Functions of Real Variables

Leopold Wuhan Zhou
School of EECS, Peking University

November 28, 2025

Abstract

This note is based on Functions of Real Variables, taught by Baoping Liu in Fall 2025 at Peking University and also refers to Theory of Functions of Real Variables by Minqiang Zhou.

Contents

1 Sets and Point Sets

1.1 Introduction

Definition 1.1. Given f bounded on $[a, b]$, consider a partition $\Delta : a = x_0 < x_1 < \cdots < x_n = b$. Define $M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$, $m_i := \inf\{f(x) : x \in [x_{i-1}, x_i]\}$ and $\Delta x_i = x_i - x_{i-1}$. The upper sum and lower sum of f with respect to Δ are defined as

$$\overline{S}_\Delta(f) := \sum_{i=1}^{n} M_i \Delta x_i, \quad \underline{S}_\Delta(f) := \sum_{i=1}^{n} m_i \Delta x_i.$$

Consider Darboux integrals

$$\overline{\int_a^b} f := \inf \overline{S}_\Delta, \quad \underline{\int_a^b} f := \sup \underline{S}_\Delta.$$

If they are equal, then f is **Riemann integrable** on $[a, b]$, and the common value is called the **Riemann integral** of f on $[a, b]$, denoted by

$$\int_a^b f(x)\mathrm{d}x.$$

In fact $\lim_{\Delta \to 0} \sum f(\xi_i)\Delta x_i, \xi_i \in [x_{i-1}, x_i]$ exists if $f \in R[a, b]$. And we have some basic facts:

1. $R[a, b]$ is a vector space.
2. $C[a, b] = \{f \text{ is continuous on } [a, b]\} \subset R[a, b]$
3. If $f \in R[a, b]$, then f is not far from a continous function in the following sense. Denote $\Omega(f) = \{x_0 \in [a, b] : f \text{ is discontinuous at } x_0\}$. $\forall \varepsilon > 0, \exists I = \bigcup I_j$ where I_j are disjoint intervals, such that $\sum |I_j| < \varepsilon$ and $\Omega(f) \subseteq I$.

Example 1.2. We give two functions that are Riemann integrable.

$$f(x) = \begin{cases} 1, & x = \frac{1}{n} \\ 0, & \text{otherwise} \end{cases}, \quad f(x) = \begin{cases} 0, & x \notin \mathbb{Q}, x = 0 \\ \frac{1}{q}, & x = \frac{p}{q} \end{cases}$$

However, Riemann integral has some problems:

(a) Some "nice" functions are not Riemann integrable. For example, Dirichlet function $f(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}$ is not Riemann integrable on any interval $[0, 1]$.

(b) Riemann integrability is not preserved by pointwise limits. For example, consider $f_n(x) = \begin{cases} 1, & x \in \{x_1, \cdots, x_n\} \\ 0, & x \notin \{x_1, \cdots, x_n\} \end{cases}$ where $\mathbb{Q} \cap [0, 1] = \{x_1, x_2, \cdots\}$. Then $f_n \to f$ pointwise where f is the Dirichlet function, but each f_n is Riemann integrable with integral 0.

(c) If we equip $R[0,1]$ with metric or distance $d(f,g) = \int_0^1 |f(x) - g(x)|dx$, then $(R[0,1], d)$ behaves like a metric space, but it is not complete. For example, $f_n = \frac{1}{\sqrt{x}} X_{[\frac{1}{n},1]}$ and $f = \begin{cases} \frac{1}{\sqrt{x}}, & x \in [0,1] \\ 0, & x \notin [0,1] \end{cases}$. Then $f_n \in R[0,1]$ and $d(f_n, f) \to 0$, but $f \notin R[0,1]$.

(d) Other problems. Let $f \in R[-\pi, \pi]$. Then its Fourier series $\sum a_n e^{inx}$ and we have Parseval's identity $\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 dx = \sum |a_n|^2$. However, it may happen that $\sum a_n e^{inx}$ diverges at some point. For example, $f(x) = x$ on $[-\pi, \pi]$ and extended periodically. Then its Fourier series diverges at $x = \pi$.

(e) How to measure the length, area and volume of a geometric object in $\mathbb{R}^n$? "Measure" means a function $m : \mathcal{P}(\mathbb{R}) = \{\text{ all subsets of } \mathbb{R}\} \mapsto [0, +\infty]$ satisfying: (1) If $I = [a,b]$, then $m(I) = b-a$. (2) If $E_1 \cap E_2 = \varnothing$, then $m(E_1 \cup E_2) = m(E_1) + m(E_2)$ (3) $m(E + h) = m(E)$ for each $h \in \mathbb{R}$. Unfortunately, such m does not exist. How to fix it? We can change the domain of m from $\mathcal{P}(\mathbb{R})$ to a smaller σ-algebra $\mathcal{M} \subset \mathcal{P}(\mathbb{R})$.

(f) Integration. Let $m = \inf_{[a,b]} f, M = \sup_{[a,b]} f$. Consider a partition of $[m, M]$: $m = y_0 < y_1 < \cdots < y_n = M$. Define $E_i = \{x \in [a,b] : y_{i-1} \leqslant f(x) < y_i\}$. Then $\sum y_i m(E_i) \to \int_a^b f(x)dx$. We'll discuss this later.

1.2 Sets

Definition 1.3. Let $\{A_k\}$ be a family of sets. If

$$A_1 \supset A_2 \supset \cdots \supset A_k \supset \cdots,$$

then we say $\{A_k\}$ is a **decreasing family of sets** and $\bigcap_{k=1}^{\infty} A_k$ is the **limit** of $\{A_k\}$, denoted by $\lim_{k\to\infty} A_k$. If

$$A_1 \subset A_2 \subset \cdots \subset A_k \subset \cdots,$$

then we say $\{A_k\}$ is an **increasing family of sets** and $\bigcup_{k=1}^{\infty} A_k$ is the **limit** of $\{A_k\}$, denoted by $\lim_{k\to\infty} A_k$.

Definition 1.4. Let $\{A_k\}$ be a family of sets. The **limit superior** and **limit inferior** of $\{A_k\}$ are defined as

$$\overline{\lim_{k\to\infty}} A_k := \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k, \quad \underline{\lim_{k\to\infty}} A_k := \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} A_k.$$

If they are equal, then we say the limit of $\{A_k\}$ exists, denoted by $\lim_{k\to\infty} A_k$.

Lemma 1.5. *1.* $\overline{\lim}_{k\to\infty} A_k := \{x : \forall j \in \mathbb{N}, \exists k \geqslant j, x \in A_k\}$, *i.e. contains those* x *that belong to infinitely many* A_k.

2. $\underline{\lim}_{k\to\infty} A_k := \{x : \exists j \in \mathbb{N}, \forall k \geqslant j, x \in A_k\}$, *i.e. contains those* x *that belong to all but finitely many* A_k.

1.3 Mappings

We notice some facts of mappings.

1. $f(\bigcup A_\alpha) = \bigcup f(A_\alpha)$, $f(\bigcap A_\alpha) \subseteq \bigcap f(A_\alpha)$. For example $f(x) = \sin x, A_n = [2n\pi - \pi, 2n\pi + \pi]$.
2. $f^{-1}(\bigcup B_\alpha) = \bigcup f^{-1}(B_\alpha)$, $f^{-1}(\bigcap B_\alpha) \supseteq \bigcap f^{-1}(B_\alpha)$.
3. $f^{-1}(B^c) = f^{-1}(B)^c$.

1.4 Orders and Axioms of Choice

The followings are equivalent:

1. (Hausdorff Maximal Principle) Every partially ordered set has a maximal linearly ordered subset.
2. (Zorn's Lemma) Every partially ordered set in which every linearly ordered subset has an upper bound has a maximal element.
3. (Well Ordering Principle) Every nonempty set can be well-ordered.
4. (Axiom of Choice) If $\{A_i\}_{i\in I}$ is a family of non-empty sets, then there exists a choice function $f : I \mapsto \bigcup_{i\in I} A_i$ such that $f(i) \in A_i$ for each $i \in I$.

Example 1.6. Let $X \times Y = \{(x, y) : x \in X, y \in Y\}$. Then $\prod_\alpha X_\alpha$ is defined as a map $f : \Lambda \to \bigcup_\alpha X_\alpha$ such that $f(\alpha) \in X_\alpha$.

Lemma 1.7 (fixed point theorem form monotonic mapping)**.** *X is a nonempty set and $f : \mathscr{P}(X) \to \mathscr{P}(X)$ is a monotonic mapping, i.e. $A \subseteq B \Rightarrow f(A) \subseteq f(B)$. There exists $T \subseteq \mathscr{P}(X)$ such that $f(T) = T$.*

Proof. Let $S = \{A : A \in \mathscr{P}(X), A \subseteq f(A)\}$ and $T = \bigcup_{A\in S} A$. Clearly, $T \subseteq f(T)$ and by $f(T) \subseteq f(f(T))$, we have $f(T) \in S$ and $f(T) \subseteq T$. Hence $f(T) = T$. □

1.5 Cardinality

Proposition 1.8. $|A| \leqslant |B| \iff |B| \geqslant |A|$.

Proof. $\Rightarrow$: Let $f : A \to B$ be an injection. Take any $a \in A$ fixed, denote $g(b) = \begin{cases} f^{-1}(b), & b \in f(A) \\ a, & b \notin f(A) \end{cases}$. Then $g : B \to A$ is a surjection.

$\Leftarrow$: Let $g : B \to A$ be a surjection, i.e. $g^{-1}(a) = \{b \in B | g(b) = a\} \neq \varnothing$. Also $g^{-1}(a)$ are disjoint for different $a \in A$. Let $f(a)$ be an element in $g^{-1}(a)$. Then $f : A \to B$ is an injection. □

Proposition 1.9. *$\forall A, B$, either $|A| \leqslant |B|$ or $|B| \leqslant |A|$.*

Proof. WLOG, assume $A, B \neq \varnothing$. Consider $f : A_0 \to B$ an injection where $A_0 \subseteq A$, then f corresponds to a set $\Omega_f = \{(a, f(a)), a \in A\} \subseteq A \times B$. Conversely, consider Ω a subset of $A \times B$ such that Consider $\mathscr{J} = \{$collection of all injections $f :$ subset of $A \to B\}$ and $\sum = \{\Omega_f | f \in \mathscr{J}\} \subseteq \mathscr{P}(A \times B)$.

Impose partial order on $\sum$ defined as set inclusion. Every linearly ordered subset of $\sum$, denoted by $\widetilde{\sum}$ has an upper bound. Hence by Zorn's lemma, there exists a maximal element Ω_f. If $A_0 \neq A$ and $B_0 \neq B$, we get contradiction with maximality. Hence $A_0 = A$ or $B_0 = B$ and either $|A| \leqslant |B|$ or $|B| \leqslant |A|$. □

Theorem 1.10 (Schröder-Bernstein)**.** *If* $|A| \leqslant |B|$ *and* $|B| \leqslant |A|$*, then* $|A| = |B|$*.*

Proof. Let $f : A \to B$ and $g : B \to A$ be injections. $\forall x \in A$, if $x \in g(B)$, then we can define $g^{-1}(x) \in B$; if $g^{-1}(x) \in f(A)$, we can define $f^{-1}(g^{-1}(x))$ and so on.

We break A into 3 disjoint sets: $A_\infty = \{x \in A |$the algorithm goes for infinite steps$\}$, $A_A = \{x \in A |$ the algorithm stops at an element $\in A \backslash g(B)\}$ and $A_B = \{x \in A |$ the algorithm stops at an element $\in B \backslash f(A)\}$. Similarly, we can break B into 3 disjoint sets: $B = B_\infty \cup B_A \cup B_B$. We can verify that there are bijections between A_∞ and B_∞, A_A and B_A, A_B and B_B. Hence there is a bijection between A and B. Define

$$h(x) = \begin{cases} f(x), & x \in A_\infty \cup A_A \\ g^{-1}(x), & x \in A_B \end{cases}.$$

Then h is a bijection. □

Corollary 1.11. $|A| \leqslant |B| \leqslant |C|$ *and* $|A| = |C| \Rightarrow |A| = |B| = |C|$*. For example,* $|(-1, 1)| = |(-1, 1]| = |[-1, 1]| = |\mathbb{R}|$*.*

Theorem 1.12 (No Maximal Cardinality)**.** $\forall A \neq \varnothing$*,* $|A| < |\mathscr{P}(A)|$*.*

Proof. Clearly, $|A| \leqslant |\mathscr{P}(A)|$. Suppose $|A| = |\mathscr{P}(A)|$, then there exists a bijection $g : A \to \mathscr{P}(A)$. Consider $B = \{x \in A : x \notin g(x)\}$ and denote $b \in g^{-1}(B)$. Both $b \in g(b) = B$ and $b \notin g(b)$ lead to contradiction. □

In practice, denote $|\mathbb{N}| = \aleph_0$ and $|\mathbb{R}| = c$.

Definition 1.13. A is **denumerable** or **countably infinite** if $A \sim \mathbb{N}$. A is **countable** if $|A| \leqslant |\mathbb{N}|$.

Theorem 1.14. *Every infinite set A must contain a denumerable subset, i.e.* $|\mathbb{N}| \leqslant |A|$*.*

Proof. Take $a_1 \in A$, $a_2 \in A \backslash \{a_1\}$, $\cdots$. Then $E = \{a_1, a_2, a_3, \cdots\} \sim \mathbb{N}$. □

Proposition 1.15. *Assume A is finite and B is denumerable. Then* $A \cup B$ *is denumerable.*

Proof. Let $A = \{a_1, a_2, \cdots, a_n\}$ and $B = \{b_1, b_2, \cdots\}$. If $A \cap B = \varnothing$, $A \cup B = \{a_1, a_2, \cdots, a_n, b_1, b_2, \cdots\}$ is denumerable. Otherwise, $A \cup B = (A \backslash B) \cup B$ repeats the above argument. □

Proposition 1.16. *Assume* A_n *is denumerable (countable) for each* $n \in \mathbb{N}$*. Then* $\bigcup_{n=1}^\infty A_n$ *is denumerable (countable).*

Proof. WLOG, assume A_n are disjoint, otherwise consider $\widetilde{A_n} = A_n \backslash \bigcup_{k=1}^{n-1} A_k$. Let $A_n = \{a_{n1}, a_{n2}, \cdots\}$. Then $\bigcup_{n=1}^\infty A_n = \{a_{11}, a_{21}, a_{12}, a_{31}, a_{22}, a_{13}, \cdots\}$ is denumerable. □

Corollary 1.17. *X, Y are countable $\Rightarrow X \times Y$ is countable.*

Proof. $X \times Y = \bigcup_{n,m}(x_n, y_m) = \bigcup_m \{(x_n, y_m) | x_n \in X\}$ is a countable union of countable sets. ☐

Corollary 1.18. *$X_1, X_2, \cdots, X_n$ are countable $\Rightarrow X_1 \times X_2 \times \cdots \times X_n$ is countable.*

Caution: Corollary 1.18 does not hold for infinite products. For example, $S = \prod_{n=1}^{\infty}\{0,1\} = \{(a_1, a_2, \cdots, a_n, \cdots)\}$ is uncountable. In fact, if S is countable, then we can list all elements of S as $s_1, s_2, \cdots$. Consider $s^* = (b_1, b_2, \cdots)$ where $b_n = 1$ if the n-th element of s_n is 0 and $b_n = 0$ otherwise. Then $t \neq s_n$ for each n, contradicting the assumption.

Corollary 1.19. *$\mathbb{Q}$ is denumerable.*

Proof. $\forall x \in \mathbb{Q}, x = \frac{p}{q}$ where $(p,q) = 1, q \neq 0$. Then $|\mathbb{Z}| \leqslant |\mathbb{Q}| \leqslant |\mathbb{Z} \times \mathbb{Z}|$. By Corollary 1.11, $|\mathbb{Q}| = |\mathbb{Z}|$. ☐

Example 1.20. $\mathscr{A} = \{$disjoint open intervals of $\mathbb{R}\}$. $|\mathscr{A}| \leqslant |\mathbb{N}|$.

Proof. For every open interval I, take $r_I \in \mathbb{Q} \cap I$. Then $f : \mathscr{A} \to \mathbb{Q}, f(I) = r_I$ is an injection. Hence $|\mathscr{A}| \leqslant |\mathbb{Q}| = |\mathbb{N}|$. ☐

Example 1.21. $f : \mathbb{R} \to \mathbb{R}$ is monotonic increasing, then $|\{$ the discontinuous point of $f\}| \leqslant |\mathbb{N}|$.

Proof. Define g, for every discontinuous point x_0 of f, mapping x_0 to $(f(x_0-), f(x_0+)) = I_{x_0}$. Since f is increasing, I_{x_0} are disjoint open intervals and by Example 1.20, $|\{$ the discontinuous point of $f\}| \leqslant |\mathbb{N}|$. ☐

Lemma 1.22. $|\mathbb{N}| < |\mathbb{R}| = |\mathscr{P}(\mathbb{N})|$.

Proof. Clearly, $|\mathbb{N}| \leqslant |\mathbb{R}|$. Consider $x \in (0,1]$, denote $x = \sum_{n=1}^{\infty} \frac{a_n}{2^n}$ where $a_n = 0$ or 1 and we require there are infinite many $a_n = 1$. Consider a subset of $\mathbb{N}$, $\Omega_x = \{n_1, n_2, \cdots, n_k, \cdots | a_j = 0, \text{ if } j = n_k \text{ and } a_j \neq 0, \text{ if } j \neq n_k\}$. Then we have an injection $x \in (0,1] \mapsto \Omega_x \in \mathbb{N}$, which implies $|\mathbb{R}| = |(0,1]| \leqslant |\mathscr{P}(\mathbb{N})|$.

Conversely, let $x = \sum_{n \geqslant 1} \frac{a_n}{3^n}$. For every $\Omega \subseteq \mathbb{N}$, there is an injection $\Omega \to x \in (0,1]$ similarly. Hence $|\mathscr{P}(\mathbb{N})| = |\mathbb{R}|$. By Theorem 1.12, $|\mathbb{N}| < |\mathscr{P}(\mathbb{N})| = |\mathbb{R}|$. ☐

Theorem 1.23 (Continuum Hypothesis)**.** *There is no set A such that $|\mathbb{N}| < |A| < |\mathbb{R}|$.*

Gödel showed that the Continuum Hypothesis is compatible with ZFC. And Cohen showed that the Continuum Hypothesis can not be deduced from ZFC.

1.6 Topology of $\mathbb{R}^n$ and Metric Space

Example 1.24. $d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$, $d_\infty(x,y) = \sup_{1 \leqslant i \leqslant n} |x_i - y_i|$ and $d(x,y) = \sum_{i=1}^{n} |x_i - y_i|$ are metrics on $\mathbb{R}^n$.

Here are some basic concepts in metric space (X, d).

1. $d(E,F) = \inf\{d(x,y) : x \in E, y \in F\}$.
2. $\text{diam}E = \sup\{d(x,y) : x, y \in E\}$ and E is bounded if $\text{diam}E < +\infty$.

3. Open ball centered at x with radius r: $B_r(x) = \{y \in X | d(x,y) < r\}$.
4. E is open if $\forall x \in E, \exists r > 0, B_r(x) \subseteq E$.
5. x is a limit point of E, denoted by $x \in E'$, if $\forall r > 0, B_r(x) \cap E \backslash \{x\} \neq \varnothing$.
6. $\overline{E} = \cap_{F \supseteq E, F \text{ is closed}} F$.

Theorem 1.25 (Cantor's Intersection Theorem). *Let $F_1 \supset F_2 \supset \cdots$ be closed and bounded. Then $\bigcap_{n=1}^{\infty} F_n \neq \varnothing$.*

Caution: Boundedness is necessary. For example, $F_n = [n, +\infty)$.

Here we give a characterization of open sets in $\mathbb{R}^n$.

Lemma 1.26. *$E \subseteq \mathbb{R}^1$ is open. Then E can be written as disjoint union of open intervals.*

Proof. For each $x \in E$, define $\beta = \sup\{b | (x,b) \subseteq E\}$ and $\alpha = \inf\{a | (a,x) \subseteq E\}$. Let $I_x = (\alpha, \beta) \subseteq E$. Then $E = \bigcup_{x \in E} I_x$. Claim that $\forall x \neq y$, either $I_x \cap I_y = \varnothing$ or $I_x = I_y$. Suppose $I_x \cap I_y \neq \varnothing$ and $z \in I_x \cap I_y$. Then $I_z = I_x = I_y$. Hence E is disjoint union of open intervals. □

Lemma 1.27. *$E \subseteq \mathbb{R}^n$ is open. Then E can be written as disjoint union of half open half closed blocks, i.e. $(a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n]$.*

Proof. Let $\Gamma_0 = \{$half open half closed unit cubes with integer vertices$\}$ and $\Gamma_1 = \{$ refinement of cubes in Γ_0 into half open half closed cubes with vertices coordinates $\frac{k}{2}, k \in \mathbb{Z}\}, \cdots, \Gamma_n = \{$ refinement of cubes in Γ_{n-1} into half open half closed cubes with vertices coordinates $\frac{k}{2^n}, k \in \mathbb{Z}\}$ and so on.

Consider $H_0 = \{I \in \Gamma_0, I \subseteq E\}, H_1 = \{I \in \Gamma_1, I \subseteq G \backslash \bigcup_{I \in H_0} I\}, \cdots, H_n = \{I \in \Gamma_n, I \subseteq G \backslash \bigcup_{j=0}^{n-1} \bigcup_{I \in H_j} I\}$ and so on. Then $\bigcup_{n=0}^{\infty} \bigcup_{I \in H_n} I \subseteq E$.

To prove $E \subseteq \bigcup_{n=0}^{\infty} \bigcup_{I \in H_n} I$, take $x \in E$, we find a $B(x, \delta) \subseteq E$. Clearly, we can find unique $I_k \in \Gamma_k$ such that $x \in I_k$ and $\text{diam} I_k = \frac{\sqrt{n}}{2^k}$. Hence $I_k \subseteq B(x, \delta)$ for sufficiently large k, say $k = k_0$. Hence $x \in \bigcup_{I \in H_j} I$ for some $j \leqslant k_0$. □

Remark. If we don't require disjointness, then we find a countable open sets B_k such that $\forall E \subseteq \mathbb{R}^n$ open, $E = \bigcup_{k=1}^{\infty} B_k$. Precisely, we denote $\mathbb{Q} = \{r_1, r_2, \cdots\}$ and $B_{k,l} = B(r_k, \frac{1}{l})$. Then $\{B_{k,l}\}$ fullfills the requirement.

Definition 1.28. $E \subseteq \mathbb{R}^n$ is **compact** if every open cover $\bigcup_{\alpha \in \Lambda} U_\alpha$ of E admits a finite subcover, i.e. $\exists N$ such that $E \subseteq \bigcup_{i=1}^{N} U_{\alpha_i}$.

Theorem 1.29 (Heine-Borel). *$E \subseteq \mathbb{R}^n$ is compact $\Longleftrightarrow$ E is closed and bounded.*

Caution: $\Omega \subseteq E$ is said open in E if $\forall x \in \Omega, \exists \delta > 0, B_\delta(x) \cap E \subseteq \Omega$. For example, $E = (0,1], \Omega = (\frac{1}{2}, 1]$ is open in E but not open in $\mathbb{R}^1$.

Lemma 1.30. *For $f \in C(E, \mathbb{R}^n)$, f maps compact set to compact set.*

Proof. $\forall \Omega \in \mathbb{R}^n$ compact, every open cover $\{U_\alpha\}_{\alpha \in \Lambda}$ of $f(\Omega)$. Then $\Omega \subseteq \bigcup_{\alpha \in \Lambda} f^{-1}(U_\alpha)$. Since Ω is compact, there exists N such that $\Omega \subseteq \bigcup_{i=1}^{N} f^{-1}(U_{\alpha_i})$. Hence $f(\Omega) \subseteq \bigcup_{i=1}^{N} U_{\alpha_i}$. □

Corollary 1.31. *If $f \in C(\mathbb{R}^n, \mathbb{R})$, then on compact set Ω, f is bounded and attains its bounds.*

Proof. Since $f(\Omega)$ is compact, it is closed and bounded. Hence f can attain its supremum and infimum. □

Lemma 1.32. *For every $E \subseteq \mathbb{R}^n$, $f(x) = d(x, E)$ is uniformly continuous. (Lipschitz with constant 1)*

Proof. Given $x \neq y$, $\varepsilon > 0$, we find $z \in E$ such that $d(x,E) \leqslant d(x,z) \leqslant d(x,E) + \varepsilon$. Then $d(y,E) \leqslant d(y,z) \leqslant d(y,x) + d(x,z) \leqslant d(x,y) + d(x,E) + \varepsilon$. Hence $d(y,E) - d(x,E) \leqslant d(x,y) + \varepsilon$. Similarly, $d(x,E) - d(y,E) \leqslant d(x,y) + \varepsilon$. Since ε is arbitrary, we have $|f(x) - f(y)| \leqslant |x - y|$. □

Lemma 1.33. *$\forall E, F$ closed and $E \cap F = \varnothing$, we can find $f \in C(\mathbb{R}^n)$ such that $0 \leqslant f \leqslant 1$, $f(x) = 1$ on E and $f(x) = 0$ on F.*

Proof. $f(x) = \frac{d(x,F)}{d(x,E)+d(x,F)}$. □

In topology, we can generalize the Lemma 1.33 to normal space (X, τ).

Theorem 1.34 (Urysohn's Theorem)**.** *$\forall E, F$ closed and $E \cap F = \varnothing$, we can find $f \in C(X, \mathbb{R})$ satisfies the same conclusion.*

Lemma 1.35 (continuous extension)**.** *$\forall E \subseteq \mathbb{R}^n$ closed, $f \in C(E)$ and bounded, i.e. $|f| \leqslant M$, we can find $\tilde{f} \in C(\mathbb{R}^n)$ such that $\tilde{f}|_E = f$ and $|\tilde{f}| \leqslant M$.*

Proof. Let $A = \{x \in E | f(x) \geqslant \frac{M}{3}\}$, $B = \{x \in E | f(x) \leqslant -\frac{M}{3}\}$ and $C = \{x \in E || f(x)| < \frac{M}{3}\}$. Then A, B are closed and disjoint. By Lemma 1.33, we can find $h \in C(\mathbb{R}^n)$ such that $0 \leqslant h \leqslant 1$, $h(x) = 1$ on A and $h(x) = 0$ on B. Take $g_0 = \frac{2M}{3}h - \frac{M}{3} \in C(\mathbb{R}^n)$. Then $-\frac{M}{3} \leqslant g_0 \leqslant \frac{M}{3}$ and $f_1 = f - g_0 \in \begin{cases} [0, \frac{2M}{3}], & x \in A \\ [-\frac{2M}{3}, 0], & x \in B \\ [-\frac{2M}{3}, \frac{2M}{3}], & x \in C \end{cases}$. $|f - g_0| \leqslant \frac{2M}{3}$ on E and $f - g_0 \in C(\mathbb{R}^n)$. Repeat the construction, we find $g_1 \in C(\mathbb{R}^n)$ and $|g_0| \leqslant \frac{2M}{3^2}$ and $|f_1 - g_1| \leqslant \frac{2^2 M}{3^2}$ on E.

Iterate the construction we get $g_k \in C(\mathbb{R}^n), k = 0, 1, \cdots$ and $|g_k| \leqslant \frac{1}{3}\left(\frac{2}{3}\right)^k M$ and $|f - \sum_{i=0}^k g_i| = |f_{N+1}| \leqslant \left(\frac{2}{3}\right)^{k+1} M$ on E. Take $g = \sum_{k \geqslant 0} \frac{1}{3}\left(\frac{2}{3}\right)^k M = M$. Take $N \to \infty$, we get $\tilde{f} = \sum_{k=0}^{\infty} g_k \in C(\mathbb{R}^n)$ and $\tilde{f}|_E = f$ and $|\tilde{f}| \leqslant M$. □

Remark. Lemma 1.35 can be generalized to normal space (X, τ). And if f is unbounded, we can take $g = \arctan f$ bounded and apply Lemma 1.35 to get $\tilde{g} \in C(\mathbb{R}^n)$ and take $\tilde{f} = \tan \tilde{g}$.

Theorem 1.36 (Tietze Extension Theorem)**.** *$\forall E \subseteq X$ closed, $f \in C(E, [-M, M])$, we can find $\tilde{f} \in C(X, [-M, M])$ such that $\tilde{f}|_E = f$.*

1.7 Borel Sets and σ-algebra

Definition 1.37. E is a F_σ **set** if $E \subseteq \mathbb{R}^n$ and $E = \bigcup_{k=1}^{\infty} F_k$ where F_k are closed. E is a G_δ **set** if $E \subseteq \mathbb{R}^n$ and $E = \bigcap_{k=1}^{\infty} U_k$ where U_k are open.

Notice the fact that $(F_\sigma \text{ set })^c = G_\delta$ set and $(G_\delta \text{ set })^c = F_\sigma$ set .

Definition 1.38. The **Borel hierarchy** is all the sets we get as below:

1. Open and closed sets in $\mathbb{R}^n$.
2. Countable union or intersection of sets in level 1, i.e. F_σ and G_δ sets.
3. Countable union or intersection of sets in level 2, i.e. $F_{\sigma\delta}, F_{\sigma\delta}, G_{\delta\sigma}, G_{\delta\delta}$ sets.
4. $\cdots$

Definition 1.39. X is a set $\Gamma \in \mathscr{P}(X)$ is a σ**-algebra** if

1. $\varnothing \in \Gamma$.
2. $A \in \Gamma \Rightarrow A^c \in \Gamma$
3. $A_n \in \Gamma, n = 1, 2, \cdots \Rightarrow \bigcup_{n=1}^\infty A_n \in \Gamma$.

We have some easy facts:

(a) $X \in \Gamma$.

(b) $A_n \in \Gamma \Rightarrow \bigcap_{n=1}^\infty A_n \in \Gamma$, $\bigcup_{n=1}^\infty A_n \in \Gamma$, $\underline{\lim}_{n\to\infty} A_n \in \Gamma$ and $\overline{\lim}_{n\to\infty} A_n \in \Gamma$.

(c) $A, B \in \Gamma \Rightarrow A \backslash B \in \Gamma$.

(d) If $\Gamma_\alpha, \alpha \in \Sigma$ are all σ-algebras, then $\bigcap_\alpha \Gamma_\alpha$ is a σ-algebra.

Definition 1.40. Given $\Sigma \subseteq \mathscr{P}(X)$, the σ**-algebra generated by** Σ, denoted by $\Gamma(\Sigma) = \bigcap_{\Gamma_\alpha \supseteq \Sigma, \Gamma_\alpha \text{ is a } \sigma\text{-algebra}} \Gamma_\alpha$. And clearly it is the smallest σ-algebra containing Σ.

Definition 1.41. $\mathscr{B}$ = σ-algebra generated by all open sets in $\mathbb{R}^n$ is called **Borel** σ**-algebra** and sets in $\mathscr{B}$ are called **Borel sets**.

Caution. $\mathscr{B} \neq$ Borel hierarchy. The construction of Borel σ-algebra relies on transfinite recursion.

Example 1.42. $\mathbb{Q}$ is a F_σ set because $\mathbb{Q} = \bigcup_{k=1}^\infty \{r_k\}$.

Example 1.43. $f : G \to \mathbb{R}$, $G \subseteq \mathbb{R}^n$ open, $S = \{\text{the continuous point of } f\}$ is G_δ set.

Proof. $\omega_f(x, \delta) = \sup_{y,z\in B_\delta(x)} |f(y) - f(z)|$. Clearly $\omega_f(x, \delta)$ is decreasing as δ decreases. f is continuous at x_0 if and only if $w_f(x_0) := \lim_{\delta\to 0} w_f(x_0, \delta) = 0$. Then, we have

$$S = \bigcap_{k=1}^\infty \{x \in G | w_f(x) < \frac{1}{k}\} = \bigcap_{k=1}^\infty G_k.$$

We are left to prove G_k is open. Given $x_0 \in G_k$, $\exists \delta_0 > 0$ such that $w_f(x_0, \delta_0) < \frac{1}{k}$. Take $\delta_1 = \frac{\delta_0}{100}$, then $\forall \tilde{x} \in B(x_0, \delta_1)$, $B(\tilde{x}, \delta_1) \subseteq B(x_0, \delta_0)$ and hence $w_f(\tilde{x}, \delta_1) \leqslant w_f(x_0, \delta_0) < \frac{1}{k}$. Hence $\tilde{x} \in G_k$ and $B(x_0, \delta_1) \subseteq G_k$, which implies G_k is open. $\square$

1.8 Baire Category Theorem

Theorem 1.44 (Baire Category Theorem). *$E \subseteq \mathbb{R}^n$ and $E = \bigcup_{k=1}^{\infty} F_k$ where F_n are closed and has no interior point. Then E has no interior point.*

Proof. Suppose E has interior point x_0, then $\overline{B(x_0, \delta_0)} \subseteq E$. Since F_1 is closed and has no interior, we can find $x_1 \in B(x_0, \delta_0) \backslash F_1$ and $\overline{B(x_1, \delta_1)} \subseteq B(x_0, \delta_0)$ such that $\overline{B(x_1, \delta_1)} \cap F_1 = \varnothing$. Repeat the process and we have F_k closed and has no interior, then we can find $x_k \in B(x_{k-1}, \delta_{k-1}) \backslash F_k$ and $\overline{B(x_k, \delta_k)} \subseteq B(x_{k-1}, \delta_{k-1})$ such that $\overline{B(x_k, \delta_k)} \cap F_k \neq \varnothing$ and $\delta_k < \frac{\delta_{k-1}}{2}$. Hence $\{x_k\}$ is a Cauchy sequence and converges to x^*. Since $\{x_k\}_{k \geqslant N}^{\infty} \subseteq \overline{B(x_N, \delta_N)}$ and $\overline{B(x_N, \delta_N)}$ is closed, $x^* \in \overline{B(x_N, \delta_N)}$. Since $\overline{B(x_N, \delta_N)} \cap F_N = \varnothing$, $x^* \in F_N$. This contradicts $x^* \in E$. □

Actually, Baire Category Theorem holds for any complete metric space.

Definition 1.45. $E \subseteq \mathbb{R}^n$ is **dense** if $\overline{E} = \mathbb{R}^n$. E is **nowhere dense** if $\overline{E}^{\circ} = \varnothing$.

E is of **first category** or a **meager set** if $E = \bigcup_{k=1}^{\infty} E_k$ where E_k are nowhere dense. E is of **second category** if it is not of first category.

Example 1.46. $\mathbb{Q}^n$ is of first category and $\mathbb{R}^n$ is of second category.

Next is a useful corollary of Baire Category Theorem.

Corollary 1.47. *If $U_n \subseteq \mathbb{R}^n$ are open and dense, then $\bigcap_{n=1}^{\infty} U_n$ is dense.*

Proof. Let $F_n = U_n^c$, then F_n are closed and has no interior point. By Baire Category Theorem, $\left(\bigcup_{n=1}^{\infty} F_n\right)^{\circ} = \varnothing$. Hence $\overline{\bigcap_{n=1}^{\infty} U_n} = \overline{\left(\bigcup_{n=1}^{\infty} F_n\right)^{c}} = \mathbb{R}^n$. □

Example 1.48. $\mathbb{Q}^n$ can not be G_δ set.

Proof. Suppose $\mathbb{Q}^n = \bigcap_{k=1}^{\infty} U_k$ where U_k are open. Since $\mathbb{Q}^n \subseteq U_k$, U_k are dense and U_k^c has no interior point. $\mathbb{R}^n = \mathbb{Q}^n \cup \bigcup_{k=1}^{\infty} U_k^c = \{r_l\}_{l \geqslant 1} \cup \bigcup_{k=1}^{\infty} U_k^c$. By Baire Category Theorem, $\mathbb{R}^n$ has no interior point, which is a contradiction. □

Example 1.49. There exists no function $f : \mathbb{R} \to \mathbb{R}$ such that f is continuous at $\mathbb{Q}$ and discontinuous at $\mathbb{Q}^c$.

Proof. By Example 1.43, {continuous points of f} is a G_δ set. So it can not be $\mathbb{Q}$. □

1.9 Cantor Set

$$F_0 = [0, 1], \quad F_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] = F_{1,1} \cup F_{1,2},$$

i.e. remove the middle $\frac{1}{3}$ of F_0 and then F_1 contains two closed intervals each of length $\frac{1}{3}$. Then

$$F_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right] = F_{2,1} \cup F_{2,2} \cup F_{2,3} \cup F_{2,4}$$

contains four intervals each of length $\frac{1}{9}$. Continue the process and we get F_n, which is closed and contains 2^n intervals each of length $\frac{1}{3^n}$.

Definition 1.50. The **Cantor set** is $C = \bigcap_{n=0}^{\infty} F_n$.

Definition 1.51. $E \subseteq \mathbb{R}^n$ is a **perfect set** if $E' = E$.

Notice some basic properties of Cantor set:

1. C is nonempty, closed and bounded.
2. C has no interior point.
3. C is not path connect.
4. C is a perfect set.
5. $|C| = |\mathbb{R}|$.

Proof. (2): For every $x \in C$ and any $\delta > 0$, consider $(x-\delta, x+\delta) \cap C$. Take n_0 large enough such that $\frac{1}{3^{n_0}} < \frac{\delta}{10}$ and $x \in F_{n_0,j}$. Hence $(x-\delta, x+\delta) \cap C^c \neq \varnothing$. (3) is similar.

(4): For every $x \in C$ and every n such that $x \in F_{n,j_n}$. Take the endpoint of F_{n,j_n}, which is not x, denoted by x_n. Then $x_n \in C$ and $|x_n - x| \leqslant \frac{1}{3^n}$. Hence x is a limit point of C and $C \subseteq C'$. Since C is closed, $C' \subseteq C$ and $C = C'$.

(5): For every $x \in C$, $x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$ where $a_n = \{0, 2\}$. Consider $f : C \to [0,1]$, $f(x) = \sum_{n=1}^{\infty} \frac{\frac{1}{2}a_n}{2^n}$, which is a surjection. Hence $|[0,1]| \leqslant |C| \leqslant |[0,1]|$ and $|C| = |\mathbb{R}|$. ☐

$F_0 = [0,1]$, F_1 is obtained by removing a middle interval of length $\frac{1}{p}$ from F_0, F_2 is obtained by removing two interval of length $\frac{1}{p^2}$ from F_1 and so on. Then F_n contains 2^n intervals each of length $\frac{1}{p^n}$.

Definition 1.52. The **Harnack set** is $\mathscr{H} = \bigcap_{k=0}^{\infty} F_k$.

The length of H is

$$|\mathscr{H}| = 1 - \sum_{i=1}^{\infty} \frac{2^{i-1}}{p^i} = 1 - \frac{1}{p-2} > 0, \quad p > 3.$$

We now construct the Cantor-Lebesgue function. Recall that we have

$$f : C \to [0,1], \quad f(x) = \sum_{n=1}^{\infty} \frac{\frac{1}{2}a_n}{2^n}.$$

Notice that $f(0) = 0$, $f(1) = 1$ and f is increasing.

Proposition 1.53. *For each interval (α, β) removed in the construction, we have $f(\alpha) = f(\beta)$.*

Proof.

$$\alpha = \sum_{k=1}^{N} \frac{a_k}{3^k} + \frac{1}{3^{N+1}} = \sum_{k=1}^{N} \frac{a_k}{3^k} + \sum_{k=N+2}^{\infty} \frac{2}{3^k}, \quad \beta = \sum_{k=1}^{N} \frac{a_k}{3^k} + \frac{2}{3^{N+1}}.$$

Then $f(\alpha) = \sum_{k=1}^{N} \frac{\frac{1}{2}a_k}{2^k} + \frac{1}{2^{N+1}}$ and $f(\beta) = \sum_{k=1}^{N} \frac{\frac{1}{2}a_k}{2^k} + \frac{1}{2^{N+1}}$. Hence f is not injective. ☐

Now we extend f to a function $F : [0,1] \to [0,1]$ such that

$$F(x) = \sup\{f(y), y \in C, y \leqslant x\}.$$

Then F is increasing, onto and continuous. F is a constant on any intervals removed in the construction and $F' = 0$ in these intervals. However, $\int_0^1 F'(x)dx = 0 \neq F(1) - F(0) = 1$.

2 Lebesgue Measure

2.1 Lebesgue Measure

Our goal is to construct measure on set $X = \mathbb{R}^n$.

Definition 2.1. Given X a set and $\mathscr{A} \subseteq \mathscr{P}(X)$ a σ-algebra, then $(X, \mathscr{A})$ is a **measurable space**. Given measurable space $(X, \mathscr{A})$, $\mu : \mathscr{A} \to [0, +\infty]$ is a **measure** if

1. $\mu(\varnothing) = 0$.
2. If $A_n \in \mathscr{A}, n = 1, 2, \cdots$ are disjoint, then $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

And $(X, \mathscr{A}, \mu)$ is called a **measure space**.

Here is a naive approach. We expect for $I = [a, b]$, (a, b), $(a, b]$ or $[a, b)$, $\mu(I) = b - a$ and for $I = I_1 \times I_2 \times \cdots \times I_n$, $\mu(I) = \prod_{i=1}^{n} |I_i|$.

Definition 2.2. Given $E \subseteq \mathbb{R}^n$, $\{I_j\}_{j \geqslant 1}$ is a **L-cover** of E if I_j are open rectangles and $E \subseteq \bigcup_{j \geqslant 1} I_j$.

Definition 2.3. $m^*(E) = \inf \left\{ \sum_{j=1}^{\infty} |I_j| : \{I_j\}_{j \geqslant 1} \text{ is a } L\text{-cover of } E \right\}$ is called the **outer Lebesgue measure** of E.

Example 2.4. We have some basic facts:

(a) $m^*(\varnothing) = 0$ and $m^*(\mathbb{R}^n) = +\infty$;

(b) $m^*(\mathbb{Q}^n) = 0$;

(c) For every open rectangle I, $m^*(I) = |I| = m^*(\overline{I})$.

Proof. (b): $\mathbb{Q}^n = \{r_j\}_{j=1}^{\infty}$. For every ε, take cube I_j centered at r_j with volume $\frac{\varepsilon}{2^j}$, then $\{I_j\}_{j \geqslant 1}$ is a L-cover of $\mathbb{Q}^n$ and $\sum_{j=1}^{\infty} |I_j| = \varepsilon$.

(c): The first equality follows from the definition. For the second equality, $\forall \lambda > 1$, $\overline{I} \subseteq \lambda I$ and $|\lambda I| = \lambda^n |I|$. Hence $m^*(\overline{I}) \leqslant |\lambda I|$ and $m^*(\overline{I}) \leqslant |I|$ as $\lambda \to 1$. □

Theorem 2.5. *Here are some properties:*

1. $m^*(E) \geqslant 0$;
2. $E_1 \subseteq E_2 \Rightarrow m^*(E_1) \leqslant m^*(E_2)$.
3. $m^*\left(\bigcup_{k=1}^{\infty} E_k\right) \leqslant \sum_{k=1}^{\infty} m^*(E_k)$;
4. *Outer regularity:* $m^*(E) = \inf\{m^*(O), O \supseteq E, O \text{ is open}\}$;
5. *If* $d(E_1, E_2) = \delta > 0$, *then* $m^*(E_1 \cup E_2) = m^*(E_1) + m^*(E_2)$;
6. $m^*(E + \{x_0\}) = m^*(E)$.
7. $m^*(\lambda E) = |\lambda|^n m^*(E)$.

Proof. (2): Any L-cover of E_2 is also a L-cover of E_1.

(3): WLOG, assume $\sum_{k=1}^{\infty} m^*(E_k) < +\infty$. Take L-cover of E_j such that

$$E_j \subseteq \bigcup_{k=1}^{\infty} I_{j,k}, \quad \sum_k |I_{j,k}| \leqslant m^*(E_j) + \frac{\varepsilon}{2^j}.$$

Then $\bigcup_{j,k} I_{j,k}$ is a L-cover of $\bigcup_{j=1}^{\infty} E_j$ and $m^*\left(\bigcup_{k=1}^{\infty} E_k\right) \leqslant \sum_{j,k} |I_{j,k}| \leqslant \sum_{j=1}^{\infty} m^*(E_j) + \varepsilon$. Since ε is arbitrary, we get the conclusion.

(4): LHS $\leqslant$ RHS is obvious. For every ε, there exists an L-cover $\{I_j\}_{j\geqslant 1}$ such that $\sum_{j=1}^{\infty} |I_j| \leqslant m^*(E)+\varepsilon$. Take $O = \bigcup_{j=1}^{\infty} I_j \supseteq E$ open, then $m^*(O) \leqslant \sum_{j=1}^{\infty} |I_j| \leqslant m^*(E)+\varepsilon$. Since ε is arbitrary, we get the conclusion.

(5): LHS $\leqslant$ RHS follows from 3. For every $\varepsilon > 0$, take L-cover $\{I_j\}_{j\geqslant 1}$ of $E_1 \cup E_2$ such that

$$\sum_{j=1}^{\infty} |I_j| \leqslant m^*(E_1 \cup E_2) + \varepsilon.$$

We refine the L-cover such that each I_j is decomposed into a collection of $I_{j,k}$ (maybe open, closed and half open and half closed) such that diam$I_{j,k} < \frac{\delta}{10}$. Consider $\widetilde{I_{j,k}} = \lambda \mathring{I}_{j,k}, \lambda > 1$, then $\bigcup_{j,k} \widetilde{I_{j,k}}$ is is a L-cover for $E_1 \cup E_2$ and diam$\widetilde{I_{j,k}} < \frac{\delta\lambda}{10}$. Then each $\widetilde{I_{j,k}}$ intersects with atmost one of E_1 or E_2. Hence $\{\widetilde{I_{j,k}}\}_{\widetilde{I_{j,k}} \cap E_1 \neq \varnothing}$ and $\{\widetilde{I_{j,k}}\}_{\widetilde{I_{j,k}} \cap E_2 \neq \varnothing}$ are L-covers of E_1 and E_2. Hence

$$m^*(E_1) + m^*(E_2) \leqslant \sum_{j,k} |\widetilde{I_{j,k}}| \leqslant \lambda^n \sum_{j,k} |I_{j,k}| = \lambda^n \sum_{j=1}^{\infty} |I_j| \leqslant \lambda^n (m^*(E_1 \cup E_2) + \varepsilon).$$

Let $\varepsilon \to 0$ and $\lambda \to 1$, we get the conclusion. □

Our motivation to construct a measure is to expect $m^*(E) = m_*(E)$, where $m^*(E) = \inf\{\sum_{j=1}^{\infty} |I_j| : \{I_j\}_{j\geqslant 1}$ is a L-cover of $E\}$ and $m_*(E) = \sup\{\sum_{j=1}^{\infty} |I_j| : I_j$ open and $\bigcup_{j=1}^{\infty} I_j \subseteq E\}$.

However, the expectation definition of "measurable set" $m^*(A) = m_*(A)$, which is called the **Jordan measure**, is bad. Consider $E = [0,1]\backslash\mathbb{Q}$, which is not measurable.

We fix the definition by $\forall I \supseteq E, m_*(E) = |I| - m^*(I\backslash E)$. We can rewrite $m^*(E) = |I| - m^*(I\backslash E)$, then $m^*(E) + m^*(I\backslash E) = |I|$. Hence we have the following definition.

Definition 2.6. $E \subseteq \mathbb{R}^n$ is called **measurable** if $\forall T \subseteq \mathbb{R}^n$,

$$m^*(T) = m^*(T \cap E) + m^*(T \cap E^c), \tag{2.1}$$

where T is called a **test set**. This is called **Carathéodory's criterion**.

$m^*(T) \leqslant m^*(T \cap E) + m^*(T \cap E^c)$ always holds, so we only need to check the other direction.

An alternative definition of measurable set is: E is measurable if for every $\varepsilon > 0$, there exists O open such that $m^*(O\backslash E) < \varepsilon$. This is called **Littlewood First Principle**. We will prove the equivalence later.

Denote $\mathscr{M} = \{$ collection of measurable sets$\}$.

Example 2.7. If $m^*(E) = 0$, then $E \in \mathscr{M}$.

Proof. For every T, we have

$$m^*(T \cap E) + m^*(T \cap E^c) \leqslant m^*(E) + m^*(T) = m^*(T).$$

□

Definition 2.8. E is a **null set** if $m^*(E) = 0$.

Example 2.9. A countable set and Cantor set are null sets.

Proposition 2.10. $|\mathscr{M}| = |\mathscr{P}(\mathbb{R})|$.

Proof. Notice $\mathscr{M} \subseteq \mathscr{P}(\mathbb{R})^n \sim \mathscr{P}(\mathbb{R})$. Also we have $C \in \mathscr{M}$ and $\mathscr{P}(C) \subseteq \mathscr{M}$. By $C \sim \mathbb{R}$, we have $|\mathscr{M}| = |\mathscr{P}(\mathbb{R})|$. □

Theorem 2.11. *$\mathscr{M}$ is a σ-algebra and $m^*|_{\mathscr{M}}$ is a measure, i.e.*

1. $\varnothing \in \mathscr{M}$.
2. $A \in \mathscr{M} \Rightarrow A^c \in \mathscr{M}$.
3. $A_j \in \mathscr{M} \Rightarrow \bigcup_{j=1}^{\infty} A_j \in \mathscr{M}$.
4. $m^*(\varnothing) = 0$.
5. $m^*\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} m^*(A_j)$ *if* $A_j \in \mathscr{M}$ *and disjoint.*

Proof. (1), (2), (4) follows from definition.

Let's first consider finite union $A_1 \cup A_2$. For every T, we have

$$\begin{aligned} m^*(T \cap (A_1 \cup A_2)) + m^*(T \cap (A_1 \cup A_2)^c) &\leqslant m^*(T \cap A_1 \cap A_2^c) + m^*(T \cap A_1 \cap A_2) \\ &+ m^*(T \cap A_1^c \cap A_2^c) + m^*(T \cap A_1^c \cap A_2) = m^*(T \cap A_1) + m^*(T \cap A_1^c) = m^*(T). \end{aligned}$$

Furthermore, if $A_1, A_2 \in \mathscr{M}$ and $A_1 \cap A_2 = \varnothing$, then $m^*(A_1 \cup A_2) = m^*(A_1) + m^*(A_2)$.

Now for countable many $A_n \in \mathscr{M}$, WLOG, assume A_n are disjoint (otherwise, consider $\widetilde{A_n} = A_n \backslash \bigcup_{k=1}^{n-1} A_k$). Denote

$$S_N = \bigcup_{n=1}^{N} A_n, \quad S_\infty = \bigcup_{n=1}^{\infty} A_n.$$

For every $T \subseteq \mathbb{R}^n$, we have

$$\begin{aligned} m^*(T) = m^*(T \cap S_N) + m^*(T \cap S_N^c) &= \sum_{j=1}^{N} m^*(T \cap A_j) + m^*(T \cap S_N^c) \\ &\geqslant \sum_{j=1}^{N} m^*(T \cap A_j) + m^*(T \cap S_\infty^c) \geqslant \sum_{j=1}^{\infty} m^*(T \cap A_j) + m^*(T \cap S_\infty^c) \\ &\geqslant m^*(T \cap S_\infty) + m^*(T \cap S_\infty^c) \geqslant m^*(T). \end{aligned}$$

Hence, $S_\infty \in \mathscr{M}$. Furthermore, take $T = S_\infty$, we have $m^*(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} m^*(A_j)$. □

Definition 2.12. Denote $m^* = m$ on $\mathscr{M}$. m is called the **Lebesgue measure**.

Theorem 2.13. *We show some properties of* m*:*

1. $E, F \in \mathscr{M}$ *and* $E \subseteq F \Longrightarrow m(E) \leqslant m(F)$.
2. *Subadditivity:* $E_j \in \mathscr{M} \Rightarrow m\left(\bigcup_{j=1}^{\infty} E_j\right) \leqslant \sum_{j=1}^{\infty} m(E_j)$.
3. *Continuous from below:* $E_j \in \mathscr{M}$ *and* $E_1 \subseteq E_2 \subseteq \cdots \subseteq E_n \cdots$*, then*
$$m(\lim_{n\to\infty} E_n) = \lim_{n\to\infty} m(E_n). \tag{2.2}$$
4. *Continuous from above:* $F_j \in \mathscr{M}$ *and* $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \cdots$ *and* $m(F_{k_0}) < +\infty$ *for some* k_0*, then*
$$m(\lim_{n\to\infty} F_n) = \lim_{n\to\infty} m(F_n). \tag{2.3}$$

Proof. (3):

$$\begin{aligned} m\left(\bigcup_{n=1}^{\infty} E_n\right) &= m\left(\bigcup_{n=1}^{N} (E_n \backslash E_{n-1})\right) = \sum_{n=1}^{N} m(E_n \backslash E_{n-1}) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} (m(E_n) - m(E_{n-1})) = \lim_{N\to\infty} m(E_N). \end{aligned}$$

(4): WLOG, assume $m(F_1) < +\infty$. Denote $E_n = F_1 \backslash F_n \in \mathscr{M}$, then E_n increases.

$$\begin{aligned} \lim_{n\to\infty} (m(F_1) - m(F_n)) &= \lim_{n\to\infty} m(E_n) = m\left(\lim_{n\to\infty} E_n\right) = m(\bigcup_{n=1}^{\infty} E_n) \\ &= m(F_1) - m\left(\bigcap_{n=1}^{\infty} F_n\right) = m(F_1) - m\left(\lim_{n\to\infty} F_n\right). \end{aligned}$$

□

Caution: Without the finiteness of $m(F_{k_0})$, the conclusion may not hold. For example, let $F_n = [n, +\infty)$, then $\bigcap_{n=1}^{\infty} F_n = \varnothing$ and $m(F_n) = +\infty$ for all n.

Corollary 2.14. *For* $E_n \in \mathscr{M}$*, we have*

$$m\left(\varliminf_{n\to\infty} E_n\right) \leqslant \varliminf_{n\to\infty} m(E_n), \quad m\left(\varlimsup_{n\to\infty} E_n\right) \geqslant \varlimsup_{n\to\infty} m(E_n), \tag{2.4}$$

provided $m(\bigcup_{n\geqslant k_0} E_n) < +\infty$ *for some* k_0*.*

Proof.

$$m\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} E_k\right) = \lim_{n\to\infty} m\left(\bigcap_{k=n}^{\infty} E_k\right) \leqslant \varliminf_{n\to\infty} m(E_n).$$

$$m\left(\bigcap_{n\geqslant 1}\bigcup_{k\geqslant n} E_k\right) = \lim_{n\to\infty} m\left(\bigcup_{k\geqslant n} E_k\right) \geqslant \varlimsup_{n\to\infty} m(E_n).$$

□

Corollary 2.15 (Borel-Cantelli Lemma)**.** *If* $E_j \in \mathscr{M}, \sum_{j=1}^{\infty} m(E_j) < +\infty$*, then*

$$m\left(\overline{\lim_{n\to\infty}} E_n\right) = m\left(\underline{\lim_{n\to\infty}} E_n\right) = 0. \tag{2.5}$$

Proof.

$$m\left(\underline{\lim_{n\to\infty}} E_n\right) \leqslant \underline{\lim_{n\to\infty}} m\left(E_n\right) = 0.$$

$$m\left(\overline{\lim_{n\to\infty}} E_n\right) = \lim_{n\to\infty} m\left(\bigcup_{k\geqslant n} E_k\right) \leqslant \lim_{n\to\infty} \sum_{k\geqslant n} m(E_k) = 0.$$

□

2.2 Measurable Set

Our goal is to show Borel sets are measurable and measurable sets are not so far from Borel sets and open sets. We begin with a useful lemma.

Lemma 2.16 (Carathéodory's Lemma)**.** *Given* $G \subset \mathbb{R}^n$ *open and* $E \subseteq G$*, denote* $E_k = \{x \in E : d(x, G^c) \geqslant \frac{1}{k}\}$*. Then*

$$\lim_{k\to\infty} m^*(E_k) = m^*(E). \tag{2.6}$$

Proof. Notice that E_k increases as k increases and $E = \bigcup_{k=1}^{\infty} E_k$. Hence $\lim_{k\to\infty} m^*(E_k) \leqslant m^*(E)$.

To prove the reverse, denote $A_j = E_j \backslash E_{j-1}$ and $E_0 = \varnothing$, then we can notice that $d(A_j, A_{j+2}) > 0$. WLOG, we assume $\lim_{n\to\infty} m^*(E_k) < \infty$. Since $E_{2k} = \bigcup_{j=1}^{k} A_j \supseteq \bigcup_{j=1}^{k} A_{2j}$, we have

$$m^*(E_{2k}) \geqslant m^*\left(\bigcup_{j=1}^{k} A_{2j}\right) = \sum_{j=1}^{k} m^*(A_{2j}). \tag{2.7}$$

Similarly, $m^*(E_{2k+1}) \geqslant \sum_{j=1}^{k+1} m^*(A_{2j-1})$.

$$\begin{aligned} m^*(E) &= m^*\left(E_{2k} \cup \left(\bigcup_{j=k}^{\infty} A_{2j}\right) \cup \left(\bigcup_{j=k}^{\infty} A_{2j+1}\right)\right) \\ &\leqslant m^*(E_{2k}) + \sum_{j=k}^{\infty} m^*(A_{2j}) + \sum_{j=k}^{\infty} m^*(A_{2j+1}) \end{aligned} \tag{2.8}$$

Let $k \to \infty$, we have $m^*(E) \leqslant \lim_{k\to\infty} m^*(E_{2k})$. □

Theorem 2.17. $\mathscr{B} \subseteq \mathscr{M}$.

Proof. We only need to prove closed sets are measurable. For every F closed and $T \subseteq \mathbb{R}^n$ a test set, apply Carathéodory's lemma, we have $\{F_k\} \subseteq T \cap F^c$, such that

$$d(F_k, F) \geqslant \frac{1}{k}, \quad \lim_{k\to\infty} m^*(F_k) = m^*(T \backslash F).$$

$$\begin{aligned} m^*(T) = m^*((T \cap F) \cup (T \cap F^c)) \geqslant m^*((T \cap F) \cup F_k) = m^*(T \cap F) + m^*(F_k) \\ \to m^*(T \cap F) + m^*(T \cap F^c), \quad k \to \infty. \end{aligned}$$

□

Lemma 2.18. *Let $E \in \mathscr{M}$, then for every $\varepsilon > 0$, there exists G open and F closed such that $F \subseteq E \subseteq G$, $m(G\backslash E) < \varepsilon$ and $m(E\backslash F) < \varepsilon$.*

Proof. First consider $m(E) < \infty$. Take L-cover $\{I_j\}$ for E such that $\sum_{j=1}^{\infty} |I_j| \leqslant m(E)+\varepsilon$. Take $G = \bigcup_{j=1}^{\infty} I_j$ open and $E \subseteq G$, then $m(G) \leqslant \sum_{j=1}^{\infty} |I_j| \leqslant m(E)+\varepsilon$ and $m(G\backslash E) < \varepsilon$.

When $m(E) = +\infty$, consider

$$E_N = E \cap [-N, N]^n.$$

For every N, take G_N open such that $G_N \supseteq E_N$ and $m(G_N\backslash E_N) < \frac{\varepsilon}{2^N}$. Take $G = \bigcup_{N=1}^{\infty} G_N$ open and $E \subseteq G$.

$$G\backslash E = \bigcup_{N=1}^{\infty}(G_N\backslash E) \subseteq \bigcup_{N=1}^{\infty}(G_N\backslash E_N), \quad m(G\backslash E) \leqslant \sum_{N=1}^{\infty} m(G_N\backslash E_N) < \varepsilon.$$

Now consider E^c. Take $\widetilde{G}$ open such that $\widetilde{G} \supseteq E^c$, $m(\widetilde{G}\backslash E^c) < \varepsilon$. Take $F = \widetilde{G}^c$ closed and $F \subseteq E$, then we have $m(E\backslash F) = m(\widetilde{G}\backslash E^c) < \varepsilon$. □

Theorem 2.19. *For $E \in \mathscr{M}$, we can find a G_δ set H and a F_σ set K such that $E = H\backslash Z_1, E = K\backslash Z_2$ with Z_1, Z_2 null sets.*

Proof. Take $\varepsilon = \frac{1}{k}$ in Lemma 2.18, we get G_k open and F_k closed such that $F_k \subseteq E \subseteq G_k$, $m(G_k\backslash E) < \frac{1}{k}$ and $m(E\backslash F_k) < \frac{1}{k}$. Take $H = \bigcap_{k=1}^{\infty} G_k$ and $K = \bigcup_{k=1}^{\infty} F_k$, then

$$m(H\backslash E) = m\left(\bigcap_{k=1}^{\infty}(G_k\backslash E)\right) \leqslant m\left(G_k\backslash E\right) < \frac{1}{k} \to 0,$$

$$m(E\backslash K) = m\left(\bigcap_{k=1}^{\infty}(E\backslash F_k)\right) \leqslant m(E\backslash F_k) < \frac{1}{k} \to 0, \quad k \to \infty.$$

□

Remark: For $E \subseteq \mathbb{R}^n$, the followings are equivalent:

(a) $\forall T \subseteq \mathbb{R}^n$, $m^*(T) = m^*(T \cap E) + m^*(T \cap E^c)$.

(b) $\forall \varepsilon > 0, \exists G \supseteq E$ open such that $m^*(G\backslash E) < \varepsilon$.

Proof. We have showed that (a) ⇒ (b). For (b) ⇒ (a): $\forall T \subseteq \mathbb{R}^n$, take $G \supseteq E$ open such that $m^*(G\backslash E) < \varepsilon$. Take $G_k = \{x \in T \cap G : d(x, G^c) \geqslant \frac{1}{k}\}$, then

$$\lim_{k\to\infty} m^*(G_k) = m^*(T \cap G), \quad m^*(T) \geqslant m^*(G_k \cup (T \cap G^c)) = m^*(G_k) + m^*(T \cap G^c).$$

Let $k \to \infty$, we get $m^*(T) \geqslant m^*(T \cap G) + m^*(T \cap G^c)$. Since $m^*(T \cap G) \geqslant m^*(T \cap E)$ and $m^*(T \cap E^c) \leqslant m^*(T \cap E^c \cap G) + m^*(T \cap E^c \cap G^c) \leqslant \varepsilon + m^*(T \cap E^c \cap G^c)$, we have

$$m^*(T) \geqslant m^*(T \cap E) + m^*(T \cap E^c) - \varepsilon.$$

Since ε is arbitrary, we get the conclusion. □

Definition 2.20. H is called a **equi-measure hull** of E and K is called a **equi-measure kernel** of E.

The concept of equi-measure hull can be generalized to arbitrary sets.

Theorem 2.21. *If $E \subseteq \mathbb{R}^n$, we can find a G_δ set $H \supseteq E$ such that $m(H) = m^*(E)$.*

Proof. For every $k \in \mathbb{N}$, we find L-cover $\{I_j^{(k)}\}_{j=1}^\infty$ of E such that $\sum_{j=1}^\infty |I_j^{(k)}| \leqslant m^*(E) + \frac{1}{k}$. Take $H = \bigcap_{k=1}^\infty \bigcup_{j=1}^\infty I_{j,k}$ a G_δ set, then

$$m^*(E) \leqslant m^*(H) \leqslant m^*\left(\bigcup_{j=1}^\infty I_j^{(k)}\right) \leqslant m^*(E) + \frac{1}{k}.$$

Let $k \to \infty$, then we get the conclusion. □

Caution: If $m^*(E) < \infty$, $m(H) - m^*(E) = 0$, but this does not implies $m^*(H \backslash E) = 0$. If $m^*(H \backslash E) = 0$, then $H \backslash E \in \mathscr{M}$ and hence $E \in \mathscr{M}$, but we are dealing with general sets.

Corollary 2.22. *For $E_k \subseteq \mathbb{R}^n$, we have*

1.
$$m^*\left(\varliminf_{k\to\infty} E_k\right) \leqslant \varliminf_{k\to\infty} m^*(E_k). \tag{2.9}$$

2. *If E_k are increasing, then*
$$m^*\left(\lim_{k\to\infty} E_k\right) = \lim_{k\to\infty} m^*(E_k). \tag{2.10}$$

Proof. (1): Take H_k the equi-measure hull of E_k. Then

$$\varliminf_{k\to\infty} E_k \subseteq \varliminf_{k\to\infty} H_k, \quad m^*\left(\varliminf_{k\to\infty} E_k\right) \leqslant m\left(\varliminf_{k\to\infty} H_k\right) \leqslant \varliminf_{k\to\infty} m(H_k) = \varliminf_{k\to\infty} m^*(E_k).$$

(2): If E_k are increasing, then

$$m^*(E_k) \leqslant m^*(\bigcup_{k=1}^\infty E_k) \leqslant \varliminf_{k\to\infty} m^*(E_k) \Rightarrow \lim_{k\to\infty} m^*(E_k) \leqslant m^*\left(\lim_{k\to\infty} E_k\right) = \lim_{k\to\infty} m^*(E_k).$$

□

Remark: $(\mathbb{R}^n, \mathscr{B}, m)$ is a measure space and here m is called Lebesgue-Borel measure.

In general, given a measure space $(X, \mathscr{A}, \mu)$ and $E \subseteq \mathscr{A}$ such that $\mu(E) = 0$, we don't know if subset of E is in $\mathscr{A}$.

Definition 2.23. A measure μ is called **complete** if its domain contains all subsets of null sets.

Theorem 2.24. *For every measure space $(X, \mathscr{A}, \mu)$, we can find a completion $(X, \widetilde{\mathscr{A}}, \widetilde{\mu})$ such that $\mathscr{A} \subseteq \widetilde{\mathscr{A}}$, $\widetilde{\mu}|_{\mathscr{A}} = \mu$ and $\widetilde{\mu}$ is complete.*

In this sense, $(\mathbb{R}^n, \mathscr{M}, m)$ is the completion of $(\mathbb{R}^n, \mathscr{B}, m)$, where $\mathscr{M}$ is the σ-algebra generated by $\mathscr{B}$ and all subsets of null sets.

Now we study sets with positive measure.

Theorem 2.25. *If $E \in \mathscr{M}$ and $m(E) > 0$, then for every $\lambda \in (0,1)$, we can find open rectangle I such that $\lambda|I| < m(E \cap I)$.*

Proof. WLOG, suppose $0 < m(E) < \infty$. Otherwise, take $\widetilde{E} = E \cap [-N, N]$ for N large enough. Suppose the statement is false. Take ε (decided later) and L-cover $\{I_j\}$ of E such that $\sum_{j=1}^{\infty} |I_j| \leqslant m(E) + \varepsilon$. Then

$$\lambda|I_j| \geqslant m(E \cap I_j) \Longrightarrow \frac{m(E)}{\lambda} = \frac{m\left(\left(\bigcup_{j=1}^{\infty} I_j\right) \cap E\right)}{\lambda} \leqslant \sum_{j=1}^{\infty} \frac{m(I_j \cap E)}{\lambda}$$
$$\leqslant \sum_{j=1}^{\infty} |I_j| \leqslant m(E) + \varepsilon.$$

Hence we can take ε such that $\frac{\lambda\varepsilon}{1-\lambda} = \frac{1}{2}m(E)$ and get the contradiction. □

Theorem 2.26 (Steinhaus)**.** *If $E \in \mathbb{R}$ measurable and $m(E) > 0$, then there exists $\delta > 0$ such that*

$$B(0, \delta) \subset E - E = \{x - y : x, y \in E\}.$$

Proof. WLOG, $0 < m(E) < \infty$. By Theorem 2.25, for $\lambda \in (0,1)$ (decided later), we can find open rectangle I such that $\lambda|I| < m(E \cap I)$ and denote $\widetilde{E} = I \cap E \subset I$.

Suppose the conclusion is false. Then we find $x_k \to 0$ such that $x_k \notin \widetilde{E} - \widetilde{E}$, i.e. $\widetilde{E} \cap \widetilde{E} + \{x_k\} = \varnothing$. Hence for k large enough, we have

$$\begin{aligned} 2m(\widetilde{E}) &= m(\widetilde{E} \cup (\widetilde{E} + \{x_k\})) \leqslant m(I \cup (I + \{x_k\})) \\ &= m(I) + m(I + \{x_k\}) - m(I \cap (I + \{x_k\})) \leqslant 2|I| - 2^{-n}|I| \leqslant \frac{2 - 2^{-n}}{\lambda} m(\widetilde{E}) \end{aligned}$$

Take $\lambda \in (0,1)$ such that $\lambda > 1 - \frac{1}{2^{n+1}}$, then we get contradiction. □

Remark: Modern Problem: For $E \in \mathscr{M}$ and $\Omega = \{|x-y| | x, y \in E\}$, when will Ω have nontrivial measure? Then we have Falconer's distance problem: If Hausdorff dimension of $E > \frac{n}{2}$, then Ω has non trivial measure.

Also, we can not take $\lambda = 1$ in Theorem 2.25. Consider Example 2.27

Example 2.27. There exists $E \subseteq [0,1]$ such that for every I, we have

$$0 < m(E \cap I) < |I|.$$

Proof. Take H_1 the Harnack set such that $m(H_1) = \frac{1}{2}$, then H_1^c is a countable union of open intevals. Change them to closed intervals. Construct Harnack set for each interval such that $m(H_{1j}) = \frac{|I_j^{(1)}|}{2^2}$ and $H_2 = \bigcup_{j=1}^{\infty} H_{1j}$. Repeat the process and let $E = \bigcup_{n=1}^{\infty} H_n$. □

Example 2.28. Suppose $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x + y) = f(x) + f(y)$$

and there exists E of positive measure such that $\forall x \in E, |f(x)| \leqslant M$. Then $f(x) = xf(1)$.

Proof. It's easy to know $\forall p \in \mathbb{Q}$, $|f(px)| = |pf(x)|$. By Steinhaus Theorem, $E - E \supset B(0, \delta)$, hence for every $x \in (-\delta, \delta)$, we have

$$x = x_1 - x_2, x_i \in E \Longrightarrow |f(x)| = |f(x_1) - f(x_2)| \leqslant 2M.$$

For any $x \in \mathbb{R}$, we can find $n \in \mathbb{N}$ large enough such that $n \geqslant \frac{1}{\delta}$ and $p \in \mathbb{Q}$ such that $|nx - p| < \delta$. Then we have

$$\begin{aligned} |f(x) - xf(1)| &= |\frac{1}{n}f(nx) - \frac{1}{n}f(p) + \frac{1}{n}f(p) - xf(1)| \leqslant \frac{1}{n}|f(nx-p)| + |\frac{p}{n} - x||f(1)| \\ &\leqslant \frac{2M}{n} + \frac{\delta}{n}|f(1)| \to 0, \quad n \to \infty. \end{aligned}$$

Hence $f(x) = xf(1)$. □

2.3 Non Measurable Set

Theorem 2.29. *There exists non-measurable sets in $\mathbb{R}^n$.*

Proof. We define an equivalence relation in $\mathbb{R}^n$: $x \sim y$ if $x - y \in \mathbb{Q}^n$, then $[x] := \{z \in \mathbb{R}^n | z \sim x\}$. By Axiom of Choice, we take $W = \{$ pick one representative from each equivalent class $\}$. We claim $W \notin \mathscr{M}$.

Suppose $W \in \mathscr{M}$, then $m(W) = 0$. Because if $m(W) > 0$, then there exists δ such that $B(0, \delta) \subset W - W$. That means there exists

$$r \in B(0, \delta) \cap \mathbb{Q}^n, \quad r \neq 0, \quad r = w_1 - w_2,$$

contradicts the construction of W. Hence $m(W) = 0$.

$$\mathbb{R}^n = \bigcup_{r \in \mathbb{Q}^n} (W + \{r\}).$$

Then $\infty = m(\mathbb{R}^n) = \sum_{r \in \mathbb{Q}^n} m(W + \{r\}) = 0$, which is a contradiction. □

Remark: Consider the one dimension version. Take $V \subset [0, 1]$ to be one such construction, then V is called a **Vitali set**. Clearly, V is non-measurable.

Lemma 2.30. *If g is continuous, then $\forall A \in \mathscr{B}, g^{-1}(A) \in \mathscr{B}$.*

Proof. Let's prove $\Omega = \{A \subseteq \mathbb{R}^n | g^{-1}(A) \in \mathscr{B}\}$ is a σ-algebra. Notice open sets are in Ω. Hence if Ω is σ-algebra, then $\mathscr{B} \subset \Omega$.

First we have $\varnothing \in \Omega$. If $A \in \Omega$, then $g^{-1}(A^c) = (g^{-1}(A))^c \in \mathscr{B}$, then $A^c \in \Omega$. If $\{A_n\} \in \Omega$, then $g^{-1}(\bigcup A_n) = \bigcup g^{-1}(A_n) \in \mathscr{B}$, then $\bigcup A_n \in \Omega$. □

Now we clarify some facts:

1. Lebesgue outer measure does not satisfy finite additivity.
2. It is impossible to find a functions $\mu : \mathscr{P}(\mathbb{R}^n) \to [0, +\infty]$ satisfying
   (a) $\mu(\varnothing) = 0$;
   (b) $\mu(\bigcup E_j) = \sum \mu(E_j)$ where E_j are disjoint;
   (c) If E congruent to F, then $\mu(E) = \mu(F)$;

(d) $\mu((0,1)^n) = 1$.

3. Any set with positive measure must contain a non-measurable subset.

4. $\mathscr{M} \backslash \mathscr{B} \neq \varnothing$.

Proof. (1): Notice the following facts:

(a) $m^*(V) > 0$;

(b) For any $E \subset V$ and $E \in \mathscr{M}$, we have $m(E) = 0$;

(c) $m^*([0,1]\backslash V) = 1$. Suppose $m^*([0,1]\backslash V) < 1$, take L-cover $\{I_j\}$ of $[0,1]\backslash V$ such that $\sum |I_j| < m^*([0,1]\backslash V) + \varepsilon < 1$ for ε small enough. Then $m([0,1]\backslash(\bigcup I_j)) = 1 - m(\bigcup I_j) \geqslant 1 - \sum |I_j| > 0$. Hence $[0,1]\backslash V \subset \bigcup I_j \Rightarrow [0,1]\backslash(\bigcup I_j) \subset V$. From (b), we get contradiction.

Combine the three facts, we get

$$m^*([0,1]) < m^*(V) + m^*([0,1]\backslash V) = 1 + m^*([0,1]\backslash V) < 1.$$

(2): We prove it for $n = 1$. (Similar proof works for any n.) Consider Vitali set V, $V_r = \{x + r | x \in V \cap [0, 1-r)\} \cup \{x + r - 1 | x \in V \cap [1-r, 1]\}$. Clearly, V_r disjoint from $r \in \mathbb{Q} \cap [0,1]$. Since $[0,1) = \bigcup_{r\in\mathbb{Q}\cap[0,1]} V_r$, we have

$$1 = \mu([0,1)) = \sum_{r\in\mathbb{Q}\cap[0,1]} \mu(V_r) = \sum_{r\in\mathbb{Q}\cap[0,1]} \mu(V).$$

Since $\mu(V_r) = \mu(V)$, if $\mu(V) > 0$, we have $\sum_{r\in\mathbb{Q}\cap[0,1]} \mu(V_r) = +\infty$; if $\mu(V) = 0$, we have $\sum_{r\in\mathbb{Q}\cap[0,1]} \mu(V_r) = 0$. Both cases lead to contradiction.

(3): $\forall E \in \mathscr{M}$ with $m(E) > 0$,

$$\mathbb{R} = \bigcup_{r\in\mathbb{Q}} W + \{r\}, \quad E = E \cap \mathbb{R} = \bigcup_{r\in\mathbb{Q}} (E \cap W_r).$$

If all $E \cap W_r \in \mathscr{M}$, then $E \cap W_r$ is a measurable subset of W_r, hence $m(E \cap W_r) = 0$ and $m(E) = 0$, which is a contradiction.

(4): Consider the Cantor-Lebesgue function $F : [0,1] \to [0,1]$. Take $h(x) = x + F(x) : [0,1] \to [0,2]$ strictly increasing, continuous and onto, hence h^{-1} exists and is continous. Notice for any interval $I \subseteq [0,1]\backslash C$, $f(x) = \text{constant}$ and $m(h(I)) = m(I)$. Hence $m(h([0,1]\backslash C)) = m([0,1]\backslash C) = 1$. Since $m(h([0,1])) = 2$, $m(h(C)) = 1$.

Take $N \subset h(C)$ non-measurable. $h^{-1}(N) \subset C$ is null, hence measurable. We claim $h^{-1}(N) \notin \mathscr{B}$. Otherwise, apply Lemma 2.30 to h^{-1}, then if $h^{-1}(N) \in \mathscr{B}$, we have $N = h(h^{-1}(N)) \in \mathscr{B}$, which is a contradiction. □

Remark:

(a) Axiom of Choice is necessary to the construction of non-measurable set. If we drop Axiom of Choice, then we can construct a model where all sets are measurable (Solovay 1970).

(b) For every $\alpha > 0$, there is a construction of W such that $m^*(W) = \alpha$.

(c) An interesting consequence of Axiom of Choice is Banach-Tarski Paradox: For $n \geqslant 3$, $A, B \subseteq \mathbb{R}^n$ nonempty, bounded and open, we can cut $A = \bigcup_{j=1}^k A_j$ and $B = \bigcup_{j=1}^k B_j$, such that A_j disjoint, B_j disjoint and A_j congruent to B_j. $n = 1, 2$ holds for replacing with countable cut.

2.4 Continuous Tranformation and Measurability

Lemma 2.31. *Let $T \in C(\mathbb{R}^n, \mathbb{R}^n)$. Then T maps F_σ sets to F_σ sets.*

Proof. First assume $F = \bigcup_{i=1}^\infty F_i$ where F_i is closed and bounded. Since F_i is compact, $T(F_i)$ is compact. Hence $T(F) = \bigcup_{i=1}^\infty T(F_i)$ is a F_σ set.

For general F_σ set $F = \bigcup_{i=1}^\infty F_i$ where F_i closed, take $F^{(N)} = F \cap [-N,N]^n = \bigcup_{i=1}^\infty (F_i \cap [-N,N]^n)$. Then $T(F^{(N)}) = \bigcup_{i=1}^\infty T(F_i \cap [-N,N]^n)$ is F_σ set and $T(F) = \bigcup_{N=1}^\infty T(F^{(N)})$ is F_σ set. □

Lemma 2.32. *Suppose $T \in C(\mathbb{R}^n, \mathbb{R}^n)$ and for every null set $Z \subseteq \mathbb{R}^n$, $T(Z)$ is null. Then T maps measurable sets to measurable sets.*

Proof. For every $E \in \mathscr{M}$, we have $E = K \cup Z$, where K is a F_σ set and Z is a null set. Then $T(E) = T(K) \cup T(Z)$, where $T(K)$ is F_σ set and $T(Z)$ is null set, which is measurable. □

Lemma 2.33. *$T : \mathbb{R}^n \to \mathbb{R}^n$ is linear, i.e. $T(x) = T_{n\times n}x$, where $T = T_{n\times n}$ is a $n \times n$ matrix with $\det T \neq 0$. Then for every $E \subseteq \mathbb{R}^n$,*

$$m^*(T(E)) = |\det T| m^*(E). \tag{2.11}$$

Proof. For every nonsingular T, T canbe written as a product of elementary matrices. Let $A_{ij} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & \\ & & 1 & \cdots & 0 & \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix}$, $B_i(\lambda) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$ and $C_{ij}(\lambda) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \lambda & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$. Hence we reduce to the proof if $I = \{x = (x_1, \cdots, x_n), x_i \in (0,1)\}$,

$$m^*(T(I)) = |\det T| m^*(I)$$

for $T = A_{ij}, B_i(\lambda), C_{ij}(\lambda)$.

□

Corollary 2.34. *If T is linear and nonsingular, then for every $E \in \mathscr{M}$, we have $T(E) \in \mathscr{M}$ and*

$$m(T(E)) = |\det T| m(E). \tag{2.12}$$

3 Measurable Function

3.1 Definition and Properties

Denote $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. Assume the basic facts

1. $x + (\pm\infty) = \pm\infty$, $x - (\pm\infty) = \mp\infty$
2. $x \cdot (\pm\infty) = \begin{cases} +\infty & x > 0 \\ -\infty & x < 0 \end{cases}$
3. $(\pm\infty) + (\pm\infty) = \pm\infty, \quad (\pm\infty)(\pm\infty) = +\infty, \quad (\pm\infty)(\mp\infty) = -\infty$

Caution: $(\pm\infty) - (\pm\infty)$ and $0 \cdot (\pm\infty)$ are undefined.

Definition 3.1. Let $E \subseteq \mathbb{R}^n$. $f : E \to \overline{\mathbb{R}}$ is a **measurable function** on E if $\forall t \in \mathbb{R}$, $\{x \in E | f(x) > t\}$ is measurable.

Theorem 3.2. *We have the fact that the followings are equivalent:*

1. $\{x \in E | f(x) > t\}$ *is measurable for all* $t \in \mathbb{R}$*;*
2. $\{x \in E | f(x) \geqslant t\}$ *is measurable for all* $t \in \mathbb{R}$*;*
3. $\{x \in E | f(x) < t\}$ *is measurable for all* $t \in \mathbb{R}$*;*
4. $\{x \in E | f(x) \leqslant t\}$ *is measurable for all* $t \in \mathbb{R}$*;*
5. $\{x \in E | f(x) > t\}$ *is measurable for all* $t \in \Omega$ *where* Ω *is dense in* $\mathbb{R}$*.*

Proof. Simply notice the facts $\{x \in E | f(x) > t\} = \bigcup_{n=1}^{\infty} \{x \in E | f(x) \geqslant t + \frac{1}{n}\}$ and $\{x \in E | f(x) \geqslant t\} = \bigcap_{n=1}^{\infty} \{x \in E | f(x) < t - \frac{1}{n}\}$.

Also if $t_n \in \Omega$ such that $t_n \to t$ and $t_n > t$, we have

$$\{x \in E | f(x) > t\} = \bigcup_{n=1}^{\infty} \{x \in E | f(x) > t_n\}.$$

□

Theorem 3.3. *We have the following properties for* $E \in \mathscr{M}$*:*

1. $f : E \to \overline{\mathbb{R}}$ *is measurable, then the following sets are measurable:*
   (a) $\{x \in E | f(x) = t\} = \{x \in E | f(x) \geqslant t\} \cap \{x \in E | f(x) \leqslant t\}$*;*
   (b) $\{x \in E | f(x) = +\infty\} = \bigcap_{n=1}^{\infty} \{x \in E | f(x) > n\}$ *or* $\{x \in E | f(x) = -\infty\} = \bigcap_{n=1}^{\infty} \{x \in E | f(x) < -n\}$*;*
   (c) $\{x \in E | f(x) < +\infty\} = \bigcup_{n=1}^{\infty} \{x \in E | f(x) < n\}$ *or* $\{x \in E | f(x) > -\infty\} = \bigcup_{n=1}^{\infty} \{x \in E | f(x) > -n\}$*;*
2. f *is measurable on* E_1 *and* E_2*, where* $E_1, E_2 \in \mathscr{M}$*, then* f *is measurable on* $E_1 \cup E_2$*;*
3. f *is measurable on* E *and* $A \subseteq E$ *with* A *measurable, then* $f|_A : A \to \overline{\mathbb{R}}$ *is measurable;*

4. *$E \in \mathscr{M} \iff \chi_E$ is measurable on $\mathbb{R}^n$. If $f : [a, b] \to \mathbb{R}$ is monotonic, then f is measurable.*

5. *$f : E \to \mathbb{R}$ measurable, $\Phi \in C(\mathbb{R}, \mathbb{R})$, then $\Phi(f) : E \to \mathbb{R}$ is measurable.*

6. *$f : E \to \mathbb{R}$ continuous, then f is measurable.*

7. *If f, g are measurable on E, then λf, $f \pm g$, f^2, $f \cdot g$ are all measurable (here $\lambda \in \mathbb{R}$).*

8. *$\{f_k\}$ are measurable on E, then $\sup_{k\to\infty} f_k$, $\inf_{k\to\infty} f_k$, $\overline{\lim}_{k\to\infty} f_k$, $\underline{\lim}_{k\to\infty} f_k$ are all measurable.*

9. *f is measurable on E. Then $f^+(x) = \max\{f(x), 0\}$, $f^-(x) = \max\{-f(x), 0\}$ are measurable. Hence $f = f^+ - f^-$ and $|f| = f^+ + f^-$ are measurable.*

   Caution: *$|f|$ is measurable $\nRightarrow$ f is measurable. Consider $f(x) = \begin{cases} 1 & x \in N \\ -1 & x \notin N \end{cases}$, where N is non-measurable.*

10. *$f_k \to f$ pointwisely and $\{f_k(x)\}$ are measurable, then f is measurable.*

Proof. (2): $\{x \in E_1 \cup E_2 | f(x) > t\} = \{x \in E_1 | f(x) > t\} \cup \{x \in E_2 | f(x) > t\}$.

(3): $\{x \in A | f(x) > t\} = \{x \in E | f(x) > t\} \cap A$.

(4): $\{x \in [a, b] | f(x) > t\}$ is one of the following: an interval, a single point or empty set, hence measurable.

(5): $\{x \in E | \Phi(f) > t\} \implies \{x \in E | f(x) \in \Phi^{-1}((t, +\infty))\} = \bigcup_{j=1}^N I_j$ where I_j is open and disjoint and N can be finite or infinite. Hence $\{x \in E | f(x) \in \Phi^{-1}((t, +\infty))\} = \bigcup_{j=1}^N \{x \in E | f(x) \in I_j = (a_j, b_j)\}$ is measurable.

(6): $\{x \in E | f(x) > t\} = f^{-1}((t, +\infty))$ is open in E.

(7):

$$\{x \in E : \lambda f > t\} = \begin{cases} \{x \in E : f > \frac{t}{\lambda}\} & \lambda > 0 \\ \{x \in E : f < \frac{t}{\lambda}\} & \lambda < 0 \end{cases}.$$

$$\{x \in E | f(x) + g(x) > t\} = \bigcup_{r_n \in \mathbb{Q}} \left(\{x \in E | f(x) > r_n\} \cap \{x \in E | g(x) > t - r_n\} \right).$$

$$\{x \in E | f(x)^2 > t\} = \begin{cases} E & t < 0 \\ \{x \in E | f(x) > \sqrt{t}\} \cup \{x \in E | f(x) < -\sqrt{t}\} & t \geqslant 0 \end{cases}.$$

Notice that $f \cdot g = \frac{(f+g)^2 - (f-g)^2}{4}$.

(8):

$$\{x \in E | \sup_{k\to\infty} \{f_k(x)\} > t\} = \bigcup_{n=1}^{\infty} \{x \in E | f_n(x) > t\}.$$

Notice that $\inf_{k\to\infty} f_k = -\sup_{k\to\infty}(-f_k)$, $\overline{\lim}_{k\to\infty} f_k = \inf_{n\geqslant 1} \sup_{k\geqslant n} f_k$ and $\underline{\lim}_{k\to\infty} f_k = \sup_{n\geqslant 1} \inf_{k\geqslant n} f_k$. □

Definition 3.4. We say a property holds **almost everywhere** on E if it holds on E_1 such that $E = E_1 \cup Z$ where Z is null.

Theorem 3.5. *$f = g$ a.e. on E. If f is measurable on E, then g is measurable on E.*

Proof. Let $E = E_1 \cup Z$ such that Z is null and $f(x) = g(x)$ for $x \in E_1$.

$$\{x \in E|g(x) > t\} = \{x \in E_1|f(x) > t\} \cup \{x \in Z|g(x) > t\}.$$

□

The main theorem here is the Approximation Theorem.

Definition 3.6. Suppose E is measurable.

$$f = \sum_{i=1}^{k} a_i \chi_{E_i} \tag{3.1}$$

where E_i are disjoint and $\bigcup_{i=1}^{k} E_i = E$. Then f is called a **simple function** on E.

If furthermore $E_i \in \mathscr{M}$, then f is called **simple measurable function**.

If E_i are rectangles (can be open/closed/half-open, can be of finite or infinite size), we call f a **step function**.

Theorem 3.7 (Approximation Theorem)**.** *1. If $f \geqslant 0$ on E and f is measurable, then there exists simple measurable functions $\{\varphi_k(x)\}$ such that*

$$0 \leqslant \varphi_k(x) \leqslant \varphi_{k+1}(x), k = 1, 2, \cdots \textit{ and } \lim_{k\to\infty} \varphi_k(x) = f(x), \forall x \in E.$$

2. If f is measurable on E, then there exists simple measurable functions $\{\varphi_k(x)\}$ such that $|\varphi_k(x)| \leqslant |f|$ and

$$\lim_{k\to\infty} \varphi_k(x) = f(x), \forall x \in E.$$

Remark: If f is bounded, then in (1) and (2), we have uniformly convergence.

Proof. (1): For every $k \in \mathbb{N}$, denote

$$E_k = \{x \in E|f(x) \geqslant 2^k\}, \quad E_k^j = \{x \in E|\frac{j}{2^k} \leqslant f(x) \leqslant \frac{j+1}{2^k}\}, j = 0, 1, \cdots, 2^{2k} - 1.$$

Take

$$\varphi_k(x) = \sum_{j=0}^{2^{2k}-1} \frac{j}{2^k} \chi_{E_k^j}(x) + 2^k \chi_{E_k}(x). \tag{3.2}$$

We have $0 \leqslant \varphi_k(x) \leqslant \varphi_{k+1}(x) \leqslant f(x)$.

If $f(x_0) = +\infty$, then $\varphi_k(x_0) = 2^k$, then $\varphi_k(x_0) \to f(x_0)$. If $f(x_0) < +\infty$, then there exists k_*, for every $k > k_*$, $|\varphi_k(x_0) - f(x_0)| < \frac{1}{2^k} \to 0$, as $k \to \infty$.

Furthermore, if f is bounded $|f| \leqslant M$, take $2^{k_*} > M$, then for every $x \in E$ we have $|\varphi_k(x) - f(x)| < \frac{1}{2^k}$ for $k > k_*$. Hence we get uniform convergence.

(2): Since f is measurable, f^+ and f^- are measurable. Apply (1) to $f^\pm$, then we get $\varphi_k^{(1)}, \varphi_k^{(2)}$. Then $\varphi_k = \varphi_k^{(1)} - \varphi_k^{(2)}$, we have $|\varphi_k(x)| \leqslant \varphi_k^{(1)}(x) + \varphi_k^{(2)}(x) \leqslant f^+(x) + f^-(x) = |f(x)|$ and $\lim_{k\to\infty} \varphi_k(x) = \lim_{k\to\infty} \varphi_k^{(1)}(x) - \lim_{k\to\infty} \varphi_k^{(2)}(x) = f^+(x) - f^-(x) = f(x)$(Notice we don't have issue with $\infty - \infty$ here). When f is bounded, we have uniform convergence from (1). □

Definition 3.8. The **support set** of $f(x)$ on $E \subseteq \mathbb{R}^n$, denoted by $\text{supp}(f)$, is the closure of

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

Corollary 3.9. *In the Approximation Theorem, we can replace φ_k by $\widetilde{\varphi_k}$ which are simple measurable functions with compact supports.*

Proof. Take $\widetilde{\varphi_k} = \varphi_k \chi_{B(0,k)}$. This is because for every $x_* \in \mathbb{R}^n$, we can find a k_0 large enough such that $x_* \in B(0,k)$ for $k \geqslant k_0$, then $\lim_{k\to\infty} \widetilde{\varphi_k}(x_*) = \lim_{k\to\infty} \varphi_k(x_*) = f(x_*)$. □

3.2 Mode of Convergence

Recall in calculus, we have different notion of convergence/Cauchy sequence.

- $f_n \to f$ pointwise (p.w.) if for every x and $\varepsilon > 0$, $\exists N = N(x)$ such that $\forall n > N$, $|f_n(x) - f(x)| < \varepsilon$.
- $f_n \to f$ uniformly if N is independent of x.
- Now we have third convergence: $f_n \to f$ a.e. on E.
- Almost uniform convergence.
- Convergence in measure.

We have the fact: If $f_n \to f$ a.e. on E and f_n are measurable on E, then f is measurable on E.

Lemma 3.10. *Suppose $\{f_k(x)\}$ are finite a.e. on E and $m(E) < +\infty$. If $f_k(x) \to f(x)$ a.e. on E, then for every $\varepsilon > 0$, denote*

$$E_\varepsilon = \{x \in E | |f_k - f| \geqslant \varepsilon\},$$

we have

$$\lim_{j\to\infty} m(\bigcup_{k\geqslant j} E_k(\varepsilon)) = 0. \tag{3.3}$$

Proof. Notice that $\bigcap_{j=1}^{\infty} \bigcup_{k\geqslant j} E_k(\varepsilon) \subseteq \{x \in E | f_k(x) \nrightarrow f(x)\}$. Because for every $x \in \bigcap_{j=1}^{\infty} \bigcup_{k\geqslant j} E_k(\varepsilon)$, $x \in \bigcup_{k\geqslant j} E_k(\varepsilon)$. Hence we find $k_1 < k_2 < \cdots$ such that $x \in E_{k_i}(\varepsilon)$, i.e., $|f_{k_i}(x) - f(x)| \geqslant \varepsilon$. By the inclusion, $\bigcap_{j=1}^{\infty} \bigcup_{k\geqslant j} E_k(\varepsilon)$ is null. Since $\bigcup_{k\geqslant j} E_k(\varepsilon) \subseteq E$ has finite measure, we have

$$\lim_{j\to\infty} m(\bigcup_{k\geqslant j} E_k(\varepsilon)) = m(\bigcap_{j=1}^{\infty} \bigcup_{k\geqslant j} E_k(\varepsilon)) = 0.$$

□

Theorem 3.11 (Egorov's Theorem)**.** *Suppose $m(E) < +\infty$ and f_k, f are finite a.e. on E. If $f_k \to f$ a.e. on E, then for every $\delta > 0$, there exists $E_\delta \subseteq E$ measurable with $m(E_\delta) \leqslant \delta$ such that $f_k \rightrightarrows f$ on $E \backslash E_\delta$.*

Proof. Take $\varepsilon = \frac{1}{i}$, we have

$$\lim_{j\to\infty} m\left(\bigcup_{k\geqslant j}^{\infty} E_k(\frac{1}{i})\right) = 0.$$

Take j_i such that $m\left(\bigcup_{k\geqslant j_i} E_k(\frac{1}{i})\right) < \frac{\delta}{2^i}$ and $E_\delta = \bigcup_{i=1}^{\infty}\bigcup_{k\geqslant j_i} E_k(\frac{1}{i})$, then $m(E_\delta) \leqslant \delta$.

$$E\backslash E_\delta = \bigcap_{i=1}^{\infty}\bigcap_{k\geqslant j_i}\{x \in E \mid |f_k(x) - f(x)| < \frac{1}{i}\}.$$

For every $\varepsilon > 0$, take i_* such that $\frac{1}{i_*} < \varepsilon$. When $k \geqslant j_{i_*}$, for all $x \in E\backslash E_\delta$,

$$|f_k(x) - f(x)| < \frac{1}{i} < \varepsilon \Longrightarrow f_k(x) \rightrightarrows f \text{ on } E\backslash E_\delta.$$

□

Remark:

(a) Theorem fails if $m(E) = +\infty$. For example, $f_k = \chi_{(0,k)}$ and $f = \chi_{(0,+\infty)}$.

(b) For $m(E) = +\infty$, we reformulate Egorov's Theorem: For every $M > 0$, there exists $E_M \subseteq E$ such that $m(E_M) > M$ and $f_k \rightrightarrows f$ on E_M.

3.3 Relation of Measurable and Continuous Functions

Theorem 3.12 (Lusin's Theorem)**.** *Suppose $f : E \subset \mathbb{R}^n \to \overline{\mathbb{R}}$ is measurable on E and finite a.e. on E. Then for every $\delta > 0$, there exists $F \subseteq E$ measurable such that $m(E\backslash F) < \delta$ and f is continuous on F.*

Corollary 3.13. *Suppose $f : E \subset \mathbb{R}^n \to \overline{\mathbb{R}}$ is measurable on E and finite a.e. on E. Then for every $\delta > 0$, there exists a continuous function $g(x)$ on $\mathbb{R}^n$ such that*

$$m(\{x \in E \mid f(x) \neq g(x)\}) < \delta.$$

Corollary 3.14. *Suppose $f : E \subset \mathbb{R}^n \to \overline{\mathbb{R}}$ is measurable on E and finite a.e. on E. Then there exists a sequence of continuous functions $\{g_k\}$ on $\mathbb{R}^n$ such that*

$$g_k \to f \text{ a.e. on } E.$$

Proof. Take $\delta = \frac{1}{k}$ in Corollary 3.13, we find g_k such that

$$m(\{x \in E \mid f(x) \neq g_k(x)\}) < \frac{1}{k} \Longrightarrow g_k \to f \text{ by measure on } E.$$

Apply Riesz's Theorem, then we can find a subsequence

$$g_{k_j} \to f \text{ a.e. on } E.$$

□

Recall that if f

Example 3.15. Let $F(x)$ be Cantor-Lebesgue function on $[0, 1]$ and $h(x) = x + F(x) : [0, 1] \to [0, 2]$ strictly increasing, continuous and onto, hence h^{-1} exists and is continuous. Take $N \subset h(C)$ non-measurable, then $h^{-1}(N) \subset C$ is null, hence measurable. Take $f = \chi_{h^{-1}((N)}$ measurable and $g = h^{-1}$ continuous. Then $f(g) = \chi_{h^{-1}(N)}(h^{-1}(x)) = \chi_N(x)$ is non-measurable.

Theorem 3.16. *Suppose $T \in C(\mathbb{R}^n, \mathbb{R}^n)$ statisfies that for every null set $Z \subseteq \mathbb{R}^n$, $T(Z)$ is null. Then for every $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is measurable, then $f(T(x))$ is measurable.*

Proof.

$$\{x : f(T(x)) > t\} = \{x : T(x) : f^{-1}((t, +\infty))\} = T^{-1}(K) \cup T^{-1}(Z).$$

□

Corollary 3.17. *If T is linear and nonsingular, then for every measurable function $f : E \to \overline{\mathbb{R}}$, $f(T(x))$ is measurable.*

Proof.

$$m^*(T(E)) = |\det T| m^*(E).$$

□

4 Lebesgue Integral

4.1 Definitions and Properties

We start with non-negative measurable functions. Denote $L^+ = \{f : \mathbb{R}^n \to \overline{\mathbb{R}}$ measurable and $f \geqslant 0\}$ and $L_E^+ = \{f : E \to \overline{\mathbb{R}}$ measurable and $f \geqslant 0\}$.

Definition 4.1. Suppose $f \in L^+$ is a simple measurable function

$$f = \sum_{i=1}^{k} a_i \chi_{A_i}, \quad \bigcup_{i=1}^{k} A_i = \mathbb{R}^n, A_i \cap A_j = \varnothing (i \neq j).$$

Given $E \subseteq \mathbb{R}^n$ measurable, we define

$$\int_E f(x)dx = \sum_{i=1}^{k} a_i m(A_i \cap E). \tag{4.1}$$

(Here we make the assumption $0 \cdot \infty = 0$.)

Definition 4.2. For every $f \in L_E^+$, we define

$$\int_E f(x)\,dx = \sup\left\{\int_E \varphi(x)\,dx \;\middle|\; 0 \leqslant \varphi \leqslant f \text{ on } E,\ \varphi \text{ simple measurable on } E\right\}. \tag{4.2}$$

Proposition 4.3. *Here we consider nonnegative simple measurable functions.*

1. $C\int_E f\mathrm{d}x = \int_E Cf\mathrm{d}x$.
2. $\int_E (f+g)\mathrm{d}x = \int_E f\mathrm{d}x + \int_E g\mathrm{d}x$.
3. $\int_{E\cup F} f\mathrm{d}x = \int_E f\mathrm{d}x + \int_F f\mathrm{d}x$ *if E, F are disjoint measurable sets.*
4. *If E_k is measurable and $\{E_k\}$ is increasing such that $E_k \to E$, then*

$$\int_E f(x)dx = \lim_{k\to\infty} \int_{E_k} f(x)dx.$$

Proof. (2): For $f = \sum_{i=1}^M a_i\chi_{A_i}$ and $g = \sum_{j=1}^N b_j\chi_{B_j}$, we have

$$f+g = \sum_{i=1}^{M}\sum_{j=1}^{N}(a_i+b_j)m(E\cap A_i\cap B_j) = \sum_{i=1}^{M} a_i m(E\cap A_i) + \sum_{j=1}^{N} b_j m(E\cap B_j) = \int_E f\mathrm{d}x + \int_E g\mathrm{d}x.$$

(4):

$$\sum_{i=1}^{k} a_i m(E_k \cap A_i) \to \sum_{i=1}^{k} a_i m(E \cap A_i).$$

□

Proposition 4.4. *Here we only consider functions in L_E^+.*

1. $0 \leqslant f \leqslant g \Longrightarrow 0 \leqslant \int_E f\mathrm{d}x \leqslant \int_E g\mathrm{d}x$.
2. $0 \leqslant f \leqslant F$ *and* $\int_E f\mathrm{d}x < +\infty \Longrightarrow \int_E f\mathrm{d}x < \infty$.

3. If f is bounded and $m(E) < +\infty$, then $\int_E f\mathrm{d}x < +\infty$.

4. If $A \subseteq E$ measurable, then $\int_A f(x)\mathrm{d}x = \int_E f(x)\chi_A(x)\mathrm{d}x$.

5. $f = 0$ a.e. on $E \iff \int_E f\mathrm{d}x = 0$.

6. If $\int_E f\mathrm{d}x < +\infty$, then $f < +\infty$ a.e. on E.

Proof. (4):

$$\begin{aligned}\int_A f(x)\mathrm{d}x &= \sup\left\{\int_A \varphi(x)\mathrm{d}x \,\middle|\, 0 \leqslant \varphi \leqslant f \text{ on } A, \varphi \text{ simple measurable on } A\right\}\\ &= \sup\left\{\int_E \varphi(x)\chi_A(x)\mathrm{d}x \,\middle|\, 0 \leqslant \varphi\chi_A \leqslant f\chi_A \text{ on } A, \varphi \text{ simple measurable on } A\right\}\\ &= \sup\left\{\int_E h\mathrm{d}x \,\middle|\, 0 \leqslant h \leqslant f\chi_A \text{ on } E, h \text{ simple measurable on } E\right\}\\ &= \int_E f(x)\chi_A(x)\mathrm{d}x.\end{aligned}$$

(5): $\Rightarrow$: For every simple measurable function $h \leqslant f$

$\Leftarrow$: Suppose statement is false, i.e. $m(\{x \in E : f(x) \neq 0\}) = \delta > 0$. $E_k = \{x \in E : f(x) > \frac{1}{k}\}$. Then E_k increases to $\{x \in E : f(x) \neq 0\}$. Hence there exists k_0 such that $m(E_{k_0}) > \frac{\delta}{2}$. Then $\int_E f\mathrm{d}x \geqslant \int_{E_{k_0}} f\mathrm{d}x \geqslant \int_{E_{k_0}} \frac{1}{k_0}\mathrm{d}x = \frac{m(E_{k_0})}{k_0} > 0$, contradiction.

(6): $\{x \in E : f = +\infty\} = \bigcap_{k=1}^{\infty}\{x \in E : f(x) > k\}$. $A = \int_E f\mathrm{d}x \geqslant \int_E f\chi_{E_k}\mathrm{d}x \geqslant km(E_k)$. Hence $m(E_k) \leqslant \frac{A}{k} \to 0$ as $k \to \infty$. □

Theorem 4.5 (Beppo Levi Monotone Convergence Theorem)**.** *Given*

$$0 \leqslant f_1 \leqslant f_2 \leqslant \cdots \leqslant f_k \leqslant \cdots$$

on E measurable and $\lim_{k\to\infty} f_k(x) = f(x)$ on E, then

$$\lim_{k\to\infty}\int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x.$$

Proof. Since $f_k \leqslant f$, we have $\int_E f_k\mathrm{d}x \leqslant \int_E f\mathrm{d}x$. Hence

$$\lim_{k\to\infty}\int_E f_k\mathrm{d}x \leqslant \int_E f\mathrm{d}x.$$

WLOG, we can assume $f < +\infty$ on E. For every simple measurable function φ on E such that $0 \leqslant \varphi \leqslant f$ and every $\alpha \in (0,1)$, denote $E_k = \{x \in E : f_k(x) \geqslant \alpha\varphi(x)\}$, then E_k increases to E.

$$\int_E f_k\mathrm{d}x \geqslant \int_{E_k} f_k\mathrm{d}x \geqslant \int_{E_k} \alpha\varphi\mathrm{d}x = \alpha\int_{E_k}\varphi\mathrm{d}x \to \alpha\int_E \varphi\mathrm{d}x \Rightarrow \lim_{k\to\infty}\int_E f_k\mathrm{d}x \geqslant \alpha\int_E\varphi\mathrm{d}x.$$

Hence

$$\lim_{k\to\infty}\int_E f_k\mathrm{d}x \geqslant \alpha\sup_{0\leqslant\varphi\leqslant f}\int_E\varphi\mathrm{d}x = \alpha\int_E f\mathrm{d}x.$$

Let $\alpha \to 1$, we get

$$\lim_{k\to\infty}\int_E f_k\mathrm{d}x \geqslant \int_E f\mathrm{d}x.$$

□

Remark:

1. In Monotone Convergence Theorem, we can consider $0 \leqslant f_1 \leqslant f_2 \leqslant \cdots$ a.e. on E and $f_k \to f$ a.e. on E and the conclusion still holds.

2. In general, without monotonicity a.e., the statement fails. Consider $f_n = \chi_{(n,n+1)}$ or $g_n = n\chi_{(0,\frac{1}{n})}$.

Corollary 4.6. *For $f, g \in L_E^+$ and $\alpha, \beta \geqslant 0$, we have*

$$\int_E \alpha f + \beta g \mathrm{d}x = \alpha \int_E f \mathrm{d}x + \beta \int_E g \mathrm{d}x. \tag{4.3}$$

Proof. Take f_n and g_n simple measurable functions such that $f_n \uparrow f$ and $g_n \uparrow g$. Then apply Monotone Convergence Theorem. □

Corollary 4.7. *For $f_n \in L_E^+$, we have*

$$\int_E \sum_{n=1}^{\infty} f_n \mathrm{d}x = \sum_{n=1}^{\infty} \int_E f_n \mathrm{d}x. \tag{4.4}$$

Proof. Take $S(x) = \sum_{n=1}^{\infty} f_n(x)$ and $S_N(x) = \sum_{n=1}^{N} f_n(x)$, then apply Monotone Convergence Theorem.

$$\lim_{N\to\infty} \sum_{n=1}^{N} \int_E f_n(x)\mathrm{d}x = \lim_{N\to\infty} \int_E S_N(x)\mathrm{d}x = \int_E S(x)\mathrm{d}x.$$

□

Corollary 4.8. *For $E = \bigcup_{k=1}^{\infty} E_k$, where E_k are measurable and disjoint, and $f \in L_E^+$, we have*

$$\int_E f(x)\mathrm{d}x = \sum_{k=1}^{\infty} \int_{E_k} f(x)\mathrm{d}x. \tag{4.5}$$

Proof.

$$\sum_{k=1}^{\infty} f\chi_{E_k} = f\chi_E.$$

□

Corollary 4.9. *Given $f_k \geqslant 0$and $f_k \downarrow f$ a.e. on E measurable, then*

$$\lim_{k\to\infty} \int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x.$$

Proof. Apply the Monotone Convergence Theorem. □

Theorem 4.10 (Fatou's Lemma). *Given $\{f_k\} \subseteq L_E^+$, then*

$$\int_E \varliminf_{k\to\infty} f_k(x)\mathrm{d}x \leqslant \varliminf_{k\to\infty} \int_E f_k(x)\mathrm{d}x. \tag{4.6}$$

Corollary 4.11. *For $\{f_n\}, f \in L^+$ and $f_n \to f$ a.e. on E, then*

$$\int_E f(x)\mathrm{d}x \leqslant \varliminf_{n\to\infty} \int_E f_n(x)\mathrm{d}x.$$

Example 4.12. For $f \in L_E^+$ such that $f < +\infty$ a.e. on E and $m(E) < +\infty$, take $0 \leqslant y_0 < y_1 < y_2 < \cdots < y_n < \cdots < +\infty$ such that $y_{k+1} - y_k \leqslant \delta$. Denote $E_k = \{x \in E : y_k \leqslant f(x) < y_{k+1}\}$, then

$$\int_E f(x)\mathrm{d}x < +\infty \Longleftrightarrow \sum_{k\geqslant 0} y_k m(E_k) < +\infty.$$

Furthermore, under this condition, $\int_E f\mathrm{d}x = \lim_{\delta\to 0}\sum_{k\geqslant 0} y_k m(E_k)$.

Proof.

$$\int_{E_k} y_k\mathrm{d}x \leqslant \int_{E_k} f(x)\mathrm{d}x \leqslant \int_{E_k} y_{k+1}\mathrm{d}x.$$

Hence

$$y_k m(E_k) \leqslant \int_{E_k} f(x)\mathrm{d}x \leqslant (y_{k+1} - y_k)m(E_k) + y_k m(E_k).$$

Summing over k, we get

$$\sum_{k\geqslant 0} y_k m(E_k) \leqslant \sum_{k=1}^{\infty}\int_{E_k} f(x)\mathrm{d}x \leqslant \sum_{k\geqslant 0} y_{k+1} m(E_k).$$

When $\delta \to 0$, $y_{k+1} - y_k \to 0$, we have $\sum_{k\geqslant 0} y_{k+1}m(E_k) - \sum_{k\geqslant 0} y_k m(E_k) = \sum_{k\geqslant 0}(y_{k+1} - y_k)m(E_k) \to 0$. □

4.2 Lebesgue Integral for General Measurabe Functions

Definition 4.13. Given f measurable on E, if $\int_E f^+\mathrm{d}x$ and $\int_E f^-\mathrm{d}x$ at least one is finite, then we define

$$\int_E f\mathrm{d}x = \int_E f^+\mathrm{d}x - \int_E f^-\mathrm{d}x. \tag{4.7}$$

If both are finite, then we say f is **Lebesgue integrable** on E.

Remark:

(a) Denote $L(E) = \{f$ measurable and Lebesgue integrable on $E\}$. Notice that $f \in L(E) \Longleftrightarrow |f| \in L(E)$ and

$$|\int_E f(x)\mathrm{d}x| \leqslant \int_E |f(x)|\mathrm{d}x.$$

We have the following basic properties.

1. f is bounded a.e. on E and $m(E) < +\infty \Longrightarrow f \in L(E)$.

2. $f \in L(E) \Longrightarrow |f| < +\infty$ a.e. on E.

3. $f = 0$ a.e. on $E \iff$
$$\int_E f(x)\mathrm{d}x = 0.$$

4. $|f| \leqslant g$ a.e. on E and $g \in L(E) \implies f \in L(E)$.

5. If $f \in L(\mathbb{R}^n)$, then
$$\lim_{N\to\infty} \int_{|x|\geqslant N} |f(x)|\mathrm{d}x = 0.$$

6. For any $\alpha, \beta \in \mathbb{R}$ and $f, g \in L(E)$, we have
$$\int_E (\alpha f(x) + \beta g(x))\mathrm{d}x = \alpha \int_E f(x)\mathrm{d}x + \beta \int_E g(x)\mathrm{d}x.$$

7. $f \leqslant g$ and $f, g \in L(E) \implies \int_E f(x)\mathrm{d}x \leqslant \int_E g(x)\mathrm{d}x$.

8. If $f, g \in L(E)$, then (a) $f = g$ a.e. on $E \iff$ (b) $\int_E |f(x) - g(x)|\mathrm{d}x = 0 \iff$ (c) For any $Omega \subset E$ measurable, $\int_\Omega f(x)\mathrm{d}x = \int_\Omega g(x)\mathrm{d}x$.

9. $E_k \in \mathscr{M}$ such that E_k disjoint and $\bigcup_{k=1}^\infty E_k = E$. For $f \in L(E)$, we have
$$\int_E f(x)\mathrm{d}x = \sum_{k=1}^\infty \int_{E_k} f(x)\mathrm{d}x.$$

Proof. (5): By Monotone Convergence Theorem,
$$\int_{|x|\geqslant N} |f(x)|\mathrm{d}x = \int_{\mathbb{R}^n} |f(x)|\chi_{B(0,N)^c}(x)\mathrm{d}x \to 0, \quad N \to \infty.$$

(6): First we prove $\int_E \alpha f\mathrm{d}x = \alpha\int_E f\mathrm{d}x$. If $\alpha = 0$, it is trivial. If $\alpha > 0$, then
$$\int_E \alpha f\mathrm{d}x = \int_E \alpha f^+\mathrm{d}x - \int_E \alpha f^-\mathrm{d}x = \alpha\left(\int_E f^+\mathrm{d}x - \int_E f^-\mathrm{d}x\right) = \alpha\int_E f\mathrm{d}x.$$
If $\alpha < 0$, we first consider $\alpha = 1$.
$$\int_E (-f)\,\mathrm{d}x = \int_E (-f)^+\mathrm{d}x - \int_E (-f)^-\mathrm{d}x = \int_E f^-\mathrm{d}x - \int_E f^+\mathrm{d}x = -\int_E f\mathrm{d}x.$$
Hence for general $\alpha < 0$,
$$\int_E \alpha f\mathrm{d}x = \int_E (-|\alpha|f)\mathrm{d}x = -|\alpha|\int_E f\mathrm{d}x = \alpha\int_E f\mathrm{d}x.$$

Next we prove $\int_E (f+g)\mathrm{d}x = \int_E f\mathrm{d}x + \int_E g\mathrm{d}x$. Notice that
$$f+g = (f+g)^+ - (f+g)^- = f^+ + g^+ - f^- - g^- \implies (f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+.$$

(7): $f^+ - f^- \leqslant g^+ - g^- \implies f^+ + g^- \leqslant g^+ + f^- \implies \int_E f^+\mathrm{d}x + \int_E g^-\mathrm{d}x \leqslant \int_E g^+\mathrm{d}x + \int_E f^-\mathrm{d}x \implies \int_E f\mathrm{d}x \leqslant \int_E g\mathrm{d}x$.

(8): (a) $\Rightarrow$ (b) and (b) $\Rightarrow$ (c): trivial. (c) $\Rightarrow$ (a): Suppose statement is false, then $\widetilde{E} = \{f \neq g\}$ is not null $\implies \widetilde{E}_1 = \{f > g\}$ or $\widetilde{E}_2 = \{f < g\}$ has positive measure.

WLOG, we assume it is $\widetilde{E}_1 \implies \widetilde{E}_1 = \bigcup_{k\geqslant 1}\{f-g>\frac{1}{k}\}$. We find $\Omega = \{f-g>\frac{1}{k_0}\}$ for some k_0, which has positive measure. Then $\int_\Omega f(x)\mathrm{d}x - \int_\Omega g(x)\mathrm{d}x \geqslant \frac{1}{k_0}m(\Omega) > 0$, contradiction.

(9): The conclusion holds for f^+ and f^-, hence

$$\begin{aligned}\int_E f(x)\mathrm{d}x &= \int_E f^+(x)\mathrm{d}x - \int_E f^-(x)\mathrm{d}x = \sum_{k=1}^{\infty}\left(\int_{E_k} f^+(x)\mathrm{d}x - \int_{E_k} f^-(x)\mathrm{d}x\right)\\ &= \sum_{k=1}^{\infty}\int_{E_k} f(x)\mathrm{d}x.\end{aligned}$$

□

Example 4.14. $f \in L([a,b])$. If $\int_{[a,c]} f(x)\mathrm{d}x = 0$ for every $c \in [a,b]$, then $f = 0$ a.e. on $[a,b]$.

Proof. Suppose statement is false. $\widetilde{E} = \{x \in [a,b] : f(x) \neq 0\} = \widetilde{E}_1 \cup \widetilde{E}_2 = \{x \in [a,b] : f(x) > 0\} \cup \{x \in [a,b] : f(x) < 0\}$ is not null. Repeat the argument in (8). Take $F \subset \Omega$ closed such that $m(F) > 0$. $G = (a,b)\backslash F$ is open and $G = \bigcup_{i=1}^{\infty}(a_i, b_i)$. Then

$$0 = \int_{[a,b]} f(x)\mathrm{d}x = \int_G f(x)\mathrm{d}x + \int_F f(x)\mathrm{d}x \geqslant \frac{1}{k_0}m(F) + \int_G f(x)\mathrm{d}x \implies \int_G f(x)\mathrm{d}x \neq 0.$$

Hence there exists (α_i, β_i) such that $\int_{(\alpha_i,\beta_i)} f(x)\mathrm{d}x \neq 0$ and at least one of $\int_{[\alpha_i,\beta_i]} f(x)\mathrm{d}x$ and $\int_{[a,\alpha_i]} f(x)\mathrm{d}x$ is not zero, contradiction. □

Theorem 4.15 (Absolute Continuity of Integration)**.** *If $f \in L(E)$, then for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every measurable set Omega $\subseteq E$ with $m(\Omega) < \delta$, we have*

$$\left|\int_\Omega f(x)\mathrm{d}x\right| < \varepsilon.$$

Proof. Take $g = |f| \in L^+(E) \cap L(E)$. then we can find $0 \leqslant \varphi_k \leqslant g$ simple measurable functions such that $\varphi_k \uparrow g$. By Monotone Convergence Theorem, $\int_E \varphi_k(x)\mathrm{d}x \to \int_E g(x)\mathrm{d}x$. In particular, for every $\varepsilon > 0$, we can take k_0 such that $\int_E (g-\varphi_{k_0})(x)\mathrm{d}x < \frac{\varepsilon}{100}$. Suppose $|\varphi_k| \leqslant M$ for some M, hence take $\delta = \frac{\varepsilon}{100M}$, we have

$$\int_\Omega g(x)\mathrm{d}x \leqslant \int_\Omega g(x)\mathrm{d}x \leqslant \frac{\varepsilon}{100} + Mm(\Omega) < \varepsilon.$$

□

Example 4.16. $f \in L(E)$ and

$$0 < A = \int_E f(x)\mathrm{d}x < +\infty.$$

Then for every $B \in (0, A)$, we can find $\Omega \subseteq E$ such that

$$\int_\Omega f(x)\mathrm{d}x = B.$$

Proof.

$$g(t) = \int_{E \cap B(0,t)} f(x)\mathrm{d}x.$$

Then $g(0) = 0$ and $\lim_{t \to +\infty} g(t) = A$. It suffices to show that g is continuous. For every fixed $t \in [0, \infty)$,

$$g(t + \Delta t) - g(t) = \int_{E \cap (B(0,t+\Delta t) \setminus B(0,t))} f(x)\mathrm{d}x \to 0, \quad \Delta t \to 0.$$

□

Lemma 4.17 (Jensen's Inequality)**.** *$\omega(x) \in L^+(E)$ and $\int_E \omega(x)\mathrm{d}x = 1$. $f : E \to [a, b]$ is measurable and $f\omega \in L(E)$. $\varphi : [a, b] \to \mathbb{R}$ is a concave up function. Then*

$$\varphi\left(\int_E f(x)\omega(x)\mathrm{d}x\right) \leqslant \int_E \varphi(f(x))\omega(x)\mathrm{d}x.$$

Remark: If we take $\omega(x) = \frac{1}{m(E)}$, then Jensen's Inequality becomes

$$\varphi\left(\frac{1}{m(E)} \int_E f(x)\mathrm{d}x\right) \leqslant \frac{1}{m(E)} \int_E \varphi(f(x))\mathrm{d}x.$$

Proof. Since $a \leqslant f(x) \leqslant b$, we have $a\omega(x) \leqslant f(x)\omega(x) \leqslant b\omega(x) \Longrightarrow a \leqslant \int_E f(x)\omega(x)\mathrm{d}x \leqslant b$. We denote

$$y_0 = \int_E f(x)\omega(x)\mathrm{d}x.$$

Case 1: $y_0 \in (a, b)$. we can find L such that

$$\varphi(y) \leqslant \varphi(y_0) + L(y - y_0), \quad \forall y \in [a, b].$$

Hence

$$\int_E \varphi(f(x))\omega(x)\mathrm{d}x \geqslant \varphi(y_0) + \int_E L(f(x) - y_0)\omega(x)\mathrm{d}x = \varphi(y_0) = \varphi(\int_E f(x)\omega(x)\mathrm{d}x).$$

Case 2: $y_0 = b$ (or $y_0 = a$). Since $(b - f(x))\omega(x) \geqslant 0$ and

$$\int_E (b - f(x))\omega(x)\mathrm{d}x = b - y_0 = 0 \Longrightarrow f(x)\omega(x) = b\omega(x) \text{ a.e. on } E.$$

Hence

$$\int_E \varphi(f(x))\omega(x)\mathrm{d}x = \int_E \varphi(b)\omega(x)\mathrm{d}x = \varphi(b) = \varphi(\int_E f(x)\omega(x)\mathrm{d}x).$$

□

Proposition 4.18 (Integration under Linear Transformation)**.** *$f \in L(\mathbb{R}^n)$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear nonsingular. Then*

$$\int_{\mathbb{R}^n} f(Tx)\mathrm{d}x = |\det T|^{-1} \int_{\mathbb{R}^n} f(x)\mathrm{d}x.$$

Proof. First we consider $f = \sum_{i=1}^{m} a_i \chi_{A_i}$ simple measurable function. Then $f(Tx) = \sum_{i=1}^{m} a_i \chi_{T^{-1}(A_i)}(x)$ and

$$\int_{\mathbb{R}^n} f(Tx)\mathrm{d}x = \sum_{i=1}^{m} a_i m(T^{-1}(A_i)) = \sum_{i=1}^{m} a_i |\det T|^{-1} m(A_i) = |\det T|^{-1} \int_{\mathbb{R}^n} f(x)\mathrm{d}x.$$

Then by standard procedure, we can prove the conclusion for $f \in L^+$ and then $f \in L(E)$. □

Example 4.19. $f \in L[0,+\infty)$, then

$$\lim_{n\to\infty} f(x+n) = 0, \quad \text{a.e. } x \in [0,+\infty).$$

Proof. We can reduce the problem to $x \in [0,1]$.

$$\begin{aligned}\int_1^{\infty} |f(x)|\mathrm{d}x &= \sum_{n=1}^{\infty} \int_n^{n+1} |f(x)|\mathrm{d}x = \sum_{n=1}^{\infty} \int_0^1 |f(x+n)|\mathrm{d}x \\ &= \int_0^1 \sum_{n=1}^{\infty} |f(x+n)|\mathrm{d}x < +\infty.\end{aligned} \tag{4.8}$$

Hence $\sum_{n=1}^{\infty} |f(x+n)| < +\infty$ a.e. $x \in [0,1] \Longrightarrow \lim_{n\to\infty} f(x+n) = 0$ a.e. $x \in [0,1]$. □

Theorem 4.20 (Dominated Convergence Theorem)**.** *$f_k \in L(E)$ and*

$$\lim_{k\to\infty} f_k(x) = f(x), \quad a.e.\ x \in E.$$

If there exists $F(x) \in L(E)$ such that

$$|f_k(x)| \leqslant F(x), \quad a.e.\ x \in E, k = 1, 2, \cdots,$$

then $f \in L(E)$ and

$$\lim_{k\to\infty} \int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x. \tag{4.9}$$

Proof. $f \in L(E)$. By Fatou's Lemma, we have

$$\int_E \underline{\lim}_{k\to\infty} (F(x) + f_k(x))\,\mathrm{d}x \leqslant \underline{\lim}_{k\to\infty} \int_E (F(x) + f_k(x))\mathrm{d}x,$$

$$\int_E \underline{\lim}_{k\to\infty} (F(x) - f_k(x))\,\mathrm{d}x \leqslant \underline{\lim}_{k\to\infty} \int_E (F(x) - f_k(x))\mathrm{d}x.$$

Hence

$$\begin{aligned}\int_E (F(x) + f(x))\,\mathrm{d}x &\leqslant \int_E F(x)\mathrm{d}x + \underline{\lim}_{k\to\infty} \int_E f_k(x)\mathrm{d}x, \\ \int_E (F(x) - f(x))\,\mathrm{d}x &\leqslant \int_E F(x)\mathrm{d}x - \overline{\lim}_{k\to\infty} \int_E f_k(x)\mathrm{d}x,\end{aligned}$$

$$\Longrightarrow \overline{\lim}_{k\to\infty} \int_E f_k(x)\mathrm{d}x \leqslant \int_E f(x)\mathrm{d}x \leqslant \underline{\lim}_{k\to\infty} \int_E f_k(x)\mathrm{d}x \Longrightarrow \lim_{k\to\infty} \int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x.$$

□

Remark: In fact, since $|f_k - f| \leqslant 2F$ and Dominated Convergence Theorem, we have

$$\lim_{k\to\infty}\int_E |f_k(x) - f(x)|\mathrm{d}x = 0.$$

We also have the convergence in measure version of Dominated Convergence Theorem.

Theorem 4.21. *$f_k \in L(E)$ and*

$$\lim_{k\to\infty} f_k(x) = f(x) \text{ in measure on } E.$$

If there exists $F(x) \in L(E)$ such that

$$|f_k(x)| \leqslant F(x), \quad a.e.\ x \in E, k = 1, 2, \cdots,$$

then $f \in L(E)$ and

$$\lim_{k\to\infty}\int_E f_k(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x. \tag{4.10}$$

Also, we have $\lim_{k\to\infty}\int_E |f_k(x) - f(x)|\mathrm{d}x = 0$.

Proof. $f_{k_j} \to f$ a.e. on E for some subsequence $\{f_{k_j}\}$. By Dominated Convergence Theorem, we have

$$\lim_{j\to\infty}\int_E f_{k_j}(x)\mathrm{d}x = \int_E f(x)\mathrm{d}x.$$

$$\begin{aligned}\int_E |f_k(x) - f(x)|\mathrm{d}x \leqslant & \int_{E\cap\{|x|\geqslant K\}} |f_k(x) - f(x)|\mathrm{d}x \\ & + \int_{E\cap\{|x|<K\}\cap|f_k(x)-f(x)|\geqslant\varepsilon_1} |f_k(x) - f(x)|\mathrm{d}x \\ & + \int_{E\cap\{|x|<K\}\cap|f_k(x)-f(x)|<\varepsilon_1} |f_k(x) - f(x)|\mathrm{d}x\end{aligned}$$

For any $\varepsilon > 0$, take K large enough such that $\int_{E\cap\{|x|\geqslant K\}} |F|\mathrm{d}x \leqslant \frac{\varepsilon}{100}$ and there exists δ sucht that whenever $m(\Omega) < \delta, \int_\Omega |F|\mathrm{d}x \leqslant \frac{\varepsilon}{100}$. For ε_1, there exists N such that for $k \geqslant N$, $m(\{x \in E : |f_k(x) - f(x)| \geqslant \varepsilon_1\}) < \delta$. Hence for $k \geqslant N$,

$$\int_E |f_k(x) - f(x)|\mathrm{d}x < \frac{\varepsilon}{50} + \frac{\varepsilon}{50} + c\varepsilon_1 K^n < \varepsilon, \quad \text{if } \varepsilon_1 \text{ is small enough.}$$

□

Remark: We say $f_k \to f$ in $L^1(E)$ if $\lim_{k\to\infty}\int_E |f_k(x) - f(x)|\mathrm{d}x = 0$.

(a) $f_k \to f$ in $L^1(E) \not\Leftrightarrow f_k \to f$ a.e. on E. Consider $n\chi_{(0,\frac{1}{n})}$ on $E = (0, 1)$ and the 'moving characteristic function with shrinking area'.

(b) $f_k \to f$ in $L^1(E) \not\Leftrightarrow f_k \to f$ a.u. on E. Consider the same examples as in (a).

(c) $f_k \to f$ in $L^1(E) \Longrightarrow f_k \to f$ in measure on E.

Proof. (c): Denote $E_\varepsilon^k = \{x \in E : |f_k(x) - f(x)| \geqslant \varepsilon\}$. Then

$$\int_E |f_k(x) - f(x)|\mathrm{d}x \geqslant \int_{E_\varepsilon^k} |f_k(x) - f(x)|\mathrm{d}x \geqslant \varepsilon m(E_\varepsilon^k)$$

$$\Longrightarrow m(E_\varepsilon^k) \leqslant \frac{1}{\varepsilon}\int_E |f_k(x) - f(x)|\mathrm{d}x \to 0, \quad k \to \infty.$$

□

Corollary 4.22. $f_k \in L(E)$ *and*

$$\int_E \sum_{k=1}^{\infty} |f_k(x)|\mathrm{d}x < +\infty,$$

then $\sum_{k=1}^{\infty} f_k(x)$ *converges a.e.* $x \in E$ *and*

$$\sum_{k=1}^{\infty}\int_E f_k(x)\mathrm{d}x = \int_E \sum_{k=1}^{\infty} f_k(x)\mathrm{d}x. \tag{4.11}$$

Proof. Denote $F(x) = \sum_{k=1}^{\infty} |f_k(x)| < +\infty$ a.e. $x \in E$, which implies that $\sum_{k=1}^{\infty} f_k(x)$ converges a.e. $x \in E$. Denote $S_n(x) = \sum_{k=1}^{n} f_k(x)$, then $|S_n(x)| \leqslant F(x)$ and by Dominated Convergence Theorem, we have

$$\lim_{n\to\infty}\sum_{k=1}^{n}\int_E f_k(x)\mathrm{d}x = \lim_{n\to\infty}\int_E S_n(x)\mathrm{d}x = \int_E \lim_{n\to\infty} S_n(x)\mathrm{d}x = \int_E \sum_{k=1}^{\infty} f_k(x)\mathrm{d}x.$$

□

Corollary 4.23 (Differentiation Theorem)**.** *$f(x, y)$ is defined on $E \times (a, b)$, where $E \in \mathscr{M}$. f is Lebesgue integrable as a function of x for every fixed $y \in (a, b)$ and f is differentiable with respect to y. If there exists $F(x) \in L(E)$ such that*

$$\left|\frac{\partial f}{\partial y}(x, y)\right| \leqslant F(x) \Longrightarrow \frac{\mathrm{d}}{\mathrm{d}y}\int_E f(x, y)\mathrm{d}x = \int_E \frac{\partial f}{\partial y}(x, y)\mathrm{d}x. \tag{4.12}$$

Proof. For every fixed $y \in (a, b)$ and for every $h_k \to 0$ such that $y + h_k \in (a, b)$, we have

$$\lim_{k\to\infty}\frac{f(x, y + h_k) - f(x, y)}{h_k} = \frac{\partial f}{\partial y}(x, y)$$

and by mean value theorem,

$$\left|\frac{f(x, y + h_k) - f(x, y)}{h_k}\right| \leqslant F(x).$$

By Dominated Convergence Theorem, we have

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_E f(x, y)\mathrm{d}x = \lim_{k\to\infty}\int_E \frac{f(x, y + h_k) - f(x, y)}{h_k}\mathrm{d}x = \int_E \frac{\partial f}{\partial y}(x, y)\mathrm{d}x.$$

□

4.3 Lebesgue Integrability and Continuity

Lemma 4.24. *$f \in L(E)$. For any $\varepsilon > 0$, we can find $g \in C_c(\mathbb{R}^n)$ ($C_c(\mathbb{R}^n) = \{$ continuous and compactly supported functions: $\mathbb{R}^n \to \mathbb{R}\}$) such that*

$$\int_E |f(x) - g(x)| \mathrm{d}x < \varepsilon.$$

Proof. We consider φ_k simple measurable functions with compact support such that $|\varphi_k| \leqslant |f|$ and $\varphi_k \to f$. By Dominated Convergence Theorem, we have

$$\int_E |f(x) - \varphi_k(x)| \mathrm{d}x \to 0, \quad k \to \infty.$$

Hence for any ε, we can take one simple measurable function φ with compact support such that

$$\int_E |f(x) - \varphi(x)| \mathrm{d}x < \frac{\varepsilon}{2}.$$

For φ, we can find $g \in C(\mathbb{R}^n)$ such that $m(\{x \in E : g(x) \neq \varphi(x)\}) \leqslant \frac{\varepsilon}{4M}$, where M is the bound of φ. Since φ has compact support, we say $\operatorname{supp}(\varphi) \subset B(0, R)$. We can take a continuous function $\widetilde{g} = gh \in C_c(\mathbb{R}^n)$, where $h = \begin{cases} 1, & |x| \leqslant R \\ 0, & |x| \geqslant 2R \end{cases}$. Then

$$\{x \in E : \varphi(x) \neq \widetilde{g}(x)\} \subseteq \{x \in E : \varphi(x) \neq g(x)\} \Longrightarrow m\left(\{x \in E : \varphi(x) \neq \widetilde{g}(x)\}\right) < \frac{\varepsilon}{4M}.$$

$$\begin{aligned} \int_E |f(x) - \widetilde{g}(x)| \mathrm{d}x &\leqslant \int_E |f(x) - \varphi(x)| \mathrm{d}x + \int_E |\varphi(x) - \widetilde{g}(x)| \mathrm{d}x \\ &< \frac{\varepsilon}{2} + 2Mm(\{x \in E : \varphi(x) \neq \widetilde{g}(x)\}) \leqslant \frac{\varepsilon}{2} + 2M \cdot \frac{\varepsilon}{4M} = \varepsilon. \end{aligned}$$

□

Remark:

1. Notice in the conclusion, $|\varphi| \leqslant M \Longrightarrow |g| \leqslant M$.
2. In particular, given $|f| \leqslant M$ and $f \in L(E)$, for any $\varepsilon > 0$, we can find $g \in C_c(\mathbb{R}^n)$ such that $|g| \leqslant M$ and $\int_E |f(x) - g(x)| \mathrm{d}x < \varepsilon$.

Corollary 4.25. *$f \in L(E)$. We can find $g_k \in C_c(\mathbb{R}^n)$ such that*

$$\lim_{k \to \infty} \int_E |f(x) - g_k(x)| \mathrm{d}x = 0, \quad \lim_{k \to \infty} g_k(x) = f(x) \text{ a.e. } x \in E.$$

Furthermore, if $|f(x)| \leqslant M$, we can require $|g_k(x)| \leqslant M$.

Corollary 4.26. *$f \in L(E)$. We can find φ_k step functions such that*

$$\lim_{k \to \infty} \int_E |f(x) - \varphi_k(x)| \mathrm{d}x = 0, \quad \lim_{k \to \infty} \varphi_k(x) = f(x) \text{ a.e. } x \in E.$$

Furthermore, if $|f(x)| \leqslant M$, we can require $|\varphi_k(x)| \leqslant M$.

Proof. It suffices to prove for any ε, we can find a step function. First, we find $g \in C_c(\mathbb{R}^n)$ such that $\int_E |f(x)-g(x)|\mathrm{d}x < \frac{\varepsilon}{2}$ and $\operatorname{supp}(g) \subset [-N,N]^n$. Since g is uniformly continuous on $[-N,N]^n$, for any $\varepsilon_1 < \frac{\varepsilon}{2(2N)^n}$, there exists $\delta > 0$ such that $|g(x)-g(y)| < \varepsilon_1$, whenever $|x-y| < \delta, x, y \in [-N,N]^n$.

Cut $[-N,N]^n = \bigcup_{j=1}^k I_j$ into a union of disjoint cubes with side length less than $\frac{\delta}{100\sqrt{n}}$. Take $\varphi = \sum_{j=1}^k a_j \chi_{I_j}$, where $a_j = g(x_j)$ for some $x_j \in I_j$. Then

$$\forall x \in [-N,N]^n, \quad |g(x)-\varphi(x)| \leqslant \varepsilon_1 \Longrightarrow \int_E |g(x)-\varphi(x)|\mathrm{d}x < (2N)^n \varepsilon_1 < \frac{\varepsilon}{2}.$$

□

Example 4.27. $f \in L(\mathbb{R}^n)$. If for any $\varphi \in C_c(\mathbb{R}^n)$, we have

$$\int_{\mathbb{R}^n} f(x)\varphi(x)\mathrm{d}x = 0 \Longrightarrow f = 0 \text{ a.e. on } \mathbb{R}^n.$$

Proof. Suppose not true. WLOG, we find a set E such that $f > \varepsilon_0$ on E for some ε_0 and $0 < m(E) < \infty$. Then

$$\int_{\mathbb{R}^n} f(x)\chi_E(x)\mathrm{d}x \geqslant \varepsilon_0 m(E) > 0.$$

Take $g_k \in C_c(\mathbb{R}^n)$ such that $g_k \to \chi_E$ in L^1 and a.e. on E and $|g_k| \leqslant 1$. By Dominated Convergence Theorem, we have

$$\int_{\mathbb{R}^n} f(x)g_k(x)\mathrm{d}x \to \int_{\mathbb{R}^n} f(x)\chi_E(x)\mathrm{d}x > 0,$$

which contradicts the assumption. □

Example 4.28. $f \in L([0,1])$ bounded and

$$\int_{[0,1]} x^n f(x)\mathrm{d}x = 0, \quad n \geqslant 0 \Longrightarrow f = 0 \text{ a.e. on } [0,1].$$

Proof. Assume $|f| \leqslant M$. For any $\varepsilon > 0$, we find $g \in C_c(\mathbb{R})$ such that $\int_{[0,1]} |f(x)-g(x)|\mathrm{d}x < \frac{\varepsilon}{10M}$. For g, we apply Weierstrass Approximation Theorem and find polynomials $p_k(x) \rightrightarrows g$ on $[-N,N] \supset \operatorname{supp}(g)$ such that

$$\int_{[0,1]} |g(x)-p_k(x)|\mathrm{d}x \to 0, \quad k \to \infty.$$

Hence we can find one polynomial $p(x)$ such that

$$\int_{[0,1]} |g(x)-p(x)|\mathrm{d}x < \frac{\varepsilon}{10M} \Longrightarrow$$

$$\int_{[0,1]} f^2(x)\mathrm{d}x \leqslant \int_{[0,1]} |f(x)-g(x)||f(x)|\mathrm{d}x + \int_{[0,1]} |g(x)-p(x)||f(x)|\mathrm{d}x + \int_{[0,1]} p(x)f(x)\mathrm{d}x < M\frac{\varepsilon}{10M} + M\frac{\varepsilon}{10M} < \varepsilon \Longrightarrow \int_{[0,1]} f^2(x)\mathrm{d}x = 0.$$

□

Lemma 4.29 (Riemann-Lebesgue, general version). *g_n are measurable on $[a,b]$ and satisfies*

1. $|g_n| \leqslant M$;
2. *For any* $c \in [a,b]$,

$$\lim_{n\to\infty} \int_{[a,c]} g_n(x)\mathrm{d}x = 0.$$

Then for any $f \in L([a,b])$, we have

$$\lim_{n\to\infty} \int_{[a,b]} f(x)g_n(x)\mathrm{d}x = 0. \tag{4.13}$$

Proof. For any $\varepsilon > 0$, take φ step function such that

$$\int_{[a,b]} |f(x) - \varphi(x)|\mathrm{d}x < \frac{\varepsilon}{2(M+1)}.$$

$$\begin{aligned}\left|\int_{[a,b]} f(x)g_n(x)\mathrm{d}x\right| &\leqslant \int_{[a,b]} |f(x)-\varphi(x)||g_n(x)|\mathrm{d}x + \left|\int_{[a,b]} \varphi(x)g_n(x)\mathrm{d}x\right| \\ &\leqslant M\int_{[a,b]} |f(x)-\varphi(x)|\mathrm{d}x + \left|\sum_{j=1}^{k} \lambda_j \int_{[\alpha_j,\beta_j]} g_n(x)\mathrm{d}x\right| \\ &< M\frac{\varepsilon}{2(M+1)} + \frac{\varepsilon}{2} < \varepsilon, \qquad \text{for } n \text{ large enough.}\end{aligned}$$

□

Lemma 4.30 (Riemann-Lebesgue, classical version). *$f \in L(\mathbb{R})$.*

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x)\cos nx\mathrm{d}x = 0, \quad \lim_{n\to\infty} \int_{\mathbb{R}} f(x)\sin nx\mathrm{d}x = 0. \tag{4.14}$$

Proof. For any ε, take φ step function with compact support. Then we have

$$\begin{aligned}\left|\int_{\mathbb{R}} f(x)\cos nx\mathrm{d}x\right| &\leqslant \int_{\mathbb{R}} |f(x)-\varphi(x)||\cos nx|\mathrm{d}x + \left|\int_{\mathbb{R}} \varphi(x)\cos nx\mathrm{d}x\right| \\ &\leqslant \int_{\mathbb{R}} |f(x)-\varphi(x)|\mathrm{d}x + \left|\sum_{j=1}^{k} \lambda_j \frac{\sin nx}{n}\Big|_{\alpha_j}^{\beta_j}\right| < \varepsilon.\end{aligned}$$

□

Example 4.31. $\lambda_n \to 0$. Denote

$$A = \{x \in \mathbb{R} : \lim_{n\to\infty} \sin\lambda_n x \text{ exists }\} \Longrightarrow m(A) = 0.$$

Proof. For any $B \subset \mathbb{R}$ bounded and measurable, define $f(x) = \lim_{n\to\infty} \chi_A(x)\sin\lambda_n x$. By Dominated Convergence Theorem, we have

$$\int_B f(x)\mathrm{d}x = \lim_{n\to\infty} \int_B \chi_A(x)\sin\lambda_n x\mathrm{d}x = 0.$$

By Riemann-Lebesgue Lemma, we have

$$\lim_{n\to\infty}\int_{[-N,N]}\chi_{A\cap B}(x)\sin\lambda_n x\mathrm{d}x \to 0 \Longrightarrow \int_B f(x)\mathrm{d}x = 0 \Longrightarrow f(x) = 0 \text{ a.e. } x\in B.$$

By Dominated Convergence Theorem again, we have

$$\begin{aligned}0 &= \int_B f^2(x)\mathrm{d}x = \lim_{n\to\infty}\int_B (\chi_A(x)\sin\lambda_n x)^2\,\mathrm{d}x = \lim_{n\to\infty}\int_B \chi_A(x)\frac{1-\cos 2\lambda_n x}{2}\mathrm{d}x\\ &= \frac{1}{2}m(A\cap B) - \lim_{n\to\infty}\frac{1}{2}\int_B \chi_A(x)\cos 2\lambda_n x\mathrm{d}x = \frac{1}{2}m(A\cap B).\end{aligned}$$

Since B is arbitrary, we have $m(A) = 0$. □

Lemma 4.32. *If $f\in L(\mathbb{R}^n)$, then*

$$\lim_{h\to 0}\int_{\mathbb{R}^n}|f(x+h)-f(x)|\mathrm{d}x = 0. \tag{4.15}$$

Proof. For every $\varepsilon > 0$, take $g\in C_c(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n}|f(x)-g(x)|\mathrm{d}x < \frac{\varepsilon}{3}$. Suppose $\mathrm{supp} g\subseteq[-N,N]^n$. Since g is uniformly continuous in $[-N,N]^n$,

$$\int_{\mathbb{R}^n}|g(x+h)-g(x)|\mathrm{d}x = \int_{[-N,N]^n}|g(x+h)-g(x)|\mathrm{d}x < \varepsilon_1(2N)^n < \frac{\varepsilon}{3},\quad h<\delta.$$

$$\int_{\mathbb{R}^n}|f(x+h)-f(x)|\mathrm{d}x \leqslant \int_{\mathbb{R}^n}|f(x+h)-g(x+h)|\mathrm{d}x+\int_{\mathbb{R}^n}|g(x+h)-g(x)|\mathrm{d}x+\int_{\mathbb{R}^n}|g(x)-f(x)|\mathrm{d}x < \varepsilon.$$

□

Corollary 4.33. *If $E\in\mathscr{M}$ and $m(E) < +\infty$, then $m(E\cap E+\{h\})\to m(E)$ as $h\to 0$.*

Proof. $m(E) = \int\chi_E\mathrm{d}x$ and $m(E\cap E+\{h\}) = \int\chi_E(x)\chi_{E+\{h\}}(x)\mathrm{d}x$. Hence

$$\begin{aligned}|m(E\cap E+\{h\}) - m(E)| &= \left|\int\chi_E(x)\chi_{E+\{h\}}(x)\mathrm{d}x - \int\chi_E^2(x)\mathrm{d}x\right|\\ &= \left|\int\chi_E(x)(\chi_{E+\{h\}}(x)-\chi_E(x))\mathrm{d}x\right|\\ &\leqslant \int|\chi_E(x)-\chi_E(x+h)|\mathrm{d}x \to 0,\quad h\to 0.\end{aligned}$$

□

4.4 Lebesgue Integral and Riemann Integral

Theorem 4.34. *Suppose f bounded on $[a,b]$. Then $f\in R[a,b]\Longleftrightarrow$ discontinuous points of f have Lebesgue measure zero.*

Theorem 4.35. *$f\in R[a,b]\Longrightarrow f\in L([a,b])$ and the two integerals take the same value.*

Recall the notations and terminologies of Riemann integral. For Riemann integral, we have a partition $\Delta : a = x_0 < x_1 < \cdots < x_n = b$ and $\Delta x_i = x_i - x_{i-1}, \Delta = \max\{\Delta x_i\}, \omega_i = \sup_{x,y\in[x_{i-1},x_i]}|f(x)-f(y)|$. We know that $f\in R[a,b]\Longleftrightarrow\sum\omega_i\Delta x_i\to 0$ as $\Delta\to 0$. $\omega_f(x) = \lim_{\delta\to 0}\sup_{y,z\in B_\delta(x)}|f(z)-f(y)|$.

Proof. Denote $E = \{\text{discontinuous points of } f\} = \{x|\omega_f(x) > 0\}$. First we show that $m(E) \Longrightarrow f \notin R[a,b]$. $E = \bigcup_{k\geqslant 1}$ where $E_k = \{x|\omega_f(x) \geqslant \frac{1}{k}\}$. Since $m(E) > 0$, there exists k_0 such that $m(E_{k_0}) > 0$. For every partition Δ, we have

$$\sum \omega_i \Delta x_i \geqslant \sum_{(x_{i-1},x_i)\cap E_{k_0}\neq\varnothing} \omega_i \Delta x_i \geqslant \frac{1}{k_0} \sum_{[x_{i-1},x_i]\cap E_{k_0}\neq\varnothing} \Delta x_i \geqslant \frac{1}{k_0} m(E_{k_0}) > 0.$$

Next, we show that $m(E) = 0 \Longrightarrow f \in R[a,b]$. Notice that $m([a,b]\backslash E) = m([a,b])$. For every $\varepsilon > 0$, take $F \subseteq [a,b]\backslash E$ closed such that $m(([a,b]\backslash E)\backslash F) < \varepsilon_1$. For every $z \in F$, f is continuous at z. For every ε_2, there exists $\delta(z)$ such that $|f(y) - f(z)| < \frac{\varepsilon_2}{2}$ when $|y - z| \leqslant \delta(z)$. In particular, for every $x, y \in B(z, \delta(z))$, we have $|f(x) - f(y)| < \varepsilon_2$. We cover F by $\bigcup_{z\in F} B(z, \frac{1}{10}\delta(z))$. Since F is compact, there is a finite subcover $F \subseteq \bigcup_{j=1}^N B(z_j, \frac{1}{10}\delta(z_j))$. If we denote $|f| \leqslant M$, then for any partition Δ,

$$\sum \omega_i \Delta x_i = \sum_{[x_{i-1},x_i]\cap F\neq\varnothing} \omega_i \Delta x_i + \sum_{[x_{i-1},x_i]\cap F=\varnothing} \omega_i \Delta x_i \leqslant \varepsilon_2(b-a) + 2M\varepsilon_1.$$

Take $\varepsilon_1 = \frac{\varepsilon}{4M}$ and $\varepsilon_2 = \frac{\varepsilon}{2(b-a)}$, we have $\sum \omega_i \Delta x_i < \varepsilon$. Hence $f \in R[a,b]$. □

Proof. Since $f \in R[a,b]$, f is bounded. We show that f is measurable. $[a,b] = E \cup ([a,b]\backslash E) = E \cup Z \cup F_k$, where E is the set of discontinuous points of f, Z is a null set and F_k .

For every $t \in \mathbb{R}$, $\{x \in [a,b] : f(x) > t\} = (\{x \in E : f(x) > t\}) \cup (\{x \in Z : f(x) > t\}) \cup (\bigcup_{k\geqslant 1}\{x \in F_k : f(x) > t\})$ is measurable. Hence f and $|f|$ is measurable and $|f|$ is Lebesgue integrable.

$$\sum_{i=1}^n m_i \Delta x_i \leqslant \int_{[a,b]} f \mathrm{d}x = \sum_{i=1}^n \int_{[x_{i-1},x_i]} f \mathrm{d}x \leqslant \sum_{i=1}^n M_i \Delta x_i, \quad m_i = \inf_{[x_{i-1},x_i]} f, M_i = \sup_{[x_{i-1},x_i]} f.$$

Since $f \in R[a,b]$, we have $\sum_{i=1}^n M_i \Delta x_i \to \int_a^b f(x)\mathrm{d}x$ and $\sum_{i=1}^n m_i \Delta x_i \to \int_a^b f(x)\mathrm{d}x$ as $\Delta \to 0$. Hence the two integrals are equal. □

Remark: In the future, we will not distinguish the notations of Riemann integral and Lebesgue integral.

Caution: The above discussion fails for singular integrals.

Example 4.36. Take $f(x) = \begin{cases} \frac{\sin x}{x}, & x \in (0, +\infty) \\ 0, & x = 0 \end{cases}$. From calculus, we know $f \in R[0, +\infty)$ and $\int_0^\infty \frac{\sin x}{x} \mathrm{d}x = \frac{\pi}{2}$. However, $f \notin L[0, +\infty)$. Because $\int_0^\infty \left|\frac{\sin x}{x}\right| \mathrm{d}x = \sum_{n\geqslant 0} \int_{n\pi}^{(n+1)\pi} \frac{|\sin y|}{n\pi + y} \mathrm{d}y \geqslant \sum_{n\geqslant 0} \frac{2}{(n+1)\pi} = +\infty$.

Example 4.37. $h(x) = \frac{1}{x}\sin\frac{1}{x}$ but $h \notin L(0,1)$. $\int_0^1 \frac{1}{x}\sin\frac{1}{x}\mathrm{d}x = \int_1^\infty y \sin y \mathrm{d}\left(\frac{1}{y}\right) = \int_1^\infty \frac{\sin y}{y}\mathrm{d}y$.

Lemma 4.38. *Suppose $E_k \uparrow E$ and $f \in L(E_k)$. If $\lim \int_{E_k} |f|\mathrm{d}x < +\infty$, then $f \in L(E)$ and $\int_E f\mathrm{d}x = \lim_{k\to\infty} \int_{E_k} f\mathrm{d}x$.*

Proof. $g_k(x) = |f(x)|\chi_{E_k} \uparrow |f(x)|$ on E. By Monotone Convergence Theorem, $\int_{E_k} |f|\mathrm{d}x = \int_E |g_k|\mathrm{d}x \to \int |f|\mathrm{d}x \Longrightarrow f \in L(E)$. Now $h_k = f\chi_{E_k} \to f$ on E. Since $|h_k| \leqslant |f|$, by Dominated Convergence Theorem, we have $\int h_k \mathrm{d}x = \int_{E_k} f\mathrm{d}x \to \int f\mathrm{d}x$. □

Remark: $|f| \in R[0,\infty) \Longrightarrow f \in L[0,\infty)$.

4.5 Tonelli and Fubini Theorems

Suppose $f(x,y)$ with $x \in \mathbb{R}^p$, $y \in \mathbb{R}^q$ and $p+q=n$. In calculus, we face the problem for Riemann integrals: whether $\iint_{\mathbb{R}^n} f(x,y)\mathrm{d}x\mathrm{d}y$, $\int_{\mathbb{R}^q}\left(\int_{\mathbb{R}^p} f(x,y)\mathrm{d}x\right)\mathrm{d}y$ and $\int_{\mathbb{R}^p}\left(\int_{\mathbb{R}^q} f(x,y)\mathrm{d}y\right)\mathrm{d}x$ are equal. We know that if $|f| \in R(\mathbb{R}^n)$, then they are well-defined and equal.

Now we want to extend the discussion to Lebesgue integrals.

Theorem 4.39 (Tonelli). *Suppose $f \in L^+(\mathbb{R}^n)$, then we have*

1. *For a.e. $y \in \mathbb{R}^q$, $f(x,y) \in L^+_x(\mathbb{R}^p)$ and for a.e. $x \in \mathbb{R}^p$, $f(x,y)$ is in $L^+(\mathbb{R}^q)$.*
2. *For a.e. $y \in \mathbb{R}^q$, $h(y) = \int_{\mathbb{R}^p} f(x,y)\mathrm{d}x$ and for a.e. $x \in \mathbb{R}^p$, $g(x) = \int_{\mathbb{R}^q} f(x,y)\mathrm{d}y$ are both in $L^+(\mathbb{R}^q)$ and $L^+(\mathbb{R}^p)$ respectively.*
3. $\iint$

Theorem 4.40 (Fubini).

The proof goes by considering $\mathscr{F} \subseteq L^+(\mathbb{R}^n)$ satisfies

(A)

we will study the properties of $\mathscr{F}$ and eventually prove $\mathscr{F} = L^+(\mathbb{R}^n)$. Step 1: Define $\mathscr{F}$.

Step 2:

Lemma 4.41.
1. *$f \in \mathscr{F} \implies af \in \mathscr{F}$ for every $a \geqslant 0$.*
2. *$f, g \in \mathscr{F} \implies f+g \in \mathscr{F}$.*
3. *$f, g \in \mathscr{F}, f-g \geqslant 0$ and $g \in L(\mathbb{R}^n) \implies f-g \in \mathscr{F}$.*
4. *$f_k \in \mathscr{F}, f_k \uparrow f$ on $\mathbb{R}^n \implies f \in \mathscr{F}$.*
5. *$f_k \in \mathscr{F}, f_k \downarrow f$ and $f_1 \in L(\mathbb{R}^n) \implies f \in \mathscr{F}$.*

Proof. (1)(2): follows from definition.

(3):

□

Step3: $\mathscr{F} = L^+(\mathbb{R}^n)$. We reduce the proof to consider only characteristic functions because of (1)(2)(4).

Now we consider $E \in \mathscr{M}$, starting with single case to the general case.

Case 1: $E = I_1 \times I_2$, where $I_1 \subset \mathbb{R}^p, I_2 \subset \mathbb{R}^q$ are both rectangles.

Case 2:

Case 3:

Case 4: E is null. $E \subset G_k$, where G_k is open and $m(G_k) < \frac{1}{k}$. Then $E \subset H = \bigcap_{k=1}^{\infty} G_k$. We verify (A)(B)(C) for H and then for E. Use Case 4 for H.

Case 5: $E \in \mathscr{M}$. Then $E = G \backslash Z$, where G is a G_δ set and $\chi_E = \chi_G - \chi_Z \geqslant 0$. Since $\chi_G, \chi_Z \in \mathscr{F}, \chi_z \in L(\mathbb{R}^n)$, we have $\chi_E \in \mathscr{F}$.

Step 5: Proof of Fubini Theorem. $f \in L(\mathbb{R}^n) \implies f^\pm \in L^+(\mathbb{R}^n) \cap L(\mathbb{R}^n)$.

Remark:

1. Tonelli Theorem fails for Riemann integrals. Consider $E \subset [0,1]^2$ such that dense, countable and intersects with horizontal and vertical lines at most one point. Take $f(x,y) = \chi_E(x,y)$, then $f \notin R([0,1]^2)$.

$$\int_0^1 f(x,y)\mathrm{d}x = 0 = \int_0^1 f(x,y)\mathrm{d}y$$

$$\implies \int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}x\right)\mathrm{d}y = 0 = \int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}y\right)\mathrm{d}x = 0.$$

Let's construct E. Denote $\Omega = \mathbb{Q} \cap [0,1]^2 = \{(x_i,y_i)\}_{i=1}^{\infty}$. Take $(\widetilde{x}_1,\widetilde{y}_1) = (x_1,y_1)$ and take $(\widetilde{x}_2,\widetilde{y}_2) \in \Omega \cap B\left((x_2,y_2),\frac{1}{2}\right)$ such taht it does not lie on the horizontal and vertical lines decided by $(\widetilde{x}_1,\widetilde{y}_1)$. Generally, take $(\widetilde{x}_n,\widetilde{y}_n) \in \Omega \cap B\left((x_n,y_n),\frac{1}{n}\right)$ such that it does not lie on the horizontal and vertical lines through $(\widetilde{x}_i,\widetilde{y}_i), i = 1,2,\cdots,n-1$.

$E = \{(\widetilde{x}_n,\widetilde{y}_n)\}_{n=1}^{\infty}$ is dense. That's because for any $(x,y) \in [0,1]^2$ and any $\varepsilon > 0$, take $n_0 > \frac{2}{\varepsilon}$.

2. Even if the iterated integral exists and equal, it is not necessarily true that $f \in L(\mathbb{R}^n)$. Sierpinski construct non-measurable set in $\mathbb{R}^2$ such that the set intersects with horizontal/vertical lines at most twice, then χ_E will be our example.

3. There exists examples of non integerable functions such that has iterated integrals but not equal. Consider $f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2}$ on $[0,1]^2$. We have

$$\int_0^1 f(x,y)\mathrm{d}y = \left(\frac{y}{x^2+y^2}\right)\Big|_0^1 = \frac{1}{x^2+1} \implies \int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}y\right)\mathrm{d}x = \frac{\pi}{4},$$

$$\int_0^1 f(x,y)\mathrm{d}x = \left(-\frac{x}{x^2+y^2}\right)\Big|_0^1 = -\frac{1}{y^2+1} \implies \int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}x\right)\mathrm{d}y = -\frac{\pi}{4}.$$

Proposition 4.42. *For any $E \in \mathscr{M}(\mathbb{R}^n)$, $f(x,y) = \chi_E(x,y) \in L^+(\mathbb{R}^n)$. Then for a.e. $y \in \mathbb{R}^q$, $\chi_E(x,y) \in L_x^+(\mathbb{R}^p)$ and for a.e. $x \in \mathbb{R}^p$, $\chi_E(x,y) \in L_y^+(\mathbb{R}^q)$.*

Proof. $E^y = \{x \in \mathbb{R}^p : (x,y) \in E\}$ and $E_x = \{y \in \mathbb{R}^q : (x,y) \in E\}$. Fix x, $\chi_E(x,y) = \chi_{E_x}(y)$. Fix y, $\chi_E(x,y) = \chi_{E^y}(x)$.

$$m(E_x) = \int_{\mathbb{R}^q} \chi_E(x,y)\mathrm{d}y \in L^+(\mathbb{R}^p), \quad m(E^y) = \int_{\mathbb{R}^p} \chi_E(x,y)\mathrm{d}x \in L^+(\mathbb{R}^q).$$

$$m(E) = \iint\limits_{\mathbb{R}^n} \chi_E(x)\mathrm{d}x = \int_{\mathbb{R}^p} m(E_x)\mathrm{d}x = \int_{\mathbb{R}^q} m(E^y)\mathrm{d}y < +\infty.$$

□

Remark: Converse is not true. Consider Sierpinski's example. E_x, E^y are sets of three kinds, $\varnothing$, one points and two points. Hence E_x, E^y are measurable for all x,y. However, E is non-measurable.

We provide an easier case. $E = [0,1] \times N$, where N is non-measurable set. $E^y = \{x \in \mathbb{R} : (x,y) \in [0,1] \times N\} = \begin{cases} [0,1], & y \in N \\ \varnothing, & y \notin N \end{cases}$, but E is not measurable, since $E_x = N$ non-measurable.

Lemma 4.43. $E_1 \in \mathscr{M}(\mathbb{R}^p), E_2 \in \mathscr{M}(\mathbb{R}^q)$. *Then*

$$E_1 \times E_2 \in \mathscr{M}(\mathbb{R}^n), \quad m(E_1 \times E_2) = m(E_1)m(E_2).$$

Proof. $E_1 = \left(\bigcup_{i=1}^{\infty} F_i\right) \cup Z, E_2 = \left(\bigcup_{j=1}^{\infty} K_j\right) \cup W$, where F_i, K_j are closed sets and Z, W are null. Then $E_1 \times E_2$ is a countable union of sets of the type $A \times B$, where A, B are closed or null.

Case 1: A, B are closed. Then $A \times B$ is closed in $\mathbb{R}^n$, hence measurable.

Case 2: $m(A) < \infty$ and $m(B) = 0$. Find L-cover for A, B such that for any ε,

$$A \subset \bigcup_{j=1}^{\infty} I_j, \quad \sum_{j=1}^{\infty} |I_j| < m(A) + \varepsilon, \quad B \subset \bigcup_{k=1}^{\infty} J_k, \quad \sum_{k=1}^{\infty} |J_k| < \varepsilon.$$

$$A \times B \subset \bigcup_{j,k=1}^{\infty} I_j \times J_k \Longrightarrow m^*(A \times B) \leqslant \sum_{j,k=1}^{\infty} |I_j||J_k| \leqslant (m(A) + 1)\varepsilon.$$

Notice

$$\chi_{E_1 \times E_2}(x, y) = \chi_{E_1}(x) \cdot \chi_{E_2}(y).$$

By Tonelli Theorem, we have the conclusion. □

Now we introduce convolution.

Definition 4.44. f, g are measurable functions on $\mathbb{R}^n$. If

$$\int_{\mathbb{R}^n} (f(x-y)g(y)) \,\mathrm{d}y$$

exists for a.e. $x \in \mathbb{R}^n$, then we define the **convolution** of f and g as

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\mathrm{d}y. \tag{4.16}$$

Remark: $f * g = g * f$.

Theorem 4.45. $f, g \in L(\mathbb{R}^n) \Longrightarrow f * g$ *exists for a.e.* x *and* $f * g \in L(\mathbb{R}^n)$. *Furthermore,*

$$\int_{\mathbb{R}^n} |(f * g)(x)|\mathrm{d}x \leqslant \left(\int_{\mathbb{R}^n} |f(x)|\mathrm{d}x\right)\left(\int_{\mathbb{R}^n} |g(y)|\mathrm{d}y\right). \tag{4.17}$$

Proof. $|f(x-y)g(y)| = F(x, y) \in L^+(\mathbb{R}^{2n})$. By Tonelli Theorem,

$$\iint_{\mathbb{R}^{2n}} F(x, y)\mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\mathrm{d}x\right)\mathrm{d}y = \int_{\mathbb{R}^n} |f(x)|\mathrm{d}x \int_{\mathbb{R}^n} |g(y)|\mathrm{d}y < +\infty.$$

Hence for a.e. x,

$$\int_{\mathbb{R}^n} |f(x-y)g(y)|\mathrm{d}y < +\infty, \quad \int_{\mathbb{R}^n} |f(x-y)g(y)|\mathrm{d}y \in L(\mathbb{R}^n).$$

$$\left|\int_{\mathbb{R}^n} f(x-y)g(y)\mathrm{d}y\right| \leqslant \int_{\mathbb{R}^n} |f(x-y)g(y)|\mathrm{d}y < +\infty \quad \text{a.e. } x.$$

Also $f * g \in L(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} |(f * g)(x)|\mathrm{d}x \leqslant \iint_{\mathbb{R}^{2n}} |f(x-y)g(y)|\mathrm{d}x\mathrm{d}y = \left(\int_{\mathbb{R}^n} |f(x)|\mathrm{d}x\right)\left(\int_{\mathbb{R}^n} |g(y)|\mathrm{d}y\right).$$

□

Lemma 4.46. *$f \in L(\mathbb{R}^n)$ and g is bounded and measurable. Then $F(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\mathrm{d}y$ is uniformly continuous.*

Proof. Assume $|g| \leqslant M$. Since $|f(x-y)g(y)| \leqslant M|f(x-y)|$, $F(x)$ is well-defined for every x.

$$\begin{aligned}|F(x+h)-F(x)| &\leqslant \int_{\mathbb{R}^n} |f(x+h-y)-f(x-y)||g(y)|\mathrm{d}y \\ &\leqslant M\int_{\mathbb{R}^n} |f(x+h-y)-f(x-y)|\mathrm{d}y \to 0, \quad h \to 0.\end{aligned}$$

□

5 Lebesgue Differentiation Theory

Recall we have the following facts for Riemann integral:

1. If $f \in R[a,b]$ and f is continuous at $x_0 \in [a,b]$, then
$$F(x) = \int_a^x f(t)\mathrm{d}t \in C[a,b]$$
and F is differentiable at x_0 with $F'(x_0) = f(x_0)$.

2. f is differentiable on $[a,b]$ and
$$f' \in R[a,b] \Longrightarrow \int_a^x f'(t)\mathrm{d}t = f(b) - f(a).$$

An easier case is $f \in C^1[a,b]$, which is the Fundamental Theorem of Calculus. In our case, for any partition $\Delta : a = x_0 < x_1 < \cdots < x_n = x$ of $[a,x]$, we have

$$f(x) - f(a) = \sum_{i=1}^n (f(x_i) - f(x_{i-1})) = \sum_{i=1}^n f'(\xi_i)\Delta x_i \to \int_a^x f'(t)\mathrm{d}t, \quad \Delta \to 0.$$

Now we want to generalize the above results to Lebesgue integral.

5.1 Lebesgue Differentiation Theorem

$$F(x) = \int_a^x f(t)\mathrm{d}t, \quad f \in L([a,b]).$$
$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(t)\mathrm{d}t = \frac{1}{h}\int_0^h f(x+t)\mathrm{d}t.$$

We can generalize the above discussion to

$$\frac{1}{|B(x,\delta)|}\int_{B(x,\delta)} f(t)\mathrm{d}t.$$

Or more generally for $x \in B_r$ and $|B_r| \to 0$ as $r \to 0$,

$$\frac{1}{|B_r|}\int_{B_r} f(t)\mathrm{d}t.$$

Theorem 5.1. $f \in L^1(\mathbb{R}^n)$. *Then for a.e.* $x \in \mathbb{R}^n$,

$$\lim_{\delta \to 0} \frac{1}{|B(x,\delta)|}\int_{B(x,\delta)} f(y)\mathrm{d}y = f(x). \tag{5.1}$$

Remark:

1. $L^1_{\text{loc}}(\mathbb{R}^n) = \{f :$ measurable and $f \in L(K)$ for every compact $K \subset \mathbb{R}^n\}$. The above theorem also holds for $f \in L^1_{\text{loc}}(\mathbb{R}^n)$.

2. If $f \in C(\mathbb{R}^n)$, the result is trivial. When δ is small enough,

$$\left|\frac{1}{|B(x,\delta)|}\int_{B(x,\delta)}(f(y)-f(x))\,\mathrm{d}y\right| \leqslant \varepsilon.$$

Definition 5.2. The **Hardy-Littlewood maximal function** of $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is defined as

$$Mf(x) = \sup_{r>0}\frac{1}{|B(x,r)|}\int_{B(x,r)}|f(y)|\mathrm{d}y. \tag{5.2}$$

Lemma 5.3.

$$A_r f(x) = \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)\mathrm{d}y$$

is a continuous function for $(x,r) \in \mathbb{R}^n \times \mathbb{R}_+$.

Proof. For any $(x_0, r_0) \in \mathbb{R}^n \times \mathbb{R}_+$, consider any $(x_k, y_k) \to (x_0, r_0)$. We have

□

Lemma 5.4. *Mf is measurable.*

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

$$\{x \in \mathbb{R} : Mf(x) > t\} = \bigcup_{r\in\mathbb{Q}^+}\{x \in \mathbb{R} : A_r|f|(x) > t\}.$$

□

Remark: Even if $f \in L(\mathbb{R}^n)$, usually $Mf \notin L(\mathbb{R}^n)$.

Lemma 5.5 (Vitali Covering Lemma, Version 1)**.** *$B_1, \cdots, B_N$ are finite many open balls in $\mathbb{R}^n$. Then there exists a subcollection of disjoint balls $B_{i_1}, B_{i_2}, \cdots, B_{i_k}$ such that*

$$\sum_{l=1}^{k} m(B_{i_l}) \geqslant \frac{1}{3^n} m\left(\bigcup_{i=1}^{N} B_i\right).$$

Proof. Notice that if $B(x_1,r_1) \cap B(x_2,r_2) \neq \varnothing$ and $r_1 > r_2$, then $B(x_2,r_2) \subset B(x_1,3r_1)$.

Take B_{i_1} to be the ball with the largest radius and kick out all balls that intersect with B_{i_1}. Repeat the process until no balls are left, then we get disjoint balls $B_{i_1}, B_{i_2}, \cdots, B_{i_k}$ such that

$$\bigcup_{i=1}^{N} B_i \subset \bigcup_{j=1}^{k} 3B_{i_j} \Longrightarrow m\left(\bigcup_{i=1}^{N} B_i\right) \leqslant \sum_{j=1}^{k} m(3B_{i_j}) = 3^n \sum_{j=1}^{k} m(B_{i_j}).$$

□

Lemma 5.6 (Vitali Covering Lemma, Version 2)**.** *$\{B_\alpha, \alpha \in \Lambda\}$ is a collection of open balls. For any $c < m(\bigcup_{\alpha\in\Lambda} B_\alpha)$, we can find finite disjoint subcollection $B_{i_1}, B_{i_2}, \cdots, B_{i_L}$ such that*

$$m\left(\bigcup_{i=1}^{L} B_{\alpha_i}\right) \geqslant \frac{c}{3^n}.$$

Proof. We can find a compact set $K \subset \bigcup_{\alpha\in\Lambda} B_\alpha$ such that $c < m(K) < m(\bigcup_{\alpha\in\Lambda} B_\alpha)$. We find finite subcover for K such that $K \subset \bigcup_{i=1}^N B_{\alpha_i}$. By Version 1, we can find disjoint subcollection $B_{i_1}, B_{i_2}, \cdots, B_{i_L}$ such that

$$\sum_{j=1}^L |B_{i_j}| \geqslant \frac{1}{3^n} m\left(\bigcup_{i=1}^N B_{\alpha_i}\right) \geqslant \frac{1}{3^n} m(K) > \frac{c}{3^n}.$$

□

Theorem 5.7 (Maximal Function Theorem)**.** *For any $f \in L(\mathbb{R}^n)$, there exists a constant C such that*

$$m(\{x \in \mathbb{R}^n : Mf(x) > \alpha\}) \leqslant \frac{C}{\alpha}\int_{\mathbb{R}^n} |f(x)|\mathrm{d}x. \tag{5.3}$$

Proof. For $\alpha > 0$, denote $E_\alpha = \{x \in \mathbb{R}^n : Mf(x) > \alpha\}$. For any $x \in E_\alpha$, there exists $r_x > 0$ such that

$$A_{r_x}|f|(x) = \frac{1}{|B(x,r_x)|}\int_{B(x,r_x)} |f(y)|\mathrm{d}y > \alpha.$$

Then $\{B(x, r_x) : x \in E_\alpha\}$ is a cover of E_α. By Vitali Covering Lemma, for any $c < m\left(\bigcup_{x\in E_\alpha} B(x, r_x)\right)$, we find disjoint balls $B(x_i, r_i)$ such that

$$\sum_{i=1}^N |B(x_i,r_i)| \geqslant \frac{c}{3^n} \Longrightarrow \int_{\bigcup_{i=1}^N B(x_i,r_i)} |f(y)|\mathrm{d}y \geqslant \alpha\sum_{i=1}^N |B(x_i,r_i)| \geqslant \frac{\alpha c}{3^n}.$$

Let $C \to m(\bigcup_{x\in E_\alpha} B(x, r_x))$, we have

$$m(E_\alpha) \leqslant m\left(\bigcup_{x\in E_\alpha} B(x,r_x)\right) \leqslant \frac{3^n}{\alpha}\int_{\mathbb{R}^n} |f(y)|\mathrm{d}y.$$

□

Theorem 5.8. *For any $f \in L(\mathbb{R}^n)$, we have $\lim_{r\to 0} A_r f(x) = f(x)$ for a.e. x.*

Remark: If

Definition 5.9. g is measurable if

$$m\left(\{\}\right)$$

.

Example 5.10. $g(x) = \frac{1}{|x|} \notin L^1(\mathbb{R})$, but $g(x) \in L^{1,\infty}(\mathbb{R})$.

Proof. Notice if $g \in C(\mathbb{R}^n)$, then $\lim_{r\to 0} A_r g(x) = g(x)$. Given $f \in L(\mathbb{R}^n)$, for any $\varepsilon > 0$, we can find $g \in C_c(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} |f(x) - g(x)|\mathrm{d}x < \varepsilon$. Then

$$\begin{aligned}
|A_r f(x) - f(x)| &= \left|\frac{1}{|B(x,r)|}\int_{B(x,r)} (f(y) - f(x))\mathrm{d}y\right| \\
&\leqslant \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y) - g(y)| + |g(y) - g(x)| + |g(x) - f(x)|\mathrm{d}y \\
&\leqslant A_r|f-g|(x) + \frac{1}{|B(x,r)|}\int_{B(x,r)} |g(y) - g(x)|\mathrm{d}y + |g(x) - f(x)|.
\end{aligned}$$

$$\limsup_{r\to 0} |A_r f(x) - f(x)| \leqslant M(f-g)(x) + |g(x) - f(x)|.$$

Denote

$$\begin{aligned} E_k &= \{x \in \mathbb{R}^n : \limsup_{r\to 0} |A_r f(x) - f(x)| > \frac{1}{k}\} \\ &\subset \{x \in \mathbb{R}^n : M(f-g)(x) > \frac{1}{2k}\} \cup \{x \in \mathbb{R}^n : |g(x) - f(x)| > \frac{1}{2k}\}. \end{aligned}$$

□

Definition 5.11. $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. A point $x \in \mathbb{R}^n$ is called a **Lebesgue point** of f if

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)| \mathrm{d}y = 0. \tag{5.4}$$

Denote L_f be the set of all Lebesgue points of f.

Theorem 5.12.

$$m(L_f^c) = 0.$$

Proof. For any $c \in \mathbb{Q}$, $|f(x) - c| \in L^1_{\text{loc}}(\mathbb{R}^n)$. By

□

Remark: For any $r > 0$, consider $E_r \subset \mathbb{R}^n$ measurable such that $E_r \subset B(x,r)$ and $m(E_r) \geqslant C|B(x,r)|$ for some constant $C > 0$. Then for a.e. $x \in \mathbb{R}^n$,

Application: Define $\text{supp}(f) = \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}$. Then $\text{supp}(f * g) \subset \text{supp}(f) + \text{supp}(g)$.

Lemma 5.13. *$f \in L(\mathbb{R}^n), g \in C^k(\mathbb{R}^n)$ and $|\partial^\alpha g| \leqslant M$ for all $|\alpha| \leqslant k$. Then $f * g \in C^k(\mathbb{R}^n)$ and*

$$\partial^\alpha (f * g) = f * (\partial^\alpha g), \quad |\alpha| \leqslant k.$$

Proof. We only prove the case $k = 1$. Fix $x \in \mathbb{R}^n$ and consider $h \to 0$,

$$\frac{(f*g)(x + he_i) - (f*g)(x)}{h} = \int_{\mathbb{R}^n} f(y) \left(\frac{g(x + he_i - y) - g(x-y)}{h} \right) \mathrm{d}y.$$

Since $g \in C^1(\mathbb{R}^n)$, we have

$$\left| \frac{g(x + he_i - y) - g(x-y)}{h} \right| = |\partial_{x_i} g(x - y + \theta h e_i)| \leqslant M$$

for some $\theta \in (0,1)$. □

Consider $\phi \in C_c^\infty(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} \phi = 1$ and $\phi_\varepsilon(x) = \frac{1}{\varepsilon^n} \phi\left(\frac{x}{\varepsilon}\right)$. This kind of function exists, for example, $\phi(x) = \begin{cases} e^{\left(\frac{1}{|x|^2-1}\right)}, & |x| < 1 \\ 0, & |x| \geqslant 1 \end{cases}$.

Theorem 5.14. *$f \in L(\mathbb{R}^n)$. Then*

$$\lim_{\varepsilon \to 0} f * \phi_\varepsilon = f$$

in $L(\mathbb{R}^n)$ and for any $x \in L_f$.

Proof.

$$\begin{aligned} f * \phi_\varepsilon(x) - f(x) &= \int_{\mathbb{R}^n} \phi_\varepsilon(y)(f(x-y) - f(x))\mathrm{d}y \\ &= \int_{\mathbb{R}^n} \phi(z)(f(x-\varepsilon z) - f(x))\mathrm{d}z. \end{aligned}$$

By Fubini Theorem,

$$\begin{aligned} \|f * \phi_\varepsilon - f\|_{L(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \phi(z)(f(x-\varepsilon z) - f(x))\mathrm{d}z \right| \mathrm{d}x \\ &\leqslant \int_{\mathbb{R}^n} |\phi(z)| \left(\int_{\mathbb{R}^n} |f(x-\varepsilon z) - f(x)|\mathrm{d}x \right) \mathrm{d}z. \end{aligned}$$

$$\lim_{\varepsilon \to 0} |\phi(z)| \int_{\mathbb{R}^n} |f(x-\varepsilon z) - f(x)|\mathrm{d}z = 0, \quad |\phi(z)| \int_{\mathbb{R}^n} |f(x-\varepsilon z) - f(x)|\mathrm{d}z \leqslant 2|\phi(z)| \int_{\mathbb{R}^n} |f(x)|\mathrm{d}x.$$

By Dominated Convergence Theorem, we have the conclusion. □