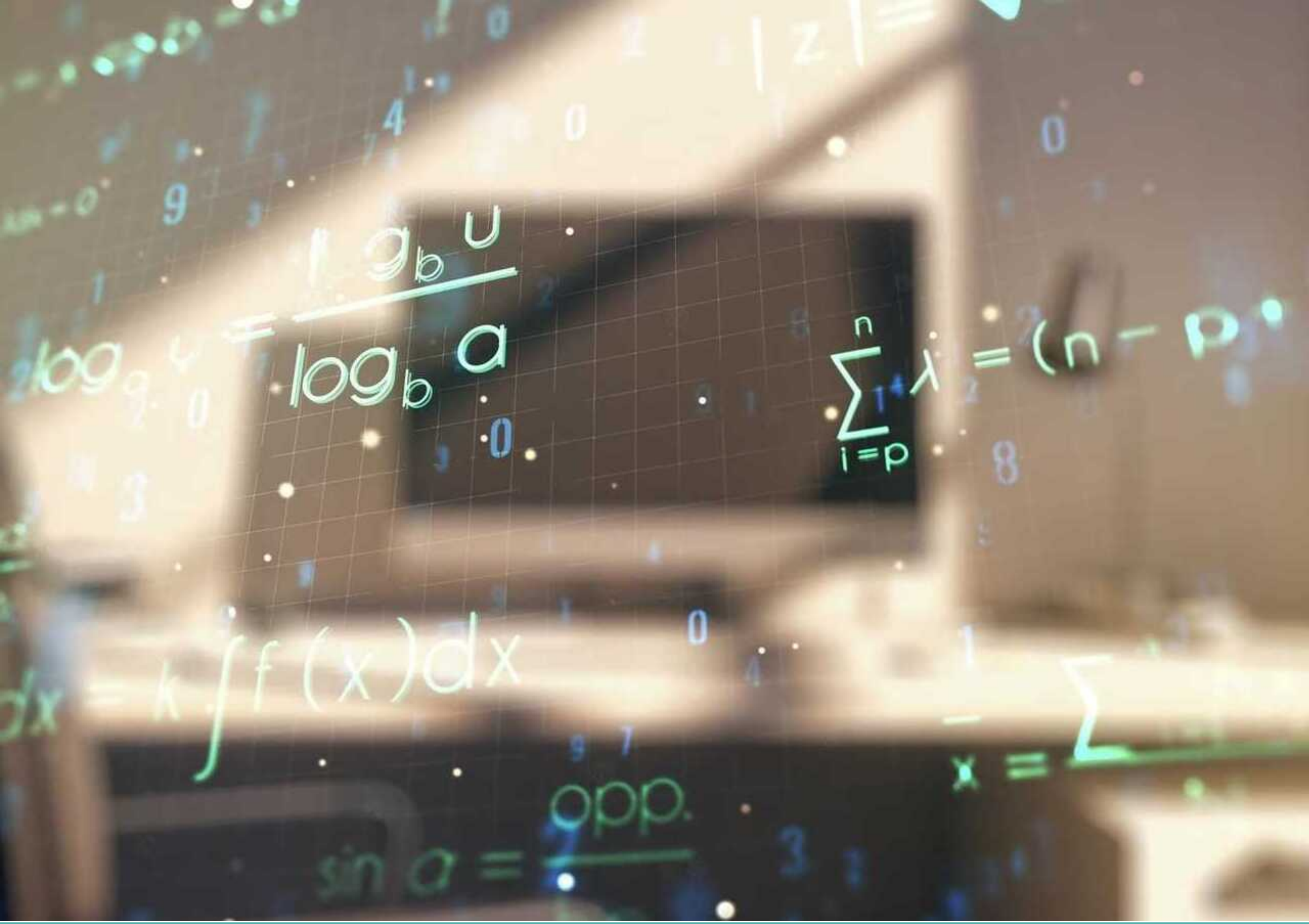


AS.110.415/416: Honors Analysis I & II

Author: Tian Zhou

Institute: Johns Hopkins University

Date: May 13, 2024

Version: 1

Year: Fall 2023

***Textbook**: Principles of Mathematical Analysis - Walter Rudin,*
Real Analysis: Measure Theory, Integration, and Hilbert Spaces - Stein

Contents

Part I

Analysis I

Chapter 1 The Real and Complex Number Systems

Introduction

- *Ordered Set and Least-upper-bound*
- *Real Field and Properties*
- *The Complex Field*

1.1 Ordered Set

Definition 1.1 (Ordered Set)

Suppose S be a set. An ***order*** *on S is a relations, denoted by $<$, with the following properties:*

(1) Trichotomy: If $x, y \in S$, then exactly one of the following $x = y$, $x < y$, $y < x$ is true.

(2) Transitivity: If $x, y, z \in S$, $x < y$, and $y < z$, then $x < z$.

The ***ordered set*** *is a set S in which an order is defined.*

♣

Definition 1.2 (Supremum, Infimum)

Suppose S is a ordered set, $E \subset S$, and E is bounded above. Suppose there exists $\alpha \in S$ such that:

(1) α is an upper bound of E.

(2) If $\gamma < \alpha$, then γ is not an upper bound of E.

Then $\alpha \in S$ is called the ***least upper bound*** *of E (or the supremum of E) and is denoted by $\alpha = \sup E$.*

The definition of ***greatest lower bound*** *(infimum) is an analogous.*

♣

Remark The second statement is equivalent to: for all upper bounds γ, we have $\gamma \geq \alpha$.

Definition 1.3 (Least Upper Bound Property)

An ordered set S has the ***least-upper-bound property*** *if for all nonempty $E \subset S$ that is bounded above, then $\sup E$ exists in S.*

♣

Theorem 1.1

Suppose S is an ordered set the least-upper-bound property, then S has the greatest-lower-bound property, that is, for all nonempty $E \subset S$ that is bounded below, then $\inf E$ exists in S.

♡

Proof Let L be the set of all lower bounds of E, $L \neq \varnothing$. Since L is bounded above by elements in E, there exists

$\alpha = \sup L$ in S. It follows that $\alpha = \inf E$ because (1) for all $x \in E$, $\alpha \leq x$ since x is an upper bound of L, it follows that $\alpha \in L$, and (2) $\gamma \leq \alpha$ for all lower bounds $\gamma \in L$ by definition. This completes the proof.

Remark We construct the set of lower bounds to convert the l.u.b. property to g.l.b. The construction of L gives the following relationships: $L \leq \alpha \leq E$ and $\sup L = \alpha = \inf E$.

1.2 The Real Field

Definition 1.4 (Field, Ordered Field)

A ***field*** $(F, +, \cdot)$ *is a set* F *such that* $(F, +)$ *and* $(F \setminus \{0\}, \cdot)$ *are abelian group, and multiplicative is distributive to addition.*

An ***ordered field*** *is a field* $(F, + \cdots)$ *which is also an ordered set such that*

1. $x + y < x + z$ *if* $x, y, z \in F$ *and* $y < z$*, and*
2. $xy > 0$ *if* $x, y \in F$*,* $x, y > 0$*.*

♣

Example 1.1 There exists no order that turns $\mathbb{C}$ into an ordered field.

Proposition 1.1 (Existence Theorem)

There exists an ordered field $\mathbb{R}$ *which has the least upper bound property.*

♠

Remark Suppose $E \subset S$, $\alpha = \sup E$ if and only if for all $\varepsilon > 0$, there exists $x \in E$ such that $\alpha - \varepsilon < x \leq \alpha$.

Note *Well-ordering principle of* $\mathbb{N}$*: if* E *is a nonempty subset of* $\mathbb{N}$*, the* E *has a least element in it.*

Theorem 1.2

1. *Archimedean property: if* $x, y \in \mathbb{R}$ *and* $x > 0$*, there exists* $N \in \mathbb{Z}_{>0}$ *such that* $nx > y$*.*
2. $\mathbb{Q}$ *is dense in* $\mathbb{R}$*: if* $x, y \in R$ *and* $x < y$*, there exists* $q \in \mathbb{Q}$ *such that* $x < q < y$*.*

♡

Proof (1) For the sake of contradiction, suppose there exists x, y such that $nx \leq y$ for all $n \in \mathbb{Z}_{>0}$. Let $E = \{nx \mid n \in \mathbb{Z}_{>0}\}$, clearly E is nonempty and bounded above by y, there exists $\alpha = \sup E$. There exists $nx \in E$ such that $\alpha - x < nx \leq \alpha$, it follows that $(n+1)x > \alpha$, contradicting the fact that $\alpha = \sup E$.

(2) There exists $n \in \mathbb{N}$ such that $n(y - x) > 1$, namely $ny - 1 > nx$. Apply the Archimedean property again, we obtain $m_1, m_2 \in \mathbb{Z}_{>0}$ such that $m_1 > nx$, $m_2 > -nx$, so $-m_2 < nx < m_1$. It follows that there exists m $(-m_2 \leq m \leq m_1)$ such that $m - 1 \leq nx < m$. Then $nx < m < ny$, so $x < m/n < y$ where $m/n \in \mathbb{Q}$. ■

Remark Indeed, the set of all irrationals $\mathbb{Q}^c$ is also dense in $\mathbb{R}$.

Theorem 1.3

For every real $x > 0$ *and every integer* $n > 0$*, there exists a unique real* $y > 0$ *such that* $y^n = x$*; in other words,* $x^{1/n}$ *exists and is unique.*

♡

Proof For the existence, let $E = \{t > 0 \mid t^n < x\}$. It is not hard to show E is nonempty (by choosing $t < \min(x, 1)$) and bounded above (by $1 + x$), so there exists $\alpha = \sup E$ by the least-upper-bound property.

We now prove $\alpha^n = x$ by contradiction. Notice that $b^n - a^n = (b-a)(b^{n-1} + ab^{n-1} + \cdots + a^{n-1}) < (b-a)nb^{n-1}$.

- Assume $\alpha^n > x$, put $h = (\alpha^n - x)/n\alpha^{n-1}$, then

$$\alpha^n - (\alpha - h)^n < h \cdot n\alpha^{n-1} \leq a^n - x.$$

  That is, $x < (\alpha - h)^n < \alpha^n$, contradicting to the fact that $\alpha = \sup E$.

- Assume $\alpha^n < x$, put $h = \min\{1, (x - \alpha^n)/n(\alpha + 1)^{n-1}\}$, then

$$(\alpha + h)^n - \alpha^n < h \cdot n(\alpha + h)^{n-1} \leq h \cdot n(\alpha + 1)^{n-1} \leq x - \alpha^n.$$

  That is, $\alpha^n < (\alpha + h)^n < x$, contradicting to the fact that α is an upper bound.

Hence, $\alpha^n = x$. ■

Definition 1.5 (Extended Real Number System)

The extended real number system consists of the real field R and two symbols, $+\infty$ and $-\infty$. We preserve the original order in $\mathbb{R}$, and define $-\infty < x < +\infty$ for every $x \in R$. ♣

The extended real number system does not form a field.

1.3 The Complex Field and The Euclidean Spaces

Definition 1.6 (Complex Number)

A ***complex number*** *is an ordered pair* (a, b) *of real numbers. Let* $x = (a, b)$ *and* $y = (c, d)$*, we define the addition and multiplication by* $x + y = (a + c, b + d)$ *and* $xy = (ac - bd, ad + bc)$*.* ♣

Remark The complex number along with addition and multiplication forms a field $\mathbb{C}$, and it contains $\mathbb{R}$ as a subfield.

We define the *conjugate* of $x = (a, b)$ by $\bar{x} = (a, -b)$ and define the absolute value $|x| = (x\bar{x})^{1/2}$. The complex numbers have the following properties:

- $\overline{z + w} = \bar{z} + \bar{w}$, $\overline{zw} = \bar{z}\bar{w}$;
- $z + \bar{z} = 2\text{Re}(z)$, $z = \bar{z} = 2\text{Im}(z)$;
- $z\bar{z}$ is real and positive (exception when $z = 0$), so $|z| > 0$;
- $|\bar{z}| = |z|$;
- $|zw| = |z||w|$;
- $|\text{Re}(z)| \leq |z|$;
- (triangle inequality) $|z + w| \leq |z| + |w|$

  Proof:
  $$\begin{aligned} |z + w|^2 &= (z + w)(\bar{z} + \bar{w}) = z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} \\ &= |z|^2 + 2\text{Re}(z\bar{w}) + |w|^2 \\ &\leq |z|^2 + 2|z\bar{w}| + |w|^2 = |z|^2 + 2|z||w| + |w|^2 \\ &= (|z| + |w|)^2, \end{aligned}$$
  it follows that $|z + w| \leq |z| + |w|$. ■

Proposition 1.2 (Schwarz Inequality)

If $a_1, \cdots, a_n$ *and* $b_1, \cdots, b_n$ *are complex numbers, then*
$$\left| \sum_{j=1}^{n} a_j \bar{b_j} \right|^2 \leq \sum_{j=1}^{n} |a_j|^2 \sum_{j=1}^{n} |b_j|^2.$$
♠

Proof Put $A = \sum |a_j|^2$, $B = \sum |b_j|^2$, and $C = \sum a_j \bar{b_j}$. If $B = 0$, $b_1 = \cdots = b_n = 0$, so the conclusion is trivial. Suppose therefore $B > 0$, then
$$\begin{aligned} 0 \leq \sum |Ba_j - Cb_j|^2 &= \sum (Ba_j - Cb_j)(B\bar{a_j} - \overline{Cb_j}) \\ &= B^2 \sum |a_j|^2 - B\bar{C} \sum a_j \bar{b_j} - BC \sum \bar{a_j} b_j + |C|^2 \sum |b_j|^2 \\ &= B^2 A - B\bar{C}C - BC\bar{C} + B|C|^2 \\ &= B(AB - |C|^2). \end{aligned}$$

Therefore, $AB - |C|^2 \geq 0$. ■

Proof (*Alternative*) We use the same notation of A, B, C as above. Notice that $AB = \sum_{i,j} a_i \bar{b_j} \bar{a}_i b_j$ and $|C|^2 = \sum_{i,j} a_i \bar{b_j} \bar{a_j} b_i$. Then

$$\begin{aligned} AB &= \Big(\sum_i a_i \bar{a_i}\Big)\Big(\sum_j b_j \bar{b_j}\Big) \\ &= \Big(\sum_i a_i b_i\Big)\Big(\sum_j \bar{a_j} \bar{b_j}\Big) + \sum_{i,j} a_i \bar{b_j}(\bar{a}_i b_j - \bar{a_j} b_i) \\ &= \Big|\sum_i a_i b_j\Big|^2 + \sum_{i \leq j}(a_i \bar{b_j} - a_j \bar{b_i})(\bar{a}_i b_j - \bar{a_j} b_i) \\ &= |C|^2 + \sum_{i \leq j} |a_i \bar{b_j} - a_j \bar{b_i}|^2 \\ &\geq |C|^2. \end{aligned}$$

Hence $AB \geq |C|^2$. ■

Definition 1.7 (Euclidean Space)

For $k \in \mathbb{Z}_{>0}$, $\mathbb{R}^k = \{\mathbf{x} : \mathbf{x} = (x_1, \cdots, x_k),\ x_i \in \mathbb{R}$ *for all* $i\}$ ♣

Let $\mathbf{x} = (x_1, \cdots, x_k)$ and $\mathbf{y} = (y_1, \cdots, y_k)$. The addition is defined by $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \cdots, x_k + y_k)$, the scalar multiplication is defined by $a\mathbf{x} = (ax_1, \cdots, ax_k)$, the inner product is defined by $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^k x_i y_i$, and the norm is defined by $|\mathbf{x}| = (\mathbf{x} \cdot \mathbf{x})^{1/2} = (\sum_{i=1}^k x_i^2)^{1/2}$.

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^k$, $\alpha \in \mathbb{R}$, the following properties holds:

- $|\mathbf{x}| \geq 0$, and the equality holds if and only if $x = 0$;
- $|a\mathbf{x}| = |a||\mathbf{x}|$;
- (Cauchy-Schwartz) $|\mathbf{x} \cdot \mathbf{y}| \leq |\mathbf{x}||\mathbf{y}|$;
- (triangle inequality) $|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|$;
- $|\mathbf{x} - \mathbf{z}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}|$.

Chapter 2 Basic Topology

Introduction

- *Countable and Uncountable Sets*
- *Limit Point, Closed Set, Closure*
- *Open Relative (Subspace topology)*
- *Heine-Borel Theorem*
- *Neighborhoods, Open Sets*
- *Bounded Set, Dense Subset*
- *Open Cover, Compact Set*
- *Perfect Set, Connected Sets*

2.1 Countable Sets

Definition 2.1 (1-1 Correspondence)

Suppose A, B are sets, we say A and B are in 1-1 correspondence if there exists a bijection $f : A \to B$. We write $A \sim B$, and this relation is a equivalence relation. ♣

Definition 2.2 (Countability)

For any $n \in \mathbb{Z}^+$, define $J_n := \{1, \cdots, n\}$, and let $J = \{1, \cdots\} = \mathbb{Z}_{>0}$. For any set A, we say:

- *A is **finite** if $A \sim J_n$ for some n (the empty set is considered finite).*
- *A is **countable** if $A \sim J$.*
- *A is **at most countable** if A is finite or countable, and A is uncountable if it is not at most countable.* ♣

Example 2.1 $\mathbb{Z}$ is countable, because $f : \mathbb{N} \to \mathbb{Z}$, defined by $f(n) = n/2$ if n is even and $f(n) = -(n-1)/2$ if n is odd, is a bijection.

Definition 2.3 (Sequence)

*A **sequence** is a function defined on $\mathbb{N}$. If f is a sequence, we denote $x_n = f(n)$, and we write f as $\{x_n\}$.* ♣

Proposition 2.1

Every infinite subset of a countable set A is countable. ♠

Proof Let $E \subset A$ be an infinite subset. Since A is countable, there exists $\{x_n\} = A$. Let n_1 be the least positive integer such that $x_{n_1} \in E$, which exists by the well-ordering principle. Recursively, choose n_i from $E \setminus \{x_1, \cdots, x_{n-1}\}$, which is nonempty since E is infinite, such that n_i is the least positive integer such that $x_{n_i} \in E$. Putting $f(k) = x_{n_k}$ $(k \in \mathbb{Z}_+)$, we obtain an 1-1 correspondence between E and J. ■

Note: That is, every subset of a countable set is at most countable.

Proposition 2.2

Let $\{E_n\}$, $n = 1, 2, \cdots$ be a sequence of countable sets, and put $S = \bigcup_{n=1}^{\infty} E_n$, then S is countable. In other words, the countable union of countable sets is countable.

Proof Let $E_n = \{x_{n,k}\}_{k=1}^{\infty}$ for all n, S can be enumerated as:

$$\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{array}$$

namely $S = \{x_{1,1}, x_{2,1}, x_{1,2}, x_{3,1}, \cdots\}$. Then S is at most countable. Since $E_1 \subset S$ is countable thus infinite, S is countable. ■

Corollary 2.1

The at most countable union of at most countable sets is at most countable.

Proposition 2.3

Let A be a countable set, and let B_n be the set of all n-tuples $(a_1, \cdots, a_n)$ where $a_k \in A$ ($k = 1, \cdots, n$), and the element need not be distinct. Then B_n is countable. In other words, the finite cartesian product of countable sets is countable.

Proof We proceed by induction on n. If $n = 1$, the statement is trivial. For $n > 1$, suppose B_{n-1} is countable. Fix $b \in B_{n-1}$, let $E_b := \{(a, b) \,|\, a \in A\}$, which is countable since A is countable. Then $B_n = \bigcup_{b \in B_{n-1}} E_b$ is a countable union of countable sets, then B_n is countable by proposition 2.2. ■

Corollary 2.2

$\mathbb{Q}$ is countable.

Proof The set of $\mathbb{Z} \times \mathbb{Z}$ is countable by proposition 2.3, and $\mathbb{Q}$ can be view as the subset of $\mathbb{Z} \times \mathbb{Z}^* \subset \mathbb{Z} \times \mathbb{Z}$ by the map $f : (x, y) \mapsto x/y$, followed by $\mathbb{Q}$ is countable by 2.2. ■

Proposition 2.4 (Cantor)

$\mathbb{R}$ is uncountable.

Proof For the sake of contradiction, suppose $\mathbb{R}$ is countable, then so is $(0, 1) \subset \mathbb{R}$. Clearly, $(0, 1)$ is infinite. We can enumerate $(0, 1)$ as $\{x_n\}_{n=1}^{\infty}$, and let $x_n = 0.x_{n1}x_{n2} \cdots$ be the decimal representation. Choose $y = 0.b_1 b_2 \cdots$

for which $b_n \neq x_{nn}$ for all n. It follows that $y \neq x_n$ for all n since $b_n \neq x_{nn}$, so $y \notin \{x_n\}_{n=1}^{\infty}$, contradicting the fact that $y \in (0, 1)$. ■

2.2 Metric Spaces Topology

Definition 2.4 (Metric Spaces)

*A **metric space** is a set X with a distance function (**metric**) $d : X \to X \to \mathbb{R}$ such that:*

(a) $d(x, y) > 0$ if $p \neq q$, and $d(p, p) = 0$;

(b) $d(p, q) = d(q, p)$;

(c) $d(p, q) \leq d(p, r) + d(r, q)$ for any $r \in X$.

♣

Example 2.2 Let $X = \mathbb{R}^k$, defined $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ to be the usual Euclidean distance. d satisfies all the conditions in the above definition, so $(\mathbb{R}^k, d)$ is a metric space, and we called the Euclidean distance the "usual" distance in $\mathbb{R}^k$.

Definition 2.5 (Neighborhood, Open Set)

Let (X, d) be a metric space,

*(a) A **neighborhood** of p is a set $N_r(p)$ consisting of all $q \in X$ such that $d(p, q) < r$, for some $r > 0$. The number r is called the radius of $N_r(p)$.*

(b) A point E is an interior point of E if there is a neighborhood N of p such that $N \subset E$.

*(c) A set E is **open** if every point of E is an interior point.*

♣

Remark A set E is ***open*** if and only if for all $p \in E$, there exists $r > 0$ such that $N_r(p) \subset E$.

Proposition 2.5

Every neighborhood is an open set.

♠

Proof Consider the neighborhood $E = N_r(p)$. For all $p' \in E$, let $h = d(p, p') < r$, then $N_{r-h}(p') \subset E$, because for all $q \in E$, $d(p, q) \leq d(p, p') + d(p', q) < h + (r - h) = r$. Hence p' is an interior point for all $p' \in E$, thus E is open. ■

Definition 2.6 (Closed Set)

Let (X, d) be a metric space, suppose $E \subset X$,

*(a) A point p is a **limit point** of the set E if every neighborhood of p contains a point $q \neq p$ such that $q \in E$. If $p \in E$ and p is not a limit point of E, then p is called the **isolated point** of E.*

*(b) E is **closed** if every limit point of E is a point of E.*

♣

Remark Equivalently, p is a limit point if and only if $N_r^*(p) \cap E \neq \varnothing$, where we denote $N_r^*(p) := N_r(p) \setminus \{p\}$.

Proposition 2.6

If p is a limit point of a set E, then every neighborhood of p contains infinitely many points of E.

Proof Proof by contradiction. Suppose p is a limit point of E, and there exists a neighborhood of p containing finitely many points $q_1, \cdots, q_n$. Put $r = \min_i d(p, q_i)$, then $r > 0$ since $\{q_i\}$ is finite. It follows that $N_r(p)$ contains no points of $E \setminus \{p\}$, contradicting that p is a limit point. ■

Remark Corollary: A finite point set has no limit points.

Definition 2.7 (Boundedness, Dense)

Let (X, d) be a metric space, suppose $E \subset X$,

*(a) The **complement** of E, denoted by E^c, is $E^c = \{p \in X \mid p \notin E\}$.*

*(b) E is **bounded** if there exists $M > 0$ and $p \in E$ such that $d(p, q) < M$ for all $q \in E$.*

*(c) E is **dense** in X if every point of X is a limit point of E or in E.*

DeMorgan's Law: Let $\{E_\alpha\}$ be a collection of sets E_α, then $(\bigcup_\alpha E_\alpha)^c = \bigcap_\alpha (E_\alpha^c)$.

Theorem 2.1

A set E is open if and only if its complement is closed.

Remark Corollary: A set F is closed if and only if its complement is open.

Proof ($\Rightarrow$) Suppose E is open, and let x be a limit point of E^c. If $x \in E$, there exists r' such that $N_{r'}(x) \subset E$, so $N_{r'}(x) \cap E^c = \varnothing$, contradicting that x is a limit point of E^c. Thus $x \in E^c$, implying that E^c is closed.

($\Leftarrow$) Suppose E^c is closed, and let $x \in E$. Since $x \notin E^c$, x is not a limit point of E^c, implying that there exists $r > 0$ such that $N_r^*(x) \cap E^c = N_r(x) \cap E^c = \varnothing$. It follows that $N_r(x) \subset E$, thus E is open. ■

Proposition 2.7

(a) Arbitrary unions and finite intersections of open sets are open.

(b) Arbitrary intersections and finite unions of closed sets are closed.

Proof (a) (i) Suppose $x \in G = \bigcup_\alpha G_\alpha$, x is a point G_β thus an interior point of G_β for some β. Then x is an interior point of G since $G_\beta \subset G$, so the arbitrary union of open sets is open. (ii) Suppose $x \in G = \bigcap_{i=1}^n G_i$, then for all i, there exists r_i such that $N_{r_i}(x) \subset G_i$. Put $r = \min\{r_i\}$, we have $N_r(x) \subset G_i$ for all i, so $N_r(x) \subset G$. Thus, x is an interior point of G, so G is open.

(b) By taking the complement and using DeMorgan's Law, we obtain (b) from (a).

Remark The infinite intersection of open sets is not necessarily open. For instance, $G_n = (-1/n, 1/n)$ for $n \in \mathbb{N}$, then $\bigcap G_n = \{0\}$ is not an open subset.

Definition 2.8 (Closure)

*If X is a metric, $E \subset X$, and E' denotes the set of all limit points of E in X, the the **closure** of E is the set $\bar{E} = E \cup E'$.*

Note *The **interior** E° is defined to be the set of all interior points of E. The **boundary** ∂E is defined to be $\partial E := \overline{E} \setminus E^\circ$.*

Proposition 2.8

If X is a metric space and $E \subset X$, then

(a) $\bar{E}$ is closed,

(b) $E = \bar{E}$ if and only if E is closed, and

(c) $\bar{E} \subset F$ for every closed set F containing E.

Remark By (a) and (c), $\bar{E}$ is the smallest closed subset that contains E.

Proof (a) Let $p \in \bar{E}^c$, p is not in E nor a limit point of E, so there exists $r > 0$ such that $N_r(p) \cap E = \varnothing$. If $q \in N_r(p) \cap E'$, let $d = d(p,q)$, then there exists $q' \in E \cap N_{r-d}(q) \subset N_r(p) \cap E = \varnothing$, contradiction. Therefore, $N_r(p) \cap \bar{E} = N_r(p) \cap (E \cup E') = \varnothing$, then p is the interior point of $\bar{E}^c$, so $\bar{E}^c$ is open, followed by $\bar{E}$ is closed.

(b) Suppose $E = \bar{E}$, then E is closed by (a). Conversely, suppose E is closed, $E \subset \bar{E} = E \cap E' = E$, so $E = \bar{E}$.

(c) Since $E \subset F$, $E' \subset F$ because E' are limit points of F and thus in F because F is closed. Therefore, $\bar{E} = E \cup E' \subset F$.

Remark For (a), we show that for $p \in \bar{E}^c$, $N_r(p)$ contains no points of E, and it contains no limit points of E, otherwise it intersects E. Then we conclude p is an interior point.

Proposition 2.9

Let E be a nonempty set of $\mathbb{R}$ which is bounded above, and let $y = \sup E$. Then $y \in \bar{E}$, thus $y \in E$ if E is closed.

Proof Suppose $y \in E$, then obviously $y \in \bar{E}$. Suppose $y \notin E$, then for all $\varepsilon > 0$, there exists y' such that $y' \in N_\varepsilon(y) \cap E = N_\varepsilon^*(y) \cap E$. It implies that $y \in E'$, so $y \in \bar{E}$. ∎

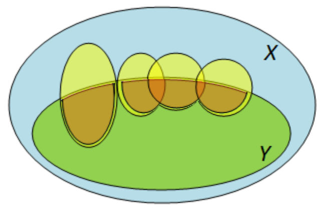
Definition 2.9 (Open Relative)

*Let $Y \subset X$ be a non-empty subset. $E \subset Y$ is **open relative** to Y if for each $p \in E$, there exists $r > 0$ such that $N_r(p) \cap Y \subset E$. Equivalently, there exists $r > 0$ such that $q \in E$ whenever $d(p,q) < r$ and $q \in Y$.*

Proposition 2.10

Suppose $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.

Remark E is open relative to $Y \subset$ means E is open in the ***subspace topology*** Y on X.



Proof ($\Rightarrow$) Suppose E is open relative to Y. To each $p \in E$ there is a positive number r_p such that $N_{r_p}(p) \cap Y \subset E$. Let $G = \bigcup_{p \in E} N_{r_p}(p)$, G is clearly open. Note that for all $p \in E$, $p \in N_{r_p}(p) \cap Y$, then $E = \bigcup_{p \in E}(N_{r_p}(p) \cap Y) = (\bigcup_{p \in E} N_{r_p}) \cap Y = G \cap Y$.

($\Leftarrow$) Suppose $E = Y \cap G$ for some open set G in X. For all $p \in E = G \cap Y$, there exists $r > 0$ such that $N_r(p) \subset G$ since G is open in X, then $N_r(p) \cap Y \subset G \cap Y = E$. Thus, E is open relative to Y. ■

Example 2.3 Consider $E = (0,1) \times \{0\}$. E is open (relative) to $Y = \mathbb{R} \times \{0\}$, considering E as a subset of Y. However, if we consider E as a subset of $X = \mathbb{R}^2$, E is not open.

2.3 Compact Space

Definition 2.10 (Open cover, Compactness)

*Suppose (X, d) is a metric space. An **open cover** of a set $E \subset X$ is a collection of open sets $\{G_\alpha \,|\, \alpha \in A\}$ such that $E \subset \bigcup_{\alpha \in A} G_\alpha$.*

*$K \subset X$ is **compact** if every open cover contains a finite subcover.* ♣

Proposition 2.11

Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y. ♠

Proof ($\Rightarrow$) Suppose K is compact relative to X, and assume $\{V_\alpha = G_\alpha \cap Y\}$ is an open cover open relative to Y. Then $\{G_\alpha\}$ is an open cover of K, so there is a finite subcover $\{G_i\}$ since K is compact relative to X. Thus, $K \subset (\bigcap_{i=1}^{n} G_i) \cap Y = \bigcap_{i=1}^{n} (G_i \cap Y) = \bigcap_{i=1}^{n} V_i$. It follows that there exists a finite subcover $\{V_i = G_i \cap Y\}$ of K open relative to Y, so K is open relative to Y.

($\Leftarrow$) The converse is an analogous. ■

Proposition 2.12

Compact subsets of a metric space are closed. ♠

Proof Suppose K is compact, we want to show K^c is open. Let $q \in K^c$ be given. For all $p \in K$, let $d = d(p, q)/2 > 0$, we define the neighborhoods $p \in U_p = N_d(p)$ and $q \in V_p = N_d(q)$. Note that $\{U_p \,|\, p \in K\}$ forms an open cover of K, so there exists a finite subcover $\{U_{p_i}\}$. Consider $V = \bigcap_{i=1}^{n} V_{p_i}$. Note that V is open, and $V \cap K = \varnothing$, since for all U_{p_i}, $K \cap U_{p_i} \subset V_{p_i} \cap U_{p_i} = \varnothing$. Therefore, p is an interior point, so K is closed since the choice of p is arbitrary. ■

Proposition 2.13

Closed subsets of compact sets are compact. ♠

Proof Suppose $F \subset K \subset X$ where F is closed relative to K and K is compact. Assume $\{U_\alpha\}$ is an open cover of F. Adding the open set F^c to $\{U_\alpha\}$ yields an open cover of K, so there exists a finite subcover $\{V_i\}$ of K since K is compact. Removing F^c from $\{V_i\}$ (if exists) gives a finite subcover of F. Hence F is compact. ■

Corollary 2.3

The intersection of a compact set and a closed set is compact. ♡

Proposition 2.14 (Finite intersection property)

If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X such that the intersection of every finite subcollection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is nonempty.

Proof For the sake of contradiction, suppose $\bigcap K_\alpha = \varnothing$. Fix $K_1 \in \{K_\alpha\}$, then $K_1 \subset \bigcup K_\alpha^c$. By the compactness, there exists $\alpha_1, \cdots, \alpha_n$ such that $K_1 \subset \bigcup_{i=1}^n K_{\alpha_i}^c = (\bigcap_{i=1}^n K_{\alpha_i})^c$. Then $K_1 \cap \bigcap_{i=1}^n K_{\alpha_i} = \varnothing$, contradicting that finite intersections are nonempty. ■

Note ***Corollary***: *If $\{K_n\}_{n\in\mathbb{N}}$ is a sequence of nonempty compact sets such that $K_n \supset K_{n+1}$ for all $n \in \mathbb{N}$, then $\bigcap_{n=1}^\infty K_n$ is nonempty.*

k-cell A **k-cell** is a set $I \subset \mathbb{R}^k$ of the fork $I = [a_1, b_1] \times \cdots \times [a_k, b_k]$ where $a_j < b_j$ for $j = 1, \cdots, k$.

Lemma 2.1

If $\{I_n\}$ is a sequence of intervals in $\mathbb{R}^1$ such that $I_n \supset I_{n+1}$, then $\bigcap_{n=1}^\infty I_n$ is nonempty.

Proof Let $I_n = [a_n, b_n]$ for all n, and put $E = \{a_n\}$. E is nonempty and bounded above by b_1, so there exists $x = \sup E$. For all m, notice that $a_1 \leq a_2 \leq \cdots a_m \leq b_m \leq \cdots \leq b_2 \leq b_1$, so $x \leq b_m$. Also note that clearly $a_m \leq x$ by the definition of supremum, thus $x \in I_m$. Hence $x \in \bigcap_{m=1}^\infty I_m$. ■

Remark It is nor hard to show the intersection of a sequence of k-cells is nonempty.

Proposition 2.15

Every k-cell is compact.

Proof Proof by contradiction. Suppose $I \subset \mathbb{R}^k$ is a k-cell and is not compact. Put $\delta = \sqrt{\sum(a_i - b_i)^2}$. Let $c_j = (a_j + b_j)/2$, dividing $[a_j, b_j]$ into $[a_j, c_j] \cup [c_j, b_j]$ determines 2^k k-cell, and at least one of the k-cells, denoted by I_1, is not compact because I is not compact.

Continuing this process we obtain a sequence $\{I_n\}$ such that (a) $I_n \supset I_{n+1}$, (b) I_n cannot be covered by any finite subcollection of an open cover $\{G_\alpha\}$, and (c) $|x - y| \leq 2^{-n}\delta$ if $x, y \in I_n$. There exists $x^* \in \bigcap I_n$ by Lemma 2.1 and $x^* \in G_\alpha$ for some α . Since G_α is open, there exists $r > 0$ such that $N_r(x^*) \subset G_\alpha$, and there exists $n \in \mathbb{Z}_{>0}$ such that $2^{-n} < r$ by the Archimedean property. This leads to a clear contradiction to (b). Hence I is compact. ■

Lemma 2.2

Suppose K is compact and $E \subset K$ is an infinite subset. Then E has a limit point in K.

Proof Proof by contradiction. Suppose E has no limit point in K, then for all $q \in E$ there exists $\varepsilon_q > 0$ such that $N^*_{\varepsilon_q}(q) \cap E = \varnothing$. That is, $N_{\varepsilon_q}(q) \cap E = \{q\}$. The collection $\{N_{\varepsilon_q}(q) \,|\, q \in E\}$ forms an open cover, there exists a finite subcover by the compactness, contradicting to the fact that E is infinite. ■

Theorem 2.2 (Heine-Borel Theorem)

Suppose E is a subset of $\mathbb{R}^k$ with Euclidean metric, then the following are equivalent:

(a) *E is closed and bounded.*

(b) *E is compact.*

(c) *Every infinite subset of E has a limit point in E.*

♡

Remark In general, $(a) \not\Rightarrow (b)$ and $(a) \not\Rightarrow (c)$.

Proof $(a) \Rightarrow (b)$: E is bounded, so there is a k-cell containing E. Then E is a closed subset of compact set, so E is compact by Proposition 2.13.

$(b) \Rightarrow (c)$: Lemma 2.2.

$(c) \Rightarrow (a)$: Suppose E is not bounded, then E contains points $S = \{x_n\}_{n=1}^{\infty}$ such that $|x_n| > n$. S has no limit points since $N_{1/2}(p) \cap E$ contains at most two points, then (c) does not hold since S is infinite.

Now suppose E is not closed, then there exists a limit point x of E such that $x \notin E$. Construct $S = \{x_n\}_{n=1}^{\infty}$ such that $x_n \in N_{1/n} \cap E$. Assume y is another limit point of E, let $d = |x-y|/2 > 0$, and choose n_0 for which $1/n_0 \leq d$. Then $|x_n - y| \geq |x-y| - |x-x_n| \geq 2d - 1/n$, so $|x_n - y| \geq d$ for $n \geq n_0$. It implies that $N_d(y)$ is contains finitely many points in E, so y is not a limit point of E by Proposition (2.6). Then S is infinite and the only limit point is x but $x \notin E$. Therefore, E is closed and bounded by contrapositive. ■

Theorem 2.3 (Weierstrass)

Every bounded infinite subset E of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.

♡

Proof E is a subset of a k-cell $I \subset \mathbb{R}^k$ by the boundedness. Since I is compact, E has a limit point in $I \subset \mathbb{R}^k$ by Lemma (2.2). ■

2.4 Perfect Sets and Connected Sets

2.4.1 Perfect Sets

Definition 2.11 (Perfect Sets)

*Suppose (X, d) is a metric space and $E \subset X$. E is **perfect** if $E = E'$, equivalently, E is closed and has no isolated points. If $p \in E$ is not a limit point of E, p is called an **isolated point** of E.* ♣

Example 2.4 For fixed $a, b \in \mathbb{R}$, the closed interval $[a, b] \subset \mathbb{R}^1$ is perfect.

Proposition 2.16

Let P be a nonempty perfect set in $\mathbb{R}^k$, then P is uncountable. ♠

Proof P is infinite because it has a limit point. Assume P is countable and $P = \{x_i\}_{i=1}^{\infty}$. Fix $r_1 > 0$, let $V_1 = N_{r_1}(x_1)$. Since x_1 is a limit point, $V_1 \cap P \neq \varnothing$. We can construct recursively a sequence of neighborhoods $V_2, V_3, \cdots$ of points in E, for which (i) $\overline{V_{n+1}} \subset V_n$ and (ii) $x_n \notin \overline{V_{n+1}}$, and we know that $V_n \cap P \neq \varnothing$ since the center of V_n is a limit point of P.

Put $K_n = \overline{V_n} \cap P$. K_n is compact since $\overline{V_n}$ is compact and P is closed. Then $\bigcap_{i=1}^{\infty} K_n$ is nonempty by the Lemma (2.1). However, $x_n \neq K_{n+1}$ implies $\bigcap_{n=1}^{\infty} K_n = \varnothing$. By contradiction, P is uncountable. ■

Remark Key Claim: Given an open set U and $x \in X$, there exists an open subset $V \subsetneq U$ such that $x \notin V$, this holds by the Hausdorff axiom.

Key idea: We can construct a strictly decreasing sequence $\{V_n\}$ of neighborhoods of points of P, for which every V_n intersects P (by perfectness) but V_n converges to points outside of P (by excluding x_n in V_{n+1}). Then there is a contradiction regards to the intersection of $\{\overline{V_n} \cap P\}$.

Corollary 2.4

Every interval $[a, b]$ ($a < b$) is uncountable. In particular, the set of all real numbers is uncountable. ♡

Example 2.5 *Cantor Set*: Let $E_0 = [0, 1]$. Recursively define E_n by removing the middle thirds of the intervals in E_{n-1}, e.g., $E_1 = [0, 1/3] \cup [2/3, 1]$. We obtain a sequence of compact sets E_n such that

(i) $E_1 \supset E_2 \supset \cdots$, and

(ii) E_n is the union of 2^n intervals, each of length 3^{-n}.

The set $P = \bigcap_{i=1}^{\infty} E_n$ is called the *Cantor Set*, and

- P is compact and P is nonempty by Lemma 2.1.
- P contains no segment. By the construction, the segment of the form $((3k+1)/3^m, (3k+2)/3^m)$ is not contained in P, but every segment (α, β) contains such segment, so P contains no segments.

- P is perfect. Let $x \in P$ and S be a segment containing x. Let I_n be the interval of E_n containing x, choose n large enough so that $I_n \subsetneq S$. Put x_n be the endpoint of of I_n such that $x_n \neq x$, it follows that $x_n \in P$ thus $x_n \in S \cap P$, so x is a limit point of P. Hence P is perfect.

The Cantor set is an example of totally disconnected, perfect, compact metric space.

2.4.2 Connected Sets

Definition 2.12 (Connectedness)

Suppose X is a metric space and $A, B \subset X$. A and B are said to be ***separated*** *if $A \cap \overline{B} = \overline{A} \cap B = \varnothing$. A set $E \subset X$ is said to be* ***connected*** *if E is not a union of two nonempty separated sets.* ♣

Proposition 2.17

Suppose $E \subset \mathbb{R}$, E is connected if and only if it has the following property: if $x, y \in E$ and $x < z < y$, then $z \in E$. ♠

Proof ($\Rightarrow$) Proof by contrapositive. Assume there exists $z \in (x, y)$ such that $z \notin E$. Then $E = A_z \cup B_z$ where $A_z := E \cap (-\infty, z)$ and $B_z := E \cap (x, \infty)$. A_z, B_z are clearly nonempty and separated, then E is not connected.

($\Leftarrow$) Proof by contrapositive. Assume E is not connected and A, B is a separation. Choose $x \in A$ and $y \in B$, assume $x < y$ without loss of generality. Let $a = \sup(A \cup [x, y])$ and $b = \inf(B \cup [x, y])$. Clearly $a \leq b$. If $a < b$, choose $c \in (a, b)$, then $c \notin A \cup B = E$ but $x < c < y$, contradiction. Otherwise if $a = b$, $a \in \overline{A} \cap \overline{B}$, it means that $a \notin A \cup B = E$ since $A \cap \overline{B} = \overline{A} \cap B = \varnothing$. Then $x < a < y$ and $a \notin E$. ■

Remark The following are criteria of connectedness:

(a) The subset set $E \subset X$ is connected if and only if there exists no disjoint nonempty open (relative to E) subsets A, B of E such that $E = A \cup B$.

(b) The subset set $E \subset X$ is connected if and only if the only subsets that are both open and closed (relative to E) are empty set and E itself.

Chapter 3 Numerical Sequences and Series

Introduction

- *Convergent Sequences*
- *Cauchy Sequences*
- *Subsequence and Subsequential Limits*

3.1 Convergent Sequences

3.1.1 Convergent Sequences

Definition 3.1 (Convergence)

*A sequence $\{p_n\}$ in a metric space X is said to **converge** if there is a point $p \in X$ such that for all $\varepsilon > 0$, there exists an integer N such that $d(p_n, p) < \varepsilon$ if $n \geq N$. We denote the convergence by $p_n \to p$ or $\lim_{n\to\infty} p_n = p$. If $\{p_n\}$ does not converge, it is said to **diverge**.*

Proposition 3.1

Let $\{p_n\}$ be a sequence in a metric space X,

(a) $\{p_n\}$ converges to $p \in X$ if and only if every neighborhood of p contains p_n for all but finitely many n.

(b) If $p, p' \in X$ and $\{p_n\}$ converges to both p and p', then $p = p'$.

(c) If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.

(d) If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = \lim_{n\to\infty} p_n$.

Proof (a) ($\Rightarrow$) The forward direction is trivial by definition, since for all $N_\varepsilon(p)$, we can choose N by definition such that $N_\varepsilon(p)$ contains all p_n for which $n \geq N$. ($\Leftarrow$) Conversely, let $\varepsilon > 0$ be given. Put $E = \{n \in \mathbb{Z}_{>0} : p_n \notin N_\varepsilon(p)\}$, E is finite. Let $N = \max E$, then $p_n \in N_\varepsilon(p)$ for all $n \geq N + 1$.

(b) Suppose $\{p_n\}$ converges to both p and p'. Assume $p \neq p'$, let $d = d(p, p')/2$. Then there exists N such that $p_n \in N_d(p) \cap N_d(p')$ for $n \geq N$, but $N_d(p) \cap N_d(p') = \varnothing$, contradiction.

(c) Suppose $p_n \to p$. There exists N such that $d(p_n, p) < 1$ for all $n \geq N$, then diameter is bounded by $M = \max\{d(p_1, p), \cdots, d(p_{N-1}, p), 1\}$.

(d) For all $n \in \mathbb{Z}_{>0}$, choose $p_n \in N_{1/n}(p)$, then the sequence $\{p_n\}$ converges to p. ■

Proposition 3.2

Suppose $\{s_n\}, \{t_n\}$ are complex sequences, and $\lim_{n\to\infty} s_n = s$, $\lim_{n\to\infty} t_n = t$. Then

(a) $\lim_{n\to\infty}(s_n + t_n) = s + t$;

(b) $\lim_{n\to\infty}(cs_n) = cs$, $\lim_{n\to\infty}(c + s_n) = c + s$, *for any number* c*;*

(c) $\lim_{n\to\infty} s_n t_n = st$*;*

(d) $\lim_{n\to\infty} 1/s_n = 1/s$, *given* $s_n \neq 0$ *and* $s \neq 0$.

Proof (d) Choose M such that $|s_n - s| < |s|/2$ if $n \geq M$, then we see $|s_n| > |s|/2$ $(n \geq m)$. Given $\varepsilon > 0$, there is an integer $N > M$ such that $n \geq N$ implies $|s_n - s| < |s|^2\varepsilon/2$, then

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| = \left|\frac{s_n - s}{s_n s}\right| < \frac{2}{|s|^2}|s_n - s| < \varepsilon.$$

■

Proposition 3.3

(a) *Suppose* $\boldsymbol{x}_n \in \mathbb{R}^k$ *and* $\boldsymbol{x_n} = (\alpha_{1,n}, \cdots, \alpha_{k,n})$. *Then* $\{\boldsymbol{x}_n\}$ *converges to* $\boldsymbol{x} = (\alpha_1, \cdots, \alpha_k)$ *if and only if* $\lim_{n\to\infty} a_{j,n} = a_j$ *for every* j.

1. *Suppose* $\{\boldsymbol{x}_n\}$ *and* $\{\boldsymbol{y}_n\}$ *are sequences in* $\mathbb{R}^k$, $\{\beta_n\}$ *is a sequence of real numbers, and* $\boldsymbol{x}_n \to \boldsymbol{x}$, $\boldsymbol{y}_n \to \boldsymbol{y}$, $\beta_n \to \beta$. *Then* $\lim_{n\to\infty}(\boldsymbol{x}_n + \boldsymbol{y}_n) = \boldsymbol{x} + \boldsymbol{y}$, $\lim_{n\to\infty}(\boldsymbol{x}_n \cdot \boldsymbol{y}_n) = \boldsymbol{x} \cdot \boldsymbol{y}$, *and* $\lim_{n\to\infty}(\beta_n \boldsymbol{x}_n) = \beta\boldsymbol{x}$,

3.1.2 Subsequence and Subsequential Limits

Definition 3.2 (Subsequence)

Given a sequence $\{p_n\}$, *consider a sequence* $\{n_k\}$ *of positive integers such that* $n_1 < n_2 < \cdots$. *The the sequence* $\{p_{n_i}\}$ *is called a* ***subsequence*** *of* $\{p_n\}$. *If* $\{p_{n_i}\}$ *converges, its limit is called a* ***subsequential limit*** *of* $\{p_n\}$.

Remark $\{p_n\}$ converges to p if and only if every subsequences of $\{p_n\}$ converge to p.

Proposition 3.4

(a) *If* $\{p_n\}$ *is a sequence in a compact metric space* X, *then some subsequence of* $\{p_n\}$ *converges to a point of* X.

(b) *Every bounded sequence in* $\mathbb{R}^k$ *contains a convergent subsequence.*

Proof (a) Let $E = \{p_n \,|\, n \in \mathbb{N}\}$. If E is finite, there is $p \in E$ appears infinitely many times, then the subsequence consisting only p converges to $p \in X$. If E is countable, E has a limit point p in X by Lemma 2.2. Choose n_i such that $d(p, p_{n_i}) < 1/i$ and $n_i > n_{i-1}$, which exists because $N_{1/i}(p) \cap E$ contains infinitely many points. Then $\{p_{n_i}\}$ converges to p.

(b) Follows directly from (a), since E bounded means it lies in some k-cell. ■

Proposition 3.5

The subsequential limits of a sequence $\{p_n\}$ in a metric space X form a closed subset of X.

Proof Let E^* be the set of all subsequential limits, and let q be a limit point of E^*. Choose $\{q_n\} \subset E^*$ such that $d(q_n, q) < 1/n$ for all n. For every $n \in \mathbb{N}$, there exists a subsequence $\{p_{n,i}\}_{i\in\mathbb{N}}$ converging to q_n, so there exists M such that $d(p_{n_i}, q_n) < 1/n$ for $i \geq M$, choose $m_n := n_i$ such that $i \geq M$ and $m_n > m_{n-1}$. Consider the subsequence $\{p_{m_i}\}$, for each $i \in \mathbb{N}$, $d(p_{m_i}, q) \leq d(p_{m_i}, q_i) + d(q_i, q) = 2/i$. Let $\varepsilon > 0$, there exists N such that $2/N < \varepsilon$, so $d(p_{m_i}, q) \leq 2/N < \varepsilon$ for $i \geq N$, hence $p_{m_i} \to q$. ■

Definition 3.3 (Upper and Lower Limits)

Let $\{s_n\}$ be a sequence, let E be the set of subsequential limits (in the extended real number system), we define the ***upper and lower limits*** *of $\{s_n\}$ to be $s^* = \sup E$ and $s_* = \inf E$, denoted by $s^* = \limsup_{n\to\infty} s_n$ and $s_* = \liminf_{n\to\infty} s_n$.*

Proposition 3.6

Let $\{s_n\}$ be a sequence of real number, let E and s^* be defined as above, then

(a) $s^* \in E$.

(b) If $x > s^*$, there is an integer N such that $n \geq N$ implies $s_n < x$.

Moreover, s^* is the unique number with both properties. The result for s_* is analogous.

Proof (a) If $s^* = +\infty$, E is not bounded, so $s^* = +\infty \in E$. If $-\infty < s^* < +\infty$, since E is closed (Proposition 3.5), $s^* \in E$. If $s^* = -\infty$, $E = \{-\infty\}$, so $s^* = -\infty$.

(b) Assume there is $x > s^*$ such that $s_n \geq x$ for infinitely many values of n, then there is a subsequential limit y such that $y \geq x > s^*$, contradiction.

Uniqueness: Assume p, q satisfy both (a) and (b) and $p \neq q$. WLOG, let $p < q$, then there is x such that $p < x < q$. Since p satisfies (b), $s_n < x$ whenever $n \geq N$ for some N, so $q \notin E$, contradiction the fact that q satisfies (a). ■

Note *Suppose $s_n \leq t_n$ for $n \geq N$, where N is fixed, then $\liminf_{n\to\infty} s_n \leq \liminf_{n\to\inf} t_n$ and $\liminf_{n\to\infty} s_n \leq \limsup_{n\to\sup} t_n$.*

3.1.3 Cauchy Sequence

Definition 3.4 (Cauchy Sequence)

A sequence $\{p_n\}_{n\in\mathbb{N}}$ in a metric space (X, d) is a ***Cauchy sequence*** *if for all $\varphi > 0$, there exists $N > 0$ such that $m, n \geq N$ implies $d(p_m, p_n) < \varepsilon$.*

Proposition 3.7

In a metric space, every convergent sequence is a Cauchy sequence.

Proof Suppose $p_n \to p$. Let $\varepsilon > 0$ be given, there exists $N > 0$ such that $d(p_n, p) < \varepsilon/2$ if $n \geq N$. Then for $n, m \geq N$, $d(p_n, p_m) \leq d(p_n, p) + d(p, p_m) = \varepsilon/2 + \varepsilon/2 = \varepsilon$, so $\{p_n\}$ is Cauchy.

Proposition 3.8

(a) If X is a compact metric space and if $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges to some point of X.

(b) In particular, every Cauchy sequence converges in $\mathbb{R}^k$.

Proof (a) By Proposition 3.4, there exists a convergent subsequence $\{p_{n_k}\}_{k\in\mathbb{N}}$ and denote by p the point it converges to. Let $\varepsilon > 0$ be given, There exists N such that $d(p_n, p_m) < \varepsilon/2$ for $n, m \geq N$ by Cauchy condition; and there exists $M > N$ and $d(p_{n_k}, p) < \varepsilon/2$ if $n_k \geq M$, by convergence of the subsequence. For $n \geq \max\{M, N\}$, choose p_{n_k} such that $n_k > M > N$, then $d(p_n, p) \leq d(p_n, p_{n_k}) + d(p_{n_k}, p) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Hence $\{p_n\}$ converges to p.

(b) Every Cauchy sequence is bounded in $\mathbb{R}^k$: diam $E_N < 1$ for some N, so the diameter of E is at most $\max\{x_1, \cdots, x_N, x_N + 1\}$. Hence E has a bounded closure in $\mathbb{R}^k$ and the proposition then follows from (a). ∎

Remark The property that used in part (a) can be stated as: every Cauchy sequence with a convergent subsequence is convergent.

Definition 3.5 (Complete)

*A metric space in which every Cauchy sequence converges is said to be **complete**.*

Example 3.1 The set of all rational, denoted by $\mathbb{Q}$, is not complete. Consider the sequence "approaching" π.

Definition 3.6 (Monotonicity)

*A sequence $\{s_n\}$ of real numbers is said to be **monotonically increasing** if $s_n \leq s_{n+1}$ $(n = 1, \cdots)$, and it is **monotonically decreasing** if $s_n \geq s_{n+1}$ $(n = 1, \cdots)$.*

Proposition 3.9

Suppose $\{s_n\}$ is monotonic in $\mathbb{R}$. Then $\{s_n\}$ converges if and only if it is bounded.

Proof One direction follows directly from Proposition 3.1. For the other direction, without loss of generality, assume $\{s_n\}$ is monotonically increasing. Consider $E = \{s_n\}$, there exists $\alpha = \sup E$ by the l.u.b. property. Let $\varepsilon > 0$ be given, there exists $N > 0$ such that $\alpha - \varepsilon < s_N \leq \alpha$, then $\alpha - \varepsilon < s_n \leq \alpha$ for $n \geq N$ by the monotonicity. Hence $\{s_n\}$ converges to α. ∎

Chapter 4 Continuity

Introduction

- *Limit of Functions*
- *Extreme Value Theorem*
- *Intermediate Value Theorem*
- *Normed Vector Space, Banach Space*
- *Continuity*
- *Uniform Continuity*
- *Discontinuity*

4.1 Limits of Functions

Definition 4.1 (Limit of Functions)

Let X, Y be metric spaces; suppose $E \subset X$, $f : E \to Y$, and p is a limit point of E. We write $f(x) \to q$ as $x \to p$ or $\lim_{x\to p} f(x) = q$ if there is a point $q \in Y$ such that: for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$0 < d_X(x, p) < \delta \Longrightarrow d_Y(f(x), q) < \varepsilon$$

♣

Proposition 4.1

Let X, Y, E, f, p be defined as above. Then $\lim_{x\to p} f(x) = q$ if and only if $\lim_{n\to\infty} f(p_n) = q$ for every sequence $\{p_n\}$ such that $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$.

♠

Proof ($\Rightarrow$) Suppose $\lim_{x\to p} f(x) = q$ and $\varepsilon > 0$, there exists $\delta > 0$ satisfying the definition above. For every sequence $\{p_n\}$ that satisfies the above properties, there exists N such that $0 < d_X(p_n, p) < \delta$ for $n \geq N$, in which $d_Y(p_n, p) < \varepsilon$. Hence $\lim_{n\to\infty} f(p_n) = q$.

($\Leftarrow$) Suppose $\lim_{x\to p} f(x) \neq q$, there exists $\varepsilon > 0$ such that for all $\delta > 0$, there is $x \in E$ such that $0 < d_X(p, x) < \delta$ but $d_Y(q, f(x)) \geq \varepsilon$. Construct a sequence $\{p_n\}$ by choosing $\delta_n = 1/n$, then it satisfies the desired properties but $d_Y(q, f(p_n)) \geq \varepsilon$, so $\lim_{n\to\infty} f(p_n) \neq q$. ■

Proposition 4.2

If f has a limit at p, the limit is unique.

♠

Proof Since the limit of a sequence $\{p_n\}$ is unique, the proposition follows directly from Proposition 4.1. ■

Binary Operations Suppose f, g are functions defined on E to $\mathbb{R}^k$, we define addition $f + g$ by $(f + g)(x) = f(x) + g(x)$ and multiplication fg by $(fg)(x) = f(x)g(x)$. Similarly, we define $f - g$ and f/g (defined only at points x such that $g(x) \neq 0$). The scalar multiplication λf is defined by $(\lambda f)(x) = \lambda f(x)$ for all $\lambda \in \mathbb{R}$. The limit laws still holds.

Remark The change of variable in limits is stated as follows: If $x = g(t)$ is an invertible function with inverse g^{-1} in the deleted neighborhood of $t = b$, and $\lim_{t\to b} g(t) = a$, $\lim_{x\to a} g^{-1}(x) = b$, then either both the limits $\lim_{x\to a} f(x)$ and $\lim_{t\to b} f(g(t))$ exist and are equal or both of them don't exist.

4.2 Continuity

4.2.1 Continuous Functions

Definition 4.2 (Continuity)

Suppose (X, d_X) and (Y, d_Y) are metric spaces. A function $f : X \to Y$ is ***continuous at p*** *if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(x), f(p)) < \varepsilon$ for all x such that $d_X(x, p) < \delta$.*

If f is continuous at every point of X, then f is ***continuous on X****.* ♣

Proposition 4.3

Suppose $f : X \to Y$ and p is a limit point of E. Then f is continuous at p if and only if $\lim_{x\to p} f(x) = f(p)$. ♠

Remark If $p \in X$ is an isolated point, then f is continuous at $p \in X$.

Proposition 4.4 (Composition of Continuous Functions)

Suppose X, Y, Z are metric spaces, and $E \subset X$. If $f : E \to Y$ is continuous at $p \in E$, and $g : f(E) \to Z$ is continuous at $f(p)$, then $g \circ f : E \to Z$ is continuous at p. ♠

Proof Let $\varepsilon > 0$ be given. Since g is continuous, there exists $\delta > 0$ such that $d_Z(g(f(p)), g(f(q))) < \varepsilon$ if $d_Y(f(p), f(q)) < \delta$. Again since f is continuous, there exists $\lambda > 0$ such that $d_Y(f(p), f(q)) < \delta$ if $d_X(p, q) < \lambda$. Hence $d_Z((g \circ f)(p), (g \circ f)(q)) < \varepsilon$ if $d_X(p, q) < \lambda$, so $g \circ f$ is continuous by definition. ■

Proposition 4.5

A mapping $f : X \to Y$ is continuous on X if and only if $f^{-1}(V)$ is open in X for every open set V in Y. ♠

Proof ($\Rightarrow$): Suppose f is continuous and V is open in Y. For every $p \in f^{-1}(V)$, there exists $\varepsilon > 0$ such that $N_\varepsilon(f(p)) \subset V$, and by continuity of f there exists $\delta > 0$ such that $d_Y(f(p), f(q)) < \varepsilon$ if $d_X(p, q) < \delta$ for all $q \in X$. It follows that $N_\delta(p) \subset f^{-1}(V)$, i.e., p is an interior point in $f^{-1}(V)$, thus $f^{-1}(V)$ is open.

($\Leftarrow$) Given $p \in X$ and $\varepsilon > 0$, let $V = N_\varepsilon(f(p))$ be the open neighborhood of $f(p)$. By the hypothesis $f^{-1}(V)$ is open, thus there exists $\delta > 0$ such that $N_\delta(p) \subset f^{-1}(V)$. In other words, $d_Y(f(p), f(q)) < \varepsilon$ if $d_X(p, q) < \delta$, so f is continuous at p. The choice of p is arbitrary implies that f is continuous on X. ■

Example 4.1 The converse does not necessarily hold. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1/(x^2 + 1)$ is continuous on $\mathbb{R}$ (which is both open and closed). However, its image $f(\mathbb{R}) = (0, 1]$ is not open nor closed.

Corollary 4.1

A mapping $f : X \to Y$ is continuous on X if and only if $f^{-1}(C)$ is open in X for every closed set C in Y. ♡

Example 4.2 Thomae's function $f : \mathbb{R} \to \mathbb{R}$ is defined by

$$f(x) = \begin{cases} 0 & \text{if } x \notin \mathbb{Q} \\ 1/n & \text{if } x = m/n \in \mathbb{Q}, \text{ where } m \in \mathbb{Z},\, n \in \mathbb{Z}_{>0},\, m, n \text{ coprime.} \end{cases}$$

This function is continuous at irrationals and discontinuous at rationals.

The function may also be continuous at finitely many points. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x$ if $x \in \mathbb{Q}$ and $f(x) = 0$ otherwise is continuous only at $x = 0$.

4.2.2 Continuity and Compactness

Definition 4.3 (Bounded Function)

A mapping $f : E \to \mathbb{R}^k$ is said to be ***bounded*** *if there is a real number M such that $|f(x)| \leq M$ for all $x \in E$.* ♣

Proposition 4.6

Suppose $f : X \to Y$ is continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact. ♠

Proof Suppose $\{V_\alpha\}$ is an open cover of $f(X)$. Since $\{f^{-1}(V_\alpha)\}$ is an open cover of X because each $f^{-1}(V_\alpha)$ is open by Proposition 4.5, the compactness implies that there is a finite subcover $\{f^{-1}(V_i)\}_{i=1}^n$ of X. Note that hence $\{V_i\}_{i=1}^n$ is a finite subcover of $f(X)$ since $f(f^{-1})(E) \subset E$, it follows that $f(X)$ is compact. ■

Corollary 4.2

Suppose $f : X \to \mathbb{R}^k$ is continuous mapping of a compact metric space X into $\mathbb{R}^k$, then $f(X)$ is closed and bounded, and f is thus bounded. ♡

Proposition 4.7 (Extreme Value Theorem)

Suppose f is a continuous real function on a compact metric space X, and $M = \sup_{p \in X} f(p)$, $m = \inf_{p \in X} f(p)$. Then there exists points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$. ♠

Proof Since $f(X)$ is closed and bounded, hence $f(X)$ contains M and m. ■

Proposition 4.8 (Inverse of Continuous Function)

Suppose f is a continuous bijective mapping of a compact metric space X into metric space Y. Then the inverse mapping f^{-1} defined on Y by $f^{-1}(f(x)) = x$ is a continuous mapping of Y onto X.

Proof For every closed set $V \subset X$, V is compact, so $(f^{-1})^{-1}(V) = f(V)$ is compact and thus closed. Therefore, f^{-1} is continuous by Corollary (4.1). ∎

Definition 4.4 (Uniform Continuity)

Suppose $f : X \to Y$ be a mapping of metric spaces, f is said to be ***uniformly continuous*** *on X if for every $\varepsilon > 0$, there exists $\delta > 0$ such that $d_Y(f(p), f(q)) < \varepsilon$ for all $p, q \in X$ such that $d_X(p, q) < \delta$.*

Example 4.3 Consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. Let $\varepsilon > 0$, given $\delta > 0$, let $p = 1/\delta$ and $q = \delta/2 + 1/\delta$. Then $|p - q| = \delta/2 < \delta$, but

$$|f(p) - f(q)| = |1/\delta^2 - (\delta^2/4 + 1 + 1/\delta^2)| = 1 + \delta^2 > \varepsilon,$$

so f is not uniformly continuous on $\mathbb{R}$. Note that the issue is that $\mathbb{R}$ is not compact.

Proposition 4.9

Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is uniformly continuous on X.

Proof Let $\varepsilon > 0$ be given, choose $\delta_p > 0$ such that $d_X(p, q) < \delta_p \Rightarrow d_Y(f(p), f(q)) < \varepsilon/2$. Since X is compact, there exists a finite cover of neighborhoods $\{N_{\delta_i/2}(p_i)\}$. Put $\delta = \min \delta_i/2$, then for all p, q such that $d_X(p, q) < \delta$, there exists p_i such that $p \in N_{\delta_i/2}(p_i)$, then $p, q \in N_{\delta_i}(p_i)$. Then $d_Y(f(p), f(q)) \le d_Y(f(p), f(p_i)) + d_Y(f(p_i), f(q)) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Hence f is uniformly continuous. ∎

Example 4.4 The compactness is essential. The continuous function f is not necessarily uniformly continuous even it is bounded. Consider $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = \sin(1/x)$, and $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = \sin(x^2)$. f and g are both bounded and continuous, yet they are not uniformly continuous.

4.2.3 Continuity and Connectedness

Proposition 4.10

Suppose $f : X \to Y$ where X, Y are metric spaces. If E is a connected subset of X, then $f(E)$ is connected.

Proof Proof by contrapositive. Suppose $f(E)$ is not connected and A, B forms a separation of $f(E)$. Let $A' = f^{-1}(A) \cap E$ and $B' = f^{-1}(B) \cap E$, then $E = A' \cup B'$. A', B' are nonempty because $A, B \subset f(E)$ are nonempty, and $\overline{A'} \cap B' \subset f^{-1}(\overline{A} \cap B) = \varnothing$ (WLOG, $A \cap \overline{B'} = \varnothing$). Therefore, A', B' form a separation of E, so E

is not connected. ■

Proposition 4.11 (Intermediate Value Theorem)

Let $f : [a,b] \to \mathbb{R}$ be continuous. If $f(a) < f(b)$ and $f(a) < c < f(b)$, then there exists a point $x \in (a,b)$ such that $f(x) = c$.

Proof Since $[a,b]$ is connected, $f([a,b])$ is connected, so $c \in f([a,b])$ by connectedness. ■

4.3 Discontinuity, Monotonicity

Definition 4.5 (One-sided Limit)

Let f be defined on (a, b). Consider any point x such that $a \le x < b$, we write $f(x+) = q$ if $f(t_n) \to q$ as $n \to \infty$ for all sequences $\{t_n\}$ in (x, b) converging to x. The definition of $f(x-)$ is analogous. ♣

Remark The limit of f at x exists if and only if the one-sided limits coincide, namely $f(x+) = f(x-)$; in this case, $\lim_{t\to x} f(x) = f(x+) = f(x-)$.

Definition 4.6 (Discontinuity)

*Let f be defined on (a, b). If f is discontinuous at a point x, and if $f(x+)$ and $f(x-)$ exist, then f is said to have a discontinuity of the **first kind** (or a simple discontinuity). Otherwise the discontinuity is said to be of the **second kind**.* ♣

Remark There is two types of simple discontinuity: (a) $f(x+) \ne f(x-)$ (removable discontinuity), and (b) $f(x+) = f(x-) \ne f(x)$ (jump discontinuity).

Definition 4.7 (Monotonicity)

*Let $f : (a, b) \to \mathbb{R}$, then f is said to be **monotonically increasing** on (a, b) if $a < x < y < b$ implies $f(a) \le f(b)$. The definition of monotonically decreasing function is analogous.* ♣

Proposition 4.12

Let f be monotonically increasing on (a, b). Then

(a) $f(x+)$ and $f(x-)$ exist at every point of $x \in (a, b)$.

(b) $\sup_{a<t<x} f(t) = f(x-) \le f(x) \le f(x+) = \inf_{x<t<b} f(t)$.

(c) If $a < x < y < b$, then $f(x+) \le f(y-)$.

Analogous results hold for monotonically decreasing functions. ♠

Proof (a) Consider $S = \{f(t) \,|\, a < t < x\}$, there exists $A := \sup S$ since S is nonempty and bounded above by $f(x)$. Let $\varepsilon > 0$ be given, there exists $t_0 \in (a, x)$ such that $A - \varepsilon < f(t_0) \le A$. Put $\delta = x - t$, then $|A - f(t)| < |A - f(t_0)| < \varepsilon$ if $|x - t| < \delta$. Thus, $f(x-) = A$ exists, and $f(x+)$ exists WLOG.

(b) By the definition of $f(x-)$ in Part (a), $\inf_{a<t<x} f(t) = f(x-)$, and $f(x-) \le f(x)$ holds by monotonicity. The inequality for $f(x+)$ holds WLOG.

(c) This assertion follows directly from the inequality $f(x+) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t) \le \sup_{x<t<y} f(t) = \sup_{x<t<y} f(t) = f(y-)$. ■

Note ***Corollary***: *Monotonic functions have no discontinuities of the second kind.*

Proposition 4.13

Let $f : (a, b) \to \mathbb{R}$ be monotonic real function, then the set of points at which f is discontinuous is at most countable.

Proof WLOG, assume f is monotonically increasing and $E = \{x \in (a, b) \mid f \text{ is discontinuous at } x\}$. Since f is increasing, $f(x-) < f(x+)$ if $x \in E$, then there exists $r_x \in \mathbb{Q}$ such that $f(x-) < r_x < f(x+)$. Define $\varphi : E \to \mathbb{Q}$ by $\varphi(x) = r_x$, then φ is clearly injective since $f(x+) \leq f(y-)$ if $x < y$. Therefore, $E \sim f(E) \subset \mathbb{Q}$, so E is at most countable. ■

4.4 Normed Vector Spaces

Definition 4.8 (Norm, Normed Vector Spaces)

*A **norm** on a vector space V is a function $\|\cdot\| : V \to [0, \infty)$ satisfying*

(i) (positivity) $0 \le \|x\| < \infty$ for all $x \in V$

(ii) (definiteness) $\|x\| = 0$ if and only if $x = 0$

(iii) (scalar multiplication) $\|\alpha x\| = |\alpha|\|x\|$ for all scalar α and $x \in V$.

(iv) (triangle inequality) $\|x + y\| \le \|x\| + \|y\|$ for all $x, y \in V$.

*The pair $(V, \|\cdot\|)$ is called a **normed vector space**.* ♣

A function $\|\cdot\| : V \to [0, \infty)$ satisfying all properties above except (ii) is called a ***pseudonorm*** on V.

If $(V, \|\cdot\|)$ is a normed vector space, then the function $d : X \times X \to [0, \infty)$ is defined by $d(x, y) := \|x - y\|$ is a metric on V. This is called the usual metric or ***induced metric*** on V.

Definition 4.9 (Convergence)

*Suppose $\{x_n\}$ is a sequence in a normed vector space $(V, \|\cdot\|)$. The series $\sum_{i=1}^{\infty} x_i$ is said to **converge** if the sequence of partial sums $\{s_n\}$, where $s_n = \sum_{i=1}^{n} x_i$, converges to some $x \in V$ in the sense that $\lim_{n\to\infty}\|x - \sum_{i=1}^{n} x_n\| = 0$. In this case, we write $\sum_{i=1}^{\infty} x_n = x$.* ♣

Definition 4.10 (Banach Space)

*A **Banach space** is a normed vector space which is complete with respect to the induced metric.* ♣

Proposition 4.14

A normed vector space $(V, \|\cdot\|)$ is Banach (namely complete) if and only if a series $\sum_{i=1}^{\infty} x_i$ converges whenever $\sum_{i=1}^{\infty}\|x_i\|$ converges. ♠

Proof ($\Rightarrow$): Let $S_n = \sum_{i=1}^{n} x_i$ and $T_n = \sum_{i=1}^{n}\|x_i\|$, suppose $\{T_n\}$ converges. Let $\varepsilon > 0$ be given. Since $\{T_n\}$ is Cauchy, so there exists $N > 0$ such that $n > m \ge N$ implies $|T_n - T_m| < \varepsilon$. Then

$$\|S_n - S_m\| = \left\|\sum_{i=n+1}^{m} x_i\right\| \le \sum_{i=n+1}^{m}\|x_i\| = |T_n - T_m| < \varepsilon.$$

Hence $\{S_n\}$ is Cauchy and thus converges since X is complete.

($\Leftarrow$): Let $\{x_n\}$ be Cauchy in X. For each $i \in \mathbb{Z}_{>0}$, choose N_i such that $N_i > N_{i-1}$ for which $n > m \ge N_i$ implies that $\|x_n - x_m\| \le 1/2^i$. Define $y_i = x_{n_{i+1}} - x_{n_i}$, then $\|y_i\| = \|x_{n_{i+1}} - x_{n_i}\| \le 1/2^i$, so $\sum_{i=1}^{k}\|y_i\|$ converges, followed by $\sum_{i=1}^{\infty} y_i$ converges. Note that $\sum_{i=1}^{n} y_i = x_{n_i+1} - x_{n_1}$, then $\{x_{n_i}\}$ is convergent. By the

Cauchy condition and the convergence of $\{x_{n_i}\}$, it is not hard to show that $\{x_n\}$ converges to $\lim_{i\to\infty} x_{n_i}$ using triangle inequality. ■

Definition 4.11 (Equivalent Norms)

*Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a vector space X are called **equivalent** if there exists $c_1, c_2 > 0$ such that $c_1\|x\|_2 \leq \|x\|_1 \leq c_2\|x\|_2$ for all $x \in X$.* ♣

Proposition 4.15

All norms on a finite dimensional vector space X are equivalent. ♠

Proof Suppose $\{e_1, \cdots, e_n\}$ is a basis of X, define $\|\sum_{i=1}^n a_i e_i\|_1 := \sum_{i=1}^n |a_i|$, and put $S := \{u \in X \mid \|u\|_1 = 1\}$. Given $\|\cdot\|_2$, we can define $f : (X, \|\cdot\|_1) \to \mathbb{R}$ by the equality $f(x) = \|x\|_2$. We now want to show f is continuous and S is compact, thus, it follows that $\operatorname{im} f|_S = \{\|u\|_2 \mid u \in S\}$ has a maximum and a minimum by the extreme value theorem. Then for $x \in X$, we can put $u = x/\|x\|_1$. As shown above, using the above inequality, multiplying by $\|x\|_1$ yields the desired result $c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_2$. ■

Chapter 5 Differentiation

5.1 Differentiation and Mean Value Theorems

Introduction

- *Differentiation*
- *Operations and Chain Rule*
- *Darboux*
- *Mean Value Theorems*
- *Taylor's Theorem*

5.1.1 Differentiation

Definition 5.1 (Differentiable)

For function $f : [a, b] \to \mathbb{R}$, we say f is differentiable at $x \in [a, b]$ if the limit of $\phi(t) := [f(t) - f(x)]/(t - x)$ exists when $t \to x$, i.e, the limit

$$\lim_{t \to x} \phi(t) = \lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

exists. In this case, we denote by the limit $f'(x) := \lim_{t \to x} \phi(t)$. ♣

Proposition 5.1

If $f : [a, b] \to \mathbb{R}$ is differentiable at $x \in [a, b]$, then f is continuous at x. ♠

Proof Suppose f is differentiable at x, then $\lim_{t \to x}(f(t) - f(x)) = \lim_{t \to x} \phi(t) \cdot (t - x) = f'(x) \lim_{t \to x}(t - x) = 0$, so f is continuous.

Property *Suppose f, g are real-valued functions differentiable at x, then $f + g$, fg, and f/g are differentiable at x, and*

(a) $(f + g)'(x) = f'(x) + g'(x)$,

(b) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$,

(c) $(f/g)'(x) = [f'(x)g(x) - f(x)g'(x)]/g(x)^2$ *if* $g(x) \neq 0$.

Theorem 5.1 (Chain Rule)

Suppose f is continuous on $[a, b]$ and differentiable at $x \in [a, b]$, and g is defined on $f([a, b])$ and differentiable at $f(x)$. If $h(t) = g(f(t))$, then h differentiable at x and

$$h'(x) = g'(f(x))f'(x).$$

♡

Proof Note that $f'(x)$ and $g'(f(x))$ exists by the differentiability, so

$$h'(x) = \lim_{t\to x} \frac{g(f(t)) - g(f(x))}{t-x} = \lim_{t\to x} \frac{g(f(t)) - g(f(x))}{f(t)-f(x)} \frac{f(t)-f(x)}{t-x} = g'(f(x))f'(x).$$

Proposition 5.2 (Derivative of Inverse Function)

Let $f : X \to Y$ ($X, Y \subseteq \mathbb{R}$) be an invertible function that is differentiable at $p \in E$. Suppose that $f^{-1} : F \to E$ is continuous at $q := f(p)$ and that $f'(p) \neq 0$. Then f^{-1} is differentiable at $q = f(p)$, and we have $(f^{-1})'(q) = 1/f'(p)$. ♠

5.1.2 Mean Value Theorems

Definition 5.2 (Local Extrema)

*Let f be a real function on a metric space X. We say that f has a **local maximum** at $p \in X$ if there exists $\delta > 0$ such that $f(q) \leq f(p)$ for all $q \in X$ with $d(p, q) < \delta$. Local minimums are defined likewise.* ♣

Proposition 5.3 (Rolle's Theorem)

Let f be defined on $[a, b]$; if f has a local maximum at a point $x \in (a, b)$ and if $f'(x)$ exists, then $f'(x) = 0$. The analogous statement for local minima also holds. ♠

Proof Since $f'(x)$ exists, $\lim_{t\to x} \phi(x)$ exists thus $\phi(x+)$ and $\phi(x-)$ exists. Note that $f(t) - f(x) \leq 0$ for all t, it follows that $\phi(x+) \leq 0$ and $\phi(x-) \geq 0$. Hence the existence of $f'(x)$ implies that $f'(x) = \lim_{t\to x} \phi = 0$. ■

Theorem 5.2 (Cauchy Mean Value Theorem)

If f and g are continuous real functions on $[a, b]$ which are differentiable in (a, b), then there is a point $x \in (a, b)$ at which

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x).$$

♡

Remark For non-degenerated cases, the condition is equivalent to: there exists x such that $g'(x)/f'(x) = [g(b) - g(a)]/[f(b) - f(a)]$.

Proof We may assume $f(b) - f(a) \neq 0$, otherwise the results follows directly from 5.3. Define $s(x) = f(x)/[f(b) - f(a)]$ and $t(x) = g(x)/[g(b) - g(a)]$, then $s(b) - s(a) = t(b) - t(a) = 1$. Notice that $(s - t)(b) = (s - t)(a)$, then by Rolle's Theorem, $(s - t)'(x) = 0$ for some $x \in (a, b)$, then $s'(x) = t'(x)$. Hence $g'(x)/f'(x) = [g(b) - g(a)]/[f(b) - f(a)]$. ■

Corollary 5.1 (Mean Value Theorem)

If f is a real continuous function on $[a,b]$ which is differentiable in (a,b), then there is a point $x \in (a,b)$ at which $f(b) - f(a) = (b-a)f'(x)$.

Proof Follows immediately from Cauchy MVT by taking $g(x) = x$. ■

Proposition 5.4

Suppose f is differentiable in (a,b). If $f'(x) \geq 0$ for all $x \in (a,b)$, then f is monotonically increasing; if $f'(x) = 0$, then f is constant; and if $f'(x) \leq 0$, then f is monotonically decreasing.

Proof Suppose $x_1 < x_2$, then $f(x_2) - f(x_1) = (x_2 - x_1)f'(x)$ for some $x \in (x_1, x_2)$ by MVT. The assertion follows immediately. ■

Proposition 5.5 (Darboux)

Suppose $f : [a,b] \to \mathbb{R}$ is differentiable, and $f'(a) < \lambda < f'(b)$. Then there exists $x \in (a,b)$ such that $f'(x) = \lambda$.

Proof Let $g(x) = f(x) - \lambda t$. Note that $g'(a) < 0 < g'(b)$, there exists t_1, t_2 such that $g(t_1) < g(a)$ and $g(t_2) < g(b)$, so $g(a)$ and $g(b)$ are not the absolute minimum. Then minimum is attained at some $x \in [t_1, t_2] \subset (a,b)$, so $g'(x) = 0$ and thus $f'(x) = \lambda$. ■

Corollary 5.2

If f is differentiable on $[a,b]$, then f' cannot have any simple discontinuity on $[a,b]$.

Example 5.1 The function f can be differentiable on $[a,b]$ but still have second kind of discontinuity. Suppose

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}, \quad \text{then } f'(x) = \begin{cases} 2x\sin(1/x) - \cos(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}.$$

f' is differentiable and has a second kind of discontinuity at $x = 0$ since $f'(0+)$ and $f'(0-)$ do not exist.

5.2 Derivative of Higher Order, Vector-Valued Functions

Theorem 5.3 (Taylor's Theorem)

Suppose $f : [a, b] \to \mathbb{R}$, $n \in \mathbb{Z}_{>0}$, $f^{(n-1)}$ *is continuous on* $[a, b]$, *and* $f^{(n)}$ *exists for every* $t \in (a, b)$. *Let* α, β *be distinct points of* $[a, b]$, *and define*

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k.$$

Then there exists $x \in (\alpha, \beta)$ *such that*

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta-\alpha)^n.$$

In general, the theorem shows that f *can be approximated by a polynomial of degree* $n-1$, *and it allows us to estimate the error, if we know bounds on* $|f^{(n)}(x)|$.

Proof Define $g(t) = f(t) - P(t) - M(t-\alpha)^n$, where M is defined for which $f(\beta) - P(\beta) + M(\beta - \alpha)$. Note that $g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0$ and $g(\beta) = 0$. Then there exists $x_1 \in (\alpha, \beta)$ such that $g'(x_1) = 0$ by MVT; continuing in this manner, we obtain $x_i \in (\alpha, x_{i-1})$ such that $g^{(i)}(x_i) = 0$. Therefore, $g^{(n)}(x_n) = 0$, thus $M = f'(x_n)/n!$.

Example 5.2 The Mean Value Theorem does not hold explicitly for vector-valued functions. Consider $F(t) = (\cos t, \sin t)$. $F(2\pi) - F(0) = (0, 0)$, but $2\pi F'(t) = 2\pi(-\sin t, \cos t) \neq (0, 0)$. It follows that $F'(t) \neq [F(2\pi) - F(0)]/(2\pi - 0)$ for all t. However, the following generalization holds.

Proposition 5.6

Suppose $\boldsymbol{f}$ *is a continuous mapping of* $[a, b]$ *into* $\mathbb{R}^k$ *and* $\boldsymbol{f}$ *is differentiable in* (a, b), *then there exists* $x \in (a, b)$ *such that* $|\boldsymbol{f}(b) - \boldsymbol{f}(b)| \leq (b-a)|\boldsymbol{f}'(x)|$.

Proof If $f(b) - f(a) = 0$, the inequality holds immediately. Suppose $z = f(b) - f(a) \neq 0$, define $\varphi(t) = z \cdot f(t)$, then φ is a real-valued function differentiable on (a, b). By MVT, $\varphi(b) - \varphi(a) = (b-a)\varphi'(x)$ for some x, so $|z|^2 = z \cdot (f(b) - f(a)) = (b-a)z \cdot f'(x)$. Then

$$|z|^2 = |(b-a)z \cdot f'(x)| \leq (b-a)|z||f'(x)|,$$

where the inequality holds by Cauchy-Schwartz. Therefore $|z| \leq (b-a)|f'(x)|$. ■

Chapter 6 Sequences and Series of Functions

Introduction

- *Pointwise Convergence, Uniform Convergence*
- *Uniform Convergence Properties*
- *Criteria of Uniform Convergence*
- *Equicontinuous Family*

6.1 Uniform Convergence

Definition 6.1 (Convergence of Sequence of Functions)

*Suppose $E \subset X$ where X is a metric space and $\{f_n\}$ is a sequence of complex-valued functions defined on E. Define a function f by $f(x) = \lim_{n\to\infty} f_n(x)$, then f is the **limit function** of $\{f_n\}$, and $\{f_n\}$ is said to **converges to f pointwise**.*

♣

Example 6.1 The double limit of a sequence of continuous function is not interchangeable, i.e., in general,

$$\lim_{t\to x}\lim_{n\to\infty} f_n(x) \neq \lim_{n\to\infty}\lim_{t\to x} f_n(x).$$

Suppose $m, n \in \mathbb{Z}_{>0}$, let $s_{m,n} = m/(m+n)$. Then for every fixed n, $\lim_{m\to\infty} s_{m,n} = 1$, so that $\lim_{n\to\infty}\lim_{m\to\infty} s_{m,n} = 1$. On the other hand, for every fixed m, $\lim_{n\to\infty} s_{m,n} = 0$ so that $\lim_{m\to\infty}\lim_{n\to\infty} s_{m,n} = 0$.

Definition 6.2 (Uniform Convergence)

*A sequence of functions $\{f_n : E \to \mathbb{R}\}$ is said to **converge uniformly** to f on E if for every $\varepsilon > 0$ there is an integer N such that $n \geq N$ implies $|f_n(x) - f(x)| < \varepsilon$ for all $x \in E$.*

♣

Remark The difference between pointwise convergence and uniform convergence is that N depends only on $\varepsilon > 0$ in uniform convergence, and N depends on $\varepsilon > 0$ and $x \in E$ in pointwise convergence.

Proposition 6.1 (Uniformly Cauchy Criterion)

The sequence of functions $\{f_n\}$ converges uniformly on E if and only if for every $\varepsilon > 0$ there exists an integer N such that $m, n \geq N$ and $x \in E$ implies $|f_n(x) - f_m(x)| < \varepsilon$.

♠

Proof ($\Rightarrow$) Suppose $\{f_n\}$ converges uniformly to f. Let $\varepsilon > 0$. There exists $N > 0$ such that $n \geq N$ implies $|f_n(x) - f(x)| < \varepsilon/2$ for all $x \in E$, so $|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon$.

($\Leftarrow$) Suppose $\{f_n\}$ is uniformly Cauchy. The sequence converges pointwise to some f because $\{f_n(x)\}$ is Cauchy for all $x \in E$. Let $\varepsilon > 0$ be given, let N be chosen so that $|f_n(x) - f_m(x)| < \varepsilon/2$. Fix n and let $m \to \infty$, then $|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \varepsilon/2 < \varepsilon$ for all x, so $\{f_n\}$ converges uniformly. ■

Proposition 6.2

Suppose $\lim_{n\to\infty} f_n(x) = f(x)$ *(*$x \in E$*), put* $M_n = \sup_{x\in E} |f_n(x) - f(x)|$. *Then* $f_n \to f$ *uniformly on* E *if and only if* $M_n \to 0$ *as* $n \to 0$.

Proposition 6.3 (Weierstrass M-test)

Suppose $\{f_n\}$ *is a sequence of functions, and* $|f_n(x)| \leq M_n$. *Then* $\sum f_n$ *converges uniformly on* E *if* $\sum M_n$ *converges.*

Proof $\{M_n\}$ is convergent and thus Cauchy. Let $\varepsilon > 0$. For n, m sufficiently large, $|\sum_{i=m}^{n} f_n(x)| \leq \sum_{i=m}^{n} M_n < \varepsilon$ for all $x \in E$, so $\{\sum_{i=1}^{n} f_n\}$ is uniformly Cauchy. Then $\sum f_n$ is uniformly convergent. ■

6.2 Uniform Convergence, Continuity, and Differentiation

6.2.1 Uniform Convergence and Continuity

Theorem 6.1

Suppose $f_n \to f$ uniformly on E in a metric space. Let x be a limit point of E, and suppose that $A_n := \lim_{t\to x} f_n(t)$. Then $\{A_n\}$ converges, and $\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n$. That is,

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

♡

Proof (a) Let $\varepsilon > 0$. For sufficiently large N, $n \geq m \geq N$ implies $|f_n(t) - f_m(t)| < \varepsilon/2$ for all $t \in E$, so $|A_n - A_m| \leq \lim_{t\to x} |f_n(t) - f_m(t)| \leq \varepsilon/2 < \varepsilon$. This implies that $\{A_n\}$ is Cauchy, so it converges.

(b) Let $A := \lim_{n\to\infty} A_n$ and $f_n \to f$ uniformly. Let $\varepsilon > 0$ be given. Notice that

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|.$$

For all $t \in E$, for sufficiently large N, we have $|f(t) - f_n(t)| < \varepsilon/3$ by the uniform convergence given that $f_n \to f$ uniformly; $|f_n(t) - A_n| < \varepsilon/3$ by the definition; and $|A_n - A|$ by the definition of A. Therefore, $\lim_{t\to x} f(t) = A$, as desired. ■

Corollary 6.1

Suppose the sequence $\{f_n\}$ is continuous on E for each n, and $f_n \to f$ uniformly on E, then f is continuous on E.

♡

Remark The converse does not hold. Consider the below example where we let f_n be defined on $(0, 1)$. Then f is continuous but f_n does not converge uniformly.

Example 6.2 Suppose $f_n : [0, 1] \to \mathbb{R}$ is defined by $f_n(x) = x^n$. The sequence $\{f_n\}$ converges pointwise to

$$f(x) = \begin{cases} 0 & \text{if } x \neq 1 \\ 1 & \text{if } x = 1 \end{cases}.$$

We see that $\lim_{t\to 1}\lim_{n\to\infty} f_n(t) = \lim_{t\to 1} f(t) = 1$, whereas $\lim_{n\to\infty}\lim_{t\to 1} f_n(t) = \lim_{n\to\infty} 1 = 1$. Indeed, there exists x such that $x^n > 1/2$ for all n by intermediate value theorem, then $M_n = \sup_{x\in E} |f_n(x) - 0| > 1/2$. It implies that M_n does not converge to 0, so $\{f_n\}$ does not converge to f uniformly.

Proposition 6.4

Suppose K is compact, and

(i) $\{f_n\}$ is a sequence of continuous functions on K,

(ii) $\{f_n\}$ converges pointwise to a continuous function f on K,

(iii) $f_n(x) \geq f_{n+1}(x)$ for all $x \in K$.

Then $f_n \to f$ uniformly on K.

Proof Set $g_n = f_n - f$ for each $n \in \mathbb{Z}_{>0}$, then g_n are continuous and $g_n \to 0$ pointwise. Let $\varepsilon > 0$ be given. Let $K_n = \{x \in X \mid g_n(x) \geq \varepsilon\}$, K_n is closed because f is continuous, thus K_n is compact by Proposition 2.13. Note that $g_n(x) \geq g_{n+1}(x)$, so $K_n \supset K_{n+1}$. Since $g_n(x) \to 0$, $x \notin K_n$ for sufficiently large n, so $\bigcap K_n = \varnothing$. Then there exists $K_n = \varnothing$ by the finite intersection property, so $|g_n(x)| < \varepsilon$ for all $x \in X$. ■

Example 6.3 Compactness is essential in the assumption of the above proposition. Consider $f_n = 1/(nx+1)$ defined on $(0,1)$, the sequence $\{f_n\}$ converges to 0. It satisfies all of the three conditions, but f_n does not converge to 0 uniformly.

Definition 6.3 ($\mathscr{C}(X)$)

If X is a metric space, then $\mathscr{C}(X)$ denotes the set of all complex-valued continuous bounded functions with domain X.

Note *$\mathscr{C}(X)$ is a normed vector space (over $\mathbb{C}$) by associating the supremum norm $\|f\| = \sup_{x\in X}|f(x)|$ to each function f.*

Proposition 6.5

The metric induced by the supremum norm makes $\mathscr{C}(X)$ into a complete metric space.

Proof Suppose $\{f_n\}$ is Cauchy in $\mathscr{C}(X)$. It is uniformly Cauchy, so it converges uniformly to a function f by Proposition 6.1. The continuity of f follows from Corollary 6.1. Since there exists N such that $\|f_n - f\| < 1$ if $n \geq N$, then f is bounded by $\|f_n\| + 1$. ■

6.2.2 Uniform Convergence and Differentiation

The goal is to investigate the relationships between differentiability and uniform convergence. Suppose $\{f_n\}$ is a sequence of differentiable function on $[a,b] \subset \mathbb{R}$, and suppose $f_n \to f$ pointwise (or uniformly). The questions are

(a) Is the limit function f differentiable?

(b) If f differentiable on $[a,b]$, do we have $f_n'(x) \to f'(x)$ for $x \in [a,b]$?

Example 6.4 Define $f_n(x) = x^n$ for $x \in [0,1]$. The limit function is given by $f(x) = 1$ if $x = 1$ and $f(x) = 0$ otherwise. f is clearly not differentiable at $x = 0$ since it is not continuous, so (a) fails if the convergence is pointwise.

Define $f_n(x) = \sqrt{x^2 + 1/n}$ for $x \in [-1,1]$. The limit function is given by $f(x) = |x|$. Although $f_n \to f$ uniformly, f is still not differentiable at $x = 0$, so (a) fails even under uniform convergence.

Example 6.5 Define $f_n : [-1,1] \to \mathbb{R}$ for each $n \in \mathbb{Z}_{>0}$ by $f_n(x) = x/[1+(n-1)x^2]$. If $x = 0$, $f_n(0) = 0$ for all n; on the other hand, $\lim_{n\to\infty} f_n(x) = 0$ for each fixed x. The limit function is thus $f(x) = 0$, and it is differentiable. However,

$$f_n'(0) = \lim_{h\to 0}\frac{f_n(h)-f_n(0)}{h} = \lim_{h\to 0}\frac{h/[1+(n-1)h^2]}{h} = \lim_{h\to 0}\frac{1}{1+(n-1)h^2} = 1,$$

so f_n' does not converge to f'.

Proposition 6.6

There exists a real-valued continuous function which is nowhere differentiable.

Example 6.6 (Weierstrass) Define $\varphi(x) = |x|$ for $-1 \le x \le 1$ and let $\varphi(x+2) = \varphi(x)$, then $\varphi(x)$ is continuous on $\mathbb{R}$. Define the function (by a series of fractal sawtooth)

$$f = \sum_{n=0}^{\infty} f_n := \sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \varphi(4^n x).$$

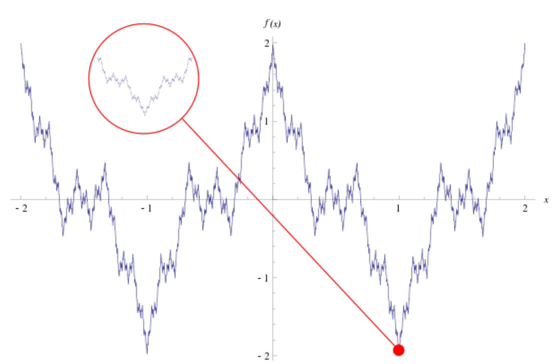
By Weierstrass M-test, the series $\sum f_n$ convergence uniformly to f, so f is continuous.

Fix $x \in \mathbb{R}$ and $m \in \mathbb{Z}_{>0}$, put $\delta_m = \pm 4^{-m}/2$, so there is no integer between $4^m x$ and $4^m(x+\delta_m)$. Define $\gamma_n = [\varphi(4^n(x+\delta_m)) - \varphi(4^n x)]/\delta_m$. Note that $|\gamma_n| \le 4^n$ because $|\varphi(s)-\varphi(t)| \le |s-t|$; in particular, if $n = m$, $|\gamma_n| = 4^n$, and if $n > m$, $\gamma_n = 0$ because $4^n\delta_m$ is even. Then

$$\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \gamma_n\right| \ge 3^m - \sum_{n=0}^{m-1} 3^n = \frac{1}{2}(3^m+1).$$

Since $\delta_m \to 0$ when $m \to \infty$, the limit of the above expression does not exist, it follows that f is nowhere differentiable on $\mathbb{R}$.

The Weierstrass function is defined as Fourier series: $f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$, where $0 < x < 1$.



Theorem 6.2

Suppose $\{f_n\}$ is a sequence of differentiable functions on $[a,b]$, such that $\{f_n(x_0)\}$ converges for some $x_0 \in [a,b]$. If $\{f_n'\}$ converges uniformly on $[a,b]$, then $\{f_n\}$ converges uniformly on $[a,b]$ to a function f, and

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

Proof (a) Let $\varepsilon > 0$ be given. Since $\{f_n(x_0)\}$ is convergent and thus Cauchy, there exists N_1 such that $n, m \geq N_1$ implies $|f_n(x_0) - f_m(x_0)| < \varepsilon/2$. Since $\{f_n'\}$ converges uniformly and thus uniformly Cauchy, there exists N_2 such that $n, m \geq N_2$ implies $|f_n'(x) - f_m'(x)| < \varepsilon/2(b-a)$ for all x.

Let $N = \max\{N_1, N_2\}$. For $n, m \geq N$, by Mean Value Theorem, there exists $x \in (x, t)$ such that

$$|(f_n - f_m)(x) - (f_n - f_m)(t)| = |(f_n' - f_m'))(c)(x-t)| < \frac{\varepsilon}{2(b-a)}|x-t| \leq \frac{\varepsilon}{2}.$$

Then the triangle inequality implies that

$$|f_n(x) - f_m(x)| \leq |(f_n - f_m)(x) - (f_n - f_m)(t)| + |f_n(x_0) - f_m(x_0)| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Hence $\{f_n\}$ is uniformly Cauchy and thus converge uniformly to some function f.

(b) Let f be the limit function of $\{f_n\}$, for fixed x, we define $\phi_n(t) := [f_n(t) - f_n(x)]/(t-x)$ for $t \neq x$, and define $\phi(t) = [f(t) - f(x)]/(t-x)$. As shown above, for $n, m \geq N$,

$$|\phi_n(t) - \phi_m(t)| < \frac{\varepsilon}{2(b-a)},$$

so $\{\phi_n\}$ converges uniformly, for $t \neq x$. Since $\{f_n\}$ converges to f, we conclude that $\lim_{n\to\infty} \phi_n(t) = \phi(t)$. Applying Theorem 6.1 yields $\lim_{n\to\infty} f_n'(x) = \lim_{t\to x} \lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{t\to x} \phi(t) = f'(x)$. ■

6.3 Equicontinuous Families of Functions

Definition 6.4 (Pointwise bounded, Uniformly bounded)

Let $\{f_n\}$ be a sequence of functions define on a set E, we say that $\{f_n\}$ is ***pointwise bounded*** *on E if $\{f_n(x)\}$ is bounded for every $x \in E$, i.e., $|f_n(x)| < \phi(x)$ for all n and some finite-valued function ϕ.*

We say that $\{f_n\}$ is ***uniformly bounded*** *on E if there exists M such that $|f_n(x)| < M$ for all n and $x \in E$.* ♣

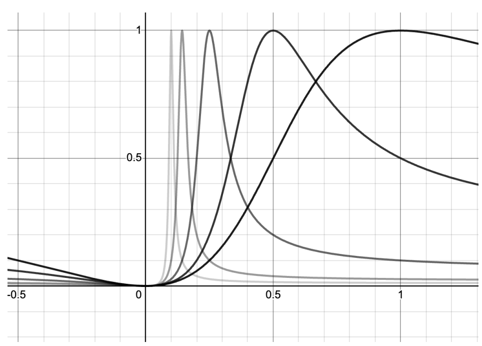
Remark In $\mathbb{C}$, every bounded sequence contains a convergent subsequence. However, the generalization fails to hold on the set of functions:

(i) It is not generally true that every sequence $\{f_n\}$ of bounded continuous functions (even if uniformly bounded on a compact set) contains a pointwise convergent subsequence. For instance, consider $f_n(x) = \sin nx$ on $[0, 2\pi]$.

However, a desired subsequence exists on a countable subset E_1 of E for the sequence of pointwise bounded functions. (See Proposition 6.7)

(ii) It is not generally true that every convergent sequence of functions $\{f_n\}$ (even if uniformly bounded on a compact set) contains a uniformly convergent subsequence? (See Example 6.7)

Example 6.7 Let $f_n(x) = x^2/[x^2+(1-nx)^2]$ for $x \in [0,1]$. $\{f_n\}$ is uniformly bounded on $[0,1]$ since $|f_n(x)| \le 1$ for all n and x, and $\lim_{n\to\infty} f_n(x) = 0$. However, $f_n(1/n) = 1$ for all n, so that no subsequence converge uniformly on $[0,1]$.



The concept needed in this connection is "equicontinuity".

Proposition 6.7

If $\{f_n\}$ is a pointwise bounded sequence of complex functions on a countable set E, then $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{f_{n_k}\}$ converges for every $x \in E$. ♠

Proof Suppose $E = \{x_i\}_{i\in\mathbb{Z}_{>0}}$. Note that $\{f_n(x_1)\}$ is bounded in $\mathbb{C}$, there is a subsequence $S_1 := \{f_{1,j}\}_j$ such that $\{f_{1,j}(x_1)\}_j$ is convergent. We define $S_i = \{f_{i,j}\}$ recursively as follows, for every $i > 1$, $S_{i-1}(x) = \{f_{i-1,j}(x_i)\}_j$ is bounded and infinite, so there is a subsequence $S_i := \{f_{i,j}\}_{j\in\mathbb{Z}_{>0}}$ of S_{i-1} for which converges at x_i.

Consider the subsequence $S = \{f_{i,i}\}$ (diagonal process). Note that S is a subsequence of S_i except for the first $i-1$

terms, so $S(x_i)$ converges for every $i = 1, \cdots$. ■

Definition 6.5 (Equicontinuity)

A family $\mathscr{F}$ of complex functions f defined on a set E in a metric space X is said to be ***equicontinuous*** *on E if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $d(x, y) < \delta$ for all $x, y \in E$ and $f \in \mathscr{F}$.* ♣

Remark The concept of equicontinuity is similar to uniform convergence. In the equicontinuous family, functions are uniformly continuous in the same extent, whereas in uniformly convergent sequence, functions converge in the same extent for every point.

Proposition 6.8

If K is a compact metric space, $f_n \in \mathscr{C}(K)$ for $n \in \mathbb{Z}_{>0}$, and $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K. ♠

Proof Let $\varepsilon > 0$ be given, and suppose the limit function is f. By uniform convergence, there exists N such that $|f_n(x) - f(x)| < \varepsilon/3$ for $n \geq N$ and $x \in K$. Since $\{f_n\}$ is continuous and K is compact, f is continuous by Corollary 6.1, and thus f is uniformly continuous by Proposition 4.9. Then there exists δ_0 such that $|f(x) - f(y)| < \varepsilon$ if $d(x, y) < \delta_0$. It follows that

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f(x)| + |f(x) - f(y)| + |f(y) - f_n(y)| = \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

for $d(x, y) < \delta_0$.

For $n < N$, f_n is uniformly continuous by Proposition 4.9 since f_n is continuous on a compact set. Then there exists $\delta_n > 0$ such that $|f_n(x) - f_n(y)| < \varepsilon$ if $d(x, y) < \delta_n$.

Hence putting $\delta = \min\{\delta_0, \delta_1, \cdots, \delta_{N-1}\}$ suffices. ■

Remark We use the uniform convergence degenerate the case into finite case. For $n \geq N$, we use triangle inequality to convert f_n to f, which is uniformly continuous, and then bound $|f_n(x) - f_n(y)|$ for all $n \geq N$. For $n < N$, we can directly use uniform continuity for each individual f_n. Then taking the minimum of δ's suffices.

Theorem 6.3 (Arzelà–Ascoli)

If K is compact, $\{f_n\} \in \mathscr{C}(K)$ for $n \in \mathbb{Z}_{>0}$, and $\{f_n\}$ is pointwise bounded and equicontinuous on K, then

(a) $\{f_n\}$ is uniformly bounded on K,

(b) $\{f_n\}$ contains a uniformly convergent subsequence. ♡

Proof (a) By equicontinuity, there exists $\delta > 0$ such that $|f_n(x) - f_n(y)| < 1$ for $d(x, y) < \delta$. By compactness, $K = \{N_\delta(x_i)\}_{1 \leq i \leq N}$. For every $1 \leq i \leq N$, there exists M_i such that $|f_n(x_i)| < M_i$ by boundedness. Put $M = \max\{M_1, \cdots, M_n\}$. For every $x \in K$, there exists x_k such that $d(x, x_k) < \delta$, then for every $n \in \mathbb{Z}_{>0}$,

$$|f_n(x)| \leq |f_n(x) - f_n(x_k)| + |f_n(x_k)| < 1 + M_k \leq 1 + M.$$

Hence $|f_n|$ is bounded uniformly by $1 + M$. ■

(b) For each n, K is covered by finite neighborhoods of the form $N_{1/n}(x_i)$, let K_n be the set of x_i's. Define $K' := \bigcup K_n$, then K' is countable, so there exists a subsequence $\{g_n\}_{n\in\mathbb{Z}_{>0}}$ of $\{f_n\}$ for which converges on K' by Proposition 6.7.

Let $\varepsilon > 0$ be given. By equicontinuity, there exists $\delta > 0$ such that $|g_n(x) - g_n(y)| < \varepsilon/3$ for $d(x, y) < \delta$ and all n. Choose $c > 0$ such that $1/c < \delta$, then for every $x \in K$, there exists $x_i \in K_c \subset K'$ such that $d(x, x_i) < \delta$ by the construction of K'. Also, by the convergence, for each $x_i \in K_n$, there exists N_i such that $|g_n(x_i) - g_m(x_i)|$ for all $n, m \geq N_i$. Put $N = \max\{N_i\}$. For $n, m \geq N$ and every $x \in E$, choose x_i as above, then

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(x_i)| + |g_n(x_i) - g_m(x_i)| + |g_m(x_i) - g_m(x)| = \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

Hence $\{g_n\}$ converges uniformly on K. ■

Remark (a) We use compactness to degenerate the problem into finite points $\{x_i\}$. For every point x_i, $\{f_n(x_i)\}$ bounded by pointwise boundedness, then the equicontinuity allows us to bound the function on the neighborhood of x_i.

(b) For each δ, we can choose a finite subset by compactness so that their neighborhoods covers K. The equicontinuity implies that bounding $|f_n - f_m|$ on the finite subset allows us to bound $|f_n - f_m|$ on their neighborhoods. Then it suffices to prove a subsequence converges on the finite subset; this can be done because there exists a countable dense subset of K and thus a subsequence converging on it.

6.4 The Stone-Weierstrass Theorem

Property *The following equalities hold by considering the binomial distribution:*

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}=1,$$
$$\sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k}=\mathbb{E}[X]=nx,$$
$$\sum_{k=0}^{n}(nx-k)^2\binom{n}{k}x^k(1-x)^{n-k}=\mathit{Var}[X]=nx(1-x).$$

Proposition 6.9 (Weierstrass Approximation Theorem)

If f is a continuous complex function on $[a,b]$, there exists a sequence of polynomials P_n such that $\lim_{n\to\infty} P_n(x)=f(x)$ uniformly on $[a,b]$. If f is real, then P_n may be take real.

Proof Without loss of generality, we may assume $[a,b]=[0,1]$. Let $B_n(f)(x)=\sum_{k=0}^{n} f(k/n)\cdot b_{k,n}(x)$ where $b_{k,n}(x)=\binom{n}{k}x^k(1-x)^{n-k}$ is the Bernstein polynomial, we therefore want to show $B_n(f)\to f$ uniformly.

Since f is continuous on a compact set, f is uniformly continuous and bounded by some M. Let $\varepsilon>0$ be given. There exists $\delta>0$ such that $|f(x)<f(y)|<\varepsilon/2$ whenever $|x-y|<\delta$. Choose $N=M/\delta^2\varepsilon$. For $n\geq N$, since $f(x)=f(x)\cdot\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}$,

$$\begin{aligned}|B_n(f)(x)-f(x)| &\leq \sum_{k=0}^{n}|f(k/n)-f(x)|\binom{n}{k}x^k(1-x)^{n-k}\\ &= \underbrace{\sum_{|x-k/n|<\delta}|f(k/n)-f(x)|\,b_{k,n}(x)}_{A}+\underbrace{\sum_{|x-k/n|\geq\delta}|f(k/n)-f(x)|\,b_{k,n}(x)}_{B}.\end{aligned} \tag{6.4.1}$$

For $|x-k/n|<\delta$, $|f(k/n)-f(x)|<\varepsilon/2$, then

$$A\leq|f(k/n)-f(x)|\sum_{k=0}^{n}b_{k,n}(x)=|f(k/n)-f(x)|=\varepsilon/2.$$

For $|x-k/n|\geq\delta$, (Chebyshev's inequality)

$$\begin{aligned}B &\leq 2M\sum_{|x-k/n|\geq\delta}b_{k,n}(x)\leq\sum_{k=0}^{n}\frac{(x-k/n)^2}{\delta^2}b_{k,n}(x)=\frac{2M}{\delta^2n^2}\sum_{k=0}^{n}(nx-k)^2b_{k,n}(x)\\ &=\frac{2M}{\delta^2n^2}\cdot nx(1-x)\leq\frac{2M}{\delta^2N}\cdot\frac{1}{4}=\frac{\varepsilon}{2}.\end{aligned}$$

Then $|B_n(f)(x)-f(x)|\leq A+B=\varepsilon$ for all $x\in[0,1]$, so $B_n(f)$ converges uniformly to f. ■

Remark The polynomial $B_n(f)(x)$ may be viewed as the weighted average of f on $[0,1]$ where the weight is given by the binomial distribution. For every x_0, when n approaches ∞, the binomial distribution is concentrated at x_0, so the term $b_{k,n}(x_0)$ vanishes when k/n is far from x, i.e, the polynomial $B_n(f)(x_0)$ converges to $f(x_0)$.

Corollary 6.2

For every interval $[-a, a]$, there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and such that $\lim_{n\to\infty} P_n(x) = |x|$ uniformly on $[-a, a]$. ♡

Definition 6.6 (Algebra of Functions, Uniform Closure)

*Let $\mathscr{A}$ be a family of of functions on a set E, then $\mathscr{A}$ is an **algebra** if $f + g, fg, cf \in \mathscr{A}$ for all $f, g \in \mathscr{A}$ and constant c.*

*If $\mathscr{A}$ has the property that $f \in \mathscr{A}$ whenever $f_n \to f$ uniformly for $f_n \in \mathscr{A}$, then $\mathscr{A}$ is said to be **uniformly closed**. The **uniform closure** of $\mathscr{A}$ is the set of all limit functions of uniformly convergent sequences in $\mathscr{A}$.* ♣

Example 6.8 The set of polynomials on $\mathbb{R}$ is an algebra. $\mathbb{C}([a, b])$ is the uniform closure of the set of all polynomials on $[a, b]$, by the Weierstrass approximation problem.

Proposition 6.10

Suppose $\mathscr{B}$ is the uniform closure of an algebra $\mathscr{A}$ of bounded functions. Then $\mathscr{B}$ is a uniformly closed algebra. ♠

Proof Sketch: Suppose $f_n \to f$ uniformly and $g_n \to g$ uniformly. It is not hard to see that f, g are (uniformly) bounded on E, and $f_n + g_n \to f + g$, $f_n g_n \to fg$, and $cf_n \to cf$. Hence $f + g, fg, cf \in \mathscr{B}$, i.e., $\mathscr{B}$ is an algebra. By Proposition 2.9, the uniform closure $\mathscr{B}$ is (uniformly) closed. ■

Definition 6.7 (Separate Points, Vanish at No Points)

*Let $\mathscr{A}$ be a family of functions on a metric space E. $\mathscr{A}$ said to **separate points** on E if to every pair of distinct $x_1, x_2 \in E$, there corresponds a function $f \in \mathscr{A}$ such that $f(x_1) = f(x_2)$.*

*If to each $x \in E$ there corresponds a function $g \in \mathscr{A}$ such that $g(x) \neq 0$, we say that $\mathscr{A}$ **vanishes at no point** of E.* ♣

Example 6.9 The algebra of all polynomials in one variables separates points and vanishes at no points. The algebra of all even polynomials on $[-1, 1]$ does not separate points on $[-1, 1]$ since $f(-1) = f(1)$ for all even polynomials f.

Proposition 6.11

Suppose $\mathscr{A}$ is an algebra of function on a set E, $\mathscr{A}$ separate points on E and vanishes at no point of E. Suppose x_1, x_2 are distinct points of E, and c_1, c_2 are constants (real if $\mathscr{A}$ is a real algebra). Then $\mathscr{A}$ contains function f such that $f(x_1) = c_1$ and $f(x_2) = c_2$. ♠

Proof Since $\mathscr{A}$ separate points and vanishes at no point of E, there exists $g, h, k \in \mathscr{A}$ such that $g(x_1) \neq g(x_2)$, $g(x_1) \neq 0$, and $g(x_2) \neq 0$. Set $u = gk - g(x_1)k$ and $v = gh - g(x_2)h$. It is not hard to show $u(x_1) = v(x_2) = 0$ and $u(x_2), v(x_1) \neq 0$. Then the function $f := c_1 v/v(x_1) + c_2 u/u(x_2)$ is the desired function. ■

Theorem 6.4 (Stone-Weierstrass Theorem)

Let $\mathscr{A}$ be an algebra of real continuous functions of a compact set K. If $\mathscr{A}$ separates points on K and if $\mathscr{A}$ vanishes at no point of K, then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all real continuous functions on K. ♡

Proof **STEP 1**: If $f \in \mathscr{B}$, then $|f| \in \mathscr{B}$.

Proof: Let $a = \sup_{x \in K} |f(x)|$, and let $\varepsilon > 0$ be given. By Corollary 6.2, there exists $c_1, \cdots, c_n \in \mathbb{R}$ such that $|\sum_{i=1}^n c_i y^i - |y|| < \varepsilon$ for all $y \in [-a, a]$. Since $\mathscr{B}$ is an algebra, $g(x) = \sum_{i=1}^n c_i f(x)^i \in \mathscr{B}$, and $|g(x) - f(x)| < \varepsilon$ for all $x \in K$. Hence $|f|$ is an uniform limit of sequence in $\mathscr{B}$, so $|f| \in \mathscr{B}$ since $\mathscr{B}$ is uniformly closed.

STEP 2: If $f, g \in \mathscr{B}$, then $\max(f, g) \in \mathscr{B}$ and $\min(f, g) \in \mathscr{B}$.

Proof: Notice that $\max(f, g) = ((f + g) + |f - g|)/2$, so $\max(f, g) \in \mathscr{B}$ follows immediately from the fact that $|f - g| \in \mathscr{B}$. The result holds for $\min(f, g)$, and the result may be extended to any finite set of functions.

STEP 3: Given a real function f, continuous on K, a point $x \in K$, and $\varepsilon > 0$, there exists a function $g_x \in \mathscr{B}$ such that $g_x(x) = f(x)$ and $g_x(t) > f(t - \varepsilon)$ for all $t \in K$.

Proof: For each $y \in K$, there exists $h_y \in \mathscr{B}$ such that $h_y(x) = f(x)$ and $h_y(y) = f(y)$ by Proposition 6.11. By the continuity of h_y there exists an open set J_y such that $h_y(t) > f(t) - \varepsilon$, and the compactness implies that $K \subset J_{y_1} \cup \cdots \cup J_{y_n}$ for some $y_1, \cdots, y_n$. Then setting $g_x := \max(h_{y_1}, \cdots, h_{y_n})$ suffices.

STEP 4: Given a real function f, continuous on K, and $\varepsilon > 0$, there exists a function $h \in \mathscr{B}$ such that $|h(x) - f(x)| < \varepsilon$ for $x \in K$.

Proof: Consider g_x for each $x \in K$. By the continuity of g_x, there exists open set V_x containing x such that $g_x(t) < f(t) + \varepsilon$. Since K is compact, $K \subset V_{x_1} \cup \cdots \cup V_{x_m}$ for some m. The setting $h := \min(g_{x_1}, \cdots, g_{x_m})$ suffices since $h(t) > f(t) - \varepsilon$ by Step 3 and the construction implies that $h(t) < f(t) + \varepsilon$. ■

Remark The Stone-Weierstrass Theorem does not hold for complex algebra.

Definition 6.8 (Self-Adjoint Algebra)

An algebra $\mathscr{A}$ of complex functions is said to be ***self-adjoint*** *if the complex conjugate $\bar{f} \in \mathscr{A}$ for all $f \in \mathscr{A}$.* ♣

Theorem 6.5

Suppose $\mathscr{A}$ is a self-adjoint algebra of complex continuous functions on a compact set K, and $\mathscr{A}$ separates points and vanishes at no point of K. Then the uniform closure $\mathscr{B}$ of $\mathscr{A}$ consists of all complex continuous functions on K. In other words, $\mathscr{A}$ is dense in $\mathscr{C}(K)$. ♡

Part II

Analysis II

Chapter 7 Measure Theory

Introduction

- *Lebesgue exterior measure*
- *Properties of exterior measure and measure*
- *Cantor set and Vitali Set*
- *Littlewood's three principles of real analysis*
- *Lebesgue measure*
- *σ-algebra and Borel set*
- *Measurable Functions*

7.1 Preliminaries

7.1.1 Riemann Integral Recitation

Suppose $f : I \to \mathbb{R}$ is abounded function defined on a rectangle $I \subset \mathbb{R}^n$. We define Riemann integral as follows:

- We partition I into a finite collection of almost disjoint rectangles $\Gamma = \{I_n\}_{n=1}^N$ (whose interiors are pairwise disjoint), select points $\xi_k \in I_k$, and define the **Riemann partial sum** be

$$R_\Gamma(\xi_1, \cdots, \xi_N) := \sum_{k=1}^{N} f(\xi_k) v(I_k)$$

  where $v(I)$ denotes the volume of the rectangle I.
- f is said **Riemann integrable** if the limit of R_Γ exists as the norm of partition Γ ($|\Gamma| := \max_k \operatorname{diam}(I_k)$) satisfies $|\Gamma| \to 0$.
- More precisely, the **Riemann integral** of f, denoted as $\int_I f(x)\, dx = A$, exists if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for any partition Γ of I such that $|\Gamma| < \delta$ with any choice of $\{\xi_k\}_{k=1}^N$, the inequality $|R_\Gamma(\xi_1, \cdots, \xi_N) - A| < \varepsilon$.

Alternatively, we may use upper and lower Riemann sum to define Riemann integral:

Definition 7.1 (Riemann Integral)

Define the upper and lower Riemann sum by

$$U_\Gamma := \sum_{k=1}^{N} v(I_k) \sup_{x \in I_k} f(x), \qquad L_\Gamma := \sum_{k=1}^{N} v(I_k) \inf_{x \in I_k} f(x),$$

then f is ***Riemann integrable*** *if and only if $\lim_{|\Gamma|\to 0} U_\Gamma = \lim_{|\Gamma|\to 0} L_\Gamma = A$, and its Riemann integral is denoted by $\int_I f(x)\, dx$.* ♣

Example 7.1 Riemann integral is very restrictive. For instance, Dirichlet function $\chi_{\mathbb{Q}\cap[0,1]}$, i.e., the characteristic function on rationals on $[0, 1]$, is not Riemann integrable, because $U_\Gamma = 1$ and $L_\Gamma = 0$ for any partition Γ.

Indeed, for a function to be Riemann integrable, it need to be *continuous "almost everywhere"*.

7.1.2 Rectangles and Cubes

A **rectangle** R in $\mathbb{R}^d$ is given by the product of d one-dimensional closed and bounded intervals $R = \prod_{1\leq j\leq d}[a_j, b_j]$ where $a_j \leq b_j$ are real numbers, and its **volume** is given by $|R| = \prod_{1\leq j\leq d}|b_j - a_j|$. A **cube** is rectangle for which the all its side lengths are equal.

Two rectangles A, B are said to **almost disjoint** (non-overlapping) if their interior are disjoint, then we use the convention $A \sqcup B$ to denote their (almost) disjoint union.

Proposition 7.1

If a rectangle I is the union of finitely many non-overlapping rectangles, i.e., $I = \bigsqcup_{n=1}^{N} I_n$, then $|I| = \sum_{n=1}^{N}|I_n|$. In particular, if rectangle $I, I_1, \cdots, I_n$ satisfy $I \subset \bigcup_{k=1}^{N} I_k$ then $|I| \leq \sum_{k=1}^{N}|I_k|$.

Proof Sketch: For each rectangle I_n, we may divide it into grids of rectangles $\{\tilde{I}_{n,j}\}_{j=1}^{M_n}$ of non-overlapping rectangles, for which $\{\tilde{I}_{n,j}\}_{n,j}$ is a grid of rectangles that forms I. Then $|I| = \sum_{n,j}|I_{n,j}| = \sum_{n=1}^{N}\sum_{j=1}^{M_n}|I_{n,j}| = \sum_{n=1}^{N}|I_n|$. The second statement follows from the first statement by breaking $I_1, \cdots, I_n$ into (not necessarily distinct) non-overlapping rectangles.

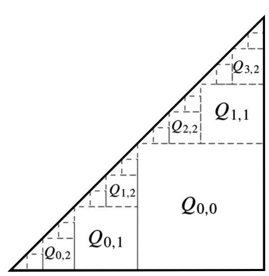
Note ***Lemma**: Every open set $G \subset \mathbb{R}$ can be written as a countable union of disjoint open intervals.*

Remark The above lemma does not hold in general Euclidean space $\mathbb{R}^n$.

Lemma 7.1

Every open set $G \subset \mathbb{R}^n$ can be written as a countable union of non-overlapping (closed) cubes.

Proof Consider the collection $\tilde{\mathcal{F}}_0$ of all cubes of side length 1 whose vertex points are integer lattice points in $\mathbb{Z}^n$, let $\mathcal{F}_0 \subset \tilde{\mathcal{F}}_0$ denotes the collection of cubes in $\tilde{\mathcal{F}}_0$ which are contained in G. Repeatedly, for step k, subdivide all cubes in $\tilde{\mathcal{F}}_k$ that are not contained in $\bigsqcup_{i=1}^{k}\mathcal{F}_i$ into 2^n cubes (i.e., divide side length by 2 for each dimension), denoted by $\tilde{\mathcal{F}}_{k+1}$. Define $\mathcal{F}_{k+1}$ to be the collection of cubes $Q \in \tilde{\mathcal{F}}_{k+1}$ contained in G. Note that for each k, F_k is a countable collection of non-overlapping cubes of side length 2^{-k}, so does their union $\mathcal{F} := \bigcup \mathcal{F}_k$. Denote by $H := \bigsqcup_{Q\in\mathcal{F}} Q$ the union of all chosen cubes.



It suffices to prove $G = H$. $H \supset G$ is obvious because $Q \subset G$ for all $Q \in \mathcal{F}$. Conversely, for every $x \in G$, there exists an open ball B of radius contained in G. By the Archimedean principle, there exists a cube Q in some $\tilde{\mathcal{F}}_k$ such that $x \in Q \subseteq B$. This contradicts to the construction of $\mathcal{F}_k$. Hence any open set G can be written as a countable union of non-overlapping cubes. ■

Remark The decomposition shown in the proof is the "dyadic decomposition of $\mathbb{R}^n$"

7.2 Lebesgue Exterior Measure

7.2.1 Lebesgue Exterior Measure

Definition 7.2 (Exterior Measure)

Let $E \subset \mathbb{R}^n$, we define the ***exterior (outer) measure*** *of E by*

$$m_*(E) := \inf \left\{ \sum_{n=1}^{\infty} |Q_j| \,\middle|\, Q_j \text{ are cubes}, E \subset \bigcup_{n=1}^{\infty} Q_j \right\}$$

where the infimum is taken over all countable covering of E by (closed) cubes and $|Q|$ denotes the volume of Q. ♣

Remark It is not suffice to allow finite sums (Jordan outer measure). In this case, we obtain $m_*(\mathbb{Z}) = \infty$ and $m_*([0,1]) = 1$, while $\mathbb{Z}$ is countable and $[0,1]$ is uncountable.

Example 7.2 Suppose $E = \{x\}$ is a singleton, the outer measure is $m_*(E) = 0$ since we can cover the point x with an arbitrarily small cube.

Suppose $E = \{x_n\}_{n=1}^{\infty}$ is countable. For all $\varepsilon > 0$, we may cover each x_n with a cube Q_n whose volume is no greater than $\varepsilon/2^{n+1}$, then $m_*(E) \le \sum_{n=1}^{\infty} |Q_n| \le \sum_{n=1}^{\infty} \varepsilon/2^{n+1} = \varepsilon$. It follows that $m_*(E) = 0$, i.e., the outer measure of countable set is zero.

Example 7.3 The outer measure of a cube Q equal to its volume $|Q|$.

Proof $m_*(Q) \le |Q|$ because Q covers itself. Suppose $\{Q_j\}$ is a countable covering of Q by cubes. For each Q_j, find Q_j^* such that $(Q_j^*)^\circ \supset Q_j$, $|Q^*| \le |Q_j| + 2^{-j}\varepsilon$. Since Q is compact and $\{(Q_j^*)^\circ\}$ forms a open covering of Q, there exists $N \in \mathbb{N}$ such that $Q \subset \bigcup_{j=1}^{N} Q_j^*$. By Proposition 7.1,

$$|Q| \le \sum_{j=1}^{N} |Q_j^*| \le \sum_{j=1}^{N} \left(|Q_j| + 2^{-j}\varepsilon\right) \le \sum_{j} |Q_j| + \varepsilon.$$

It follows that $|Q| \le \sum_j |Q_j|$ for arbitrary covering $\{Q_j\}$, so $m_*(Q) \ge |Q|$. Hence $m_*(Q) = |Q|$. ■

Remark The approach is to prove $|Q| \le \sum_{k=1}^{\infty} |Q_k|$ for any covering $\{Q_k\}$, which we proved by reduce it to finite cubes by compactness (through enlarging each Q_j to a slightly larger open cube) and apply Proposition 7.1.

Note *It is valid to replace the coverings by cubes with rectangles or closed balls.*

Proof Denote by $m_*^{\mathcal{R}}$ the outer measure by rectangles. Notice that every cube is a rectangle, $m_*^{\mathcal{R}}(E) \le m_*(E)$. Suppose $\{R_j\}_{j=1}^{\infty}$ is a covering of E by rectangles. For each j, there exists a countable covering $\{Q_{j,k}\}_{k=1}^{\infty}$ of R_j by cubes such that $\sum_{k=1}^{\infty} |Q_{j,k}| < |R_j| + 2^{-j}\varepsilon$. Since $\{Q_{j,k}\}_{j,k}$ form a covering,

$$m_*(E) \le \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} |Q_{j,k}| < \sum_{j=1}^{\infty} \left(|R_j| + 2^{-k}\varepsilon\right) = \sum_{j=1}^{\infty} |R_j| + \varepsilon.$$

It follows that $m_*(E) \le \sum_{j=1}^{\infty} |R_j|$, then $m_*(E) \le m_*^{\mathcal{R}}(E)$. Hence $m_*(E) = m_*^{\mathcal{R}}(E)$. ■

Remark One direction is trivial, and for the other direction, we use the fact that every rectangle is almost a countable union of cubes.

Example 7.4 For a rectangle I, $m_*(I) = |I|$.

Proof The proof is analogous to the previous example since cubes can be replaced by rectangles. ■

Example 7.5 $m_*(\mathbb{R}^n) = +\infty$.

Proof Since cubes $Q \subset \mathbb{R}^n$ can have arbitrarily large volume, we must have $m_*(\mathbb{R}^n) = +\infty$ by monotonicity (see Proposition 7.2 (a)). ■

7.2.2 Properties of The Exterior Measure

Proposition 7.2

(a) ***Monotonicity****: If $E_1 \subset E_2$, then $m_*(E_1) \le m_*(E_2)$.*

(b) ***Countable sub-additivity****: If $E = \bigcup_{j=1}^\infty E_j$, then $m_*(E) \le \sum_{j=1}^\infty m_*(E_j)$.*

Proof (a) This proposition follows directly from the fact that every covering of E_2 covers E_1.

(b) Let $\varepsilon > 0$ be given. For each j, there exists a covering $\{Q_{j,k}\}_k$ such that $\sum_{k=1}^\infty |Q_{j,k}| < m_*(E_j) + 2^{-j}\varepsilon$. Notice that $\{Q_{j,k}\}_{j,k}$ forms a covering, then

$$m_*(E) \le \sum_{j=1}^\infty \sum_{k=1}^\infty |Q_{j,k}| \le \sum_{j=1}^\infty \left(m_*(E_j) + 2^{-j}\varepsilon\right) = \sum_{j=1}^\infty m_*(E_j) + \varepsilon.$$

Hence $m_*(E) \le \sum_{j=1}^\infty m_*(E_j)$. ■

Corollary 7.1

(a) Completeness: If $m_*(F) = 0$ and $E \subset F$, then $m_*(E) = 0$.

(b) If $m_*(E_k) = 0$ for all k, then $m_*(\bigcup E_k) = 0$.

Proposition 7.3 (Approximation by Open Sets)

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

Proof $m_*(E) \le \inf m_*(\mathcal{O})$ follows immediately from the monotonicity. Conversely, let $\varepsilon > 0$ be given, there exists a covering $\{Q_j\}$ such that $\sum |Q_j| < m_*(E) + \varepsilon/2$. For each Q_j, there exists Q_j^* such that $(Q_j^*)^\circ \supset Q_j$ and $|Q_j^*| < |Q_j| + 2^{-(j+1)}\varepsilon$. Let $\mathcal{O} = \bigcup (Q_j^*)^\circ$, then $\mathcal{O}$ is open and $E \subset \mathcal{O} \subset \bigcup Q_j^*$. Note that

$$m_*(\mathcal{O}) \le \sum_{j=1}^\infty |Q_j^*| \le \sum_{j=1}^\infty \left(|Q_j| + 2^{-(j+1)}\varepsilon\right) \le \sum_{j=1}^\infty |Q_j| + \varepsilon/2 < m_*(E) + \varepsilon,$$

it follows that $\inf m_*(\mathcal{O}) \leq m_*(E)$. Hence $m_*(E) = \inf m_*(\mathcal{O})$. ■

Remark It suffices to prove $\sum |Q_j| \geq \inf m_*(\mathcal{O}) - \varepsilon$ for all covering $\{Q_j\}$ and $\varepsilon > 0$ by the properties of infimum. Note that we have a slightly larger open cube Q_j^* contains Q_j for each j, adjoining Q_j^*'s gives an open set that is slightly larger than $\bigcup_j Q_j \supseteq E$.

Proposition 7.4

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

Proof By subadditivity, $m_*(E_1 \cup E_2) \leq m_*(E_1) + m_*(E_2)$. Conversely, for any covering $\{Q_j\}$ of $E_1 \cup E_2$, we can subdivide each Q_j into finitely many nonoverlapping cubes whose diameter is less than $d(E_1, E_2)$, forming a covering $\{I_k\}$ of $E_1 \cup E_2$. Let S_1 and S_2 denotes all I_k's which intersects E_1 and E_2, respectively. Notice that no cubes I_k intersects both E_1 and E_2, namely $S_1 \cap S_2 = \varnothing$, and S_1, S_2 form a covering of E_1, E_2, resp. Then

$$\sum |Q_j| = \sum |I_k| \geq \sum_{I_k \in S_1} |I_k| + \sum_{I_k \in S_2} |I_k| \geq m_*(E_1) + m_*(E_2),$$

followed by $m_*(E_1 \cup E_2) \geq m_*(E_1) + m_*(E_2)$. Hence $m_*(E) = m_*(E_1) + m_*(E_2)$. ■

Remark The approach is to divide each cube in the covering to smaller cubes so that no cubes intersect both E_1 and E_2, then it is not hard to show the equality.

Proposition 7.5

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

Proof $m_*(E) \leq \sum_{j=1}^{\infty} |Q_j|$ by monotonicity. Conversely, let $\varepsilon > 0$ be given. For each Q_j, choose Q_j^* be a cube contained in Q_j such that $|Q_j^*| > |Q_j| - 2^{-j}\varepsilon$. Note that $d(Q_j^*, Q_k^*) > 0$ for $j \neq k$; by Proposition 7.4, for every N,

$$m_*(E) \geq \sum_{j=1}^{N} m_*(Q_j^*) = \sum_{j=1}^{N} |Q_j^*| \geq \sum_{j=1}^{M} (|Q_j| - 2^{-j}\varepsilon) = \sum_{j=1}^{N} |Q_j| - \varepsilon.$$

Let $n \to \infty$, $m_*(E) \geq \sum_{j=1}^{\infty} |Q_j| - \varepsilon$, followed by $m_*(E) \geq \sum_{j=1}^{\infty} |Q_j|$. Hence $m_*(E) = \sum_{j=1}^{\infty} |Q_j| - \varepsilon$. ■

Remark The approach is to take slightly smaller cube Q_j^* for each Q_j (so they have positive distance), then applying Proposition 7.4 gives the desired result for finite case, thus letting $N \to +\infty$ yields the desired equality.

7.3 Lebesgue Measure

Definition 7.3 (Lebesgue Measure)

*A set $E \subset \mathbb{R}^n$ is said to be **(Lebesgue) measureable** if for all $\varepsilon > 0$, there exists an open set G such that $E \subset G$ and $m_*(G \setminus E) < \varepsilon$. If E is measurable, we define its **(Lebesgue) measure** to be $m(E) := m_*(E)$.* ♣

That is, E is measurable if it can be approximated by open set from above.

Property *Every open set and every rectangle are measurable.*

Proof It is trivial that open set is open by definition. For rectangle R, there exists rectangle R^* such that $m_*(R^*) < m_*(R) + \varepsilon$, then $(R^*)^\circ$ is the desired open set. ■

Property *Every set with zero outer measure (aka, **null sets**) is measurable.*

Proof Suppose $m_*(E) = 0$, Proposition 7.3 implies that for all $\varepsilon > 0$, there exists open $G \supset E$ such that $m_*(G) < m_*(E) + \varepsilon = \varepsilon$. Then $m_*(G \setminus E) \le m_*(G) < \varepsilon$, so E is measurable. ■

Property *A countable union of measurable sets is also measurable.*

Proof Let $E_1, \cdots$ be measurable sets and $E = \bigcup E_k$. For all $\varepsilon > 0$, for each E_k, there is an open set G_k such that $m_*(G_k \setminus E_k) < 2^{-k}\varepsilon$. Let $G = \bigcup G_k$, then G is open, and $m_*(G \setminus E) \le \sum m_*(G_k \setminus E_k) \le \varepsilon$. Hence E is measurable. ■

Property *Every closed set is measurable.*

Proof Suppose F is closed. Proposition 7.3 implies that for all $\varepsilon > 0$, there exists G open such that $m_*(G) < m_*(F) + \varepsilon$. Suppose F is bounded and thus compact. Note that $G \setminus F$ can be written as the countable union of non-overlapping cubes, $\bigcup_{j=1}^\infty Q_j$. For each N, note that $\bigcup_{j=1}^N Q_j$ is compact and F is closed implies their distance is positive, Proposition 7.4 implies that $m_*(G) = m_*(F) + m_*(\bigsqcup_{j=1}^N Q_j) = m_*(F) + \sum_{j=1}^N |Q_j|$, then $\sum_{j=1}^N |Q_j| \le m_*(G) - m_*(F) \le \varepsilon$. Let $N \to +\infty$, the subadditivity yields $m_*(G \setminus F) \le \sum_{j=1}^\infty |Q_j| \le \varepsilon$.
If F is not bounded, let Q_k be cubes of side length k, note that $\mathbb{R}^n = \bigcup Q_k$. Since $F \sqcup Q_k$ is closed and bounded, thus measurable, $F = \bigcup (F \cap Q_k)$ is also measurable. ■

Proof Consider the bounded case, for any open set $G \supseteq F$, we see that $m(G) = m(F) + m(G \setminus F)$ by the property of compactness, thus $m_*(G \setminus F)$ can be arbitrarily small.

Property *The complement of any measurable set is measurable.*

Proof For every $k \in \mathbb{N}$, there exists $G_k \supset E$ open such that $m_*(G_k \setminus E) \le 1/n$, and the complement G_k^c is closed and thus measurable. Then $S = \bigcup_{n=1}^\infty G_n^c$ is measurable. Note that $S \subset E^c$ and $E^c \setminus S \subset E^c \setminus G_k^c = G_k \setminus E$, then $m_*(E^c \setminus S) \le 1/n$ for all n, followed by $m_*(E^c \setminus S) = 0$ and $E^c \setminus S$ is measurable. Hence $E^c = S \cup (E^c \setminus S)$ is measurable. ■

Property *A countable intersection of measurable sets is also measurable.*

Proof Follows immediately from property 3 and 5. ■

Theorem 7.1 (Countable Additivity)

If $E_1, E_2, \cdots$ are disjoint measurable sets, then $m(\bigsqcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} m(E_k)$. ♡

Proof Let $E := \bigsqcup_{k=1}^{\infty} E_k$. One direction $m(E) \leq \sum_k m(E_k)$ follows immediately from countable subadditivity. Conversely, suppose all E_k are bounded, let $\varepsilon > 0$ be given. For each k, there exists a closed set $F_k \subset E_k$ such that $m_*(E \setminus F) \leq 2^{-k}\varepsilon$ (since E^c is measurable). Note that F_k's have positive distance since they are disjoint, then for each $N \in \mathbb{N}$, $m(\bigsqcup_{k=1}^{N} E_k) \geq m(\bigsqcup_{k=1}^{N} F_k) = \sum_{k=1}^{N} m(F_k) \geq \sum_{k=1}^{N} m(E_k) - \varepsilon$. Let $N \to +\infty$, we have $m(E) \geq \sum_k m(E_k) - \varepsilon$, followed by $m(E) \geq \sum_k m(E_k)$. Hence $m(E) = \sum_k m(E_k)$.

For E_k that is not bounded, choose cubes $Q_j \nearrow \mathbb{R}^n$. Then $\{E_{k,j} := E_k \cap (Q_j \setminus Q_{j-1})\}_j$ is a pairwise disjoint collection of bounded sets, for which $E_k = \bigsqcup_j E_{k,j}$. Apply the previous result, we have $m(E_k) = \sum_j m(E_{k,j})$, hence $m(E) = \sum_{k,j} m(E_{k,j}) = \sum_k m(E_k)$. ■

Remark Consider the bounded case, note that E can be approximated by closed set from inside, this gives sets with positive distance. We obtain the desired equality for finite case by applying Proposition 7.4, and this can be easily extended to the countable case.

Corollary 7.2

Let $\{I_k\}$ be a countable collection of non-overlapping rectangles, then $m(\bigsqcup_k I_k) = \sum_k m(I_k)$. ♡

Definition 7.4 (Monotonicity)

If $E_1, E_2, \cdots$ is a countable collection of subsets that ***increases*** *to E, i.e., $E_k \subset E_{k+1}$ for all k and $E = \bigcup_k E_k$, we write $E_k \nearrow E$. Similarly, if $E_1, E_2, \cdots$ is a countable collection of subsets that* ***decreases*** *to E, i.e., $E_k \supset E_{k+1}$ for all k and $E = \bigcap_k E_k$, we write $E_k \searrow E$.* ♣

Theorem 7.2 (Continuity from above/below)

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

(a) If $E_k \nearrow E$, then $m(E) = \lim_{k\to\infty} m(E_k)$.

(b) If $E_k \searrow E$ and $m(E_k) < \infty$ for some k, then $m(E) = \lim_{k\to\infty} m(E_k)$. ♡

Proof (a) We may assume E_k's have finite measure, otherwise the equality holds obviously. Denote by $G_1 = E_1$ and $G_k = E_k \setminus E_{k-1}$ for $k > 1$, then $E = \bigsqcup_k G_k$. It follows that

$$m(E) = \sum_{k=1}^{\infty} m(G_k) = \lim_{k\to\infty} m\left(\bigsqcup_{n=1}^{k} G_n\right) = \lim_{k\to\infty} m(E_k).$$

(b) Without loss of generality, we may assume $m(E_1) < \infty$, then $m(E_k) < \infty$ for all E_k. Define G_k's as above, then $E_1 = E \sqcup (\bigsqcup_k G_k)$. By the previous result,

$$m(E_1) = m(E) + \lim_{n\to\infty} \sum_{k=1}^{n} m\left(E_k \setminus E_{k-1}\right) = m(E) + m(E_1) - \lim_{k\to\infty} m(E_k),$$

hence $m(E) = \lim_{k\to\infty} m(E_k)$. ■

7.4 σ-Algebra and Borel Sets

Definition 7.5 (σ-Algebra)

*A collection Σ of subsets of some universal set U is called a **σ-algebra** if it satisfies:*

(1) $U \in \Sigma$

(2) Closed under complement: If $E \in \Sigma$, then $E^c \in \Sigma$

(3) Closed under countable union: If $E_k \in \Sigma$ for all $k \in \mathbb{N}$, then $\bigcup_k E_k \in \Sigma$

Example 7.6 The collection of all (Lebesgue) measurable sets in $\mathbb{R}^n$ is a σ-algebra.

Definition 7.6 (Borel σ-Algebra)

*The smallest σ-algebra containing all open sets in $\mathbb{R}^n$ is called the **Borel σ-algebra**, denoted $\mathscr{B}$, and the sets in $\mathscr{B}$ are **Borel sets**.*

Example 7.7 All open sets, closed set, F_σ-sets (countable union of closed sets), and G_δ-sets (countable intersection of open sets) are in Borel σ-algebra.

Remark $\mathscr{B}$ is a proper subset of $\mathscr{M}$ (the collection of Lebesgue measurable sets).

Proposition 7.6

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

(a) if and only if E differs from a G_δ by a null set (set of measure zero),

(b) if and only if E differs from a F_σ by a null set.

Proof Suppose H is a G_δ set and $Z = H \setminus E$ is a null set for some Z, H. Since H and Z are measurable, then $E = H \setminus Z$ is measurable. Conversely, for all k, there exists $G_k \supset E$ such that $m(G_k \setminus E) < 1/k$. Let $H = \bigcap_k G_k$ be a G_δ set, then $m(H \setminus E) = 0$. The second statement is analogous. ■

7.5 Vitali Sets

Lemma 7.2 (Invariance of Lebesgue measure)

Translation invariance: Suppose $E \in \mathscr{M}$ and $h \in \mathbb{R}^n$, then $E + h = \{x + h \mid x \in E\}$ is measurable and the measure is $m(E + h) = m(E)$.

Dilation invariance: Suppose E is measurable, $\delta E = \{(\delta_1 x_1, \cdots, \delta_n x_n) \mid x \in E\}$, then δE is measurable and $m(\delta E) = \delta_1 \cdots \delta_n m(E)$. ♡

Define an equivalence relation on $[0, 1]$ as follows: $x \sim y$ if and only if $x - y \in \mathbb{Q}$. The equivalence classes $[x] = \{x + q \in [0, 1] \mid q \in \mathbb{Q}\}$ are either disjoint or coincide. They form a partition of $[0, 1] = \bigsqcup_{\alpha \in V} [x_\alpha]$ (under the axiom of choice), where V consists of one representative from each class.

Theorem 7.3

The Vitali set V is not measurable. ♡

Proof Let $V_q := V + q = \{x + q \,|\, x \in V\}$, and denote by $Q = [-1, 1] \cap \mathbb{Q}$. We have the following three observations:

(a) $[0, 1] \subset \bigcup_{q \in K} V_q$: suppose $x \in [0, 1]$, then $x \sim y$ for some $y \in V$ by the definition of V. That is, $y - x \in \mathbb{Q}$, in particular $x - y \in [-1, 1] \cap \mathbb{Q} = Q$, so there exists $q \in Q$ such that $x = y + q$, i.e., $x \in V_q$.

(b) V_q's are disjoint: If $x \in V_p \cap V_q$, then $x = y + p = y' + q$ for some $y, y' \in V$, $p, q \in Q$. Then $y' - y \in [-1, 1] \cap \mathbb{Q} = Q$, so $y = y'$ by the definition of V, followed by $p = q$, hence $V_p = V_q$.

(c) $\bigsqcup_{q \in Q} V_q \subset [-1, 2]$: the statement is trivial since $V \subset [0, 1]$ and $q \in [-1, 1]$.

Combining the above observations, we have the following claim:

$$[0, 1] \subset \bigsqcup_{q \in Q} V_q \subset [-1, 2]. \tag{7.5.1}$$

Assume FSOC that V is measurable, then V_q is measurable and $m(V_q) = m(V)$ for all $q \in Q$. By the Equation (7.5.1) and monotonicity, given that $m(\bigsqcup_{q \in Q} V_q) = \sum_{q \in Q} m(V)$ by additivity,

$$1 = m([0, 1]) \leq m\left(\bigsqcup_{q \in Q} V_q\right) = \sum_{q \in Q} m(V) \leq m([-1, 2]) = 3$$

which is impossible since $\sum_{q \in Q} m(V) = |Q| \cdot m(V)$ and Q is countable. Hence V is not countable. ■

7.6 Measurable Functions

7.6.1 Measurable Functions

Definition 7.7 (Measurable Functions)

*Consider the real-valued function f defined on a measurable set $E \subset \mathbb{R}^n$, $f : E \to \mathbb{R} \cup \{\pm\infty\}$ (NB: we say f is finite-valued if $-\infty < f(x) < \infty$ for all $x \in E$). The function f is **measurable**, if for all $\alpha \in R$, $\{f < \alpha\} := \{x \in E \,|\, f(x) < \alpha\}$ is measurable.* ♣

Proposition 7.7

The equivalent definition/characterization of measurable functions includes: f is measurable if and only if $\{f \leq a\}$ (or $\{f > a\}$, or $\{f \geq a\}$) is measurable for all $a \in \mathbb{R}$.

In particular, if f is finite-valued, then f is measurable if and only if $\{a < f < b\}$ is measurable for all $a, b \in \mathbb{R}$. ♠

Proof The equivalence between $\{f < a\}$ and $\{f \leq a\}$ follows from $\{f \leq a\} = \bigcap_{n \in \mathbb{N}} \{f < a + 1/n\}$ and $\{f < a\} = \bigcup_{n \in \mathbb{N}} \{f \leq a - 1/n\}$. The equivalence between $\{f < a\}$ and $\{f \geq a\}$ and between $\{f \leq a\}$ and $\{f > a\}$ follows directly from taking complement.

For finite-valued f, note that $\{f < a\} = \bigcup_{b \in \mathbb{Z}} \{b < f < a\}$ and $\{b < f < a\} = \{f < a\} - \bigcup_{q \in \mathbb{Q}, q \leq b} \{f < b\}$, therefore f is measurable if and only if $\{a < f < b\}$ is measurable. ■

Proposition 7.8

A finite valued function f is measurable if and only if $f^{-1}(G)$ is measurable for every open set $G \subset \mathbb{R}$. ♠

Proof ($\Rightarrow$) Suppose G is open in $\mathbb{R}$, G can be written as a union of open intervals $G = \bigsqcup_k (a_k, b_k)$. Since the preimage of intervals $f^{-1}((a_k, b_k)) = \{a_k < f < b_k\}$ are measurable, $f^{-1}(G) = \bigsqcup_k f^{-1}((a_k, b_k))$ is measurable.

($\Leftarrow$) Conversely, the statement follows immediately from the fact that every $\{f < a\}$ is a preimage of an open interval. ■

Proposition 7.9

(a) If f is continuous on $\mathbb{R}^n$, then f is measurable.

(b) If f is measurable and finite-valued, and φ is continuous on $\mathbb{R}$, then $\varphi \circ f$ is measurable. ♠

Proof (a) follows immediately from the fact that $f^{-1}(G)$ is open and thus measurable for every open set G by continuity.

(b) Note that $\{\varphi \circ f < a\} = \{x \in E \,|\, f(x) \in \varphi^{-1}((-\infty, a))\}$, and $G := \varphi^{-1}((-\infty, a))$ is open by the continuity of φ, then $\{\varphi \circ f < a\} = f^{-1}(G)$ is measurable by Proposition 7.8. ■

Proposition 7.10

Suppose $\{f_k\}$ is a sequence of measurable functions f defined on E, then $\sup_k f_k$, $\inf_k f_k$, $\limsup_k f_k$, and $\liminf_k f_k$ are all measurable functions.

Proof Note that $\{\sup_n f_k > a\} = \bigcup_n \{f_n > a\}$, so $\sup f_k$ is measurable, and similarly $\inf f_k$ is measurable. The result holds for $\limsup$ since $\limsup_{k\to\infty} f_k(x) = \inf_n\{\sup_{k\geq n} f_n\}$, and it holds for $\liminf$ wlog. ∎

Corollary 7.3

Suppose $\{f_k\}$ is a sequence of measurable functions, and $f(x) = \lim_{k\to\infty} f_k(x)$, then $f(x)$ is measurable.

Proposition 7.11

Suppose f, g are finite-valued measurable functions, then $f+g$ and fg are measurable.

Proof Note that $\{f+g>a\} = \{f > a-g\} = \bigcup_{q\in\mathbb{Q}}\{f > q > a-g\}$, and $\{f > q > a-g\} = \{f > q\} \cap \{g > a-q\}$ is measurable, so $\{f+g>a\}$ is measurable, followed by $f+g$ is measurable.

Notice that $fg = \frac{1}{4}[(f+g)^2 + (f-g)^2]$, and $f+g, f-g$ are measurable thus $(f+g)^2, (f-g)^2$ are measurable, then fg is measurable. ∎

Definition 7.8 (Almost Everywhere)

*A property is said to hold **almost everywhere** in E (abbreviated as a.e.) if it holds in E except for a subset of E with zero measure.*

Proposition 7.12

If f is measurable, $f = g$ a.e., then g is measurable and $m(\{g<a\}) = m(\{f<a\})$ for all $a \in \mathbb{R}$.

7.6.2 Approximate Measurable Functions by Simple Functions

Definition 7.9 (Characteristic Function, Simple Function)

The characteristic function (indicator function) of a set A is defined as $\chi_A(x) = 1$ if $x \in A$ and otherwise $\chi_A(x) = 0$.

*A **simple function** is a function of the form $f(x) = \sum_{k=1}^N a_k \chi_{E_k}$ where $a_k \in \mathbb{R}$ and E_k is measurable of finite measure for all k.*

Remark Without loss of generality, we may assume a_k's are distinct and E_k's are disjoint in a simple function.

Proposition 7.13

(a) *Suppose f is a non-negative measurable function, then there exits an increasing sequence of non-negative simple functions f_k converge to f pointwise.*

(b) *Suppose f is a measurable function, then there exists an increasing sequence of simple functions f_k such that $|f_k(x)| \leq |f_{k+1}(x)|$ and f_k converge to f pointwise.*

Proof (a) Suppose f is a non-negative measurable function; for each $k \in \mathbb{N}$, define

$$f_k(x) = \begin{cases} 0 & \text{if } |x| > k \\ k & \text{if } f(x) \geq k \\ \dfrac{j-1}{2^k} & \text{if } f(x) \in \left[\dfrac{j-1}{2^k}, \dfrac{j}{2^k}\right) \text{ for } j \in \{1, \cdots, k \cdot 2^k\}. \end{cases}$$

- Each $f_k(x)$ is simple because $f_k = k\chi_{\{f>k\}} + \sum_{j=1}^{k \cdot 2^k} \left(\frac{j-1}{2^k}\right) \chi_{E_{j,k}}$, where each $E_{j,k} = [(j-1)/2^k, j/2^k)$ is measurable with finite measure.
- f_k is clearly nonnegative.
- To prove $f_k(x) \leq f_{k+1}(x)$ is increase, consider the following three cases:
  (i) If $|x| > k$, $f_{k+1}(x) \geq 0 = f_k(x)$;
  (ii) If $f(x) \geq k$, assume $f(x) \in \left[\frac{j-1}{2^{k+1}}, \frac{j}{2^{k+1}}\right)$, then $f_{k+1}(x) \geq \min(k+1, \frac{j-1}{2^{k+1}}) \geq k = f_k(x)$;
  (iii) If $f(x) < k$, assume $f(x) \in \left[\frac{j-1}{2^k}, \frac{j}{2^k}\right)$, then $f_{k+1}(x) \geq \frac{2j-2}{2^{k+1}} = \frac{j-1}{2^k} = f_k(x)$.
- Lastly, we want to prove $f_k \to f$ pointwise. If $f(x) = +\infty$, then for $k \geq x$, $f_k(x) = k$, so $f_k(x) \nearrow +\infty$. Suppose $f(x) < +\infty$, let $\varepsilon > 0$ be given. There exists N such that $1/2^N < \varepsilon$, then for $k \geq \max(f(x), N)$, $f(x) - f_k(x) \leq 1/2^k < \varepsilon$, so $\lim_{k\to\infty} f_k(x) = f(x)$.

Hence there exists an increasing sequence of nonnegative simple functions that converge to f.

(b) Suppose f is a measurable function, let $f^+ = \max(f, 0)$ and $f^- = -\min(f, 0)$. By part (a), there exists sequences of simple functions $g_k \to f^+$ and $h_k \to f^-$. Let $f_k = g_k - h_k$. $|f_k(x)| \leq |f_{k+1}(x)|$ because $|f_k| = g_k(x) + h_k(x) \leq g_{k+1}(x) + h_{k+1}(x) = |f_{k+1}(x)|$ since either $g_k(x)$ or $h_k(x)$ will be zero for every x. In addition, $\lim_{k\to\infty} f_k(x) = \lim_{k\to\infty}[g(x) - h(x)] = f^+(x) - f^-(x) = f(x)$. Hence f_k converges to f. ∎

7.7 Littlewood's Three Principles of Real Analysis

Theorem 7.4 (Littlewood's Three Principles of Real Analysis)

(a) Every (measurable) set is nearly a finite union of cubes. [Proposition 7.14]

(b) Every (measurable) function is nearly continuous. [Lusin's Theorem, 7.15]

(c) Every convergent sequence (convergent almost everywhere) of function is nearly uniformly convergent. [Egorov's Theorem, 7.16]

♡

Note that "nearly" means the condition holds for $E \setminus N$ where N is a set with small measure.

Proposition 7.14

*If $m(E)$ is finite, then there exists a finite union $F = \bigcup_{j=1}^{N} Q_j$ of closed cubes such that $m(E \triangle F) \leq \varepsilon$ (where $E \triangle F := (E \setminus F) \cup (F \setminus E)$ is the **symmetric difference**).*

♠

Proof Suppose E is measurable with finite measure. Choose $\{Q_j\}$ such that $\sum_{j=1}^{\infty} m(Q_j) \leq m(E) + \varepsilon/2$. Since $\bigcup_{j=1}^{n} Q_j \nearrow E$, there exists N s.t. $\sum_{j=N+1}^{\infty} m(Q_j) < \varepsilon/2$. Define $F = \bigcup_{j=1}^{N} Q_j$, then

$$m(E \triangle F) = m(E \setminus F) + m(F \setminus E) \leq m\left(\bigcup_{j=N+1}^{\infty} Q_j\right) + \left(\sum_{j=1}^{\infty} m(Q_j) - m(E)\right) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Proposition 7.15 (Egorov's Theorem)

Suppose $\{f_k\}$ is a sequence of measurable functions that converge (a.e.) to a finite-valued function f on a measurable set E of finite measure. Then for all $\eta > 0$, there exists closed $F \supset E$ such that $m(E \setminus F) < \eta$ and $f_k \to f$ uniformly on F.

♠

Lemma: Under the same hypothesis, for all $\varepsilon > 0$ and $\eta > 0$, there exists a closed set $F \supset E$ and $N \in \mathbb{N}$ such that $m(E \setminus F) < \eta$ and $|f(x) - f_k(x)| < \varepsilon$ for all $x \in F$ and $k \geq N$.

Proof: Define $E_n = \bigcap_{k=n}^{\infty} \{f(x) - \varepsilon < f_k < f(x) + \varepsilon\}$, then E_n is measurable. Note that $E_n \nearrow E$, so $m(E) = \lim_{n\to\infty} m(E_n)$, followed by there exists N such that $m(E \setminus E_N) < \eta/2$. We may choose a closed set $F \subset E_N$ such that $m(E_N \setminus F) < \eta/2$. Therefore, $m(E \setminus F) < \eta$ and $|f(x) - f_k(x)| < \varepsilon$ for all $x \in F$ and $k \geq N$.

Proof For all $n \in \mathbb{N}^+$, there exists a closed set F_n such that $m(E \setminus F_n) < \eta/2^n$ such that $|f(x) - f_k(x)| < 1/n$ on F_n for $k \geq N_n$. Put $F = \bigcap_{n=1}^{\infty} F_n$, then F is closed and $m(E \setminus F) < \eta$. For all $\varepsilon > 0$, there exists N s.t. $1/N < \varepsilon$, so $|f(x) - f_k(x)| < 1/N < \varepsilon$ on $F \subset F_N$ for $n \geq N_N$. Therefore, $f_k \to f$ uniformly on F. ■

Proposition 7.16 (Lusin's Theorem)

Let f be a finite-valued measurable function defined on a measurable set E, then for all $\varepsilon > 0$, there is a closed set $F \subset E$ such that $m(E \setminus F) < \varepsilon$, and $f|_F$ is continuous.

Lemma: A simple measurable function f defined on E satisfies: for all $\varepsilon > 0$, there is a closed set $F \subset E$ such that $m(E \setminus F) < \varepsilon$ and $f|_F$ is continuous.

Proof: Suppose $f = \sum_{k=1}^{N} a_k \chi_{E_k}$. We may choose a closed set $F_k \subset E_k$ s.t. $m(E_k \setminus F_k) < 2^{-k}\varepsilon$ for each k, and let $F = \bigsqcup_{k=1}^{N} F_k$. Then $m(E \setminus F) = \sum_k m(E_k \setminus F_k) < \varepsilon$. It suffices to prove $f|_F$ is continuous: note that the $F_k$for all $\{x_i\} \subset F$ such that $x_i \to x \in F_K$, there exists N such that $x_i \in F_K$ for all $i \geq N$ since $\{F_k\}$ have positive distance pairwise, so $f(x_i) \to a_K = f(x)$.

Proof There exists a sequence of simple functions f_k converges to f pointwise. Suppose $m(E) < +\infty$. By the above lemma, for each k, there is a closed set F_k such that $f_k|_{F_k}$ is continuous and $m(E \setminus F_k) < 2^{-(k+1)\varepsilon}$. Then $m(E \setminus \bigcap F_k) \leq \sum m(E \setminus F_k) \leq \varepsilon/2$, and $f_k|_{\bigcap F_k}$ is continuous for all k. By Egorov's Theorem, there is a closed set F' such that $f_k \to f$ uniformly on F' and $m(E \setminus F') < \varepsilon/2$. Let $F := F' \cap \bigcap_{k=1}^{\infty} F_k$. Then F is closed, $m(E \setminus F) \leq m(E \setminus F') + m(E \setminus \bigcap F_k) = \varepsilon$. In addition, since $f_k|_F$ is continuous for all k and $f_k \to f$ converges uniformly, $f|_F$ is continuous.

On the other hand, suppose $m(E) = +\infty$. Let $E_k = E \cap \{x \mid k \leq |x| < k+1\}$, and choose a closed set $F_k \subset E_k$ s.t.$f_k|_{F_k}$ is continuous and $m(E_k \setminus F_k) < 2^{-k}\varepsilon$ for all k. Let $F = \bigcup_{k=1}^{\infty} F_k$, then $m(E \setminus F) < \varepsilon$. Note that F is closed (by proving every point in F^c is open since F_k's are closed sets with positive distance pairwise), and $f|_F$ is continuous (since $f_k|_{F_k}$ is continuous and F_k's have positive distance pairwise). Hence the statement holds even if $m(E) = +\infty$. ■

Chapter 8 Lebesgue Integration Theory

Introduction

❑

❑

8.1 The Lebesgue Integral

We are going to define integration progressively on (i) simple function, (ii) bounded functions supported on a set of finite measure, (iii) non-negative functions, and then (iv) integrable functions (the general case).

8.1.1 Stage One: Simple Functions

Suppose φ is a simple function with canonical form $\varphi(x) = \sum_{k=1}^{N} a_k \chi_{E_k}$ where a_k's are distinct and nonzero, and F_k's are disjoint. Then we define the **Lebesgue integral** of φ by

$$\int \varphi(x)\, dx := \sum_{j=1}^{N} a_j m(E_j).$$

If E is measurable with finite measure, then we define the integral on E by $\int_E \varphi = \int \varphi \chi_E$.

Property

(a) ***Independence of the representation***: *If $\varphi = \sum_{k=1}^{N} a_k \chi_{E_k}$ is any representation of φ, then $\int \varphi = \sum_{j=1}^{N} a_j m(E_j)$.*

(b) ***Linearity***: *If φ and ψ are simple and $a, b \in \mathbb{R}$, then $\int (a\varphi + b\psi) = a \int \varphi + b \int \psi$.*

(c) ***Additivity***: *If E and F are disjoint subsets of $\mathbb{R}^d$ with finite measure, then $\int_{E \cup F} \varphi = \int_E \varphi + \int_F \varphi$.*

(d) ***Monotonicity***: *If $\varphi \leq \psi$ are simple, then $\inf \varphi \leq \int \psi$.*

(e) ***Triangle inequality***: *If φ is a simple function, then so is $|\varphi|$, and $\left|\int \varphi\right| \leq \int |\varphi|$.*

8.1.2 Stage Two: Bounded Functions Supported on a Set of Finite Measure

Definition 8.1 (Support)

*The **support** of a measurable function f is defined to be the set of all points where f does not vanish, $supp(f) = \{x \mid f(x) \neq 0\}$; f is said to be supported on E if $supp(f) \subset E$.* ♣

Suppose f is a bounded function supported on a set E with finite measure, there is a sequence of simple functions $\{\varphi_k\}$ such that $\varphi_k \to f$. The goal is to define $\int f := \lim_{k\to\infty} \varphi_k$.

Lemma 8.1 (Well-definedness of Lebesgue Integral)

Let f be a bounded function supported on a set E of finite measure. Suppose $\{\varphi_k\}$ is a sequence of simple function bounded by M, support on E, and $\varphi_k \to f$ a.e.. Then

(a) *The limit $\lim \int \varphi_k$ exists.*

(b) *If $f = 0$ a.e., $\lim \int \varphi_k = 0$.*

Proof (a) By Egorov's Theorem, $\varphi_k \to f$ uniformly on some $F_\eta \subset E$ s.t. $m(E \setminus F_\eta) < \eta$, then there exists N s.t. $|\varphi_k - f| < \varepsilon/2$ for $k \geq N$. Then for $k, j \geq N$,

$$\int |\varphi_k - \varphi_j| = \int_{F_\eta} |\varphi_k - \varphi_j| + \int_{E \setminus F_\varepsilon} |\varphi_k - \varphi_j| \leq m(E)\varepsilon + 2M\eta.$$

By choosing appropriate ε and η, we can bound $|\varphi_k - \varphi_j|$ by an arbitrarily small number. Therefore $\{\int \varphi_k\}$ is Cauchy thus converges.

(b) We may choose F_η such that $m(E \setminus F_\eta) < \eta$, and $\varphi_k|_{F_\eta} < \varepsilon$ for sufficiently large k, applying the same argument as above yields that $\int |\varphi_k| \leq m(E)\varepsilon + M\eta$, hence we see that $\lim \int \varphi_k = 0$. ∎

Remark Consequently, it is valid to define the Lebesgue integral $\int f = \lim_{k\to\infty} \int \varphi_k$. The linearity, additivity monotonicity, and triangle inequality holds.

Proposition 8.1

Let f be a nonnegative bounded function supported on a set of finite measure. If $\int f = 0$, then $f = 0$ a.e.

Proof For an arbitrary α, $\int f \geq \int \alpha \chi_{\{f \geq \alpha\}} = \alpha m(\{f \geq \alpha\})$ by monotonicity, then $m(\{f \geq \alpha\}) \leq \frac{1}{\alpha} \int f$ [Chebyshev's Inequality]. Since $\int f = 0$, $m(\{f \geq 1/k\}) = 0$ for all k, then $\{f > 0\} = \bigcup_k \{f \geq 1/k\}$ has measure 0, followed by $f = 0$ a.e. ∎

Theorem 8.1 (Bounded convergence theorem (B.C.T.))

Suppose $\{f_k\}$ is a sequence of measurable bounded by M and supported on a set E of finite measure, $f_k \to f$ a.e. Then f is measurable, bounded and support on E a.e., and $\int |f_k - f| \to 0$. Consequently $\int f_k \to \int f$.

Proof Similar to Lemma 8.1, there is a F_η such that $m(E \setminus F_\eta) < \eta$ and $f_k \to f$ uniformly on F_η, then $|f_k - f| < \varepsilon$ on F_η for sufficiently large k. Then $\int |f_k - f| \leq \int_{F_\eta} |f_k - f| + \int_{E \setminus F_\eta} |f_k - f| \leq m(E)\varepsilon + 2M\eta$, followed by $\lim \int |f_k - f| = 0$, and $\int f_k \to \int f$ follows immediately from Proposition 8.1. ∎

Remark The bounded convergence theorem implies the validity of interchanging the integral and limit: $\lim \int f_n = \int \lim f_n$.

Proposition 8.2

Suppose f is Riemann integrable in $[a, b]$. Then f is Lebesgue measurable, and $\int_{[a,b]} f = \int_{[a,b]}^{\mathcal{R}} f$, namely two integrals agree over $[a, b]$.

Proof For each lower Riemann sum, we may write it as an integral of simple functions

$$L_\Gamma(f) = \sum_{k=1}^{N} \inf_{[x_{k-1},x_k]} f(x) \cdot (x_k - x_{k-1}) = \int_{[a,b]} \varphi \qquad \text{where } \varphi = \sum_{k=1}^{N} \inf f(x) \cdot \chi_{[x_{k-1},x_k]}.$$

By taking the refinement, we have a sequence $\varphi_1 \leq \varphi_2 \leq \cdots \leq f$. Analogously, we have a sequence $\psi_1 \geq \psi_2 \geq \cdots \geq f$ corresponding to upper Riemann sums. The sequences $\{\varphi_k\}, \{\psi_k\}$ are bounded. Then the Riemann integrablity implies that $\lim \int \varphi_k = \lim \int \psi_k$. Let $\tilde{\varphi}$, $\tilde{\psi}$ be the limits of $\{\varphi_k\}$, $\{\psi_k\}$, resp (the limit exists because they are monotonic and bounded). Note that $\int(\tilde{\psi} - \tilde{\varphi}) = \lim \int \psi - \lim \int \varphi = 0$, then $f = \tilde{\varphi} = \tilde{\psi}$ a.e. by Proposition 8.1. Hence $\int_{[a,b]} f = \int_{[a,b]} \tilde{\varphi} = \int_{[a,b]}^{\mathcal{R}} f$. ■

8.1.3 Stage Three: Nonnegative Measurable Functions

Definition 8.2 (Lebesgue Integral)

Let $f \geq 0$ be a measurable function. Define the (extended) ***Lebesgue integral*** $\int f(x)\,dx := \sup \int g(x)\,dx$ *where the supremum is taken over all measurable functions g such that $0 \leq g \leq f$, and where g is bounded and supported on a set of finite measure. We say f is* ***Lebesgue integrable*** *if $\int f(x)\,dx < +\infty$.* ♣

Proposition 8.3

The integral of non-negative measurable functions enjoys the following properties:

(a) *Linearity: If $a, b \geq 0$, f, g are nonnegative measurable functions, then $\int(af + bg) = a\int f + b\int g$.*

(b) *Additivity: If E and F are disjoint subsets of $\mathbb{R}^d$, and $f \geq 0$, then $\int_{E \sqcup F} = \int_E f + \int_F f$.*

(c) *Monotonicity: If $0 \leq f \leq g$, then $\int f \leq \int g$.*

(d) *If h is integrable, and $0 \leq f \leq h$, then f is integrable.*

(e) *If f is integrable, then $f < +\infty$ a.e.*

(f) *If $\int f = 0$, then $f = 0$ a.e.* ♠

Example 8.1 The analogy of bounded convergence theorem does not necessarily hold, i.e. $f_k \to f$ a.e. $\not\Rightarrow \int f_k \to \int f$. Consider the sequence of functions

$$f_k = n\chi_{(0,1/k)}.$$

Note that $f_k \to f := 0$, yet $\int f_k = 1$ for all k, it follows that $\int f = 0 \neq 1 = \lim \int f_k$.

Lemma 8.2 (Fatou)

Suppose $\{f_n\}$ is a sequence of measurable functions with $f_n \geq 0$. If $\lim_{n\to\infty} f_n(x) = f(x)$ for a.e. x, then $\int f \leq \liminf_{n\to\infty} \int f_n$. ♡

Proof Choose an arbitrary g for which $0 \leq g \leq f$ are bounded function supported on a finite measure set. Let

$g_k := \min\{g, f_k\} \leq g$, then note that $g_k \to g$ a.e. By the bounded convergence theorem (8.1), $\int g = \lim \int g_k$, then we have $\int g = \lim \int g_k \leq \liminf \int f_k$ since $g_k \leq f_k$, hence $\int f = \sup \int g \leq \liminf \int f_k$

Corollary 8.1

Suppose f is a non-negative measurable function, and $\{f_n\}$ a sequence of non-negative measurable functions with $f_n(x) \leq f(x)$ and $f_n(x) \to f(x)$ for almost every x. Then $\lim_{n\to\infty} \int f_n = \int f$. ♡

Proof Since $f_n(x) \leq f(x)$, we have $\int f_n \leq \int f$, it follows that $\limsup \int f_n \leq \int f$. Combined with Fatou (Lemma 8.2), we have $\int f = \lim \int f_n$. ■

Corollary 8.2 (Monotone convergence theorem (M.C.T))

(a) *Suppose $\{f_n\}$ is a sequence of non-negative measurable functions with $f_n \nearrow f$. Then $\lim_{n\to\infty} \int f_n = \int f$.*

(b) *Consider the series $\sum_k a_k(x)$ where $a_k \geq 0$ is measurable for all k, then $\int \sum_k a_k(x)\,dx = \sum_k \int a_k(x)\,dx$. Moreover, if $\sum \int a_k$ is finite, the series $\sum a_k(x)$ is convergent for a.e. x.* ♡

Proof (a) follows immediately from Corollary 8.1.

(b) The first statement follows by taking $f_j = \sum_{k=1}^{j} a_k(x)$ and note that $f_j \nearrow \sum_k a_k$. The second statement follows from Proposition 8.3 (e). ■

8.1.4 Stage Four: General Case

Definition 8.3 (Lebesgue Integral)

*Let f be a real-valued measurable function on $\mathbb{R}^d$, we say that f is **Lebesgue integrable** if $|f|$ is integrable as a nonnegative function.*

*If f is Lebesgue integrable, let $f^+(x) = \max(f, 0)$ and $f^-(x) = \max(-f, 0)$, and define the **Lebesgue integral** of f by $\int f = \int f^+ - \int f^-$.* ♣

Property *The integral of Lebesgue integrable functions is linear, additive, monotonic, and satisfies the triangle inequality.*

Proposition 8.4

Suppose f is integrable on $\mathbb{R}^d$, then for every $\varepsilon > 0$:

(a) *There exists a set of finite measure B such that $\int_{B^c} |f| < \varepsilon$.*

(b) *There is a $\delta > 0$ such that $\int_E |f| < \varepsilon$ whenever $m(E) < \delta$, i.e., the absolute continuity holds.* ♠

Remark (a) implies that integrable function vanish near ∞ (however, is is not true that $\lim_{|x|\to\infty} f(x) = 0$).
(b) implies that the map $\mathcal{M} \to \mathbb{R}^+$ defined by $E \mapsto \int_E |f|$ is (absolute) continuous.

Proof (a) WLOG, assume $f \geq 0$. Let B_k be the ball centered at the origin with radius k, and put $f_k = f\chi_{B_k} \nearrow f$. By monotone convergence theorem (Corollary 8.2), $\lim_k \int f_k = \int f$. Then there is N such that $|\int f - \int f_k| < \varepsilon$ for $k \geq N$, followed by B_N is the desired set.

(b) Let $E_k = \{f \leq k\}$ and $f_k = f\chi_{E_k}$, then $f_k \nearrow f$. Then for sufficiently large k, $|\int f - \int f_k| < \varepsilon/2$. Choose $\delta = \varepsilon/2k$, then for E such that $m(E) < \delta$,

$$\int_E f \leq \int_E f_k + \int_E (f - f_k) \leq m(E)k + \frac{\varepsilon}{2} < \varepsilon. \qquad \blacksquare$$

Theorem 8.2 (Dominated convergence theorem (D.C.T))

Suppose $\{f_n\}$ is a sequence of measurable functions such that $f_n(x) \to f(x)$ a.e. x. If $|f_n(x)| \leq g(x)$ a.e., where g is integrable, then $\lim \int f_k = \int f$. ♡

Remark In fact, $\lim \int |f_k - f| \to 0$.

Proof Note that $-g \leq f_k \leq g$ for all k, then each f_k and f are integrable. By Fatou (Lemma 8.2), since $g \pm f_k, g \pm f \geq 0$ a.e. and $g \pm f_k \to g \pm f$,

$$\begin{aligned}
\int (g+f) \leq \liminf \int (g + f_k) = \int g + \liminf \int f_k \quad &\Longrightarrow \quad \int f \leq \liminf \int f_k, \\
\int (g-f) \leq \liminf \int (g - f_k) = \int g - \limsup \int f_k \quad &\Longrightarrow \quad \int f \geq \limsup \int f_k.
\end{aligned}$$

Combining both inequalities gives $\lim \int f_k = \int f$. $\blacksquare$

8.1.5 Complex-valued Functions

A complex-valued function $f : \mathbb{R}^d \to \mathbb{C}$ may be written as $f(x) = u(x) + iv(x)$ where u, v are the real and imaginary parts of f, resp. The complex-valued function f is ***Lebesgue integrable*** if $|f| := \sqrt{|u|^2 + |v|^2}$ is integrable, if and only if u, v are integrable. In such case, we define its ***Lebesgue integral*** as

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

Note that $|(f+g)(x)| \leq |f(x)| + |g(x)|$, the monotonicity yields that $f + g$ is integrable if f, g are integrable; and similarly af is integrable if f is integrable. Therefore, the integral is linear over $\mathbb{C}$.

8.2 The Space L1 of Integrable Functions

8.2.1 L1 Space

Definition 8.4 (Norm, L^1 Space)

For any integrable function f on $\mathbb{R}^d$ (over $\mathbb{C}$) we define the norm of f, $\|f\|_{L^1} = \int_{\mathbb{R}^d} |f(x)|\,dx$. The space $\mathbf{L^1}(\mathbb{R}^\mathbf{d})$ is the space of equivalence classes of integrable functions, where the define $f \sim g$ if $f = g$ a.e. ♣

Property *$L^1(\mathbb{R}^d)$ inherits the property of vector spaces: suppose f, g are two functions in $L^1(\mathbb{R}^d)$,*

(i) $\|\alpha f\|_{L^1(\mathbb{R}^d)} = |a| \|f\|_{L^1(\mathbb{R}^d)}$ *for all $a \in \mathbb{C}$.*

(ii) $\|f + g\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} + \|g\|_{L^1(\mathbb{R}^d)}$.

(iii) $\|f\|_{L^1(\mathbb{R}^d)} = 0$ *if and only if $f = 0$ a.e.*

(iv) $d(f, g) = \|f - g\|_{L^1(\mathbb{R}^d)}$ *defines a metric on $L^1(\mathbb{R}^d)$.*

Theorem 8.3 (Riesz-Fischer)

*The vector space L^1 is **complete** in its metric.* ♡

Proof Suppose $\{f_n\}$ is Cauchy in $L^1(\mathbb{R}^d)$. For each k, we may choose f_{n_k} such that $n_k > n_{k-1}$ and $\|f_{n_k} - f_m\| < 2^{-k}$ for $m \geq n_k$; then the subsequence $\{f_{n_k}\}$ satisfies that $\|f_{n_k} - f_{n_k+1}\| < 2^{-k}$. Define $f = f_{n_1} + \sum_{k=1}^{\infty} (f_{n_{k+1}} - f_{n_k})$ and $g = |f_{n_1}| + \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|$.
By M.C.T. (monotone convergence theorem), $\int g = \int |f_{k_1}| + \sum_{j=1}^{\infty} |f_{k_{j+1}} - f_{k_j}| \leq \int |f_{n_1}| + \sum_{k=1}^{\infty} 2^{-k} < \infty$, so g is integrable. By D.C.T. (dominated convergence theorem), since $|f_{n_k}| \leq g$ and the partial sum of f is simply f_{n_k}, i.e., $f_{n_k} \nearrow f$, then $\|f_{n_k} - f\| = \int |f_{n_k} - f| \to 0$, namely f_{n_k} converges to f both pointwise a.e. and in L^1. Finally, $\{f_k\} \to f$ in L^1 since $\{f_k\}$ is Cauchy and contains a convergence subsequence. ■

Remark Summary: We first find a subsequence $\{f_{n_k}\}$ whose norm stabilizes rapidly, then apply D.C.T. to prove the convergence of $\{f_{n_k}\} \to f$ in L^1, finally we show that containing a convergent subsequence implies the convergence of $\{f_n\}$.

Corollary 8.3

If $\{f_n\}_{n=1}^{\infty}$ converges to f in L^1, then there exists a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \to f(x)$ a.e. x. ♡

Remark Note that $\{f_n\}$ converges in L^1 does not implies that $f_n \to f$ a.e., indeed, f_n may converge nowhere to f.

Proposition 8.5

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

(a) simple functions

(b) *step functions*

(c) *the continuous functions of compact support (denoted by $C_c(\mathbb{R}^d)$)*

Proof It suffices to consider nonnegative real-valued functions f.

(a) By Proposition (7.13), there exists a sequence of nonnegative simple functions $\varphi_k \nearrow f$. Note that $\lim \int \varphi_k = \int f$ by M.C.T, then $\|f - \varphi_k\| = \int (f - \varphi_k) \to 0$.

(b) It suffices to approximate χ_E by step functions. By Littlewood's first principle, there exists rectangles $\{Q_j\}_j$ such that $m(E \triangle Q) < \varepsilon$ where $Q := \bigcup Q_j$. Then $\|\chi_Q - \chi_E\| \leq m(E \triangle Q) < \varepsilon$.

(c) It suffices to approximate χ_Q by functions in $C_c(\mathbb{R}^d)$. Let $Q' \supseteq Q$ be a rectangle such that $m(Q' \setminus Q) < \varepsilon$. Define f such that $f|_Q = 1$, $f|_{Q'^c} = 0$, and $f|_{Q' \setminus Q}$ is linear, then f is continuous and supported on Q'. Note that $\|f - \chi_Q\| = \|f|_{Q' \setminus Q}\| \leq m(Q' \setminus Q) < \varepsilon$. ■

8.2.2 Invariance Properties, Translation and Continuity

Proposition 8.6 (Invariance Properties)

(a) ***Translation Invariance***: *For all $f \in L^1(\mathbb{R}^d)$, $\int f(x-h)\,dx = \int f(x)\,dx$.*

(b) ***Dilation Invariance***: *$\delta^d \int_{\mathbb{R}^d} f(\delta x)\,dx = \int_{\mathbb{R}^d} f(x)\,dx$ for $\delta > 0$.*

(c) ***Reflection Invariance***: *$\int f(-x)\,dx = \int f(x)\,dx$*

Proof By Proposition 8.5, the family of simple functions are dense, it suffices to show the translation invariance for χ_E: $\int \chi_E(x-h)\,dx = \int_{x-h \in E} dx = \int_{x \in E+h} dx = m(E+h) = m(E) = \int \chi_E\,dx$. The proof for (b) and (c) are analogous. ■

In particular, suppose $f, g \in L^1(\mathbb{R}^d)$ such that $y \mapsto f(x-y)g(y)$ is integrable for some fixed x. In such case, $\int f(x-y)g(y)\,dx = \int f(y)g(x-y)\,dx$.

Definition 8.5 (Convolution)

*Suppose $f, g \in L^1(\mathbb{R}^d)$, we define the **convolution** for f, g by $(f * g)(x) := \int f(y)g(x-y)\,dy$.*

Proposition 8.7

Let $f_h(x) := f(x-h)$. Suppose $f \in L^1(\mathbb{R}^d)$, then $\|f_h - f\|_{L^1} \to 0$ as $h \to 0$.

Proof Let $g \in C_c(\mathbb{R}^d)$ (continuous function of compact support), clearly $\|g_h - g\| \to 0$ as $|h| \to 0$. For every $f \in L^1(\mathbb{R}^d)$, $\|g - f\|$ can be bounded arbitrarily small for some $g \in C_c(\mathbb{R}^n)$ by Proposition 8.5. Then the triangle inequality

$$\|f_h - f\| \leq \|f_h - g_h\| + \|g_h - g\| + \|g - f\| = 2\|g - f\| + \|g_h - g\|$$

implies that $\|f_h - f\| \to 0$ as $|h| \to 0$. ■

8.3 Fubini's Theorem

8.3.1 Fubini's Theorem and Tonelli's Theorem

For a set $E \subseteq \mathbb{R}^m \times \mathbb{R}^n$, we define its **slices** by $E_y = \{x \in \mathbb{R}^m \mid (x, y) \in E\}$. Suppose $f(x, y)$ is a function defined on $\mathbb{R}^m \times \mathbb{R}^n$, we then define its **slices** corresponding to $y \in \mathbb{R}^n$ by $f_x(y) : x \mapsto f(x, y)$.

Example 8.2 It is not necessarily true that E being measurable in $\mathbb{R}^{m+n}$ implies that E_x if measurable for all $x \in \mathbb{R}^m$. Consider the Vitali set V and let $E = V \times \{0\}$. E is clearly measurable in $\mathbb{R}^2$ because E is the subset of a null set $\mathbb{R} \times \{0\}$, but the slice E_y with $y = 0$ is not measurable.

Note that the above statement holds for "almost every" $x \in \mathbb{R}^m$. This is an immediately corollary of Fubini's Theorem by taking the function $f = \chi_E$.

Theorem 8.4 (Fubini's Theorem)

Suppose $f(x, y)$ is integrable on $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$, then for almost every $x \in \mathbb{R}^m$:

(a) The slice $f_x : y \mapsto f(x, y)$ is (measurable and) integrable on $\mathbb{R}^n$ for each $x \in \mathbb{R}^m$ fixed.

(b) The function defined by $\int_{\mathbb{R}^n} f_x(y)\, dy$ is (measurable and) integrable on $\mathbb{R}^m$.

Moreover,

(c) $\int_{\mathbb{R}^{m+n}} f(x, y)\, dx\, dy = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y)\, dy\right) dx$.

♡

Proof Let $\mathcal{F}$ denotes the family of functions in $L^1(\mathbb{R}^{m+n})$ who satisfy the above conditions.

Step 1: ($\mathcal{F}$ is closed under linear combination): Suppose $\{f_k\}_{k=1}^N$ is a finite collection of functions in $\mathcal{F}$, then $\sum_{k=1}^N a_k f_k \in \mathcal{F}$ for $a_k \in \mathbb{R}$.

Proof: The above three conditions follow from the linearity of Lebesgue integral.

Step 2: ($\mathcal{F}$ is closed under (monotonic) limit) Suppose $\{f_k\}_k$ is a sequence of functions in $\mathcal{F}$ such that $f_k \nearrow f$ (or respectively $f_k \searrow f$) where $f \in L^1(\mathbb{R}^{m+n})$, then $f \in \mathcal{F}$.

Proof: (i) For a.e. x (on where all f_k satisfy the above conditions), note that $f_x : y \mapsto f(x, y) = \sup_k f_k(x, y)$ and $f_k(x, y)$ is measurable, then f_x measurable. Since $f_k(x, \cdot) \nearrow f(x, \cdot)$, by M.C.T. (NB: the nonnegative condition can be easily satisfied by subtracting each function by $f_1(x, y)$), we see that $h_k(x) := \int f_k(x, y)\, dy \nearrow \int f(x, y)\, dy =: h(x)$. We will prove the integrability after part (iii).

(ii) For a.e. x, since $x \mapsto h(x) = \sup h_k(x)$ and $h_k(x)$ is measurable, then $x \mapsto h(x)$ is measurable.

(iii) For a.e. x, apply M.C.T. (to $h_k - h_1$ as above), we have $\int h_k(x)\, dx \nearrow \int h(x)\, dx$. It suffices to show that $\int h_k(x) \nearrow \iint f(x, y)\, dx\, dy$. Apply the third condition for f_k and M.C.T again, we see that $\int h_k(x)\, dx = \iint f_k(x, y)\, dx\, dy \nearrow \iint f(x, y)\, dx\, dy$.

Note that $\int h(x)\, dx = \iint f(x, y)\, dx\, dy < +\infty$, this proves the integrability of $h(x) = \int f_x(y)\, dy$ in (ii); indeed, $\int f_x(y)\, dy = h(x) < +\infty$ for a.e. x, proving the integrability of f_x in (i).

Remark Step reduced the proof to $0 \leq f \in L^1(\mathbb{R}^{m+n})$. Step 2 and the fact that any $f \geq 0$ can be approximated by $\{\varphi_k\}$ simple functions s.t. $\varphi_k \nearrow f$ a.e. implies that if we know $\varphi_k \in \mathcal{F}$ and $f \in L^1(\mathbb{R}^{m+n})$, then $f \in \mathcal{F}$.

Therefore, we may reduce the proof to characteristic functions χ_E where E is measurable with finite measure, which can be further deduced to G_δ set by Proposition 7.6.

Step 3: Let E be a G_δ set in $\mathbb{R}^{m+n}$ with finite measure, then $\chi_E \in \mathcal{F}$.

Proof: It suffices to show the statement for open set G of finite measure, by step 2. We can write G as $G = \bigsqcup Q_j$ where Q_j's are almost disjoint cubes, thus $G = (\bigsqcup Q_j^\circ) \sqcup (\bigsqcup F_j)$ where $E_j \subseteq \partial Q_j$.

(a) In order to show $\chi_{\bigsqcup Q_j^\circ} \in \mathcal{F}$, by step 1 and 2, it suffices to show $\chi_{Q_j^\circ} \in \mathcal{F}$. Suppose $\chi_{I_1 \times I_2} \in \mathcal{F}$ where I_1 and I_2 are open cubes in $\mathbb{R}^m$ and $\mathbb{R}^n$, resp. (i) For $x \in \mathbb{R}^m$, if $x \in I_1$, $(I_1 \times I_2)_x = I_2$ is finite-measurable in $\mathbb{R}^n$; on the other hand if $x \notin I_1$, $(I_1 \times I_2)_x = \varnothing$. (ii) $x \mapsto m_{\mathbb{R}^n}((I_1 \times I_2)_x)$ is $v(I_2)\chi_{I_1}$, thus it is measurable. (iii) $\iint_{\mathbb{R}^{m+n}} \chi_{I_1 \times I_2} = m(I_1 \times I_2) = |I_1 \times I_2| = \int_{\mathbb{R}^n} m(I_1)\chi_{I_2} = \int_{\mathbb{R}^m} (\int_{\mathbb{R}^n} \chi_{(I_1 \times I_2)_x} dy) dx$.

(b) We want to show that $E \subseteq \partial Q$ where Q is a cube in $\mathbb{R}^{m+n}$, then $\chi_E \in \mathcal{F}$. It is not hard to show this statement since E is a hyperplane in $\mathbb{R}^{m+n}$

Hence $G \in \mathcal{F}$ and thus we proved the statement.

Step 4: (null set belongs to $\mathcal{F}$) Let N be a null set in $\mathbb{R}^{m+n}$, then $\chi_N \in \mathcal{F}$. In particular, the slice N_x is a null set for a.e. $x \in \mathbb{R}^n$.

Proof: There exists a G_δ set H s.t. $N \subseteq H$ and $m_{\mathbb{R}^{m+n}}(H) = 0$. Then $N_x \subseteq H_x$ by definition. By step 3, $m_{\mathbb{R}^{m+n}}(H) = \int_{\mathbb{R}^n} m_{\mathbb{R}^n}(H_x)\, dx$ implies $m_{\mathbb{R}^n}(H_x) = 0$ for a.e. $x \in \mathbb{R}^m$, therefore $m_{\mathbb{R}^n}(N_x) = 0$ a.e.. Therefore, $N \in \mathcal{F}$.

Step 5: $\chi_E \in \mathcal{F}$ if E is finite-measurable in $\mathbb{R}^{m+n}$.

Proof: The statement follows immediately from Proposition 7.6, E differs from a G_δ set by a null set.

Step 6: Every function $f \in L^1(\mathbb{R}^{m+n})$ belongs to $\mathcal{F}$, namely Fubini's Theorem holds.

Proof: For all $f \in L^1(\mathbb{R}^{m+n})$, $\varphi_k \nearrow f$ for some increasing sequence of simple functions $\{\varphi_k\}$, and $\varphi_k \in \mathcal{F}$ by step 6 and step 1, then $f \in \mathcal{F}$ by step 2. ■

Remark The converse does not necessarily hold, i.e., it is not the case that f measurable in $\mathbb{R}^{m+n}$ and $\int_{\mathbb{R}^m} (\int_{\mathbb{R}^n} f_x(y)\, dy)\, dx$ being finite implies that $f \in L^1(\mathbb{R}^{m+n})$.

Example 8.3 Consider the union of cubes $Q_j \subseteq \mathbb{R}^2$ aligning in the diagonal with side length 2^{-j}. For each cube, we subdivide it into 4 sub-cubes, and assign $f(x) = 1/|Q_j|$ for upper left and lower right sub-cubes and $f(x) = -1/|Q_j|$ for the other two. It is not hard to show that the slices are zero everywhere. However, $\iint_{\mathbb{R}^2} f$ is not defined because $\iint_{\mathbb{R}^2} |f| = \sum_j m(Q_j) \cdot (1/|Q_j|) = +\infty$.

Theorem 8.5 (Tonelli's Theorem)

Let $f(x, y)$ be a nonnegative measurable function in $\mathbb{R}^{m+n}$. Then

(i) for a.e. $x \in \mathbb{R}^m$, the slice $f_x : y \mapsto f(x, y)$ is measurable in $\mathbb{R}^n$;

(ii) the function $x \mapsto \int_{\mathbb{R}^n} f_x\, dy$ is measurable in $\mathbb{R}^m$ (in the extended real number system); and moreover,

(iii) $\int_{\mathbb{R}^{m+n}} f(x, y)\, dx\, dy = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x, y)\, dy \right) dx$ (in the extended real number system). ♡

Remark Fubini-Tonelli's Theorem: we commonly apply Tornell's theorem to $|f|$ to check $f \in L^1(\mathbb{R}^{m+n}) \Leftrightarrow |f| \in L^1(\mathbb{R}^{m+n})$, then we may compute $\iint_{\mathbb{R}^{m+n}} f(x,y)\,dy\,dx$ using Fubini's Theorem.

Proof Cosntruct $f_k \in L^1(\mathbb{R}^{m+n})$ s.t. $f_k \nearrow f$ by defining

$$f_k(x,y) := \begin{cases} 0 & \text{if } |(x,y)| > k \\ \min\{f(x,y), k\} & \text{if } |(x,y)| \le k \end{cases}$$

Since f_k is nonnegative, bounded and supported on a set of finite measure, f_k is integrable. Moreover, $f_k(x,y) \nearrow f(x,y)$ for all (x,y), and $f_k \in \mathcal{F}$ for all k. M.C.T. implies that (i) $\iint_{\mathbb{R}^{m+n}} f_k \nearrow \iint_{\mathbb{R}^{m+n}} f$. For a.e. x, $(f_k)_x$ is measurable and integrable, then $(f_k)_x \nearrow f_x$ for a.e. x, so $h_k(x) := \int_{\mathbb{R}^n} f_k(x,y)\,dy \nearrow \int_{\mathbb{R}^n} f(x,y)\,dy =: h(x)$, and thus (ii) $\int_{\mathbb{R}^n} h_k(x)\,dx = \int_{\mathbb{R}^n} h(x)\,dx$. Note that $\iint_{\mathbb{R}^{m+n}} f_k = \int_{\mathbb{R}^n} h_k(x)\,dx$ by Fubini's Theorem, hence $\int_{\mathbb{R}^{m+n}} f(x,y)\,dx\,dy = \int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^n} f(x,y)\,dy\right) dx$ follows from (i), (ii), and the uniqueness of limit. ■

8.3.2 Application of Fubini's Theorem

Proposition 8.8

Let E_1, E_2 be measurable sets in $\mathbb{R}^m, \mathbb{R}^n$, resp. Then $E = E_1 \times E_2$ is measurable in $\mathbb{R}^{m+n}$ with $m_{\mathbb{R}^{m+n}}(E) = m_{\mathbb{R}^m}(E_1) m_{\mathbb{R}^n}(E_2)$, with the understanding that if one of the sets has measure 0, then $m(E) = 0$. ♠

Proof If E is measurable in $\mathbb{R}^{m+n}$, apply Tonelli's theorem to χ_E, we see that

$$m_{\mathbb{R}^{m+n}}(E) = \int_{\mathbb{R}^m} m_{\mathbb{R}^n}(E_x)\,dx \int_{E_1} m_{\mathbb{R}^n}(E_2)\,dx = m_{\mathbb{R}^m}(E_1) m_{\mathbb{R}^n}(E_2).$$

It suffices to show E is measurable. Note that there is G_δ set H_1 s.t. $H_1 \supseteq E_1$ and $m_{\mathbb{R}^m}(H_1 \setminus E_1) = 0$, and analogously there is H_2 corresponding to E_2. $E \subseteq H_1 \times H_2$ is a G_δ set in $\mathbb{R}^{m+n}$, and

$$(H_1 \times H_2) \setminus E = (H_1 \times H_2) \setminus (E_1 \times E_2) \subseteq ((H_1 \setminus E_1) \times H_2) \cup (H_1 \times (H_2 \setminus E_2)).$$

Lemma 1: $m_*(A_1 \times A_2) \le m_*(A_1) \times m_*(A_2)$ where $A_1 \subseteq \mathbb{R}^m$ and $A_2 \subseteq \mathbb{R}^n$, with the understanding that if one of the sets has exterior measure 0, then $m_*(A_1 \times A_2) = 0$..

Proof: Let $\varepsilon > 0$. By the definition of outer measure, there exists $\{Q_j^1\}, \{Q_j^2\}$ covering of A_1, A_2 by cubes s.t. $\sum_j |Q_j^1| \le m_*(A_1) + \varepsilon$, and similarly for A_2. Then $A_1 \times A_2 \subseteq \left(\bigcup_j Q_j^1\right) \times \left(\bigcup_j Q_j^2\right) = \bigcup_{i,j} (Q_i^1 \times Q_j^2)$, followed by

$$m_*(A_1 \times A_2) \le \sum_{i,j} |Q_i^1 \times Q_j^2| \le \sum_j |Q_j^1| \cdot \sum_j |Q_j^2| \le (m_*(A) + \varepsilon)(m_*(B) + \varepsilon).$$

If $m_*(A_1), m_*(A_2) < +\infty$, pass $\varepsilon \to 0$, we see that $m_*(A_1 \times A_2) \le (m_*(A)+\varepsilon)(m_*(B)+\varepsilon)$. On the other hand, if $m_*(A_1) = 0$ and $m_*(A_2) = +\infty$, then $A_2^j := A_2 \cap \{y \in \mathbb{R}^n \mid |y| \le j\} \nearrow A_2$, we see that $A_1 \times A_2 = \bigcup_j (A_1 \times A_2^j)$ is a null set. ■

Apply the lemma to the above equality, we see that $(H_1 \setminus E_1) \times H_2$ and $H_1 \times (H_2 \setminus E_2)$ are null sets, so E is measurable since it differs from a G_δ set by a null set. ■

Proposition 8.9

Let f be a nonnegative function on $\mathbb{R}^n$ and $\mathcal{A} := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \,|\, 0 \leq y \leq f(x)\}$. Then

(i) f is measurable on $\mathbb{R}^n$ if and only if $\mathcal{A}$ is measurable in $\mathbb{R}^{n+1}$.

(ii) If the condition in (i) holds, $\int_{\mathbb{R}^n} f(x)\, dx = m_{\mathbb{R}^{n+1}}(\mathcal{A})$.

Remark The Riemann integral of $f \geq 0$ can be viewed as area below the graph of f, we generalize it to Lebesgue integral to be the measure below the graph.

Proof (ii) Apply Tonelli's theorem to $\chi_{\mathcal{A}}$,

$$m_{\mathbb{R}^{n+1}}(\mathcal{A}) = \int_{\mathbb{R}^n} m_{\mathbb{R}}(\mathcal{A}_x)\, dx = \int_{\mathbb{R}^n} f(x)\, dx,$$

where the second equality holds by the fact that $\mathcal{A}_x := \{y \in \mathbb{R} \,|\, (x, y) \in \mathcal{A}\} = [0, f(x)]$. ∎

Chapter 9 Lebesgue Differentiation Theory

9.1 Differentiation of the Integral

Theorem 9.1 (Lebesgue Differentiation Theorem (L.D.T.))

If $f \in L^1(\mathbb{R}^n)$, then

$$\lim_{Q \to x} \frac{1}{m(Q)} \int_Q f = f(x) \quad \text{for a.e. } x \in \mathbb{R}^n,$$

where Q's are cubes containing x passing to the limit $m(Q) \to 0$. ♡

Proof The desired statement holds for all $g \in C_C(\mathbb{R}^n)$ (continuous function on compact support) because they are uniformly continuous.

Let $f \in L^1(\mathbb{R}^n)$ and $\varepsilon > 0$. There exists $g \in C_C(\mathbb{R}^n)$ such that $\|f - g\|_{L^1} < \varepsilon$ since $\mathbb{C}_C(\mathbb{R}^n)$ is dense. Note that $\frac{1}{m(Q)} \int_Q f - f(x) = [\frac{1}{m(Q)} \int_Q (f-g)] + [\frac{1}{m(Q)} \int_Q g - g(x)] + [g(x) - f(x)]$, by taking the limit,

$$\limsup_{Q \to x} \left| \frac{1}{m(Q)} \int_Q f - f(x) \right| \le \limsup_{Q \to x} \left| \frac{1}{m(Q)} \int_Q (f-g) \right| + \underbrace{\limsup_{Q \to x} \left| \frac{1}{m(Q)} \int_Q g - g(x) \right|}_{=0} + |g(x) - f(x)| .$$

For any $\alpha > 0$, define

$$E_\alpha := \left\{ x \in \mathbb{R}^n : \limsup_{Q \to x} \left| \frac{1}{m(Q)} \int_Q f - f(x) \right| > 2\alpha \right\},$$

it suffices to prove $m(E_\alpha) \to 0$ as $\alpha \to 0$. Note that

$$E_\alpha \subseteq \underbrace{\left\{ x \in \mathbb{R}^n : \limsup_{Q \to x} \left| \frac{1}{m(Q)} \int_Q (f-g) \right| > \alpha \right\}}_{A_\alpha} \cup \underbrace{\left\{ x \in \mathbb{R}^n : |g(x) - f(x)| > \alpha \right\}}_{B_\alpha}.$$

By Tchebychev's inequality, $m(B_\alpha) \le \frac{1}{\alpha} \int |g(x) - f(x)| \, dx = \|f - g\|/\alpha < \varepsilon/\alpha$, thus $m(B_\alpha) = 0$ as ε can be arbitrarily small. We now consider the measure of A_α, let's first define the maximal function.

Definition 9.1 (Hardy-Littlewood Maximal Function)

Let $h \in L^1(\mathbb{R}^n)$, we define its ***Hardy-Littlewood maximal function*** *of h as*

$$h^*(x) := \sup_{Q \ni x} \frac{1}{m(Q)} \int_Q |h|.$$

♣

Lemma 9.1 (Elementary version of Vitali lemma)

Suppose $\mathcal{F} = \{Q_1, \cdots, Q_N\}$ is a finite collection of cubes in $\mathbb{R}^d$, then there exists a disjoint subcollection $\{Q_{i_j}\}_j$ of $\mathcal{F}$ s.t. $m(\bigcup_{i=1}^N Q_i) \leq 3^d \sum_{j=1}^l m(Q_{i_j})$. ♡

Proof We claim that if two cubes Q, R intersect, where $l(Q) \leq l(R)$, then $Q \subseteq 3R$ (where $3R$ is defined to be the cube centered at the center of R with triple side length). Let Q_{i_1} be the cube in $\mathcal{F}$ with largest side length, and define $\mathcal{F}_1 = \{Q_k \in \mathcal{F} : Q_k \cap Q_{i_1} = \varnothing\}$ to be the set of cubes which does not intersects Q_{i_1}. We then recursively choose Q_{i_j} to be the largest cube in $\mathcal{F}_{j-1}$ and define the corresponding $\mathcal{F}_j$. Then we see $\{Q_{i_j}\}$ are pairwise disjoint and covers all cubes in $\mathcal{F}$, thus $m(\bigcup_i Q_i) \leq m(\bigcup_j 3Q_{i_j}) \leq 3^d \sum_j m(Q_{i_j})$ as desired. ■

Lemma 9.2 (Hardy-Littlewood)

Suppose h is integrable on $\mathbb{R}^d$. Then

(a) h^* is measurable.

*(b) h^* belongs to **weak-$L^1(\mathbb{R}^d)$**: for some constant C, h^* satisfies $m(\{h^* > \alpha\}) \leq (C/\alpha) \cdot \|h\|_{L^1}$ for all $\alpha > 0$.* ♡

Proof (a) For any $\lambda > 0$. For x s.t. $f^*(x) > \lambda$, there exists $Q \ni x$ s.t. $\frac{1}{m(Q)} \int_Q |f| > \lambda$. Then for all $y \in Q$, $f^*(y) \geq \frac{1}{m(Q)} \int_Q |f| > \lambda$, i.e., $Q \subseteq \{f^* > \lambda\}$, hence $\{f^* > \lambda\}$ is open thus f^* is measurable.

(b) For each $x \in \{f^* > \alpha\}$, there is $Q_x \ni x$ s.t. $\frac{1}{m(Q_x)} \int_{Q_x} |f| > \alpha$. Let $K \subseteq \{h^* > \alpha\}$ be an arbitrary compact subset, there exists $\{x_1, \cdots, x_N\}$ such that $K \subseteq \bigcup_{i=1}^N Q_i^\circ \subseteq \bigcup_{i=1}^N Q_i$ where $Q_i := Q_{x_i}$ by the compactness. Apply Lemma 9.1, there exists pairwise disjoint collection $\{Q_{i_j}\}_j$ s.t. $m(\bigcup_i Q_i) \leq 3^d \sum_j m(Q_{i_j})$, therefore $m(K) \leq 3^d \sum_j m(Q_{i_j})$.

Recall that each Q_x satisfies that $m(Q_x) < \frac{1}{\alpha} \int_{Q_x} |f|$, then $m(Q_{i_j}) < (1/\alpha) \cdot \int_{Q_{i_j}} |f|$, followed by

$$\sum_{j=1}^l m(Q_{i_j}) \leq \frac{1}{\alpha} \int_{\bigsqcup_{j=1}^l Q_{i_j}} |f| \leq \frac{1}{\alpha} \int_{\mathbb{R}^d} |f|.$$

Combining the above inequality $m(K) \leq 3^d \sum_j m(Q_{i_j})$, we see $m(K) \leq (3^d/\alpha) \cdot \|f\|_{L^1}$ for any compact $K \subseteq \{f^* > \alpha\}$, hence the desired statement holds. ■

Remark It is not necessary that $f \in L^1(\mathbb{R}^d)$ implies $f^* \in L^1(\mathbb{R}^d)$. (Ex. 4)

Proof (cont. Theorem 9.1) By Lemma 9.2, we see that $m(A_\alpha) \leq (C/\alpha) \cdot \|f - g\|_{L^1} < C\varepsilon/\alpha$, then $m(A_\alpha) = 0$ since ε can be arbitrarily small. Therefore, $m(E) \leq m(A_\alpha) + m(B_\alpha) = 0$, completing the proof. ■

Definition 9.2 (Locally integrable)

*A measurable function f is **locally integrable**, denoted $f \in L_{loc}(\mathbb{R}^d)$, if $f \in L^1(B)$ for all balls B in $\mathbb{R}^d$.* ♣

Remark

(a) L.D.T holds even if $f \in L_{loc}(\mathbb{R}^d)$: suppose $f \in L_{loc}(\mathbb{R}^d)$, then $\lim_{Q \to x} \frac{1}{m(Q)} \int_Q f = f(x)$ for a.e. x.

(b) For any measurable $E \subseteq \mathbb{R}^d$, $\chi_E \in L_{loc}(\mathbb{R}^d)$. Apply L.D.T. to χ_E we obtain $\lim_{Q \to x} \frac{1}{m(Q)} \int_Q \chi_E = \chi_E(x)$ for a.e. $x \in \mathbb{R}^d$. Therefore,

$$\lim_{Q \to x} \frac{m(E \cap Q)}{m(Q)} = \lim_{Q \to x} \frac{1}{m(Q)} \int_Q \chi_E = \begin{cases} 1 & \text{for a.e. } x \in E, \\ 0 & \text{for a.e. } x \notin E, \end{cases}$$

and we refer to a point such that $\lim_{Q \to x} \frac{m(E \cap Q)}{m(Q)} = 1$ as **Lebesgue density point** of E.

(c) We have shown $\lim_{Q \to x} \frac{1}{m(Q)} \int_Q (f - f(x))\, dy = 0$ for a.e. x. In fact, $\lim_{Q \to x} \frac{1}{m(Q)} \int_Q |f(y) - f(x)|\, dy = 0$ for a.e. $x \in \mathbb{R}^n$; a point satisfies the above equality is called a **Lebesgue point** of f.

Corollary 9.1

If $f \in L_{loc}(\mathbb{R}^d)$, then $\lim_{Q \to x} \frac{1}{m(Q)} \int_Q |f(y) - f(x)|\, dy = 0$ for a.e. $x \in \mathbb{R}^n$. That is, almost every point is a Lebesgue point. ♡

Proof For each $q \in \mathbb{Q}$, apply L.D.T. to $|f(y) - q|$ gives that $\lim_{Q \to x} \frac{1}{m(Q)} \int_Q |f(y) - q| = |f(x) - q|$ for $x \in \mathbb{R}^d \setminus Z_q$, where Z_q is a null set; then the above equality holds for a.e. x (i.e., $x \notin \bigcup_{q \in \mathbb{Q}} Z_q$). Let $\varepsilon > 0$ be given. For such $x \in \mathbb{R}^d$, there is $q \in \mathbb{Q}$ such that $|f(x) - q| < \varepsilon$, then

$$\lim_{Q \to x} \frac{1}{m(Q)} \int_Q |f - f(x)| \leq \lim_{Q \to x} \frac{1}{m(Q)} \int_Q |f - q| + |q - f(x)| < 2\varepsilon.$$ ■

9.2 Approximations to the Identity

Let k be a bounded integrable function in $\mathbb{R}^n$ s.t. $\int k = 1$. Let $k_\delta(x) := \frac{1}{\delta^n} k(\frac{x}{\delta})$. We obtain the following observations

- $\int_{\mathbb{R}^n} k_\delta(x)\,dx = \int_{\mathbb{R}^n} k(x)\,dx$ (analogous for $|k_\delta(x)|$) by the dilation invariance.
- If k has compact support, denoted B_{R_0}, then k_δ is supported in $B_{\delta R_0}$.

For any $f \in L^1(\mathbb{R}^n)$, consider the convolution $f * k_\delta(x) := \int f(y)\delta_k(x-y)\,dy$. Recall that $\|f * k_\delta\|_{L^1} \leq \|f\|_{L^1}\|k_\delta\|_{L^1} = \|f\|_{L^1}\|k\|_{L^1}$.

Remark Under some additional assumptions on k, $f * k_\delta \to f$, this is known as the ***approximation to the identity***.

Proposition 9.1

Let k be a bounded integrable function in $\mathbb{R}^n$ s.t. $\int k = 1$. Suppose k has compact support, then $f * k_\delta(x) \to f(x)$ as $\delta \to 0$, for any x that is a Lebesgue point of f (in particular, for a.e. x). ♠

Remark Approximation to the identity: the proposition asserts that the map $f \mapsto f * k_\delta$ converges to the identity map $f \mapsto f$ as $\delta \to 0$.

Proof Suppose $|f(x)|$ is bounded by M, and supported on B_{R_0}. Note that

$$\begin{aligned} f * k_\delta(x) - f(x) &= \int f(x-y)k_\delta(y)\,dy - f(x) = \int f(x-y)k_\delta(y)\,dy - f(x)\int k_\delta(y)dy \\ &= \int [f(x-y) - f(x)]\,k_\delta(y)\,dy \end{aligned} \tag{9.2.1}$$

Suppose x is a Lebesgue points, the above equality yields

$$\begin{aligned} |f * k_\delta(x) - f(x)| &\leq \int_{|y|\leq \delta R_0} |f(x-y) - f(x)||k_\delta(y)|\,dy \leq \frac{M}{\delta^n}\int_{|y|\leq\delta R_0} |f(x-y)-f(x)|\,dy \\ &= \frac{M}{\delta^n} w_0(\delta R_0)^n \cdot \frac{1}{w_0(\delta R_0)^n}\int_{|z-x|\leq \delta R_0} |f(z) - f(x)|, \end{aligned}$$

where $w_0(\delta R_0)^n$ represents the volume of the ball $B_{\delta R_0}$. The first part of the expression is equivalent to $M w_0 R_0^n$, which is a constant independent of δ, and the second part converges to 0 as $\delta \to 0$ by the definition of Lebesgue point, hence we see $f * k_\delta(x) - f(x) \to 0$ as $\delta \to 0$. ■

Proposition 9.2

Let k be a bounded integrable function in $\mathbb{R}^n$ s.t. $\int k = 1$. Then $f * k_\delta(x) \to f(x)$ in L^1 as $\delta \to 0$. ♠

Remark Since CV in L^1, there is $\{\delta_j\} \to 0^+$ s.t. $f * k_\delta(x) \to f(x)$ for a.e. x (this does not imply $f * k_\delta(x) \to f(x)$ for a.e. x).

Proof Apply Equation (9.2.1), $f * k_\delta(x) - f(x) = \int [f(x-y) - f(x)]k_\delta(y)\,dy$. Apply Tonelli's theorem, we see that

$$\int |f * k_\delta(x) - f(x)|\ dx \leq \iint |f(x-y) - f(x)| \cdot |k_\delta(y)|\, dx\, dy\ = \int |k_\delta(y)| \cdot a(y)\, dy,$$

where $a(y) := \int |f(x-y) - f(x)|\, dx$, and it is clear that $\|a(y)\|_{L^1} \leq 2\|f\|_{L^1}$. Fix $\varepsilon > 0$, there is $\eta > 0$ s.t. $a(y) < \varepsilon$ for $y < \eta$ by Proposition 8.7. Then

$$\begin{aligned}\int |k_\delta(y)| \cdot a(y)\, dy &\leq \int_{|y|<\eta} |k_\delta(y)| \cdot a(y)\, dy + \int_{|y|>\eta} |k_\delta(y)| \cdot a(y)\, dy \\ &\leq \varepsilon \int |k_\delta(y)|\, dy + 2\|f\|_{L^1} \int_{|y|>\eta} |k_\delta(y)|\, dy = \varepsilon\|k\|_{L^1} + 2\|f\|_{L^1} \int_{|y|>\eta} |k_\delta(y)|\, dy.\end{aligned}$$

Note that $\int_{|y|>\eta} |k_\delta(y)|\, dy = \delta^{-n} \int_{|y|>\eta} |k(y/\delta)|\, dy = \int_{|z|>\eta/\delta} |k(z)|\, dz < \varepsilon$ for sufficiently small δ. Therefore, $\int |f * k_\delta(x) - f(x)|\ dx \leq \varepsilon\|k\|_{L^1} + 2\varepsilon\|f\|_{L^1}$, it follows that $f * k_\delta \to f(x)$ in L^1. ■

Lemma 9.3

Let $f \in L^1(\mathbb{R}^n)$ and x be a Lebesgue point of f. Let

$$a(r) = \frac{1}{r^n} \int_{|y|<r} |f(x-y) - f(y)|\, dy.$$

Then $a(r) \to 0$ as $r \to 0$; and moreover, $a(r)$ is bounded for all $r > 0$. ♡

Proof The convergence follows immediately by the definition of Lebesgue point. Then there exists $r_0 > 0$ s.t. $a(r) \leq 1$ whenever $r < r_0$. For $r \geq r_0$,

$$\begin{aligned} a(r) &\leq \frac{1}{r^n} \int_{|y|<r} |f(x-y)|\, dy + \frac{1}{r^n} \int_{|y|<r} |f(x)|\, dy \\ &\leq \frac{1}{r_0^n} \int |f(x-y)|\, dy + \frac{1}{r^n} \cdot (w_n r^n)|f(x)| = \frac{1}{r_0^n}\|f\|_{L^1} + w|f(x)|,\end{aligned}$$

so $a(r)$ is bounded. ■

Proposition 9.3

Let k be a bounded integrable function in $\mathbb{R}^n$ s.t. $\int k = 1$. Suppose $k(x) \in O(1/|x|^{n+\lambda})$ for some $\lambda > 0$, then $f * k_\delta(x) \to f(x)$ as $\delta \to 0$. ♠

Proof Similarly to previous two propositions,

$$|f * k_\delta(x) - f(x)| \leq \int |f(x-y) - f(x)||k_\delta(y)|\, dy = \underbrace{\int_{|y|<\delta}}_{(1)} + \underbrace{\int_{|y|>\delta}}_{(2)}.$$

Let k be bounded by M, then

$$(1) = \int_{|y|<\delta} |f(x-y) - f(x)| \cdot \frac{1}{\delta^n}|k(y/\delta)|\, dy \leq \frac{M}{\delta^n} \int_{|y|<\delta} |f(x-y) - f(x)|\, dy = Ma(\delta).$$

We denote by $\int_{(a)}$ the integration on annulus $2^k\delta < |y| < 2^{k+1}\delta$. For sufficiently small δ, $k(y/\delta) < c/|y/\delta|^{n+\lambda}$, then

$$
\begin{aligned}
(2) &= \sum_{k=0}^{\infty} \int_{2^k\delta<|y|<2^{k+1}\delta} |f(x-y)-f(x)||k_\delta(y)|\,dy \le \sum_k \int_{(a)} |f(x-y)-f(x)| \cdot \frac{1}{\delta^n} \cdot \frac{c}{|y/\delta|^{n+\lambda}}\,dy \\
&= c\delta^\lambda \sum_k \int_{(a)} |f(x-y)-f(x)| \cdot \frac{1}{|y|^{n+\lambda}}\,dy \le c\delta^\lambda \sum_k \frac{1}{(2^k\delta)^{n+\lambda}} \int_{(a)} |f(x-y)-f(x)|\,dy \\
&\le c\delta^\lambda \sum_k \frac{(2^{k+1}\delta)^n}{(2^k\delta)^{n+\lambda}} \cdot \frac{1}{(2^{k+1}\delta)^n} \int_{|y|<2^{k+1}\delta} |f(x-y)-f(x)|\,dy \\
&= c2^n \sum_k 2^{-k\lambda} a(2^{k+1}\delta).
\end{aligned}
$$

Combining both parts and the fact that $a(x)$ is bounded by some A, we see that

$$
|f * k_\delta(x) - f(x)| \le (1) + (2) \le Ma(\delta) + c2^n \sum_k 2^{-k\lambda} a(2^{k+1}\delta).
$$

For sufficiently large N, we have $\sum_{k\ge N} 2^{-k\lambda} < \varepsilon$; and for sufficiently small δ, we have $a(\delta) < \varepsilon$, and $a(2^{k+1}\delta) < \varepsilon / \sum_{k<N} 2^{-k\lambda}$ for $k < N$. Therefore,

$$
\begin{aligned}
|f * k_\delta(x) - f(x)| &\le MA + \sum_{k<N} 2^{-k\lambda} a(2^{k+1}\delta) + \sum_{k\ge N} 2^{-k\lambda} a(2^{k+1}\delta) \\
&\le M\varepsilon + \sum_{k<N} 2^{-k\lambda} \frac{\varepsilon}{\sum_{k<N} 2^{-k\lambda}} + A \sum_{k\ge N} 2^{-k\lambda} = M\varepsilon + \varepsilon + A\varepsilon,
\end{aligned}
$$

i.e., $|f * k_\delta(x) - f(x)| \to 0$ as $\delta \to 0$. ■

Proposition 9.4

If $k \in C_c^m(\mathbb{R}^n)$ is m-th order differentiable, then $f * k \in C^m(\mathbb{R}^n)$, with bounded derivatives.

Proof Claim: If $k \in C_c(\mathbb{R}^n)$, then $f * k$ is continuous and bounded (HW question). It thus suffices to show $\frac{\partial}{\partial x_i}(f * k(x)) = f * \frac{\partial}{\partial x_i} k(x)$, then the desired statement follows from induction. Note that

$$
\frac{f * k(x + he_i) - f * k(x)}{h} = \int f(y) \cdot \frac{k(x + he_i - y) - k(x - y)}{h}\,dy.
$$

Since the integrand is bounded above by $|f(y)| \sup |\frac{\partial}{\partial x_i} k|$ (the supremum exists because k is continuous on a compacts set), which is independent of h. Apply D.C.T., we see that as $h \to 0$,

$$
\frac{f * k(x + he_i) - f * k(x)}{h} \to \int f(y) \frac{\partial}{\partial x_i} k(x - y)\,dy = f * \frac{\partial}{\partial x_i} k.
$$

■

Chapter 10 Hilbert Spaces

10.1 The Hilbert Space L2

Definition 10.1 ($L^2(\mathbb{R}^n)$)

$L^2(\mathbb{R}^n)$ is the collection of complex-valued measurable function in $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} |f(x)|^2\,dx < +\infty$. We define the L^2-norm of f as

$$\|f\|_{L^2} := \left(\int_{\mathbb{R}^n} |f(x)|^2\,dx\right)^{1/2}$$

♣

Remark We take $1/2$-th power on L^2-norm to preserve the linearity of the operator.

Remark

(i) Suppose $f, p \in L^2(\mathbb{R}^n)$ for which $f = p$ a.e., then $\|f - g\|_{L^2} = 0$, we may identify them as the same element in L^2.

(ii) We say $f \in L^2(E)$ if $f\chi_E \in L^2(\mathbb{R}^n)$.

(iii) For $1 \le p < +\infty$, we define ***L^p-norm*** by $\|f\|_{L^p} := (\int |f(x)|^p\,dx)^{1/p}$.

Definition 10.2 (Inner product in L^2)

We define the inner product for any $f, g \in L^2(\mathbb{R}^n)$ by $\langle f, g\rangle := \int f(x)\overline{g(x)}\,dx$.

♣

Proof The inner product is well-defined because $f\bar{g}$ is integrable:

$$\int |f\bar{g}| = \int |f||g| \le \int \frac{1}{2}(|f|^2 + |g|^2) < +\infty \tag{10.1.1}$$

where the first inequality follows from $ab \le (a^2 + b^2)/2$ for $a, b \ge 0$ by AM-GM inequality, and the second follows from the fact that $f, g \in L^2(\mathbb{R}^n)$.

Proposition 10.1

*(a) The inner product $\langle\cdot,\cdot\rangle$ in $L^2(\mathbb{R}^n)$ satisfies the **Cauchy-Schwartz inequality**: $|\langle f, g\rangle| \le \|f\|\|g\|$.*

(b) For any $g \in L^2(\mathbb{R}^n)$ fixed, $f \in L^2(\mathbb{R}^n) \mapsto \langle f, g\rangle$ is linear, and $\langle g, f\rangle = \overline{\langle f, g\rangle}$.

(c) $L^2(\mathbb{R}^n)$ is a vector space over $\mathbb{C}$, and $\|\cdot\|_{L^2}$ is a norm of L^2.

♠

Proof (a) If $\|f\| = 0$ or $\|g\| = 0$, wlog, $\|f\| = 0$, then $f = 0$ a.e., then the statement is trivial. On the other hand, suppose $\|f\| = \|g\| = 1$. Apply inequality (10.1.1), $\left|\int f\bar{g}\right| \le \int \frac{1}{2}\left(|f|^2 + |g|^2\right) = 1$. Then for f, g, consider $f/\|f\|$ and $g/\|g\|$, we see that

$$\left|\int \frac{f}{\|f\|}\frac{\bar{g}}{\|g\|}\right| \le 1 \quad\Longrightarrow\quad \left|\int f\bar{g}\right| \le \|f\|\|g\|.$$

(b) follows from the linearity of the integral.

(c) It suffices to prove the triangle inequality: for $f, g \in L^2$,

$$\begin{aligned}\|f+g\|^2 &= \langle f+g, f+g\rangle = \|f\|^2 + \langle f, g\rangle + \overline{\langle f, g\rangle} + \|g\|^2 \\ &\leq \|f\|^2 + 2\operatorname{Re}\langle f, g\rangle + \|g\|^2 \leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 \leq (\|f\| + \|g\|)^2,\end{aligned}$$

where the second last inequality in line 2 holds by Cauchy-Schwartz inequality. Taking the square root gives the desired statement. ■

Theorem 10.1

The space $L^2(\mathbb{R}^n)$ is complete with respect to the metric $d(f, g) = \|f - g\|_{L^2}$ induced by the L^2-norm. ♡

Proof Let $\{f_k\}$ be a Cauchy sequence, we want to show $\exists f \in L^2(\mathbb{R}^n)$ s.t. $d(f_k, f) = \|f_k - f\| \to 0$. Choose a subsequence $\{f_{k_i}\}_i$ of $\{f_k\}$ s.t. $k_{i+1} > k_i$ and $\|f_{k_{i+1}} - f_{k_i}\| < 2^{-i}$ for all i. Define $f(x) := f_{k_1} + \sum_{k=1}^{\infty}(f_{k_{i+1}} - f_{k_i})$ and $g(x) := |f_{k_1}| + \sum_{k=1}^{\infty}|f_{k_{i+1}} - f_{k_i}|$.

Step 1: $g \in L^2(\mathbb{R}^n)$, then $f \in L^2(\mathbb{R}^n)$.

Denote by $P_N(f)(x)$, $P_N(g)(x)$ the partial sum of $f(x)$ and $g(x)$, resp. Then

$$\|P_N(g)\| \leq \|f_{k_1}\| + \|P_N(g) - |f_{k_1}|\| \leq \|f_{k_1}\| + \sum_{k=1}^{N}\|f_{k_{i+1}} - f_{k_i}\| \leq \|f_{k_1}\| + \sum_{i=1}^{N} 2^{-i} < +\infty.$$

Apply M.C.T., $\int |g|^2 = \lim_{N\to\infty}\int |P_N(g)|^2 < +\infty$, therefore $g \in L^2(\mathbb{R}^n)$. Hence $|f| \leq g$ implies $f \in L^2(\mathbb{R}^n)$.

Step 2: $\|f_{k_i} - f\| \to 0$, i.e., $f_{k_i} \to f$ in L^2.

Since $P_N(f)(x) = f_{k_{N+1}}(x)$ by telescoping series, we see $f_{k_N}(x) = P_{k_{N-1}}(f)(x) \to f(x)$ a.e. Note that $|f - f_{k_i}| = |f - P_i(f)|^2 \leq (2g)^2$. Apply D.C.T, $\|f - f_{k_i}\| = \int |f - f_{k_i}|^2 \to 0$, namely $\|f_{k_i} - f\| \to 0$.

Step 3: $f_k \to f$ in L^2, namely f_k converges.

Given $\varepsilon > 0$. There is N s.t. $\|f_n - f_m\| < \varepsilon/2$ for $n \geq m \geq N$, and $\|f_{k_n} - f\| < \varepsilon/2$ for $n \geq N$. Then for $n \geq N$, $\|f_n - f\| \leq \|f_n - f_{k_n}\| + \|f_{k_n} - f\| < \varepsilon$. ■

Theorem 10.2

The space $L^2(\mathbb{R}^n)$ is separable, i.e., it contains a countable dense subset. ♡

Proof Let $\mathcal{C}$ be the collection of all finite linear combinations of χ_D where D is a dyadic cube in $\mathbb{R}^n$, with the coefficients being complex numbers whose real and imaginary parts are rational (i.e., $\mathbb{Q}(i)$), then $\mathcal{C}$ is countable.

It suffices to prove $\mathcal{C}$ is dense in $L^2(\mathbb{R}^n)$.

(1) Given $f \in L^2(\mathbb{R}^n)$. Let

$$g_k(x) := \begin{cases} f(x) & \text{if } |x| \leq k \text{ and } |f(x)| \leq k \\ 0 & \text{otherwise} \end{cases}$$

Then $g_k(x) \to f(x)$ a.e., and $|g_k - f|^2 \leq |f|^2$. Apply D.C.T., $\int |g_k - f|^2 \to 0$; in particular, there exists g_N s.t.

$\|g_N - f\|_{L^2} < \varepsilon$.

(2) Let $g := g_N$, then $g \in L^1(\mathbb{R}^n)$ since it is bounded and supported on a compact set. Then there exists a step function φ s.t. $\int |g - \varphi| < \varepsilon^2/2N$. Therefore, $\int |g - \varphi|^2 \leq 2N \int |g - \varphi| < \varepsilon^2$, hence $\|g - \varphi\| < \varepsilon$.

(3) Note that open sets can be decompose into dyadic cubes, there exists $\psi \in \mathcal{C}$ such that $\|\varphi - \psi\|$.

Consequently $\|f - \psi\| < 3\varepsilon$, hence $\mathcal{C}$ is dense in $L^2(\mathbb{R}^n)$. ■

10.2 Hilbert Spaces

10.2.1 Hilbert Spaces

Definition 10.3 (Hilbert space)

*A set $\mathcal{H}$ is a **Hilbert space** if it satisfies the following:*

(H1) $\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$).

(H2) $\mathcal{H}$ is equipped with an innter product $\langle f, g\rangle$ so that

- *$f \mapsto \langle f, g\rangle$ is linear on $\mathcal{H}$ for every fixed $g \in \mathcal{H}$,*
- *$\langle g, f\rangle = \overline{\langle f, g\rangle}$, and*
- *$\langle f, f\rangle \geq 0$ for all $f \in \mathcal{H}$, with equality hold iff $f = 0$.*

We let $\|f\| = \langle f, f\rangle^{1/2}$.

(H3) $\mathcal{H}$ is complete with respect to the metric $d(f, g) = \|f, g\|$.

**(H4) $\mathcal{H}$ is separable.* ♣

Remark Cauchy-Schwarz inequality and the triangle inequality follows from (H1) and (H2).

Example 10.1 $(L^2(\mathbb{R}^n), \langle\cdot,\cdot\rangle)$ is a Hilbert space over $\mathbb{C}$.

Example 10.2 Finite dimensional vector space $\mathbb{C}^N = \{(z_1, \cdots, z_N) \,|\, z_i \in \mathbb{C}\}$, equipped the inner product $\langle z, w\rangle = \sum_{i=1}^N z_i \overline{w_i}$, is a Hilbert space over $\mathbb{C}$.
$\mathbb{R}^N$ with the standard Euclidean inner product is a Hilbert space over $\mathbb{R}$.

Example 10.3 Denote $l^2 := l_2(\mathbb{N}) = \{(x_1, x_2, \cdots) \,|\, x_i \in \mathbb{C}, \sum_i |x_i|^2 < +\infty\}$. Define the inner product $\langle x, y\rangle = \sum_i x_i \overline{y_i}$. Then $(l^2, \langle\cdot,\cdot\rangle)$ is a Hilbert space over $\mathbb{C}$.

Example 10.4 (Supplementary example) Denote $W^{1,2}(\mathbb{R}^n) := \{f \in L^2(\mathbb{R}^n) \,|\, |\nabla f| \in L^2(\mathbb{R}^n)\}$. Define the inner product $\langle f, g\rangle := \langle f, g\rangle_{L^2} + \sum_{i=1}^n \langle \partial_i f, \partial_i g\rangle_{L^2}$. Then $(W^{1,2}, \langle\cdot,\cdot\rangle)$ is a Hilbert space over $\mathbb{C}$.

10.2.2 Orthogonality

Remark A ***Banach space*** is a normed vector space + (H3), thus all Hilbert spaces are Banach. The advantage of Hilbert space is that it equips an inner product, containing the notion of orthogonality.

Definition 10.4 (Orthogonality)

*Two elements f and g in a Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$ are **orthogonal** or perpendicular if $\langle f, g\rangle = 0$, and we then write $f \perp g$.* ♣

Remark The ***Pythagorean Theorem*** holds: if $f \perp g$, then $\|f+g\|^2 = \|f\|^2 + \|g\|^2$.

Definition 10.5 (Orthonormal)

*A collection $\{e_\alpha\}_{\alpha \in A}$ in $\mathcal{H}$ is **orthonormal** if $\langle e_\alpha, e_\beta \rangle = 1$ if $\alpha = \beta$ and otherwise $\langle e_\alpha, e_\beta \rangle = 0$.* ♣

Remark Any orthonormal collection in $\mathcal{H}$ is at most countable, since $\mathcal{H}$ has a countable dense subset. Therefore, we may use $\mathbb{N}$ as the index set A.

Proposition 10.2

If $\{e_k\}$ is orthonormal in $\mathcal{H}$, and $f = \sum_{k=1}^N a_k e_k \in \mathcal{H}$, then $\|f\|^2 = \sum_{k=1}^N |a_k|^2$. ♠

Proof The proof follows from Pythagorean theorem and the fatc that $a_k = \langle f, e_k \rangle$. ■

Definition 10.6 (Orthonormal basis)

*We say an orthonormal collection $\{e_k\}$ of $\mathcal{H}$ is an **orthonormal basis** if the finite linear combination of e_k's over $\mathbb{C}$ is dense in $\mathcal{H}$.* ♣

Theorem 10.3

Let $\{e_k\}$ be a orthonormal collection $\{e_k\}$ in $\mathcal{H}$, then the following are equivalent:

(a) *Finite linear combinations of elements in $\{e_k\}$ are dense in $\mathcal{H}$ (i.e., $\{e_k\}$ is a orthonormal basis).*

(b) *If $f \in \mathcal{H}$ and $\langle f, e_j \rangle = 0$ for all j, then $f = 0$.*

(c) *If $f \in \mathcal{H}$, and $S_N(f) := \sum_{k=1}^N a_k e_k$, where $a_k = \langle f, e_k \rangle$, then $S_N(f) \to f$ in the norm as $N \to \infty$; i.e., $\sum_{k=1}^N \langle f, e_k \rangle e_k \to f$.*

(d) *If $a_k = \langle f, e_k \rangle$, then $\|f\|^2 = \sum_k |a_k|^2$. (Parseval's identity)* ♡

Proof $(a) \Rightarrow (b)$: Let $\varepsilon > 0$ be given. Suppose $f \perp e_j$ for all j. By (a), there exists $\{a_k\}_{k=1}^n$ s.t. $\|f - \sum_{k=1}^N a_k e_k\| < \varepsilon$. Then

$$\|f\|^2 = \langle f, f \rangle = \left\langle f - \sum_{k=1}^n a_k e_k, f \right\rangle + \cancelto{0}{\left\langle \sum_{k=1}^n a_k e_k, f \right\rangle} \le \left\| f - \sum_{k=1}^n a_k e_k \right\| \|f\| \le \varepsilon \|f\|,$$

so either $\|f\| = 0$ or $\|f\| < \varepsilon$. Hence $\|f\| = 0$ since the choice of ε is arbitrary, so $f = 0$.

$(b) \Rightarrow (c)$: For any k, by orthonormal condition, $\langle S_N(f), e_k \rangle = a_k = \langle f, e_k \rangle$. Then $f - S_N(f) \perp e_k$ for each k, followed by $f - S_N(f) \perp S_N(f)$. By Pythagorean theorem,

$$\|f\|^2 = \|f - S_N(f)\|^2 + \|S_N(f)\|^2 = \|f - S_N(f)\|^2 + \sum_{k=1}^N |a_k|^2. \tag{10.2.1}$$

To prove $S_N(f)$ converges, it suffices to prove it is Cauchy. For N, M, $\|S_N(f) - S_M(f)\|^2 = \sum_{M<k\le N} |a_k|^2$. By

(10.2.1), we have ***Bessel's inequality***

$$\sum_{k=1}^{\infty} |a_k|^2 = \lim_{N\to\infty} \sum_{k=1}^{N} |a_k|^2 \le \|f\|^2. \tag{10.2.2}$$

Then we see $\|S_N(f) - S_M(f)\|^2 = \sum_{M<k\le N} |a_k|^2$ can be arbitrarily small for sufficiently large M, N, thus $S_N(f)$ is Cauchy thus convergence in $\mathcal{H}$.

Lastly, we prove $S_N(f) \to f$. Let k be fixed, $\langle f - S_N(f), e_k\rangle = a_k - \langle S_N, e_k\rangle = 0$ for $N > k$, then the hypothesis (b) implies that $f - S_N(f) \to 0$ in $\mathcal{H}$, as desired.

(*c*) ⇒ (*d*): By (c), $S_N(f) \to f$ in $\mathcal{H}$. Apply (10.2.1) as $N \to \infty$ yields $\sum_{k=1}^{N} |a_k|^2 \to \|f\|$ as $N \to \infty$.

(*d*) ⇒ (*a*): Let $\varphi := \sum_{k=1}^{N} \langle f, e_k\rangle e_k \in \mathcal{H}$, let $\varepsilon > 0$. Apply (10.2.1), then $\|f - \varphi\|^2 = \|f\|^2 - \sum_{k=1}^{N} |\langle f, e_k\rangle|^2 < \varepsilon^2$ by the hypothesis, by choosing sufficiently large N. ■

Theorem 10.4

Any Hilbert space has an orthonormal basis. ♡

Proof $\mathcal{H}$ has a countable dense subset $\mathcal{F} = \{f_k\}$ by definition. We may assume $\mathcal{F}$ is linearly independent by removing elements that are linearly dependent with previous terms. We then apply the ***Gram-Schmidt*** algorithm: Let $e_1 = f_1/\|f_1\|$. For each $k > 1$, define recursively

$$e_k = \frac{f_k - \sum_{j<k} \langle f_k, e_j\rangle e_j}{\|f_k - \sum_{j<k} \langle f_k, e_j\rangle e_j\|}.$$

Then $e_1, \cdots, e_N$ are orthonormal for all N, because $\|e_k\| = 1$ and $\langle e_k, e_l\rangle = \mathfrak{c}\langle f_k - \sum_{j<k} \langle f_k, e_j\rangle e_j, e_l\rangle = \langle f_k, e_l\rangle - \langle \langle f_k, e_l\rangle e_l, e_l\rangle = 0$ for $k > l$ (where $\mathfrak{c}$ denotes a constant).

It suffices to show $e_1, \cdots, e_N$ has the same span as $f_1, \cdots, f_N$. Note that $e_N = \mathfrak{c} \cdot e_N = \mathfrak{c}(f_N - \sum \lambda_i e_i)$, then f_N may be written as the linear combination of $e_1, \cdots, e_N$. By induction, we see that the span remains the same. ■

Remark We say $\mathcal{H}$ is finite-dimensional if there exists a finite orthonormal basis, i.e., the Gram-Schmidt algorithm terminates.

Definition 10.7 (Unitary isomorphism)

*Let $\mathcal{H}$ and $\mathcal{H}'$ be Hilbert spaces, we say a mapping $T : \mathcal{H} \to \mathcal{H}'$ is a **unitary isomorphism** if*

(i) T is a linear map, i.e., $T(\alpha f + \beta g) = \alpha T(f) + \beta T(g)$,

(ii) T is a bijection, and

(iii) $\|Tf\|_{\mathcal{H}'} = \|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$. ♣

Remark The condition (iii) implies that inner products are preserved under T, namely $\langle Tf, Tg\rangle_{\mathcal{H}'} = \langle f, g\rangle_{\mathcal{H}}$.

Proposition 10.3

Any two infinite dimensional Hilbert spaces are unitary equivalent.

Proof Let $\{e_1, \cdots\}$ and $\{e'_1, \cdots\}$ denote the orthonormal basis of $\mathcal{H}_1$ and $\mathcal{H}_2$, resp. Define $T : \mathcal{H}_1 \to \mathcal{H}_2$ by $T(e_i) = e'_i$. Suppose $f \in \mathcal{H}$, we can identify it with $\sum_{i \in \mathbb{N}} a_i e_i$ where $a_i = \langle f, e_i \rangle_{\mathcal{H}_1}$ by Theorem 10.3 (c), then by the definition $T(f) = \sum_i a_i e'_i$.

It suffices to show T is a unitary isomorphism. (1) T is linear. (2) T is bijective follows from that $T^{-1} : e'_i \mapsto e_i$ is the inverse of T. (3) For all $f = \sum_i a_i e_i \in \mathcal{H}$, since $\|T(f)\| = \|\sum_i a_i e_i\| = \sum_i |a_i|^2 < +\infty$, then $T(f)$ is well-defined (as a convergent series) and $\|T(f)\| = \|f\|$ by Parseval identity. ■

Remark Any infinite dimensional Hilbert space is equivalent to l^2.

10.3 Fourier Series

Consider $L^2([-\pi,\pi])$ with inner product $\langle f,g\rangle := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$.

Theorem 10.5

$\{e^{ikx}\}_{k\in\mathbb{Z}}$ is an orthonormal basis for $L^2([-\pi,\pi])$.

Remark Suppose $k,j\in\mathbb{Z}$, then

$$\langle e^{ikx}, e^{ijx}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ikx}\overline{e^{ijx}} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-j)x} = \begin{cases} 0 & \text{if } k\neq j \\ 1 & \text{if } k=j \end{cases}$$

where the last equality follows from the periodicity of sin and cos. Therefore, we see that $\{e^{ikx}\}$ is orthonormal.

Example 10.5 Suppose f is Riemann integrable or piecewise continuous function on $[-\pi,\pi]$. Define $a_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\,dx$ and $b_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\,dx$. Then we may write f as

$$f = \sum_{k\in\mathbb{N}} a_k\sin(kx) + b_k\cos(kx).$$

Here the basis is $\{\sin(kx),\cos(kx)\}_{k\in\mathbb{N}}$.

Remark 1: $e^{-ikx} = \cos kx - i\sin kx$ for $k\in\mathbb{N}$, we may establish that the two bases are approximately identical: $\{e^{-ikx}, e^{ikx}\}_{k\in\mathbb{N}} \approx \{\sin kx, \cos kx\}_{k\in\mathbb{N}}$.

Remark 2: If f is Riemann integrable or piecewise continuous function (they are pre-Hilbert space), then $f\in L^2$.

Let $f\in L^2([-\pi,\pi])$, we extend f to be defined on $\mathbb{R}$, and we define the **Fourier coefficient** to be

$$a_k := \langle f, e^{-kx}\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{e^{ikx}}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}\,dx.$$

Proposition 10.4

(a) If $a_k = 0$ for all $k\in\mathbb{Z}$, then $f(x)=0$ for a.e. x.

**(b) $\sum_{k\in\mathbb{Z}} a_k r^{|k|}e^{-ikx} \to f(x)$ for a.e. x as $r\to 1$.*

Proof (b) is beyond the scope of this class. Note that $\sum_{k\geq 0} z^k = 1/(1-z)$ from *Harmonic Analysis*, then letting $z = re^{ix}$ yields

$$\begin{aligned}\sum_{k\in\mathbb{Z}} r^{|k|}e^{ikx} &= \sum_{k\geq 0} r^k e^{ikx} + \sum_{k<0} r^{-k}e^{ikx} = \sum_{k\geq 0}(re^{ix})^k + \sum_{k<0}(re^{-ix})^{-k} \\ &= \frac{1}{1-z} + \frac{\bar{z}}{1-\bar{z}} = \frac{1-z\bar{z}}{(1-z)(1-\bar{z})} = \frac{1-r^2}{1-2r\cos x + r^2}.\end{aligned}$$

We claim that $P_r(x) = \frac{1}{2\pi}\frac{1-r^2}{1-2r\cos x+r^2}$ is a good kernel with $\delta = 1-r$. The proof is omitted.

(a) follows immediately from (b). ■

Remark It follows from this proposition that $\{e^{ikx}\}_k$ forms an orthonormal basis of $L^2([-\pi, \pi])$, proving Theorem 10.5.

Corollary 10.1

For any $f \in L^2([-\pi, \pi])$,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 = \sum_{k \in \mathbb{Z}} |a_k|^2.$$

Moreover, the Fourier series of f *converges to* f *in* L^2 *norm, i.e.,* $S_N(f)(x) := \sum_{|k| \leq N} a_k e^{ikx}$ *in* L^2. ♡

Chapter 11 Abstract Measure Theory

11.1 Abstract Measure

Definition 11.1 (Measure space)

*A **measure space** on a set X is a triple $(X, \mathcal{M}, \mu)$ where*

*(1) $\mathcal{M}$ is a **σ-algebra**, which is a (i) non-empty collection of subsets of X closed under (ii) complements and (iii) countable unions (thus countable intersections). We refer to elements in $\mathcal{M}$ as **measurable sets**.*

(2) $\mu : \mathcal{M} \to [0, \infty]$ is a function satisfying countable additivity: for any countable collection of disjoint sets in $\mathcal{M}$, $E_1, E_2, \cdots$,

$$\mu\left(\bigsqcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

*We refer to $\mu(E)$ as the **measure** of E.* ♣

Example 11.1 Consider Lebesgue measure, $X = \mathbb{R}^n$, $\mathcal{M} =$ collection of Lebesgue measurable sets, and $\mu = m$ is the Lebesgue measure.

Example 11.2 Let $X = \{x_k\}_k$, $\mathcal{M}$ be all subsets of X, and define $\mu(\{x_k\}) = \mu_k$ where $\{\mu_k\}$ is a sequence of numbers in $[0, \infty]$. Then for any $E \in \mathcal{M}$, $\mu(E) = \sum_{x_k \in E} \mu_k$ (intuition: $\mu(E)$ is the weighted sum of the entries in E). The triple $(X, \mathcal{M}, \mu)$ is measure space. In particular, if $\mu_k \equiv 1$, then μ is the counting measure.

Example 11.3 Let $X = \mathbb{R}^n$, $\mathcal{M}$ be all Lebesgue measurable sets in $\mathbb{R}^n$, and for any $E \in \mathcal{M}$, define $\mu(E) = \int_E f$ where f is a nonnegative measurable function on $\mathbb{R}^n$. The triple $(X, \mathcal{M}, \mu)$ is measure space. In particular, if $f \equiv 1$, the measure corresponds the Lebesgue measure.

Remark Lebesgue-Radon–Nikodym theorem implies that any measure on $\mathbb{R}^n$ must be a combination of measurable spaces in example 11.2 and 11.3.

More precisely, let μ be a measure on $\mathbb{R}^n$. Then $\mu = \mu_{as} + \mu_s$, where $\mu_{as}(E) = \int_E f\, dx$ where f is some non-negative integrable function, and μ_s is singular w.r.t. m (i.e., μ_s and m are supported on disjoint sets of $\mathbb{R}^n$).

Definition 11.2 (Outer measure)

*An **outer measure** on a set X is a function μ^* from all subsets of X to $[0, +\infty]$ satisfies*

(1) $\mu_*(\varnothing) = 0$

(2) Monotonicity: If $E_1 \subseteq E_2$, then $\mu_*(E_1) \leq \mu_*(E_2)$.

(3) Countable subadditivity: For any countable collection of sets $E_1, E_2, \cdots$ in X, we have $\mu_*(\bigcup E_k) \leq$

$\sum \mu_*(E_k)$.

Definition 11.3 (Caratheodory measurable sets)

*A set $E \subseteq X$ is **(Caratheodory) measurable** if for any $A \subseteq X$,*

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

Remark

(i) By the countable subadditivity, the condition can be reduced to one direction $\mu_*(A) \geq \mu_(A \cap E) + \mu_*(A \cap E^c)$.

(ii) The definition of Lebesgue measurable set is equivalent to Caratheodory criterion in Lebesgue measure space.

Theorem 11.1

Given an outer measure μ_* on a set X, the collection $\mathcal{M}$ of all measurable sets form a σ-algebra. Moreover, μ_* restricted to $\mathcal{M}$ is a measure.

Proof (i) $\varnothing \in \mathcal{M}$ because $\mu_*(E \cap \varnothing) + \mu_*(E \cap \varnothing^c) = \mu_*(\varnothing) + \mu(E) = \mu(E)$ for all E. (ii) $\mathcal{M}$ is closed under complement because the criterion is symmetry to complement.

(iii) *Claim*: $\mathcal{M}$ is closed under finite unions and is finite additive.

Proof: Let $E_1, E_2 \in \mathcal{M}$, then

$$\begin{aligned}\mu_*(A) = \mu_*(E_1 \cap A) + \mu_*(E_1^c \cap A) &= \mu_*(E_1 \cap A) + \mu_*(E_1^c \cap E_2 \cap A) + \mu_*(E_1^c \cap E_2^c \cap A) \\ &\geq \mu_*((E_1 \cup E_2) \cap E) + \mu_*(E_1^c \cap E_2^c \cap A),\end{aligned}$$

where the inequality follows from the countable subadditivity; it follows that $E_1 \cup E_2$ is measurable. If E_1 and E_2 are disjoint, then

$$\mu^*(E_1 \sqcup E_2) = \mu^*((E_1 \sqcup E_2) \cap E_1) + \mu_*((E_1 \sqcup E_2) \cap E_1^c) = \mu_*(E_1) + \mu_*(E_2),$$

followed by the finite additivity. *

Claim: $\mathcal{M}$ is closed under countable union and is countable additive.

Proof: Suppose $G = \bigcup E_k$ where $E_k \in \mathcal{M}$, we may assume E_k's are disjoint WLOG. For any N, $\mu_*(A) \geq \sum_{k=1}^{N} \mu_*(A \cap E_k) + \mu_*(A \cap G^c)$ by Caratheodory criterion. Let $N \to +\infty$, then

$$\mu_*(A) \geq \sum_{k=1}^{\infty} \mu_*(A \cap E_k) + \mu_*(A \cap G^c) \geq \mu_*(A \cap G) + \mu_*(A \cap G^c) \tag{11.1.1}$$

where the second inequality follows from countable subadditivity. It follows that (11.1.1) becomes an equality, and $G \in \mathcal{M}$. Take $A = G$ in (11.1.1), we see that $\mu_*(G) = \sum_{k=1}^{\infty} \mu(E_k) + 0$, followed by the countable additivity property. ■

Definition 11.4 (σ-finite)

*Ww say a measure space $(X, \mathcal{M}, \mu)$ is **σ-finite** if X can be written as the union of countable many measurable sets of finite measure.* ♣

11.2 Metric Outer Measure

Definition 11.5 (Borel σ-algebra)

*The **Borel σ-algebra** $\mathcal{B}_X$ is the smallest σ-algebra containing all open sets.* ♣

Definition 11.6 (Metric outer measure)

*We say an outer measure μ_* on (X, d) is a **metric outer measure** if $\mu_*(A \cup B) = \mu_*(A) + \mu_*(B)$ for any $A, B \subseteq X$ such that $d(A, B) := \inf\{d(x, y) \mid x \in A, y \in B\} > 0$.* ♣

Theorem 11.2

If μ_* is a metric outer measure on (X, d), then Borel sets in X are Caratheodory measurable and μ_* restricted to $\mathcal{B}_x$ is a measure. ♡

Proof By Theorem 11.1, $\mathcal{M}$ is a σ-algebra and $\mu_*|_{\mathcal{M}}$ is a measure. To show $\mathcal{B}_X \subseteq \mathcal{M}$, it suffices to show all open/closed sets are measurable.

Let F be a closed set in X and $A \subseteq X$. We may assume $\mu_*(A) < +\infty$, otherwise the statement is trivial. Define $E_k := \{x \in A \cap F^c \mid d(x, F) > 1/k\}$. Then by monotonicity and the definition of metric outer measure, $\mu_*(A) \geq \mu_*((A \cap F) \cup E_k) = \mu_*(A \cap F) + \mu_*(E_k)$.

It suffices to show $\lim_{k\to\infty} \mu_*(E_k) = \mu_*(A \cap F^c)$, and $\leq$ direction is trivial. Let $C_k = E_{k+1} \setminus E_k$. Note that $d(E_k, C_{k+1}) > 0$, then

$$\mu_*(E_{k+2}) \geq \mu_*(C_{k+1} \cup E_k) = \mu_*(C_{k+1}) + \mu_*(E_k).$$

Inductively, $\mu_*(E_{2k}) \geq \sum_{i=1}^{k} \mu_*(C_{2i-1})$ and $\mu_*(E_{2k+1}) \geq \sum_{i=1}^{k} \mu_*(C_{2i})$; it follows that the series $\sum_{i=1}^{\infty} \mu_*(C_i)$ is bounded thus convergent. Notice that

$$\mu_*(E_k) \leq \mu_*(A \cap F^c) \leq \mu_*(E_k) + \sum_{i=k}^{\infty} \mu_*(C_i),$$

and the $\sum_{i=k}^{\infty} \mu_*(C_i)$ can be arbitrarily small as $k \to \infty$. Hence $\mu_*(A \cap F^c) = \lim_{k\to\infty} \mu_*(E_k)$. ■

Proposition 11.1

Suppose the Borel measure μ is finite on all balls in X with finite radii. Then for any Borel set E, any $\varepsilon > 0$, there exists an open set $G \supseteq E$, closed set $F \subseteq E$ such that $\mu(G \setminus E) < \varepsilon$ and $\mu(E \setminus F) < \varepsilon$. ♠

Lemma: Let $(X, \mathcal{M}, \mu)$ be a measure space. If measurable sets $E_k \nearrow E$, then $\mu(E_k) \nearrow \mu(E)$.

Proof Let $\mathcal{F}$ be the collection of Borel sets satisfying these properties, $\mathcal{F}$ is nonempty because $\varnothing \in \mathcal{F}$.

(1) We first show $\mathcal{F}$ is a σ-algebra. (i) $\varnothing \in \mathcal{F}$. (ii) If $E \in \mathcal{F}$ then $E^c \in \mathcal{F}$. (iii) Suppose $E_k \in \mathcal{F}$, and let $E = \bigcup E_k$. Let $G = \bigcup G_k$ where $G_k \supseteq E_k$ is the open set s.t. $\mu(G_k \backslash E_k) < 2^{-k}\varepsilon$. Then $\mu(G \backslash E) \leq \mu(\bigcup_k (G_k \backslash E_k)) \leq \sum \mu(G_k \backslash E_k) < \varepsilon$. On the other hand, we may choose $F = \bigcup F_k$ where $F_k \subseteq E_k$ is closed and $\mu(E_k \setminus F_k) < 2^{-k}\varepsilon$, then $\mu(E \backslash F) < \varepsilon$. By the continuity from below, $F_N := \bigcup_{k=1}^N F_k \nearrow F$ implies $\mu(F_N) \nearrow \mu(F)$; then choosing a sufficiently large N yields a closed set F_N s.t. $\mu(E \backslash F_N) \leq \mu(E \backslash F) + \mu(F \backslash F_N) < 2\varepsilon$. Hence $E = \bigcup E_k \in \mathcal{F}$.

(2) It suffices to show $\mathcal{F}$ contains all open sets. It is clear that open sets can be approximated by themselves, so it suffices to show that an open set G can be approximated by a closed set $F \subseteq G$ s.t. $\mu(G \backslash F) < \varepsilon$. Let $F_k := \{x \mid d(x, G^c) \geq 1/k\}$ be a closed set, and put $G = \bigcup E_k$. Then $F_k \nearrow G$, and thus $\mu(F_k) \nearrow \mu(G)$; taking sufficiently large k yields E_k s.t. $\mu(G \backslash F) < \varepsilon$. ■

11.3 Premeasure and The Extension Theorem

Definition 11.7 (Algebra)

*Given a set X, an **algebra** in X is a non-empty collection of subsets of X which are (i) closed under complement and (ii) closed under finite unions (and thus intersections).* ♣

Definition 11.8 (Premeasure)

*A **premeasure** on an algebra $\mathcal{A}$ is a function $\mu_0 : \mathcal{A} \to [0, +\infty]$ such that*

(1) $\mu_0(\varnothing) = 0$,

(2) If $A_1, A_2, \cdots$ is a countable collection of disjoint sets in $\mathcal{A}$ such that $\bigsqcup_{k=1}^{\infty} A_k \in \mathcal{A}$, then $\mu_0(\bigsqcup_k A_k) = \sum_k \mu_0(A_k)$. ♣

Remark The monotonicity follows immediately from (2).

Example 11.4 Consider the Lebesgue premeasure: let $X = \mathbb{R}^n$, μ_0 defined rectangles is their the volume, and $\mathcal{A}$ is the algebra generated by rectangles. The definition above give rise to a premeasure because (1) is obvious and (2) follows from Proposition 7.5.

By the premeasure, it give rises to the Lebesgue outer measure $\mu_* = m_*$ (we are going to justify the extension in the following lemma) defined on all subsets of $\mathbb{R}^n$ and satisfies the countable sub-additivity. Then we can extend it to the Lebesgue measure $\mu = m$ using Caratheodory criterion.

Lemma 11.1

If μ_0 is a premeasure on an algebra $\mathcal{A}$, define an outer measure μ_* on any subset E of X by

$$\mu_*(E) = \inf \left\{ \sum_{k=1}^{\infty} \mu_0(E_k) \;\middle|\; E \subseteq \bigcup_{k=1}^{\infty} E_k, \text{ where } E_j \in \mathcal{A} \text{ for all } j \right\}.$$

Then μ_* satisfies:

(a) μ_* is an outer measure on X.

(b) $\mu_*(A) = \mu_0(A)$ for all $A \in \mathcal{A}$.

(b) Any set in $\mathcal{A}$ is Caratheodory measurable w.r.t. μ_*. ♡

Proof (a) It is clear that $\mu_*(\varnothing) = 0$, and the monotonicity and countable subadditivity follow from the corresponding conditions in premeasure.

(b) By the definition of μ_*, $\mu_*(A) \leq \mu_0(A)$. For all $A_j \in \mathcal{A}$ s.t. $A \subseteq \bigcup A_j$, we map assume $\{A_j\}$ are pairwise disjoint, then $A = \bigsqcup_j (A_j \cap A)$. By the definition of μ_0 and monotonicity, we see that $\mu_0(A) = \sum_j \mu_0 * (A \cap A_j) \leq \sum_j \mu_0(A_j)$. Therefore, $\mu_0(A) \leq \mu_*(A)$.

(c) Let $A \in \mathcal{A}$ and $B \subseteq X$. Let $A_j \in \mathcal{A}$ s.t. $B \subseteq \bigcup A_j$. Then $B \cap A \subseteq \bigcup (A_j \cap A)$, followed by

$$\sum_j \mu_0(A_j) = \sum_j \mu_0(A_j \cap A) + \sum_j \mu_0(A_j \setminus A) \geq \mu_*(B \cap A) + \mu_*(B \backslash A).$$

Therefore, $\mu_*(A) \geq \mu_(B \cap A) + \mu_*(B \backslash A)$. ■

Remark The above extension is **unique**: Let $\mathcal{M}$ be a σ-algebra containing $\mathcal{A}$, let μ be the measure generated from μ_*. Assume that μ is σ-finite, then for any other measure ν defined on $\mathcal{M}$ s.t. $\nu|_{\mathcal{A}} = \mu_0$, the two measure are identical, i.e., $\nu(E) = \mu(E)$ for any $E \in \mathcal{M}$.

Example 11.5 Let $(X_1, \mathcal{M}_1, \mu_1)$, $(X_2, \mathcal{M}_2, \mu_2)$ be two σ-finite measure space. Construct a measure space on $X := X_1 \times X_2$. Define the premeasure μ_0 as: for sets of the form $A \times B$ ("measurable rectangles") where $A \in \mathcal{M}_1, B \in \mathcal{M}_2$, we defined $\mu_0(A \times B) = \mu_1(A) \cdot \mu_2(B)$. Let $\mathcal{A}$ be the algebra generated by measurable rectangles.

We may extend this premeasure into a measure of the product space $X = X_1 \times X_2$.