Measure Theory Notes

Lou Yi

May 7, 2025

Contents

1 Lebesgue Measure on $\mathbb{R}^d$

1.1 Preliminary

Definition 1.1 *A (closed)* ***rectangle*** *R in $\mathbb{R}^d$ is given by the product of d one-dimensional closed and bounded intervals:*

$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_d, b_d],$$

where $a_j \leq b_j$ are real numbers, $j = 1, 2, \cdots, d$. A rectangle is a ***cube*** *if all its side length are equal. The* ***volume*** *of the rectangle R, denoted by $|R|$, is defined to be*

$$|R| = (b_1 - a_1) \cdots (b_d - a_d).$$

we define the volume of open rectangles and half-open rectangles in the same way. A union of rectangles is said to be ***almost disjoint*** *if the interiors of the rectangles are disjoint.*

Lemma 1.2 *If a rectangle is almost disjoint union of finitely many other rectangles, say $R = \bigcup_{k=1}^{N} R_k$, then*

$$|R| = \sum_{k=1}^{N} |R_k|.$$

Lemma 1.3 *If $R, R_1, \cdots, R_N$ are rectangles, and $R \subset \bigcup_{k=1}^{N} R_k$, then*

$$|R| \leq \sum_{k=1}^{N} |R_k|.$$

Theorem 1.4 *Every open subset $\mathcal{O}$ of $\mathbb{R}$ can be written uniquely as the countable union of disjoint open intervals.*

Theorem 1.5 *Every open subset $\mathcal{O}$ of $\mathbb{R}^d$, $d \geq 1$ can be written as a countable union of almost disjoint closed cubes.*

Lemma 1.6 *Let $\mathscr{A} \subset \mathcal{P}(\mathbb{R}^n)$, then the following are equivalent:*

1. *$\mathscr{A}$ is the algebra generated by open rectangles, that is the set of the form $I_1 \times I_2 \times \cdots \times I_d$ where I_k is an open interval, that is, either $(-\infty, a_k)$, (a_k, b_k), (b_k, ∞) or $(-\infty, \infty)$.*

2. *$\mathscr{A}$ is the algebra generated by closed rectangles.*

3. *$\mathscr{A}$ consists of sets which are finite disjoint union of rectangles, sets of the form $I_1 \times \cdots \times I_n$.*

Lemma 1.7 *Let $\mathscr{A} \subset \mathcal{P}(\mathbb{R}^d)$. Then the following are equivalent:*

1. *$\mathscr{A}$ is the σ-algebra generated by open rectangles.*

2. *$\mathscr{A}$ is the σ-algebra generated by closed rectangles.*

3. *$\mathscr{A}$ is the σ-algebra generated by bounded open rectangles.*

4. $\mathscr{A}$ is the σ-algebra generated by bounded closed rectangles.

5. $\mathscr{A}$ is the σ-algebra generated by open sets.

6. $\mathscr{A}$ is the σ-algebra generated by closed sets.

1.2 Outer Measure

Definition 1.8 (Outer Measure on $\mathbb{R}^d$) *If E is any subset of $\mathbb{R}^d$, the* ***outer measure*** *of E is defined to be*

$$m^*(E) = \inf \sum_{j=1}^{\infty} |Q_j|,$$

where the infimum is taken over all countable coverings $E \subset \bigcup_{j=1}^{\infty} Q_j$ by closed cubes. The outer measure is always non-negative but could be infinite, that is $0 \leq m^*(E) \leq \infty$.

Remark 1.8.1 *One can replace the covering by cubes, with covering by rectangles; or with coverings by balls.*

Remark 1.8.2 *From the definition, one can easily show that a finite set have an outer measure of 0.*

Proposition 1.9 (Properties of Outer Measure)

1. *For every $\epsilon > 0$, there exists a covering $E \subset \bigcup\limits_{j=1}^{\infty} Q_j$ with*

$$\sum_{j=1}^{\infty} |Q_j| \leq m^*(E) + \epsilon.$$

2. *(Monotonicity) If $E_1 \subset E_2$, then $M^*(E_1) \leq M^*(E_2)$.*

3. *(Translation) Let $x \in R^d$ and $E_1 \subset R^d$, then $M^*(E_1) = M^*(x + E_1)$.*

4. *(Countable sub-additivity) If $E = \bigcup_{j=1}^{\infty} E_j$, then $m^*(E) \leq \sum_{j=1}^{\infty} m^*(E_j)$.*

5. *If $E \subset \mathbb{R}^d$, then $m^*(E) = \inf m^*(\mathcal{O})$, where the infimum is taken over all open sets $\mathcal{O}$ containing E.*

6. *If $E = E_1 \cup E_2$ and $d(E_1, E_2) > 0$, then*

$$m^*(E) = m^*(E_1) + m^*(E_2).$$

7. *If a set E is the countable union of almost disjoint cubes, $E = \bigcup_{j=1}^{\infty} Q_j$, then*

$$m^*(E) = \sum_{j=1}^{\infty} |Q_j|.$$

Proof:

1. This follows from the definition of outer measure.

2. This is true as any covering of E_2 is a covering of E_1.

3. One can establish a one to one correspondence between the cover of E_1 and $x + E_1$.

4. First, we may assume that each $m^*(E_j) < \infty$, as otherwise the inequality clearly holds. For any $\epsilon > 0$, the definition of outer measures yields for each j a covering $E_j \subset \bigcup_{k=1}^{\infty} Q_{k,j}$ by closed cubes with

$$\sum_{k=1}^{\infty} |Q_{k,j}| \leq m^*(E_j) + \frac{\epsilon}{2^j}.$$

Then $E \subset \bigcup_{j,k=1}^{\infty} Q_{k,j}$ is a covering of E by closed cubes, so

$$m^*(E) \leq \sum_{j=1}^{\infty} m^*(E_j) + \epsilon.$$

As $\epsilon > 0$ is arbitrary, then the desired statement follows.

5. By Monotonicity, it is clear that the inequality $m^*(E) \leq \inf m^*(\mathcal{O})$ holds. For the reverse inequality, let $\epsilon > 0$ and choose cubes Q_j such that $E \subset \bigcup_{j=1}^{\infty} Q_j$ with

$$\sum_{j=1}^{\infty} |Q_j| \leq m^*(E) + \frac{\epsilon}{2}.$$

Then replace each Q_j by an open cube Q_j^0 containing Q_j, with $|Q_j^0| \leq |Q_j| + \frac{\epsilon}{2^{j+1}}$. Take $\mathcal{O}$ to be the union of these open cubes, then by countable sub-additivity, we get the desired result.

6. By countable sub-additivity, we have $m^*(E) \leq m^*(E_1) + m^*(E_2)$, we need to show the reverse inequality. First select δ such that $d(E_1, E_2) > \delta > 0$. Next choose a covering $E \subset \bigcup_{j=1}^{\infty} Q_j$ by closed cubes with $\sum_{j=1}^{\infty} |Q_j| \leq m^*(E) + \epsilon$. We may, after subdividing the cubes, assume that each Q_j has a diameter less than δ. In this case, each Q_j can intersect at most one of the two sets E_1 or E_2. If we denote by J_1 and J_2 the sets of those those indices j for which Q_j intersects E_1 and E_2, respectively, then $J_1 \cap J_2$ is empty, and we have

$$E_1 \subset \bigcup_{j \in J_1}^{\infty} Q_j, \quad E_2 \subset \bigcup_{j \in J_2}^{\infty} Q_j.$$

Therefore, $m^*(E_1) + m^*(E_2) \leq m^*(E) + \epsilon$. Since ϵ is arbitrary, then we have the desired result.

7. Let $\tilde{Q}_j$ denote a cube strictly contained in Q_j such that $|Q_j| \leq |\tilde{Q}_j| + \frac{\epsilon}{2^j}$, where ϵ is arbitrary but fixed. Then for every N, the cubes $\tilde{Q}_1, \tilde{Q}_2, \cdots, \tilde{Q}_N$ are disjoint, hence at a finite distance from one another. Then repeated application of Property (5) gives that

$$m^*\left(\bigcup_{j=1}^{N} \tilde{Q}_j\right) = \sum_{j=1}^{N} |\tilde{Q}_j| \geq \sum_{j=1}^{N} \left(|Q_j| - \frac{\epsilon}{2^j}\right).$$

Then we conclude that $m^*(E) \geq \sum_{j=1}^{N} |Q_j| - \epsilon$ for arbitrary ϵ. Let $N \to \infty$, then we get $\sum_{j=1}^{\infty} |Q_j| \leq m^*(E)$. By countable sub-additivity, we have the other inclusion.

$\square$

Corollary 1.9.1 *The outer measure of any cubes and rectangles agrees with its volume. The outer measure of an open set equals the sum of the volumes of the cubes in a decomposition.*

Proposition 1.10 *Every countable subset of R^d has outer measure* 0.

Proof: Suppose $A = \{a_1, a_2, \cdots,\}$ is a countable subset of $\mathbb{R}$. Let $\epsilon > 0$. For $j \in \mathbb{Z}^+$, let Q_j be a closed cube that contains a_j and has volume $\frac{\epsilon}{2^j}$. Then $\{Q_1, Q_2, \cdots,\}$ is a countable collection of closed cubes that contains A, thus by countable sub-additivity, we get $m^*(A) \leq \epsilon$. As ϵ is arbitrary, we get that $m^*(A) = 0$. $\square$

Lemma 1.11 *There exist disjoint subsets A and B of $\mathbb{R}$ such that*

$$m^*(A \cup B) \neq m^*(A) + m^*(B).$$

Proof: For $a \in [-1, 1]$, let $\tilde{a}$ be the set of numbers in $[-1, 1]$ that differs from a by a rational number. Thus

$$[-1, 1] = \bigcup_{a \in [-1,1]} \tilde{a}.$$

Let V be a set that contains exactly one element in each of the distinct sets in $\{\tilde{a} : a \in [-1, 1]\}$. Let $r_1, r_2, \cdots$ be a sequence of distinct rational numbers such that

$$[-2, 2] \cap \mathbb{Q} = \{r_1, r_2, \cdots,\}.$$

Then

$$[-1, 1] \subseteq \bigcup_{k=1}^{\infty} (r_k + V),$$

where the set inclusion above holds because if $a \in [-1, 1]$, then letting v be the unique element of $V \cap \tilde{a}$, we have $a - v \in [-2, 2] \cap \mathbb{Q}$, which implies that $a = r_k + v \in r_k + V$ for some $k \in \mathbb{Z}^+$.
By monotonicity and countable sub-additivity implies that

$$m^*([-1, 1]) \leq \sum_{k=1}^{\infty} m^*(r_k + V).$$

We know that $m^*([-1, 1]) = 2$, so $2 \leq \sum_{k=1}^{\infty} m^*(V)$, in particular, $m^*(V) > 0$.
Note that the sets $r_1 + V, r_2 + V, \cdots$ are disjoint and have finite outer measure. And notice for any $n \in \mathbb{Z}^+$, $\bigcup_{k=1}^{n} r_k + V \subset [-3, 3]$ which has a finite outer measure. Then we have

$$m^*\left(\bigcup_{k=1}^{n} r_k + V\right) \leq m^*([-3, 3]) \leq 6.$$

Now if we have $m^*(A \cup B) = m^*(A) + m^*(B)$ for all disjoint subsets A, B of $\mathbb{R}$. This would imply

$$m^*\left(\bigcup_{k=1}^{n} r_k + V\right) = \bigcup_{k=1}^{n} m^*(r_k + V) = n \cdot m^*(V).$$

As $m^*(V) > 0$, $\exists n$, such that $n \cdot m^*(V) > 6$, which gives a contradiction. $\square$

1.3 Measurable Sets and the Lebesgue Measure

Definition 1.12 *A subset E of $\mathbb{R}^d$ is* ***Lebesgue measurable*** *of simply* ***measurable****, if for any $\epsilon > 0$, there exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and*

$$m^*(\mathcal{O} - E) \leq \epsilon.$$

If E is measurable, we define its ***Lebesgue measure*** *$m(E)$ by*

$$m(E) = m^*(E).$$

We denote the set of all Lebesgue measurable sets by the symbol $\mathcal{L}$.

Remark 1.12.1 *Clearly, the Lebesgue measure inherits all the Properties of the outer measure.*

Proposition 1.13 (Properties of Measurable sets)

1. *Every open set in $\mathbb{R}^d$ is measurable.*
2. *If $m^*(E) = 0$, then E is measurable. In particular, if F is a subset of a set of exterior measure 0, then F is measurable.*
3. *A countable union of measurable sets is measurable.*
4. *Closed sets are measurable.*
5. *The complement of a measurable set if measurable.*
6. *A countable intersection of measurable set is measurable.*

Remark 1.13.1 *This shows that the collection of Lebesgue measurable sets forms a σ-algebra.*

Proof:

1. Take $\mathcal{O}$ to be itself.
2. By property (5) of outer measures (see Proposition 1.9), we know that for every $\epsilon > 0$, there exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and $m^*(\mathcal{O}) \leq \epsilon$. Since $(\mathcal{O} - E) \subset \mathcal{O})$, then monotonicity implies $m^*(\mathcal{O} - E) \leq \epsilon$, as desired.

3. Suppose $E = \bigcup_{j=1}^{\infty} E_j$, where E_j is measurable. Given $\epsilon > 0$, we may choose for each j an open set $\mathcal{O}_j$ with $E_j \subset \mathcal{O}_j$ and $m^*(\mathcal{O}_j - E_j) \leq \epsilon/2^j$. Then the union $\mathcal{O}$ of the O_j's is open, with $E \subset \mathcal{O}$, and $(\mathcal{O} - E) \subset \bigcup_{j=1}^{\infty}(\mathcal{O}_j - E_j)$. So by monotonicity and sub-additivity of the outer measure, we have the desired result.

4. Note any closed set F can be written as countable union of compact sets; so it suffices to show that compact sets are measurable.

   Now suppose F is compact, (in particular $m^*(F) < \infty$), and let $\epsilon > 0$. By Property (5) of outer measure (Proposition 1.9), we can select an open set $\mathcal{O}$ with $F \subset \mathcal{O}$ and $m^*(\mathcal{O}) \leq m^*(F) + \epsilon$. Since F is closed, the difference $\mathcal{O} - F$ is open, so we can write this difference as a countable union of almost disjoint cubes, that is,

$$\mathcal{O} - F = \bigcup_{j=1}^{\infty} Q_j.$$

   For a fixed N, the finite union $K = \bigcup_{j=1}^{N} Q_j$ is compact; therefore $d(K, F) > 0$. Since $(K \cup F) \subset \mathcal{O}$, then we conclude that

$$\begin{aligned} m^*(\mathcal{O}) &\geq m^*(F) + m^*(K) \\ &= m^*(F) + \sum_{j=1}^{N} m^*(Q_j). \end{aligned}$$

   Letting $N \to \infty$, we get $\sum_{j=1}^{\infty} m^*(Q_j) \leq \epsilon$. Thus by sub-additivity,

$$m^*(\mathcal{O} - F) = m^*\left(\bigcup_{j=1}^{\infty} m^*(Q_j)\right) \leq \sum_{j=1}^{\infty} m^*(Q_j) \leq \epsilon$$

   as desired.

5. If E is measurable, then for every positive integer n we may choose an open set $\mathcal{O}_n$ with $E \subset \mathcal{O}_n$ and $m^*(\mathcal{O}_n - E) \leq 1/n$. The complement $\mathcal{O}_n^c$ is closed, hence measurable, which implies that the union $S = \bigcup_{n=1}^{\infty} \mathcal{O}_n^c$ is also measurable by Property (3). Now note $S \subset E^c$, and

$$(E^c - S) \subset (\mathcal{O}_n - E),$$

   such that $m^*(E^c - S) \leq 1/n$ for all n. Therefore, $m^*(E^c - S) = 0$, and $E^c - S$ is measurable by Property (2). Therefore E^c is measurable since S is measurable.

6. This follows from Property (3) and (5) and De Morgan's Law.

$\square$

Theorem 1.14 (Countable Additivity of Measure) *If $E_1, E_2, \cdots$, are disjoint measurable sets, and $E =$*

$\bigcup_{j=1}^{\infty} E_j$, *then*

$$m(E) = \sum_{j=1}^{\infty} m(E_j).$$

Proof: First, we assume further that E_j is bounded, . Then for each j, by applying the definition of measurability to E_j^c, we can choose a closed subset F_j of E_j with $m^*(E_j - F_j) \leq \epsilon/2^j$. For each fixed N, the sets $F_1, \cdots, F_n$ are compact and disjoint, so by Property (6) of outer measure hence measure (Proposition 1.9) that

$$m\left(\bigcup_{j=1}^{N} F_j\right) = \sum_{j=1}^{N} m(F_j).$$

Since $\bigcup\limits_{j=1}^{N} F_j \subset E$, we must have

$$m(E) \geq \sum_{j=1}^{N} m(F_j) \geq \sum_{j=1}^{N} m(E_j) - \epsilon$$

Letting N tend to infinity, and since ϵ was arbitrary, we find that

$$m(E) \geq \sum_{j=1}^{\infty} m(E_j).$$

Note the reverse inequality always holds by countable sub-additivity.

In the general case, we select any sequence of cubes $\{Q_k\}_{k=1}^{\infty}$ that increases to $\mathbb{R}^d$, in this sense that $Q_k \subset Q_{k+1}$ for all $k \geq 1$ and $\bigcup_{k=1}^{\infty} Q_k = \mathbb{R}^d$. We then let $S_1 = Q_1$ and $S_k = Q_k - Q_{k-1}$ for $k \geq 2$. If we define measurable sets by $E_{j,k} = E_j \cap S_k$, then

$$E = \bigcup_{j,k} E_{j,k}.$$

The union above is disjoint and every $E_{j,k}$ is bounded. Moreover $E_j = \bigcup_{k=1}^{\infty} E_{j,k}$, and this union is also disjoint. So we obtain

$$m(E) = \sum_{j,k} m(E_{j,k}) = \sum_j \sum_k m(E_{j,k}) = \sum_j m(E_j)$$

as desired. □

Corollary 1.14.1 *Suppose A and B are measurable sets, then $A - B$ is measurable, $A \Delta B$ is also measurable.*

Proof: Note $A\Delta B = (A - B) \cup (B - A)$, so we only need to show $A - B$ is measurable. Since $A - B = A \cap B^c$, then $A - B$ is measurable. □

Notation: if $E_1, E_2, \cdots$ is a countable collection of subsets of $\mathbb{R}^d$ that increases to E in the sense that $E_k \subset E_{k+1}$ for all k, and $E = \bigcup_{k=1}^{\infty} E_k$, then we write $E_k \nearrow E$. Similarly, if $E_1, E_2, \cdots$ decreases to E in the sense that $E_k \subset E_{k+1}$ for all K, and $E = \bigcap_{k=1}^{\infty} E_k$, we write $E_k \searrow E$.

Corollary 1.14.2 *Suppose $E_1, E_2, \cdots$ are measurable subsets of $\mathbb{R}^d$.*

- *If $E_k \nearrow E$, then $m(E) = \lim_{N\to\infty} m(E_N)$.*
- *If $E_k \searrow E$ and $m(E_k) < \infty$ for some k, then*

$$m(E) = \lim_{N\to\infty} m(E_N).$$

Proof: For the first part, let $G_1 = E_1$, and $G_k = E_k - E_{k-1}$ for $k \geq 2$. Then the sets G_k are measurable, disjoint, and $E = \bigcup_{k=1}^{\infty} G_k$. Hence

$$m(E) = \sum_{k=1}^{\infty} m(G_k) = \lim_{N\to\infty} \sum_{k=1}^{N} m(G_k) = \lim_{N\to\infty} m\left(\bigcup_{k=1}^{N} G_k\right),$$

and since $\bigcup_{k=1}^{N} G_k = E_n$, we get the desired limit.

For the second part, we may clearly assume that $m(E_1) < \infty$. Let $G_k = E_k - E_{k+1}$ for each k, so that

$$E_1 = E \cup \bigcup_{k=1}^{\infty} G_k$$

is a disjoint union of measurable sets. As a result, we find that

$$\begin{aligned} m(E_1) &= m(E) + \lim_{N\to\infty} \sum_{k=1}^{N-1} (m(E_k) - m(E_{k+1})) \\ &= m(E) + m(E_1) - \lim_{N\to\infty} m(E_N). \end{aligned}$$

Since $m(E_1) < \infty$, we see that $m(E) = \lim_{N\to\infty} m(E_N)$. □

Theorem 1.15 (The Borel-Cantelli Lemma) *Suppose $\{E_k\}_{k=1}^{\infty}$ is a countable family of measurable subsets of $\mathbb{R}^d$ and that*

$$\sum_{k=1}^{\infty} m(E_k) < \infty.$$

Let E be the sets of elements $x \in \mathbb{R}^d$, such that $x \in E_k$ for infinitely many k. Then E is measurable and $m(E) = 0$.

Proof: Since $E = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} E_k$.Then by Lemma 1.14.2, the result follows easily. □

Theorem 1.16 *Suppose E is a measurable subset of $\mathbb{R}^d$. Then, for every $\epsilon > 0$:*

1. *There exists an open set $\mathcal{O}$ with $E \subset \mathcal{O}$ and $m(\mathcal{O} - E) \leq \epsilon$.*
2. *There exists a closed set F with $F \subset E$ and $m(E - F) \leq \epsilon$.*
3. *If $m(E)$ is finite, there exists a compact set with $K \subset E$ and $m(E - K) \leq \epsilon$.*

4. *if $m(E)$ is finite, there exists a finite union $F = \bigcup_{j=1}^{N} Q_j$ of closed cubes such that*

$$m(E\Delta F) \le \epsilon.$$

Proof:

1. Definition of measurability.

2. Since E^c is measurable, there exists an open set $\mathcal{O}$ with $E^c \subset \mathcal{O}$ and $m(\mathcal{O} - E^c) \le \epsilon$. If we let $F = \mathcal{O}^c$, then F is closed, $F \subset E$, and $E - F = \mathcal{O} - E^c$. Hence $m(E - F) \le \epsilon$ as desired.

3. We first pick a closed set F so that $F \subset E$ and $m(E - F) \le \epsilon/2$. For each n, we let B_n denote the ball centered at the origin of radius n, and define compact sets $K_n = F \cap B_n$. Then $E - K_n$ is a sequence of measurable sets that decreases to $E - F$, and since $m(E) < \infty$, we conclude that for all large n one has $m(E - K_n) \le \epsilon$.

4. Choose a family of closed cubes $\{Q_j\}_{j=1}^{\infty}$ so that

$$E \subset \bigcup_{j=1}^{\infty} Q_j \quad \text{and} \quad \sum_{j=1}^{\infty} |Q_j| \le m(E) + \frac{\epsilon}{2}.$$

This is possible because we can first find an open $\mathcal{O}$ that contains E, and $m(\mathcal{O} - E) \le \epsilon/2$. Then decompose $\mathcal{O}$ into countable union of almost disjoint cubes, we have the desired Q_j's.

Since $m(E) < \infty$, the series converges and there exits $N > 0$ such that $\sum_{j=N+1}^{\infty} |Q_j| < \epsilon/2$. If $F = \bigcup_{j=1}^{N} Q_j$, then

$$\begin{aligned}
m(E\Delta F) &= m(E - F) + m(F - E) \\
&\le m\left(\bigcup_{j=N+1}^{\infty} Q_j\right) + m\left(\bigcup_{j=1}^{\infty} Q_j - E\right) \\
&\le \sum_{j=N+1}^{\infty} |Q_j| + \sum_{j=1}^{\infty} |Q_j| - m(E) \\
&\le \epsilon.
\end{aligned}$$

□

Proposition 1.17 (Invariance Property of Lebesgue Measure)

1. *If E is a measurable set and $h \in \mathbb{R}^d$, then $E_h = h + E$ is also measurable, and $m(E + h) = m(E)$.*

2. *If $\delta = (\delta_1, \cdots, \delta_d) \in \mathbb{R}^d$, then the set $\delta E = \{(\delta_1 x_1, \cdots, \delta_d x_d) : (x_1, \cdots, x_d) \in E\}$ is measurable and $m(\delta E) = \delta_1 \cdots \delta_d m(E)$.*

3. *In particular, if E is measurable, then $-E = \{-x : x \in E\}$ and $m(-E) = m(E)$.*

Proof: Property (1) is clear, and (3) follows from (2). For (2), it suffices to show for outer measures. Firstly, suppose all of the δ_i's are nonzero, the if $O \subset \mathbb{R}^d$ is an open set that contains E, then δO is an open set that contains δE; conversely, if O_δ is an open set that contains δE, then $\frac{1}{\delta}O = \left(\frac{1}{\delta_1}, \cdots, \frac{1}{\delta_d}\right) O$ is an open set that contains E. Since $m^*(E) = \inf m^*(O)$, and by decomposition into almost disjoint cubes, we have $m^*(\delta O) = \delta_1 \cdots \delta_d m^*(O)$, then we have $\delta_1 \cdots \delta_d m^*(E) = m^*(\delta E)$.
Now if any of δ_i is zero, then we show that $m^*(\delta E) = 0$, then it would mean that δE is measurable. WLOG, let $\delta_1 = 0$. Then for any open set O, covering E, decompose it into almost disjoint union of cubes, and change the first coordinate of these cubes by first translating to be centred at 0, then scale by a factor of ϵ. Then by taking the union of these rectangles, we would obtain an open set O' that covers δE. Furthermore, we can show that $m^*(O') = \epsilon\delta_2 \cdots \delta_d m^*(O)$. Now letting $\epsilon \to 0$, we get the desired result. $\square$

Notation: we denote the Borel σ-algebra on $\mathbb{R}^d$ by $\mathcal{B}_{\mathbb{R}^d}$. Note since all open sets are measurable, then $\mathcal{B}_{\mathbb{R}^d}$ is a subset of the set of all Lebesgue measurable sets. We call a set that is a countable intersection of open sets to be an G_δ set; we call a set that is the countable union of closed sets to be a F_σ set.

From the point of view of the Borel sets, the Lebesgue sets arises as the completion of the σ-algebra of Borel sets, the Lebesgue sets, that is, by adjoing all subsets of Borel sets of measure zero. Of course, we also know that not all sets are measurable by Vitali's Theorem.

Proposition 1.18 *A subset E of $\mathbb{R}^d$ is measurable*

1. *if and only if E differs from a G_δ by a set of measure zero,*
2. *if and only if E differs from an F_σ by a set of measure zero.*

Proof: If E satisfies any one (1) or (2), then as F_σ, G_δ and the sets of measure zero are measurable, then E is measurable.

Conversely, if E is measurable, then for each integer $n \geq 1$, we may select an open set $\mathcal{O}_n$ that contains E, and such that $m(\mathcal{O}_n - E) \leq 1/n$. Then $S = \bigcap \mathcal{O}_n$ is a G_δ that contains E, and $(S - E) \subset (\mathcal{O}_n - E)$ for all n, so $m^*(S - E) = 0$, so $S - E$ is measurable and $m(S - E) = 0$. Similarly, we can show the second part of the proposition. $\square$

Proposition 1.19 *Suppose L is a linear transformation in $\mathbb{R}^d$, then if E is a measurable subset of $\mathbb{R}^d$, then so is $L(E)$, with $m(L(E)) = |\det L| m(E)$.*

Proof: We will only prove the first half of the statement, as the second half requires more tools.

Since L is linear on a finite dimensional vector space, it is Lipschitiz continuous. So if G is compact, then $L(G)$ is compact. Hence if G is an F_σ set, so is $L(G)$, (can decompose all the closed sets into compact ones). Next L maps any cube of side length ℓ int oa cube of side length $c_d M\ell$, where M is the Lipschitiz constant, and $c_d = 2\sqrt{d}$. Then if E is measurable, then $m(E - G) = 0$ for some F_δ set G, so $E - G \subset \bigcup Q_j$ for some countable collection of cubes $\{Q_j\}$, with $\sum m(Q_j) < \epsilon$. Then $m^*(L(E) - L(G)) \leq C\epsilon$, for some constant C, so $m(L(E) - L(G)) = m^*(L(E) - L(G)) = 0$, and hence $L(E)$ is measurable. $\square$

1.4 Lebesgue Measure Using Extension Theorem

Of course, we could also use the Caratheéodory Extension Theorem to obtain the so called Lebesgue Measure, however we need to proceed with more caution. In this case, the definition is given below:

Definition 1.20 (Lebesgue Measure) *The **Lebesgue Measure** λ on $\mathbb{R}^d$ is the unique measured defined on the σ-algebra $\mathcal{L}(\mathbb{R}^d)$ satisfies the following property:*

1. *(Open/Borel Measurability) Every Borel sets is measurable.*
2. *(Translation Invariant) For $X \in \mathcal{L}(\mathbb{R}^d)$ and $a \in \mathbb{R}^d$, $\lambda(a + X) = \lambda(X)$.*
3. *(Inner and Outer Regularity) if $X \in \mathcal{L}(\mathbb{R}^d)$, then $\exists X_1$ that is the countable union of compact sets and X_2 that is the countable intersection of open sets, such that*

$$X_1 \subseteq X \subseteq X_2.$$

   and $\lambda(X_2 \setminus X_1) = 0$.

4. *(Normalized on Unit Cube) $\lambda([0,1]^d) = 1$.*
5. *(Complete) If $N \in \mathcal{L}(\mathbb{R}^d)$ has $\lambda(N) = 0$, then every $X \subseteq N$ is an element of $\mathcal{L}(\mathbb{R}^d)$ with $\lambda(X) = 0$.*

Remark 1.20.1 *The inner and outer regularity condition described in the definition is slightly stronger than what these terms usually means. However, for the special case of Lebesgue measure, it satisfies this condition and is a much more useful property than the standard inner and outer regularity (see Appendix).*

Remark 1.20.2 *If we construct the natural premeasure and use the Carathéodory to extend it to the Borel sets, we get a measure μ that satisfy everything but completeness. This is because the size of the Borel set is $2^{\aleph_0}$, but the measurable set in a complete measure is $2^{\mathbb{R}}$. We can do a completion of this measure to get Lebesuge measure. Nevertheless, the σ-algebra $\mathcal{S}$ that consists of all elements satisfying the Carathéodory splitting condition with respect to the outer measure is the desired $\mathcal{L}(\mathbb{R}^d)$.*

We first define the natural premeasure on the Algebra formed by rectangles.

Proposition 1.21 *Let $\mathscr{A} = \mathscr{A}(\mathbb{R}^d)$ be the algebra of $\mathbb{R}^d$ consisting of finite union of disjoint rectangles. Then there is a unique premeasure μ_0 defined on $\mathscr{A}$ such that*

$$\mu_0([a_1, b_1] \times \cdots \times [a_d, b_d]) = \prod_{i=1}^{d} (b_i - a_i).$$

Proof: Set

$$\mu_0(I_1 \times \cdots \times I_d) = \prod_{i=1}^{d} \ell(I_i)$$

where I_i are intervals, and the length $\ell(I)$ of an interval is defined in the obvious way. Well-definedness of μ_0 follows from 1.2.

Next, we need to show restricted countable additivity. Let $A = \bigsqcup_{n=1}^{\infty} A_n$, where A is the countable union of rectangles, and A_n are rectangles. Then we want to show

$$\mu_0(A) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

WLOG, we can only consider the case where A is a single rectangle. Since μ_0 is finitely additive (clear), then

$$\mu_0(A) \geq \sum_{n=1}^{N} \mu_0(A_n)$$

for any $N \in \mathbb{N}$. So $\mu_0(A) \geq \sum \mu_0(A_n)$. Conversely, we show $\mu_0(A) \leq \sum_{n=1}^{\infty} \mu_0(A_n) + \epsilon$ for any ϵ. Let $\bar{A}$ be the closure of A and $\bar{A_n}$ be the closure of A_n. It is enough to show $\mu_0(\bar{A}) \leq \sum_{n=1}^{\infty} \mu_0(\bar{A_n}) + \epsilon$. Now suppose $\bar{A}$ is bounded, slightly enlarge each $\bar{A_n}$ into an open set A'_n by $\frac{\epsilon}{2^n}$. Then by compactness we have the desired claim. Now if $\bar{A}$ is not bounded, then by consider $\bar{A} \cap [-k, k]^d$ and letting $k \to \infty$, we again will get the desire result.

Uniqueness is clear. $\square$

Lemma 1.22 *Let μ be a measure on a σ-algebra $\mathscr{A}$ of subsets of Ω. There is a unique complete measure $\tilde{\mu}$ on σ-algebra $\tilde{\mathscr{A}}$ such that $\mathscr{A} \subseteq \tilde{\mathscr{A}}$, and $\mu = \tilde{\mu}|_{\tilde{\mathscr{A}}}$ and $\tilde{\mu}$ is minimized with respect to this inclusion property. Moreover, if $\tilde{X} \in \tilde{\mathscr{A}}$, there is X_1 and $X_2 \in \mathscr{A}$ such that $X_1 \subseteq \tilde{X} \subseteq X_2$, $\mu(X_2 \setminus X_1) = 0$.*

Remark 1.22.1 *We call the $\tilde{\mu}$ to be the **completion** of μ.*

Proof: Basically if $\mu(X) = 0$, then we add every subset of X into $\tilde{\mathscr{A}}$. $\square$

Theorem 1.23 *The Lebesgue measure defined in Definition 1.20 on $\mathbb{R}^d$ exists and is unique.*

Proof: (Existence): Let μ_0 be the natural premeasure on the algebra $\mathscr{A}$ containing all the rectangles in $\mathbb{R}^d$. Then use Carathéodory extension theorem to get a measure μ on $\sigma(\mathscr{A}) = \mathcal{B}$. Use Lemma 1.22 to get a complete measure $\tilde{\mu}$ on $\tilde{A}$. We will show it satisfies all the properties.

It is clear that $\tilde{\mu}$ satisfies the Open/Borel measurability, normalized and complete. It remains to show translation invariant and regularity.

Claim: $\tilde{\mu}$ satisfies translation invariant.
Indeed, set μ'_0 to be the set function given by $\mu'_0(X) = \mu_0(a + x)$ for fixed a. μ'_0 is a premeasure and

$$\mu'([a_1, b_1] \times \cdots \times [a_d, b_d]) = \prod_{i=1}^{d} (b_i - a_i).$$

Then by uniqueness $\mu'_0 = \mu_0$. Next, we verify μ is translation invariant. Set $\mu'(x) = \mu(a + x)$ for fixed a. Both μ' and μ are measure on $\mathcal{B}$ and extension on μ_0. Then by the uniqueness part of the Carathéodory extension theorem, we have $\mu' = \mu$. So μ is translational invariant. Finally, in the same way, we can show $\tilde{\mu}$ is translational invariant.

Claim: $\tilde{\mu}$ satisfies inner and outer regularity.
We will prove this statement by induction that every Borel set satisfies this property.
(Base case): take compact box $X = I_1 \times \cdots \times I_d$. Take $X_1 = X$ and $X_2 = \bigcap_{n \in N}(I_1 + (-1/n, 1/n)) \times \cdots \times (I_d + (-1/n, 1/n))$. Then $X_1 \subset X \subset X_2$, so $\mu(X_2) = \mu(x) = \mu(x_1) < \infty$.
Now we verify $\mathscr{C} = \{X \subseteq \mathcal{B}(\mathbb{R}^d) : X \text{ is inner and outer regular}\}$. We show $\mathscr{C}$ is a σ-algebra. Clearly $\emptyset, \Omega = \mathbb{R}^d$ is inside $\mathscr{C}$. If $X, X' \subseteq \mathscr{C}$, then $X_1 \subset X \subset X_2$ and $X_1' \subset X' \subset X_2'$. Then

$$X_1 \setminus X_2' \subseteq X \setminus X' \subseteq X_2 \setminus X_1',$$

where the left hand side is the countable union of closed sets (hence compact sets), and the right hand side the countable intersection of open sets. We can also show that these three sets have the same measure μ.
Now suppose $(X^{(i)})$ is a sequence of disjoint sets in $\mathscr{C}$. Then

$$X_1^{(i)} \subseteq X^{(i)} \subseteq X_2^{(i)}.$$

Then if $X = \bigsqcup X^{(i)}$, take $X_1 = \bigsqcup X_1^{(i)}$. For each i and k, choose open set $U_k^{(i)} \supset X^{(i)}$, s.t., $\mu(U_k^{(i)} \setminus X^{(i)}) < \frac{1}{k \cdot 2^i}$. Set $U_k^i = \bigcup_i U_k^{(i)}$, and set $X_2 = \bigcap_k U_k$. Then one can verify that $X_1 \subseteq X \subseteq X_2$ and $\mu(X_2 \setminus X_1) = \inf_k \mu(U_k \setminus X_1) = 0$.
Thus we conclude that every set in $\mathcal{B}(\mathbb{R}^d)$ satisfies inner and outer regularity. Lastly, we just need to show that for any X which is $\tilde{\mu}$-measurable. Since $\tilde{\mu}$ is the completion of μ, we can approximate X by $B_1, B_2 \in \mathcal{B}(\mathbb{R}^d)$, such that $B_1 \subseteq X \subseteq B_2$, and $\mu(B_2 \setminus B_1) = 0$. Then it is clear that X satisfies inner and outer regularity.

(Uniqueness): Suppose $\nu = \tilde{\mu}$, then $\nu|_{\mathscr{A}} = \mu_0$. Then by the Uniqueness in the Carathéodry Extension Theorem, $\nu|_{\mathcal{B}} = \mu$. By the Uniqueness of the completion, we have $\nu|_{\tilde{\mathcal{B}}} = \tilde{\mu}$. □

1.5 Lebesgue Measurable Functions

In this section, we will look at some property of Lebesgue Measurable Functions.

Proposition 1.24 *Every continuous function $f : \mathbb{R}^d \to \mathbb{R}$, is measurable. Similarly, if $E \subset \mathbb{R}^d$ is Lebesgue measurable, then every continuous function $f : E \to \mathbb{R}$ is measurable.*

Proof: Suppose f is continuous, then $U_a = \{x \in \mathbb{R}^d : f(x) > a\}$ is open for all $a \in \mathbb{R}$. Since every open set is measurable, then f is measurable. □

Proposition 1.25 *Let f and g be measurable real-valued functions (cannot take infinity as a value) defined on a measurable subset E of $\mathbb{R}^d$. Let F be real and continuous on $\mathbb{R}^2$, and put*

$$h(x) = F(f(x), g(x)) \quad (x \in E).$$

Then h is measurable.
In particular, cf, $f \pm g$, fg and f/g $(g \neq 0)$ are measurable.

Proof: Let

$$G_a = \{(u,v) \,|\, F(u,v) > a\}.$$

Then G_a is an open subset of $\mathbb{R}^2$, so we can write as countable union of almost disjoint cubes

$$G_a = \bigcup_{n=1}^{\infty} Q_n,$$

where

$$Q_n = \{(u,v) \,|\, a_n \le u \le a_n + d_n,\ b_n \le v \le b_n + d_n\}.$$

Since

$$\{x|a_n \le f(x) \le a_n + d_n\} = \{x|f(x) \ge a_n\} \cap \{x|f(x) \le b_n\}$$

is measurable, it follows that the set

$$\{x|(f(x), g(x)) \in Q_n\} = \{x|a_n \le f(x) \le a_n + d_n\} \cap \{x|b_n \le g(x) \le b_n + d_n\}$$

is measurable. Hence the same is true of

$$\begin{aligned}\{x|h(x) > a\} &= \{x|(f(x), g(x)) \in G_a\} \\ &= \bigcup_{n=1}^{\infty} \{x|(f(x), g(x)) \in Q_n\}.\end{aligned}$$

□

Corollary 1.25.1 *Suppose $f : X \to \mathbb{R}$ is a measurable function, where X is an arbitrary measurable space, $\Phi : \mathbb{R} \to \mathbb{R}$ is a continuous function, then $\Phi \circ f$ is measurable.*

Proof: Consider the preimage of $(a, -\infty)$ under Φ, since Φ is continuous, then $\Phi^{-1}((a, -\infty))$ is open, which can be represented as the countable union of open intervals. The preimage of each under f is measurable, hence $f^{-1}(\Phi^{-1}((a, -\infty)))$ is measurable for each $a \in \mathbb{R}$. Hence $\Phi \circ f$ is measurable. □

Definition 1.26 *The **characteristic function** of a set E is defined to be*

$$\chi_E(x) = \begin{cases} 1 & x \in E, \\ 0 & x \notin \mathbb{E}. \end{cases}.$$

*Let s be a real-valued function defined on a measurable space Ω. If the range of s is finite, we say that s is a **simple function**.*

Note that if the range of s consists of the distinct numbers $c_1, \cdots, c_n$. Let

$$E_i = \{x|s(x) = c_i\} \quad (i = 1, \cdots, n).$$

Then the E_i's are clearly pairwise disjoint and

$$s(x) = \sum_{i=1}^{n} c_i K_{E_i}(x).$$

This is known as the ***canonical form*** *of* s.

Lemma 1.27 *The sum and product of simple functions are simple functions. In particular, if $s(x)$ is a simple function then $cs(x)$ is also simple, where c is a scalar.*

Proposition 1.28 *Suppose $s : X \to \mathbb{R}$, where X is an arbitrary measurable space is given by*

$$s = \sum_{i=1}^{n} c_i \chi_{E_i}.$$

Then s is measurable if and only if the sets $E_1, \cdots, E_n$ are measurable in X.

Proof: $\Rightarrow$: suppose s is measurable, then WLOG, let the set of c_i's be listed in ascending order, and let $c_0 = -\infty$, and $c_{n+1} = +\infty$. Then for every $i = 1, 2, \cdots, n$, we have

$$\{x \in X \,:\, c_{i-1} < s(x) < c_{i+1}\}$$

is measurable. However, note that $c_{i-1} < s(x) < c_{i+1}$ if and only if $s(x) = c_i$, so $x \in E_i$. Thus E_i is measurable for each i.

$\Leftarrow$: conversely, suppose each set $E_1, \cdots, E_n$ are measurable, then for every $\alpha \in \mathbb{R}$, consider the set

$$\{x \in X \,:\, s(x) > \alpha\} = \bigcup_{a_j > \alpha} E_j.$$

Since each E_j is measurable, then the finite union of measurable sets are measurable, thus s is measurable. $\square$

Theorem 1.29 *Let $f : X \to [-\infty, \infty]$ be a function, X is an arbitrary measurable function. There exists a sequence $\{s_n\}$ of simple functions such that $s_n(x) \to f(x)$ as $n \to \infty$, for every $x \in X$. If f is measurable, $\{s_n\}$ may also be chosen to be a sequence of measurable functions. In addition, if f is bounded, then we can choose $\{s_n\}$ to be uniformly bounded that converges uniformly to f, and at the same time $\{s|_n|\}$ is monotonically increasing, i.e., $\forall x \in X$, $s_1(x) \leq s_2(x) \cdots \leq s_n(x) \leq \cdots$.*

Proof: First we consider the case $f \geq 0$ and bounded. Define

$$E_{ni} = \left\{ x \;\middle|\; \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n} \right\}, \quad F_n = \{x | f(x) \geq n\}$$

for $n = 1, 2, 3, \cdots$, $i = 1, 2, \cdots, n2^n$. Put

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{ni}} + n\chi_{F_n}.$$

Then we have constructed such a sequence $\{s_n\}$ converging to f. Note that this $\{s_n\}$ we constructed is nonnegative, uniformly bounded, converging to f uniformly and monotonically increasing.

It is clear that if f is measurable, then each s_n is measurable.

If f is bounded, i.e., $f(X) \subset (-M, M)$, then consider $F(x) = f(x) + M$. Then $F \geq 0$. So we exists $\{s_n\}$ such that $\{s_n\}$ is monotonically increasing, uniformly bounded, and converging to $F(x)$. Thus $\{s_n - M\}$ has the desired property for f.

Next, if $f \geq 0$ and is unbounded, then the same construction works, but we lose the property of uniformly boundedness and uniform convergence.

Lastly, for the general case, let $f = f^+ - f^-$, and apply the preceding construction to f^+ and to f^-. In fact, in this case, we get that $|s_n| \leq |s_{n+1}|$ for all $n \in \mathbb{N}$. □

Corollary 1.29.1 *Suppose $f : X \to [-m, +\infty]$, $m \in \mathbb{R}$, then there exists a sequence $\{s_n\}$ of simple functions that is monotonically increasings, $s_n \geq -m$ and converges to f pointwise.*

Proof: The construction is clear from the proof of the theorem. We first show for $f : X \to [0, +\infty]$, then consider $f - m$, $m \in \mathbb{R}$. □

Then letting $X = \mathbb{R}^d$, we get the following results:

Theorem 1.30 *Suppose f is a non-negative measurable function on $\mathbb{R}^d$, then there exists an increasing sequence of non-negative measurable simple functions $\{s_n\}_{n=1}^{\infty}$ that converges pointwise to f, that is,*

$$s_n(x) \leq s_{n+1}(x) \quad \textit{and} \quad \lim_{n\to\infty} s_n(x) = f(x), \quad \textit{for all } x.$$

Theorem 1.31 *Suppose f is measurable on $\mathbb{R}^d$, there exists a sequence of measurable simple functions $\{s_n\}_{n=1}^{\infty}$ that satisfies*

$$|s_n(x)| \leq |s_{n+1}(x)| \quad \textit{and} \quad \lim_{n\to\infty} s_n(x) = f(x), \quad \textit{for all } x.$$

Remark 1.31.1 *By restricting the domain of each simple function, we can in particular take each of the simple functions to only take nonzero values on a finite measure.*

Theorem 1.32 *Suppose f is measurable on $\mathbb{R}^d$. Then there exists a sequence of step function $\{\psi_n\}_{n=1}^{\infty}$ that converges pointwise to $f(x)$ for almost every x. Here step functions are of the form*

$$f = \sum_{k=1}^{N} c_k \chi_{R_k},$$

where each R_k is a rectangle in $\mathbb{R}^d$ and c_k are constants.

Proof: It suffices to show that if E is a measurable set with finite measure, then $f = \chi_E$ can be approximated by step functions. Then the rest follows from Theorem1.31 and remark of Theorem 1.31.

Now by Theorem 1.16, item (4), we have that for every ϵ, there exists finite cubes $Q_1, \cdots, Q_N$ such that

$$m\left(E\Delta\bigcup_{j=1}^{N} Q_j\right) \leq \epsilon.$$

By considering the grid formed by extending the sides of these cubes, we see that there exists almost disjoint rectangles, $\tilde{R}_1, \cdots, \tilde{R}_M$ such that $\bigcup_{j=1}^{N} Q_j = \bigcup_{j=1}^{M} \tilde{R}_j$. By taking rectangles R_j contained in $\tilde{R}_j$, and slight smaller in size, we find a collection of disjoint rectangles that satisfy

$$m\left(E\Delta\bigcup_{j=1}^{M} R_j\right) \leq 2\epsilon.$$

Therefore

$$f(x) = \sum_{j=1}^{M} \chi_{R_j}(x),$$

except possibly on a set with measure less than or equal to 2ϵ. Consequently, for every $n \geq 1$, there exists a step function $\psi_n(x)$ such that if

$$E_n = \{x : f(x) \neq \psi_n(x)\},$$

then $m(E_n) \leq 2^{-k}$. If we let $F_n = \bigcup_{j=n+1}^{\infty} E_j$ and $F = \bigcap_{n=1}^{\infty} F_n$, then $m(F) = 0$ since $m(F_n) \leq 2^{-n}$, and $\psi_n(x) \to f(x)$ for all x in the complement of F, which is the desired result. $\square$

1.6 Littlewood's Three Principle

The Littlewood's Three Principle States:

1. Every set is nearly a finite union of intervals.
2. Every function is nearly continuous.
3. Every convergent sequence is nearly uniformly convergent.

However, this is an informal version of three theorems. The precise version of the first principle is item (4) of Theorem 1.16. The second and the third principle are given by Lusin's Theorem and Egorov's Theorem respectively.

Theorem 1.33 (Egorov) *Suppose $\{f_k\}_{k=1}^{\infty}$ is a sequence of measurable functions defined on a measurable set E (Not necessarily subset of $\mathbb{R}^d$) with $m(E) < \infty$, and assume that $f_k \to f$ almost everywhere on E. Given $\epsilon > 0$, we can find a closed set (or more strongly a compact set) $A_\epsilon \subset E$ such that $m(E - A_\epsilon) \leq \epsilon$ and $f_k \to f$ uniformly on A_ϵ.*

Proof: WLOG, let $f_k(x) \to f(x)$ for every $x \in E$, since a measure zero set does not affect the result. For each pair of non-negative integers n and k, let

$$E_k^n = \{x \in E \,:\, |f_j(x) - f(x)| < 1/n, \text{ for all } j > k\}.$$

Now fix n and note that $E_k^n \subset E_{k+1}^n$, and $E_k^n \nearrow E$ as k tends to infinity. This implies that there exists k_n such that $m(E - E_{k_n}^n) < \frac{1}{2^n}$. By construction, we then have

$$|f_j(x) - f(x)| < 1/n, \text{ whenever } j > k_n \text{ and } x \in E_{k_n}^n.$$

We choose N so that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon/2$, and let

$$\tilde{A}_\epsilon = \bigcap_{n \geq N} E_{k_n}^n.$$

We first observe that

$$m(E - \tilde{A}_\epsilon) \leq \sum_{n=N}^{\infty} m(E - E_{k_n}^n) < \frac{\epsilon}{2}.$$

Next, if $\delta > 0$, we choose $n \geq N$ such that $1/n < \delta$, and note that $x \in \tilde{A}_\epsilon$ implies $x \in E_{k_n}^n$. Then $|f_j(x) - f(x)| < \delta$ whenever $j > k_n$. Hence f_k converges uniformly to f on $\tilde{A}_\epsilon$.
Finally using Theorem 1.16, we can choose a closed subset $A_\epsilon \subset \tilde{A}_\epsilon$ with $m(\tilde{A}_\epsilon - A_\epsilon) < \frac{\epsilon}{2}$. As a result, we have $m(E - A_\epsilon) < \epsilon$ and the theorem is proved. Of course, by approximating a closed set with compact set using regularity, we can conclude that the closed set in the statement of the theorem can be replaced by compact set. □

Theorem 1.34 (Lusin) *Suppose f is measurable and finite valued on $E \subset \mathbb{R}^d$ with E of finite measure. Then for every $\epsilon > 0$ there exists a closed set F_ϵ, with*

$$F_\epsilon \subset E, \text{ and } m(E - F_\epsilon) \leq \epsilon$$

and such that $f|_{F_\epsilon}$ is continuous. Again by regularity, we can replace the closed set F_ϵ with a compact one.

Proof: Let f_n be a sequence of step functions so that $f_n \to f$ a.e., which exists by Theorem 7.49. Then we may find sets E_n so that $m(E_n) < \frac{1}{2^n}$ and f_n is continuous outside E_n. By Egorov's Theorem, we may find a set $A_{\epsilon/3}$ on which $f_n \to f$ uniformly and $m(E - A_{\epsilon/3}) \leq \epsilon/3$. Then we consider

$$F' = A_{\epsilon/3} - \bigcup_{n \geq N} E_n$$

for N so large that $\sum_{n \geq N} \frac{1}{2^n} < \frac{\epsilon}{3}$. Now for every $n \geq N$ the function f_n is continuous on F'; thus f is also continuous on F'. To finish the proof, we merely need to approximate the set F' by a closed set $F_\epsilon \subset F'$ such that $m(F' - F_\epsilon) < \frac{\epsilon}{3}$. □

Corollary 1.34.1 *Suppose f is measurable on $(\mathbb{R}^d, \mathcal{L}, \lambda)$. Then f is the a.e. pointwise limit of continuous functions.*

Proof: Note $f\chi_{[-n,n]^d}$ converges pointwise to f when $n \to \infty$. So we just need to construct sequence that converges pointwise to $f\chi_{[-n,n]^d}$, then by diagonalization argument, we will get a sequence that converges pointwise to f. Choose a sequence of compact sets $K_{m,n} \subseteq [-m,m]^d$, such that $\mu([-m,m]^d \setminus K_{m,n}) < \frac{1}{n}$, and f is continuous on $K_{m,n}$ with respect to the subspace topology. But then by the Tietze Extension Theorem, we can extend $f|_{K_{m,n}}$ to a continuous function on $[-m,m]^d$and set $f_{m,n}$ to be such continuous extension. Then we can check that $(f_{n,n})$ converges to f pointwise almost everywhere. $\square$

Theorem 1.35 (The Brunn-Minkowski Inequality) *Suppose A and B are measurable sets in $\mathbb{R}^d$ and the sum $A + B = \{x' + x'' \,|\, x' \in A,\, x'' \in B\}$ is also measurable, then*

$$m(A+B)^{\frac{1}{d}} \geq m(A)^{\frac{1}{d}} + m(B)^{\frac{1}{d}}.$$

2 Integration

In this section, we will introduce the integration theory for the Lebesgue Theory. However, all of these can be generalized to general measure space $(\Omega, \mathscr{A}, \mu)$.

2.1 Basic Definitions and Properties

Definition 2.1 (Lebesgue Integral of Simple Functions) *Let φ be a measurable simple function on $\mathbb{R}^d$ written in its canonical form, that is*

$$\varphi(x) = \sum_{k=1}^{N} c_k \chi_{E_k}(x),$$

where the E_k are measurable sets and a_k are constants. Then we define the ***Lebesgue integral of*** *φ by*

$$\int_{\mathbb{R}^d} \varphi(x)dx = \sum_{k=1}^{M} c_k \lambda(E_k).$$

If E is a measurable subset of $\mathbb{R}^d$, then $\varphi(x)\chi_E(x)$ is also a simple function, and we define

$$\int_E \varphi(x)dx = \int_{\mathbb{R}^d} \varphi(x)\chi_E(x)dx.$$

Notice that the integral may have value $+\infty$.

Remark 2.1.1 *To emphasize the choice of the Lebesgue measure λ in the definition of the integral, one sometimes writes*

$$\int_{\mathbb{R}^d} \varphi(x)d\lambda(x) \quad or \quad \int \varphi d\lambda$$

for the Lebesgue integral of φ. Sometimes, for convenience sake, we may simple denote the integral to be $\int \varphi$.

Lemma 2.2 (Independence of the representation) *If $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$ is any representation of φ then*

$$\int \varphi = \sum_{k=1}^{N} a_k \lambda(E_k).$$

Proof: First, we suppose that $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$ where the sets E_k are disjoint, but we do not suppose that the numbers a_k are distinct. For each distinct non-zero value a among the $\{a_k\}$,we define $E'_a = \bigcup E_k$, where the union is taken over those indices k such that $a_k = a$. Note that the sets E'_a are disjoint, and $\lambda(E'_a) = \sum \lambda(E_k)$, where the sum is taken over the same sets of k's. Then clearly $\varphi = \sum a\chi_{E'_a}$, where the sum is over the distinct non-zero values of $\{a_k\}$. Thus

$$\int \varphi = \sum a\lambda(E'_a) = \sum_{k=1}^{N} a_k \lambda(E_k).$$

Next, suppose $\varphi = \sum_{k=1}^{n} a_k \chi_{E_k}$, were we no longer asume that the E_k are disjoint. Then we can "refine" the decomposition $\bigcup_{k=1}^{N} E_k$ by finding sets E_1^*, $E_2^*, \cdots, E_n^*$ with the property that $\bigcup_{k=1}^{N} E_k = \bigcup_{j=1}^{n} E_j^*$; the sets E_j^* $(j = 1, \cdots, n)$ are mutually disjoint; and for each k, $E_k = \bigcup E_j^*$, where the union is taken over those E_j^* that are contained in E_k. For each j, let now $a_j^* = \sum a_k$, with the summation taken over all k such that E_k contains

E_j^*. Then clearly $\varphi = \sum_{j=1}^{n} a_j^* \chi_{E_j^*}$. However, this is a decomposition already dealt with above because the E_j^* are disjoint. Thus

$$\int \varphi = \sum a_j^* m(E_j^*) = \sum \sum_{E_k \supset E_j^*} a_k m(E_j^*) = \sum a_k m(E_k).$$

□

Definition 2.3 (Lebesgue Integral of Nonnegative Functions) *Let f be a measurable and nonnegative function defined on a measurable set $E \subseteq \mathbb{R}^d$, we define the* ***Lebesgue integral of*** *f to be*

$$\int_E f d\lambda = \sup_s \int_E s d\lambda$$

where the supremum is taken over all measurable simple functions s such that $0 \leq s \leq f$. Notice that the integral may have value $+\infty$.

Remark 2.3.1 *It easy to see that for every nonnegative simple measurable function s,*

$$\int_E s d\lambda = \sum_{i=1}^{n} c_i \lambda(E \cap E_i)$$

which agrees with our previously defined definition for the Lebesgue integral of simple functions.

Definition 2.4 (Lebesgue Integral of General Functions) *Let f be an arbitrary measurable function defined on a measurable set $E \subseteq \mathbb{R}^d$, and consider the two integrals*

$$\int_E f^+ d\lambda, \quad \int_E f^- d\lambda$$

where f^+ and f^- are defined as

$$f^+ = \max(f, 0), \quad f^- = -\min(f, 0).$$

If at least one of the integral is finite, we define the ***Lebesgue integral of*** *f by*

$$\int_E f d\lambda = \int_E f^+ d\lambda - \int_E f^- d\lambda.$$

If both $\int_E f^+ d\lambda$ and $\int_E f^- d\lambda$ are finite, then $\int_E f d\lambda$ is finite. In this case we say that f is ***integrable or summable*** *on E in the Lebesgue sense, with respect to the Lebesgue measure λ. We write $f \in \mathscr{L}(\lambda)$ on E.*

Remark 2.4.1 *If $\int_E f d\lambda$ is $\pm\infty$, then the integral of f over E is defined, although f is not integrable; f is integrable on E only if its integral over E is finite.*

Proposition 2.5 (Properties of Lebesgue Integrals)

1. *If f is measurable and bounded on E, and if $\lambda(E) < +\infty$, then $f \in \mathscr{L}(\lambda)$ on E.*

2. *If $a \leq f(x) \leq b$ for all $x \in E$, and $\lambda(E) < +\infty$, then*

$$a\lambda(E) \leq \int_E f d\lambda \leq b\lambda(E).$$

3. *If f and g are measurable functions on E, and if $f(x) \leq g(x)$ for every $x \in E$, then*

$$\int_E f d\lambda \leq \int_E g d\lambda.$$

4. *If $f \in \mathscr{L}(\lambda)$ on E, then $cf \in \mathscr{L}(\lambda)$ on E. Moreover, for every finite constant c, and any function f,*

$$\int_E c f d\lambda = c \int_E f d\lambda.$$

5. *If $\lambda(E) = 0$, and f is measurable, then*

$$\int_E f d\lambda = 0.$$

6. *Suppose $X, Y \subseteq E$ are measurable sets, and $X \cap Y = \emptyset$. Then*

$$\int_{X \cup Y} f d\lambda = \int_X f d\lambda + \int_Y f d\lambda.$$

7. *If $f \in \mathscr{L}(\lambda)$ on E, A is Lebesgue Measurable, and $A \subset E$, then $f \in \mathscr{L}(\lambda)$ on A.*

8. *If f is a measurable function on E, then*

$$\left| \int_E f d\lambda \right| \leq \int_E |f| d\lambda.$$

Remark 2.5.1 *Since the integral of a measurable function on a measure zero set always give 0. Then the assumptions can always be weakened to the case where the conditions only holds almost everywhere.*

Proof:

1. Let us first consider the case where $f \geq 0$. Since f is bounded, then $|f| \leq M$ for some $M \in \mathbb{R}$. Then let $s : E \to \mathbb{R}$ be the simple function defined by $s(x) = M\chi_E(x)$. Then for any simple function $t : E \to \mathbb{R}$, such that $t \leq f$, it is clear that $t \leq M$. Hence

$$\int_E f d\lambda \leq \int_E s d\lambda = M\lambda(E) \leq \infty.$$

Similarly, we can show that $\int_E f d\lambda \geq 0$, so $f \in \mathscr{L}(\lambda)$ on E.

Next, suppose we have a general function f. Then $f = f^+ - f^-$. Since f is bounded, then so must be f^+ and f^-. Therefore f^+, $f^- \in \mathscr{L}(\lambda)$ on E. Then it is clear that

$$\int_E f d\lambda = \int_E f^+ d\lambda - \int_E f^- d\lambda$$

is also finite, so $f \in \mathscr{L}(\lambda)$ on E.

2. We can so this by considering the 4 cases based on the value of a and b. The proof of each case is similar to that of 1.

3. We consider the two cases:
   Case 1: $f, g \geq 0$ and $f \leq g$ on E. For a simple function s, s.t., $0 \leq s \leq f$ on E, then $0 \leq s \leq g$ on X. By definition of $\int_E g d\lambda$, $\int_E s d\lambda \leq \int_E g d\lambda$, this is true for all s satisfying $0 \leq s \leq g$. Hence $\int_E g d\lambda$ is an upper bound for the set of $I_E(s)$, where $0 \leq s \leq f$. Since $\int_E f d\lambda$ is the least upper bound of the set of $I_E(s)$, then $\int_E f d\lambda \leq \int_E g d\lambda$.
   Case 2: we consider general functions f, g, such that $f \leq g$. Then $f^+ \leq g^+$ and $g^- \leq f^-$, Thus $\int_E f^+ \leq \int_E g^+ d\lambda$ and $\int_E g^- d\lambda \leq \int_E f^- d\lambda$. Suppose the integral of f and g are defined, then

$$\int_E f d\lambda = \int_E f^+ d\lambda - \int_E f^- d\lambda \leq \int_E g^+ \lambda - \int_E g^- d\lambda = \int_E g d\lambda.$$

4. First, we consider simple measurable function $s = \sum_{i=1}^{n} c_i X_{E_i}$. Then $as = \sum_{i=1}^{n} (c \cdot c_i) \chi_{E_i}$. Hence

$$\int_E cf d\lambda = \sum_{i=1}^{n} (c \cdot c_i) \lambda(E_i \cap E) = c \sum_{i=1}^{n} c_i \lambda(E_i \cap E) = c \int_E f d\lambda.$$

   Suppose $c = 0$, then the statement is clearly true for any measurable function f.
   Next, consider the case where $f \geq 0$ and $c > 0$. Then for any simple measurable function $0 \leq s \leq cf$ on E, it is the case that $0 \leq \frac{s}{c} \leq f$. Since by definition, $\frac{1}{c} \int_E s d\lambda = \int_E \frac{s}{c} d\lambda \leq \int_E f d\lambda$, then $\int_E s d\lambda \leq c \int_E f d\lambda$, hence $\int_E cf d\lambda \leq c \int_E f d\lambda$. On the other hand, suppose s is a simple measurable function satisfying $0 \leq s \leq f$ on E, then $0 \leq cs \leq cf$ on E. Therefore, $c \int_E s d\lambda = \int_E cs d\lambda \leq \int_E cf d\lambda$. Then $\int_E s d\lambda \leq \frac{1}{c} \int_E cf d\lambda$, so $\int_E f d\lambda \leq \frac{1}{c} \int_E cf d\lambda \Rightarrow c \int_E f d\lambda \leq \int_E cf d\lambda$. Hence we conclude that $c \int_E f d\lambda = \int_E cf d\lambda$.

   Now we consider the case that $f \geq 0$, and $c < 0$. But since

$$\int_E (-f) d\lambda = \int_E (-f)^+ d\lambda - \int_E (-f)^- d\lambda = 0 - \int_E f d\lambda.$$

   Then

$$\int_E (cf) d\lambda = \int_E |a|(-1) f d\lambda = |a| \int_E (-f) d\lambda = -|a| \int_E f d\lambda = a \int_E f d\lambda.$$

   Lastly, we consider the case where f is any measurable function and $c > 0$. Since one can clearly verify that for any real value a, $(af)^+ = af^+$ and $(af)^- = af^-$, then

$$\begin{aligned} \int_E (cf) d\lambda &= \int_E (cf)^+ d\lambda - \int_E (cf)^- d\lambda \\ &= \int_E cf^+ d\lambda - \int_E cf^- d\lambda \end{aligned}$$

$$= c\left(\int_E f^+ d\lambda - \int_E f^- d\lambda\right)$$
$$= c\int_E d\lambda.$$

The case for $c < 0$ follows similarly.

5. Suppose $f \geq 0$. And let s be any simple function, so it will have the form

$$s(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x) \quad (x \in E,\, c_i > 0)$$

for each $i \in \{1, \cdots, n\}$, E_i is measurable. And since $E \cap E_i \subseteq E$ and it is also measurable, then $\lambda(E \cap E_i) \leq \lambda(E) = 0$. Hence, for every such s, we have

$$\int_E s d\lambda = \sum_{i=1}^{n} c_i \lambda(E \cap E_i) = 0.$$

Thus $\int_E f d\lambda = 0$, as it is the sup of the set of zeros.

Next, suppose f is any general function, we have that $f = f^+$ and f^-. Then by our previous analysis, we would be able to get that $\int_E f^+ E d\lambda = \int_E f^- E d\lambda = 0$, hence we can conclude that

$$\int_E d\lambda = 0.$$

6. For any simple function $s = \sum_{i=1}^{n} c_i \chi_{E_i}(x)$, we have

$$\begin{aligned}
\int_{X\cup Y} s d\lambda &= \sum_{i=1}^{n} c_i \lambda(E_i \cap (X \cup Y)) \\
&= \sum_{i=1}^{n} c_i \lambda((E_i \cap X) \cup (E_i \cap Y)) \\
&= \sum_{i=1}^{n} c_i [\lambda(E_i \cap X) + \lambda(E_i \cap Y)] \\
&= \sum_{i=1}^{n} c_i \lambda(E_i \cap X) + \sum_{i=1}^{n} c_i \lambda(E_i \cap Y) \\
&= \int_X s d\lambda + \int_Y s d\lambda.
\end{aligned}$$

Next for any $f \geq 0$. For any simple measurable function s such that $0 \leq s \leq f$ on $X \cup Y$, we have $0 \leq s \leq f$ on X and $0 \leq s \leq f$ on Y. Thus

$$\begin{aligned}
&\int_{X\cup Y} s d\lambda = \int_X s d\lambda + \int_Y s d\lambda \leq \int_X f d\lambda + \int_Y f d\lambda \\
\Rightarrow &\int_{X\cup Y} f d\lambda \leq \int_X f d\lambda + \int_Y f d\lambda.
\end{aligned}$$

On the other hand, for any simple measurable functions s, t, s.t., $0 \leq s \leq f$ on X and $0 \leq t \leq f$ on Y, consider $p = s\chi_X + t\chi_Y$. It is a simple function such that $0 \leq p \leq f$ on $X \cup Y$. Then

$$\begin{aligned}\int_{X\cup Y} pd\lambda &= \int_X pd\lambda + \int_Y d\lambda \\ &= \int_X sd\lambda + \int_Y td\lambda \\ \Rightarrow \int_X s\lambda + \int_Y td\lambda &\leq \int_{X\cup Y} fd\lambda \\ \Rightarrow \int_X fd\lambda + \int_Y fd\lambda &\leq \int_{X\cup Y} fd\lambda\end{aligned}$$

Since if for all $a \in A$ and $b \in B$, $a + b \leq c$ then $\sup A + \sup B \leq c$. Hence we conclude that $\int_{X\cup Y} fd\lambda = \int_X fd\lambda + \int_Y fd\lambda$.
Lastly, we consider a general function, but this just follows from the fact that we can decompose any function f into f^+ and f^-.

7. Consider f^+ and f^-. We know $\int_E f^+d\lambda$ and $\int_E f^-d\lambda$ are both finite, it suffices to show that $\int_A f^+d\lambda$ and $\int_A f^-d\lambda$ are both finite.

   So let us only consider the case of $\int_A f^+d\lambda$. Suppose we have any simple function s, such that $s \leq f^+$ on A, then $s \leq f^+$ on E, and as A is measurable, we if let s' be defined to be s for every $x \in A$ and 0. Then $s' \leq f^+$ in E.
   Hence we can easily see that
   $$\int_A sd\lambda = \int_A s'd\lambda = \int_E s'd\lambda \leq \int_E fd^+d\lambda.$$
   Thus $\int_A f^+d\lambda$ is finite. Similarly, we can show that $\int_A f^-d\lambda$ is also finite. Thus we conclude that $f \in \mathscr{L}(\lambda)$ on A.

8. Suppose f is not integrable, then the statement clearly holds. On the other hand, suppose f is integrable, then $\left|\int_E fd\lambda\right|$ is finite. Let $X = \{x \in E : f(x) \geq 0\}$ and $Y = \{x \in E : f(x) < 0\}$. Then $\int_X f$ and $\int_Y f$ are finite. Clearly,
   $$\left|\int_X f\right| = \int_X f \leq \int_X |f|$$
   and
   $$\left|\int_Y f\right| = -\int_Y f \leq \int_Y |f|.$$
   So
   $$\left|\int_E fd\lambda\right| \leq \left|\int_X fd\lambda\right| + \left|\int_Y fd\lambda\right| \leq \int_X |f|d\lambda + \int_Y |f|d\lambda = \int_E |f|d\lambda.$$

□

Lemma 2.6 *If the measurable function $f : \mathbb{R}^d \to [-\infty; \infty]$ is integrable, then $|f(x)| < \infty$ for almost every x.*

Proof: Clear from the definition. □

Proposition 2.7

1. *Suppose f is measurable and nonnegative on $\mathbb{R}^d$. For every Lebesgue Measurable set A, define*

$$\phi(A) = \int_A f d\lambda.$$

Then ϕ is countably additive $\mathcal{L}$.

2. *The same conclusion holds if $f \in \mathscr{L}(\lambda)$ on $\mathbb{R}^d$.*

Proof: Firstly, notice that (2) follows from (1) if we write $f = f^+ - f^-$ and apply (1) to f^+ and to f^-. Hence we only prove (1), i.e., we have to show that

$$\phi(A) = \sum_{n=1}^{\infty} \phi(A_n)$$

if $A_n \in \mathcal{L}\,(n = 1, 2, 3, \cdots)$, $A_i \cap A_j = 0$ for $i \neq j$, and $A = \bigcup_{n=1}^{\infty} A_n$.

If f is a characteristic function, then the countable additivity of ϕ is precisely the same as the countable additivity of λ, since

$$\int_A \chi_E d\lambda = \lambda(A \cap E).$$

If f is simple, then f is of the form

$$f(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x) \quad (x \in X,\, c_i > 0).$$

Then the conclusion again holds.

In the general case, for every measurable simple function s such that $0 \leq s \leq f$,

$$\int_A s d\lambda = \sum_{n=1}^{\infty} \int_{A_n} s d\lambda \leq \sum_{n=1}^{\infty} \phi(A_n).$$

Note the middle equality follows from the countable additivity of measure. Therefore, we have

$$\phi(A) \leq \sum_{n=1}^{\infty} (A_n).$$

Now if $\phi(A_n) = +\infty$ for some n, then we clearly have $\phi(A) = \sum_{n=1}^{\infty} A_n$, since $\phi(A) \geq \phi(A_n)$. Suppose $\phi(A_n) < +\infty$ for every n. Given $\epsilon > 0$, we can choose simple functions s such that $0 \leq s \leq f$, and such that

$$\int_{A_i} s_i d\lambda \geq \int_{A_i} f d\lambda - \frac{\epsilon}{2^i}.$$

This is done by defining s piecewise on each A_i. Hence

$$\phi(A) \geq \int_A s d\lambda = \sum_{n=1}^{\infty} \int_{A_n} s d\lambda \geq \sum_{n=1}^{\infty} \phi(A_n) - \epsilon.$$

Letting $\epsilon \to 0$, we conclude that $\phi(A) = \sum_{n=1}^{\infty} \phi(A_n)$. □

Corollary 2.7.1 *Suppose (A_n) is a non-decreasing sequence of measurable set, and $A = \bigcup_{n=1}^{\infty} A_n$. Let φ be a measurable simple function, then*

$$\int_A \varphi d\lambda = \lim_{n\to\infty} \int_{A_n} \varphi d\lambda.$$

Notation: we denote $f \sim g$ on E if the set

$$\{x|f(x) \neq g(x)\} \cap E$$

has measure zero.

Corollary 2.7.2 *$\sim$ is an equivalence relation. If $f \sim g$ on E, then*

$$\int_A f d\lambda = \int_A g d\lambda,$$

provided the integral exists, for every measurable subset A of E.

Proof: $f \sim f$ since it is the set $\{x|f(x) \neq f(x)\}$ is an empty set. It is clear that $f \sim g$ implies $g \sim f$. Now, suppose $f \sim g$, $g \sim h$, then $\{x|f(x) \neq h(x)\} \subset \{x|f(x) \neq g(x)\} \cup \{x|g(x) \neq h(x)\}$, hence $f(x) \sim h(x)$. So $\sim$ is an equivalence relation. The second statement follows clearly from the fact that for every set $A \subset E$, $f = g$ almost everywhere on A. □

Proposition 2.8 *Suppose f is measurable on E, $|f| \leq g$ a.e. on E, and $g \in \mathscr{L}(\lambda)$ on E. Then $f \in \mathscr{L}(\lambda)$ on E.*

Proof: We have $f^+ \leq g$ and $f^- \leq g$ a.e. on E. □

2.2 Convergence Theorems

Theorem 2.9 (Bounded Convergence Theorem) *Suppose that $\{f_n\}$ is a sequence of measurable functions that are all bounded by M, are supported on a set E of finite measure, and $f_n(x) \to f(x)$ almost everywhere as $n \to \infty$. Then f is measurable, bounded, supported on E for almost every x, and*

$$\int |f_n - f| \to 0, \quad \textit{as} \;\; n \to \infty.$$

Consequently,

$$\int f_n \to \int f \;\; \textit{as} \;\; n \to \infty.$$

Proof: From the assumptions, one sees at once that f is bounded by M almost everywhere and vanishes outside E except possibly one a set of measure zero. Clearly, the triangle inequality for the integral implies that is suffices to prove that $\int |f_n - f| \to 0$ as n tends to infinity.

Given $\epsilon > 0$, we may find, by Egorov's theorem, a measurable subset $A|_\epsilon$ of E such that $m(E - A_\epsilon) \leq \epsilon$ and $f_n \to f$ uniformly on A_ϵ. Then we know that for all sufficiently large n we have $|f_n(x) - f(x)| \leq \epsilon$ for all $x \in A_\epsilon$. Then

$$\begin{aligned}\int |f_n(x) - f(x)|dx &\leq \int_{A_\epsilon} |f_n(x) - f(x)|dx + \int_{E-A_\epsilon} |f_n(x) - f(x)|dx \\ &\leq \epsilon m(E) + 2Mm(E - A_\epsilon) \\ &\leq \epsilon(m(E) + 2M)\end{aligned}$$

for all large n. Since ϵ is arbitrary, the proof of the theorem is complete. □

Corollary 2.9.1 *If $f \geq 0$ on $\mathbb{R}^d$. If $\int f = 0$, then $f = 0$ almost everywhere.*

Proof: If for each integer $k \geq 1$, we set $E_k = \{x \in E : f(x) \geq \frac{1}{k}\}$, then the fact that $k^{-1}\chi_{E_k}(x) \leq f(x)$ implies

$$k^{-1}m(E_k) \leq \int f$$

by monotonicity of the integral. Thus $m(E_k) = 0$ for all k, and since

$$\{x : f(x) > 0\} = \bigcup_{k=1}^{\infty} E_k,$$

we see that $f = 0$ almost everywhere. □

Theorem 2.10 (Lebesgue's Monotone Convergence Theorem) *Suppose $E \in \mathcal{L}$. Let $\{f_n\}$ be a sequence of measurable functions such that*

$$0 \leq f_1(x) \leq f_2(x) \leq \cdots \quad (x \in E).$$

Let f be defined by

$$f_n(x) \to f(x) \quad (x \in E)$$

as $n \to \infty$. Then

$$\int_E f_n d\lambda \to \int_E f d\lambda \quad (n \to \infty).$$

Proof: Since $0 \leq f_1(x) \leq f_2(x) \leq \cdots$, $x \in E$, then as $n \to \infty$,

$$\int_E f_n d\lambda \to \alpha$$

for some $\alpha \in [0, \infty]$ by monotone convergence theorem of real sequences. Since $f_n \leq f$, then $\int f_n \leq \int f$. So if $\alpha = \infty$, then it is clear that $\int f_n = \int f$. Thus assume that $\alpha < +\infty$, then we have $\alpha \leq \int_E f d\lambda$ and we want to show that $\alpha \geq \int_E f d\lambda$.

Choose $c \in [0,1)$, and let s be any simple measurable function such that $0 \leq s \leq f$. Put

$$E_n = \{x | f_n(x) \geq c \cdot s(x)\} \quad (n = 1, 2, 3, \cdots).$$

Then $E_1 \subset E_2 \subset E_3 \subset \cdots$; and since $f_n(x) \to f(x)$ pointwise, we get

$$E = \bigcup_{n=1}^{\infty} E_n.$$

So for every n, we have

$$\int_E f_n d\lambda \geq \int_{E_n} f_n d\lambda \geq c \int_{E_n} s d\lambda.$$

We let $n \to \infty$ in the integral, since the integral is countably additive set function (by Proposition 2.7), and by making the E_n's disjoint, we obtain

$$\alpha \geq c \int_E s d\lambda.$$

Letting $c \to 1$, we see that

$$\alpha \geq \int_E s d\lambda,$$

so

$$\alpha \geq \int_E f d\lambda.$$

Hence the theorem follows. (Note one can also do this using a difference of ϵ between f and f_n, instead of c. The alternative way is more natural, yet it involves much more work.) $\square$

Proposition 2.11 *Suppose $f = f_1 + f_2$, where $f_i \in \mathscr{L}(\lambda)$ on E $(i = 1, 2)$. Then $f \in \mathscr{L}(\lambda)$ on E, and*

$$\int_E f d\lambda = \int_E f_1 d\lambda + \int_E f_2 d\lambda.$$

Proof: First,suppose $f_1 \geq 0$, $f_2 \geq 0$. If f_1 and f_2 are simple, then $f = f_1 + f_2$ clearly makes the statement true. Hence, we assume that at least one of f_1, f_2 is not simple and $f_1, f_2 \geq 0$. Then we can choose monotonically increasing sequences $\{s'_n\}$, $\{s''_n\}$ of nonneagtive measurable simple functions which converge to f_1, f_2 (we know this is possible). Put $s_n = s'_n + s''_n$. Then

$$\int_E s_n d\lambda = \int_E s'_n d\lambda + \int_E s''_n d\lambda.$$

If we let $n \to \infty$, and use Lebesgue's Monotone Convergence Theorem, we have

$$\int_E f d\lambda = \int_E f_1 d\lambda + \int_E f_2 d\lambda.$$

Next, suppose $f_1 \geq 0$, $f_2 \leq 0$. We put

$$A = \{x | f(x) \geq 0\}, \quad B = \{x | f(x) < 0\}.$$

Then f, f_1, and $-f_2$ are nonnegative on A. Hence

$$(1) \qquad \int_A f_1 d\lambda = \int_A f d\lambda + \int_A (-f_2) d\lambda = \int_A f d\lambda - \int_A f_2 d\lambda.$$

Similarly, $-f$, f_1, and $-f_2$ are nonnegative on B, so that

$$\int_B (-f_2) d\lambda = \int_B f_1 d\lambda + \int_B (-f) d\lambda,$$

or

$$(2) \qquad \int_B f_1 d\lambda = \int_B f d\lambda - \int_B f_2 d\lambda,$$

and the desired equality follows from equation (1) and (2).
Lastly, for the general case, E can be decomposed into four sets E_i on each of which $f_1(x)$ and $f_2(x)$ are of constant sign. The two cases we have proved would imply

$$\int_{E_i} f d\lambda = \int_{E_i} f_1 d\lambda + \int_{E_i} f_2 d\lambda \quad (i = 1, 2, 3, 4),$$

hence we have

$$\int_E f d\lambda = \int_E f_1 d\lambda + \int_E f_2 d\lambda$$

in general. □

Corollary 2.11.1 *Suppose $E \in \mathcal{L}$. If $\{f_n\}$ is a sequence of nonnegative measurable function and*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (x \in E),$$

then

$$\int_E f d\lambda = \sum_{n=1}^{\infty} \int_E f_n d\lambda.$$

Proof: The partial sums of f forms a monotonically increasing sequence. □

Theorem 2.12 (Fatou's Theorem) *Suppose $E \in \mathcal{L}$. If $\{f_n\}$ is a sequence of nonnegative measurable functions and*

$$f(x) = \liminf_{n\to\infty} f_n(x) \quad (x \in E),$$

then

$$\int_E f d\lambda \leq \liminf_{n\to\infty} \int_E f_n d\lambda.$$

Strict inequality can occur in the above inequality.

Proof: For $n = 1, 2, 3, \cdots$ and $x \in E$, put

$$g_n(x) = \inf f_i(x) \quad (i \geq n).$$

Then g_n is measurable on E and

$$0 \le g_1(x) \le g_2(x) \le \cdots \quad (1)$$
$$g_n(x) \le f_i(x) \quad (i \ge n) \quad (2)$$
$$g_n(x) \to f(x) \quad (n \to \infty) \quad (3)$$

Then by inequality (1), (3) and Lebesgue's Monotone Convergence Theorem, we have

$$\int_E g_n d\lambda \to \int_E f d\lambda \quad (4)$$

as $n \to \infty$. So by (2), we get

$$\liminf_{n\to\infty} \int_E g_n d\lambda \le \liminf_{n\to\infty} \int_E f_n d\lambda$$

then by (4) we have

$$\int_E f d\lambda \le \liminf_{n\to\infty} \int_E f_n d\lambda.$$

□

Corollary 2.12.1 *Suppose f is a non-negative measurable function, and $\{f_n\}$ is a sequence of non-negative measurable functions with $f_n(x) \le f(x)$ and $f_n(x) \to f(x)$ for almost every x. Then*

$$\lim_{n\to\infty} \int f_n = \int f.$$

Proof: Since $f_n(x) \le f(x)$ a.e., then $\int f_n \le \int f$ so $\limsup \int f_n \le \int f$. And by Fatou's Lemma, $\int f \le \liminf \int f_n$. So the conclusion follows. □

Theorem 2.13 (Lebesgue's Dominated Convergence Theorem) *Suppose $E \in \mathcal{L}$. Let $\{f_n\}$ be a sequence of measurable functions such that*

$$f_n(x) \to f(x) \quad (x \in E)$$

as $n \to \infty$. If there exists a function $g \in \mathscr{L}(\lambda)$ on E, such that

$$|f_n(x)| \le g(x) \quad (n = 1, 2, 3, \cdots, x \in E),$$

then

$$\lim_{n\to\infty} \int_E f_n d\lambda = \int_E f d\lambda.$$

Hence if $\{f_n\}$ is dominated by g, then we have the dominated convergence. If $f_n(x) \to f(x)$ almost everywhere on E, the same conclusion also holds.

Proof: First, suppose $|f_n(x)| \le g(x)$, then by proposition 2.8, we have that $f_n \in \mathscr{L}(\lambda)$ and $f \in \mathscr{L}(\lambda)$ on E. Since $f_n + g \ge 0$, Fatou's theorem shows that

$$\int_E (f+g) d\lambda \le \liminf_{n\to\infty} \int (f_n + g) d\lambda,$$

or

$$(1) \qquad \int_E f d\lambda \leq \liminf_{n\to\infty} \int_E f_n d\lambda.$$

Since $g - f_n \geq 0$, we see similarly that

$$\int_E (g-f)d\lambda \leq \liminf_{n\to\infty} \int_E (g-f_n)d\lambda,$$

so that

$$-\int_E f d\lambda \leq \liminf_{n\to\infty} \left[-\int_E f_n d\lambda\right],$$

which is the same as

$$(2) \qquad \int_E f d\lambda \geq \limsup_{n\to\infty} \int_E f d\lambda.$$

The existence of the limit and the equality in $\lim\limits_{n\to\infty} \int_E f_n d\lambda = \int_E f d\lambda$ follows from (1) and (2). □

Corollary 2.13.1 *If $\lambda(E) < +\infty$, $\{f_n\}$ is uniformly bounded on E, and $f_n(x) \to f(x)$ on E, then*

$$\lim_{n\to\infty} \int_E f_n d\lambda = \int_E f d\lambda.$$

Hence a uniformly bounded convergent sequence is often said to be boundedly convergent.

Remark 2.13.1 *So we can see that the Lebesgue Bounded Convergence Theorem is just a special case of the Lebesgue Dominated Convergence.*

Corollary 2.13.2 (Generalized DCT) *Let $\{f_n\}$ be a sequence of measurable functions and $f_n \to f$ a.e. Moreover, let $\{g_n\}$ be a sequence of integrable functions and $g_n \to g$ a.e., if $|f_n| \leq g_n$, and $\lim \int g_n = \int g$, then*

$$\lim \int f_n = \int f.$$

2.3 Comparison With The Riemann Integral

Proposition 2.14 *Suppose f is Riemann integrable on the closed interval $[a, b]$. Then f is measurable, and*

$$\int_{[a,b]}^{\mathcal{R}} f(x)dx = \int_{[a,b]}^{\mathcal{L}} f(x)dx$$

where the integral on the left-hand side is the standard Riemann integral, and that on the right-hand side is the Lebesgue integral.

Proof: By definition, a Riemann integrable function is bounded, say $|f(x)| \leq M$, so we need to prove that f is measurable, and then establish the quality of integrals.

Again, by definition of Riemann integrability, we may construct two sequences of step functions $\{\varphi_k\}$ and $\{\psi_k\}$ that satisfy the following properties: $|\varphi_k(x)| \leq M$ and $|\psi_k(x)| \leq M$ for all $x \in [a,b]$ and

$$\varphi_1(x) \leq \varphi_2(x) \leq \cdots \leq f \leq \cdots \leq \psi_2(x) \leq \psi_1(x),$$

and

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{R}} \varphi_k(x)dx = \lim_{k\to\infty} \int_{[a,b]}^{\mathcal{R}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{R}} f(x)dx.$$

It follows immediately from their definition that for step functions, the Riemann and Lebesgue integrals agree; therefore,

$$\int_{[a,b]}^{\mathcal{R}} \varphi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)dx \quad \text{and} \quad \int_{[a,b]}^{\mathcal{R}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \psi_k(x)dx.$$

Next, if we let

$$\tilde{\varphi}(x) = \lim_{k\to\infty} \varphi_k(x) \quad \text{and} \quad \tilde{\psi}(x) = \lim_{k\to\infty} \psi_k(x),$$

we have $\tilde{\varphi} \leq f \leq \tilde{\psi}$. Moreover, both $\tilde{\varphi}$ and $\tilde{\psi}$ are measurable, and the bounded convergence theorem yields

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \varphi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \tilde{\varphi}(x)dx$$

$$\lim_{k\to\infty} \int_{[a,b]}^{\mathcal{L}} \psi_k(x)dx = \int_{[a,b]}^{\mathcal{L}} \tilde{\psi}(x)dx$$

Hence

$$\int_{[a,b]}^{\mathcal{L}} (\tilde{\psi}(x) - \tilde{\varphi}(x))dx = 0,$$

and since $\psi_k - \varphi_k \geq 0$, we must have $\tilde{\psi} - \tilde{\varphi} \geq 0$. So $\tilde{\psi} - \tilde{\varphi} = 0$, a.e., and therefore $\tilde{\varphi} = \tilde{\psi} = f$ a.e. The rest of the proof is trivial. $\square$

Theorem 2.15 *Let $[a,b]$ be a bounded interval, then f is Riemann-integrable if and only if it is bounded and the set of points where f is discontinuous has Lebesgue measure zero.*

Proof: Define

$$W_f(a) = \lim_{\delta\to 0} \sup\{|f(x) - f(x')|,\, |\, x, x' \in [a-\delta, a+\delta]\}.$$

For any $t > 0$, $\{a : w_f(a) > t\}$ is open, and $w_f(x)$ is bounded and measurable.
We claim that

$$\int^{UR} f(x)dx - \int^{LR} f(x)dx = \int w_f(x)d\mu.$$

Where UR and LR stands for upper and lower Riemann sum respectively. For partition $P = \{a_0, \cdots, a_n\}$, set

$$w_P(x) = \begin{cases} M_i - m_i & \text{for } x \in (a_{i-1}, a_i) \\ 0 & x = a_i \end{cases}.$$

Choose a sequence of partition $\{P_n\}$ of finer and finer partition so that

$$\int^{UR} f(x)dx = \lim_{n\to\infty} U(P_n, f) \text{ and } \int^{LR} f(x)dx = \lim_{n\to\infty} L(P_n, f) \text{ and } |\Delta_i| \to 0.$$

Then we will have $w_{P_n}(x) \to w_f(x)$ almost everywhere. Now

$$\sum w_{P_n}(x_i)\Delta x_i \to \int^{UR} f(x)dx - \int^{LR} f(x)dx$$

And

$$\sum w_{P_n}(x_i)\Delta x_i = \int w_{P_n} d\mu.$$

By the Lebesgue Dominated convergence, note w_P and w_f is bounded by the constant function $2\sup\|f\|$, then we have

$$\int^{UR} f(x)dx - \int^{LR} f(x)dx = \int w_f(x)d\mu.$$

Now f is Riemann integrable iff $\int^{UR} f(x)dx - \int^{LR} f(x)dx = 0$ iff $\int w_f(x)d\mu = 0$. That is $w_f(x)$ is zero almost everywhere. Interpreting w_f, $w_f(x)$ is zero iff f is continuous at x. So f is continuous for almost all x. $\square$

However, notice that a Riemann integrable function does not have to be Lebesgue integrable. For instance the function

$$\int_{\mathbb{R}} \frac{\sin(x)}{x} dx$$

is Riemann integrable (is equal to π) and is not Lebesgue integrable.

2.4 Complex-valued functions

If f is a complex-valued function on $\mathbb{R}^d$, we may write it as

$$f(x) = u(x) + iv(x),$$

where u and v are real-valued functions called the real and imaginary parts of f, repsectively. The function f is measurable if and only if both u and v are measurable. We then say that f is **Lebesgue integrable** if the function $|f(x)| = (u(x)^2 + v(x)^2)^{\frac{1}{2}}$ is Lebesgue integrable.

Now since

$$|u(x)| \leq |f(x)| \text{ and } |v(x)| \leq |f(x)|.$$

and

$$|f(x)| \leq |u(x)| + |v(x)|.$$

Then a complex-valued function is integral if and only if both its real and imaginary parts are integrable. Then the Lebesgue integral of f is defined by

$$\int f(x)dx = \int u(x)dx + i\int v(d)dx.$$

Finally, if E is measurable subset of $\mathbb{R}^d$, and f is a complex-valued measurable function on E, we say that f is Lebesgue integral on E if $f\chi_E$ is integrable on $\mathbb{R}^d$, and we define $\int_E f = \int f\chi_E$.

The collection of all complex-valued integrable function on a measurable subset $E \subseteq \mathbb{R}^d$ forms a vector space over $\mathbb{C}$. If f and g are integrable, then so is $f+g$, as $|(f+g)(x)| \leq |f(x)| + |g(x)|$. It is also clear that the integral is still linear over $\mathbb{C}$.

2.5 L^1 space

Proposition 2.16 (Characterization of Lebesgue Integrable Functions) *Suppose f is integrable on $\mathbb{R}^d$. Then for every $\epsilon > 0$:*

1. *There exists a set of finite measure B, such that*

$$\int_{B^c} |f| < \epsilon.$$

2. *(Absolute Continuity) There is a $\delta > 0$ such that*

$$\int_E |f| < \epsilon, \quad \textit{whenever } m(E) < \delta.$$

Proof: By replacing f with $|f|$ we may assume that $f \geq 0$.

For the first part, let B_N denote the ball of radius N centered at the origin, and note that if $f_N(x) = f(x)\chi_{B_N}(x)$, then $f_N \geq 0$ is measurable, $f_N(x) \leq f_{N+1}(x)$, and $f_N(x) \to \infty$ pointwise. Then by the Monotone Convergence Theorem, we have

$$\lim_{N\to\infty} \int f_N = \int f.$$

In particular, for some large N,

$$0 \leq \int f - \int f\chi_{B_N} < \epsilon,$$

and since $1 - \chi_{B_N} = \chi_{B_N^c}$, this implies that

$$\int_{B_N^c} f < \epsilon$$

for some large enough N.

For the second part, assuming that $f \geq 0$, we let $f_N(x) = f(x)\chi_{E_N}$ where

$$E_N = \{x : f(x) \leq N\}.$$

Note $f_N \geq 0$ is measurable, $f_N(x) \leq f_{N+1}(x)$, and given $\epsilon > 0$, then by the monotone convergence theorem, there exists an integer $N > 0$ such that

$$\int (f - f_N) < \frac{\epsilon}{2}.$$

We now pick $\delta > 0$ so that $N\delta < \frac{\epsilon}{2}$. If $m(E) < \delta$, then

$$\begin{aligned}\int_E f &= \int_E (f - f_N) + \int_E f_N \\ &\leq \int (f - f_N) + \int_E f_N \\ &\leq \int (f - f_N) + Nm(E) \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

$\square$

Lemma 2.17 *A measurable function $f : \mathbb{R}^d \to [-\infty; \infty]$ is Lebesgue integrable if and only if $|f|$ is Lebesgue integrable.*

Proof: By the Lebesgue Convergence Theorem, if $|f|$ is integrable, then f is integrable. Conversely, if f is integrable, then $\int f^+$ and $\int f^-$ are finite. Then $\int |f| = \int f^+ + \int f^-$ is also finite, hence integrable. $\square$

Definition 2.18 (L^1-norm) *For any integrable function f (real or complex), we define the L^1**-norm** of f to be*

$$\|f\|_1 = \|f\|_{L^1} = \|f\|_{L^1(E)} = \int_E |f(x)| dx.$$

Definition 2.19 ($L^1(\mathbb{R}^d)$) *Recall the equivalence relation $\sim$, where $f \sim g$ if $f = g$ for almost all $x \in \mathbb{R}^d$. Then we define the space $L^1(\mathbb{R}^d)$ to be the set of equivalence class of the Lebesgue integrable functions under the equivalence relation $\sim$.*

Remark 2.19.1 *For convenience sake, we will sometimes retain the terminology that an element $f \in L^1(\mathbb{R}^d)$ is an integrable function. Note then addition and scalar multiplication is well-defined, the norm $\|f\|$ of an element $f \in L^1(\mathbb{R}^d)$ is also well-defined. We will verify below that with the $L^1(\mathbb{R}^d)$ norm, the space $L^1(\mathbb{R}^d)$ is a Banach space.*

Proposition 2.20 ($L^1(\mathbb{R}^d)$ is a norm) *Suppose f and g are two functions in $L^1(\mathbb{R}^d)$, then*

1. $\|af\|_{L^1(\mathbb{R}^d)} = |a|\|f\|_{L^1(\mathbb{R}^d)}$ *for all* $a \in \mathbb{C}$;
2. $\|f + g\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} + \|g\|_{L^1(\mathbb{R}^d)}$;
3. $\|f\|_{L^1(\mathbb{R}^d)} = 0$ *if and only if* $f = 0$ *a.e.;*

Proof: Clear. $\square$

Theorem 2.21 (Riesz-Fischer) *The vector space L^1 is complete under the metric induced by the L^1 norm.*

Proof: Suppose $\{f_n\}$ is a Cauchy sequence in the norm, so that $\|f_n - f_m\|_{L^1(\mathbb{R}^d)} \to 0$ as $n, m \to \infty$. We extract a subsequence of $\{f_n\}$ that converges to f, both pointwise almost everywhere and in the norm.

Let $\{f_{n_k}\}_{k=1}^{\infty}$ be a subsequence of $\{f_n\}$ with the following property:

$$\|f_{n_{k+1}} - f_{n_k}\| \leq 2^{-k} \text{ for all } k \geq 1.$$

Such subsequence exists, since that for every 2^{-k}, $\exists N(2^{-k})$ such that whenever $n, m \geq N(\epsilon)$, $\|f_n - f_m\| \leq \epsilon$.

Now consider the series

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty}(f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty}|f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

By the monotone convergence theorem, g is integrable, and since $|f| \leq g$, hence so is f, since $|f| \leq g$. Note then that the series defining g converges almost everywhere, hence the series defining f converges almost everywhere, and since partial sums of this series are precisely the f_{n_k}, we find that

$$f_{n_k}(x) \to f(x), \quad \text{a.e.}\, x.$$

Next we show that $f_{n_k} \to f$ in L^1. Note that $|f - f_{n_k}| \leq g$ for all k, then by Lebesgue dominated convergence theorem, we get $\|f_{n_k} - f\|_{L^1} \to 0$ as k tends to infinity.

Lastly, since $\{f_n\}$ is Cauchy, and it has a convergent subsequence, then $\{f_n\}$ converges. Since L^1 is a metric space, then the limit of $\{f_n\}$ is the f we found. Thus the proof is complete. □

Corollary 2.21.1 *If $\{f_n\}_{n=1}^{\infty}$ converges to f in L^1, then there exists a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ such that*

$$f_{n_k}(x) \to f(x), \quad a.e.\ x.$$

That is convergence in L^1 implies existence of almost everywhere pointwise convergence subsequence.

Theorem 2.22 *The following families of functions are dense in $L^1(\mathbb{R}^d)$:*

- *The simple functions.*
- *The step functions.*
- *The continuous functions of compact support.*

Proof: Let f be an integrable function on $\mathbb{R}^d$. First, we may assume that f is real-valued, because we approximate its real and imaginary parts independently. If this is the case, we may then write $f = f^+ - f^-$ where $f^+, f^- \geq 0$, and it now suffices to prove the theorem when $f \geq 0$.

For (1), we know that there exists a sequence of monotone nonnegative simple function $\{\varphi_k\}$ that converges to f pointwise. By the Lebesgue dominated convergence, we have

$$\|f - \varphi_k\|_{L^1} \to 0, \quad \text{as } k \to \infty.$$

Thus there are simple functions that are arbitrarily close to f in the L^1 norm.

For (2), we first note that by (1) it suffices to approximate simple functions with compact support by step functions (Since for any $\epsilon > 0$, there exists $r > 0$, such that $\int_{B_r^c} f < \epsilon$). Then, we recall that such a simple function is a finite linear combination of characteristic functions of sets of finite measure, so it suffices to show that if E is s such a set, then there is a step function ψ so that $\|\chi_E - \psi\|_{L^1}$ is small. However, we have shown this.

By (2), it suffices to establish (3) when f is characteristic function of a rectangle bounded rectangle. Then using a piecewise linear function we can approximate such a function using a continuous function with compact support. $\square$

Theorem 2.23 (Chebyshev's Inequality) *Suppose $f \geq 0$, and f is integrable. If $\alpha > 0$ and $E_a = \{x : f(x) > a\}$. Then*

$$m(E_a) \leq \frac{1}{a} \int f.$$

Proof: Note the simple function $a\chi_{E_a} \leq f$, hence the result follows. $\square$

Lemma 2.24 *Suppose $f \geq 0$, and let $E_{2^k} = \{x : f(x) > 2^k\}$ and $F_k = \{x : 2^k < f(x) \leq 2^{k+1}\}$. If f is finite almost everywhere, then*

$$\bigcup_{k=-\infty}^{\infty} F_k = \{f(x) > 0\},$$

and the sets F_k are disjoint. Then f is integrable if and only if

$$\sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty, \quad \textit{if and only if} \quad \sum_{k=-\infty}^{\infty} 2^k m(E_{2^k}) < \infty.$$

Proof: By simple function approximation and monotone convergence theorem, we have f is integrable if and only if

$$\sum_{k=-\infty}^{\infty} 2^k m(F_k) < \infty.$$

Next, $\sum 2^k m(E_{2^k}) < \infty$ clearly implies $2^k m(F_k) < \infty$. Conversely, if $\sum 2^k m(F_k)$ converges, then

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} 2^k m(E_{2^k}) &= \sum_{k=-\infty}^{\infty} \sum_{j=k}^{\infty} 2^k m(F_j) \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{j} 2^k m(F_j)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=-\infty}^{\infty} 2^{j+1} m(F_j) \\
&= 2 \sum_{j=-\infty}^{\infty} 2^{j} m(F_j) \\
&< \infty.
\end{aligned}
$$

□

Corollary 2.24.1 *Define $f, g : \mathbb{R}^d \to \mathbb{R}$ by*

$$
f(x) = \begin{cases} |x|^{-a} & \text{if } |x| \le 1, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad g(x) = \begin{cases} |x|^{-b} & \text{if } |x| > 1, \\ 0 & \text{otherwise.} \end{cases}.
$$

Then f is integrable on $\mathbb{R}^d$ if and only if $a < d$ and g is integrable on $\mathbb{R}^d$ if and only if $b > d$.

The results shown in this section works if we replace $L^1(\mathbb{R}^d)$ with $L^1(E)$, where E is any fixed subset of $\mathbb{R}^d$ that has positive measure.

2.6 Invariance Properties of Integrals

Definition 2.25 (Translation of Functions) *If f is a function defined on $\mathbb{R}^d$, the* ***translation*** *of f by a vector $h \in \mathbb{R}^d$ is the function f_h, defined by $f_h(x) = f(x - h)$.*

Remark 2.25.1 *For any given $x \in \mathbb{R}^d$, the statement that $f_h(x) \to f(x)$ as $h \to 0$ is the same as the continuity of f at the point x.*

Proposition 2.26 *Let f be an integrable function defined on $\mathbb{R}^d$, and $h \in R^d$, then*

$$
\int_{\mathbb{R}^d} f(x-h)dx = \int_{\mathbb{R}^d} f_h(x)dx = \int_{\mathbb{R}^d} f(x)dx. \tag{2.1}
$$

Proof: Firstly, the assertion holds when $f = \chi_E$, where E is measurable. As a result of linearity, the Identity (2.1) holds for all simple functions. Now if f is nonnegative and $\{\varphi_n\}$ is a sequence of simple functions that increase pointwise a.e. to f, then $\{(\varphi_n)_h\}$ is a sequence of simple function that increase to f_h pointwise a.e. and the monotone convergence theorem implies Identity (2.1) in this case. Thus if f is complexed valued and integrable, f_h is integrable and we see that $\|f_h\|_{L^1} = \|f\|_{L^1}$. □

Similarly, we can show the following property:

Proposition 2.27 *Suppose f is an integrable function defined on $\mathbb{R}^d$, and $\delta > 0$, then*

$$
\begin{aligned}
\delta^d \int_{\mathbb{R}^d} f(\delta x)dx &= \int_{\mathbb{R}^d} f(x)dx \\
\int_{\mathbb{R}^d} f(-x)dx &= \int_{\mathbb{R}^d} f(x)dx.
\end{aligned}
$$

Corollary 2.27.1 *Suppose that f and g are a pair of measurable functions on $\mathbb{R}^d$ so that for some fixed $x \in \mathbb{R}^d$ the function $y \mapsto f(x-y)g(y)$ is integrable. Then the function $y \mapsto f(y)g(x-y)$ is then also integrable and we have*

$$\int_{\mathbb{R}^d} f(x-y)g(y)dy = \int_{\mathbb{R}^d} f(y)g(x-y)dy.$$

Proof: This follows from a change of variable which replaces y by $x-y$, which is a combination of translation and a reflection. $\square$

Definition 2.28 (Convolution) *Let f and g be two complex functions defined on $\mathbb{R}^d$, we define the **convolution of** f **and** g $f * g$ to be*

$$(f * g)(x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy = \int_{\mathbb{R}^d} f(y)g(x-y)dy.$$

Remark 2.28.1 *Note that $f * g = g * f$.*

Proposition 2.29 *Suppose $f \in L^1(\mathbb{R}^d)$, then*

$$\|f_h - f\|_{L^1} \to 0 \quad as \quad h \to 0.$$

Proof: This is true since we can approximate any integrable function with continuous functions of compact support (which is uniformly continuous). $\square$

3 Product Measure

3.1 Products of Measurable Spaces

Definition 3.1 (General Rectangles) *Suppose X and Y are sets, a* ***rectangle*** *in $X \times Y$ is a set of the form $A \times B$ where $A \subseteq X$ and $B \subseteq Y$.*

Definition 3.2 (Product of Measurable Spaces) *Let $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ be two measure space. The* ***product σ-algebra*** *$\mathscr{A}_1 \otimes \mathscr{A}_2$ is the smallest σ-algebra containing all rectangles $A_1 \times A_2$, where $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$.*

A ***measurable rectangle*** *in $\mathscr{A}_1 \otimes \mathscr{A}_2$ is a set of the form $A_1 \times A_2$, where $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$.*

Lemma 3.3 *Let $\mathscr{A}$ be the collection of unions of finitely many disjoint "rectangles", i.e., set of the form $A_1 \times A_2$, with $A_1 \in \mathscr{A}_1$, $A_2 \in \mathscr{A}_2$. Then $\mathscr{A}$ is an algebra and $\mathscr{A}_1 \otimes \mathscr{A}_2$ is the σ-algebra generated by $\mathscr{A}$.*

Proof: Clearly $\mathscr{A}$ is closed under finite unions. The collection $\mathscr{A}$ is also closed under finite intersections. To verify this claim, note that if $A_1, \cdots, A_n, C_1, \cdots, C_m \in \mathscr{A}_1$ and $B_1, \cdots, B_n, D_1, \cdots, D_n \in \mathscr{A}_2$. Then

$$
\begin{aligned}
&((A_1 \times B_1) \cup \cdots \cup (A_n \times B_n)) \cap ((C_1 \times D_1) \cup \cdots \cup (C_m \times D_m)) \\
=& \bigcup_{j=1}^{n} \bigcup_{k=1}^{m} ((A_j \times B_j) \cap (C_k \times D_k)) \\
=& \bigcup_{j=1}^{n} \bigcup_{k=1}^{m} ((A_j \cap C_k) \times (B_j \cap D_k)).
\end{aligned}
$$

Next, if $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$, then

$$(\Omega_1 \times \Omega_2) \setminus (A_1 \times A_2) = ((\Omega_1 \setminus A_1) \times \Omega_2) \cup (\Omega_1 \times (\Omega_2 \setminus A_2)).$$

Hence the complement of each measurable rectangle is in $\mathscr{A}$. Thus the complement of a finite union of measurable rectangle is in $\mathscr{A}$ by De Morgan's Laws). Hence $\mathscr{A}$ is an algebra. Then by definition $\mathscr{A}_1 \otimes \mathscr{A}_2$ is the σ-algebra generated by $\mathscr{A}$. $\square$

Definition 3.4 (Cross Sections of Sets) *Suppose X and Y are sets and $E \subseteq X \times Y$. Then for $a \in X$ and $b \in Y$, the* ***cross sections*** *$[E]_a$ and $[E]^b$ are defined by*

$$[E]_a = \{y \in Y \,:\, (a, y) \in E\} \text{ and } [E]^b = \{x \in X \,:\, (x, b) \in E\}.$$

Lemma 3.5 *Suppose $\mathscr{A}_1$ is a σ-algebra on Ω_1, and $\mathscr{A}_2$ is a σ-algebra on Ω_2. If $E \in \mathscr{A}_1 \otimes \mathscr{A}_2$, then*

- *$[E]_a \in \mathscr{A}_2$ for every $a \in \Omega_1$;*
- *$[E]^b \in \mathscr{A}_1$ for every $b \in \Omega_2$.*

Proof: Let $\mathscr{C}$ denote the collection of subsets E of $\Omega_1 \times \Omega_2$ for which the conclusion of this result holds. Then $A_1 \times A_2 \in \mathscr{C}$ for all $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$.

The collection $\mathscr{C}$ is closed under complement and countable union because

$$[(\Omega_1 \times \Omega_2) \setminus E]_a = \Omega_2 \setminus [E]_a.$$

and

$$[E_1 \cup E_2 \cup \cdots]_a = [E_1]_a \cup [E_2]_a \cup \cdots$$

for all subsets $E_1, E_2, \cdots$ of $\Omega_1 \times \Omega_2$, and all $a \in \Omega_1$. The similar statement holds for cross section with respect to all $b \in \Omega_2$.

Then $\mathscr{C}$ is a σ-algebra, hence it contains $\mathscr{A}_1 \otimes \mathscr{A}_2$. $\square$

Definition 3.6 (Cross Section of Functions) *Suppose X and Y are sets, $f : X \times Y \to \mathbb{R}$ is a function. Then for any $a \in X$ and $b \in Y$, the cross section $[f]_a : Y \to \mathbb{R}$ and $[f]^b : X \to \mathbb{R}$ are defined by*

$$[f]_a(y) = f(a, y) \text{ for } y \in Y \quad \text{and} \quad [f]^b(x) = f(x, b) \text{ for } x \in X.$$

Lemma 3.7 *Suppose $(\Omega_1, \mathscr{A}_1)$ and $(\Omega_2, \mathscr{A}_2)$ are measurable spaces, and $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ is an $\mathscr{A}_1 \otimes \mathscr{A}_2$-measurable function. Then*

- *$[f]_a$ is a $\mathscr{A}_2$ measurable function on Ω_2 for every $a \in \Omega_1$;*
- *$[f]^b$ is a $\mathscr{A}_1$ measurable function on Ω_1 for every $b \in \Omega_2$.*

Proof: Suppose D is a Borel subset of R and $a \in \Omega_1$. If $y \in \Omega_2$, then

$$\begin{aligned} y \in ([f]_a)^{-1}(D) &\Leftrightarrow [f]_a(y) \in D \\ &\Leftrightarrow f(a, y) \in D \\ &\Leftrightarrow (a, y) \in f^{-1}(D) \\ &\Leftrightarrow y \in [f^{-1}(D)]_a \end{aligned}$$

Thus

$$([f]_a)^{-1}(D) = [f^{-1}(D)]_a.$$

Because f is an $\mathscr{A}_1 \otimes \mathscr{A}_2$-measurable function, $f^{-1}(D) \in \mathscr{A}_1 \otimes \mathscr{A}_2$. Thus $([f]_a)^{-1}(D) \in \mathscr{A}_2$. Hence $[f]_a$ is a $\mathscr{A}_2$-measurable function for every $a \in \Omega_1$. The same argument works for showing $[f]^b$ is a $\mathscr{A}_1$-measurable function for every $b \in \Omega_2$. $\square$

3.2 Products of Measures

Proposition 3.8 *Suppose $(\Omega_1, \mathscr{A}_1, \mu_1)$, $(\Omega_2, \mathscr{A}_2, \mu_2)$ are two σ-finite measure space. If $E \in \mathscr{A}_1 \otimes \mathscr{A}_2$, then*

- *$x \mapsto \mu_2([E]_x)$ is an $\mathscr{A}_1$-measurable function on Ω_1;*
- *$y \mapsto \mu_1([E]^y)$ is an $\mathscr{A}_2$-measurable function on Ω_2.*

Definition 3.9 (Iterated Integrals) *Suppose (X, S, μ) and (Y, T, v) are measurable spaces and $f : X \times Y \to \mathbb{R}$ is a function. Then*

$$\int_X \int_Y f(x,y)dv(y)d\mu(x) \ \text{ means } \ \int_X \left(\int_Y f(x,y)dv(y) \right) d\mu(x).$$

Definition 3.10 *Suppose $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ are σ-finite measure spaces. Then we define the **product measure** $\mu = \mu_1 \otimes \mu_2 : \mathscr{A}_1 \otimes \mathscr{A}_2 \to \mathbb{R} \cup \{\infty\}$ to be the measure given by*

$$\mu(E) = \int_{\Omega_1} \int_{\Omega_2} \chi_E(x,y)d\mu_2(y)d\mu_1(x),$$

for $E \in \mathscr{A}_1 \otimes \mathscr{A}_2$.

Remark 3.10.1 *The measure on $\mathscr{A}_1 \otimes \mathscr{A}_2$ is the unique function $\mu_0 : \mathscr{A} \to \mathbb{R} \cup \{\infty\}$ such that*

$$\mu(A_1 \times A_2) = \mu_1(A)\mu_2(A)$$

and

$$\mu(A \sqcup A') = \mu(A) + \mu(A').$$

In fact, we can also get this measure from Caracthéodory's Extension theorem by checking that $\mu(A_1 \times A_2) = \mu_1(A_1) \times \mu_2(A_2)$ for all $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$ defines a premeasure on the algebra of measurable rectangles. The uniqueness also follows from the Extension Theorem.

Theorem 3.11 *Suppose $(\Omega_1, \mathscr{A}_1, \mu_1)$ and $(\Omega_2, \mathscr{A}_2, \mu_2)$ are σ-finite measure spaces, then $\mu_1 \otimes \mu_2$ is a measure on $(\Omega_1 \times \Omega_2, \mathscr{A}_1 \otimes \mathscr{A}_2)$.*

Proof: Clearly $(\mu_1 \otimes \mu_2)(\emptyset) = 0$. Suppose $E_1, E_2, \cdots$ is a disjoint sequence of sets in $\mathscr{A}_1 \otimes \mathscr{A}_2$. Then

$$\begin{aligned}
(\mu_1 \otimes \mu_2)\left(\bigcup_{k=1}^{\infty} E_k \right) &= \int_{\Omega_1} \mu_2 \left(\left[\bigcup_{k=1}^{\infty} E_k \right]_x \right) d\mu_1(x) \\
&= \int_X \mu_2 \left(\bigcup_{k=1}^{\infty} ([E_k]_x) \right) d\mu_1(x) \\
&= \int_X \left(\sum_{k=1}^{\infty} \mu_2([E_k]_x) \right) d\mu_1(x) \\
&= \sum_{k=1}^{\infty} \int_X \mu_2([E_k]_x) d\mu_1(x) \\
&= \sum_{k=1}^{\infty} (\mu_1 \otimes \mu_2)(E_k),
\end{aligned}$$

where the fourth equality follows from the Monotone Convergence Theorem. Thus $\mu_1 \otimes \mu_2$ is countable additive. $\square$

Definition 3.12 (Probability Space) *A measure space $(\Omega, \mathscr{A}, \mu)$ is a **probability space** if $\mu(\Omega) = 1$. We call the σ-algebra in this case as a σ-field, the measurable sets to be events, and we call measurable functions as random variables, and we call the measure function μ to be P.*

Definition 3.13 (Infinite Product) *Let $(\Omega_n, \mathscr{A}_n, \mu_n)_{n\in\mathbb{N}}$ be a countable collection of measure spaces. Then we define*

$$\bigotimes_{n\in\mathbb{N}} \mathscr{A}_n = \sigma\left(\{A_1 \times \cdots \times A_n \times \Omega_{n+1} \times \Omega_{n+2} \times \cdots \mid A_i \in \mathscr{A}_i\}\right).$$

If each $(\Omega_n, \mathscr{A}, \mu_n)$ as a product space, then we define the measure $\bigotimes \mu_n : \bigotimes_{n\in\mathbb{N}} \mathscr{A}_n \to \mathbb{R}$ by

$$\bigotimes \mu_n(A_1 \times \cdots \times A_n \times \Omega_{n+1} \times \cdots) = \prod_{i=1}^{n} \mu_i(A_i).$$

3.3 Iterated Integrals

Theorem 3.14 (Tonelli's Theorem) *Suppose (X, S, μ) and (Y, T, ν) are σ-finite measure spaces. Suppose $f : X \times Y \to [0, \infty]$ is $S \otimes T$ measurable. Then*

1. *$x \mapsto \int_Y f(x, y)d\nu(y)$ is an S measurable function on X;*
2. *$y \mapsto \int_X f(x, y)d\mu(x)$ is a T measurable function on Y*

and

$$\int_{X\times Y} fd(\mu \otimes \nu) = \int_X \int_Y f(x, y)d\nu(y)d\mu(x) = \int_Y \int_X f(x, y)d\mu(x)d\nu(y).$$

Theorem 3.15 (Fubini's Theorem) *Suppose (X, S, μ) and (Y, T, ν) are σ-finite measure spaces. Suppose $f : X \times Y \to [-\infty, \infty]$ is $S \otimes T$ measurable and $\int_{X\times Y} |f|d(\mu \otimes \nu) < \infty$. Then*

$$\int_Y |f(x, y)|d\nu(y) < \infty \text{ for almost every } x \in X$$

and

$$\int_X |f(x, y)|d\mu(x) < \infty \text{ for almost every } y \in Y.$$

Furthermore,

$$x \mapsto \int_Y f(x, y)d\nu(y) \text{ is an } S\text{-measurable function on } X,$$

and

$$y \mapsto \int_X f(x, y)d\mu(x) \text{ is a } T\text{-measurable function on } Y,$$

and

$$\int_{X\times Y} fd(\mu \otimes \nu) = \int_X \int_Y f(x, y)d\nu(y)d\mu(x) = \int_Y \int_X f(x, y)d\mu(x)d\nu(y).$$

Remark 3.15.1 *Note that $\int_Y f(, y)d\nu(y)$ and $\int_X f(x, y)d\mu(x)$ may not be defined everywhere (it is defined almost everywhere). So when it is undefined, we can set its value to be 0.*

Remark 3.15.2 *The Lebesgue measure on $\mathbb{R}^2$ is not the product measure of Lebesgue measures on $\mathbb{R}$.*

Next, we want to apply Fubini's Theorem for $\mathbb{R}^d$. We may write $\mathbb{R}^d$ as a product $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, where $d = d_1 + d_2$, and $d_1, d_2 \geq 1$. However, we need to note that the Lebesgue measure on $\mathbb{R}^d$ is not the product measure of the Lebesgue measure on $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$. This is because the product measure is not complete, while the Lebesgue measure is complete, so we only have

$$\mathcal{L}(\mathbb{R}^{d_1}) \otimes \mathcal{L}(\mathbb{R}^{d_2}) \subset \mathcal{L}(\mathbb{R}^d).$$

Then if a function is measurable on $\mathbb{R}^d$, it might not necessarily be measurable with respect to the product measure $\lambda(\mathbb{R}^{d_1}) \otimes \lambda(\mathbb{R}^{d_2})$. However, what is true is that the slice function will be measurable almost everywhere, which will not affect integrability.

Proposition 3.16 *The restriction of Lebesgue Measure on $\mathbb{R}^d$ agrees with the product of Lebesgue Measure on $\mathbb{R}^{d_1}$ and $\mathbb{R}^{d_2}$ on the product space $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$. Moreover, the completion of the product measure agrees with the Lebesgue measure on $\mathbb{R}^d$.*

Proof: Show $\lambda_{\mathbb{R}^d} = \lambda_{\mathbb{R}^{d_1}} \otimes \lambda_{\mathbb{R}^{d_2}}$ using the Carathéodory Extension Theorem. Next, by the uniqueness of completion, we conclude that the completion of the product measure agrees with the Lebesgue measure on $\mathbb{R}^d$. $\square$

Corollary 3.16.1 (Fubini's Theorem for $\mathbb{R}^d$) *Suppose $f(x, y)$ is an Lebesgue integrable function on $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ (can be complex), and*

$$\int_{\mathbb{R}^d} |f| d\lambda < \infty.$$

Then

$$\int_{\mathbb{R}^{d_1}} |f(x, y)| dy < \infty \quad \text{for almost every } x \in \mathbb{R}^{d_1}$$

and

$$\int_{\mathbb{R}^{d_2}} |f(x, y)| dx < \infty \quad \text{for almost every } y \in \mathbb{R}^{d_2}.$$

Furthermore,

$$x \mapsto \int_{\mathbb{R}^{d_2}} f(x, y) dy \quad \text{is an Lebesgue integrable function on } \mathbb{R}^{d_1},$$

and

$$y \mapsto \int_{\mathbb{R}^{d_1}} f(x, y) dx \quad \text{is a Lebsgue integrable function on } \mathbb{R}^{d_2},$$

and

$$\int_{\mathbb{R}^d} f d\lambda = \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} f(x, y) dy dx = \int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_1}} f(x, y) dx dy.$$

Furthermore, if f is non-negative, then

$$\int_{\mathbb{R}^d} f d\lambda = \int_{\mathbb{R}^{d_1}} \int_{\mathbb{R}^{d_2}} f(x, y) dy dx = \int_{\mathbb{R}^{d_2}} \int_{\mathbb{R}^{d_1}} f(x, y) dx dy$$

in the extended sense.

Corollary 3.16.2 *If E is a measurable set in $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, then for every $y \in \mathbb{R}^{d_2}$, the cross section*

$$[E]^y = \{x \in \mathbb{R}^{d_1} \,:\, (x, y) \in E\}$$

is a measurable subset of $\mathbb{R}^{d_1}$. And $m([E]^y)$ is a measurable function of y and

$$m(E) = \int_{\mathbb{R}^{d^2}} m([E]^y) dy.$$

Lemma 3.17 *Suppose f is a measurable function on $\mathbb{R}^{d_1}$. Then the function $\tilde{f}$ defined by $\tilde{f}(x, y) = f(x)$ is measurable on $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$.*

Suppose f is a measurable function on $\mathbb{R}^d$, then the function $\tilde{f}(x, y) = f(x - y)$ is a measurable on $\mathbb{R}^d \times \mathbb{R}^d$.

Proof: For the first statement, the function $F((a, b)) \to a$ is a real and continuous function on $\mathbb{R}^2$. Since f is measurable and the constant function is measurable, then

$$\tilde{f}(x, y) = F(f(x), y)$$

is measurable by Proposition 1.25.

The proof for the second statement is similar, but we also need to use the fact that the composition of a continuous function with a measurable function is measurable. □

Corollary 3.17.1 (Area of a graph) *Suppose $f(x)$ is a non-negative function on $\mathbb{R}^d$, and let*

$$\mathscr{A} = \{(x, y) \in \mathbb{R}^d \times \mathbb{R} \,:\, 0 \le y \le f(x)\}.$$

Then

- *f is measurable on $\mathbb{R}^d$ if and only if $\mathscr{A}$ is measurable in $\mathbb{R}^{d+1}$.*
- *If f is measurable, then*
$$\int_{\mathbb{R}^d} f(x) dx = m(\mathscr{A}).$$

Proof: If f is measurable on $\mathbb{R}^d$, then $F(x, y) = y - f(x)$ is measurable on $\mathbb{R}^{d+1}$. Since $\mathscr{A} = \{y \ge 0\} \cap \{F \le 0\}$, then $\mathscr{A}$ is also measurable. Conversely, if $\mathscr{A}$ is measurable. We note that for each $x \in \mathbb{R}^{d_1}$, the slice $[A]_x$ is a closed segment, namely $[A]_x = [0, f(x)]$. By Corollary 3.16.2, we have the measurability of $m([\mathscr{A}]_x) = f(x)$. Moreover,

$$m(\mathscr{A}) = \int \chi_{\mathscr{A}}(x, y) dx dy = \int_{\mathbb{R}^{d_1}} m(\mathscr{A}_x) dx = \int_{\mathbb{R}^{d_1}} f(x) dx.$$

□

Corollary 3.17.2 *Suppose f is integrable on $\mathbb{R}^d$. For each $\alpha > 0$, let $E_\alpha = \{x \,:\, |f(x)| > \alpha\}$, then*

$$\int_{\mathbb{R}^d} |f(x)| dx = \int_0^\infty m(E_\alpha) d\alpha.$$

Proof: Follows directly from Corollary 3.16.2 and 3.17.1. $\square$

3.4 Volume of d-balls

For $d = 0$, we set the volume of the unit 0-ball to be 1, for $d = 1$, a unit 1-ball is a segment with length 2, hence has measure 2. For $d = 2$, by Corollary 3.16.2, we have

$$\alpha(2) = 2\int_{-1}^{1} (1 - x^2)^{\frac{1}{2}} dx.$$

Not that $(1 - x^2)^{\frac{1}{2}}$ is Riemann Measurable on $[-1; 1]$, so

$$\int_{-1}^{1} (1 - x^2)^{\frac{1}{2}} dx = \left[\frac{1}{2}(\sin^{-1} x + x\sqrt{1 - x^2})\right]_{x=-1}^{x=1} = \frac{\pi}{2}.$$

Thus $\alpha(2) = \pi$. Next, by Corollary 3.16.2 again, for $d > 2$ we have

$$\alpha(d) = 2\int_{0}^{1} (1 - y^2)^{\frac{d-1}{2}} \alpha(d-1) dy.$$

This is because a $d - 1$ ball with radius r has measure $r^{d-1}\alpha(d - 1)$. So

$$\begin{aligned}
\alpha(d) &= 2\alpha(d-1)\int_{0}^{1} (1 - y^2)^{\frac{d-1}{2}} dy \\
&= \alpha(d-1)\int_{0}^{1} z^{\frac{1}{2}-1}(1 - z)^{\frac{d+1}{2}-1} dz \\
&= \alpha(d-1)B\left(\frac{1}{2}, \frac{d+1}{2}\right) \\
&= \alpha(d-1)\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}+1\right)} \\
&= \alpha(d-1)\frac{\pi}{\cdot}\frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}+1\right)}
\end{aligned}$$

Then for all $d \geq 3$, we have

$$\alpha(d) = \alpha(2)\sqrt{\pi}^{d-2} \cdot \frac{1}{\Gamma(\frac{d}{2}+1)} = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}.$$

Note this formula also works for $d = 0, 1$ and $d = 2$. Hence for all $d \in \mathbb{N}_{\geq 0}$, we have

$$\alpha(d) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}.$$

Then we have the following formula:

$$\lambda(B_d(R)) = \begin{cases} 1 & d = 0 \\ 2R & d = 1 \\ \frac{2\pi}{n} R^2 \times \lambda(B_{d-2}(R)) & d \geq 2 \end{cases}.$$

3.5 Integration Formula for Polar Coordinates

The polar coordinates of a point $x \in \mathbb{R}^d \setminus \{0\}$ are the pair (r, γ), where $0 < r < \infty$ and γ belongs to the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d,\ |x| = 1\}$. These are determined by

$$r = |x|, \quad \gamma = \frac{x}{|x|} \quad \Leftrightarrow \quad x = r\gamma. \tag{3.1}$$

In this section, we aim to prove the formula

$$\int_{\mathbb{R}^d} f(x)dx = \int_{S^{d-1}} \left(\int_0^\infty f(r\gamma) r^{d-1} dr \right) d\sigma(\gamma) \tag{3.2}$$

under appropriate hypotheses. For this, we consider the following pair of measure space. First, $(X_1, \mathscr{A}_1, \mu_1)$, where $X_1 = (0, \infty)$, $\mathscr{A}_1$ is the collection of Lebesgue measurable sets in $(0, \infty)$, and $d\mu_1(r) = r^{d-1}dr$ in the sense that $\mu_1(E) = \int_E r^{d-1}dr$. Next, $X-2$ is the unit sphere S^{d-1}. Next, we let X_2 be the unit sphere S^{d-1}, and the measure μ_2 is determined by Equation (3.2) with $\mu_2 = \sigma$. Given any set $E \subset S^{d-1}$ we let $\tilde{E} = \{x \in \mathbb{R}^d \, : \, x/|x| \in E,\, 0 < |x| < 1\}$ be the "sector" in the unit ball whose "end-points" are in E. We shall say $E \in \mathscr{A}_2$ exactly when $\tilde{E}$ is a Lebesgue measurable subset of $\mathbb{R}^d$, and define $\mu_2(E) = \sigma(e) = d \cdot m(\tilde{E})$, where m is the Lebesgue measure in $\mathbb{R}^d$. In this way both $(X_1, \mathscr{A}_1, \mu_1)$ and $(X_2, \mathscr{A}_2, \mu_2)$ satisfies all the properties of complete and σ-finite measure spaces. We note also that the sphere S^{d-1} has a metric on it given by $d(\gamma, \gamma') = |\gamma - \gamma'|$. If E is an open set with respect to this metric in S^{d-1}, then $\tilde{E}$ is open in $\mathbb{R}^d$, and hence E is a measurable set in S^{d-1}.

Theorem 3.18 *Suppose f is an integrable function on $\mathbb{R}^d$. Then for almost every $\gamma \in S^{d-1}$ the slice f^γ defined by $f^\gamma(r) = f(r\gamma)$ is an integrable function with respect to the measure $r^{d-1}dr$. Moreover, $\int_0^\infty f^\gamma(r) r^{d-1} dr$ is integrable on S^{d-1} and*

$$\int_{\mathbb{R}^d} f(x)dx = \int_{S^{d-1}} \left(\int_0^\infty f(r\gamma) r^{d-1} dr \right) d\sigma(\gamma).$$

Alternatively, we can reverse the order of integration.

Proof: We consider the product measure $\mu = \mu_1 \times \mu_2$ on $X_1 \times X_2$. Since the space $X_1 \times X_2 = \{(r, \gamma) \, : \, 0 < r < \infty \text{ and } \gamma \in S^{d-1}\}$ can be identified with $\mathbb{R}^d \setminus \{0\}$, we can think of μ as a measure of the latter space, and our main task is to identify it with the (restriction of) Lebesgue measure on that space. We claim first that whenever $E = E_1 \times E_2 \subset X_1 \times X_2$ is a measurable rectangle, we have

$$m(E) = \mu(E) = \mu_1(E_1)\mu_2(E_2).$$

In fact this holds for E_2 an arbitrary measurable subset of S^{d-1} and $E_1 = (0, 1)$, because then $E = E_1 \times E_2$ is the sector $\tilde{E}_2$, while $\mu_1(E_1) = \frac{1}{d}$.

Because of the relative dilation-invariant of Lebesegue measure, $m(E) = \mu(E)$ holds whenever $E = (0, b) \times E_2$, $b > 0$. Then a simple limiting argument proves the same identity for the case $E_1 = (0, a]$ and subtraction to all open intervals $E_1 = (a, b)$ and thus for all open sets. Thus we have $m(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2)$ for all open sets E_1, and hence for all closed sets, and therefore for all Lebesgue measurable sets. Again by countable additivity, we have $m(E) = \mu(E)$ when E is a countable union of measurable rectangles. Now we note that any open set in $\mathbb{R}^d \setminus \{0\}$ can be written as a countable union of rectangles in $\bigcup_{j=1}^{\infty} A_j \times B_j$, where A_j and B_j are open in $(0, \infty)$ and S^{d-1}, respectively. This shows that any open set in $\mathbb{R}^d$ is in $\mathscr{A}_1 \otimes \mathscr{A}_2$ and the Lebesgue measure agrees with the product measure, hence this also holds for all Borel sets in $\mathbb{R}^d$. For Lebesgue measurable sets, by completing $\mathscr{A}_1 \otimes \mathscr{A}_2$, we see that the identity $m(E) = \mu(E)$ also holds.

Hence, we have proven the formula for integration in polar coordinates for the special case $f = \chi_E$. Then by an standard argument, the result holds for all Lebesgue measurable functions. $\square$

4 Differentiation

4.1 Differentiation Of The Integral

In this section, let f be an integrable function on $\mathbb{R}^d$, we want to study whether it is true that

$$\lim_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)}\int_B f(y)dy = f(x), \quad \text{for a.e. } x.$$

We refer to this question as the averaging problem.

Firstly, if f is continuous, then the result is clear, as

$$f(x) - \frac{1}{m(B)}\int_B f(y)dy = \frac{1}{m(B)}\int_B (f(x)-f(y))dy \leq \frac{1}{m(B)}\int_B |f(x)-f(y)|dy.$$

As f is continuous, then the RHS can be made arbitrarily small as $m(B)\to 0$.

Definition 4.1 (Maximal Function) *If f is integrable on $\mathbb{R}^d$, we define its **maximal function** f^* by*

$$f^*(x) = \sup_{x\in B} \frac{1}{m(B)}\int_B |f(y)|dy, \quad x\in\mathbb{R}^d,$$

where the supremum is taken over all balls containing the point x.

Theorem 4.2 *Suppose f is integrable on $\mathbb{R}^d$. Then:*

1. *f^* is measurable.*
2. *$f^*(x) < \infty$ for a.e. x.*
3. *f^* satisfies*
$$m(\{x\in\mathbb{R}^d : f^*(x) > \alpha\}) \leq \frac{A}{\alpha}\|f\|_{L^1(\mathbb{R}^d)} \tag{4.1}$$
for all $\alpha > 0$, where $A = 3^d$.

Proof:

1. We show the set $E_\alpha = \{x\in\mathbb{R}^d : f^*(x) > \alpha\}$ is open, then by the second property, we have f^* is measurable. If $x\in E_\alpha$, then there exists a ball B such that $x\in B$ and
$$\frac{1}{m(B)}\int_B |f(y)|dy > \alpha.$$
Now any point y close enough to x will also belong to B, hence $x\in E_\alpha$.

2. Note
$$\{x : f^*(x) = \infty\} \subset \{x : f^*(x) > \alpha\}$$
for all α. Taking the limit as α tends to infinity, the thrid property yields $m(\{x : f^*(x) = \infty\}) = 0$.

3. Let $E_\alpha = \{x : f^*(x) > \alpha\}$, then for each $x \in E_\alpha$ there exists a ball B_x that contains x and such that

$$\frac{1}{m(B_x)} \int_{B_x} |f(y)|dy > \alpha.$$

Therefore, for each ball B_x we have

$$m(B_x) < \frac{1}{\alpha} \int_{B_x} |f(y)|dy. \tag{4.2}$$

Fix a compact subset K of E_α. Since K is covered by $\bigcup_{x \in E_\alpha} B_x$, we may select a finite subcover of K, say $K \subset \bigcup_{l=1}^N B_l$. The covering lemma (4.3) guarantees the existence of a sub-collection $B_{i_1}, \cdots, B_{i_k}$ of disjoint balls with

$$m\left(\bigcup_{l=1}^N B_l\right) \leq 3^d \sum_{j=1}^k m(B_{i_j}). \tag{4.3}$$

Since the balls $B_{i_1}, \cdots, B_{i_k}$ are disjoint and satisfy (4.2) and (4.3), then

$$\begin{aligned} m(K) &\leq m\left(\bigcup_{l=1}^N B_l\right) \\ &= 3^d \sum_{j=1}^k m(B_{i_j}) \\ &= \frac{3^d}{\alpha} \sum_{j=1}^k \int_{B_{i_j}} |f(y)|dy \\ &= \frac{3^d}{\alpha} \int_{\bigcup B_{i_j}} |f(y)|dy \\ &\leq \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(y)|dy. \end{aligned}$$

Since this inequality is true for all compact subset K of E_α, the proof of the weak type inequality for the maximal operator is complete.

□

Lemma 4.3 (Vitali Covering Lemma) *Suppose $\mathcal{B} = \{B_1, B_2, \cdots, B_N\}$ is a finite collection of open balls in $\mathbb{R}^d$. Then there exists a disjoint sub-collection $B_{i_1}, B_{i_2}, \cdots, B_{i_k}$ of $\mathcal{B}$ that satisfies*

$$m\left(\bigcup_{l=1}^N B_l\right) \leq 3^d \sum_{j=1}^k m(B_{i_j}).$$

Proof: Firstly, pick a ball B_{i_1} in $\mathcal{B}$ with maximal radius, and then delete from $\mathcal{B}$ the ball B_{i_1}, as well as any balls that intersection B_{i_1}. Thus all the balls that are deleted are contained in the ball $\tilde{B}_{i_1}$ concentric with B_{i_1}, but with 3 times its radius.

The remaining balls yield a new collection B', for which we repeat the procedure. We pick B_{i_2} with largest radius in B' and delete from B' the ball B_{i_2} and any ball that intersects with it. Continuing this way we find, after at most N steps, a collection of disjoint balls $B_{i_1}, \cdots, B_{i_k}$. It is clear that the collection of balls we selected satisfies the desired inequality. $\square$

Using the same proof, we can in fact prove the following more general result:

Lemma 4.4 (Vitali Covering Lemma: Metric Space Finite Version) *Let $B_1, \cdots, B_n$ be any finite collection of balls contained in an arbitrary metric space. Then there exists a subcollection $B_{j_1}, \cdots, B_{j_m}$ of these balls which are disjoint and satisfy*

$$B_1 \cup B_2 \cdots \cup B_n \subseteq 3B_{j_1} \cup \cdots \cup 3B_{j_m}$$

where if $B = B(x, r)$, then $3B(x, r) = B(x, 3r)$.

The story becomes a little bit different when the collection of balls are infinite, we have the following theorem:

Theorem 4.5 (Vitali Infinite Covering Lemma) *Let $\mathcal{F}$ be an arbitray collection of balls in a separable metric space such that*

$$R := \sup\{\text{diam}(B) : B \in \mathcal{F}\} < \infty.$$

Then there exists a countable subcollection $\mathcal{G} \subset \mathcal{F}$ such that the balls of $\mathcal{G}$ are pairwise disjoint and satisfy

$$\bigcup_{B \in \mathcal{F}} B \subseteq \bigcup_{C \in \mathcal{G}} 5C.$$

Moreover, each $B \in \mathcal{F}$ intersects some $C \in \mathcal{G}$ with $B \subseteq 5C$.

Proof: Consider the partition of $\mathcal{F}$ into subcollections F_n, $n \geq 0$, defined by

$$F_n = \{B \in \mathcal{F} : 2^{-n-1}R < \text{diam}(B) \leq 2^{-n}R\}.$$

We define a sequence G_n with $G_n \subset F_n$ inductively:

- Set $H_0 = F_0$ and let G_0 be the maximal disjoint subcollection of H_0 (such subcollection exists by Zorn's Lemma).
- Assuming that $G_0, \cdots, G_n$ have been defined, let

$$H_{n+1} = \{B \in F_{n+1} : B \cap C = \emptyset, \ \forall C \in G_0 \cup \cdots \cup G_n\}.$$

- Define G_{n+1} to be the maximal disjoint subcollection of H_{n+1}.

Finally, we define

$$\mathcal{G} := \bigcup_{n=0}^{\infty} G_n.$$

Then $\mathcal{G} \subset \mathcal{F}$, and it is a disjoint collection, so it would be countable since the given metric space is separable. Next, let $B \in \mathcal{F}$, there must be some n such that B belongs to F_n, either B belongs to H_n, which implies $n > 0$ and means that B intersects a ball from the union of $G_0, \cdots, G_{n-1}$ or $B \in H_n$ and by maximality of G_n, B intersects a ball in G_n. In any case, B intersects a ball C that belongs to the union of $G_0, \cdots, G_n$. Such a ball C must have a radius larger than $2^{-n-1}R$. Since the radius of B is less than or equal to $2^{-n}R$¡ we conclude that $B \subset 5C$. $\square$

4.2 Lebesgue Differentiation Theorem

Theorem 4.6 (Lebesgue Theorem) *If f is integrable on $\mathbb{R}^d$, then*

$$\lim_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)} \int_B f(y)dy = f(x) \quad \textit{for a.e. } x. \tag{4.4}$$

Proof: It suffices to show that for each $\alpha > 0$, the set

$$E_\alpha = \left\{ x \, : \, \limsup_{\substack{m(B)\to 0\\ x\in B}} \left| \frac{1}{m(B)} \int_B f(y)dy - f(x) \right| > 2\alpha \right\}$$

has measure zero.

Fix α, for each $\epsilon > 0$, we may select a continuous function g of compact support with $\|f - g\|_{L^1(\mathbb{R}^d)} < \epsilon$. The continuity of g implies that

$$\lim_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)} \int_B g(y)dy = g(x), \quad \text{for all x}$$

Since we may write the difference $\frac{1}{m(B)} \int_B f(y)dy - f(x)$ as

$$\frac{1}{m(B)} \int_B (f(y) - g(y))dy + \frac{1}{m(B)} \int_B g(y)dy - g(x) + g(x) - f(x)$$

we find that

$$\limsup_{\substack{m(B)\to 0\\ x\in B}} \left| \frac{1}{m(B)} \int_B f(y)dy - f(x) \right| \le (f-g)^*(x) + |g(x) - f(x)|,$$

where the symbol $*$ indicates the maximal function. Consequently, if

$$F_\alpha = \{x \, : \, (f-g)^*(x) > \alpha\} \quad \text{and} \quad G_\alpha = \{x \, : \, |f(x) - g(x)| > \alpha\}$$

Then $E_\alpha \subseteq (F_\alpha \cup G_\alpha)$. Because if u_1 and u_2 are positive, then $u_1 + u_2 > 2\alpha$ only if $u_i > \alpha$ for at least one u_i. On the one hand, Chebyshev's Inequality yields

$$m(G_\alpha) \le \frac{1}{\alpha} \|f - g\|_{L^1(\mathbb{R}^d)},$$

and on the other hand, the weak type estimate for the maximal function gives

$$m(F_\alpha) \le \frac{A}{\alpha} \|f - g\|_{L^1(\mathbb{R}^d)}.$$

The function g was selected so that $\|f - g\|_{L^1(\mathbb{R}^d)} < \epsilon$. Hence we get

$$m(E_\alpha) \le \frac{A}{\alpha}\epsilon + \frac{1}{\alpha}\epsilon.$$

Since ϵ is arbitrary, we must have $m(E_\alpha) = 0$, and the proof is complete. $\square$

Corollary 4.6.1 *If f is integrable on $\mathbb{R}^d$, then $f^*(x) \geq |f(x)|$ for almost every x.*

Proof: Since for almost every $x \in \mathbb{R}^d$, we have

$$|f(x)| = \lim_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)} \int_B |f(y)|dy \leq f^*(x)$$

by the definition of maximal function. □

Definition 4.7 (Locally Integrable) *We say a measurable function f on $\mathbb{R}^d$ is **locally integrable**, if for every ball B the function $f(x)\chi_B(x)$ is integrable. We shall denote by $L^1_{loc}(\mathbb{R}^d)$ the space of all locally integrable function.*

Clearly, theorem 4.6 holds for $f \in L^1_{loc}(\mathbb{R}^d)$.

Theorem 4.8 *If $f \in L^1_{loc}(\mathbb{R}^d)$, then*

$$\lim_{\substack{m(B)\to 0\\ x\in B}} \frac{1}{m(B)} \int_B f(y)dy = f(x) \quad \textit{for a.e. } x.$$

Corollary 4.8.1 *If f is integrable on $[a,b]$, and*

$$\int_a^x f(t)dt = 0$$

for all $x \in [a,b]$, then $f(t) = 0$ a.e.

Corollary 4.8.2 (Fundamental Theorem Of Calculus Part 1) *If $f : [a,b] \to \mathbb{R}$ is integrable on $[a,b]$, and*

$$F(x) = \int_a^x f(t)dt + F(a),$$

then $F'(x) = f(x)$ for a.e. $x \in [a,b]$.

Definition 4.9 (Lebesgue Density) *If E is a measurable set and $x \in \mathbb{R}^d$, we say that x is a point of **Lebesgue density** of E if*

$$\lim_{\substack{m(B)\to 0\\ x\in B}} \frac{m(B\cap E)}{m(B)} = 1.$$

Remark 4.9.1 *Loosely speaking, this condition says that small balls around x are almost entirely covered by E. More precisely, for every $\alpha < 1$ close to 1, and every ball of sufficiently small radius containing x, we have*

$$m(B \cap E) \geq \alpha m(B).$$

Corollary 4.9.1 *Suppose E is a measurable subset of $\mathbb{R}^d$. Then:*

- *Almost every $x \in E$ is a point of density of E.*
- *Almost every $x \notin E$ is not a point of density of E.*

Proof: Consider the characteristic function of E and apply the Lebesgue Theorem to χ_E. $\square$

Corollary 4.9.2

Definition 4.10 (Lebesgue Set) *If f is locally integrable on $\mathbb{R}^d$, the **Lebesgue set** of f consists of all points $\bar{x} \in \mathbb{R}^d$ for which $f(\bar{x})$ is finite and*

$$\lim_{\substack{m(B)\to 0 \\ \bar{x}\in B}} \frac{1}{m(B)} \int_B |f(y) - f(\bar{x})| dy = 0.$$

Remark 4.10.1 *Note $\bar{x}$ belongs to the Lebesgue set of f whenever f is continuous at $\bar{x}$. Also if $\bar{x}$ is in the Lebesgue set of f, then*

$$\lim_{\substack{m(B)\to 0 \\ x\in B}} \frac{1}{m(B)} \int_B f(y) dy = f(\bar{x}).$$

Corollary 4.10.1 *If f is locally integrable on $\mathbb{R}^d$, then almost every point belongs to the Lebesgue set of f.*

Remark 4.10.2 *As elements of $L^1(\mathbb{R}^d)$ are actually equivalence classes, where two functions are equivalent if they differ on a set of measure zero. However, we see that the limit in Equation (4.4) is independent of the representation of f chosen, because the integral $\int_B f(y)dy = \int_B g(y)dy$ if $f \cong g$. However, the Lebesgue set of f depends on the particular representative of f that we consider.*

Proof: Apply Theorem 4.8 to the function $|f(y) - r|$ shows that for each rational r, there exists a set E_r of measure zero, such that

$$\lim_{\substack{m(B)\to 0 \\ x\in B}} \frac{1}{m(B)} \int_B |f(y) - r| dy = |f(x) - r| \quad \text{whenever } x \notin E_r.$$

If $E = \bigcup_{r\in\mathbb{Q}} E_r$, then $m(E) = 0$. Now suppose that $\bar{x} \notin E$ and $f(\bar{x})$ is finite. Given $\epsilon > 0$, there existsa a rational r such that $|f(\bar{x}) - r| < \epsilon$. Since

$$\frac{1}{m(B)} \int_B |f(y) - f(\bar{x})| dy \leq \frac{1}{m(B)} \int_B |f(y) - r| dy + |f(\bar{x}) - r|,$$

we must have

$$\limsup_{\substack{m(B)\to 0 \\ \bar{x}\in B}} \frac{1}{m(B)} \int_B |f(y) - f(\bar{x})| dy \leq 2\epsilon,$$

and thus $\bar{x}$ is in the Lebesgue set of f. $\square$

Definition 4.11 (Bounded Eccentricity) *A collection of sets $\{U_\alpha\}$ is said to be **shrink regularly** to $\bar{x}$ or has **bounded eccentricity** at $\bar{x}$ if there is a constant $c > 0$ such that for each U_α containing $\bar{x}$ there is a ball B with*

$$\bar{x} \in B, \quad U_\alpha \subset B, \quad \textit{and} \quad m(U_\alpha) \geq c\, m(B).$$

Thus U_α is contained in B, but its measure is comparable to the measure of B.

Example: the set of all open cubes containing $\bar{x}$ shrink regularly to $\bar{x}$. However in $\mathbb{R}^d$ with $d \geq 2$, the collection of all open rectangles containing $\bar{x}$ does not shrink regularly to $\bar{x}$.

Corollary 4.11.1 *Suppose f is locally integrable on $\mathbb{R}^d$. If $\{U_\alpha\}$ shrinks regular to $\bar{x}$, then*

$$\lim_{\substack{m(U_\alpha)\to 0\\ x\in U_\alpha}} \frac{1}{m(U_\alpha)} \int_{U_\alpha} f(y)dy = f(\bar{x})$$

for every point $\bar{x}$ in the Lebesgue set of f.

Proof: Note that if $\bar{x} \in B$ with $U_\alpha \subset B$ and $m(U_\alpha) \geq cm(B)$, then

$$\frac{1}{m(U_\alpha)} \int_{U_\alpha} |f(y) - f(\bar{x})| dy \leq \frac{1}{cm(B)} \int_B |f(y) - f(\bar{x})| dy.$$

And the limit for the right hand side as $m(B) \to 0$ is $f(\bar{x})$. $\square$

4.3 Differentiability of functions

4.3.1 Bounded Variation

Definition 4.12 (Rectifiable Curves) *Let γ be a parametrized curve in the plane given by $z(t) = (x(t), y(t))$, where $a \leq t \leq b$. Here $x(t)$ and $y(t)$ are continuous real-valued functions on $[a, b]$. The curve γ is* ***rectifiable*** *if there exists $M < \infty$ such that, for any partition $a = t_0 < t_1 < \cdots < t_N = b$ of $[a, b]$,*

$$\sum_{j=1}^{N} |z(t_j) - z(t_{j-1})| \leq M.$$

In this case, the ***length*** *$L(\gamma)$ of the curve is defined to the supremum over all partitions of the sum on the left-hand side. Alternatively, $L(\gamma)$ is the infimum of all M that forms an upper bound of the sum.*

Definition 4.13 (Bounded Variation) *Suppose $F(t)$ is a complex-valued function defined on $[a, b]$, and $a = t_0 < t_1 < \cdots < t_N = b$ is a partition of this interval. The* ***variation*** *of F on this partition is defined by*

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|.$$

The function F is said to be of ***bounded variation*** *if the variation of F over all partitions are bounded, that is, there exists $M < \infty$ so that*

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| \leq M$$

for all partitions $a = t_0 < t_1 < \cdots < t_N = b$.

Remark 4.13.1 *We observe that if a partition $\tilde{P}$ given by $a = \tilde{t_0} < \tilde{t_1} < \cdots < \tilde{t}_M = b$ is a refinement of a partition P given by $a = t_0 < t_1 < \cdots < t_N = b$, then the variation of F on $\tilde{P}$ is greater than or equal to the variation of F on P.*

Theorem 4.14 *A curve parametrized by $(x(t), y(t))$, $a \le t \le b$, is rectifiable if and only if both $x(t)$ and $y(t)$ are of bounded variation.*

Proof: Since if we let $F(t) = x(t) + iy(t)$, then

$$F(t_j) - F(t_{j-1}) = (x(t_j) - x(t_{j-1})) + i(y(t_j) - y(t_{j-1})),$$

and if a and b are real, then $|a + ib| \le |a| + |b| \le 2|a + ib|$. $\square$

Definition 4.15 (Total Variation) *The **total variation** of a function F (can be complex) on $[a, x]$ (where $a \le x \le b$) is defined by*

$$T_F(a, x) = \sup \sum_{j=1}^{N} |F(t_j) - F(t_{j-1})|,$$

*where the sup is over all partitions of $[a, x]$. The **positive variation** of a real valued function F on $[a, x]$ is*

$$P_F(a, x) = \sup_{(+)} F(t_j) - F(t_{j-1}),$$

*where the sum is over all j such that $F(t_j) \ge F(t_{j-1})$, and the supremum is over all partitions of $[a, x]$. The **negative variation** of a real valued function F on $[a, x]$ is defined by*

$$N_F(a, x) = \sup_{(-)} -[F(t_j) - F(t_{j-1})],$$

where the sum is over all j such that $F(t_j) \le F(t_{j-1})$, and the supremum is over all partitions of $[a, x]$.

Lemma 4.16 *Suppose F is real-valued and of bounded variation on $[a, b]$. Then for all $a \le x \le b$ one has*

$$F(x) - F(a) = P_F(a, x) - N_F(a, x).$$

and

$$T_F(a, x) = P_F(a, x) + N_F(a, x).$$

Proof: Given $\epsilon > 0$ there exists a partition $a = t_0 < \cdots < t_N = x$ of $[a, x]$, such that

$$\left| P_F - \sum_{(+)} (F(t_j) - F(t_{j-1})) \right| < \epsilon \quad \text{and} \quad \left| N_F - \sum_{(-)} (-(F(t_j) - F(t_{j-1}))) \right| < \epsilon.$$

Since by direct computation,

$$F(x) - F(a) = \sum_{(+)} (F(t_j) - F(t_{j-1})) - \sum_{(-)} (-(F(t_j) - F(t_{j-1}))),$$

we find that $|F(x) - F(a) - [P_F - N_F]| < 2\epsilon$, which proves the first identity.

For the second identity, we also note that for any partition of $a = t_0 < \cdots < t_N = x$ of $[a, x]$ we have

$$\sum_{j=1}^{N} |F(t_j) - F(t_{j-1})| = \sum_{(+)} (F(t_j) - F(t_{j-1})) + \sum_{(-)} (-(F(t_j) - F(t_{j-1}))),$$

hence $T_F \le P_F + N_F$. Also, the above implies

$$\sum_{(+)} (F(t_j) - F(t_{j-1})) + \sum_{(-)} (-(F(t_j) - F(t_{j-1}))) \le T_F.$$

Once again, one can argue using common refinements of partitions in the definitions of P_F and N_F to deduce the inequality $P_F + N_F \le T_F$, and the lemma is proved. □

Theorem 4.17 (Jordan's Decomposition Theorem) *A real-valued function F on $[a, b]$ is of bounded variation if and only if F is the difference of two increasing bounded functions.*

Proof: Clearly, if $F = F_1 - F_2$, where each F_j is bounded and increasing, then F is of bounded variation (Simple argument using triangle inequality).
Conversely, suppose F is of bounded variation. Then we let $F_1(x) = P_F(a, x) + F(a)$ and $F_2(x) = N_F(a, x)$. Clearly, both F_1 and F_2 are increasing, of bounded variation and by the previous Lemma, $F(x) = F_1(x) - F_2(x)$. □

Remark 4.17.1 *A complex-valued function, hence, is of bounded variation if it is the complex linear combination of four increasing functions.*

For a curve γ parametrized by a continuous function $z(t) = x(t) + iy(t)$. Assume the curve is rectifiable, we define $L(A, B)$ as the length of the segment of γ that arises as the image of those t for which $A \le t \le B$, with $a \le A \le B \le b$. Note that $L(A, B) = T_F(A, B)$, where $F(t) = z(t)$. We see that

$$L(A, C) + L(C, B) = L(A, B) \text{ if } A \le C \le B.$$

Moreover, we observe that $L(A, B)$ is a continuous function (as z is continuous). Then we also observe that if a function of bounded variation is continuous, then so is its total variation.

4.3.2 Derivative For Continuous Functions

Lemma 4.18 (Rising Sun Lemma) *Suppose G is real-valued and continuous on $\mathbb{R}$. Let E be the set of points x such that*

$$G(x + h) > G(x), \quad \text{for some } h = h_x > 0.$$

If E is non-empty, then it must be open, and hence can be written as a countable disjoint union of open intervals $E = \bigcup (a_k, b_k)$. If (a_k, b_k) is a finite interval in this union, then

$$G(b_k) - G(a_k) = 0.$$

Proof: Since G is continuous, it is clear that E is open whenever it is non-empty and can therefore be written as a disjoint union of countably many open intervals. If (a_k, b_k) denotes a finite interval in this decomposition, then $a_k \notin E$; therefore we cannot have $G(b_k) > G(a_k)$. We now suppose that $G(b_k) < G(a_k)$. By continuity, there exists $a_k < c < b_k$, so that

$$G(c) = \frac{G(a_k) + G(b_k)}{2},$$

and in fact we may choose c furthest to the right in the interval (a_k, b_k). Since $c \in E$, there exists $d > c$ such that $G(d) > G(c)$. Since $b_k \notin E$, we must have $G(x) \leq G(b_k)$ for all $x \geq b_k$; therefore $d < b_k$. Since $G(d) > G(c)$, there exists $c' > d$ with $c' < b_k$ and $G(c') = G(c)$, which contradicts the fact that c was chosen furthest to the right in (a_k, b_k). This shows that we must have $G(a_k) = G(b_k)$. □

Corollary 4.18.1 *Suppose G is real-valued and continuous on a closed interval $[a, b]$. If E denotes the set of points x in (a, b) so that $G(x + h) > G(x)$ for some $h > 0$, then E is either empty or open. In the latter case, it is a disjoint union of countably many intervals (a_k, b_k), and $G(a_k) = G(b_k)$, except possibly when $a = a_k$, in which case we only have $G(a_k) \leq G(b_k)$.*

Theorem 4.19 *If F is of bounded variation on $[a, b]$, then F is differentiable almost everywhere. I.e., the quotient*

$$\lim_{h\to 0} \frac{F(x + h) - F(x)}{h}$$

exists for almost every $x \in [a, b]$.

Proof: It suffices to consider the case when F is bounded increasing, as if any F with bounded variation can be decomposed into the difference of two bounded increasing functions. We will first prove the case where F is assumed to be continuous, then in the later section we will show this is also true for general functions.

We first define

$$\Delta_h(F)(x) = \frac{F(x + h) - F(x)}{h}.$$

We also consider the four **Dini Number or Dini Derivative** at x defined by

$$D^+(F)(x) = \limsup_{\substack{h\to 0\\ h>0}} \Delta h(F)(x)$$

$$D_+(F)(x) = \liminf_{\substack{h\to 0\\ h>0}} \Delta h(F)(x)$$

$$D^-(F)(x) = \limsup_{\substack{h\to 0\\ h<0}} \Delta h(F)(x)$$

$$D_-(F)(x) = \liminf_{\substack{h\to 0\\ h<0}} \Delta h(F)(x)$$

Then $D_+ \leq D^+$ and $D_- \leq D^-$. To prove the theorem, it suffices to show that

(i) $D^+(F) < \infty$ for a.e. x and

(ii) $D^+(F)(x) \leq D_-(F)(x)$ for a.e. x.

If these results hold, then by applying (ii) to $-F(-x)$ instead of $F(x)$, we obtain $D^{-}(F)(x) \leq D_{+}(F)(x)$ for a.e. x. Therefore

$$D^{+} \leq D_{-} \leq D^{-} \leq D_{+} \leq D^{+} < \infty$$

for a.e. x. Thus all four Dini numbers are finite and equal almost everywhere, hence $F'(x)$ exists for almost every point x.

Since F is increasing, bounded and continuous on $[a, b]$. For a fixed $\gamma > 0$, let

$$E_{\gamma} = \{x : D^{+}(F)(x) > \gamma\}.$$

Since F is continuous, then each E_{γ} is the limit of measurable sets, hence measurable. Next, we apply Corollary 4.18.1 to the function $G(x) = F(x) - \gamma x$, and note that we have $E_{\gamma} \subset \bigcup_k (a_k, b_k)$, where $F(b_k) - F(a_k) \geq \gamma(b_k - a_k)$. Consequently,

$$\begin{aligned} m(E_{\gamma}) &\leq \sum_k m((a_k, b_k)) \\ &\leq \frac{1}{\gamma} \sum_k F(b_k) - F(a_k) \\ &\leq \frac{1}{\gamma}(F(b) - F(a)). \end{aligned}$$

Therefore, $m(E_{\gamma}) \to 0$ as γ tends to infinity, and since $\{D^{+}(F)(x) < \infty\} \subseteq E_{\gamma}$ for all γ, this shows that $D^{+}(F)(x) < \infty$ almost everywhere.

Next, let

$$E = \{x \in (a, b) : D^{+}(F)(x) > R \text{ and } r > D_{-}(F)(x)\}.$$

We show that $D^{+}(F)(x) \leq D_{-}(F)(x)$ a.e. by showing that $m(E) = 0$ for arbitrary choice of r and R such that $R > r$.

Suppose towards a contradiction that $m(E) > 0$. Because $\frac{R}{r} > 1$, we can find an open set O such that $E \subset O \subset (a, b)$, yet $m(O) < m(E) \cdot \frac{R}{r}$. Now O can be written as $\bigcup I_n$, where I_n are disjoint open intervals. Fix n and apply Corollary 4.18.1 to the function $G(x) = -F(-x) + rx$ on the interval $-I_n$. Reflecting through the origin again yields an open set $\bigcup_k (a_k, b_k)$ contained in I_n, where the intervals (a_k, b_k) are disjoint, with

$$F(b_k) - F(a_k) \leq r(b_k - a_k).$$

However, on each interval (a_k, b_k), we apply Corollary 4.18.1 to the function $G(x) = F(x) - Rx$, which obtains an open set $V_n = \bigcup_{k,j} (a_{k,j},\, b_{k,j})$ of disjoint open intervals $(a_{k,j}, b_{k,j})$ with $(a_{k,j}, b_{k,j}) \subset (a_k, b_k)$ for every j, and

$$F(b_{k,j}) - F(a_{k,j}) \geq R(b_{k,j} - a_{k,j}).$$

Then using the fact that F is increasing, we find that

$$m(V_n) = \sum_{k,j} (b_{k,j} - a_{k,j})$$

$$
\begin{aligned}
&\leq \frac{1}{R}\sum_{k,j} F(b_{k,j}) - F(a_{k,j}) \\
&\leq \frac{1}{R}\sum_{k} F(b_k) - F(a_k) \\
&\leq \frac{r}{R}\sum_{k} (b_k - a_k) \\
&\leq \frac{r}{R} m(I_n).
\end{aligned}
$$

Note that $E \cap I_n \subset V_n$, since $D^+(F)(x) > R$ and $r > D_-(F)(x)$ for each $x \in E$, then $V_n \subset I_n$. We now sum in n:

$$m(E) = \sum_n m(E \cap I_n) \leq \sum_n m(V_n) \leq \frac{r}{R}\sum m(I_n) = \frac{r}{R} m(O) < m(E).$$

The strict inequality gives a contradiction so we need to have $m(E) = 0$. So we have proven the theorem for when F is continuous. □

Corollary 4.19.1 *If F is increasing and continuous, then F' exists almost everywhere. Moreover F' is measurable, non-negative, and*

$$\int_a^b F'(x)dx \leq F(b) - F(a).$$

In particular, if F is bounded on $\mathbb{R}$, then F' is integrable on $\mathbb{R}$.

Remark 4.19.1 *For points $x \in [a, b]$ such that $F'(x)$ is not defined, we can set it to be 0, which does not affect the value of the integral.*

Remark 4.19.2 *We can only establish*

$$\int_a^b F'(x)dx \leq F(b) - F(a).$$

As strict inequality can indeed happen, as we will see in the example after the corollary.

Proof: For $n \geq 1$, we consider the quotient

$$G_n(x) = \frac{F(x + 1/n) - F(x)}{1/n},$$

By Theorem 4.19, we have $G_n(x) \to F'(x)$ for a.e. x, which shows in particular that F' is measurable and non-negative.

We now extend F as a continuous function on all of $\mathbb{R}$, by Fatou's lemma, we know that

$$\int_a^b F'(x)dx \leq \liminf_{n\to\infty} \int_a^b G_n(x)dx.$$

To complete the proof, it suffices to note that

$$\int_a^b G_n(x)dx = \frac{1}{1/n}\int_a^b F(x + 1/n)dx - \frac{1}{1/n}\int_a^b F(x)dx$$

$$= \frac{1}{1/n}\int_{a+1/n}^{b+1/n} F(y)dy - \frac{1}{1/n}\int_a^b F(x)dx$$
$$= \frac{1}{n}\int_{b}^{b+1/n} F(x)dx - \frac{1}{1/n}\int_a^{a+1/n} F(x)dx.$$

Since F is continuous, the first and second terms converge to $F(b)$ and $F(a)$ respectively as $n \to \infty$. To prove the last statement, use Monotone convergence theorem by considering $F'(x)\chi_{[-n;n]}$. □

Example:(The Cantor-Lebesgue Function)

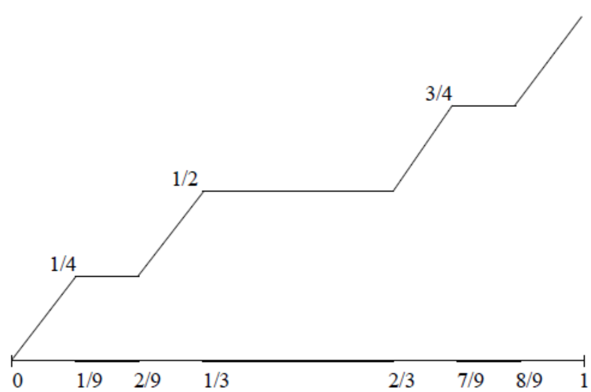
We now construct an increasing function $F : [0,1] \to [0,1]$ with $F(0) = 0$ and $F(1) = 1$, but $F'(x) = 0$ almost everywhere. Hence F is of bounded variation, but

$$0 = \int_a^b F'(x)dx < F(b) - F(a) = 1.$$

Let C denote the triadic Cantor set, and recall that

$$C = \bigcap_{k=0}^{\infty} C_k,$$

where each C_k is a disjoint union of 2^k closed intervals each of length $\frac{1}{3^k}$. Let $F_1(x)$ be the continuous increasing function on $[0,1]$ that satisfies $F_1(0) = 0$, $F_1(x) = \frac{1}{2}$ if $1/3 \le x \le 2/3$, $F_1(1) = 1$, and F_1 is linear on C_1. Similarly, let F_2 be the continuous and increasing function shown in the image:



This process yields a sequence of continuous increasing functions $\{F_n\}_{n=1}^{\infty}$ such that

$$|F_{n+1}(x) - F_n(x)| \le 2^{-n-1}.$$

Hence $\{F_n\}$ converges uniformly to a continuous limit F called the **Cantor-Lebesgue function**. By construction, F is increasing, $F(0) = 0$, $F(1) = 1$, and we see that F is constant on each interval of the complement of the Cantor set. Since $m(C) = 0$, we find that $F'(x) = 0$ almost everywhere.

Proposition 4.20 *Suppose a complex valued f is of bounded variation on $[a,b]$, then*

$$\int_a^b |f'(x)| d\lambda \leq T_f(a,b).$$

Proof: Let $g(x) = T_f(a,x)$. Then $g(a) = 0$ and $g(b) = T_f(a,b)$. g is increasing and bounded, so $g'(x)$ exists a.e. We have

$$\int_a^b g'(x)dx \leq g(b) - g(a) = T_f(a,b).$$

Since f is of bounded variation, f' exists a.e. Since

$$|f'| = \lim_{n\to\infty} n|f(x+\frac{1}{n}) - f(x)| \leq \lim_{n\to\infty} nT_f(x, x+\frac{1}{n}).$$

One also have

$$\lim_{n\to\infty} nT_f(x, x+\frac{1}{n})(f) = \lim_{n\to\infty} \frac{g(x+\frac{1}{n}) - g(x)}{\frac{1}{n}} = g'(x).$$

So,

$$\int_a^b |f'(x)| \leq \int_a^b g'(x) = T_f(a,b).$$

□

Definition 4.21 (Absolutely Continuous) *A function F defined on $[a,b]$ is* ***absolutely continuous*** *if for any $\epsilon > 0$, there exists $\delta > 0$ so that*

$$\sum_{k=1}^N |F(b_k) - F(a_k)| < \epsilon \quad \textit{whenver} \quad \sum_{k=1}^N (b_k - a_k) < \delta,$$

and the intervals (a_k, b_k), $k = 1, \cdots, N$ are disjoint.

Remark 4.21.1 *From the definition, it is clear that absolutely continuous functions are uniformly continuous, and if F is absolutely continuous on a bounded interval, then it is also bounded variation on the same interval. Moreover, its total variation is absolutely continuous.*

Proposition 4.22 *Consider the following function on defined on $[0,1]$:*

$$f(x) = \begin{cases} x^\alpha \sin\frac{1}{x^\beta}, & x \neq 0, \\ 0 & x = 0 \end{cases}$$

Then if $\alpha \leq \beta$, f is uniformly continuous on $[0,1]$ but not absolutely continuous (in fact not even of bounded variation). If $\alpha > \beta$, then f is absolutely continuous on $[0,1]$.

Proof: Apply Mean Value Theorem. □

Proposition 4.23 *Suppose $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous. Then f maps sets of zero measure to sets of zero measure. Moreover, f will map measurable sets to measurable sets.*

Proof: For the first half of the statement, approximate the null set using outer regularity and break the approximating open set into disjoint intervals. Then apply the absolutely continuous condition. The second statement follows from the inner regularity and the first statement. $\square$

Definition 4.24 (Vitali Covering) *A collection $\mathcal{B}$ of balls $\{B\}$ is said to be a* ***Vitali covering*** *of a set E if for every $x \in E$ and any $\eta > 0$, there is a ball $B \in \mathcal{B}$, such that $x \in B$ and $m(B) < \eta$. Thus every point is covered by a ball of arbitrarily small measure.*

Theorem 4.25 (Vitali Covering Theorem) *Suppose $E \subset \mathbb{R}^d$ is a set of finite measure and $\mathcal{B}$ is a Vitali covering of E. For any $\delta > 0$ we can find finitely many balls $B_1, \cdots, B_n$ in $\mathcal{B}$ that are disjoint and so that*

$$\sum_{i=1}^{N} m(B_i) \geq m(E) - \delta.$$

In particular, we can arrange the choice of the balls so that

$$m\left(E - \bigcup_{i=1}^{N} B_i\right) < 2\delta.$$

Theorem 4.26 *If F is absolutely continuous on $[a, b]$, then $F'(x)$ exists almost everywhere. Moreover, if $F'(x) = 0$ for a.e. x, then F is constant.*

Proof: If F is absolutely continuous, then it is continuous and of bounded variation, hence $F'(x)$ exists almost everywhere. So we only need to show the second statement.

We just need to show that $F(a) = F(b)$, since we can replace the interval $[a, b]$ by any sub-interval. Now let E be the set of those $x \in (a, b)$ where $F'(x)$ exists and is zero. By our assumption $m(E) = b - a$. Next, momentarily fix $\epsilon > 0$. Since for each $x \in E$, we have

$$\lim_{h \to 0} \left| \frac{F(x+h) - F(x)}{h} \right| = 0$$

then for each $\eta > 0$, we have an open interval $I = (a_x, b_x) \subset [a, b]$ containing x, with

$$|F(b_x) - F(a_x)| \leq \epsilon(b_x - a_x), \quad \text{and} \quad b_x - a_x < \eta.$$

The collection of these intervals forms a Vitali covering of E, and hence by the lemma, for $\delta > 0$, we can select finitely many I_i, $1 \leq i \leq N$, $I_i = (a_i, b_i)$, which are disjoint and such that

$$\sum_{i=1}^{N} m(I_i) \geq m(E) - \delta = (b - a) - \delta.$$

However, $|F(b_i) - F(a_i)| \leq \epsilon(b_i - a_i)$, and upon adding these inequality, we get

$$\sum_{i=1}^{N} |F(b_i) - F(a_i)| \leq \epsilon(b - a),$$

since the intervals I_i are disjoint and lie in $[a,b]$. Next consider the complement of $\bigcup_{j=1}^{N} I_j$ in $[a,b]$. It consists of finitely many closed intervals $\bigcup_{k=1}^{M}[\alpha_k,\beta_k]$ with total length $\leq \delta$. Thus by the absolutely continuity of F, $\sum_{k=1}^{M}|F(\beta_k)-F(\alpha_k)|\leq \epsilon$. Hence

$$|F(b)-F(a)| \leq \sum_{i=1}^{N}|F(b_i)-F(a_i)| + \sum_{k=1}^{M}|F(\beta_k)-F(\alpha_k)| \leq \epsilon(b-a)+\epsilon.$$

Since $\epsilon > 0$ is arbitrary, we conclude that $F(b)-F(a)=0$. $\square$

Theorem 4.27 (Fundamental Theorem of Calculus) *Suppose F is absolutely continuous on $[a,b]$. Then F' exists almost everywhere and is integrable. Moreover*

$$F(x)-F(a) = \int_a^x F'(y)dy, \quad \text{for all } a \leq x \leq b.$$

Conversely, if f is integrable on $[a,b]$, then there exists an absolutely continuous function F such that $F'(x)=f(x)$ almost everywhere, and in fact, we make take $F(x)=\int_a^x f(y)dy$.

Proof: Since we know that a real-valued absolutely continuous function is the difference of two continuous increasing functions, then F' is integrable on $[a,b]$. Now let $G(x)=\int_a^x F'(y)dy$, G is then absolutely continuous; hence so is the difference $G(x)-F(x)$. By the Lebesgue Differentiation Theorem, we know that $G'(x)=F'(x)$ for a.e. x. Hence the difference $F-G$ has derivative 0 almost everywhere, which concludes that $F-G$ is constant. Evaluating this expression at $x=a$ gives the desired result.

The converse is a consequence of the observation that $\int_a^x f(y)dy$ is absolutely continuous, and the Lebesgue differentiation theorem gives $F'(x)=f(x)$ a.e. $\square$

Corollary 4.27.1 *The following are equivalent:*

- *$f:\mathbb{R}\to\mathbb{R}$ is Lipschitz with $|f(x)-f(y)|\leq M|x-y|$ for some constant $M>0$.*
- *f is absolutely continuous and $|f'(x)|\leq M$ a.e.*

Proof: Clear. $\square$

4.3.3 Differentiability of Jump Functions

We first recall the following lemma.

Lemma 4.28 *A bounded monotone function F on $[a,b]$ has at most countably many discontinuities, and all such the discontinuities are jump discontinuities.*

Now let F be a bounded increasing function, $\{x_n\}_{n=1}^{\infty}$ denote the points where F is discontinuous, and let α_n denote the jump of F at x_n, that is $\alpha_n = F(x_n^+) - F(x_n^-)$, and

$$F(x_n) = F(x_n^-) + \theta_n \alpha_n$$

for some θ_n, with $0 \leq \theta_n \leq 1$. If we let

$$j_n(x) = \begin{cases} 0 & \text{if } x < x_n, \\ \theta_n & \text{if } x = x_n, \\ 1 & \text{if } x > x_n \end{cases} .$$

then we define the **jump function** associated to F by

$$J_F(x) = \sum_{n=1}^{\infty} \alpha_n j_n(x).$$

Immediately, we note that if F is bounded, then

$$\sum_{n=1}^{\infty} \alpha_n \leq F(b) - F(a) < \infty$$

and hence the series defining J converges absolutely and uniformly.

The following lemma is easy to prove.

Lemma 4.29 *If F is increasing and bounded on $[a, b]$, then:*

1. *$J(x)$ is discontinuous precisely at the points $\{x_n\}$ and has a jump at x_n equal to that of F.*
2. *The difference $F(x) - J(x)$ is increasing and continuous.*

Theorem 4.30 *If J is the jump function associated to an increasing and bounded function F on $[a, b]$, then $J'(x)$ exists and vanishes almost everywhere.*

Remark 4.30.1 *This proves the remaining case of Theorem 4.19.*

Proof: Given any $\epsilon > 0$, we note that the set E of those x where

$$\limsup_{h \to 0} \frac{j(x+h) - j(x)}{h} > \epsilon \tag{4.5}$$

is a measurable set. Suppose $\delta = m(E)$. We need to show that $\delta = 0$. Now observe that since the series $\sum a_n$ arising in the definition of J converges, then for any η, we can find an N so large that $\sum_{n>N} \alpha_n < \eta$. We then write

$$J_0(x) = \sum_{n>N} \alpha_n j_n(x),$$

and because of our choice of N we have

$$J_0(b) - J_0(a) < \eta.$$

However, $J - J_0$ is a finite sum of terms $a_n j_n(x)$, and therefore, the set of points where Equation (4.5) holds, with J replaced by J_0, differs from E by at most a finite set: the points $\{x_1, x_2, \cdots, x_N\}$. Thus we can find a compact set K, with $m(K) \geq \frac{\delta}{2}$, so that

$$\limsup_{h \to 0} \frac{J_0(x+h) - J_0(x)}{h} > \epsilon$$

for each $x \in K$. Hence there are intervals (a_x, b_x) containing x, $x \in K$, so that $J_0(b_x) - J_0(a_x) > \epsilon(b_x - a_x)$. We can first choose a finite collection of these intervals that covers K, and then apply the Vitali Covering Lemma to select intervals I_1, I_2, $\cdots$, I_n which are disjoint, and for which $\sum_{j=1}^{n} m(I_j) \geq m(K)/3$. The intervals $I_j = (a_j, b_j)$ of course satisfy

$$J_0(b_j) - J_0(a_j) > \epsilon(b_j - a_j).$$

Now,

$$J_0(b) - J_0(a) \geq \sum_{j=1}^{N} J_0(b_j) - J_0(a_j) > \epsilon \sum (b_j - a_j) \geq \frac{\epsilon}{3} m(K) \geq \frac{\epsilon}{6} \delta.$$

Thus, $\epsilon\delta/6 < \eta$, and as η is arbitrary, it follows that $\delta = 0$. □

Theorem 4.31 (Lebesgue Decomposition) *Let F be an increasing function on $[a, b]$. Then we can write $F = F_A + F_C + F_J$, where each of the functions F_A, F_C and F_J is increasing and*

- *F_A is absolutely continuous.*
- *F_C is continuous, but $F_C'(x) = 0$ for a.e. x.*
- *F_J is a jump function.*

Moreover, each component F_A, F_C, F_J is uniquely determined up to additive constant.

Remark 4.31.1 *There is also a Lebesgue decomposition Theorem for measures.*

Proof: The jump function F_j is defined as in the beginning of the section. Then $F - F_j$ is is continuous and F_j is unique up to constants. Next, let $G = F - F_j$, then G is continuous and increasing, so G' exists a.e. Define $F_A(x) = G(a) + \int_a^x G'(t)dt$. Then F_A is absolutely continuous and increasing. Define $F_C = F - F_A - F_J$. Then $F_C'(x) = 0$ a.e. and F_C is continuous. We show F_C is increasing. Since $\int_x^y G'(t) \leq G(y) - G(x)$. Then if $a \leq x < y \leq b$, we have $F_C(y) - F_C(x) \geq G(y) - G(x) \geq 0$, as desired. Lastly, from the unique up to constant of G, we get the unique up to constant of F_A and hence F_C. □

4.3.4 Change of Variable Formula and Integration By Parts

Let F be an absolutely continuous function on $[a, b]$, and $A = F(a)$, $B = F(b)$. Then there exists an F that is also strictly increasing, but $F'(x) = 0$ on a set of positive measure. In fact we can construct this function as follows: Let C be a generalized cantor set obtained in the the same manner as the Cantor set except that each of the interval removed at n^{th} iteration has length $\frac{\alpha}{3^n}$, $0 < \alpha < 1$. In this way, we have that the $m(C) = 1 - \alpha$. Let K be $[0, 1] \setminus C$ and define $F : [0, 1] \to \mathbb{R}$ by $F(x) = \int_a^x \chi_K(x)dx$. Since F is given by an integral, it is absolutely continuous. Furthermore, if $x_1, x_2 \in [0, 1]$, $x_1 < x_2$, then we have $F(x_2) - F(x_1) = \int_{x_1}^{x_2} \chi_K(x)dx > 0$ by the definition of K

(the generalized cantor set contains no intervals). Hence, it follows that $F : [0, 1] \to [0, \alpha]$ is a strictly increasing bijection. Lastly, by the Fundamental Theorem of Calculus (4.27), we know that $F'(x) = \chi_K(x)$ a.e., so $F'(x) = 0$ on a set of positive measure.

Next, we will show that there exists B, such that $m(B) = 0$ and $F^{-1}(B)$ is not measurable. In particular, this also shows that if F is an absolutely continuous and increasing on $[a, b]$ and f is any measurable function on $[A, B]$, then $f(F(x))$ does not need to be measurable (Let f be χ_B).

Lemma 4.32 *For the function F defined above and any Borel subset B of $[0, 1]$, $m(F(B)) = \int_B \chi_K(x)dx$.*

Proof: Let $P = \{[0, x] \,|\, x \in [0, 1]\}$. Let $\mathcal{L}$ be the collection of all Borel subsets B of $[0, 1]$ such that $m(f(B)) = \int_B \chi_K(x)dx$. We will verify that $P \subseteq \mathcal{L}$, P is a π-class and $\mathcal{L}$ is a λ-system (with respect to $[0, 1]$).

Clearly, P is non-empty and for any $B_1, B_2 \in P$, $B_1 \cap B_2 \in P$ so P is a π-class.

Let $G : [0, \alpha] \to [0, 1]$ be the inverse of F. Then G is strictly increasing and hence Borel. If $B \subseteq [0, 1]$ is Borel, then $F(B) = g^{-1}(B)$ is a Borel subset of $[0, \alpha]$, hence $m(F(B))$ is well-defined. Obviously $B \mapsto m(F(B))$ is σ-additive, so it is a finite measure. Denote the finite measure $B \mapsto m(F(B))$ and $B \mapsto \int_B \chi_K$ by μ and ν respectively. Then $\mu([0, 1]) = \alpha = \nu([0, 1])$. So $[0, 1] \in \mathcal{L}$. It is easy to show that if $B \in \mathcal{L}$, then $[0, 1] \setminus B \in \mathcal{L}$ and $\mathcal{L}$ is closed under countable disjoint unions. Therefore $\mathcal{L}$ is a λ-system. Finally, one can easily check that $P \subseteq \mathcal{L}$, so $\sigma(P) \subseteq \mathcal{L}$, which implies $\mathcal{L} = \mathcal{B}([0, 1])$ since $\sigma(P) = \mathcal{B}([0, 1])$. So the claim we want to prove holds. $\square$

Now since $m(C) > 0$ and $F' = 0$ on C a.e., then exists Lebesgue measurable set B_1 with $m(B_1) > 0$ and $F' = 0$ on B_1. Then by the completeness of Lebesgue Measure, we can choose a Borel set B_2, such that $m(B_2 \Delta B_1) = 0$. Then $m(B_2) = m(B_1) > 0$ and $F' = 0$ a.e. on B_2. Now every Borel set that has a positive measure contains a non-Lebesgue measurable set. Choose a non-measurable set $B \subseteq B_2$, we have $F(B) \subseteq F(B_2)$, and since $m(F(B_2)) = \int_{B_2} \chi_K = 0$, then $f(B)$ is Lebesgue measurable with $m(F(D)) = 0$. Finally as F is a bijection, we have $F^{-1}(F(D)) = D$ is not measurable.

However, the following fact is true. If F is absolutely continuous and monotone increasing on $[a, b]$. If E is a measurable subset of $[A, B]$, the set $F^{-1}(E) \cap \{F'(X) > 0\}$ is measurable by first showing that $m(O) = \int_{F^{-1}(O)} F'(x)dx$ for any open set O, then approximate any measurable set with a G_δ set. Using this fact, we may prove that if F is absolutely continuous and monotone increasing on $[a, b]$, with $F(a) = A$ and $F(b) = B$. Suppose f is any measurable function on $[A, B]$. then $f(F(x))F'(x)$ is measurable on $[a, b]$. Since F is monotone increasing, then $F'(x) \geq 0$. Let $I_a = \{x \,|\, f(F(x))F'(x) > a\}$. When $a > 0$, we have

$$I_a = \bigcup_{q \in \mathbb{Q}^+} \left[\left\{ x \,|\, f(F(x)) > \frac{a}{q} \right\} \cap \{x \,|\, F'(x) > 0\} \right] \cap \{x \,|\, F'(x) > q\}.$$

The terms in the brackets are measurable by the fact. We know that F' is a measurable function. Similarly, we can show for the case $a \leq 0$. Hence it follows that $f(F(x))F'(x)$ is measurable.

Theorem 4.33 (Change of Variable Formula) *Suppose F is absolutely continuous and monotone increasing* *on $[a,b]$, with $F(a) = A$ and $F(b) = B$. If f is integrable on $[A,B]$, then*

$$\int_A^B f(y)dy = \int_a^b f(F(x))F'(x)dx.$$

Proof: Let $G(x) = \int_A^x f(y)dy$ and $H(x) = \int_a^x f(F(x))F'(x)dx$. Then by the Fundamental Theorem of Calculus, we have $G'(x) = f(x)$, $H'(x) = f(F(x))F'(x)$ a.e.

Next since F is differentiable on $[a,b]$ a.e. and G is differentiable on $[A,B]$ a.e., then $G \circ F$ is differentiable on $[a,b]$ a.e. By chain rule we have $(G \circ F)'(x) = f(F(x))F'(x)$ a.e. So $G'(F(x)) - H'(x) = 0$ a.e., so $G(F(x)) - H(x)$ is a constant. Substitute $x = a$, we get $G(F(x)) - H(x) \equiv 0$, i.e., the change of variable formula holds. □

Proposition 4.34 (Integration By Parts) *Suppose F and G are absolutely continuous on $[a,b]$, then*

$$\int_a^b F'(x)G(x)dx = -\int_a^b F(x)G'(x)dx + [F(x)G(x)]_a^b.$$

Proof: Since F and G are absolutely continuous, they are bounded, so clearly $F \cdot G$ is absolutely continuous. Then by the product rule and the Fundamental Theorem of calculus, the formula holds. □

4.4 More On Rectifiable Curves

4.4.1 Length of Rectifiable Curves

Proposition 4.35 *Suppose F is a complex-valued and absolutely continuous on $[a,b]$. Then*

$$T_F(a,b) = \int_a^b |F'(t)|dt.$$

Remark 4.35.1 *In fact we have the converse of the statement. If F is of bounded variation and $T_f(a,b) = \int_a^b |F'(t)|dt$, then F is absolutely continuous on $[a,b]$.*

Proof: For any partition $a = t_0 < t_1 < \cdots < t_N = b$, we have

$$\begin{aligned}\sum_{j=1}^N |F(t_j) - F(t_{j-1})| &= \sum_{j=1}^N \left| \int_{t_{j-1}}^{t_j} F'(t)dt \right| \\ &\le \sum_{j=1}^N \int_{t_{j-1}}^{t_j} |F'(t)|dt \\ &= \int_a^b |F'(t)|dt.\end{aligned}$$

So this proves

$$T_F(a,b) \le \int_a^b |F'(t)|dt. \tag{4.6}$$

The reverse inequality follows from Proposition 4.20. □

Theorem 4.36 *Suppose $(x(t), y(t)$ is a curve defined for $a \le t \le b$. If both $x(t)$ and $y(t)$ are absolutely continuous, then the curve is rectifiable, and if L denotes its length, we have*

$$L = \int_a^b (x'(t)^2 + y'(t)^2)^{1/2} dt.$$

Proof: Note that if $F(t) = x(t) + iy(t)$ is absolutely continuous then it is automatically of bounded variation, and hence the curve is rectifiable. Then the identity follows from the Proposition 4.35. $\square$

Note that any curve have infinitely many different parametrizations. A rectifiable curve, however, has a unique natural parametrization: the **arc length parametrization**. Let $L(A, B)$ denote the length function, and for the variable $t \in [a, b]$, set $s = s(t) = L(a, t)$. Then $s(t)$, the arc-length, is a continuous increasing function which map $[a, b]$ to $[0, L]$, where L is the length of the curve. The arc length parametrization of the curve is now given by the pair $\tilde{z}(s) = \tilde{x}(s) + i\tilde{y}(s)$, where $\tilde{z}(s) = z(t)$, for $s = s(t)$. Notice that in this way, the function $\tilde{z}(s)$ is well-defined on $[0, L]$. Since if $s(t_1) = s(t_2)$, $t_1 < t_2$, then in fact $z(t)$ does not vary in the interval $[t_1, t_2]$ and thus $z(t_1) = z(t_2)$. Moreover, $|\tilde{z}(s_1) - \tilde{z}(s_2)| \le |s_1 - s_2|$, for all pairs $s_1, s_2 \in [0, L]$., since the left-hand side of the inequality is the distance between two points on the curve, while the right-hand side is the length of the portion of the curve joining these two points. Also, as s varies from 0 to L, $\tilde{z}(s)$ traces out the same points (in the same order) that $z(t)$ does as t varies from a to b.

Theorem 4.37 *Suppose $(x(t), y(t))$, $a \le t \le b$, is a rectifiable curve that has length L. Consider the arc-length parametrization $\tilde{z}(s) = (\tilde{x}(s), \tilde{y}(s))$ described above. Then $\tilde{x}$ and $\tilde{y}$ are absolutely continuous, $|\tilde{z}'(s)| = 1$ for almost every $s \in [0, L]$, and*

$$L = \int_0^L (\tilde{x}'(s)^2 + \tilde{y}'(s)^2)^{1/2} ds.$$

Proof: We noted that $|\tilde{z}(s_1) - \tilde{z}(s_2)| \le |s_1 - s_2|$, so it follows immediately that $\tilde{z}(s)$ is absolutely continuous, hence differentiable almost everywhere. Moreover, this inequality also proves that $|\tilde{z}'(s)| \le 1$ for almost every s. By definition the total variation of $\tilde{z}$ equals L, and by the previous theorem, we must have $L = \int_0^L |\tilde{z}'(s)| ds$, which is possible only when $|\tilde{z}'(s)| = 1$ a.e. $\square$

Definition 4.38 *A curve parametrized by $z(t) = (x(t), y(t))$, $a \le t \le b$, is said to be **simple** if the mapping is injective. It is a closed simple curve if the mapping $t \mapsto z(t)$ is injective for t in $[a, b)$ and $z(a) = z(b)$. A curve is **quasi-simple** if the mapping is injective for t in the complement of finitely many points in $[a, b]$.*

Notation, for a curve $z(t)$, we let Γ denote the track of $z(t)$, that is

$$\Gamma := \{z(t) \,:\, a \le t \le b\}$$

And for any set $E \subset \mathbb{R}^d$, we denote the δ neighbourhood of E to be the set

$$E^\delta := \{x \in \mathbb{R}^d \,:\, d(x, E) < \delta\}.$$

Definition 4.39 (Minkowski Content) *We say that the set K has **Minkowski Content** if the limit*

$$\lim_{\delta\to 0} \frac{m(K^\delta)}{2\delta}$$

exists. When this limit exists, we denote it by $\mathcal{M}(K)$.

Remark 4.39.1 *If we denote*

$$\mathcal{M}^*(K) := \limsup_{\delta\to 0} \frac{m(K^\delta)}{2\delta} \quad \text{and} \quad \mathcal{M}_*(K) = \liminf_{\delta\to 0} \frac{m(K^\delta)}{2\delta}.$$

Then K has Minkowski content if and only if $\mathcal{M}^*(K) < \infty$ and $\mathcal{M}_*(K) = \mathcal{M}^*(K)$.

We have an important result regaring to the Minkowski content of a quasi-simple curve, whose proof will be omitted.

Theorem 4.40 *Suppose Γ is the tract of a quasi-simple curve. The Minkowski content of Γ exists if and only if Γ is rectifiable. When this is the case, and L is the length of the curve, then $\mathcal{M}(\Gamma) = L$.*

4.4.2 Isoperimetric Inequality

The isoperimetric inequality in the plane states, in effect, that among all curves of a given length it is the circle that encloses the maximum area. In this section we aim to solve this problem.

We suppose that Ω is a bounded open subset of $\mathbb{R}^2$, and that its boundary $\bar{\Omega} - \Omega$ is a rectifiable curve Γ, with length $\ell(\Gamma)$. We do not require that Γ be a simple closed curve. Then we have the following Isoperimetric Theorem:

Theorem 4.41 (Isoperimetric Theorem) $4\pi m(\Omega) \le \ell(\Gamma)^2$.

Remark 4.41.1 *A similar result holds even without the assumption that the boundary is a (rectifiable) curve. In fact, the proof will show that for any bounded open set Ω whose boundary is Γ, we have*

$$4\pi m(\Omega) \le \mathcal{M}^*(\Gamma)^2.$$

Proof: For each $\delta > 0$, we consider the outer set

$$\Omega_+(\delta) = \{x \in \mathbb{R}^2 \,:\, d(x, \bar{\Omega}) < \delta\},$$

and the inner set

$$\Omega_-(\delta) = \{x \in \mathbb{R}^2 \,:\, d(x, \Omega^c) \ge \delta\}.$$

Thus $\Omega_-(\delta) \subset \Omega \subset \Omega_+(\delta)$.
We notice that

$$\Omega_+(\delta) = \Omega_-(\delta) \cup \Gamma^\delta$$

and that this union is disjoint. Moreover, if $D(\delta)$ is the open ball (disc) of radius δ centered at the origin, $D(\delta) = \{x \in \mathbb{R}^2 : |x| < \delta\}$, then clearly

$$\begin{cases} \Omega_+(\delta) & \supset \Omega + D(\delta) \\ \Omega & \supset \Omega_-(\delta) + D(\delta). \end{cases} \tag{4.7}$$

We now apply the Brunn-Minkowski inequality to the first inclusion, which gives

$$m(\Omega_+(\delta)) \geq (m(\Omega)^1 2 + m(D(\delta))^{1/2})^2.$$

Since $m(D(\delta)) = \pi\delta^2$, and $(A + B)^2 \geq A^2 + 2AB$ whenever A and B are nonnegative, we find that

$$m(\Omega_+(\delta)) \geq m(\Omega) + 2\pi^{1/2}\delta m(\Omega)^{1/2}.$$

Similarly, $m(\Omega) \geq m(\Omega_-(\delta)) + 2\pi^{1/2}\delta m(\Omega_(\delta))^{1/2}$ using the second inclusion, which implies

$$-m(\Omega_-(\delta)) \geq -m(\Omega) + 2\pi^{1/2}\delta m(\Omega_-(\delta))^{1/2}$$

Next, since $m(\Gamma^\delta) = m(\Omega_+(\delta)) - m(\Omega_-(\delta))$, then we have

$$m(\Gamma^\delta) \geq 2\pi^{1/2}\delta(m(\Omega)^{1/2} + m(\Omega_-(\delta))^{1/2}).$$

We now divide both sides by 2δ and take the limsup as $\delta \to 0$, this gives

$$\mathcal{M}^*(\Gamma) \geq \pi^1 2(2m(\Omega)^{1/2}).$$

Since $\Omega_-(\delta) \nearrow \Omega$ as $\delta \to 0$, and $\ell(\Gamma) \geq \mathcal{M}^*(\Gamma)$, so

$$\ell(\Gamma) \geq 2\pi^{1/2}m(\Omega)^{1/2},$$

and the theorem follows by squaring both sides. $\square$

5 L^p spaces

5.1 General Theory

Definition 5.1 (L^p Spaces) *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space.*

- *If $1 \leq p < \infty$, the L^p-**space**, $L^p(\Omega, \mathscr{A}, \mu)$ on $(\Omega, \mathscr{A}, \mu)$ consists of measurable f on Ω such that*

$$\int |f(x)|^p d\mu < \infty$$

*or equivalently, the L^p **norm** is finite:*

$$\|f\|_p := \left(\int |f(x)|^p d\mu\right)^{1/p} < \infty.$$

- *If $p = \infty$, the L^∞-**space**, $L^\infty(\Omega, \mathscr{A}, \mu)$ consists of measurable f such that there is $M > 0$, $f(x) \leq M$ a.e. $x \in \Omega$, equivalently the L^∞ **norm** is finite:*

$$\|f\|_\infty = \operatorname{esssup}|f| = \inf\{M \,:\, \exists N \in \mathscr{A},\, \mu(N) = 0,\, |f(x)| < M \text{ for all } x \notin N\} < \infty.$$

Proposition 5.2 *For all $1 \leq p \leq \infty$, $L^p(\Omega, \mathscr{A}, \mu)$ is a vector space.*

Proof: Note the space of all real valued / complex valued functions form a vector space. Hence suffices to check that $L^p(\Omega, \mathscr{A}, \mu)$ is a vector subspace. But note that for $1 \leq p < \infty$

$$\begin{aligned}
\int |f+g|^p d\mu &\leq \int (|f| + |g|)^p d\mu \\
&\leq \int (2\max\{|f|, |g|\})^p d\mu \\
&= \int 2^p \max\{|f|^p, |g|^p\} d\mu \\
&\leq \int 2^p (|f|^p + |g|^p) d\mu < \infty.
\end{aligned}$$

For the case $p = \infty$.

$$\|f+g\|_\infty = \operatorname{esssup}\|f+g\| \leq \operatorname{esssup}\|f\| + \operatorname{esssup}\|g\| < \infty.$$

It is clear that L^p is closed under scalar multiplication and inversion for $1 \leq p \leq \infty$, hence the proposition follows. $\square$

Definition 5.3 (Norm And Seminorm) *Let V be a vector space, a **norm** $\|\cdot\|$ on V is a map $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ such that*

1. *$\|v\| = 0$ iff $v = 0$.*
2. *$\|u + v\| \leq \|u\| + \|v\|$.*
3. *$\|\alpha v\| = |\alpha|\|v\|$, for any scalar α.*

If condition (1) is relaxed to $\|v\| = 0$ if $v = 0$, then $\|\cdot\|$ is a ***seminorm****.*

A ***semi-inner product*** *$\langle\cdot,\cdot\rangle$ in V is a map from $V \times V \to R$ satisfies*

1. $\langle v, v\rangle \geq 0$
2. $\langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle$
3. $\langle u, v_1 + v_2\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle$
4. $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$
5. $\langle u, v\rangle = \overline{\langle v, u\rangle}$.

Remark 5.3.1 *A semi-inner product will always induce a seminorm by $\|v\| = \sqrt{\langle v, v\rangle}$, an inner product will always induce a norm. Each norm will induce a metric and each semi norm will induce a pseudo metric. But not all norm comes from an inner product but those satisfies the parallelogram law:*

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + 2\|g\|^2).$$

Lemma 5.4 (Weighted AM-GM / Young's Inequality) *Let $p, q \in \mathbb{R}$, s.t., $1 < p, q < \infty$, and $\frac{1}{p} + \frac{1}{q} = 1$. Then for all $A, B \in \mathbb{R}_{\geq 0}$, we have*

$$AB \leq \frac{A^p}{p} + \frac{B^q}{q}.$$

With equality if and only if $a^p = b^q$.

Proof: It is clear that when $A = 0$ or $B = 0$, the inequality holds. So assume $A > 0$ and $B > 0$, let $t = \frac{1}{p}$ and $\frac{1}{q} = 1 - t$. Since the log function is concave, we have

$$\ln(tA^p + (1-t)B^q) \geq t\ln(A^p) + (1-t)\ln(B^q) = \ln(A) + \ln(B) = \ln(AB).$$

Since ln is a strictly increasing function, we have $tA^p + (1-t)B^q \geq AB$, i.e.,

$$AB \leq \frac{A^p}{p} + \frac{B^q}{q}.$$

With equality holding if and only if $A^p = B^q$. $\square$

Lemma 5.5 (Hölder's Inequality) *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. Then for $1 \leq p \leq \infty$, if f and g are measurable (real or complex) on Ω, then*

$$\|fg\|_1 \leq \|f\|_p\|g\|_q.$$

When $p, q \in (1, \infty)$, and $f \in L^p$, $g \in L^q$, the equality holds iff $|f|^p$ and $|g|^q$ are linearly dependent in L^1,i.e., $\exists\alpha, \beta \geq 0$, not both zero such that $\alpha|f|^p = \beta|g|^q$ a.e.

Remark 5.5.1 *The numbers p and q are said to be Hölder conjugates of each other. The special case $p = q = 2$ gives a form of the Cauchy-Schwarz inequality. Notice that the inequality holds even if $\|fg\|_1$ is infinite. Conversely, if f is in L^p and g is in L^q, then fg is in L^1.*

Proof: If $p = 1$, and $q = \infty$, then $|fg| \le \|f\|_\infty |g|$ almost everywhere, so the Hölder's inequality in this case follows from the monotonicity of the Lebesgue integral. Similarly we have the case for $p = \infty$ and $q = 1$.

Hence we assume $p, q \in (1, \infty)$. The case $\|f\|_p = 0$, $\|f\|_p = \infty$ or $\|g\|_q = 0$, $\|g\|_q = \infty$ are trivial. So WLOG, assume $\|f\|_p, \|g\|_q \in (0, \infty)$. Then dividing f and g by $\|f\|_p$ and $\|g\|_q$ respectively, we may assume

$$\|f\|_p = \|g\|_q = 1.$$

By Young's inequality, we have

$$|f(s)g(s)| \le \frac{|f(s)|^p}{p} + \frac{|g(s)|^q}{q}, \quad \forall s \in \Omega.$$

Integrating both sides gives

$$\|fg\|_1 \le \frac{\|f\|_p^p}{p} + \frac{\|g\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1,$$

which proves the claim.

Under the assumption $p \in (1, \infty)$, and $\|f\|_p = \|g\|_q$ are finite (can scale), equality holds if and only if $|f|^p = |g|^q$ almost everywhere. More generally, if $\|f\|_p$ and $\|g\|_q$ are in $(0, \infty)$, then Hölder's inequality becomes an equality if and only if there exists a real number $\alpha, \beta > 0$, namely

$$\alpha = \|g\|_q^q, \quad \beta = \|f\|_p^p,$$

such that

$$\alpha |f|^p = \beta |g|^q$$

almost everywhere. $\square$

Corollary 5.5.1 (General Hölder's Inequality) *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. Assume that $r \in (0, \infty]$ and $p_1, \cdots, p_n \in (0, \infty]$ such that*

$$\sum_{k=1}^{n} \frac{1}{p_k} = \frac{1}{r}$$

where $\frac{1}{\infty}$ is interpreted as 0. Then for all measurable real or complex valued functions $f_1, \cdots, f_n$ defined on Ω,

$$\left\| \prod_{k=1}^{n} f_k \right\|_r \le \prod_{k=1}^{n} \|f_k\|_{p_k}.$$

In particular, if $f_k \in L^{p_k}(\mu)$ for all $k \in \{1, \cdots, n\}$ then $\prod_{k=1}^{n} f_k \in L^r(\mu)$.

Remark 5.5.2 *For $r \in (0, 1)$, $\|\cdot\|_r$ is in general not a norm as it doesn't satisfy the triangle inequality. However, we still adopt the notation.*

Proof: We prove by induction, it is clear that the statement holds for $n = 1$. Now for general n, suppose the statement holds for $n - 1$ and WLOG let $p_1 \le \cdots \le p_n$.

Case 1: if $p_n = \infty$, then

$$\sum_{k=1}^{n-1} \frac{1}{p_k} = \frac{1}{r}.$$

Pulling out the essential supremum of $|f_n|$ and using the induction hypothesis, we get

$$\begin{aligned}\|f_1 \cdots f_n\|_r &\leq \|f_1 \cdots f_{n-1}\|_r \|f_n\|_\infty \\ &\leq \|f_1\|_{p_1} \cdots \|f_{n-1}\|_{p_{n-1}} \|f_n\|_\infty.\end{aligned}$$

Case 2: if $p_n < \infty$, then necessarily $r < \infty$ as well. Then set

$$p := \frac{p_n}{p_n - r}, \quad q := \frac{p_n}{r},$$

note (p, q) are Hölder conjugates in $(1, \infty)$. Then the Hölder inequality gives

$$\||f_1 \cdots f_{n-1}|^r |f_n|^r\|_1 \leq \||f_1 \cdots f_{n-1}|^r\|_p \||f_n|^r\|_q.$$

Raising to the power $\frac{1}{r}$, we get

$$\|f_1 \cdots f_n\|_r \leq \|f_1 \cdots f_{n-1}\|_{pr} \|f_n\|_{qr}.$$

Since $qr = p_n$, and

$$\sum_{k=1}^{n-1} \frac{1}{p_k} = \frac{1}{r} - \frac{1}{p_n} = \frac{p_n - r}{r p_n} = \frac{1}{pr},$$

then the desired statement follows. $\square$

Lemma 5.6 (Jensen's Inequality) *Let $(\Omega, \mathscr{A}, \mu)$ be a probability space. Let $g : \Omega \to \mathbb{R}$ be a μ measurable function and $\varphi : \mathbb{R} \to \mathbb{R}$ is convex. Then*

$$\varphi\left(\int_\Omega g d\mu\right) \leq \int_\Omega \varphi \circ g d\mu.$$

Proof: We define

$$x_0 := \int_\Omega g \, d\mu,$$

because of the existence of subderivatives for convex functions, we may choose a and b such that

$$ax + b \leq \varphi(x),$$

for all real x and

$$ax_0 + b = \varphi(x_0).$$

But then we have that

$$\varphi \circ g(\omega) \geq a g(\omega) + b$$

for almost all $\omega \in \Omega$. Since we have a probability measure, the integral is monotone with $\mu(\Omega) = 1$ so that

$$\int_\Omega \varphi \circ g \, d\mu \geq \int_\Omega (ag + b) \, d\mu = a \int_\Omega g \, d\mu + b \int_\Omega d\mu = a x_0 + b = \varphi(x_0) = \varphi\left(\int_\Omega g \, d\mu\right),$$

as desired. $\square$

Theorem 5.7 (Minkowski's Inequality) *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. Then for $1 \leq p \leq \infty$, $f, g \in L^p$, then*

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

When $1 < p < \infty$, equality holds if and only if $f \sim \lambda g$ for some $\lambda \geq 0$ or $g \sim 0$.

Proof: One can verify directly that the statement holds for $p = 1$ or $p = \infty$, so we assume that $1 < p < \infty$.

If $\|f+g\|_p = 0$, then the statement holds trivially. So also assume $\|f+g\|_p \neq 0$. Note

$$\begin{aligned}
|f+g|^p &= |f+g||f+g|^{p-1} \\
&\leq (|f|+|g|)|f+g|^{p-1} \\
&= |f||f+g|^{p-1} + |g||f+g|^{p-1}
\end{aligned}$$

Apply Holder's inequality, we have

$$\int |f||f+g|^{p-1} d\mu \leq \|f\|_p \cdot \|(f+g)^{p-1}\|_q,$$

where q is the Hölder conjugate of p. Then $p = (p-1)q$. So

$$\|(f+g)^{p-1}\|_q = \|f+g\|_p^{p-1}.$$

Then

$$\|f+g\|_p^p = \int |f+g|^p d\mu \leq \|f\|_p \cdot \|f+g\|_p^{p-1} + \|g\|_p \cdot \|f+g\|_p^{p-1}$$

Then as $\|f+g\|_p \neq 0$, then

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

$\square$

Corollary 5.7.1 *For $1 \leq p \leq \infty$, $\|\cdot\|_p$ is a seminorm. Moreover, when $p = 2$, setting*

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

is a semi-inner product.

Definition 5.8 *If $1 \leq p \leq \infty$, the Lebesgue space $\hat{L}^p(\Omega, \mathscr{A}, \mu)$ is the quotient space of $L^p(\Omega, \mathscr{A}, \mu)$ with respect to the a.e.-equiavlence relation. On the space $\hat{L}^p$, we define addition and scalar multiplication by*

$$[f] + [g] = [f+g], \quad \textit{and} \quad \alpha[f] = [\alpha f].$$

Theorem 5.9 *For $1 \leq p \leq \infty$, the Lebesgue space $\hat{L}^p$ with the defined addition, scalar multiplication and p-norm form a normed vector space. When $p = 2$, there is an obvious way to define the inner product.*

Theorem 5.10 (Riesz Fischer) *For $1 \le p \le \infty$, the Lebesgue space $\hat{L}^p$ is complete with respect to the p-norm.*

Proof: We consider two cases, p if finite and p is infinite.

Case 1: p is finite.
Suppose $\{f_n\}$ is a Cauchy sequence in the norm, so that $\|f_n - f_m\|_{L^p(\mathbb{R}^d)} \to 0$ as $n, m \to \infty$. We extract a subsequence of $\{f_n\}$ that converges to f, both pointwise almost everywhere and in measure.

Let $\{f_{n_k}\}_{k=1}^{\infty}$ be a subsequence of $\{f_n\}$ with the following property:

$$\|f_{n_{k+1}} - f_{n_k}\|_{L^p} \le 2^{-k} \ \text{ for all } k \ge 1.$$

Such subsequence exists, since that for every 2^{-k}, $\exists N(2^{-k})$ such that whenever $n, m \ge N(\epsilon)$, $\|f_n - f_m\|_{L^p} \le \epsilon$.

Now consider the series

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty} (f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

Note that by the monotone convergence theorem and Minkowski's inequality, we have

$$\int |f(x)|^p d\mu \le \|g(x)\|_p^p \le (\|f_{n_1}(x)\|_p + 1)^p < \infty,$$

So $g \in L^p$ and $f \in L^p$. Then the series defining g converges almost everywhere, hence the series defining f converges almost everywhere, and since partial sums of this series are precisely the f_{n_k}, we find that

$$f_{n_k}(x) \to f(x), \quad \text{a.e. } x.$$

Next we show that $f_{n_k} \to f$ in L^p. Note that $|f - f_{n_k}| \le g$ for all k, then by Lebesgue dominated convergence theorem, we get $\|f_{n_k} - f\|_{L^p} \to 0$ as k tends to infinity.

Lastly, since $\{f_n\}$ is Cauchy, and it has a convergent subsequence, then $\{f_n\}$ converges. Since L^p is a metric space, then the limit of $\{f_n\}$ is the f we found.

Case 2: $p = \infty$. This is very similar to the case when p is finite. However, instead of using the monotone convergence theorem, and dominated convergence theorem, one can simply make using of the comparison of L^∞ norms, that is if $|f| \le |g|$ a.e., then $\|f\|_{L^\infty} \le \|g\|_{L^\infty}$. □

Definition 5.11 (ℓ^p Spaces) *Let $1 \le p < \infty$, the space ℓ^p is the space of all infinite sequence with finite p-norm, i.e., for any $(x_n) \in \ell^p$, we have*

$$\|x_n\|_p = \left(\sum |x_n|^p\right)^{\frac{1}{p}} < \infty.$$

If $p = \infty$*, then for any* $x_n \in \ell^\infty$*, we have*

$$\|x_n\|_\infty = \sup_n |x_n| < \infty.$$

Remark 5.11.1 *This is basically a special case where* $\Omega = \mathbb{N}$ *and* μ *is the counting measure.*

Consequently, we have the following Hölder's inequality and Minkowski Inequality for sequences:

Theorem 5.12 *Let* $(x_n) \in \ell^p$ *and* $y_n \in \ell^q$*, then*

1. $(x_n y_n) \in \ell^1$ *and*
$$\|x_n y_n\|_1 \leq \|x_n\|_p \cdot \|y_n\|_q$$
for $q > 1$.

2. *If* $q = p$*, then*
$$\|x_n + y_n\|_p \leq \|x_n\|_p + \|y_n\|_p$$
for $p \geq 1$.

5.2 Convergence and Inclusion of L^p spaces

Definition 5.13 (Types Of Convergence) *A sequence* $\{f_n\}$ *of measurable functions* ***converges in measure*** *to a measurable function, if for any* $\epsilon > 0$,

$$\lim_{n\to\infty} \mu(\{x \in \Omega : |f_n(x) - f| \geq \epsilon\}) \to 0.$$

We say $\{f_n\}$ *is Cauchy in measure, if for any* $\epsilon, \delta > 0$*, there is an* $N_{\epsilon,\delta}$*, s.t., for all* $m, n \geq N_{\epsilon,\delta}$,

$$\mu(\{x \in \Omega : |f_m - f_n| \geq \epsilon\}) < \delta.$$

Lemma 5.14 *On a finite measure space, a.e. pointwise convergence implies convergence in measure. Cauchy in measure (not necessarily on a finite measure space) implies a.e. pointwise convergence subsequence, moreover, we can also arrange the subsequence to be convergent in measure.*

Proof: For the first statement, given $\epsilon > 0$. Let $\{f_n\} \subset M(\Omega, \mathscr{A}, \mu)$ that converges to f pointwise a.e. Define $E_n = \{x : |f_n - f| > \epsilon\}$. Define $F_n = \bigcup_{k=n}^\infty A_k$. Then $\{F_n\}$ is a decreasing sequence, and $\mu(F_1) < \infty$ as the measure space is finite. So

$$\mu\left(\bigcap_{n=1}^\infty F_n\right) = \lim_{n\to\infty} \mu(B_n) = 0.$$

Thus

$$\lim_{n\to\infty} E_n \leq \lim_{n\to\infty} F_n = 0.$$

From the proof, we also see that we only require $\mu(F_n) < \infty$.

For the second statement, let $\{f_n\} \subset M(\Omega, \mathscr{A}, \mu)$. For each $k > 0$, let N_k be such that if $m, n > N_k$, then

$$\mu(\{x \in \Omega : |f_m(x) - f_n(x)| > 2^{-k}\} < 2^{-k}.$$

We choose a subsequence (g_k) of (f_n), where $g_k = f_{N_k}$. Then we show this subsequence convergences pointwise a.e. Set

$$E_k = \{x \in \Omega : |g_k(x) - g_{k+1}(x) > 2^{-k}\}$$

then

$$\mu(E_k) < 2^{-k},$$

and

$$F_k := \bigcup_{j=k}^{\infty} E_j \implies \mu(E_k) < 2^{-k+1}.$$

If $i \geq j \geq k$ and $x \notin F_k$, then

$$|g_i(x) - g_j(x)| \leq \frac{1}{2^{j+1}} + \frac{1}{2^{j+2}} + \cdots + \frac{1}{2^{i-1}} < \frac{1}{2^{j-1}}.$$

Take $F = \bigcap_{k=1}^{\infty} F_k$, $\mu(F) = \inf \mu(F_k) = 0$. We show that if $x \notin F$, then the sequence $\{g_n(x)\}$ is Cauchy. As $x \notin F_k$ for some K if $x \notin F$, and the previous calculation implies $g_n(x)$ is Cauchy. Define

$$g(x) = \begin{cases} \lim g_n(x), & \text{if } x \notin F \\ 0, & \text{otherwise} \end{cases}$$

Then $g_n \to g$ a.e. And since we can show for every $\epsilon > 0$, the set

$$\bigcup_{n=1}^{\infty} \{x : |g_n(x) - g(x)| > \epsilon\}$$

is finite, then by the first statement, we can conclude that the sequence converges in measure. □

Corollary 5.14.1 *If $\{f_n\}$ is Cauchy in measure, then there is f such that f_n converges to f in measure. Moreover, f is unique up to a set of measure zero.*

Proof: Take subsequence $\{g_n\}$ such that $g_n \to f$ pointwise a.e. g_k differs from f only on a small set, and f_n differs from f_{n_k} only on a small set. Hence we can get that f_n differs from f on a very small set. □

Proposition 5.15 *If $\{f_n\} \subset M(\Omega, \mathscr{A}, \mu)$ converges uniformly to measurable f or doing so in L^p-fashion. Then f_n converges in measure to f.*

Remark 5.15.1 *The same result holds if we replace convergence by Cauchy.*

Proof: Suppose $f_n \to f$, uniformly. Then given any $\epsilon > 0$, $\exists N \in \mathbb{N}$, s.t., if $n > N$, then $|f_n - f| < \epsilon$. Hence

$$\mu(\{x \in \Omega : |f_n(x) - f| \geq \epsilon\}) = 0$$

for large enough n.

For L^p-convergence, the case where $p = \infty$ is clear. So assume $p \neq \infty$. If $f_n \to f$ in L^p. Then, suppose

$$\limsup_{n\to\infty} \mu(\{x \in \Omega : |f_n(x) - f| \geq \epsilon\}) = c > 0.$$

Then

$$\limsup_{n\to\infty} \|f_n - f\|_{L^p} \geq \left(\epsilon^p \frac{c}{2}\right)^{1/p} > 0$$

which is a contradiction. $\square$

Proposition 5.16 *Suppose $f \in L^p \cap L^q$, then $f \in L^r$, where $p < r < q$.*

Proof: Suppose $q = \infty$, then it is trivial. Otherwise,

$$\begin{aligned}\int_\Omega |f|^r &= \int_{|f|\geq 1} |f|^r + \int_{|f|<1} |f|^r \\ &\leq \int_{|f|\geq 1} |f|^q + \int_{|f|<1} |f|^p \\ &< \infty.\end{aligned}$$

$\square$

Theorem 5.17 (Lebesgue Dominated Convergence Theorem: Convergence In Measure) *Suppose $\{f_n\} \subset M(\Omega, \mathscr{A}, \mu)$, and f_n converges to f in measure. Moreover,*

$$|f_n(x)| < g(x) \quad a.e.$$

with $g \in L^1(\Omega, \mu)$. Then

$$\int_\Omega f d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu.$$

Proof: Let $\alpha_n := \int_\Omega f_n$, we show $\{\alpha_n\}$ converges and is equal to $\int_\Omega f d\mu$ (this might not be well-defined yet).

To show $\{\alpha_n\}$ converges to some L, it suffices to show that for every subsequence of $\{\alpha_n\}$, there is a subsequence that converges to L. Let $\{\alpha_{n_k}\}$ be a subsequence, then $\{f_{n_k}\}$ is a subsequence of $\{f_n\}$. Since f_n converges to f in measure, then $\{f_{n_k}\}$ has a a.e. convergence subsequence that also converges in measure. Since f_n converges in measure to f, then the limit of a.e. convergent subsequence equals to f a.e. Then by the standard version of LDCT, we have the limit of the integral is $\int f d\mu$ as desired. $\square$

Theorem 5.18 (p-norm Dominated Convergence Theorem) *Let $(f_n) \subset L^p(\Omega)$, where $1 \leq p < \infty$. If $f_n \to f$, and f_n is dominated by some $g \in L^p(\Omega)$, i.e., $|f_n| \leq g$ a.e. for all n. Then $f \in L^p(\Omega)$ and $f_n \to f$ in p-norm.*

Proof: Note that

$$|f_n - f| \le |f_n| + |f| \le 2g.$$

So

$$|f_n - f|^p \le 2^p g^p \in L^p$$

and apply the L.D.C.T to get the desired result. $\square$

Proposition 5.19 *Suppose $1 \le p < \infty$. Let $f_n \to f$ a.e., and $f_n, f \in L^p$. Then $f_n \to f$ in p-norm iff $\|f_n\|_p - \|f\|_p \to 0$.*

Proof: Suppose $f_n \to f$ in p-norm, then $\|f_n - f\|_p \to 0$, so by Minkowski Inequality, we have $\|f_n\|_p \to \|f\|_p$.

Conversely, suppose $\|f_n\|_p \to \|f\|_p$ and $|f_n|^p \to |f|^p$ a.e. Since

$$|f_n - f|^p \le 2^p(|f_n|^p + |f|^p),$$

write

$$h_n = |f_n - f|^p \text{ and } g_n = 2^p(|f_n|^p + |f|^p).$$

Then clearly $g_n \in L^1$, and

$$g_n \to g := 2^{p+1}|f|^p \in L^1.$$

Apply the Generalized D.C.T on h_n and g_n gives

$$\int |f_n - f|^p \to 0.$$

$\square$

Lemma 5.20 *Let (f_n) be a sequence of measurable function and f is a measurable function. The following are equivalent:*

1. *(f_n) converges to f in L^∞.*
2. *There is a null set N such that on $\Omega \setminus N$, f_n converges uniformly to f.*

Proof: It is clear that (2) implies (1). We show (1) implies (2). Suppose (f_n) converges to f in L^∞, then

$$\|f_n - f\|_{L^\infty} \to 0$$

Then for each $n \in \mathbb{N}$, let E_n consists of all $x \in \Omega$ such that $|f_n(x) - f(x)| > \|f_n - f\|_{L^\infty}$, then E_n is a null set. Hence $N = \bigcup_{n=1}^\infty E_n$ has measure zero by countable additivity. Thus on $\Omega \setminus N$, f_n converges uniformly to f. $\square$

Definition 5.21 *A sequence of functions converges* ***nearly uniformly*** *to f if for each $\delta > 0$, one can find E_δ with $\mu(E_\delta) < \delta$ such that f_n converges uniformly to f in $\Omega \setminus E_\delta$.*

Lemma 5.22 *If a sequence (f_n) of measurable functions converges in L^∞ to f, then (f_n) converges nearly uniformly to f.*

Proof: Clear. $\square$

Lemma 5.23 *If a sequence (f_n) of measurable function converges nearly uniformly to f, then (f_n) converges in measure to f.*

Proof: Suppose $f_n \to f$ converges nearly uniformly. Let $\delta > 0$ be arbitrary, then $f_n \to f$ on $\Omega \setminus E_\delta$. For any ϵ, one can choose N_ϵ, such that

$$|f|_{\Omega \setminus E_\delta} - f_n|_{\Omega \setminus E - \delta}| < \epsilon, \;\; n \geq N_\epsilon.$$

Then the measure such that $|f_n(x) - f(x)| > \epsilon$ can be arbitrarily small as n is large enough. $\square$

5.3 L^p Spaces In Finite Measure Space

Finite measure space is analogous to the notion of compactness. Finite positive measure space can be scaled to have a measure of 1. Then, this is more of less the setting of probability.

Theorem 5.24 *Suppose $\mu(\Omega) < \infty$, then $L^\infty \subseteq L^p$, $1 \leq p \leq \infty$. Moreover if (f_n) is a sequence which converges in L^∞, then (f_n) converges in L^p for $1 \leq p < \infty$, in particular, if (f_n) are also a sequence in $L^p(\Omega, \mathscr{A}, \mu)$, then f is in $L^p(\Omega, \mathscr{A}, \mu)$.*

Remark 5.24.1 *Notice that converges in L^p does not imply each f_n is in L^p, nor does it mean f is in L^p.*

Proof: Suppose $f \in L^\infty$. Then there is M such that

$$N := \{x \in \Omega \,|\, |f(x)| > M\}$$

is a null set. Then

$$\begin{aligned} \int_\Omega |f|^p d\mu &= \int_{\Omega \setminus N} |f|^p d\mu + \int_N |f|^p d\mu \\ &\leq M^p \mu(\Omega) + 0 \\ &\leq \infty. \end{aligned}$$

Next, let (f_n) be a sequence that converges in L^∞ to f. We show that $\|f_n - f\|_{L^p} \to 0$. For any $\epsilon > 0$, there is N_ϵ, s.t., $\|f_n - f\| < \epsilon$ for $n \geq N_\epsilon$. Then

$$\|f_n - f\|_{L^p}^p \leq \epsilon^p \mu(\Omega)$$

for $n \geq N_\epsilon$. As ϵ is arbitrary, then $f_n \to f$ in L^p. $\square$

Corollary 5.24.1 *If $\mu(\Omega) < \infty$, and (f_n) converges uniformly to f, then f_n converges in L^p to f.*

Proposition 5.25 *Suppose $\mu(\Omega) < \infty$, and $1 \leq p < q < \infty$. Then $L^q \subset L^p$. Moreover, if (f_n) is a sequence converges to f in L^q, then (f_n) converges to f in L^p. In particular, if each f_n is also in L^p, then f is also in L^p.*

Remark 5.25.1 *The inclusion is usually proper, i.e. there are functions in L^p but not in L^q.*

Proof: Let $\alpha = \frac{q}{p}$ and $\beta = \frac{q}{q-p}$, then $\frac{1}{\alpha} + \frac{1}{p} = 1$. Then by Hölder's Inequality, we have

$$\int |f|^p d\mu = \||f|^p\|_1 \leq \||f|^p\|_\alpha \|\chi_\Omega\|_\beta = \left(\int |f|^q\right)^{\frac{p}{q}} \cdot C < \infty.$$

□

Theorem 5.26 *Suppose $\mu(\Omega) < \infty$ and (f_n) is a sequence of measurable functions converging a.e. to f. Then (f_n) converges nearly uniformly and also in measure to f.*

Proof: It suffices to show that $(f_n) \to f$ a.e. implies (f_n) converges nearly uniformly to f. Then the second part follows immediate from Lemma 5.23. WLOG, we can assume that $f_n \to f$ pointwise, then for $m, n > 0$, set

$$E_n(m) = \bigcup_{k=n}^{\infty} \{x : |f_k(x) - f(x)| \geq \frac{1}{m}\}.$$

So $E_n(m)$ is measurable, $E_{n+1}(m) \subseteq E_n(m)$, since $f_n \to f$ pointwise. $\bigcap E_n(m) = \emptyset$, $\mu(\Omega) < \infty$, hence $\mu(\bigcap_n E_n(m)) = 0$. Then for each m, there exists k_m such that

$$\mu(E_{k_m}(m)) < \frac{\delta}{2^m}.$$

Set

$$E_\delta = \bigcup_{m=1}^{\infty} E_{k_m}(m),$$

Then f converges uniformly on $\Omega \setminus E_\delta$ and $\mu(E_\delta) \leq \delta$. □

Proposition 5.27 *If $1 \leq p < r < \infty$, then*

$$\ell^p \subset \ell^r \subset \ell^\infty.$$

Proof: The inequality $\ell^r \subset \ell^\infty$ is clear. Let $(x_n) \in \ell^p$, so $x_n \to 0$, i.e., there exists $N \in \mathbb{N}$ such that $|x_n| < 1$ for all $n \geq N$. Hence

$$\begin{aligned}\sum |x_n|^r &= \sum_{|x_n|\geq 1} |x_n|^r + \sum_{|x_n|<1} |x_n|^r \\ &\leq \sum_{|x_n|>1} |x_n|^r + \sum_{|x_n|^p} \\ &< \infty.\end{aligned}$$

□

5.4 Application To Probability Space

Recall a space $(\Omega, \mathscr{A}, \mu)$ is called a **probability space** if $\mu(\Omega) = 1$. A **random variable** X is a measurable function from Ω to $\mathbb{R}$. The **expectation** $E[X]$ of such X is just $\int_\Omega X d\mu$. The **variance** of X is given by

$$\text{Var(X)} = \int_\Omega (X - E[X])^2 d\mu = \|X - E[X]\|_{L^2}^2$$

For random variables X, Y, their **covariance** is given by

$$\text{Cov}(X_n, X_m) = \int_\Omega (X - E[X])(Y - E[Y]) d\mu = \langle X - E[X], Y - E[Y] \rangle.$$

Definition 5.28 (Uncorrelated) *A sequence (X_n) of random variables is* ***uncorrelated*** *if the covariance* $\text{Cov}(X_n, X_m) =$ 0 *for all* $m \neq n$.

A sequence (X_n) of random variables is ***covariance stationary*** *if $(E[X_n])$ is constant and for each $k \geq 0$, $(\text{Cov}(X_n, X_{n+k}))$ is a constant. In particular, $\text{Var}(X_n) = \text{Cov}(X_n, X_n)$ is a constant.*

Theorem 5.29 (Chebyshev Weak Law of Large Number) *If (X_n) is an uncorrelated covariance stationary sequence of random variable. Let*

$$\overline{X_n} = \frac{X_1 + \cdots + X_n}{n}.$$

Then $(\overline{X_n})$ converges in probability to $E[X_i]\chi_\Omega$ (which is just $E[X_1]\chi_\Omega$).

Proof: We will show $(\overline{X_n})$ converges in L^2 to $E[X_1]\chi_\Omega$, this would imply $(\overline{X_n})$ converges to $E[X_1]\chi_\Omega$ in measure. Let $\alpha = E[X_1]$.

$$\begin{aligned}
\int (\overline{X_n} - \alpha)^2 d\mu &= \int \left(\frac{X_1 + \cdots + X_n}{n} - \alpha \right)^2 d\mu \\
&= \int \left[\frac{(X_1 - \alpha) + \cdots + (X_n - \alpha)}{n} \right]^2 d\mu \\
&= \frac{1}{n^2} \sum_{1 \leq i,j \leq n} \int (X_i - \alpha)(X_j - \alpha)] d\mu \\
&= \frac{1}{n^2} \sum_{i=1}^n \int (X_i - \alpha)^2 d\mu \\
&= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) d\mu \\
&= \frac{\text{Var}(X_1)}{n} \to 0
\end{aligned}$$

□

Definition 5.30 (Independent Events) *Let X be a random variable, its* ***distribution function*** *F_X, is defined to be*

$$F_X(a) := \mu\{w \in \Omega : X(\omega) \geq \alpha\}.$$

*Two events E, F in Ω (measurable subsets of Ω) are called **independent** if*

$$\mu(E \cap F) = \mu(E)\mu(F).$$

Two random variables X, Y are independent if for all α and β,

$$E_\alpha := \{\omega \in \Omega \,:\, X(\omega) \geq \alpha\}$$
$$F_\beta := \{\omega \in \Omega \,:\, Y(\omega) \geq \beta\}$$

are independent events.

Proposition 5.31 *If X and Y are independent random variables, then $E[XY] = E[X]E[Y]$.*

Proof: In the special case of charactersitic functions, $X = \chi_A$ and $Y = \chi_B$, then

$$\begin{aligned}\int \chi_A \cdot \chi_B d\mu &= \int \chi_{A\cap B} \\ &= \mu(A \cap B) \\ &= \mu(A) \cdot \mu(B) \\ &= \int \chi_A d\mu \cdot \int \chi_B d\mu\end{aligned}$$

since X, Y are independent.

Then by linearity, this is true for all nonnegative simple functions. By MCT, this is true for all nonnegative functions. Lastly, by linearity, the statement is true for all general functions. □

Corollary 5.31.1 *If X and Y are independent, then they are uncorrelated, i,.e.,* $\mathrm{Cov}(X, Y) = 0$.

Proof:

$$\begin{aligned}&\int (X - E[X]\cdot\chi_\Omega)(Y - E[Y]\cdot\chi_\Omega) d\mu \\ &= \int XY d\mu - E[Y]\int X d\mu - E[X]\int Y d\mu + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] \\ &= 0\end{aligned}$$

□

Definition 5.32 (Identically Distributed) *A sequence (X_n) of random variable is **identically distributed**, if $F_{X_i} = F_{X_j}$ for all $i, j \in \mathbb{N}$.*

Corollary 5.32.1 *Suppose (X_n) is a sequence of identically distributed independent random variables (i.i.d.). Assume that X_n has finite mean and variance. Take*

$$\overline{X_n} = \frac{X_1 + \cdots + X_n}{n} \tag{5.1}$$

Then $(\overline{X_n})$ converges in measure to $\alpha \cdot \chi_N$, with $\alpha = E[X_1] = \int X_1 d\mu$.

Proof: By Corollary 5.31.1, the X_n's are uncorrelated. We need to check $E[X_i] = E[X_j] = \alpha$:

$$\begin{aligned} E[X_i] &= \int X_i d\mu \\ &= \int_{-\infty}^{\infty} sF'_{X_i}(s)ds \\ &= \int_{-\infty}^{\infty} sF'_{X_j}(s)ds \\ &= E[X_j] \end{aligned}$$

Then by similar argument, we can show $\mathrm{Var}(X_i) = \mathrm{Var}(X_j)$. So the sequence (X_i) is covariance stationary. Apply the weak law of large number, we get the desired result. □

Proposition 5.33 (Chebyshev's Inequality) *Suppose X is a random variable with mean α and variance σ^2, then for all $\gamma > 0$, then*

$$\mu\{\omega \in \Omega : |X(\omega) - \alpha| > \gamma\sigma\} \leq \frac{1}{\gamma^2}$$

Proof: By the definition of variance, we have

$$\begin{aligned} \sigma^2 &= \int_\Omega (x-\alpha)^2 d\mu \\ &\geq \int_{|x-\alpha|\geq\gamma\sigma} (x-\alpha)^2 d\mu \\ &\geq \int_{|x-\alpha|\geq\gamma\sigma} \gamma^2\sigma^2 d\mu \\ &\geq \gamma^2\sigma^2\mu\{\omega \in \Omega : |X(\omega) - \alpha| > \delta\alpha\} \end{aligned}$$

Dividing by $\gamma^2\sigma^2$ we get the desired result. □

Theorem 5.34 (Strong Law of Large Number) *Suppose (X_n) is an i.i.d. sequence of random variables of finite measure and variance. Let*

$$\overline{X_n} = \frac{X_1 + \cdots + X_n}{n}.$$

Then $(\overline{X_n})$ converge a.e. (almost surely) to $\alpha\chi_\Omega$, where $\alpha = E[X_1]$.

Remark 5.34.1 *Since the space is finite. Then convergence a.e. implies convergence in measure. So the strong law of large number implies the weak law of large number.*

5.5 Riesz Representation Theorem for L^p Spaces

Definition 5.35 (Linear Functional) *A **real linear functional** on $L^p(\Omega, \mathscr{A}, \mu)$ is a mapping $G : L^p \to \mathbb{R}$, s.t.,*

$$G(af + bg) = aG(f) + bG(b)$$

for all $a, b \in \mathbb{R}$ and $f, g \in L^p$. A linear functional is ***bounded*** *if there is a constant M, such that*

$$\|G(f)\| \leq M\|f\|_p, \ \ \forall \|f\|_p \leq 1.$$

If G is bounded, then we define the ***norm*** *of G to be*

$$\|G\| = \sup\{|G(f)| \, : \, \|f\|_p \leq 1\}.$$

Lemma 5.36 *Let $g \in L^1(\Omega)$ and $\int g\varphi \leq M\|\varphi\|_p$ for $1 \leq p < \infty$ for every simple function $\varphi \in L^p(\Omega)$. Then $g \in L^q(\Omega)$ and $\|g\|_q \leq M$, where q is the Hölder conjugate of p.*

Proof: Since g is measurable then there exists a nondecreasing sequence of measurable simple function $\{\varphi_n\} \subset M(\Omega, \mathscr{A}, \mu)$ such that $\varphi_n \to |g|$ pointwise, hence $\varphi_n^q \to |g|^q$ pointwise. Let

$$h_n(x) = \operatorname{sgn}(g)\varphi_n^{q-1},$$

then $h_n \in L^p$ by L.D.C.T. and

$$\int \varphi_n^q \leq \int g h_n \leq M\|h_n\|_p = M\left(\int \varphi_n^q\right)^{\frac{1}{p}}.$$

Dividing both sides by $(\int \varphi_n^q)^{1/p}$ gives

$$\|\varphi_n\|_q \leq M.$$

Hence the result follows from the MCT. □

Lemma 5.37 *Suppose $1 \leq p \leq \infty$ and q is the Hölder conjugate of p. If $g \in L^q$, then the functional defined by*

$$\ell(f) = \int_X f(x)g(x)d\mu(x)$$

is bounded with $\|\ell\| \leq \|g\|_{L^q}$.

Proof: Hölder's Inequality. □

Corollary 5.37.1 *For $1 \leq p \leq \infty$, $L^q \subset (L^p)^*$ where q is the Hölder conjugate of p.*

Lemma 5.38 *Suppose $1 \leq p, q \leq \infty$ are Hölder conjugates, then if $g \in L^q$, then $\|g\|_{L^q} = \sup\limits_{\|f\|_{L^p} \leq 1} \left|\int fg\right|$.*

Proof: If $g = 0$, there is nothing to prove, so we may assume that g is not 0 a.e., and hence $\|g\|_{L^q} \neq 0$. By Hölder's inequality, we have that

$$\|g\|_{L^q} \geq \sup_{\|f\|_{L^p} \leq 1} \left|\int fg\right|.$$

To prove the reverse inequality we consider several cases:

- First, if $q = 1$ and $p = \infty$, we may take $f(x) = \text{sign } g(x)$. Then, we have $\|f\|_{L^\infty} = 1$, and clearly, $\int fg = \|g\|_{L^1}$.
- If $1 < p, q < \infty$, then we set $f(x) = \frac{|g(x)|^{q-1}\text{sign } g(x)}{\|g\|_{L^q}^{q-1}}$. We observe that $\|f\|_{L^p}^p = \int |g(x)|^{p(q-1)}d\mu / \|g\|_{L^q}^{p(q-1)} = 1$ since $p(q-1) = q$, and that $\int fg = \|g\|_{L^q}$.

- Finally, if $q = \infty$ and $p = 1$, let $\epsilon > 0$, and E a set of finite positive measure, where $|g(x)| \geq \|g\|_{L^\infty} - \epsilon$. (Such a set exists by the definition of $\|g\|_{L^\infty}$ and the fact that the measure μ is σ - finite.) Then, if we take $f(x) = \frac{\chi_E(x)\text{sign } g(x)}{\mu(E)}$, where χ_E denotes the characteristic function of the set E, we see that $\|f\|_{L^1} = 1$, and also
$$\left|\int fg\right| = \frac{1}{\mu(E)}\int_E |g| \geq \|g\|_\infty - \epsilon.$$

□

Lemma 5.39 *Suppose $1 \leq p, q \leq \infty$ are Hölder conjugates (When $p = 1$, also impose the condition that X is σ-finite), and g is integrable on all sets of finite measure with*

$$\sup_{\substack{\|f\|_{L^p} < 1 \\ f \text{ simple}}} \left|\int fg\right| = M < \infty.$$

Then $g \in L^q$ and $\|g\|_{L^q} = M$.

Proof: Let g_n be a sequence of simple functions so that $|g_n(x)| \leq |g(x)|$ and $g_n(x) \to g(x)$ for each x. When $p > 1$ ($q < \infty$), let $f_n(x) = [|g_n(x)|^{q-1}\text{sgn}g(x)]/\|g_n\|_{L^q}^{1-1}$. We note $\|f_n\|_{L^p} = 1$ and

$$\int f_n g = \frac{\int |g_n(x)|^q}{\|g_n\|_{L^q}^{q-1}} = \|g_n\|_{L^q}$$

which does not exceed M. By Fatou's lemma, it follows that $\int |g|^q \leq M^q$, so $g \in L^q$, with $\|g\|_{L^q} \leq M$. The direction $\|g\|_{L^q} \geq M$ follows from Hölder's inequality.

When $p = 1$, we can consider $f_n(x) = (\text{sgn}g(x))\chi_{E_n}(x)$, where E_n is an increasing sequence of sets of finite measure whose union is X. □

Theorem 5.40 (Riesz Representation Theorem for L^p spaces) *Suppose $(\Omega, \mathscr{A}, \mu)$ is an arbitrary measure space if $1 < p < \infty$, and let it be a σ-finite measure space if $p = 1$. Let $G : L^p \to \mathbb{R}$ be a bounded linear functional. Then there is $g \in L^q$ such that*
$$G(f) = \int f \cdot g d\mu$$
where q is the Hölder conjugate of p. Moreover,

$$\|G\| = \|g\|_{L^q}.$$

Proof: First suppose that the underlying space has finite measure. Then let ℓ be the given functional on L^p, we can then define a set function ν by
$$\nu(E) = \ell(\chi_E),$$

where E is any measurable set. This is well-defined since χ_E is L^p as the space has finite measure. We observe that

$$|\nu(E)| \leq c(\mu(E))^{1/p},$$

where c is the norm of the linear functional. Now the linearity of ℓ implies that ν is finitely-additive. Moreover, if $\{E_n\}$ is a countable collection of disjoint measurable sets, and we put $E = \bigcup_{n=1}^{\infty} E_n$, $E_N^* = \bigcup_{n=N+1}^{\infty} E_n$, then

$$\chi_E = \chi_{E_N^*} + \sum_{n=1}^{N} \chi_{E_n}.$$

Thus $\nu(E) = \nu(E_N^*) + \sum_{n=1}^{N} \nu(E_n)$. However $\nu(E_N^*) \to 0$ as $N \to \infty$, because $|\nu(E_N^*)| \leq c(\mu(E_N^*)^{1/p}$, where c is the norm of the linear functional, and $p < \infty$. This shows that ν is countably additive. In particular, we also observe that ν is absolutely continuous with respect to μ. Hence, by the Randon Nykodim Theorem (7.49), after decomposing ν into the positive and negative part, we can find an integrable function g so that

$$\nu(E) = \int_E g d\mu$$

for every measurable set E. Thus we have $\ell(\chi_E) = \int \chi_E g d\mu$. The presentation $\ell(f) = \int f g d\mu$ then extends immediately to simple functions f and by passage to the limit to all $f \in L^p$. (Since the simple functions are dense in L^p for $1 \leq p < \infty$). Also by Lemma (5.39), we see that $\|g\|_{L^q} = \|\ell\|$).

Now suppose we have a general measure space. We again invoke Lemma (5.39) to complete the proof. □

The dual of L^∞ is not (at least) L^1. We can consider the following functional $F : L^\infty([0,1]) \to \mathbb{R}$ defined by

$$F(y) = y(0)$$

for every continuous function y, where $y \in L^\infty([0,1])$. Then by continuous extension, we indeed get a well-defined linear map that is bounded. Now suppose there exists $g \in L^1$

$$F(y) = \int_0^1 y g dx.$$

Let (y_n) be a sequence of continuous function on $[0,1]$ defined as $y_n(0) = 1$ and $y_n(x) \searrow 0$ on $[0,1]$. Then $|y_n| \leq 1$, and

$$F(y_n) = 1 = \int_0^1 y_n(x) g(x) dx.$$

Note that $y_n(x)g(x) \to 0$ a.e. and

$$|g(x)y_n(x)| \leq g(x) \in L^1[0,1].$$

So by the L.D.C.T, we have

$$\int_0^1 y_n(x) g(x) dx \to 0$$

which is a contradiction.

5.6 L^p spaces with $0 < p < 1$

In this section we consider what happens to the space L^p if $0 < p < 1$. Again, the space will consists of all measurable functions $f : \Omega \to \mathbb{R}$ such that

$$\int_\Omega |f|^p d\mu < \infty.$$

This space is indeed a linear space, that is if $f, g \in L^p$, then $f + cg \in L^p$. However. the nature way of defining the norm fails due to the following two results. Nevertheless, we will still use the symbol $\|\cdot\|_p$ to denote

$$\|f\|_p = \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}}.$$

Proposition 5.41 (Reverse Hölder Inequality) *Let $0 < p < 1$, and q be the Hölder Conjugate of p. Suppose $f \in L^p$ and $g \in L^q$*

$$\int |fg| \geq \|f\|_p \|g\|_q.$$

Proof: Since $0 < p < 1$, then $q < 0$. Let $r = \frac{1}{p}$ and $t = -\frac{q}{p}$. Then r, t are Hölder conjugates and $1 < r < \infty$. Moreover, $(fg)^p \in L^r$ and $g^{-p} \in L^1$. Using Hölder's Inequality, we get

$$\begin{aligned}
\int ||fg|^p |g|^{-p}| &\leq \left(\int (|fg|^p)^r\right))^{\frac{1}{r}} \cdot \left(\int (|g|^{-p})^t\right)^{\frac{1}{t}} \\
&= \left(\int |fg|\right)^p \left(\int |g|^q\right)^{-\frac{p}{q}}
\end{aligned}$$

Hence

$$\left(\int |f|^p\right)^{\frac{1}{p}} \leq \left(\int |fg|\right)\left(\int |g|^q\right)^{-\frac{1}{q}}.$$

Multiplying both sides by $(\int |g|^q)^{1/q}$ we get the desired result. □

Corollary 5.41.1 (Reverse Minkowski Inequality) *Let $f, g \in L^p$ for $0 < p < 1$, then*

$$\|f\|_p + \|g\|_p \leq \|f + g\|_p.$$

5.7 Density Results

Similar to Theorem 2.22, we can establish the following results.

Proposition 5.42 *Let $(\Omega, \mathscr{A}, \mu)$ be an arbitrary measure space. The collection of simple functions in $L^p(\Omega)$ is dense in L^p for $1 \leq p \leq \infty$.*

Proposition 5.43 *Suppose $\Omega \subseteq \mathbb{R}^d$, then step functions in $L^p(\Omega)$ are dense in $L^p(\Omega)$ for $1 \leq p < \infty$.*

Proposition 5.44 *Suppose $\Omega \subseteq \mathbb{R}^d$, then the family of all continuous function with compact support in Ω is dense in $L^p(\Omega)$ for $1 \leq p < \infty$.*

Proposition 5.45 *If $\mu(\Omega) < \infty$, then simple functions in $L^\infty(\Omega)$ are dense in $L^\infty(\Omega)$.*

Proposition 5.46 *For $1 \le p < \infty$, the space $L^p(\mathbb{R}^n)$ is separable. In fact, the finite rational linear combination of the family $\{\chi_R(x)\}$ where R is a rectangle in $\mathbb{R}^d$ with rational coordinates is dense in $L^p(\mathbb{R}^d)$.*

6 Hausdorff Dimension

6.1 Hausdorff Measure

The theory begins with the introduction of a new notion of volume or mass. This "measure" is closely tied with the idea of dimension which prevails throughout the subject. More precisely, following Hausdorff, one considers for each appropriate set E and each $\alpha > 0$ the quantity $m_\alpha(E)$, which can be interpreted as the α-dimensional mass of E among sets of dimension α, where the word "dimension" carries (for now) onlyan intuitive meaning. Then, if α is larger than the dimension of the set E, the set has a negligible mass, and we have $m_\alpha(E) = 0$. If α is smaller than the dimension of E, then E is very large (comparatively), hence $m_\alpha(E) = 1$. For the critical case when α is the dimension of E, the quantity $m_\alpha(E)$ describes the actual α-dimensional size of the set.

Example: Recall that the standard Cantor set $\mathcal{C}$ in $[0, 1]$ has zero Lebesgue measure. This statement expresses the fact that $\mathcal{C}$ has one-dimensional mass or length equal to zero. We shall prove that $\mathcal{C}$ has a well-defined fractional Hausdorff dimension of $\log 2/ \log 3$, and the corresponding Hausodrff measure of the Cantor set is positive and finite. Another illustration of the theory developed below consists of starting with Γ, a rectifiable curve in the plane. Then Γ has zero two-dimensional Lebesgue measure. This is intuitively clear, since Γ is a one-dimensional object in a two-dimensional space. This is where the Hausdorff measure comes into play: the quantity $m_1(\Gamma)$ is not only finite, but precisely equal to the length of Γ whose formula is given in the previous sections.

Definition 6.1 *For any subset E of $\mathbb{R}^d$, we define the **exterior α-dimensional Hausdorff measure** of E by*

$$m_\alpha^*(E) := \lim_{\delta\to 0} \inf \left\{ \sum_k (\operatorname{diam} F_k)^\alpha \;\; E \subset \bigcup_{k=1}^\infty F_k, \;\operatorname{diam} F_k \leq \delta \textit{ for all } k \right\}.$$

Remark 6.1.1 *We note that the quantity*

$$\mathcal{H}_\alpha^\delta(E) := \lim_{\delta\to 0} \inf \left\{ \sum_k (\operatorname{diam} F_k)^\alpha \;\; E \subset \bigcup_{k=1}^\infty F_k, \;\operatorname{diam} F_k \leq \delta \textit{ for all } k \right\}$$

is increasing as δ decreases, so that the limit

$$m_\alpha^*(E) = \lim_{\delta\to 0} \mathcal{H}_\alpha^\delta(E)$$

exits and could be $+\infty$. We also note that in particular, one has $\mathcal{H}_\alpha^\delta(E) \leq m_\alpha^*(E)$ for all $\delta > 0$.

Remark 6.1.2 *Scaling is the key notion that appears at the heart of the definition of the exterior Hausdorff measure. Loosely speaking, the measure of a set scales according to its dimension. For instance, if Q is a cube in $\mathbb{R}^d$, the volume of rQ is $r^d|Q|$. Now in the definition of exterior Hausdorff measure, if the set F is scaled by r, then $(\operatorname{diam} F)^\alpha$ scales by r^α, which captures this scaling with respect to dimension idea.*

Proposition 6.2 (Properties of Hausdorff Exterior Measure)

1. $m_\alpha^*(\emptyset) = 0$.

2. *(Monotonicity) If $E_1 \subset E_2$, then $m_\alpha^*(E_1) \leq m_\alpha^*(E_2)$.*

3. *(Sub-additivity)* $m_\alpha^*\left(\bigcup_{j=1}^{\infty} E_j\right) \le \sum_{j=1}^{\infty} m_\alpha^*(E_j)$ *for any countable family* $\{E_j\}$ *of sets in* $\mathbb{R}^d$.

4. *If* $d(E_1, E_2) > 0$, *then* $m_\alpha^*(E_1 \cup E_2) = m_\alpha^*(E_1) + m_\alpha^*(E_2)$. *In particular, we see* m_α^* *is a metric exterior measure.*

Proof: (1) and (2) is clear. We prove (3) and (4).

For the proof of (3), fix δ, and choose for each j a cover $\{F_{j,k}\}_{k=1}^{\infty}$ of E_j by sets of diameter less than δ such that $\sum_k (\operatorname{diam} F_{j,k})^\alpha \le \mathcal{H}_\alpha^\delta(E_j) + \epsilon/2^j$. Since $\bigcup_{j,k} F_{j,k}$ is a cover of E by sets of diameter less than δ, we must have

$$\begin{aligned}
\mathcal{H}_\alpha^\delta(E) &\le \sum_{j=1}^{\infty} \mathcal{H}_\alpha^\delta(E_j) + \epsilon \\
&\le \sum_{j=1}^{\infty} m_\alpha^*(E_j) + \epsilon.
\end{aligned}$$

Since ϵ is arbitrary, the inequality $\mathcal{H}_\alpha^\delta(E) \le \sum m_\alpha^*(E_j)$ holds, and we let δ tend to 0 to prove the countable sub - additivity of m_α^*.

For the proof of (4), by sub-additivity, it suffices to prove $m_\alpha^*(E_1 \cup E_2) \ge m_\alpha^*(E_1) + m_\alpha^(E_2)$. By since $d(E_1, E_2) > 0$, then when δ is small enough, we clearly see

$$H_\alpha^\delta(E_1 \cup E_2) \ge H_\alpha^\delta(E_1) + H_\alpha^\delta(E_2).$$

Hence we prove the desired result by letting $\delta \to 0$. □

Proposition 6.3 *The restriction of* m_α^* *to the Borel sets of* $\mathbb{R}^d$ *is a countably additive measure on the Borel sets.*

Remark 6.3.1 *The restricted measure is denoted* m_α, *and is called the* ***α-dimensional Hausdorff measure***.

Proof: Since m_α^* is a metric exterior measure, this follows directly from Theorem (7.26). □

Corollary 6.3.1 (Property of Hausdorff Measure)

1. *If* $\{E_j\}$ *is a countable family of disjoint Borel sets, and* $E = \bigcup_{j=1}^{\infty} E_j$, *then*

$$m_\alpha(E) = \sum_{j=1}^{\infty} m_\alpha(E_j).$$

2. *If* $h \in \mathbb{R}^d$ *and* $\lambda \in \mathbb{R}$, *then*

$$m_\alpha(E + h) = m_\alpha(E)$$

and

$$m_\alpha(\lambda E) = |\lambda|^\alpha m_\alpha(E).$$

3. *Suppose r is a rotation map on $\mathbb{R}^d$, then*

$$m_\alpha(r(E)) = m_\alpha(E).$$

Proposition 6.4 *The quantity $m_0(E)$ counts the number of points in E. $m_1(E) = m(E)$ for all Borel sets $E \subset \mathbb{R}$, where m is the standard Lebesgue measure. If E is a Borel subset of $\mathbb{R}^d$, then $c_d m_d(E) = m(E)$ for some constant c_d that depends only on the dimension d, in fact, we have*

$$c_d = \frac{m(B)}{(\operatorname{diam} B)^d},$$

where B is the unit ball in $\mathbb{R}^d$.

Proof: For the first assertion, it is clear that $m_0(E) = |E|$ if E is finite, and $m_0(E) = \infty$ if E is infinite. We also have by definition that $m_1(E) = m(E)$ if $E \subset \mathbb{R}$. The proof of the last assertion follows from iso-diametric inequality. □

Proposition 6.5 *If $m_\alpha^*(E) < \infty$ and $\beta > \alpha$, then $m_\beta^*(E) = 0$. If $m_\alpha^*(E) > 0$ and $\beta < \alpha$, then $m_\beta^*(E) = \infty$.*

Proof: If $\operatorname{diam} F \leq \delta$, and $\beta > \alpha$, then

$$(\operatorname{diam} F)^\beta = (\operatorname{diam} F)^{\beta-\alpha}(\operatorname{diam} F)^\alpha \leq \delta^{\beta-\alpha}(\operatorname{diam} F)^\alpha.$$

Consequently,

$$\mathcal{H}_\beta^\delta(E) \leq \delta^{\beta-\alpha}\mathcal{H}_\alpha^\delta(E) \leq \delta^{\beta-\alpha} m_\alpha^*(E).$$

Since $m_\alpha^*(E) < \infty$ and $\beta - \alpha > 0$, we find in the limit as δ tends to 0 that $m_\beta^*(E) = 0$. The contrapositive of the first statement gives the second statement. □

Remark 6.5.1 1. *If I is a finite line segment in $\mathbb{R}^d$, then $0 < m_1(I) < \infty$.*

2. *If Q is a k-rectangle in $\mathbb{R}^d$, then $0 < m_k(Q) < \infty$.*

3. *If $\mathcal{O}$ is a non-empty open set in $\mathbb{R}^d$, then $m_\alpha(\mathcal{O}) = \infty$ whenever $\alpha < d$. Indeed, because we have $m_d(\mathcal{O}) > 0$.*

6.2 Hausdorff Dimension

Definition 6.6 (Hausdorff Dimension) *Given a Borel subset E of $\mathbb{R}^d$, we know that there exists a unique $\alpha \in \mathbb{R}_{\geq 0}$ such that*

$$m_\beta(E) = \begin{cases} \infty & \text{if } \beta < \alpha \\ 0 & \text{if } \alpha < \beta. \end{cases}$$

That is

$$\alpha = \sup\{\beta : m_\beta(E) = \infty\} = \inf\{\beta : m_\beta(E) = 0\}.$$

*We say that E has **Hausdorff dimension** α and write $\dim E = \alpha$. If $0 < m_\alpha(E) < \infty$, we say that E has **strict Hausdorff dimension** α. If E has fractional dimension, it is sometimes referred to as a **fractal**.*

Definition 6.7 (Lipschitz Continuity) *A function f defined on a subset E of $\mathbb{R}^d$ satisfies a* ***Lipschtiz condition with exponent*** *γ if*

$$|f(x) - f(y)| \leq M|x-y|^\gamma$$

for any $x, y \in E$. In practice, we usually assume $0 < \gamma \leq 1$, and for the case $\gamma = 1$, we say the function satisfies the Lipschtiz condition.

Lemma 6.8 *Suppose a function f defined on a compact set E satisfies a Lipschitz condition with exponent γ. Then*

1. *$m_\beta(f(E)) < M^\beta m_\alpha(E)$ if $\beta = \frac{\alpha}{\gamma}$.*
2. *$\dim f(E) \leq \frac{1}{\gamma} \dim E$.*

Proof: Suppose $\{F_k\}$ is a countable family of sets that covers E. Then $\{f(E \cap F_k)\}$ covers $f(E)$, and moreover, $f(E \cap F_k)$ has diameter less than $M(\operatorname{diam} F_k)^\gamma$. Hence

$$\sum_k (\operatorname{diam} f(E \cap F_k))^{\alpha/\gamma} \leq M^{\alpha/\gamma} \sum_k (\operatorname{diam} F_k)^\alpha,$$

which proves part (1). Then part (2) immediately follows from part (1). □

Lemma 6.9 *The Cantor - Lebesgue function F on $\mathcal{C}$ satisfies a Lipschitz condition with exponent $\gamma = \log 2/\log 3$.*

Proof: Recall that the function F is constructed as the pointwise limit of$\{F_n\}$ of piecewise linear functions. The function F_n increases by at most 2^{-n} on each interval of length 3^{-n}. So the slope of F_n is always bounded by $(3/2)^n$, and hence

$$|F_n(x) - F_n(y)| \leq \left(\frac{3}{2}\right)^n |x-y|.$$

Moreover, the approximating sequence also satisfies $|F(x) - F_n(x)| \leq 1/2^n$. These two estimates together with an application of the triangle inequality give

$$\begin{aligned} |F(x) - F(y)| &\leq |F_n(x) - F_n(y)| + |F(x) - F_n(x)| + |F(y) - F_n(y)| \\ &\leq \left(\frac{3}{2}\right)^n |x-y| + \frac{2}{2^n}. \end{aligned}$$

Having fixed x and y, we then minimize the right hand side by choosing n so that both terms have the same order of magnitude. This is achieved by taking n so that $3^n|x-y|$ is between 1 and 3. Then, we see that

$$|F(x) - F(y)| \leq c2^{-n} = c(3^{-n})^\gamma \leq M|x-y|^\gamma,$$

since $3^\gamma = 2$ and 3^{-n} is not greater than $|x-y|$. □

More generally, we have the following lemma:

Lemma 6.10 *Suppose $\{f_j\}$ is a sequence of continuous functions on the interval $[0,1]$ that satisfy*

$$|f_j(t) - f_j(s)| \leq A^j |t-s| \quad \text{for some } A > 1,$$

and

$$|f_j(t) - f_{j+1}(t)| \le B^{-j} \quad \textit{for some } B > 1.$$

Then the limit $f(t) = \lim_{j\to\infty} f_j(t)$ *exists and satisfies*

$$|f(t) - f(s)| \le M|t-s|^\gamma,$$

where $\gamma = \log B/\log(AB)$.

Proof: The continuous limit f is given by the uniformly convergent series

$$f(t) = f_1(t) + \sum_{k=1}^{\infty}(f_{k+1}(t) - f_k(t)),$$

and therefore

$$|f(t) - f_j(t)| \le \sum_{k=j}^{\infty} |f_{k+1}(t) - f_k(t)| \le \sum_{k=j}^{\infty} B^{-k} \le cB^{-j}.$$

The triangle inequality, an application of the inequality just obtained, and the inequality in the statement of the lemma give

$$\begin{aligned}|f(t) - f(s)| &\le |f_j(t) - f_j(s)| + |(f-f_j)(t)| + |(f-f_j)(s)| \\ &\le c(A^j|t-s| + B^{-j}).\end{aligned}$$

For a fixed pair of numbers t and s with $t \ne s$, we choose j to minimize the sum $A^j|t-s| + B^{-j}$. This is essentially achieved by picking j so that two terms $A^j|t-s|$ and B^{-j} are comparable. More precisely, we choose a j that satisfies

$$(AB)^j|t-s| \le 1 \quad \text{and} \quad 1 \le (AB)^{j+1}|t-s|.$$

Since $|t-s| \le 2$ and $AB > 1$, such a j must exist. The first inequality then gives

$$A^j|t-s| \le B^{-j},$$

while raising the second inequality to the power γ, and using the fact that $(AB)^\gamma = B$ gives

$$1 \le B^j|t-s|^\gamma.$$

Thus $B^{-j} \le |t-s|^\gamma$, and consequently

$$|f(t) - f(s)| \le c(A^j|t-s| + B^{-j}) \le M|t-s|^\gamma,$$

as was to be shown. □

Theorem 6.11 *The Cantor set* $\mathcal{C}$ *has strict Hausdorff dimension* $\alpha = \log 2/\log 3$.

Proof: The inequality

$$m_\alpha(\mathcal{C}) \le 1$$

follows from the construction of $\mathcal{C}$ and the definitions. Indeed, recall from Chapter 1 that $\mathcal{C} = \bigcap C_k$, where each C_k is a finite union of 2^k intervals of length 3^{-k}. Given $\delta > 0$, we first choose K so large that $3^{-K} < \delta$. Since the set C_K covers $\mathcal{C}$ and consists of 2^K intervals of diameter $3^{-K} < \delta$, we must have

$$\mathcal{H}_\alpha^\delta(\mathcal{C}) \leq 2^K (3^{-K})^\alpha.$$

However, α satisfies precisely $3^\alpha = 2$, hence $2^K(3^{-K})^\alpha = 1$, and therefore $m_\alpha(\mathcal{C}) \leq 1$.

The reverse inequality, which consists of proving that $0 < m_\alpha(\mathcal{C})$. Take f as the Cantor-Lebesgue function, and $E = \mathcal{C}$ and $\alpha = \gamma = \log 2/\log 3$. Then by Lemma (6.8) and Lemma (6.9), we conclude that

$$0 < m_1([0,1]) \leq M^\beta m_\alpha(\mathcal{C}).$$

Thus $\dim \mathcal{C} = \log 2/\log 3$. $\square$

Similarly, we can prove the following interesting results:

Theorem 6.12 *Suppose the curve γ is continuous and quasi-simple. Then γ is rectifiable if and only if $\Gamma = \{\gamma(t) : \alpha \leq t \leq b\}$ has strict Hausdorff dimension one. Moreover, in this case the length of the curve is precisely its one-dimensional measure $m_1(\Gamma)$.*

Theorem 6.13 *The Sierpinski Triangle $\mathcal{S}$ has strict Hasudorff dimension $\alpha = \log 3/\log 2$.*

Theorem 6.14 *The von Koch Curve has a strict dimension of $\alpha = \log 4/\log 3$.*

Definition 6.15 (Self-Similarity) *A mapping $S : \mathbb{R}^d \to \mathbb{R}^d$ is said to be a **similarity** with **ratio** $r > 0$ if*

$$|S(x) - S(y)| = r|x - y|.$$

*Given finitely many similarities $S_1, \ldots, S_m$ with the same ratio r, we say that the set $F \subset \mathbb{R}^d$ is **self-similar** if*

$$F = S_1(F) \cup \cdots \cup S_m(F).$$

We point out the relevance of the various examples we have already seen.

Remark 6.15.1 *It can be shown that every similarity of $\mathbb{R}^d$ is the composition of a translation, a rotation, and a dilation by r.*

Lemma 6.16 *Let $S_1, \cdots, S_m$ be similarities with ratio r on $\mathbb{R}^d$. Then there exists a closed ball B so that $S_j(B) \subset B$ for all $j = 1, \cdots, m$.*

Proof: Indeed, we note that if S is a similarity with ratio r, then

$$\begin{aligned} |S(x)| &\leq |S(x) - S(0)| + |S(0)| \\ &\leq r|x| + |S(0)|. \end{aligned}$$

If we require that $|x| \leq R$ implies $|S(x)| \leq R$, it suffices to choose R so that $rR + |S(0)| \leq R$, that is, $R \geq |S(0)|/(1-r)$. In this fashion, we obtain for each S_j a ball B_j centered at the origin that satisfies $S_j(B_j) \subset B_j$. If B denotes the ball among the B_j with the largest radius, then the above shows that $S_j(B) \subset B$ for all j. $\square$

Definition 6.17 (Hasudorff Distance) *Let A and B be two compact sets, we define the* ***Hausdorff distance*** *as*

$$\mathrm{dist}(A, B) = \inf\{\delta \,:\, B \subset A^\delta \;\; and \;\; A \subset B^\delta\}.$$

The following is easy to verify for Hasudorff distances:

Lemma 6.18 $\mathrm{dist}(\cdot,\cdot)$ *is a metric on the space of all compact subsets of $\mathbb{R}^d$. Moreover, if $S_1, \cdots, S_m$ are similarities with the same ratio r, then let $\tilde{S}(A)$ denote the set*

$$\tilde{S}(A) = S_1(A) \cup \cdots \cup S_m(A),$$

, we have

$$\mathrm{dist}(\tilde{S}(A), \tilde{S}(B)) \leq r\mathrm{dist}(A, B).$$

Remark 6.18.1 *Since we see that each S_j is continuous, they map compact sets to compact sets, so* $\mathrm{dist}(\tilde{S}(A), \tilde{S}(B))$ *is well-defined. Another observation is that, if A contains more than one point, then so does $\tilde{S}(A)$.*

Theorem 6.19 *Suppose $S_1, S_2, \cdots, S_m$ are m similarities, each with same ratio r that satisfies* $0 < r < 1$. *Then there exists a unique nonempty set F such that*

$$F = S_1(F) \cup \cdots \cup S_m(F).$$

Proof: We first choose a big enough closed ball B such that $\tilde{S}(B) \subset B$, this is possible by Lemma (6.16). Set $F_k = \tilde{S}^k(B)$, then each F_k is compact, non-empty, and $F_k \subset F_{k-1}$, since $\tilde{S}(B) \subset B$. If we let

$$F = \bigcap_{k=1}^{\infty} F_k,$$

then F is compact, non-empty and clearly $\tilde{S}(F) = F$.

We prove the uniqueness of F. Suppose G is another compact set so that $\tilde{S}(G) = G$. Then, by Lemma 6.18, we have that

$$\mathrm{dist}(\tilde{S}(F), \tilde{S}(G)) = \mathrm{dist}(F, G) \leq r\mathrm{dist}(F, G).$$

Since $r < 1$, this forces $\mathrm{dist}(F, G) = 0$, so $F = G$. $\square$

Under an additional technical condition, one can calculate the precise Hausdorff dimension of the self - similar set F. Loosely speaking, the restriction holds if the sets $S_1(F), \ldots, S_m(F)$ do not overlap too much. Indeed, if these sets were disjoint, then we could argue that

$$m_\alpha(F) = \sum_{j=1}^{m} m_\alpha(S_j(F)).$$

Since each S_j scales by r, we would then have $m_\alpha(S_j(F)) = r^\alpha m_\alpha(F)$. Hence

$$m_\alpha(F) = mr^\alpha m_\alpha(F).$$

If $m_\alpha(F)$ were finite, then we would have that $mr^\alpha = 1$; thus

$$\alpha = \frac{\log m}{\log 1/r}.$$

Definition 6.20 (Separated Similarities) *We say that the similarities $S_1, \cdots, S_m$ are **separated** if there is an bounded open set $\mathcal{O}$ so that*

$$\mathcal{O} \supset S_1(\mathcal{O}) \cup \cdots \cup S_m(\mathcal{O}),$$

and the $S_j(\mathcal{O})$ are disjoint. It is not assumed that $\mathcal{O}$ contains the set F given by Theorem (6.19).

Theorem 6.21 *Suppose $S_1, S_2, \cdots, S_m$ are m separated similarities with the common ratio r that satisfies $0 < r < 1$. Then the set F has Hausdorff dimension equal to $\log m/\log(1/r)$.*

Proof: If $\alpha = \log m/\log(1/r)$, we claim that $m_\alpha(F) < \infty$, hence $\dim F \leq \alpha$. In fact, this inequality holds without the separation assumption.

Let $F_k = \tilde{S}^k(B)$, and $\tilde{S}^k(B)$ is the union of m^k sets of diameter less tan cr^k, where $c = \operatorname{diam} B$, each of the form

$$S_{n_1} \circ S_{n_2} \circ \cdots \circ S_{n_k}(B), \quad \text{where } 1 \leq n_i \leq m \text{ and } 1 \leq i \leq k.$$

Consequently, if $cr^k \leq \delta$, then

$$\begin{aligned} \mathcal{H}_\alpha^\delta(F) &\leq \sum_{n_1, \cdots, n_k} (\operatorname{diam} S_{n_1} \circ \cdots \circ S_{n_k}(B))^\alpha \\ &\leq c' m^k r^{\alpha k} \end{aligned}$$

since $\alpha = \log m/\log(1/r)$. Since c' is independent of δ, we get $m_\alpha(F) \leq c'$.

To prove $m_\alpha(F) > 0$, we need the separation condition. Fix a point $\bar{x}$ in F, we define the "vertices" of the k^{th} generation as the m^k points that lie in F and are given by

$$S_{n_1} \circ \cdots \circ S_{n_k}(\bar{x}), \quad \text{where } 1 \leq n_1 \leq m, \cdots, 1 \leq n_k \leq m.$$

Each vertex is labelled by $(n_1, \cdots, n_k)$. Vertices need not be distinct, so they are counted with their multiplicities.

Similarly, we define the "open sets" of the k^{th} generation to be the m^k sets given by

$$S_{n_1} \circ \cdots \circ S_{n_k}(\mathcal{O}), \quad \text{where } 1 \leq n_1 \leq m, \cdots, 1 \leq n_k \leq m,$$

and where $\mathcal{O}$ is fixed and chosen to satisfy the separation condition. Such open sets are again labelled by multi-indices $(n_1, n_2, \cdots, n_k)$ with $1 \leq n_j \leq m$, $1 \leq j \leq k$.

Then the open sets of the k^{th} generations are disjoint, since those of the first generation are disjoint. Moreover, if $k \geq \ell$, each open set of the ℓ^{th} generation contains $m^{k-\ell}$ open sets of the k^{th} generation.

Suppose v is a vertex of the k^{th} generation, and let $\mathcal{O}(v)$ denote the open set in the k^{th} generation which is associated to v, that is v and $\mathcal{O}(v)$ carry the same label $(n_1, n_2, \cdots, n_k)$. Since $\bar{x}$ is at a fixed distance from the original open set $\mathcal{O}$, and $\mathcal{O}$ has a finite diameter, we find that

1. $d(v, \mathcal{O}(v)) \leq cr^k$.
2. $c'r^k \leq \operatorname{diam} \mathcal{O}(v) \leq cr^k$.

It suffices to prove that if $\mathcal{B} = \{B_j\}_{j=1}^N$ is a finite collection of balls whose diameters are less than δ and whose union covers F, then

$$\sum_{j=1}^{N} (\operatorname{diam} B_j)^\alpha \geq c > 0.$$

Suppose we have such a covering by balls, and choose k so that

$$r^k \leq \min_{1 \leq j \leq N} \operatorname{diam} B_j < r^{k-1}.$$

Let N_ℓ denote the number of balls in $\mathcal{B}$ so that

$$r^\ell \leq \operatorname{diam} B_j \leq r^{\ell-1}.$$

By lemma (6.22), We see that the total number of vertices of the k^{th} generation that can be covered by the collection $\mathcal{B}$ can be no more than $c \sum_\ell N_\ell m^{k-\ell}$. Since all m^k vertices of the k^{th} generation belong to F, we must have $c \sum_\ell N_\ell m^{k-\ell} \geq m^k$, and hence

$$\sum_\ell N_\ell m^{-\ell} \geq c.$$

The definition of α gives $r^{\ell\alpha} = m^{-\ell}$, and therefore

$$\sum_{j=1}^{N} (\operatorname{diam} B_j)^\alpha \geq \sum_\ell N_\ell r^{\ell\alpha} \geq c.$$

□

Lemma 6.22 *Suppose B is a ball in the covering $\mathcal{B}$ that satisfies*

$$r^\ell \leq \operatorname{diam} B < r^{\ell-1} \quad \text{for some } \ell \leq k.$$

Then B contains at most $cm^{k-\ell}$ vertices of the k^{th} generation.

Proof: If v is a vertex of the k^{th} generation with $v \in B$, and $\mathcal{O}(v)$ denotes the corresponding open set of the k^{th} generation, then, for some fixed dilate B^* of B, then since $d(v, \mathcal{O}(v)) \leq cr^k$ and $c'r^k \leq \operatorname{diam} \mathcal{O}(v) \leq cr^k$, we have $\mathcal{O}(v) \subset B^*$, and B^* also contains the open set of generations ℓ that contains $\mathcal{O}(v)$.

Since B^* has volume $cr^{d\ell}$, and each open set in the ℓ^{th} generation has volume comparable to $r^{d\ell}$. B^* can contain at most c open sets of generation ℓ. Hence B^* contains at most $cm^{k-\ell}$ open sets of the k^{th} generation. Consequently, B can contain at most $cm^{k-\ell}$ vertices of the k^{th} generation. $\square$

7 Appendix - Abstract Measure Theory

7.1 "Total" measure and why it does not work

Convention: we define some arithmetic rules for $\pm\infty$,

$$\begin{aligned}
\infty + \infty = x + \infty = \infty + x = \infty, \quad \forall x \in \mathbb{R} \\
(-\infty) + (-\infty) = x + (-\infty) = (-\infty) + x = -\infty, \quad \forall x \in \mathbb{R} \\
\infty \cdot \infty = \infty, \quad (-\infty) \cdot \infty = -\infty \\
(-\infty) \cdot (\infty) = -\infty, \quad (-\infty) \cdot (-\infty) = \infty \\
x \cdot \infty = \infty \cdot x = \begin{cases} \infty, & x > 0 \\ -\infty, & x < 0 \\ 0, & x = 0 \end{cases}, \quad x \in \mathbb{R}.
\end{aligned}$$

Definition 7.1 *A function $\mu : \mathcal{P}(\Omega) \to \bar{\mathbb{R}}$ is a* ***"total" measure*** *on Ω if it satisfies the following:*

1. *$\mu(\emptyset) = 0$;*
2. *$\mu(X) \geq 0$ for all $x \in \mathcal{P}(\Omega)$;*
3. *μ is countably additive in the sense that if $(X_n)_{n=0}^{\infty}$ is a sequence of disjoint subset of Ω, then*

$$\mu\left(\bigcup_{n=0}^{\infty} X_n\right) = \sum_{n=0}^{\infty} \mu(X_n) = \lim_{N\to\infty} \sum_{n=0}^{N} \mu(X_n).$$

Examples:

1. $\mu(X) = 0$ for all $X \subset \Omega$.
2. The counting measure.
3. The Dirac measure.

Theorem 7.2 (Vitali) *There is no total measure that is translational invariant (or simply invariant) on $\mathbb{R}$ such that $\mu([0,1]) = 1$.*

Remark 7.2.1 *It is for this reason, we cannot define a measure on all subsets of $\mathbb{R}$ which extends our intuitive notion of "volume".*

Proof: See Lemma 1.11. Countable additivity fails. □

7.2 Sigma-algebra

Definition 7.3 (σ-algebra) *A subcollection $\mathscr{A}$ of $\mathcal{P}(\Omega)$ is a **σ-algebra** if the following conditions are satisfied:*

1. $\emptyset, \Omega \in \mathscr{A}$;
2. *If $X, Y \in \mathscr{A}$, then $X - Y \in \mathscr{A}$.*
3. *If $(X_n)_{n=1}^{\infty}$ is a sequence of disjoint members of $\mathscr{A}$, then $\bigcup_{n=1}^{\infty} X_n \in \mathscr{A}$.*

If condition (3) is replaced by a weak one:

3'. *If X_1 and X_2 are disjoint elements of $\mathscr{A}$, then $X_1 \cup X_2 \in \mathscr{A}$.*

*then $\mathscr{A}$ is known as a **Boolean algebra**.*

Remark 7.3.1 *Note that there are equivalent definitions of this definition.*

Remark 7.3.2 *From the definition, it is clear that any finite Boolean algebra is a σ-algebra. Also any σ-algebra is closed under countable intersection*

Lemma 7.4 *The arbitrary intersection of σ-algebras is a σ-algebra.Then if $\mathscr{C} \subseteq \mathcal{P}(\Omega)$, there is a smallest σ-algebra containing $\mathscr{C}$, this is called the **σ-algebra generated by** $\mathscr{C}$ and denoted by $\sigma(\mathscr{C})$.*

Definition 7.5 (Borel σ-algebra) *Let $(X, \mathcal{T})$ be a topological space, then we define the **Borel σ-algebra** on X to be the the smallest σ-algebra that contains all open sets, i.e., it is the σ-algebra generated by $\mathcal{T}$ and denoted $\mathcal{B}(T)$. Elements of this σ-algebra are called **Borel sets**, or we say the element is **Borel**.*

7.3 Measurable Spaces and Measurable Functions

Definition 7.6 (Premeasure) *A **premeasure** μ is an extended real valued function on an algebra $\mathscr{A}$, satisfying*

1. $\mu(\emptyset) = 0$;
2. $\mu(X) \geq 0$ *for all $x \in \mathscr{A}$;*
3. *If $\{x_n\}$ is a sequence of disjoint sets in $\mathscr{A}$, such that $\bigsqcup_{n=1}^{\infty} X_n \in \mathscr{A}$, then*

$$\mu\left(\bigcup_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} \mu(x_n).$$

Definition 7.7 (Measurable Space) *Suppose $\mathscr{A}$ is an σ-algebra. We say ϕ is a **set function** defined on $\mathscr{A}$ if ϕ assigns every $A \in \mathscr{A}$ a number $\phi(A)$ of the extended real number system.*

*Suppose Ω is an arbitrary set, Ω is said to be a **measure space** if there exists a $\sigma-$algebra $\mathscr{A}$ on Ω (which are called **measurable sets**) and a non-negative (can take infinity) countably additive set function μ (which is called a **measure**), defined on $\mathscr{A}$. If $S \in \mathscr{A}$, then we say S is μ **measurable**. If $\mathscr{A}$ is a σ-algebra on Ω, then we say $(\Omega, \mathscr{A})$ is a **measurable space**. Commonly, for a measure space, we will use $(\Omega, \mathscr{A}, \mu)$ or $(X, \mathcal{S}, \mu)$ to* *denote the space.*

Example:

1. Any total measure is a premeasure on $\mathcal{P}(\Omega)$.
2. We can define a premeasure on the algebra that contains all open rectangles in $\mathbb{R}^d$ by the function. We will extend this to a unique measure on $\mathcal{B}(\mathbb{R}^d)$, using the Caratheodory extension theorem. In addition, we will also construct this measure explicitly.

Lemma 7.8 *Let $(\Omega, \mathscr{A}, \mu)$ is a measurable space. If*

- *$X, Y \in \mathscr{A}$, $X \subset Y$, then $\mu(X) \leq \mu(Y)$;*
- *Furthermore, if $\mu(X) < \infty$, then $\mu(Y \setminus X) = \mu(Y) - \mu(X)$.*

Definition 7.9 (Measurable functions) *Let f be a function defined on a measurable space Ω, with values in the extended real number system. Then the function f is said to be **measurable** if the set*

$$\{x \in X \,:\, f(x) > a\}$$

is measurable for every real number a.

Proposition 7.10 *Each of the following five conditions implies the other four, i.e., the five statements are equivalent:*

1. *$\{x|f(x) > a\}$ is measurable for every real a.*
2. *$\{x|f(x) \geq a\}$ is measurable for every real a.*
3. *$\{x|f(x) < a\}$ is measurable for every real a.*
4. *$\{x|f(x) \leq a\}$ is measurable for every real a.*
5. *If further $f : X \to \mathbb{R}$, then $\{x|a < f(x) < b\}$ is measurable for every real numbers a, b.*

Proof: The relations

$$\begin{aligned}
\{x|f(x) \geq a\} &= \bigcap_{n=1}^{\infty} \left\{x|f(x) > a - \frac{1}{n}\right\}, \\
\{x|f(x) < a\} &= X - \{x|f(x) \geq a\}, \\
\{x|f(x) \leq a\} &= \bigcap_{n=1}^{\infty} \left\{x|f(x) < a + \frac{1}{n}\right\}, \\
\{x|f(x) > a\} &= X - \{x|f(x) \leq a\}
\end{aligned}$$

shows successively that (1) implies (2) implies (3) implies (4) implies (1).
Now since (2) implies (4), then (2) would imply (5). Lastly, we show that (5) implies (1), this is clear as

$$\{x|f(x) \geq a\} = \bigcup_{n=1}^{\infty} \{x|a \leq f(x) \leq a+n\}.$$

As we imposed the condition that $f(x) \neq \infty$. □

Remark 7.10.1 *From the proof, we can see that the additional assumption in condition* (5) *which says f only takes real values can be removed if we know that $f^{-1}(\infty)$ and $f^{-1}(\infty)$ are measurable.*

Notation: we denote $L_f^{\geq}(\alpha)$ as the set $\{x \in \Omega : f(x) \geq \alpha\}$. And similarly, we can define the "super level set" of other types.

Corollary 7.10.1 *The finite-valued function f is measurable if and only if $f^{-1}(\mathcal{O})$ is measurable for every open set $\mathcal{O}$, and if and only if $f^{-1}(F)$ is measurable for every closed set F. Thus the finite-valued function f is measurable if and only if $f^{-1}(B)$ is measurable for every Borel set in $\mathbb{R}$.*

Remark 7.10.2 *Of course, if we make the additional hypothesis that both $f^{-1}(\infty)$ and $f^{-1}(-\infty)$ are measurable sets, then we can again drop the finite-valued assumption.*

Proof: Note every open set $\mathcal{O} \subset \mathbb{R}$ can be written as countable union of open intervals and every closed set is the complement of an open set. Lastly do an induction on the complexity of Borel sets, we get the final statement. □

Corollary 7.10.2 *If f is measurable, then $|f|$ is measurable.*

Proof: Since $\{x||f(x)| < a\} = \{x|f(x) < a\} \cap \{x|f(x) > -a\}$, then $|f|$ is measurable if f is measurable. □

Proposition 7.11 *Let $\{f_n\}$ be a sequence of measurable functions from X to $\mathbb{R} \cup \{-\infty, \infty\}$. For $x \in X$, put*

$$g(x) = \sup f_n(x) \quad (n = 1, 2, 3, \cdots),$$
$$h(x) = \limsup_{n \to \infty} f_n(x).$$

Then g and h are measurable. The same of holds if we replace $\sup$ *and* $\limsup$ *with* $\inf$ *and* $\liminf$*.*

Proof: The proposition follows from the following observation:

$$\{x|g(x) > a\} = \bigcup_{n=1}^{\infty} \{x|f_n(x) > a\},$$
$$h(x) = \inf g_m(x),$$

where $g_m(x) = \sup f_n(x)\, (n \geq m)$. □

Corollary 7.11.1

1. *If f and g are measurable, then max (f, g) and min (f, g) are measurable. In particular, if*

$$f^+ = \max(f, 0), \quad f^- = -\min(f, 0),$$

it follows that f^+ and f^- are measurable.

2. *The limit of a convergent sequence of measurable functions is measurable. The is, if $\{f_n\}$ is a sequence of measurable functions from $X \to \mathbb{R} \cup \{-\infty, \infty\}$, then $\lim_{n\to\infty} f_n$ is measurable.*

Proof:

1. Consider the sequence consisting of f and g, then $\sup(f, g) = \max(f, g)$ and $\inf(f, g) = \min(f, g)$.

2. This is clear, as if $\lim f_n$ exists, then $\lim f_n = \limsup f_n = \liminf f_n$.

$\square$

Lemma 7.12 *If $f : X \to \mathbb{R} \cup \{-\infty, \infty\}$ is measurable, then f^k for $k \in \mathbb{Z}^+$ is measurable.*

Proof: If k is odd, then $\{f^k > a\} = \{f > a^{1/k}\}$ and if k is even and $a \geq 0$, then $\{f^k > a\} = \{f > a^{1/k}\} \cup \{f < -a^{1/k}\}$. $\square$

Definition 7.13 (Almost everywhere) *Suppose f, g are two functions defined on a measurable subset E of the measurable space $(\Omega, \mathscr{A}, \mu)$, then we say that f and g are equal* ***almost everywhere****, and write*

$$f(x) = g(x), \quad a.e.$$

if the set $\{x \in E : f(x) \neq g(x)\}$ has measure zero.

Proposition 7.14 *Suppose f is measurable, and $f(x) = g(x)$ a.e. Then g is measurable.*

Proof: Since $\{f > a\}$ and $\{g > a\}$ differs only by a set of measure zero, so either both are measurable or both are not. $\square$

Remark 7.14.1 *This allows us to replace all the results before with a weaker assumption if the measure we considering is* ***complete*** *(See next subsection), that is the assumption only holds almost everywhere on the domain of the function.*

7.4 Carathéodory Extension Theorem

Definition 7.15 (Finite and σ-finite measures) *Suppose $(\Omega, \mathscr{A}, \mu)$ is triple such that $\mathscr{A}$ is a algebra and μ is a premeasure defined on members of $\mathscr{A}$. We say that μ is **finite** if $\mu(\Omega) < \infty$. We say μ is σ-**finite** if $\Omega = \bigcup_n^\infty \Omega_n$, with each $\Omega_n \in \mathscr{A}$, and $\mu(\Omega_n) < \infty$.*

Remark 7.15.1 *It is clear that all finite measures are σ-finite.*

Theorem 7.16 (Carathéodory Extension Theorem) *Suppose $(\Omega, \mathscr{A}, \mu)$ is a triple where $\mathscr{A}$ is an algebra, μ is a premeasure defined on members of $\mathscr{A}$. Then μ can be extended to a measure μ' on the $\sigma-$algebra $\sigma(\mathscr{A})$. Moreover, if μ is σ-finite, then μ' is the unique such extension.*

The proof of the extension theorem has 2 steps:

- Step 1: we construct an outer measure from μ.
- Step 2: This outer measure can be restricted to a measure.

We will use this entire chapter to prove the existence and the uniqueness of this Extension Theorem.

Definition 7.17 (Outer Measure) *Let Ω be a domain, $\mu^* : \mathcal{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is a (total) **outer measure**, if*

1. *$\mu^*(\emptyset) = 0$.*
2. *If $X \subseteq Y \subseteq \Omega$, then $0 \le \mu^*(X) \le \mu^*(Y)$.*
3. *(Countably Subadditivity) Suppose $A = \bigcup A_n$, where $A_n \in \mathcal{P}(\Omega)$, then*

$$\mu^*(A) \le \sum_{n=1}^{\infty} \mu^*(A_n).$$

Lemma 7.18 *Suppose μ is a premeasure defined on some algebra $\mathscr{A}$ of subsets of Ω. Then*

$$\mu^*(X) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) \,|\, (A_n)_{n=1}^{\infty} \subset \mathscr{A} \text{ that covers } X \right\}$$

is an outer measure. Moreover, μ^* extends μ, i.e., for all $A \subseteq \mathscr{A}$, $\mu^*(A) = \mu(A)$.

Proof: One can clearly see that μ^* satisfies property 1 and 2 in the definition of an outer measure. To verify property 3, let $X = \bigcup_n X_n$, where $X_n \in \mathcal{P}(\Omega)$. We show $\mu^*(X) \le \epsilon + \sum_{n=1}^{\infty} \mu^*(X_n)$ for any $\epsilon > 0$. By the definition of $\mu^*(X_n)$, $\exists (A_{n,k}) \subset \mathscr{A}$ which cover X_n, and

$$\sum_{k=1}^{\infty} \mu(A_{n,k}) \le \mu^*(X_n) + \frac{\epsilon}{2^n}.$$

Then we have

$$\mu^*(X) \le \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(A_{n,k}) \le \sum_{n=1}^{\infty} \left(\mu^*(X_n) + \frac{\epsilon}{2^n} \right) \le \sum_{n=1}^{\infty} \mu^*(X_n) + \epsilon.$$

Lastly, it is clear that $\mu^*(X) \leq \mu(X)$ for any $X \in \mathscr{A}$. On the other hand, let (A_n) be an $\mathscr{A}$-cover of X. We define $B_0 = A_0 \cap X$ and, for $n \geq 0$,

$$B_{n+1} = (A_{n+1} \cap X) \setminus \bigcup_{k \leq n} (A_k \cap X) .$$

Then clearly $B_n \in \mathscr{A}$, the B_n's are disjoint with $\bigsqcup_n B_n = X$ and $\mu(B_n) \leq \mu(A_n)$. Since μ is a premeasure on $\mathcal{A}$ we have $\mu(X) = \sum_{n=1}^{\infty} \mu(B_n)$ which is less than or equal to $\sum_{n=1}^{\infty} \mu(A_n)$. Since this holds for any $\mathscr{A}$-cover of X we have $\mu(X) \leq \mu^*(X)$. □

Definition 7.19 (Carathéodory Splitting Condition) *We say $S \subseteq \Omega$ satisfies the Carathéodory splitting condition (with respect to the outer measure μ^*), if $\mu^*(X) = \mu^*(X \cap S) + \mu^*(X \cap S^c)$ for all $X \subseteq \Omega$, where $S^c = \Omega \setminus S$.*

Lemma 7.20 *Suppose $\mathscr{A}$ is an algebra of the subsets of Ω, and μ is a premeasure on $\mathscr{A}$. Let μ^* be the outer measure defined by*

$$\mu^*(X) = \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) \,|\, (A_n)_{n=1}^{\infty} \subset \mathscr{A} \text{ that covers } X\right\}.$$

Then for each $A \in \mathscr{A}$, A satisfies the Carathéodory splitting condition.

Proof: We show that for any $A \in \mathscr{A}$ and $X \subseteq \Omega$,

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

By subadditivity, we have $\mu^*(X) \leq \mu^*(X \cap A) + \mu^*(X \cap A^c)$. For the converse direction,, if $\mu^*(X) = \infty$, then it is clear that at least one of the summand on the right must be ∞, otherwise, we can find a cover for X whose premeasure sums to a finite number. So we assume that $\mu^*(X) < \infty$. It is enough to show that for all $\epsilon > 0$,

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

By the definition of μ^*, we get a sequence $(A_n)_{n=1}^{\infty}$ in $\mathscr{A}$ such that $X \subseteq \bigcup_{n=1}^{\infty} A_n$ and

$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(X) + \epsilon.$$

Note that both $A_n \cap A$ and $A_n \cap A^c$ are still in $\mathscr{A}$, and

$$\mu(A_n) = \mu(A_n \cap A) + \mu(A_n \cap A^c)$$

with $(A_n \cap A)_{n=1}^{\infty}$ is an $\mathscr{A}$-cover of $X \cap A$ and $(A_n \cap A^c)$ is an $\mathscr{A}$-cover of $X \cap A^c$. Thus the desired result easily follows. □

Proposition 7.21 *Suppose μ^* is an arbitrary outer measure on Ω. Let $\mathcal{S} \subset \mathcal{P}(\Omega)$ be the collection of all sets that satisfies the Carathéodory splitting condition with respect to μ^*.Then $\mathcal{S}$ is a σ-algebra that contains $\mathscr{A}$. Moreover, $\mu^*|_{\mathcal{S}}$ is a measure.*

Proof: It is clear that $\emptyset$ and Ω are elements of $\mathcal{S}$ and $\mathcal{S}$ is closed under complement.

We will first prove that $\mathcal{S}$ is closed under finite intersections and unions.

Let $S_1, S_2 \in \mathcal{S}$ and let X be any subset of Ω. We have $X \cap S_1^c = X \cap (S_1 \cap S_2)^c \cap S_1^c$ since $(S_1 \cap S_2)^c \supseteq S_1^c$. On the other hand we have $(S_1 \cap S_2)^c = S_1^c \cup S_2^c$ and hence

$$X \cap (S_1 \cap S_2)^c \cap S_1 = (X \cap S_1 \cap S_2^c) \cup (X \cap S_1 \cap S_1^c) = X \cap S_1 \cap S_2^c.$$

Therefore, we have

$$\begin{aligned}\mu^* \left(X \cap (S_1 \cap S_2)^c\right) &= \mu^* \left(X \cap (S_1 \cap S_2)^c \cap S_1\right) + \mu^* \left(X \cap (S_1 \cap S_2)^c \cap S_1^c\right) \\ &= \mu^* \left(X \cap S_1^c\right) + \mu^* \left(X \cap S_1 \cap S_2^c\right).\end{aligned}$$

Note that the first equality follows from applying the Carathéodory's Splitting condition for $X \cap (S_1 \cap S_2)^c$. Now adding $\mu^*(X \cap S_1 \cap S_2)$ and noting that

$$\mu^*(X \cap S_1) = \mu^*(X \cap S_1 \cap S_2) + \mu^* \left(X \cap S_1 \cap S_2^c\right),$$

one obtains $\mu^*(X \cap S_1 \cap S_1) + \mu^* \left(X \cap (S_1 \cap S_2)^c\right) = \mu^*(X)$. Hence, we get $S_1 \cap S_2 \in \mathcal{S}$.
Since $S_1 \cup S_2 = (S_1^c \cap S_2^c)^c$ and $S_1 \backslash S_2 = S_1 \cap S_2^c$ we see that $\mathcal{S}$ is closed under finite unions and the set-theoretic difference.

Next we claim that if $(S_n)_{n \leq N}$, $n \in \mathbb{N}_{\geq 0}$ is a disjoint finite sequence in $\mathcal{S}$, then

$$\mu^* \left(X \cap \left(\bigsqcup_{n \leq N} S_n\right)\right) = \sum_{n \leq N} \mu^*(X \cap S_n).$$

which in particular implies that μ^* is finitely additive when restricted to $\mathcal{S}$.
By induction, it is enough to show in the case where $N = 1$. One has

$$\begin{aligned}\mu^*(X \cap (S_0 \sqcup S_1)) &= \mu^*((X \cap (S_0 \sqcup S_1)) \cap S_0) + \mu^*((X \cap (S_0 \sqcup S_1)) \cap S_0^c) \\ &= \mu^*((X \cap S_0) + \mu^*((X \cap S_1).\end{aligned}$$

Note that the second equality holds because $S_1 \subseteq S_0^c$ coming from the fact that S_0 and S_1 are disjoint. This completes the proof of the claim.

Now we fix a sequence (S_n) in $\mathcal{S}$ which the S_n are pairwise disjoint and we denote the union $\bigsqcup_n S_n$ by S. We will simultaneously show that $S \in \mathcal{S}$ and $\mu^*(S) = \sum_n \mu^*(S_n)$. This implies that $\mathcal{S}$ is a σ-algebra and $\mu^* |_{\mathcal{S}}$ is a measure because we already know that $\mathcal{S}$ is an algebra, $\mu^*(\emptyset) = 0$, and $\mu^*(X) \geq 0$ for all $X \in \mathcal{S}$.

By countable subadditivity, we have $\mu^*(X \cap S) \leq \sum_n \mu^*(X \cap S_n)$. Let N be a large natural number. Since

$X \cap S^c \subseteq X \cap \left(\bigsqcup_{n\le N} S_n\right)^c$ and $\mathcal{S}$ is closed under finite unions we have

$$\mu^*(X \cap S^c) + \mu^*\left(X \cap \left(\bigsqcup_{n\le N} S_n\right)\right) \le \mu^*\left(X \cap \left(\bigsqcup_{n\le N} S_n\right)^c\right) + \mu^*\left(X \cap \left(\bigsqcup_{n\le N} S_n\right)\right) = \mu^*(X).$$

Recall that $\mu^*\left(X \cap \left(\bigsqcup_{i=0}^n S_i\right)\right) = \sum_{i=0}^n \mu^*(X \cap S_i)$ by the earlier claim. So

$$\mu^*(X \cap S^c) + \sum_{n=0}^{N} \mu^*(X \cap S_n) \le \mu^*(X).$$

But N is arbitrary in this equation and we can safely let it go to ∞ and obtain

$$\mu^*(X \cap S^c) + \sum_{n} \mu^*(X \cap S_n) \le \mu^*(X).$$

On the other hand, by countably subadditivity we have

$$\mu^*(X) \le \mu^*\left(X \cap S^c\right) + \mu^*(X \cap S) \le \mu^*(X \cap S^c) + \sum_{n} \mu^*(X \cap S_n).$$

Thus, the equality must happen everywhere. In particular, $\mu^*(X) = \mu^*\left(X \cap S^c\right) + \mu^*\left(X \cap S\right)$, implying $S \in \mathcal{S}$. Moreover, applying this in the special case with $X = S$, we recover $\mu^*(S) = \sum_n \mu^*(S_n)$ as desired. $\square$

Corollary 7.21.1 *Suppose $\mathscr{A}$ is an algebra of the subsets of Ω, and μ is a premeasure on $\mathscr{A}$. Let μ^* be the outer measure defined by*

$$\mu^*(X) = \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) \,|\, (A_n)_{n=1}^{\infty} \subset \mathscr{A} \text{ that covers } X\right\}.$$

Then $\sigma(\mathscr{A}) \subseteq \mathcal{S}$ and $\mu^*|_{\sigma(\mathscr{A})}$ is a measure.

Remark 7.21.1 *With this we can show the existence part of the Carathéodory's extension theorem. Let $\nu := \mu^*|_{\mathcal{S}}$, we know it is a measure on $\mathcal{S}$. Then since $\sigma(\mathscr{A}) \subseteq \mathcal{S}$, and $\mu = \nu|_{\sigma(\mathscr{A})}$ is a measure.*

Next, we proceed to show the uniqueness of the Carathéodory Extension Theorem. But first, we need to introduce a lemma.

Lemma 7.22 (Dynkin's Theorem) *Suppose $\mathscr{A}$ is an algebra of subsets of Ω. Let $\sigma(\mathscr{A})$ be the σ-algebra generated by $\mathscr{A}$. Then $\sigma(\mathscr{A})$ is the smallest λ-system containing $\mathscr{A}$, where a collection $\mathscr{C}$ of subsets of Ω is called a λ-system if it satisfies the following conditions:*

1. *$\Omega \in \mathscr{C}$;*
2. *(Closed under taking complement) For $X \in \mathscr{C}$, its complement $X^c = \Omega \setminus X$ is also in $\mathscr{C}$.*
3. *(Closed under taking disjoint union) For a disjoint sequence $(X_n)_{n=1}^{\infty}$ in $\mathscr{C}$, we have $\bigcup X_n \in \mathscr{C}$.*

Remark 7.22.1 *We have a stronger result known as the Dynkin's $\pi - \lambda$ Theorem. Which states that if $\mathscr{A}$ is a π-system (non-empty and closed under intersection), then the smallest σ-algebra generated by $\mathscr{A}$ coincides with the smallest λ-system containing $\mathscr{A}$.*

Proof: Note that the intersection of two λ-systems is a λ-system. Moreover, $\sigma(\mathscr{A})$ is a λ-system with $\mathscr{A} \subseteq \sigma(\mathscr{A})$. Hence, there is a smallest λ-system containing $\mathscr{A}$, which we denote by $\mathscr{C}_0$. Since $\mathscr{C}_0$ is the smallest such system, we have $\mathscr{C}_0 \subseteq \sigma(\mathscr{A})$. To show that $\sigma(\mathscr{A}) = \mathscr{C}_0$, it is enough to show that $\mathscr{C}_0$ is a σ-algebra, as it then implies that $\sigma(\mathscr{A}) \subseteq \mathscr{C}_0$. In particular, it is enough to show that $\mathscr{C}_0$ is closed unders intersection.
For each $X \in \mathcal{P}(\Omega)$, set

$$\mathscr{C}_X = \{Y \subseteq \Omega : X \cap Y \in \mathscr{C}_0\}.$$

We first note that for all $X \subseteq \mathscr{C}_0$, the collection $\mathscr{C}_X$ satisfy (2) and (3). Indeed, for $Y \in \mathscr{C}_X$, we have $X \cap Y$ is in the λ-system $\mathscr{C}_0$, so

$$X \cap Y^c = (X^c \cup Y)^c = (X^c \cup (X \cap Y))^c$$

is also in $\mathscr{C}_0$. This shows that $\mathscr{C}_X$ satisfies (ii). Now, suppose (Y_n) is a disjoint sequence such that $X \cap Y_n$ is in the λ-system $\mathscr{C}_0$ for each n. Then $X \cap (\bigsqcup_n Y_n) = \bigsqcup_n (X \cap Y_n)$ is in $\mathscr{C}_0$. Hence, $\mathscr{C}_X$ satisfies (iii).
If we also know $X \in \mathscr{C}_0$, the collection $\mathscr{C}_X$ is a λ-system. Indeed, as $X \cap \Omega = X$ is in $\mathscr{C}_0$, we have $\Omega \in \mathscr{C}_X$, so $\mathscr{C}_X$ satisfy (1).
Now, we show for $Y \in \mathscr{A}$, that $\mathscr{C}_0 \subseteq \mathscr{C}_Y$. Recall that $\mathscr{C}_0$ is the smallest λ-system containing $\mathscr{A}$. Hence, it is enough to show that $\mathscr{A} \subseteq \mathscr{C}_Y$, or equivalently, an arbitrary $Z \in \mathscr{A}$ is in $\mathscr{C}_Y$. This holds because $Y \cap Z \in \mathscr{A} \subseteq \mathscr{C}_0$.
Next, we show for $X \in \mathscr{C}_0$, that $\mathscr{C}_0 \subseteq \mathscr{C}_X$. Similar as in the preceding paragraph, we need to show that an arbitrary $Y \in \mathscr{A}$ is in $\mathscr{C}_X$. This is the same as $X \cap Y \in \mathscr{C}_0$, or $X \in \mathscr{C}_Y$, which is true because $\mathscr{C}_0 \subseteq \mathscr{C}_Y$.
We are finally ready to show that $\mathscr{C}_0$ is closed under intersection. For $X, Y \in \mathscr{C}_0$, we need $X \cap Y \in \mathscr{C}_0$. This is the same as $Y \in \mathscr{C}_X$, which is true because $\mathscr{C}_0 \subseteq \mathscr{C}_X$. $\square$

Finally we show the uniqueness part in the Carathéodory Extension Theorem. Suppose $(\mathcal{S}, \nu)$ and $(\mathcal{S}', \nu')$ are two pairs of σ-algebras and measures as in the statement of the theorem. We need to prove that $\mathcal{S}' = \mathcal{S}$ and $\nu' = \nu$. By symmetry, it suffices to show that $\mathcal{S} \subseteq \mathcal{S}'$ and $\nu = \nu'|_{\mathcal{S}}$.

Using the fact that μ is σ-finite, we obtain a sequence $(\Omega_n)_{n=0}^{\infty}$ in $\mathscr{A}$ such that $\Omega = \bigcup \Omega_n$ and $\mu(\Omega_n) < \infty$. In particular, we can make the Ω_n's to be disjoint. It suffices to show for any arbitrary $S \in \mathcal{S}$ that for each n, the set $S_n = S \cap \Omega_n$ is in S', and $\nu(S_n) = \nu'(S_n)$. Indeed, if we have done so, then $S = \bigcup S_n$ is in $\mathcal{S}'$ because $\mathcal{S}'$ is a σ-algebra, and $\nu(S) = \sum \nu(S_n) = \sum \nu'(S_n) = \nu'(S)$.

We first consider the case where S is in the σ-algebra $\sigma(\mathscr{A})$ generated by $\mathscr{A}$. Set

$$\mathscr{C}_n := \{Y \in \sigma(\mathscr{A}) \,:\, Y \cap \Omega_n \in \mathcal{S}' \text{ and } \nu(Y \cap \Omega_n) = \nu'(Y \cap \Omega_n)\}.$$

We verify that $\mathscr{C}_n$ is a λ-system. It is immediate that $\mathscr{C}_n$ satisfy (1). For (2), because $Y^c \cap \Omega_n = \Omega_n \setminus (Y \cap \Omega_n)$, and

$$\nu(Y^c \cap \Omega_n) = \nu(\Omega_n) - \nu(Y \cap \Omega_n) = \nu'(\Omega_n) - \nu'(Y \cap \Omega_n) = \nu'(Y^c \cap \Omega_n).$$

Finally, (3) follows from the fact that both ν and ν' are countably additive. Then by the preceding lemma, we have $\sigma(\mathscr{A}) \subseteq \mathscr{C}_n$. In particular, X is in $\mathscr{C}_n$, which implies $X_n \in \mathscr{C}'s$ and $\nu(X \cap \Omega_n) = \nu'(X \cap \Omega_n)$.

Now consider an arbitrary $S \in \mathcal{S}$. Using the outer approximation property of ν, there is $U_n \in \sigma(\mathscr{A})$ such that $S_n \subseteq U_n$ and $\nu(S_n) = \nu(I_n)$. Replacing U_n with $U_n \cap \Omega_n$ if necessary, we can assume that $U_n \subseteq \Omega_n$. Likewise, there is $V_n \in \sigma(\mathscr{A})$ such that $\Omega_n \setminus S_n \subseteq V_n$, $V_n \subseteq \Omega_n$, and $\nu(\Omega_n \setminus S_n) = \nu(V_n)$. Set $L_n = \Omega_n \setminus V_n$, we then have

$$L_n \subseteq S_n \subseteq U_n \text{ and } \nu(L_n) = \nu(S_n) = \nu(U_n).$$

By the earlier special case, we know that $\nu(L_n) = \nu'(L_n)$ and $\nu(U_n) = \nu'(U_n)$. It follows from the completeness of ν' that $S_n \in \mathcal{S}'$ and $\nu(S_n) = \nu'(S_n)$, which completes the proof.

Definition 7.23 (Complete) *A measure μ defined on a σ-algebra $\mathscr{A}$ of subsets of Ω is **complete** if for all $X \subseteq U \subseteq \Omega$ with $U \in \mathscr{A}$ and $\mu(U) = 0$, we also have $X \in \mathscr{A}$ and $\mu(X) = 0$.*

Lemma 7.24 *Suppose μ is a complete measure defined on a σ-algebra $\mathscr{A}$ of subsets of Ω. Then for all $L \subseteq X \subseteq U \subseteq \Omega$ with $L, U \in \mathscr{A}$ and $\mu(L) = \mu(U) < \infty$, we also have $X \in \mathscr{A}$ and $\mu(L) = \mu(X) = \mu(U)$.*

Proof: One has $\emptyset \subseteq X \setminus L \subseteq U \setminus L$. Then $\emptyset, U \setminus L \in \mathscr{A}$, and $\mu(\emptyset) = \mu(U \setminus L) = 0$. From the definition, we have $X \setminus L \in \mathscr{A}$ and $\mu(X \setminus L) = 0$. Then $X \in cA$ and $\mu(X) = \mu(L)$. $\square$

7.5 Metric Exterior Measures

We know that given a metric space X, a natural σ-algebra to consider on X is the σ-algebra generated by the open sets with respect to the topology induced by the metric. Recall the elements of this σ-algebra is called Borel sets.

We now turn our attention to those exterior/outer measures on X with the special property of being additive on sets that are "well-separated". We show that his property guarantees that this exterior measure defines a measure on the Borel σ-algebra.

Definition 7.25 (Metric Exterior Measure) *Given two sets A and B in a metric space (X, d), the **distance** between A and B is given by*

$$d(A, B) := \inf\{d(x, y) \,:\, x \in A \text{ and } y \in B\}.$$

*An exterior measure μ_* on X is a **metric** exterior measure if it satisfies*

$$\mu_*(A \cup B) = \mu_*(A) + \mu_*(B)$$

whenever $d(A, B) > 0$.

Theorem 7.26 *If μ_* is a metric exterior measure on a metric space X, then all the Borel sets on X satisfies the Carathéodory splitting condition. Hence μ_* restricted to the Borel σ-algebra $\mathcal{B}$ is a measure.*

Proof: Suffices to show that all the closed sets satisfies the Carathéodory splitting condition since $\mathcal{B}$ is the smallest σ-algebra containing all the closed sets.

Let F be an arbitrary closed sets and A be a subset of X with $\mu_*(A) < \infty$. For each $n > 0$, let

$$A_n := \{x \in F^c \cap A \,:\, d(x, F) \geq \frac{1}{n}\}.$$

Then $A_n \subset A_{n+1}$, and since F is closed, we have $F^c \cap A = \bigcup_{n=1}^{\infty} A_n$. Also the distance between $F \cap A$ and A_n is A_n is at least $1/n$, and since μ_* is a metric exterior measure, we have

$$\mu_*(A) \geq \mu_*((F \cap A) \cup A_n) = \mu_*(F \cap A) + \mu_*(A_n).$$

Next, we claim that

$$\lim_{n\to\infty} \mu_*(A_n) = \mu_*(F^c \cap A). \tag{7.1}$$

To see this, let $B_n = A_{n+1} \cap A_n^c$ and note that

$$d(B_{n+1}, A_n) \geq \frac{1}{n(n+1)}$$

since if $x \in B_{n+1}$ and $d(x, y) < \frac{1}{n(n+1)}$, the triangle inequality shows that $d(y, F) < \frac{1}{n}$, hence $y \notin A_n$. Therefore, we have

$$\mu_*(A_{2k+1}) \geq \mu_*(B_{2k} \cup A_{2k-1}) = \mu_*(B_{2k}) + \mu_*(A_{2k-1}),$$

and this implies that

$$\mu_*(A_{2k+1}) \geq \sum_{j=1}^{k} \mu_*(B_{2j}).$$

Similarly, we can show $\mu_*(A_{2k}) \geq \sum_{j=1}^{k} \mu_*(B_{2j-1})$. Since $\mu_*(A)$ is finite, we find that both series $\sum \mu_*(B_{2j})$ and $\sum \mu_*(B_{2j-1})$ are convergent. Finally, we note that

$$\mu_*(A_n) \leq \mu_*(F^c \cap A) \leq \mu_*(A_n) + \sum_{j=n+1}^{\infty} \mu_*(B_j).$$

So by letting $n \to \infty$, we prove (7.1) and we find that $\mu_*(A) \geq \mu_*(F \cap A) + \mu_*(F^c \cap A)$. The other direction is a direct consequence of subadditivity. Hence we conclude that all closed sets satisfies Carathéodory splitting condition. Hence by Proposition (7.21), the statement in the theorem follows. $\square$

Definition 7.27 (Borel Measure) *Given a metric space X, a measure μ defined on the Borel sets of X is referred to as a* ***Borel measure***

Remark 7.27.1 *In practice, we often consider Borel measures that assigns all finite balls a finite measure, as this ensures the Borel regularity condition which is described in the following proposition.*

Proposition 7.28 (Inner and Outer Regularity) *Suppose the Borel measure μ is finite on all balls in X of finite radius. Then for any Borel set E and any $\epsilon > 0$, there are open set $\mathcal{O}$ and a closed set F such that $E \subset \mathcal{O}$ and $\mu(\mathcal{O} - E) < \epsilon$, while $F \subset E$ and $\mu(E - F) < \epsilon$.*

Proof: We first observe the following: suppose that $F^* = \bigcup_{k=1}^{\infty} F_k$, where F_k are closed sets. Then for any $\epsilon > 0$, we can find a closed set $F \subset F^*$ such that $\mu(F^* - F) < \epsilon$. To prove this, we can assume that the sets $\{F_k\}$ are increasing. Fix a point $x_0 \in X$, and let B_n denote the ball $\{x : d(x, x_0) < n\}$, with $B_0 = \{\emptyset\}$. Since $\bigcup_{n=1}^{\infty} B_n = X$, we have that

$$F^* = \bigcup_{n \in \mathbb{N}} F^* \cap (\bar{B}_n - B_{n-1}).$$

Now for each n, $F^* \cap (\bar{B}_n - B_{n-1})$ is the limit as $k \to \infty$ of the increasing sequence of closed sets $F_k \cap (\bar{B}_n - B_{n-1})$. Since $\bar{B}_n$ has finite measure, we can find an $N = N(n)$ so that $(F^* - F_{N(n)}) \cap (\bar{B}_n - B_{n-1})$ has measure less than $\frac{\epsilon}{2^n}$. If we now let

$$F = \bigcup_{n=1}^{\infty} (F_{N(n)} \cap (\bar{B}_n - B_{n-1}),$$

it follows that the measure of $F^* - F$ is at most ϵ. F is closed because its intersection with $\bar{B}_n$ is closed, this shows that F contains all its limit points.

Let let $\mathcal{C}$ be the collection of all sets that satisfy the conclusion of the proposition. Notice first that if E belongs to $\mathcal{C}$ then so does its complement. We can also show $\mathcal{C}$ is closed under countable union. Hence by Dynkin's Theorem, suffices to show that $\mathcal{C}$ contains all the opens sets are in $\mathcal{C}$. This directly follows from the fact that for any $\mathcal{O}$ open in X, we can let

$$F_k = \{x \in \bar{B}_k : d(x, \mathcal{O}^c \geq 1/k\}$$

then $\mathcal{O} = \bigcup\limits_{k=1}^{\infty} F_k$. Apply the previous observation, we get the desired result. $\square$

7.6 Monotone Class Theorem

Definition 7.29 (Monotone Class) *Suppose X is a set and $\mathscr{C}$ is a set of subsets of X. Then $\mathscr{C}$ is called a* ***monotone class*** *on X if the following two conditions are satisfied:*

- *If $E_1 \subseteq E_2 \subseteq \cdots$ is an increasing sequence of sets in $\mathscr{C}$, then $\bigcup_{k=1}^{\infty} E_k \in \mathscr{C}$;*
- *If $E_1 \supseteq E_2 \supseteq \cdots$ is a decreasing sequence of sets in $\mathscr{C}$, then $\bigcap_{k=1}^{\infty} E_k \in \mathscr{C}$.*

Remark 7.29.1 *Every σ-algebra is a monotone class. However, some monotone class are not closed even under finite unions.*

Lemma 7.30 *If $\mathscr{A}$ is a collection of subsets of some set X, then the intersection of all monotone classes on W that contain $\mathscr{A}$ is a monotone class that contains $\mathscr{A}$. Thus this intersection is the smallest monotone class on X that contain $\mathscr{A}$.*

Theorem 7.31 (Monotone Class Theorem) *Suppose $\mathscr{A}$ is an algebra on a set X. Then the smallest σ-algebra containing $\mathscr{A}$ is the smallest monotone class containing $\mathscr{A}$.*

Proof: Let $\mathscr{C}$ denote the smallest Monotone class containing $\mathscr{A}$. Because every σ-algebra is a monotone class, then $\mathscr{C}$ is contained in the smallest σ-algebra containing $\mathscr{A}$.

We prove the inclusion in other direction, first suppose $A \in \mathscr{A}$. Let

$$\Sigma = \{E \in \mathscr{C} : A \cup E \in \mathscr{C}\}$$

Then $\mathscr{A} \subseteq \Sigma$. We can also check that Σ is a monotone class, thus the smallest monotone class that contains $\mathscr{A}$ is contained in Σ, which means that $\mathscr{C} \subseteq \Sigma$. Hence we have prove that $A \cup E \in \mathscr{A}$ for every $E \in \mathscr{C}$.

Now let

$$\Lambda = \{D \in \mathscr{C} \,:\, D \cup E \in \mathscr{C} \text{ for all } E \in \mathscr{C}\}.$$

It is clear that $\mathscr{A} \subseteq \Lambda$. We can also show that Λ is a monotone class, hence for any $D, E \in \mathscr{C}$, $D \cup E \in \mathscr{C}$.

So the monotone class $\mathscr{C}$ is closed under finite unions. Now if $E_1, E_2, \cdots, \in \mathscr{C}$, then

$$E_1 \cup E_2 \cup \cdots = E_1 \cup (E_1 \cup E_2) \cup (E_1 \cup E_2 \cup E_3) \cup \cdots$$

which is an increasing union of sequence of sets in $\mathscr{C}$, hence is in $\mathscr{C}$. Therefore, $\mathscr{C}$ is closed under countable unions.

Finally, let

$$\mathscr{C}' = \{E \in \mathscr{C} \,:\, X \setminus E \in \mathscr{C}\}.$$

Then $\mathscr{A} \subseteq \mathscr{C}'$. We can again verify that $\mathscr{C}'$ is a monotone class. Thus $\mathscr{C}$ is closed under complementation. It is also clear that $X \in \mathscr{A} \subseteq \mathscr{C}$. Thus $\mathscr{C}$ is a σ-algebra, and in particular, it must contain $\sigma(\mathscr{A})$. $\square$

7.7 Signed Measure

Definition 7.32 (Signed Measure) *Let (X, Σ) be a measurable space. A* ***signed measure*** *on (X, Σ) is a function $\nu : \Sigma \to [-\infty; \infty]$ such that*

- *ν assumes at most one of the values $+\infty$, $-\infty$.*
- *$\ni (\emptyset) = 0$;*
- *If $\{E_j\}$ is any countable collection of pairwise disjoint measurable sets, then*

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

  in addition, the series converges absolutely if the left hand side is finite.

Remark 7.32.1 *If ν is a signed measure and $A \subseteq B$, then it is not necessarily true that $\nu(A) \leq \nu(B)$.*

Definition 7.33 (Positive and Negative Sets) *Let ν be a signed measure on a measurable space (X, Σ) and let $A \in \Sigma$. Then:*

1. *If $\nu(E) \geq 0$ for every $E \in \Sigma$ and $E \subseteq A$, then we say that A is* ***positive with respect to*** *ν. We will denote a positive set by P.*
2. *If $\nu(E) \leq 0$ for every $E \in \Sigma$ and $E \subseteq A$, then we say that A is* ***negative with respect to*** *ν. We will denote a negative set by N.*
3. *If $\nu(E) = 0$ for every $E \in \Sigma$ and $E \subseteq A$, then we say that A is* ***null with respect to*** *ν.*

In particular, the empty set is an example of all three cases.

Lemma 7.34 *If a set A is both positive and negative with respect to a signed measure ν, then A is null with respect to ν.*

Proof: Clear from the definition. $\square$

Proposition 7.35 *Let ν be a signed measure on a measurable space (X, Σ). Then*

1. *Every measurable subset of a positive set in Σ is positive, and the countable union of positive set in Σ is positive.*
2. *Every measurable subset of a negative set in Σ is negative, and the countable union of negative set in Σ is negative.*
3. *very measurable subset of a null set in Σ is null, and the countable union of null set in Σ is null.*

Proof: We will prove the first statement, as the rest are similar. The first half of that statement is immediate from the definition. Let $\{P_n\}$ be a countable family of positive sets in Σ, and let A be their union. Now suppose $E \subseteq A$ is measurable, define $E_1 = E \cap P_1$ and

$$E_n = E \cap \left(P_n \setminus \bigcup_{i=1}^{n-1} P_i \right).$$

Then $\bigcup E_n = E$, and $E_n \subseteq P_n$, so $\nu(E_n) \geq 0$ for all n. Moreover, E_n's are disjoint, hence by countable additivity, we have the desired result. $\square$

Theorem 7.36 (Hahn's Lemma) *Let ν be a signed measure on a measurable space (X, Σ) and let $E \in \Sigma$ such that $0 < \nu(E) < \infty$. There there exists a measurable subset of E that is positive and of positive measure.*

Proof: We will assume that E is not a positive set, otherwise just take E as the desired subset. Then E contains a set of negative measure. Let n_1 be the smallest integer for which there exists a set $E_1 \subset E$ such that

$$\nu(E_1) < -\frac{1}{n_1}.$$

If $E \setminus E_1$ is a positive set, then we are done. Otherwise, we repeat the argument for $E \setminus E_1$. Let n_2 be the smallest integer for which there exists a set $E_2 \subset E \setminus E_1$ such that

$$\nu(E_2) < -\frac{1}{n_2}.$$

We repeat the argument for $E \setminus (E_1 \cup E_2)$. If this process stops after k steps, then $E \setminus (\bigcup_{j=1}^{k} E_j)$ is positive and is subset of E and we are done. Otherwise continue by induction to define the set

$$P = E \setminus \left(\bigcup_{j=1}^{\infty} E_j \right)$$

where $\{E_j\}$ is a pairwise disjoint subsets of E such that

$$\nu(E_k) < -\frac{1}{n_k}.$$

This will also produce an infinite sequence $n_1, n_2, \cdots$. Then

$$\nu(P) = \upsilon(E) - \sum_{k=1}^{\infty} \nu(E_k).$$

Since $\nu(P) < \infty$, we conclude that the series $\sum_{k=1}^{\infty} \frac{1}{n_k}$ absolutely converges, so $\frac{1}{n_k} \to 0$. This implies that

$$\nu(P) \geq \nu(E) + \sum_{k=1}^{\infty} \frac{1}{n_k} > \nu(E) > 0.$$

We have proved that P is a set of positive measure. To show that P is positive, let $F \subseteq P$. Then for each k we have

$$F \subseteq P \subseteq E \setminus \left(\bigcup_{j=1}^{k-1} E_j \right).$$

This implies that

$$\nu(F) \geq -\frac{1}{n_k - 1}.$$

The result follows since $n_k \to \infty$. $\square$

7.8 Decomposition of Measures

Theorem 7.37 (Hahn Decomposition Theorem) *Let ν be a signed measure on a measurable space (X, Σ). Then there exists a positive set P and a negative set N such that $X = P \cup N$ and $P \cap N = \emptyset$. Moreover, this decomposition is unique up to null sets.*

Proof: We can assume that WLOG, $-\infty \leq \nu < \infty$. Let $M = \sup\{\nu(A)$, where A is positive set with respct to $\nu\}$, note $0 \leq M < \infty$. If $M = 0$, then we are done by Hahn's Lemma. So let $0 < M < \infty$.

Let $\{P_k\}$ be a sequence of positive set such that $\nu(P_n) \to M$. Let

$$P = \bigcup_{n=1}^{\infty} P_n.$$

Then P is positive and

$$\nu(P) = \nu(P - P_k) + \nu(P_k) \geq \nu(P_k).$$

So

$$M = \lim_{k \to \infty} \nu(P_k) \leq \nu(P) \leq M$$

and consequently $\nu(P) = M$. Let $N = X \setminus P$, we show N is a negative set. If not, then there exists a set $A \subseteq N$ of positive measure, by Hahn's lemma, there is a positive set $Q \subseteq N$. Then $Q \cap P = \emptyset$ and $Q \cup P$ is positive, and so

$$\nu(Q \cup P) \geq \nu(P) = M$$

which is a contradiction. This establishes the existence of the decomposition.

Suppose $\{P, N\}$ and $\{P', N'\}$ are two Hahn decompositions of X. Then note that

$$P \setminus P' \subset P, \text{ and } P \setminus P' \subseteq N'.$$

Hence $P \setminus P'$ is both negative and positive set, so it must be the null set. Similarly, we have $P' \setminus P$ is the null set. The same argument can be applied for N and N'. □

Theorem 7.38 (Jordan Decomposition Theorem) *Let ν be a signed measure on a measurable space (X, Σ). Then there exists unique positive measures ν^+, ν^-, such that*

$$\nu = \nu^+ - \nu^-.$$

Proof: Existence: By Hahn's Lemma, we have $X = P \cup N$, where P is positive and N is negative, $P \cap N = \emptyset$. Then for any $E \subseteq X$ measurable, define $\nu^+(E) = \nu(E \cap P)$ and $\nu^-(E) = -\nu(E \cap N)$. Then $\nu = \nu^+ - \nu^-$.

Uniqueness: let $\nu = \mu^+ - \mu^-$. Let $X = P \cup N$, where P is positive and N is negative. Since for any $E \subseteq X$,

$$\nu(E) = \nu(E \cap P) + \nu(E \cap N)$$

and $\nu(E \cap P)$ is positive $\nu(E \cap N)$ is negative, we have no choice but to define the measures ν^+ and ν^- in the way described above. Since by Hahn's Lemma, the sets P and N are unique up to null sets, then we can see ν^+ and ν^- are unique. □

Definition 7.39 (Total Variation of Measures) *Suppose ν is a signed measure on the measurable space (X, Σ). Then we define the **total variation** of the measure ν by*

$$|\nu|(E) = \nu^+(E) + \nu^-(E), \ \forall E \in \Sigma.$$

7.9 Radon-Nikodym Theorem

Definition 7.40 (Finite and Σ-finite Measures) *A measure ν is **finite** if it only takes values in $\mathbb{R}$ for all $E \subseteq \Omega$. A measure ν is σ-**finite** if it is a countable sum of finite measures.*

Definition 7.41 (Support) *We say a measure ν is **supported** on a set A, if $\nu(E) = \nu(E \cap A)$ for all $E \in \mathscr{A}$.*

Definition 7.42 (Mutually Singular Measures) *Let ν be a signed measure on $(X, \mathscr{A}, \mu)$, then we say ν and μ are **mutually singular** if there are disjoint subsets A and B in $\mathscr{A}$, such that ν is supported on A and μ is supported on B. In this case, we write $\nu \perp \mu$.*

Definition 7.43 (Absolutely Continuous Finite Measure) *Let ν be a finite signed measure on a signed measure space (X, Σ, μ), and $E \in \Sigma$, we say that ν is absolutely continuous with respect to μ if for every $\epsilon > 0$, there exists $\delta > 0$ such that if $\mu(E) < \delta$ then $\nu(E) < \epsilon$.*

Definition 7.44 (Absolutely Continuity of Measure) *Let μ and ν be any two signed measures on Ω. We say ν is **absolutely continuous** with respect to μ if for all measurable $E \subseteq \Omega$, $\mu(E) = 0$ implies $\nu(E) = 0$. In this case, we write $\nu \ll \mu$.*

Lemma 7.45 *Let f be a measurable function from Ω to $\mathbb{R}$, whose Lebesgue integral is defined on Ω. Then the signed measure μ_f defined by*

$$\mu_f(A) := \int_A f d\mu$$

is absolutely continuous with respect to μ.

Proof: Clear from the definition. □

Proposition 7.46 *Let ν be a finite signed measure on a measure space (X, Σ, μ). If $E \in \Sigma$ and $\nu(E) = 0$ whenever $\nu(E) = 0$, then for every $\epsilon > 0$, there exists $\delta > 0$ such that if $\nu(E) < \delta$ then $\nu(E) < \epsilon$.*

Remark 7.46.1 *It is clear that the condition in Definition 7.43 always implies the condition in Definition 7.44. Hence by this proposition, it shows the two definitions are actually equivalent for ν being a finite measure.*

Proof: Suppose not, then there is $\epsilon > 0$ and $E_k \in \Sigma$ such that

$$\mu(E_k) < \frac{1}{2^k} \text{ and } \nu(E_k) \geq \epsilon$$

for every $k \in \mathbb{N}$. Define

$$F_m := \bigcup_{k=m}^{\infty} E_k.$$

F_m is a decreasing sequence, and

$$\mu(F_m) \leq \sum_{k=m} \frac{1}{2^k} = \frac{1}{2^{m-1}}.$$

Now let

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

By the property of σ-algebra, we have $F \in \Sigma$. Then

$$\mu(F) \leq \lim_{m\to\infty} \frac{1}{2^{m-1}} = 0.$$

But on the other hand, we have

$$\nu(F_1) \leq \nu(X) < \infty$$

and $\mu(F_m) \geq \epsilon$ for all m, hence $\nu(F) \geq \epsilon$, which is a contradiction. □

Proposition 7.47 *Let ν be a signed measure and μ is a measure. Then $\nu \ll \mu$ if and only if $\nu^+ \ll \mu$ and $\nu^- \ll \mu$.*

Proof: Follows from Jordan Decomposition Theorem. □

Lemma 7.48 *Let ν and μ be finite measure (not signed) on a measurable space (X, Σ), and $\nu \ll \mu$. Then there exists $A \in \Sigma$ and $\epsilon > 0$ such that $\mu(A) > 0$ and $\nu(E) \geq \epsilon\mu(E)$ for all $E \subseteq A$.*

Theorem 7.49 (Radon-Nikodym Theorem) *Suppose μ is a σ-finite positive measure on the measure space (X, Σ) and ν is a σ-finite signed measure on Σ. Then there exists unique signed measures ν_α and ν_s such that $\nu_a \ll \mu$, $\nu_s \perp \mu$ and $\nu = \nu_a + \nu_s$. In addition, the measure ν_a takes the form $d\nu_a = fd\mu$; that is*

$$\nu_a(E) = \int_E f(x)d\mu(x)$$

for some extended μ-integrable function f. In the special case where $\nu \ll \mu$, we have $d\nu = fd\mu$.

Proof: We introduce a proof that is an application of the Hilbert space idea. We start with the case when both ν and μ are positive and finite. Let $\rho = \nu + \mu$, and consider the transformation on $L^2(X, \rho)$ defined by

$$\ell(\psi) = \int_X \psi(x)d\nu(x).$$

The mapping ℓ defines a bounded linear functional on $L^2(X, \rho)$, since

$$\begin{aligned} |\ell(\psi)| &\leq \int_X |\psi(x)|d\nu(x) \\ &\leq \int_X |\psi(x)|d\rho(x) \\ &\leq (\rho(X))^{1/2} \left(\int_X |\psi(x)|^2 d\rho(x) \right)^{1/2}. \end{aligned}$$

Since $L^2(X, \rho)$ is a Hilbert space, so the Riesz representation theorem tell us that there exists a $g \in L^2(X, \rho)$ such that

$$\int_X \psi(x)d\nu(x) = \int_X \psi(x)g(x)d\rho(x)$$

for all $\psi \in L^2(X, \rho)$. If $E \in \Sigma$ with $\rho(E) > 0$, then we set $\psi = \chi_E$ to see that

$$0 \leq \frac{1}{\rho(E)} \int_E g(x)d\rho(x) \leq 1, \tag{7.2}$$

since $\rho \geq \nu$, so $\rho(E) \geq \nu(E)$. From this, we conclude that

$$0 \leq g(x) \leq 1 \quad \text{a.e.}\, x$$

with respect to the measure ρ; WLOG, we assume that $0 \leq g(x) \leq 1$ for all x. Then Equation (7.2) tells us that

$$\int \psi(1 - g)d\nu = \int \psi g d\mu.$$

Consider now the two sets

$$A = \{x \in X \,:\, 0 \leq g(x) < 1\} \quad \text{and} \quad B = \{x \in X \,:\, g(x) = 1\},$$

and define two measures ν_a and ν_s on Σ by

$$\nu_a(E) = \nu(A \cap E) \quad \text{and} \quad \nu_s(E) = \nu(B \cap E). \tag{7.3}$$

Then it is clear that $\nu_s \perp \mu$, as μ is supported on B and μ is supported on B^c, since

$$0 = \int \chi_B d\mu = \mu(B).$$

Finally, we set $\psi = \chi_E(1 + g + \cdots + g^n)$ in Equation (7.3) we obtain

$$\int_E (1 - g^{n+1}) d\nu = \int_E g(1 + \cdots + g^n) d\mu. \tag{7.4}$$

Since $(1 - g^{n+1})(x) = 0$ in $x \in B$, and $(1 - g^{n+1})(x) \to 1$ if $x \in A$, then the dominated convergence theorem implies that the left hand side of Equation (7.4) converges to $\nu(A \cap E) = \nu_a(E)$. Also $1 + g + \cdots + g^n$ converges to $\frac{1}{1-g}$, so we find in the limit that

$$\nu_a(E) = \int_E f d\mu, \quad \text{where } f = \frac{g}{1-g}.$$

Note that $f \in L^1(X, \mu)$, since $\nu_a(X) \leq \nu(X) < \infty$.

If ν and μ are σ-finite and positive, we may clearly find sets $E_j \in \Sigma$ such that $X = \bigcup E_j$ and $\mu(E_j) < \infty$, $\nu(E_j) < \infty$ for all j. We may define positive and finite measures on Σ by

$$\mu_j(E) = \mu(E \cap E_j) \quad \text{and} \quad \nu_j(E) = \nu(E \cap E_j),$$

and then we can write for each j, $\nu_j = \nu_{j,a} + \nu_{j,s}$ where $\nu_{j,s} \perp \mu_j$ and $\nu_{j,a} = f_j d\mu_j$. Then by setting

$$f = \sum f_j, \quad \nu_s = \sum \nu_{j,s}, \quad \text{and} \quad \nu_a = \sum \nu_{j,a},$$

we have the desired decomposition and f.

Finally if ν is signed, then by Jordan Decomposition Theorem (7.38), we can decompose ν uniquely into $\nu^+ - \nu^-$ where ν^+ and ν^- are positive measures. Hence we can apply the previous case. □

Definition 7.50 (Radon-Nikodym Derivative) *The function f predicted by the Radon-Nikodym Theorem is refered to as the* ***Radon-Nikodym derivative*** *of ν with respect to μ, and is denoted by*

$$f = \frac{d\nu}{d\mu}, \, d\nu = f d\mu.$$

Theorem 7.51 (Fundamental Theorem of Calculus for RN derivatives) *Let ν be a σ-finite signed measure and μ be a σ-finite measure on a measurable space (X, Σ). Then $\nu \ll \mu$ if and only if $\frac{d\nu}{d\mu}$ exists and*

$$\nu(A) = \int_A \frac{d\nu}{d\mu} d\mu.$$

Proof: The forward direction follows from Theorem 7.50, and reverse direction follows immediately from the definition of absolutely continuity. $\square$

Proposition 7.52 (Calculus of RN Derivative) *Let ν, λ be two σ-finite signed measures and μ is a σ-finite measure on measurable space (X, Σ). Then*

1. *If $\frac{d\nu}{d\mu}$ exists and $c \in \mathbb{R}$, then*
$$\frac{dc\nu}{d\mu} = c\frac{d\nu}{d\mu}$$
a.e. on X.

2. *If $\nu + \lambda$ is a defined measure on (X, Σ) such that $\frac{d\nu}{d\mu}$ and $\frac{d\lambda}{d\mu}$ both exist, then*
$$\frac{d(\nu + \lambda)}{d\mu} = \frac{d\nu}{d\mu} + \frac{d\lambda}{d\mu}$$
a.e. on X.

3. *If $\lambda \geq 0$ (i.e., not signed) and $\frac{d\lambda}{d\mu}$ exists, then*
$$\frac{d\lambda}{d\mu} \geq 0.$$

Theorem 7.53 *Let ν be a signed σ-finite measure and μ be a σ-finite measure on a measurable space (X, Σ). If $\frac{d\nu}{d\mu}$ exists, then for every μ-measurable function on X, we have*
$$\int_X f d\nu = \int_X f \frac{d\nu}{d\mu} d\mu.$$

Theorem 7.54 (Chain Rule of RN Derivatives) *Let μ, ν, λ be σ-finite measures on a measurable space (X, Σ). If $\frac{d\nu}{d\mu}, \frac{d\mu}{d\lambda}$ exist on X. Then $\frac{d\nu}{d\lambda}$ exists and*
$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \cdot \frac{d\mu}{d\lambda}.$$

7.10 Lebesgue Decomposition of Measures

Definition 7.55 (Mutually Singular Measures) *Let μ and λ be two measure on a measurable space (X, Σ). Then μ and λ are said to be* ***mutually singular****, denoted $\nu \perp \lambda$ on (X, Σ) if there exists a decomposition $\{A, B\}$ such that*
$$\mu(A) = \lambda(B) = 0.$$

Corollary 7.55.1 *Let ν be a signed measure, then the two Jordan decomposition measure of ν are mutually singular, i.e., $\nu^+ \perp \nu^-$.*

Theorem 7.56 (Lebesgue Decomposition Theorem) *Let ν be a σ-finite measure on a measurable space $(X, \Sigma\mu)$. Then there exists unique measure λ, ρ such that $\nu = \lambda + \rho$, where $\lambda \ll \mu$ and $\rho \perp \mu$.*