ f)(x)$$

We are given that $f : X \to \mathbb{C}$ is $\mathcal{M}$-measurable (which means $(\mathcal{M}, \mathcal{B}_{\mathbb{C}})$-measurable), and we have shown that $g : \mathbb{C} \to \mathbb{R}$ is $(\mathcal{B}_{\mathbb{C}}, \mathcal{B}_{\mathbb{R}})$-measurable. Since the composition of two measurable mappings is always measurable, it follows that $|f| = g \circ f$ is an $\mathcal{M}$-measurable function. □

Example 10. *Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to \mathbb{C}$ is an $\mathcal{M}$-measurable function, then its complex signum component function* $\operatorname{sgn} f : X \to \mathbb{C}$*, defined pointwise by:*

$$(\operatorname{sgn} f)(x) = \operatorname{sgn}\big(f(x)\big) \quad \textit{where} \quad \operatorname{sgn} z = \begin{cases} \frac{z}{|z|} & \textit{if } z \neq 0, \\ 0 & \textit{if } z = 0, \end{cases}$$

is also $\mathcal{M}$-measurable.

Proof. Unlike the absolute value function, the standalone complex signum function $\operatorname{sgn} : \mathbb{C} \to \mathbb{C}$ is *not* continuous on the entire complex plane because it has a sharp jump discontinuity at the origin $z = 0$. Therefore, we cannot invoke Corollary 2.2 directly on $\mathbb{C}$. Instead, we must prove that sgn is Borel measurable by looking at the preimages of open sets.

Let $U \subset \mathbb{C}$ be an arbitrary open target set in the complex plane. We analyze the behavior of the preimage $\operatorname{sgn}^{-1}(U)$ under two distinct conditions based on whether the origin is inside our target window:

1. **Case 1:** $0 \notin U$
   If the target open set U does not contain 0, then any point z that maps into U must satisfy $z \neq 0$. On the punctured plane $\mathbb{C} \setminus \{0\}$, the mapping $z \mapsto \frac{z}{|z|}$ is a continuous algebraic combination of continuous functions. Because the restriction of sgn to $\mathbb{C} \setminus \{0\}$ is continuous, the preimage of the open set U must be completely open in $\mathbb{C} \setminus \{0\}$. Since $\mathbb{C} \setminus \{0\}$ is itself an open subset of $\mathbb{C}$, $\operatorname{sgn}^{-1}(U)$ is open in $\mathbb{C}$. Since every open set is a Borel set, $\operatorname{sgn}^{-1}(U) \in \mathcal{B}_{\mathbb{C}}$.

2. **Case 2:** $0 \in U$
   If the target open set U contains 0, then by definition, the origin $z = 0$ is pulled back into our preimage set because $\operatorname{sgn}(0) = 0 \in U$. We can split the preimage into the origin piece and everything else:
   $$\operatorname{sgn}^{-1}(U) = \operatorname{sgn}^{-1}\big(U \setminus \{0\}\big) \cup \{0\}$$
   Let $V = \operatorname{sgn}^{-1}\big(U \setminus \{0\}\big)$. Since $U \setminus \{0\}$ does not contain 0, it falls under the exact logic of **Case 1**, meaning V is an open set in $\mathbb{C}$. Therefore, the total preimage can be written as:
   $$\operatorname{sgn}^{-1}(U) = V \cup \{0\}$$
   where V is open. Since open sets are Borel sets ($V \in \mathcal{B}_{\mathbb{C}}$) and singleton points are closed and thus Borel sets ($\{0\} \in \mathcal{B}_{\mathbb{C}}$), their union must belong to the Borel σ-algebra $\mathcal{B}_{\mathbb{C}}$.

In both cases, the preimage of an arbitrary open set U under sgn belongs to $\mathcal{B}_{\mathbb{C}}$. By Proposition 2.1, because the open sets generate the Borel σ-algebra, the mapping $\operatorname{sgn} : \mathbb{C} \to \mathbb{C}$ is $(\mathcal{B}_{\mathbb{C}}, \mathcal{B}_{\mathbb{C}})$-measurable.

Finally, since $\operatorname{sgn} f$ is the composition of the Borel measurable function sgn with the $\mathcal{M}$-measurable function f:

$$\operatorname{sgn} f = \operatorname{sgn} \circ f$$

the composition of these two measurable mappings ensures that $\operatorname{sgn} f$ is $\mathcal{M}$-measurable. □

Definition 34 (Characteristic Function / Indicator Function)**.** *Let $(X, \mathcal{M})$ be a measurable space and let $E \subset X$. The **characteristic function** (also known as the **indicator function**) of the set E, denoted by χ_E or 1_E, is a function $\chi_E : X \to \{0, 1\}$ defined pointwise by:*

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

An essential fact for students to understand is that χ_E is an $\mathcal{M}$-measurable function if and only if the underlying set E is a measurable set (i.e., $E \in \mathcal{M}$).

Definition 35 (Simple Function)**.** *A function $f : X \to \mathbb{C}$ is called a **simple function** if its range is a finite subset of the complex numbers $\mathbb{C}$. Structurally, any such measurable function can be expressed as a finite linear combination, with complex coefficients, of the characteristic functions of sets belonging to our σ-algebra $\mathcal{M}$:*

$$f = \sum_{j=1}^{n} z_j \chi_{E_j}$$

where $z_j \in \mathbb{C}$ and $E_j \in \mathcal{M}$ for all $j \in \{1, \ldots, n\}$.

Definition 36 (Standard Representation of a Simple Function)**.** *Let $f : X \to \mathbb{C}$ be a simple function whose distinct non-zero and zero values in its finite range are given by the set $\{z_1, z_2, \ldots, z_n\}$. The **standard representation** of f is the unique configuration:*

$$f = \sum_{j=1}^{n} z_j \chi_{E_j}$$

where the sets E_j are explicitly defined as the preimages of each distinct value in the range:

$$E_j = f^{-1}(\{z_j\})$$

In this standard representation, the coefficients z_j are strictly distinct, and the corresponding measurable sets E_j are pairwise disjoint and collectively exhaust the domain space, meaning $\bigcup_{j=1}^{n} E_j = X$.

Remark (Crucial Pedagogical Notes for Integration Theory)**.** *For students transitioning into the theory of Lebesgue integration, the following structural constraints placed on simple functions are highly important:*

1. ***Measurability of Underlying Sets and Functions:*** *In our pedagogical context, whenever we define a simple function as a linear combination $\sum z_j \chi_{E_j}$, we implicitly require every single component set E_j to be a measurable set ($E_j \in \mathcal{M}$). This crucial requirement guarantees that the simple function f is strictly measurable. This is foundational because when we later define the Lebesgue integral of a simple function, the definition relies on computing the measures of these individual sets:*

$$\int_X f \, d\mu = \sum_{j=1}^{n} z_j \mu(E_j)$$

   If the sets E_j were not elements of $\mathcal{M}$, the measure $\mu(E_j)$ would be completely undefined, breaking our integration framework.

2. ***Exclusion of Infinite Values:*** *By definition, simple functions are strictly maps into the complex numbers $\mathbb{C}$ or real numbers $\mathbb{R}$, and **are not allowed to assume the values** $\pm\infty$. Keeping the range of simple functions entirely finite prevents indeterminate mathematical forms (such as $\infty \times \infty$) from appearing when computing simple linear combination steps during integration.*

3. ***Domain Exhaustion and Zero Coefficients:*** *For a simple function to satisfy the basic set-theoretic definition of a function on X, it must assign a clear value to every single element $x \in X$. Consequently, in its standard representation, the union of the disjoint sets E_j must equal the full domain space X ($\bigcup_{j=1}^{n} E_j = X$).*

   To achieve this total coverage, ***one of the coefficients z_j may well be*** *0. Students must not discard the term $0 \cdot \chi_{E_j}$ even though it evaluates to zero; it must be kept as an active structural part of the standard representation because that partition set E_j represents exactly where the function vanishes, and it plays a vital role when f interacts with or is restricted by other functions.*

Example 11. *Let $(X, \mathcal{M})$ be a measurable space. If $f, g : X \to \mathbb{C}$ are simple functions, then their sum $f + g$ and their product fg are also simple functions.*

Proof. To make this proof completely transparent for students, we must show two things for each combination: first, how the functions look explicitly when combined on a shared partition of the domain, and second, that their ranges remain strictly finite.

Step 1: Constructing a Joint Partition of the Domain Space

By definition, f and g are simple functions, meaning they have finite ranges and can be written in their standard representations:

$$f = \sum_{j=1}^{n} z_j \chi_{E_j} \quad \text{and} \quad g = \sum_{k=1}^{m} w_k \chi_{F_k}$$

where $\{z_1, \dots, z_n\} \subset \mathbb{C}$ is the finite range of f, and $\{w_1, \dots, w_m\} \subset \mathbb{C}$ is the finite range of g. The sets $E_j = f^{-1}(\{z_j\})$ partition X, and the sets $F_k = g^{-1}(\{w_k\})$ also partition X. All sets E_j and F_k belong to the σ-algebra $\mathcal{M}$.

To analyze $f + g$ and fg together, we construct a new, finer joint partition of X by intersecting these two families of sets. For every pair of indices (j, k), define:

$$A_{j,k} = E_j \cap F_k$$

Since $\mathcal{M}$ is a σ-algebra, it is closed under finite intersections, meaning each $A_{j,k} \in \mathcal{M}$. Furthermore, because $\{E_j\}$ and $\{F_k\}$ are individual partitions of X, the collection of all non-empty intersections $\{A_{j,k} : 1 \le j \le n, 1 \le k \le m\}$ forms a pairwise disjoint partition that perfectly exhausts the full domain space X:

$$\bigcup_{j=1}^{n} \bigcup_{k=1}^{m} A_{j,k} = \bigcup_{j=1}^{n} \bigcup_{k=1}^{m} (E_j \cap F_k) = \left(\bigcup_{j=1}^{n} E_j \right) \cap \left(\bigcup_{k=1}^{m} F_k \right) = X \cap X = X$$

On any specific sub-set $A_{j,k}$, if an element $x \in A_{j,k}$, then $x \in E_j$ (forcing $f(x) = z_j$) and $x \in F_k$ (forcing $g(x) = w_k$).

—

Step 2: Analysis of the Sum Function $(f + g)$

Let us observe how the sum function looks. By evaluating $f(x) + g(x)$ across our joint partition, we can express the entire sum function as a linear combination over the sets $A_{j,k}$:

$$(f + g)(x) = \sum_{j=1}^{n} \sum_{k=1}^{m} (z_j + w_k) \chi_{A_{j,k}}(x)$$

Checking the criteria for a simple function:

1. **Measurability of the component sets:** Every underlying subset $A_{j,k} = E_j \cap F_k$ belongs to $\mathcal{M}$. Thus, $\chi_{A_{j,k}}$ is a measurable indicator function for each term.

2. **Finiteness of the range:** The coefficients in this new expression are of the form $(z_j + w_k)$. Since there are exactly n values for z_j and m values for w_k, the total number of distinct possible values in the range of $f + g$ is at most $n \times m$, which is strictly finite.

Because $f + g$ is a finite linear combination of characteristic functions of measurable sets, it is an $\mathcal{M}$-measurable function with a finite range in $\mathbb{C}$, matching the definition of a simple function.

—

Step 3: Analysis of the Product Function (fg)

Similarly, let us observe how the product function looks when evaluated over our joint partition sets. For any point $x \in A_{j,k}$, the product multiplies their respective constant values:

$$(fg)(x) = f(x) \cdot g(x) = z_j \cdot w_k$$

This allows us to write the total product function explicitly as:

$$(fg)(x) = \sum_{j=1}^{n} \sum_{k=1}^{m} (z_j w_k) \chi_{A_{j,k}}(x)$$

Checking the criteria for a simple function:

1. **Measurability of the component sets:** Just as before, each set $A_{j,k}$ is a measurable set in $\mathcal{M}$, ensuring that every component function $\chi_{A_{j,k}}$ is measurable.
2. **Finiteness of the range:** The elements making up the range of fg are formed by the products $(z_j w_k)$. The maximum number of distinct values this product range can hold is bounded above by $n \times m$, which is strictly finite.

Since fg can be written as a finite linear combination of indicator functions belonging to $\mathcal{M}$ and does not assume infinite values, it is a simple function.

Note on Standard Representation

Note to students: The joint expressions derived above,

$$\sum_{j,k} (z_j + w_k) \chi_{A_{j,k}} \quad \text{and} \quad \sum_{j,k} (z_j w_k) \chi_{A_{j,k}}$$

are structurally valid linear combinations, but they might not be the official **standard representation** of $f + g$ or fg. If different index pairs yield identical values (for instance, if $z_1 + w_2 = z_3 + w_1$), those matching coefficients must be grouped together, and their corresponding disjoint sets $A_{j,k}$ must be combined using unions to yield the unique standard representation where all coefficients are strictly distinct. Regardless of this final consolidation step, the functions are proved to be simple. ☐

Theorem 15. *a. If $f : X \to [0, \infty]$ is measurable, there is a sequence $\{\phi_n\}$ of simple functions such that $0 \le \phi_1 \le \phi_2 \le \cdots \le f$, $\phi_n \to f$ pointwise, and $\phi_n \to f$ uniformly on any set on which f is bounded.*

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

Proof. To make this challenging proof completely transparent for your students, we will break the entire explanation into structured, digestible phases.

—

Part (a): Non-negative Extended Functions ($f : X \to [0, \infty]$)

The core strategy here is to approximate the codomain $[0, \infty]$ by chopping it into finite horizontal strips, seeing where the function lands, and building a simple step-function from the preimages. As n increases, our grid gets taller and finer.

Phase 1: Step-by-Step Evolution of the Construction

Let us build the approximating functions ϕ_n explicitly for the first few stages to see the pattern.

Stage 1: Building ϕ_0

We look at the codomain and establish a ceiling at $2^0 = 1$.

- **The Cutoff:** We divide the codomain into two primary zones: values below $2^0 = 1$, and values above or equal to 1.
- **The Grid Resolution:** For the region $[0, 1)$, we partition it into intervals of width $\frac{1}{2^0} = 1$. This gives us exactly one interval: $[0, 1)$.
- **The Assignment Rule:**
  - If $f(x) \in [0, 1)$, we assign $\phi_0(x) = 0$ (the lower bound of the strip).
  - If $f(x) \in [1, \infty]$, we assign $\phi_0(x) = 1 = 2^0$ (our current ceiling).

Stage 2: Building ϕ_1

Now we double our height ceiling to $2^1 = 2$, and sharpen our grid resolution to step sizes of $\frac{1}{2^1} = \frac{1}{2}$.

- **The Cutoff:** The infinite tail now starts at $2^1 = 2$. If $f(x) \geq 2$, we set $\phi_1(x) = 2$.
- **The Grid Resolution:** The bounded region $[0, 2)$ is chopped into steps of width $\frac{1}{2}$. This yields $2 \times 2^1 = 4$ strips:
$$\left[0, \frac{1}{2}\right), \quad \left[\frac{1}{2}, 1\right), \quad \left[1, \frac{3}{2}\right), \quad \left[\frac{3}{2}, 2\right)$$
- **The Assignment Rule:** For any point x, if $f(x)$ lands in one of these bounded intervals, $\phi_1(x)$ takes the value of the lower edge of that interval (e.g., if $f(x) \in [\frac{1}{2}, 1)$, then $\phi_1(x) = \frac{1}{2}$).

Stage 3: Building ϕ_2

We raise the ceiling to $2^2 = 4$, and make the grid resolution twice as fine: step sizes of $\frac{1}{2^2} = \frac{1}{4}$.

- **The Cutoff:** If $f(x) \geq 4$, we set $\phi_2(x) = 4$.
- **The Grid Resolution:** The bounded range $[0, 4)$ is sliced into sub-intervals of width $\frac{1}{4}$. This creates $4 \times 2^2 = 16$ distinct horizontal strips.
- **The Assignment Rule:** $\phi_2(x)$ tracks the lower endpoint of whichever of the 16 strips $f(x)$ occupies.

General Stage: Building ϕ_n

Following this exact progression, for any arbitrary integer n, our ceiling becomes 2^n and our resolution width shrinks to $\frac{1}{2^n}$. The total number of bounded horizontal slices is $2^n \times 2^n = 2^{2n}$.

Let us formally define the preimages of these slices in the domain space X:

- For the bounded strips, where the index k runs from 0 up to $2^{2n} - 1$:
$$E_n^k = f^{-1}\left(\left(\frac{k}{2^n}, \frac{k+1}{2^n}\right]\right)$$
- For the unbounded tail region beyond our ceiling:
$$F_n = f^{-1}\big((2^n, \infty]\big)$$

Since f is given to be a measurable function, the preimages of these Borel intervals are guaranteed to be measurable sets in X ($E_n^k, F_n \in \mathcal{M}$).

By multiplying the characteristic indicator functions of these sets by their corresponding lower-edge values, we arrive at the general formula for ϕ_n:

$$\phi_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \chi_{E_n^k} + 2^n \chi_{F_n}$$

Because this is a finite linear combination of characteristic functions of measurable sets, ϕ_n is a mathematically valid simple function for every n.

—

Phase 2: Verifying the Monotone Growth ($\phi_n \leq \phi_{n+1}$)

To show that our sequence grows monotonically towards f, let us look at what happens when we move from stage n to stage $n+1$.

When moving to stage $n+1$, the grid spacing cuts exactly in half, from $\frac{1}{2^n}$ to $\frac{1}{2^{n+1}}$. This means each single horizontal strip from stage n is split perfectly into two sub-strips at stage $n+1$:

$$\text{Strip } k \text{ of stage } n : \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) \implies \left[\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}}\right) \cup \left[\frac{2k+1}{2^{n+1}}, \frac{2k+2}{2^{n+1}}\right)$$

Now let us check the value of our function at a point x:

- If $f(x)$ lands in the bottom half of the old strip, its assigned lower-bound value remains exactly the same: $\phi_{n+1}(x) = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = \phi_n(x)$.
- If $f(x)$ lands in the top half of the old strip, its new lower-bound value steps upward: $\phi_{n+1}(x) = \frac{2k+1}{2^{n+1}} > \frac{k}{2^n} = \phi_n(x)$.
- If $f(x)$ was originally trapped by the ceiling 2^n (meaning $\phi_n(x) = 2^n$), the new ceiling at stage $n+1$ has moved up to 2^{n+1}, allowing $\phi_{n+1}(x)$ to measure higher values up to 2^{n+1}.

In all possible scenarios, the value never drops. Thus, $0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f$ holds for all $x \in X$.

—

Phase 3: Setting Up the Error Bound ($0 \leq f - \phi_n \leq 2^{-n}$)

By design, if the value of our function $f(x)$ is bounded below our ceiling 2^n, it must fall into exactly one of the finite strips $\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$.

On that strip, our simple function rounds the value down to the nearest lower edge: $\phi_n(x) = \frac{k}{2^n}$. Since the total width of this strip is precisely $\frac{1}{2^n}$, the absolute error between the true value $f(x)$ and our rounded value $\phi_n(x)$ cannot exceed the width of that strip. Therefore:

$$\text{If } f(x) \leq 2^n, \quad \text{then } 0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n} = 2^{-n}$$

—

Phase 4: Proving Pointwise Convergence

Let $x \in X$ be any fixed arbitrary point in our domain. We consider two possibilities for the true value of $f(x)$:

- **Case 1: $f(x)$ is finite ($f(x) < \infty$)**
  Since $f(x)$ is a finite number, by the Archimedean property of real numbers, we can choose a large integer N such that $2^N \geq f(x)$. For all indices $n \geq N$, the ceiling 2^n has surpassed the value of $f(x)$. We can then confidently apply our error bound equation from Phase 3:

  $$0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n} \quad \text{for all } n \geq N$$

  Taking the limit as $n \to \infty$, the squeeze theorem forces the error to zero:

  $$\lim_{n\to\infty} \big(f(x) - \phi_n(x)\big) = 0 \implies \lim_{n\to\infty} \phi_n(x) = f(x)$$

- **Case 2: $f(x)$ is infinite ($f(x) = \infty$)**
  If $f(x) = \infty$, then for every single index n, the value of $f(x)$ always sits past our ceiling 2^n. According to our construction rule, if a point lies in the tail F_n, the function outputs the ceiling value:

  $$\phi_n(x) = 2^n \quad \text{for all } n$$

  Taking the limit of this sequence gives:

  $$\lim_{n\to\infty} \phi_n(x) = \lim_{n\to\infty} 2^n = \infty = f(x)$$

Thus, $\phi_n \to f$ pointwise across the entire domain space X.

—

Phase 5: Proving Uniform Convergence on Bounded Sets

Let $B \subset X$ be any subset of our domain where the function f is bounded. This means there exists some fixed finite constant $M > 0$ such that:

$$f(x) \leq M \quad \text{for all } x \in B$$

Let us choose a fixed integer N large enough such that $2^N \geq M$. This means that for any index $n \geq N$, the entire set B fits completely below our moving ceiling ($f(x) \leq M \leq 2^N \leq 2^n$).

Because every single point $x \in B$ lies below the ceiling for any $n \geq N$, the uniform error bound derived in Phase 3 applies globally across the entire set B at the same time:

$$\sup_{x \in B} |f(x) - \phi_n(x)| \leq \frac{1}{2^n} \quad \text{for all } n \geq N$$

Since $\frac{1}{2^n} \to 0$ as $n \to \infty$, the supremum of the error over the entire set B vanishes. This proves that $\phi_n \to f$ uniformly on B.

—

Part (b): Complex-Valued Functions ($f : X \to \mathbb{C}$)

Now we extend our proof to a complex-valued measurable function f. We do this by breaking the complex function down into its real component building blocks, applying part (a), and recombining them.

Step 1: Splitting into Non-negative Real Functions

Let us decompose f into its real and imaginary parts:

$$f(x) = g(x) + ih(x) \quad \text{where } g = \operatorname{Re} f \text{ and } h = \operatorname{Im} f$$

By Corollary 2.5, since f is complex-measurable, both g and h are real-valued measurable functions on X.

Next, we split these real functions into their corresponding positive and negative parts:

$$g(x) = g^+(x) - g^-(x) \qquad \text{and} \qquad h(x) = h^+(x) - h^-(x)$$

By our earlier work on functional decompositions, all four functions—g^+, g^-, h^+, and h^-—are non-negative, finite, measurable functions on X.

Step 2: Constructing Approximating Sequences for Each Component

Since g^+, g^-, h^+, h^- are non-negative and measurable, we can apply the results from **Part (a)** to each component individually. This yields four independent sequences of non-negative simple functions that climb monotonically to their targets:

$$\psi_n^+ \nearrow g^+, \quad \psi_n^- \nearrow g^-, \quad \zeta_n^+ \nearrow h^+, \quad \zeta_n^- \nearrow h^-$$

Now, we combine these four separate simple sequences back into a single complex-valued sequence $\{\phi_n\}$ defined by:

$$\phi_n = \left(\psi_n^+ - \psi_n^-\right) + i\left(\zeta_n^+ - \zeta_n^-\right)$$

Because the sum, difference, and scalar multiplication of simple functions yield simple functions, ϕ_n is a complex simple function for every n.

Step 3: Verifying Pointwise and Uniform Convergence

Since limits commute with basic linear combinations:

$$\begin{aligned}\lim_{n\to\infty} \phi_n &= \left(\lim_{n\to\infty} \psi_n^+ - \lim_{n\to\infty} \psi_n^-\right) + i\left(\lim_{n\to\infty} \zeta_n^+ - \lim_{n\to\infty} \zeta_n^-\right)\\ &= (g^+ - g^-) + i(h^+ - h^-) = g + ih = f\end{aligned}$$

This proves $\phi_n \to f$ pointwise. Furthermore, if f is bounded on a set B, then its components g and h must be bounded on B, forcing their positive and negative parts to be bounded. Since each component sequence converges uniformly on B by Part (a), their linear combination ϕ_n converges uniformly to f on B.

Step 4: Proving Monotone Growth of the Absolute Values ($|\phi_n| \leq |\phi_{n+1}|$)

To complete the proof, we must show that the magnitudes grow monotonically: $|\phi_n| \leq |\phi_{n+1}| \leq |f|$.

Recall that for our constructed sequences, at any point x, the functions $\psi_n^+(x)$ and $\psi_n^-(x)$ cannot be non-zero at the same time because $g^+(x)$ and $g^-(x)$ are disjointly supported. Therefore, the absolute value of the real part simplifies cleanly:

$$|\operatorname{Re}\phi_n| = |\psi_n^+ - \psi_n^-| = \psi_n^+ + \psi_n^-$$

Because $\psi_n^+ \leq \psi_{n+1}^+$ and $\psi_n^- \leq \psi_{n+1}^-$, it follows that:

$$|\operatorname{Re}\phi_n| \leq |\operatorname{Re}\phi_{n+1}| \leq \operatorname{Re} f^+ + \operatorname{Re} f^- = |g|$$

By identical logic applied to the imaginary components:

$$|\operatorname{Im}\phi_n| \leq |\operatorname{Im}\phi_{n+1}| \leq |h|$$

Using the definition of the complex modulus ($|z| = \sqrt{(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2}$), since the absolute values of both components grow monotonically at every step, the total magnitude must grow monotonically:

$$|\phi_n| = \sqrt{|\operatorname{Re}\phi_n|^2 + |\operatorname{Im}\phi_n|^2} \leq \sqrt{|\operatorname{Re}\phi_{n+1}|^2 + |\operatorname{Im}\phi_{n+1}|^2} = |\phi_{n+1}|$$

And since components are bounded by their targets:

$$|\phi_{n+1}| \leq \sqrt{|g|^2 + |h|^2} = |f|$$

This establishes $0 \leq |\phi_1| \leq |\phi_2| \leq \cdots \leq |f|$, completes the proof. □

Proposition (Proposition 2.11). *The following implications are valid iff the measure μ is complete:*

a. If f is measurable and $f = g$ μ-a.e., then g is measurable.

b. If f_n is measurable for $n \in \mathbb{N}$ and $f_n \to f$ μ-a.e., then f is measurable.

Proof. To make this proof fully transparent for your students, we will split our argument into two directional halves: the forward direction ($\Longrightarrow$), where we assume completeness to prove statements (a) and (b), and the backward direction ($\Longleftarrow$), where we show that if either statement holds, the measure space must be complete.

—

Direction 1: Assume μ is Complete ($\Longrightarrow$)

Let $(X, \mathcal{M}, \mu)$ be a complete measure space. Recall that a measure space is **complete** if every subset of a set of measure zero is itself measurable (i.e., if $N \in \mathcal{M}$ with $\mu(N) = 0$, then any $E \subset N$ satisfies $E \in \mathcal{M}$, which forces $\mu(E) = 0$).

Proof of Part (a)

We are given that f is an $\mathcal{M}$-measurable function, and $f = g$ μ-almost everywhere (μ-a.e.).

Step 1: **Isolate the exceptional set:** By definition of μ-a.e., the set of points where f and g differ is contained within a measurable set of measure zero. Let:

$$N = \{x \in X : f(x) \neq g(x)\}$$

Since $f = g$ μ-a.e., there exists a set $M \in \mathcal{M}$ such that $N \subset M$ and $\mu(M) = 0$.

Step 2: **Analyze preimages under g:** To prove that g is measurable, we must show that for any Borel set $B \in \mathcal{B}_{\mathbb{R}}$ (or $\mathcal{B}_{\overline{\mathbb{R}}}$), its preimage under g is a member of $\mathcal{M}$. Let us partition the preimage $g^{-1}(B)$ into two distinct parts: points outside our exceptional set M, and points inside M.

$$g^{-1}(B) = \left(g^{-1}(B) \cap M^c\right) \cup \left(g^{-1}(B) \cap M\right)$$

Step 3: **Evaluate each component:**

- **Outside the exceptional set (M^c):** On the set M^c, the functions f and g are identical. Therefore, $g^{-1}(B) \cap M^c = f^{-1}(B) \cap M^c$. Because f is measurable, $f^{-1}(B) \in \mathcal{M}$. Since $M \in \mathcal{M}$, its complement $M^c \in \mathcal{M}$. The intersection of two measurable sets is measurable, so $\left(f^{-1}(B) \cap M^c\right) \in \mathcal{M}$.
- **Inside the exceptional set (M):** Let $E = g^{-1}(B) \cap M$. By construction, E is a subset of M ($E \subset M$). Since our measure space is **complete** and M is a measurable set of measure zero ($\mu(M) = 0$), the subset E is automatically guaranteed to be a measurable set ($E \in \mathcal{M}$).

Step 4: **Recombine the components:** We can now express $g^{-1}(B)$ as the union of two verified elements of $\mathcal{M}$:

$$g^{-1}(B) = \underbrace{\left(f^{-1}(B) \cap M^c\right)}_{\in \mathcal{M}} \cup \underbrace{\left(g^{-1}(B) \cap M\right)}_{\in \mathcal{M}}$$

Since $\mathcal{M}$ is a σ-algebra, it is closed under countable unions. Therefore, $g^{-1}(B) \in \mathcal{M}$ for any Borel set B, proving that g is an $\mathcal{M}$-measurable function.

Proof of Part (b)

We are given a sequence $\{f_n\}$ of measurable functions such that $f_n \to f$ μ-a.e..

Step 1: **Isolate the non-convergence set:** Let $N \subset X$ be the set of points where the sequence $f_n(x)$ fails to converge to $f(x)$. Since $f_n \to f$ μ-a.e., there exists a measurable set $M \in \mathcal{M}$ such that $N \subset M$ and $\mu(M) = 0$.

Step 2: **Construct a modified sequence:** Let us define a new sequence of functions $\{g_n\}$ that forces convergence everywhere by zeroing out the sequence on our exceptional set M:

$$g_n(x) = f_n(x)\chi_{M^c}(x) = \begin{cases} f_n(x) & \text{if } x \notin M \\ 0 & \text{if } x \in M \end{cases}$$

Because each f_n is measurable and M^c is a measurable set, the characteristic function χ_{M^c} is measurable. The product of two measurable functions is measurable, so each g_n is strictly an $\mathcal{M}$-measurable function.

Step 3: **Analyze the pointwise limit of g_n:** Because we explicitly zeroed out the sequence on M, the sequence $g_n(x)$ converges pointwise at *every single point* in X to a limit function $g(x)$, defined by:

$$g(x) = \lim_{n\to\infty} g_n(x) = \begin{cases} f(x) & \text{if } x \notin M \\ 0 & \text{if } x \in M \end{cases}$$

By Proposition 2.7 (or Corollary 2.9), the pointwise limit of a sequence of measurable functions is always measurable. Therefore, g is a measurable function.

Step 4: **Link g back to f:** By comparing the definitions of $f(x)$ and $g(x)$, we see that they are identical everywhere except possibly on our exceptional set M. Because M has measure zero, we have:

$$f = g \quad \mu\text{-a.e.}$$

We have already established that g is a measurable function. By applying our result from **Part (a)**, since f equals a measurable function g μ-almost everywhere, f is guaranteed to be measurable.

—

Direction 2: Prove Completeness is Necessary ($\Longleftarrow$)

To establish the "iff" condition, we must show that if $(X, \mathcal{M}, \mu)$ is *not* complete, then implications (a) and (b) can fail. We do this by demonstrating that if statement (a) holds, the space must be complete.

Suppose statement (a) holds valid on our measure space. Let $M \in \mathcal{M}$ be an arbitrary set with $\mu(M) = 0$, and let $E \subset M$ be an arbitrary subset of M. To prove completeness, we must show that $E \in \mathcal{M}$.

Step 1: **Define a tracking function for E:** Let $g = \chi_E$ be the characteristic indicator function of our subset E:

$$g(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Step 2: **Define a baseline comparison function:** Let $f = 0$ be the zero function on X. The constant function 0 is trivially measurable.

Step 3: **Show that $f = g$ μ-almost everywhere:** Let us find the set of points where $f(x) \neq g(x)$:

$$\{x \in X : f(x) \neq g(x)\} = \{x \in X : 0 \neq \chi_E(x)\} = E$$

By our initial choice, $E \subset M$ and $\mu(M) = 0$. Therefore, the set of points where f and g differ is contained inside a set of measure zero. This satisfies the definition of almost everywhere equivalence:

$$f = g \quad \mu\text{-a.e.}$$

Step 4: **Invoke Statement (a):** By our hypothesis, statement (a) is valid. Since f is measurable and $f = g$ μ-a.e., statement (a) forces the function $g = \chi_E$ to be an $\mathcal{M}$-measurable function.

Step 5: **Conclude completeness:** Recall the fundamental property of indicator functions: a characteristic function χ_E is measurable if and only if its underlying set E belongs to the σ-algebra $\mathcal{M}$. Because $g = \chi_E$ is proven to be measurable, it follows directly that $E \in \mathcal{M}$.

Since an arbitrary subset E of any null set M is a member of $\mathcal{M}$, the measure space $(X, \mathcal{M}, \mu)$ is complete. (Note: An identical argument can be made using statement (b) by setting a constant sequence $f_n = 0$, which converges μ-a.e. to χ_E). □

Proposition (Proposition 2.12). *Let $(X, \mathcal{M}, \mu)$ be a measure space and let $(X, \overline{\mathcal{M}}, \overline{\mu})$ be its completion. If f is an $\overline{\mathcal{M}}$-measurable function on X, there is an $\mathcal{M}$-measurable function g such that $f = g$ $\overline{\mu}$-almost everywhere.*

Proof. To ensure total clarity for students, we will break this proof into three progressive stages: first establishing the result for characteristic functions, expanding it linearly to simple functions, and finally taking the pointwise limit to cover arbitrary measurable functions.

Throughout this proof, we recall the definition of the completion $\overline{\mathcal{M}}$: a set $E \in \overline{\mathcal{M}}$ if and only if it can be decomposed as $E = A \cup E_0$, where $A \in \mathcal{M}$ and $E_0 \subset N_0$ for some $N_0 \in \mathcal{M}$ with $\mu(N_0) = 0$. By definition, $\overline{\mu}(E) = \mu(A)$.

—

Part 1: The Base Case for Characteristic Functions

Let $f = \chi_E$ be an $\overline{\mathcal{M}}$-measurable characteristic function, which means that the underlying set E belongs to the completed σ-algebra $\overline{\mathcal{M}}$.

Step 1: **Decompose the completed set:** By the definition of a completion, since $E \in \overline{\mathcal{M}}$, we can write $E = A \cup E_0$, where $A \in \mathcal{M}$ and $E_0 \subset N_0$ for some symmetric null set $N_0 \in \mathcal{M}$ satisfying $\mu(N_0) = 0$.

Step 2: **Construct the target function:** We define our candidate function $g : X \to \mathbb{R}$ to be the characteristic function of the clean, original measurable set A:

$$g = \chi_A$$

Since $A \in \mathcal{M}$, the function g is strictly $\mathcal{M}$-measurable.

Step 3: **Verify almost-everywhere equality:** Let us find where $f(x)$ and $g(x)$ fail to match:

$$\{x \in X : \chi_E(x) \neq \chi_A(x)\} = E \setminus A = E_0 \subset N_0$$

Since $N_0 \in \mathcal{M}$ and $\mu(N_0) = 0$, the definition of the completion ensures that $\overline{\mu}(N_0) = \mu(N_0) = 0$. Since the exceptional set is contained within a set of $\overline{\mu}$-measure zero, we conclude:

$$f = g \quad \overline{\mu}\text{-almost everywhere}$$

—

Part 2: Extension to Simple Functions

Let f be an $\overline{\mathcal{M}}$-measurable simple function.

Step 1: **Represent the function:** We can write f using its standard representation:

$$f = \sum_{j=1}^{n} z_j \chi_{E_j}$$

where $z_j \in \mathbb{C}$ and each component set $E_j \in \overline{\mathcal{M}}$.

Step 2: **Apply the base case term-by-term:** By applying **Part 1** to each individual completed set E_j, we find corresponding sets $A_j \in \mathcal{M}$ such that $E_j \setminus A_j \subset N_j$ for some $N_j \in \mathcal{M}$ with $\mu(N_j) = 0$.

Step 3: **Construct the simple target function:** We piece together these separate components to build a clean function g:

$$g = \sum_{j=1}^{n} z_j \chi_{A_j}$$

Since every single $A_j \in \mathcal{M}$, the linear combination g is strictly $\mathcal{M}$-measurable.

Step 4: **Verify almost-everywhere equality:** The functions f and g can only differ if at least one of their component pairs differ. Therefore, the total exceptional set is bounded by the finite union of the individual exceptional sets:

$$\{x \in X : f(x) \neq g(x)\} \subset \bigcup_{j=1}^{n} (E_j \setminus A_j) \subset \bigcup_{j=1}^{n} N_j$$

Let $N_{\text{total}} = \bigcup_{j=1}^{n} N_j$. Since $\mathcal{M}$ is a σ-algebra, $N_{\text{total}} \in \mathcal{M}$. By subadditivity:

$$\mu(N_{\text{total}}) \leq \sum_{j=1}^{n} \mu(N_j) = \sum_{j=1}^{n} 0 = 0$$

Since the subset of disagreement is trapped inside a set of $\overline{\mu}$-measure zero, $f = g$ $\overline{\mu}$-a.e. holds for all simple functions.

—

Part 3: The General Case via Pointwise Limits

Let f be an arbitrary general $\overline{\mathcal{M}}$-measurable function on X.

Step 1: **Approximate with simple functions:** By Theorem 2.10, any measurable function can be realized as the pointwise limit of a sequence of simple functions. Therefore, there exists a sequence $\{\phi_n\}$ of $\overline{\mathcal{M}}$-measurable simple functions such that:

$$\phi_n(x) \to f(x) \quad \text{pointwise for all } x \in X$$

Step 2: **Replace the sequence components:** By applying our results from **Part 2** to each term in the sequence, for every index n, there exists an $\mathcal{M}$-measurable simple function ψ_n that matches ϕ_n everywhere except on an exceptional set $E_n \in \overline{\mathcal{M}}$ with $\overline{\mu}(E_n) = 0$. By definition, each E_n is contained within a measurable null set $M_n \in \mathcal{M}$ such that $\mu(M_n) = 0$.

Step 3: **Construct a unified exceptional set N:** We collect all these separate problem sites into a single master set:

$$N = \bigcup_{n=1}^{\infty} M_n$$

Why does this N work?

- Since each $M_n \in \mathcal{M}$ and a σ-algebra is closed under countable unions, our master set N is guaranteed to belong to $\mathcal{M}$.

- By countable subadditivity of measures, $\mu(N) \leq \sum_{n=1}^{\infty} \mu(M_n) = \sum_{n=1}^{\infty} 0 = 0$. Thus, N is a strictly valid measurable set of measure zero.
- Crucially, because N contains every single M_n, it contains all individual exceptional sets ($N \supset \bigcup_1^{\infty} E_n$). Therefore, outside of this single set N, **every single function ψ_n matches its counterpart ϕ_n perfectly simultaneously.**

Step 4: **Construct the final function g:** We define our final target function g by taking the pointwise limit of our clean sequence, while explicitly forcing it to zero on our master exceptional set to protect against divergence:

$$g(x) = \lim_{n\to\infty} \chi_{X\setminus N}(x)\psi_n(x)$$

Step 5: **Verify that g is measurable:** Since $N \in \mathcal{M}$, its complement $X \setminus N \in \mathcal{M}$, meaning the indicator function $\chi_{X\setminus N}$ is $\mathcal{M}$-measurable. Because each ψ_n is $\mathcal{M}$-measurable, the products are $\mathcal{M}$-measurable. Finally, by Corollary 2.9, the pointwise limit of a sequence of measurable functions is always measurable. Therefore, g is an $\mathcal{M}$-measurable function.

Step 6: **Verify that $f = g$ $\overline{\mu}$-almost everywhere:** Let us evaluate the behavior of our functions on the complement set N^c: For any point $x \in N^c$:

- The indicator function evaluates to $\chi_{X\setminus N}(x) = 1$.
- Since $x \notin N$, it belongs to none of the exceptional sets E_n, meaning $\psi_n(x) = \phi_n(x)$ for every single n.

Substituting these values into the definition of g:

$$g(x) = \lim_{n\to\infty} (1 \cdot \psi_n(x)) = \lim_{n\to\infty} \phi_n(x) = f(x)$$

This demonstrates that $g(x) = f(x)$ for every single point in N^c. Because the complement set N has a $\overline{\mu}$-measure of zero, we have established that:

$$f = g \quad \overline{\mu}\text{-almost everywhere}$$

This completes the proof for the general case. □

2.2 Integration of Nonnegative Functions

Definition 37 (The Space L^+)**.** *We fix a measure space $(X, \mathcal{M}, \mu)$. We define:*

$$L^+ = \{f : X \to [0, \infty] \mid f \text{ is } \mathcal{M}\text{-measurable}\}$$

representing the space of all non-negative, extended real-valued measurable functions on X.

Definition 38 (Integral of a Nonnegative Simple Function)**.** *If $\phi \in L^+$ is a simple function with its standard representation given by:*

$$\phi = \sum_{j=1}^{n} a_j \chi_{E_j}$$

*where $a_j \in [0, \infty)$ are distinct values and the sets $E_j = \phi^{-1}(\{a_j\})$ are pairwise disjoint elements of $\mathcal{M}$ whose union is X, we define the **integral of ϕ with respect to** μ by:*

$$\int \phi \, d\mu = \sum_{j=1}^{n} a_j \mu(E_j)$$

Explanatory Remarks

Remark (The Necessity of the $0 \cdot \infty = 0$ Convention)**.** *In measure theory, we strictly enforce the structural convention that:*

$$0 \cdot \infty = 0$$

This algebraic rule is necessary because a simple function can take the value 0 on a set of infinite measure. Without this convention, computing $\int \phi \, d\mu$ using the definition above would leave the term $0 \cdot \mu(E_j) = 0 \cdot \infty$ completely undefined, paralyzing our arithmetic on unbounded domains.

Example: *Consider the real line with the Lebesgue measure, $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, m)$. Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined as:*

$$f(x) = \begin{cases} 1 & \text{if } 0 < x \le 2 \\ 2 & \text{if } 2 < x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

The standard representation of this simple function partitions the entire space $X = \mathbb{R}$ into three disjoint measurable sets:

- $E_1 = (0, 2]$ *with value* $a_1 = 1$*, where* $m(E_1) = 2$.
- $E_2 = (2, 4]$ *with value* $a_2 = 2$*, where* $m(E_2) = 2$.
- $E_3 = (-\infty, 0] \cup (4, \infty)$ *with value* $a_3 = 0$*, where* $m(E_3) = \infty$.

When we calculate the total integral over $\mathbb{R}$ according to our definition:

$$\begin{aligned} \int f\, dm &= a_1 m(E_1) + a_2 m(E_2) + a_3 m(E_3) \\ &= (1 \cdot 2) + (2 \cdot 2) + \underbrace{(0 \cdot \infty)}_{=0} \\ &= 2 + 4 + 0 = 6 \end{aligned}$$

By assigning $0 \cdot \infty = 0$, the infinite tails of the real line where the function is dead drop away cleanly, matching our geometric intuition that the total area under this step curve is finite.

Remark (Integration over Measurable Subsets)**.** *If $A \in \mathcal{M}$ is a measurable subset of X, we define the integration of a simple function ϕ over the sub-domain A by restricting the function using the characteristic indicator χ_A:*

$$\int_A \phi\, d\mu = \int \phi \chi_A\, d\mu$$

This modification preserves our operational framework because the product $\phi\chi_A$ remains a strictly valid simple function. Given the standard representation $\phi = \sum a_j \chi_{E_j}$, distributing χ_A reveals:

$$\phi\chi_A = \left(\sum_{j=1}^{n} a_j \chi_{E_j} \right) \chi_A = \sum_{j=1}^{n} a_j \chi_{E_j \cap A}$$

Because $\mathcal{M}$ is a σ-algebra, each intersection $(E_j \cap A)$ is a measurable set, meaning $\phi\chi_A$ is a finite linear combination of measurable indicator functions—confirming it is simple. Thus, its subset integral can be computed directly as:

$$\int_A \phi\, d\mu = \sum_{j=1}^{n} a_j \mu(E_j \cap A)$$

Remark (Notational Variations)**.** *When there is no risk of ambiguity regarding the underlying measure space, several standard shorthand notations are used interchangeably in analytical literature:*

- $\int \phi$*: The most minimalist representation, omitting the measure symbol entirely when the background context implies μ.*
- $\int \phi(x)\, d\mu(x)$*: Used to explicitly track the dummy variable of integration, which becomes helpful when handling compound functions or evaluating nested multiple integrals.*
- $\int \phi(x)\, \mu(dx)$*: A notation common in probability theory and functional analysis that treats the measure expression as an operator acting on changing differential increments.*

Proposition (Proposition 2.13)**.** *Let ϕ and ψ be simple functions in L^+.*

a. *If $c \ge 0$, $\int c\phi = c \int \phi$.*

b. $\int(\phi + \psi) = \int \phi + \int \psi$.

c. *If $\phi \le \psi$, then $\int \phi \le \int \psi$.*

d. *The map $A \mapsto \int_A \phi\, d\mu$ is a measure on $\mathcal{M}$.*

Proof. **Detailed Proof of Part (a)**

Let $\phi = \sum_{j=1}^{n} a_j \chi_{E_j}$ be the standard representation of the simple function ϕ, where $a_j \in [0, \infty)$ are the distinct values taken by ϕ, and $E_j = \phi^{-1}(\{a_j\})$ are pairwise disjoint measurable sets whose union is X.

Step 1: Handle the case $c = 0$. If $c = 0$, then the function $c\phi = 0 \cdot \phi$ is the constant zero function on X. The standard representation of the zero function is simply $0 \cdot \chi_X$. By definition of the simple integral:

$$\int 0\, d\mu = 0 \cdot \mu(X)$$

Under the core conventions of measure theory, if $\mu(X) = \infty$, the product evaluates via $0 \cdot \infty = 0$. Thus, $\int 0 \cdot \phi = 0$. On the other side, $c \int \phi = 0 \cdot \int \phi = 0$. Hence, equality holds when $c = 0$.

Step 2: Handle the case $c > 0$. If $c > 0$, the function $c\phi$ scales every value output. Since the original constants a_j are distinct, the scaled values ca_j remain distinct. The preimage sets are completely unaffected because the set of points where $c\phi(x) = ca_j$ is exactly the set of points where $\phi(x) = a_j$:

$$(c\phi)^{-1}(\{ca_j\}) = \phi^{-1}(\{a_j\}) = E_j$$

Therefore, the standard representation of the new function $c\phi$ is exactly $\sum_{j=1}^{n} (ca_j) \chi_{E_j}$.

Step 3: Apply the definition of the integral. Using the definition of the simple function integral directly on this standard representation:

$$\int c\phi\, d\mu = \sum_{j=1}^{n} (ca_j) \mu(E_j)$$

Since c is a fixed real constant factor in each term of this finite algebraic sum, we can factor it out using standard distributive laws:

$$\int c\phi\, d\mu = c \sum_{j=1}^{n} a_j \mu(E_j) = c \int \phi\, d\mu$$

Detailed Proof of Part (b)

Let $\phi = \sum_{j=1}^{n} a_j \chi_{E_j}$ and $\psi = \sum_{k=1}^{m} b_k \chi_{F_k}$ be the standard representations of ϕ and ψ respectively. This establishes that $\{E_j\}_{j=1}^{n}$ is a partition of X and $\{F_k\}_{k=1}^{m}$ is another partition of X.

Step 1: Set-Theoretic breakdown via common refinement. Because the sets $\{E_j\}$ cover the entire space X, intersecting any set F_k with the entire collection decomposes it into localized, non-overlapping patches:

$$F_k = F_k \cap X = F_k \cap \left(\bigcup_{j=1}^{n} E_j \right) = \bigcup_{j=1}^{n} (E_j \cap F_k)$$

Similarly, each set E_j can be decomposed across the partition components of F_k:

$$E_j = E_j \cap X = E_j \cap \left(\bigcup_{k=1}^{m} F_k \right) = \bigcup_{k=1}^{m} (E_j \cap F_k)$$

Because these components are pairwise disjoint, applying the property of finite additivity of the measure μ converts these set unions into clear numerical sums:

$$\mu(E_j) = \sum_{k=1}^{m} \mu(E_j \cap F_k) \quad \text{and} \quad \mu(F_k) = \sum_{j=1}^{n} \mu(E_j \cap F_k)$$

Step 2: Express independent integrals using the refined sets. We can now expand the definition of $\int \phi$ and $\int \psi$ by substituting these measure decompositions:

$$\int \phi = \sum_{j=1}^{n} a_j \mu(E_j) = \sum_{j=1}^{n} a_j \left(\sum_{k=1}^{m} \mu(E_j \cap F_k) \right) = \sum_{j,k} a_j \mu(E_j \cap F_k)$$

$$\int \psi = \sum_{k=1}^{m} b_k \mu(F_k) = \sum_{k=1}^{m} b_k \left(\sum_{j=1}^{n} \mu(E_j \cap F_k) \right) = \sum_{j,k} b_k \mu(E_j \cap F_k)$$

Adding these two decoupled double sums together allows us to group terms by their shared intersection measure $\mu(E_j \cap F_k)$:

$$\int \phi + \int \psi = \sum_{j,k} a_j \mu(E_j \cap F_k) + \sum_{j,k} b_k \mu(E_j \cap F_k) = \sum_{j,k} (a_j + b_k) \mu(E_j \cap F_k)$$

Step 3: Match with the integral of the sum function. Now let's look directly at the combined function $(\phi+\psi)$. For any arbitrary point $x \in X$, that point must fall into exactly one specific intersection tile $E_j \cap F_k$. At that point, $\phi(x) = a_j$ and $\psi(x) = b_k$, which means:

$$(\phi + \psi)(x) = a_j + b_k \quad \text{for all } x \in E_j \cap F_k$$

Thus, we can represent the combined function as a linear combination over the refinement grid:

$$\phi + \psi = \sum_{j,k} (a_j + b_k) \chi_{E_j \cap F_k}$$

Although this layout may not be a *standard representation* (since different grid tiles might accidentally share the same total value, e.g., $a_j + b_k = a_{j'} + b_{k'}$), grouping identical coefficients together proves that the calculated integral is invariant under refined representations. Calculating the integral from this form yields:

$$\int (\phi + \psi) = \sum_{j,k} (a_j + b_k) \mu(E_j \cap F_k)$$

This matches our final equation from Step 2 identically, completing the proof of additivity.

Detailed Proof of Part (c)

Let $\phi = \sum_{j=1}^{n} a_j \chi_{E_j}$ and $\psi = \sum_{k=1}^{m} b_k \chi_{F_k}$ be the standard representations. We are given that $\phi(x) \leq \psi(x)$ for all $x \in X$.

Step 1: Pointwise constraint on overlapping zones. Consider any refined intersection domain $E_j \cap F_k$. If this set is empty, its measure is zero and it does not affect our calculations. If it is non-empty, there is a point $x \in E_j \cap F_k$. At this point, $\phi(x) = a_j$ and $\psi(x) = b_k$. Because $\phi \leq \psi$ everywhere on X, it follows that:

$$a_j \leq b_k \quad \text{whenever } E_j \cap F_k \neq \varnothing$$

Step 2: Comparison of the expanded sums. Using the refinement mesh formulas derived in Part (b), we express both integrals over the identical base sets:

$$\int \phi = \sum_{j,k} a_j \mu(E_j \cap F_k) \quad \text{and} \quad \int \psi = \sum_{j,k} b_k \mu(E_j \cap F_k)$$

Since the scalar measures $\mu(E_j \cap F_k)$ are strictly non-negative, and $a_j \leq b_k$ for every single non-empty tile, we can perform a direct term-by-term inequality comparison:

$$a_j \mu(E_j \cap F_k) \leq b_k \mu(E_j \cap F_k) \quad \text{for all } j, k$$

Summing up these individual inequalities preserves the overall relationship, confirming:

$$\sum_{j,k} a_j \mu(E_j \cap F_k) \leq \sum_{j,k} b_k \mu(E_j \cap F_k) \implies \int \phi \leq \int \psi$$

Detailed Proof of Part (d)

Let $\phi = \sum_{j=1}^{n} a_j \chi_{E_j}$ be the standard representation of our simple function on X. We define a set function $\nu : \mathcal{M} \to [0, \infty]$ by $\nu(A) = \int_A \phi \, d\mu$. To prove that ν constitutes a valid measure, we must verify the two basic axioms of a measure.

Step 1: Verify that the empty set maps to zero. Let $A = \varnothing$. By definition of subset integration over a simple function:

$$\nu(\varnothing) = \sum_{j=1}^{n} a_j \mu(E_j \cap \varnothing) = \sum_{j=1}^{n} a_j \mu(\varnothing)$$

Since the underlying μ is a valid measure, $\mu(\varnothing) = 0$. Since all coefficients a_j are finite non-negative numbers, the product $a_j \cdot 0$ is 0 for every term. Thus, $\nu(\varnothing) = 0$.

Step 2: Verify Countable Additivity. Let $\{A_k\}_{k=1}^{\infty}$ be a sequence of pairwise disjoint sets in $\mathcal{M}$, and denote their union as $A = \bigcup_{k=1}^{\infty} A_k$. We apply the definition of subset integration to evaluate $\nu(A)$:

$$\nu(A) = \sum_{j=1}^{n} a_j \mu(A \cap E_j)$$

We expand the set intersection inside the measure expression using distributive laws:

$$A \cap E_j = \left(\bigcup_{k=1}^{\infty} A_k \right) \cap E_j = \bigcup_{k=1}^{\infty} (A_k \cap E_j)$$

Because the sets A_k are mutually disjoint, their individual intersections with E_j are also completely disjoint. Since the background measure μ is countably additive, we can rewrite this measure of a countable union as a convergent infinite series:

$$\mu(A \cap E_j) = \sum_{k=1}^{\infty} \mu(A_k \cap E_j)$$

Substituting this series expression back into our formula for $\nu(A)$ gives:

$$\nu(A) = \sum_{j=1}^{n} a_j \left(\sum_{k=1}^{\infty} \mu(A_k \cap E_j) \right) = \sum_{j=1}^{n} \sum_{k=1}^{\infty} a_j \mu(A_k \cap E_j)$$

Step 3: Change the summation order. The expression above consists of a finite sum of infinite series containing purely non-negative terms. By standard analytical rules governing series with values in $[0, \infty]$, we can change the order of summation without altering the final value:

$$\nu(A) = \sum_{k=1}^{\infty} \left(\sum_{j=1}^{n} a_j \mu(A_k \cap E_j) \right)$$

We recognize that the inner finite sum is exactly the definition of the subset integral over the set A_k:

$$\sum_{j=1}^{n} a_j \mu(A_k \cap E_j) = \int_{A_k} \phi \, d\mu = \nu(A_k)$$

Replacing the inner block with $\nu(A_k)$ leaves us with:

$$\nu(A) = \sum_{k=1}^{\infty} \nu(A_k)$$

This proves countable additivity, establishing that the assignment $A \mapsto \int_A \phi \, d\mu$ is a fully qualified measure. $\square$

Definition of the Integral for General Nonnegative Functions

We extend the definition of the integral from simple functions to all nonnegative measurable functions $f \in L^+$. If $f : X \to [0, \infty]$ is an $\mathcal{M}$-measurable function, we define its **integral** with respect to μ by:

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

—

Remark on the Case when f is Simple

An important check for consistency is ensuring that this new definition matches our original definition when f itself happens to be a simple function.

Proposition. *If f is a simple function in L^+, then the two definitions of $\int f \, d\mu$ agree.*

Proof. Let $f \in L^+$ be a simple function. Let $I_{\text{old}}(f)$ denote the integral of f calculated using its standard representation, and let $I_{\text{new}}(f)$ denote the value given by the supremum formula above:

$$I_{\text{new}}(f) = \sup \left\{ \int \phi \, d\mu : 0 \leq \phi \leq f, \, \phi \text{ is simple} \right\}$$

To establish complete clarity for students, we show that $I_{\text{new}}(f) = I_{\text{old}}(f)$ by proving both inequalities:

Step 1: **Proving $I_{\text{old}}(f) \leq I_{\text{new}}(f)$:** The set over which the supremum is taken consists of all simple functions ϕ that satisfy the constraint $0 \leq \phi \leq f$. Since f is assumed to be a simple function, and it trivially satisfies $0 \leq f \leq f$, the function f **is itself a member** of this family of functions.

Because the supremum of a set is greater than or equal to any individual element belonging to that set, evaluating the set at $\phi = f$ gives:

$$I_{\text{new}}(f) \geq I_{\text{old}}(f)$$

Step 2: **Proving $I_{\text{new}}(f) \leq I_{\text{old}}(f)$:** By Proposition 2.13c, the simple integral preserves order: if ϕ and f are simple functions and $\phi \leq f$, then $\int \phi \, d\mu \leq \int f \, d\mu$.

Therefore, $I_{\text{old}}(f)$ acts as an upper bound for every single numerical value in the set $\left\{ \int \phi \, d\mu : 0 \leq \phi \leq f, \, \phi \text{ simple} \right\}$. Since the supremum is the *least* upper bound, it cannot exceed this value:

$$I_{\text{new}}(f) \leq I_{\text{old}}(f)$$

Since both $I_{\text{old}}(f) \leq I_{\text{new}}(f)$ and $I_{\text{new}}(f) \leq I_{\text{old}}(f)$ are true, we conclude that the two definitions agree seamlessly. □

Exercise 7. *Let $f, g \in L^+$ and let $c \in [0, \infty)$. Show that:*

1. *$\int f \leq \int g$ whenever $f \leq g$.*
2. *$\int cf = c \int f$.*

Proof. Recall that for any general nonnegative measurable function $h \in L^+$, its integral is defined by the supremum over all lower simple functions:

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

Proof of Part 1 (Monotonicity):
Let S_f and S_g denote the families of simple functions bounded above by f and g respectively:

$$S_f = \{\phi : 0 \leq \phi \leq f, \, \phi \text{ is simple}\}$$

$$S_g = \{\phi : 0 \leq \phi \leq g, \, \phi \text{ is simple}\}$$

Suppose $\phi \in S_f$. By definition, ϕ is a simple function satisfying $0 \leq \phi(x) \leq f(x)$ for all $x \in X$. We are given the hypothesis that $f(x) \leq g(x)$ for all $x \in X$. By the transitive property of real inequalities, since $\phi(x) \leq f(x)$ and $f(x) \leq g(x)$, it must hold that:

$$\phi(x) \leq g(x) \quad \text{for all } x \in X$$

Therefore, ϕ automatically satisfies the conditions to belong to S_g. This directly establishes the set containment:

$$S_f \subseteq S_g$$

In set theory and real analysis, if a set of real numbers A is a subset of a set B ($A \subseteq B$), then the supremum of the larger set B must be greater than or equal to the supremum of the smaller set A. Taking the supremum of the integrals over both collections gives:

$$\sup_{\phi \in S_f} \int \phi \, d\mu \leq \sup_{\phi \in S_g} \int \phi \, d\mu$$

By the definition of the integral for general nonnegative functions, this inequality is exactly:

$$\int f \, d\mu \leq \int g \, d\mu$$

Proof of Part 2 (Homogeneity):

We analyze this property by dividing it into two cases based on the value of c.

Case 2a: $c = 0$.

If $c = 0$, the function $cf = 0 \cdot f$ is identically the zero function 0. The only non-negative simple function satisfying $0 \leq \phi \leq 0$ is $\phi = 0$. Thus, its supremum definition yields:

$$\int 0 \cdot f \, d\mu = \sup\{0\} = 0$$

On the other hand, multiplying the scalar constant gives $c \int f \, d\mu = 0 \cdot \int f \, d\mu$. Under standard measure-theoretic conventions, even if $\int f \, d\mu = \infty$, the product evaluates via $0 \cdot \infty = 0$. Thus, the identity holds perfectly when $c = 0$.

Case 2b: $c > 0$.

When c is strictly positive, we establish equality by demonstrating both directions of the inequality.

Step 1: Proving $\int cf \, d\mu \geq c \int f \, d\mu$.

Let ϕ be any simple function satisfying $0 \leq \phi \leq f$. Multiplying this pointwise inequality by the positive real constant $c > 0$ preserves the inequality direction:

$$0 \leq c\phi \leq cf$$

Since ϕ is simple, the scaled function $c\phi$ is also a simple function. Because $c\phi$ is a valid lower simple function for cf, the definition of the general integral as a supremum implies:

$$\int cf \, d\mu \geq \int c\phi \, d\mu$$

By Proposition 2.13a (homogeneity for simple functions), we have $\int c\phi \, d\mu = c \int \phi \, d\mu$. Substituting this yields:

$$\int cf \, d\mu \geq c \int \phi \, d\mu \implies \frac{1}{c} \int cf \, d\mu \geq \int \phi \, d\mu$$

Since this upper bound holds uniformly for every simple function $\phi \leq f$, it must also bound their least upper bound:

$$\frac{1}{c} \int cf \, d\mu \geq \sup_{0 \leq \phi \leq f} \int \phi \, d\mu = \int f \, d\mu$$

Multiplying both sides by c gives our first directional inequality:

$$\int cf \, d\mu \geq c \int f \, d\mu$$

Step 2: Proving $\int cf\,d\mu \le c\int f\,d\mu$.
Let ψ be any simple function satisfying $0 \le \psi \le cf$. Because $c > 0$, we can divide the pointwise inequality by c to obtain:

$$0 \le \frac{1}{c}\psi \le f$$

The function $\frac{1}{c}\psi$ is simple because it is a scalar multiple of the simple function ψ. Therefore, $\frac{1}{c}\psi$ belongs to the family of lower simple functions for f. By definition of the supremum:

$$\int f\,d\mu \ge \int \left(\frac{1}{c}\psi\right) d\mu$$

Applying simple function homogeneity (Proposition 2.13a) shows $\int(\frac{1}{c}\psi)d\mu = \frac{1}{c}\int \psi\,d\mu$:

$$\int f\,d\mu \ge \frac{1}{c}\int \psi\,d\mu \implies c\int f\,d\mu \ge \int \psi\,d\mu$$

Because the constant value $c\int f\,d\mu$ acts as a uniform upper bound for all choices of simple functions $\psi \le cf$, it must be greater than or equal to their supremum:

$$c\int f\,d\mu \ge \sup_{0\le\psi\le cf}\int \psi\,d\mu = \int cf\,d\mu$$

Combining the conclusions of Step 1 and Step 2 establishes that $\int cf = c\int f$ holds perfectly for all parameters. □

Theorem 16 (The Monotone Convergence Theorem)**.** *If $\{f_n\}$ is a sequence in L^+ such that $f_j \le f_{j+1}$ for all j, and $f = \lim_{n\to\infty} f_n$ $(= \sup_n f_n)$, then $\int f = \lim_{n\to\infty}\int f_n$.*

Proof. We break down the proof into clear, digestible steps to ensure maximum conceptual clarity.

Step 1: Establishing the first inequality ($\lim_{n\to\infty}\int f_n \le \int f$)

Because the sequence $\{f_n\}$ is monotonic, we are given that $f_n(x) \le f_{n+1}(x)$ for all n and all $x \in X$. Since the sequence increases pointwise to f, it is clear that for each fixed n, we have $f_n \le f$ everywhere on X.

By the monotonicity property of the integral for general nonnegative functions, $f_n \le f$ implies that:

$$\int f_n \le \int f \quad \text{for all } n$$

Furthermore, since $f_n \le f_{n+1}$, the sequence of real numbers $\{\int f_n\}$ is a non-decreasing sequence in $[0,\infty]$. Every non-decreasing sequence in the extended real numbers has a well-defined limit (which may be ∞). Taking the limit on both sides preserves the inequality, giving us:

$$\lim_{n\to\infty}\int f_n \le \int f$$

Step 2: Setting up the reverse direction with a fractional parameter α

To establish the reverse inequality $\lim_{n\to\infty}\int f_n \ge \int f$, it is sufficient to show that $\lim_{n\to\infty}\int f_n \ge \int \phi$ for any arbitrary simple function ϕ that sits entirely below f $(0 \le \phi \le f)$. Once that is shown, taking the supremum over all such simple functions ϕ will yield $\int f$ by definition.

However, because f_n converges to f only in the limit, $f_n(x)$ might be strictly smaller than $\phi(x)$ for many values of n. To create a "safety margin" that allows f_n to eventually overshoot the simple function benchmark, we introduce a strict scaling fraction $\alpha \in (0,1)$.

Let ϕ be a simple function such that $0 \le \phi \le f$. For each $n \in \mathbb{N}$, we construct the set E_n containing all points where f_n has successfully grown larger than this scaled benchmark $\alpha\phi$:

$$E_n = \{x \in X : f_n(x) \ge \alpha\phi(x)\}$$

Because f_n and ϕ are measurable functions, each E_n is a measurable set ($E_n \in \mathcal{M}$).

Step 3: Analyzing the sets E_n

We observe two crucial set-theoretic properties of the sequence $\{E_n\}$:

1. **The sequence is increasing ($E_n \subseteq E_{n+1}$):** If a point x satisfies $f_n(x) \geq \alpha\phi(x)$, then because $f_{n+1}(x) \geq f_n(x)$, it must also satisfy $f_{n+1}(x) \geq \alpha\phi(x)$. Thus, $x \in E_{n+1}$.
2. **The sets cover the entire space ($\bigcup_{n=1}^{\infty} E_n = X$):** Let $x \in X$.
   - If $\phi(x) = 0$, then $\alpha\phi(x) = 0$. Since $f_n \in L^+$, $f_1(x) \geq 0 = \alpha\phi(x)$, meaning $x \in E_1$.
   - If $\phi(x) > 0$, then because $\alpha < 1$, we have $\alpha\phi(x) < \phi(x) \leq f(x)$. Since $\lim_{n\to\infty} f_n(x) = f(x)$, the sequence of numbers $f_n(x)$ must eventually strictly exceed the value $\alpha\phi(x)$. Thus, there exists some index n such that $f_n(x) \geq \alpha\phi(x)$, meaning $x \in E_n$.

Therefore, $\{E_n\}$ is an increasing sequence of measurable sets whose union is X.

Step 4: Estimating the integral bounded over E_n
For any measurable subset, integrating over the whole space is always greater than or equal to integrating over a localized patch because the function is entirely non-negative ($f_n \geq 0$). Using this and the defining condition of E_n, we can write a chain of inequalities:

$$\int f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq \int_{E_n} \alpha\phi \, d\mu = \alpha \int_{E_n} \phi \, d\mu$$

Summarizing this relationship, we have:

$$\int f_n \, d\mu \geq \alpha \int_{E_n} \phi \, d\mu$$

Step 5: Passing to the limit using continuity of measure
We now take the limit as $n \to \infty$ on both sides of our inequality:

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \lim_{n\to\infty} \alpha \int_{E_n} \phi \, d\mu = \alpha \lim_{n\to\infty} \int_{E_n} \phi \, d\mu$$

Recall that the assignment $A \mapsto \int_A \phi \, d\mu$ is a valid measure on $\mathcal{M}$. For any valid measure, the measure of a countable union of increasing sets is equal to the limit of the measures of the individual sets. Applying this basic property of measures directly to our set sequence $\{E_n\}$ yields:

$$\lim_{n\to\infty} \int_{E_n} \phi \, d\mu = \int_{\bigcup E_n} \phi \, d\mu = \int_X \phi \, d\mu = \int \phi \, d\mu$$

Substituting this limit back into our main line of estimation gives:

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \alpha \int \phi \, d\mu$$

Step 6: Eliminating the parameter α and taking the supremum
Why is this inequality sufficient? Notice that the left-hand side, $\lim_{n\to\infty} \int f_n \, d\mu$, is a fixed real number that has no dependence on our choice of α. Because the inequality holds true for any arbitrary fraction α strictly less than 1, it must remain true in the limit as $\alpha \to 1$ from below. Letting $\alpha \to 1$ yields:

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \int \phi \, d\mu$$

This inequality holds for *every* simple function ϕ that satisfies $0 \leq \phi \leq f$. Since the left side is a uniform upper bound for all such values, it must be greater than or equal to their supremum:

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \sup_{0\leq\phi\leq f} \int \phi \, d\mu$$

By definition, this supremum is exactly the integral of f:

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \int f \, d\mu$$

Combining Step 1 and Step 6, we conclude that $\int f = \lim_{n\to\infty} \int f_n$. $\square$

Remark. *Let us analyze why introducing the scaling factor $\alpha < 1$ is absolutely necessary for the set construction step in this proof.*

If we omitted α (which is equivalent to setting $\alpha = 1$), our sets would look like:

$$E_n = \{x \in X : f_n(x) \geq \phi(x)\}$$

Without α, it is no longer guaranteed that $\bigcup_{n=1}^{\infty} E_n = X$.

Consider a space consisting of a single point $X = \{x\}$, where the background measure is just a point mass. Let $f(x) = c$ for some positive constant $c > 0$. Let us define our increasing sequence of functions to be:

$$f_n(x) = c - \frac{1}{n}$$

It is clear that $f_n(x)$ increases pointwise to $f(x) = c$.

Now, let us choose our lower simple function ϕ to be the constant function $\phi(x) = c$. This is a perfectly valid choice because ϕ is simple and satisfies $0 \leq \phi \leq f$ (since $c \leq c$ is true).

If we try to find the sets E_n for $\alpha = 1$, we check the condition:

$$f_n(x) \geq \phi(x) \iff c - \frac{1}{n} \geq c \iff -\frac{1}{n} \geq 0$$

This statement is false for every single value of $n \in \mathbb{N}$. Therefore, if $\alpha = 1$, every set in our sequence is empty ($E_n = \varnothing$), which means their union is empty ($\bigcup E_n = \varnothing \neq X$). The continuity of measure would yield a limit of 0 instead of $\int \phi$, causing the proof to completely break down.

By introducing $\alpha < 1$, the condition becomes $c - \frac{1}{n} \geq \alpha c$. As n grows large, $-\frac{1}{n}$ vanishes, and since $c > \alpha c$, this inequality is guaranteed to become true for all sufficiently large n. This ensures that the point x eventually enters the sets E_n, validating that $\bigcup E_n = X$.

Remark. *The Monotone Convergence Theorem serves as an essential, practical tool for calculating integrals of general nonnegative functions.*

By its foundational definition, computing $\int f\, d\mu$ requires finding the supremum over a massive, typically uncountable family of lower simple functions:

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

Evaluating such a supremum directly from this collection is often exceptionally difficult, if not practically impossible.

The Monotone Convergence Theorem completely bypasses this obstacle. It guarantees that to calculate the integral of f, we do not need to look at every possible lower simple function. Instead, it is entirely sufficient to construct just a single sequence $\{\phi_n\}$ of simple functions that increases pointwise to f, and then compute a straightforward limit:

$$\int f\, d\mu = \lim_{n \to \infty} \int \phi_n\, d\mu$$

Since we are already guaranteed that such increasing sequences of simple functions always exist for any measurable function in L^+, this result transforms an abstract supremum over an unmanageable set into a concrete, sequential limiting process. This immediate significance provides the groundwork for establishing key architectural properties of integration, such as additivity.

Theorem 17. *If $\{f_n\}$ is a finite or infinite sequence in L^+ and $f = \sum_n f_n$, then $\int f = \sum_n \int f_n$.*

Proof. We establish the theorem by first proving the result for two functions, extending it to any finite sum via induction, and finally taking the limit to reach infinite sums.

Part 1: The Case for Two Functions (f_1 and f_2)

Let $f_1, f_2 \in L^+$. By the standard construction theorem for nonnegative measurable functions, we can find two sequences $\{\phi_j\}$ and $\{\psi_j\}$ of nonnegative simple functions that increase pointwise to f_1 and f_2 respectively:

$$\phi_j \to f_1 \quad \text{as } j \to \infty, \quad \text{with } \phi_j \leq \phi_{j+1}$$

$$\psi_j \to f_2 \quad \text{as } j \to \infty, \quad \text{with } \psi_j \leq \psi_{j+1}$$

Now, consider the combined sequence formed by adding these simple functions together: $\{\phi_j + \psi_j\}$. Because both underlying sequences are non-decreasing, their sum is also a non-decreasing sequence of nonnegative simple functions:

$$(\phi_j + \psi_j) \leq (\phi_{j+1} + \psi_{j+1})$$

Furthermore, by the standard algebraic properties of limits, this combined sequence increases pointwise to the sum of the two target functions:

$$\lim_{j\to\infty}(\phi_j + \psi_j) = \lim_{j\to\infty}\phi_j + \lim_{j\to\infty}\psi_j = f_1 + f_2$$

Since $\{\phi_j + \psi_j\}$ is a sequence of functions in L^+ that increases monotonically to $f_1 + f_2$, we can apply the Monotone Convergence Theorem to swap the integral and the limit:

$$\int (f_1 + f_2) = \lim_{j\to\infty}\int(\phi_j + \psi_j)$$

For simple functions, we already know that the integral is completely additive. Therefore, we can break apart the term inside the limit:

$$\lim_{j\to\infty}\int(\phi_j + \psi_j) = \lim_{j\to\infty}\left(\int \phi_j + \int \psi_j\right)$$

Because both $\{\int \phi_j\}$ and $\{\int \psi_j\}$ are non-decreasing sequences of real numbers, their individual limits exist in $[0, \infty]$. We can distribute the limit over the addition:

$$\lim_{j\to\infty}\left(\int \phi_j + \int \psi_j\right) = \lim_{j\to\infty}\int \phi_j + \lim_{j\to\infty}\int \psi_j$$

Finally, applying the Monotone Convergence Theorem separately back to $\phi_j \to f_1$ and $\psi_j \to f_2$, we see that these two limits are exactly the definitions of $\int f_1$ and $\int f_2$. This completes our chain of equalities:

$$\int (f_1 + f_2) = \int f_1 + \int f_2$$

Part 2: Extension to Any Finite Sum via Mathematical Induction
We now extend this property to a sum of any finite number of functions N.

- **Base Case:** For $N = 1$, $\int f_1 = \int f_1$ is trivially true. For $N = 2$, the statement was verified in Part 1.
- **Inductive Step:** Assume that the linearity holds for a finite collection of size N, meaning $\int \sum_{n=1}^{N} f_n = \sum_{n=1}^{N}\int f_n$. We examine a collection of size $N + 1$:

$$\int \sum_{n=1}^{N+1} f_n = \int \left(\left(\sum_{n=1}^{N} f_n\right) + f_{N+1}\right)$$

Since $\sum_{n=1}^{N} f_n$ is a measurable function in L^+, we can treat it as a single unit and apply our two-function additivity rule from Part 1:

$$\int \left(\left(\sum_{n=1}^{N} f_n\right) + f_{N+1}\right) = \int \left(\sum_{n=1}^{N} f_n\right) + \int f_{N+1}$$

Now, applying our induction hypothesis to the first term allows us to break it apart into a sum of individual integrals:

$$\int \left(\sum_{n=1}^{N} f_n\right) + \int f_{N+1} = \sum_{n=1}^{N}\int f_n + \int f_{N+1} = \sum_{n=1}^{N+1}\int f_n$$

Thus, by induction, the theorem holds for any finite N:

$$\int \sum_{n=1}^{N} f_n = \sum_{n=1}^{N} \int f_n$$

Part 3: The Case for Infinite Sums ($N \to \infty$)
Now consider the infinite series case where $f = \sum_{n=1}^{\infty} f_n$. We can define a sequence of partial sums $\{g_N\}$ by:

$$g_N = \sum_{n=1}^{N} f_n$$

Let us verify the conditions of the partial sums sequence $\{g_N\}$:

1. **It belongs to L^+:** Since each $f_n \geq 0$, every finite sum g_N is non-negative and measurable.
2. **It is monotonically increasing:** Since we add a non-negative term $f_{N+1} \geq 0$ at each step, we have:
$$g_{N+1} = g_N + f_{N+1} \geq g_N$$
3. **It converges pointwise to f:** By definition of an infinite series, the sequence of partial sums converges to the total sum:
$$\lim_{N\to\infty} g_N = \sum_{n=1}^{\infty} f_n = f$$

Since $\{g_N\}$ is a sequence in L^+ that increases monotonically to f, we can safely apply the Monotone Convergence Theorem to state that:

$$\int f = \lim_{N\to\infty} \int g_N$$

Substituting our definition of g_N into this equation gives:

$$\int f = \lim_{N\to\infty} \int \left(\sum_{n=1}^{N} f_n \right)$$

Using our finite summation result proved via induction in Part 2, we can rewrite the integral of the finite sum as the finite sum of the integrals:

$$\lim_{N\to\infty} \int \left(\sum_{n=1}^{N} f_n \right) = \lim_{N\to\infty} \sum_{n=1}^{N} \int f_n$$

By standard definition, the limit of finite partial sums of numbers is exactly the definition of an infinite series of numbers:

$$\lim_{N\to\infty} \sum_{n=1}^{N} \int f_n = \sum_{n=1}^{\infty} \int f_n$$

This yields our final targeted identity for infinite series collections:

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n$$

□

Proposition. *If $f \in L^+$, then $\int f = 0$ iff $f = 0$ a.e.*

Proof. The statement is an "if and only if" (iff) condition, so we split our verification into two logical directions. We first prove the result for simple functions, and then apply it to general functions in L^+.

Part 1: The Case for Simple Functions
Let ϕ be a non-negative simple function with its standard canonical representation given by:

$$\phi = \sum_{j=1}^{n} a_j \chi_{E_j}$$

where the sets E_j are pairwise disjoint and the coefficients $a_j \geq 0$ are non-negative distinct values. By definition, the integral of this simple function is the finite sum:

$$\int \phi \, d\mu = \sum_{j=1}^{n} a_j \mu(E_j)$$

Since each individual term $a_j \mu(E_j)$ in this summation is non-negative, the total sum can equal exactly zero if and only if every single term in the sum is individually zero:

$$\sum_{j=1}^{n} a_j \mu(E_j) = 0 \iff a_j \mu(E_j) = 0 \quad \text{for every } j \in \{1, \dots, n\}$$

For a product of two non-negative extended real numbers to be zero ($a_j \mu(E_j) = 0$), at least one of the factors must be zero. Therefore, for each index j, either the coefficient value $a_j = 0$ or the localized patch has measure $\mu(E_j) = 0$.

This condition means that $\phi(x)$ can only take non-zero values on sets whose measure is zero. Thus, the function is equal to zero almost everywhere ($\phi = 0$ a.e.). This fully establishes the proposition for simple functions.

Part 2: General Case — Forward Direction ($\implies$)

Now, let $f \in L^+$ be a general non-negative measurable function. For this half, we assume that $f = 0$ a.e., and we want to deduce that $\int f = 0$.

Recall that the general integral is defined as the supremum over all lower simple functions:

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

Let ϕ be any simple function satisfying $0 \leq \phi \leq f$. Because $0 \leq \phi(x) \leq f(x)$ holds pointwise, and we are given that $f(x) = 0$ for almost every x, it forces $\phi(x) = 0$ for almost every x.

Since ϕ is a simple function that equals 0 a.e., we can apply our conclusion from Part 1 to state that its integral must be zero ($\int \phi = 0$). Because *every* valid lower simple function ϕ has an integral of exactly 0, the collection of numbers we are taking the supremum over contains nothing but the value 0. Therefore, their least upper bound must be zero:

$$\int f = \sup_{\phi \leq f} \int \phi = \sup\{0\} = 0$$

Part 3: General Case — Backward Direction ($\impliedby$)

For the final half, we assume that $\int f = 0$, and we must prove that $f = 0$ a.e.. We show this by proving the contrapositive: if it is false that $f = 0$ a.e., then it must be true that $\int f > 0$.

Suppose $f = 0$ a.e. is false. This means the set of points where $f(x)$ is strictly positive has a strictly positive measure:

$$\mu(\{x \in X : f(x) > 0\}) > 0$$

We can decompose this set of positive points into a countable union of nested tracking sets, categorized by how far away from zero the function values are. For each $n \in \mathbb{N}$, define:

$$E_n = \left\{ x \in X : f(x) > \frac{1}{n} \right\}$$

If a point has a function value $f(x) > 0$, there must exist some sufficiently large integer n such that $f(x) > \frac{1}{n}$. Thus, we can write the positive set as a countable union:

$$\{x \in X : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$$

By the countable subadditivity property of measures, if a countable union of sets has a strictly positive total measure, it is impossible for all individual component sets to have a measure of zero. Therefore, there must exist at least one specific index n such that:

$$\mu(E_n) > 0$$

Let us fix this specific set E_n. For any point $x \in E_n$, we know that $f(x) > \frac{1}{n}$ by definition. For any point $x \notin E_n$, since f is non-negative, $f(x) \geq 0$. We can compactly construct a simple function that acts as a lower benchmark across the whole space:

$$f \geq \frac{1}{n}\chi_{E_n}$$

Because $\phi = \frac{1}{n}\chi_{E_n}$ is a valid non-negative simple function lying entirely below f, the supremum definition of the general integral ensures that the integral of f must be greater than or equal to the integral of this simple benchmark:

$$\int f\, d\mu \geq \int \left(\frac{1}{n}\chi_{E_n}\right) d\mu = \frac{1}{n}\mu(E_n)$$

Since we established that $\mu(E_n) > 0$, multiplying it by the positive fraction $\frac{1}{n}$ yields a strictly positive value:

$$\int f\, d\mu \geq \frac{1}{n}\mu(E_n) > 0$$

Thus, if f is not zero almost everywhere, its integral must be strictly greater than zero. By contraposition, if $\int f = 0$, it is a mathematical certainty that $f = 0$ a.e. □

Corollary 18. *If $\{f_n\} \subset L^+$, $f \in L^+$, and $f_n(x)$ increases to $f(x)$ for a.e. x, then $\int f = \lim \int f_n$.*

Proof. We break down the proof into clear steps to see how we can reduce this "almost everywhere" statement to the standard Monotone Convergence Theorem.

Step 1: Setting up the exceptional set of measure zero

We are given that $f_n(x)$ increases to $f(x)$ for almost every $x \in X$. By definition, this means there exists a set $E \in \mathcal{M}$ where the convergence holds everywhere on E, and the complement set E^c (where convergence might fail) has a measure of zero:

$$\mu(E^c) = 0$$

Step 2: Modifying our functions using characteristic functions

To ignore the troublesome points in E^c, we multiply our functions by the indicator function χ_E. Let's examine the new functions $f\chi_E$ and $f_n\chi_E$:

- For any point $x \in E$, we have $\chi_E(x) = 1$. Therefore, $f_n(x)\chi_E(x) = f_n(x)$ and $f(x)\chi_E(x) = f(x)$. Since we know $f_n(x) \to f(x)$ monotonically on E, the modified sequence $\{f_n\chi_E\}$ increases pointwise to $f\chi_E$ on E.
- For any point $x \in E^c$, we have $\chi_E(x) = 0$. Therefore, $f_n(x)\chi_E(x) = 0$ and $f(x)\chi_E(x) = 0$. On this set, the sequence is constantly 0 and converges trivially to 0.

Because it increases on both pieces, the sequence of functions $\{f_n\chi_E\}$ is a sequence in L^+ that increases **everywhere** on X to the function $f\chi_E$.

Step 3: Applying the Monotone Convergence Theorem

Since $\{f_n\chi_E\}$ is monotonic and converges pointwise at every single point in X, it satisfies the strict conditions of the standard Monotone Convergence Theorem. This allows us to exchange the limit and the integral safely:

$$\int f\chi_E = \lim_{n\to\infty} \int f_n\chi_E$$

Step 4: Relating modified functions back to the originals via measure zero sets

Now we look at the difference between our original functions and the modified ones:

- The function $f - f\chi_E$ is equal to 0 on E and equal to f on E^c. Since $\mu(E^c) = 0$, this function is non-zero only on a set of measure zero. Thus, $f - f\chi_E = 0$ a.e..
- Similarly, for each fixed n, the function $f_n - f_n\chi_E$ is equal to 0 on E and equal to f_n on E^c. Thus, $f_n - f_n\chi_E = 0$ a.e..

Recall the general measure theory principle: *if two non-negative functions are equal almost everywhere, their integrals are exactly equal.* Applying this fact gives us two crucial identities:

$$\int f = \int f\chi_E$$

$$\int f_n = \int f_n\chi_E \quad \text{for all } n$$

Step 5: Final string of equalities
Combining the identities from Step 4 with our limit result from Step 3, we obtain:

$$\int f = \int f\chi_E = \lim_{n\to\infty} \int f_n\chi_E = \lim_{n\to\infty} \int f_n$$

Therefore, $\int f = \lim \int f_n$. □

Remark. *The structural requirement that the sequence of functions $\{f_n\}$ must be monotonically increasing (at least almost everywhere) is an absolute necessity for the validity of the Monotone Convergence Theorem. If we drop this condition, the limit of the integrals may fail to equal the integral of the pointwise limit.*

To see exactly why this breaks down, let our space be the real numbers $X = \mathbb{R}$ equipped with the standard Lebesgue measure μ. We analyze two classic counterexamples where the sequence converges pointwise but fails to be monotonic:

1. ***The Moving Block Example (Horizontal Escape):*** *Consider the sequence of moving characteristic indicator functions:*

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

For any fixed point $x \in \mathbb{R}$, as n grows larger than x, the interval shifts completely past that point, meaning $f_n(x) = 0$ for all sufficiently large n. Thus, the sequence converges pointwise to zero everywhere:

$$f(x) = \lim_{n\to\infty} f_n(x) = 0 \implies \int f\, d\mu = 0$$

However, calculating the area under each individual function yields a constant value of 1 for every index n:

$$\int f_n\, d\mu = \mu\big((n, n+1)\big) = 1 \implies \lim_{n\to\infty} \int f_n\, d\mu = 1$$

Here, $1 \neq 0$, showing the theorem fails because the sequence is not increasing.

2. ***The Shrinking Tall Spike Example (Vertical Escape):*** *Consider a sequence of shrinking but tall spikes focused near the origin:*

$$g_n = n\chi_{(0,1/n)}$$

If we pick $x \leq 0$, then $g_n(x) = 0$ for all n. If we pick any positive point $x > 0$, by the Archimedean property, we can find a large enough index n such that $\frac{1}{n} < x$, dropping the function value down to 0 permanently. Hence, this sequence also converges pointwise to zero everywhere:

$$g(x) = \lim_{n\to\infty} g_n(x) = 0 \implies \int g\, d\mu = 0$$

Yet, the area of each individual spike remains perfectly constant due to the scaling cancellation between width and height:

$$\int g_n\, d\mu = n \cdot \mu\big((0, 1/n)\big) = n \cdot \frac{1}{n} = 1 \implies \lim_{n\to\infty} \int g_n\, d\mu = 1$$

Once again, $1 \neq 0$, showing a complete mismatch.

The geometric intuition underlying both failures becomes clear when looking at their graphs: the total area under the curve "escapes to infinity" as $n \to \infty$. In the first example, the area escapes horizontally along the real line; in the second example, it escapes vertically by squeezing into an infinitely thin, infinitely tall spike. Because this mass disappears into the limit of the domain rather than building up locally, the area in the limit is strictly less than the limit of the individual areas. The monotonicity constraint prevents this escape by forcing the functions to grow everywhere, pinning the mass down from below.

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

$$\int (\liminf_{n\to\infty} f_n) \leq \liminf_{n\to\infty} \int f_n$$

Proof. We break down the proof into three main structural stages: the construction of a monotonic sequence of functions, the derivation of a pointwise integral bound, and the application of the Monotone Convergence Theorem(MCT).

Step 1: Constructing the Increasing Sequence $\{g_k\}$

Recall the mathematical definition of the limit inferior ($\liminf$) of a sequence of functions. It is defined by looking at the infimum of the remaining "tails" of the sequence, and then taking the limit as we look further down those tails:

$$\liminf_{n\to\infty} f_n(x) = \lim_{k\to\infty} \left(\inf_{n\geq k} f_n(x) \right)$$

Let us isolate this inner tail-infimum term and define it as a new function, $g_k(x)$:

$$g_k(x) = \inf_{n\geq k} f_n(x)$$

Let's carefully analyze the behavior of this sequence of functions $\{g_k\}$:

- **Nonnegativity ($g_k \in L^+$):** Since every $f_n \geq 0$ for all n, their lower bounds must also be nonnegative ($g_k \geq 0$).
- **Monotonicity ($g_k \leq g_{k+1}$):** Notice what happens when we move from index k to $k+1$. For g_k, we take the infimum over the set $\{f_k, f_{k+1}, f_{k+2}, \dots\}$. For g_{k+1}, we drop the first element f_k and take the infimum over a smaller pool of options: $\{f_{k+1}, f_{k+2}, \dots\}$. Squeezing an infimum over a smaller collection of elements can only make the value grow larger or stay the same. Thus, $\{g_k\}$ forms a **monotonically increasing sequence**:

$$g_1(x) \leq g_2(x) \leq g_3(x) \leq \cdots \leq g_k(x) \leq g_{k+1}(x) \leq \dots$$

- **Pointwise Convergence:** By definition, as $k \to \infty$, this increasing sequence converges pointwise exactly to our target function:

$$\lim_{k\to\infty} g_k(x) = \liminf_{n\to\infty} f_n(x)$$

Step 2: Establishing a Finite Lower Bound via Monotonicity of the Integral

By our definition of g_k, it serves as a pointwise lower bound for every individual function f_j sitting inside its tail collection ($j \geq k$):

$$g_k(x) \leq f_j(x) \quad \text{for all } j \geq k$$

Because integration preserves pointwise inequalities, we can take the integral of both sides across this structural threshold:

$$\int g_k \leq \int f_j \quad \text{for all } j \geq k$$

Since $\int g_k$ is a single fixed real number that is smaller than or equal to every number in the numerical tail set $\left\{ \int f_k, \int f_{k+1}, \int f_{k+2}, \dots \right\}$, it must be less than or equal to their infimum:

$$\int g_k \leq \inf_{j\geq k} \int f_j$$

Now, let us take the limit of both sides of this numerical inequality as $k \to \infty$:

$$\lim_{k\to\infty} \int g_k \leq \lim_{k\to\infty} \left(\inf_{j\geq k} \int f_j \right)$$

By definition, the limit on the right side is exactly the limit inferior of our sequence of integrals:

$$\lim_{k\to\infty} \int g_k \leq \liminf_{n\to\infty} \int f_n$$

Step 3: Where and How the Monotone Convergence Theorem (MCT) is Applied

Look at the left side of our inequality: $\lim_{k\to\infty} \int g_k$. ¿ **Student Crucial Alert:** Because $\{g_k\}$ is a sequence of nonnegative functions that is **monotonically increasing** ($g_k \leq g_{k+1}$), it satisfies the requirements of the Monotone Convergence Theorem! The MCT guarantees that for an increasing sequence, we can swap the limit outside the integral to the inside:

$$\lim_{k\to\infty} \int g_k = \int \left(\lim_{k\to\infty} g_k \right)$$

Since we already established in Step 1 that $\lim_{k\to\infty} g_k = \liminf_{n\to\infty} f_n$, we can substitute this inside the integral:

$$\lim_{k\to\infty} \int g_k = \int \left(\liminf_{n\to\infty} f_n \right)$$

Finally, replacing the left side of our inequality from Step 2 with this MCT identity yields our final result:

$$\int (\liminf_{n\to\infty} f_n) \leq \liminf_{n\to\infty} \int f_n$$

□

Corollary 19. *If* $\{f_n\} \subset L^+$, $f \in L^+$, *and* $f_n \to f$ *a.e., then* $\int f \leq \liminf \int f_n$.

Proof. We break down the verification into clear, structured steps to demonstrate how we can handle the "almost everywhere" condition and apply Fatou's Lemma seamlessly.

Step 1: Isolate the Exceptional Null Set

We are given that $f_n \to f$ almost everywhere (*a.e.*). By definition, this means there exists a measurable set $E \in \mathcal{M}$ where pointwise convergence holds for every single element, and its complement set E^c has a measure of zero:

$$\mu(E^c) = 0$$

Step 2: Construct Globally Well-Behaved Functions

To extend the convergence behavior to the entire space without altering our integrals, we mask the behavior on E^c by multiplying our functions by the indicator function χ_E:

$$\tilde{f}_n = f_n \chi_E \quad \text{and} \quad \tilde{f} = f \chi_E$$

Let us verify how this modified sequence behaves pointwise across the whole domain X:

- For any point $x \in E$, we have $\chi_E(x) = 1$. Thus, $\tilde{f}_n(x) = f_n(x)$ and $\tilde{f}(x) = f(x)$. Since we know $f_n(x) \to f(x)$ on E, it follows that $\tilde{f}_n(x) \to \tilde{f}(x)$.
- For any point $x \in E^c$, we have $\chi_E(x) = 0$. Thus, $\tilde{f}_n(x) = 0$ and $\tilde{f}(x) = 0$. The sequence becomes a constant sequence of zeros, which trivially converges to 0.

Consequently, the modified sequence $\{\tilde{f}_n\}$ converges pointwise to $\tilde{f}$ **everywhere** on X.

Step 3: Preserve Integrals via Almost Everywhere Equivalence

Because the original functions and our newly constructed functions differ only on the set E^c, and $\mu(E^c) = 0$, they are equal almost everywhere:

$$\tilde{f}_n = f_n \text{ a.e.} \quad \text{and} \quad \tilde{f} = f \text{ a.e.}$$

Recall Proposition 2.16, which states that functions that are equal almost everywhere yield identical integrals. This gives us two vital identities:

$$\int \tilde{f}_n = \int f_n \quad \text{for all } n$$

$$\int \tilde{f} = \int f$$

Step 4: Apply Fatou's Lemma to the Everywhere-Convergent Sequence

Since the sequence $\{\tilde{f}_n\}$ converges pointwise everywhere on X to $\tilde{f}$, its limit inferior is simply equal to the pointwise limit function itself:

$$\liminf_{n\to\infty} \tilde{f}_n = \tilde{f}$$

We now apply the standard version of Fatou's Lemma (Theorem 2.18) directly to this well-behaved sequence:

$$\int \left(\liminf_{n\to\infty} \tilde{f}_n\right) \leq \liminf_{n\to\infty} \int \tilde{f}_n$$

Substituting $\liminf_{n\to\infty} \tilde{f}_n = \tilde{f}$ into the left-hand side transforms the inequality into:

$$\int \tilde{f} \leq \liminf_{n\to\infty} \int \tilde{f}_n$$

Step 5: Final Substitution

Finally, we substitute our preservation identities from Step 3 ($\int \tilde{f} = \int f$ and $\int \tilde{f}_n = \int f_n$) into our inequality to return to our original functions:

$$\int f \leq \liminf_{n\to\infty} \int f_n$$

This completes the proof. □

Proposition. *If $f \in L^+$ and $\int f < \infty$, then the set $\{x : f(x) = \infty\}$ is a null set and the set $\{x : f(x) > 0\}$ is σ-finite.*

Proof. This proposition consists of two separate assertions. We will prove each one step-by-step with maximum clarity.

Part 1: Proving $\{x : f(x) = \infty\}$ is a Null Set

Let us define E to be the set of points where the function evaluates to infinity:

$$E = \{x \in X : f(x) = \infty\}$$

We want to prove that $\mu(E) = 0$.

1. Consider any positive constant scalar $c > 0$.
2. By the definition of our set E, for every point $x \in E$, the value $f(x)$ is infinitely larger than c, so $f(x) \geq c$.
3. For any point $x \notin E$, since $f \in L^+$, we trivially have $f(x) \geq 0$.
4. Therefore, we can bound f from below across the entire space using a simple indicator function scaled by c:
$$f \geq c\chi_E$$
5. Since integration preserves pointwise inequalities, taking the integral on both sides gives:
$$\int f \, d\mu \geq \int c\chi_E \, d\mu = c\mu(E)$$

6. We can isolate the measure term by dividing both sides by our positive constant c:

$$\mu(E) \leq \frac{1}{c} \int f \, d\mu$$

Since this structural inequality remains valid for any choice of $c > 0$, we can examine the behavior as $c \to \infty$:

$$\lim_{c \to \infty} \left(\frac{1}{c} \int f \, d\mu \right)$$

We are given by hypothesis that the total mass of the integral is strictly finite ($\int f < \infty$). Dividing a fixed, finite number by a value c that grows arbitrarily large forces the fraction down to zero:

$$\mu(E) \leq 0$$

Because a measure is non-negative by definition ($\mu(E) \geq 0$), it is trapped. Hence, $\mu(E) = 0$, proving that $\{x : f(x) = \infty\}$ is a null set.

Part 2: Proving $\{x : f(x) > 0\}$ is σ-Finite
Let us denote the support set where our function is strictly positive as A:

$$A = \{x \in X : f(x) > 0\}$$

To show that A is σ-finite, we must express it as a countable union of measurable sets, where every individual component set has a strictly finite measure.

1. Let us slice the set A into a countable sequence of sets $\{A_n\}_{n=1}^{\infty}$ based on how far away from zero the function values are:
$$A_n = \left\{ x \in X : f(x) > \frac{1}{n} \right\}$$
2. If any point x has a strictly positive value $f(x) > 0$, by the Archimedean property, there must exist a sufficiently large integer $n \in \mathbb{N}$ such that $f(x) > \frac{1}{n}$.
3. This implies that every point in A belongs to at least one of the component sets A_n. Thus, we can write A as a countable union:
$$A = \bigcup_{n=1}^{\infty} A_n$$

Now, we must verify that each individual set A_n has a finite measure ($\mu(A_n) < \infty$):

- For any point $x \in A_n$, we have $f(x) > \frac{1}{n}$ by definition.
- For any point $x \notin A_n$, we have $f(x) \geq 0$ since $f \in L^+$.
- Therefore, we can establish a lower benchmark function across the entire domain:
$$f \geq \frac{1}{n} \chi_{A_n}$$
- Integrating both sides of this inequality yields:
$$\int f \, d\mu \geq \int \frac{1}{n} \chi_{A_n} \, d\mu = \frac{1}{n} \mu(A_n)$$
- Multiplying both sides by the integer n, we isolate our measure:
$$\mu(A_n) \leq n \int f \, d\mu$$

Since n is a finite integer, and $\int f \, d\mu$ is a finite value by our hypothesis, their product must be strictly finite:

$$\mu(A_n) \leq n \int f \, d\mu < \infty$$

Because we have represented $A = \{x : f(x) > 0\}$ as a countable union of sets A_n, each of which satisfies $\mu(A_n) < \infty$, the set is σ-finite by definition. $\square$

Integration of Real-Valued Functions

We extend the definition of the integral from non-negative measurable functions to general real-valued measurable functions by decomposing each function into its positive and negative components. Throughout this section, we work over a fixed measure space $(X, \mathcal{M}, \mu)$.

Definition 39 (Positive and Negative Parts). *Let $f : X \to \mathbb{R}$ be a real-valued function. The* ***positive part*** *f^+ and the* ***negative part*** *f^- of f are defined pointwise for each $x \in X$ by:*

$$f^+(x) = \max\{f(x), 0\}$$

$$f^-(x) = \max\{-f(x), 0\}$$

Properties of the Decomposition: By construction, if f is measurable, then both f^+ and f^- belong to L^+. Furthermore, they satisfy the following algebraic identities for all $x \in X$:

1. $f(x) = f^+(x) - f^-(x)$
2. $|f(x)| = f^+(x) + f^-(x)$

Definition 40 (Integration of Real-Valued Functions). *Let $f : X \to \mathbb{R}$ be a measurable function such that at least one of the integrals $\int f^+ \, d\mu$ or $\int f^- \, d\mu$ is finite. The* ***integral*** *of f with respect to the measure μ is defined by:*

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$$

Definition 41 (Integrable Function). *A real-valued measurable function f is said to be* ***integrable*** *if both $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are strictly finite:*

$$\int f^+ \, d\mu < \infty \quad \textit{and} \quad \int f^- \, d\mu < \infty$$

Exercise 8. *Prove that a real-valued measurable function f is integrable if and only if its absolute value $|f|$ is integrable, meaning:*

$$\int |f| \, d\mu < \infty$$

Solution. We establish the equivalence by verifying both implications directly from the properties of the function decomposition.

Forward Direction ($\Longrightarrow$): Assume that f is integrable. By definition, this implies that $\int f^+ \, d\mu < \infty$ and $\int f^- \, d\mu < \infty$. Utilizing the identity $|f| = f^+ + f^-$ and the additivity of the integral on L^+, we obtain:

$$\int |f| \, d\mu = \int (f^+ + f^-) \, d\mu = \int f^+ \, d\mu + \int f^- \, d\mu$$

Since the sum of two finite real numbers is finite, it follows that $\int |f| \, d\mu < \infty$.

Backward Direction ($\Longleftarrow$): Assume that $\int |f| \, d\mu < \infty$. Pointwise, we have the inequalities $f^+(x) \leq |f(x)|$ and $f^-(x) \leq |f(x)|$ for all $x \in X$. By the monotonicity property of the integral on L^+, we can bound the integrals of the components:

$$\int f^+ \, d\mu \leq \int |f| \, d\mu < \infty$$

$$\int f^- \, d\mu \leq \int |f| \, d\mu < \infty$$

Because both components have finite integrals, f satisfies the definition of an integrable function. $\square$

Proposition. *The set of integrable real-valued functions on X is a real vector space, and the integral is a linear functional on it.*

Proof. To establish this proposition completely, we must verify two distinct structural properties: that the space is closed under scalar multiplication and addition (making it a real vector space), and that the integration operator preserves these operations linearly.

Step 1: Closure Under Scalar Multiplication

Let f be an integrable real-valued function, and let $a \in \mathbb{R}$ be an arbitrary scalar multiplier. To confirm that the scaled function af remains within the space of integrable functions, we verify that its absolute value has a finite integral:

$$\int |af|\, d\mu = \int |a||f|\, d\mu = |a| \int |f|\, d\mu$$

Since f is integrable, $\int |f|\, d\mu < \infty$. Multiplying this finite value by the finite constant $|a|$ yields a finite result, demonstrating that af is indeed integrable.

Furthermore, checking the scalar homogeneity of the integration operator ($\int af = a \int f$) splits into two natural cases based on the sign of a:

- If $a \geq 0$, then $(af)^+ = af^+$ and $(af)^- = af^-$. Applying the definition of real-valued integration and homogeneity on L^+ yields:

$$\int af\, d\mu = \int af^+\, d\mu - \int af^-\, d\mu = a\left(\int f^+\, d\mu - \int f^-\, d\mu\right) = a \int f\, d\mu$$

- If $a < 0$, multiplying by a negative number swaps the positive and negative components: $(af)^+ = -af^-$ and $(af)^- = -af^+$. Noting that $-a > 0$, we have:

$$\int af\, d\mu = \int (-af^-)\, d\mu - \int (-af^+)\, d\mu = -a \int f^-\, d\mu - (-a) \int f^+\, d\mu$$

  Factoring out a from this expression yields:

$$\int af\, d\mu = a \int f^+\, d\mu - a \int f^-\, d\mu = a\left(\int f^+\, d\mu - \int f^-\, d\mu\right) = a \int f\, d\mu$$

Step 2: Closure Under Addition

Let f and g be two integrable real-valued functions, and let $h = f + g$ denote their sum. By the triangle inequality, we can bound the absolute value pointwise:

$$|h| = |f + g| \leq |f| + |g|$$

Integrating both sides and using the additivity property of integrals for non-negative functions shows:

$$\int |h|\, d\mu \leq \int (|f| + |g|)\, d\mu = \int |f|\, d\mu + \int |g|\, d\mu$$

Because f and g are integrable, both terms on the right-hand side are finite. Therefore, $\int |h|\, d\mu < \infty$, confirming that $h = f + g$ is integrable. This establishes that the set of integrable real-valued functions forms a genuine real vector space.

Step 3: Additivity of the Integral

To prove that $\int (f+g)\, d\mu = \int f\, d\mu + \int g\, d\mu$, we cannot simply write $h^+ - h^- = (f^+ - f^-) + (g^+ - g^-)$ and split them directly, as some of these parts might evaluate to $\infty - \infty$ pointwise before integration. Instead, we rearrange the equation algebraically to eliminate all subtraction signs.

1. We begin with the pointwise decomposition identity for $h = f + g$:

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

2. We move all negative terms to the opposite sides of the equality to secure a purely additive pointwise relationship:

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

3. Since every term in this expression is non-negative, we can safely apply the linear additivity theorem for L^+ functions (Theorem 2.15) to integrate both sides:

$$\int h^+ + \int f^- + \int g^- = \int h^- + \int f^+ + \int g^+$$

4. Because f, g, and h are verified to be integrable, every individual integral in this equation is a finite real number. This finiteness permits standard algebraic subtraction. We group the terms belonging to each function together by moving them back:

$$\left(\int h^+ - \int h^-\right) = \left(\int f^+ - \int f^-\right) + \left(\int g^+ - \int g^-\right)$$

5. Applying the standard definition of the real-valued integral to each grouped component yields the final linear result:

$$\int h\, d\mu = \int f\, d\mu + \int g\, d\mu$$

Combining Step 1 and Step 3 shows that the integration operator is a linear functional on this vector space. □

Definition 42 (Integration of Complex-Valued Functions)**.** *Let $f : X \to \mathbb{C}$ be a complex-valued measurable function. We say that f is **integrable** if its absolute value satisfies:*

$$\int |f|\, d\mu < \infty$$

*In this case, the **complex integral** of f over the space X is defined in terms of its real and imaginary parts by:*

$$\int f\, d\mu = \int \operatorname{Re} f\, d\mu + i \int \operatorname{Im} f\, d\mu$$

Definition 43 (Integral over a Measurable Set)**.** *Let $E \in \mathcal{M}$ be a measurable set, and let f be a complex-valued measurable function. We say that f is **integrable on** E if:*

$$\int_E |f|\, d\mu < \infty$$

where $\int_E |f|\, d\mu = \int |f|\chi_E\, d\mu$. Under this condition, the integral of f over the set E is defined as:

$$\int_E f\, d\mu = \int f\chi_E\, d\mu = \int_E \operatorname{Re} f\, d\mu + i \int_E \operatorname{Im} f\, d\mu$$

Remark. *The definition of complex integration relies implicitly on the well-definedness of the individual real integrals $\int \operatorname{Re} f\, d\mu$ and $\int \operatorname{Im} f\, d\mu$. Pointwise, a complex function and its components are related by the following classical modulus inequalities:*

$$|\operatorname{Re} f| \le |f|, \quad |\operatorname{Im} f| \le |f|, \quad \text{and} \quad |f| \le |\operatorname{Re} f| + |\operatorname{Im} f|$$

By the monotonicity property of the integral, these geometric bounds imply that f is integrable ($\int |f|\, d\mu < \infty$) if and only if both its real part $\operatorname{Re} f$ and its imaginary part $\operatorname{Im} f$ are integrable real-valued functions ($\int |\operatorname{Re} f|\, d\mu < \infty$ and $\int |\operatorname{Im} f|\, d\mu < \infty$). This equivalence ensures that the component integrals on the right-hand side of the definition are always finite real numbers, rendering the complex integral fully stable under standard linear combinations.

Proposition. *The space of complex-valued integrable functions is a complex vector space and the integral is a complex-linear functional on it.*

Exercise 9. *Prove that the space of complex-valued integrable functions forms a complex vector space and show that the integration operator is complex-linear.*

Solution. Let $f, g : X \to \mathbb{C}$ be complex-valued integrable functions and let $c = \alpha + i\beta \in \mathbb{C}$ be an arbitrary complex scalar.

1. Vector Space Properties: By properties of the complex modulus and the triangle inequality, we have $|f+g| \le |f| + |g|$ and $|cf| = |c||f|$. Integrating these expressions preserves the inequalities:

$$\int |f+g|\, d\mu \le \int |f|\, d\mu + \int |g|\, d\mu < \infty$$

$$\int |cf|\, d\mu = |c| \int |f|\, d\mu < \infty$$

Since their integrals remain finite, the space is closed under addition and complex scalar multiplication, forming a complex vector space.

2. Complex Linearity: Additivity follows directly from the real-valued case by expanding the real and imaginary parts:

$$\int (f+g)\, d\mu = \int (\operatorname{Re} f + \operatorname{Re} g)\, d\mu + i \int (\operatorname{Im} f + \operatorname{Im} g)\, d\mu$$

$$= \left(\int \operatorname{Re} f\, d\mu + i \int \operatorname{Im} f\, d\mu \right) + \left(\int \operatorname{Re} g\, d\mu + i \int \operatorname{Im} g\, d\mu \right) = \int f\, d\mu + \int g\, d\mu$$

For scalar homogeneity, it suffices to test multiplication by the imaginary unit i since real scalar homogeneity is already established. Note that $if = i(\operatorname{Re} f + i \operatorname{Im} f) = -\operatorname{Im} f + i \operatorname{Re} f$. Applying the definition yields:

$$\int if\, d\mu = \int (-\operatorname{Im} f)\, d\mu + i \int \operatorname{Re} f\, d\mu = -\int \operatorname{Im} f\, d\mu + i \int \operatorname{Re} f\, d\mu$$

$$= i \left(\int \operatorname{Re} f\, d\mu + i \int \operatorname{Im} f\, d\mu \right) = i \int f\, d\mu$$

Combining additivity with complex scalar homogeneity demonstrates that the integral is a complex-linear functional. □

Remark (Notations for Integrable Spaces)**.** *Depending on the level of specificity required by the context, the space of complex-valued integrable functions on a given measure space is represented by various standard notations. It is provisionally denoted by $L^1(\mu)$, or more explicitly by $L^1(X, \mathcal{M}, \mu)$ and $L^1(X)$ to highlight the underlying set and σ-algebra. When the measure space is understood from context, it is simply abbreviated as L^1.*

Proposition. *If $f \in L^1$, then $\left|\int f\, d\mu\right| \le \int |f|\, d\mu$.*

Proof. We break down the verification of this fundamental integral inequality into three logical structural cases: the trivial zero case, the real-valued case, and the general complex-valued case.

Case 1: The Trivial Case

If the total integral of our function evaluates exactly to zero, meaning $\int f\, d\mu = 0$, then the left-hand side of the inequality becomes:

$$\left| \int f\, d\mu \right| = |0| = 0$$

Since $f \in L^1$, its modulus $|f|$ is a non-negative function belonging to L^+, which means its integral is naturally non-negative:

$$\int |f|\, d\mu \ge 0$$

Thus, the inequality $0 \le \int |f|\, d\mu$ holds trivially.

Case 2: The Real-Valued Case

Assume that f is a real-valued integrable function. By definition, we can decompose f into its positive and negative components via $f = f^+ - f^-$, where $|f| = f^+ + f^-$. Applying the definition of real integration and the classical numerical triangle inequality, we find:

$$\left|\int f\,d\mu\right| = \left|\int f^+\,d\mu - \int f^-\,d\mu\right| \leq \left|\int f^+\,d\mu\right| + \left|\int f^-\,d\mu\right|$$

Because f^+ and f^- are non-negative, their integrals are already non-negative real values, making the outer absolute value brackets redundant:

$$\left|\int f^+\,d\mu\right| + \left|\int f^-\,d\mu\right| = \int f^+\,d\mu + \int f^-\,d\mu$$

Using the additive property of non-negative integration to merge the integrals back together establishes the statement for real functions:

$$\int f^+\,d\mu + \int f^-\,d\mu = \int (f^+ + f^-)\,d\mu = \int |f|\,d\mu$$

Case 3: The Complex-Valued Case

Now suppose that f is a general complex-valued integrable function such that $\int f\,d\mu \neq 0$. The integral $\int f\,d\mu$ yields a definite non-zero complex scalar value $z \in \mathbb{C}$.

1. We can express any non-zero complex number in polar form, or extract its phase by taking the complex conjugate of its signum. Let us define a complex scalar $\alpha \in \mathbb{C}$ as:

$$\alpha = \overline{\operatorname{sgn}\left(\int f\,d\mu\right)}$$

2. By this exact phase-rotational setup, multiplying our complex integral value by α rotates it directly onto the positive real axis, isolating its absolute value:

$$\left|\int f\,d\mu\right| = \alpha \int f\,d\mu$$

3. Since integration is a complex-linear functional, we can move the constant scalar multiplier α inside the integral operator smoothly:

$$\alpha \int f\,d\mu = \int \alpha f\,d\mu$$

4. Because the left side of our equation ($\left|\int f\,d\mu\right|$) is purely real, the right side ($\int \alpha f\,d\mu$) must be a purely real number as well. A real number is completely identical to its own real part, meaning:

$$\int \alpha f\,d\mu = \operatorname{Re}\left(\int \alpha f\,d\mu\right)$$

5. By the definition of complex integration, the real part of an integral is precisely the integral of the function's real component:

$$\operatorname{Re}\left(\int \alpha f\,d\mu\right) = \int \operatorname{Re}(\alpha f)\,d\mu$$

6. Pointwise everywhere on the domain, the real component of any complex number is bounded above by its full geometric absolute modulus: $\operatorname{Re}(\alpha f(x)) \leq |\alpha f(x)|$. Due to the basic monotonicity of the integral, this gives:

$$\int \operatorname{Re}(\alpha f)\,d\mu \leq \int |\alpha f|\,d\mu$$

7. Since the signum phase factor has a unit modulus ($|\alpha| = 1$), the absolute value splits multiplicatively into $|\alpha f| = |\alpha||f| = |f|$. Substituting this back into our expression provides our final boundary:

$$\int |\alpha f|\,d\mu = \int |f|\,d\mu$$

Collating this chain of modifications link by link confirms the validity of the inequality for all complex scenarios:

$$\left|\int f\,d\mu\right| = \int \operatorname{Re}(\alpha f)\,d\mu \leq \int |f|\,d\mu$$

□

Proposition. *Let $f, g \in L^1$.*

(a) *If $f \in L^1$, then the set $\{x : f(x) \neq 0\}$ is σ-finite.*

(b) *$\int_E f\, d\mu = \int_E g\, d\mu$ for all $E \in \mathcal{M}$ if and only if $\int |f - g|\, d\mu = 0$, which holds if and only if $f = g$ a.e.*

Proof. We address each part of the proposition independently, breaking down the bidirectional equivalences into explicit, pedagogical steps.

Proof of Part (a)

We wish to show that the support set $S = \{x \in X : f(x) \neq 0\}$ is σ-finite. 1. Since $f \in L^1$ is a complex-valued function, its absolute value $|f|$ is a non-negative measurable function satisfying $\int |f|\, d\mu < \infty$. 2. Pointwise, the complex modulus $|f(x)|$ is non-zero if and only if $f(x)$ itself is non-zero. Thus, we have the set identity:

$$\{x \in X : f(x) \neq 0\} = \{x \in X : |f(x)| > 0\}$$

3. Since $|f| \in L^+$ and $\int |f|\, d\mu < \infty$, we directly invoke Proposition 2.20, which states that the support of a non-negative function with a finite integral is σ-finite. Therefore, the set $\{x : f(x) \neq 0\}$ is σ-finite.

Proof of Part (b)

This assertion contains a three-way logical equivalence. We establish it by proving two separate equivalence links: first, that $\int |f - g|\, d\mu = 0 \iff f = g$ a.e., and second, that $\int |f - g|\, d\mu = 0 \iff \int_E f = \int_E g$ for all $E \in \mathcal{M}$.

Link 1: Verifying $\int |f - g|\, d\mu = 0 \iff f = g$ a.e.

- Define the auxiliary function $h(x) = |f(x) - g(x)|$. Since $f, g \in L^1$, their difference is integrable, meaning $h \in L^+$ and $\int h\, d\mu < \infty$.
- By Proposition 2.16, for any non-negative function $h \in L^+$, the integral $\int h\, d\mu = 0$ if and only if $h = 0$ almost everywhere.
- Noting that $|f(x) - g(x)| = 0 \iff f(x) = g(x)$, this means $\int |f - g|\, d\mu = 0 \iff f = g$ a.e., establishing the second equivalence.

Link 2: Forward Implication ($\int |f - g|\, d\mu = 0 \implies \int_E f\, d\mu = \int_E g\, d\mu$ for all E)

1. Assume that $\int |f - g|\, d\mu = 0$. Let $E \in \mathcal{M}$ be any arbitrary measurable set.
2. We look at the absolute modulus of the difference between the two set integrals. By complex linearity of the integral, we merge them:

$$\left| \int_E f\, d\mu - \int_E g\, d\mu \right| = \left| \int_E (f - g)\, d\mu \right| = \left| \int (f - g)\chi_E\, d\mu \right|$$

3. Applying Proposition 2.22 ($\left|\int \psi\right| \leq \int |\psi|$) to the function $\psi = (f - g)\chi_E$, we bound the expression from above:

$$\left| \int (f - g)\chi_E\, d\mu \right| \leq \int |(f - g)\chi_E|\, d\mu = \int |f - g|\chi_E\, d\mu$$

4. Since the indicator function satisfies $\chi_E(x) \leq 1$ everywhere, the monotonicity of the non-negative integral gives:

$$\int |f - g|\chi_E\, d\mu \leq \int |f - g|\, d\mu$$

5. By our initial assumption, $\int |f - g|\, d\mu = 0$. Stringing the inequalities together reveals:

$$0 \leq \left| \int_E f\, d\mu - \int_E g\, d\mu \right| \leq \int |f - g|\, d\mu = 0$$

6. The absolute value is forced to be exactly zero, which implies that $\int_E f\,d\mu = \int_E g\,d\mu$ for every measurable set E.

Link 3: Backward Implication ($\int_E f\,d\mu = \int_E g\,d\mu$ for all $E \implies f = g$ a.e.)

1. We prove this implication via contraposition. Suppose it is false that $f = g$ a.e.We will construct a specific set E where the integrals cannot match.
2. Let us define the difference function $D = f - g$. By assumption, D does not equal zero almost everywhere. Let us break D into its real and imaginary parts:
$$u = \mathrm{Re}(f - g), \quad v = \mathrm{Im}(f - g)$$
3. Since $D \neq 0$ on a set of positive measure, at least one of its components (u or v) must be non-zero on a set of positive measure. Furthermore, decomposing these real components into their positive and negative parts ($u = u^+ - u^-$ and $v = v^+ - v^-$), it follows that at least one of the four non-negative components (u^+, u^-, v^+, v^-) must be strictly positive on a set of positive measure.
4. Without loss of generality, let us assume that u^+ is non-zero on a set of positive measure. We define our target measurable test set E as the support of u^+:
$$E = \{x \in X : u^+(x) > 0\}$$
By our assumption, $\mu(E) > 0$.
5. Let us examine the behavior of u^- on this specific set E. By the construction of positive/negative parts, if $u^+(x) > 0$, then $u(x) > 0$, which automatically means $u^-(x) = 0$. Thus, $u^- = 0$ identically on E.
6. Now, we calculate the real part of the integrated difference over this set E:
$$\mathrm{Re}\left(\int_E f\,d\mu - \int_E g\,d\mu\right) = \int_E \mathrm{Re}(f - g)\,d\mu = \int_E u\,d\mu$$
7. Expanding $u = u^+ - u^-$ and using the fact that $u^- = 0$ on E, the equation simplifies to:
$$\int_E u\,d\mu = \int_E (u^+ - u^-)\,d\mu = \int_E u^+\,d\mu$$
8. Since $u^+ \in L^+$ and its support set E has a strictly positive measure, Proposition 2.16 guarantees that its integral over its support must be strictly greater than zero:
$$\int_E u^+\,d\mu > 0$$
9. Because its real part is strictly positive, the full complex value of the difference cannot be zero: $\int_E f\,d\mu - \int_E g\,d\mu \neq 0$.

An identical construction holds if u^-, v^+, or v^- were chosen as the non-zero component. Thus, if $f \neq g$ on a set of positive measure, there exists some $E \in \mathcal{M}$ such that $\int_E f \neq \int_E g$. This completes the proof of the backward implication by contraposition. □

Remark (Null Sets and the Structural Redefinition of L^1)**.** *The preceding proposition underscores a fundamental theme in measure theory: for the purposes of integration, behavior on null sets can be completely disregarded. This leads to two critical insights regarding the domain and structural nature of integrable functions:*

Part 1: Invariance Under Null Set Modifications

As demonstrated by Proposition 2.23(b), if two integrable functions f and g satisfy $f = g$ almost everywhere (a.e.), their integrals are identical over every measurable set $E \in \mathcal{M}$. This mathematical fact *has an immediate practical utility: we can integrate functions that are not even defined on the entire space X, provided their domain of definition covers a set E whose complement E^c is a null set ($\mu(E^c) = 0$).*

We can arbitrarily extend f to the remaining null set E^c by assigning $f(x) = 0$ (or any other value) for all $x \in E^c$ without altering any integration values. Similarly, this allows us to smoothly treat extended real-valued functions ($\overline{\mathbb{R}}$-valued functions) that take values of $\pm\infty$ on a null set as standard real-valued functions for integration purposes.

Part 2: The Rigorous Definition of $L^1(\mu)$ via Equivalence Classes

While it is intuitive to think of $L^1(\mu)$ as a collection of individual functions, doing so creates a serious algebraic defect if we wish to treat it as a normed vector space. In a true normed linear space, the norm of an element is zero ($\|f\|_1 = \int |f| \, d\mu = 0$) if and only if that element is the literal zero element ($f \equiv 0$ everywhere). However, any non-zero function that vanishes merely almost everywhere also satisfies $\int |f| \, d\mu = 0$.

*To restore this topological requirement, we redefine $L^1(\mu)$ as a space of **equivalence classes** under the equivalence relation of pointwise convergence almost everywhere:*

$$f \sim g \iff f(x) = g(x) \quad \text{for } \mu\text{-a.e. } x \in X$$

An element within this redefined space $L^1(\mu)$ is no longer a solitary function, but rather an entire equivalence class $[f]$, defined by:

$$[f] = \{g : X \to \mathbb{C} \mid g \text{ is measurable and } g = f \text{ a.e.}\}$$

Algebraic and Notational Consistency:

- ***Vector Space Operations:*** *Vector addition and scalar multiplication remain completely well-defined on these classes. If $f_1 \sim f_2$ and $g_1 \sim g_2$, then $(f_1 + g_1) \sim (f_2 + g_2)$ and $(cf_1) \sim (cf_2)$. Thus, $L^1(\mu)$ forms a legitimate complex vector space under pointwise a.e. operations.*
- ***Abuse of Notation:*** *To avoid cumbersome notation, we almost never write $[f] \in L^1(\mu)$ in practice. Instead, we employ the standard notation $f \in L^1(\mu)$, implicitly understanding that f represents its underlying equivalence class of a.e.-defined integrable functions. This universally accepted convention simplifies prose and rarely induces mathematical confusion.*

Remark (Integration on Non-Complete Measure Spaces and Functional Identification)**.** *When dealing with a measure space $(X, \mathcal{M}, \mu)$ that is not complete, a technical nuisance can arise: a function might be altered on a subset of a null set $N \in \mathcal{M}$, but if that subset is not itself in $\mathcal{M}$, the altered function could theoretically lose its measurability.*

However, because the integral completely ignores behavior on null sets, we can circumvent this difficulty entirely by working directly with the equivalence classes of almost everywhere-defined functions. Whether we formally take the completion of the measure space $(\mathcal{M}, \overline{\mu})$ or simply identify functions that match μ-almost everywhere, the resulting space of integrable equivalence classes $L^1(\mu)$ is exactly the same. Consequently, we can freely alter, restrict, or extend our functions on null sets without any risk of exiting the well-defined boundaries of our linear space.

Definition 44 (The L^1 Distance Metric)**.** *Let $f, g \in L^1(\mu)$ be two equivalence classes of integrable functions. We define the L^1 **distance function** $d : L^1(\mu) \times L^1(\mu) \to [0, \infty)$ by:*

$$d(f, g) = \int |f - g| \, d\mu$$

Proposition. *The function $d(f, g) = \int |f - g| \, d\mu$ is a well-defined metric on the space $L^1(\mu)$.*

Proof. To prove that d is a metric, we must verify that it satisfies the three fundamental axioms of a metric space for any arbitrary elements $f, g, h \in L^1(\mu)$:

1. **Positivity and Identity of Indiscernibles ($d(f, g) = 0 \iff f = g$):**

- Since $|f(x)-g(x)| \geq 0$ pointwise everywhere, the monotonicity of the integral ensures that $d(f,g) = \int |f-g|\,d\mu \geq 0$.
- If we evaluate functions individually, $\int |f-g|\,d\mu = 0$ only guarantees that $f(x) = g(x)$ almost everywhere. However, because $L^1(\mu)$ is explicitly defined as the space of equivalence classes under the relation of a.e. equality, the condition $f = g$ a.e. means that f and g represent the exact same equivalence class element. Therefore, $d(f,g) = 0 \iff f = g$ as elements of $L^1(\mu)$.

2. Symmetry ($d(f,g) = d(g,f)$): By the symmetric property of the absolute value modulus, $|f(x)-g(x)| = |g(x)-f(x)|$ pointwise for all $x \in X$. Integrating both sides yields:

$$d(f,g) = \int |f-g|\,d\mu = \int |g-f|\,d\mu = d(g,f)$$

3. The Triangle Inequality ($d(f,g) \leq d(f,h) + d(h,g)$): For any intermediary function h, we apply the classical numerical triangle inequality pointwise to the integrands:

$$|f(x)-g(x)| = |(f(x)-h(x)) + (h(x)-g(x))| \leq |f(x)-h(x)| + |h(x)-g(x)|$$

By applying the monotonicity and linearity properties of the integral on these non-negative functions, we maintain the inequality:

$$\int |f-g|\,d\mu \leq \int (|f-h| + |h-g|)\,d\mu = \int |f-h|\,d\mu + \int |h-g|\,d\mu$$

Substituting our definition of d gives:

$$d(f,g) \leq d(f,h) + d(h,g)$$

Since all three axioms hold, d is verified to be a genuine metric, making $(L^1(\mu), d)$ a legitimate metric space. □

Definition 45 (L^1 Convergence)**.** *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions in $L^1(\mu)$, and let $f \in L^1(\mu)$. We say that the sequence **converges to f in L^1** (or converges in the mean) if the metric distance between f_n and f tends to zero as $n \to \infty$:*

$$\lim_{n\to\infty} d(f_n, f) = \lim_{n\to\infty} \int |f_n - f|\,d\mu = 0$$

We denote this convergence behavior by $f_n \xrightarrow{L^1} f$.

Theorem 20 (The Dominated Convergence Theorem)**.** *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence in L^1 such that:*

(a) *$f_n \to f$ almost everywhere (a.e.) as $n \to \infty$, and*

(b) *there exists a non-negative function $g \in L^1$ such that $|f_n| \leq g$ a.e. for all n.*

Then $f \in L^1$ and the limit commutes with the integral operator:

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

Proof. We break down the verification of this foundational convergence result into explicit, structured steps.

Step 1: Measurability and Integrability of the Limit f

1. Pointwise limits almost everywhere preserve measurability. Since each f_n is measurable and $f_n \to f$ a.e., the function f is guaranteed to be measurable (if necessary, after redefining it on a null set where convergence fails).

2. Taking the absolute limit of the hypothesis $|f_n| \leq g$ a.e., it follows that the pointwise limit also obeys the same dominant bound:
$$|f| \leq g \quad \text{a.e.}$$
3. Since $g \in L^1$ is integrable, its integral is finite ($\int g \, d\mu < \infty$). By the monotonicity property of non-negative integrals, we find:
$$\int |f| \, d\mu \leq \int g \, d\mu < \infty$$
By definition, this confirms that $f \in L^1$.

Step 2: Reduction to Real-Valued Functions

- Suppose the theorem holds for all real-valued sequences. Since f_n and f can be decomposed into real and imaginary parts ($f_n = \operatorname{Re} f_n + i \operatorname{Im} f_n$), we can examine them separately.
- The complex modulus bounds satisfy $|\operatorname{Re} f_n| \leq |f_n| \leq g$ and $|\operatorname{Im} f_n| \leq |f_n| \leq g$. Thus, both component sequences are dominated by the same $g \in L^1$.
- By complex linearity, if the real and imaginary integrals converge independently, the full complex integral converges. Therefore, it suffices to carry out the remaining proof assuming f_n and f are purely real-valued.

Step 3: Constructing Non-Negative Sequences for Fatou's Lemma Because the functions f_n can take negative values, we cannot apply Fatou's lemma directly to them. However, since $-g \leq f_n \leq g$ a.e., we can manipulate the terms to isolate two separate, non-negative sequences:
$$\text{Sequence 1: } \psi_n = g + f_n \geq 0 \quad \text{a.e.}$$
$$\text{Sequence 2: } \phi_n = g - f_n \geq 0 \quad \text{a.e.}$$
Since $\psi_n, \phi_n \in L^+$, they are perfectly suited for Fatou's Lemma ($\int \liminf h_n \leq \liminf \int h_n$).

Step 4: Applying Fatou's Lemma to the First Sequence ($g + f_n$)

1. We take the limit inferior of the integrals of ψ_n. By linearity, we expand the expression:
$$\liminf_{n\to\infty} \int (g + f_n) \, d\mu = \liminf_{n\to\infty} \left(\int g \, d\mu + \int f_n \, d\mu \right)$$
Since $\int g \, d\mu$ is a constant real number independent of n, it factors out of the limit inferior directly:
$$\liminf_{n\to\infty} \left(\int g \, d\mu + \int f_n \, d\mu \right) = \int g \, d\mu + \liminf_{n\to\infty} \int f_n \, d\mu$$
2. Now we evaluate the other side of Fatou's inequality. Pointwise, $\lim f_n = f$ a.e., meaning $\liminf(g + f_n) = g + f$ a.e.İntegrating this limit yields:
$$\int \liminf_{n\to\infty} (g + f_n) \, d\mu = \int (g + f) \, d\mu = \int g \, d\mu + \int f \, d\mu$$
3. Combining these two components via Fatou's inequality gives:
$$\int g \, d\mu + \int f \, d\mu \leq \int g \, d\mu + \liminf_{n\to\infty} \int f_n \, d\mu$$
4. Subtracting the finite constant $\int g \, d\mu$ from both sides establishes our lower bound:
$$\int f \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu \qquad \text{(Inequality 1)}$$

Step 5: Applying Fatou's Lemma to the Second Sequence ($g - f_n$)

1. We mirror the same process for the non-negative sequence $\phi_n = g - f_n$. Expanding the limit inferior of the integrals yields:
$$\liminf_{n\to\infty} \int (g - f_n)\, d\mu = \liminf_{n\to\infty} \left(\int g\, d\mu - \int f_n\, d\mu \right)$$
2. When a negative sign is factored out of a limit inferior, it flips into a limit superior: $\liminf(-x_n) = -\limsup(x_n)$. Thus:
$$\liminf_{n\to\infty} \left(\int g\, d\mu - \int f_n\, d\mu \right) = \int g\, d\mu - \limsup_{n\to\infty} \int f_n\, d\mu$$
3. Pointwise, the integral of the limit inferior is:
$$\int \liminf_{n\to\infty} (g - f_n)\, d\mu = \int (g - f)\, d\mu = \int g\, d\mu - \int f\, d\mu$$
4. Stringing them together through Fatou's Lemma provides:
$$\int g\, d\mu - \int f\, d\mu \leq \int g\, d\mu - \limsup_{n\to\infty} \int f_n\, d\mu$$
5. Subtracting $\int g\, d\mu$ and multiplying the entire inequality by -1 reverses the inequality direction, establishing our upper bound:
$$\int f\, d\mu \geq \limsup_{n\to\infty} \int f_n\, d\mu \qquad \text{(Inequality 2)}$$

Step 6: Squeezing the Limits Together We combine Inequality 1 and Inequality 2 into a single chain:
$$\limsup_{n\to\infty} \int f_n\, d\mu \leq \int f\, d\mu \leq \liminf_{n\to\infty} \int f_n\, d\mu$$
For any sequence of numbers, the limit inferior is naturally bounded above by its limit superior ($\liminf x_n \leq \limsup x_n$). Thus, the sandwich/squeeze theorem forces these values to collapse into a singular equality:
$$\liminf_{n\to\infty} \int f_n\, d\mu = \limsup_{n\to\infty} \int f_n\, d\mu = \int f\, d\mu$$
Since the limit inferior and limit superior are equal, the ordinary limit exists and matches this shared value:
$$\lim_{n\to\infty} \int f_n\, d\mu = \int f\, d\mu$$
This completes the proof. □

Theorem 21. *Suppose that $\{f_j\}_{j=1}^{\infty}$ is a sequence of functions in L^1 such that:*
$$\sum_{j=1}^{\infty} \int |f_j|\, d\mu < \infty$$
Then the series $\sum_{j=1}^{\infty} f_j$ converges almost everywhere to a function in L^1, and the infinite summation commutes with the integral operator:
$$\int \sum_{j=1}^{\infty} f_j\, d\mu = \sum_{j=1}^{\infty} \int f_j\, d\mu$$

Proof. We break down the proof of this theorem into logical, independent steps, utilizing foundational results from integration theory without relying on textbook numbering.

Step 1: Constructing a Dominating Non-Negative Function

1. Let us define a function $g : X \to [0, \infty]$ as the infinite sum of the absolute values of our sequence components:
$$g(x) = \sum_{j=1}^{\infty} |f_j(x)|$$
By construction, g is a non-negative, measurable function since it is the pointwise limit of a monotonically increasing sequence of non-negative measurable functions (the partial sums).

2. We calculate the integral of g. By the Monotone Convergence Theorem (which permits the interchanging of integrals and infinite sums for non-negative functions), we can pull the summation out of the integral:
$$\int g\, d\mu = \int \left(\sum_{j=1}^{\infty} |f_j| \right) d\mu = \sum_{j=1}^{\infty} \int |f_j|\, d\mu$$

3. By our initial hypothesis, this infinite sum of integrals is strictly finite:
$$\int g\, d\mu < \infty$$
Therefore, by definition, g belongs to the space of integrable functions, L^1.

Step 2: Proving Almost Everywhere Convergence of the Series

1. Since $g \in L^1$ and $\int g\, d\mu < \infty$, the set of points where g takes an infinite value must be a null set. That is:
$$g(x) = \sum_{j=1}^{\infty} |f_j(x)| < \infty \quad \text{for } \mu\text{-almost every } x \in X$$

2. For any point x where this sum is finite, the numerical complex series $\sum_{j=1}^{\infty} f_j(x)$ is **absolutely convergent** in $\mathbb{C}$.

3. Since absolute convergence implies ordinary convergence for series of complex numbers, the pointwise limit function:
$$f(x) = \sum_{j=1}^{\infty} f_j(x)$$
exists and is well-defined for μ-almost every $x \in X$. We can set $f(x) = 0$ on the remaining null set where the series fails to converge, preserving its overall measurability.

Step 3: Verification that the Limit Function is in L^1

1. By the triangle inequality for complex numbers applied pointwise, the modulus of the limit function is bounded by the sum of the moduli:
$$|f(x)| = \left| \sum_{j=1}^{\infty} f_j(x) \right| \leq \sum_{j=1}^{\infty} |f_j(x)| = g(x) \quad \text{a.e.}$$

2. Applying the monotonicity property of the integral, we integrate this inequality:
$$\int |f|\, d\mu \leq \int g\, d\mu < \infty$$
Since its absolute integral is finite, $f \in L^1$.

Step 4: Applying the Dominated Convergence Theorem to Interchanging Operations

1. Let us define the sequence of partial sums $S_n(x)$ by:

$$S_n(x) = \sum_{j=1}^{n} f_j(x)$$

By definition, $S_n \to f$ almost everywhere as $n \to \infty$.

2. We check the dominance condition for this sequence of partial sums. Using the triangle inequality once more, each partial sum satisfies:

$$|S_n(x)| = \left| \sum_{j=1}^{n} f_j(x) \right| \leq \sum_{j=1}^{n} |f_j(x)| \leq \sum_{j=1}^{\infty} |f_j(x)| = g(x) \quad \text{a.e.}$$

Thus, $|S_n| \leq g$ almost everywhere for all n, where $g \in L^1$ is our non-negative dominating function.

3. Since both conditions of the Dominated Convergence Theorem are satisfied ($S_n \to f$ a.e. and $|S_n| \leq g$ a.e.), we conclude that the limit commutes with the integral operator:

$$\int f \, d\mu = \lim_{n\to\infty} \int S_n \, d\mu$$

4. Expanding f and S_n back into their series representations, we get:

$$\int \left(\sum_{j=1}^{\infty} f_j \right) d\mu = \lim_{n\to\infty} \int \left(\sum_{j=1}^{n} f_j \right) d\mu$$

5. By the linearity of the integral for finite combinations, we can swap the finite sum and the integral on the right side:

$$\lim_{n\to\infty} \int \left(\sum_{j=1}^{n} f_j \right) d\mu = \lim_{n\to\infty} \sum_{j=1}^{n} \int f_j \, d\mu$$

6. By the definition of an infinite series of numbers, the limit of the finite partial sums is exactly the infinite sum:

$$\lim_{n\to\infty} \sum_{j=1}^{n} \int f_j \, d\mu = \sum_{j=1}^{\infty} \int f_j \, d\mu$$

Combining these links yields our final desired identity:

$$\int \sum_{j=1}^{\infty} f_j \, d\mu = \sum_{j=1}^{\infty} \int f_j \, d\mu$$

□

1 Modes of Convergence and Counterexamples

When analyzing sequences of functions on a measure space $(X, \mathcal{M}, \mu)$, convergence can be understood in several distinct mathematical frameworks. Before exploring the detailed counterexamples, let us explicitly state the classical directional hierarchy connecting the standard pointwise modes of convergence:

Uniform Convergence $\implies$ Pointwise Convergence $\implies$ Almost Everywhere (a.e.) Convergence

Importantly, **none of these standard pointwise modes imply L^1 convergence**, nor does L^1 convergence inherently imply any of them.

To rigorously illustrate these gaps, we analyze four classical counterexamples defined on the real line equipped with the standard Lebesgue measure, $(\mathbb{R}, \mathcal{L}, m)$.

—

1.1 Example I: The Escape to Horizontal Infinity (The Dying Moving Rectangles)

Let $f_n = n^{-1}\chi_{(0,n)}$ for $n \in \mathbb{N}$.

Analysis of Pointwise and Uniform Convergence: For any fixed $x \in \mathbb{R}$, we look at the tail limit of the sequence:

- If $x \leq 0$, then $f_n(x) = 0$ for all n, so $\lim_{n\to\infty} f_n(x) = 0$.
- If $x > 0$, by the Archimedean property, there exists an index $N \in \mathbb{N}$ such that $n > x$ for all $n \geq N$. Thus, for all $n \geq N$, the point x falls inside the interval $(0, n)$, giving $f_n(x) = \frac{1}{n}$. Fixing x and taking the limit yields $\lim_{n\to\infty} \frac{1}{n} = 0$.

Hence, $f_n \to 0$ pointwise everywhere. Furthermore, because the supremum error across the entire real line satisfies:

$$\sup_{x\in\mathbb{R}} |f_n(x) - 0| = \frac{1}{n} \xrightarrow{n\to\infty} 0$$

this sequence converges to 0 **uniformly** on $\mathbb{R}$.

Analysis of L^1 Convergence: We compute the L^1 norm of f_n directly by evaluating its integral:

$$\|f_n - 0\|_1 = \int_{\mathbb{R}} \frac{1}{n}\chi_{(0,n)}\, dm = \frac{1}{n} \cdot m\big((0, n)\big) = \frac{1}{n} \cdot n = 1$$

Taking the limit as $n \to \infty$:

$$\lim_{n\to\infty} \|f_n - 0\|_1 = 1 \neq 0$$

Conclusion: Despite converging uniformly and pointwise everywhere to 0, $f_n \stackrel{L^1}{\not\longrightarrow} 0$. Uniform convergence on an infinite measure space does not guarantee L^1 convergence.

—

1.2 Example II: The Moving Train (Escape to Spatial Infinity)

Let $f_n = \chi_{(n,n+1)}$ for $n \in \mathbb{N}$.

Analysis of Pointwise Convergence: Fix any arbitrary point $x \in \mathbb{R}$. As n grows, the interval $(n, n + 1)$ shifts progressively out to positive infinity. Choosing an integer $N > x$ ensures that for all $n \geq N$, $x \notin (n, n + 1)$, meaning $f_n(x) = 0$. Therefore:

$$\lim_{n\to\infty} f_n(x) = 0 \quad \text{for every } x \in \mathbb{R}$$

Thus, $f_n \to 0$ pointwise everywhere (and consequently, almost everywhere). Note that this convergence is not uniform, because $\sup_{x\in\mathbb{R}} |f_n(x)| = 1$ for every n.

Analysis of L^1 Convergence: Calculating the total area under each moving profile reveals:

$$\|f_n - 0\|_1 = \int_{\mathbb{R}} \chi_{(n,n+1)}\, dm = m\big((n, n + 1)\big) = 1$$

Taking the limit as $n \to \infty$:

$$\lim_{n\to\infty} \|f_n - 0\|_1 = 1 \neq 0$$

Conclusion: Pointwise convergence everywhere does not imply L^1 convergence, as a fixed chunk of mass can completely escape to infinity.

—

1.3 Example III: The Expanding Spike (Escape to Vertical Infinity)

Let $f_n = n\chi_{(0,1/n)}$ for $n \in \mathbb{N}$.

Analysis of Pointwise and a.e. Convergence: Let us evaluate the limit at individual points $x \in \mathbb{R}$:

- If $x \leq 0$, then $f_n(x) = 0$ for all n, so $\lim_{n\to\infty} f_n(x) = 0$.
- If $x > 0$, by the Archimedean property, we can find an index $N \in \mathbb{N}$ such that $\frac{1}{N} < x$. For all indices $n \geq N$, it follows that $\frac{1}{n} < x$, meaning $x \notin (0, \frac{1}{n})$. Consequently, $f_n(x) = 0$ for all $n \geq N$, yielding $\lim_{n\to\infty} f_n(x) = 0$.

Thus, $f_n(x) \to 0$ for every single $x \in \mathbb{R}$. The sequence converges pointwise everywhere (and hence a.e.) to the zero function.

Analysis of L^1 Convergence: Evaluating the L^1 norm metric yields:

$$\|f_n - 0\|_1 = \int_{\mathbb{R}} n\chi_{(0,1/n)}\, dm = n \cdot m\big((0, 1/n)\big) = n \cdot \frac{1}{n} = 1$$

As $n \to \infty$, the limit of the integral values remains exactly 1:

$$\lim_{n\to\infty} \|f_n - 0\|_1 = 1 \neq 0$$

Conclusion: Even on a bounded domain of finite total measure (the unit interval), pointwise convergence everywhere does not imply L^1 convergence if mass concentrates vertically into a sharpening singularity.

—

1.4 Example IV: The Typewriter Sequence (The Converse Failure)

We define a sequence of indicator functions over the compact interval $[0, 1]$ by slicing it into finer binary subintervals. Any index $n \in \mathbb{N}$ can be uniquely expressed as:

$$n = 2^k + j \quad \text{where } k \geq 0 \text{ and } 0 \leq j < 2^k$$

Using this parameterization, we define our sequence as:

$$f_n = \chi_{I_n} \quad \text{where } I_n = \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right]$$

The first few terms of this sequence are:

$$f_1 = \chi_{[0,1]}, \quad f_2 = \chi_{[0,1/2]}, \quad f_3 = \chi_{[1/2,1]}, \quad f_4 = \chi_{[0,1/4]}, \quad f_5 = \chi_{[1/4,1/2]}, \; \ldots$$

Analysis of L^1 Convergence: Let us calculate the L^1 distance between f_n and the zero function 0:

$$\|f_n - 0\|_1 = \int_{[0,1]} \chi_{I_n}\, dm = m(I_n) = \frac{1}{2^k}$$

As $n \to \infty$, the structural generation index $k \to \infty$ as well. Taking the limit shows:

$$\lim_{n\to\infty} \|f_n - 0\|_1 = \lim_{k\to\infty} \frac{1}{2^k} = 0$$

Therefore, by definition, $f_n \xrightarrow{L^1} 0$.

Analysis of Pointwise and a.e. Convergence: Let us track the behavior of the sequence at any fixed point $x \in [0, 1]$. As the indices progress, the subintervals I_n repeatedly sweep across the entire interval $[0, 1]$ like a typewriter carriage.

- For each level k, the collection of intervals $\{I_n\}_{j=0}^{2^k-1}$ forms a complete cover of $[0, 1]$. This guarantees that there is always a choice of j such that $x \in I_n$, forcing $f_n(x) = 1$ infinitely many times.
- Simultaneously, for each level k, there are remaining subintervals that do not contain x, which means $f_n(x) = 0$ infinitely many times as well.

Because the numerical sequence $\{f_n(x)\}_{n=1}^{\infty}$ constantly oscillates between 0 and 1, the limit $\lim_{n\to\infty} f_n(x)$ does not exist for any point $x \in [0, 1]$.
Conclusion: The sequence f_n converges to 0 in L^1, but it does not converge pointwise anywhere on $[0, 1]$ (meaning it also fails to converge almost everywhere).

—

1.5 Summary of Convergence Insights

By examining these observations together, we get a clear picture of how these topologies are completely independent:

- **Example I** demonstrates that uniform convergence on an infinite measure space is insufficient to ensure L^1 convergence because mass can flatten out across horizontal infinity.
- **Example II** demonstrates that pointwise convergence everywhere is insufficient to ensure L^1 convergence because a fixed lump of mass can escape to spatial infinity.
- **Example III** demonstrates that pointwise convergence everywhere on a bounded domain is insufficient to ensure L^1 convergence because mass can escape vertically into a singular spike.
- **Example IV** demonstrates the complete failure of the converse; a sequence can converge perfectly to zero in the L^1 metric space while failing to settle on a pointwise limit at any point whatsoever.

This collective behavior proves that L^1 convergence is a global metric property governed by integrated area, making it separate from the local constraints of pointwise and uniform convergence systems.

Remark. *While it is generally true that almost everywhere (a.e.) convergence does not imply L^1 convergence, we can bridge this gap by introducing a dominating baseline. Specifically, if a sequence of functions $\{f_n\}$ satisfies $f_n \to f$ a.e. and there exists a dominating function $g \in L^1$ such that $|f_n| \leq g$ for all n, then $f_n \to f$ in L^1.*

This relationship follows directly as an application of the Dominated Convergence Theorem. Because the triangle inequality gives the pointwise bound:

$$|f_n - f| \leq |f_n| + |f| \leq g + g = 2g$$

and $2g$ remains integrable, the area under the sequence cannot escape to infinity, forcing the integrated limit to vanish:

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

Conversely, if a sequence is known to converge in L^1, it is guaranteed that some subsequence will converge to f a.e..

Definition 46. *Let $\{f_n\}$ be a sequence of measurable, complex-valued functions on a measure space $(X, \mathcal{M}, \mu)$.*

- *The sequence $\{f_n\}$ is said to be **Cauchy in measure** if for every $\epsilon > 0$:*

$$\mu\big(\{x \in X : |f_n(x) - f_m(x)| \geq \epsilon\}\big) \xrightarrow{m,n\to\infty} 0$$

- *The sequence $\{f_n\}$ **converges in measure** to a measurable function f if for every $\epsilon > 0$:*

$$\mu\big(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}\big) \xrightarrow{n\to\infty} 0$$

Remark. *To highlight how convergence in measure interacts with other topological structures, we evaluate its behavior across the four classical counterexamples defined on the standard real line measure space $(\mathbb{R}, \mathcal{L}, m)$ targeting the zero function ($f = 0$):*

1. ***Example I (The Escape to Horizontal Infinity):*** *Let $f_n = n^{-1}\chi_{(0,n)}$. For any fixed deviation threshold $\epsilon > 0$, we can choose $N \in \mathbb{N}$ such that $\frac{1}{N} < \epsilon$. For all $n \geq N$, the maximum vertical height of f_n is $\frac{1}{n} \leq \frac{1}{N} < \epsilon$. Consequently, the set where the function is at least ϵ is completely empty, yielding $m(\{x : |f_n(x) - 0| \geq \epsilon\}) = m(\emptyset) = 0$. Thus, this sequence **converges to zero in measure**.*
2. ***Example II (The Moving Train):*** *Let $f_n = \chi_{(n,n+1)}$. If we choose $\epsilon = \frac{1}{2}$, then for any two distinct indices m and n with $|n-m| \geq 1$, the underlying support intervals $(n, n+1)$ and $(m, m+1)$ are disjoint. On $(n, n+1)$, we have $|f_n(x) - f_m(x)| = |1 - 0| = 1 \geq \frac{1}{2}$, and symmetrically the same holds on $(m, m+1)$. Thus, $m(\{x : |f_n(x) - f_m(x)| \geq \frac{1}{2}\}) \geq 1 + 1 = 2$. Because this error set does not shrink as $m, n \to \infty$, the sequence is **not Cauchy in measure** (and hence cannot converge in measure).*

3. ***Example III (The Expanding Spike):*** *Let $f_n = n\chi_{(0,1/n)}$. For any fixed choice of $\epsilon > 0$, as soon as $n \geq \epsilon$, the spike height exceeds our threshold precisely over its support zone $(0, \frac{1}{n})$. Measuring this region gives $m(\{x : |f_n(x) - 0| \geq \epsilon\}) = m((0, \frac{1}{n})) = \frac{1}{n}$. Taking the limit yields $\lim_{n\to\infty} \frac{1}{n} = 0$, meaning the sequence* ***converges to zero in measure****.*

4. ***Example IV (The Typewriter Sequence):*** *Let $f_n = \chi_{I_n}$ over $[0,1]$, where $m(I_n) = \frac{1}{2^k}$ for $n = 2^k + j$. For any fixed tolerance level $0 < \epsilon \leq 1$, the function $f_n(x)$ matches or exceeds ϵ exactly when $x \in I_n$. Calculating the size of this region gives $m(\{x : |f_n(x) - 0| \geq \epsilon\}) = m(I_n) = \frac{1}{2^k}$. As $n \to \infty$, the generation index $k \to \infty$, forcing $\lim_{n\to\infty} \frac{1}{2^k} = 0$. Thus, despite completely failing to converge pointwise at any individual point in $[0,1]$, the sequence* ***converges to zero in measure****.*

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

Proof. For any given tolerance threshold $\epsilon > 0$, let us define the set $E_{n,\epsilon}$ containing all points where the absolute deviation between $f_n(x)$ and $f(x)$ is at least ϵ:

$$E_{n,\epsilon} = \{x \in X : |f_n(x) - f(x)| \geq \epsilon\}$$

We can establish a lower bound on the total L^1 integral distance by restricting the domain of integration from the entire space X down to this specific subset $E_{n,\epsilon}$:

$$\int_X |f_n - f| \, d\mu \geq \int_{E_{n,\epsilon}} |f_n - f| \, d\mu$$

Because every point x inside $E_{n,\epsilon}$ satisfies the property $|f_n(x) - f(x)| \geq \epsilon$ by definition, we can replace the integrand with this uniform lower bound:

$$\int_{E_{n,\epsilon}} |f_n - f| \, d\mu \geq \int_{E_{n,\epsilon}} \epsilon \, d\mu = \epsilon \cdot \mu(E_{n,\epsilon})$$

Combining these two inequalities yields the standard Chebyshev-type bound for integrals:

$$\int_X |f_n - f| \, d\mu \geq \epsilon \cdot \mu(E_{n,\epsilon}) \implies \mu(E_{n,\epsilon}) \leq \frac{1}{\epsilon} \int_X |f_n - f| \, d\mu$$

By hypothesis, the sequence converges in L^1, meaning $\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0$. Taking the limit on both sides as $n \to \infty$ forces the measure of the set to vanish:

$$\mu(E_{n,\epsilon}) \leq \frac{1}{\epsilon} \cdot 0 \implies \lim_{n\to\infty} \mu\big(\{x \in X : |f_n(x) - f(x)| \geq \epsilon\}\big) = 0$$

Since this holds for every arbitrary $\epsilon > 0$, we conclude that $f_n \to f$ in measure. $\square$

Remark. *The converse of the proposition is false: a sequence of functions can converge in measure without converging in the L^1 norm. This directional breakdown can be clearly demonstrated using two of our earlier classical counterexamples defined on the real line with the standard Lebesgue measure $(\mathbb{R}, \mathcal{L}, m)$ aiming toward the zero function ($f = 0$):*

- ***Failure via Example I (Escape to Horizontal Infinity):*** *Let $f_n = n^{-1}\chi_{(0,n)}$. As shown previously, for any fixed $\epsilon > 0$, choosing any index $n > \frac{1}{\epsilon}$ bounds the vertical profile entirely below the error threshold, making $m(\{x : |f_n(x) - 0| \geq \epsilon\}) = 0$ for all large n. Thus, $f_n \to 0$ in measure. However, evaluating its integrated mass gives:*

$$\|f_n - 0\|_1 = \int_{\mathbb{R}} \frac{1}{n}\chi_{(0,n)} \, dm = \frac{1}{n} \cdot n = 1 \implies \lim_{n\to\infty} \|f_n - 0\|_1 = 1 \neq 0$$

  Therefore, the sequence fails to converge to 0 in L^1 because the integrated area flattens but stretches out indefinitely across horizontal infinity.

- ***Failure via Example III (Escape to Vertical Infinity):*** *Let $f_n = n\chi_{(0,1/n)}$. For any fixed threshold $\epsilon > 0$, the function value matches or exceeds ϵ strictly over the interval $(0, \frac{1}{n})$ as long as $n \geq \epsilon$. Measuring this support zone gives:*

$$m\big(\{x : |f_n(x) - 0| \geq \epsilon\}\big) = m\left(\left(0, \frac{1}{n}\right)\right) = \frac{1}{n} \xrightarrow{n\to\infty} 0$$

This confirms convergence to 0 *in measure. Yet, calculating the total integrated mass underneath this moving profile shows:*

$$\|f_n - 0\|_1 = \int_{\mathbb{R}} n\chi_{(0,1/n)}\, dm = n \cdot \frac{1}{n} = 1 \implies \lim_{n\to\infty} \|f_n - 0\|_1 = 1 \neq 0$$

The sequence fails to converge to 0 *in* L^1 *because the mass concentrates vertically into an unbounded singular spike over a domain of shrinking width.*

Theorem 22. *Suppose that* $\{f_n\}$ *is a sequence of measurable, complex-valued functions that is Cauchy in measure. Then there exists a measurable function* f *such that* $f_n \to f$ *in measure, and there is a subsequence* $\{f_{n_j}\}$ *that converges to* f *almost everywhere (a.e.). Moreover, if there is another function* g *such that* $f_n \to g$ *in measure, then* $g = f$ *a.e.*

Corollary 23. *If* $f_n \to f$ *in* L^1*, then there is a subsequence* $\{f_{n_j}\}$ *such that* $f_{n_j} \to f$ *almost everywhere (a.e.).*

Proof. Suppose the sequence $\{f_n\}$ converges to f in L^1. By the previously established proposition, any sequence that converges in L^1 must also converge to the same limit in measure. Therefore, we have:

$$f_n \to f \text{ in measure}$$

Every sequence that converges in measure is structurally Cauchy in measure. Thus, by applying the theorem stated directly above, the existence of a subsequence $\{f_{n_j}\}$ that converges to f almost everywhere (a.e.) is guaranteed. □

Remark. *In general, almost everywhere (a.e.) convergence does not imply convergence in measure. A classical illustration of this failure is given by* ***Example II (The Moving Train)*** *defined on the real line with the standard Lebesgue measure* $(\mathbb{R}, \mathcal{L}, m)$*:*

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

As established previously, for any fixed $x \in \mathbb{R}$*, selecting an integer* $N > x$ *ensures that* $f_n(x) = 0$ *for all* $n \geq N$*, meaning* $f_n \to 0$ *pointwise everywhere (and thus a.e.).*

However, if we test convergence in measure to the zero function ($f = 0$) *with an error tolerance* $\epsilon = \frac{1}{2}$*, the set of points where the function is at least* ϵ *is exactly the shifting interval itself:*

$$m\big(\{x \in \mathbb{R} : |f_n(x) - 0| \geq \tfrac{1}{2}\}\big) = m\big((n, n+1)\big) = 1$$

Taking the limit as $n \to \infty$*, this measure remains constantly* 1 *and fails to approach* 0*. Thus,* f_n *does not converge to* 0 *in measure because the unit block of mass escapes completely to spatial infinity.*

Theorem 24 (Egoroff's Theorem)**.** *Suppose that* $\mu(X) < \infty$*, and* $f_1, f_2, \ldots$ *and* f *are measurable complex-valued functions on* X *such that* $f_n \to f$ *a.e. Then for every* $\epsilon > 0$ *there exists a measurable set* $E \subset X$ *such that* $\mu(E) < \epsilon$ *and* $f_n \to f$ *uniformly on* E^c*.*

Remark. *Egoroff's Theorem reveals that on a finite measure space, almost everywhere convergence is close to being uniform convergence. Specifically, while true uniform convergence across the entire space can still fail, we can strip away an arbitrarily small piece of the domain* (E) *with* $\mu(E) < \epsilon$*, leaving a large remaining ground* (E^c) *where the function sequence behaves uniformly.*

Definition 47. *A sequence of measurable functions* $\{f_n\}$ *is said to converge* ***almost uniformly*** *to a function* f *if for every* $\epsilon > 0$*, there exists a measurable set* E *such that* $\mu(E) < \epsilon$ *and* $f_n \to f$ *uniformly on* E^c*.*

2 Product Measures

When dealing with two separate measure spaces, we can construct a unified measure space on their Cartesian product. This process mimics how the area of a plane region can be built from elementary rectangles whose sides are measurable sets. The construction proceeds systematically from rectangles to a global measure using the framework of premeasures and outer measures.

2.1 Step 1: Defining Measurable Rectangles

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two arbitrary measure spaces. The foundational building blocks for our product space are the Cartesian products of measurable sets from each individual space.

- A subset $A \times B \subset X \times Y$ is called a **measurable rectangle** if $A \in \mathcal{M}$ and $B \in \mathcal{N}$.
- The geometric and algebraic interaction between these rectangles satisfies simple set-theoretic identities under intersection and complementation:

$$(A \times B) \cap (E \times F) = (A \cap E) \times (B \cap F)$$
$$(A \times B)^c = (X \times B^c) \cup (A^c \times B)$$

2.2 Step 2: Constructing the Algebra of Rectangles

While the collection of all measurable rectangles is stable under intersections, it is not closed under finite unions or complements. To remedy this, we define a larger collection:

- Let $\mathcal{A}$ be the collection consisting of all finite disjoint unions of measurable rectangles.
- Due to the intersection and complement rules established in Step 1, the collection $\mathcal{A}$ forms an **algebra** of sets.
- The σ-algebra generated by this algebra $\mathcal{A}$ (or equivalently, generated by the collection of all individual measurable rectangles) is called the **product σ-algebra** on $X \times Y$, denoted by:

$$\mathcal{M} \otimes \mathcal{N}$$

2.3 Step 3: Defining a Premeasure via Characteristic Functions

To define a consistent notion of "volume" or "area" on the algebra $\mathcal{A}$, we must assign a size to any element $E \in \mathcal{A}$. If $A \times B$ is a rectangle, its natural measure should be $\mu(A)\nu(B)$. For a general set $E \in \mathcal{A}$ expressed as a countable (or finite) disjoint union of rectangles $E = \bigcup_{j=1}^{\infty}(A_j \times B_j)$, we must ensure that adding up individual rectangle areas is a well-defined operation.

We verify this using the properties of characteristic functions. For any point $x \in X$ and $y \in Y$, the characteristic function of a disjoint union splits pointwise:

$$\chi_E(x, y) = \chi_{\bigcup(A_j \times B_j)}(x, y) = \sum_{j=1}^{\infty} \chi_{A_j \times B_j}(x, y) = \sum_{j=1}^{\infty} \chi_{A_j}(x)\chi_{B_j}(y)$$

To find the measure, we integrate this identity with respect to the individual measures step-by-step:

1. Integrating first with respect to x over the space X, and interchanging the order of summation and integration, we obtain:

$$\int_X \chi_E(x, y)\, d\mu(x) = \sum_{j=1}^{\infty} \int_X \chi_{A_j}(x)\chi_{B_j}(y)\, d\mu(x) = \sum_{j=1}^{\infty} \mu(A_j)\chi_{B_j}(y)$$

2. Next, we integrate this resulting function with respect to y over the space Y. Interchanging the summation and integration once more yields:

$$\int_Y \left(\int_X \chi_E(x, y)\, d\mu(x) \right) d\nu(y) = \sum_{j=1}^{\infty} \mu(A_j) \int_Y \chi_{B_j}(y)\, d\nu(y) = \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j)$$

This double-integration process proves that the value $\sum \mu(A_j)\nu(B_j)$ depends only on the set E itself, not on the choice of how E is decomposed. Thus, we can unambiguously define a set function π on the algebra $\mathcal{A}$. For any finite disjoint union $E = \bigcup_1^n (A_j \times B_j)$, we define:

$$\pi(E) = \sum_{j=1}^{n} \mu(A_j)\nu(B_j)$$

The analytical interchange of summation and integration guarantees that π is countably additive on $\mathcal{A}$, meaning π is a valid **premeasure** on the algebra $\mathcal{A}$.

2.4 Step 4: Extension to the Product Measure

With a well-behaved premeasure π established on the algebra $\mathcal{A}$, we can leverage Carathéodory's general extension procedure to expand it across the entire space:

1. We first construct an **outer measure** $(\mu \times \nu)^*$ on all subsets of $X \times Y$ by covering any arbitrary set with countable collections of elements from the algebra $\mathcal{A}$.
2. Carathéodory's theorem states that the collection of all sets that split every other set cleanly with respect to this outer measure forms a σ-algebra, and the restriction of the outer measure to that σ-algebra yields a fully complete measure.
3. Since the algebra of rectangles $\mathcal{A}$ generates the product σ-algebra $\mathcal{M} \otimes \mathcal{N}$, this product σ-algebra is automatically contained inside the family of outer-measurable sets.

By restricting this outer measure specifically to $\mathcal{M} \otimes \mathcal{N}$, we obtain our final destination: the **product measure**, denoted by $\mu \times \nu$. This unique measure extends our premeasure and satisfies the expected geometric property for every basic measurable rectangle:

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B) \quad \text{for all } A \in \mathcal{M}, B \in \mathcal{N}$$

If both underlying spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite, this extension is completely unique.

Remark. *The systematic construction established above for two measure spaces generalizes naturally to any finite number of factors. Specifically, let $(X_j, \mathcal{M}_j, \mu_j)$ be measure spaces for $j = 1, \ldots, n$.*

*By defining a general **measurable rectangle** to be a subset of the Cartesian product of the form:*

$$A_1 \times \cdots \times A_n \quad \text{where } A_j \in \mathcal{M}_j \text{ for each } j$$

the collection $\mathcal{A}$ consisting of all finite disjoint unions of these n-dimensional rectangles forms a valid algebra. Following the identical premeasure extension and Carathéodory restriction scheme, we obtain a unique product measure $\mu_1 \times \cdots \times \mu_n$ on the product σ-algebra $\mathcal{M}_1 \otimes \cdots \otimes \mathcal{M}_n$ such that:

$$(\mu_1 \times \cdots \times \mu_n)(A_1 \times \cdots \times A_n) = \prod_{j=1}^{n} \mu_j(A_j)$$

Furthermore, if each individual component measure μ_j is σ-finite, this finite product satisfies standard associative properties. For instance, a three-factor space can be regrouped step-by-step as an iterative pair without changing the underlying outcome:

$$\mathcal{M}_1 \otimes \mathcal{M}_2 \otimes \mathcal{M}_3 = (\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3$$

$$\mu_1 \times \mu_2 \times \mu_3 = (\mu_1 \times \mu_2) \times \mu_3$$

This allows calculations on higher finite dimensions to be broken down recursively into operations on pairs of spaces.

2.5 Sections of Sets and Functions

When working with product spaces $X \times Y$, it is essential to analyze higher-dimensional subsets and functions by looking at their lower-dimensional "slices" or cross-sections. This technique allows us to fix a coordinate in one space and study the behavior on the remaining space.

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two measure spaces, and let E be a subset of the Cartesian product space $X \times Y$.

Definition 48 (Sections of Sets)**.** *For any point $x \in X$ and $y \in Y$, we define:*

- *The x-**section** of E (denoted E_x) is the subset of Y consisting of all elements whose first coordinate is x:*
$$E_x = \{y \in Y : (x, y) \in E\}$$
- *The y-**section** of E (denoted E^y) is the subset of X consisting of all elements whose second coordinate is y:*
$$E^y = \{x \in X : (x, y) \in E\}$$

Definition 49 (Sections of Functions)**.** *Similarly, if f is a function defined on the product space $X \times Y$, we track its profile slices across individual lines by fixing a variable:*

- *The x-**section** of f is the function f_x defined on Y by varying the second component:*
$$f_x(y) = f(x, y)$$
- *The y-**section** of f is the function f^y defined on X by varying the first component:*
$$f^y(x) = f(x, y)$$

Remark. *An immediate, practical application of these definitions arises when dealing with characteristic functions of product sets. Taking a section of a characteristic function of a set E is completely identical to evaluating the characteristic function of that set's corresponding geometric section:*

$$(\chi_E)_x = \chi_{E_x} \quad \text{and} \quad (\chi_E)^y = \chi_{E^y}$$

Exercise 10. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be measure spaces, and let $\mathcal{M} \otimes \mathcal{N}$ denote the product σ-algebra on $X \times Y$. Prove the following properties regarding sections of sets and functions:*

a. If $E \in \mathcal{M} \otimes \mathcal{N}$, then for every $x \in X$, the x-section E_x belongs to $\mathcal{N}$, and for every $y \in Y$, the y-section E^y belongs to $\mathcal{M}$.

b. If f is a $\mathcal{M} \otimes \mathcal{N}$-measurable function on $X \times Y$, then for every $x \in X$, the section f_x is $\mathcal{N}$-measurable, and for every $y \in Y$, the section f^y is $\mathcal{M}$-measurable.

Proof. **Part (a): Measurability of Set Sections**

To prove that sections of any set in the product σ-algebra remain measurable in their respective component spaces, we employ the standard monotone class or σ-algebra generation argument. Let $\mathcal{R}$ be the collection of all subsets $E \subset X \times Y$ whose sections are appropriately measurable:

$$\mathcal{R} = \{E \subset X \times Y : E_x \in \mathcal{N}\ \forall x \in X, \text{ and } E^y \in \mathcal{M}\ \forall y \in Y\}$$

We proceed step-by-step to show that $\mathcal{R}$ is a σ-algebra containing all measurable rectangles:

- **Step 1: Base case for basic measurable rectangles.** Let $E = A \times B$ where $A \in \mathcal{M}$ and $B \in \mathcal{N}$. Fix any $x \in X$. By definition of the x-section:
$$(A \times B)_x = \{y \in Y : (x, y) \in A \times B\} = \begin{cases} B & \text{if } x \in A \\ \emptyset & \text{if } x \notin A \end{cases}$$
Since $B \in \mathcal{N}$ and $\emptyset \in \mathcal{N}$, it follows that $(A \times B)_x \in \mathcal{N}$ for all $x \in X$. Symmetrically, fixing any $y \in Y$ yields:
$$(A \times B)^y = \{x \in X : (x, y) \in A \times B\} = \begin{cases} A & \text{if } y \in B \\ \emptyset & \text{if } y \notin B \end{cases}$$
Since $A \in \mathcal{M}$ and $\emptyset \in \mathcal{M}$, we have $(A \times B)^y \in \mathcal{M}$ for all $y \in Y$. Thus, every basic measurable rectangle belongs to $\mathcal{R}$.

- **Step 2: Closure under complements.** Suppose $E \in \mathcal{R}$. Consider its complement E^c. For any fixed $x \in X$, a point $y \in (E^c)_x$ if and only if $(x, y) \notin E$, which means $y \notin E_x$. This gives the set identity:
$$(E^c)_x = (E_x)^c$$
Since $E \in \mathcal{R}$, we know $E_x \in \mathcal{N}$. Because $\mathcal{N}$ is a σ-algebra, its complement $(E_x)^c$ also belongs to $\mathcal{N}$. By identical logic for y-sections, $(E^c)^y = (E^y)^c \in \mathcal{M}$ because $E^y \in \mathcal{M}$. Thus, $E^c \in \mathcal{R}$.

- **Step 3: Closure under countable unions.** Let $\{E_j\}_{j=1}^{\infty}$ be a countable sequence of sets in $\mathcal{R}$. For any fixed $x \in X$, a point y is in the section of the union if and only if $(x, y) \in \bigcup_1^\infty E_j$. This occurs if $(x, y) \in E_j$ for some j, meaning $y \in (E_j)_x$ for some j. Hence, we have the identity:
$$\left(\bigcup_{j=1}^{\infty} E_j\right)_x = \bigcup_{j=1}^{\infty} (E_j)_x$$

Since each $E_j \in \mathcal{R}$, each component $(E_j)_x \in \mathcal{N}$. Because $\mathcal{N}$ is closed under countable unions, the resulting section belongs to $\mathcal{N}$. An identical breakdown applies to the y-sections:

$$\left(\bigcup_{j=1}^{\infty} E_j\right)^y = \bigcup_{j=1}^{\infty} (E_j)^y \in \mathcal{M}$$

Therefore, $\bigcup_{j=1}^{\infty} E_j \in \mathcal{R}$.

Since $\mathcal{R}$ contains the empty set, is closed under complements, and is closed under countable unions, it forms a σ-algebra. Because $\mathcal{R}$ contains all measurable rectangles, it must contain the smallest σ-algebra generated by them:

$$\mathcal{R} \supset \mathcal{M} \otimes \mathcal{N}$$

This proves that for every $E \in \mathcal{M} \otimes \mathcal{N}$, $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$.

—

Part (b): Measurability of Function Sections

Let $f : X \times Y \to \mathbb{C}$ (or $\overline{\mathbb{R}}$) be a measurable function with respect to the product σ-algebra $\mathcal{M} \otimes \mathcal{N}$. To establish that the section functions f_x and f^y are measurable on their respective domains, we analyze the pre-image of a measurable target set $B \subset \mathbb{C}$.

Fix any $x \in X$. By definition, the section function satisfies $f_x(y) = f(x, y)$. The pre-image of B under f_x is calculated as:

$$(f_x)^{-1}(B) = \{y \in Y : f_x(y) \in B\} = \{y \in Y : f(x, y) \in B\}$$

Now consider the global pre-image of B under the original function f, which we will denote as the full product set $E \subset X \times Y$:

$$E = f^{-1}(B) = \{(x', y') \in X \times Y : f(x', y') \in B\}$$

Since f is a measurable function on the product space, the set E belongs to $\mathcal{M} \otimes \mathcal{N}$ by definition of measurability.

Taking the x-section of this global pre-image set E yields:

$$E_x = \{y \in Y : (x, y) \in E\} = \{y \in Y : f(x, y) \in B\}$$

Comparing our expressions, we discover the structural identity:

$$(f_x)^{-1}(B) = (f^{-1}(B))_x$$

By Part (a), since $f^{-1}(B) \in \mathcal{M} \otimes \mathcal{N}$, its x-section must belong to the component σ-algebra $\mathcal{N}$. Therefore, $(f_x)^{-1}(B) \in \mathcal{N}$, proving that f_x is a $\mathcal{N}$-measurable function for every $x \in X$.

By a completely symmetrical computation for the y-section, we find:

$$(f^y)^{-1}(B) = (f^{-1}(B))^y$$

Since $f^{-1}(B) \in \mathcal{M} \otimes \mathcal{N}$, Part (a) guarantees that its y-section belongs to $\mathcal{M}$. This ensures that $(f^y)^{-1}(B) \in \mathcal{M}$, confirming that f^y is a $\mathcal{M}$-measurable function for every $y \in Y$. □

Theorem 25 (The Fubini-Tonelli Theorem)**.** *Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces.*

a. (Tonelli) *If $f \in L^+(X \times Y)$ is a non-negative measurable function, then the section functions defined by $g(x) = \int_Y f_x \, d\nu$ and $h(y) = \int_X f^y \, d\mu$ are in $L^+(X)$ and $L^+(Y)$, respectively, and the iterative integral equality holds:*

$$\begin{aligned} \int_{X \times Y} f \, d(\mu \times \nu) &= \int_X \left[\int_Y f(x, y) \, d\nu(y)\right] d\mu(x) \\ &= \int_Y \left[\int_X f(x, y) \, d\mu(x)\right] d\nu(y) \end{aligned}$$

b. (Fubini) *If $f \in L^1(\mu \times \nu)$ is an integrable function, then the section function f_x is in $L^1(\nu)$ for almost every (a.e.) $x \in X$, and the section function f^y is in $L^1(\mu)$ for almost every (a.e.) $y \in Y$. Moreover, the almost everywhere defined integral functions given by $g(x) = \int_Y f_x \, d\nu$ and $h(x) = \int_Y f^y \, d\nu$ belong to $L^1(\mu)$ and $L^1(\nu)$, respectively, and the identical iterative integral relations expressed in equation in (a) hold.*

Theorem 26 (The Fubini-Tonelli Theorem for Complete Measures)**.** *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be complete, σ-finite measure spaces, and let $(X \times Y, \mathcal{L}, \lambda)$ be the completion of the product space $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. If f is a $\mathcal{L}$-measurable function on $X \times Y$ and satisfies either:*

- ***(a)*** *$f \geq 0$ is a non-negative function, or*
- ***(b)*** *$f \in L^1(\lambda)$ is an integrable function,*

*then the x-section f_x is $\mathcal{N}$-measurable for almost every (a.e.) $x \in X$, and the y-section f^y is $\mathcal{M}$-measurable for almost every (a.e.) $y \in Y$. In case **(b)**, f_x and f^y are also integrable for a.e. x and y, respectively. Moreover, the coordinate mapping functions $x \mapsto \int_Y f_x \, d\nu$ and $y \mapsto \int_X f^y \, d\mu$ are measurable (and also integrable in case **(b)**), and the iterated integrals satisfy:*

$$\int_{X \times Y} f \, d\lambda = \int_X \left[\int_Y f(x,y) \, d\nu(y) \right] d\mu(x) = \int_Y \left[\int_X f(x,y) \, d\mu(x) \right] d\nu(y)$$

Remark. *This complete measure variation follows as a direct analytical corollary of the standard Fubini-Tonelli theorem established on the pre-completion space $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$.*

By the fundamental properties of completion spaces, any function f that is measurable with respect to the completed σ-algebra $\mathcal{L}$ can be uniquely written as a combination:

$$f = g + h$$

where g is a structurally clean, $\mathcal{M} \otimes \mathcal{N}$-measurable function, and h is a null function that vanishes almost everywhere with respect to the product measure $\mu \times \nu$ (meaning it is supported entirely within a set of product outer measure zero).

Applying the standard theorem to g immediately yields the desired measurability, integrability, and iterated integration equality properties. Meanwhile, the null remainder component h behaves perfectly under slicing because, by the standard Fubini-Tonelli theorem applied to its non-negative envelope, the cross-sections of a product null set are themselves null sets in the component spaces for almost every coordinate. Consequently, the contribution of h vanishes identically inside both the global and iterated integral channels, directly proving the theorem as a straightforward corollary.

Remark (Cardinality, Null Sets, and the Pathology of Projections)**.** *To conclude this chapter, we reflect on the structural boundaries of σ-algebras, their size, and how standard operations like projections can break measurability. These counterexamples highlight why completing a measure space is practically necessary.*

1. Cardinality of σ-Algebras and Trivial Non-Measurability

Let us contrast the sizes of the Borel and Lebesgue σ-algebras on $\mathbb{R}$:

- *The Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$ is generated by the countable collection of open intervals with rational endpoints. By transfinite induction up to the first uncountable ordinal ω_1, it can be shown that the cardinality of $\mathcal{B}_{\mathbb{R}}$ is exactly the cardinality of the continuum:*

$$|\mathcal{B}_{\mathbb{R}}| = \mathfrak{c} = 2^{\aleph_0}$$

- *On the other hand, the Lebesgue σ-algebra $\mathcal{L}_{\mathbb{R}}$ is the completion of $\mathcal{B}_{\mathbb{R}}$. It contains the Cantor ternary set C, which is closed (hence Borel) and has Lebesgue measure zero ($m(C) = 0$). Since C is uncountable, its cardinality is $|C| = \mathfrak{c}$.*
- *Because $\mathcal{L}_{\mathbb{R}}$ contains every subset of a null set, the power set $\mathcal{P}(C)$ is entirely contained inside $\mathcal{L}_{\mathbb{R}}$. The cardinality of $\mathcal{L}_{\mathbb{R}}$ must therefore be at least the cardinality of $\mathcal{P}(C)$:*

$$|\mathcal{L}_{\mathbb{R}}| \geq |\mathcal{P}(C)| = 2^{\mathfrak{c}} > \mathfrak{c}$$

*This cardinal mismatch yields an important insight: **the existence of non-Borel sets is mathematically trivial.** Because $|\mathcal{B}_{\mathbb{R}}| = \mathfrak{c}$ while the total number of subsets of $\mathbb{R}$ is $2^{\mathfrak{c}}$, the overwhelming majority of subsets of $\mathbb{R}$ are not Borel.*

*By comparison, finding a non-Lebesgue measurable set is much more difficult. It cannot be done by a cardinality argument alone, because $|\mathcal{L}_{\mathbb{R}}| = |\mathcal{P}(\mathbb{R})| = 2^{\mathfrak{c}}$. To construct a non-Lebesgue measurable set, we must rely on the Axiom of Choice to produce a translation-invariant anomaly, such as the **Vitali Set**, which is not even Lebesgue measurable.*

2. Non-Borel Subsets of Borel Null Sets

Because $|\mathcal{B}_{\mathbb{R}}| = \mathfrak{c}$ and $|\mathcal{P}(C)| = 2^{\mathfrak{c}}$, it is mathematically guaranteed that there must exist subsets of the Cantor set C that are not Borel sets. Let $E \subset C$ be one such non-Borel subset ($E \notin \mathcal{B}_{\mathbb{R}}$).

*Since E is contained within the Cantor set ($E \subset C$) and $m(C) = 0$, the Lebesgue outer measure of E is identically zero ($m^*(E) = 0$). In the Lebesgue framework, E is a Lebesgue measurable null set ($E \in \mathcal{L}_{\mathbb{R}}$). However, in the Borel framework, E is a non-measurable subset of a Borel null set. This shows that **$\mathcal{B}_{\mathbb{R}}$ is not complete**, as it contains null sets with subsets that fall outside the σ-algebra.*

3. The Breakdown of Projections under $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$

Consider the product space $\mathbb{R}^2$ equipped with the product Borel σ-algebra $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^2}$. Take our non-Borel set $E \subset C$ from above, and embed it into the horizontal axis of $\mathbb{R}^2$ by defining:

$$K = E \times \{0\} \subset \mathbb{R}^2$$

Let us look at how this set behaves under the product structure:

- *The set K is contained within the larger rectangle $C \times \{0\}$.*
- *Since $C \in \mathcal{B}_{\mathbb{R}}$ and $\{0\} \in \mathcal{B}_{\mathbb{R}}$, their product $C \times \{0\}$ is a Borel set in $\mathbb{R}^2$, and its product Borel measure is $m(C) \times m(\{0\}) = 0 \times 0 = 0$.*
- *Because K is a subset of this Borel null set, K is automatically a Lebesgue measurable set in the completed space $(\mathbb{R}^2, \mathcal{L}_{\mathbb{R}^2}, \bar{m})$. In fact, its completed Lebesgue measure is zero.*
- *However, because $E \notin \mathcal{B}_{\mathbb{R}}$, the set K itself does not belong to the uncompleted product σ-algebra $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}}$.*

Now, let $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ be the projection map onto the first coordinate, defined by $\pi_1(x, y) = x$. When we apply this projection map to our set K, we get:

$$\pi_1(K) = \pi_1(E \times \{0\}) = E$$

*This reveals the structural mismatch. The set K belongs to the completion of the product σ-algebra ($\mathcal{L}_{\mathbb{R}^2}$) because it lives inside a product null rectangle. Yet, its linear projection $\pi_1(K) = E$ is **not a Borel set** in $\mathbb{R}$. Thus, passing a set through the completion and then projecting it can map it completely outside the original component σ-algebra.*

4. Projections of Lebesgue Measurable Sets Need Not Be Lebesgue Measurable

We can take this pathological behavior a step further to show that projections can fail even within the Lebesgue framework.

Let $F \subset \mathbb{R}$ be a truly non-Lebesgue measurable set (for example, the Vitali Set, so $F \notin \mathcal{L}_{\mathbb{R}}$). Let us embed F into the plane along the horizontal axis by defining:

$$W = F \times \{0\} \subset \mathbb{R}^2$$

Just like before, W is a subset of the product rectangle $\mathbb{R} \times \{0\}$. While $\mathbb{R}$ has infinite measure, the single point $\{0\}$ has a Lebesgue measure of zero. The standard product premeasure on rectangles dictates:

$$(m \times m)(\mathbb{R} \times \{0\}) = m(\mathbb{R}) \cdot m(\{0\}) = \infty \cdot 0 = 0$$

Because W is a subset of a set with product measure zero, it has a Lebesgue outer measure of zero in $\mathbb{R}^2$. This means W is a Lebesgue measurable set in the completed plane space ($W \in \mathcal{L}_{\mathbb{R}^2}$).

However, if we project W onto the first coordinate, we get:

$$\pi_1(W) = \pi_1(F \times \{0\}) = F$$

Remember that F was chosen specifically to be non-Lebesgue measurable ($F \notin \mathcal{L}_{\mathbb{R}}$).

*This gives us a major counterexample to keep in mind: **The projection of a Lebesgue measurable set from $\mathbb{R}^2$ onto $\mathbb{R}$ does not have to be Lebesgue measurable.** Even though W is a null set in the plane, its projection stretches out across the axis in a way that completely destroys its measurability. This shows that geometric projections do not preserve measurability.*