NOTES ON MEASURE THEORY

Abstract. Notes on the Measure Theory course taken at IITM by Prof. S. Kesavan

1. Riemann Integration

Let $[a,b]$ be a closed interval in $\mathbb{R}$. Let $\mathcal{R}$ be the class of Riemann integrable functions on $[a,b]$.

Consider a partition of the interval $P = \{a = x_1, x_2, \cdots, x_n = b\}$. Let $t_i \in [x_{i-1}, x_i]$ and $S(P,f) = \sum_{i=1}^{n} f(t_i)\Delta x_i$. Let $\mu(P) = \max_{i\in\{i,\cdots,n\}} \Delta x_i$.

We say that $\lim_{\mu(P)\to 0} S(P,f) = A$ if $\forall \epsilon > 0, \exists \delta > 0$ such that $\forall P$ such that $\mu(P) < \delta$, and for **all choices** of $t_i \in [x_{i-1}, x_i], 1 \le i \le n$, $|S(P,f) - A| < \epsilon$.

Theorem 1.1. *(1) If $\lim_{\mu(P)\to 0} S(P,f) = A$, then $f \in \mathcal{R}$ and $\int_a^b f\,dx = A$.*

(2) If f is continuous, then $f \in \mathcal{R}$.

(3) If $f \in \mathcal{R}$, then $\lim_{\mu(P)\to 0} S(P,f) = \int_a^b f\,dx$.

Proof. (1) Let $\epsilon > 0$. Then $\exists \delta > 0$ such that

$$\mu(P) < \delta \implies A - \frac{\epsilon}{2} < S(P,f) < A + \frac{\epsilon}{2}.$$

Choose such a partition. Letting t_i vary, moving towards the points of infimum and supremum, we get

$$A - \frac{\epsilon}{2} \le L(P,f) \le U(P,f) \le A + \frac{\epsilon}{2}$$
$$\implies U - L \le \epsilon \implies f \in \mathcal{R}.$$

Also, $L(P,f) \le \int_a^b f\,dx \le U(P,f) \implies \int_a^b f\,dx = A$.

(2) If f is continuous on $[a,b]$, then f is uniformaly continuous on $[a,b]$. Hence for $\epsilon > 0, \exists \delta > 0$ such that $|x-y| < \delta \implies |f(x) - f(y)| < \epsilon$. Hence, if $\mu(P) < \delta$, then $M_i - m_i < \epsilon$.

$$\implies U(P,f) - L(P,f) < \epsilon(b-a).$$

(3) Let $\epsilon > 0$ and $M = \sup_{x\in[a,b]} |f(x)|$. $f \in \mathcal{R} \implies \exists P*$ such that $U(P*,f) < \int_a^b f\,dx + \frac{\epsilon}{4}$. Assume that $P*$ has n subintervals and choose $\delta_1 < \frac{\epsilon}{4Mn}$. Choose P such that $\mu(P) < \delta_1$. Now $U(P,f)$ can be written as two sums, one that depends on the contributions of subintervals containing nodes of $P*$ and those that do not contain nodes of $P*$. The first sum is less than $\frac{M\epsilon(n-1)}{4Mn}$, while the second sum is less that $U(P*,f)$. Hence $U(P,f) < \frac{\epsilon}{2} + \int_a^b f\,dx$. Similarly, find $\delta_2 > 0$ such that $\mu(P) < \delta_2 \implies L(P,f) >$

$\int_a^b f\,dx - \frac{\epsilon}{2}$. Let $\delta = \min(\delta_1, \delta_2)$. $\implies \int_a^b f\,dx - \frac{\epsilon}{2} < L \leq S \leq U < \int_a^b f\,dx + \frac{\epsilon}{2}$.

□

Example 1.2. Let $r_1, r_2, \cdots, r_n, \cdots$ be an enumeration of the rationals in $[a, b]$. Let

$$f_n(x) = \begin{cases} 1, & x = x_1, \cdots, x_n, \\ 0, & \text{otherwise}. \end{cases}$$

Then for a partition P with n subintervals, $L(P, f) = 0$ and $U(P, f) < n\mu(P)$. Taking $\mu(P)$ as small as required, we get that $f_n \in \mathcal{R}$ and $\int_a^b f_n\,dx = 0$.

Example 1.3. Let

$$f(x) = \begin{cases} 1, & x \text{ is rational}, \\ 0, & \text{otherwise}. \end{cases}$$

Then for every partition P, $L(P, f) = 0$ and $U(P, f) = b - a \neq 0$. Hence $f \notin \mathcal{R}$. We observe that the f_n of the previous example converge pointwise to f.

This is why the need for a different kind fo integration arises. Consider a step function $f = \sum_{i=1}^n \alpha_i \chi_{I_i}$, where I_i are intervals. Then we define $\int_a^b f = \sum_{i=1}^n \alpha_i l(I_i)$.

We note that the integral is also equal to

$$\sum_{i=1}^n \alpha_i(\text{ total length of intervals where } f = \alpha_i).$$

Next, we ask whether the intervals I_i can be replaced by arbitrary sets E_i. How do we talk about the 'length' of an arbitrary set? Hence, the need for measures arises.

2. Measures

Let X be a non-empty set and $\mathcal{P}(X)$ denote the power set of X.

Definition 2.1. A **ring** $\mathcal{R}$ is a subset of $\mathcal{P}(X)$ which is closed under unions and differences. That is,

$$E, F \in \mathcal{R} \implies E \cup F, E \setminus F \in \mathcal{R}.$$

Remark 2.2. Let $\mathcal{R}$ be a ring. Then

(1) $\emptyset \in \mathcal{R}$.
(2) $E, F \in R \implies E \cap F = E \setminus (E \setminus F) \in \mathcal{R}$.
(3) $E \triangle F = (E \setminus F) \cup (F \setminus E) \in \mathcal{R}$.
Hence a ring contains the empty set, is closed under intersections and symmetric differences. It is also clear that a ring is closed under finite unions and intersections.
(4) If $\mathcal{R}$ is closed under unions and finite differences, then it is a ring since $E \setminus F = (E \cup F) \triangle F$.
(5) If $\mathcal{R}$ is closed under intersections and finite differences, then it is a ring since $E \cup F = (E \triangle F) \triangle (E \cap F)$.

Example 2.3. (1) $\mathcal{P}(X)$ is trivially a ring.
(2) $\mathcal{R} = \{\emptyset\}$ is a ring.
(3) Let $X = \mathbb{Z}$ and $\mathcal{R} = \{A \subset \mathbb{Z} : A \text{ is finite or } \phi\}$.
(4) Let $X = \mathbb{R}$ and $\mathcal{P} = \{[a, b) : a, b \in \mathbb{R}, a \leq b\}$ and $\mathcal{R}$ = set of finite unions of elements of $\mathcal{P}$.

Definition 2.4. A ring $\mathcal{R}$ is said to be an algebra if $X \in \mathcal{R}$.

Hence an algebra is closed under unions and complementations. Conversely, if $\mathcal{R}$ is closed under unions and complementations, it is an algebra, since

$$X = E \cup E^c, E \setminus F = E \cap F^c = (E^c \cup F^c)^c.$$

Let $\mathcal{E} \subseteq \mathcal{P}(X)$. Then $\mathcal{P}(X)$ is a ring containing $\mathcal{E}$. The intersection of rings is again a ring. Hence, the intersection of all rings containing $\mathcal{E}$ is the smallest ring containing $\mathcal{E}$ and is called the ring generated by $\mathcal{E}$. We denote it by $\mathcal{R}(\mathcal{E})$

Let $\mathcal{R}' = \{\text{all subsets which covered by a finite number of members of } \mathcal{E}\}$. Then $\mathcal{E} \subseteq \mathcal{R}'$ and $E_1, E_2 \in \mathcal{R}' \implies E_1 \cup E_2, E_1 \setminus E_2 \in \mathcal{R}'$. Hence $\mathcal{R}'$ is a ring containing $\mathcal{E}$. Hence, $\mathcal{R}(\mathcal{E}) \subseteq \mathcal{R}'$. This shows that every member of the ring generated by $\mathcal{E}$ can be covered by finitely many members of $\mathcal{E}$.

Definition 2.5. A σ-ring $\mathcal{S}$ is a collection of subsets of X which is closed under differences and countable unions. That is,

(1) $E, F \in \mathcal{S} \implies E \setminus F \in \mathcal{S}$.
(2) $\{E_i\}_{i=1}^{\infty} \subseteq \mathcal{S} \implies \cup_{i=1}^{\infty} E_i \in \mathcal{S}$.

Let $E = \cup_{i=1}^{\infty} E_i$. Then $\cap_{i=1}^{\infty} = E \setminus (\cup_{i=1}^{\infty}(E \setminus E_i)) \in \mathcal{S}$.

As before, for $\mathcal{E} \subseteq \mathcal{P}(X)$, one can talk of the σ-ring generated by $\mathcal{E}$ and denote it by $\mathcal{S}(\mathcal{E})$.

Definition 2.6. $\mathcal{S}$ is called a σ-algebra if it si a σ-ring such that $X \in \mathcal{S}$.

The above definition can be shown to be equivalent to $\mathcal{S}$ being closed under countable unions and complementations.

Let (X, τ) be a topological space. Then $\mathcal{S}(\tau)$, the σ-ring generated by τ is actually a σ-algebra and is called the Borel σ-algebra generated by (X, τ).

Let X be a non-empty set and $\mathcal{R}$ a ring on X.

Definition 2.7. A measure μ on $\mathcal{R}$ is an extended real-valued function such that:

(1) $\mu(E) \geq 0 \, \forall E \in \mathcal{R}$
(2) $\mu(\emptyset) = 0$
(3) μ satisfies countable additivity, that is, if $\{E_i\}_{i-1}^{\infty}$ is a collection of mutually disjoint sets in $\mathcal{R}$ and if $E = \cup_{i=1}^{\infty} E_i \in \mathcal{R}$, then $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$.

Here, we need not worry about convergence of the series as μ is an **extended** real valued function, and rearrangements do not affect the summation since each term is positive.

Remark 2.8. If we assume that there exists at least one $E \in \mathcal{R}$ such that $\mu(E) < \infty$, then (2) follows from (1) and (3), for

$$E = E \cup \emptyset \cup \emptyset \cup \cdots$$

and hence

$$\mu(E) = \mu(E) + \mu(\emptyset) + \mu(\emptyset) + \cdots$$

giving $\mu(\emptyset) = 0$ since $\mu(E)$ is finite.

Example 2.9. (1) Let $X \neq \emptyset$ and $\mathcal{R} = \mathcal{P}(X)$. For $E \subseteq X$, define

$$\mu(E) = \begin{cases} \text{number of elements in } E, & \text{if } E \text{ is finite,} \\ \infty, & \text{otherwise.} \end{cases}$$

(2) **Dirac** measure: $X \neq \emptyset, \mathcal{R} = \mathcal{P}(X)$. Fix $x_0 \in X$. Let

$$\delta_{x_0}(E) = \begin{cases} 1, & \text{if } x_0 \in E, \\ 0, & \text{otherwise.} \end{cases}$$

(3) $X \neq \emptyset$, $\mathcal{R}$=the ring of finite subsets of X. Let $f : X \to \mathbb{R}$ be a non-negative function. Let $E = \{x_1, x_2, \cdots, x_n\}$. Define $\mu(E) = \sum_{i=1}^{n} f(x_n)$.

Let X be a non-empty set, $\mathcal{R}$ a ring and μ a measure on $\mathcal{R}$.

Proposition 2.10. *(1) μ is monotone, i.e.,*

$$E \subset F, E, F \in \mathcal{R} \implies \mu(E) \leq \mu(F).$$

(2) μ is subtractive, i.e.,

$$\mu(F \setminus E) = \mu(F) - \mu(E) \quad \forall E \subset F,\, E, F \in \mathcal{R},\, \mu(E) < \infty.$$

Proof. (1) If $E \subset F$, then $F = E \cup (F \setminus E)$. Hence $\mu(F) = \mu(E) + \mu(F \setminus E)$. Hene $\mu(E) \leq \mu(F)$.

(2) Further, if $\mu(E) < \infty$, then $\mu(F \setminus E) = \mu(F) - \mu(E)$.

□

Proposition 2.11. *Let $E, E_i \in \mathcal{R}, \cup_{i=1}^{\infty} E_i \subset E$ and E_i disjoint. Then*

$$\mu(E) \geq \sum_{i=1}^{\infty} \mu(E_i).$$

Proof. $\forall n, \cup_{i=1}^{n} E_i \subset E$. Hence $\mu(E) \geq \mu(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} \mu(E_i)\, \forall n$. Hence $\mu(E) \geq \sum_{i=1}^{\infty} \mu(E_i)$. □

Proposition 2.12 (Continuity from below)**.** *Let $E_i \in \mathcal{R}, \{E_i\}$ be an increasing sequence. Let $E = \cup_{i=1}^{\infty} E_i \in \mathcal{R}$. Then $\mu(E) = \lim_{i \to \infty} \mu(E_i)$.*

Proof. $\mu(\cup_{i=1}^{\infty}) = \mu(\cup_{i=1}^{\infty}(E_i \setminus E_{i-1})) = \sum_{i=1}^{\infty} \mu(E_i \setminus E_{i-1}) = \lim_{n \to \infty} \sum_{i=1}^{n} \mu(E_i \setminus E_{i-1}) = \lim_{n \to \infty} \mu(\cup_{i=1}^{n}(E_i \setminus E_{i-1})) = \lim_{n \to \infty} \mu(E_n)$. □

Proposition 2.13. *Let $\{E_i\}$ be a decreasing sequence in $\mathcal{R}$, $E = \cap_{i=1}^{\infty} E_i \in \mathcal{R}$. Suppose $\exists m \in \mathbb{N}$ such that $\mu(E_m) < \infty$. Then $\mu(E) = \lim_{n \to \infty} \mu(E_n)$.*

Proof.

$$\mu(E_m) < \infty \implies \mu(E_n) < \infty \, \forall n \geq m.$$

$$\mu(E_m) - \mu(E) = \mu(E_m) - \mu(\cap_{n\geq m} E_n) = \mu(E_m \setminus \cap_{n\geq m} E_n)$$
$$= \mu(\cup_{n\geq m}(E_m \setminus E_n)) = \lim_{n\to\infty} \mu(E_m \setminus E_n)$$
$$= \lim_{n\to\infty}(\mu(E_m) - \mu(E_n)) = \mu(E_m) - \lim_{n\to\infty} \mu(E_n).$$

Hence $\mu(E) = \lim_{n\to\infty} \mu(E_n)$. □

Example 2.14. Let $X = \mathbb{N}$, $\mathcal{R} = \mathcal{P}(\mathbb{N})$ and μ be the counting measure. Let $E_n = \{n, n+1, \cdots\}$. Then $\mu(E_n) = \infty \forall n$, $\cap_{n=1}^{\infty}(E_n) = \emptyset$ and $\mu(\emptyset) = 0$. Hence the condition that $\mu(E_m) < \infty$ for some $m \in \mathbb{N}$ cannot be removed.

Definition 2.15. Let $X \neq \emptyset$, $\mathcal{R}$ be a ring and μ a measure. $\mathcal{H}(\mathcal{R})$ is defined as the smallest hereditary σ-ring containing $\mathcal{R}$. That is, $\mathcal{H}(\mathcal{R})$ is a σ-ring containing $\mathcal{R}$, and if $E \in \mathcal{H}(\mathcal{R}), F \subset E$, then $F \in \mathcal{H}(\mathcal{R})$.

Let $\mathcal{E}$ be any collection of sets and $\mathcal{S}(\mathcal{E})$ be the smallest σ-ring containing $\mathcal{E}$. Let $f \in \mathcal{S}(\mathcal{E})$. Then $\exists \{E_i\}_{i=1}^{\infty}$ in $\mathcal{E}$ such that $F \subset \cup_{i=1}^{\infty} E_i$. Let $\mathcal{H}(\mathcal{R})$ be the collection of all sets which can be covered by a countable number of elements of $\mathcal{R}$. Then $\mathcal{H}(\mathcal{R})$ is hereditary.

Definition 2.16 (Outer measure)**.** An outer measure μ^* on $\mathcal{H}$ is an extended real valued function $\mu^* : \mathcal{H} \to \mathbb{R} \cup \{\infty\}$ such that

(1) $\mu^*(E) \geq 0 \, \forall E \in \mathcal{H}$.
(2) $\mu^*(\emptyset) = 0$.
(3) μ^* is monotone, i.e., $E \subset F, E, F \in \mathcal{H} \implies \mu^*(E) \leq \mu^*(F)$.
(4) μ^* is countably additive, i.e. if $E = \cup_{i=1}^{\infty} E_i, E, E_i \in \mathcal{H}$, then $\mu^*(E) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$.

Proposition 2.17. *Let $X \neq \emptyset, \mathcal{R}$ be a ring and μ a measure. Let $E \in \mathcal{H}$. Then we know that $\exists \{E_i\}, E_i \in \mathcal{R}$ such that $E \subset \cup_{i=1}^{\infty} E_i$. Define*

$$\mu^*(E) = inf \, \{\sum_{i=1}^{\infty} \mu(E_i) | E \subset \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{R}\}$$

Then μ^* is an outer measure on $\mathcal{H}(\mathcal{R})$ which extends μ, i.e. if $E \in \mathcal{R}$, then $\mu(E) = \mu^*(E)$.

Proof. Let $E \in \mathcal{R}$. Now, $E \subset E$, hence $\mu^*(E) \leq \mu(E)$. Let $E \subset \cup_{i=1}^{\infty} E_i$. Then $\mu(E) \leq \sum_{i=1}^{\infty} \mu(E_i)$. This is true for every cover $\{E_i\}$ of E and hence $\mu(E) \leq \mu^*(E)$. In particular, $\mu^*(\emptyset) = 0$. Clearly, $\mu^* \geq 0$.

Let $E \subset F$. Then every cover of F is also a cover of E. Hence $\mu^*(E) \leq \mu^*(F)$.

Let $E, E_i \in \mathcal{H}(\mathcal{R}), E = \cup_{i=1}^{\infty} E_i$. Assume $\exists i$ such that $\mu^*(E_i) = \infty$. Then $\sum_{i=1}^{\infty} \mu^*(E_i) = \infty \geq \mu^*(E)$. So assume that $\mu^*(E_i) < \infty \forall i$. Let $\epsilon > 0$. Then $\exists \{E_{ij}\}_{j=1}^{\infty}$ in $\mathcal{R}$ such that $E_i \subset \cup_{i=1}^{\infty} E_{ij}$ and $\sum_{i=1}^{\infty} \mu(E_{ij}) < \mu^*(E_i) + \frac{\epsilon}{2^i}$. This is true since $\mu(E_i) < \infty$. Hence

$$E \subset \cup_{i=1}^{\infty} \cup_{j=1}^{\infty} E_{ij}$$

$$\implies \mu^*(E) \leq \sum_{i,j=1}^{\infty} \mu(E_{ij}) \leq \sum_{i=1}^{\infty} (\sum_{j=1}^{\infty} \mu(E_{ij}))$$

$$\leq \sum_{i=1}^{\infty} \mu^*(E_i) + \epsilon.$$

Hence $\mu^*(E) \leq \sum_{i=1}^{\infty} \mu^*(E_i)$.

□

Example 2.18. Let $X = \mathbb{N}, \mathcal{R}$ be the set of all finite subsets in$\mathbb{N}$ and μ be the counting measure. Then $\mathcal{H}(\mathcal{R}) = \mathcal{P}(\mathbb{N})$. $\mu^*(E) = \infty$ if E is infinite. So the outer measure may be infinite even though the measure is finite.

Definition 2.19. A measure μ (or an outer measure μ^*) is called finite if $\mu(E) < \infty$ $(\mu^*(e) < \infty), \forall E \in \mathcal{R}$.

A measure μ (or an outer measure μ^*) is called σ-finite if $\forall E \in \mathcal{R}$, $\exists \{E_i\} \subset \mathcal{R}$ such that $E \subseteq \cup_{i=1}^{\infty} E_i$ and $\mu(E_i) < \infty \, \forall i$ (respectively $\mu^*(E_i) < \infty \, \forall i$).

Proposition 2.20. *If μ is σ-finite on $\mathcal{R}$, then μ^* is also sigma-finite on $\mathcal{H}(\mathcal{R})$.*

Proof. Let $E \in \mathcal{H}(\mathcal{R})$. $\exists E_i \in \mathcal{R}$ such that $E \subset \cup_{i=1}^{\infty} E_i$. Since μ is σ-finite, $E_i \subset \cup_{j=1}^{\infty} E_{ij}$, with $E_{ij} \in \mathcal{R}$ and $\mu(E_{ij}) < \infty$. Hence $E \subset \cup_{i=1}^{\infty} \cup_{j=1}^{\infty} E_{ij}$ and $\mu^*(E_{ij}) = \mu(E_{ij}) < \infty$. □

Definition 2.21. Let $\mathcal{H}$ be any hereditary σ-ring and μ^* be an outer measure on $\mathcal{H}$. A set E is said to be μ^*-measurable if $\forall A \in \mathcal{H}$, we have

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

We want to obtain a measure from the outer measure μ^*, hence we need to build our way towards countable additivity.

Let $\bar{\mathcal{S}}$ be the collection of all μ^*-measurable sets in $\mathcal{H}$.

Proposition 2.22. *Let $\mathcal{H}$ be a hereditary σ-ring, μ^* an outer measure on $\mathcal{H}$ and $\bar{\mathcal{S}}$ the collection of μ^*-measurable sets in $\mathcal{H}$. Then $\bar{\mathcal{S}}$ is a ring.*

Proof. Let $E, F \in \bar{\mathcal{S}}$. We must show the following:

(1) $E \cup F \in \bar{\mathcal{S}}$
(2) $E \setminus F \in \bar{\mathcal{S}}$.

Let $A \in \mathcal{H}$ be arbitrary.

$$E \in \bar{\mathcal{S}} \implies \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$
$$F \in \bar{\mathcal{S}} \implies \mu^*(A \cap E) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^c)$$

and

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

So

$$\mu^*(A) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^c) + \mu^*(A \cap E^c \cap F) + \mu^*(A \cap E^c \cap F^c). \quad (1)$$

Replacing A by $A \cap (A \cap (E \cup F))$, we get:

$$\mu^*(A \cap (E \cup F)) = \mu^*((A \cap (E \cup F) \cap E \cap F) + \mu^*((A \cap (E \cup F) \cap E \cap F^c)$$
$$+ \mu^*((A \cap (E \cup F) \cap E^c \cap F) + \mu^*((A \cap (E \cup F) \cap E^c \cap F^c).$$

This reduces to:

$$\mu^*(A \cap (E \cup F)) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^c) + \mu^*(A \cap E^c \cap F). \quad (2)$$

Hence (1) becomes:

$$\mu^*(A) = \mu^*(A \cap (E \cup F)) + \mu^*(A \cap (E \cup F)^c).$$

So $E \cup F \in \bar{\mathcal{S}}$.

Next, we replace A in (1) by $A \cap (E \setminus F)^c = A \cap (E \cap F^c)^c = A \cap (E^c \cup F)$ to get

$$\mu^*(A \cap (E \setminus F)^c) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E^c \cap F) + \mu^*(A \cap E^c \cap F^c).$$

Substituting in (1), we get

$$\mu^*(A) = \mu^*(A \cap (E \setminus F)^c) + \mu^*(A \cap (E \setminus F)).$$

So $E \setminus F \in \bar{\mathcal{S}}$ and hence $\bar{\mathcal{S}}$ is a ring.

□

Proposition 2.23. *$\bar{\mathcal{S}}$ is a σ-ring. Further, if $\{E_i\}_{i=1}^{\infty}$ is a disjoint sequence and $E = \cup_{i=1}^{\infty} E_i$, then*

$$\mu^*(A \cap E) = \sum_{i=1}^{\infty} \mu^*(A \cap E_i) \, \forall A \in \mathcal{H}. \tag{3}$$

Proof. Consider E_1 and E_2. Since $E_1 \cap E_2 = \emptyset, E_1 \subseteq E_2^c$ and $E_2 \subseteq E_1^c$. Taking $E = E_1$ and $F = E_2$ in (2), we get:

$$\mu^*(A \cap (E_1 \cup E_2)) = \mu^*(A \cap E_1) + \mu^*(A \cap E_2).$$

By induction, we get:

$$\mu^*(A \cap (\cup_{i=1}^{n} E_i)) = \sum_{i=1}^{n} \mu^*(A \cap E_i).$$

Let $F_n = \cup_{i=1}^{n} E_i$. $F_n \in \bar{\mathcal{S}}$ since $\bar{\mathcal{S}}$ is a ring.

Let $A \in \mathcal{H}$ be arbitrary. Then

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap F_n) + \mu^*(A \cap F_n^c) \\ &= \sum_{i=1}^{n} \mu^*(A \cap E_i) + \mu^*(A \cap F_n^c) \\ &\geq \sum_{i=1}^{n} \mu^*(A \cap E_i) + \mu^*(A \cap E^c) \text{ by monotonicity.}\end{aligned} \tag{4}$$

The above is true for every $n \in \mathbb{N}$. Hence

$$\mu^*(A) \geq \sum_{i=1}^{\infty} \mu^*(A \cap E_i) + \mu^*(A \cap E^c). \tag{5}$$

Now replace A by $A \cap E$ in (5) to get:

$$\mu^*(A \cap E) \geq \sum_{i=1}^{\infty} \mu^*(A \cap E_i).$$

But by subadditivity, we already have

$$\mu^*(A \cap E) \leq \sum_{i=1}^{\infty} \mu^*(A \cap E_i).$$

Now, substituting in (4), we get

$$\mu^*(A) \geq \mu^*(A \cap E).$$

By subadditivty,

$$\mu^*(A) \le \mu^*(A \cap E).$$

Hence $E \in \bar{\mathcal{S}}$. Thus, $\bar{\mathcal{S}}$ is closed under countable disjoint unions. But this is sufficient to show that it is a σ-ring, since it is already known to be a ring. $\square$

Theorem 2.24 (Caratheodory Extension)**.** *Let $\mathcal{H}$ be a hereditary σ-ring and μ^* be an outer measure on $\mathcal{H}$. Let $\bar{\mathcal{S}}$ be the σ-ring of μ^* measurable sets. Define for $E \in \bar{\mathcal{S}}, \bar{\mu}(E) = \mu^*(E)$. Then $\bar{\mu}$ is a measure on $\bar{\mathcal{S}}$, which is complete. That is, if $\bar{\mu}(E) = 0$ and $F \subseteq E$, then $F \in \bar{\mathcal{S}}$ and $\bar{\mu}(F) = 0$.*

Proof. Let $\{E_i\}_{i=1}^{\infty}$ be a disjoint sequence in $\bar{\mathcal{S}}, E = \cup_{i=1}^{\infty} E_i$. Taking $A = E$ in (3), we get

$$\bar{\mu}(E) = \mu^*(E) = \sum_{i=1}^{\infty} u^*(E_i) = \sum_{i=1}^{\infty} \bar{u}(E_i).$$

So $\bar{\mu}$ is a measure on $\bar{\mathcal{S}}$. Next, let $\mu^*(E) = 0$ for some $E \in \mathcal{H}$. Let $A \in \mathcal{H}$ be arbitrary. Then

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

But already, $\mu^*(A) \le \mu^*(A \cap E^c) + \mu^*(A \cap E)$. Hence $E \in \bar{\mathcal{S}}$. By monotonicity, if $F \subseteq E$, $\mu^*(F) = 0$ and so $\bar{\mu}$ is complete. $\square$

Theorem 2.25. *Let $\mathcal{R}$ be a ring, μ a measure on $\mathcal{R}$. Let $\mathcal{H}(\mathcal{R})$ be the hereditary σ-ring generated by $\mathcal{R}$. Let μ^* be the canonical outer measure on $\mathcal{H}(\mathcal{R})$. Let $\bar{\mathcal{S}}$ be the σ-ring of μ^* measurable sets. Then $\mathcal{S}(\mathcal{R})$, the σ-ring generated by $\mathcal{R}$, is contained in $\bar{\mathcal{S}}$.*

Proof. It suffices to show that $\mathcal{R} \subset \bar{\mathcal{S}}$. Let $E \in \mathcal{R}, A \in \mathcal{H}(\mathcal{R})$ be arbitrary. We must show that $\mu^*(A) \ge \mu^*(A \cap E) + \mu^*(A \cap E^c)$. If $\mu^*(A) = \infty$, there is nothing to prove. Assume $\mu^*(A) < \infty$. Let $\epsilon > 0$. Then $\exists \{E_i\}_{i=1}^{\infty} \subseteq \mathcal{R}$ such that $A \subset \cup_{i=1}^{\infty} E_i$ and $\sum_{i=1}^{\infty} \mu(E_i) < \mu^*(A) + \epsilon$. We have

$$\begin{aligned}\mu^*(A) + \epsilon &> \sum_{i=1}^{\infty} \mu(E_i) \\ &= \sum_{i=1}^{\infty} \mu(E_i \cap E) + \sum_{i=1}^{\infty} \mu(E_i \cap E^c) \\ &= \sum_{i=1}^{\infty} \mu^*(E_i \cap E) + \sum_{i=1}^{\infty} \mu^*(E_i \cap E^c) \\ &\ge \mu^*(A \cap E) + \mu^*(A \cap E^c) \text{ by subadditivity.}\end{aligned}$$

Hence $E \in \bar{\mathcal{S}} \implies \mathcal{R} \subseteq \bar{\mathcal{S}}$. $\square$

Proposition 2.26. *Let $E \in \mathcal{H}(\mathcal{R})$. Then*

$$\begin{aligned}\mu^*(E) &= \mathit{inf}\,\{\bar{\mu}(F) : F \in \bar{\mathcal{S}}, E \subset F\} \\ &= \mathit{inf}\,\{\bar{\mu}(F) : F \in \mathcal{S}(\mathcal{R}), E \subset F\}\end{aligned}$$

Proof.
$$\begin{aligned}
\mu^*(E) &= \inf \{\sum_{i=1}^{\infty} \mu(E_i) : E \subset \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{R}\} \\
&= \inf \{\sum_{i=1}^{\infty} \bar{\mu}(E_i) : E \subset \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{R}\} \\
&\geq \inf \{\sum_{i=1}^{\infty} \bar{\mu}(E_i) : E \subset \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{S}(\mathcal{R})\} \\
&\geq \inf \{\bar{\mu}(\cup_{i=1}^{\infty}(E_i)) : E \subset \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{S}(\mathcal{R})\} \\
&= \inf \{\bar{\mu}(F) : E \subset F, F \in \mathcal{S}(\mathcal{R})\} \\
&\geq \inf \{\bar{\mu}(F) : E \subset F, F \in \bar{\mathcal{S}}\} \\
&\geq \mu^*(E) \quad (\text{since } \bar{\mu}(F) = \mu^*(F) \geq \mu^*(E)).
\end{aligned}$$

□

Thus equality must hold in every step. In particular, we record the following equality:
$$\mu^*(E) = \inf \{\bar{\mu}(\cup_{i=1}^{\infty}(E_i)) : E \subset \cup_{i=1}^{\infty} E_i, E_i \in \mathcal{S}(\mathcal{R})\}.$$
But the right hand side of the above equality, by the definition of an outer measure, is equal to $(\bar{\mu})^*(E)$. Thus, if we start with the measure $\bar{\mu}$ on the σ-ring $\mathcal{S}(\mathcal{R})$, the corresponding outer measure is μ^*.

Definition 2.27. Let $E \in \mathcal{H}(\mathcal{R})$. When $F \supset E, F \in \mathcal{S}(\mathcal{R})$, we say that F is a measurable cover of E if whenever $G \in \mathcal{S}(\mathcal{R})$ and $G \subseteq F \setminus E$, then $\bar{\mu}(G) = 0$.

Proposition 2.28. *Let $E \in \mathcal{H}(\mathcal{R})$ and $\mu^*(E) < \infty$. Then there exists a measurable cover of E.*

Proof. Let $n \in \mathbb{N}$. Since $\mu^*(E) < \infty$, by Proposition 2.26, $\exists F_n \in \mathcal{S}(\mathcal{R})$ such that $E \subseteq F_n$ and $\bar{\mu}(F_n) < \mu^*(E) + \frac{1}{n}$. Let $F = \cap_{n=1}^{\infty}$. Then $F \in \mathcal{S}(\mathcal{R})$ and $E \subseteq F$, and
$$\mu^*(E) \leq \mu^*(F) \leq \mu^*(F_n) = \bar{\mu}(F_n) < \mu^*(E) + \frac{1}{n} \quad \forall n \in \mathbb{N}.$$
Hence
$$\mu^*(E) = \mu^*(F) = \bar{\mu}(F).$$
Let $G \subseteq F \setminus E$, $G \in \mathcal{S}(\mathcal{R})$. Then $E \subseteq F \setminus G$ and $\mu(G) < \infty$.
$$\bar{\mu}(F) = \mu^*(E) \leq \mu^*(F \setminus G) = \bar{\mu}(F \setminus G) = \bar{\mu}(F) - \bar{\mu}(G).$$
By finiteness of $\bar{\mu}(F), \bar{\mu}(G) = 0$. □

Remark 2.29. The above is true if μ is σ-finite.

Remark 2.30. If μ is σ-finite, then μ^* is σ-finite. Hence $\bar{\mu}$ is σ-finite on $\mathcal{S}(\mathcal{R})$ and $\bar{\mathcal{S}}$.

3. Lebesgue Measure

Let $\mathcal{P} = \{[a, b) : a, b \in \mathbb{R}, a \leq b\}$ and $\mathcal{R}$ be the set of finite unions of elements from $\mathcal{P}$. $\mathcal{R}$ can also be shown to be equal to the set of finite **disjoint** unions of elements of $\mathcal{P}$.

Define $\mu([a, b)) := b - a, \mu(\emptyset) = 0$.

Proposition 3.1. *Let $\{E_1, \cdots, E_n\}$ be a finite disjoint collection of elements from $\mathcal{P}$, all of them contained in $E_0 \in \mathcal{P}$. Then $\sum_{i=1}^n \mu(E_i) \leq \mu(E_0$.*

Proof. Let $E_i = [a_i, b_i), i = 0, 1, \cdots, n$. If necessary, renumber the sets so that, by virtue for disjointness, $a_0 \leq a_1 < b_1 \leq a_2 < b_2 \leq \cdots \leq b_0$.

$$\begin{aligned}\sum_{i=1}^n \mu(E_i) &= \sum_{i=1}^n (b_i - a_i) \\ &\leq \sum_{i=1}^n (b_i - a_i) + \sum_{i=1}^{n-1} (a_{i+1} - b_i) \\ &= b_n - a_1 \\ &\leq b_0 - a_0 \\ &= \mu(E_0)\end{aligned}$$

□

Proposition 3.2. *Let $F_0 = [a_0, b_{0]}]$ be a closed interval contained in the union of open intervals $U_i = (a_i, b_i), i = 1, 2, \cdots, n$. Then $b_0 - a_0 \leq \sum_{i=1} 6n(b_i - a_i)$.*

Proof. Renumber and get rid of superfluous open intervals if necessary so that $b_i \in (a_{i+1}, b_{i+1}) \, \forall i$. Then $a_0 \in U_1, b_0 \in U_n$ and

$$b_0 - a_0 < b_n - a_1 = (b_1 - a_1) + \sum_{i=1}^{n-1} (b_{i+1} - b_i) \leq \sum_{i=1}^n (b_i - a_i).$$

□

Proposition 3.3. *Let $\{E_0, E_1, \cdots, E_n, \cdots\}$ be a sequence in $\mathcal{P}$ such that $E_0 \subset \cup_{i=1}^\infty E_i$. Then $\mu(E_0) \leq \sum_{i=1}^n \mu(E_i)$.*

Proof. The result is trivial if $E_0 = \emptyset$. So assume $E_0 \neq \emptyset$, i.e., $b_0 - a_0 > 0$. Choose $\epsilon > 0$ such that $0 < \epsilon < b_0 = a_0$. Let $\delta > 0$ be arbitrary. Then $F_0 = [a_0, b_0 - \epsilon] \subset E_0$ Similarly, $U_i = (a_i - \frac{\delta}{2}, b_i)$. So $E_i \subseteq U_i$. Hence, $F_0 \subseteq \cup_{i=1}^\infty U_i$. As F_0 is compact, $\exists n \in \mathbb{N}$ such that $F_0 \subseteq \cup_{i=1}^n U_i$. By Proposition 3.2,

$$b_0 - \epsilon - a_0 \leq \sum_{i=1}^n (b_i - a_i + \frac{\delta}{2^i}) \leq \sum_{i=1}^\infty (b_i - a_i) + \delta.$$

Since ϵ and δ are arbitrary, we get $b_0 - a_0 \leq \sum_{i=1}^n (b_i - a_i)$, i.e., $\mu(E_0) \leq \sum_{i=1}^\infty \mu(E_i)$. □

Proposition 3.4. *μ is countably additive on $\mathcal{P}$. That is, if $\{E_i\} \subseteq \mathcal{P}$, E_i disjoint, $E = \cup_{i=1}^\infty E_i \in \mathcal{P}$, then $\mu(E) = \sum_{i=1}^\infty \mu(E_i)$.*

Proof. By Proposition 3.3, $\mu(E) \leq \sum_{i=1}^\infty \mu(E_i)$. By Proposition 3.1, $\sum_{i=1}^n \mu(E_i) \leq \mu(E) \forall n \in \mathbb{N}$. Hence $\sum_{i=1}^\infty \mu(E_i) \leq \mu(E)$ □

Theorem 3.5. *$\exists$ a measure $\tilde{\mu}$ on $\mathcal{R}$ such that*

$$\tilde{\mu}([a, b)) = \mu([a, b)) = b - a \, \forall [a, b) \in \mathcal{P}.$$

Proof. Let $E \in \mathcal{R}, E = \cup_{i=1}^{n} = \cup_{j=1}^{m}$, with $E_i, F_j \in \mathcal{P}$, E_i, F_j disjoint. Then $E_i = \cup_{j=1}^{m}(E_i \cap F_j)\,\forall i$ and $F_j = \cup_{i=1}^{n}(F_j \cap E_i)\,\forall m$. So $\mu(E_i) = \sum_{j=1}^{m} \mu(E_i \cap F_j)$ and $\mu(F_j) = \sum_{i=1}^{n} \mu(F_j \cap E_i)$. Hence $\sum_{i=1}^{n} \mu(E_i) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mu(E_i \cap F_j) = \sum_{j=1}^{m} \mu(F_j)$. So we define $\tilde{\mu}$ as follows: Let $E \in \mathcal{R}, E = \cup_{i=1}^{n}, E_i \in \mathcal{P}, E_i$ disjoint. Let $\tilde{\mu}(E) = \sum_{i=1}^{n} \mu(E_i)$. Then

(1) $\tilde{\mu}(\emptyset) = 0$
(2) $\tilde{\mu} \geq 0$
(3) $\tilde{\mu}$ is finitely additive.
(4) $\tilde{\mu}$ is countably additive.

We now prove that $\tilde{\mu}$ is countably additive. Let $E = \cup_{i=1}^{\infty} E_i, E \in \mathcal{R}, E_i \in \mathcal{R}, E_i$ disjoint. Then for each $i, E_i = \cup_{k=1}^{n_i} E_{ik}, E_{ik} \in \mathcal{P}, E_{ik}$ disjoint.

(1) If $E \in \mathcal{P}$, then $E = \cup_{i=1}^{\infty} \cup_{k=1}^{n_i} E_{ik}, E_{ik} \in \mathcal{P}, E \in \mathcal{P}$. By Proposition 3.4, $\tilde{\mu}(E) = \sum_{i=1}^{\infty}\sum_{k=1}^{n_i} \mu(E_{ik}) = \sum_{i=1}^{\infty} \mu(E_{ik})$.
(2) If $E \in \mathcal{R}$, then $E = \cup_{i=1}^{n} F_j, F_j \in \mathcal{P}$ disjoint. Then $F_j = E \cap F_j = \cup_{i=1}^{\infty} E_i \cap F_j$, where the $E_i \cap F_j$ are in $\mathcal{R}$ and disjoint. Now,

$$\begin{aligned}\tilde{\mu}(E) &= \sum_{j=1}^{n} \mu(F_j)\\ &= \sum_{j=1}^{n}(\sum_{i=1}^{\infty} \mu(E_i \cap F_j))\\ &= \sum_{i=1}^{\infty}(\sum_{j=1}^{n} \mu(E_i \cap F_j))\end{aligned}$$

$E_i = E \cap E_i = \cup_{j=1}^{n} E_i \cap F_j$. Hence by finite additivity of $\tilde{\mu}$,

$$\tilde{\mu}(E) = \sum_{i=1}^{\infty} \mu(E_i) = \sum_{i=1}^{\infty} \tilde{\mu}(E_i).$$

□

Hereafter we refer to $\tilde{\mu}$ as merely μ. Thus we now have a measure μ on the ring $\mathcal{R}$. By the Caratheodory extension theorem, we get the completion of μ, μ^* from which we get $\bar{\mu}$ on the σ-rings $\mathcal{S}(\mathcal{R})$ and $\bar{\mathcal{S}}$. $\bar{\mu}$ is called the Lebesgue measure on $\mathbb{R}$, and sets in $\bar{\mathcal{S}}$ are called Lebesgue measurable sets.

Since $\mathbb{R} = \cup_{n \in \mathbb{Z}}[n, n+1), \mathbb{R} \in \mathcal{H}(\mathcal{R})$. Hence $\mathcal{H}(\mathcal{R})$ is in fact a σ-algebra and is equal to all the subsets of $\mathbb{R}$. $\bar{\mu}(\mathbb{R}) = \infty$, but $\bar{\mu}$ is a σ-finite measure. Henceforth we drop the bar and simply write μ.

Suppose we take $\mathbb{R}^n$ instead of $\mathbb{R}$. Let $\mathcal{P} = \{\prod_{i=1}^{n}[a_i, b_i) : a_i < b_i\}$. Then Propositions 3.1, 3.2, 3.3 and 3.4 are true, so that all of the above is true.

What is $\mu(\mathbb{R})$ in $(\mathbb{R}^2, \mu)$, where μ is the Lebesgue measure in $\mathbb{R}^2$? It is 0. To prove this, let $[a, b) \subset \mathbb{R}$. Let $E_\epsilon = [a, b) \times [0, \epsilon)$. Then $\mu(E_\epsilon) = \epsilon(b-a)$. Since $[a, b) \times \{0\} = \cap_{n=1}^{\infty} E_{\frac{1}{n}}$, $\mu([a, b) \times \{0\}) = \lim_{n\to\infty} \frac{1}{n}(b-a) = 0$. Now $\mathbb{R} = \mathbb{R} \times \{0\} = \cup_{n \in \mathbb{Z}}[n, n+1) \times \{0\}$. Hence $\mu(\mathbb{R}) = 0$.

In general, if μ is the Lebesgue measure on $\mathbb{R}^n$, and if E is a subspace of $\mathbb{R}^n$ with dimension strictly less than n, then $\mu(E) = 0$.

Proposition 3.6. *$\mathcal{S}(\mathcal{R}) = \mathcal{S}(U)$, where U is the collection of all open subsets of $\mathbb{R}$. That is, $\mathcal{S}(\mathcal{R})$ is the Borel σ- algebra on $\mathbb{R}$.*

Proof. Let $a, b \in \mathbb{R}$. $(a, b) = [a, b) \setminus \{a\}$ and $\{a\} = \cap_{n=1}^{\infty}[a, a + \frac{1}{n}) \in \mathcal{S}(\mathcal{R})$. Hence $(a, b) \in \mathcal{S}(\mathcal{R})$. Since every open set is the countable union of open intervals, $U \subseteq \mathcal{S}(\mathcal{R})$. Hence $\mathcal{S}(U) \subseteq \mathcal{S}(\mathcal{R})$.

Conversely, $[a, b) = (a, b) \cup \{a\}$ and $\{a\} = \cap_{n=1}^{\infty}(a - \frac{1}{n}, a + \frac{1}{n}) \in \mathcal{S}(U)$. Thus $[a, b) \in \mathcal{S}(U) \implies \mathcal{P} \subseteq \mathcal{S}(U) \implies \mathcal{R} \subseteq \mathcal{S}(U) \implies \mathcal{S}(\mathcal{R}) \subseteq \mathcal{S}(U)$. □

Corollary 3.7. *Every countable set is Borel measurable and its measure is 0.*

Proof. We have seen that $\{a\} = \cap_{n=1}^{\infty}(a - \frac{1}{n}, a + \frac{1}{n}) \in \mathcal{S}(U) = \mathcal{S}(\mathcal{R})$ and $\mu(\{A\}) = \lim_{n\to\infty} \frac{2}{n} = 0$. By countable additivity, the emasure of a countable set is 0. □

Proposition 3.8. *Let $E \subseteq \mathbb{R}$. Then $\mu^*(E) = inf\, \{\mu^*(U) : U \text{ open}, E \subset U\}$.*

Proof. The proposition is trivially true if $\mu^*(E) = \infty$. Assume $\mu^*(E) < \infty$. If $E \subset U$, then $\mu^*(E) \le \mu^*(U)$. Hence $\mu^*(E) \le \inf \{\mu^*(U) : U \text{ open}, E \subset U\}$. Let $\epsilon > 0$. Then $\exists E_i = [a_i, b_i)$ such that $E \subset \cup_{i=1}^{\infty} E_i$ and $\sum_{i=1}^{n} \mu(E_i) < \mu^*(E) + \frac{\epsilon}{2}$. i.e., $\sum_{i=1}^{\infty}(b_i - a_i) < \mu^*(E) + \frac{\epsilon}{2}$.

$E_i \subset U_i = (a_i - \frac{\epsilon}{2^{i+1}}, b_i)$. Let $U = \cup_{i=1}^{\infty} U_i$, an open set. Then $E \subset U$ and $\mu^*(U) \le \sum_{i=}^{\infty}(b_i - a_i) + \frac{\epsilon}{2} < \mu^*(E) + \epsilon$. Hence $\mu^*(E) = \inf \{\mu^*(U) : U \text{ open}, E \subset U\}$. □

Proposition 3.9. *Let $E \subseteq \mathbb{R}$. Then the following are equivalent:*

(1) E is (Lebesgue) measurable.
(2) Given $\epsilon > 0$, $\exists$ an open set U such that $E \subset U$, $\mu^*(U \setminus E) < \epsilon$.
(3) Given $\epsilon > 0$, $\exists$ a closed set F such that $F \subset E$, $\mu^*(E \setminus F) < \epsilon$.
(4) $\exists G$, a G_δ set such that $E \subset G$, $\mu^*(G \setminus E) = 0$.
(5) $\exists F$, an F_σ set such that $F \subset E$, $\mu^*(E \setminus F) = 0$.

Proof. We first show that (1) $\implies$ (2) $\implies$ (4) $\implies$ (1). Suppose (1) holds. First assume that $\mu^*(E) < \infty$. For $\epsilon > 0, \exists U$ open such that $E \subset U$ and $\mu^*(U) < \mu^*(E) + \epsilon$. Since E is measurable, $\mu^*(U \setminus E) = \mu^*(U) - \mu^*(E) < \epsilon$. If $\mu^*(E) = \infty$, we can write $E = \cup_{i=1}^{\infty} E_i, \mu^*(E_i) < \infty$, by σ finiteness. For each E_i, $\exists U_i$ such that $\mu^*(U_i \setminus E_i) < \frac{\epsilon}{2^i}$. Let $U = \cup_{i=}^{\infty} U_i$, which is open. A=Then, $E \subset U$ and $\mu^*(U \setminus E) \le \sum_{i=1}^{\infty} \mu^*(U_i \setminus E_i) < \sum_{i=1}^{\infty} \frac{\epsilon}{2^i} = \epsilon$. Hence (2) holds.

Next, suppose (2) holds. Given E, choose U_n open such that $\mu^*(U_n \setminus E) < \frac{1}{n}$. Let $G = \cap_{n=1}^{\infty} U_n$. Then $E \subset G$ and $\mu^*(G \setminus E) \le \mu^*(U_n \setminus E) < \frac{1}{n} \forall n$. Hence $\mu^*(G \setminus E) = 0$, i.e., (4) holds.

Finally, suppose (4) holds. $\mu^*(G \setminus E) = 0 \implies G \setminus E$ is measurable, since μ^* is a complete measure. G is Borel measurable since it is a G_δ set, and hence measurable. Hence $E = G \setminus (G \setminus E)$ is measurable, i.e. (1) holds.

Finally, we show that (1) $\implies$ (3) $\implies$ (4). Suppose E is measurable. Then E^c is measurable. Since (1) $\implies$ (2), for $\epsilon > 0$, $\exists U$ open such that $E^c \subseteq U$ and $\mu^*(U \setminus E^c) < \epsilon$. Let $F = U^c$. Then F is closed and $F \subseteq E$. Since $U \setminus E^c = E \setminus U^c$, we get $\mu^*(E \setminus F) = \mu^*(E \setminus U^c) = \mu^*(U \setminus E^c) < \epsilon$.

Suppose (3) holds. Given E, choose F_n closed such that $F_n \subseteq E$ and $\mu^*(E \backslash F_n) < \frac{1}{n}$. Let $F = \cup_{n=1}^{\infty} F_n$. Then $F \subseteq E$ and $\mu^*(E \setminus F) \leq \mu^*(E \setminus F_n) < \frac{1}{n}\, \forall n$. Hence $\mu^*(E \setminus F) = 0$.

Suppose (5) holds. Since $\mu^*(E \setminus F) = 0$, $E \setminus F$ is measurable. F is measurable as it is an F_σ set. Hence $E = (E \setminus F) \cup F$ is measurable. □

Let $X = [0,1]$, $X_1 = (\frac{1}{3}, \frac{2}{3}))$, $X_2 = (\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9})$, X_3 the union of the four middle third intervals of $X \setminus (X_1 \cup X_2)$. Let $C = X \setminus (\cup_{n=1}^{\infty} X_n)$. Then the following hold:

1. Each X_n is open and hence C is closed.
2. $\mu(X_1) = \frac{1}{3}, \mu(X_2) = \frac{2}{9}, \mu(X_3) = \frac{4}{27}$. In general, $\mu(X_n) = \frac{2^{n-1}}{3^n}$. Hence $\mu(\cup_{n=1}^{\infty} X_n = 1$ and $\mu(C) = 0$. C can be thought of as the set of numbers between 0 and 1 such that 1 does not occur anywhere in its ternary expansion.
3. C is uncountable by Cantor's diagonalisation,
4. C is nowhere dense.

$\mu(C) = 0$ implies that every subset of C is measurable. The cardinality of C is c (continuum), hence the cardinality of Lebesgue measurable sets is 2^c. However, it can be shown that the cardinality of Borel measurable sets is c. Hence Borel measurable sets are properly contained in Lebesgue measurable sets. In particular, this means that the Borel measure on $\mathcal{S}(\mathcal{R})$ on $\mathbb{R}$ is not complete.

Let $T : \mathbb{R} \to \mathbb{R}$ be given by $Tx = \alpha x + \beta$, $\alpha > 0$. Then T is a bijection and $T^{-1}x = \frac{x-\beta}{\alpha}$.

Let $\mathcal{S} = \mathcal{S}(\mathcal{R})$ be the Borel σ-algebra. Let $T(\mathcal{S}) = \{T(E) : E \in \mathcal{S}\}$. Then $T(\mathcal{S})$ is a σ-algebra. Since T is a bijection, $T(\mathbb{R}) = \mathbb{R}$. $T(A \setminus B) = T(A) \setminus T(B)$ and $T(\cup_i E_i) = \cup_i T(E_i)$.

Suppose $E = [a,b) = T(F)$, where $F = [\frac{a-\beta}{\alpha}, \frac{b-\beta}{\alpha})$. Hence $\mathcal{P} \subset T(\mathcal{S})$ and thus $\mathcal{S} \subset T(\mathcal{S})$. Similarly, $\mathcal{S} \subset T^{-1}(\mathcal{S})$ and thus $T(\mathcal{S}) = \mathcal{S}$. Thus, E is a Borel set iff $T(E)$ is a Borel set.

Proposition 3.10. *Let $E \subset \mathbb{R}$. Then $\mu^*(T(E)) = \alpha\mu^*(E)$.*

Proof. $\mu^*(T(E)) = \inf\{\sum \mu(F_i) : F_i \in \mathcal{P}, T(E) \subset \cup F_i\}$. But $F_i = T(E_i)$ for some $E_i \in \mathcal{P}$ and $E_i \supset E$. Hence $\mu^*(T(E)) = \inf\{\sum \mu(T(E_i)) : E_i \in \mathcal{P}, E \subset \cup E_i\}$. If $E_i = [a,b), T(E_i) = [\alpha a + \beta, \alpha b + \beta)$ and $\mu(T(E_i)) = \alpha(b-a) = \alpha\mu^*(E)$. Hence $\mu^*(T(E)) = \alpha\mu^*(E)$. □

Now assume $E \subset \mathbb{R}$ is Lebesgue measurable. We claim that $T(E)$ is Lebesgue measurable. Let $A \subset \mathbb{R}$. Then

$$\begin{aligned}\mu^*(A \cap T(E)) + \mu^*(A \cap (T(E))^c) &= \mu^*(T(T^{-1} \cap E)) + \mu^*(T(T^{-1}(A) \cap E^c)) \\ &= \alpha[\mu^*(T^{-1}(A) \cap E) + \mu^*(T^{-1}(A) \cap E^c)] \\ &= \alpha\mu^*(T^{-1}(A)) \text{ since } E \text{ is Lebesgue measurable} \\ &= \alpha\frac{1}{\alpha}\mu^*(A).\end{aligned}$$

Remark 3.11. If $\alpha = 1$, then all of the above goes through in $\mathbb{R}^n$ as well.

Remark 3.12. If $\alpha = 1, T(E + x) = T(E)\, \forall E, \forall x$. That is, the Lebesgue measure is translation invariant.

Theorem 3.13. *Let ν be a Borel measure on $\mathbb{R}^N$ such that*

(1) $\nu(K) < \infty \quad \forall K \subset \mathbb{R}^n$ compact
(2) $\nu(E) = \inf\{\nu(V) : V \text{open}, E \subset V\} \quad \forall E$ Borel
(3) ν is translation invariant.

Then $\exists$ a constant $c \in \mathbb{R}$ such that $\nu = c\mu$, i.e. for all Borel measurable sets $E, \nu(E) = c\mu(E)$.

Proof. Let $a = (a_1, \cdots, a_N) \in \mathbb{R}^N$ and $\delta > 0$. Let $Q(a, \delta) = \prod_{i=1}^N [a_i, a_i + \delta)$. Denote by Ω_n the collection of all boxes of this form where $\delta = 2^{-n}$ and a has coordinates which are integral multiples of 2^{-n}.

(1) If $x \in \mathbb{R}^N$, then x belongs to exactly one box of $\Omega_n\, \forall n$.
(2) Every open set is the countable disjoint union of boxes taken from $\Omega_1 \cup \Omega_2 \cup \cdots$.
(3) Let $Q = Q(a, 1)$. Then Q is the disjoint union of 2^{Nn} identical boxes $\tilde{Q}$ from Ω_n.

Since ν and μ are translation invariant, all these boxes have the same measure. $\mu(Q) = 1$. Let $\nu(Q) = c$. Then

$$2^{Nn}\nu(\tilde{Q}) = \nu(Q) = c = c\mu(Q) = c2^{Nn}\mu(\tilde{Q}).$$

Hence $\nu(\tilde{Q}) = c\mu(\tilde{Q})$. By (2), it follows that if U is open, $\nu(U) = c\mu(Q)$. By the second hypothesis, $\nu(E) = c\mu(E)$. □

Theorem 3.14. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be linear. Then for all Borel sets E in $\mathbb{R}^n$, $\mu(A(E)) = |\det A|\mu(E)$.*

Proof.
(1) If A is singular, then $A(E) \subset A(\mathbb{R}^n) =$ a lower dimensional subspace of $\mathbb{R}^n$. Then $\mu(A(\mathbb{R}^n)) = 0$. Hence $A(E)$ is Lebesgue measurable and $\mu(A(E)) = 0 = |\det A|\mu(E)$.
(2) Let A be non-singular. Now E is Borel iff $A(E)$ is Borel.
$\mathcal{S} = \{E : A(E) \text{ is Borel}\}$ is a σ-algebra since E open iff $A(E)$ open, $E = \cup E_i, E_i$ disjoint, then $A(E) = \cup A(E_i)$. Define $\nu(E) = \mu(A(E))$. Then ν is a measure on $\mathcal{S}$. K compact implies that $A(E)$ is compact. Hence $\nu(K) = \mu(A(K)) < \infty$ and $\nu(E + x) = \mu(A(E) + Ax) = \mu(A(E)) = \nu(E)$.

Thus ν is translation invariant. $E \subset V$ is open iff $A(E) \subset A(V)$ is open. Hence

$$\begin{aligned}\inf\{\nu(V) : V \supset E, V \text{ open }\} &= \inf\{\mu(A(V)) : V \supset E\} \\ &= \inf\{\mu(A(V)) : A(V) \supset A(E)\} \\ &= \mu(A(E)) \\ &= \nu(E).\end{aligned}$$

Hence $\nu = c_A \mu$ by Theorem 3.13. Claim: $c_{AB} = c_A c_B$.

Exercise 3.15. (a) If A is orthogonal and E the unit ball. Show that $A(E) = E \implies c_A = 1 = \det A$.

(b) A diagonal, $A = \text{diag }(\lambda_1, \cdots, \lambda_n)$, $\lambda_i > 0\,\forall i$ and $E = [0,1]^N$. Then $A(E) = \prod_{i=1}^N [0, \lambda_i]$, and $\mu(A(E)) = \prod_{i=1}^N \lambda_i = \det A$.

(c) A non-singular. Then $A = RQ$, where R is positive definite and Q orthogonal. R can further be written as $P^T DP$, where P is orthogonal and D is diagonal. Then $A = P^T DPQ$ and $c_A = \det D = |\det A|$.

□

We now construct a non-(Lebesge) measurable set. Let $x, y \in [0,1)$. Then

$$x + y = \begin{cases} x + y, & x + y < 1 \\ x + y - 1, & x + y \geq 1 \end{cases}$$

Let $E \subset [0,1)$. Then $E + y = \{x + y : x \in E\}$.

Lemma 3.16. *If $E \subset [0,1)$ is measurable, then $E + y$ is measurable for all $y \in [0,1]$ and $\mu(E + y) = \mu(E)$.*

Proof. For each $y \in Y$, let $E_1 = E \cap [0, 1-y)$ and $E_2 = E \cap [1-y, 1)$. They are disjoint and $\mu(E) = \mu(E_1 + \mu(E_2)$. $E_1 + y = E_1 + y$ and $E_2 + y = E_2 + (y-1)$. Claim: $(E_1 + y) \cap (E_2 + y) =$. If not, suppose $\exists a, b \in [0,1)$ such that $a \in E_1, b \in E_2$ and $a + y = b + y - 1$. Then $b - a = 1$ which is a contradiction, since $a, b \in [0,1)$. Hence

$$\begin{aligned}\mu(E + y) &= \mu(E_1 + y) + \mu(E_2 + y) \\ &= \mu(E_1 + \mu(E_2) \\ &= \mu(E).\end{aligned}$$

□

Let $x, y \in [0,1)$. Define a relation on $[0,1)$ by: $x \sim y \iff x - y \in \mathbb{Q}$. This is an equivalence relation. Thus, $[0,1)$ can be written as the disjoint union of equivalence classes. Let P be the set made up of exactly one representative from each equivalence class, using the axiom of choice.

Claim: P is not measurable. Let $\{r_i\}$ be an enumeration of the set in $[0,1)$ such that $r_0 = 0$. Let $P_i = P + r_i$. The P_is are mutually disjoint, for if $x \in P_i \cap P_j$, then $x = r_i + p_i = r_j + p_j, p_i, p_j \in P \implies p_i - p_j \in \mathbb{Q} \implies p_i \sim p_j$. For each $i, \mu)P_i) = \mu(P)$ and $\cup_{i=1}^\infty P_i = [0,1)$. If P is measurable, then $\sum \mu(P_i) = 1$, a contradiction. Hence $\bar{\mathcal{S}} \subsetneq \mathcal{P}(\mathbb{R})$.

Remark 3.17. (1) Let $E \subset P$ be measurable and $E_i = E + r_i$. The E_is are disjoint measurable sets and $\mu(E_i) = \mu(E)$. Then $1 \geq \mu(E) = \sum \mu(E_i)$. Hence the only measurable subsets of P are of emasure 0.

(2) Let $A \subset [0,1), \mu(A) > 0$ and $E_i = A \cap P_i$. If all E_is are measurable, then $\mu(A) = \sum \mu(E_i)$. But $\mu(E_i)0$ bu the previous remark. Hence, one of the E_is is not measurable. So every set of positive measure has a non-measurable set.

4. Measurable Functions

Let X be a non-empty set and $\mathcal{S}$ be a σ-algebra on X. Then $(X, \mathcal{S})$ is a measurable spaces and elements of $\mathcal{S}$ are called measurable sets. Suppose μ is a measure on $\mathcal{S}$. Then $(X, \mathcal{S}, \mu)$ is called a measure space.

Definition 4.1. Let $X, \mathcal{S})$ be a measurable space, and $f : X \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function. Then f is said to be measurable if $\forall \alpha \in \mathbb{R}$, the set $\{x : f(x) > \alpha\} \in \mathcal{S}$, i.e., $f^{-1}((\alpha, \infty]) \in \mathcal{S}\, \forall \alpha \in \mathbb{R}$.

Remark 4.2. Let $X = \mathbb{R}$ and $f : X \to \mathbb{R} \cup \{\pm\infty\}$. Then f is said to be Lebesgue (respectively Borel) measurable if $f^{-1}((-\infty, \infty))$ is a Lebesgue (respectively Borel) measurable set.

Proposition 4.3. *Let f be an extended real-valued function on a measurable space $(X, \mathcal{S})$. Then the following are equivalent:*

(1) $\forall \alpha \in \mathbb{R}, f^{-1}((\alpha, \infty]) \in \mathcal{S}$.
(2) $\forall \alpha \in \mathbb{R}, f^{-1}([\alpha, \infty]) \in \mathcal{S}$.
(3) $\forall \alpha \in \mathbb{R}, f^{-1}([-\infty, \alpha)) \in \mathcal{S}$.
(4) $\forall \alpha \in \mathbb{R}, f^{-1}([-\infty, \alpha]) \in \mathcal{S}$.

Proof. (1) $\implies$ (2) follows from $f^{-1}([\alpha, \infty]) = \cap_n f^{-1}((\alpha - \frac{1}{n}, \infty])$. (2) $\implies$ (3) follows from $f^{-1}([-\infty, \alpha)) = f^{-1}([\alpha, \infty]^c) = f^{-1}([\alpha, \infty])^c$. (3) $\implies$ (4) follows from $f^{-1}([\infty, \alpha]) = \cap_n f^{-1}([-\infty, \alpha + \frac{1}{n}))$. (4) $\implies$ (1) follows from $f^{-1}((\alpha, \infty]) = f^{-1}([-\infty, \alpha]^c) = f^{-1}([-\infty, \alpha])^c$. □

Corollary 4.4. *If f is measurable, then:*

(1) For $\alpha \in \mathbb{R} \cup \{\infty\}, f^{-1}(\{\alpha\})$ is measurable.
(2) For V open in $\mathbb{R}$, $f^{-1}(V)$ is measurable.

Proof. (1) $\{\infty\} = \cap_n [n.\infty]$ and $\{-\infty\} = cap_n[-\infty, -n]$. For $\alpha \in \mathbb{R}, \{\alpha\} = (-\infty, \alpha] \cap [\alpha, \infty)$.
(2) $(a, b) = [-\infty, b) \cap (a, \infty] \implies f^{-1}((a, b)) \in \mathcal{S} \implies f^{-1}(V) \in \mathcal{S}$.

□

Remark 4.5. Let $f : X \to \mathbb{R}$. Then f is measurable iff $f^{-1}(U)$ is measurable for all open sets U. This holds as $f^{-1}([-\infty, \alpha)) - f^{-1}((-\infty, \alpha)) \in \mathcal{S}$. But $f^{-1}(\{\alpha\}) \in \mathcal{S}\, \forall \alpha \in \mathbb{R} \not\Longrightarrow f$ is measurable.

Example 4.6. Let $X = \mathbb{R}$ and μ be the Lebesgue measure. Let E be a non-measurable subset of $[0, 1)$. Define

$$f(x) = \begin{cases} x, & x \in E \\ -x, & x \in [0,1) \setminus E \\ -2, & x \notin [0,1) \end{cases}$$

Then

$$f^{-1}(\{\alpha\}) = \begin{cases} \mathbb{R} \setminus [0,1), & \alpha = -2 \\ \{\alpha\}, & -\alpha \in [0,1) \setminus E \\ \{\alpha\}, & \alpha \in E \\ \emptyset, & \text{otherwise} \end{cases}$$

So the inverse of each singleton set is measurable. But $\{x : f(x) > 0\} = f^{-1}((0,\infty]) = E$ which is not measurable.

Exercise 4.7. (1) Let $A \subset X$. Define

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

Then χ_A is measurable iff A is measurable.

(2) Every continuous function $f : \mathbb{R} \to \mathbb{R}$ is both Lebesgue and Borel measurable.

Proposition 4.8. *Let f, g be measurable functions on $(X, \mathcal{S})$ and $c \in \mathbb{R}$. Then $cf, f+g$ and fg are all measurable.*

Proof.

$$\{x : xf(x) < \alpha\} = \begin{cases} \{x : f(x) < \frac{\alpha}{c}\}, & c > 0 \\ \{x : f(x) > \frac{\alpha}{c}\}, & c < 0 \end{cases}$$

In particular, $-f$ is measurable.

$$\begin{aligned} \{x : f(x) + g(x) < \alpha\} &= \{x : f(x) < \alpha - g(x)\} \\ &= \cup_{r \in \mathbb{Q}}(\{x : f(x) < r\}) \cap \{x : g(x) < \alpha + r\} \end{aligned}$$

Finally, we show that f measurable implies that f^2 is measurable. This shows that fg is measurable as $fg = \frac{(f+g)^2 - (f-g)^2}{4}$.

$$\{x : f^2 > \alpha\} = \begin{cases} \emptyset, & \alpha \leq 0 \\ \{x : f(x) > \sqrt{\alpha}\} \cup \{x : f(x) < -\sqrt{\alpha}\}, & \alpha > 0 \end{cases}$$

□

Proposition 4.9. *If f is measurable then so is $|f|$.*

Proof. $\{x : |f(x)| < \alpha\} = \{x : -\alpha < f(x) < \alpha\}$.

□

But the converse is not true.

Example 4.10. Let E be a non-measurable set. Let

$$f(x) = \begin{cases} 1, & x \in E \\ -1, & x \notin E \end{cases}$$

Definition 4.11. Max $\{f, g\} = \frac{1}{2}(f+g+|f-g|\}$ and Min $\{f, g\} = \frac{1}{2}(f+g-|f-g|\}$. Let $f^+ = \max\ \{f, 0\}$ and $f^- = -\min\ \{f, 0\}$. Then $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

Proposition 4.12. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be Borel measurable and $f : X \to \mathbb{R}$ be a real-valued measurable function. Then $\phi \circ f : X \to \mathbb{R}$ is measurable.*

Proof. $\{x : \phi(f(x)) > \alpha\} = f^{-1}(\phi^{-1}((\alpha, \infty]))$. Now, $\phi^{-1}((\alpha, \infty])$ is a Borel set. Hence it suffices to show that f measurable implies that $f^{-1}(E)$ is measurable for all Borel sets E. Let $\tilde{\mathcal{S}} = \{E : f^{-1}(E) \text{ is measurable}\}$. Then $\tilde{\mathcal{S}}$ is a σ-algebra containing open sets and hence contains all Borel sets. □

Proposition 4.13. *Suppose $\{f_n\}$ is a sequence of measurable functions. Let $h(x) = \sup_n f_n(x)$ and $g(x) = \inf_n f_n(x)$. Then h and g are measurable.*

Proof. The proof follows as

$$\{x : h(x) > c\} = \cup_{n=1}^{\infty}\{x : f_n(x) > c\}$$

and

$$\{x : g(x) < c\} = \cup_{n=1}^{\infty}\{x : f_n(x) < c\}.$$

□

Corollary 4.14. *Suppose $\{f_n\}$ is a sequence of measurable functions. Then $\lim\sup_{n\to\infty} f_n$ and $\lim\inf_{n\to\infty} f_n$ are measurable.*

Proof. The proof follows as

$$\limsup_{n\to\infty} f_n(x) = \inf_n \sup_{m\geq n} f_m(x)$$

and

$$\liminf_{n\to\infty} f_n(x) = \sup_n \inf_{m\geq n} f_m(x)$$

□

Corollary 4.15. *If $\{f_n\}$ is a sequence of measurable functions and $f_n \to f$, then f is measurable.*

Definition 4.16. A function of the form $f = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ is said to be a simple function.

Theorem 4.17. *Let f be an extended real valued measurable function which is non-negative. Then f is the increasing limit of non-negative simple functions, i.e. $\exists f_n$ simple such that $f_n \geq 0$, $f_n \leq f_{n+1}\,\forall n$ and $\lim_{n\to\infty} f_n(x) = f(x)\,\forall x$.*

Proof. Let $n \in \mathbb{N}$. For i such that $1 \leq i \leq n\,2^n$, define $E_{n,i} = f^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}))$ and $F_n = f^{-1}([n, \infty])$. Then $E_{n,i}, F_n$ are all measurable. Define

$$f_n = (\sum_{i=1}^{n2^n} \frac{i-1}{2^n}\chi_{E_{n,i}}) + \chi_{F_n}.$$

Then $f_n \geq 0, f_n$ simple and $f_n \leq f$. In fact,

$$f_n(x) = \begin{cases} n, & f(x) \geq n \\ \frac{i-1}{2^n}, & f(x) < n \text{ and } f(x) \in [\frac{i-1}{2^n}, \frac{i}{2^n}). \end{cases}$$

Hence $f_n \leq f_{n+1}$ and $\lim_{n\to\infty} f_n = f$. □

Since every measurable function f can be written as the difference of two non-negative functions, i.e. $f = f^+ - f^-$, we get the folloowing corollary.

Corollary 4.18. *Every measurable function is the pointwise limit of simple functions.*

4.1. **The Cantor Function.** Let $C = [0,1]$ and C be the Cantor set. Define

$$f_0(x) = x \quad \forall x \in X,$$

$$f_1(x) = \begin{cases} x, & x \in [0, \frac{1}{3}] \\ \frac{1}{3}, & x \in [\frac{1}{3}, \frac{2}{3}] \\ 2(x - \frac{2}{3}) + \frac{1}{3}, & x \in [\frac{2}{3}, 1] \end{cases}$$

$$f_2(x) = \begin{cases} x, & x \in [0, \frac{1}{9}] \\ \frac{1}{9}, & x \in [\frac{1}{9}, \frac{2}{9}] \\ 2(x - \frac{1}{3}) + \frac{1}{3}, & x \in [\frac{2}{9}, \frac{1}{3}] \\ \frac{1}{3}, & x \in [\frac{1}{3}, \frac{2}{3}] \\ 2(x - \frac{2}{3}) + \frac{1}{3}, & x \in [\frac{2}{3}, \frac{7}{9}] \\ \frac{5}{9}, & x \in [\frac{7}{9}, \frac{8}{9}] \\ 4(x - \frac{2}{3}) - \frac{1}{3}, & x \in [\frac{8}{9}, 1] \end{cases}$$

Continuing in this fashion we get a sequence $\{f_n\}$ such that $f_n \leq f_{n+1} \, \forall n$, and $\max |f_n(x) - f_{n+1}(x)| \leq \frac{2}{3}^{n+1}$. Hence $\|f_n - f_{n+1}\|_\infty \leq \frac{2}{3}^{n+1}$, Hence $\{f_n\}$ is uniformly Cauchy. Thus $f_n \to f$ uniformly to a continuous function f. f is non-decreasing and is constant on each interval in the complement of the Cantor set. f is called the Cantor function.

Define $\psi(y) = y + f(y)$. Then ψ is continuous, non-negative and strictly monotonic. $psi(0) = 0, \psi(1) = 2$, and hence $\psi : [0,1] \to [0,2]$ is $1{-}1$ and onto. Let ϕ be its inverse, i.e., $x = \phi(x) + f(\phi(x))$. Then ϕ is also monotonic; $x \geq y \implies \phi(x) \geq \phi(y)$. Now,

$$x - y = \phi(x) - \phi(y) + f(\phi(x)) - f(\phi(y)),$$

and $f(\phi(x)) - f(\phi(y)) \geq 0$. Hence $|\phi(x) - \phi(y)| \leq |x - y|$, i.e. ϕ is Lipschitz continuous. In particular, it is continuous.

Since ψ is $1 - 1$, it maps disjoint sets into disjoint sets. Let I be an interval in C^c. Let $x \in I$. Then $\psi(x) = x + c_I$, where f takes the value c_I on I. Hence $\psi(I)$ is a translate of I. Thus

$$\begin{aligned} \mu(C^c) &= \mu(\psi(C^c)) \\ &= 1. \end{aligned}$$

Hence $\mu(\psi(C)) = 1$ as $\mu(\psi(X)) = 2$. Let S be a non-measurable set contained in $\psi(C)$. Let $M = \psi^{-1}(S) \subset C$. Then M is Lebesgue measurable since μ is complete and $\mu(C) = 0$. Claim: M is not Borel measurable. Suppose M were Borel measurable.Then $\phi^{-1}(M)$ is Borel measurable as ϕ is continuous. Then $\phi^{-1}(M)$ is Lebesgue measurable. But $\phi^{-1}(M) = S$. Hence we get a contradiction.

Let $\Phi = \chi_M$. Then Φ is Lebesgue measurable. Define $\eta = \Phi \circ \phi$. η. Since ψ is continuous, it is Lebesgue measurable. If η were measurable, then $\eta^{-1}(\{1\})$ is measurable. But

$$\begin{aligned} \eta^{-1}(\{1\}) &= \{x : \eta(x) = 1\} \\ &= \{x : \Phi(\phi(x)) = 1\} \\ &= \phi^{-1}(M) \\ &= S, \end{aligned}$$

which is not measurable. Hence the composition of two measurable functions need not be measurable.

Definition 4.19. An occurrence is said to happen **almost everywhere** if it occurs on a set E such that $\mu(E^c) = 0$.

Exercise 4.20. Are the following statements equivalent? Or does any one imply the other?

(1) f is continuous almost everywhere (a.e).
(2) $f = g$ almost everywhere, and g is continuous.

In fact, neither statement implies the other. To see that (1) $\not\Longrightarrow$ (2), let

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Then f is continuous almost everywhere. But for any g which is equal to f a.e., g is not continuous. For (2) $\not\Longrightarrow$ (1), let $f = \chi_{\mathbb{Q}}$ and $g \equiv 0$ on $\mathbb{R}$. Then $f = g$ a.e. and g is continuous, but f is discontinuous everywhere.

Theorem 4.21 (Egoroff)**.** *Let $(X, \mathcal{S}, \mu)$ be a measure space and $\mu(X) < \infty$. Suppose $\{f_n\}$ is a sequence of measurable real valued functions converging pointwise to a real-valued measurable function. Then given $\epsilon > 0, \exists F \in \mathcal{S}$ wtth $\mu(F) < \epsilon$ such that $f_n \to f$ uniformly on F^c.*

Proof. Let $n, m \in \mathbb{N}$. Define

$$E_{n,m} = \bigcap_{i=n}^{\infty} \{x : |f_i(x) - f(x)| < \frac{1}{m}\}$$

$x \in E_{n,m} \implies \forall i \geq n, |f_i(x) - f(x)| < \frac{1}{m}$. Hence for each $m \in \mathbb{N}$,

$$E_{1,m} \subset E_{2,m} \subset \cdots \subset E_{n,m} \subset E_{n+1,m} \subset \cdots .$$

For $x \in X$ and $m \in \mathbb{N}, \exists\, n_m$ such that $n \geq n_m \implies |f_n(x) - f(x)| < \frac{1}{m}$.

$$\implies x \in E_{n_m,m} \implies X = \bigcup_{n=1}^{\infty} E_{n,m} \quad \forall m \in \mathbb{N}.$$

Now, $\mu(X) < \infty \implies \exists\, n_m$ such that

$$\mu(X \setminus E_{n_m,m}) = \mu(X) - \mu(E_{n_m,m}) < \frac{\epsilon}{2^m}.$$

Let

$$F = \bigcup_{m=1}^{\infty} (X \setminus E_{n_m,m}).$$

Then $mu(F) < \infty$ and

$$F^c = \bigcap_{m=1}^{\infty} E_{nm,m}.$$

Guven $\eta > 0$ choose m such that $\frac{1}{m} < \eta$. Since $x \in E_{n_m,m}$, $\forall n \geq n_m$,

$$|f_n(x) - f(x)| < \frac{1}{m} < \eta \quad \forall x \in F^c.$$

□

Remark 4.22. Egoroff's theorem is not true if $\mu(X) = \infty$. Here is a counter example.

Example 4.23. Let $X = \mathbb{N}, \mathcal{S} = \mathcal{P}(\mathbb{N})$ and μ be the counting measure. The only measure zero set in this measure space is the empty set. Hence if Egoroff's theorem were true, pointwise convergence would imply uniform convergence everywhere. Let $f_n = \chi_{\{1,2,\cdots,n\}}$. Then pointwise, $f_n \to f \equiv 1$, a constant. But for every $n \in \mathbb{N}$ and $m > n$, $f_n(m) = 0$ and $f_m(m) = 1$. Hence $f_n \not\to f$ uniformly.

Definition 4.24. A sequence f_n is said to converge to f almost uniformly if $\forall \epsilon > 0, \exists F \in \mathcal{S}$ such that $\mu(F) < \epsilon$ and $f_N \to f$ uniformly on F^c.

Theorem 4.25. *$f_n \to f$ almost uniformly $\implies$ $f_n \to f$ pointwise almost everywhere.*

Proof. For each $m \in \mathbb{N}, \exists\, F_m \in \mathcal{S}$ with $\mu(F_m) < \frac{1}{m}$ such that $f_n \to f$ uniformly on F_m^c. Let

$$F = \bigcap_{m=1}^{\infty} F_m.$$

Then $\mu(F) = 0$.

$$F^c = \bigcap_{m=1}^{\infty} F_m^c.$$

Hence $f_n(x) \to f(x)\, \forall x \in F^c$, i.e., $f_n \to f$ almost everywhere. □

Theorem 4.26 (Lusin)**.** *Let f be a Lebesgue measurable function on $[a, b] \subset \mathbb{R}$. Then given $\epsilon > 0, \exists$ a continuous function ϕ such that $\mu(\{x : f(x) \neq \phi(x)\}) < \epsilon$ and sup $|\phi| \leq$ sup $|f|$.*

Proof. We first prove the theorem for characteristic functions. Let $f = \chi_E, E \in \mathcal{S}$. Then $\exists\, U$ open, F closed such that $F \subset E \subset U$ and $\mu(U \setminus E) < \frac{\epsilon}{2}$ and $\mu(E \setminus F) < \frac{\epsilon}{2}$. By Urysohn's lemma, $\exists \phi$ continuous function with $0 \leq \phi \leq 1$ and

$$\phi = \begin{cases} 1, & \text{on } F \\ 0, & \text{on } U^c \end{cases}$$

Hence $\{x : f(x) \neq \phi(x)\} \subset U \setminus F$ and $\mu(U \setminus F) < \epsilon$. Next, let f be a simple function, i.e.,

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

For each $i, \exists \phi_i$ such that $\mu(\{x : \phi_i \neq \chi_{A_i}\}) < \frac{\epsilon}{n}$. Let

$$\phi = \sum_{i=1}^{n} \alpha_i \phi_i.$$

Then ϕ is continuous and

$$\{x : \phi(x) \neq f(x)\} \subset \bigcup_{i=1}^{n} \{x : \phi_i(x) \neq f_i(x)\}.$$

Hence $\mu(\{x : \phi(x) \neq f(x)\}) < \epsilon$. Next we consider a non-negative measurable function. Then by Theorem 4.17, $\exists f_n \geq 0$ simple functions such that $f_n \nearrow f$. BY Egoroff's theorem, $\exists F$ with $\mu(F) < \frac{\epsilon}{4}$ and such that $f_n \to f$ uniformly on F^c. By regularity of the Lebesgue measure, $\exists C_0 \subset F \subset [a, b]$ compact such that $\mu(F^c \setminus C) < \frac{\epsilon}{4}$. Hence $\mu([a, b] \setminus C_0) < \frac{\epsilon}{2}$ and $f_n \to f$ uniformly on C_0. Since each f_n is simple, $\exists \phi_n$ continuous such that $\mu(\{\phi_n \neq f_n\}) < \frac{\epsilon}{2^{n+1}}$. Hence $\exists C_n$ closed such that $\{\phi_n \neq f_n\} = [a, b] \setminus C_n$ and $\mu([a, b] \setminus C_n) < \frac{\epsilon}{2^{n+1}}$. Let $C = \bigcap_{n=0}^{\infty} C_n$. Then

$\mu([a,b]\setminus C) \leq \epsilon$. C is compact and ϕ_n is continuous on C for each n. Also, $\phi_n = f_n$ for each n on C. Hence ϕ_n converges uniformly on C to a continuous function ϕ, with $\phi = f$ on C. By Tietze extension theorem, ϕ can be extended to a continuous function on $[a,b]$. Hence, $\mu(\{\phi \neq f\}) \leq \mu([a,b]\setminus C) < \epsilon$. □

Definition 4.27 (Convergence in measure)**.** A sequence $\{f_n\}$ is said to converge to f in measure if

$$\forall \epsilon > 0, \lim_{n\to\infty} \mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) = 0.$$

$\{f_n\}$ is said to be Cauchy in measure if $\forall \epsilon > 0$ and $\delta > 0$, $\exists N$ such that $\forall n, m \geq N, \mu(\{x : |f_n(x) - f_m(x)| \geq \epsilon\}) < \delta$.

Proposition 4.28. *Let $f_n \xrightarrow{\mu} f$ and $g_n \xrightarrow{\mu} g$. Then $\forall \alpha, \beta \in \mathbb{R}$, $\alpha f_n + \beta g_n \xrightarrow{\mu} \alpha f + \beta g$.*

Proof.

$$|(\alpha f_n(x) + \beta g_n(x)) - (\alpha f(x) + \beta g(x))| \leq |\alpha||f_n(x) - f(x)| + |\beta||g_n(x) - g(x)|.$$

Hence

$$\mu(\{x : |(\alpha f_n(x)+\beta g_n(x))-(\alpha f(x)+\beta g(x))| \geq \epsilon\}) \leq \mu(\{x : |f_n(x)-f(x)| \geq \frac{\epsilon}{2\alpha}\})+\mu(\{x : |g_n(x)-g(x)| \geq \frac{\epsilon}{2\beta}\}).$$

□

Proposition 4.29. *Let $f_n \to f$. Then $|f_n| \to |f|$.*

Proof. The proof follows since $||f_n(x) - |f(x)|| \leq |f_n(x) - f(x)|$. □

Proposition 4.30. *Let $\mu(X) < \infty$ and f_n, f be real valued functions. Suppose $f_n \xrightarrow{\mu} f$ and $g_n \xrightarrow{\mu} g$. Then $f_n g_n \xrightarrow{\mu} fg$.*

Proof. It suffices to show that $f_n^2 \xrightarrow{\mu} f^2$. Suppose first that $f_n \xrightarrow{\mu} 0$.

$$\mu(\{x : |f_n(x)|^2) \geq \epsilon\} \leq \mu(\{x : |f_n(x)| \geq \sqrt{\epsilon}\}).$$

In general, if $f_n \xrightarrow{\mu} f$, then $f_n - f \xrightarrow{\mu} 0$. Let $E_n = \{x : |f(x)| > n\}$. Then $E_n \searrow \emptyset$. Since $\mu(X) < \infty, \lim_{n\to\infty} \mu(E_n) = 0$. Given $\delta > 0$, choose $m \in \mathbb{N}$ such that $\mu(E_n) < \delta \, \forall n \geq m$. On $E_m^c, |f| \geq m$. Now,

$$\{x : |f_n(x)f(x)-f^2(x)| \geq \epsilon\} = \{x : |f_n f(x)-f^2(x)| \geq \epsilon\}\cap E_m \cup \{x : |f_n f(x)-f^2(x)| \geq \epsilon\}\cap E_m^c.$$

On $E_m^c, |f_n f - f^2| \leq |f||f_n - f| \leq m|f_n - f|$. Hence, $|f_n - f| \geq \frac{\epsilon}{m}$ on the set $E_m^c \cap \{x : |f_n f - f^2| \geq \epsilon\}$. Hence,

$$\mu(\{x : |f_n f - f^2| \geq \epsilon\}) < \delta + \mu(\{x : |f_n - f| \geq \frac{\epsilon}{m}\}).$$

Hence $f_n f \xrightarrow{\mu} f^2$. Now $(f_n - f)^2 = f_n^2 - 2f_n f + f^2 = (f_n^2 - f^2) + 2(-f_n f + f^2)$. Hence $f_n^2 - f^2 \xrightarrow{\mu} 0$. □

Example 4.31. Proposition 4.30 is not true if $\mu(X) = \infty$. Let $X = \mathbb{N}, \mathcal{S} = \mathcal{P}(\mathbb{N})$ and μ be the counting emasure. Let

$$f_n(k) = \begin{cases} \frac{1}{n}, \, 1 \leq k \leq n \\ 0, \, k > n \end{cases}$$

Then $f_n \to 0$ uniformly, hence $f_n \xrightarrow{\mu} 0$. (This is true since $\{x : |f_n(x) - f(x)| \geq \epsilon\} = \phi$ for n large. Next, let $g(x) \equiv n$. Then for $\epsilon < 1$, $f_n g(n) = 1$. Hence $\mu(\{x : |f_n g(x)| \geq \epsilon\}) \geq 1 \,\forall n$. Hence $f_n g$ does not converge to 0 in measure.

Proposition 4.32. *Let $\mu(X) < \infty$. Suppose $\{f_n\}$ is a sequence of real valued functions and $f_n \to f$ almost everywhere, f a real valued function. Then $f_n \xrightarrow{\mu} f$.*

Proof. Let $D = \{x : f_n(x) \nrightarrow f(x)\}$. Then $\mu(D) = 0$.

Let $\epsilon > 0$ and $E_m(\epsilon) = \{x : |f_m(x) - f(x)| \geq \epsilon\}$. Then

$$\begin{aligned} D &= \bigcup_{\epsilon \geq 0} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m(\epsilon) \\ &= \bigcup_{\epsilon \geq 0} \limsup_{n\to\infty} E_n(\epsilon). \end{aligned}$$

Since $\mu(D) = 0$, $\forall \epsilon > 0, \mu(\{\limsup_{n\to\infty} E_n(\epsilon)\}) = 0$. Since $\mu(X) < \infty$,

$$\begin{aligned} 0 &= \mu(\{\limsup_{n\to\infty} E_n(\epsilon)\}) \\ &\geq \limsup \mu(E_n(\epsilon)) \\ &\geq \liminf \mu(E_n(\epsilon)) \\ &\geq 0. \end{aligned}$$

Hence $\lim_{n\to\infty} \mu(E_n(\epsilon)) = 0$. That is, $f_n \xrightarrow{\mu} f$. □

Example 4.33. Proposition 4.32 is not true if $\mu(X) = infty$. Let $X = \mathbb{N}, \mathcal{S} = \mathcal{P}(\mathbb{N})$ and μ be the counting measure. $f_n \xrightarrow{\mu} f \iff f_n \to f$ uniformly. Let $f_n = \chi_{\{1,2,\cdots,n\}}$. Then $f_n \to f \equiv 1$ pointwise, but not uniformly.

Example 4.34. Convergence in measure does not imply convergence almost everywhere, even in a finite measure space. Let $X = [0,1]$, mu be the Lebesgue measure. Let $\chi_n^i = \chi_{[\frac{i-1}{n}, \frac{i}{n}]}$. Let $x \in [0,1]$ and $n \in \mathbb{N}$. Consider the sequence $\{\chi_1^1, \chi_2^1, \chi 2^2, \chi 3^1, \chi_3^2, \chi_3^3, \cdots\}$. Then $\exists i$ such that $\chi_n^i(x) = 1$ and $\exists j$ such that $\chi_n^j(x) = 0$. Hence this sequence does not converge for any x. But it converges in measure since $\mu(\{x : |\chi_n^i(x)| \geq \epsilon\}) = \frac{1}{n} \to 0$.

Lemma 4.35 (Borel-Cantelli)**.** *Let $\{E_k\}$ be a sequence of measurable sets such that $\sum_{i=1}^{\infty} \mu(E_i) < \infty$. Then except on a set of measure 0, every x belongs to at most finitely many E_ks.*

Proof. Let $E = \{x : x$ belongs to infinitely many $E_k\}$. Then $E = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m$. Hence $\mu(E) \leq \mu(\bigcup_{m=n}^{\infty} E_m) \leq \sum_{m=n}^{\infty} \mu(E_m) \,\forall n$. Hence $\mu(E) = 0$. □

Proposition 4.36. *Let $f_n \xrightarrow{\mu} f$. Then there exists a subsequence which converges to f almost everywhere.*

Proof. Let $E_{n,m} = \{x : |f_n(x) - f(x)| \geq \frac{1}{m}\}$. Then $\forall m, \exists n_0(m)$ such that $\forall n \geq n_0(m), \mu(E_{n,m}) < \frac{1}{2^m}$. Then $\sum_{m=1}^{\infty} \mu(E_{n_0(m),m}) < \sum_{m=1}^{\infty} \frac{1}{2^m} < \infty$. By the Borel-Cantelli lemma, $\exists E$ with $\mu(E) = 0$ such that $\forall x \in E^c$, x belongs to at most finitely many sets $E_{n_0(m),m}$. That is, $\forall x \in E^c, \exists N$ (that depends on x) such that if $m \geq N$, then $x \notin E_{n_0(m),m}$. That is, $\forall m \geq N, |f_{n_0(m)}(x) - f(x)| < \frac{1}{m}$. That is, $f_{n_0(m)}(x) \to f(x) \, \forall x \in E^c, \mu(E) = 0$. □

Proposition 4.37. *If $f_n \xrightarrow{\mu} f$, then $\{f_n\}$ is Cauchy in measure. If also $f_n \xrightarrow{\mu} g$, then $f = g$ almost everywhere.*

Proof. The proof of the first part follows since $\{x : |f_n(x) - f_m(x)| \geq \epsilon\} \subseteq \{x : |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\} \cup \{x : |f_m(x) - f(x)| \geq \frac{\epsilon}{2}\}$. Suppose $f_n \xrightarrow{\mu}$ to f and g. Then for $\epsilon >$, $\mu(\{x : |f_n(x) - g(x)| \geq \epsilon\}) \leq \mu(\{x : |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}) + \mu(\{x : |f_n(x) - g(x)| \geq \frac{\epsilon}{2}\}) \, \forall n$. Taking the limit as $n \to \infty$, we get $\mu(\{x : |f(x) - g(x)| \geq \epsilon\}) = 0 \, \forall \epsilon > 0$. □

Proposition 4.38. *Suppose $\{f_n\}$ is Cauchy in measure. Then $\exists$ a subsequence such that $\{f_{n_k}\}$ is almost uniformly Cauchy.*

Proof. Given $k \in \mathbb{N}$, $\exists n(k)$ such that $\forall n, m \geq n(k)$, $\mu(\{x : |f_n(x) - f_m(x)| \geq \frac{1}{2^k}\}) < \frac{1}{2^k}$. Let

$$n_1 = n(1) \geq 1$$
$$n_2 = \max\{n_1 + 1, n(2)\} \geq 2$$
$$n_3 = \max\{n_2 + 1, n(3)\} \geq 3$$

and so on. Let $E_k = \{x : |f_{n_k}(x) - f_{n_{k+1}}(x)| \geq \frac{1}{2^k}\}$. Then $\mu(E_k) < \frac{1}{2^k}$. Given $\delta > 0$, choose k such that $\frac{1}{2^{k-1}} < \delta$. Let $F = E_k \cup E_{k+1} \cup \cdots$. Then $\mu(F) \leq \sum_{i=1}^{\infty} \mu(E_i) = \frac{1}{2^{k-1}} < \delta$. Let $\epsilon > 0$. Choose $i \geq k$ such that $\frac{1}{2^{i-1}} < \epsilon$. Let $i \leq l \leq m, x \in F^c = \bigcap_{j=k} \infty E_j^c$. Then

$$\begin{aligned}
|f_{n_l}(x) - f_{n_m}(x)| &\leq \sum_{j=l}^{m-1} |f_{n_j}(x) - f_{n_{j+1}}(x)| \\
&< \sum_{j=l}^{\infty} \frac{1}{2^j} \\
&= \frac{1}{2^{l-1}} \\
&< \frac{1}{2^{i-1}} \\
&< \epsilon.
\end{aligned}$$

Hence $\{f_{n_k}\}$ is Cauchy on F^c and $\mu(F) < \delta$.

□

Proposition 4.39. *Suppose $\{f_n\}$ is Cauchy in measure. Then there exists f measurable such that $f_n \xrightarrow{\mu} f$.*

Proof. By Proposition 4.1, let $\{f_{n_k}\}$ be a subsequence which is almost uniformly Cauchy. Then $\{f_{n_k}\}$ is Cauchy almost everywhere. Hence there exists a measurable function f such that $f_{n_k} \to f$ almsot everywhere. But this implies that $f_{n_k} \to f$ almost uniformly. Now, $\{x : |f_n(x) - f(x)| \geq \epsilon\} \subseteq \{x : |f_n(x) - f_{n_k}(x)| \geq \frac{\epsilon}{2}\} \cup \{x : |f_n(x) - f(x)| \geq \frac{\epsilon}{2}\}$. Let $\eta > 0$. Then $\exists N_1 \in \mathbb{N}$ such that $n, n_k \geq N_1 \Rightarrow \mu(\{x : |f_n(x) - f_{n_k}(x)| \geq \epsilon\}) < \frac{\eta}{2}$ since $\{f_n\}$ is Cauchy in measure. Also, $f_{n_k} \to f$ almost uniformly implies that there exists a set E of measure less than $\frac{\eta}{2}$ and $N_2 \in \mathbb{N}$ such that $\forall n_k \geq N_2, |f_{n_k}(x) - f(x)| < \epsilon$ on E^c. That is, $\forall n_k \geq N_2, \{x : |f_{n_k}(x) - f(x)| \geq \epsilon\} \subset E$. Take $N = \max\{N_1, N_2\}$. Then $\forall n \geq N$, $\mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) < \eta$. Hence $f_n \xrightarrow{\mu} f$. □

Proposition 4.40. *Suppose $f_n \to f$ almost uniformly. Then $f_n \xrightarrow{\mu} f$.*

Proof. Let $\epsilon, \delta > 0$. Then $\exists F$, $\mu(F) < \delta$ such that $f_n \to f$ uniformly on F^c. Then $\exists n_0(\epsilon)$ such that $\forall n \geq n_0(\epsilon)$ and $\forall x \in F^c$, $|f_n(x) - f(x)| < \epsilon$. That is, $\forall n \geq n_0(\epsilon), \{x : |f_n(x) - f(x)| \geq \epsilon\} \subseteq F$. Hence $\mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) < \delta$. Hence $\lim\limits_{n\to\infty} \mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) = 0$. That is, $f_n \xrightarrow{\mu} f$. □

Finally we have the following implications:

Convergence almost everywhere $\xRightarrow[\text{Egoroff}]{\mu(X)<\infty}$ Convergence almost uniformly$\Rightarrow$ Convergence almost everywhere.
Convergence almost everywhere $\xRightarrow{\mu(X)<\infty}$ Convergence in measure $\xRightarrow{\text{subsequence}}$ Convergence almost everywhere.
Convergence in measure $\xRightarrow{\text{subsequence}}$ Convergence almost uniformly $\Rightarrow$ Convergence in measure.

5. Integration

Let $(X, \mathcal{S}, \mu)$ be a measure space. Suppose ϕ is a function on X such that its range is a finite set $\{\alpha_1, \cdots, \alpha_n\}$ with each $\alpha_i \geq 0$. Let $A_i = \phi^{-1}(\alpha_i)$. Then $\phi = \sum\limits_{i=1}^{n} \alpha_i \chi_{A_i}$ where the A_is are disjoint. Such a function is called a simple function. Define the integral of ϕ as $\int\limits_X \phi\, d\mu = \sum\limits_{i=1}^{n} \alpha_i \mu(A_i)$.

If $\phi = \sum\limits_{j=1}^{m} \beta_j \chi_{B_j}$, where the B_js are disjoint, then each B_j is contained in some A_i and in the case $\beta_j = \alpha_i$. $A_i = \bigcup\limits_{B_j \subset A_i} B_j$. Hence $\mu(A_i) = \sum\limits_{B_j \subset A_i} \mu(B_j)$. Hence $\sum\limits_{i=1}^{n} \alpha_i \mu(A_i) = \sum\limits_{j=1}^{m} \beta_j \mu(B_j)$.

Now, let us consider the case of general E_i which need not be disjoint. Suppose $\phi = \sum\limits_{i=1}^{N} c_i \chi_{E_i}$. Let $A_i, \cdots, A_n \subset X$. Let $A^\epsilon = \begin{cases} A, \, \epsilon = 1 \\ A^c, \, \epsilon = -1 \end{cases}$

Let $\epsilon = (\epsilon_1, \cdots, \epsilon_n)$, $\epsilon_i = \pm 1$ and $\epsilon_0 = (-1, \cdots, -1)$. Let $A_\epsilon = \bigcap\limits_{i=1} 6n A_i^{\epsilon_1}$, $A_{\epsilon_0} = \bigcap\limits_{i=1}^{n} A_i^c = \bigcup\limits_{i=1}^{n} (A_i)^c$. Now $\epsilon = \eta \iff \epsilon_i = \eta_i \, \forall i$. $\epsilon \neq \eta \Rightarrow A_\epsilon \cap A_\eta = \emptyset$. We claim that $A_i = \bigcup\limits_{\epsilon_i = 1 \text{ and } \epsilon \neq \epsilon_0} A_\epsilon$. Clearly, $\bigcup\limits_{\epsilon_i = 1} A_\epsilon \subset A_i$. Conversely, if $x \in A_i$. Let $\epsilon_k = \begin{cases} +1, \, x \in A_k \text{ (Hence } \epsilon_i = 1) \\ -1, \, x \notin A_k. \end{cases}$.

Let $\epsilon = (\epsilon_k)$. Then $x \in A_\epsilon$. A_ϵ are disjoint. $\phi = \sum\limits_{i=1}^{m} c_i \chi_{E_i}$. Now, $E_i = \bigcup\limits_{\epsilon_i=1 \text{ and } \epsilon\neq\epsilon_0} E_\epsilon$, where the E_ϵ are disjoint. Hence $\chi_{E_i} = \sum\limits_{\epsilon_i=1 \text{ and } \epsilon\neq\epsilon_0} \chi_{E_\epsilon}$. Hence $\phi = \sum\limits_{i+1}^{m} c_i \sum\limits_{\epsilon_i=1 \text{ and } \epsilon\neq\epsilon_0} \chi_{E_\epsilon} = \sum\limits_{\epsilon\neq\epsilon_0} \sum\limits_{i=1 \text{ and } \epsilon_i=1}^{m} \chi_{E_\epsilon}$.

Call $\sum\limits_{i=1,\epsilon_i=1}^{m} c_i$ as $c(\epsilon)$. We now have a disjoint partition. Hence, by definition,

$$\begin{aligned}
\int\limits_X \phi\, d\mu &= \sum_{\epsilon\neq\epsilon_0} c(\epsilon)\mu(E_\epsilon) \\
&= \sum_{\epsilon\neq\epsilon_0} (\sum_{i=1,\epsilon_i=1}^{m} c_i)\mu(E_\epsilon) \\
&= \sum_{i=1}^{m} c_i \sum_{\epsilon_i=1} \mu(E_\epsilon) \\
&= \sum_{i=1}^{m} c_i\mu(E_i).
\end{aligned}$$

So we have proved that the definition holds even when E_i are not disjoint. If $\phi = \sum\limits_{i=1}^{n} \alpha_i \chi_{A_i}$ and $E \subset X$, then $\int\limits_E \phi\, d\mu = \sum\limits_{i=1}^{n} \alpha_i \mu(A_i \cap E)$. This can be shown to be equal to $\int\limits_X \phi\chi_E\, d\mu$.

Now, let $f \geq 0$ be a measurable function. Let $E \subset X$. Define $\int\limits_E f\, d\mu = \sup\limits_{\phi \text{ simple}, 0\leq\phi\leq f} \int\limits_E \phi\, d\mu$. It can be seen that the two definitions of integration coincide for simple functions.

Proposition 5.1. *Let $E \subseteq X$. Then*

(1) $0 \leq f \leq g \Rightarrow \int\limits_E f\, d\mu \leq \int\limits_E g\, d\mu$.
(2) $A \subseteq B, f \geq 0 \Rightarrow \int\limits_A f\, d\mu \leq \int\limits_B f\, d\mu$.
(3) $f \geq 0, 0 \leq c < \infty \Rightarrow \int\limits_E cf\, d\mu = c\int\limits_E f\, d\mu$.
(4) $f \equiv 0$ *on* $E \Rightarrow \int\limits_E f\, d\mu$.
(5) $\mu(E) = 0, f \geq 0 \Rightarrow \int\limits_E f\, d\mu = 0$.
(6) $\int\limits_E f\, d\mu = \int\limits_X f\chi_E\, d\mu$.

Proposition 5.2. *(1) Let $\phi \geq 0$ be simple. Define $\nu(E) = \int\limits_E \phi\, d\mu$, $E \in \mathcal{S}$. Then ν is a measure on $\mathcal{S}$.*
(2) If $\phi, \psi \geq 0$ are simple functions, then $\int\limits_X (\phi + \psi)\, d\mu = \int\limits_X \phi\, d\mu + \int\limits_X \psi\, d\mu$.

Proof. (1) It is clear that $\nu \geq 0$ and $\nu(\emptyset) = 0$. We need only prove countable additivity. Let $E = \bigcup_{i=1}^{\infty} E_i$, disjoint union. Let $\phi = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$. Then

$$\begin{aligned}
\nu(E) &= \int_E \phi \, d\mu \\
&= \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E) \\
&= \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{\infty} \mu(A_i \cap E_j) \\
&= \sum_{j=1}^{\infty} \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E_j) \\
&= \sum_{j=1}^{\infty} \int_{E_j} \phi \, d\mu \\
&= \sum_{j=1}^{\infty} \nu(E_j).
\end{aligned}$$

(2) Let $\phi = \sum_{i=1}^{n} \alpha_i \chi_{A_i}$ and $\psi = \sum_{j=1}^{m} \beta_j \chi_{B_j}$. Let $E_{ij} = A_i \cap B_j$. Then

$$\begin{aligned}
\int_{E_{ij}} (\phi + \psi) \, d\mu &= (\alpha_i + \beta_j)\mu(E_{ij}) \\
&= \int_{E_{ij}} \phi \, d\mu + \int_{E_{ij}} \psi \, d\mu.
\end{aligned}$$

Since X is the disjoint union of sets of the form E_{ij}, the result follows from (1).

□

Theorem 5.3 (Lebesgue Monotone Convergence)**.** *Let $\{f_n\}$ be a sequence of non-zero measurable functions such that*

(1) $0 \leq f_1(x) \leq f_2(x) \leq \cdots \ \forall x$
(2) $f_n \to f$ *pointwise almost everywhere.*

Then $\int_X f_n \, d\mu \to \int_X f \, d\mu$.

Proof. Since $f_n \leq f \quad \forall n$, by (1) of Proposition 5.1, $\int_X f_n \, d\mu \leq \int_X f \, d\mu \quad \forall n$. Let $\alpha = \sup_n \int_X f_n \, d\mu$. Then $\alpha \leq \int_X f \, d\mu$. Now, let ϕ be a simple function such that $0 \leq \phi \leq f$. Let $0 < c < 1$. Define $E_n = \{x : f_n(x) \geq c\phi(x)\}$. Then $E_i \subset E_2 \subset \cdots$. If $f(x) = 0$, then $\phi(x) = 0 \Rightarrow x \in E_1$. If $f(x) > 0$, then $c\phi(x) < f(x)$. Since $f_n(x) \to f(x), \exists n$ such that $c\phi(x) \leq f_n(x) \leq f(x)$. Hence $x \in E_n$. So $X = \bigcup_{n=1}^{\infty} E_n$. Now $\int_X f_n \, d\mu \geq \int_{E_n} f_n \, d\mu \geq c \int_{E_n} \phi \, d\mu = c\nu(E_n)$. where ν is a measure as defined in Proposition 5.2. Hence $\alpha \geq c \lim_{n\to\infty} \nu(E_n) = c\nu(X) = c\int_X \phi\, d\mu$. Since this is true for every simple function ϕ such that $0 \leq \phi \leq f$, we get $\alpha \geq c\int_X f\, d\mu$. As $c \to \infty, \alpha \geq \int_X f\, d\mu$. Hence $\lim_{n\to\infty} \int_X f_n\, d\mu = \sup_n \int_X f_n\, d\mu = \int_X f\, d\mu$. □

Proposition 5.4. *Let $f_n \geq 0$ be measurable. Then $f(x) = \sum_{n=1}^{\infty} f_n(x)$ is measurable and non-negative. Further, $\int_X f\, d\mu = \sum_{n=1}^{\infty} \int_X f_n\, d\mu$.*

Proof. Let $\phi_n^1 \nearrow f_1, 0 \leq \phi_n^1$ simple, and $\phi_n^2 \nearrow f_2, 0 \leq \phi_n^2$, simple. Then $\phi_n^1 + \phi_n^2 \nearrow f_1 + f_2$. Since $\int_X (\phi_n^1 + \phi_n^2)\, d\mu = \int_X \phi_n^1\, d\mu + \int_X \phi_n^2\, d\mu$, we get $\int_X (f_1 + f_2)\, d\mu = \int_X f_1\, d\mu + \int_X f_2\, d\mu$. By induction, $\int_X (f_1 + \cdots + f_n)\, d\mu = \sum_{i=1}^{n} \int_X f_i\, d\mu$. Let $g_n = \sum_{i=1}^{n}$. Then $g_n \geq 0$ and $g_n \nearrow f$. By the monotone convergence theorem, $\int_X f\, d\mu = \sum_{i=1}^{\infty} \int_X f_i\, d\mu$. □

Example 5.5. Let $X = \mathbb{N}, \mathcal{S} = \mathcal{P}(\mathbb{N})$ and μ be the counting measure. Let $E = \{n_1, \cdots, n_k\}$ be a finite set. Let $f : X \to \mathbb{R}$ a be a non-negative function. It is then a sequence. $f = \sum_{i=1}^{k} f(n_i)\chi_{\{n_i\}}$. Then $\int_E f\, d\mu = \sum_{i=1}^{k} f(n_i)$. Suppose $f \geq 0$ on $\mathbb{N}$. Then we can think of $f_n = f \upharpoonright \{1, \cdots, n\}$. Then $f_n \nearrow f$. By the monotone convergence theorem,

$$\int_X f\, d\mu = \sum_{n=1}^{\infty} f(n).$$

Example 5.6. Fix $x_0 \in X$. Let

$$\delta_{x_0}(E) = \begin{cases} 1, \ , x_0 \in W \\ 0, \ x_0 \notin E \end{cases}$$

Let $\phi = \sum_{i=1}^{n} \alpha_i \chi_{E_i}$, E_i disjoint. Then x_0 belongs to at most one E_i. In this case, $\int_X \phi\, d\mu = \alpha_i$, i.e., $\int_X \phi\, d\mu = \phi(x_0)$.

For $f \geq 0$, consider a sequence of simple functions $\{\phi_n\}$ that increase to f. Then $\int_X f\, d\mu = \lim \int_X \phi_n\, d\mu = \lim \phi_n(x_0) = f(x_0)$.

Example 5.7. Let $\{a_{ij}\}$ be a double sequence of non-negative numbers. Then

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}.$$

Let $X = \mathbb{N}, \mathcal{S} = \mathcal{P}(\mathbb{N})$ and μ be the counting measure. Let $f_i(j) = a_{ij}$ and $f = \sum_{i=1}^{\infty} f_i$. By Proposition 5.4,

$$\int_X f\, d\mu = \sum_{i=1}^{\infty} \int_X f_i\, d\mu.$$

That is,

$$\sum_{j=1}^{\infty} f(j) = \sum_{i=1}^{\infty} \int_X f_i\, d\mu.$$

Hence

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} f_i(j) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} f_i(j).$$

Example 5.8. Consider $I = \int\limits_0^1 \frac{\sin^{-1} x}{\sqrt{1-x^2}}\, dx$. Taking the transformation $y = \sin^{-1} x$, we get $I = \frac{\pi^2}{8}$.

Now, the Taylor expansion of $\sin^{-1} x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 2}{2\cdot 4}\frac{x^5}{5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7} + \cdots$ By Proposition 5.4, we get

$$\begin{aligned} I &= \int\limits_0^1 \frac{x}{\sqrt{1-x^2}} + \sum_{n=1}^{\infty} \frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\frac{1}{2n+1}\int\limits_0^1 \frac{x^{2n+1}}{\sqrt{1-x^2}}\, dx \\ &= 1 + \sum_{n=1}^{\infty} \frac{1\cdot 3\cdots(2n-1)}{2\cdot 4\cdots 2n}\frac{1}{2n+1}\frac{2n\cdot(2n-1)\cdots 2}{(2n+1)\cdot(2n-1)\cdots 3} \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n+1)^2} \end{aligned}$$

.

Hence $\frac{\pi^2}{8} = \sum\limits_{n=0}^{\infty} \frac{1}{(2n+1)^2}$. If $S = \sum\limits_{n=1}^{\infty}$, then we get $S = \frac{\pi^2}{6}$.

Theorem 5.9 (Fatou's Lemma)**.** *Let $f_n \geq 0$ be a sequence of measurable functions. Then*

$$\int\limits_X (\mathit{lim}\ \mathit{inf}_{n\to\infty} f_n)\, d\mu \leq \mathit{lim}\ \mathit{inf}_{n\to\infty} \int\limits_X f_n\, d\mu.$$

Proof. Let $g_k(x) = \inf\limits_{i\geq k} f_i(x)$. Then $g_k \geq 0$ and $\lim\limits_{k\to\infty} g_k = \lim\limits_{n} \inf f_n$. By the Monotone convergence theorem, we get

$$\int\limits_X (\liminf_{n\to\infty} f_n)\, d\mu = \lim_{n\to\infty} \int\limits_X g_k\, d\mu.$$

But

$$\int\limits_X g_k\, d\mu \leq \inf_{i\geq k} \int\limits_X f_i\, d\mu.$$

Hence the result follows. □

Example 5.10. Strict inequality can occur in Fatou's lemma. Let $X = \mathbb{R}, \mu$ be the Lebesgue measure. Let $f_n = \chi_{[n,n+1]}$. Then $\lim\limits_{n\to\infty} f_n(x) = 0\, \forall x \in \mathbb{R}$, but $\int\limits_X f_n\, d\mu = 1\, \forall n$.

Proposition 5.11. *Let $f \geq 0$ be measurable. Define*

$$\nu(E) = \int\limits_E f\, d\mu.$$

Then ν is a measure. Further, if $g \geq 0$,

$$\int\limits_X g\, d\nu = \int\limits_X fg\, d\mu.$$

The notation used is $d\nu = f d\mu$, or $\frac{d\nu}{d\mu} = f$.

Proof. Clearly, $\nu \geq 0$ and $\nu(\emptyset) = 0$. Suppose $E = \bigcup_{i=1}^{\infty} E_i$ disjoint union. Then $\nu(E) = \int_E f\,d\mu = \int_X f\chi_E\,d\mu$. Now, $\chi_E = \sum_{i=1}^{\infty} \chi_{E_i}$. Hence

$$\begin{aligned}\int_X f\chi_E\,d\mu &= \int_X \sum_{i=1}^{\infty} f\chi_{E_i}\,d\mu \\ &= \sum_{i=1}^{\infty} \int_X f\chi_{E_i}\,d\mu \\ &= \sum_{i=1}^{\infty} \int_{E_i} f\,d\mu \\ &= \sum_{i=1}^{\infty} \nu(E_i).\end{aligned}$$

Next, let $g = \chi_E$. Then

$$\begin{aligned}\int_X g\,d\nu &= \int_X \chi_E\,d\nu \\ &= \nu(E) \\ &= \int_E f\,d\mu \\ &= \int_X f\chi_E\,d\mu \\ &= \int_X fg\,d\mu.\end{aligned}$$

By linearity, the above also holds for simple functions, and by the Monotone convergence theorem, it holds for all positive measurable functions. □

Example 5.12. Let $f \geq 0$ and $\int_X f\,d\mu = 0$. Then $f = 0$ almost everywhere. Let $E_n = \{x : f(x) \geq \frac{1}{n}\}$, Then $E = \{x : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$. Now $0 = \int_X f\,d\mu \geq \int_{E_n} f_n\,d\mu \geq \frac{1}{n}\mu(E_n)$. Hence $\mu(E_n) = 0\,\forall n \Rightarrow \mu(E) = 0$.

Definition 5.13. A function f is said to be integrable if $\int_X |f|\,d\mu < \infty$. A real-valued f can be written as $f^+ - f^-$, where $f^+, f^- \geq 0$. Hence we define the integral of f as follows:

$$\int_X f\,d\mu = \int_X f^+\,d\mu - \int_X f^-\,d\mu.$$

Suppose f is a complex valued function. Then $f = u + iv$, where u, v are real-valued functions. Then $\int_X |f|\,d\mu < \infty \iff \int_X |u|\,d\mu < \infty$ and $\int_X |v|\,d\mu < \infty$. Define

$$\int_X f\,d\mu = \int_X u\,d\mu + i\int_X v\,d\mu.$$

Theorem 5.14. *Let f, g be complex integrable functions and $\alpha, \beta \in \mathbb{C}$. Then $\alpha f + \beta g$ is integrable and*

$$\int_X \alpha f + \beta g \, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu.$$

Proof. $|\alpha f + \beta g| \leq |\alpha||f| + |\beta||g| \Rightarrow \alpha f + \beta g$ is integrable. It suffices to show the following:

(1) $\int_X (f+g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu \, \forall f, g$ real
(2) $\int_X \alpha f \, d\mu = \alpha \int_X f \, d\mu \, \forall \alpha \in \mathbb{C}$.

Let f, g be real valued and $h = f + g$. Then $h^+ - h^- = f^+ - f^- + g^+ - g^-$. That is, $h^+ + f^- + g^- = h^- + f^+ + g^+$, where each term is non-negative. Hence,

$$\int_X h^+ \, d\mu + \int_X f^- \, d\mu + \int_X g^- \, d\mu = \int_X h^- \, d\mu + \int_X f^+ \, d\mu + \int_X g^+ \, d\mu.$$

That is,

$$\int_X h^+ \, d\mu - \int_X h^- \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu + \int_X g^+ \, d\mu - \int_X g^- \, d\mu$$

as required. Next, let $\alpha \geq 0$. Then $\int_X \alpha f \, d\mu = \int_X (\alpha f)^+ - (\alpha f)^- \, d\mu = \int_X \alpha f^+ \, d\mu - \int_X \alpha f^- \, d\mu = \alpha(\int_X f^+ \, d\mu - \int_X f^- \, d\mu) = \alpha \int_X f \, d\mu$.

Now, $-f = f^- - f^+$. Hence $\int_X -f \, d\mu = \int_X f^- \, d\mu - \int_X f^+ \, d\mu = -\int_X f \, d\mu$.

Finally,

$$\begin{aligned}
\int_X if \, d\mu &= \int_X i(u+iv) \, d\mu \\
&= \int_X iu - v \, d\mu \\
&= -\int_X v \, d\mu + i \int_X u \, d\mu \\
&= i(\int_X u \, d\mu + i \int_X v \, d\mu) \\
&= i \int_X f \, d\mu.
\end{aligned}$$

□

Theorem 5.15. $|\int_X f \, d\mu| \leq \int_X |f| \, d\mu$.

Proof. Let $\alpha = \int_X f \, d\mu$. Then $|\alpha| e^{i\theta} = \alpha$, for some $\theta \in [0, 2\pi]$. That is,

$$|\int_X f \, d\mu| = e^{-i\theta} \int_X f \, d\mu = \int_X e^{-i\theta} f \, d\mu. \quad (6)$$

Let $g = e^{-i\theta} f = u + iv$. But by (6), we get $\int_X g\, d\mu = \int_X e\, d\mu$. Also, $u \leq |u| \leq |g| = |f|$. Hence $|\int_X f\, d\mu| \leq \int_X |f|\, d\mu$. □

Theorem 5.16 (Dominated Convergence Theorem)**.** *Let $f_n \to f$ almost everywhere and suppose $|f_n| \leq g$ almost everywhere, where g is an integrable function. Then*

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

Hence,

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

Proof. $\|f_n| \leq g\, \forall n \Rightarrow |f| \leq g$. Hence $|f_n - f| \leq 2g$. Thus, $2g - |f_n - f| \geq 0$. By Fatou's lemma,

$$\int_X \lim_{n\to\infty} (2g - |f_n - f|\, d\mu \leq \liminf_{n\to\infty} 2g - |f_n - f|\, d\mu.$$

Hence

$$\begin{aligned}\int_X 2g\, d\mu &\leq \liminf_{n\to\infty} \int_X 2g\, d\mu - \int_X |f_n - f|\, d\mu \\ &= \int_X 2g\, d\mu - \limsup_{n\to\infty} \int_X |f_n - f|\, d\mu.\end{aligned}$$

This implies that $\limsup_{n\to\infty} \int_X |f_n - f|\, d\mu \leq 0$, since g is integrable. Hence we get

$$0 \leq \liminf_{n\to\infty} \int_X |f_n - f|\, d\mu \leq \limsup_{n\to\infty} \int_X |f_n - f|\, d\mu \leq 0.$$

Hence

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

□

Example 5.17. Consider $\mathbb{N}, \mathcal{P}(\mathbb{N})$ and the counting measure.

(1) Let

$$f_n(k) = \begin{cases} \frac{1}{n},\ 1 \leq k \leq n \\ 0,\ k > n \end{cases}$$

Then $f_n \to 0$ uniformly, but $\int_{\mathbb{N}} f_n\, d\mu = 1\, \forall n$. But in this case, f_ns are not bounded by an integrable function.

Let $f_n(k) = \begin{cases} \frac{1}{k},\ 1 \leq k \leq n \\ 0,\ k > n \end{cases}$ Then $f_n \to f$, where $f(k) = \frac{1}{k}\, \forall k$, which is not integrable. However, by Montonoce convergence theorem, $f_n \nearrow f$.

Theorem 5.18. *Let f be a bounded function on $[a, b]$. Then*

(1) *If f is Riemann integrable, then it is Lebesgue integrable, and*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f\,d\mu$$

where μ is the Lebesgue measure.

(2) *f is Riemann integrable iff f is continuous almost everywhere.*

Proof. Let $\{P_k\}$ be a sequence of partitions of $[a,b]$ such that the mesh $\Delta P_k = \max_{1\le i\le n} |x_k - x_{k-1}| \to 0$. Each P_{k+1} is a refinement of P_k. Let $P_k = \{a = x_0 < x_1 < \cdots < x_n = b\}$. Define functions U_k, L_k as follows:

$$U_k(a) = L_k(a) = f(a)$$

and

$$U_k(x) = M_i$$
$$L_k(x) = m_i$$

if $x \in [x_{i-1}, x_i]$ and $M_i = \sup_{x\in[x_{i-1},x_i]} f(x)$ and $m_i = \inf_{x\in[x_{i-1},x_i]} f(x)$.

Now $\int_{[a,b]} U_k\,d\mu = U(P,f)$ and $\int_{[a,b]} L_k\,d\mu = L(P,f)$. Also,

$$U_1(x) \ge U_2(x) \ge \cdots \ge f(x) \ge \cdots \ge L_2(x) \ge L_1(x). \tag{7}$$

Since f is bounded, we have $\int_{[a,b]} U_1\,d\mu < \infty$. Now, $U_k \to U \ge f$, $L_k \to L \le f$, and by the dominated convergence theorem, all the functions in (7) are integrable. Further, $\int_{[a,b]} U_k\,d\mu \to \int_{[a,b]} U\,d\mu$ and $\int_{[a,b]} L_k\,d\mu \to \int_{[a,b]} L\,d\mu$. Since f is Riemann integrable, $\int_{[a,b]} U_k\,d\mu = U(P_k,f) \to \int_a^b f(x)\,dx$ as $\Delta P_k \to 0$ and $\int_{[a,b]} L_k\,d\mu = L(P_k,f) \to \int_a^b f(x)\,dx$ as $\Delta P_k \to 0$. Hence

$$\int_a^b f(x)\,dx = \int_{[a,b]} U\,d\mu \ge \int_{[a,b]} f\,d\mu \ge \int_{[a,b]} L\,d\mu = \int_a^b f(x)\,dx.$$

Hence, $\int_{[a,b]} f\,d\mu = \int_a^b f(x)\,dx$ if f is Riemann integrable. Also, $f = U = L$ almost everywhere. By further throwing away the partition points, f is continuous almost everywhere.

Conversely, if f is continuous almost everywhere, we get $f = U = L$ at all those points. f continuous at a point x implies that for $\epsilon > 0, \exists\delta > 0$ such that $|x - z| < \delta \Rightarrow |f(x) - f(z)| < \frac{\epsilon}{2}$. Choose k such that $\Delta P_k < \delta$. Take x a point that is not a partition point of $P_k = \{a = x_0^k < x_1^k < \cdots < x^k n = b\}$. Then $x \in [x_{i-1}^k, x_i^k]$ for some i. Now, for each $z \in [x_{i-1}^k, x_i^k], |f(x) - f(z)| < \frac{\epsilon}{2}$. Hence $|M_i - f(x)| \le \frac{\epsilon}{2}$ and $|m_i - f(x)| \le \frac{\epsilon}{2}$. So, $|M_i - m_i| \le \epsilon$. This, U_k, L_k tend to each other and $U = L = f$. If $f = U = L$, then by the dominated convergence theorem, $\int_{[a,b]} U_k\,d\mu \to \int_{[a,b]} f\,d\mu$ and $\int_{[a,b]} L_k\,d\mu \to \int_{[a,b]} f\,d\mu$. Hence, $|\int_{[a,b]} U_k\,d\mu - \int_{[a,b]} L_k\,d\mu| \to 0$. □

Example 5.19. Let $X = [0, \infty)$ with Lebesgue measure. Let $f(x) = \frac{\sin x}{x}$. Then the Riemann integral of f exists and is finite (equal to $\frac{\pi}{2}$). However, f is not Lebesgue integrable, i.e., $\int\limits_0^\infty |\frac{\sin x}{x}|\, dx = \infty$. $\int\limits_0^\infty |\frac{\sin x}{x}|\, dx \geq \sum\limits_{n=1}^\infty \int\limits_{n\pi+\frac{\pi}{4}}^{n\pi+\frac{\pi}{2}} |\frac{\sin x}{x}|\, dx$. On $[n\pi+\frac{\pi}{4}, n\pi+\frac{\pi}{2}]$, $|\sin x| \geq \frac{1}{\sqrt{2}}$ and $x = |x| \leq n\pi + \frac{\pi}{2} = (2n+1)\frac{\pi}{2}$, hence $\frac{1}{|x|} \geq \frac{2}{\pi(2n+1)}$. Thus

$$\int\limits_0^\infty |\frac{\sin x}{x}|\, dx \geq \sum_{n=1}^\infty \frac{\sqrt{2}}{\pi} \frac{1}{(2n+1)} \frac{\pi}{4} = \frac{1}{2\sqrt{2}} \sum_{n=1}^\infty \frac{1}{2n+1} = \infty.$$

Example 5.20. (1) Let $X = (0,1)$ and $f(x) = \frac{1}{\sqrt{x}}$. Let

$$f_n(x) = \begin{cases} 0,\ 0 < x < \frac{1}{n} \\ \frac{1}{\sqrt{x}},\ \frac{1}{n} < x < 1 \end{cases}$$

Then $\{f_n\}$ increases to f. By the monotone convergence theorem,

$$\begin{aligned} \int\limits_0^1 f(x)\, dx &= \lim_{n\to\infty} \int\limits_0^1 f_n(x)\, dx \\ &= \lim_{n\to\infty} \int\limits_{\frac{1}{n}}^1 \frac{1}{\sqrt{x}}\, dx \\ &= \lim_{n\to\infty} [2\sqrt{x}]_{\frac{1}{n}}^1 \\ &= \lim_{n\to\infty} [2 - \frac{2}{\sqrt{n}}] \\ &= 2. \end{aligned}$$

(2) Let $X = (1, \infty)$ and $f(x) = \frac{1}{\sqrt{x}}$. Let

$$f_n(x) = \begin{cases} \frac{1}{\sqrt{x}},\ x \in (1, n) \\ 0,\ x \in [n, \infty) \end{cases}$$

Then $\{f_n\}$ increases to f. Hence $\int\limits_1^\infty \frac{1}{\sqrt{x}}\, dx = \lim\limits_{n\to\infty} \int\limits_1^n \frac{1}{\sqrt{x}}\, dx = \lim\limits_{n\to\infty} [2\sqrt{x}]_1^n = \lim\limits_{n\to\infty} [2\sqrt{n} - 2] = \infty$.

Example 5.21. Let $f : \mathbb{R} \to \mathbb{R}$ be integrable and $t \in \mathbb{R}$ be fixed. Then

$$\int\limits_\infty^\infty f(x+t)\, dx = \int\limits_\infty^\infty f(x)\, dx.$$

First, let $f = \chi_E$. Then

$$f(x+t) = \begin{cases} 1,\ x+t \in E \\ 0,\ x+t \notin E \end{cases} = \chi_{E-t}(x)$$

Hence

$$\int\limits_\infty^\infty f(x+t)\, dx = \mu(E-t) = \mu(E) = \int\limits_\infty^\infty f(x)\, dx.$$

By linearity, the result is true for simple functions, by the monotone convergence theorem for non-negative functions, and then by linearity for complex valued functions.

Example 5.22. Let $f : [0,1] \times [0,1] \to \mathbb{R}$. Suppose

(1) f is measurable with respect to x for each fixed t.
(2) $|f(x,t)| \leq g(x)$, where g is integrable.
(3) $f(x,t) \to \phi(x)$ as $t \to 0$.

Then

$$\lim_{t\to 0} \int_0^1 f(x,t)\,dx = \int_0^1 \phi(x)\,dx.$$

Proof. Let $t_n \to 0$. Let $\phi_n(x) = f(x,t_n)$. Then $phi_n(x) \to \phi(x)$, and $|\phi(x)| \leq g(x)$, g integrable. By the dominated convergence theorem, $\int_0^1 \phi_n(x)\,dx \to \int_0^1 \phi(x)\,dx$. □

Example 5.23. Let $f : [0,1] \times [0,1] \to \mathbb{R}$. Suppose

(1) f is measurable with respect to x for each fixed t.
(2) f is continuous with respect to t for each fixed x.
(3) $|f(x,t)| \leq g(x)$, where g is integrable.
(4) $f(x,t) \to \phi(x)$ as $t \to 0$.

Further, suppose $\frac{\partial f}{\partial t}(x,t)$ exists for all $x, t \in [0,1]$, is continuous with respect to t and $|\frac{\partial f}{\partial t}(x,t)| \leq M$. Let $h(t) = \int_0^1 f(x,t)\,dx$. Then h is differentiable and $\frac{dh}{dt} = \int_0^1 \frac{\partial f}{\partial t}(x,t)\,dx$. That is,

$$\frac{d}{dt} \int_0^1 f(x,t)\,dx = \int_0^1 \frac{\partial f}{\partial t}(x,t)\,dx.$$

Proof. $\frac{h(t+\tau)-h(t)}{\tau} = \int_0^1 \frac{f(x,t+\tau)-f(x,t)}{\tau}\,dx = \int_0^1 \frac{\partial f(x,t+\theta\tau)}{\partial t}\,dx$, by the mean value theorem. Now, $\frac{\partial f(x,t+\theta\tau)}{\partial t} \xrightarrow{\tau\to 0} \frac{\partial f}{\partial t}(x,t)$, and $|\frac{\partial f(x,t+\theta\tau)}{\partial t}| \leq M$. By the dominated convergence theorem, as $\tau_n \to 0$, we get $\frac{h(t+\tau_n)-h(t)}{\tau_n} \to \int_0^1 \frac{\partial f(x,t)}{\partial t}\,dx$. □

6. Differentiation

Theorem 6.1. *Let $f : [a,b] \to \mathbb{R}$ be monotonically increasing. Then f is differentiable almost everywhere.*

Theorem 6.2. *Let f be monotonically increasing on $[a,b]$. Then*

$$\int_a^b f'(t)\,dt \leq f(b) - f(a).$$

Before the proof, we do the following exercise.

Exercise 6.3. Let $f : \mathbb{R} \to \mathbb{R}$, t fixed. Then

$$\int_a^b f(x+t)\,dx = \int_{a+t}^{b+t} f(y)\,dy.$$

First, let $f = \chi_E$. $x \in [a,b], x+t \in E \Rightarrow x+t \in [a+t, b+t], x+t \in [a+t,b+t] \cap E$. Hence $\int_a^b \chi_E(x+t)\,dx = \mu(E \cap [a+t, b+t]) = \int_{a+t}^{b+t} \chi_E\,dx$.

Proof. By Theorem 6.1, f is differentiable almost everywhere. Define

$$g(x) = \begin{cases} \lim\limits_{h\to 0} \frac{f(x+h)-f(x)}{h}, & \text{if the limit exists} \\ 0, & \text{otherwise} \end{cases}$$

Then $g = f'$ almost everywhere. Let $g_n = n(f(x + frac1n) - f(x))$. Then $g_n \to f'$ almost everywhere and $g_n \geq 0$. By Fatou's lemma,

$$\int_a^b \liminf_{n\to\infty} g_n\,dt \leq \liminf_{n\to\infty} \int_a^b g_n(t)\,dt.$$

That is,

$$\int_a^b f'(t)\,dt \leq \liminf_{n\to\infty} \int_a^b g_n(t)\,dt.$$

Now,

$$\begin{aligned}
\int_a^b g_n(t)\,dt &= n\Big(\int_a^b f(t+\frac{1}{n})\,dt - \int_a^b f(t)\,dt\Big) \\
&= n\Big(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} f(t)\,dt - \int_a^b f(t)\,dt\Big) \\
&= n\Big(\int_a^{b+\frac{1}{n}} f(t)\,dt - \int_a^{a+\frac{1}{n}} f(t)\,dt\Big) - \int_a^b f(t)\,dt\Big) \\
&= n\Big(\int_b^{b+\frac{1}{n}} f(t)\,dt + \int_a^b f(t)\,dt - \int_a^{a+\frac{1}{n}} f(t)\,dt\Big) - \int_a^b f(t)\,dt\Big) \\
&= n\Big(\int_b^{b+\frac{1}{n}} f(t)\,dt - \int_a^{a+\frac{1}{n}} f(t)\,dt\Big) \\
&= f(b) - n\int_a^{a+\frac{1}{n}} f(t)\,dt.
\end{aligned}$$

Hence $\int_a^b f'(t)\,dt \leq f(b) - \limsup\limits_{n\to\infty} n \int_a^{a+\frac{1}{n}} f(t)\,dt$. Now, $\int_a^{a+\frac{1}{n}} f(t)\,dt \geq \frac{1}{n} f(a)$. Hence $\int_a^b f'(t)\,dt \leq f(b) - \limsup\limits_{n\to\infty} n \cdot \frac{1}{n} f(a) = f(b) - f(a)$. □

Remark 6.4. Strict inequality can hold in Theorem 6.2 as in the case of the Cantor function.

Definition 6.5. Let $f : [a, b] \to \mathbb{R}$, and $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[a, b]$. Let $t(\mathcal{P}, f) = \sum_{i=1}^{n} |f(x_i) - f(x_{i=1})|$ and $T_a^b(f) = \sup_{\mathcal{P}} t(\mathcal{P}, f)$. $T_a^b(f)$ is called the total variation of f over $[a, b]$, and f is said to be of bounded variation over $[a, b]$ if $T_a^b(f) < \infty$.

Example 6.6.
(1) If f is Lipschitz continuous, it is of bounded variation.
(2) Any monotonic function is of bounded variation with $T_a^b(f) = f(b) - f(a)$ or $f(a) - f(b)$.
(3) Any continuously differentiable function or differentiable function with bounded derivative is of bounded variation.

Example 6.7. Let

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^2},\ 0 \le x \le 1 \\ 0,\ x = 0 \end{cases}$$

Note that f' is not bounded and blows up near 0. Claim: f is not of bounded variation. Let $\mathcal{P} = \{0, 1\} \cup \{\sqrt{\frac{2}{\pi(2k+1)}} : 0 \le k \le n\}$. For each k,

$$\begin{aligned} |f(x_k) - f(x_k)| &= |\frac{2}{\pi(2k+1)} \sin \frac{(2k+1)\pi}{2} - \frac{2}{\pi(2k-1)} \sin \frac{(2k-1)\pi}{2}| \\ &= \frac{2}{\pi(2k+1)} + \frac{2}{\pi(2k-1)} \\ &= \frac{2}{\pi}(\frac{2k-1+2k+1}{(2k-1)(2k+1)}) \\ &= \frac{2}{\pi}\frac{4k}{4k^2-1} \\ &\ge \frac{8}{\pi}\frac{k}{4k^2} \\ &= \frac{2}{\pi k}. \end{aligned}$$

Hence $\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \ge \frac{2}{\pi} \sum_{k=1}^{n} \frac{1}{k} \to \infty$ as $n \to \infty$.

Proposition 6.8.
(1) If f is of bounded variation, then it is bounded.
(2) Let If f, g are of bounded variation, the so are $f + g$ and fg.

Proof.
(1) Let $t \in [a, b]$, and $P = \{a, t, b\}$. Then $|f(t) - f(a)| + |f(b) - f(t)| \le T_a^b(f)$. Hence $|f(x)| \le |f(a)| + T_a^b(f) < \infty$, $\forall t$.
(2) $f + g$ is of bounded variation by the triangular inequality. Consider fg.

$$\begin{aligned} |fg(x_i) - fg(x_{i-1})| &\le |f(x_i)(g(x_i) - g(x_{i-1}))| + |(f(x_i) - f(x_{i-1}))g(x_{i-1})| \\ &\le \|f\|_\infty |g(x_i) - g(x_{i-1})| + \|g\|_\infty |f(x_i) - f(x_{i-1})| \end{aligned}$$

□

For $r \in \mathbb{R}^+$, let $r^+ = \max\ \{r, 0\}$ and $r^- = -\min\ \{r, 0\}$. Then $r = r^+ - r^-$ and $|r| = r^+ + r^-$.

For a partition $\mathcal{P}$, Let

$$P(\mathcal{P}, f) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))^+$$

and

$$N(\mathcal{P}, f) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))^-.$$

Then $P(\mathcal{P}, f) + N(\mathcal{P}, f) = t(\mathcal{P}, f)$, and $P(\mathcal{P}, f) - N(\mathcal{P}, f) = f(b) - f(a)$. Let $P_a^b(f) = \sup_{\mathcal{P}} P(\mathcal{P}, f)$ and $N_a^b(f) = \sup_{\mathcal{P}} N(\mathcal{P}, f)$.

Proposition 6.9. *Let f be of bounded variation on $[a, b]$. Then*

$$T_a^b(f) = P_a^b(f) + N_a^b(f)$$

and

$$f(b) - f(a) = P_a^b(f) - N_a^b(f).$$

Proof. Fix a partition $\mathcal{P}$. Then $P_a^b - N_a^b = f(b) - f(a)$. Hence

$$P = N + (f(b) - f(a)) \leq N_a^b(f) + f(b) - f(a).$$

Hence

$$P_a^b(f) \leq N_a^b(f) + f(b) - f(a).$$

Similarly,

$$N_a^b(f) - P_a^b(f) \leq f(a) - f(b).$$

Hence

$$P_a^b(f) - N_a^b(f) = f(b) - f(a).$$

Now, $t(\mathcal{P}, f) = P(\mathcal{P}, f) + N(\mathcal{P}, f) \leq P_a^b(f) + N_a^b(f)$. Hence

$$T_a^b(f) \leq P_a^b(f) + N_a^b(f).$$

Now,

$$\begin{aligned} T_a^b(f) &\geq t(\mathcal{P}, f) \\ &= P(\mathcal{P}, f) + N(\mathcal{P}, f) \\ &= P(\mathcal{P}, f) + (P(\mathcal{P}, f) - (f(b) - f(a))) \\ &= 2P(\mathcal{P}, f) - (f(b) - f(a)) \\ &= 2P(\mathcal{P}, f) - (P_a^b(f) - N_a^b(f)). \end{aligned}$$

Taking the supremum over all partitions $\mathcal{P}$, we get

$$\begin{aligned} T_a^b(f) &\geq 2P_a^b(f) - P_a^b(f) + N_a^b(f) \\ &= P_a^b(f) + N_a^b(f). \end{aligned}$$

□

Consider the interval $[a, b]$ and $x, y \in (a, b)$ such that $x < y$. Then $[a, x], [a, y] \subseteq [a, b]$ and $T_a^x(f) \leq T_a^y(f)$, $P_a^x(f) \leq P_a^y(f)$, and $N_a^x(f) \leq N_a^y(f)$. Now, $f(a) - f(a) = P_a^x(f) - N_a^x(f)$. Hence $f(x) = (P_a^x(f) + f(a)) - N_a^x(f)$, where both $(P_a^x(f) + f(a))$ and $N_a^x(f)$ are monotonic. Hence we have the following theorem:

Theorem 6.10. *f is of bounded variation on $[a, b]$ iff f is the difference of two mononically increasing functions.*

Corollary 6.11. *If f is of bounded variation, then f' exists almost everywhere.*

Proposition 6.12. *If f is of bounded variation, then*

$$\int_a^b |f'|\,dx \le T_a^b(f).$$

If f is continuously differentiable, then

$$\int_a^b |f'|\,dx = T_a^b(f).$$

Proof. We have $f(x)-f(a) = P_a^x(f) - N_a^x(f)$. Hence $f'(x) = P_a^x(f)' - N_a^x()f'$. Hence $|f'(x)| = |P_a^x(f)'| + |N_a^x()f'|$. But as $P_a^x(f), N_a^x(f)$ are increasing, their derivatives are non-negative. Hence $|f(x)'| \le P_a^x(f)' + N_a^x(f)'$. Hence $|f(x)'| \le T_a^x(f)'$. So we have

$$\begin{aligned}\int_a^b |f'|\,dx &\le \int_a^b (T_a^x)'\,dx\\ &\le T_a^b(f) - T_a^a(f)\\ &= T_a^b(f).\end{aligned}$$

Next, let f be continuously differentiable. Let $\mathcal{P}$ be any partition. Then

$$f(x_i) - f(x_{i-1}) = \int_{x_i}^{x_i} f'(t)\,dt$$

and

$$|f(x_i) - f(x_{i-1})| \le \int_{x_i}^{x_i} |f'(t)|\,dt.$$

Summing over i, we get

$$t(\mathcal{P}, f) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \le \int_a^b |f'(t)|\,dt.$$

Taking supremum over all partitions, we get

$$T_a^b(f) \le \int_a^b |f'(t)|\,dt.$$

□

6.1. **Vector Valued Maps.** Let $f:[a,b]\to\mathbb{R}^N, f=(f_1,f_2,\cdots,f_N)$ and $|f(x)| = (\sum_{i=1}^{N} |f_i(x)|^2)^{\frac{1}{2}}$. Let $\mathcal{P}$ be a partition of $[a,b]$ and $t(\mathcal{P},f) = \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$. Then f is of bounded variation iff $\sup_{\mathcal{P}} t(\mathcal{P},f) < \infty$.

Lemma 6.13. *Let $f:[a,b]\to\mathbb{R}^N$. Then*

$$|\int_a^b f\,dx| \le \int_a^b |f|\,dx.$$

Proof. Let $y_i = \int\limits_a^b f_i\,dx$. Then $\int\limits_a^b f\,dx = y = (y_1, \cdots, y_N)$. $|\int\limits_a^b f\,dx|^2 = |y|^2 = \sum\limits_{i=1}^N |y_i|^2 = \sum\limits_{i=1}^N y_i \int\limits_a^b f_i\,dx = \int\limits_a^b \sum\limits_{i=1}^N y_i f_i(x)\,dx \le \int\limits_a^b |y||f|\,dx = |y| \int\limits_a^b |f|\,dx$. If $y = 0$, the proof is clear. Else, by dividing, we get

$$|\int\limits_a^b f\,dx| \le \int\limits_a^b |f|\,dx.$$

□

Theorem 6.14. *If $f : [a, b] \to \mathbb{R}^N$ is of bounded variation and continuously differentiable, then*

$$T_a^b(f) = \int\limits_a^b |f'(x)|\,dx.$$

Proof. Consider any partition $\mathcal{P}$. Then $\sum\limits_{i=1}^n |f(x_i) - f(x_{i-1})| = \sum\limits_{i=1}^n |\int\limits_{x_{i-1}}^{x_i} f'(t)\,dt| \le \sum\limits_{i=1}^n \int\limits_{x_{i-1}}^{x_i} |f'(t)|\,dt = \int\limits_a^b |f'(t)|\,dt$. Hence, $T_a^b(f) \le \int\limits_a^b |f'(t)|\,dt$.

Since f is continuously differentiable on $[a, b]$, f' is uniformly continuous on $[a, b]$. Hence for $\epsilon > 0, \exists \delta > 0$ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$. Choose $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ to be a partition such that $\forall i, \Delta x_i = x_i - x_{i-1} < \delta$. Let $t \in [x_{i-1}, x_i]$. Then $|f'(t)| < |f'(x_i)| + \epsilon$. Hence

$$\int\limits_{x_{i-1}}^{X_i} |f'(t)|\,dt \le |f'(x_i)|\Delta x_i + \epsilon \Delta x_i.$$

That is,

$$\begin{aligned}
\int\limits_{x_{i-1}}^{X_i} |f'(t)|\,dt - \epsilon\Delta x_i \le |f'(x_i)|\Delta x_i \\
&= |\int\limits_{x_{i-1}}^{x_i} f'(x_i)\,dt| \\
&= |\int\limits_{x_{i-1}}^{x_i} (f'(t) + f'(x_i) - f'(t))\,dt| \\
&\le |\int\limits_{x_{i-1}}^{x_i} f'(t)\,dt| + |\int\limits_{x_{i-1}}^{x_i} f'(x_i) - f'(t)\,dt| \\
&\le |f(x_i) - f(x_{i-1})| + \int\limits_{x_{i-1}}^{x_i} |f'(x_i) - f(t)|\,dt \\
&\le |f(x_i) - f(x_{i-1})| + \epsilon\Delta x_i.
\end{aligned}$$

Hence

$$\int_a^b |f'(t)|\,dt - \epsilon(b-a) \le T(\mathcal{P}, f) + \epsilon(b-a).$$

Thus,

$$\int_a^b |f'(t)|\,dt \le T(\mathcal{P}, f) + 2\epsilon(b-a) \le T_a^b(f) + 2\epsilon(b-a).$$

Hence, $\int_a^b |f'(t)|\,dt \le T_a^b(f)$. □

Let γ be a curve in $\mathbb{R}^2$. That is, $\gamma : [0,1] \to \mathbb{R}^2$. γ is a rectifiable arc iff γ is of bounded variation and the length of γ is $T_0^1(\gamma)$. If γ is a continuously differentiable function, then its length$= \int_0^1 |\gamma'(t)|\,dt$.

Let $\gamma(t) = (x(t), y(t))$. Then $\gamma'(t) = (x'(t), y'(t))$ and $|\gamma'(t)| = \sqrt{x'^2 + y'^2}$. Hence length of $\gamma = \int_0^1 \sqrt{x'^2 + y'^2}\,dt = \int_0^1 \sqrt{1 + (\frac{dy}{dx})^2}\,dx$

Proposition 6.15. *Let $f \ge 0$ be integrable. Then given $\epsilon > 0, \exists \delta > 0$ such that if $\mu(E) < \delta$, then $\int_E f\,d\mu < \epsilon$.*

Proof. (1) First, assume f is bounded. THen $\exists K > 0$ such that $|f| \le K$. Then $\int_E f\,d\mu \le K\mu(E)$. So, in this case, we choose $\delta < \frac{\epsilon}{K}$.

(2) In the general case, Let

$$f_n(x) = \begin{cases} f(x),\ f(x) \le n \\ n,\ f(x) > n \end{cases}$$

Then f_n is bounded, $f_n \le f$ and $f_n \nearrow f$. By the monotonone convergence theorem, $\int f_n \to \int f$. So given $\epsilon > 0$, choose N such that $\forall n \ge N, \int_X f - \int_X f_n < \frac{\epsilon}{2}$. Consider the corresponding f_N. By the bounded case, $\exists \delta > 0$ such that $\mu(E) < \delta \Rightarrow \int_E f_N\,d\mu < \frac{\epsilon}{2}$. Hence, if $\mu(E) < \delta$,

$$\begin{aligned} \int_E f\,d\mu &= \int_E f - f_N\,d\mu + \int_E f_N\,d\mu \\ &\le \int_X f - f_N\,d\mu + \int_E f_N\,d\mu \\ &< \epsilon. \end{aligned}$$

□

Proposition 6.16. *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Then*

$$\lim_{h\to 0} \frac{1}{h} \int_x^{x+h} f(t)\,dt = f(x).$$

Proof. Let $\epsilon > 0$ and $x \in \mathbb{R}$. Then $\exists \delta > 0$ such that $|t - x| < \delta \Rightarrow |f(t) - f(x) < \epsilon$.

$$\begin{aligned} |\frac{1}{h} \int_x^{x+h} f(t)\,dt - f(x)| &= |\frac{1}{h} \int_x^{x+h} f(t) - f(x)\,dt| \\ &\leq \frac{1}{h} \int_x^{x+h} |f(t) - f(x)|\,dt \\ &\leq \frac{1}{h} \epsilon h, \end{aligned}$$

by choosing h small enough. □

Proposition 6.17. *Let f be integrable on $[a, b]$. Define $F(x) = \int_a^x f(t)\,dt$. Then F is a uniformly continuous function of bounded variation.*

Proof. Let $x < y$. Then $|F(x) - F(y)| = |\int_x^y f(t)\,dt| \leq \int_x^y |f(t)|\,dt$. By Proposition 6.15, given $\epsilon > 0, \exists \delta > 0$ such that $|x - y| < \delta \Rightarrow \int_x^y |f(t)|\,dt < \epsilon$. Hence F is uniformly continuous. Next, let $\mathcal{P}$ be a partition of $[a, b]$. THen

$$\begin{aligned} \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| &= \sum_{i=1}^{n} |\int_{x_{i-1}}^{x_i} f(t)\,dt| \\ &\leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |f(t)|\,dt \\ &= \int_a^b |f(t)|\,dt. \end{aligned}$$

Hence F is of bounded variation. □

Proposition 6.18. *Let f be integrable on $[a, b]$. Assume that $\forall x \in [a, b], \int_a^x f(t)\,dt = 0$. Then $f = 0$ almost everywhere.*

Proof. Let $E_+ = \{x : f(x) > 0\}$. We will show that $\mu(E_+) = 0$. Similarly, we show that $E_- = \{x : f(x) < 0\}$ has measure zero. Assume $\mu(E_+) > 0$. Then $\exists F$ closed, $F \subseteq E_+$ and $\mu(F) > 0$. Let $U = (a, b) \setminus F$. Then U is open and $U \cap F = \emptyset$.

Also,

$$\int_a^b f(t)\,dt = \int_U f(t)\,dt + \int_F f(t)\,dt.$$

Hence

$$\int_U f(t)\,dt = -\int_F f(t)\,dt \neq 0.$$

Let $U = \bigcup_{n=1}^{\infty} [a_n, b_n)$, a disjoint union. Let $g_n = f \restriction \bigcup_{i=1}^{n} [a_i, b_i)$. Then $|g_n| \leq |f|$, f integrable and $g_n \to f \restriction U$. By the dominated convergence theorem, $\int_U g_n\,d\mu \to$

$\int\limits_U f\, d\mu$. Now, $0 \neq \int\limits_U f(t)\, dt = \sum\limits_{n=1}^{\infty} \int\limits_{a_n}^{b_n} f(t)\, dt$. Hence $\exists n$ such that $\int\limits_{a_n}^{b_n} f(t)\, dt \neq 0$. That is, $\int\limits_a^{b_n} f(t)\, dt - \int\limits_a^{a_n} f(t)\, dt \neq 0$, a contradiction. □

Proposition 6.19. *Let f be a bounded measurable function on $[a,b]$ and let $F(x) = \int\limits_a^x f(t)\, dt$. Then $F' = f$ almost everywhere.*

Proof. By Proposition 6.17, F is of bounded variation. Hence F' exists almost everywhere. Define

$$f_n(x) = n(F(x + \frac{1}{n}) - F(x)) = n \int\limits_x^{x+\frac{1}{n}} f(t)\, dt.$$

Suppse $|f| \leq K$. Then $f| \leq K$ and $f \to F'$ almost everywhere. By the dominated convergence theorem, $\forall c \in [a,b]$,

$$\begin{aligned}
\int\limits_a^c F'(t)\, dt &= \lim_{n\to\infty} \int\limits_a^c f_n(t)\, dt \\
&= \lim_{n\to\infty} n \int\limits_a^c (F(t + \frac{1}{n}) - F(t))\, dt \\
&= \lim_{n\to\infty} n(\int\limits_c^{c+\frac{1}{n}} F(t)\, dt - \int\limits_a^{a+\frac{1}{n}} F(t)\, dt) \\
&= F(c) - F(a) \\
&= \int\limits_a^c f(t)\, dt \quad \forall c \in [a,b].
\end{aligned}$$

By Proposition 6.18, $F' = f$ almost everywhere. □

Theorem 6.20. *Suppose f is integrable on $[a,b]$ and $F(x) = \int\limits_a^x f(t)\, dt$. Then $F' = f$ almost everywhere.*

Proof. Assume without loss of generality that $f \geq 0$. Define

$$f_n(x) = \begin{cases} f(x),\ f(x) \leq n \\ n,\ f(x) > 0 \end{cases}$$

Then $f_n \nearrow f$ and $f - f_n \geq 0$, f_n bounded for all n. Let $G_n(x) = \int\limits_a^x (f - f_n)(t)\, dt$. Then $\forall n, G_n$ is monotonically increasing and thus differentiable almost everywhere. By Proposition 6.1, $(\int\limits_a^x f_n(t)\, dt)' = f_n$ almost everywhere. Now, $F(x) = \int_a^x (f - f_n)(t)\, dt + \int\limits_a^x f_n(t)\, dt = G_n(x) + \int\limits_a^x f_n(t)\, dt$. Hence, $F' = G' + f_n \geq f_n\ \forall n$ since G_n' is positive. Hence $F \geq f$. Now, $F(b) \geq \int\limits_a^b F'(x)\, dx \geq \int\limits_a^b f(x)\, dx = F(b) - F(a) = F(b)$, where the first inequality holds since F is monotonically increasing due to the fact that $f \geq 0$. Hence, equality holds and $\int_a^b F'(x)\,dx = \int_a^b f(x)\,dx$. But $F' - f \geq 0 \Rightarrow F' = f$ almost everywhere. $\square$

Definition 6.21 (Absolutely continuous)**.** Suppose F is a function such that given $\epsilon > 0, \exists \delta > 0$ such that if $\{(x_k, y_k)\}$ is a disjoint collection of intervals such that $\sum_{k=1}^{n} y_k - x_k < \delta$, then $\sum_{k=1}^{n} |F(y_k) - F(x_k)| < \epsilon$. Then F is said to be absolutely continuous. Any indefinite integral is absolutely continuous.

Proposition 6.22. *If f is absolutely continuous, then f is of bounded variation.*

Theorem 6.23. *Suppose f is absolutely continuous and $f' = 0$ almost everywhere. Then f is a constant.*

Theorem 6.24. *Suppose F is absolutely continuous. Then*

$$F(x) = F(a) + \int_a^x F'(t)\,dt.$$

Proof. F is of bounded variation and hence $F = F_1 - F_2$, where F_i are monotonically increasing.

$$\begin{aligned}\int_a^b |F'| &\leq \int_a^b |F_1'| + |F_2'| \\ &= \int_a^b F_1' + \int_a^b F_2' \\ &\leq F_1(b) - F_1(a) + F_2(b) - F_2(a) \\ &= F(b) - F(a).\end{aligned}$$

Hence F' is integrable. Let $G(x) = \int_a^x F'(t)\,dt$. Then $G' = F'$ almost everywhere. Also, $F - G$ is absoltuely continuous. By Theorem 6.23, $F = G + c$. Hence $F(x) = \int_a^x F'(t)\,dt + \text{ constant }$. Take $x = a$ to get $F(x) = \int_a^x F'(t)\,dt + F(a)$. $\square$

7. Product Spaces

8. L^p Spaces