CHAPTER 1

Measure Theory

1. Topology and Metric Space

Definition 1. *A collection τ of subsets of a set X is said to be a topology in X if τ has the following three properties:*
(i). $\emptyset \in \tau$ *and* $X \in \tau$.
(ii). If $V_i \in \tau$ *for* $i = 1, \cdots, n$, *then* $V_1 \cap V_2 \cap \cdots \cap V_n \in \tau$.
(iii). If $\{V_\alpha\}$ *is an arbitrary collection of members of* τ, *then* $\cup_\alpha V_\alpha \in \tau$.

If τ is a topology in X, then X is called a topological space, and the members of τ are called the open sets in X. A set E is said to be closed if E^c is open.

Example 1. *(i).* $\tau = \{\emptyset, X\}$ *is a topology.*
(ii). $\tau = 2^X$ *is a topology.*
(iii). Let E_1, E_2 *be two nonempty subsets of* X, *then*

$$\tau = \{\emptyset, E_1, E_2, E_1 \cap E_2, E_1 \cup E_2, X\}$$

is a topology of X. *This topology has* 6 *open sets if all the listed subsets are different.*

Definition 2. *A metric space is a set* X *in which a distance function* $\rho : X \times X \to \mathbb{R}^+$ *is defined with the following properties:*
(a). $0 \leq \rho < \infty$ *for all* $x, y \in X$.
(b). $\rho(x, y) = 0$ *if and only if* $x = y$.
(c). $\rho(x, y) = \rho(y, x)$ *for all* $x, y \in X$.
(d). $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ *for all* $x, y, z \in X$.

A metric space (X, ρ) has a natural topology which is the collection of all sets $E \subset X$ which are arbitrary unions of open balls of the form

$$B(x, r) = \{y \in X : \rho(x, y) < r\}.$$

2. σ-algebra

Definition 3. *A collection $\mathcal{M}$ of subsets of a set X is said to be a σ-algebra in X if $\mathcal{M}$ has the following three properties:*
(i). $X \in \mathcal{M}$.
(ii). If $A \in \mathcal{M}$, then $A^c \in \mathcal{M}$.
(iii). If $A_n \in \mathcal{M}$, $n = 1, 2, \cdots$, then $A = \cup_{n=1}^{\infty} A_n \in \mathcal{M}$.

If $\mathcal{M}$ is a σ-algebra in X, then $(X, \mathcal{M})$ is called a measurable space, and the members of $\mathcal{M}$ are called the measurable sets in X.

Example 2. *(i). $\mathcal{M} = \{\emptyset, X\}$ is a σ-algebra.*
(ii). $\mathcal{M} = 2^X$ is a σ-algebra.
(iii). Let $E \neq X$ be a nonempty subset of X, then

$$\mathcal{M} = \{\emptyset, E, X \backslash E, X\}$$

is a σ-algebra of X.

Proposition 1. *Let $(X, \mathcal{M})$ be a measurable space. Then*
(i). $\emptyset \in \mathcal{M}$.
(ii). $A, B \in \mathcal{M}$ implies $A \backslash B \in \mathcal{M}$.
(iii). $\mathcal{M}$ is closed under finite unions, finite or countable intersections.

Proof. (i) $\emptyset = X^c \in \mathcal{M}$.
(ii). $A \backslash B = A \cap B^c = (A^c \cup B)^c$.
(iii). Let $A_k \in \mathcal{M}$, $1 \leq k \leq n$. We define $A_k = \emptyset$ for $k \geq n+1$. Then the finite union

$$\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{\infty} A_k \in \mathcal{M}$$

and the finite intersection

$$\bigcap_{k=1}^{n} A_k = \left(\bigcup_{k=1}^{n} A_k^c \right)^c \in \mathcal{M}.$$

Let $A_k \in \mathcal{M}$, $k \geq 1$. Then the countable intersection

$$\bigcap_{k=1}^{\infty} A_k = \left(\bigcup_{k=1}^{\infty} A_k^c \right)^c \in \mathcal{M}.$$

□

It is easy to verify from the definition of σ-algebra that any intersection of σ-algebras is a σ-algebra:

Proposition 2. *If $\{\mathcal{M}_\alpha\}_{\alpha \in \Lambda}$ is a family of σ-algebras in a set X, then $\cap_{\alpha \in \Lambda} \mathcal{M}_\alpha$ is also a σ-algebra in X.*

Theorem 1. *If $\mathcal{F}$ is any collection of subsets of X , there exists a smallest σ-algebra $\mathcal{M}^*$ in X such that $\mathcal{F} \subset \mathcal{M}^*$.*

We say $\mathcal{M}^*$ is the σ-algebra generated by $\mathcal{F}$, which can be defined as the intersection of all $\mathcal{F}$ containing σ-algebras in X.

Let (X, τ) be a topological space. Let $\mathcal{B}$ be the σ-algebra generated by τ. The members of $\mathcal{B}$ are called the Borel sets of X. In particular, closed sets are Borel sets, and so are all countable unions of closed sets and all countable intersections of open sets.

3. Measure

DEFINITION 4. *A positive measure is a function* μ*, defined on a* σ*-algebra* $\mathcal{M}$*, whose range is in* $[0,\infty]$*, which satisfies*
(i). $\mu(\emptyset)=0$.
(ii). (Countable additivity) If $\{A_i\}_{i=1}^{\infty}$ *is a pairwise disjoint collection of members of* $\mathcal{M}$*, then*

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

The triple $(X,\mathcal{M},\mu)$ *is called measure space.*

If $\mu(X)<\infty$, then μ is said to be a finite measure.
If $\mu(X)=1$, then μ is a probability measure.
If $X=\displaystyle\bigcup_{i=1}^{\infty}A_i$ where $A_i\in\mathcal{M}$ and $\mu(A_i)<\infty$ for all $i=1,2,3,\cdots$, then we say that μ is σ-finite.

EXAMPLE 3. *(i). Trivial measure* $\mu\equiv 0$.
(ii). Counting measure: For any $A\in\mathcal{M}$*,* $\mu(A)=\sharp$ *of elements in* A.
(iii). Dirac measure: Given $x_0\in X$*, we have for any* $A\in\mathcal{M}$*,* $\mu(A)=1$ *if* $x_0\in A$ *and* $\mu(A)=0$ *if* $x_0\notin A$.

A complex (real) measure is a complex-valued (real-valued) countably additive function defined on a σ-algebra.

THEOREM 2. *Let* $(X,\mathcal{M},\mu)$ *be a measure space.*
(i). (Finite additivity) If $\{A_i\}_{i=1}^{n}$ *is a pairwise disjoint collection of members of* $\mathcal{M}$*, then*

$$\mu\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mu(A_i).$$

(ii). (Monotonicity) $A\subset B$ *implies* $\mu(A)\le\mu(B)$ *if* $A,B\in\mathcal{M}$.
(iv). Let $A_1\subset A_2\subset A_3\subset\cdots$ *be members in* $\mathcal{M}$ *and* $A=\bigcup_{i=1}^{\infty}A_i$*, then*

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

(v). Let $A_1\supset A_2\supset A_3\cdots$ *be members in* $\mathcal{M}$*,* $\mu(A_1)$ *is finite and* $A=\bigcap_{i=1}^{\infty}A_i$*, then*

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

PROOF. (i). Using countable additivity on $\{A_i\}_{i=1}^{\infty}$ with $A_i=\emptyset$ for $i\ge n+1$.
(ii). Write $B=A\cup(B\backslash A)$, then use finite additivity.
(iii). Let $B_1=A_1$, $B_i=A_i\backslash A_{i-1}$ for $i\ge 2$, then $\{B_i\}_{i=1}^{\infty}$ are pairwise disjoint and $A=\bigcup_{i=1}^{\infty}B_i$. Hence countable additivity implies

$$\mu(A)=\sum_{i=1}^{\infty}\mu(B_i)=\lim_{n\to\infty}\sum_{i=1}^{n}\mu(B_i)=\lim_{n\to\infty}\mu(A_n)$$

where, in the last step, we used finite additivity.

(iv). Consider $C_i = A_1 \backslash A_i$, then we have $C_i \nearrow$ and $A_1 \backslash A = \bigcup\limits_{i=1}^{\infty} C_i$, so (iv) implies

$$\mu\left(A_1 \backslash A\right) = \lim_{n \to \infty} \mu\left(C_i\right).$$

Since $\mu\left(A_1\right) < \infty$, $\mu\left(A_1 \backslash A\right) = \mu\left(A_1\right) - \mu\left(A\right)$ and $\mu\left(C_i\right) = \mu\left(A_1\right) - \mu\left(A_i\right)$, we have

$$\mu\left(A\right) = \lim_{n \to \infty} \mu\left(A_n\right).$$

□

REMARK 1. *The assumption "$\mu\left(A_1\right)$ is finite" is necessary in part (v) of the above theorem. For example, let μ be the Lebesgue measure and $A_n = (n, \infty)$, then we have for each n, $\mu\left(A_n\right) = \infty$ while $\mu\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = 0$.*

4. Outer Measure and Caratheodory Construction

Definition 5. *Let X be a set. A function*

$$\mu^* : 2^X \to [0, \infty]$$

is called an outer measure if
(i). $\mu^*(\emptyset) = 0$;
(ii). (Monotonicity) $A \subset B$ implies $\mu^*(A) \leq \mu^*(B)$;
(iii). (Countable Sub-additivity) for all sets $A_i \subset X$,

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

We call a set $E \subset X$ μ^*-measurable if for any $A \subset X$,

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

This criterion is called Caratheodory condition. Since

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \backslash E)$$

holds for any $A, E \subset X$, Caratheodory condition is equivalent to the one-sided inequality

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

Theorem 3. *The μ^*-measurable sets form a σ-algebra $\mathcal{M}^*$ and μ^* restricted to $\mathcal{M}^*$ is a measure.*

Proof. Step 1. $X \in \mathcal{M}^*$.
Step 2. If $E \in \mathcal{M}^*$, then $E^c \in \mathcal{M}^*$ since Caratheodory condition is symmetric in E and E^c.
Step 3. If $E, F \in \mathcal{M}^*$, then $E \cup F \in \mathcal{M}^*$. To see this, for any $A \subset X$, we have from Caratheodory condition,

$$\begin{aligned} \mu^*(A) &= \mu^*(A \cap E) + \mu^*(A \backslash E), \\ \mu^*(A \backslash E) &= \mu^*((A \backslash E) \cap F) + \mu^*((A \backslash E) \backslash F) \\ &= \mu^*((A \backslash E) \cap F) + \mu^*(A \backslash (E \cup F)), \\ \mu^*(A \cap (E \cup F)) &= \mu^*((A \cap (E \cup F)) \cap E) + \mu^*((A \cap (E \cup F)) \backslash E) \\ &= \mu^*(A \cap E) + \mu^*((A \backslash E) \cap F). \end{aligned}$$

which yields $\mu^*(A) = \mu^*(A \cap (E \cup F)) + \mu^*(A \backslash (E \cup F))$. Hence, $E \cup F \in \mathcal{M}^*$.
Step 4. If $E, F \in \mathcal{M}^*$, then $E \backslash F \in \mathcal{M}^*$ since $E \backslash F = (E^c \cup F)^c$.
Step 5. If $E_n \in \mathcal{M}^*$, $n = 1, 2, \cdots$ are pairwise disjoint, then we have for any $A \subset X$,

$$\sum_{n=1}^{\infty} \mu^*(A \cap E_n) = \mu^*\left(A \cap \left(\bigcup_{n=1}^{\infty} E_n\right)\right).$$

First, from the Caratheodory condition and the measurability of E_1, we have

$$\begin{aligned} \mu^*(A \cap (E_1 \cup E_2)) &= \mu^*((A \cap (E_1 \cup E_2)) \cap E_1) + \mu^*((A \cap (E_1 \cup E_2)) \backslash E_1) \\ &= \mu^*(A \cap E_1) + \mu^*(A \cap E_2). \end{aligned}$$

Hence,

$$\sum_{i=1}^{n} \mu^*(A \cap E_i) = \mu^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)\right) \leq \mu^*\left(A \cap \left(\bigcup_{i=1}^{\infty} E_i\right)\right).$$

Let $n \to \infty$, we have

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

On the other hand,

$$\sum_{i=1}^{\infty} \mu^* (A \cap E_i) \geq \mu^* \left(A \cap \left(\bigcup_{i=1}^{\infty} E_i \right) \right)$$

follows from countable sub-additivity.
Step 6. If $E_n \in \mathcal{M}^*$, $n = 1, 2, \cdots$ are pairwise disjoint, then $E = \cup_{n=1}^{\infty} E_n \in \mathcal{M}^*$. For any $A \subset X$, we need to show

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

We could assume $\mu^* (A) < \infty$, otherwise the inequality is trivial. Since $\cup_{i=1}^{n} E_i \in \mathcal{M}^*$, we have

$$\begin{aligned} \mu^* (A) &\geq \mu^* \left(A \cap \left(\bigcup_{i=1}^{n} E_i \right) \right) + \mu^* \left(A \backslash \left(\bigcup_{i=1}^{n} E_i \right) \right) \\ &\geq \sum_{i=1}^{n} \mu^* (A \cap E_i) + \mu^* \left(A \backslash \left(\bigcup_{i=1}^{\infty} E_i \right) \right). \end{aligned}$$

Let $n \to \infty$, we have

$$\begin{aligned} \mu^* (A) &\geq \sum_{i=1}^{\infty} \mu^* (A \cap E_i) + \mu^* \left(A \backslash \left(\bigcup_{i=1}^{\infty} E_i \right) \right) \\ &= \mu^* \left(A \cap \left(\bigcup_{i=1}^{\infty} E_i \right) \right) + \mu^* \left(A \backslash \left(\bigcup_{i=1}^{\infty} E_i \right) \right). \end{aligned}$$

Step 7. If $E_n \in \mathcal{M}^*$, $n = 1, 2, \cdots$, then $E = \cup_{n=1}^{\infty} E_n \in \mathcal{M}^*$. To see this, let $F_1 = E_1$ and

$$F_n = \left(\bigcup_{i=1}^{n} E_i \right) \backslash F_{n-1} \text{ for } n \geq 2.$$

Then $F_n \in \mathcal{M}^*$ are pairwise disjoint and hence from step 6,

$$E = \bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n \in \mathcal{M}^*.$$

Step 8. From Steps 1,2,7, we see $\mathcal{M}^*$ is a σ-algebra. μ^* is a measure on $\mathcal{M}^*$ follows from step 5 with $A = X$. □

A measure μ is said to be complete if every subset of zero measure set is measurable. It is easy to see that μ^* constructed above is complete.

Let (X, ρ) be a metric space. For $E, F \subset X$, we define

$$\operatorname{dist} (E, F) = \inf_{x \in E, y \in F} \rho (x, y).$$

Definition 6. *An outer measure μ^* defined on subsets of a metric space (X, ρ) is called metric outer measure if*

$$\mu^* (E \cup F) = \mu^* (E) + \mu^* (F)$$

whenever $E, F \subset X$ *and* $\operatorname{dist}(E, F) > 0$.

THEOREM 4. *If* μ^* *is a metric outer measure, then all Borel sets are* μ^**-measurable.*

PROOF. We only need to show that any open set is μ^*-measurable. Let O be an open set. For any $A \subset X$, we need to prove

$$\mu^*(A) \geq \mu^*(A \cap O) + \mu^*(A \backslash O). \tag{4.1}$$

We assume $\mu^*(A) < \infty$, otherwise (4.1) is trivial. Let

$$O_n = \left\{ x \in O : \operatorname{dist}(x, O^c) > \frac{1}{n} \right\}.$$

Then

$$\operatorname{dist}(A \cap O_n, A \backslash O) > 0,$$

and we have

$$\begin{aligned} \mu^*(A) &\geq \mu^*(A \backslash (O \backslash O_n)) = \mu^*(A \cap O_n) + \mu^*(A \backslash O) \\ &\geq \mu^*(A \cap O) - \mu^*(A \cap (O \backslash O_n)) + \mu^*(A \backslash O) \end{aligned} \tag{4.2}$$

where the last inequality follows from subadditivity. Next, for any $n \geq 1$, we define $D_n = O_{n+1} \backslash O_n$. Then

$$O \backslash O_n = \bigcup_{k=n}^{\infty} D_k.$$

Since mutual distances between $D_1, D_3, D_5, \cdots$ are positive, we have for any $n \geq 1$,

$$\sum_{k=1}^{n} \mu^*(A \cap D_{2k-1}) = \mu^*\left(A \cap \left(\bigcup_{k=1}^{n} D_{2k-1}\right)\right) \leq \mu^*(A).$$

Let $n \to \infty$, we have

$$\sum_{k=1}^{\infty} \mu^*(A \cap D_{2k-1}) \leq \mu^*(A).$$

Similarly,

$$\sum_{k=1}^{\infty} \mu^*(A \cap D_{2k}) \leq \mu^*(A).$$

Thus,

$$\sum_{n=1}^{\infty} \mu^*(A \cap D_n) \leq 2\mu^*(A) < \infty.$$

So we have

$$\mu^*(A \cap (O \backslash O_n)) = \mu^*\left(\bigcup_{k=n}^{\infty} (A \cap D_k)\right) = \sum_{k=n}^{\infty} \mu^*(A \cap D_k) \to 0 \text{ as } n \to \infty. \tag{4.3}$$

Finally, (4.1) follows from (4.2) and (4.3). □

5. Lebesgue Measure

A bounded closed interval of $\mathbb{R}^n$ is a set of the form

$$I = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

where $a_i \leq b_i$, $1 \leq i \leq n$. The volume of I is defined by

$$|I| = \prod_{k=1}^{n} (b_k - a_k).$$

We define a set function L_n^* such that for any subset $A \subset \mathbb{R}^n$,

(5.1)

$$L_n^*(A) = \inf \left\{ \sum_{i=1}^{\infty} |I_i| : \text{each } I_i \text{ is a bounded closed interval and } A \subset \bigcup_{i=1}^{\infty} I_i \right\}.$$

Remark 2. *We can replace the closed intervals with open intervals and for each $A \subset \mathbb{R}^n$,*

$$L_n^*(A) = \inf \left\{ \sum_{i=1}^{\infty} |I_i| : \textit{each } I_i \textit{ is a bounded open interval and } A \subset \bigcup_{i=1}^{\infty} I_i \right\}. \tag{5.2}$$

Theorem 5. *L_n^* is a metric outer measure.*

Proof. (i). $\mu^*(\emptyset) = 0$ and monotonicity is trivial.
(ii). Let $A_i \subset \mathbb{R}^n$ and $A = \cup_{i=1}^{\infty} A_i$, we need to show that

$$L_n^*(A) \leq \sum_{i=1}^{\infty} L_n^*(A_i). \tag{5.3}$$

We could assume $\sum_{i=1}^{\infty} L_n^*(A_i) < \infty$, otherwise it is trivial. Given $\varepsilon > 0$, for each A_i, there exists bounded open intervals $\{I_{ij}\}_{j=1}^{\infty}$, such that $A_i \subset \bigcup\limits_{j=1}^{\infty} I_{ij}$ and

$$\sum_{j=1}^{\infty} |I_{ij}| \leq L_n^*(A_i) + \frac{\varepsilon}{2^i}.$$

Since $A \subset \bigcup\limits_{i,j=1}^{\infty} I_{ij}$, we have

$$L_n^*(A) \leq \sum_{i,j=1}^{\infty} |I_{ij}| \leq \sum_{i=1}^{\infty} \left(L_n^*(A_i) + \frac{\varepsilon}{2^i} \right) \leq \sum_{i=1}^{\infty} L_n^*(A_i) + \varepsilon.$$

Let $\varepsilon \to 0$, (5.3) follows.
(iii). Let $A_1, A_2 \subset \mathbb{R}^n$ and $d = \operatorname{dist}(A_1, A_2) > 0$. We need to show

$$L_n^*(A_1 \cup A_2) \geq L_n^*(A_1) + L_n^*(A_2).$$

We first observe that we could limit the size of I_i in the covering, in particular, for any set $A \subset \mathbb{R}^n$,

$$L_n^*(A) = \inf \left\{ \sum_{i=1}^{\infty} |I_i| : \text{each } I_i \text{ is bounded close interval, } \operatorname{diam} I_i < d \text{ and } A \subset \bigcup_{i=1}^{\infty} I_i \right\}.$$

For any $\varepsilon > 0$, there exists bounded close intervals $\{I_i\}_{i=1}^{\infty}$, such that $\operatorname{diam} I_i < d$, $A \subset \bigcup\limits_{i=1}^{\infty} I_i$ and

$$L_n^*(A_1 \cup A_2) \geq \sum_{i=1}^{\infty} |I_i| - \varepsilon.$$

For each i, since $\operatorname{diam} I_i < d = \operatorname{dist}(A_1, A_2)$, we have either $I_i \cap A_1 = \emptyset$ or $I_i \cap A_2 = \emptyset$ so we can separate $\{I_i\}_{i=1}^{\infty}$ as two disjoint subcollections $\{I_i\}_{i\in\Lambda_1}$ and $\{I_i\}_{i\in\Lambda_2}$such that $A_j \subset \bigcup\limits_{i=\Lambda_j}^{\infty} I_i$, $j = 1, 2$. Hence, we have

$$L_n^*(A_1) + L_n^*(A_2) \leq \sum_{i=1}^{\infty} |I_i| \leq L_n^*(A_1 \cup A_2) + \varepsilon.$$

Theorem is proved by letting $\varepsilon \to 0$. $\square$

Definition 7. *L_n^* is called outer Lebesgue measure on $\mathbb{R}^n$. L_n^*-measurable sets are said to be Lebesgue measurable. The restriction of L_n^* to Lebesgue measurable sets is called Lebesgue measure and is denoted by L_n.*

Corollary 1. *All Borel sets are Lebesgue measurable.*

Lebesgue measure is complete: All sets with $L_n^*(A) = 0$ are Lebesgue measurable and have zero Lebesgue measure.

Theorem 6. *Let I be a bounded closed (or open) interval*
(i). $L_n(I) = |I|$.
(ii). $L_n(\partial I) = 0$.

Proof. We first assume I is a bounded closed interval. Then since I is a covering of itself, we have from the definition,

$$L_n(I) \leq |I|.$$

Next, for any $\varepsilon > 0$, there exists open intervals $\{I_i\}$ which covers I , such that

$$L_n(I) \geq \sum_{i=1}^{\infty} |I_i| - \varepsilon.$$

Since I is compact, there exists a finite subcollection, which we denote by $\{I_i\}_{i=1}^{k}$ such that $I \subset \bigcup\limits_{i=1}^{k} I_i$. We claim

$$|I| \leq \sum_{i=1}^{k} |I_i|. \tag{5.4}$$

To see this, we observe that

$$\chi_I \leq \sum_{i=1}^{k} \chi_{I_i},$$

so the Riemann integral

$$\int \chi_I dx \leq \int \sum_{i=1}^{k} \chi_{I_i} dx \leq \sum_{i=1}^{k} \int \chi_{I_i} dx,$$

and (5.4) follows. Hence, we have for any $\varepsilon > 0$,

$$L_n(I) \geq |I| - \varepsilon.$$

So we conclude $L_n(I) = |I|$. If $I = \prod_{i=1}^{n} (a_i, b_i)$ is open, we observe that for any $0 < \varepsilon < \frac{1}{3} \min_i |b_i - a_i|$,

$$\prod_{i=1}^{n} [a_i + \varepsilon, b_i - \varepsilon] \subset I \subset \prod_{i=1}^{n} [a_i - \varepsilon, b_i + \varepsilon],$$

hence,

$$L^n\left(\prod_{i=1}^{n} [a_i + \varepsilon, b_i - \varepsilon]\right) \leq L^n(I) \leq L^n\left(\prod_{i=1}^{n} [a_i - \varepsilon, b_i + \varepsilon]\right),$$

and

$$\prod_{i=1}^{n} (b_i - a_i - 2\varepsilon) \leq L^n(I) \leq \prod_{i=1}^{n} (b_i - a_i + 2\varepsilon).$$

Let $\varepsilon \to 0$, we obtain

$$L_n(I) = \prod_{i=1}^{n} (b_i - a_i) = |I|.$$

Finally,

$$L_n(\partial I) = L_n\left(\overline{I}\right) - L_n\left(\mathring{I}\right) = 0.$$

□

Theorem 7. *An arbitrary open set in $\mathbb{R}^n$ is a union of closed dyadic cubes with pairwise disjoint interiors. Hence Lebesgue measure of the open set equals the sum of the measures of these cubes.*

Theorem 8. *For an arbitrary set E,*

$$L_n^*(E) = \inf\left\{L_n(U) : U \textit{ is open and } E \subset U\right\}.$$

Proof. First, from monotonicity property of outer measure, we have for each open set U such that $E \subset U$,

$$L_n^*(E) \leq L_n(U),$$

hence

$$L_n^*(E) \leq \inf\left\{L_n(U) : U \text{ is open and } E \subset U\right\}.$$

Next, we use the definition (5.2) of outer measure. For any $\varepsilon > 0$, there exists a covering of E by bounded open intervals $\{I_i\}_{i=1}^{\infty}$, such that

$$L_n^*(E) \geq \sum_{i=1}^{\infty} |I_i| - \varepsilon \geq L_n(U) - \varepsilon$$

where $U = \bigcup_{i=1}^{\infty} I_i$ is open. Hence,

$$L_n^*(E) \geq \inf\left\{L_n(U) : U \text{ is open and } E \subset U\right\} - \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we have

$$L_n^*(E) \geq \inf\left\{L_n(U) : U \text{ is open and } E \subset U\right\}$$

which completes the proof. □

Theorem 8 holds for a much bigger class of measures. Actually, we have the following general result:

THEOREM 9. *Let X be a metric space and μ a measure in $\mathcal{B}(X)$. Suppose that X is a union of countably many open sets of finite measure. Then for any $E \in \mathcal{B}(X)$,*

$$\begin{aligned}\mu(E) &= \inf\{\mu(G) : G \textit{ is open and } E \subset G\} \\ &= \sup\{\mu(F) : F \textit{ is closed and } F \subset E\}.\end{aligned}$$

PROOF. We define a set function μ^* such that for any $E \subset X$,

$$\mu^*(E) = \inf\{\mu(G) : G \text{ is open and } E \subset G\}.$$

Then following the proof of Theorem 5, we can show that μ^* is a metric outer measure. Hence, any Borel sets are μ^*-measurable. We only need to show $\mu(E) = \mu^*(E)$ for any Borel set E.

Step 1: For any Borel set E, $\mu^*(E) \geq \mu(E)$. To see this, from the definition of μ^*, for any $\varepsilon > 0$, there exists open set G, $E \subset G$ and

$$\mu^*(E) \geq \mu(G) - \varepsilon.$$

Since $\mu(G) \geq \mu(E)$, we have $\mu^*(E) \geq \mu(E) - \varepsilon$. And $\mu^*(E) \geq \mu(E)$ follows by letting $\varepsilon \to 0$.

Step 2. For any open set E, $\mu^*(E) = \mu(E)$. Since E is open, from the definition of μ^*, we have $\mu^*(E) \leq \mu(E)$ Combining step 1, we have $\mu^*(E) = \mu(E)$.

Step 3. Let $X = \cup_{k=1}^{\infty} V_k$ where each V_k is an open set with finite μ-measure. Define, for each k, $U_k = \cup_{i=1}^{k} V_i$, then U_k is an open set with finite μ-measure and $X = \cup_{k=1}^{\infty} U_k$. Furthermore, U_k is monotone increasing in k. For any set E, from step 1, we have for each k,

$$\begin{aligned}\mu^*(U_k) &= \mu^*(U_k \cap E) + \mu^*(U_k \backslash E) \\ &\geq \mu(U_k \cap E) + \mu(U_k \backslash E) \\ &= \mu(U_k).\end{aligned}$$

However, we also have from step 2 $\mu^*(U_k) = \mu(U_k) < \infty$, hence, $\mu^*(U_k \cap E) = \mu(U_k \cap E)$. Let $k \to \infty$, Theorem 2 implies $\mu^*(E) = \mu(E)$.

$$\mu(E) = \sup\{\mu(F) : F \text{ is closed and } F \subset E\}$$

can be proved similarly. □

It is easy to see that any set of the form $B \cup E$, where $B \in \mathcal{B}(X)$ and $L_n^*(E) = 0$, is Lebesgue measurable. We will show that any Lebesgue measurable set can be expressed in this form. Let's recall the definition of G_δ and F_δ sets:

DEFINITION 8. *Let X be a metric space. By a G_δ set we mean a set of the form*

$$A = \bigcap_{i=1}^{\infty} G_i,$$

where the sets $G_i \subset X$ are open and by F_σ set we mean a set of the form

$$B = \bigcup_{i=1}^{\infty} F_i,$$

where the sets $F_i \subset X$ are closed.

Clearly all G_δ and F_δ sets are Borel.

THEOREM 10. *Let $A \subset \mathbb{R}^n$. Then the following statements are equivalent:*
1). A is Lebesgue measurable;
2). For every $\varepsilon > 0$ there is an open set G such that $A \subset G$ and $L_n^*(G \backslash A) < \varepsilon$;
3). There is a G_δ set H such that $A \subset H$ and $L_n^*(H \backslash A) = 0$;
4). For every $\varepsilon > 0$ there is a closed set F such that $F \subset A$ and $L_n^*(A \backslash F) < \varepsilon$;
5). There is a F_σ set M such that $M \subset A$ and $L_n^*(A \backslash M) = 0$;
6). For every $\varepsilon > 0$, there is an open set G and a closed set F such that $F \subset A \subset G$ and $L_n^*(G \backslash F) < \varepsilon$.

PROOF. Step 1: 2),3),4),5),6) implies 1) follows from the homework problem.
Step 2: 1) implies 2). Let A be Lebesgue measurable, we first assume $L_n(A) < \infty$. For any $\varepsilon > 0$, Theorem 8 implies

$$L_n(A) > L_n(G) - \varepsilon$$

for some open set G such that $A \subset G$. Since $L_n(G) = L_n(A) + L_n(G \backslash A)$, we have $L_n(G \backslash A) < \varepsilon$. If $L_n(A) = \infty$, we define $A_k = A \cap B_k$ where B_k is the open ball with radius k centered at the origin. Then A_k is Lebesgue measurable and $L_n(A_k) < \infty$. So for any $\varepsilon > 0$, there exists open set G_k such that $A_k \subset G_k$ and $L_n(G_k \backslash A_k) < \frac{\varepsilon}{2^k}$. Let $G = \bigcup_{k=1}^{\infty} G_k$. G is open and $A \subset G$. Since $G \backslash A \subset \bigcup_{k=1}^{\infty} (G_k \backslash A_k)$, we have

$$L_n(G \backslash A) \le \sum_{k=1}^{\infty} L_n(G_k \backslash A_k) < \varepsilon.$$

Step 3. 2) imples 3). For each $k \ge 1$, let G_k be an open set containing A such that $L_n(G_k \backslash A) < \frac{1}{k}$. Let $H = \bigcap_{k=1}^{\infty} G_k$, then H is a G_δ set and

$$L_n(G \backslash A) \le L_n(G_k \backslash A) < \frac{1}{k}$$

for each k, hence $L_n(G \backslash A) = 0$.
Step 4. 1) implies 4).

Step 5. 4) implies 5).

Step 6. 2) and 4) implies 6). □

THEOREM 11. *A set $E \subset \mathbb{R}^n$ has Lebesgue measure zero if and only if for every $\varepsilon > 0$, E can be covered by a family of open balls $\{B_{r_i}(x_i)\}$ such that*

$$\sum_{i=1}^{\infty} r_i^n < \varepsilon.$$

PROOF. "$\Longleftarrow$": Since $E \subset \bigcup_{i=1}^{\infty} B_{r_i}(x_i)$, we have

$$L_n^*(E) \le \sum_{i=1}^{\infty} L_n^*(B_{r_i}(x_i)) = \sum_{i=1}^{\infty} \omega_n r_i^n < \varepsilon \omega_n.$$

Let $\varepsilon \to 0$, we have
"$\Longrightarrow$": If $L_n^*(E) = 0$, there exists an open set U, such that $E \subset U$ and $L_n(U) < \varepsilon$. Since U is a union of closed dyadic cubes with pairwise disjoint interiors and each

of cube with side length l can be covered by an open ball with radius $r = \sqrt{n}l$. Hence we can cover U by a family of open balls $\{B_{r_i}(x_i)\}$ such that

$$\sum_{i=1}^{\infty}\left(\frac{r_i}{\sqrt{n}}\right)^n < \varepsilon,$$

i.e.,

$$\sum_{i=1}^{\infty} r_i^n < \left(\sqrt{n}\right)^n \varepsilon.$$

□

Theorem 12. *If* $f : \mathbb{R}^n \to \mathbb{R}^n$ *is Lipschitz continuous and the direct image of Lebesgue measure zero set has Lebesgue measure zero.*

Proof. Since $f : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous, there exists L such that for any $x, y \in \mathbb{R}^n$,

$$|f(x) - f(y)| \le L|x-y|.$$

Let $E \subset \mathbb{R}^n$ be a set with Lebesgue measure zero. Then E can be covered by a family of open balls $\{B_{r_i}(x_i)\}$ such that

$$\sum_{i=1}^{\infty} r_i^n < \varepsilon.$$

Since, $f(B_{r_i}(x_i)) \subset B_{Lr_i}(f(x_i))$, we have $f(E) \subset \bigcup_{i=1}^{\infty} B_{Lr_i}(f(x_i))$ and

$$\sum_{i=1}^{\infty}(Lr_i)^n < L^n\varepsilon.$$

Hence, $f(E)$ has Lebesgue measure zero. □

The Lebesgue measure is invariant under translations. The following theorem says that this property essentially defines Lebesgue measure.

Theorem 13. *If* μ *is a measure on* $\mathcal{B}(\mathbb{R}^n)$ *such that* $\mu(a+E) = \mu(E)$ *for all* $a \in \mathbb{R}^n$*,* $E \in \mathcal{B}(\mathbb{R}^n)$ *and* $\mu([0,1]^n) = 1$*. Then* $\mu(E) = L_n(E)$ *on* $\mathcal{B}(\mathbb{R}^n)$*.*

Proof. From Theorem 9, it suffices to show $\mu(U) = L_n(U)$ for any open set U. An arbitrary open set in $\mathbb{R}^n$ is a union of closed dyadic cubes with pairwise disjoint interiors. Each of the cube is a translation of a cube of the form $\left[0, \frac{1}{2^k}\right]^n$ for some k. Let $Q_k = \left(0, \frac{1}{2^k}\right)^n$. Since, $\mu(Q_1) \le 1$ and Q_1 contains the disjoint union of 2^{kn} cubes, each of the cubes is a translation of Q_k, we have

$$1 = \mu(Q_1) \ge 2^{kn}\mu(Q_k).$$

Hence, $\mu(Q_k) \le \frac{1}{2^{kn}}$. For each $k \ge 2$, since ∂Q_1 has $2n$ sides and each side can be covered by $2^{k(n-1)}$ cubes, each of the cubes is a translation of Q_{k-1}, we have

$$\mu(\partial Q_1) \le 2n \cdot 2^{k(n-1)} \cdot \frac{1}{2^{(k-1)n}} = \frac{n}{2^{k-n-1}}.$$

Let $k \to \infty$, we have $\mu(\partial Q_1) = 0$. Similarly, we can show $\mu(\partial Q_k) = 0$ for each k and hence $\mu\left(\overline{Q_k}\right) \le \frac{1}{2^{kn}}$. On the other hand, since Q_1 is a union of 2^{kn} closed cubes, each of the cubes is a translation of $\overline{Q_k}$, we have

$$1 = \mu(Q_1) \le 2^{kn}\mu\left(\overline{Q_k}\right)$$

and $\mu\left(\overline{Q_k}\right) \geq \frac{1}{2^{kn}}$. Hence $\mu\left(\overline{Q_k}\right) = \frac{1}{2^{kn}} = L_n\left(\overline{Q_k}\right)$. $\mu(U) = L_n(U)$ follows from the proof of Theorem 7. $\square$

The translation invariant property of the Lebesgue measure can be used to construct set which is not measurable.

THEOREM 14 (Vitali). *There is a set $E \subset \mathbb{R}^n$ which is not measurable.*

PROOF. We define an equivalence relation such that for any $x, y \in [0,1]^n$, $x \sim y$ if each component of $y - x$ is rational. Then $[0,1]^n$ is the union of a family of pairwise disjoint sets of the form

$$[x] = \{y \sim x : y \in [0,1]^n\}.$$

It follows from the Axiom of Choice that there is a set $E \subset [0,1]^n$ which contains exactly one element from each set in the family. We claim E is not measurable. Let Q be the collection of all points in $[-1,1]^n$ with rational components. For any $r \in Q$, we define $E_r = r + E$. If E is measurable, then E_r is measurable for each r and $L_n(E_r) = L_n(E)$. Now $\{E_r\}_{r\in Q}$ is pairwise disjoint and we have

$$[0,1]^n \subset \bigcup_{r\in Q} E_r \subset [-1,2]^n.$$

Hence

$$1 \leq \sum_{r\in Q} L_n(E_r) \leq 3^n$$

which is impossible. $\square$

REMARK 3. *Any set with positive Lebesgue outer measure contains a subset which is not measurable.*

Lebesgue measure is also rotational invariant. Actually, we have the following general result:

THEOREM 15. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a nondegenerate linear transformation represented by the matrix A, then $f(E)$ is Lebesgue measurable if and only if E is Lebesgue measurable. Moreover,*

$$L_n(f(E)) = |\det A|\, L_n(E). \tag{5.5}$$

PROOF. Since f is a homeomorphism, it preserves the class of Borel sets. Since both f and f^{-1} are Lipschitz continuous, f preserves the class of sets of measure zero. Therefore f preserves the class of Lebesgue measurable sets. We only need to verify (5.5) for Borel sets since $L_n(E) = 0$ implies $L_n(f(E)) = 0$. Define a new set function

$$\mu(E) = L_n(f(E)).$$

Then it is easy to verify that μ is a translation invariant measure on Borel sets of $\mathbb{R}^n$. Let $a = \mu([0,1]^n)$. Then $\frac{1}{a}\mu$ satisfies

$$\left(\frac{1}{a}\mu\right)([0,1]^n) = 1.$$

So (13) implies $\frac{1}{a}\mu = L_n$. We need to check $a = |\det A|$. We view $a : GL(n) \to (0,\infty)$ as a function defined on the class of invertible matrices. We have

$$a(A_1 A_2) = a(A_1)\, a(A_2).$$

Any invertible matrix A can be decomposed into a product of elementary matrices

$$A = T_1 T_2 T_3 \cdots T_k$$

where T_i is one of the following three types of matrices

$$\text{Type I: } \begin{pmatrix} c & \\ & I_{n-1} \end{pmatrix},$$

$$\text{Type II: } \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & \\ 0 & 1 & & & 0 & \\ \vdots & & \ddots & & \vdots & \\ 0 & & & 1 & 0 & \\ 1 & 0 & \cdots & 0 & 0 & \\ & & & & & I_{n-i} \end{pmatrix}$$

$$\text{Type III: } \begin{pmatrix} 1 & 1 & \\ 0 & 1 & \\ & & I_{n-2} \end{pmatrix}.$$

Here Type I matrices multiply the first row of a matrix by a nonzero number c; type II matrices interchange the first row of a matrix with its i-th row; the type III matrix adds the second row of a matrix to its first row. It suffices to prove $a(A) = |\det A|$ for these three types of elementary matrices. If T is of type I, then $a(T) = |\det T| = |c|$ follows from the fact that $T([0,1]^n)$ is a rectangular box with Lebesgue measure $|c|$. If T is of type II, since $T^2 = I$, we have $a^2(T) = a\left(T^2\right) = 1$, so we have $a(T) = 1 = |\det T|$ also. If T is of type III, we see $T([0,1]^n) = D \times [0,1]^{n-2}$ where

$$D = \left\{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_2 \leq 1, x_2 \leq x_1 \leq x_2 + 1\right\}.$$

Using the translation invariant property of Lebesgue measure, we can show

$$L_n\left(D \times [0,1]^{n-2}\right) = L_n([0,1]^n) = 1.$$

Hence $a(T) = 1 = |\det T|$. $\square$

CHAPTER 2

Integration

1. Measurable Functions

We first recall the continuity of functions:

Definition 9. *Let X, Y be two topological spaces, we say a function $f : X \to Y$ is continuous if $f^{-1}(U)$ is open in X whenever U is open in Y.*

Measurable functions are defined in a similar way.

Definition 10. *Let $(X, \mathfrak{M})$ be a measurable space and Y be a topological space, we say a function $f : X \to Y$ is measurable if $f^{-1}(U) \in \mathfrak{M}$ whenever U is open in Y.*

Example 4. *If $E \subset X$ is measurable, then the characteristic function*

$$\chi_E : E \to \mathbb{R}$$

is measurable since for any open set $U \subset \mathbb{R}$, $f^{-1}(U)$ must be one of the four measurable set $\emptyset, E, E^c, X$.

If $E \subset X$ is a measurable set, then we say that $f : E \to Y$ is measurable if $f^{-1}(U) \in \mathfrak{M}$ whenever U is open in Y. It is easy to check that

$$\mathfrak{M}_E = \{A \cap E : A \in \mathfrak{M}\},$$

is a σ-algebra and measurability of $f : E \to Y$ is equivalent to measurability of f with respect to $\mathfrak{M}_E$.

If X itself is a topological space and $\mathfrak{M} = \mathfrak{B}(X)$, then a measurable function $f : X \to Y$ is said to be Borel. i.e., f is a Borel function if $f^{-1}(U) \in \mathfrak{B}(X)$ whenever U is open in Y. Especially, any continuous function is Borel.

Theorem 16. *Let $(X, \mathfrak{M})$ be a measurable space, $f : X \to Y$ a measurable function and $g : Y \to Z$ a continuous function. Then the function $g \circ f : X \to Z$ is measurable.*

Proof. For any open set U of Z, we have

$$(g \circ f)^{-1}(U) = f^{-1}\left(g^{-1}(U)\right)$$

which is measurable. $\square$

Example 5. *If $f : X \to \mathbb{C}$ is measurable, then the functions $|f|$ is measurable.*

Example 6. *If $f : X \to \mathbb{R}$ is measurable, then for any polynomial $P : \mathbb{R} \to \mathbb{R}$, $P(f) : X \to \mathbb{R}$ is measurable.*

THEOREM 17. *Let* $f : X \to Y$ *where* $(X, \mathfrak{M})$ *is a measurable space and* Y *is a topological space and*
(a). The set

$$\Omega = \left\{E \subset Y : f^{-1}(E) \in \mathfrak{M}\right\}$$

is a σ*-algebra.*
(b). If f *is measurable and* $E \subset Y$ *is a Borel set, then* $f^{-1}(E) \in \mathfrak{M}$*.*
(c). If $Y = \mathbb{R}$ *and* $f^{-1}((\alpha, \infty)) \in \mathfrak{M}$ *for any* $\alpha \in \mathbb{R}$*, then* f *is measurable.*
(d). If f *is measurable, if* Z *is another topological space and* $g : Y \to Z$ *is a Borel mapping, then* $g \circ f : X \to Z$ *is measurable.*

PROOF. (a). Direct verification.

(b). Since f is measurable, Ω contains all open sets in Y. Since Ω is a σ-algebra, it contains all Borel sets. Hence, $E \subset Y$ is a Borel set implies $f^{-1}(E) \in \mathfrak{M}$.

(c). Ω is a σ-algebra containing all sets of the form (α, ∞) with $\alpha \in \mathbb{R}$. For any $\beta > \alpha$, we have

$$(\alpha, \beta] = (\alpha, \infty) \setminus (\beta, \infty) \in \Omega,$$

and

$$(\alpha, \beta) = \bigcup_{k=1}^{\infty} \left(\alpha, \beta - \frac{1}{k}\right] \in \Omega,$$

$$(-\infty, \alpha) = \bigcup_{k=1}^{\infty} (-k, \alpha) \in \Omega.$$

Hence Ω contains all open intervals. Since any open set in $\mathbb{R}$ is a countable union of open intervals, Ω contains all open sets. Hence, f is measurable.

(d). For any open set U of Z, we have

$$(g \circ f)^{-1}(U) = f^{-1}\left(g^{-1}(U)\right)$$

which is measurable since $g^{-1}(U)$ is Borel and (b). □

REMARK 4. *Let* $f : X \to Y$ *where* $(X, \mathfrak{M})$ *is a measurable space and* (Y, τ) *is a topological space. If* $\mathfrak{F}$ *is a collection of subsets of* Y *such that*

$$\mathfrak{F} \subset \Omega = \left\{E \subset Y : f^{-1}(E) \in \mathfrak{M}\right\} \text{ and } \tau \subset \sigma(\mathfrak{F}),$$

then f *is measurable.*

THEOREM 18. *Let* $(X, \mathfrak{M})$ *be a measurable space and* Y *a topological space.Let* $u, v : X \to \mathbb{R}$ *be measurable and* $\Phi : \mathbb{R}^2 \to Y$ *be continuous. then the function* $h(x) = \Phi(u(x), v(x)) : X \to Y$ *is measurable.*

PROOF. Let $f(x) = (u(x), v(x)) : X \to \mathbb{R}^2$. Since $h = \Phi \circ f$, it suffices to prove that f is measurable. For any open rectangle $R = (a, b) \times (c, d)$, we have

$$f^{-1}(R) = u^{-1}((a, b)) \cap v^{-1}((c, d)) \in \mathfrak{M}.$$

Since any open set in $\mathbb{R}^2$ is a countable union of bounded open rectangles, we conclude f is measurable. □

COROLLARY 2. *If* $f, g : X \to \mathbb{R}$ *are measurable,then* $f \pm g$ *and* fg *are measurable.*

THEOREM 19. *If* $f : X \to \mathbb{C}$ *is measurable, then there exists a complex measurable function* $\alpha : X \to \mathbb{C}$ *such that* $|\alpha| = 1$ *and* $f = \alpha |f|$*.*

The set $[-\infty,\infty]=\mathbb{R}\cup\{-\infty,\infty\}$ is called the extended real line. $[-\infty,\infty]$ is a topological space such that a set $E\in[-\infty,\infty]$ is said to be open if it is of the form $(a,b)\,,[-\infty,a)\,,(a,\infty]$ or any union of segments of the type.

THEOREM 20. *Let* $(X,\mathfrak{M})$ *be a measurable space. Extended real function* $f: X\to[-\infty,\infty]$ *is measurable if and only if* $f^{-1}\left((\alpha,\infty]\right)\in\mathfrak{M}$ *for any* $\alpha\in\mathbb{R}$.

DEFINITION 11. *Let* $\{a_n\}$ *be a sequence in* $[-\infty,\infty]$ *and put*

$$b_k=\sup\{a_i\}_{i=k}^{\infty}\,,\beta=\inf\{b_k\}_{k=1}^{\infty}\,.$$

We call β *the upper limit of* $\{a_n\}$ *and write*

$$\beta=\limsup_{n\to\infty}a_n.$$

Since b_k is monotone non-increasing in k, we have $\beta=\lim_{n\to\infty}b_k$.

EXERCISE 1. *There exists a subsequence* $\{a_{n_i}\}$ *of* $\{a_n\}$*, such that* $\beta=\lim_{i\to\infty}a_{n_i}$.

Similarly, we have the following definition:

DEFINITION 12. *Let* $\{a_n\}$ *be a sequence in* $[-\infty,\infty]$ *and put*

$$b_k=\inf\{a_i\}_{i=k}^{\infty}\,,\beta=\sup\{b_k\}_{k=1}^{\infty}\,.$$

We call β *the lower limit of* $\{a_n\}$ *and write*

$$\beta=\liminf_{n\to\infty}a_n.$$

Since b_k is monotone nondecreasing in k, we have $\beta=\lim_{n\to\infty}b_k$.

REMARK 5. *Let* $\{a_n\}$ *be a sequence in* $[-\infty,\infty]$*, we have*

$$\liminf_{n\to\infty}a_n=-\limsup_{n\to\infty}\left(-a_n\right).$$

REMARK 6. *Let* $\{a_n\}$ *be a sequence in* $[-\infty,\infty]$*. Assume* $\{a_{n_i}\}$ *is a subsequence such that* $\lambda=\lim_{i\to\infty}a_{n_i}$ *is defined. Then we have*

$$\liminf_{n\to\infty}a_n\le\lambda\le\limsup_{n\to\infty}a_n.$$

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of extended real functions on X. We can define functions

$$\left(\sup_n f_n\right)(x)=\sup_n f_n(x)\,,\left(\inf_n f_n\right)(x)=\inf_n f(x)$$

$$\left(\limsup_{n\to\infty}f_n\right)(x)=\limsup_{n\to\infty}f_n(x)\,,\left(\liminf_{n\to\infty}f_n\right)(x)=\liminf_{n\to\infty}f_n(x)\,.$$

If

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

exists for every $x\in X$, then we call f the pointwise limit of the sequence $\{f_n\}$. In this case,

$$f=\limsup_{n\to\infty}f_n=\liminf_{n\to\infty}f_n.$$

THEOREM 21. *If* $f_n: X\to[-\infty,\infty]$ *is measurable, then*

$$\sup_n f_n,\ \inf_n f_n,\ \limsup_{n\to\infty}f_n\ \text{and}\ \liminf_{n\to\infty}f_n$$

are measurable.

Proof. For any $\alpha \in \mathbb{R}$,

$$\left(\sup_n f_n\right)^{-1}((\alpha,\infty]) = \bigcup_{n=1}^{\infty} f_n^{-1}((\alpha,\infty])$$

is measurable. Hence, $\sup_n f_n$ is measurable. Since $\inf_n f_n = -\sup_n(-f_n)$, $\inf_n f_n$ is measurable. Now,

$$\limsup_{n\to\infty} f_n = \inf_{n\geq 1}\left(\sup_{k\geq n} f_k\right), \liminf_{n\to\infty} f_n = \sup_{n\geq 1}\left(\inf_{k\geq n} f_k\right).$$

So $\limsup_{n\to\infty} f_n$ and $\liminf_{n\to\infty} f_n$ are measurable. □

Corollary 3. *If f, g are measurable extended real functions on X, then so are* $\max\{f,g\}$ *and* $\inf\{f,g\}$. *In particular,*

$$f^+ = \max\{f,0\} \text{ and } f^- = -\min\{f,0\}$$

are measurable.

For pointwise limit of measurable functions, we have a very general result.

Theorem 22. *The pointwise limit of a sequence of measurable functions from a measurable space into a metric space is measurable.*

Proof. Let $f_n : X \to Y$ and Y be a metric space. For any open set $U \subset Y$, we define for each $n \geq 1$, the set

$$U_n = \left\{x \in Y : \text{dist}(x, U^c) > \frac{1}{n}\right\}.$$

Then U_n is open and $U_n \subset U$. We claim

$$f^{-1}(U) = \bigcup_{n=1}^{\infty}\bigcup_{l=1}^{\infty}\bigcap_{k=l}^{\infty} f_k^{-1}(U_n).$$

To see this, $x \in f^{-1}(U)$ if and only if $f(x) \in U$ and $f(x) \in U$ if and only if there exists n, l such that $f_k(x) \in U_n$ for each $k \geq l$. Since f_k is measurable, $f_k^{-1}(U_n)$ is a measurable set and hence $f^{-1}(U)$ is measurable. □

Definition 13. *A complex function s on a measurable space X whose range consists of only finitely many points will be called a simple function.*

If $\alpha_i, 1 \leq i \leq n$ are the distinct nonzero values of a simple function s, if we set

$$A_i = \{x : s(x) = \alpha_i\},$$

then

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

s is measurable if and only if each A_i is measurable.

Theorem 23. *Let $f : X \to [0,\infty]$ be measurable. There exist simple measurable functions s_n on X such that*
(a). $0 \leq s_1 \leq s_2 \leq \cdots \leq f$,
(b). $s_n(x) \to f(x)$ *as* $n \to \infty$, *for every* $x \in X$.

Proof. We define

$$\varphi_n(t) = \begin{cases} \frac{[2^n t]}{2^n} & \text{if} \quad 0 \le t < n, \\ n & \text{if} \quad n \le t \le \infty \end{cases}$$

where $[x]$ is the integer part of x. Then $\varphi_n(t)$ is monotone increasing in n and $\lim_{n\to\infty} \varphi_n(t) = t$ for any $t \in [0,\infty]$. Furthermore, $\varphi_n : [0,\infty] \to [0,\infty]$ is a Borel function. Let

$$s_n(x) = \varphi_n(f(x)),$$

then s_n is a measurable simple function which satisfies all the requirement. □

Remark 7. *If f is bounded, then s_n constructed above converges to f uniformly.*

2. Lebesgue Integrals

Throughout integration theory, one inevitably encounters ∞. We define

$$a+\infty=\infty+a=\infty \text{ for any } a\in[0,\infty],$$

and

$$a\cdot\infty=\infty\cdot a=\begin{cases}\infty & \text{if} \quad 0<a\leq\infty\\ 0 & \text{if} \quad a=0.\end{cases}$$

Under this definition, one can check that if as $n\to\infty$, $0\leq a_n\nearrow a$ and $0\leq b_n\nearrow b$, then

$$a_nb_n\to ab \text{ as } n\to\infty.$$

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

Definition 14. *Let*

$$s=\sum_{i=1}^{n}\alpha_i\chi_{A_i} \tag{2.1}$$

be a measurable simple function from X into $[0,\infty)$. Here $\alpha_i, 1\leq i\leq n$ are distinct values of s. If $E\in\mathfrak{M}$, then we define the integral of s over E

$$\int_E sd\mu\equiv\sum_{i=1}^{n}\alpha_i\mu\left(A_i\cap E\right).$$

If $f:X\to[0,\infty]$ is measurable, and $E\in\mathfrak{M}$, we define the Lebesgue integral of f over E w.r.t. μ,

$$\int_E fd\mu=\sup\int_E sd\mu.$$

where the supremum being taken on all simple measurable functions s such that $0\leq s\leq f$.

The following properties of Lebesgue integral can be easily deduced from the definition:

Proposition 3. *(a). If $0\leq f\leq g$, then*

$$\int_E fd\mu\leq\int_E gd\mu.$$

(b). If $A\subset B$ and $f\geq0$, then

$$\int_A fd\mu\leq\int_B fd\mu.$$

(c). If $f\geq0$ and $c\in[0,\infty)$ is a constant, then

$$\int_E cfd\mu\leq c\int_E fd\mu.$$

(d). If $f=0$ for all $x\in E$, then

$$\int_E fd\mu=0.$$

(e). If $\mu(E)=0$, then

$$\int_E fd\mu=0.$$

(f). If $f \geq 0$*, then*

$$\int_E f d\mu = \int_X f \chi_E d\mu.$$

LEMMA 1. *Let* $s, t : X \to [0, \infty)$ *be two simple measurable functions. For* $E \in \mathfrak{M}$*, we define*

$$\varphi(E) = \int_E s d\mu.$$

Then φ *is a measure on* $\mathfrak{M}$*. Furthermore,*

$$\int_X (s+t)\, d\mu = \int_X s d\mu + \int_X t d\mu. \tag{2.2}$$

PROOF. Let

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

We have

$$\varphi(E) = \int_E s d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E).$$

Obviously $\varphi(\emptyset) = 0$. So it suffices to check countable additivity. Let $\{E_k\}_{k=1}^{\infty}$ be a disjoint collection of measurable set, we have from the countable additivity of μ,

$$\begin{aligned}
\varphi\left(\bigcup_{k=1}^{\infty} E_k\right) &= \sum_{i=1}^{n} \alpha_i \mu\left(A_i \bigcap \left(\bigcup_{k=1}^{\infty} E_k\right)\right) \\
&= \sum_{i=1}^{n} \alpha_i \sum_{k=1}^{\infty} \mu\left(A_i \bigcap E_k\right) = \sum_{k=1}^{\infty} \sum_{i=1}^{n} \alpha_i \mu\left(A_i \bigcap E_k\right) \\
&= \sum_{k=1}^{\infty} \varphi(E_k).
\end{aligned}$$

Hence, φ is a measure on $\mathfrak{M}$. Next, let

$$t = \sum_{j=1}^{m} \beta_j \chi_{B_j}.$$

Then we have $s + t = \alpha_i + \beta_j$ on $E_{ij} = A_i \cap B_j$, and

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

Since X is the disjoint union of $E_{ij}, 1 \leq i \leq n, 1 \leq j \leq m$, (2.2) follows from the additivity of φ w.r.t. $s + t, s$ and t. □

THEOREM 24 (Lebesgue's Monotone Convergence Theorem). *Let* $f_n : X \to [0, \infty]$ *be a sequence of measurable functions such that*

$$f_n \nearrow f \text{ pointwisely as } n \to \infty.$$

Then f *is measurable, and*

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

PROOF. f is measurable because it is pointwise limit of measurable functions. Since $\int_X f_n d\mu$ is monotone increasing,

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu$$

is a well defined extended real number. And since $f_n \leq f$ for each n, we have

$$\alpha \leq \int_X f d\mu. \tag{2.3}$$

Let s be any simple measurable function such that $0 \leq s \leq f$, let $c \in (0,1)$ be a constant. Define for each n,

$$E_n = \{x : f_n(x) \geq cs(x)\}.$$

Then E_n is measurable, monotone increasing in n and $X = \bigcup_{n=1}^{\infty} E_n$. And we have

$$\int_X f_n d\mu \geq \int_{E_n} f d\mu \geq c \int_{E_n} s d\mu.$$

Let $n \to \infty$, we have

$$\lim_{n\to\infty} \left(\int_X f_n d\mu\right) \geq c \int_X s d\mu$$

here we used Lemma 1. Taking supremum on all such s, we have

$$\lim_{n\to\infty} \left(\int_X f_n d\mu\right) \geq c \int_X f d\mu.$$

Since c can be arbitrarily close to 1, we have

$$\lim_{n\to\infty} \left(\int_X f_n d\mu\right) \geq \int_X f d\mu. \tag{2.4}$$

The Theorem follows from (2.3) and (2.4). □

THEOREM 25. *Let $f_n : X \to [0,\infty]$, $n \in \mathbb{N}$ be a sequence of measurable functions and for each $x \in X$,*

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

Then

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

PROOF. See Theorem 1.27 in Rudin's book. □

THEOREM 26 (Fatou's Lemma). *Let $f_n : X \to [0,\infty]$, $n \in \mathbb{N}$ be a sequence of measurable functions. Then*

$$\int_X \left(\liminf_{n\to\infty} f_n\right) d\mu \leq \liminf_{n\to\infty} \left(\int_X f_n d\mu\right).$$

PROOF. See Theorem 1.28 in Rudin's book. □

Theorem 27. *Let $f : X \to [0, \infty]$ be measurable. For $E \in \mathfrak{M}$, we define*

$$\varphi(E) = \int_E f d\mu.$$

Then φ is a measure on $\mathfrak{M}$. And

$$\int_X g d\varphi = \int_X g f d\mu$$

for every measurable function $g : X \to [0, \infty]$.

Proof. See Theorem 1.29 in Rudin's book. $\square$

Definition 15. *We define $L^1(\mu)$ to be the collection of all complex measurable functions f on X for which*

$$\|f\|_1 = \int_X |f| \, d\mu < \infty.$$

We say f is Lebesgue integrable if $f \in L^1(\mu)$.

Remark 8. *We could use $L^1(\mu, [-\infty, \infty])$ to denote all extended real measurable functions f on X for which*

$$\|f\|_1 = \int_X |f| \, d\mu < \infty.$$

Definition 16. *If $f \in L^1(\mu)$, we can write $f = u + iv$ where u, v are real functions. We define for any $E \in \mathfrak{M}$,*

$$\int_E f d\mu = \int_E u^+ d\mu - \int_E u^- d\mu + i \left(\int_E v^+ d\mu - \int_E v^- d\mu \right).$$

Remark 9. *Here $\int_E u^+ d\mu, \int_E u^- d\mu, \int_E v^+ d\mu, \int_E v^- d\mu$ are all finite when $f \in L^1(\mu)$.*

Theorem 28. *Suppose $f, g \in L^1(\mu)$ and $\alpha, \beta \in \mathbb{C}$ are constants. Then $\alpha f + \beta g \in L^1(\mu)$, and*

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

Proof. See Theorem 1.32 in Rudin's book. $\square$

Similarly we define $L^1(\mu, \mathbb{R}^n)$ as the class of all measurable mappings $f = (f_1, f_2, \cdots, f_n) : X \to \mathbb{R}^n$ such that

$$\|f\|_1 = \int_X |f| \, d\mu < \infty.$$

And we set

$$\int_X f d\mu = \left(\int_X f_1 d\mu, \int_X f_2 d\mu, \cdots, \int_X f_n d\mu \right).$$

Theorem 29. *If $f \in L^1(\mu, \mathbb{R}^n)$ or $f \in L^1(\mu, \mathbb{C})$, then*

$$\left| \int_X f d\mu \right| \le \int_X |f| \, d\mu.$$

PROOF. See Theorem 1.33 in Rudin's book for $f \in L^1(\mu, \mathbb{C})$. If $f \in L^1(\mu, \mathbb{R}^n)$, we define $\alpha = \int_X f d\mu$, then

$$|\alpha|^2 = \left\langle \alpha, \int_X f d\mu \right\rangle = \int_X \langle \alpha, f \rangle \, d\mu \leq |\alpha| \int_X |f| \, d\mu = |\alpha| \int_X |f| \, d\mu.$$

□

THEOREM 30 (Lebesgue's Dominated Convergence Theorem). *Let $f_n : X \to \mathbb{C}$, $n \in \mathbb{N}$ be a sequence of measurable functions such that*

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

exists for every $x \in X$. If for some $g \in L^1(\mu)$,

$$|f_n| \leq g$$

holds for every n, then $f \in L^1(\mu)$,

$$\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0, \tag{2.5}$$

and

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \left(\lim_{n\to\infty} f_n \right) d\mu = \int_X f d\mu. \tag{2.6}$$

PROOF. Since $|f| \leq g$ and $g \in L^1(\mu)$, we have $f \in L^1(\mu)$. Since $|f_n - f| \leq 2g$, we can apply Fatou's lemma to $2g - |f_n - f|$ and yields

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

Hence,

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

which implies (2.5). Finally,

$$0 \leq \left| \int_X (f_n - f) \, d\mu \right| \leq \int_X |f_n - f| \, d\mu \to 0$$

which implies (2.6). □

3. Sets of Zero or Small Measure

Definition 17. *Let $E \in \mathfrak{M}$. We say "Property P holds almost everywhere on E" if there exists a measurable subset N of E, such that $\mu(N) = 0$ and P holds everywhere on $E \backslash N$.*

Example 7. *We say $f_n \to f$ a.e. on X if there is a set $N \in \mathfrak{M}$, $\mu(N) = 0$ and $f_n(x) \to f(x)$ for all $x \in X \backslash N$.*

Example 8. *We say two measurable functions $f = g$ a.e. on X if the set $\{x \in X : f(x) \neq g(x)\}$ has measure zero. We write $f \sim g$ which defines an equivalence relation.*

Definition 18. *μ is said to be a complete measure, if all subsets of zero measure set are measurable.*

Any measure constructed from an outer measure is complete. More generally, any measure can be transformed into a complete measure through expanding its measurable space.

Theorem 31. *Let $(X, \mathfrak{M}, \mu)$ be a measurable space, let $\mathfrak{M}^*$ be the collection of all $E \subset X$ for which there exist sets $A, B \in \mathfrak{M}$ such that $A \subset E \subset B$ and $\mu(B \backslash A) = 0$, and define $\mu(E) = \mu(A)$ in this situation. Then $\mathfrak{M}^*$ is a σ-algebra, and μ is a complete measure on $\mathfrak{M}^*$.*

Hence, we may assume a measure is complete whenever it is convenient. The general rule for completed measure is that sets of measure zero are usually negligible.

For example, we have the following version of Lebesgue's Dominated Convergence Theorem in which exceptional sets of measure zero are admitted.

Theorem 32 (Lebesgue's Dominated Convergence Theorem). *Let $f_n : X \to \mathbb{C}$, $n \in \mathbb{N}$ be a sequence of measurable functions such that*

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

exists for almost every $x \in X$. If for some $g \in L^1(\mu)$,

$$|f_n| \leq g$$

holds a.e. in X for every n, then $f \in L^1(\mu)$,

$$\lim_{n\to\infty} \int_X |f_n - f| \, d\mu = 0, \tag{3.1}$$

and

$$\int_X f d\mu = \int_X \left(\lim_{n\to\infty} f_n\right) d\mu = \lim_{n\to\infty} \int_X f_n d\mu. \tag{3.2}$$

If $f = g$ a.e. on X, then for any $E \in \mathfrak{M}$,

$$\int_E f d\mu = \int_E g d\mu.$$

On the other hand, we have

Theorem 33. *(a). Suppose $f : X \to \mathbb{C}$ is measurable and $E \in \mathfrak{M}$, and $\int_E |f| \, d\mu = 0$. Then $f = 0$ a.e. on E.*
(b). Suppose $f \in L^1(\mu, \mathbb{C})$ and $\int_E f d\mu = 0$ for any $E \in \mathfrak{M}$. Then $f = 0$ a.e. on X.

Proof. (a). Define for each $n \in \mathbb{N}$,

$$E_n = \left\{ x \in E : |f| \geq \frac{1}{n} \right\}.$$

Then we have

$$\int_E |f| \, d\mu \geq \int_{E_n} |f| \, d\mu \geq \int_{E_n} \frac{1}{n} d\mu = \frac{1}{n} \mu(E_n).$$

Hence $\mu(E_n) = 0$ for each n, and which implies

$$\mu\left(\bigcup_{n=1}^{\infty} E_n \right) = 0.$$

Since $f = 0$ on $X \backslash \bigcup\limits_{n=1}^{\infty} E_n$, we have $f = 0$ a.e. on E.

(b). We write $f = u + iv$. Let $E_1 = \{x \in X : u(x) \geq 0\}$, then we have

$$0 = \int_{E_1} f d\mu = \int_{E_1} u d\mu + i \int_{E_1} v d\mu.$$

Hence we have $\int_{E_1} u d\mu = 0$. Since $u \geq 0$ on E_1, (a) implies $u = 0$ a.e. on E_1. Similarly, one can prove $u = 0$ a.e.on E_1^c since $u \leq 0$ on E_1^c. Hence, $u = 0$ a.e. on X. $v = 0$ a.e. on X can be proved similarly. Hence $f = u + iv = 0$ a.e. on X. ☐

A topological space is called separable if it contains a countable dense subset.

Theorem 34. *If Y is a separable metric space and $f : \mathbb{R}^n \to Y$ is Lebesgue measurable, then there is a Borel mapping $g : \mathbb{R}^n \to Y$ such that $f = g$ a.e..*

Next, we define convergence of functions in measure which is also called convergence in probability when μ is a probability measure.

Definition 19. *We say that a sequence of measurable functions $f_n : X \to \mathbb{R}$ converges in measure to a measurable function $f : X \to \mathbb{R}$ if for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mu(\{x \in X : |f(x) - f_n(x)| \geq \varepsilon\}) = 0.$$

And we write

$$f_n \xrightarrow{\mu} f.$$

Theorem 35. *If $\mu(X) < \infty$ and a sequence of measurable functions $f_n : X \to \mathbb{R}$ converges to f a.e., then it also converges in measure.*

Proof. Given $\varepsilon > 0$, we define

$$E_n = \{x \in X : |f(x) - f_n(x)| \geq \varepsilon\}.$$

If $f_n(x) \to f(x)$ as $n \to \infty$, then

$$x \in \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} E_k^c.$$

Since $f_n \to f$ a.e., we have

$$0 = \mu\left(X \backslash \left(\bigcup_{n=1}^{\infty} \bigcap_{k\geq n} E_k^c \right) \right) = \mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k\geq n}^{\infty} E_k \right).$$

Since $\mu(X) < \infty$, Theorem 2 implies

$$\lim_{n\to\infty} \mu\left(\bigcup_{k\geq n}^{\infty} E_k\right) = 0.$$

Since $E_n \subset \bigcup\limits_{k\geq n}^{\infty} E_k$, we have

$$\lim_{n\to\infty} \mu(E_n) = 0.$$

i.e., $f_n \xrightarrow{\mu} f$. $\square$

Remark 10. *$\mu(X) < \infty$ is necessary. $f_n = \chi_{[n,n+1]}$ provides a counter example when $X = \mathbb{R}$ and μ is Lebesgue measure.*

Theorem 36 (Lebesgue). *If a sequence of measurable functions*

$$f_n \xrightarrow{\mu} f.$$

Then there is a subsequence $f_{n_i} \to f$ a.e..

Proof. For every i, there exists n_i s.t. $\mu(E_i) \leq \frac{1}{2^i}$ where

$$E_i = \left\{x \in X : |f(x) - f_{n_i}(x)| \geq \frac{1}{i}\right\}.$$

We can assume that n_i is monotone increasing. Now we define for each k the set

$$F_k = \bigcap_{i=k}^{\infty} E_i^c = X \backslash \bigcup_{i=k}^{\infty} E_i.$$

Then $f_{n_i}(x) \to f(x)$ uniformly on each F_k, and hence $f_{n_i}(x) \to f(x)$ pointwisely on $\bigcup\limits_{k=1}^{\infty} F_k$. Since

$$\mu(X\backslash F_k) = \mu\left(\bigcup_{i=k}^{\infty} E_i\right) \leq \sum_{i=k}^{\infty} \frac{1}{2^i} = \frac{1}{2^{k-1}},$$

we have $\mu\left(X\backslash\left(\bigcup\limits_{k=1}^{\infty} F_k\right)\right) = 0$. Hence, $f_{n_i} \to f$ a.e.. $\square$

Remark 11. *In general, $f_n \xrightarrow{\mu} f$ may not imply $f_n \to f$ a.e.. Intuitively, $f_n \xrightarrow{\mu} f$ allows the bad set jumping around to cover whole X which is not permitted if $f_n \to f$ a.e..*

Theorem 37 (Egorov's Theorem). *If $\mu(X) < \infty$ and a sequence of real valued measurable functions f_n converges to f a.e. on X. Then for any $\varepsilon > 0$, there exists a measurable set E such that $\mu(X\backslash E) \leq \varepsilon$ and $f_n \Longrightarrow f$ uniformly on E.*

Proof. Let

$$E_{n,k} = \left\{x \in X : |f(x) - f_n(x)| \geq \frac{1}{k}\right\}$$

and

$$F_{n,k} = \bigcup_{m\geq n} E_{m,k}.$$

Then $F_{n,k}$ is monotone decreasing in n. Now $f_n(x) \to f(x)$ implies

$$x \notin \bigcap_{n=1}^{\infty} F_{n,k}.$$

Since $f_n \to f$ a.e. and since $\mu(X) < \infty$, we have

$$\lim_{n\to\infty} \mu(F_{n,k}) = \mu\left(\bigcap_{n=1}^{\infty} F_{n,k}\right) = 0.$$

Hence, given $\varepsilon > 0$, there exists n_k such that

$$\mu(F_{n_k,k}) \le \frac{\varepsilon}{2^k}.$$

Now let

$$E = X \backslash \left(\bigcup_{k=1}^{\infty} F_{n_k,k}\right) = \bigcap_{n=1}^{\infty} F_{n_k,k}^c.$$

Then $\mu(X\backslash E) \le \varepsilon$. Now for $x \in E$, we have

$$x \in F_{n_k,k}^c = \bigcap_{m \ge n_k}^{\infty} E_{m,k}^c$$

for each k, i.e., for any $m \ge n_k$,

$$|f(x) - f_m(x)| < \frac{1}{k}.$$

Hence, $f_n \Longrightarrow f$ uniformly on E. □

Egorov's Theorem is of the same spirit as Lusin's Theorem. Both of them can be generalized to the situation when the target space is a separable metric space. For example, we have

Theorem 38 (Lusin's Theorem). *Let X be a metric space and μ a measure in $\mathfrak{B}(X)$ such that X is a union of countably many open sets of finite measure. If $f : X \to Y$ is a Borel mapping with values in a separable metric space, then for every $\varepsilon > 0$ there is a closed set $F \subset X$ such that $\mu(X\backslash F) < \varepsilon$ and f is continuous on F.*

We leave the proof to interested students.

4. Riesz Representation Theorem

Let X be a topological space.

Definition 20. *X is a Hausdorff space if the following is true: If $p \in X$, $q \in X$ and $p \neq q$, then p has a neighborhood U and q has a neighborhood V s.t. $U \cap V = \emptyset$.*

Any metric space is a Hausdorff space.

Definition 21. *A set $K \subset X$ is compact if every open cover of K contains a finite subcover. X is said to be locally compact if every point of X has a neighborhood whose closure is compact.*

The Euclidean space $\mathbb{R}^n$ is locally compact since any closed and bounded subset of $\mathbb{R}^n$ is compact.

Definition 22. *The support of a real function f on X, denoted by* $\operatorname{supp} f$, *is the closure of the set*

$$\{x : f(x) \neq 0\}.$$

The collection of all continuous real functions on X whose support is compact is denoted by $C_c(X)$.

$C_c(X)$ is a vector space. We say a linear functional $\Lambda : C_c(X) \to \mathbb{R}$ is positive if

$$\Lambda f \geq 0 \text{ whenever } f \geq 0.$$

Definition 23. *Let X be a Hausdorff space. A measure μ on the Borel sets $\mathfrak{B}(X)$ that satisfies $\mu(K) < \infty$ for each compact set K is called a Borel measure. A Borel measure is said to be regular if*
(i). For every $E \in \mathfrak{B}(X)$,

$$\mu(E) = \inf\{\mu(V) : E \subset V, V \text{ open}\}.$$

(iii). For every open set E,

$$\mu(E) = \sup\{\mu(K) : K \subset E, K \text{ compact}\}.$$

Remark 12. *Theorem 9 implies that every Borel measure in $\mathbb{R}^n$ is regular.*

Theorem 39 (Riesz Representation Theorem). *Let X be a locally compact Hausdorff space, and let $\Lambda : C_c(X) \to \mathbb{R}$ be a positive linear functional. Then there exists a unique regular Borel measure μ such that for any $f \in C_c(X)$,*

$$\Lambda f = \int_X f d\mu.$$

Let $V \subset X$ be open and $f \in C_c(X)$, we write $f \prec V$ if $0 \leq f(x) \leq 1$ on X and $\operatorname{supp} f \subset V$. We need the following theorem

Theorem 40 (Partition of Unity). *Let X be a locally compact Hausdorff space and let A be a compact subset of X. If $V_1, \cdots, V_n$ are open sets such that $A \subset \bigcup_{i=1}^n V_i$, then there exist continuous functions $f_i \in C_c(X)$ such that $f_i \prec V$ and*

$$\sum_{i=1}^n f_i = 1 \text{ on } A.$$

Proof. Let $x \in A$. Then there exists some i, $1 \leq i \leq n$ such that $x \in V_i$. Since X is a locally compact Hausdorff space, there exists a neighborhood U_x of x with compact closure such that $\overline{U_x} \subset V_i$. Since A is compact, there exists x_k, $1 \leq k \leq m$ such that

$$A \subset \bigcup_{i=1}^{m} U_{x_k}.$$

For each i, $1 \leq i \leq n$, we define G_i to be the union of all those U_{x_k} such that $U_{x_k} \subset V_i$. Then $A \subset \bigcup\limits_{i=1}^{n} G_i$. For each i, there exists an open set B_i with compact closure such that $\overline{G_i} \subset B_i \subset \overline{B_i} \subset V_i$. And from Urysohn's theorem, there exists a continuous function $g_i : X \to [0,1]$, such that $g_i(x) = 1$ for all $x \in \overline{G_i}$ and $g_i(x) = 0$ for all $x \notin B_i$. There also exists a continuous function $h : X \to [0,1]$, such that $h(x) = 1$ for all $x \in A$ and $h(x) = 0$ for all $x \notin \bigcup\limits_{i=1}^{n} G_i$. Set

$$g(x) = (1-h) + \sum_{i=1}^{n} g_i,$$

then g is continuous and $g(x) > 0$ for all $x \in X$. Set $f_i = \frac{g_i}{g}$, then f_i, $1 \leq i \leq n$, satisfy the desired properties. □

Here we present the proof of Riesz Representation Theorem when $X = \mathbb{R}^n$. Interested students should read the full proof of Theorem 2.14 in Rudin's book.

Proof of Riesz Representation Theorem. For each open subset V of X, we define

$$\lambda(V) = \sup\{\Lambda f : f \prec V\}.$$

And for any subset $E \subset X$, we define

$$\mu(E) = \inf\{\lambda(V) : E \subset V, V \text{ open}\}.$$

We claim that μ is an outer measure. $\mu(\emptyset) = 0$ follows from the fact that $f \prec \emptyset$ if and only if $f \equiv 0$. Also, from the definition, μ is monotonic. So we only need to check the countable subadditivity. Let E_n be a sequence of subsets of X, and $E = \bigcup\limits_{n=1}^{\infty} E_n$. We assume

$$\sum_{n=1}^{\infty} \mu(E_n) < \infty$$

otherwise the subadditivity is trivial. For each $\varepsilon > 0$, and for each $n \in \mathbb{N}$, there exists open set V_n such that $E_n \subset V_n$ and

$$\lambda(V_n) \leq \mu(E_n) + \frac{\varepsilon}{2^n}.$$

Set $V = \bigcup\limits_{n=1}^{\infty} V_n$. If $f \prec V$ holds, then $K = \operatorname{supp} f \subset \bigcup\limits_{n=1}^{\infty} V_n$. Since K is compact, there exists some m such that $K \subset \bigcup\limits_{n=1}^{m} V_n$. Now Theorem of Partition of Unity implies the existence of functions $f_1, \cdots, f_m \in C_c(X)$, such that $f_n \prec V_n$ for $n = 1, \cdots, m$ and

$$\sum_{n=1}^{m} f_n = 1 \text{ on } K.$$

Since $f \leq \sum_{n=1}^{m} f_n$, we have

$$\Lambda f \leq \sum_{n=1}^{m} \Lambda f_n \leq \sum_{n=1}^{m} \lambda\left(V_n\right) \leq \sum_{n=1}^{\infty} \mu\left(E_n\right) + \varepsilon.$$

Taking supremum on all $f \prec V$, we have

$$\lambda\left(V\right) \leq \sum_{n=1}^{\infty} \mu\left(E_n\right) + \varepsilon.$$

Since $E \subset V$, we have

$$\mu\left(E\right) \leq \lambda\left(V\right) \leq \sum_{n=1}^{\infty} \mu\left(E_n\right) + \varepsilon.$$

Since ε can be arbitrary, subadditivity follows.

When $X = \mathbb{R}^n$, we can show that μ is a metric outer measure. Let E, F be two subsets of X such that

$$\operatorname{dist}\left(E, F\right) > 0,$$

we need to show

$$\mu\left(E \cup F\right) \geq \mu\left(E\right) + \mu\left(F\right). \tag{4.1}$$

For any open set V such that $E \cup F \subset V$, we define open sets

$$V_E = V \cap \left\{x \in X : \operatorname{dist}\left(x, E\right) < \frac{1}{3} \operatorname{dist}\left(E, F\right)\right\},$$
$$V_F = V \cap \left\{x \in X : \operatorname{dist}\left(x, F\right) < \frac{1}{3} \operatorname{dist}\left(E, F\right)\right\}.$$

Then $V_E \cap V_F = \emptyset$, $E \subset V_E$ and $F \subset V_F$. Since $f_1 \prec V_E$, $f_2 \prec V_F$ implies $f_1 + f_2 \prec V$, we have

$$\lambda\left(V\right) \geq \lambda\left(V_E\right) + \lambda\left(V_F\right) \geq \mu\left(E\right) + \mu\left(F\right).$$

Taking infimum over all such V, (4.1) follows.

Since μ is a metric outer measure, all Borel sets are measurable. Now we show that $\mu\left(K\right) < \infty$ for each compact set K. When $X = \mathbb{R}^n$, $K \subset B_R\left(0\right)$ for some $R > 0$. Let $f \in C_c\left(X\right)$ be such that $f \equiv 1$ on $\overline{B_R\left(0\right)}$, then we have

$$\mu\left(K\right) \leq \lambda\left(B_R\left(0\right)\right) \leq \Lambda f < \infty.$$

Hence μ is a Borel measure. And since $X = \mathbb{R}^n$, μ is a regular Borel measure.

Next, we show

$$\Lambda f = \int_X f d\mu$$

for any $f \in C_c\left(X\right)$. Fix an open set V such that $K = \operatorname{supp} f \subset V$ and $\mu\left(V\right) < \infty$. Choose $c > 0$ such that $\left|f\left(x\right)\right| < c$ for all $x \in X$. Given $\varepsilon > 0$, pick n s.t. $\frac{2c}{n} < \varepsilon$. Let $y_i = -c + \frac{2c}{n} i$, $i = 0, 1, \cdots, n$. For each $i = 1, \cdots, n$, let

$$A_i = \left\{x \in K : y_{i-1} < f\left(x\right) \leq y_i\right\}$$
$$\subset W_i = \left\{x \in V : y_{i-1} - \varepsilon < f\left(x\right) < y_i + \varepsilon\right\}.$$

By the regularity of μ, for each $1 \leq i \leq n$, there exists an open set V_i such that $A_i \subset V_i \subset W_i$ and

$$\mu\left(V_i\right) - \mu\left(A_i\right) < \frac{\varepsilon}{n}.$$

We have

$$K \subset \bigcup_{i=1}^{n} V_i \subset V.$$

There exist partition of unity $g_1, \cdots, g_n \in C_c(X)$ such that $g_i \prec V_i$ and $\sum_{i=1}^{n} g_i = 1$ on K. Since $fg_i \le (y_i + \varepsilon) g_i$, we have

$$\begin{aligned} F(f) - \int_X f d\mu &= \sum_{i=1}^{n} F(fg_i) - \sum_{i=1}^{n} \int_{A_i} f d\mu \\ &\le \sum_{i=1}^{n} (y_i + \varepsilon) F(g_i) - \sum_{i=1}^{n} (y_i - \varepsilon) \mu(A_i) \\ &\le \sum_{i=1}^{n} (y_i + \varepsilon) \mu(V_i) - \sum_{i=1}^{n} (y_i - \varepsilon) \mu(A_i) \\ &= \sum_{i=1}^{n} (y_i + \varepsilon) (\mu(V_i) - \mu(A_i)) + 2\varepsilon \sum_{i=1}^{n} \mu(A_i) \\ &\le \sum_{i=1}^{n} (c + \varepsilon) \frac{\varepsilon}{n} + 2\varepsilon\mu(K) = \varepsilon(c + \varepsilon + 2\mu K). \end{aligned}$$

Hence, letting $\varepsilon \to 0$, we have $F(f) - \int_X f d\mu \le 0$. Now replace f by $-f$, we have

$$F(-f) - \int_X (-f) d\mu \le 0,$$

i.e., $F(f) - \int_X f d\mu \ge 0$. So $F(f) = \int_X f d\mu$ holds.

Finally, we show the uniqueness of μ. Let μ, ν be two Borel regular measures such that

$$\int_X f d\mu = \int_X f d\nu$$

for every $f \in C_c(X)$. Let K be any compact set. Given $\varepsilon > 0$, there exists an open set V such that $K \subset V$ and $\mu(V) \le \mu(K) + \varepsilon$. Uryson's Theorem implies the existence of $f \in C_c(X)$ such that $f = 1$ on K and $f \prec V$. Since $\chi_K \le f \le \chi_V$, we have

$$\begin{aligned} \nu(K) &= \int_X \chi_K d\nu \le \int_X f d\nu \\ &= \int_X f d\mu \le \int_X \chi_V d\mu = \mu(V) \le \mu(K) + \varepsilon. \end{aligned}$$

Since ε is arbitrary, we have $\nu(K) \le \mu(K)$. Same proof also yields $\mu(K) \le \nu(K)$, and so $\mu(K) = \nu(K)$. Since μ, ν are inner regular, we have $\mu(E) = \nu(E)$ for any open set E. And since μ, ν are outer regular, we have $\mu = \nu$ on $\mathfrak{B}(X)$. $\square$

REMARK 13. *We can expand the measurable space to make* μ *a complete measure.*

REMARK 14. *Let* $X = \mathbb{R}^n$. *We define functional*

$$\Lambda f = \int_{\mathbb{R}^n} f(x) dx$$

for any $f \in C_c(\mathbb{R}^n)$ *where the integral is Riemann integral. Riesz representation implies the existence of a Borel regular measure* μ *such that for any* $f \in C_c(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}^n} f d\mu.$$

One can check that μ *is the restriction of Lebesgue measure on* $\mathfrak{B}(\mathbb{R}^n)$.

5. Product Spaces and Fubini Theorem

Let $(X,\mathcal{S})$ and $(Y,\mathcal{T})$ be two measurable spaces. A measurable rectangle is a set of the form $A\times B$, where $A\in\mathcal{S}$ and $B\in\mathcal{T}$.

Definition 24. *$\mathcal{S}\times\mathcal{T}$ is defined to be the smallest σ-algebra in $X\times Y$ which contains every measurable rectangle.*

If $E\subset X\times Y$, $x\in X$, $y\in Y$, the x-section and y-section of E is defined by

$$E_x=\{y:(x,y)\in E\}\,,\ E^y=\{x:(x,y)\in E\}\,.$$

Theorem 41. *If $E\in\mathcal{S}\times\mathcal{T}$, then $E_x\in\mathcal{S}$ and $E^y\in\mathcal{S}$ for every $x\in X$ and $y\in Y$.*

Proof. Let

$$\Omega=\{E\in\mathcal{S}\times\mathcal{T}:E_x\in\mathcal{S}\text{ for every }x\in X\}\,.$$

It is easy to verify that Ω is a σ-algebra containing every measurable rectangle. Hence, $\Omega=\mathcal{S}\times\mathcal{T}$. □

With each function f on $X\times Y$ and with each $x\in X$ we define function f_x on Y by

$$f_x\,(y)=f\,(x,y)\,,\ y\in Y.$$

Similarly, with each $y\in Y$, we can define function f^y on Y by

$$f^y\,(y)=f\,(x,y)\,,\ x\in X.$$

Theorem 42. *Let f be an $(\mathcal{S}\times\mathcal{T})$-measurable function on $X\times Y$. Then*
(a). For each $x\in X$, f_x is a $\mathcal{T}$-measurable function on Y.
(b). For each $y\in Y$, f^y is a $\mathcal{S}$-measurable function on X.

Proof. For any open set V, put

$$Q=\{(x,y):f\,(x,y)\in V\}\,.$$

Then $Q\in\mathcal{S}\times\mathcal{T}$ and

$$Q_x=\{y:f_x\,(y)\in V\}\in\mathcal{T},$$

hence f_x is a $\mathcal{T}$-measurable function on Y. (b) can be proved similarly. □

We need Dynkin's $\pi\lambda$-system lemma.

Definition 25. *Let $\mathcal{A}$ be a collection of subsets of X. We say that $\mathcal{A}$ is a π-system if $\emptyset\in\mathcal{A}$ and $A\cap B\in\mathcal{A}$ for any $A,B\in\mathcal{A}$.*

Definition 26. *Let $\mathcal{A}$ be a collection of subsets of X. We say that $\mathcal{A}$ is a λ-system if $\mathcal{A}$ has the following three properties:*
(i). $X\in\mathcal{A}$.
(ii). $A\in\mathcal{A}$ implies $A^c\in\mathcal{A}$.
(iii). If $A_n\in\mathcal{A}$, $n=1,2,\cdots$ are disjoint, then $\bigcup\limits_{n=1}^{\infty}A_n\in\mathcal{A}$.

If $\mathcal{A}$ is both a λ-system and a d-system, then $\mathcal{A}$ is a σ-algebra.

Lemma 2 (Dynkin's lemma). *Let $\mathcal{A}$ be a π-system. Then any λ-system containing $\mathcal{A}$ contains also the σ-algebra generated by $\mathcal{A}$.*

PROOF. Denote by $\mathcal{B}$ be the intersection of all λ-systems containing $\mathcal{A}$. Then $\mathcal{B}$ is itself a λ-system. We shall show that $\mathcal{B}$ is also a π-system and hence a σ-algebra, thus proving the lemma. Consider

$$\mathcal{B}_1 = \{B \in \mathcal{B} : A \cap B \in \mathcal{B} \text{ for any } A \in \mathcal{A}\},$$

then $\mathcal{A} \subset \mathcal{B}_1$. Now we check that $\mathcal{B}_1$ is also a λ-system.

(i). $X \in \mathcal{B}_1$ since $X \in \mathcal{B}$ and for any $A \in \mathcal{A}$, $A \cap X = A \in \mathcal{A} \subset \mathcal{B}$.

(ii). If $B \in \mathcal{B}_1$, for any $A \in \mathcal{A}$,

$$A \cap B^c = A \cap (A \cap B)^C = (A^c \cup (A \cap B))^c \in \mathcal{B}.$$

So $B^c \in \mathcal{B}_1$.

(iii). If $B_n \in \mathcal{B}_1$, $n = 1, 2, \cdots$ are disjoint, then for any $A \in \mathcal{A}$,

$$A \cap \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} (A \cap B_n) \in \mathcal{B}.$$

Hence, $\bigcup\limits_{n=1}^{\infty} B_n \in \mathcal{B}_1$.

So we conclude $\mathcal{B}_1 = \mathcal{B}$. Next, we define

$$\mathcal{B}_2 = \{B \in \mathcal{B} : A \cap B \in \mathcal{B} \text{ for any } A \in \mathcal{B}\}.$$

We can prove $\mathcal{B}_2$ is also a λ-system. Now $\mathcal{A} \subset \mathcal{B}$, so we again have $\mathcal{B}_2 = \mathcal{B}$. Hence, $\mathcal{B}$ is a π-system and hence a σ-algebra. ☐

THEOREM 43. *Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \lambda)$ be two σ-finite measure spaces. Suppose $Q \in \mathcal{S} \times \mathcal{T}$. If*

$$\varphi(x) = \lambda(Q_x)$$

for every $x \in X$ and

$$\psi(y) = \mu(Q^y)$$

for every $y \in Y$, then φ is $\mathcal{S}$-measurable and ψ is $\mathcal{T}$-measurable, and

$$\int_X \varphi d\mu = \int_Y \psi d\lambda.$$

PROOF. Let Ω be the class of all $Q \in \mathcal{S} \times \mathcal{T}$ for which the conclusion of the theorem holds.

Step 1. Ω contains all measurable rectangles. Let $Q = A \times B$, where $A \in \mathcal{S}$, $B \in \mathcal{T}$, then we have

$$\varphi(x) = \lambda(Q_x) = \begin{cases} \lambda(B) & \text{if} \quad x \in A, \\ 0 & \text{if} \quad x \notin A \end{cases}$$

and

$$\psi(y) = \mu(Q^y) = \begin{cases} \mu(A) & \text{if} \quad y \in B, \\ 0 & \text{if} \quad y \notin B. \end{cases}$$

Hence, φ and ψ are measurable and

$$\int_X \varphi d\mu = \lambda(B)\mu(A) = \int_Y \psi d\lambda,$$

i.e., $Q \in \Omega$.

Step 2. If $Q_n \in \Omega$, $n = 1, 2, \cdots$ are disjoint, then $Q = \bigcup_{n=1}^{\infty} Q_n \in \Omega$. Let

$$\begin{aligned} \varphi_n(x) &= \lambda\left(\left(Q_n\right)_x\right) \text{ for any } x \in X, \\ \psi_n(y) &= \mu\left(Q_n^y\right) \text{ for any } y \in Y. \end{aligned}$$

Then

$$\begin{aligned} \varphi(x) &= \lambda\left(Q_x\right) = \sum_{n=1}^{\infty} \varphi_n(x), \\ \psi(y) &= \mu\left(Q^y\right) = \sum_{n=1}^{\infty} \psi_n(y). \end{aligned}$$

Since φ_n and ψ_n are measurable, we have φ and ψ are measurable, and

$$\int_X \varphi d\mu = \sum_{n=1}^{\infty} \int_X \varphi_n d\mu = \sum_{n=1}^{\infty} \int_Y \psi_n d\lambda = \int_Y \psi d\lambda.$$

So $Q \in \Omega$.

Step 3. Let X be the disjoint union of $\mathcal{S}$-measurable sets X_m with $\mu(X_m) < \infty$ and Y be the disjoint union of $\mathcal{T}$-measurable sets Y_n with $\lambda(Y_n) < \infty$. For any $m, n \in \mathbb{N}$ and $Q \in \mathcal{S} \times \mathcal{T}$, we claim $Q \cap (X_m \times Y_n) \in \Omega$. Let Ω_{mn} be the collection of all $Q \in \mathcal{S} \times \mathcal{T}$ such that $Q \cap (X_m \times Y_n) \in \Omega$.

(a). Ω_{mn} contains all measurable rectangles. Let $Q = A \times B$, , where $A \in \mathcal{S}$, $B \in \mathcal{T}$, then from step 1,

$$Q \cap (X_m \times Y_n) = (A \cap X_m) \times (B \cap Y_n) \in \Omega,$$

hence $Q \in \Omega_{mn}$.

(b). $X \times Y \in \Omega_{mn}$ since it is a measurable rectangle.

(c). Ω_{mn} is closed under complement. Let $Q \in \Omega_{mn}$, i.e.,

$$Q_{mn} = Q \cap (X_m \times Y_n) \in \Omega.$$

We have

$$\varphi(x) = \lambda\left(\left[Q \cap (X_m \times Y_n)\right]_x\right) = \left\{ \begin{array}{cl} \lambda\left(Q_x \cap Y_n\right) & \text{if} \quad x \in X_m, \\ 0 & \text{if} \quad x \notin X_m. \end{array} \right.$$

is measurable. Since

$$\begin{aligned} \varphi_1(x) = \lambda\left(\left[Q^c \cap (X_m \times Y_n)\right]_x\right) &= \left\{ \begin{array}{cl} \lambda\left(Q_x^c \cap Y_n\right) & \text{if} \quad x \in X_m, \\ 0 & \text{if} \quad x \notin X_m \end{array} \right. \\ &= \left\{ \begin{array}{cl} \lambda(Y_n) - \varphi(x) & \text{if} \quad x \in X_m, \\ 0 & \text{if} \quad x \notin X_m \end{array} \right. = \lambda(Y_n)\chi_{X_m} - \varphi(x), \end{aligned}$$

we conclude $\varphi_1(x)$ is measurable. Similarly, since

$$\psi(y) = \lambda\left(\left[Q \cap (X_m \times Y_n)\right]^y\right) = \left\{ \begin{array}{cl} \mu\left(Q^y \cap X_m\right) & \text{if} \quad y \in Y_n, \\ 0 & \text{if} \quad y \notin Y_n. \end{array} \right.$$

is measurable, we have

$$\psi_1(y) = \lambda\left(\left[Q^c \cap (X_m \times Y_n)\right]^y\right) = \mu(X_m)\chi_{Y_n} - \psi(y)$$

is also measurable. Moreover, we have

$$\begin{aligned}\int_X \varphi_1 d\mu &= \int_X \left(\lambda\left(Y_n\right)\chi_{X_m} - \varphi\right) d\mu = \mu\left(X_m\right)\lambda\left(Y_n\right) - \int_X \varphi d\mu \\ &= \mu\left(X_m\right)\lambda\left(Y_n\right) - \int_Y \psi d\lambda = \int_Y \left(\mu\left(X_m\right)\chi_{Y_n} - \psi\right) d\lambda = \int_Y \psi_1 d\lambda.\end{aligned}$$

Hence, we have $Q^c \cap (X_m \times Y_n) \in \Omega$, i.e., $Q^c \in \Omega_{mn}$.

(d). Ω_{mn} is closed under countable disjoint union. This follows from step 2.

(e). Let $\mathcal{A}$ be the collection of all measurable rectangles. Then $\mathcal{A}$ is a π-system. Since Ω_{mn} is a λ-system containing $\mathcal{A}$, Dynkin's lemma implies $\Omega_{mn} \supset \sigma(\mathcal{A}) = \mathcal{S} \times \mathcal{T}$. Hence, $\Omega_{mn} = \mathcal{S} \times \mathcal{T}$.

Step 4. For any $Q \in \mathcal{S} \times \mathcal{T}$, $Q_{mn} = Q \cap (X_m \times Y_n) \in \Omega$. From step 2,

$$Q = \bigcup_{m,n=1}^{\infty} Q_{mn} \in \Omega.$$

Hence $\Omega = \mathcal{S} \times \mathcal{T}$. □

Definition 27. *Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \lambda)$ be two σ-finite measure spaces. We define for any $Q \in \mathcal{S} \times \mathcal{T}$,*

$$(\mu \times \lambda)(Q) = \int_X \lambda(Q_x)\, d\mu = \int_Y \mu(Q^y)\, d\lambda.$$

Then $(X \times Y, \mathcal{S} \times \mathcal{T}, \mu \times \lambda)$ is a measure space.

Theorem 44 (The Fubini Theorem). *Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \lambda)$ be two σ-finite measure spaces and let f be an $(\mathcal{S} \times \mathcal{T})$-measurable function on $X \times Y$. Then*
(a). If $0 \le f \le \infty$ and if

$$\begin{aligned}\varphi(x) &= \int_Y f_x d\lambda, x \in X, \\ \psi(y) &= \int_Y f^y d\mu, y \in Y,\end{aligned} \tag{5.1}$$

then φ is $\mathcal{S}$-measurable, ψ is $\mathcal{T}$-measurable and

$$\int_X \varphi d\mu = \int_Y \psi d\lambda = \int_{X \times Y} f d(\mu \times \lambda). \tag{5.2}$$

(b). If f is complex and if

$$\varphi^*(x) = \int_Y |f|_x\, d\lambda \text{ and } \int_X \varphi^* d\mu < \infty,$$

then $f \in L^1(\mu \times \lambda)$.
(c). If $f \in L^1(\mu \times \lambda)$, then $f_x \in L^1(\lambda)$ for almost all $x \in X$, $f^y \in L^1(\mu)$ for almost all $y \in Y$, φ, ψ defined by (5.1) *a.e. are in $L^1(\mu)$ and $L^1(\lambda)$, respectively, and* (5.2) *holds.*

Proof. (a). From Theorem 43, the conclusion holds for $f = \chi_Q$ for any $Q \in \mathcal{S} \times \mathcal{T}$. Hence, (a) holds for any measurable simple function s. Now let s_n, $n \in \mathbb{N}$ be a monotone increasing sequence of measurable simple functions such that

$$\lim_{n \to \infty} s_n(x, y) = f(x, y) \text{ for any } (x, y) \in X \times Y.$$

Let

$$\varphi_n(x) = \int_Y (s_n)_x \, d\lambda, x \in X,$$
$$\psi_n(y) = \int_Y (s_n)^y \, d\mu, y \in Y,$$

then φ_n is $\mathcal{S}$-measurable, ψ_n is $\mathcal{T}$-measurable and

$$\int_X \varphi_n d\mu = \int_Y \psi_n d\lambda = \int_{X\times Y} s_n d(\mu \times \lambda). \tag{5.3}$$

And that

$$\varphi(x) = \int_Y f_x d\lambda = \lim_{n\to\infty} \varphi_n(x), x \in X,$$
$$\psi(y) = \int_Y f^y d\mu = \lim_{n\to\infty} \psi_n(x), y \in Y,$$

hence φ is $\mathcal{S}$-measurable, ψ is $\mathcal{T}$-measurable. Finally, Monotone Convergence Theorem applied to (5.3) yields

$$\int_X \varphi d\mu = \int_Y \psi d\lambda = \int_{X\times Y} f d(\mu \times \lambda).$$

(b). Applying (a) to $|f|$, (b) follows.

(c). We assume f is real. Let $f = f^+ - f^-$, then $f^+, f^- \in L^1(\mu \times \lambda)$. Let

$$\varphi^{\pm}(x) = \int_Y \left(f^{\pm}\right)_x d\lambda, x \in X,$$
$$\psi^{\pm}(y) = \int_Y \left(f^{\pm}\right)^y d\mu, y \in Y,$$

then $\varphi^{\pm}$ are $\mathcal{S}$-measurable, $\psi^{\pm}$ are $\mathcal{T}$-measurable and

$$\int_X \varphi^{\pm} d\mu = \int_Y \psi^{\pm} d\lambda = \int_{X\times Y} f^{\pm} d(\mu \times \lambda) < \infty.$$

Hence, $\varphi^{\pm}, \psi^{\pm} < \infty$, a.e., and so

$$\varphi(x) = \varphi^+(x) - \varphi^-(x) = \int_Y f_x d\lambda, \text{ a.e. } x \in X,$$
$$\psi(y) = \psi^+(x) - \psi^-(x) = \int_Y f^y d\mu, \text{ a.e. } y \in Y,$$

are well defined, and they are in $L^1(\mu)$ and $L^1(\lambda)$, respectively, and (5.2) holds. □

Remark 15. *The students are encouraged to check the counterexamples in 8.9 of Rudin's book.*

Remark 16. *The product of complete measures may not be complete.*

Let $(X, \mathfrak{M}, \mu)$ be a measurable space. We recall the construction of the completion of measure μ in Theorem 31: Let $\mathfrak{M}^*$ be the collection of all $E \subset X$ for which there exist sets $A, B \in \mathfrak{M}$ such that $A \subset E \subset B$ and $\mu(B \backslash A) = 0$, and define $\mu(E) = \mu(A)$ in this situation. Then $\mathfrak{M}^*$ is a σ-algebra, and μ is a complete measure on $\mathfrak{M}^*$.

Theorem 45. *Let μ_k denote Lebesgue measure on $\mathbb{R}^k$. If $k = r + s$, $r, s \geq 1$, then μ_k is the completion of the product measure $\mu_r \times \mu_s$.*

Proof. Let $\mathfrak{M}_k$ be the collection of Lebesgue measurable sets in $\mathbb{R}^k$ and $\mathfrak{B}_k$ be the collection of Borel sets in $\mathbb{R}^k$. Then the homework problem implies $\mathfrak{B}_k \subset \mathfrak{M}_r \times \mathfrak{M}_s$. From the property of $\mathfrak{M}_k$, we see $\mathfrak{M}_k$ is the completion of $\mathfrak{B}_k$. And since $\mu_r \times \mu_s = \mu_k$ on rectangular boxes, we claim μ_k is the completion of the product measure $\mu_r \times \mu_s$. $\square$

Theorem 46. *Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{T}, \lambda)$ be two complete σ-finite measure spaces. Let $(\mathcal{S} \times \mathcal{T})^*$ be the completion of $\mathcal{S} \times \mathcal{T}$, relative to the measure $\mu \times \lambda$. Let f be an $(\mathcal{S} \times \mathcal{T})^*$-measurable function on $X \times Y$. Then*
(a). If $0 \leq f \leq \infty$, then f_x is $\mathcal{T}$ measurable for almost all $x \in X$, and f^y is $\mathcal{S}$ measurable for almost all $y \in Y$. Let

$$\varphi(x) = \int_Y f_x d\lambda, \text{ a.e. } x \in X, \tag{5.4}$$

$$\psi(y) = \int_Y f^y d\mu, \text{ a.e. } y \in Y,$$

then φ is $\mathcal{S}$-measurable, ψ is $\mathcal{T}$-measurable and

$$\int_X \varphi d\mu = \int_Y \psi d\lambda = \int_{X \times Y} f d(\mu \times \lambda). \tag{5.5}$$

(b). If f is complex and if

$$\varphi^*(x) = \int_Y |f|_x d\lambda \text{ and } \int_X \varphi^* d\mu < \infty,$$

then $f \in L^1(\mu \times \lambda)$.
(c). If $f \in L^1(\mu \times \lambda)$, then $f_x \in L^1(\lambda)$ for almost all $x \in X$, $f^y \in L^1(\mu)$ for almost all $y \in Y$, φ, ψ defined by (5.4) a.e. are in $L^1(\mu)$ and $L^1(\lambda)$, respectively, and (5.5) holds.

Interested student should check Rudin's book for proof.

CHAPTER 3

L^p-Spaces

1. Introduction

Let (X,μ) be a measure space.

Definition 28. *For* $0<p<\infty$, $\tilde{L}^p(\mu)$ *denotes the collection of all complex valued measurable functions such that*

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

and we define $L^p(\mu)=\tilde{L}^p(\mu)/\sim$, *where* $f\sim g$ *if* $f=g$ *a.e.. For* $p=\infty$, *we define* $\tilde{L}^\infty(\mu)$ *to be the collection of essentially bounded measurable functions. A function* f *is said to be essentially bounded if there is* $M>0$ *with*

$$|f(x)|\le M \ \text{a.e.}. \tag{1.1}$$

We use $\|f\|_\infty$ *to denote the smallest* M *satisfying* (1.1). *We define* $L^\infty(\mu)=\tilde{L}^\infty(\mu)/\sim$.

If μ is a counting measure on $\mathbb{N}$, then each function $f:\mathbb{N}\to\mathbb{C}$ can be viewed as a sequence $\{f(n)\}$. And we have

$$\|f\|_p=\left(\sum_{n=1}^{\infty}|f(n)|^p\right)^{\frac{1}{p}} \text{ for } 0<p<\infty,$$

and

$$\|f\|_\infty=\sup_{n\in\mathbb{N}}|f(n)|.$$

And we denote $l^p=L^p(\mu)$ in this situation.

We say a function φ is convex on interval (a,b) if for any $x,y\in(a,b)$ and for any $\lambda\in(0,1)$,

$$\varphi((1-\lambda)x+\lambda y)\le(1-\lambda)\varphi(x)+\lambda\varphi(y).$$

Geometrically, this condition means that if $a<x<t<y<b$, then $(t,\varphi(t))$ lies below or on the line connecting the points $(x,\varphi(x))$ and $(y,\varphi(y))$. It is easy to see that φ is convex on (a,b) if and only if for any $a<s<t<u<b$,

$$\frac{\varphi(t)-\varphi(s)}{t-s}\le\frac{\varphi(u)-\varphi(t)}{u-t}.$$

Hence if φ is pointwisely differentiable, then φ is convex if and only if φ' is monotone increasing.

Lemma 3. *A convex function* φ *on* (a,b) *is continuous on* (a,b).

Proof. Given $x_0 \in (a,b)$, we pick $s,t \in (a,b)$ such that $s < x_0 < t$. Then for any $x \in (s,t)$, $x \neq x_0$, we have

$$\frac{\varphi(x_0) - \varphi(s)}{x_0 - s} \leq \frac{\varphi(x) - \varphi(x_0)}{x - x_0} \leq \frac{\varphi(t) - \varphi(x_0)}{t - x_0},$$

which implies the continuity of φ at x_0. □

Theorem 47 (Jensen's Inequality). *Let (X,μ) be a measure space with $\mu(X) = 1$. Let $(a,b) \subset \mathbb{R}$ be an open interval, possibly unbounded. If f is a real valued function in $L^1(\mu)$ s.t. $f(x) \in (a,b)$ for all $x \in X$ and if φ is convex on (a,b), then*

$$\varphi\left(\int_X f d\mu\right) \leq \int_X \varphi(f)\, d\mu. \tag{1.2}$$

Proof. Let

$$t = \int_X f d\mu.$$

Then $t \in (a,b)$. Let

$$\beta = \sup_{a<s<t} \frac{\varphi(t) - \varphi(s)}{t - s}.$$

Then for any $s \in (a,b)$,

$$\varphi(s) \geq \varphi(t) + \beta(s - t).$$

Hence, for any $x \in X$,

$$\varphi(f(x)) \geq \varphi(t) + \beta(f(x) - t).$$

Integrating over X and using $\mu(X) = 1$, we have (1.2). □

Definition 29. *Two positive real numbers $p,q \in (1,\infty)$ are called a pair of conjugate exponents if*

$$\frac{1}{p} + \frac{1}{q} = 1.$$

As $p \to 1$, we have $q \to \infty$. Consequently 1 and ∞ are also regarded as a pair of conjugate exponents.

Theorem 48 (Young's Inequality). *If $p,q \in (1,\infty)$ are a pair of conjugate exponents. Then for any $a,b \geq 0$,*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Theorem 49 (Holder's Inequality). *Let (X,μ) be a measure space. If $p,q \in (1,\infty)$ are a pair of conjugate exponents. Then for any measurable functions $f,g : X \to \mathbb{C}$ measurable,*

$$\int_X |fg|\, d\mu \leq \|f\|_p \|g\|_q.$$

Theorem 50 (Minkowski's Inequality). *Let (X,μ) be a measure space and $p \in [1,\infty]$. Then for any measurable functions $f,g : X \to \mathbb{C}$ measurable,*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

DEFINITION 30. *Let K denote $\mathbb{R}$ or $\mathbb{C}$. The normed space is a pair $(X, \|\cdot\|)$, where X is a linear space over K and*

$$\|\cdot\| : X \to [0, \infty] .$$

is a function satisfying
(i). $\|x+y\| \leq \|x\| + \|y\|$ holds for all $x, y \in X$;
(ii). $\|\alpha x\| \leq |\alpha| \|x\|$ holds for all $x \in X$, $\alpha \in K$;
(iii). $\|x\| = 0$ if and only if $x = 0$.

REMARK 17. *A normed space X is a metric space with $d(x, y) = \|x - y\|$.*

DEFINITION 31. *A normed space $(X, \|\cdot\|)$ is called a Banach space if it is complete with respect to the metric $d(x, y) = \|x - y\|$.*

THEOREM 51. *$L^p(\mu)$ is a Banach space for all $1 \leq p \leq \infty$.*

PROOF. It is easy to check that $L^p(\mu)$ is a normed space for any $p \in [1, \infty]$. Now we prove it's complete. We first assume $p < \infty$. Let $\{f_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $L^p(\mu)$. Then there is a subsequence $\{f_{n_i}\}_{i=1}^{\infty}$ such that

$$\left\|f_{n_{i+1}} - f_{n_i}\right\|_p < \frac{1}{2^i}.$$

Define

$$g_k = \sum_{i=1}^{k} \left|f_{n_{i+1}} - f_{n_i}\right| \text{ and } g = \sum_{i=1}^{\infty} \left|f_{n_{i+1}} - f_{n_i}\right| .$$

Then we have from Minkowski's inequality

$$\|g_k\| \leq 1 \text{ for each } k \in \mathbb{N}$$

and g_k is monotone increasing with

$$\lim_{k\to\infty} g_k(x) = g(x) .$$

Fatou's lemma then implies

$$\int_X g^p d\mu \leq \lim_{k\to\infty} \int_X g_k^p d\mu = 0.$$

Hence, $g(x) < \infty$ a.e., i.e. and

$$f_{n_1} + \sum_{i=1}^{\infty} \left(f_{n_{i+1}} - f_{n_i}\right)$$

is absolute convergent for almost every $x \in X$. Let

$$f = f_{n_1} + \sum_{i=1}^{\infty} \left(f_{n_{i+1}} - f_{n_i}\right) \text{ a.e..}$$

Then

$$f(x) = \lim_{i\to\infty} f_{n_i}(x) \text{ a.e..}$$

For any $\varepsilon > 0$, there exists K s.t. $\|f_n - f_m\|_p < \varepsilon$ for any $n, m > K$. Hence Fatou's lemma implies for any $m > K$,

$$\int_X |f - f_m|^p d\mu \leq \lim_{i\to\infty} \int_X |f_{n_i} - f_m|^p d\mu \leq \varepsilon.$$

Hence $f - f_m \in L^p$, $f \in L^p$ and

$$\lim_{m\to\infty} \left\| f - f_m \right\|_p = 0.$$

When $p = \infty$, assume $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^{\infty}(\mu)$. Let

$$A_k = \left\{ x \in X : f_k(x) > \left\| f_k \right\|_{\infty} \right\},$$
$$B_{m,n} = \left\{ x \in X : \left| f_m(x) - f_n(x) \right| > \left\| f_m(x) - f_n(x) \right\|_{\infty} \right\}.$$

Then $A_k, B_{m,n}$ are measure zero set and there union E also has measure zero. For any $x \in E^c$, $\{f_n(x)\}_{n=1}^{\infty}$ is a Cauchy sequence in X, hence

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

is defined for almost every $x \in X$. For any $\varepsilon > 0$, there exists $K > 0$, such that $\left\| f_n - f_m \right\|_{\infty} < \varepsilon$ for any $n, m > K$. Hence, for any $x \in E^c$,

$$\left| f(x) - f_m(x) \right| = \lim_{n\to\infty} \left| f_n(x) - f_m(x) \right| \le \lim_{n\to\infty} \left\| f_n - f_m \right\|_{\infty} < \varepsilon.$$

So we conclude

$$\lim_{m\to\infty} \left\| f - f_m \right\|_{\infty} = 0.$$

□

As a byproduct of the above proof, we have

Corollary 4. *If $1 \le p \le \infty$, $\{f_n\}_{n=1}^{\infty} \subset L^p(\mu)$, $f \in L^p(\mu)$, and*

$$\lim_{n\to\infty} \left\| f - f_n \right\|_{\infty} = 0.$$

Then there exists a subsequence $\{f_{n_i}\}_{i=1}^{\infty}$, such that $f_{n_i} \to f$ a.e..

Theorem 52. *Let S be the class of all complex, measurable, simple functions on X s.t.*

$$\mu\{x : s(x) \neq 0\} < \infty.$$

If $1 \le p < \infty$, then S is dense in $L^p(\mu)$.

When μ is the Lebesgue measure on $\mathbb{R}^n$, we write $L^p(\mu) = L^p(\mathbb{R}^n)$

Theorem 53. *If $1 \le p < \infty$, then $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.*

Remark 18. *This density theorem holds when μ is a regular Borel measure on a locally compact Hausdorff space X.*

Theorem 54. *The completion of $C_c(\mathbb{R}^n)$ with respect to $\left\|\cdot\right\|_{\infty}$ is $C_0(\mathbb{R}^n)$, the space of continuous functions which converges to zero at ∞.*

2. Convolution of Functions

THEOREM 55. *Suppose* $f, g \in L^1(\mathbb{R}^n)$, *then*

$$\int_{\mathbb{R}^n} |f(x-y)\, g(y)|\, dy < \infty$$

for almost all x. *For these* x, *define*

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

Then $h \in L^1(\mathbb{R}^n)$, *and*

$$\|h\|_1 \le \|f\|_1 \|g\|_1$$

PROOF. One can verify that $f(x-y)\, g(y)$ is Lebesgue measurable in $\mathbb{R}^{2n}$. Fubini Theorem implies

$$\begin{aligned} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)\, g(y)|\, dy \right) dx &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)\, g(y)|\, dx \right) dy \\ &= \int_{\mathbb{R}^n} (|g(y)| \cdot \|f\|_1)\, dy = \|f\|_1 \|g\|_1 . \end{aligned}$$

Hence

$$\int_{\mathbb{R}^n} |f(x-y)\, g(y)|\, dy < \infty$$

for almost all x and $h \in L^1(\mathbb{R}^n)$,

$$\|h\|_1 \le \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)\, g(y)|\, dy \right) dx \le \|f\|_1 \|g\|_1 .$$

□

We say h is the convolution of f and g, and we write $h = f * g$. It is easy to see

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

We say f is locally integrable in $\mathbb{R}^n$ if f is integrable on any bounded subset of $\mathbb{R}^n$, and we write $f \in L^1_{loc}(\mathbb{R}^n)$. The convolution can also be defined for $f \in L^1_{loc}(\mathbb{R}^n)$ and $g \in C_c(\mathbb{R}^n)$.

THEOREM 56. *Let* $f \in L^1_{loc}(\mathbb{R}^n)$ *and* $g \in C_c(\mathbb{R}^n)$. *Then the convolution*

$$h(x) = \int_{\mathbb{R}^n} f(x-y)\, g(y)\, dy$$

is continuous in $\mathbb{R}^n$.

PROOF. Assume $\operatorname{supp} g \subset B_R(0)$, for some $R > 0$. Given $x \in \mathbb{R}^n$, for any $x' \in B_1(x)$, we have

$$\begin{aligned} |h(x) - h(x')| &= \left| \int_{\mathbb{R}^n} f(y)\, g(x-y)\, dy - \int_{\mathbb{R}^n} f(y)\, g(x'-y)\, dy \right| \\ &\le \int_{\mathbb{R}^n} |f(y)|\, |g(x-y) - g(x'-y)|\, dy \\ &\le \left(\sup_y |g(x-y) - g(x'-y)| \right) \int_{B_{R+1}(x)} |f(y)|\, dy. \end{aligned}$$

Since g is uniformly continuous, we have

$$\lim_{x'\to x}\sup_{y}\left|g\left(x-y\right)-g\left(x'-y\right)\right|=0,$$

hence,

$$\lim_{x'\to x}h\left(x'\right)=h\left(x\right)$$

and h is continous. $\square$

THEOREM 57. *Let $f\in L^1_{loc}\left(\mathbb{R}^n\right)$ and $g\in C_c\left(\mathbb{R}^n\right)\cap C^r\left(\mathbb{R}^n\right)$ for some $r\geq 1$. Then the convolution $h=f*g\in C^r\left(\mathbb{R}^n\right)$ and*

$$D^\alpha h\left(x\right)=f*D^\alpha g$$

for any index $|\alpha|\leq r$.

PROOF. Let e_i be a unit vector in x_i direction. Fix $x\in\mathbb{R}$. For any $t\in\mathbb{R}$, $|t|\leq 1$,

$$\frac{h\left(x+te_i\right)-h\left(x\right)}{t}=\int_{\mathbb{R}^n}f\left(y\right)\frac{g\left(x-y+te_i\right)-g\left(x-y\right)}{t}dy,$$

Since $f\in L^1_{loc}\left(\mathbb{R}^n\right)$, $f\left(y\right)\frac{g(x-y+te_i)-g(x-y)}{t}$ vanishes away a bounded subset E_x, we have

$$\left|f\left(y\right)\frac{g\left(x-y+te_i\right)-g\left(x-y\right)}{t}\right|\leq\left\|\frac{\partial}{\partial x_i}g\right\|_\infty\left|f\left(y\right)\right|\chi_{E_x}\in L^1\left(\mathbb{R}^n\right)\text{ in }y.$$

Lebesgue dominated convergence implies

$$\lim_{t\to 0}\int_{\mathbb{R}^n}f\left(y\right)\frac{g\left(x-y+te_i\right)-g\left(x-y\right)}{t}dy=\int_{\mathbb{R}^n}f\left(y\right)\frac{\partial}{\partial x_i}g\left(x-y\right)dy=f*\frac{\partial}{\partial x_i}g.$$

So $\frac{\partial}{\partial x_i}h\left(x\right)$ exists, and $\frac{\partial}{\partial x_i}h\left(x\right)=f*\frac{\partial g}{\partial x_i}$. Higher order derivatives follows from mathematical induction. $\square$

Let $\varphi\in C_0^\infty\left(\mathbb{R}^n\right)$ be a function satisfying $\varphi\geq 0$, $\operatorname{supp}\varphi\subset\overline{B_1\left(0\right)}$ and

$$\int_{\mathbb{R}^n}\varphi\left(x\right)dx=1. \tag{2.1}$$

For example, we could take

$$\varphi\left(x\right)=\begin{cases} ce^{\frac{1}{|x|^2-1}} & \text{if } \; |x|<1,\\ 0 & \text{if } \; |x|\geq 1\end{cases}$$

where $c>0$ is a constant chosen to gurantee (2.1). For $\varepsilon>0$, we define

$$\varphi_\varepsilon\left(x\right)=\varepsilon^{-n}\varphi\left(\frac{x}{\varepsilon}\right).$$

Then $\operatorname{supp}\varphi\subset\overline{B_\varepsilon\left(0\right)}$ and

$$\int_{\mathbb{R}^n}\varphi_\varepsilon\left(x\right)dx=1.$$

For any $f\in L^1_{loc}\left(\mathbb{R}^n\right)$, we define

$$f_\varepsilon=f*\varphi_\varepsilon\in C^\infty\left(\mathbb{R}^n\right).$$

We say φ_ε is a mollifier and f_ε is a mollification of f.

THEOREM 58. *If f is continuous in $\mathbb{R}^n$, then $f_\varepsilon\Longrightarrow f$ uniformly on any compact subset as $\varepsilon\to 0$.*

PROOF.

$$\begin{aligned} f(x) - f_\varepsilon(x) &= f(x) - \int_{\mathbb{R}^n} f(x-y)\,\varphi_\varepsilon(y)\,dy \\ &= \int_{B_\varepsilon(0)} \left(f(x) - f(x-y)\right)\varphi_\varepsilon(y)\,dy. \end{aligned}$$

Since f is uniform continuous on any compact set, $f_\varepsilon \Longrightarrow f$ uniformly on any compact subset as $\varepsilon \to 0$. □

REMARK 19. *If $f \in C_c(\mathbb{R}^n)$, then $f_\varepsilon \in C_c(\mathbb{R}^n)$ and $f_\varepsilon \Longrightarrow f$ uniformly on $\mathbb{R}^n$ as $\varepsilon \to 0$.*

LEMMA 4. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f_\varepsilon\|_p \le \|f\|_p$.*

PROOF. $p = 1$ follows from Theorem 55. When $1 < p < \infty$, let $q = p'$, we have

$$\begin{aligned} |f_\varepsilon(x)|^p &= \left(\int_{\mathbb{R}^n} f(x-y)\,\varphi_\varepsilon(y)\,dy\right)^p = \left(\int_{\mathbb{R}^n} f(x-y)\,\varphi_\varepsilon^{\frac{1}{p}}(y)\,\varphi_\varepsilon^{\frac{1}{q}}(y)\,dy\right)^p \\ &\le \int_{\mathbb{R}^n} |f(x-y)|^p\,\varphi_\varepsilon(y)\,dy \left(\int_{\mathbb{R}^n} \varphi_\varepsilon(y)\,dy\right)^{\frac{p}{q}} = |f|^p * \varphi_\varepsilon, \end{aligned}$$

hence $\|f_\varepsilon\|_p \le \|f\|_p$ follows from $p = 1$ case. □

THEOREM 59. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$, then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f - f_\varepsilon\|_p \to 0$ as $\varepsilon \to 0$.*

PROOF. For any $\delta > 0$, there exists $g \in C_c(\mathbb{R}^n)$,

$$\|f - g\|_p < \frac{\delta}{3}.$$

And for such g, there exists ε_0, such that for any $\varepsilon \le \varepsilon_0$,

$$\|g - g_\varepsilon\|_p < \frac{\delta}{3}.$$

Hence, for any $\varepsilon \le \varepsilon_0$,

$$\|f - f_\varepsilon\|_p \le \|f - g\|_p + \|g - g_\varepsilon\|_p + \|g_\varepsilon - f_\varepsilon\|_p < \delta.$$

□

COROLLARY 5. *$C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$.*

PROOF. Since continuous functions with compact support are dense in $L^p(\mathbb{R}^n)$, it suffices to prove that every compactly supported continuous function can be approximated in $L^p(\mathbb{R}^n)$ by $C_c^\infty(\mathbb{R}^n)$. This, however, immediately follows from Theorem 59, because if f vanishes outside a bounded set, then f_ε has compact support. □

CHAPTER 4

Integration and Differentiation

1. Signed Measures

Definition 32. *Let $(X, \mathfrak{M})$ be a measurable space. A set function $\mu : \mathfrak{M} \to [-\infty, \infty]$ is called a signed measure if*
(i). μ assumes at most one of the values ∞ and $-\infty$;
(ii). $\mu(\emptyset) = 0$;
(iii). μ is σ-additive.

Let μ_1, μ_2 be two positive measures on $(X, \mathfrak{M})$ such that one of them is finite. Then $\mu_1 - \mu_2$ is a signed measure.

Definition 33. *A measurable set A is said to be positive if $\mu(A \cap E) \geq 0$ for every $E \in \mathfrak{M}$. A measurable set A is said to be negative if $\mu(A \cap E) \leq 0$ for every $E \in \mathfrak{M}$.*

Lemma 5. *Let μ be a signed measure on $(X, \mathfrak{M})$ and let $E \in \mathfrak{M}$ with $\mu(E) > 0$. Then there exists a positive set $A \subset E$ such that $\mu(A) > 0$.*

Proof. If E is positive, then there is nothing to prove. Otherwise, there exists some $B \in \mathfrak{M}$ with $B \subset E$ and $-\infty < \mu(B) < 0$. Let n_1 be the smallest positive integer such that there is B_1 with $B_1 \subset E$ and

$$\mu(B_1) \leq -\frac{1}{n_1}.$$

If $E_1 = E \backslash B_1$ is positive, then we are done, since $\mu(E_1) > 0$ follows from

$$\mu(E \backslash B_1) + \mu(B_1) = \mu(E) > 0.$$

Otherwise, let n_2 be the smallest positive integer such that there is B_2 with $B_2 \subset E_1 \subset E$ and

$$\mu(B_2) \leq -\frac{1}{n_2}.$$

And we define $E_2 = E_1 \backslash B_2$. If E_2 is positive, then we are done, since $\mu(E_2) > 0$. Otherwise, we can continue this process. If this process stops at E_m, then E_m is a positive set with $\mu(E_m) > 0$. Otherwise, we find a sequence of positive integers n_k and measurable sets $B_k \subset E$, $k \in \mathbb{N}$, such that for $E_m = E \backslash \left(\bigcup_{k=1}^{m} B_k \right)$, n_{m+1} is the smallest positive integer such that there is $B_{m+1} \in \mathfrak{M}$ with $B_{m+1} \subset E_m$ and

$$\mu(B_{m+1}) \leq -\frac{1}{n_{m+1}}.$$

Since $\mu(E) > 0$, we have

$$-\infty < \mu\left(\bigcup_{k=1}^{\infty} B_k \right) \leq \sum_{k=1}^{\infty} -\frac{1}{n_k}.$$

Hence,

$$\sum_{k=1}^{\infty} \frac{1}{n_k} < \infty.$$

Since n_k is monotone increasing, we conclude $n_k \to \infty$ as $k \to \infty$. Let $A = E \backslash \left(\bigcup_{k=1}^{\infty} B_k \right)$. For any set $B \subset A$, since $A \subset E_k$, we have

$$\mu(B) > -\frac{1}{n_{k+1} - 1},$$

hence $\mu(B) \geq 0$, so A is positive. And $\mu(A) > 0$ follows from $\mu(B_k) < 0$ for each k and

$$\mu(A) + \sum_{k=1}^{\infty} \mu(B_k) = \mu(A) > 0.$$

□

THEOREM 60 (Hahn's Decomposition Theorem). *Let μ be a signed measure on $(X, \mathfrak{M})$. Then there exist positve set A and negative set B such that $X = A \cup B$ and $A \cap B = \emptyset$.*

PROOF. We can assume without loss of generality that μ doesn't assume $+\infty$ value, otherwise we could consider $-\mu$. Let

$$M = \sup \{\mu(E) : E \in \mathfrak{M} \text{ is positive}\}.$$

Then there exists a sequence of positive sets E_n, such that

$$\lim_{n \to \infty} \mu(E_n) = M.$$

Then $A = \bigcup_{n=1}^{\infty} E_n$ is a positive set and $\mu(A) = M < \infty$. Let $B = X \backslash A$. We claim B is a negative set. Otherwise, there exists a set $E \subset B$ with $\mu(E) > 0$, and Lemma 5 implies the existence of positive set $E' \subset E$ with $\mu(E') > 0$. Then $A \cup E'$ will be a positive set with $\mu(A \cup E') > M$, a contradiction. □

We say the pair (A, B) is a Hahn's decomposition of X with respect to μ. Let $f \in L^1(\mu)$. Then

$$\lambda(E) = \int_E f d\mu$$

defines a signed measure. It is easy to see $A = \{x \in X : f(x) \geq 0\}$, $B = X \backslash A$ defines a Hahn's decomposition of X with respect to λ.

Let μ be a signed measure on $(X, \mathfrak{M})$ and (A, B) be a Hahn's decomposition of X with respect to μ. We can define for any $E \in \mathfrak{M}$,

$$\mu^+(E) = \mu(A \cap E), \mu^-(E) = -\mu(B \cap E),$$
$$|\mu|(E) = \mu^+(E) + \mu^-(E).$$

Then μ^+, μ^- are positive measures and $\mu(E) = \mu^+(E) - \mu^-(E)$ for any $E \in \mathfrak{M}$. The pair (μ^+, μ^-) is called the Jordan's decomposition of μ. μ^+, μ^- and $|\mu|$ are called the positive variation, the negative variation and the total variation of μ, respectively.

REMARK 20. *Hahn's decomposition may not be unique. But Jordan's decomposition is uniquely defined.*

Remark 21. *Hahn's decomposition shows that any signed measure is the difference of two positive measures.*

Let $M(X)$ be the collection of all finite signed measures on $(X, \mathfrak{M})$. Then $M(X)$ is a linear space and

$$\|\mu\| = |\mu|(X)$$

defines a norm on $M(X)$.

Theorem 61. *$M(X)$ is a Banach space.*

We leave the proof to interested students.

Remark 22. *Let μ be a positive measure on $(X, \mathfrak{M})$, then $L^1(\mu)$ is a closed supspace of $M(X)$.*

2. Absolute Continuity and Radon-Nikodym Theorem

Definition 34. *Let* μ, ν *be two signed measure on* $(X, \mathfrak{M})$*.* ν *is said to be absolutely continuous with respect to* μ *if for any* $E \in \mathfrak{M}$*,* $|\mu|(E) = 0$ *implies* $\nu(E) = 0$*. And we denote* $\nu \ll \mu$*.*

Proposition 4. $\nu \ll \mu$ *if and only if* $|\nu| \ll |\mu|$*.*

Remark 23. *Let* μ, ν *be two positive measure on* $(X, \mathfrak{M})$*. Then* $\nu \leq \mu$ *implies* $\nu \ll \mu$*.*

Let μ be a measure on $(X, \mathfrak{M})$ and $f \in L^1(\mu)$. Then

$$\nu(E) = \int_E f d\mu \text{ for } E \in \mathfrak{M}$$

defines a finite signed measure which is absolutely continuous with respect to μ. The converse statement is also true if we require that μ is σ-finite.

Theorem 62 (Radon-Nikodym). *Let* ν *be a finite signed measure that is absolutely continuous with respect to a* σ*-finite measure* μ*. Then there exists a unique function* $f \in L^1(\mu)$ *such that*

$$\nu(E) = \int_E f d\mu$$

holds for all $E \in \mathfrak{M}$*.*

Proof. This is a corollary of the following version on positive measures. ☐

Theorem 63 (Radon-Nikodym). *Let* μ, ν *be two* σ*-finite measures on* $(X, \mathfrak{M})$ *and* $\nu \ll \mu$*. Then there exists a unique nonnegative measurable function* f *such that*

$$\nu(E) = \int_E f d\mu$$

holds for all $E \in \mathfrak{M}$*.*

Proof. We first assume that $\nu(X) < \infty$ and $\mu(X) < \infty$. Define

$$\Sigma = \left\{ f \in L^1(\mu) : f \geq 0, \ \int_E f d\mu \leq \nu(E) \ \text{ for any } E \in \mathfrak{M} \right\}.$$

Since $0 \in \Sigma$, Σ is nonempty. Also for any $f, g \in \Sigma$, we have $h = \max\{f, g\} \in \Sigma$. Indeed, for any $E \in \mathfrak{M}$, let

$$\begin{aligned} E_1 &= \{x \in E : f(x) \geq g(x)\} \text{ and} \\ E_2 &= \{x \in E : f(x) < g(x)\} = E \backslash E_1. \end{aligned}$$

Then

$$\begin{aligned} \int_E h d\mu = \int_{E_1} h d\mu + \int_{E_2} h d\mu &= \int_{E_1} f d\mu + \int_{E_2} g d\mu \\ &\leq \nu(E_1) + \nu(E_2) = \nu(E). \end{aligned}$$

Let

$$a = \sup_{f \in \Sigma} \int_X f d\mu,$$

then we have $a \leq \nu(X) < \infty$. So there exists $f_n \in \Sigma$, such that

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

Now let g_n be such that

$$g_n(x) = \max_{1\leq k\leq n} f_k(x),$$

then $g_n \in \Sigma$. Since g_n is monotone increasing, we can define

$$g(x) = \lim_{n\to\infty} g_n(x).$$

Now Lebesgue's monotone increasing theorem implies that for any $E \in \mathfrak{M}$,

$$\int_E g d\mu = \lim_{n\to\infty} \int_E g_n d\mu \leq \nu(E).$$

We also have

$$\int_X g d\mu = a.$$

Now we claim $\int_E g d\mu = \nu(E)$ for any $E \in \mathfrak{M}$. We define for any $E \in \mathfrak{M}$,

$$\lambda(E) = \nu(E) - \int_E g d\mu.$$

Then λ is a measure. If $\lambda(A) > 0$ for some $A \in \mathfrak{M}$. Since $\mu(A) < \infty$, there exists $\varepsilon > 0$ such that

$$\lambda(A) - \varepsilon\mu(A) > 0.$$

Since $\lambda - \varepsilon\mu$ is a signed measure with $(\lambda - \varepsilon\mu)(A) > 0$, there exists a measurable set $B \subset A$ which is positive with respect to $\lambda - \varepsilon\mu$. Now for any $E \in \mathfrak{M}$,

$$\nu(E) - \int_E (g + \varepsilon\chi_B) d\mu = \lambda(E) - \varepsilon\mu(E \cap B) \geq (\lambda - \varepsilon\mu)(E \cap B) \geq 0.$$

So $g + \varepsilon\chi_B \in \Sigma$. Now

$$(\lambda - \varepsilon\mu)(B) > 0 \Rightarrow \lambda(B) > 0 \Rightarrow \nu(B) > 0 \Rightarrow \mu(B) > 0$$

since $\nu \ll \mu$. Hence

$$\int_X (g + \varepsilon\chi_B) d\mu = \int_X g d\mu + \varepsilon\mu(B) > a$$

which is a contradiction. Hence we have proved $\int_E g d\mu = \nu(E)$ for any $E \in \mathfrak{M}$.

Now if ν, μ are σ-finite. Let $X = \bigcup_{n=1}^{\infty} E_n$ where $E_n \in \mathfrak{M}$ are disjoint and $\nu(E_n) < \infty$, $\mu(E_n) < \infty$. Then $\nu \ll \mu$ on E_n, hence there exists $f_n \in L^1(E_n, \mu)$ such that for any $E \in \mathfrak{M}$, $E \subset E_n$,

$$\nu(E) = \int_E f_n d\mu.$$

Let $f = f_n$ on E_n, then we have for any $E \in \mathfrak{M}$, $\nu(E) = \int_E f_n d\mu$.

The uniqueness is easy to proof. □

DEFINITION 35. *Two signed measures μ and ν are said to be singular (or orthogonal), in symbols $\mu \perp \nu$, if there exist $A, B \in \mathfrak{M}$, $A \cap B = \emptyset$, $A \cup B = X$ and $|\mu|(A) = |\nu|(B) = 0$.*

THEOREM 64 (Lebesgue Decomposition Theorem). *Let μ and ν be two σ-finite measures on $(X, \mathfrak{M})$. Then there exist two unique measures ν_1 and ν_2, such that*

$$\nu = \nu_1 + \nu_2,$$

$$\nu_1 \ll \mu \text{ and } \nu_2 \perp \mu.$$

PROOF. Let $\lambda \equiv \mu + \nu$. Since $\mu, \nu \ll \lambda$, there exist nonnegative measurable functions f and g such that

$$\mu(E) = \int_E f d\lambda \text{ and } \nu(E) = \int_E g d\lambda$$

holds for all $E \in \mathfrak{M}$. Let $X_0 = \{x \in X : f(x) = 0\}$ and for any $E \in \mathfrak{M}$, we define

$$\nu_1(E) = \nu(E \backslash X_0), \ \nu_2(E) = \nu(E \cap X_0).$$

Then ν_1, ν_2 are two measures s.t. $\nu = \nu_1 + \nu_2$. Since $\mu(X_0) = 0$ and $\nu_2(X \backslash X_0) = 0$, we have $\nu_2 \perp \mu$. On the other hand, if $\mu(E) = 0$ for some $E \in \mathfrak{M}$, then $f = 0$ λ-a.e. on E, hence, $\lambda(E \backslash X_0) = 0$, so we have

$$\nu_1(E) = \nu(E \backslash X_0) = \int_{E \backslash X_0} g d\lambda = 0.$$

Hence, $\nu_1 \ll \mu$. □

3. Bounded Linear Functionals on L^p

Let X, Y be two normed linear space over $\mathbb{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. A linear transformation $F : X \to Y$ is said to be bounded if there exists $M \geq 0$, such that for any $x \in X$,

$$\|Fx\|_Y \leq M \|x\|_X .$$

We define $\|F\|$ to be the smallest such M, then

$$\|F\| = \sup \{\|Fx\|_Y : x \in X, \|x\|_X \leq 1\} = \sup_{\|x\|_X = 1} \|Fx\|_Y .$$

Theorem 65. *The following three conditions are equivalent:*
(1). F is bounded.
(2). F is continuous.
(3). F is continuous at one point of X.

Proof. (1)⇒(2): Let F be bounded. For any $x_1 \neq x_2$, we have

$$\|Fx_1 - Fx_2\|_Y = \|F(x_1 - x_2)\|_Y \leq \|F\| \|x_1 - x_2\|_X .$$

So F is continuous.

(2)⇒(3) is trivial.

(3)⇒(1): Assume F is continuous at x_0, then for any $\varepsilon > 0$, there exists $\delta > 0$ such that $\|x - x_0\|_X \leq \delta$ implies

$$\|Fx - Fx_0\|_Y \leq \varepsilon.$$

From the linearity of F, $\|x\|_X \leq \delta$ implies

$$\|Fx\|_Y \leq \varepsilon.$$

Hence, $\|F\| \leq \frac{\varepsilon}{\delta}$ and F is bounded. □

When $Y = \mathbb{K}$, a linear transformation $F : X \to \mathbb{K}$ is called a linear functional.

Let (X, μ) be a measure space and $1 \leq p \leq \infty$. Then for each $g \in L^q(\mu)$, with $\frac{1}{p} + \frac{1}{q} = 1$, here $q = \infty$ if $p = 1$ and $q = 1$ if $p = \infty$, since

$$\left| \int_X fg d\mu \right| \leq \int_X |fg| d\mu \leq \|f\|_p \|g\|_q$$

for any $f \in L^p(\mu)$, the functional F_g defined by

$$F_g(f) = \int_X fg d\mu, \text{ for any } f \in L^p(\mu),$$

is a bounded linear functional with

$$\|F_g\| \leq \|g\|_q .$$

On the other hand, we have

Theorem 66 (Riesz). *Let (X, μ) be a σ-finite measure space and $1 \leq p < \infty$. Let F be a bounded linear functional on $L^p(\mu)$. Then there exists a unique $g \in L^q(\mu)$, where $\frac{1}{p} + \frac{1}{q} = 1$, such that*

$$F(f) = \int_X fg d\mu$$

holds for any $f \in L^p(\mu)$ and $\|F\| = \|g\|_q$.

Proof. We first assume $\mu(X) < \infty$. Define for any $A \in \mathfrak{M}$,

$$\nu(A) = F(\chi_A).$$

Then $\nu(\emptyset) = 0$. And if $A = \bigcup_{k=1}^{\infty} A_k$ is a disjoint union of measurable set, then we have

$$\left\| \sum_{i=1}^{k} \chi_{A_i} - \chi_A \right\|_p = \left(\sum_{i=1}^{k} \mu(A_i) - \mu(A) \right)^{\frac{1}{p}} \to 0 \tag{3.1}$$

as $k \to \infty$. Since F is linear and continuous, we have

$$\lim_{k\to\infty} \left| \sum_{i=1}^{k} \nu(A_i) - \nu(A) \right| = 0,$$

i.e.,

$$\nu(A) = \sum_{i=1}^{\infty} \nu(A_i).$$

Hence, ν is a finite signed measure. If $\mu(A) = 0$ for some $A \in \mathfrak{M}$, then $\chi_A = 0$ a.e., and hence $\nu(A) = F(\chi_A) = 0$. So we have $\nu \ll \mu$. Radon-Nikodym theorem implies the existence of $g \in L^1(\mu)$, such that

$$\nu(A) = F(\chi_A) = \int_A g d\mu.$$

From the linearity of F, we have for any simple function s,

$$F(s) = \int_X sg d\mu.$$

Since every $f \in L^\infty(\mu)$ is a uniform limit of simple functions, we have

$$F(f) = \int_X fg d\mu.$$

Now we claim $f \in L^q(\mu)$. When $p = 1$, let $M > 0$ be such that

$$E_M = \{x \in X : |g| \geq M\}$$

has positive μ-measure. Let $f(x) = \alpha(x)\chi_{E_M}$ where

$$\alpha(x) = \begin{cases} \frac{\overline{g(x)}}{|g(x)|} & \text{if } g(x) \neq 0, \\ 0 & \text{if } g(x) = 0. \end{cases}$$

We have $\|f\|_1 = \mu(E_M)$. Since $f \in L^\infty(\mu)$,

$$|F(f)| = \left| \int_X fg d\mu \right| \geq M\mu(E_M).$$

On the other hand, $|F(f)| \leq \|F\| \|f\|_1 = \|F\| \mu(E_M)$, so we have $M \leq \|F\|$. Hence $g \in L^\infty(\mu)$, and $\|g\|_\infty \leq \|F\|$. Since $L^\infty(\mu)$ is dense in $L^1(\mu)$, and $\int_X fg d\mu$ is continuous for $f \in L^1(\mu)$,

$$F(f) = \int_X fg d\mu$$

holds for any $f \in L^1(\mu)$. We also have $\|F\| \leq \|g\|_\infty$, hence $\|F\| = \|g\|_\infty$.

When $1 < p < \infty$, let

$$A_n = \{x \in X : |g| \leq n\}$$

and

$$f(x)=\alpha|g|^{\frac{q}{p}}\chi_{A_n}\in L^{\infty}(\mu).$$

Then

$$|F(f)|=\left|\int_X fgd\mu\right|=\int_{A_n}|g|^q\,d\mu,$$

and we also have

$$|F(f)|\le\|F\|\,\|f\|_p=\|F\|\left(\int_{A_n}|g|^q\,d\mu\right)^{\frac{1}{p}}.$$

Hence,

$$\int_{A_n}|g|^q\,d\mu\le\|F\|^q.$$

Lebesgue monotone increasing theorem implies

$$\int_X|g|^q\,d\mu\le\|F\|^q.$$

Hence, $g\in L^q(\mu)$ and $\|g\|_q\le\|F\|$. Since $L^{\infty}(\mu)$ is dense in $L^p(\mu)$, and $\int_X fgd\mu$ is continuous for $f\in L^p(\mu)$,

$$F(f)=\int_X fgd\mu$$

holds for any $f\in L^p(\mu)$. We also have $\|F\|\le\|g\|_q$, hence $\|F\|=\|g\|_q$. We leave the uniqueness of g as an exercise.

Now if $\mu(X)=\infty$, we write $X=\bigcup\limits_{n=1}^{\infty}X_n$ where X_n is monotone increasing and $\mu(X_n)<\infty$. Then for each n, there exists $g_n\in L^q(X_n,\mu)$, such that for any $f\in L^p(X,\mu)$, $f=0$ a.e. on $X\backslash X_n$,

$$F(f)=\int_X fg_n d\mu.$$

From the uniqueness, the function $g=g_n$ on X_n is well defined. Since $\|g_n\|_q\le\|F\|$ and $|g_n|$ is monotone increasing, we conclude $\|g\|_q\le\|F\|$. For any $f\in L^p(X,\mu)$, since

$$F(f\chi_{X_n})=\int_X f\chi_{X_n}gd\mu$$

and $f\chi_{X_n}\to f$ in $L^p(X,\mu)$, we conclude

$$F(f)=\int_X fgd\mu.$$

Finally, we also have $\|F\|\le\|g\|_q$, hence $\|F\|=\|g\|_q$. □

REMARK 24. *The proof fails when $p=\infty$ since (3.1) doesn't hold for $p=\infty$. Actually, consider Lebesgue measure on $\mathbb{R}^n$, then $C_0(\mathbb{R}^n)$ is a closed subspace of $L^{\infty}(\mathbb{R}^n)$. Define a bounded linear functional $\delta: C_0(\mathbb{R}^n)\to\mathbb{K}$ by*

$$\delta(f)=f(0).$$

Then functional analysis implies that δ can be extended to a bounded linear functional F_δ on $L^{\infty}(\mathbb{R}^n)$, apparently F_δ can't be represented by a function in $L^1(\mathbb{R}^n)$.

REMARK 25. *The dual space of $L^{\infty}(\mu)$ is the space of all finitely additive bounded measures that are absolutely continuous with respect to μ.*

4. Bounded Linear Functionals on $C_0(X)$

Let X be a locally compact Hausdorff space. Theorem 39 characterizes the positive linear functionals on $C_c(X)$. Since $C_c(X)$ is a dense subspace of $C_0(X)$, any bounded linear functional on $C_c(X)$ can be uniquely extended to a bounded linear functional on the Banach space $C_0(X)$.

Let μ be a Borel signed measure. Define for any $f \in C_0(X)$,

$$\int_X f d\mu \equiv \int_X f d\mu^+ - \int_X f d\mu^-.$$

Then $f \to \int_X f d\mu$ is a bounded linear functional on $C_0(X)$.

Theorem 67. *Let X be a locally compact Hausdorff space and F be a bounded linear functional on the Banach space $C_0(X)$. Then there is a unique regular signed measure μ, such that*

$$F(f) = \int_X f d\mu$$

for any $f \in C_0(X)$. Moreover,

$$\|F\| = \|\mu\|.$$

Proof. Define for any $f \in C_c(X)$, $f \geq 0$,

$$G(f) = \sup\{F(h) : h \in C_c(X), |h| \leq f\}.$$

And for general $f \in C_c(X)$, write $f = f^+ - f^-$, we define

$$G(f) = G(f^+) - G(f^-).$$

Now we verify that G is linear. For any $c \in R$, $G(cf) = cG(f)$ is obvious. For any $f, g \in C_c(X)$, $f, g \geq 0$, since $|h_1| \leq f$, $|h_2| \leq g$ imply $|h_1 + h_2| \leq f + g$, we have

$$G(f+g) \geq G(f) + G(g).$$

On the other hand, if $h \in C_c(X), |h| \leq f + g$, let

$$V = \{x : f(x) + g(x) > 0\},$$

we define

$$h_1(x) = \frac{f(x)}{f(x)+g(x)} h(x), h_2(x) = \frac{g(x)}{f(x)+g(x)} h(x) \text{ for } x \in V,$$

and $h_1(x) = h_2(x) = 0$ for $x \notin V$.

Then $|h_1| \leq f$, $|h_2| \leq g$ and we have

$$G(f+g) \leq G(f) + G(g).$$

Hence $G(f+g) = G(f) + G(g)$. Now for any $f, g \in C_c(X)$, we have

$$(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+,$$

hence

$$G\left((f+g)^+\right) + G(f^-) + G(g^-) = G\left((f+g)^-\right) + G(f^+) + G(g^+)$$

which implies

$$G(f+g) = G(f) + G(g).$$

Hence G is a positive linear functional on $C_c(X)$. There exists a regular Borel measure λ such that

$$G(f) = \int_X f d\lambda$$

for any $f \in C_c(X)$. Moreover, we have

$$\begin{aligned} \lambda(X) &= \sup\{G(f) : f \in C_c(X), 0 \le f \le 1\} \\ &= \sup\{F(h) : h \in C_c(X), |h| \le 1\} = \|F\|. \end{aligned}$$

Now

$$|F(f)| \le G(|f|) = \int_X |f| \, d\lambda = \|f\|_1.$$

Since $C_c(X)$ is dense in $L^1(\lambda)$, F extends to a bounded linear functional on $L^1(\lambda)$ with norm bounded by 1, hence, there exists a Borel function $g \in L^\infty(\lambda)$, $|g| \le 1$ such that

$$F(f) = \int_X fg d\lambda.$$

Let $\mu = gd\lambda$, then μ is a signed measure and for any $f \in C_c(X)$,

$$F(f) = \int_X f d\mu.$$

Now

$$|F(f)| \le \int_X |f| \, d|\mu| \le \|f\|_\infty |\mu|(X),$$

we have

$$\|F\| \le |\mu|(X).$$

On the other hand,since $|g| \le 1$, $|\mu|(X) \le \lambda(X) = \|F\|$. Hence we have

$$\|\mu\| = |\mu|(X) = \|F\|.$$

Finally, we show μ is unique. It suffices to show that if $\int_X f d\mu = 0$ holds for all $f \in C_c(X)$, then $\mu = 0$. We write $\mu = gd|\mu|$ for some g satisfying $|g| = 1$, pick $f_n \in C_c(X)$ such that $f_n \to g$ in $L^1(|\mu|)$, then we have

$$0 = \lim_{n\to\infty} \int_X f_n d\mu = \lim_{n\to\infty} \int_X f_n g d|\mu| = \int_X g^2 d|\mu| = |\mu|(X).$$

Hence $\mu = 0$.

□

5. Derivatives of Measure and Lebesgue point

To motivate the definition of derivative of measures, we first look at a simple theorem for measures on real line:

THEOREM 68. *Suppose* μ *is a signed Borel measure on* $\mathbb{R}$ *and*

$$f(x) = \mu((-\infty, x)).$$

Then the following two statements are equivalent:
(i). f *is differentiable at* x *and* $f'(x) = A$.
(ii). To every $\varepsilon > 0$, *there exists* $\delta > 0$, *s.t.,*

$$\left| \frac{\mu(I)}{m(I)} - A \right| < \varepsilon$$

for every open interval I *satisfying* $x \in I$ *and* $|I| < \delta$. *Here* m *is the Lebesgue measure on* $\mathbb{R}$.

DEFINITION 36. *Let* μ *be a signed Borel measure on* $\mathbb{R}^k$. *The symmetric derivative of* μ *at* x *is defined by*

$$(D\mu)(x) = \lim_{r \to 0} \frac{\mu(B_r(x))}{|B_r(x)|}$$

whenever the limit exists. Here

$$B_r(x) = \left\{ y \in \mathbb{R}^k : |y - x| < r \right\}$$

and $|B_r(x)| = m(B_r(x))$ *where* $m = L_k$ *is the Lebesgue measure on* $\mathbb{R}^k$.

DEFINITION 37. *Let* μ *be a signed Borel measure on* $\mathbb{R}^k$. *The maximal function of* μ *is defined by*

$$(M\mu)(x) = \sup_{r>0} \frac{|\mu|(B_r(x))}{|B_r(x)|}.$$

Since for each $\alpha \in \mathbb{R}$,

$$\left\{ x \in \mathbb{R}^k : M\mu > \alpha \right\} = \bigcup_{r>0} \left\{ x \in \mathbb{R}^k : \frac{|\mu|(B_r(x))}{|B_r(x)|} > \alpha \right\}$$

is open, the maximal function $M\mu : \mathbb{R}^k \to [0, \infty]$ is Borel measurable.

LEMMA 6. *If* W *is the union of a finite collection of balls* $B_{r_i}(x_i)$, $1 \le i \le N$, *then there is a set* $S \subset \{1, 2, \cdots, N\}$ *so that*
(a). the balls $B_{r_i}(x_i)$ *with* $i \in S$ *are disjoint,*
(b). $W \subset \bigcup_{i \in S} B_{3r_i}(x_i)$ *and*
(c). $m(W) \le 3^k m(B_{r_i}(x_i))$.

PROOF. We assume $r_1 \ge r_2 \ge \cdots \ge r_N$ and we write $B_i = B_{r_i}(x_i)$. Put $i_1 = 1$. Discard all B_k that intersect B_{i_1}. Let B_{i_2} be the first of the remaining B_k if there are any. Discard all B_k that intersect B_{i_2}. Let B_{i_3} be the first of the remaining B_k if there are any. This process stops after a finite number of steps and gives $S = \{i_1, i_2, \cdots\}$. (a) is clear to hold. Since each discarded B_k intersect with some $B_{r_i}(x_i)$, $i \in S$ with larger radius, $B_k \subset B_{3r_i}(x_i)$. Hence (b) holds. Finally, (c) follows from (b). □

THEOREM 69. *If μ is a signed Borel measure on $\mathbb{R}^k$, then for any $t > 0$,*

$$m\left(\{M\mu > t\}\right) \leq \frac{3^k}{t} \left\|\mu\right\|. \tag{5.1}$$

PROOF. We assume $\|\mu\| < \infty$. Given $t > 0$. Let K be a compact subset of $\{M\mu > t\}$. Then each $x \in K$ is the center of an open ball B for which $|\mu|(B) > t|B|$. Some finite collection of these B's covers K, and the covering lemma gives us a disjoint subcollection, say $\{B_1, \cdots, B_n\}$, such that

$$m(K) \leq 3^k \sum_{i=1}^{n} m(B_i) < \frac{3^k}{t} \sum_{i=1}^{n} |\mu|(B_i) \leq \frac{3^k}{t} |\mu|\left(\mathbb{R}^k\right) = \frac{3^k}{t} \|\mu\|.$$

Taking supremum over all such K, we have (5.1). □

Let $f \in L^1\left(\mathbb{R}^k\right)$. Then $\mu = f dm$ defines a signed Borel measure on $\mathbb{R}^k$ with $\|\mu\| = \|f\|_1$. And we can define the maximal function of f

$$(Mf)(x) = (M\mu)(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f| \, dm.$$

COROLLARY 6. *For any $t > 0$,*

$$m\left(\{Mf > t\}\right) \leq \frac{3^k}{t} \|f\|_1.$$

DEFINITION 38. *A Lebesgue measurable function f is said to belong to weak L^1, if*

$$tm\left(\{|f| > t\}\right)$$

is a bounded function of t on $(0, \infty)$.

Hence $f \in L^1\left(\mathbb{R}^k\right)$ implies that Mf belongs to weak L^1.

DEFINITION 39. *Let $f \in L^1\left(\mathbb{R}^k\right)$. A point $x \in \mathbb{R}^k$ is said to be a Lebesgue point of f if*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \, dy = 0.$$

If x is a Lebesgue point of f, then

$$f(x) = \lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} f dm.$$

THEOREM 70. *Let $f \in L^1\left(\mathbb{R}^k\right)$. Then almost every point $x \in \mathbb{R}^k$ is a Lebesgue point of f.*

PROOF. Define for $x \in \mathbb{R}^k$, $r > 0$,

$$(T_r f)(x) = \frac{1}{|B_r|} \int_{B_r(x)} |f(y) - f(x)| \, dy.$$

Let

$$(Tf)(x) = \limsup_{r\to 0} (T_r f)(x).$$

For any $n \in \mathbb{N}$, there exists $g \in C_c\left(\mathbb{R}^k\right)$, such that $\|f-g\|_1 < \frac{1}{n}$. Let $h = f - g$. Since g is continuous, $Tg \equiv 0$. Since

$$\begin{aligned}(T_r h)(x) &= \frac{1}{|B_r|}\int_{B_r(x)} |h(y) - h(x)|\, dy \le \frac{1}{|B_r|}\int_{B_r(x)} |h(y)|\, dy + |h(x)| \\ &\le Mh(x) + |h(x)|,\end{aligned}$$

we have

$$(Th)(x) \le Mh(x) + |h(x)|.$$

Since $T_r f \le T_r g + T_r h$, we have

$$(Tf)(x) \le (Th)(x) \le Mh(x) + |h(x)|.$$

Now for any $\varepsilon > 0$,

$$\{Tf > 2\varepsilon\} \subset \{Mh > \varepsilon\} \bigcup \{|h| > \varepsilon\} \equiv E_\varepsilon^n,$$

and hence

$$\{Tf > 2\varepsilon\} \subset \bigcap_{n=1}^{\infty} E_\varepsilon^n.$$

Since

$$m\left(E_\varepsilon^n\right) \le \left(3^k + 1\right)\frac{1}{n\varepsilon},$$

$$m\left(\bigcap_{n=1}^{\infty} E_\varepsilon^n\right) = 0.$$

Since Lebesgue measure is complete, $\{Tf > 2\varepsilon\}$ is measurable with Lebesgue measure zero. Since ε is arbitrary, $Tf = 0$ m-a.e.. □

Let E be a Lebesgue measurable subset of $\mathbb{R}^k$. The metric density of E at x is defined by

$$\lim_{r\to 0} \frac{m(E \cap B_r(x))}{m(B_r(x))}$$

if the limit exists.

Corollary 7. *The metric density of E*

$$\lim_{r\to 0} \frac{m(E \cap B_r(x))}{m(B_r(x))} = \chi_E$$

a.e..

Theorem 71. *Suppose μ is a signed Borel measure on $\mathbb{R}^k$, and $\mu \ll m$. Let f be the Radon-Nikodym derivative of μ w.r.t. m. Then $D\mu = f$ m-a.e..*

Proof. At any Lebesgue point of f,

$$f(x) = \lim_{r\to 0} \frac{1}{|B_r|}\int_{B_r(x)} f\, dm = \lim_{r\to 0} \frac{\mu(B_r)}{|B_r|} = D\mu.$$

□

Theorem 72. *Let μ be a Borel measure on $\mathbb{R}^k$, and let A be a Borel set such that $\mu(A) = 0$. Then*

$$D\mu(x) = 0 \text{ } m\text{-a.e. on } A.$$

PROOF. If $D\mu(x) \neq 0$, then we must have

$$(D^*\mu)(x) = \limsup_{r \to 0} \frac{\mu(B_r(x))}{|B_r(x)|} > 0.$$

We only need to show that for any $\varepsilon > 0$, the set

$$E_\varepsilon = \{x \in A : (D^*\mu)(x) > \varepsilon\}$$

has Lebesgue measure zero. Let K be a compact subset of E_ε and V be an open set containing A, then for each $x \in K$, there exists a ball $B_r(x) \subset V$ such that

$$\frac{\mu(B_r(x))}{|B_r(x)|} > \varepsilon.$$

Since K is compact, some finite collection of these balls covers K, and the covering lemma gives us a disjoint subcollection, say $\{B_1, \cdots, B_n\}$, such that

$$m(K) \leq 3^k \sum_{i=1}^{n} m(B_i) < \frac{3^k}{\varepsilon} \sum_{i=1}^{n} \mu(B_i) \leq \frac{3^k}{\varepsilon} \mu(V).$$

Since μ is regular and $\mu(A) = 0$, we can choose V with arbitrarily small $\mu(V)$, hence $m(K) = 0$. Finally, since m is regular, $m(K) = 0$ for all compact $K \subset A$ implies $m(E_\varepsilon) = 0$. □

COROLLARY 8. *Let μ be a Borel measure on $\mathbb{R}^k$ such that $\mu \perp m$. Then*

$$D\mu(x) = 0 \ m\text{-a.e. on } \mathbb{R}^k.$$

PROOF. Since $\mu \perp m$, $\mu(A) = 0$ for some Borel set A with $m(A^c) = 0$. Now $\mu(A) = 0$ implies that $D\mu(x) = 0$ m-a.e. on A and $m(A^c) = 0$ then implies $D\mu(x) = 0$ m-a.e. on $\mathbb{R}^k$. □

THEOREM 73. *If $f \in L^1(\mathbb{R})$ and*

$$F(x) = \int_{-\infty}^{x} f\,dm,$$

then $F'(x) = f(x)$ at every Lebesgue point of f, hence m-a.e..

PROOF. Let x be a Lebesgue point, then for any $h \neq 0$,

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| = \left| \frac{1}{h} \int_x^{x+h} (f(y) - f(x))\,dy \right|$$

$$\leq \frac{1}{|h|} \left| \int_x^{x+h} |f(y) - f(x)|\,dy \right| \leq 2 \frac{1}{2|h|} \int_{x-|h|}^{x+|h|} |f(y) - f(x)|\,dy.$$

Hence

$$\lim_{h \to 0} \left| \frac{F(x+h) - F(x)}{h} - f(x) \right| = 0,$$

i.e., $F'(x) = f(x)$. □

6. Differentiability of monotone functions

Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function. Then the set of all discontinuities of f is at most countable, moreover, we have

THEOREM 74. *Every monotone function is differentiable almost everywhere.*

PROOF. Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function. By replacing f with $-f$ if necessary, we could assume that f is monotone increasing. Now we define for any $x \in \mathbb{R}$,

$$f^*(x) = \lim_{y \to x^+} f(y).$$

Then f^* is also increasing. Now for any $x \in \mathbb{R}$, let $x_n \in \mathbb{R}$ be such that f is continuous at x_n and $x_n \searrow x$, then we have

$$f^*(x+) = \lim_{n\to\infty} f^*(x_n) = \lim_{n\to\infty} f(x_n) = f^*(x).$$

Hence, f^* is right continuous. Moreover, we can check that for any $x \in \mathbb{R}$,

$$f^*(x+) - f^*(x-) = f(x+) - f(x-).$$

The third homework shows that f^* induces a Borel measure $\mu = \mu_{f^*}$, such that for any $a < b$,

$$\mu((a,b)) = f^*(b-) - f^*(a),$$
$$\mu([a,b]) = f^*(b) - f^*\left(a^-\right).$$

Let $\mu = \mu_1 + \mu_2$ be the Lebesgue decomposition with respect to Lebesgue measure m such that $\mu_1 \ll m$ and $\mu_2 \perp m$. Let h be the Radon-Nikodym derivative of μ_1 w.r.t. m. Let

$$A = \{x \in \mathbb{R} : x \text{ is Lebesgue point of } h \text{ and } (D\mu_2)(x) = 0\}.$$

Then Theorem 70 and Corollary 8 implies $m(A^c) = 0$. Let $x \in A$, we claim that for any $\varepsilon > 0$, there exists $\delta > 0$, such that

$$\left|\frac{\mu(a,b)}{b-a} - h(x)\right| < \varepsilon \tag{6.1}$$

whenever $\frac{1}{3}|I_{ab}| \le b - a < \delta$, where $I_{ab} = (x-t, x+t)$ and $t = \max(|b-x|, |a-x|)$. To see this,

$$\begin{aligned}\left|\frac{\mu(a,b)}{b-a} - h(x)\right| &\le \left|\frac{\mu_1(a,b)}{b-a} - h(x)\right| + \left|\frac{\mu_2(a,b)}{b-a}\right| \\ &\le \frac{1}{b-a}\int_a^b |h(y) - h(x)| + \left|\frac{\mu_2(I_{ab})}{b-a}\right| \\ &\le \frac{3}{|I_{ab}|}\int_{I_{ab}} |h(y) - h(x)| + 3\left|\frac{\mu_2(I_{ab})}{|I_{ab}|}\right|\end{aligned}$$

which approaches zero as $|I_{ab}| \to 0$, and hence as $b - a \to 0$. Now for any x' such that $0 < x' - x < \delta$, we can find a_n,b_n such that $x < a_n < b_n < x'$, $b - a \ge \frac{1}{3}|I_{ab}|$ and $\lim a_n = x$, $\lim b_n = x'$, then we have

$$\begin{aligned}\frac{f(x') - f(x)}{x' - x} - h(x) &\ge \frac{\mu(a_n, b_n)}{x' - x} - h(x) \\ &\ge \frac{(b_n - a_n)(h(x) - \varepsilon)}{x' - x} - h(x).\end{aligned}$$

Letting $n \to \infty$, we have

$$\frac{f(x') - f(x)}{x' - x} - h(x) \geq -\varepsilon.$$

Similarly, choosing a_n,b_n such that $a_n < x < x' < b_n$, $|b_n - a_n| < \delta$, $b_n - a_n \geq \frac{1}{3}|I_{ab}|$ and $\lim a_n = x$, $\lim b_n = x'$, we can prove

$$\frac{f(x') - f(x)}{x' - x} - h(x) \leq \varepsilon.$$

Hence

$$\lim_{x' \to x^+} \frac{f(x') - f(x)}{x' - x} = h(x).$$

Similarly, we can show

$$\lim_{x' \to x^-} \frac{f(x') - f(x)}{x' - x} = h(x).$$

Hence

$$f'(x) = h(x).$$

So we have proved $f'(x) = h(x)$ for any $x \in A$. Now $m(A^c) = 0$ implies $f'(x) = h(x)$ m-a.e.. □

Remark 26. *From the proof, we see $f'(x)$ captures the absolute continuous part of μ. It is not difficult to see that each jump discontinuity contributes a delta function to the singular part of μ. However, in general, μ_2 is not the summation of delta functions from jump discontinuities. Actually, there exists continuous monotone function such that the singular part μ_2 is not zero.*

Proposition 5. *Let $f : \mathbb{R} \to \mathbb{R}$ be a monotone function, then f' exists a.e., and $f' \in L^1_{\text{loc}}(\mathbb{R})$. Moreover,*

$$f(b) - f(a) \geq \int_b^a f'(x)\,dx \tag{6.2}$$

for any $a < b$.

Proof. Since $f' = h$, and μ_1 is locally finite measure, $f' \in L^1_{\text{loc}}(\mathbb{R})$. Now for any c, d such that $a < c < d < b$, we have

$$f(b) - f(a) \geq \mu(c, d) \geq \mu_1(c, d) = \int_c^d f'(x)\,dx,$$

let $c \to a$ and $d \to b$, we have (6.2). □

Remark 27. *There exists a strictly increasing function whose derivative equals to zero a.e.. Hence*

$$f(b) - f(a) \neq \int_b^a f'(x)\,dx$$

in general. On the other hand, if $f : \mathbb{R} \to \mathbb{R}$ is differentiable at every point of $[a, b]$ and $f' \in L^1$ on $[a, b]$, then

$$f(b) - f(a) = \int_b^a f'(x)\,dx$$

for any $a < b$. This is Theorem 7.21 in Rudin's book.

7. Functions of Bounded Variation

Let $[a,b]$ be a bounded interval of $\mathbb{R}$. $P = \{t_0, t_1, \cdots, t_n\}$ is said to be a partition of $[a,b]$ if

$$a = t_0 < t_1 < \cdots < t_n = b.$$

Let $f : [a,b] \to \mathbb{R}$ be a function. Then the total variation of f over $[a,b]$ is defined to be

$$\bigvee_a^b f = \sup \left\{ \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| : P = \{t_0, t_1, \cdots, t_n\} \text{ is a partition of } [a,b] \right\}.$$

Definition 40. *$f : [a,b] \to \mathbb{R}$ is said to be of bounded variation if* $\bigvee_a^b f < \infty$. *We write* $f \in \mathrm{BV}[a,b]$.

Proposition 6. *If $f, g \in BV[a,b]$ and $\alpha \in \mathbb{R}$, then so are $f+g, \alpha f, fg$ and $|f|$.*

Proof. It is easy to see that

$$\bigvee_a^b (f+g) \leq \bigvee_a^b f + \bigvee_a^b g,$$

$$\bigvee_a^b \alpha f = |\alpha| \bigvee_a^b f,$$

$$\bigvee_a^b (fg) \leq \bigvee_a^b f \sup_{x\in[a,b]} |g(x)| + \bigvee_a^b g \sup_{x\in[a,b]} |f(x)|,$$

$$\bigvee_a^b |f| \leq \bigvee_a^b f.$$

Hence $f+g, \alpha f, |f| \in BV[a,b]$. And since functions of bounded variation are bounded, $fg \in BV[a,b]$ as well. □

Proposition 7. *If $f \in BV[a,b]$, and F is a Lipschitz function on the range of f, then $F(f) \in \mathrm{BV}[a,b]$.*

Proof. Let L be the Lipschitz constant for F. Then

$$\bigvee_a^b F(f) \leq L \bigvee_a^b f.$$

□

Example 9. *If $f \in BV[a,b]$, then $\sin f \in BV[a,b]$. And if $f \in BV[a,b]$, $f \geq c > 0$, then $\frac{1}{f} \in \mathrm{BV}[a,b]$.*

Clearly, a monotone function f on $[a,b]$ is of bounded variation with

$$\bigvee_a^b f = |f(b) - f(a)|.$$

On the other hand, we have

THEOREM 75. *If $f : [a, b] \to \mathbb{R}$ is of bounded variation, then f is the difference of two increasing functions.*

PROOF. It is easy to check that $\bigvee_a^x f$, as a function of x, is increasing on $[a, b]$. Actually for any $x < y$,

$$\bigvee_a^x f + \bigvee_x^y f = \bigvee_a^y f.$$

Now define

$$g(x) = \bigvee_a^x f - f(x),$$

then for any $x < y$,

$$\begin{aligned} g(y) - g(x) &= \bigvee_x^y f - [f(y) - f(x)] \\ &\geq |f(y) - f(x)| - [f(y) - f(x)] \geq 0. \end{aligned}$$

Hence g is also increasing. Since $f(x) = \bigvee_a^x f - g(x)$, theorem follows. $\square$

COROLLARY 9. *If $f : [a, b] \to \mathbb{R}$ is of bounded variation, then*
(i). The discontinuity of f is at most countable;
(ii). f is differentiable almost everywhere.

8. Absolutely Continuous Functions

Definition 41. *A function $f : [a,b] \to \mathbb{R}$ is said to be absolutely continuous if for every $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $(a_1, b_1), \cdots, (a_n, b_n)$ are disjoint open intervals in $[a,b]$ satisfying*

$$\sum_{i=1}^{n} (b_i - a_i) < \delta,$$

we have

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| < \varepsilon.$$

And we write $f \in AC[a,b]$.

Example 10. *Lipschitz functions are absolute continuous.*

Theorem 76. *$AC[a,b] \subset BV[a,b]$.*

Proof. Assume $f \in AC[a,b]$. Then there exists $\delta > 0$ such that whenever $(a_1, b_1), \cdots, (a_n, b_n)$ are disjoint open intervals in $[a,b]$ satisfying

$$\sum_{i=1}^{n} (b_i - a_i) < \delta,$$

we have

$$\sum_{i=1}^{n} |f(b_i) - f(a_i)| < 1.$$

Now let n be the integer part of

$$\frac{b-a}{\delta} + 1,$$

then we have $\frac{b-a}{n} < \delta$. let $P = \{t_0, t_1, \cdots, t_n\}$ be a partition of $[a,b]$ such that

$$t_i = a + i\frac{b-a}{n}, 0 \le i \le n.$$

Then since $t_i - t_{i-1} < \delta$, from the definition of total variation, we have

$$\bigvee_{t_{i-1}}^{t_i} f \le 1.$$

Hence,

$$\bigvee_a^b f = \sum_{i=1}^{n} \bigvee_{t_{i-1}}^{t_i} f \le n.$$

So $f \in BV[a,b]$. □

Proposition 8. *If $f \in AC[a,b]$, then f is continuous.*

Proposition 9. *If $f, g \in AC[a,b]$ and $\alpha \in \mathbb{R}$, then so are $f + g, \alpha f, fg$ and $|f|$.*

Proof. This is similar to the proof of Proposition 6 for $BV[a,b]$. □

Proposition 10. *If $f \in AC[a,b]$, and F is a Lipschitz function on the range of f, then $F(f) \in AC[a,b]$.*

If $f \in L^1[a,b]$, then

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

is absolutely continuous on $[a,b]$. On the other hand, we have

Theorem 77. *If $F \in AC[a,b]$, then $f = F'$ is defined a.e., and $f \in L^1[a,b]$. Moreover, for each $x \in [a,b]$, we have*

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

Proof. Let μ_F be the finite Borel regular signed measure generated by function F. Since F is continuous, we have

$$\mu_F((x,y)) = F(y) - F(x)$$

for any $a \leq x < y \leq b$. Now we claim $\mu_F \ll m$ where m is the Lebesgue measure. Let $\varepsilon > 0$ and $\delta > 0$ be the number in the definition of absolutely continuous functions. For any Borel set $A \subset (a,b)$ with $m(A) = 0$, there exists an open set $U \subset (a,b)$ such that $A \subset U$, $m(U) < \delta$ and

$$|\mu_F(U) - \mu_F(A)| < \varepsilon.$$

Let $U = \bigcup_{i=1}^{\infty} (a_i, b_i)$, then we have from $m(U) < \delta$,

$$|\mu_F(U)| \leq \sum_{i=1}^{\infty} |\mu_F(a_i, b_i)| = \sum_{i=1}^{\infty} |F(b_i) - F(a_i)| \leq \varepsilon.$$

Hence $|\mu_F(A)| < 2\varepsilon$. Since ε is arbitrary, we have $\mu_F(A) = 0$. So we have $\mu_F \ll m$ and $f = F'$ is the Radon-Nikodym derivative of μ_F w.r.t. m. Hence $f' \in L^1[a,b]$ and for each $x \in [a,b]$, we have

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

$\square$

Theorem 78. *Let $f, g \in AC[a,b]$. Then $fg \in AC[a,b]$,*

$$(fg)' = f'g + fg'$$

almost everywhere on $[a,b]$ and the integration by parts formula

$$\int_a^b f' g\,dm = fg|_a^b - \int_a^b g' f\,dm$$

holds.

Theorem 79. *If $f \in AC[a,b]$, then*

$$\bigvee_a^b f = \int_a^b |f'|\,dm.$$

Proof. Step 1:

$$\bigvee_a^b f \leq \int_a^b |f'|\,dm.$$

Step 2: If $f' \in C\left[a,b\right]$, then $\bigvee\limits_a^b f = \int_a^b \left|f'\right| dm$.

Step 3: For general $f \in AC\left[a,b\right]$, let $g_n \in C\left[a,b\right]$ and $g_n \to f'$ in L^1. And we define

$$G_n\left(x\right) = \int_a^x g_n\left(x\right).$$

Then we have

$$\bigvee_a^b f \geq \bigvee_a^b G_n - \bigvee_a^b \left(G_n - f\right) \geq \int_a^b \left|g_n\right| dm - \int_a^b \left|g_n - f'\right| dm.$$

Let $n \to \infty$, we have

$$\bigvee_a^b f \geq \int_a^b \left|f'\right| dm.$$

The conclusion follows by combining step 1 and 3. $\square$

9. Lipschitz continuous functions

Let $(X, d_X), (Y, d_Y)$ be two metric spaces and $A \subset X$. $f : A \to Y$ is said to be Lipschitz continous if there exists constant $\tilde{L} > 0$ such that for any $x_1, x_2 \in A$,

$$d_Y (f (x_1), f (x_2)) \leq \tilde{L} d_X (x_1, x_2).$$

The smallest such $\tilde{L}$,

$$L = \operatorname{Lip} f$$

is called the Lipschitz constant of f.

It is easy to see Lipschitz continous function is continuous.

Theorem 80 (Mcshane-Whitney). *Let (X, d_X) be a metric space and $A \subset X$. Let $f : A \to \mathbb{R}$ be Lipschitz continous with*

$$\operatorname{Lip} f = L.$$

Then there exists a Lipschitz continuous function $\tilde{f} : X \to \mathbb{R}$ with $\operatorname{Lip} \tilde{f} = L$ and $\tilde{f}\Big|_A = f$.

Proof. It is easy to verify that

$$\tilde{f}(x) = \inf_{a \in A} (f(a) + L d_X (x, a))$$

satisfies all requirements. □

Remark 28. *Let*

$$f^* (x) = \sup_{a \in A} (f(a) - L d_X (x, a)).$$

Then if g is a Lipschitz extension of f with $\operatorname{Lip} g = L$. Then for any $x \in X$,

$$f^* (x) \leq g (x) \leq \tilde{f}(x).$$

Corollary 10. *Let (X, d_X) be a metric space and $A \subset X$. Let $f : A \to \mathbb{R}^n$ be Lipschitz continuous with*

$$\operatorname{Lip} f = L.$$

Then there exists a Lipschitz continuous function $\tilde{f} : X \to \mathbb{R}^n$ with $\operatorname{Lip} \tilde{f} \leq \sqrt{n} L$ and $\tilde{f}\Big|_A = f$.

Proof. We can extend each component of f. □

The extra factor $\sqrt{n}$ can be dropped if $X = \mathbb{R}^m$.

Theorem 81 (Kirszbraun's theorem). *Let $A \subset \mathbb{R}^m$ and $f : A \to \mathbb{R}^n$ be Lipschitz continuous. Then there exists a Lipschitz function $\tilde{f} : \mathbb{R}^m \to \mathbb{R}^n$ such that $\operatorname{Lip} \tilde{f} = \operatorname{Lip} f$ and $\tilde{f}\Big|_A = f$.*

Now we discuss the differentiability of Lipschitz functions. Since a Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous, it is differentiable a.e.. In higher dimensional case, let's recall the definition of differentiability: is said to be differentiable at $a \in \mathbb{R}^m$ if there exists a constant vector, denoted by $Df(a) \in \mathbb{R}^m$, such that

$$\lim_{x \to a} \frac{|f(x) - f(a) - Df(a) \cdot (x - a)|}{|x - a|} = 0.$$

THEOREM 82. *Let* Ω *be a domain in* $\mathbb{R}^m$ *and* $f:\Omega\to\mathbb{R}$ *be Lipschitz continuous. Then* f *is differentiable a.e. in* Ω.

PROOF. By extending f if necessary, we can assume $f:\mathbb{R}^m\to\mathbb{R}$ is Lipschitz continuous with $\operatorname{Lip} f = L$.

Step 1: For any $v\in\mathbb{S}^{n-1}$, the directional derivative $D_v f(x)$ is defined a.e.. Let

$$B=\{x\in\mathbb{R}^m : D_v f \text{ doesn't exist at } x\}.$$

Then

$$B=\left\{x\in\mathbb{R}^m : \liminf_{t\to 0}\frac{f(x+tv)-f(x)}{t}<\limsup_{t\to 0}\frac{f(x+tv)-f(x)}{t}\right\}$$

is Borel measurable since Borel measurability of both $\liminf_{t\to 0}\frac{f(x+tv)-f(x)}{t}$ and $\limsup_{t\to 0}\frac{f(x+tv)-f(x)}{t}$ can be proved using the same argument as Theorem 21. Since on each line l parallel to v, f is absolutely continuous and hence $D_v f$ exists a.e. on l, so $B\cap l$ has Lebesgue measure zero. Fubini theorem then implies $m(B)=0$.

Step 2: From step 1, $\nabla f(x)\equiv\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_m}\right)$ is defined a.e.. Now we check that for any $v\in\mathbb{S}^{n-1}$,

$$D_v f(x)=v\cdot\nabla f(x) \text{ a.e..}$$

Let $\varphi\in C_0^1(\mathbb{R}^m)$. Since $|D_v f(x)|\le L$ a.e., we have

$$\begin{aligned}&\int_{\mathbb{R}^m} D_v f(x)\varphi(x)\,dx=\int_{\mathbb{R}^m}\lim_{t\to 0}\frac{f(x+tv)-f(x)}{t}\varphi(x)\,dx\\&=\lim_{t\to 0}\int_{\mathbb{R}^m}\frac{f(x+tv)-f(x)}{t}\varphi(x)\,dx=\lim_{t\to 0}\int_{\mathbb{R}^m}\frac{\varphi(x-tv)-\varphi(x)}{t}f(x)\,dx\\&=-\int_{\mathbb{R}^m} f(x)D_v\varphi(x)\,dx=-v\cdot\int_{\mathbb{R}^m} f(x)\nabla\varphi(x)\,dx=v\cdot\int_{\mathbb{R}^m}\nabla f(x)\varphi(x)\,dx\\&=\int_{\mathbb{R}^m} v\cdot\nabla f(x)\varphi(x)\,dx.\end{aligned}$$

Hence from Lemma 7, $D_v f(x)=v\cdot\nabla f(x)$ a.e..

Step 3: For almost every $x\in\mathbb{R}^m$, f is differentiable at x and $Df(x)=\nabla f(x)$. Let $\{v_i\}_{i=1}^\infty$ be a countable dense subset of $\mathbb{S}^{n-1}$ and

$$A=\{x\in\mathbb{R}^m : \nabla f(x) \text{ and } D_{v_i}f(x) \text{ exists for every } i=1,2,\cdots\}.$$

For any $\varepsilon>0$, since $\mathbb{S}^{n-1}$ is compact, there exists N, such that for any $v\in\mathbb{S}^{n-1}$, there exists $1\le i\le N$ such that $|v_i-v|<\varepsilon$. Given $x\in A$, there exists $\delta>0$, such that for any $0<|t|<\delta$

$$\left|\frac{f(x+tv_i)-f(x)}{t}-D_{v_i}f(x)\right|<\varepsilon$$

holds for $1\le i\le N$. For any $v\in\mathbb{S}^{n-1}$, let $1\le i\le N$ be such that $|v_i-v|<\varepsilon$, we have for any $0<|t|<\delta$,

$$\begin{aligned}&\left|\frac{f(x+tv)-f(x)}{t}-v\cdot\nabla f(x)\right|\\&\le\left|\frac{f(x+tv_i)-f(x)}{t}-v_i\cdot\nabla f(x)\right|+L|v_i-v|+|\nabla f(x)||v_i-v|\\&<\varepsilon+L\varepsilon+\sqrt{m}L\varepsilon.\end{aligned}$$

Hence, f is differentiable at x and $Df\left(x\right)=\nabla f\left(x\right)$. Since A^c has measure zero, f is differentiable a.e.. ☐

PROOF.

LEMMA 7. *If* $g\in L^1_{loc}\left(\mathbb{R}^m\right)$ *satisfies, for any* $\varphi\in C^1_0\left(\mathbb{R}^m\right)$*,*

$$\int_{\mathbb{R}^m} g\left(x\right)\varphi\left(x\right)dx=0,$$

then $g=0$ *a.e. on* $\mathbb{R}^m$*.*

☐

PROOF. Let φ_ε be the standard mollifier. Then we have for any $x\in\mathbb{R}^m$,

$$g_\varepsilon\left(x\right)=\int_{\mathbb{R}^m} g\left(y\right)\varphi\left(x-y\right)dy=0.$$

Theorem 59 implies $g_\varepsilon\to g$ locally in L^1, hence $g=0$ a.e. on $\mathbb{R}^m$.

□

10. The change-of-variables theorem

Let V be an open set in $\mathbb{R}^k$. A mapping $T : V \to \mathbb{R}^k$ is said to be differentiable at $x \in V$ if there exists a linear mapping $A : \mathbb{R}^k \to \mathbb{R}^k$, such that

$$\lim_{h \to 0} \frac{\left| T(x+h) - T(x) - Ah \right|}{|h|} = 0.$$

Linear mapping A can be represented by a matrix which is called Jacobian matrix and we have

$$A = \left(\frac{\partial T_i}{\partial x_j} \right)_{k \times k} \equiv T'(x).$$

Its determinant is called the Jacobian and we write $J_T(x) = \det T'(x)$.

Theorem 83. *Let V be an open set in $\mathbb{R}^k$ and $T : V \to \mathbb{R}^k$ is one-to-one and differentiable at every point of V. Then for every measurable function $f : \mathbb{R}^k \to [0, \infty]$,*

$$\int_{T(V)} f dm = \int_V f \left| J_T \right| dm.$$

11. Distribution functions and maximal functions

Let (X, μ) be a σ-finite positive measure. Let $f : X \to [0, \infty]$ be measurable. The function

$$\mu\{f > t\} = \mu(\{x \in X : f(x) > t\})$$

is called the distribution function of f. $\mu\{f > t\}$ is a decreasing function and hence it is Borel measurable.

Theorem 84. *Let $\varphi : [0, \infty] \to [0, \infty]$ be monotone increasing, absolutely continuous on $[0, T]$ for any $T < \infty$ and $\varphi(0) = 0$, $\lim_{t\to\infty} \varphi(t) = \varphi(\infty)$. Then*

$$\int_X \varphi \circ f d\mu = \int_0^\infty \mu\{f > t\} \varphi'(t) dt. \tag{11.1}$$

Proof. Let

$$E = \{(x, t) \in X \times [0, \infty) : f(x) > t\}.$$

Then E is measurable. For any $t \in [0, \infty)$, we define

$$E^t = \{x \in X : (x, t) \in E\}.$$

Then

$$\mu\{f > t\} = \mu\left(E^t\right) = \int_X \chi_E(x, t) d\mu(x).$$

So we have by applying Fubini's theorem

$$\begin{aligned}
&\int_0^\infty \mu\{f > t\} \varphi'(t) dt \\
&= \int_0^\infty \left[\int_X \chi_E(x, t) d\mu(x)\right] \varphi'(t) dt \\
&= \int_X d\mu(x) \int_0^\infty \chi_E(x, t) \varphi'(t) dt.
\end{aligned}$$

Since

$$\int_0^\infty \chi_E(x, t) \varphi'(t) dt = \int_0^{f(x)} \varphi'(t) dt = \varphi(f(x)),$$

we have (11.1). □

A special case is when $\varphi(t) = t$:

Corollary 11. *Let $f : X \to [0, \infty]$ be measurable. Then*

$$\int_X f d\mu = \int_0^\infty \mu\{f > t\} dt.$$

This corollary implies that any Lebesgue integral can be replaced by a Lebesgue integral on $\mathbb{R}$ with Lebesgue measure.

Recall that for any locally integrable function $f : \mathbb{R}^k \to \mathbb{R}$, the maximal function of f is defined by

$$(Mf)(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f| dm.$$

And we show that $f \in L^1\left(\mathbb{R}^k\right)$ implies that Mf belongs to weak L^1, i.e., $tm(\{|f| > t\})$ is a bounded function of t on $(0, \infty)$. On the other hand, if $f \in L^\infty\left(\mathbb{R}^k\right)$, it is easy to see $Mf \in L^\infty\left(\mathbb{R}^k\right)$. In general, we have

THEOREM 85 (Hardy-Littlewood). *If* $1 < p < \infty$ *and* $f \in L^p\left(\mathbb{R}^k\right)$, *then* $Mf \in L^p\left(\mathbb{R}^k\right)$.

PROOF. Without loss of generality, we assume $f \geq 0$. Let $c \in (0,1)$ which will be chosen later. For each $t \in (0,\infty)$, we define

$$g_t(x) = f(x)\,\chi_{\{f(x)>ct\}},\, h_t(x) = f(x) - g_t(x) = f(x)\,\chi_{\{f(x)\leq ct\}}.$$

Then we have $f = g_t + h_t$ and $h_t \leq ct$. Hence $g_t \in L^1$ and $h_t \in L^\infty$. Since

$$Mf \leq Mg_t + Mh_t \leq Mg_t + ct,$$

we have

$$\begin{aligned} m\left(\{Mf > t\}\right) &\leq m\left(\{Mg_t > (1-c)\,t\}\right) \\ &\leq \frac{3^k}{(1-c)\,t}\left\|g_t\right\|_1 = \frac{3^k}{(1-c)\,t}\int_{\{f(x)>ct\}} f(x)\,dx. \end{aligned}$$

Now

$$\begin{aligned} \int_{\mathbb{R}^k} (Mf)^p\,dm &= p\int_0^\infty m\left(\{Mf > t\}\right) t^{p-1}dt \\ &\leq \frac{3^k p}{(1-c)}\int_0^\infty \left(\int_{\{f(x)>ct\}} f(x)\,dx\right) t^{p-2}dt \\ &= \frac{3^k p}{(1-c)}\int_{\mathbb{R}^k} f(x)\,dx \int_0^{\frac{f(x)}{c}} t^{p-2}dt \\ &= \frac{3^k p}{(1-c)\,(p-1)\,c^{p-1}}\int_{\mathbb{R}^k} f^p(x)\,dx. \end{aligned}$$

When $c = \frac{p-1}{p} = \frac{1}{q}$, the constant in the right hand side achieves minimum, and we have

$$\left\|Mf\right\|_p \leq C_p\left\|f\right\|_p$$

where

$$C_p = \left(\frac{3^k p}{(1-c)\,(p-1)\,c^{p-1}}\right)^{\frac{1}{p}} = \left(3^k p q^p\right)^{\frac{1}{p}}.$$

We remark that $C_p \to 1$ as $p \to \infty$ and $C_p \to \infty$ as $p \to 1$. $\square$

REMARK 29. *In general,* $f \in L^1\left(\mathbb{R}^k\right)$ *doesn't imply* $Mf \in L^1\left(\mathbb{R}^k\right)$. *For example,* $f = \chi_{B_1(0)}$, *then* Mf *decays like* $\frac{1}{r^k}$ *at* ∞, *hence* $Mf \notin L^1\left(\mathbb{R}^k\right)$.

CHAPTER 5

Hausdorff Measure

1. Introduction

Let X be a metric space. Let $s \geq 0$. Given $\varepsilon > 0$, we define for any $E \subset X$ a set function

$$\mathcal{H}_{\varepsilon}^{s}(E) = \inf \left\{ \sum_{i=1}^{\infty} \omega_s \left(\frac{\operatorname{diam} A_i}{2} \right)^{s} : E \subset \bigcup_{i=1}^{\infty} A_i \text{ and } \operatorname{diam} A_i < \varepsilon \right\}.$$

Here

$$\omega_s = \frac{\pi^{\frac{s}{2}}}{\Gamma\left(1 + \frac{s}{2}\right)}$$

is the volume of s-dimensional unit ball where we recall the Gamma function is defined by

$$\Gamma(s) = \int_0^{\infty} t^{s-1} e^{-t} dt.$$

Since for each $E \subset X$, $H_{\varepsilon}^{s}(E)$ is monotone decreasing in ε, we can define a new set function

$$\mathcal{H}^{s}(E) = \lim_{\varepsilon \to 0} \mathcal{H}_{\varepsilon}^{s}(E).$$

Theorem 86. *$\mathcal{H}^s$ is a metric outer measure on X.*

Proof. Step 1: $\mathcal{H}^s$ is an outer measure. It is easy to see $\mathcal{H}^s(E) \in [0, \infty]$ for any $E \subset X$, $\mathcal{H}^s(\emptyset) = 0$ and monotonicity property holds. So we only need to check countable subadditivity. Let $E \subset \bigcup\limits_{n=1}^{\infty} E_n$. We can assume

$$\sum_{n=1}^{\infty} \mathcal{H}^{s}(E_n) < \infty.$$

For each $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mathcal{H}_{\varepsilon}^{s}(E_n) \leq \sum_{n=1}^{\infty} \mathcal{H}^{s}(E_n) < \infty.$$

Given $\delta > 0$, for each n, there exists a covering of E_n, $E_n \subset \bigcup\limits_{i=1}^{\infty} A_{n,i}$ with $\operatorname{diam} A_{n,i} < \varepsilon$, such that

$$\mathcal{H}_{\varepsilon}^{s}(E_n) \geq \sum_{i=1}^{\infty} \omega_s \left(\frac{\operatorname{diam} A_{n,i}}{2} \right)^{s} - \frac{\delta}{2^n}.$$

Then we have $E \subset \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} A_{n,i}$, and hence

$$\mathcal{H}_{\varepsilon}^{s}(E) \leq \sum_{n=1}^{\infty} \sum_{i=1}^{\infty} \omega_s \left(\frac{\operatorname{diam} A_{n,i}}{2} \right)^s \leq \sum_{n=1}^{\infty} \mathcal{H}_{\varepsilon}^{s}(E_n) + \delta \leq \sum_{n=1}^{\infty} \mathcal{H}^{s}(E_n) + \delta.$$

Letting $\delta \to 0$, we have

$$\mathcal{H}_{\varepsilon}^{s}(E) \leq \sum_{n=1}^{\infty} \mathcal{H}^{s}(E_n).$$

Hence,

$$\mathcal{H}^{s}(E) = \lim_{\varepsilon \to 0} \mathcal{H}_{\varepsilon}^{s}(E) \leq \sum_{n=1}^{\infty} \mathcal{H}^{s}(E_n).$$

Step 2: Let $E, F \subset X$ with $\operatorname{dist}(E, F) = \delta > 0$. Then for any $0 < \varepsilon < \delta$, since any set A with $\operatorname{diam} A < \varepsilon$ can't intersect both E and F, we have

$$\mathcal{H}_{\varepsilon}^{s}(E \cup F) = \mathcal{H}_{\varepsilon}^{s}(E) + \mathcal{H}_{\varepsilon}^{s}(F).$$

Letting $\varepsilon \to 0$, we obtain

$$\mathcal{H}^{s}(E \cup F) = \mathcal{H}^{s}(E) + \mathcal{H}^{s}(F).$$

□

DEFINITION 42. *$\mathcal{H}^s$ is called the s-dimensional Hausdorff measure on X.*

PROPOSITION 11. *(i). $\mathcal{H}^0$ is the counting measure.*
(ii). If $s < t$, then $\mathcal{H}^s(E) < \infty$ implies $\mathcal{H}^t(E) = 0$.
(iii). If $s < t$, then $\mathcal{H}^t(E) > 0$ implies $\mathcal{H}^s(E) = \infty$.

PROOF. (i). Trivial.
(ii) and (iii) follows from the observation that for any $\varepsilon > 0$,

$$\mathcal{H}_{\varepsilon}^{t}(E) \leq \frac{\omega_t}{\omega_s} \left(\frac{\varepsilon}{2} \right)^{t-s} \mathcal{H}_{\varepsilon}^{s}(E).$$

□

DEFINITION 43. *For any nonempty set E, the Hausdorff dimension of E is defined by*

$$\dim_{\mathcal{H}} E = \sup \{ s : \mathcal{H}^s(E) = 0 \} = \inf \{ s : \mathcal{H}^s(E) = \infty \}.$$

THEOREM 87. *Let $E \subset \mathbb{R}^n$.*
(i). $\mathcal{H}^s$ is rotation and translation invariant;
(ii). $\mathcal{H}^s(\lambda E) = \lambda^s \mathcal{H}^s(E)$ for any $\lambda > 0$;
(iii). $\mathcal{H}^s(E) = 0$ for any $s > n$.

REMARK 30. *If we use only balls to cover the set E in the definition of Hausdorff measure, the resulting measure is called s-dimensional spherical measure $\mathcal{S}^s$. It is easy to check $\mathcal{H}^s \leq \mathcal{S}^s \leq 2^s \mathcal{S}^s$. However, $\mathcal{H}^s \neq \mathcal{S}^s$ in general.*

2. Isodiametric Inequality and $\mathcal{H}^n = L_n$

THEOREM 88 (Isodiametric Inequality). *For any* $A \subset \mathbb{R}^n$,

$$L_n^*(A) \leq \omega_n \left(\frac{\operatorname{diam} A}{2}\right)^n.$$

PROOF. We will use Steiner symmetrization. □

LEMMA 8. *Let* $Q = [0,1]^n$. *For any* $\delta > 0$, *there exists disjoint balls* $\{B_{r_i}(x_i)\}_{i=1}^{\infty}$ *such that* $B_{r_i}(x_i) \subset Q$, $r_i < \delta$ *and*

$$L_n\left(Q \backslash \bigcup_{i=1}^{\infty} B_{r_i}(x_i)\right) = 0. \tag{2.1}$$

PROOF. Choose $B_{r_1}(x_1) \subset Q$ such that $r_1 = \min\left\{\frac{\delta}{2}, \frac{1}{2}\right\}$. Assuming $\{B_{r_i}(x_i)\}_{i=1}^{k}$ are chosen, we define

$$r_{k+1}^* = \sup\left\{r : B_r(x) \subset Q \backslash \bigcup_{i=1}^{k} B_{r_i}(x_i) \text{ for some } x \in Q\right\}.$$

And we choose $B_{r_{k+1}}(x_{k+1}) \subset Q \backslash \bigcup_{i=1}^{k} B_{r_i}(x_i)$ such that $r_{k+1} = \min\left\{\frac{\delta}{2}, \frac{r_{k+1}^*}{2}\right\}$. Then since $\bigcup_{i=1}^{\infty} B_{r_i}(x_i) \subset Q$, we have

$$\sum_{i=1}^{\infty} \omega_n r_i^n \leq 1.$$

Especially, we have $\lim_{i\to\infty} r_i = 0$, so there exists $K > 0$ such that $r_k = \frac{r_k^*}{2}$ for $k \geq K$. From the definition of r_k^*, we have for any $k \geq K$,

$$Q \subset \bigcup_{i=1}^{k} B_{r_i + 2r_{k+1}^*}(x_i).$$

Hence,

$$\sum_{i=1}^{k} \omega_n \left(r_i + 2r_{k+1}^*\right)^n \geq 1.$$

Now for any $\varepsilon > 0$, we choose $K_\varepsilon > K$ so that

$$\sum_{i=K_\varepsilon+1}^{\infty} \omega_n r_i^n < \varepsilon.$$

Hence, for any $k > K_\varepsilon$, we have

$$\begin{aligned}
1 &\leq \sum_{i=1}^{k} \omega_n \left(r_i + 2r_{k+1}^*\right)^n = \sum_{i=1}^{K_\varepsilon} \omega_n \left(r_i + 2r_{k+1}^*\right)^n + \sum_{i=K_\varepsilon+1}^{k} \omega_n \left(r_i + 2r_{k+1}^*\right)^n \\
&\leq \sum_{i=1}^{K_\varepsilon} \omega_n \left(r_i + 2r_{k+1}^*\right)^n + \sum_{i=K_\varepsilon+1}^{k} \omega_n \left(r_i + 2r_{k+1}^*\right)^n \\
&\leq \sum_{i=1}^{K_\varepsilon} \omega_n \left(r_i + 2r_{k+1}^*\right)^n + \sum_{i=K_\varepsilon+1}^{k} \omega_n (5r_i)^n \leq \sum_{i=1}^{K_\varepsilon} \omega_n \left(r_i + 2r_{k+1}^*\right)^n + 5^n\varepsilon.
\end{aligned}$$

Let $k \to \infty$, we have

$$\sum_{i=1}^{K_\varepsilon} \omega_n r_i^n + 5^n \varepsilon \geq 1,$$

and hence

$$\sum_{i=1}^{\infty} \omega_n r_i^n \geq \sum_{i=1}^{K_\varepsilon} \omega_n r_i^n \geq 1 - 5^n \varepsilon.$$

Since ε can be arbitrarily small,

$$\sum_{i=1}^{\infty} \omega_n r_i^n \geq 1.$$

Hence

$$\sum_{i=1}^{\infty} \omega_n r_i^n = 1$$

and (2.1) holds. □

THEOREM 89. *$\mathcal{H}^n = L_n^*$ on $\mathbb{R}^n$.*

PROOF. Step 1. For any $A \subset \mathbb{R}^n$, $L_n(A) \leq \mathcal{H}^n(A)$. Fix $\varepsilon > 0$ and let $\{A_i\}_{i=1}^{\infty}$ be covering of A such that $\operatorname{diam} A_i < \varepsilon$. Applying isodiametric inequality,

$$\sum_{i=1}^{\infty} \omega_n \left(\frac{\operatorname{diam} A_i}{2} \right)^n \geq \sum_{i=1}^{\infty} L_n^*(A_i) \geq L_n^*(A).$$

Taking infimum over all such covering, we obtain

$$\mathcal{H}_\varepsilon^n(A) \geq L_n^*(A).$$

Letting $\varepsilon \to 0$, we have $\mathcal{H}^n(A) \geq L_n^*(A)$.

Step 2. For any $A \subset \mathbb{R}^n$, $\mathcal{H}^n(A) \leq c(n) L_n^*(A)$ for some constant $c(n)$. For any $\varepsilon > 0$, we have

$$\begin{aligned}
\mathcal{H}_\varepsilon^n(A) &\leq \inf \left\{ \sum_{i=1}^{\infty} \omega_n \left(\frac{\operatorname{diam} Q_i}{2} \right)^n : A \subset \bigcup_{i=1}^{\infty} Q_i \text{ and } \operatorname{diam} Q_i < \varepsilon \right\} \\
&\leq c(n) \inf \left\{ \sum_{i=1}^{\infty} L_n(Q_i) : A \subset \bigcup_{i=1}^{\infty} Q_i \text{ and } \operatorname{diam} Q_i < \varepsilon \right\} \\
&= c(n) L_n^*(A).
\end{aligned}$$

Here the last identity is obvious when A is open, and it is true for general A since L_n is regular.

Step 3. $\mathcal{H}^n(A) \leq L_n^*(A)$. We first assume $A = Q$ is a cube. For any $\delta > 0$, from Lemma 8, there exists disjoint balls $\{B_{r_i}(x_i)\}_{i=1}^{\infty}$ such that $B_{r_i}(x_i) \subset Q$, $r_i < \frac{\delta}{2}$ and

$$L_n \left(Q \backslash \bigcup_{i=1}^{\infty} B_{r_i}(x_i) \right) = 0.$$

Hence from step 2,

$$\mathcal{H}^n \left(Q \backslash \bigcup_{i=1}^{\infty} B_{r_i}(x_i) \right) = 0.$$

Now,

$$\mathcal{H}_\delta^n(Q) \leq \sum_{i=1}^{\infty} \mathcal{H}_\delta^n(B_{r_i}(x_i)) \leq \sum_{i=1}^{\infty} \omega_n r_i^n = L_n(Q).$$

Letting $\delta \to 0$, $\mathcal{H}^n(Q) \leq L_n(Q)$. For open A, $\mathcal{H}^n(A) \leq L_n(A)$ follows from cube decomposition of A. And for general A, $\mathcal{H}^n(A) \leq L_n^*(A)$ again follows from the fact that L_n is regular. □

Similarly, we can prove

THEOREM 90. *$\mathcal{S}^n = L_n^*$ on $\mathbb{R}^n$.*