

# Analysis (I)

by Poshen Chang

v.1

2025-12-11

# Contents

<b>1 Topological Spaces</b>	<b>4</b>
1.1 Metric Spaces . . . . .	4
1.2 Continuity . . . . .	5
1.3 The Topology of a Metric Space . . . . .	6
1.4 Hausdorff Space . . . . .	9
1.5 Subspaces and Product Spaces . . . . .	10
1.6 Completion . . . . .	13
1.7 Connectedness . . . . .	17
1.8 Other Metric Space Concepts . . . . .	19
<b>2 Compactness</b>	<b>21</b>
2.1 Continuity and Compactness . . . . .	23
2.2 Subspaces, Products, and Hausdorff Spaces . . . . .	24
2.3 Finite Intersection Property . . . . .	25
2.4 Tychonoff Theorem . . . . .	26
2.5 Cantor Sets . . . . .	28
<b>3 Integration Theory</b>	<b>32</b>
3.1 Riemann Integrals on $\mathbb{R}$ . . . . .	32
3.2 Fundamental Theorem of Calculus . . . . .	41
3.3 Multiple Integrals . . . . .	42

---

3.4	Change of Variables . . . . .	45
<b>4</b>	<b>Function Spaces</b>	<b>50</b>
4.1	Uniform Convergence and Completeness . . . . .	50
4.2	Compactness and Equicontinuity in $C^0$ . . . . .	54
4.3	Uniform Approximation in $C^0$ . . . . .	59
4.4	Banach Fixed Point Theorem . . . . .	63
4.5	Analytic Functions . . . . .	66
4.6	Nowhere Differentiable Continuous Functions . . . . .	69
4.7	Function Approximation and Artificial Neural Networks . . . . .	72
<b>5</b>	<b>Fourier Analysis</b>	<b>75</b>
5.1	Uniqueness of Fourier Series . . . . .	76
5.2	Convolution . . . . .	78
5.3	Summability and Convergence . . . . .	82
5.4	$L^2$ Convergence . . . . .	87
5.5	Continuous function with divergent Fourier series . . . . .	91
5.6	Applications . . . . .	93
5.7	Heat equation on the circle . . . . .	102
<b>6</b>	<b>Fourier Transform</b>	<b>107</b>
6.1	Inverse Fourier Transform . . . . .	109
6.2	Convolutions . . . . .	112
6.3	Partial Differential Equations . . . . .	114
6.4	Laplace equation in the upper half plane . . . . .	116
6.5	Poisson summation formula . . . . .	120
6.6	Heisenberg uncertainty principle . . . . .	123
6.7	Generalizations of Fourier transform . . . . .	125

---

A.1	Separation Axioms . . . . .	127
A.2	Euclid's Theorem . . . . .	128
A.3	Quotient Topology . . . . .	129
A.4	Paracompactness . . . . .	130
A.5	More on Cantor Set . . . . .	132
A.6	Local Compactness and Topological Groups . . . . .	134
A.7	Oct 14, 2025 . . . . .	136
A.8	Applications of Arzela-Ascoli Theorem . . . . .	138
A.9	Nowhere Analytic Smooth Functions . . . . .	140
A.10	Riemann-Lebesgue Lemma . . . . .	140

# CHAPTER 1

## Topological Spaces

In the language of calculus, we defined convergence, continuity and such concepts on the real numbers  $\mathbb{R}$ . We would like to generalize this concept.

### 1.1 Metric Spaces

**Definition 1.1.1** A metric space is a nonempty set  $M$ , together with a metric  $d : M \times M \rightarrow \mathbb{R}$ .  $d(x, y)$  is a real number defined for all  $x, y \in M$ , which can be thought of as the distance between  $x$  and  $y$ .  $d$  satisfies the following properties:

- a) (Positive definiteness)  $d(x, y) \geq 0$ , and  $d(x, y) = 0$  iff  $x = y$ .
- b) (Symmetry)  $d(x, y) = d(y, x)$ .
- c) (Triangle inequality)  $d(x, z) \leq d(x, y) + d(y, z)$ .

We say that the pair  $(M, d)$  is a metric space. The metric  $d$  can be omitted if it is clear from the context.

**Example 1.1.1** The following are some examples of metric spaces:

- $(\mathbb{R}, d)$ , where  $d(x, y) = |x - y|$ .
- $(\mathbb{R}^n, d)$ , where

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

- $(M, d)$ , where  $M \neq \emptyset$ , and

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

This metric is called the **discrete metric** on  $M$ .

---

If  $\emptyset \neq A \subseteq M$ , and if  $M$  is a metric, then  $(A, d)$  is also a metric space. We call  $A$  a metric subspace of  $M$ , or  $A$  inherits its metric from  $M$ .

**Definition 1.1.2** We say that a sequence  $(x_n)$  in  $M$  converges to the limit  $x$  in  $M$  if for any given  $\epsilon > 0$ ,  $\exists N \in \mathbb{N}$  such that  $n \geq N$  and  $n \in \mathbb{N}$  implies  $d(x_n, x) < \epsilon$ .

It's easy to check that limits are unique. Also, every subsequence of a convergent sequence converges, and it converges to the same limit as the original sequence.

## 1.2 Continuity

**Definition 1.2.1** Let  $(M, d_M)$  and  $(N, d_N)$  be metric spaces. We say that  $f : M \rightarrow N$  is continuous if it preserves sequential convergence, i.e. for each  $(x_n)$  in  $M$  which converges to  $x$  in  $M$ , the image sequence  $(f(x_n))$  converges to  $f(x)$ .

**Proposition 1.2.1** Composition of continuous functions is continuous

*Proof.* Let  $M, N, P$  be metric spaces, and let  $f : M \rightarrow N$ ,  $g : N \rightarrow P$  be continuous functions.

Let  $(x_n)$  be a convergent sequence in  $M$  with limit  $x$ . We have

$$\lim_{n \rightarrow \infty} x_n = x \xrightarrow{f \text{ conti.}} \lim_{n \rightarrow \infty} f(x_n) = f(x) \xrightarrow{g \text{ conti.}} \lim_{n \rightarrow \infty} g(f(x_n)) = g(f(x)),$$

thus  $g \circ f$  is continuous. □

**Example 1.2.1** The following are some examples of continuous functions:

- The identity map  $\text{id} : M \rightarrow M$  is continuous.
- Every constant function  $f : M \rightarrow N$  is continuous.
- Every function  $f : M \rightarrow N$  is continuous if  $M$  is equipped with the discrete metric.

**Definition 1.2.2** If  $f : M \rightarrow N$  is a bijection such that  $f$  and  $f^{-1} : N \rightarrow M$  are continuous, then we say that  $f$  is a homeomorphism. If there exists a homeomorphism between  $M$  and  $N$ , we say that  $M, N$  are homeomorphic, denoted by  $M \cong N$ .

Intuitively, a homeomorphism is a bijection that can bend, twist, stretch the space  $M$  to make it coincide with  $N$ , but it cannot rip, puncture or shred  $M$  etc.

**Example 1.2.2** Let  $\mathbb{S}^1$  be the unit circle in the plane. Consider the interval  $[0, 2\pi)$ . Define  $f : [0, 2\pi) \rightarrow \mathbb{S}^1$  to be the function  $f(\theta) = (\cos \theta, \sin \theta)$ .  $f$  is continuous and bijective, but  $f^{-1}$  is not continuous (consider a sequence in  $\mathbb{S}^1$  approaching  $(1, 0)$  from the lower plane).

---

**Proposition 1.2.2**  $f : M \rightarrow N$  is continuous iff it satisfies the following:  $\forall \epsilon > 0$  and  $x \in M$ ,  $\exists \delta > 0$  such that if  $y \in M$  and  $d_M(x, y) < \delta$  then  $d_N(f(x), f(y)) < \epsilon$ .

*Proof.* " $\Rightarrow$ ": Suppose that  $f$  fails to satisfy the  $\epsilon$ - $\delta$  condition at some  $x \in M$ .  $\exists \epsilon > 0$  such that  $\forall \delta > 0$ ,  $\exists y \in M$  such that  $d_M(x, y) < \delta$  but  $d_N(f(x), f(y)) \geq \epsilon$ . Take  $\delta = \frac{1}{n}$ . By our assumption, we can obtain a sequence  $(y_n)$  with  $d_M(x, y_n) < \frac{1}{n}$  but  $d_N(f(x), f(y_n)) \geq \epsilon$ , so  $(y_n)$  converges to  $x$  but  $f(y_n)$  does not approach  $f(x)$ , contradicting the continuity of  $f$ .

" $\Leftarrow$ ": Suppose that  $f$  satisfies the  $\epsilon$ - $\delta$  condition at  $x$ . Let  $(x_n)$  be a sequence in  $M$  such that  $(x_n)$  converges to  $x$ . Let  $\epsilon > 0$ .  $\exists \delta > 0$  such that  $d_M(x, y) < \delta \Rightarrow d_N(f(x), f(y)) < \epsilon$ . Since  $(x_n)$  approaches  $x$ ,  $\exists K \in \mathbb{N}$  such that if  $n \geq K$ , then  $d_M(x_n, x) < \delta$ , hence  $d_N(f(x_n), f(x)) < \epsilon$  for all  $n \geq K$ . That is,  $(f(x_n))$  converges to  $f(x)$ .  $\square$

### 1.3 The Topology of a Metric Space

**Definition 1.3.1** Let  $M$  be a metric space,  $S \subseteq M$ .

- We say  $S$  is open if for each  $x \in S$ ,  $\exists r > 0$  such that  $d(x, y) < r$  implies  $y \in S$ .
- We say  $S$  is closed if its complement is open.
- We say that a point  $x \in M$  is a limit of  $S$  if there exists a sequence in  $S$  that converges to  $x$ .

**Proposition 1.3.1** A set  $S \subseteq M$  is closed iff it contains all its limits.

*Proof.* " $\Rightarrow$ ": Suppose that  $S$  is closed. Let  $(x_n)$  be a sequence in  $S$  such that  $(x_n)$  converges to  $x$ ,  $x \in M$ . Suppose  $x \notin S$ , then since  $S^c$  is open,  $\exists r > 0$  such that  $d(x, y) < r$  implies  $y \in S^c$ . Since  $(x_n)$  converges to  $x$ , we have  $d(x_n, x) < r$  for all  $n$  large enough, implying that  $x_n \in S^c$ , contradiction. Therefore, for any limit point  $x$  of  $S$ ,  $x \in S$ .

" $\Leftarrow$ ": Suppose that  $S$  contains all its limits. If  $S^c$  is not open, then  $\exists x \in S^c$  such that  $\forall n \in \mathbb{N}$ ,  $\exists x_n \in (S^c)^c = S$  such that  $d(x, x_n) < \frac{1}{n}$ . We have now constructed a sequence  $(x_n)$  in  $S$ , but converges to a point in  $S^c$ , contradiction. Thus  $S^c$  is open, i.e.  $S$  is closed.  $\square$

**Remark** Sets can be neither open nor closed, or they can also be both open and closed.

**Definition 1.3.2** The topology  $\mathcal{T}$  of  $M$  is the collection of all open subsets of  $M$ .

**Proposition 1.3.2**  $\mathcal{T}$  is closed under arbitrary union, finite intersection, and  $\mathcal{T}$  contains  $\emptyset$  and  $M$ .

*Proof.* Clearly,  $\emptyset$  and  $M$  are open.

---

Let  $(\mathcal{U}_\alpha)$  be a collection of open subsets of  $M$ . Define  $V = \bigcup_\alpha \mathcal{U}_\alpha$ . For any  $x \in V$ ,  $x \in \mathcal{U}_\alpha$  for some  $\alpha$ . Since  $\mathcal{U}_\alpha$  is open,  $\exists r > 0$  such that  $d(x, y) < r \implies y \in \mathcal{U}_\alpha \subseteq V$ , so  $V$  is open.

Define  $W = \bigcap_{i=1}^n \mathcal{U}_i$ . Given  $x \in W$ , for each  $1 \leq i \leq n$ ,  $\exists r_i > 0$  such that  $d(x, y) < r_i \implies y \in \mathcal{U}_i$ . Take  $r = \min r_1, \dots, r_n$ , then for any  $y$  satisfying  $d(x, y) < r$ , we have  $y \in \mathcal{U}_i$  for all  $1 \leq i \leq n$ , thus  $y \in W$ . Hence  $W$  is open.  $\square$

**Definition 1.3.3** Let  $X$  be a set. A topology  $\mathcal{T}$  of  $X$  is a collection of subsets of  $X$  that satisfies the following:

- a)  $\mathcal{T}$  is closed under arbitrary union,
- b)  $\mathcal{T}$  is closed under finite intersection, and
- c)  $\emptyset, X \in \mathcal{T}$ .

We say that  $(X, \mathcal{T})$  is a topological space if  $\mathcal{T}$  is a topology of  $X$ . The elements of  $\mathcal{T}$  are called open sets. We define  $S \subseteq X$  to be closed if  $S^c$  is open.

**Example 1.3.1** The following are some examples of topological spaces:

- A metric space is a topological space.
- Let  $X$  be a set, and  $\mathcal{T} = \{\emptyset, X\}$ . Then  $(X, \mathcal{T})$  is a topological space, which is known as the trivial topology.
- Let  $X$  be a set, and let  $\mathcal{T}$  be the power set of  $X$ , then  $(X, \mathcal{T})$  is a topological space, which is known as the discrete topology.

**Remark** By De Morgan's law, closed sets are closed under arbitrary intersection and finite union, also  $\emptyset, X$  are closed.

In general, infinite union of closed sets may not be closed.

**Definition 1.3.4** Let  $M$  be a metric space and  $S \subseteq M$ . Define

$$\overline{S} := \{x \in M \mid x \text{ is a limit of } S\}$$

to be the closure of  $S$ . For  $x \in M$ ,  $r > 0$ , define

$$B(x, r) := \{y \in M \mid d(x, y) < r\},$$

which is the ball centered at  $x$  with radius  $r$ , or the  $r$ -neighborhood of  $x$ .

**Proposition 1.3.3**  $\overline{S}$  is closed and  $B(x, r)$  is open.

*Proof.* If  $S = \emptyset$ , then  $\overline{S} = \emptyset$ , which is closed.

---

Suppose that  $S \neq \emptyset$ , and let  $(x_n) \rightarrow x$  be a convergence sequence in  $\overline{S}$ . We wish to prove that  $x \in \overline{S}$ , i.e. there exists a sequence in  $S$  that approaches  $x$ . Since  $x_n \in \overline{S}$ , there exists sequence  $(x_{n,k})$  in  $S$  that approaches  $x_n$  as  $k \rightarrow \infty$ . For each  $n$ , there exists a term  $x_{n,k_n}$  satisfying  $d(x_{n,k_n}, x_n) < \frac{1}{n}$ , picking these terms forms a new sequence  $(x_{n,k_n})$ , moreover,

$$d(x_{n,k_n}, x) \leq d(x_{n,k_n}, x_n) + d(x_n, x) < \frac{1}{n} + d(x_n, x),$$

which approaches 0 as  $n \rightarrow \infty$ , thus  $(x_{n,k_n})$  approaches  $x$ .

Fix  $x \in M$  and  $r > 0$ . Let  $y \in B(x, r)$  and pick  $s = r - d(x, y)$ . If  $z$  satisfies  $d(y, z) < s$ , then

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + s = r \implies B(y, s) \subseteq B(x, r),$$

so  $B(x, r)$  is open.  $\square$

**Corollary 1.3.5**  $\overline{S}$  is the smallest closed set that contains  $S$ , i.e. if  $K \supseteq S$  and  $K$  is closed then  $K \supseteq \overline{S}$ .

*Proof.*  $K$  contains the limit of each sequence in  $K$ , in particular  $S \subseteq K$  so it contains all sequence in  $S$  that converges in  $M$ , but these are precisely  $\overline{S}$ .  $\square$

With the observation from Corollary 1.3.5, we may define the closure for topological spaces in general.

**Definition 1.3.6** Let  $X$  be a topological space, and let  $S \subseteq X$ . We define  $\overline{S}$  to be the smallest closed set that contains  $S$ .  $\overline{S}$  always exists; take

$$\overline{S} = \bigcap \{E \subseteq X \mid E \text{ is closed in } X \text{ and } E \supseteq S\}.$$

Likewise, we also define the interior of  $S$ ,  $\text{int}(S)$ , to be the largest open set contained in  $S$

**Definition 1.3.7** Let  $X$  be a topological space,  $x \in X$ . A neighborhood of  $x$  is an open set containing  $x$ .

**Definition 1.3.8** Let  $X$  be a topological space and let  $(x_n)$  be a sequence in  $X$ . We say that  $(x_n)$  converges to  $x \in X$  if for all neighborhood  $\mathcal{U}$  of  $x$ ,  $\exists N \in \mathbb{N}$  such that  $x_n \in \mathcal{U}$  for all  $n \geq N$ .

**Example 1.3.2** Limits in general are not unique in topological spaces. Let  $X$  be a set with at least 2 points endowed with the trivial topology. Then every sequence in  $X$  converges to every point in  $X$ .

**Definition 1.3.9** Let  $X, Y$  be topological spaces. We say a function  $f : X \rightarrow Y$  continuous if for any open set  $V \subseteq Y$ , the preimage  $f^{-1}(V)$  is open in  $X$ .

---

**Proposition 1.3.4** A function  $f : X \rightarrow Y$  is continuous iff  $\forall x \in M$  and any neighborhood  $\mathcal{V}$  of  $f(x)$ ,  $\exists$  a neighborhood  $\mathcal{U}$  of  $x$  such that  $f(\mathcal{U}) \subseteq \mathcal{V}$ .

*Proof.* " $\Rightarrow$ ": Let  $\mathcal{V} \subseteq Y$  is open. We need to show  $f^{-1}(\mathcal{V})$  is open. Let  $x \in f^{-1}(\mathcal{V})$ , then  $f(x) \in \mathcal{V}$ . By definition,  $\exists$  a neighborhood  $\mathcal{U}_x$  such that  $f(\mathcal{U}_x) \subseteq \mathcal{V}$ . Take the union of all such  $\mathcal{U}_x$  over  $x \in f^{-1}(\mathcal{V})$ , then

$$\begin{aligned} \bigcup_{x \in f^{-1}(\mathcal{V})} \mathcal{U}_x &\subseteq f^{-1}(\mathcal{V}) \quad \text{and} \quad \forall x \in f^{-1}(\mathcal{V}), x \in \mathcal{U}_x \subseteq f^{-1}(\mathcal{V}) \\ \Rightarrow f^{-1}(\mathcal{V}) &= \bigcup_{x \in f^{-1}(\mathcal{V})} \mathcal{U}_x \end{aligned}$$

and thus  $f^{-1}(\mathcal{V})$  is open.

" $\Leftarrow$ ": Let  $x \in X$  and let  $\mathcal{V}$  be a neighborhood of  $f(x)$ . By definition  $x \in f^{-1}(\mathcal{V})$ , which is an open set by assumption. Also,  $f(f^{-1}(\mathcal{V})) \subseteq \mathcal{V}$ , hence  $f^{-1}(\mathcal{V})$  is a neighborhood of  $x$  such that  $f(f^{-1}(\mathcal{V})) \subseteq \mathcal{V}$ .  $\square$

**Definition 1.3.10** A homeomorphism is a continuous bijection between topological spaces.

**Corollary 1.3.11** A homeomorphism  $f : X \rightarrow Y$  bijects the corresponding topologies  $\mathcal{T}_X$  and  $\mathcal{T}_Y$ .

## 1.4 Hausdorff Space

**Definition 1.4.1** A topological space  $X$  is said to be Hausdorff if given any pair of distinct points  $x_1, x_2 \in X$ ,  $\exists$  neighborhoods  $\mathcal{U}_1$  of  $x_1$  and  $\mathcal{U}_2$  of  $x_2$  such that  $\mathcal{U}_1 \cap \mathcal{U}_2 = \emptyset$ .

A metric space is always a Hausdorff space.

**Lemma 1.4.2** Let  $X$  be Hausdorff.

- a) Every one-point set is closed.
- b) If a sequence  $(x_n)$  in  $X$  converges, then the limit is unique.

*Proof.* a) Pick  $x \in X$ . For any  $y$  distinct from  $x$ , there exists disjoint neighborhoods  $\mathcal{U}_x$  of  $x$  and  $\mathcal{V}_y$  of  $y$ . We have

$$\{x\}^c = \bigcup_{y \in X \setminus \{x\}} \mathcal{V}_y$$

which is an open set.

- b) Suppose that  $x, x'$  are distinct limits of  $(x_n)$ .  $\exists$  disjoint neighborhoods  $\mathcal{U}$  of  $x$  and  $\mathcal{U}'$

---

of  $x'$ .  $\exists N, N' \in \mathbb{N}$  such that " $n \geq N$  implies  $x_n \in \mathcal{U}$ " and " $n \geq N'$  implies  $x_n \in \mathcal{U}'$ ". If  $n \geq \max(N, N')$ , then  $x_n \in \mathcal{U} \cap \mathcal{U}' = \emptyset$ , contradiction. Therefore, any converging sequence has a unique limit.

□

## 1.5 Subspaces and Product Spaces

**Definition 1.5.1** Let  $X$  be a topological space, let  $A \subseteq X$ . We define the subspace topology  $\mathcal{T}_A$  of  $A$  by

$$\mathcal{T}_A = \{\mathcal{U} \subseteq A \mid \mathcal{U} = A \cap \mathcal{V} \text{ for some open } \mathcal{V} \subseteq X\}.$$

**Remark** Openness and closedness are not just properties of a set itself, but rather a set in a relation to a particular topological space.

**Proposition 1.5.1** Let  $M$  be a metric space, and let  $N \subseteq M$  be a nonempty subset. The subspace topology on  $N$  is the same as the metric topology obtained by restricting the metric of  $M$  to  $N$ .

*Proof.* Suppose that  $\mathcal{V}$  is an open set in  $M$  and let  $\mathcal{U} = N \cap \mathcal{V} \in \mathcal{T}_N$ .

We first prove that  $\mathcal{U}$  belongs to the metric topology of  $N$ . Let  $x \in \mathcal{U}$ . Since  $x \in N \cap \mathcal{V} \subseteq \mathcal{V}$ , there exists  $r > 0$  such that  $B_M(x, r) \subseteq \mathcal{V}$ , and

$$B_N(x, r) = N \cap B_M(x, r) \subseteq N \cap \mathcal{V} = \mathcal{U},$$

hence  $\mathcal{U}$  is open in  $N$  (in the metric topology).

Conversely, let  $\mathcal{U}$  be an open set in the metric topology of  $N$ .  $\forall x \in \mathcal{U}, \exists r_x > 0$  such that  $B_N(x, r_x) \subseteq \mathcal{U}$ . Note that

$$\mathcal{U} = \bigcup_{x \in \mathcal{U}} B_N(x, r_x) = \bigcup_{x \in \mathcal{U}} N \cap B_M(x, r_x) = N \cap \bigcup_{x \in \mathcal{U}} B_M(x, r_x),$$

hence  $\mathcal{U}$  belongs to the subspace topology of  $N$ . □

**Definition 1.5.2** Let  $X$  be a set. A basis in  $X$  is a collection  $\mathcal{B}$  of subsets of  $X$  satisfying

- a) Every element of  $X$  is in some element in  $\mathcal{B}$ . That is,

$$X = \bigcup_{B \in \mathcal{B}} B.$$

- b) If  $B_1, B_2 \in \mathcal{B}$  and  $x \in B_1 \cap B_2$ , there exists  $B_3 \in \mathcal{B}$  such that  $x \in B_3 \subseteq B_1 \cap B_2$ .

---

**Definition 1.5.3** Given  $X$  and a collection  $\mathcal{B}$  of subsets of  $X$ , we say that  $\mathcal{U} \subseteq X$  satisfies the basis criterion with respect to  $\mathcal{B}$  if  $\forall x \in \mathcal{U}, \exists B \in \mathcal{B}$  such that  $x \in B \subseteq \mathcal{U}$

**Lemma 1.5.4** Suppose that  $\mathcal{B}$  is a basis in  $X$ , and let  $\mathcal{T}$  be the collection of all unions of elements of  $\mathcal{B}$ . Then  $\mathcal{T}$  is precisely the collection of all subsets of  $X$  that satisfy the basis criterion w.r.t.  $\mathcal{B}$ .

*Proof.* Let  $\mathcal{U} \subseteq X$ . Suppose that  $\mathcal{U}$  satisfies the basis criterion. Let

$$\mathcal{V} = \bigcup \{B \in \mathcal{B} \mid B \subseteq \mathcal{U}\},$$

then  $\mathcal{V} \in \mathcal{T}$ . We want to show that  $\mathcal{U} = \mathcal{V}$ . Clearly,  $\mathcal{V} \subseteq \mathcal{U}$ . Let  $x \in \mathcal{U}$ . Since  $\mathcal{U}$  satisfies the basis criterion,  $\exists B \in \mathcal{B}$  such that  $x \in B \subseteq \mathcal{U}$ , so  $x \in \mathcal{V}$ , therefore  $\mathcal{U} \subseteq \mathcal{V}$ .

Conversely, suppose that  $\mathcal{U} \in \mathcal{T}$ , then  $\mathcal{U}$  is a union of elements of  $\mathcal{B}$ , say  $\mathcal{U} = \bigcup_{B \in \mathcal{A}} B$  where  $\mathcal{A} \subseteq \mathcal{B}$ . For any  $x \in \mathcal{U}$ ,  $x \in B$  for some  $B \in \mathcal{A}$ , also  $B \subseteq \mathcal{U}$ , so  $\mathcal{U}$  satisfies the basis criterion.  $\square$

**Proposition 1.5.2** Let  $\mathcal{B}$  be a basis in  $X$ , and let  $\mathcal{T}$  be collection of all unions of elements of  $\mathcal{B}$ . Then  $\mathcal{T}$  is a topology on  $X$ . This is called the topology generated by  $\mathcal{B}$ .

*Proof.* First of all,  $\emptyset = \bigcup_{B \in \emptyset} B \in \mathcal{T}$  and  $X = \bigcup_{B \in \mathcal{B}} B \in \mathcal{T}$ .

Let  $(\mathcal{U}_\alpha)$  be a collection of elements in  $\mathcal{T}$ , then the union  $\mathcal{U} = \bigcup_\alpha \mathcal{U}_\alpha$  is also a union of elements in  $\mathcal{B}$ , hence  $\mathcal{U} \in \mathcal{T}$ .

Let  $\mathcal{U}_1, \mathcal{U}_2 \in \mathcal{T}$ . For any  $x \in \mathcal{U}_1 \cap \mathcal{U}_2$ , since  $\mathcal{U}_1, \mathcal{U}_2$  satisfy the basis criterion,  $\exists B_1, B_2 \in \mathcal{B}$  such that  $x \in B_1 \subseteq \mathcal{U}_1$  and  $x \in B_2 \subseteq \mathcal{U}_2$ . By the definition of basis,  $\exists B_3 \in \mathcal{B}$  such that  $x \in B_3 \subseteq B_1 \cap B_2 \subseteq \mathcal{U}_1 \cap \mathcal{U}_2$ . Therefore,  $\mathcal{U}_1 \cap \mathcal{U}_2$  satisfies the basis criterion w.r.t.  $\mathcal{B}$ , hence  $\mathcal{U}_1 \cap \mathcal{U}_2 \in \mathcal{T}$ .  $\square$

Now, we can define product topology.

**Definition 1.5.5** Let  $X_1, \dots, X_n$  be topological spaces. Define a basis in  $X_1 \times \dots \times X_n$  by

$$\mathcal{B} = \{\mathcal{U}_1 \times \dots \times \mathcal{U}_n \mid \mathcal{U}_i \text{ is open in } X_i, 1 \leq i \leq n\}.$$

The product topology on  $X_1 \times \dots \times X_n$  is the topology generated by  $\mathcal{B}$ .

**Proposition 1.5.3** If  $X_1, \dots, X_n, Y_1, \dots, Y_n$  are topological spaces, and if  $f_i : X_i \rightarrow Y_i$  are continuous functions for  $1 \leq i \leq n$ , then the function

$$f : X_1 \times \dots \times X_n \rightarrow Y_1 \times \dots \times Y_n, \quad f(x_1, \dots, x_n) = (f_1(x_1), \dots, f_n(x_n))$$

---

is continuous.

*Proof.* We first prove that if the preimages of basis elements are open, then  $f$  is continuous. Let  $\mathcal{U}$  be an open set in  $Y_1 \times \cdots \times Y_n$ . By definition,  $\mathcal{U}$  is a union of basis elements, say  $\mathcal{U} = \bigcup_{\alpha} \mathcal{V}_{\alpha}$ , where each  $\mathcal{V}_{\alpha}$  is a basis element. We have

$$f^{-1}(\mathcal{U}) = f^{-1}\left(\bigcup_{\alpha} \mathcal{V}_{\alpha}\right) = \bigcup_{\alpha} f^{-1}(\mathcal{V}_{\alpha}),$$

which is open if each  $f^{-1}(\mathcal{V}_{\alpha})$  is open.

Let  $\mathcal{V} = \mathcal{V}_1 \times \cdots \times \mathcal{V}_n$  be a basis element in  $Y_1 \times \cdots \times Y_n$ , where  $\mathcal{V}_i$  is open in  $Y_i$ . Then

$$f^{-1}(\mathcal{V}) = f^{-1}(\mathcal{V}_1 \times \cdots \times \mathcal{V}_n) = f_1^{-1}(\mathcal{V}_1) \times \cdots \times f_n^{-1}(\mathcal{V}_n),$$

which is a basis element in  $X_1 \times \cdots \times X_n$  since each  $f_i$  is continuous. Therefore,  $f$  is continuous.  $\square$

**Definition 1.5.6** Let  $(M, d_M), (N, d_N)$  be metric spaces. Define the  $p$ -metric on  $M \times N$  by

$$d_p((x_1, y_1), (x_2, y_2)) = (d_M(x_1, x_2)^p + d_N(y_1, y_2)^p)^{\frac{1}{p}}, \quad 1 \leq p < \infty,$$

and the  $\infty$ -metric by

$$d_{\infty}((x_1, y_1), (x_2, y_2)) = \max \{d_M(x_1, x_2), d_N(y_1, y_2)\}.$$

**Proposition 1.5.4** The following properties hold for metric spaces  $M, N$ :

- $d_{\infty} \leq d_p \leq 2^{\frac{1}{p}} d_{\infty}$
- The metric topologies induced by  $d_p$  and  $d_{\infty}$  are the same.
- The metric topology coincides with the product topology.

**Proposition 1.5.5** Let  $M$  be a metric space, then the metric  $d : M \times M \rightarrow \mathbb{R}$  is continuous.

*Proof.* We use the metric  $d_1$  on  $M \times M$ .

Let  $(x_1, y_1), (x_2, y_2) \in M \times M$ . We have

$$\begin{aligned} |d(x_1, y_1) - d(x_2, y_2)| &\leq |d(x_1, y_1) - d(x_1, y_2)| + |d(x_1, y_2) - d(x_2, y_2)| \\ &\leq d(y_1, y_2) + d(x_1, x_2). \end{aligned}$$

Given  $\epsilon > 0$ , take  $\delta = \frac{\epsilon}{2}$ . If  $d_1(x_1, y_1), d_1(x_2, y_2) < \delta$ , then

$$|d(x_1, y_1) - d(x_2, y_2)| < 2\delta = \epsilon,$$

---

hence  $d$  is continuous. □

So far, we've defined product topology for finite products. We can also define product topology for infinite products.

**Definition 1.5.7** Let  $(X_i, \mathcal{T}_i)$  be topological spaces for  $i \in I$ , where  $I$  is an index set. We define the box topology on  $\prod_{i \in I} X_i$  to be the topology generated by the basis

$$\mathcal{B} = \left\{ \prod_{i \in I} \mathcal{U}_i \mid \mathcal{U}_i \text{ is open in } X_i, i \in I \right\}.$$

The product topology on  $\prod_{i \in I} X_i$  is the topology generated by the basis

$$\mathcal{B}' = \left\{ \prod_{i \in I} \mathcal{U}_i \mid \mathcal{U}_i \text{ is open in } X_i, \mathcal{U}_i = X_i \text{ for all but finitely many } i \right\}.$$

**Remark** For finite products, the box topology and the product topology coincide. However, for infinite products, the box topology is strictly finer than the product topology, i.e. every open set in the product topology is also open in the box topology, but not vice versa.

We will prefer the product topology over the box topology, because a number of important theorems about finite products still hold for infinite products under the product topology.

**Example 1.5.1** Let  $\mathbb{R}^\omega = \prod_{n \in \mathbb{N}} \mathbb{R}$ . Consider the function

$$f : \mathbb{R} \rightarrow \mathbb{R}^\omega, \quad f(x) = (x, x, x, \dots).$$

$f$  is continuous under the product topology, but not continuous under the box topology. To see this, let  $\mathcal{U} = \prod_{n \in \mathbb{N}} (-\frac{1}{n}, \frac{1}{n})$ , which is open in the box topology. We have

$$f^{-1}(\mathcal{U}) = \bigcap_{n \in \mathbb{N}} \left( -\frac{1}{n}, \frac{1}{n} \right) = \{0\},$$

which is not open in  $\mathbb{R}$ .

## 1.6 Completion

**Definition 1.6.1** A sequence  $(x_n)$  in a metric space  $M$  is called a Cauchy sequence if  $\forall \epsilon > 0$ ,  $\exists N \in \mathbb{N}$  such that  $m, n \geq N$  implies  $d(x_n, x_m) < \epsilon$ .

Every convergent sequence is a Cauchy sequence. The converse is not true in general, e.g. in  $\mathbb{Q}$ .

---

**Definition 1.6.2** A metric space  $M$  is said to be complete if every Cauchy sequence in  $M$  converges to a point in  $M$ .

Completeness is not a topological property, i.e. two homeomorphic metric spaces may not be both complete or both incomplete.

**Example 1.6.1** Define two metric on  $\mathbb{N}$  by

$$d_1(m, n) = |m - n|, \quad d_2(m, n) = \left| \frac{1}{m} - \frac{1}{n} \right|.$$

We can verify that  $(\mathbb{N}, d_1)$  and  $(\mathbb{N}, d_2)$  are homeomorphic, and  $(\mathbb{N}, d_1)$  is complete. However,  $(\mathbb{N}, d_2)$  is not complete since the sequence  $(n)$  is Cauchy but does not converge in  $\mathbb{N}$ .

**Proposition 1.6.1** Every closed subset of a complete metric space is a complete metric subspace.

**Corollary 1.6.3** Every closed subset of the Euclidean space  $\mathbb{R}^n$  is a complete metric subspace.

**Theorem 1.6.4** Every metric space can be completed. That is, a metric space  $M$  is always a metric subspace of a complete metric space  $\widehat{M}$ .

*Proof.* Let  $\mathcal{C}$  be the collection of all Cauchy sequences in  $M$ . Define a relation  $\sim$  on  $\mathcal{C}$  by

$$(x_n) \sim (y_n) \iff \lim_{n \rightarrow \infty} d(x_n, y_n) = 0.$$

We can verify that  $\sim$  is an equivalence relation. Let  $\widehat{M} = \mathcal{C}/\sim$ , and denote the equivalence class of  $(x_n)$  by  $[(x_n)]$ . Define a metric  $D$  on  $\widehat{M}$  by

$$D(X, Y) = \lim_{n \rightarrow \infty} d(x_n, y_n),$$

where  $X = [(x_n)], Y = [(y_n)]$ .

We first verify that  $D$  is well-defined and is a metric on  $\widehat{M}$ . By triangle inequality,

$$\begin{aligned} |d(x_n, y_n) - d(x_m, y_m)| &\leq |d(x_n, y_n) - d(x_m, y_n)| + |d(x_m, y_n) - d(x_m, y_m)| \\ &\leq d(x_n, x_m) + d(y_n, y_m), \end{aligned}$$

and since  $(x_n), (y_n)$  are Cauchy sequences,  $\forall \epsilon > 0, \exists N \in \mathbb{N}$  such that  $m, n \geq N$  implies  $|d(x_n, y_n) - d(x_m, y_m)| < \epsilon$ , hence  $(d(x_n, y_n))$  is a Cauchy sequence in  $\mathbb{R}$ , which converges to a limit in  $\mathbb{R}$ . Now consider another representative  $(x'_n), (y'_n)$  of  $X, Y$  respectively, and let

$$L = \lim_{n \rightarrow \infty} d(x_n, y_n), \quad L' = \lim_{n \rightarrow \infty} d(x'_n, y'_n).$$

---

By triangle inequality again, we have

$$\begin{aligned} |d(x_n, y_n) - d(x'_n, y'_n)| &\leq |d(x_n, y_n) - d(x'_n, y_n)| + |d(x'_n, y_n) - d(x'_n, y'_n)| \\ &\leq d(x_n, x'_n) + d(y_n, y'_n). \end{aligned}$$

Since  $(x_n) \sim (x'_n)$  and  $(y_n) \sim (y'_n)$ ,  $\forall \epsilon > 0$ ,  $\exists N \in \mathbb{N}$  such that  $n \geq N$  implies  $|d(x_n, y_n) - d(x'_n, y'_n)| < \epsilon$ , hence  $L = L'$ . Therefore,  $D$  is well-defined.

It is clear that  $D(X, Y) \geq 0$  and  $D(X, Y) = D(Y, X)$ . If  $D(X, Y) = 0$ , then  $d(x_n, y_n) \rightarrow 0$ , so  $(x_n) \sim (y_n)$ , i.e.  $X = Y$ . Conversely, if  $X = Y$ , then  $(x_n) \sim (y_n)$ , hence  $D(X, Y) = 0$ . Finally, for any  $X = [(x_n)], Y = [(y_n)], Z = [(z_n)] \in \widehat{M}$ , by triangle inequality,

$$D(X, Z) = \lim_{n \rightarrow \infty} d(x_n, z_n) \leq \lim_{n \rightarrow \infty} d(x_n, y_n) + \lim_{n \rightarrow \infty} d(y_n, z_n) = D(X, Y) + D(Y, Z).$$

Therefore,  $D$  is a metric on  $\widehat{M}$ .

For each  $x \in M$ , define

$$\bar{x} = (x, x, x, \dots) \in \mathcal{C},$$

then  $\bar{x}$  is a Cauchy sequence, and  $D(\bar{x}, \bar{y}) = d(x, y)$  for any  $x, y \in M$ . It is clear that the mapping  $x \mapsto \bar{x}$  is an isometric embedding of  $M$  into  $\widehat{M}$ , so we may identify  $M$  as a metric subspace of  $\widehat{M}$ .

Finally, we need to show that  $\widehat{M}$  is complete. Let  $(X_n)$  be a Cauchy sequence in  $\widehat{M}$ , where  $X_n = [(x_{n,k})]$ . We need to find  $Y \in \widehat{M}$  such that  $X_n \rightarrow Y$ .

Note that if  $(x_n)$  is Cauchy in  $M$ , then any subsequence is equivalent to  $(x_n)$ , hence WLOG we can assume that  $d(x_{n,k}, x_{n,l}) < 1/n$  for all  $k, l \geq N$ .

Set  $y_n = x_{n,n}$ . Let  $\epsilon > 0$ . Pick  $N \in \mathbb{N}$  such that  $N \geq 3/\epsilon$  and  $m, n \geq N$  implies  $D(X_n, X_m) < \epsilon/3$ , hence

$$\begin{aligned} d(y_m, y_n) &= d(x_{m,m}, x_{n,n}) \\ &\leq d(x_{m,m}, x_{m,k}) + d(x_{m,k}, x_{n,k}) + d(x_{n,k}, x_{n,n}) \\ &< \frac{1}{m} + d(x_{m,k}, x_{n,k}) + \frac{1}{n} < \frac{2\epsilon}{3} + d(x_{m,k}, x_{n,k}). \end{aligned}$$

As  $n$  approaches infinity,  $d(x_{m,k}, x_{n,k})$  approaches  $D(X_m, X_n) < \epsilon/3$ , hence  $d(y_m, y_n) < \epsilon$  for all  $m, n \geq N$ . Therefore,  $(y_n)$  is a Cauchy sequence in  $M$ .

Let  $Y = [(y_n)] \in \widehat{M}$ . It remains to show that  $X_n \rightarrow Y$ . Given  $\epsilon > 0$ , pick  $N \in \mathbb{N}$  such that  $N \geq 2/\epsilon$  such that  $k, n \geq N$  implies  $d(y_k, y_n) < \epsilon/2$ , then

$$\begin{aligned} d(x_{n,k}, y_k) &= d(x_{n,k}, x_{k,k}) \\ &\leq d(x_{n,k}, x_{n,n}) + d(x_{n,n}, x_{k,k}) < \frac{1}{n} + d(y_n, y_k) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

---

Taking  $n \rightarrow \infty$ , we have  $D(X_n, Y) \leq \epsilon$ . Therefore,  $X_n \rightarrow Y$  in  $\widehat{M}$ , and  $\widehat{M}$  is complete.  $\square$

We might wish to construct the real numbers  $\mathbb{R}$  by completing the rational numbers  $\mathbb{Q}$ . However, this approach results in some sort of circular reasoning.

We can still define concepts such as convergence and cauchy sequence, by restricting the  $\epsilon$ 's to rational numbers. Then we can still consider the completion  $\widehat{\mathbb{Q}}$  under this definition. We wish to prove that  $\widehat{\mathbb{Q}}$  is a ordered complete field.

For  $X = (x_n)$  and  $Y = (y_n)$  in  $\widehat{\mathbb{Q}}$ , define

$$X + Y = (x_n + y_n), \quad X - Y = (x_n - y_n), \quad XY = (x_n y_n), \quad \frac{X}{Y} = \left( \frac{x_n}{y_n} \right).$$

If  $X \in \widehat{\mathbb{Q}}$  has a representation  $(x_n)$  such that for some  $\epsilon > 0$ ,  $x_n \geq \epsilon$  for all  $n$ , then we say  $X$  is positive. If  $-X$  is positive, then we say  $X$  is negative. Define  $X < Y$  if  $Y - X$  is positive.

**Proposition 1.6.2** For  $X = [(x_n)]$  and  $Y = [(y_n)]$  in  $\widehat{\mathbb{Q}}$ ,  $X < Y$  iff  $\exists$  (rational)  $\epsilon > 0$  and  $N \in \mathbb{N}$  such that  $y_m - x_n > \epsilon$  for all  $m, n \geq N$ .

It remains to show that every nonempty subset of  $\widehat{\mathbb{Q}}$  that is bounded above has a least upper bound. Define the least upper bound for  $\mathcal{A}$  to be the equivalence class  $Y$  of the sequence  $(y_n)$ , where

1.  $y_0$  is the smallest integer such that  $[(y_0, y_0, y_0, \dots)]$  is an upper bound for  $\mathcal{A}$ .
2. For  $n \geq 1$ ,  $y_n$  is the smallest rational with denominator  $2^n$  such that  $[(y_n, y_n, y_n, \dots)]$  is an upper bound for  $\mathcal{A}$ .

We can check that  $(y_n)$  is monotone decreasing and  $y_n - y_{n+1} \leq 2^{-n}$ . For  $m, n \in \mathbb{N}$  with  $m > n$ , we have

$$|y_m - y_n| \leq \sum_{j=n}^{m-1} |y_{j+1} - y_j| \leq \sum_{j=n}^{m-1} 2^{-j} < \sum_{j=n}^{\infty} 2^{-j} = 2^{1-n},$$

so  $(y_n)$  is Cauchy and thus in  $\widehat{\mathbb{Q}}$ . Now we check that  $Y = \sup \mathcal{A}$ . Suppose that  $Y$  is not an upper bound for  $\mathcal{A}$ , then  $\exists X = [(x_n)] \in \mathcal{A}$  such that  $X > Y$ . By Proposition 1.6.2,  $\exists \epsilon > 0$  and  $N \in \mathbb{N}$  such that  $x_n - y_N > \epsilon$  for all  $n \geq N$ , then

$$[(y_N, y_N, y_N, \dots)] < X,$$

contradicting the construction of  $y_n$ . Therefore,  $Y$  is an upper bound for  $\mathcal{A}$ .

Now suppose that  $Z = [(z_n)]$  is another upper bound for  $\mathcal{A}$  such that  $Z < Y$ . By Proposition 1.6.2,  $\exists \epsilon > 0$  and  $N \in \mathbb{N}$  such that  $y_m - z_n > \epsilon$  for all  $m, n \geq N$ .

---

Pick  $k \geq N$  such that  $1/2^{k-1} < \epsilon$ , then for all  $m \geq N$ , we have

$$\begin{aligned} z_m + \frac{\epsilon}{2} &< y_k - \frac{\epsilon}{2} < y_k - \frac{1}{2^k} \\ \implies Z &< [(y_k - \frac{1}{2^k}, y_k - \frac{1}{2^k}, y_k - \frac{1}{2^k}, \dots)], \end{aligned}$$

but  $y_k$  is the smallest fraction with denominator  $2^k$  such that  $[(y_k, y_k, y_k, \dots)]$  is an upper bound for  $\mathcal{A}$ , contradiction. Therefore,  $Y = \sup \mathcal{A}$ .

## 1.7 Connectedness

**Definition 1.7.1** If  $X$  is a topological space, a separation of  $X$  is a pair of disjoint nonempty open sets  $U, V \subseteq X$  such that  $X = U \cup V$ . A topological space  $X$  is connected if there does not exist a separation of  $X$ .

**Proposition 1.7.1** A topological space  $X$  is connected iff the only clopen subsets of  $X$  are  $\emptyset$  and  $X$ .

**Proposition 1.7.2** Let  $X, Y$  be topological spaces and let  $f : X \rightarrow Y$  be continuous. If  $X$  is connected, then  $f(X)$  is connected.

*Proof.* Suppose that  $f(X)$  is not connected, then  $\exists$  a separation  $U, V$  of  $f(X)$ , i.e. exists nonempty open sets  $U, V$  in  $Y$  such that  $U \cap V \cap f(X) = \emptyset$  and  $U \cup V \supseteq f(X)$ . Then  $f^{-1}(U)$  and  $f^{-1}(V)$  are disjoint nonempty open sets in  $X$  such that  $X = f^{-1}(U) \cup f^{-1}(V)$ , contradicting the connectedness of  $X$ . Therefore,  $f(X)$  is connected.  $\square$

**Corollary 1.7.2** If  $X$  is connected and if  $X$  is homeomorphic to  $Y$ , then  $Y$  is connected.

**Corollary 1.7.3** (Intermediate Value Theorem) Let  $X$  be a connected topological space. If  $f : X \rightarrow \mathbb{R}$  is continuous, then if  $a, b \in f(X)$  and if  $c$  is between  $a$  and  $b$ , then  $c \in f(X)$ .

*Proof.* Suppose  $f$  attains  $a, b$  but does not attain  $c$  such that  $a < c < b$ . Then

$$X = \{x \in X \mid f(x) < c\} \cup \{x \in X \mid f(x) > c\}$$

is a separation of  $X$ , contradicting the connectedness of  $X$ . Therefore,  $c \in f(X)$ .  $\square$

**Proposition 1.7.3**  $\mathbb{R}$  is connected.

*Proof.* Let  $\mathcal{U} \subseteq \mathbb{R}$  be nonempty and clopen. Let  $x \in \mathcal{U}$  and consider

$$A = \{u \in \mathcal{U} \mid \text{the open interval } (x, u) \subseteq \mathcal{U}\}.$$

---

$A$  is nonempty since  $\mathcal{U}$  is open. Let  $s = \sup A$ . If  $s < \infty$ , then  $s$  is a limit of  $A$ . Since  $\mathcal{U}$  is closed and  $A \subseteq \mathcal{U}$ ,  $s \in \mathcal{U}$ . Since  $\mathcal{U}$  is open,  $\exists \epsilon > 0$  such that  $(s - \epsilon, s + \epsilon) \subseteq \mathcal{U}$ , but then  $(x, s + \epsilon) \subseteq \mathcal{U}$ , contradicting the fact that  $s$  is an upper bound of  $A$ . Therefore,  $s = \infty$ , and  $(x, \infty) \subseteq \mathcal{U}$ . Similarly, we can show that  $(-\infty, x) \subseteq \mathcal{U}$ . Therefore,  $\mathcal{U}$  must equal  $\mathbb{R}$ , and  $\mathbb{R}$  is connected.  $\square$

**Corollary 1.7.4**  $(a, b)$ ,  $[a, b]$ , and  $\mathbb{S}^1$  are connected.

*Proof.*  $(a, b)$  is homeomorphic to  $\mathbb{R}$  via  $f(x) = \tan(\pi(x - \frac{a+b}{2})/(b-a))$ .  $[a, b]$  is connected since it is the continuous image of  $\mathbb{R}$  via

$$g(x) = \begin{cases} a & x < 0, \\ x & 0 \leq x \leq 1, \\ b & x > 1. \end{cases}$$

.  $\mathbb{S}^1$  is connected since it is the continuous image of  $[0, 1]$  via  $h(x) = (\cos x, \sin x)$ .  $\square$

**Remark**  $[a, b]$  is not homeomorphic to  $\mathbb{S}^1$ . Removing a point  $x \in (a, b)$  disconnects  $[a, b]$ , but removing any point from  $\mathbb{S}^1$  leaves it connected.

**Proposition 1.7.4** If  $X$  is a topological space and  $A \subseteq X$  is connected, then the closure  $\overline{A}$  is connected.

*Proof.* Suppose that  $\mathcal{U}, \mathcal{V}$  is a separation of  $\overline{A}$ . If  $A$  has nonempty intersection with both  $\mathcal{U}$  and  $\mathcal{V}$ , then  $A = (A \cap \mathcal{U}) \cup (A \cap \mathcal{V})$  is a separation of  $A$ , contradicting the connectedness of  $A$ . Therefore, WLOG,  $A \subseteq \mathcal{U}$ . Each point in  $V$  has a neighborhood that is disjoint from  $A$  (just pick  $V$ ), hence  $V \cap \overline{A} = \emptyset$ , contradicting the fact that  $\mathcal{U}, \mathcal{V}$  is a separation of  $\overline{A}$ . Therefore,  $\overline{A}$  is connected.  $\square$

**Proposition 1.7.5** The union of connected sets sharing a common point is connected.

*Proof.* Let  $S = \bigcup_{\alpha \in I} S_\alpha$ , where each  $S_\alpha$  is connected and  $x \in S_\alpha$  for all  $\alpha \in I$ .

Suppose that  $\mathcal{U}, \mathcal{V}$  is a separation of  $S$ , then  $x$  must be in either  $\mathcal{U}$  or  $\mathcal{V}$ , WLOG, say  $x \in \mathcal{U}$ . Note that  $\mathcal{U} \cap S_\alpha$  is clopen in  $S_\alpha$  for all  $\alpha \in I$ , and it is nonempty since it contains  $x$ . Therefore,  $\mathcal{U} \cap S_\alpha = S_\alpha$  for all  $\alpha \in I$ , hence  $S \subseteq \mathcal{U}$ , contradicting the fact that  $\mathcal{U}, \mathcal{V}$  is a separation of  $S$ . Therefore,  $S$  is connected.  $\square$

**Example 1.7.1** The 2-sphere  $\mathbb{S}^2$  is connected, since it is the surface of revolution of the semicircle, which is connected. Convex subsets of  $\mathbb{R}^n$  since it is the union of line segments sharing a common point.

---

**Definition 1.7.5** A path joining  $x$  to  $y$  in a topological space  $X$  is a continuous map  $f : [0, 1] \rightarrow X$  such that  $f(0) = x$  and  $f(1) = y$ . If each pair of points in  $X$  can be joined by a path, then  $X$  is path-connected.

**Theorem 1.7.6** A path-connected topological space is connected.

*Proof.* Suppose that  $X$  is path-connected but not connected. Then  $\exists$  a separation  $\mathcal{U}, \mathcal{V}$  of  $X$ . Pick  $x \in \mathcal{U}$  and  $y \in \mathcal{V}$ . Since  $X$  is path-connected,  $\exists$  a path  $f : [0, 1] \rightarrow X$  such that  $f(0) = x$  and  $f(1) = y$ . Then

$$[0, 1] = f^{-1}(\mathcal{U}) \cup f^{-1}(\mathcal{V})$$

is a separation of  $[0, 1]$ , contradicting the connectedness of  $[0, 1]$ . Therefore,  $X$  is connected.  $\square$

**Example 1.7.2**  $\mathbb{R}^n$ , convex subsets of  $\mathbb{R}^n$ , and  $\mathbb{S}^n$  are path-connected, hence connected.

**Example 1.7.3** Define subsets of  $\mathbb{R}^2$  by

$$\begin{aligned} A &= \{(x, y) \mid x = 0, y \in [-1, 1]\}, \\ B &= \{(x, y) \mid y = \sin(1/x), x \in (0, 1]\}. \end{aligned}$$

Let  $X = A \cup B$ .  $X$  is connected since  $B$  is connected and  $X = \overline{B}$ . However,  $X$  is not path-connected.

## 1.8 Other Metric Space Concepts

**Definition 1.8.1** We say  $x$  is an accumulation point (cluster point) of  $S$  if every neighborhood of  $x$  contains at least one point of  $S$  different from  $x$ . We say  $x$  is a condensation point of  $S$  if every neighborhood of  $x$  contains uncountably many points of  $S$ .

**Proposition 1.8.1** Let  $M$  be a metric space and let  $S \subseteq M$ . Then  $x$  is an accumulation point of  $S$  iff  $\exists$  a sequence of distinct points in  $S$  that converges to  $x$ .

*Proof.* " $\Rightarrow$ ": Let  $x$  be an accumulation point of  $S$ . Pick  $x_1 \in S$  such that  $0 < d(x_1, x) < 1$ . Set  $r_2 = d(x_1, x)/2$ , and choose  $x_2 \in S$  such that  $0 < d(x_2, x) < r_2$ . Note that  $x_2 \neq x_1$  since  $d(x_2, x) < d(x_1, x)$ . Continuing this process, we obtain a sequence of distinct points  $(x_n)$  in  $S$  such that  $d(x_n, x) < 1/2^{n-1}$ . Therefore,  $x_n \rightarrow x$ .

" $\Leftarrow$ ": Let  $(x_n)$  be a sequence of distinct points in  $S$  such that  $x_n \rightarrow x$ . Then  $\forall r > 0$ ,  $\exists N \in \mathbb{N}$  such that  $\forall n > N$ ,  $d(x_n, x) < r$ . Since  $(x_n)$  are distinct points in  $S$ , the neighborhood  $B(x, r)$  contains at least one point of  $S$  different from  $x$ . Therefore,  $x$  is an accumulation point of  $S$ .  $\square$

---

In the following, let  $S'$  be the set of accumulation points of  $S$ .

**Proposition 1.8.2** Let  $X$  be a topological space and let  $S \subseteq X$ . Then  $S \cup S' = \overline{S}$ .

*Proof.* " $\subseteq$ ": Let  $x \in S \cup S'$ . If  $x \in S$ , then clearly  $x \in \overline{S}$ . If  $x \in S'$ , then every neighborhood of  $x$  contains a point of  $S$  different from  $x$ , hence every neighborhood of  $x$  intersects  $S$ , so  $x \in \overline{S}$ . Therefore,  $S \cup S' \subseteq \overline{S}$ .

" $\supseteq$ ": Let  $x \in \overline{S}$ . If  $x \in S$ , then  $x \in S \cup S'$ . If  $x \notin S$ , then every neighborhood of  $x$  contains a point of  $S$ , which must be different from  $x$ , hence  $x$  is an accumulation point of  $S$ , so  $x \in S'$ . Therefore,  $x \in S \cup S'$ .  $\square$

**Corollary 1.8.2**  $S$  is closed iff  $S' \subseteq S$ .

**Definition 1.8.3** Let  $X$  be a topological space. A subset  $S \subseteq X$  is perfect if  $S' = S$ .

**Example 1.8.1** The closed interval  $[a, b]$  is perfect in  $\mathbb{R}$ .  $\mathbb{Q}$  is perfect in  $\mathbb{Q}$  but not perfect in  $\mathbb{R}$ .  $\mathbb{N}$  is not perfect in the discrete topology.

**Theorem 1.8.4** Every perfect, complete metric space is uncountable.

*Proof.* Suppose not. Let  $M$  be a perfect, complete metric space that is countable. Since  $M$  is a metric space and  $M$  consists of accumulation points, it has to be infinite, hence countably infinite, so we can write  $M = \{x_1, x_2, \dots\}$ .

Define the closed ball

$$\overline{B}(x, r) = \{y \in M \mid d(x, y) \leq r\}.$$

Choose any  $y_1 \in M$  with  $y_1 \neq x_1$  and choose  $r_1 > 0$  such that  $x_1 \notin \overline{B}(y_1, r_1) =: Y_1$ . WLOG assume  $r_1 < 1$ . Since  $y_1$  is an accumulation point of  $M$ ,  $\exists y_2 \in M$  such that  $y_2 \neq x_2$  and  $y_2 \in B(y_1, r_1)$ . Choose  $r_2 > 0$  such that  $x_2 \notin \overline{B}(y_2, r_2) =: Y_2$ . WLOG assume  $r_2 < 1/2$  and  $Y_2 \subseteq Y_1$ . Continuing this process, we obtain a nested sequence of closed balls  $Y_1 \supset Y_2 \supset \dots$  such that  $x_1, x_2, \dots, x_n \notin Y_n$  and  $r_n < 1/n$ . Since the centers  $y_n$  form a Cauchy sequence, by completeness,  $y_n \rightarrow y$  for some  $y \in M$ . Since  $y \in Y_n$  for all  $n$ ,  $y \neq x_n$  for all  $n$ . Therefore,  $y \in M$  but  $y \notin \{x_1, x_2, \dots\}$ , contradicting the fact that  $M = \{x_1, x_2, \dots\}$ . Therefore,  $M$  is uncountable.  $\square$

**Corollary 1.8.5**  $\mathbb{R}$  and  $[a, b]$  are uncountable.

**Corollary 1.8.6** Each  $r$ -neighborhood of a perfect complete metric space is uncountable.

*Proof.* The closure of an open ball  $\overline{B(x, r/2)}$  is a perfect complete metric space, hence uncountable. Since  $\overline{B(x, r/2)} \subseteq B(x, r)$ ,  $B(x, r)$  is uncountable.  $\square$

## CHAPTER 2

# Compactness

A set  $K \subseteq \mathbb{R}^n$  is compact if it is closed and bounded. We wish to generalize this definition to arbitrary topological spaces.

**Definition 2.0.1** In a topological space  $X$ , a subset  $K \subseteq X$  is sequentially compact if every sequence in  $K$  has a subsequence that converges to a point in  $K$ .

**Proposition 2.0.1** A sequentially compact subset of a metric space is closed and bounded.

*Proof.* Since every convergent sequence has its limit in  $K$ ,  $K$  is closed.

Suppose that  $K$  is unbounded.  $\forall n \in \mathbb{N}$ ,  $\exists x_n, y_n \in K$  such that  $d(x_n, y_n) \geq n$ . By sequential compactness, exists convergent subsequences  $(x_{n_k})$  and  $(y_{n_k})$  that converge to  $x, y \in K$  respectively, then

$$n_k \leq d(x_{n_k}, y_{n_k}) \leq d(x_{n_k}, x) + d(x, y) + d(y, y_{n_k}).$$

As  $k \rightarrow \infty$ , the right hand side approaches  $d(x, y)$ , which is finite, but the left hand side approaches infinity, contradiction. Therefore,  $K$  is bounded.  $\square$

**Remark** The converse holds in  $\mathbb{R}^n$ , but not in general metric spaces. Consider

$$\begin{aligned} C([0, 1], \mathbb{R}) &= \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ is continuous}\}, \\ d(f, g) &= \max \{|f(x) - g(x)| \mid x \in [0, 1]\}. \end{aligned}$$

A closed unit ball is not sequentially compact. Consider the sequence of functions  $f_n(x) = x^n$ . It is clear that  $(f_n)$  is in the closed unit ball, but it does not have a convergent subsequence in  $C([0, 1], \mathbb{R})$ .

In a general topological space, it is sometimes inconvenient to use sequences to describe compactness.

---

**Definition 2.0.2** An open cover of a topological space  $X$  is a collection  $\mathcal{U}$  of open sets in  $X$  whose union contains  $X$ . A subcover of an open cover  $\mathcal{U}$  is a subcollection of  $\mathcal{U}$  that still covers  $X$ . A topological space  $X$  is compact if every open cover of  $X$  has a finite subcover.

**Example 2.0.1**  $(0, 1]$  is not compact in  $\mathbb{R}$ . Consider the open cover  $\mathcal{U} = \left\{ \left( \frac{1}{n}, 2 \right) \mid n \in \mathbb{N} \right\}$ , which does not have a finite subcover.

In general, compactness and sequential compactness are not equivalent. However, they are equivalent in metric spaces.

We first introduce some auxiliary definitions.

**Definition 2.0.3** A Lebesgue number for an open cover  $\mathcal{U}$  of a subset  $A$  of a metric space  $M$  is a positive real number  $\lambda$  such that for every  $a \in A$ ,  $\exists U \in \mathcal{U}$  such that  $B(a, \lambda) \subseteq U$ .

**Lemma 2.0.4** (Lebesgue number lemma) Every open cover of a sequentially compact set (in a metric space) has a positive Lebesgue number.

*Proof.* Suppose otherwise, so that there exists some open cover  $\mathcal{U}$  of a sequentially compact set  $K$  such that  $\forall \lambda > 0$ ,  $\exists k \in K$  such that no  $U \in \mathcal{U}$  contains  $B(k, \lambda)$ .

Take  $\lambda = 1/n$  and let  $k_n \in K$  be the corresponding point such that no  $U \in \mathcal{U}$  contains  $B(k_n, 1/n)$ . By sequential compactness,  $(k_n)$  has a convergent subsequence  $(k_{n_j})$  that converges to some  $x \in K$ . Since  $\mathcal{U}$  is an open cover of  $K$ ,  $\exists U \in \mathcal{U}$  such that  $x \in U$ . Since  $U$  is open,  $\exists r > 0$  such that  $B(x, r) \subseteq U$ . Pick  $j$  large enough such that  $1/n_j < r/2$  so that

$$d(k_{n_j}, x) < r/2 \rightsquigarrow B(k_{n_j}, 1/n_j) \subseteq B(x, r) \subseteq U,$$

contradicting the assumption that no  $U \in \mathcal{U}$  contains  $B(k_{n_j}, 1/n_j)$ . Therefore,  $\mathcal{U}$  has a positive Lebesgue number.  $\square$

If  $A$  is noncompact then it may have open covers with no positive Lebesgue number. For example, consider the set  $A = (0, 1) \subseteq \mathbb{R}$  and the open cover  $\mathcal{U} = \{A\}$ . It is clear that we cannot find  $\lambda > 0$  such that  $(a - \lambda, a + \lambda) \subseteq A$  for all  $a \in A$ .

**Theorem 2.0.5** Let  $M$  be a metric space. A subset  $K \subseteq M$  is compact iff it is sequentially compact.

*Proof.* " $\Rightarrow$ ": If  $K$  is finite then it is trivially sequentially compact (since any sequence in  $K$  must contain some element infinitely many times).

Suppose that  $K$  is infinite. Assume that  $K$  is compact but not sequentially compact, then  $\exists$  a sequence  $(x_n)$  in  $K$  that does not have a convergent subsequence. For each  $k \in K$ ,  $\exists r_k > 0$  such that  $B(k, r_k)$  contains only finitely many terms of  $(x_n)$ . Otherwise, if every neighborhood

---

of  $k$  contains infinitely many terms of  $(x_n)$ , then we can construct a subsequence of  $(x_n)$  that converges to  $k$ . The collection  $\{B(k, r_k) \mid k \in K\}$  is an open cover of  $K$ , hence it has a finite subcover

$$\{B(k_1, r_{k_1}), \dots, B(k_m, r_{k_m})\}.$$

Each  $x_n$  appears in each  $B(k_i, r_{k_i})$  only finitely many times, so  $(x_n)$  can only have finitely many values, but any such sequence must have a constant subsequence, which is convergent, contradiction. Therefore,  $K$  is sequentially compact.

” $\Leftarrow$ ”: Let  $\mathcal{U}$  be an open cover of  $K$ . By the Lebesgue number lemma,  $\mathcal{U}$  has a positive Lebesgue number  $\lambda$ . Choose  $x_1 \in K$ , then  $\exists \mathcal{U}_1 \in \mathcal{U}$  such that  $B(x_1, \lambda) \subseteq \mathcal{U}_1$ . If  $K \subseteq \mathcal{U}_1$ , then we are done. Otherwise, for any point in  $K$  not covered, we can again find  $x_2 \in K$  such that  $x_2 \notin \mathcal{U}_1$ , then  $\exists \mathcal{U}_2 \in \mathcal{U}$  such that  $B(x_2, \lambda) \subseteq \mathcal{U}_2$ . If  $K \subseteq \mathcal{U}_1 \cup \mathcal{U}_2$ , then we are done. Otherwise, we can continue this process to find  $x_3, x_4, \dots$ . If the process continues indefinitely, then we have constructed a sequence  $(x_n)$  in  $K$  such that  $x_m \notin \mathcal{U}_n$  for all  $m > n$ . By sequential compactness,  $(x_n)$  has a convergent subsequence  $(x_{n_k})$  that converges to some  $x \in K$ . Then there exists a large  $k$  such that  $d(x_{n_k}, x) < \lambda$ , which implies that  $x \in B(x_{n_k}, \lambda) \subseteq \mathcal{U}_{n_k}$ . But by construction,  $x_{n_m} \notin \mathcal{U}_{n_k}$  for all  $m > k$ , contradicting the fact that  $x_{n_m} \rightarrow x$ . Therefore, the process must terminate within finitely many steps, and we have found a finite subcover of  $\mathcal{U}$ . Hence,  $K$  is compact.  $\square$

**Corollary 2.0.6** (Heine-Borel Theorem) A subset of  $\mathbb{R}^n$  is compact iff it is closed and bounded.

**Definition 2.0.7** A metric space is totally bounded if for every  $\epsilon > 0$ , the space can be covered by finitely many balls of radius  $\epsilon$ .

**Theorem 2.0.8** A subset of a complete metric space is compact iff it is closed and totally bounded.

## 2.1 Continuity and Compactness

**Proposition 2.1.1** Let  $X, Y$  be topological spaces and let  $f : X \rightarrow Y$  be continuous. If  $X$  is compact, then  $f(X)$  is compact.

*Proof.* Let  $\mathcal{U}$  be an open cover of  $f(X)$ , then  $\mathcal{V} = \{f^{-1}(U) \mid U \in \mathcal{U}\}$  is an open cover of  $X$ . Since  $X$  is compact,  $\exists U_1, \dots, U_n \in \mathcal{U}$  such that

$$X \subseteq f^{-1}(U_1) \cup \dots \cup f^{-1}(U_n) \implies f(X) \subseteq U_1 \cup \dots \cup U_n,$$

hence  $f(X)$  is compact.  $\square$

---

**Corollary 2.1.1** (Extreme Value Theorem) Let  $X$  be a metric space. If  $X$  is compact and if  $f : X \rightarrow \mathbb{R}$  is continuous, then  $f$  is bounded and attains its maximum and minimum.

*Proof.* Since  $f(X)$  is compact in  $\mathbb{R}$ , it is closed and bounded, hence  $f$  is bounded, and since  $f(X)$  is closed,  $\sup f(X), \inf f(X) \in f(X)$ , so  $f$  attains its maximum and minimum.  $\square$

**Corollary 2.1.2** If  $X$  is compact and if  $X$  is homeomorphic to  $Y$ , then  $Y$  is compact.

**Corollary 2.1.3**  $[0, 1]$  and  $\mathbb{R}$  are not homeomorphic.

**Proposition 2.1.2** Every continuous function defined on a compact metric space is uniformly continuous.

*Proof.* Suppose not. Let  $f : M \rightarrow N$  be continuous, where  $M$  is compact, but  $f$  is not uniformly continuous. Then  $\exists \epsilon > 0$  such that  $\forall \delta > 0$ ,  $\exists x, y \in M$  such that  $d(x, y) < \delta$  but  $d(f(x), f(y)) \geq \epsilon$ . Take  $\delta = 1/n$  and let  $x_n, y_n \in M$  be the corresponding points such that  $d(x_n, y_n) < 1/n$  but  $d(f(x_n), f(y_n)) \geq \epsilon$ . By compactness,  $(x_n)$  has a convergent subsequence  $(x_{n_k})$  that converges to some  $x \in M$ . Since  $d(x_{n_k}, y_{n_k}) < 1/n_k$ , we have  $y_{n_k} \rightarrow x$  as well. Since  $f$  is continuous,  $f(x_{n_k}) \rightarrow f(x)$  and  $f(y_{n_k}) \rightarrow f(x)$ , contradicting the fact that  $d(f(x_{n_k}), f(y_{n_k})) \geq \epsilon$ . Therefore,  $f$  is uniformly continuous.  $\square$

## 2.2 Subspaces, Products, and Hausdorff Spaces

**Proposition 2.2.1** If  $A$  is compact in  $X$  and if  $B$  is compact in  $Y$ , then  $A \times B$  is compact in  $X \times Y$ .

*Proof.* Let  $\mathcal{U}$  be an open cover of  $A \times B$ . For each  $a \in A$ ,  $\{a\} \times B$  is homeomorphic to  $B$ , hence it is compact. Therefore,  $\exists U_1, U_2, \dots, U_n \in \mathcal{U}$  such that

$$\{a\} \times B \subseteq U_1 \cup U_2 \cup \dots \cup U_n.$$

For each  $b \in B$ , we can find a product open set  $V \times W \subseteq X \times Y$  containing  $(a, b)$ . These  $V \times W$  form an open cover for  $\{a\} \times B$ . We can extract  $V_1 \times W_1, \dots, V_m \times W_m$  such that they cover  $\{a\} \times B$ .

Define  $Z_a = V_1 \cap \dots \cap V_m$ , then  $Z_a \times B \subseteq U_1 \cup \dots \cup U_n$ . We have shown that for any  $a \in A$ ,  $\exists$  open cover  $Z_a$  such that  $Z_a \times B$  is covered by finitely many open sets in  $\mathcal{U}$ .

Since  $\{Z_a\}_{a \in A}$  is an open cover of  $A$  in  $X$  (each  $Z_a$  is open and contains  $a$ ), and  $A$  is compact, there exist finitely many points  $a_1, \dots, a_k \in A$  such that  $A \subseteq Z_{a_1} \cup \dots \cup Z_{a_k}$ . Then

$$A \times B \subseteq (Z_{a_1} \times B) \cup \dots \cup (Z_{a_k} \times B)$$

---

and each  $Z_{a_i} \times B$  is covered by finitely many sets from  $\mathcal{U}$ . Therefore,  $A \times B$  is covered by finitely many sets from  $\mathcal{U}$ , so  $A \times B$  is compact.  $\square$

**Proposition 2.2.2** Every closed subset of a compact space is compact.

*Proof.* Let  $\mathcal{U}$  be an open cover for  $E$ , then  $\mathcal{U} \cup \{X \setminus E\}$  is an open cover for  $X$ . Since  $X$  is compact, we can extract a finite subcover. If  $X \setminus E$  is in the subcover, we can discard it since it does not cover any point in  $E$ . Therefore,  $E$  is compact.  $\square$

**Proposition 2.2.3** Every compact subset of a Hausdorff space is closed.

*Proof.* Let  $X$  be a Hausdorff space.

**Claim** If  $A, B \subseteq X$  are disjoint compact subsets, then  $\exists$  disjoint open sets  $U, V \subseteq X$  such that  $A \subseteq U$  and  $B \subseteq V$ .

*Proof.* Suppose that  $B = \{b\}$  is a singleton.  $\forall a \in A, \exists$  disjoint open sets  $\mathcal{U}_a$  and  $\mathcal{V}_a$  such that  $a \in \mathcal{U}_a$  and  $b \in \mathcal{V}_a$ .  $\{\mathcal{U}_a \mid a \in A\}$  is an open cover for  $A$ , hence we can extract a finite subcover  $\mathcal{U}_{a_1}, \dots, \mathcal{U}_{a_n}$ . Define

$$\widetilde{U} = \bigcup_{i=1}^n \mathcal{U}_{a_i}, \quad \widetilde{V} = \bigcap_{i=1}^n \mathcal{V}_{a_i},$$

then  $\widetilde{U}$  and  $\widetilde{V}$  are disjoint open sets such that  $A \subseteq \widetilde{U}$  and  $b \in \widetilde{V}$ .

For general  $B$ , we have shown that  $\forall b \in B, \exists$  disjoint open sets  $\widetilde{U}_b, \widetilde{V}_b$  such that  $A \subseteq \widetilde{U}_b$  and  $b \in \widetilde{V}_b$ .  $\{\widetilde{V}_b \mid b \in B\}$  is an open cover for  $B$ , hence we can extract a finite subcover  $\widetilde{V}_{b_1}, \dots, \widetilde{V}_{b_m}$ . Define

$$\widetilde{U} = \bigcap_{j=1}^m \widetilde{U}_{b_j}, \quad \widetilde{V} = \bigcup_{j=1}^m \widetilde{V}_{b_j},$$

then  $\widetilde{U}$  and  $\widetilde{V}$  are disjoint open sets such that  $A \subseteq \widetilde{U}$  and  $B \subseteq \widetilde{V}$ .  $\square$

Now, let  $K \subseteq X$  be compact. We need to show that  $X \setminus K$  is open. For any  $x \in X \setminus K$ ,  $\{x\}$  and  $K$  are disjoint compact subsets, hence by the claim,  $\exists$  disjoint open sets  $U, V \subseteq X$  such that  $\{x\} \subseteq U$  and  $K \subseteq V$ . In other words, for all points  $x \in X \setminus K$ , we can find an open neighborhood  $U$  of  $x$  that is contained in  $X \setminus K$ . Therefore,  $X \setminus K$  is open, and  $K$  is closed.  $\square$

## 2.3 Finite Intersection Property

**Definition 2.3.1** A collection  $\mathcal{C}$  of subsets of  $X$  is said to have the finite intersection property if the intersection of every finite subcollection of  $\mathcal{C}$  is nonempty.

---

**Theorem 2.3.2** Let  $X$  be a topological space. Then  $X$  is compact iff every collection  $\mathcal{C}$  of nonempty closed sets in  $X$  with the finite intersection property has a nonempty intersection.

*Proof.* Given a collection  $\mathcal{U}$  of subsets of  $X$ , define

$$\mathcal{C} = \{X \setminus U \mid U \in \mathcal{U}\}.$$

The following statements hold:

1.  $\mathcal{U}$  is a collection of open sets  $\iff \mathcal{C}$  is a collection of closed sets.
2. The collection  $\mathcal{U}$  covers  $X \iff \bigcap_{C \in \mathcal{C}} C = \emptyset$ .
3. The finite subcollection  $\{U_1, \dots, U_n\}$  of  $\mathcal{U}$  covers  $X \iff \bigcap_{i=1}^n (X \setminus U_i) = \emptyset$ .

Now, if  $X$  is compact, we have the contrapositive statement that given any collection  $\mathcal{U}$  of open sets, if no finite subcollection of  $\mathcal{U}$  covers  $X$ , then  $\mathcal{U}$  does not cover  $X$ . This is directly equivalent to the statement we wish to prove by the above observations.  $\square$

**Corollary 2.3.3** Let  $X$  be a compact topological space. Suppose that  $(F_n)$  is a sequence of nonempty closed subsets of  $X$  that are nested decreasing, i.e.  $F_{n+1} \subseteq F_n$  for all  $n$ . Then  $\bigcap_{n=1}^{\infty} F_n$  is nonempty.

**Definition 2.3.4** Let  $M$  be a metric space. For a subset  $A \subseteq M$ , the diameter of  $A$  is defined to be

$$\sup \{d(x, y) \mid x, y \in A\}.$$

**Proposition 2.3.1** If in addition to the conditions in the previous corollary,

$$\lim_{n \rightarrow \infty} \text{diam}(F_n) = 0,$$

then  $\bigcap_{n=1}^{\infty} F_n$  is a singleton.

*Proof.* Write  $F = \bigcap_{n=1}^{\infty} F_n$ . We have shown that  $F$  is nonempty. Since  $F \subseteq F_n$  for all  $n$ , we have  $\text{diam}(F) = 0$ . If  $F$  contains two distinct points  $x, y$ , then  $d(x, y) > 0$ . Therefore,  $F$  must contain exactly one point.  $\square$

## 2.4 Tychonoff Theorem

**Lemma 2.4.1** Let  $X$  be a set, and let  $\mathcal{A}$  be a collection of subsets of  $X$  that has the FIP. Then  $\exists$  a collection  $\mathcal{D}$  of subsets of  $X$  such that  $\mathcal{D}$  contains  $\mathcal{A}$ ,  $\mathcal{D}$  has the FIP, and  $\mathcal{D}$  is maximal with respect to inclusion.

---

*Proof.* We will construct  $\mathcal{D}$  using Zorn's lemma. Let

$$\mathbb{A} = \{\mathcal{B} \mid \mathcal{B} \text{ is a collection of subsets of } X \text{ that has the FIP and } \mathcal{A} \subseteq \mathcal{B}\}.$$

We need to show that every chain in  $\mathbb{A}$  has an upper bound in  $\mathbb{A}$ . Let  $\mathbb{B}$  be a chain in  $\mathbb{A}$ . Define

$$\mathcal{C} = \bigcup_{\mathcal{B} \in \mathbb{B}} \mathcal{B}.$$

Clearly  $\mathcal{C}$  is an upper bound for  $\mathbb{B}$  in  $\mathbb{A}$ . We need to show that  $\mathcal{C} \in \mathbb{A}$ . It is clear that  $\mathcal{A} \subseteq \mathcal{C}$ , hence it is left to show that  $\mathcal{C}$  has the FIP. Given any finite subcollection  $\mathcal{C}_1, \dots, \mathcal{C}_n$  of  $\mathcal{C}$ ,  $\exists \mathcal{B}_1, \dots, \mathcal{B}_n \in \mathbb{B}$  such that  $\mathcal{C}_i \in \mathcal{B}_i$ . Since  $\mathbb{B}$  is a chain, there exists an element  $\mathcal{B}_k$  that contains all  $\mathcal{B}_i$ 's, hence  $\mathcal{C}_1, \dots, \mathcal{C}_n \in \mathcal{B}_k$ . Since  $\mathcal{B}_k$  has the FIP,  $\bigcap_{i=1}^n \mathcal{C}_i \neq \emptyset$ . Therefore,  $\mathcal{C}$  has the FIP, and  $\mathcal{C} \in \mathbb{A}$ . By Zorn's lemma,  $\mathbb{A}$  has a maximal element  $\mathcal{D}$ , which is what we want.  $\square$

**Lemma 2.4.2** Let  $X$  be a set, and let  $\mathcal{D}$  be a collection of subsets of  $X$  that has the FIP and is maximal with respect to inclusion. Then any finite intersection of elements of  $\mathcal{D}$  is in  $\mathcal{D}$ . Moreover, if  $A \subseteq X$  intersects every element of  $\mathcal{D}$ , then  $A \in \mathcal{D}$ .

*Proof.* Let  $B$  be an intersection of finitely many elements of  $\mathcal{D}$ . Since  $\mathcal{D}$  has the FIP,  $B \neq \emptyset$ . Define  $\mathcal{E} = \mathcal{D} \cup \{B\}$ . Pick finitely elements of  $\mathcal{E}$ . If none of them is  $B$ , then their intersection is nonempty since  $\mathcal{D}$  has the FIP. If one of them is  $B$ , then the intersection is also nonempty since  $B$  is an intersection of finitely many elements of  $\mathcal{D}$  and  $\mathcal{D}$  has the FIP. If  $B \notin \mathcal{D}$ , then  $\mathcal{E}$  has the FIP and strictly contains  $\mathcal{D}$ , contradicting the maximality of  $\mathcal{D}$ . Therefore,  $B \in \mathcal{D}$ .

Now, let  $A \subseteq X$  be such that  $A$  intersects every element of  $\mathcal{D}$ . Define  $\mathcal{F} = \mathcal{D} \cup \{A\}$ . Pick finitely elements of  $\mathcal{F}$ . If none of them is  $A$ , then their intersection is nonempty since  $\mathcal{D}$  has the FIP. If one of them is  $A$ , then the intersection is also nonempty since  $A$  intersects every element of  $\mathcal{D}$  and  $\mathcal{D}$  has the FIP. If  $A \notin \mathcal{D}$ , then  $\mathcal{F}$  has the FIP and strictly contains  $\mathcal{D}$ , contradicting the maximality of  $\mathcal{D}$ . Therefore,  $A \in \mathcal{D}$ .  $\square$

**Theorem 2.4.3** (Tychonoff) An arbitrary product of compact topological spaces is compact in the product topology.

*Proof.* Let  $X = \prod_{\alpha \in I} X_\alpha$ , where each  $X_\alpha$  is compact.

Let  $\mathcal{A}$  be a collection of subsets of  $X$  that has the FIP. We will prove that

$$\bigcap_{A \in \mathcal{A}} \overline{A} \neq \emptyset.$$

The result then follows from Theorem 2.3.2.

By 2.4.1,  $\exists$  a collection  $\mathcal{D} \supset \mathcal{A}$  maximal with respect to the FIP. It suffices to show that

---

$\bigcap_{D \in \mathcal{D}} \overline{D} \neq \emptyset$ . Given  $\alpha \in I$ , let  $\pi_\alpha : X \rightarrow X_\alpha$  be the projection map. Consider

$$\{\pi_\alpha(D) \mid D \in \mathcal{D}\},$$

which has the FIP since  $\mathcal{D}$  has the FIP. Since  $X_\alpha$  is compact, by Theorem 2.3.2, we have

$$\bigcap_{D \in \mathcal{D}} \overline{\pi_\alpha(D)} \neq \emptyset.$$

Pick an element  $x_\alpha$  in the intersection. Define  $x = (x_\alpha)_{\alpha \in I} \in X$ . We will show that  $x \in \overline{D}$  for all  $D \in \mathcal{D}$ , hence  $x \in \bigcap_{D \in \mathcal{D}} \overline{D}$ , implying that  $\bigcap_{D \in \mathcal{D}} \overline{D}$  is nonempty.

Let  $\beta \in I$  and let  $U_\beta$  be an open neighborhood of  $x_\beta$  in  $X_\beta$ . Then  $\pi_\beta^{-1}(U_\beta)$  is an open neighborhood of  $x$  in  $X$ . We claim that  $\pi_\beta^{-1}(U_\beta)$  intersects every element of  $\mathcal{D}$ . Given any  $D \in \mathcal{D}$ , by definition of  $x$ ,

$$\begin{aligned} x_\beta \in \overline{\pi_\beta(D)} &\implies U_\beta \text{ intersects } \pi_\beta(D) \text{ in some point } \pi_\beta(y) \\ &\implies y \in D \cap \pi_\beta^{-1}(U_\beta) \\ &\implies D \cap \pi_\beta^{-1}(U_\beta) \neq \emptyset. \end{aligned}$$

By Lemma 2.4.2, every set of the form  $\pi_\beta^{-1}(U_\beta)$  that contains  $x$  is in  $\mathcal{D}$ .

Recall that every basis element of the product topology is a finite intersection of sets of the form  $\pi_\beta^{-1}(U_\beta)$ . By Lemma 2.4.2, every basis element that contains  $x$  is in  $\mathcal{D}$ . Therefore, every open neighborhood of  $x$  intersects every element of  $\mathcal{D}$ , hence  $x \in \overline{D}$  for all  $D \in \mathcal{D}$ . This completes the proof.  $\square$

## 2.5 Cantor Sets

The standard middle-third Cantor set is constructed as follows. Start with the closed interval  $C_0 = [0, 1]$ . Remove the open middle third  $(1/3, 2/3)$  to obtain  $C_1 = [0, 1/3] \cup [2/3, 1]$ . Remove the open middle third of each remaining interval to obtain  $C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$ . Continuing this process indefinitely, we obtain a nested sequence of closed sets  $C_0 \supseteq C_1 \supseteq C_2 \supseteq \dots$ . The Cantor set  $C$  is defined to be the intersection of all  $C_n$ :

$$C = \bigcap_{n=0}^{\infty} C_n.$$

Note that  $C$  is nonempty since it contains the endpoints of the intervals in each  $C_n$ .

**Definition 2.5.1** A metric space  $M$  is totally disconnected if for all  $x \in M$  and  $\forall \epsilon > 0$ ,  $\exists$  a clopen set  $U$  such that  $x \in U \subseteq B(x, \epsilon)$ .

---

**Example 2.5.1** Any discrete metric space is totally disconnected.  $\mathbb{Q}$  is totally disconnected.

**Remark** The usual definition of totally disconnected is that the only connected subsets are singletons. The two definitions are equivalent for compact metric spaces.

**Theorem 2.5.2** The Cantor set  $C$  is a compact, nonempty, perfect, totally disconnected metric space.

*Proof.* The metric on  $C$  is inherited from  $\mathbb{R}$ . Since  $C$  is a closed subset of the compact set  $[0, 1]$ ,  $C$  is compact. We have already shown that  $C$  is nonempty.

Pick  $x \in C$  and  $\epsilon > 0$ . Choose  $n$  such that  $1/3^n < \epsilon$ . Then  $x$  is contained in one of the  $2^n$  closed intervals  $I$  of length  $1/3^n$  that make up  $C_n$ . Let

$$E = \left\{ 0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \dots \right\}$$

be the set of all endpoints of  $C$ . Since  $E \cap I$  is an infinite set, and  $(x - \epsilon, x + \epsilon)$  contains a point distinct from  $x$  in  $E \cap I$ ,  $x$  is an accumulation point of  $C$ . Since  $x$  is arbitrary,  $C$  is perfect.

Let  $x \in C$  and  $\epsilon > 0$ . Choose  $n$  such that  $1/3^n < \epsilon$ . Then  $x$  is contained in one of the  $2^n$  closed intervals  $I$  of length  $1/3^n$  that make up  $C_n$ . Note that both  $I$  and its complement  $J = C_n \setminus I$  are closed in  $C_n$ , hence they are clopen in  $C_n$ . Since  $C \subseteq C_n$ ,  $I \cap C$  and  $J \cap C$  are clopen in  $C$ . Therefore, we have  $x \in I \subseteq B(x, \epsilon)$ , where  $I$  is clopen in  $C$ . Since  $x$  and  $\epsilon$  are arbitrary,  $C$  is totally disconnected.  $\square$

**Corollary 2.5.3** Cantor set  $C$  is uncountable.

*Proof.* Since compact spaces are complete, by 1.8.4,  $C$  is uncountable.  $\square$

**Remark** In fact, one can show that every compact, nonempty, perfect, totally disconnected metric space is homeomorphic to the Cantor set. If  $M$  is a compact metric space then  $\exists$  a continuous surjection  $f : C \rightarrow M$ . This implies that every compact metric space has cardinality at most that of  $\mathbb{R}$ .

We can also show that the Cantor set is uncountable by constructing a bijection between  $C$  and the set of all sequences of 0's and 2's. Consider the map  $f : C \rightarrow \{0, 2\}^{\mathbb{N}}$  defined as follows. For each  $x \in C$ , we can write  $x$  in base 3:

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}, \quad a_n \in \{0, 1, 2\}.$$

If  $x \in C$ , then  $a_n \neq 1$  for all  $n$ . Clearly  $f$  is injective. To see that  $f$  is surjective, given any sequence  $(b_n) \in \{0, 2\}^{\mathbb{N}}$ , we can construct  $x \in C$  by going through the construction of  $C$ . At

---

the  $n$ -th step, if  $b_n = 0$ , we choose the left interval; if  $b_n = 2$ , we choose the right interval. The intersection of all these intervals is a singleton  $\{x\}$  by 2.3.1, thus  $f(x) = (b_n)$ . Since  $\{0, 2\}^{\mathbb{N}}$  is uncountable,  $C$  is uncountable as well.

The Cantor set is large in terms of cardinality, but it is small in terms of topology and geometry.

**Definition 2.5.4** Let  $X$  be a topological space. A set  $S \subseteq X$  is dense in  $X$  if  $\overline{S} = X$ .  $S$  is somewhere dense in  $X$  if there exists a nonempty open set  $U \subseteq X$  such that  $\overline{S \cap U} = U$ .  $S$  is nowhere dense in  $X$  if it is not somewhere dense.

A nowhere dense set is "small" in terms of topology.

**Proposition 2.5.1** The Cantor set  $C$  contains no interval and is nowhere dense in  $\mathbb{R}$ .

*Proof.* Suppose that  $C$  contains an interval  $(a, b)$ . Then  $\exists n$  such that  $1/3^n < b - a$ . But  $C_n$  is the disjoint union of closed intervals of length  $1/3^n$ , so  $(a, b)$  cannot be contained in  $C_n$ , contradicting the fact that  $C \subseteq C_n$ . Therefore,  $C$  contains no interval.

Suppose that  $C$  is dense in some open set  $U$ , then

$$C = \overline{C} \supset \overline{C \cap U} \supset U$$

which contains an open interval, contradicting the fact that  $C$  contains no interval. Therefore,  $C$  is nowhere dense in  $\mathbb{R}$ .  $\square$

We will show that  $C$  has measure 0. A set with measure 0 is "small" in terms of geometry.

**Definition 2.5.5** A set  $Z \subseteq \mathbb{R}$  has measure zero if for every  $\epsilon > 0$ ,  $\exists$  a countable cover of  $Z$  by open intervals  $\{(a_k, b_k)\}_{k \in \mathbb{N}}$  such that the total length of the cover

$$\sum_{k=1}^{\infty} (b_k - a_k) \leq \epsilon.$$

**Proposition 2.5.2** The following sets have measure zero:

- Any finite set.
- Any subset of a measure zero set.
- Any countable union of measure zero sets.

*Proof.* Let  $Z_1, Z_2, \dots$  be a sequence of measure zero sets and  $Z = \bigcup_{n=1}^{\infty} Z_n$ . Given  $\epsilon > 0$ ,

---

for each  $n$ ,  $\exists$  a countable cover of  $Z_n$  by open intervals  $\{(a_k^{(n)}, b_k^{(n)})\}_{k \in \mathbb{N}}$  such that

$$\sum_{k=1}^{\infty} (b_k^{(n)} - a_k^{(n)}) \leq \frac{\epsilon}{2^n}.$$

Then  $\{(a_k^{(n)}, b_k^{(n)}) \mid n, k \in \mathbb{N}\}$  is a countable cover of  $Z$  by open intervals such that

$$\sum_{n=1}^{\infty} \sum_{k=1}^{\infty} (b_k^{(n)} - a_k^{(n)}) \leq \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon.$$

Therefore,  $Z$  has measure zero.  $\square$

- Any countable set.
- The middle-third Cantor set.

*Proof.* Let  $\epsilon > 0$ . Choose  $n$  such that  $2^n/3^n < \epsilon$ . Then  $C \subseteq C_n$ , and  $C_n$  is the disjoint union of  $2^n$  closed intervals of length  $1/3^n$ . Cover each closed interval by an open interval of length  $1/3^n + \delta$ , where  $\delta > 0$  is chosen such that

$$2^n \delta < \epsilon - \frac{2^n}{3^n}.$$

Then we have found a cover of  $C$  by  $2^n$  open intervals such that the total length of the cover is

$$2^n \left( \frac{1}{3^n} + \delta \right) = \frac{2^n}{3^n} + 2^n \delta < \epsilon.$$

Therefore,  $C$  has measure zero.  $\square$

**Remark** It is nontrivial to prove that a nonempty open interval is not of measure zero.

It is also possible to define a Cantor set that is not of measure zero, which is the *Fat Cantor set*: in each step, we discard the middle  $1/n!$  portion of each remaining interval at the  $n$ -th step. The resulting Cantor set is still compact, nonempty, perfect, and totally disconnected, but it has positive measure. This also implies that measure is not a topological property, since the Fat Cantor set is homeomorphic to the standard Cantor set, but they have different measures.

## CHAPTER 3

# Integration Theory

### 3.1 Riemann Integrals on $\mathbb{R}$

Let  $f : [a, b] \rightarrow \mathbb{R}$  be a function. Intuitively, the integral of  $f$  is the area under the curve  $y = f(x)$ .

A partition pair consists of two finite sets of points in  $P, T \subseteq [a, b]$  where  $P = \{x_0, x_1, \dots, x_n\}$  and  $T = \{t_1, t_2, \dots, t_n\}$  are interlaced as

$$a = x_0 \leq t_1 \leq x_1 \leq t_2 \leq x_2 \leq \dots \leq t_n \leq x_n = b.$$

Assume that the points  $x_0, x_1, \dots, x_n$  are distinct. The Riemann sum of  $f$  with respect to the partition pair  $(P, T)$  is defined to be

$$R(f, P, T) = \sum_{i=1}^n f(t_i)(x_i - x_{i-1}).$$

The mesh of the partition  $P$  is defined to be

$$\text{mesh}(P) = \max_{1 \leq i \leq n} (x_i - x_{i-1}).$$

A real number  $I$  is the Riemann integral of  $f$  over  $[a, b]$  if for every  $\epsilon > 0$ ,  $\exists \delta > 0$  such that for every partition pair  $(P, T)$  with  $\text{mesh}(P) < \delta$ , we have

$$|R(f, P, T) - I| < \epsilon.$$

If such an  $I$  exists, it must be unique, and we write

$$\int_a^b f(x) dx = I = \lim_{\text{mesh}(P) \rightarrow 0} R(f, P, T).$$

---

In this case, we say that  $f$  is Riemann integrable on  $[a, b]$  with Riemann integral  $I$ .

- If  $f$  is Riemann integrable on  $[a, b]$ , then  $f$  is bounded on  $[a, b]$ .
- $f \mapsto \int_a^b f(x) dx$  is linear.
- If  $f, g$  are Riemann integrable on  $[a, b]$  and if  $f(x) \leq g(x)$  for all  $x \in [a, b]$ , then

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

- If  $f$  is Riemann integrable on  $[a, b]$  and if  $|f(x)| \leq M$  for all  $x \in [a, b]$ , then

$$\left| \int_a^b f(x) dx \right| \leq M(b - a).$$

Let  $f : [a, b] \rightarrow [-M, M]$  be a bounded function. For a partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[a, b]$ , define

$$L(f, P) = \sum_{i=1}^n m_i(x_i - x_{i-1}), \quad U(f, P) = \sum_{i=1}^n M_i(x_i - x_{i-1}),$$

where

$$m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}, \quad M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$

For any partition pair  $(P, T)$ , we have

$$L(f, P) \leq R(f, P, T) \leq U(f, P).$$

Define the lower integral and upper integral of  $f$  over  $[a, b]$  to be

$$\underline{I} = \sup_P L(f, P), \quad \bar{I} = \inf_P U(f, P).$$

If  $\underline{I} = \bar{I}$ , then we say that  $f$  is Darboux integrable on  $[a, b]$  with Darboux integral  $I = \underline{I} = \bar{I}$ .

We wish to show that Riemann integrability is equivalent to Darboux integrability, and in this case, the two integrals are equal.

**Definition 3.1.1** A partition  $P'$  refines another partition  $P$  if  $P \subseteq P'$ .

**Lemma 3.1.2** If  $P'$  refines  $P$ , then  $L(f, P) \leq L(f, P')$  and  $U(f, P') \leq U(f, P)$ .

*Proof.* By induction, it suffices to consider  $P' = P \cup \{w\}$ , where  $w \in (x_{i-1}, x_i)$  for some  $i$ . Lower sums for  $P$  and  $P'$  are the same except that the term  $m_i(x_i - x_{i-1})$  in  $L(f, P)$  is replaced by

$$m_i(w - x_{i-1}) + m_{i+1}(x_i - w) \geq m_i(x_i - x_{i-1})$$

in  $L(f, P')$ . Therefore,  $L(f, P) \leq L(f, P')$ . The proof for upper sums is similar.  $\square$

---

If  $P, P'$  are partitions of  $[a, b]$ , then for the common refinement  $P^* = P \cup P'$ , we have

$$L(f, P) \leq L(f, P^*) \leq U(f, P^*) \leq U(f, P'),$$

hence each lower sum is a lower bound for each upper sum. Therefore, a bounded function  $f$  is Darboux integrable on  $[a, b]$  iff for every  $\epsilon > 0$ ,  $\exists$  a partition  $P$  of  $[a, b]$  such that

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

**Theorem 3.1.3** Riemann integrability is equivalent to Darboux integrability, and when  $f$  is integrable in either sense, its lower integral, upper integral, and Riemann integral are equal.

*Proof.* Let  $f : [a, b] \rightarrow \mathbb{R}$  be Riemann integrable with Riemann integral  $I$ .  $\forall \epsilon > 0$ ,  $\exists \delta > 0$  such that for every partition pair  $(P, T)$  with  $\text{mesh}(P) < \delta$ , we have

$$|R(f, P, T) - I| < \epsilon/4.$$

Fix such a partition  $P = \{x_0, x_1, \dots, x_n\}$ , and choose  $T = \{t_1, t_2, \dots, t_n\}$  such that

$$t_i \in [x_{i-1}, x_i] \quad \text{and} \quad f(t_i) - m_i < \frac{\epsilon}{4(b-a)} \quad \text{for all } 1 \leq i \leq n.$$

Then

$$\begin{aligned} R(f, P, T) - L(f, P) &= \sum_{i=1}^n (f(t_i) - m_i)(x_i - x_{i-1}) \\ &< \sum_{i=1}^n \frac{\epsilon}{4(b-a)}(x_i - x_{i-1}) = \frac{\epsilon}{4(b-a)}(b - a) = \frac{\epsilon}{4}. \end{aligned}$$

Similarly, choose  $T' = \{t'_1, t'_2, \dots, t'_n\}$  such that

$$U(f, P) - R(f, P, T') < \frac{\epsilon}{4}.$$

Therefore,

$$U(f, P) - L(f, P) < \frac{\epsilon}{2} + |R(f, P, T') - I| + |R(f, P, T) - I| < \epsilon,$$

thus  $f$  is Darboux integrable.

Conversely, let  $f : [a, b] \rightarrow [-M, M]$  be Darboux integrable.  $\forall \epsilon > 0$ ,  $\exists$  a partition  $P_1$  of  $[a, b]$  such that

$$U(f, P_1) - L(f, P_1) < \epsilon/3.$$

Fix  $\epsilon > 0$  and choose

$$\delta = \frac{\epsilon}{12n_1M},$$

where  $n_1$  is the number of points in  $P_1$ . Choose any partition  $P, T$  with  $\text{mesh}(P) < \delta$ . We

---

want to show that

$$|R(f, P, T) - I| < \epsilon.$$

Define  $P^* = P \cup P_1$ , which refines both  $P$  and  $P_1$ , then

$$U(f, P^*) - L(f, P^*) \leq U(f, P_1) - L(f, P_1) < \epsilon/3.$$

Now, we compare the upper integrals  $U(f, P^*)$  and  $U(f, P)$ . Except for at most  $2n_1$  many exceptional  $P^*$ -intervals that have endpoints in  $P_1$ , we have  $I_j^* = I_j$  and  $M_j^* = M_j$  for all other  $P^*$ -intervals. Set

$$\begin{aligned} \mathcal{I} &= \{i \mid I_i \text{ contains exceptional subintervals}\}, \\ \mathcal{J}(i) &= \{j \mid I_j^* \text{ is an exceptional subinterval of } I_i\}, \end{aligned}$$

then

$$\begin{aligned} U(f, P) - U(f, P^*) &= \sum_{i \in \mathcal{I}} M_i(x_i - x_{i-1}) - \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}(i)} M_j^*(x_j^* - x_{j-1}^*) \\ &= \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}(i)} (M_i - M_j^*)(x_j^* - x_{j-1}^*) \end{aligned}$$

Note that  $M_i - M_j^* \leq 2M$  and

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}(i)} (x_j^* - x_{j-1}^*) \leq 2n_1 \text{mesh}(P) < 2n_1 \delta = \frac{\epsilon}{6M},$$

hence

$$U(f, P) - U(f, P^*) < 2M \cdot \frac{\epsilon}{6M} = \frac{\epsilon}{3}.$$

Similarly, we can show that

$$L(f, P^*) - L(f, P) < \frac{\epsilon}{3}.$$

Therefore,

$$\begin{aligned} U(f, P) - L(f, P) &< \frac{\epsilon}{3} + (U(f, P^*) - L(f, P^*)) + \frac{\epsilon}{3} \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

Note that both  $R(f, P, T)$  and  $I$  lie between  $L(f, P)$  and  $U(f, P)$ , hence

$$|R(f, P, T) - I| < \epsilon.$$

This completes the proof.  $\square$

**Example 3.1.1** Any continuous function on  $[a, b]$  is Riemann integrable.

---

*Proof.* Let  $\epsilon > 0$  be given. Then  $\exists \delta > 0$  such that for all  $x, y \in [a, b]$ , if  $|x - y| < \delta$ , then  $|f(x) - f(y)| < \epsilon/(b - a)$ . For any partition  $P$  with  $\text{mesh}(P) < \delta$ , we have

$$M_i - m_i < \frac{\epsilon}{b - a} \quad \text{for all } 1 \leq i \leq n,$$

hence

$$U(f, P) - L(f, P) = \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < \sum_{i=1}^n \frac{\epsilon}{b - a} (x_i - x_{i-1}) = \epsilon.$$

Thus  $f$  is Riemann integrable.  $\square$

**Proposition 3.1.1** A function  $f : [a, b] \rightarrow \mathbb{R}$  is Riemann integrable if for all  $\epsilon > 0$ ,  $\exists$  functions  $g, h : [a, b] \rightarrow \mathbb{R}$  such that  $g(x) \leq f(x) \leq h(x)$  for all  $x \in [a, b]$ ,  $g$  and  $h$  are Riemann integrable, and

$$\int_a^b (h(x) - g(x)) dx < \epsilon.$$

*Proof.* Let  $\epsilon > 0$ . Choose Riemann integrable  $g, h$  such that  $g(x) \leq f(x) \leq h(x)$  for all  $x \in [a, b]$  and

$$\int_a^b (h(x) - g(x)) dx < \epsilon/3.$$

For all partition  $P$ , we have

$$L(g, P) \leq L(f, P) \leq U(f, P) \leq U(h, P).$$

Since  $g$  and  $h$  are Riemann integrable,  $\exists$  a partition  $P$  such that

$$\int_a^b g(x) dx - L(g, P) < \epsilon/3, \quad U(h, P) - \int_a^b h(x) dx < \epsilon/3.$$

Combining these inequalities, we have

$$\begin{aligned} U(f, P) - L(f, P) &< U(h, P) - L(g, P) \\ &< \int_a^b h(x) dx - \int_a^b g(x) dx + 2\epsilon/3 \\ &= \int_a^b (h(x) - g(x)) dx + 2\epsilon/3 < \epsilon. \end{aligned}$$

Therefore,  $f$  is Riemann integrable.  $\square$

**Definition 3.1.4** Given a set  $E \subset \mathbb{R}$ , the characteristic function of  $E$  is defined to be

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

---

**Example 3.1.2** (Thomae's Function) Define  $[0, 1] \rightarrow \mathbb{Q}$  by

$$f(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \text{ in lowest terms} \\ 0, & x \text{ is irrational} \end{cases}$$

$f$  is discontinuous at every rational point but continuous at every irrational point. For  $\epsilon > 0$ , define

$$S_\epsilon(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \text{ in lowest terms and } q < \frac{1}{\epsilon} \\ 0, & \text{otherwise} \end{cases}$$

$S_\epsilon$  is Riemann integrable with integral 0 since  $S_\epsilon$  is nonzero at only finitely many points. Take  $g(x) = 0$  and  $h(x) = \epsilon + S_\epsilon(x)$ , then  $g(x) \leq f(x) \leq h(x)$  for all  $x \in [0, 1]$ , thus  $f$  is Riemann integrable with integral 0 by the previous proposition.

If a property holds for all points except for a set of measure zero, then we say that the property holds almost everywhere.

**Theorem 3.1.5** (Riemann's Integrability Criterion)  $f : [a, b] \rightarrow \mathbb{R}$  is Riemann integrable iff  $f$  is bounded and continuous almost everywhere.

*Proof.* Define the oscillation of  $f$  at  $x \in [a, b]$  to be

$$\text{osc}_x(f) = \limsup_{t \rightarrow x} f(t) - \liminf_{t \rightarrow x} f(t).$$

Note that  $f$  is continuous at  $x$  iff  $\text{osc}_x(f) = 0$ . If  $I$  is any interval containing  $x$ , then

$$\sup_{t \in I} f(t) - \inf_{t \in I} f(t) \geq \text{osc}_x(f).$$

" $\implies$ ": Suppose that  $f$  is Riemann integrable, then  $f$  is bounded, say  $|f(x)| \leq M$  for all  $x \in [a, b]$ . Let  $\mathcal{D}$  denote the set of discontinuities of  $f$ . Write

$$\mathcal{D} = \bigcup_{k=1}^{\infty} \mathcal{D}_k, \quad \mathcal{D}_k = \left\{ x \in [a, b] \mid \text{osc}_x(f) \geq \frac{1}{k} \right\}.$$

It suffices to show that each  $\mathcal{D}_k$  has measure zero. Given  $\epsilon > 0$  and  $k \in \mathbb{N}$ , there exists a partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[a, b]$  such that

$$U(f, P) - L(f, P) = \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < \frac{\epsilon}{k}.$$

We say that a  $\mathcal{P}$ -interval  $I_i = (x_{i-1}, x_i)$  is bad if it contains a point  $x \in \mathcal{D}_k$  in its interior. On

---

a bad interval  $I_i$ , we have  $M_i - m_i \geq \text{osc}_x(f) \geq 1/k$ . Therefore,

$$\sum_{I_i \text{ is bad}} (x_i - x_{i-1}) \leq k \sum_{i=1}^n (M_i - m_i)(x_i - x_{i-1}) < k \cdot \frac{\epsilon}{k} = \epsilon.$$

The collection of all bad  $\mathcal{P}$ -intervals covers  $\mathcal{D}_k$  except possibly the endpoints of these intervals, which is a finite set. Therefore,  $\mathcal{D}_k$  has measure zero.

” $\Leftarrow$ ”: Suppose that  $f$  is bounded and continuous almost everywhere, say  $|f(x)| \leq M$  for all  $x \in [a, b]$ . Let  $\mathcal{D}$  denote the set of discontinuities of  $f$ , and define  $\mathcal{D}_k$  as above. Given  $\epsilon > 0$ , pick  $k \in \mathbb{N}$  such that

$$\frac{1}{k} < \frac{\epsilon}{2(b-a)}$$

Since  $\mathcal{D}_k$  has measure zero,  $\exists$  a countable cover  $\mathcal{J}$  of  $\mathcal{D}_k$  by open intervals  $J_j = (a_j, b_j)$  such that

$$\sum_{j=1}^{\infty} (b_j - a_j) < \frac{\epsilon}{4M}.$$

On the other hand, for each  $x \in [a, b] \setminus \mathcal{D}_k$ ,  $\exists$  an open interval  $I_x$  containing  $x$  such that

$$\sup_{t \in I_x} f(t) - \inf_{t \in I_x} f(t) < 1/k.$$

These intervals  $I_x$  form an open cover  $\mathcal{I}$  of  $[a, b] \setminus \mathcal{D}_k$ , hence  $\mathcal{U} := \mathcal{I} \cup \mathcal{J}$  is an open cover of  $[a, b]$ . By compactness,  $\mathcal{U}$  has a positive Lebesgue number  $\lambda > 0$ .

Let  $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$  be any partition of  $[a, b]$  with  $\text{mesh}(\mathcal{P}) < \lambda$ . We claim that  $U(f, P) - L(f, P) < \epsilon$ . By definition of Lebesgue number, each  $\mathcal{P}$ -interval is contained entirely in some  $I_x$  or in some  $J_j$ . Define

$$\mathbb{I} = \{i \mid I_i \subseteq J_j \text{ for some } j\}.$$

Since  $\mathbb{I}$  contains at most finitely many indices, there exists a finite  $m$  such that

$$\bigcup_{i \in \mathbb{I}} I_i \subseteq J_1 \cup J_2 \cup \dots \cup J_m.$$

Therefore,

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_{i \in \mathbb{I}} (M_i - m_i)(x_i - x_{i-1}) + \sum_{i \notin \mathbb{I}} (M_i - m_i)(x_i - x_{i-1}) \\ &\leq 2M \sum_{i \in \mathbb{I}} (x_i - x_{i-1}) + \frac{1}{k} \sum_{i \notin \mathbb{I}} (x_i - x_{i-1}) \\ &\leq 2M \sum_{j=1}^m (b_j - a_j) + \frac{1}{k} (b - a) < 2M \cdot \frac{\epsilon}{4M} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

---

This completes the proof.  $\square$

**Remark** Being equal to a continuous function almost everywhere does not imply it is continuous almost everywhere, in particular does not imply Riemann integrability. For example, the characteristic function of  $\mathbb{Q}$  is equal to the zero function almost everywhere, but it is not Riemann integrable.

**Corollary 3.1.6** Every bounded piecewise continuous function is Riemann integrable.

*Proof.* The set of discontinuities of a piecewise continuous function is finite, hence has measure zero.  $\square$

**Corollary 3.1.7** The characteristic function of a set  $S \subseteq [a, b]$  is Riemann integrable iff the boundary  $\partial S$  has measure zero.

*Proof.* Note that  $\chi_S$  is continuous at  $x$  iff  $x \notin \partial S$ . Therefore, the set of discontinuities of  $\chi_S$  is exactly  $\partial S$ .  $\square$

**Corollary 3.1.8** Every monotone function on  $[a, b]$  is Riemann integrable.

*Proof.* A monotone function can have only jump discontinuities, hence the set of discontinuities is at most countable, thus has measure zero.  $\square$

**Corollary 3.1.9** The product of Riemann integrable functions is Riemann integrable.

*Proof.* The set of discontinuities of the product is contained in the union of the sets of discontinuities of the factors, hence has measure zero.  $\square$

**Corollary 3.1.10** If  $f : [a, b] \rightarrow [c, d]$  is Riemann integrable and  $\phi : [c, d] \rightarrow \mathbb{R}$  is continuous, then  $\phi \circ f$  is Riemann integrable.

*Proof.* The set of discontinuities of  $\phi \circ f$  is contained in the set of discontinuities of  $f$ , hence has measure zero. Since  $[c, d]$  is compact,  $\phi$  is bounded, thus  $\phi \circ f$  is bounded as well.  $\square$

**Corollary 3.1.11** If  $f$  is Riemann integrable then so is  $|f|$ .

*Proof.* Take  $\phi(x) = |x|$  in the previous corollary.  $\square$

**Corollary 3.1.12** If  $a < c < b$  and  $f : [a, b] \rightarrow \mathbb{R}$  is Riemann integrable, then  $f$  is Riemann integrable on  $[a, c]$  and  $[c, b]$ , and

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

---

Conversely, if  $f$  is Riemann integrable on  $[a, c]$  and  $[c, b]$ , then  $f$  is Riemann integrable on  $[a, b]$  and the same equality holds.

*Proof.* The set of discontinuities of  $f$  on  $[a, b]$  is the union of the sets of discontinuities of  $f$  on  $[a, c]$  and  $[c, b]$  (and possibly the point  $c$ ), hence  $f$  is Riemann integrable iff  $f|_{[a,c]}$  and  $f|_{[c,b]}$  are Riemann integrable. Note that

$$f = f\chi_{[a,c]} + f\chi_{[c,b]}.$$

Then the formula follows from linearity of the integral.  $\square$

**Corollary 3.1.13** A Riemann integrable function  $f : [a, b] \rightarrow [0, M]$  has integral 0 iff  $f(x) = 0$  almost everywhere.

*Proof.* " $\Rightarrow$ ": Suppose that  $f$  is Riemann integrable with integral 0. If  $f$  is nonzero at some continuity point  $x_0$ , then  $\exists$  an open interval  $(x_0 - \delta, x_0 + \delta) \subseteq [a, b]$  such that  $f(x) > M/2$  for all  $x$  in this interval. Therefore,

$$\int_a^b f(x) dx \geq \int_{x_0-\delta}^{x_0+\delta} f(x) dx \geq \int_{x_0-\delta}^{x_0+\delta} \frac{f(x_0)}{2} dx = f(x_0)\delta > 0,$$

contradicting the assumption that the integral is 0. Therefore,  $f(x) = 0$  at every continuity point of  $f$ , hence  $f(x) = 0$  almost everywhere.

" $\Leftarrow$ ": Suppose that  $f(x) = 0$  almost everywhere, then every interval contains a point  $x$  such that  $f(x) = 0$ . Therefore, for every partition  $P$ , we have  $m_i = 0$  for all  $1 \leq i \leq n$ , hence

$$0 \leq \int_a^b f(x) dx = \underline{I} = \sup_P L(f, P) = 0.$$

Thus the integral is 0.  $\square$

**Corollary 3.1.14** If  $f : [a, b] \rightarrow \mathbb{R}$  is Riemann integrable and if  $\psi : [c, d] \rightarrow [a, b]$  is a homeomorphism such that  $\psi^{-1}$  is Lipschitz, then  $f \circ \psi$  is Riemann integrable.

*Proof.* Clearly  $f \circ \psi$  is bounded. Let  $\mathcal{D}$  denote the set of discontinuities of  $f$ , then the set of discontinuities of  $f \circ \psi$  is  $\psi^{-1}(\mathcal{D})$ . Since  $\mathcal{D}$  has measure zero, for every  $\epsilon > 0$ ,  $\exists$  a countable cover of  $\mathcal{D}$  by open intervals  $\{(a_k, b_k)\}_{k \in \mathbb{N}}$  such that

$$\sum_{k=1}^{\infty} (b_k - a_k) < \epsilon.$$

Since  $\psi^{-1}$  is Lipschitz,  $\exists C > 0$  such that for all  $x, y \in [a, b]$ ,

$$|\psi^{-1}(x) - \psi^{-1}(y)| \leq C|x - y|.$$

---

Therefore,  $\{(\psi^{-1}(a_k), \psi^{-1}(b_k))\}_{k \in \mathbb{N}}$  is a countable cover of  $\psi^{-1}(\mathcal{D})$  by open intervals of total length at most  $C\epsilon$ . Since  $\epsilon$  is arbitrary,  $\psi^{-1}(\mathcal{D})$  has measure zero. Hence  $f \circ \psi$  is Riemann integrable.  $\square$

**Remark** If  $f$  is Riemann integrable and  $\psi$  is continuous but not Lipschitz, then  $f \circ \psi$  may not be Riemann integrable. In particular, composition of two Riemann integrable functions may not be Riemann integrable.

## 3.2 Fundamental Theorem of Calculus

Let  $f : [a, b] \rightarrow \mathbb{R}$  be Riemann integrable. Define the indefinite integral (or integral function) of  $f$  to be

$$F(x) = \int_a^x f(t) dt \quad \text{for all } x \in [a, b].$$

$F$  is continuous on  $[a, b]$  since for all  $x, y \in [a, b]$ . Moreover, the derivative of  $F$  exists and  $F'(x) = f(x)$  at every point where  $f$  is continuous.

An antiderivative of  $f$  is a function  $F : [a, b] \rightarrow \mathbb{R}$  such that  $F'(x) = f(x)$  for all  $x \in [a, b]$ .

**Corollary 3.2.1** Every continuous function on  $[a, b]$  has an antiderivative.

The second part of the Fundamental Theorem of Calculus states that an antiderivative of a Riemann integrable function, if it exists, differs from the indefinite integral by a constant.

**Theorem 3.2.2** There exists a continuous  $H : [0, 1] \rightarrow \mathbb{R}$  whose derivative exists and equals 0 almost everywhere, but such that  $H$  is not constant.

*Proof.* Each  $x \in [0, 1]$  has a unique ternary expansion

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}, \quad a_n \in \{0, 1, 2\}.$$

Define

$$H(x) = \sum_{n=1}^{\infty} \frac{b_n/2}{2^n},$$

for  $x \in C$ , where  $C$  is the Cantor set. Note that  $H$  has equal values at the endpoints of each removed interval, hence  $H$  can be extended to be constant on each removed interval.

Let  $\epsilon > 0$ , and choose  $n$  such that  $1/2^n < \epsilon$ . Take  $x, x' \in C$  such that  $|x - x'| < 3^{-n}$ . Since each discarded open interval has length  $3^{-n}$  in  $C_n$ ,  $x$  and  $x'$  must lie in the same closed

---

interval of  $C_n$ , hence the ternary expansions of  $x$  and  $x'$  agree up to the  $n$ -th digit. Therefore,

$$|H(x) - H(x')| \leq \sum_{k=n+1}^{\infty} \frac{1/2}{2^k} = \frac{1}{2^n} < \epsilon.$$

This shows that  $H$  is uniformly continuous on  $C$ . Since  $H$  is constant on  $[0, 1] \setminus C$ ,  $H$  is uniformly continuous on  $[0, 1]$ , thus  $H([0, 1]) = [0, 1]$ , also  $H$  is differentiable with derivative 0 at every point in  $[0, 1] \setminus C$ . Therefore,  $H$  is not constant but  $H' = 0$  almost everywhere.  $\square$

**Remark** It is even possible to construct a strictly increasing continuous function whose derivative is 0 almost everywhere.

### 3.3 Multiple Integrals

For notational simplicity, we will only consider functions of two variables, but the definitions and results can be easily generalized to functions of more than two variables.

Consider a rectangle  $R = [a, b] \times [c, d] \subseteq \mathbb{R}^2$ . Consider a pair of partitions  $P = \{x_0, x_1, \dots, x_m\}$  of  $[a, b]$  and  $Q = \{y_0, y_1, \dots, y_n\}$  of  $[c, d]$ . Let  $G$  be the grid partition  $G$  consisting of the rectangles

$$R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j], \quad 1 \leq i \leq m, \quad 1 \leq j \leq n.$$

Let  $S$  be a set of sample points  $s_{ij} \in R_{ij}$  for all  $1 \leq i \leq m$  and  $1 \leq j \leq n$ . The Riemann sum of a bounded function  $f : R \rightarrow \mathbb{R}$  with respect to the grid partition  $G$  and the sample points  $S$  is defined to be

$$R(f, G, S) = \sum_{i=1}^m \sum_{j=1}^n f(s_{ij}) \Delta x_i \Delta y_j,$$

where  $\Delta x_i = x_i - x_{i-1}$  and  $\Delta y_j = y_j - y_{j-1}$ . The mesh of the grid partition  $G$  is defined to be

$$\text{mesh}(G) = \max_{1 \leq i \leq m, 1 \leq j \leq n} \text{diam}(R_{ij}).$$

The function  $f$  is Riemann integrable on  $R$  if the limit of the Riemann sums exists as the mesh of the partition goes to zero:

$$\int_R f(x, y) dA = \lim_{\text{mesh}(G) \rightarrow 0} R(f, G, S).$$

The lower sum and upper sum of  $f$  with respect to the grid partition  $G$  are defined to be

$$L(f, G) = \sum_{i=1}^m \sum_{j=1}^n m_{ij} \Delta x_i \Delta y_j, \quad U(f, G) = \sum_{i=1}^m \sum_{j=1}^n M_{ij} \Delta x_i \Delta y_j,$$

---

where

$$m_{ij} = \inf_{(x,y) \in R_{ij}} f(x,y), \quad M_{ij} = \sup_{(x,y) \in R_{ij}} f(x,y).$$

The lower integral and upper integral of  $f$  on  $R$  are defined to be

$$\underline{\int}_R f = \sup_G L(f, G), \quad \overline{\int}_R f = \inf_G U(f, G),$$

where the supremum and infimum are taken over all grid partitions  $G$  of  $R$ . The function  $f$  is Darboux integrable on  $R$  if  $\underline{\int}_R f = \overline{\int}_R f$ .

**Proposition 3.3.1**

- If  $f$  is Riemann integrable on  $R$ , then  $f$  is bounded.

- The set of Riemann integrable functions on  $R$  is a vector space, and the integral is a linear functional on this space.
- Constant functions  $f \equiv c$  are integrable and its integral is  $c \cdot |R|$ .
- If  $f, g$  are Riemann integrable on  $R$  and if  $f(x, y) \leq g(x, y)$  for all  $(x, y) \in R$ , then

$$\int_R f(x, y) dA \leq \int_R g(x, y) dA.$$

- If  $f$  is Riemann integrable on  $R$ , then

$$\underline{\int}_R f = \overline{\int}_R f = \int_R f.$$

**Definition 3.3.1** A set  $Z \subseteq \mathbb{R}^2$  is of measure zero if for every  $\epsilon > 0$ ,  $\exists$  a countable cover of  $Z$  by open rectangles  $S_l$  whose total area is less than  $\epsilon$ :

$$\sum_{l=1}^{\infty} |S_l| < \epsilon.$$

Using the same argument as in the one-dimensional case, we can show that a countable union of measure zero sets has measure zero. For  $z \in R$ , define the oscillation of  $f$  at  $z$  to be

$$\text{osc}_z(f) = \lim_{r \rightarrow 0^+} \text{diam}(f(B(z, r) \cap R)).$$

Denote the set of discontinuities of  $f$  by

$$\mathcal{D} = \bigcup_{k=1}^{\infty} \mathcal{D}_k, \quad \mathcal{D}_k = \left\{ z \in R \mid \text{osc}_z(f) \geq \frac{1}{k} \right\}.$$

Since  $\mathcal{D}_k$  is closed, it is compact.

**Theorem 3.3.2** (Riemann Integrability Criterion in Higher Dimensions) A bounded function

---

$f : R \rightarrow \mathbb{R}$  is Riemann integrable iff the set of discontinuities of  $f$  has measure zero.

Suppose that  $f : R \rightarrow \mathbb{R}$  is Riemann integrable. For each fixed  $y \in [c, d]$ , define the lower slice integral and upper slice integral of  $f$  at  $y$  to be

$$\underline{F}(y) = \int_a^b f(x, y) dx, \quad \overline{F}(y) = \int_a^b f(x, y) dx.$$

Note that these are just the lower and upper integrals of the one-variable function  $f_y(x) = f(x, y)$  on  $[a, b]$ .

**Theorem 3.3.3** (Fubini's Theorem) If  $f : R \rightarrow \mathbb{R}$  is Riemann integrable, then so are  $\underline{F}$  and  $\overline{F}$ . Moreover,

$$\int_R f(x, y) dA = \int_c^d \underline{F}(y) dy = \int_c^d \overline{F}(y) dy.$$

*Proof.* Let  $G = P \times Q$  be a grid partition of  $R$ , where  $P = \{x_0, x_1, \dots, x_m\}$  is a partition of  $[a, b]$  and  $Q = \{y_0, y_1, \dots, y_n\}$  is a partition of  $[c, d]$ .

We first show that  $L(f, G) \leq L(\underline{F}, Q)$ . Fix a partition  $J_j \subseteq [c, d]$  of  $Q$ . If  $y \in J_j$ ,

$$m_{ij} = \inf \{f(s, t) \mid (s, t) \in R_{ij}\} \leq \inf \{f(s, y) \mid s \in I_i\} = m_i(f_y),$$

hence for all  $y \in J_j$ ,

$$\sum_{i=1}^m m_{ij} \Delta x_i \leq \sum_{i=1}^m m_i(f_y) \Delta x_i = L(f_y, P) \leq \int_a^b f_y(x) dx = \underline{F}(y).$$

Therefore,

$$\sum_{i=1}^m m_{ij} \Delta x_i \leq \inf_{y \in J_j} \underline{F}(y) = m_j(\underline{F}),$$

and

$$L(f, G) = \sum_{j=1}^n \sum_{i=1}^m m_{ij} \Delta x_i \Delta y_j \leq \sum_{j=1}^n m_j(\underline{F}) \Delta y_j = L(\underline{F}, Q).$$

Similarly, we can show that  $U(f, G) \geq U(\overline{F}, Q)$ , thus

$$L(f, G) \leq L(\underline{F}, Q) \leq U(\overline{F}, Q) \leq U(f, G).$$

Since  $f$  is Riemann integrable, we know that

$$\sup_Q L(\underline{F}, Q) = \inf_Q U(\overline{F}, Q) = \int_R f(x, y) dA,$$

---

hence  $\underline{F}$  and  $\overline{F}$  are Riemann integrable with

$$\int_c^d \underline{F}(y) dy = \int_c^d \overline{F}(y) dy = \int_R f(x, y) dA.$$

□

**Corollary 3.3.4** If  $f : R \rightarrow \mathbb{R}$  is Riemann integrable, then the order of integration does not affect the value of the integral.

**Corollary 3.3.5** (Cavalier's Principle) The area of a region  $S \subseteq R$  is the integral with respect to  $x$  of the lengths of the vertical slices:

$$\text{area}(S) = \int_a^b \text{length}(S_x) dx,$$

provided that  $\partial S$  has measure zero, where  $S_x = \{y \in [c, d] \mid (x, y) \in S\}$ .

*Proof.* Take  $f = \chi_S$  in Fubini's Theorem. □

**Definition 3.3.6** If  $S \subseteq \mathbb{R}^2$  is bounded, we define its area (or Jordan content) to be  $\int \chi_S dA$  if exists. If the integral exists, we say that  $S$  is Riemann measurable. Denote the area by

$$|S| = \text{area}(S) = \int \chi_S dA.$$

### 3.4 Change of Variables

**Proposition 3.4.1** If  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is a linear transformation, then for every Riemann measurable set  $S \subseteq \mathbb{R}^2$ ,  $T(S)$  is Riemann measurable and

$$|T(S)| = |\det(T)| \cdot |S|.$$

*Proof.* If  $A$  is a matrix representation of  $T$ , then we can decompose  $A$  into a product of elementary matrices. It suffices to prove the proposition for each type of elementary matrices.

Consider first that  $S = [a, b] \times [c, d]$  is a rectangle. If  $T$  is a scaling matrix, then the result is clear. If  $T$  is a permutation matrix, then  $|T(S)| = |S|$ . If  $T$  is a shear matrix, then it converts  $S$  into a parallelogram

$$\mathcal{P} = \{(x, y) \mid a + \sigma y \leq x \leq b + \sigma y, c \leq y \leq d\}.$$

We can show that  $\mathcal{P}$  is Riemann measurable, whose area can be computed using Cavalier's

Principle:

$$|\mathcal{P}| = \int_c^d \int_{a+\sigma y}^{b+\sigma y} 1 \, dx \, dy = \int_c^d (b-a) \, dy = (b-a)(d-c) = |S|.$$

Now let  $S$  be a general Riemann measurable set. Given  $\epsilon > 0$ , let  $H \subseteq S$  be a rectangle. Take a grid  $G$  on  $H$  such that

$$|S| - \epsilon \leq \sum_{R \subseteq S} |R| \leq \sum_{R: R \cap S \neq \emptyset} |R| \leq |S| + \epsilon.$$

This is always possible since  $\chi_S$  is Riemann integrable and

$$\sum_{R \subseteq S} |R| = L(\chi_S, G), \quad \sum_{R: R \cap S \neq \emptyset} |R| = U(\chi_S, G).$$

Since the interior of the inner rectangles are disjoint, we have

$$\sum_{R \subseteq S} \chi_{\text{int}(R)}(z) \leq \chi_S(z) \rightsquigarrow \sum_{R \subseteq S} \chi_{\text{int}(T(R))}(w) \leq \chi_{T(S)}(w).$$

Therefore,

$$\begin{aligned} |\det(E)|(|S| - \epsilon) &\leq |\det(E)| \sum_{R \subseteq S} |R| \\ &= \sum_{R \subseteq S} |T(R)| = \sum_{R \subseteq S} \int \chi_{\text{int}(T(R))} \, dA \leq \underline{\int} \chi_{T(S)}. \end{aligned}$$

Similarly,

$$\begin{aligned} |\det(E)|(|S| + \epsilon) &\geq |\det(E)| \sum_{R: R \cap S \neq \emptyset} |R| \\ &= \sum_{R: R \cap S \neq \emptyset} |T(R)| = \sum_{R: R \cap S \neq \emptyset} \int \chi_{T(R)} \, dA \geq \overline{\int} \chi_{T(S)}. \end{aligned}$$

Since  $\epsilon$  is arbitrary, we have

$$\underline{\int} \chi_{T(S)} = |\det(E)||S| = \overline{\int} \chi_{T(S)},$$

hence proved.  $\square$

**Remark** The proposition can be generalized to higher dimensions  $\mathbb{R}^n$ , and is called the volume multiplier formula.

**Definition 3.4.1** A norm on a vector space  $V$  (over  $\mathbb{R}$ ) is a function  $\|\cdot\| : V \rightarrow [0, \infty)$  such that

- 
1.  $\|v\| \geq 0$  for all  $v \in V$ , and  $\|v\| = 0$  iff  $v = 0$ .
  2.  $\|\lambda v\| = |\lambda| \|v\|$  for all  $\lambda \in \mathbb{R}$  and  $v \in V$ .
  3.  $\|v + w\| \leq \|v\| + \|w\|$  for all  $v, w \in V$  (triangle inequality).

A vector space with a norm is called a normed space. A norm gives rise to a metric by defining  $d(v, w) = \|v - w\|$ . If  $V, W$  are normed spaces and if  $T : V \rightarrow W$  is a linear transformation, then we can define the operator norm of  $T$  to be

$$\|T\| = \sup_{\|v\|=1} \|T(v)\|.$$

Consider  $V = W = \mathbb{R}^2$  with the sup-norm

$$\|(x, y)\| = \max\{|x|, |y|\}.$$

The closed  $r$ -neighborhood of  $z \in \mathbb{R}^2$  is a square  $S(z, r)$  under this norm. Write  $S(r) := S(0, r)$ .

**Lemma 3.4.2** Let  $\mathcal{U}$  be open with  $0 \in \mathcal{U}$ , and let  $\sigma \in (0, 1)$ . Suppose that  $\psi : \mathcal{U} \rightarrow \mathbb{R}^2$  is  $C^1$  such that  $\psi(0) = 0$ , and for all  $u \in \mathcal{U}$ , we have  $\|(D\psi)_u - I\| \leq \sigma$ . If  $S(r) \subseteq \mathcal{U}$ , then

$$S(r(1 - \sigma)) \subseteq \psi(S(r)) \subseteq S(r(1 + \sigma)).$$

*Proof.* Define  $\rho(u) = \psi(u) - u$ , then  $\rho$  is  $C^1$ ,  $\rho(0) = 0$ , and  $\|(D\rho)_u\| \leq \sigma$  for all  $u \in \mathcal{U}$ . For any  $\|u\| \leq r$ , by the Mean Value Theorem,

$$\|\psi(u)\| \leq \|u\| + \|\rho(u)\| \leq r + \sigma r = r(1 + \sigma) \rightsquigarrow \psi(S(r)) \subseteq S(r(1 + \sigma)).$$

For the other side, we wish to show that for every  $\|v\| \leq r(1 - \sigma)$ , there exists  $u \in S(r)$  such that  $\psi(u) = v$ . Define  $K_v : S(r) \rightarrow \mathbb{R}^2$  by  $K_v(u) = v - \rho(u)$ . Note that  $K_v$  is continuous and  $K_v(S(r)) \subseteq S(r)$  since for all  $\|u\| \leq r$ ,

$$\|K_v(u)\| \leq \|v\| + \|\rho(u)\| \leq r(1 - \sigma) + \sigma r = r.$$

By Banach Fixed Point Theorem,  $K_v$  has a unique fixed point  $u \in S(r)$  such that  $K_v(u) = u$ , hence  $\psi(u) = v$ . Therefore,  $S(r(1 - \sigma)) \subseteq \psi(S(r))$ .  $\square$

**Lemma 3.4.3** The Lipschitz image of a set of measure zero has measure zero.

*Proof.* Let  $Z \subseteq \mathbb{R}^2$  be a set of measure zero, and let  $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  be Lipschitz with Lipschitz constant  $L > 0$ . Given  $\epsilon > 0$ ,  $\exists$  a countable cover of  $Z$  by open squares  $S_k$  such that

$$\sum_{k=1}^{\infty} |S_k| < \frac{\epsilon}{L^2}.$$

---

Note that

$$\text{diam}(h(S_k \cap Z)) \leq L \text{diam}(S_k \cap Z) \leq L \text{diam}(S_k)$$

w.r.t. the sup-norm. Therefore,  $h(S_k \cap Z)$  is contained in an open square  $S'_k$  of side length  $L \text{diam}(S_k)$ , hence  $\{S'_k\}$  is a countable cover of  $h(Z)$  by open squares such that

$$\sum_{k=1}^{\infty} |S'_k| = \sum_{k=1}^{\infty} L^2 |S_k| < \epsilon.$$

Since  $\epsilon$  is arbitrary,  $h(Z)$  has measure zero.  $\square$

**Theorem 3.4.4** Let  $f$  be Riemann integrable, and let  $\varphi : R \rightarrow \mathbb{R}^2$  be a  $C^1$ -diffeomorphism, then

$$\int_R f \circ \varphi \cdot |\det(D\varphi)| dA = \int_{\varphi(R)} f dA.$$

*Proof.* We first verify that the integrands on both sides are Riemann integrable.

Let  $\mathcal{D}'$  be the set of discontinuities of  $f$ . Define  $\mathcal{D} = \varphi^{-1}(\mathcal{D}')$  to be the set of discontinuities of  $f \circ \varphi$ . Since  $\varphi$  is a  $C^1$ -diffeomorphism,  $\varphi^{-1}$  is Lipschitz, thus by Lemma 3.4.3,  $\mathcal{D}$  has measure zero, thus  $f \circ \varphi$  is Riemann integrable. The function  $|\det(D\varphi)|$  is continuous, hence Riemann integrable, and the product of two Riemann integrable functions is Riemann integrable, thus the integrand on the left-hand side is Riemann integrable.

Since  $\varphi$  is a  $C^1$ -diffeomorphism, in particular it is a homeomorphism, hence  $\varphi(\partial R) = \partial(\varphi(R))$ . Since  $\partial R$  has measure zero, so does  $\varphi(\partial R)$  by Lemma 3.4.3 again, hence  $\varphi(R)$  is Riemann measurable and  $\chi_{\varphi(R)}$  is Riemann integrable.

We may assume that  $f \geq 0$  since we can decompose  $f$  into

$$f = f^+ - f^-, \quad f^+ = \max\{f, 0\}, \quad f^- = \max\{-f, 0\}.$$

Take any grid partition of  $R$  into  $R_{ij}$  of radius at most  $r$ . Let  $z_{ij}$  be the center of  $R_{ij}$  and define

$$A_{ij} = (D\varphi)_{z_{ij}}, \quad \varphi(z_{ij}) = w_{ij}, \quad \varphi(R_{ij}) = W_{ij}.$$

The affine approximation of  $\varphi$  at  $z_{ij}$  is

$$\phi_{ij}(z) = A_{ij}(z - z_{ij}) + w_{ij}.$$

Define  $\psi = \phi_{ij}^{-1} \circ \varphi$ , then  $\psi_{z_{ij}} = z_{ij}$  and  $(D\psi)_{z_{ij}} = I$ . Since  $(D\psi)_z$  is uniformly continuous on the compact set  $R$ ,  $\forall R_{ij}$  and for all  $z \in R_{ij}$ , if  $r$  is sufficiently small, then

$$\|(D\psi)_z - I\| < \sigma,$$

for arbitrarily small  $\sigma > 0$ . By 3.4.2,  $\psi(R_{ij})$  is sandwiched between the squares  $S(z_{ij}, r(1-\sigma))$

and  $S(z_{ij}, r(1 + \sigma))$ . Applying  $\phi_{ij}$ , we see that

$$\begin{aligned} \phi_{ij}(S(z_{ij}, r(1 - \sigma))) &\subseteq W_{ij} \subseteq \phi_{ij}(S(z_{ij}, r(1 + \sigma))) \\ \implies (1 - \sigma)^2 |\det(A_{ij})| |R_{ij}| &\leq |W_{ij}| \leq (1 + \sigma)^2 |\det(A_{ij})| |R_{ij}|. \end{aligned}$$

Let  $m_{ij} = \inf_{R_{ij}} f \circ \varphi$  and  $M_{ij} = \sup_{R_{ij}} f \circ \varphi$ , then for all  $w \in \varphi(R)$ ,

$$\sum_{i,j} m_{ij} \chi_{\text{int}(W_{ij})}(w) \leq f(w) \leq \sum_{i,j} M_{ij} \chi_{W_{ij}}(w).$$

By integrating over  $\varphi(R)$ , we have

$$\begin{aligned} \sum_{i,j} m_{ij} |W_{ij}| &\leq \int_{\varphi(R)} f(w) dA \leq \sum_{i,j} M_{ij} |W_{ij}| \\ \implies (1 - \sigma)^2 \sum_{i,j} m_{ij} |\det(A_{ij})| |R_{ij}| &\leq \int_{\varphi(R)} f(w) dA \leq (1 + \sigma)^2 \sum_{i,j} M_{ij} |\det(A_{ij})| |R_{ij}|. \end{aligned}$$

By uniform continuity of  $\det(D\varphi)$ , if  $r$  is sufficiently small, then for all  $z \in R_{ij}$ ,

$$1 - \tau \leq \frac{|\det(D\varphi)_z|}{|\det(A_{ij})|} \leq 1 + \tau,$$

for arbitrarily small  $\tau > 0$ . Let  $L$  be the lower sum of  $f \circ \varphi \cdot |\det(D\varphi)|$  with respect to the partition  $R_{ij}$ , and let  $U$  be the corresponding upper sum, then

$$\begin{aligned} (1 - \tau)L &\leq \sum_{i,j} m_{ij} |\det A_{ij}| |R_{ij}|, \quad \sum_{i,j} M_{ij} |\det A_{ij}| |R_{ij}| \leq (1 + \tau)U \\ \implies (1 - \sigma)^2(1 - \tau)L &\leq \int_{\varphi(R)} f(w) dA \leq (1 + \sigma)^2(1 + \tau)U. \end{aligned}$$

Since  $\sigma, \tau$  can be arbitrarily small when  $r$  approaches 0, we have

$$\underline{\int}_R f \circ \varphi(z) |\det(D\varphi)_z| dA \leq \int_{\varphi(R)} f(w) dA \leq \overline{\int}_R f \circ \varphi(z) |\det(D\varphi)_z| dA,$$

hence proved.  $\square$

## CHAPTER 4

# Function Spaces

### 4.1 Uniform Convergence and Completeness

A sequence of functions  $f_n : [a, b] \rightarrow \mathbb{R}$  converges uniformly to a function  $f : [a, b] \rightarrow \mathbb{R}$  if for every  $\epsilon > 0$ ,  $\exists N \in \mathbb{N}$  such that for all  $n \geq N$  and for all  $x \in [a, b]$ ,

$$|f_n(x) - f(x)| < \epsilon.$$

We say that  $f$  is the uniform limit of  $f_n$ , and write  $f_n \rightrightarrows f$ . Uniform convergence implies pointwise convergence, but not vice versa.

**Proposition 4.1.1** If  $f_n \rightrightarrows f$  and each  $f_n$  is continuous at  $x$ , then  $f$  is continuous at  $x$ .

*Proof.* Given  $\epsilon > 0$ , choose  $N$  such that for all  $n \geq N$  and for all  $y \in [a, b]$ ,

$$|f_n(y) - f(y)| < \frac{\epsilon}{3}.$$

Since  $f_N$  is continuous at  $x$ ,  $\exists \delta > 0$  such that for all  $y \in [a, b]$  with  $|y - x| < \delta$ ,

$$|f_N(y) - f_N(x)| < \frac{\epsilon}{3}.$$

Therefore, if  $|y - x| < \delta$ , then

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Hence proved.  $\square$

**Remark** Even if  $f_n \rightarrow f$  pointwise and the limit  $f$  is continuous,  $f_n$  may not converge

---

uniformly to  $f$ . For example, let

$$f_n(x) = \begin{cases} nx, & 0 \leq x \leq \frac{1}{n}, \\ 2 - nx, & \frac{1}{n} < x \leq \frac{2}{n}, \\ 0, & \text{otherwise,} \end{cases}$$

then  $f_n \rightarrow 0$  pointwise, but not uniformly.

Let

$$C_b = C_b([a, b], \mathbb{R}) = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ is bounded}\}.$$

Define the supremum norm on  $C_b$  by

$$\|f\| = \sup_{x \in [a, b]} |f(x)|.$$

In order to distinguish from other norms, sometimes we write  $\|f\|_\infty$  for the supremum norm. Define a metric on  $C_b$  by  $d(f, g) = \|f - g\|$ .

**Proposition 4.1.2** Convergence with respect to the supremum norm is equivalent to uniform convergence.

*Proof.*  $f_n \rightharpoonup f$  iff for every  $\epsilon > 0$ ,  $\exists N \in \mathbb{N}$  such that for all  $n \geq N$  and for all  $x \in [a, b]$ ,

$$|f_n(x) - f(x)| < \epsilon \iff \sup_{x \in [a, b]} |f_n(x) - f(x)| < \epsilon \iff \|f_n - f\| < \epsilon.$$

Hence proved. □

**Theorem 4.1.1**  $C_b$  is a complete metric space with respect to the supremum norm.

*Proof.* Let  $(f_n)$  be a Cauchy sequence in  $C_b$ . For each fixed  $x \in [a, b]$ ,  $(f_n(x))$  is a Cauchy sequence in  $\mathbb{R}$  since

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|.$$

So for each  $x \in [a, b]$ , we can define

$$f(x) := \lim_{n \rightarrow \infty} f_n(x).$$

Clearly  $f_n \rightarrow f$  pointwise. Given  $\epsilon > 0$ , choose  $N$  such that for all  $m, n \geq N$ ,

$$\|f_n - f_m\| < \frac{\epsilon}{2}.$$

---

For each fixed  $x \in [a, b]$ , exists  $m(x) \geq N$  depending on  $x$  such that

$$|f_m(x) - f(x)| < \frac{\epsilon}{2}.$$

If  $n \geq N$  and  $x \in [a, b]$ , then

$$|f_n(x) - f(x)| \leq |f_n(x) - f_{m(x)}(x)| + |f_{m(x)}(x) - f(x)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Therefore, for all  $n \geq N$ ,

$$\|f_n - f\| = \sup_{x \in [a, b]} |f_n(x) - f(x)| \leq \epsilon,$$

hence  $f_n \rightarrow f$  uniformly. Since each  $f_n$  is bounded, so is  $f$ , thus  $f \in C_b$ . Hence proved.  $\square$

Let  $C^0 = C^0([a, b], \mathbb{R})$  be the space of continuous functions on  $[a, b]$ . Usually, people write  $C([a, b])$  or  $C([a, b], \mathbb{R})$  for  $C^0([a, b], \mathbb{R})$ . Since every continuous function on a compact set has a compact image,  $C^0 \subseteq C_b$ .

**Corollary 4.1.2**  $C^0$  is a closed subset of  $C_b$ , hence a complete metric space with respect to the supremum norm.

*Proof.* Converging sequences in  $C^0$  have continuous limits by Proposition 4.1.1.  $\square$

**Proposition 4.1.3** (Weierstrass M-test) If  $\sum_{n=1}^{\infty} M_n$  converges and if  $|f_n(x)| \leq M_n$  for all  $x \in [a, b]$  and for all  $n \in \mathbb{N}$ , then  $\sum_{n=1}^{\infty} f_n$  converges uniformly and absolutely on  $[a, b]$ .

*Proof.* Write  $F_n = \sum_{k=1}^n f_k$ . If  $n > m$ , then

$$\begin{aligned} d(F_n, F_m) &\leq d(F_n, F_{n-1}) + d(F_{n-1}, F_{n-2}) + \cdots + d(F_{m+1}, F_m) \\ &= \sum_{k=m+1}^n \|f_k\| \leq \sum_{k=m+1}^n M_k. \end{aligned}$$

Since  $\sum_{n=1}^{\infty} M_n$  converges,  $(F_n)$  is a Cauchy sequence in  $C_b$ , hence converges uniformly to some  $F \in C_b$ . Therefore,  $\sum_{n=1}^{\infty} f_n$  converges uniformly to  $F$ .  $\square$

**Proposition 4.1.4** The uniform limit of Riemann integrable functions is Riemann integrable, and the limit of the integrals is the integral of the limit, i.e.

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx.$$

*Proof.* Let  $f_n \in \mathcal{R}$  and assume that  $f_n \rightrightarrows f$ . Each  $f_n$  is bounded and exists  $Z_n$  of measure zero such that  $f_n$  is continuous on  $[a, b] \setminus Z_n$ . Let  $Z = \bigcup_{n=1}^{\infty} Z_n$ , then  $Z$  has measure zero and

---

$f$  is continuous on  $[a, b] \setminus Z$  by Proposition 4.1.1, hence  $f$  is Riemann integrable. To prove the equality, note that

$$\left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| \leq \int_a^b |f_n(x) - f(x)| dx \leq (b-a)\|f_n - f\| \rightarrow 0.$$

□

The above proposition tells us that the space of Riemann integrable functions  $\mathcal{R}$  is closed in  $C_b$ . Moreover, the functional

$$f \mapsto \int_a^b f(x) dx$$

is a continuous mapping from  $(\mathcal{R}, d)$  to  $\mathbb{R}$ .

**Proposition 4.1.5** The uniform limit of a sequence of differentiable functions is differentiable provided that the sequence of derivatives converges uniformly.

*Proof.* Suppose that  $f_n : [a, b] \rightarrow \mathbb{R}$  is a sequence of differentiable functions such that  $f_n \rightrightarrows f$  and  $f'_n \rightrightarrows g$ .

If each  $f'_n$  is continuous, then  $f'_n, g \in \mathcal{R}$  and we can apply the Fundamental Theorem of Calculus to get

$$f_n(x) = f_n(a) + \int_a^x f'_n(t) dt \rightrightarrows f(a) + \int_a^x g(t) dt.$$

Since  $f_n \rightrightarrows f$ , we have  $f(x) = f(a) + \int_a^x g(t) dt$ , hence  $f$  is differentiable with  $f' = g$ .

In general, fix some  $x \in [a, b]$  and define

$$\begin{aligned} \phi_n(t) &= \begin{cases} \frac{f_n(t) - f_n(x)}{t - x}, & t \neq x, \\ f'_n(x), & t = x, \end{cases} \\ \phi(t) &= \begin{cases} \frac{f(t) - f(x)}{t - x}, & t \neq x, \\ g(x), & t = x. \end{cases} \end{aligned}$$

Note that each  $\phi_n$  is continuous and  $\phi_n \rightarrow \phi$  pointwise. We claim that  $\phi_n \rightrightarrows \phi$ . For all  $m, n \in \mathbb{N}$  and  $t \neq x$ , we have

$$\phi_m(t) - \phi_n(t) = \frac{(f_m(t) - f_n(t)) - (f_m(x) - f_n(x))}{t - x}.$$

Apply the Mean Value Theorem to the function  $f_m - f_n$ , we see that there exists  $\xi$  between  $x$  and  $t$  such that

$$\phi_m(t) - \phi_n(t) = (f_m - f_n)'(\xi).$$

Since  $f'_n \rightrightarrows g$ ,  $(f'_n)$  is a Cauchy sequence in  $C_b$ , hence  $\phi_n$  is a Cauchy sequence in  $C_b$  as well. By

---

completeness of  $C_b$ ,  $\phi_n \rightrightarrows \psi$  for some continuous function  $\psi$ . Since  $\phi_n \rightarrow \phi$  pointwise, we have  $\phi = \psi$ , in particular,  $\phi$  is continuous. Therefore,  $f$  is differentiable at  $x$  with  $f'(x) = g(x)$ .  $\square$

**Remark** If we do not assume the uniform convergence of the derivatives, then the limit function may not be differentiable. For example, let  $f_n(x) = \sqrt{x^2 + \frac{1}{n}}$ , then  $f_n \rightarrow |x|$  uniformly, but  $|x|$  is not differentiable at 0. The derivatives converge pointwise but not uniformly.

## 4.2 Compactness and Equicontinuity in $C^0$

Recall the Heine-Borel Theorem: a subset of  $\mathbb{R}^n$  is compact iff it is closed and bounded. In general, this is not true for  $C^0$ . Consider the closed unit ball

$$B = \{f \in C^0([0, 1], \mathbb{R}) \mid \|f\| \leq 1\}.$$

$B$  is closed and bounded, but not compact. The problem is that  $C^0$  is infinite-dimensional.

**Definition 4.2.1** A family of functions  $(f_\alpha)_{\alpha \in I}$  in  $C^0$  is equicontinuous if for every  $\epsilon > 0$ ,  $\exists \delta > 0$  such that for all  $\alpha \in I$  and for all  $|x - y| < \delta$ ,

$$|f_\alpha(x) - f_\alpha(y)| < \epsilon.$$

Note that  $\delta$  does not depend on  $\alpha$ .

**Theorem 4.2.2** (Arzelá-Ascoli) Every bounded equicontinuous sequence of functions in  $C^0([a, b], \mathbb{R})$  has a subsequence that converges uniformly on  $[a, b]$ .

*Proof.* Let  $\mathcal{D} = \{d_1, d_2, \dots\}$  be a countable dense subset of  $[a, b]$ , for example, the set of rational numbers in  $[a, b]$ . Let  $(f_n)$  be a bounded equicontinuous sequence in  $C^0([a, b], \mathbb{R})$ .

There exists some  $M > 0$  such that  $|f_n(x)| \leq M$  for all  $x \in [a, b]$  and for all  $n \in \mathbb{N}$ . Since  $(f_n(d_1))$  is a bounded sequence in  $\mathbb{R}$ , it has a convergent subsequence  $(f_{n(1,k)}(d_1)) \rightarrow y_1$  as  $k \rightarrow \infty$ . Write  $f_k^{(1)}$  for  $f_{n(1,k)}$ . Find a subsequence  $(f_k^{(2)})$  of  $(f_k^{(1)})$  such that  $(f_k^{(2)}(d_2))$  converges to some  $y_2$ . Continuing this process, we obtain a nested sequence of subsequences  $(f_k^{(m)})$ , and if  $j \leq m$  then  $f_k^{(m)}(d_j)$  converges to  $y_j$  as  $k \rightarrow \infty$ . Define  $g_m = f_m^{(m)}$ . We claim that  $g_m$  converges pointwise on  $\mathcal{D}$ .

Fix  $j \in \mathbb{N}$ . Since  $(f_k^{(m)})$  is a nested sequence of subsequences, for all  $m \geq j$ , we have

$$f_m^{(m)} = f_{r_1}^{(m-1)} = f_{r_2}^{(m-2)} = \cdots = f_r^{(j)},$$

where  $r \geq r_{m-j-1} \geq \cdots \geq r_1 \geq m$ . Then

$$g_m(d_j) = f_m^{(m)}(d_j) = f_r^{(j)}(d_j) \rightarrow y_j \text{ as } m \rightarrow \infty,$$

---

hence  $g_m$  converges pointwise on  $\mathcal{D}$ .

We then show that  $(g_m)$  is a Cauchy sequence in  $C^0$ . Given  $\epsilon > 0$ , by equicontinuity,  $\exists \delta > 0$  such that for all  $x, y \in [a, b]$  with  $|x - y| < \delta$  and for all  $m \in \mathbb{N}$ ,

$$|g_m(x) - g_m(y)| < \frac{\epsilon}{3}.$$

Since  $[a, b]$  is compact and  $\mathcal{D}$  is dense in  $[a, b]$ , we can find a finite subset  $\{d_1, d_2, \dots, d_J\} \subseteq \mathcal{D}$  such that every  $x \in [a, b]$  is within  $\delta$  of some  $d_j$ . Since  $g_m(d_j) \rightarrow y_j$  for each  $j$ ,  $\exists N \in \mathbb{N}$  such that for all  $l, m \geq N$  and for all  $1 \leq j \leq J$ ,

$$|g_m(d_j) - g_l(d_j)| < \frac{\epsilon}{3}.$$

Thus, if  $l, m \geq N$  and  $x \in [a, b]$ , then choose some  $d_j$  such that  $|x - d_j| < \delta$ , we have

$$\begin{aligned} |g_m(x) - g_l(x)| &\leq |g_m(x) - g_m(d_j)| + |g_m(d_j) - g_l(d_j)| + |g_l(d_j) - g_l(x)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Therefore,  $(g_m)$  is a Cauchy sequence in  $C^0$ , hence converges in  $C^0$ , i.e. uniformly on  $[a, b]$ .  $\square$

**Remark** The theorem can be regarded as a characterization of compactness on  $C^0$ . If  $(f_n)$  is a sequence of bounded equicontinuous functions, then the set  $\{f_n \mid n \in \mathbb{N}\}$  is precompact, i.e. its closure is compact.

**Proposition 4.2.1** If  $(f_n)$  is an equicontinuous sequence in  $C^0$  that converges pointwise on a dense subset of  $[a, b]$ , then it converges uniformly on  $[a, b]$ .

*Proof.* Similar to the second part of the proof of the Arzelá-Ascoli Theorem.  $\square$

**Corollary 4.2.3** Assume that  $f_n : [a, b] \rightarrow \mathbb{R}$  is a sequence of differentiable functions such whose derivatives are uniformly bounded, i.e.  $\exists M > 0$  such that  $|f'_n(x)| \leq M$  for all  $n \in \mathbb{N}$  and for all  $x \in [a, b]$ . If for a point  $x_0 \in [a, b]$ , the sequence  $(f_n(x_0))$  is bounded, then  $(f_n)$  has a uniformly convergent subsequence.

*Proof.* To use Arzelá-Ascoli, we need to show that  $(f_n)$  is bounded and equicontinuous. Let  $M$  be a bound of the derivatives. For any  $x, y \in [a, b]$ , by the Mean Value Theorem,  $\exists \xi$  between  $x$  and  $y$  such that

$$|f_n(x) - f_n(y)| = |f'_n(\xi)||x - y| \leq M|x - y|.$$

Thus,  $(f_n)$  is equicontinuous.

Let  $K$  be a bound of  $|f_n(x_0)|$ . Then for any  $x \in [a, b]$ , by the Mean Value Theorem again,

---

$\exists \xi$  between  $x$  and  $x_0$  such that

$$\begin{aligned}|f_n(x)| &\leq |f_n(x_0)| + |f_n(x) - f_n(x_0)| \\&= |f_n(x_0)| + |f'_n(\xi)||x - x_0| \\&\leq K + M(b - a).\end{aligned}$$

Therefore,  $(f_n)$  is bounded. By Arzelá-Ascoli,  $(f_n)$  has a uniformly convergent subsequence.  $\square$

**Theorem 4.2.4** A subset  $\mathcal{E} \subseteq C^0$  is compact iff it is closed, bounded and equicontinuous.

*Proof.* Assume that  $\mathcal{E}$  is compact. Then  $\mathcal{E}$  is closed and bounded. Let  $\epsilon > 0$ . Since  $\mathcal{E}$  is compact, there exists a finite cover of  $\mathcal{E}$  by open balls  $B(f_i, \epsilon/3)$ ,  $i = 1, 2, \dots, N$ . Each  $f_i$  is uniformly continuous (continuous functions on compact sets are uniformly continuous), hence  $\exists \delta > 0$  such that for all  $x, y \in [a, b]$  with  $|x - y| < \delta$ ,

$$|f_i(x) - f_i(y)| < \frac{\epsilon}{3}, \quad i = 1, 2, \dots, N.$$

If  $f \in \mathcal{E}$ , then  $f \in B(f_k, \epsilon/3)$  for some  $k$ , hence for all  $x, y \in [a, b]$  with  $|x - y| < \delta$ ,

$$\begin{aligned}|f(x) - f(y)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

Therefore,  $\mathcal{E}$  is equicontinuous.

Conversely, assume that  $\mathcal{E}$  is closed, bounded and equicontinuous. Let  $(f_n)$  be a sequence in  $\mathcal{E}$ . By Arzelá-Ascoli,  $(f_n)$  has a uniformly convergent subsequence  $(f_{n_k})$  that converges to some  $f \in C^0$ . Since  $\mathcal{E}$  is closed,  $f \in \mathcal{E}$ . Therefore, every sequence in  $\mathcal{E}$  has a convergent subsequence whose limit is in  $\mathcal{E}$ , hence  $\mathcal{E}$  is compact.  $\square$

We can generalize the above theorem to general metric spaces. Let  $X$  be a metric space, and  $Y$  be a complete metric space. Let  $\mathcal{F}$  denote the set of functions from  $X$  to  $Y$ . Consider

$$\rho(y, y') = \frac{d_Y(y, y')}{1 + d_Y(y, y')}, \quad y, y' \in Y,$$

then  $\rho$  is a bounded metric on  $Y$  that induces the same topology as  $d_Y$ . In particular,  $(Y, \rho)$  is also complete. Define a metric on  $\mathcal{F}$  by

$$d(f, g) = \sup_{x \in X} \frac{\rho(f(x), g(x))}{1 + \rho(f(x), g(x))}, \quad f, g \in \mathcal{F}.$$

We say that a function  $f \in \mathcal{F}$  is bounded with respect to  $d_Y$  iff for any constant function  $c$ , we have  $\sup_{x \in X} d_Y(f(x), c) < \infty$ , or equivalently,  $d(f, c) < 1$ . A function  $f \in \mathcal{F}$  is unbounded if  $d(f, c) = 1$  for some constant function  $c$ .

---

**Proposition 4.2.2** In  $(\mathcal{F}, d)$ , the following holds:

- (a) Uniform convergence of  $(f_n)$  is equivalent to convergence with respect to  $d$ .
- (b) Completeness of  $Y$  implies completeness of  $\mathcal{F}$ .
- (c) The set  $\mathcal{F}_b$  of bounded functions in  $\mathcal{F}$  is closed in  $\mathcal{F}$ .
- (d) The set  $C^0(X, Y)$  of continuous functions in  $\mathcal{F}$  is closed in  $\mathcal{F}$ .

*Proof.*

- (a)  $f$  being the uniform limit of  $(f_n)$  is equivalent to  $d_Y(f_n(x), f(x)) \rightharpoonup 0$ , which is equivalent to  $d$ -convergence.
- (b) Similar to the proof of completeness of  $C_b$ . Let  $(f_n)$  be a Cauchy sequence in  $\mathcal{F}$ . Completeness of  $Y$  implies  $f(x) := \lim_{n \rightarrow \infty} f_n(x)$  exists for each  $x \in X$ . Cauchyness of  $(f_n)$  implies uniform convergence and  $f_n \rightarrow f$  with respect to  $d$ .
- (c) If  $f_n \in \mathcal{F}_b$  and  $d(f_n, f) \rightarrow 0$ , then for any constant function  $c$ ,

$$d(f, c) \leq d(f, f_n) + d(f_n, c) < 1,$$

for sufficiently large  $n$ , hence  $f \in \mathcal{F}_b$ .

- (d) If  $f_n \in C^0(X, Y)$  and  $d(f_n, f) \rightarrow 0$ , then for each fixed  $x \in X$ ,  $f_n(x) \rightarrow f(x)$ . Given  $\epsilon > 0$ , choose  $N$  such that for all  $n \geq N$  and for all  $x \in X$ ,  $d(f_n(x), f(x)) < \epsilon/3$ . Since  $f_N$  is continuous at  $x$ ,  $\exists \delta > 0$  such that for all  $y \in X$  with  $d_X(x, y) < \delta$ ,  $d_Y(f_N(x), f_N(y)) < \epsilon/3$ . Therefore, if  $d_X(x, y) < \delta$ , then for all  $n \geq N$ ,

$$\begin{aligned} d_Y(f(x), f(y)) &\leq d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(y)) + d_Y(f_n(y), f(y)) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

Hence  $f$  is continuous at  $x$ .

□

### Definition 4.2.5

- A family of functions  $\mathcal{E} \subseteq \mathcal{F}$  is equicontinuous if for every  $\epsilon > 0$ , exists  $\delta > 0$  such that for all  $f \in \mathcal{E}$  and for all  $x, x' \in X$  with  $d_X(x, x') < \delta$ ,

$$d_Y(f(x), f(x')) < \epsilon.$$

- $\mathcal{E}$  is pointwise equicontinuous if the  $\delta$  above may depend on  $x$ , i.e. for every  $\epsilon > 0$  and for

---

every  $x \in X$ , exists  $\delta > 0$  such that for all  $f \in \mathcal{E}$  and for all  $x' \in X$  with  $d_X(x, x') < \delta$ ,

$$d_Y(f(x), f(x')) < \epsilon.$$

**Proposition 4.2.3** If  $X$  is compact, then equicontinuity and pointwise equicontinuity of  $\mathcal{E}$  are equivalent.

*Proof.* Suppose not, then  $\exists \epsilon > 0$  such that for all  $\delta = 1/n$ ,  $\exists x_n, x'_n \in X$  and functions  $f_n \in \mathcal{E}$  such that  $d_X(x_n, x'_n) < 1/n$  but  $d_Y(f_n(x_n), f_n(x'_n)) \geq \epsilon$ . By compactness of  $X$ , exists a subsequence  $(x_{n_k})$  that converges to some  $x_0 \in X$ . Then  $x'_{n_k} \rightarrow x_0$  as well. By pointwise equicontinuity,  $\exists \delta > 0$  such that  $d_X(x, x_0) < \delta$  implies  $d_Y(f(x), f(x_0)) < \epsilon/2$  for all  $f \in \mathcal{E}$ . Take  $x = x_{n_k}, x'_{n_k}$  and  $f = f_{n_k}$  for sufficiently large  $k$ , then

$$\begin{aligned} \epsilon &\leq d_Y(f_{n_k}(x_{n_k}), f_{n_k}(x'_{n_k})) \\ &\leq d_Y(f_{n_k}(x_{n_k}), f_{n_k}(x_0)) + d_Y(f_{n_k}(x_0), f_{n_k}(x'_{n_k})) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned}$$

a contradiction.  $\square$

**Proposition 4.2.4** If the sequence of functions  $(f_n)$  in  $\mathcal{F}$  is uniformly equicontinuous,  $X$  is compact, and for each  $x \in X$ , the sequence  $(f_n(x))$  lies in a compact subset of  $Y$ , then  $(f_n)$  has a uniformly convergent subsequence.

*Proof.* Since  $X$  is compact, it is separable, i.e. has a countable dense subset. The rest is similar to the proof of the Arzelá-Ascoli Theorem.  $\square$

**Definition 4.2.6** A topological space  $X$  is  $\sigma$ -compact if it is a countable union of compact subsets.

**Example 4.2.1**  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$  are  $\sigma$ -compact. Uncountable sets equipped with the discrete topology are not  $\sigma$ -compact.

**Proposition 4.2.5** If  $X$  is  $\sigma$ -compact and if  $(f_n)$  is a sequence of pointwise equicontinuous functions in  $\mathcal{F}$  such that for each  $x \in X$ , the sequence  $(f_n(x))$  lies in a compact subset of  $Y$ , then  $(f_n)$  has a subsequence that converges uniformly on every compact subset of  $X$ .

*Proof.* Write  $X = \bigcup_{i=1}^{\infty} X_i$ , where each  $X_i$  is compact. For each  $i$ ,  $(f_n|_{X_i})$  is a sequence of uniformly equicontinuous functions. By the previous proposition,  $(f_n|_{X_1})$  has a uniformly convergent subsequence  $(f_n^{(1)})$ . This subsequence has a further subsequence  $(f_n^{(2)})$  that converges uniformly on  $X_2$ . Continuing this process, we obtain a nested sequence of subsequences  $(f_n^{(i)})$ . Define  $g_n = f_n^{(n)}$ , then  $g_n$  converges uniformly on each  $X_i$ .

---

Note that  $g_n$  converges pointwise on  $X$  since for each fixed  $x \in X$ , exists some  $X_i$  containing  $x$ . Let  $A \subseteq X$  be any compact subset, then  $(g_n|_A)$  is uniformly equicontinuous and converges pointwise on  $A$ . We can upgrade the pointwise convergence to uniform convergence on  $A$  as in the proof of the Arzelá-Ascoli Theorem (the  $\epsilon/3$ -argument).  $\square$

**Corollary 4.2.7** If  $(f_n)$  is a sequence of pointwise equicontinuous functions from  $\mathbb{R}$  to  $\mathbb{R}$ , and exists  $x_0 \in \mathbb{R}$  such that  $(f_n(x_0))$  is bounded, then  $(f_n)$  has a subsequence that converges uniformly on every compact subset of  $\mathbb{R}$ .

*Proof.* By the previous proposition, it suffices to show that for each  $x \in \mathbb{R}$ ,  $(f_n(x))$  is bounded and hence lies in a compact subset of  $\mathbb{R}$ .

Let  $[a, b]$  be any closed interval containing  $x_0$ , then  $(f_n)|_{[a, b]}$  is uniformly equicontinuous. Exists  $\delta > 0$  such that for all  $s, t \in [a, b]$  with  $|s - t| < \delta$  and for all  $n \in \mathbb{N}$ ,  $|f_n(s) - f_n(t)| < 1$ . Choose points  $a = x_0 < x_1 < x_2 < \dots < x_m = b$  such that  $|x_i - x_{i-1}| < \delta$  for all  $1 \leq i \leq m$ . Since  $(f_n(x_0))$  is bounded, exists  $K > 0$  such that  $|f_n(x_0)| \leq K$  for all  $n \in \mathbb{N}$ . Then for all  $n \in \mathbb{N}$  and for all  $1 \leq i \leq m$ ,

$$|f_n(x_i)| \leq |f_n(x_{i-1})| + |f_n(x_i) - f_n(x_{i-1})| < K + i.$$

Therefore, for all  $x \in [a, b]$  and for all  $n \in \mathbb{N}$ ,

$$|f_n(x)| \leq |f_n(x_i)| + |f_n(x) - f_n(x_i)| < K + m + 1,$$

where  $x_i$  is chosen such that  $|x - x_i| < \delta$ . Hence  $(f_n(x))$  is bounded for all  $x \in \mathbb{R}$ .  $\square$

### 4.3 Uniform Approximation in $C^0$

Given a function  $f \in C^0([a, b], \mathbb{R})$ , we would like to approximate it  $f$  by simpler functions that we can work with more easily.

**Theorem 4.3.1** (Weierstrass Approximation Theorem) The set of polynomial functions is dense in  $C^0([a, b], \mathbb{R})$  with respect to the supremum norm, i.e. for every  $f \in C^0([a, b], \mathbb{R})$  and for every  $\epsilon > 0$ , exists a polynomial function  $P$  such that

$$\|f - P\| < \epsilon.$$

*Proof.* WLOG, assume that  $[a, b] = [0, 1]$ . We may also assume that  $f(0) = f(1) = 0$  by considering  $f(x) - (f(0)(1-x) + f(1)x)$  instead of  $f$ . Extend  $f$  to  $\mathbb{R}$  by defining  $f(x) = 0$  for  $x \notin [0, 1]$ . Consider

$$\beta_n(t) = b_n(1-t^2)^n, \quad -1 \leq t \leq 1,$$

---

where  $b_n$  is chosen such that  $\int_{-1}^1 \beta_n(t) dt = 1$ . For  $x \in [0, 1]$ , define

$$P_n(x) = \int_{-1}^1 f(x+t) \beta_n(t) dt.$$

We claim that  $P_n$  is a polynomial function and  $P_n \rightrightarrows f$  on  $[0, 1]$  as  $n \rightarrow \infty$ .

Let  $u = x + t$ , then

$$\begin{aligned} P_n(x) &= \int_{x-1}^{x+1} f(u) \beta_n(u-x) du \\ &= \int_0^1 f(u) \beta_n(u-x) du, \end{aligned}$$

since  $f(u) = 0$  for  $u \notin [0, 1]$ . Note that  $\beta_n(u-x) = b_n(1-(u-x)^2)^n$  is a polynomial in  $x$  and  $u$ , hence after integrating with respect to  $u$ ,  $P_n(x)$  is a polynomial in  $x$ .

Note that

$$\begin{aligned} 1 &= \int_{-1}^1 \beta_n(t) dt \\ &\geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} b_n(1-t^2)^n dt \\ &= 2b_n \int_0^{1/\sqrt{n}} (1-t^2)^n dt \\ &\geq 2b_n \int_0^{1/\sqrt{n}} \left(1 - \frac{1}{n}\right)^n dt \\ &= \frac{2b_n}{\sqrt{n}} \left(1 - \frac{1}{n}\right)^n, \end{aligned}$$

so  $\exists c > 0$  such that  $b_n \leq c\sqrt{n}$  for all  $n \in \mathbb{N}$ . For any  $\delta > 0$ , if  $\delta \leq |t| \leq 1$ , then

$$\beta_n(t) = b_n(1-t^2)^n < c\sqrt{n}(1-\delta^2)^n \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Let  $\epsilon > 0$ . Since  $f$  is uniformly continuous on  $[0, 1]$ ,  $\exists \delta > 0$  such that for all  $x, y \in [0, 1]$  with  $|x-y| < \delta$ ,  $|f(x)-f(y)| < \epsilon/2$ . For the above  $\delta$ , choose  $N$  such that for all  $n \geq N$  and for

---

all  $|t| \geq \delta$ ,  $\beta_n(t) < \frac{\epsilon}{4M}$ , where  $M = \sup_{x \in [0,1]} |f(x)|$ . Then for all  $n \geq N$  and for all  $x \in [0, 1]$ ,

$$\begin{aligned} |P_n(x) - f(x)| &= \left| \int_{-1}^1 (f(x+t) - f(x))\beta_n(t) dt \right| \\ &= \left| \int_{-1}^1 (f(x+t) - f(x))\beta_n(t) dt \right| \\ &\leq \int_{|t|<\delta} |f(x+t) - f(x)|\beta_n(t) dt + \int_{|t|\geq\delta} |f(x+t) - f(x)|\beta_n(t) dt \\ &\leq \frac{\epsilon}{2} \cdot 1 + 2M \cdot \frac{\epsilon}{4M} = \epsilon. \end{aligned}$$

Therefore,  $P_n \rightrightarrows f$  on  $[0, 1]$ . □

We now extend this to functions defined on a compact Hausdorff space  $X$ .

#### Definition 4.3.2

- A subset  $\mathcal{A}$  of  $C^0(X) = C^0(X, \mathbb{R})$  is an algebra if it is closed under addition, scalar multiplication and function multiplication.
- An algebra  $\mathcal{A} \subset C^0(X)$  vanishes at a point  $x \in X$  if  $f(x) = 0$  for all  $f \in \mathcal{A}$ .
- An algebra  $\mathcal{A} \subset C^0(X)$  separates points if for any distinct points  $x_1, x_2 \in X$ , exists some  $f \in \mathcal{A}$  such that  $f(x_1) \neq f(x_2)$ .

**Lemma 4.3.3** If  $\mathcal{A}$  is an algebra in  $C^0(X)$  that vanishes nowhere and separates points, then for any distinct points  $x_1, x_2 \in X$  and for any  $c_1, c_2 \in \mathbb{R}$ , exists some  $f \in \mathcal{A}$  such that  $f(x_1) = c_1$  and  $f(x_2) = c_2$ .

*Proof.* Let  $x_1, x_2 \in X$  and  $c_1, c_2 \in \mathbb{R}$ . Since  $\mathcal{A}$  vanishes nowhere, exists  $g_1, g_2 \in \mathcal{A}$  such that  $g_1(x_1) \neq 0$  and  $g_2(x_2) \neq 0$ , then  $g_1^2 + g_2^2 \in \mathcal{A}$  and  $(g_1^2 + g_2^2)(x_i) > 0$  for  $i = 1, 2$ . Since  $\mathcal{A}$  separates points, exists  $h \in \mathcal{A}$  such that  $h(x_1) \neq h(x_2)$ . Note that the matrix

$$\begin{pmatrix} g(x_1) & g(x_1)h(x_1) \\ g(x_2) & g(x_2)h(x_2) \end{pmatrix}$$

is invertible, hence exists  $a, b \in \mathbb{R}$  such that

$$\begin{cases} g(x_1)a + g(x_1)h(x_1)b = c_1 \\ g(x_2)a + g(x_2)h(x_2)b = c_2 \end{cases}.$$

Define  $f = ag + bgh \in \mathcal{A}$ , then  $f(x_1) = c_1$  and  $f(x_2) = c_2$ . □

**Lemma 4.3.4** Let  $\mathcal{A}$  be an algebra in  $C^0(X)$ . If  $f \in \overline{\mathcal{A}}$ , then  $|f| \in \overline{\mathcal{A}}$ .

---

*Proof.* We need to show that  $|f|$  can be uniformly approximated by functions in  $\overline{\mathcal{A}}$ . Let  $\epsilon > 0$ . By the Weierstrass Approximation Theorem, exists a polynomial  $p$  such that

$$|p(y) - y| < \frac{\epsilon}{2}$$

for all  $y \in [-M, M]$ , where  $M = \|f\|$ . Note that  $|p(0)| = |p(0) - 0| < \epsilon/2$ . Set  $q(y) = p(y) - p(0)$ , then  $q(0) = 0$  and

$$\begin{aligned} |q(y) - |y|| &= |p(y) - p(0) - |y|| \\ &\leq |p(y) - y| + |y| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

for all  $y \in [-M, M]$ . Define  $g(x) = q(f(x))$  for all  $x \in X$ , then  $g \in \overline{\mathcal{A}}$  since  $\mathcal{A}$  is an algebra and  $q$  is a polynomial with zero constant term. If  $x \in X$ , then

$$|g(x) - |f(x)|| = |q(f(x)) - |f(x)|| < \epsilon,$$

thus  $|f| \in \overline{\mathcal{A}} = \overline{\mathcal{A}}$ . □

**Lemma 4.3.5** Let  $\mathcal{A}$  be an algebra in  $C^0(X)$ . If  $f, g \in \overline{\mathcal{A}}$ , then  $\max(f, g), \min(f, g) \in \overline{\mathcal{A}}$ .

*Proof.* Note that

$$\max(f, g) = \frac{f + g + |f - g|}{2}, \quad \min(f, g) = \frac{f + g - |f - g|}{2}.$$

By the previous lemma,  $|f - g| \in \overline{\mathcal{A}}$ , hence  $\max(f, g), \min(f, g) \in \overline{\mathcal{A}}$ . □

**Theorem 4.3.6** (Stone-Weierstrass Theorem) Let  $X$  be a compact Hausdorff space. If  $\mathcal{A}$  is an algebra of  $C^0(X)$  that vanishes nowhere and separates points, then  $\mathcal{A}$  is dense in  $C^0(X)$  with respect to the supremum norm.

*Proof.* Let  $\mathcal{A}$  be such an algebra. Given  $f \in C^0(X)$  and  $\epsilon > 0$ , we want to find some  $g \in \mathcal{A}$  such that  $\|f - g\| < \epsilon$ .

Fix distinct  $x, y \in X$ . By 4.3.3, exists some  $h_{x,y} \in \mathcal{A}$  such that

$$h_{x,y}(x) = f(x), \quad h_{x,y}(y) = f(y).$$

Define

$$\mathcal{U}_{x,y} = \{z \in X \mid h_{x,y}(z) < f(z) + \epsilon\}.$$

Then  $\mathcal{U}_{x,y}$  is an open set containing  $y$ . The collection  $\{\mathcal{U}_{x,y} \mid y \in X \setminus \{x\}\}$  is an open cover of

---

$X$ , hence exists a finite subcover  $\{\mathcal{U}_{x,y_1}, \mathcal{U}_{x,y_2}, \dots, \mathcal{U}_{x,y_n}\}$ . Define

$$g_x = \min \{h_{x,y_1}, \dots, h_{x,y_n}\} \in \overline{\mathcal{A}}.$$

For  $x \in X$ , define

$$V_x = \{z \in X \mid g_x(z) > f(z) - \epsilon\}.$$

Then  $V_x$  is an open set containing  $x$ . The collection  $\{V_x \mid x \in X\}$  is an open cover of  $X$ , hence exists a finite subcover  $\{V_{x_1}, V_{x_2}, \dots, V_{x_m}\}$ . Define

$$\begin{aligned} g &= \max \{g_{x_1}, g_{x_2}, \dots, g_{x_m}\} \in \overline{\mathcal{A}}, \\ \implies f(z) - \epsilon &< g(z) < f(z) + \epsilon \text{ for all } z \in X. \end{aligned}$$

Since any function in  $\overline{\mathcal{A}}$  can be uniformly approximated by functions in  $\mathcal{A}$ , this proves the theorem.  $\square$

**Remark** The Hausdorff condition is necessary to ensure that there exists a subalgebra that separates points. Suppose that  $X$  is not Hausdorff, and let  $x_1, x_2 \in X$  be distinct points that cannot be separated by open neighborhoods. Let  $\mathcal{A}$  be a subalgebra of  $C^0(X)$  that separates points. Then exists some  $f \in \mathcal{A}$  such that  $f(x_1) \neq f(x_2)$ . Since  $\mathbb{R}$  is Hausdorff, exists open neighborhoods  $U_1, U_2$  of  $f(x_1), f(x_2)$  respectively such that  $U_1 \cap U_2 = \emptyset$ . Then  $f^{-1}(U_1), f^{-1}(U_2)$  are open neighborhoods of  $x_1, x_2$  respectively such that  $f^{-1}(U_1) \cap f^{-1}(U_2) = \emptyset$ , a contradiction.

**Corollary 4.3.7** Any  $2\pi$ -periodic continuous function can be uniformly approximated by trigonometric polynomials, i.e. finite linear combinations of  $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots\}$ .

*Proof.* Every  $2\pi$ -periodic continuous function can be regarded as a continuous function on the unit circle  $S^1$ . The set of trigonometric polynomials forms an algebra that vanishes nowhere and separates points on  $S^1$ . By the Stone-Weierstrass Theorem, the set of trigonometric polynomials is dense in  $C^0(S^1)$  with respect to the supremum norm.  $\square$

## 4.4 Banach Fixed Point Theorem

**Definition 4.4.1** Let  $M$  be a metric space and  $f$  be a function from  $M$  to itself. We say that  $x \in M$  is a fixed point of  $f$  if  $f(x) = x$ .

**Definition 4.4.2** A function  $f : M \rightarrow M$  is a contraction mapping if exists some constant  $0 \leq \alpha < 1$  such that for all  $x, y \in M$ ,

$$d(f(x), f(y)) \leq \alpha d(x, y).$$

---

**Theorem 4.4.3** (Banach fixed point theorem) Suppose that  $f : M \rightarrow M$  is a contraction and  $M$  is complete. Then  $f$  has a unique fixed point in  $M$ . Moreover, for any  $y \in M$ , the sequence defined by  $y_0 = y$  and  $y_{n+1} = f(y_n)$  converges to the fixed point.

*Proof.* Choose any  $y \in M$  and define the sequence  $(y_n)$  by  $y_0 = y$  and  $y_{n+1} = f(y_n)$ , then

$$\begin{aligned} d(y_n, y_{n+1}) &= d(f(y_{n-1}), f(y_n)) \\ &\leq \alpha d(y_{n-1}, y_n) \\ &\leq \dots \\ &\leq \alpha^n d(y_0, y_1). \end{aligned}$$

Let  $\epsilon > 0$ . Choose  $N$  such that  $\alpha^N d(y_0, y_1) < (1 - \alpha)\epsilon$ . Then for all  $N \leq m < n$ ,

$$\begin{aligned} d(y_m, y_n) &\leq d(y_m, y_{m+1}) + d(y_{m+1}, y_{m+2}) + \dots + d(y_{n-1}, y_n) \\ &\leq \sum_{k=m}^{n-1} \alpha^k d(y_0, y_1) \\ &\leq \frac{\alpha^m}{1-\alpha} d(y_0, y_1) < \epsilon. \end{aligned}$$

Therefore,  $(y_n)$  is a Cauchy sequence in  $M$ . By completeness of  $M$ , exists some  $x \in M$  such that  $y_n \rightarrow x$ . Since  $f$  is continuous,

$$f(x) = f\left(\lim_{n \rightarrow \infty} y_n\right) = \lim_{n \rightarrow \infty} f(y_n) = \lim_{n \rightarrow \infty} y_{n+1} = x,$$

hence  $x$  is a fixed point of  $f$ . For uniqueness, suppose that  $x'$  is another fixed point of  $f$ , then

$$d(x, x') = d(f(x), f(x')) \leq \alpha d(x, x') \rightsquigarrow d(x, x') = 0,$$

hence  $x = x'$ . □

Let  $\mathcal{U} \subseteq \mathbb{R}^n$  be an open set. A vector ODE on  $\mathcal{U}$  is given as  $m$  simultaneous scalar equations

$$\begin{cases} x'_1(t) = f_1(x_1(t), \dots, x_m(t)), \\ \vdots \\ x'_m(t) = f_m(x_1(t), \dots, x_m(t)), \end{cases}$$

where each  $f_i : \mathcal{U} \rightarrow \mathbb{R}$ . We seek  $m$  real-valued functions  $x_1(t), \dots, x_m(t)$  such that the above equations hold for all  $t$  and simultaneously. If such functions exist, we say that they solve the ODE with initial condition  $(x_1(0), \dots, x_m(0))$ .

$$F(x) = (f_1(x), f_2(x), \dots, f_m(x))$$

---

is called a vector field on  $\mathcal{U}$ . We would like to find a curve  $\gamma : (a, b) \rightarrow \mathcal{U}$  such that  $a < 0 < b$  and  $\gamma'(t) = F(\gamma(t))$  for all  $t \in (a, b)$ , with initial condition  $\gamma(0) = p \in \mathcal{U}$ . Components of  $\gamma$  give a solution to the vector ODE with initial condition  $p$ .

Physically,  $\gamma(t)$  represents the position of a particle at time  $t$  moving in the vector field  $F$ . The equation  $\gamma'(t) = F(\gamma(t))$  means that the velocity of the particle at time  $t$  is given by the vector  $F(\gamma(t))$ .

Assume that  $F$  is Lipschitz continuous, i.e. exists some constant  $L > 0$  such that for all  $x, y \in \mathcal{U}$ ,

$$\|F(x) - F(y)\| \leq L\|x - y\|.$$

**Theorem 4.4.4** (Picard) Given  $p \in \mathcal{U}$ , exists an  $F$ -trajectory  $\gamma(t)$  in  $\mathcal{U}$  through  $p$ . Locally,  $\gamma$  is unique.

*Proof.* We express  $\gamma'(t) = F(\gamma(t))$  as an integral equation. Given  $g(x) = (g_1(x), \dots, g_m(x))$  for  $x \in [a, b]$ , define

$$\int_a^b g(x) dx = \left( \int_a^b g_1(x) dx, \dots, \int_a^b g_m(x) dx \right).$$

Consider

$$\gamma(t) = p + \int_0^t F(\gamma(s)) ds.$$

Note that indefinite integrals are always continuous, hence  $\gamma$  is continuous. Since  $F$  is Lipschitz continuous, it is also continuous, hence  $F \circ \gamma$  is continuous and  $\gamma$  is differentiable with  $\gamma'(t) = F(\gamma(t))$ . Therefore, a solution to the integral equation is also a solution to the ODE.

Clearly  $F$  is continuous, hence exists some compact neighborhood  $K := \overline{B(p, r)} \subseteq \mathcal{U}$  and  $M > 0$  such that  $\|F(x)\| \leq M$  for all  $x \in K$ . Choose  $\tau > 0$  small such that  $\tau M \leq r$  and  $\tau L < 1$ . Let  $\mathcal{C} = C^0([-\tau, \tau], K)$ . Define a metric on  $\mathcal{C}$  by

$$d(\gamma, \sigma) = \sup \{ \|\gamma(t) - \sigma(t)\| \mid t \in [-\tau, \tau] \}.$$

Then  $(\mathcal{C}, d)$  is a complete metric space. Given  $\gamma \in \mathcal{C}$ , define

$$\Phi(\gamma)(t) = p + \int_0^t F(\gamma(s)) ds.$$

We first show that  $\Phi(\mathcal{C}) \subseteq \mathcal{C}$ . For  $\gamma \in \mathcal{C}$  and for  $t \in [-\tau, \tau]$ ,

$$\|\Phi(\gamma)(t) - p\| = \left\| \int_0^t F(\gamma(s)) ds \right\| \leq \int_0^{|t|} \|F(\gamma(s))\| ds \leq M|t| \leq M\tau \leq r,$$

---

hence  $\Phi(\gamma)(t) \in K$  and  $\Phi(\gamma) \in \mathcal{C}$ . Next, we show that  $\Phi$  is a contraction mapping. For  $\gamma, \sigma \in \mathcal{C}$ ,

$$\begin{aligned} d(\Phi(\gamma), \Phi(\sigma)) &= \sup_{t \in [-\tau, \tau]} \left\| \int_0^t (F(\gamma(s)) - F(\sigma(s))) ds \right\| \\ &\leq \sup_{t \in [-\tau, \tau]} \int_0^{|t|} \|F(\gamma(s)) - F(\sigma(s))\| ds \\ &\leq \sup_{t \in [-\tau, \tau]} \int_0^{|t|} L \|\gamma(s) - \sigma(s)\| ds \\ &\leq L\tau d(\gamma, \sigma). \end{aligned}$$

Since  $\tau L < 1$ ,  $\Phi$  is a contraction mapping. By the Banach Fixed Point Theorem, exists a unique fixed point  $\gamma \in \mathcal{C}$  such that  $\Phi(\gamma) = \gamma$ , i.e.,

$$\gamma(t) = p + \int_0^t F(\gamma(s)) ds.$$

Any other solution defined on  $[-\tau, \tau]$  must also be a fixed point of  $\Phi$ , hence  $\gamma$  is the unique solution on  $[-\tau, \tau]$ .  $\square$

For simplicity, assume  $\mathcal{U} = \mathbb{R}^m$ . Then  $\gamma$  can be defined for all  $t \in \mathbb{R}$ . Let  $\gamma(t, p)$  be the  $F$ -trajectory through  $p$  at time  $t$ . Define

$$\varphi_t(p) = \gamma(t, p),$$

then

$$\varphi_{t+s}(p) = \varphi_t(\varphi_s(p))$$

for all  $s, t \in \mathbb{R}$ . This is called the flow generated by the vector field  $F$ .  $t \mapsto \varphi_t(p)$  is a group homomorphism from  $(\mathbb{R}, +)$  to the group of motions on  $\mathbb{R}^m$  under composition.

## 4.5 Analytic Functions

Recall that the radius of convergence of a power series  $\sum_{n=0}^{\infty} c_n x^n$  is given by

$$R = \frac{1}{\limsup_{n \rightarrow \infty} |c_n|^{1/n}}.$$

A function  $f : (a, b) \rightarrow \mathbb{R}$  is analytic if it can be represented by a power series locally, i.e. for every  $x_0 \in (a, b)$ , exists some  $r > 0$  and coefficients  $(c_n)$  such that for all  $x \in (x_0 - r, x_0 + r)$ ,

$$f(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^n,$$

---

where the power series has radius of convergence at least  $r$ .

Let  $f : (a, b) \rightarrow \mathbb{R}$  be a smooth function. The Taylor series of  $f$  at  $x_0 \in (a, b)$  is given by

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n.$$

Denote

$$M_r = \sup_{x \in (a, b)} |f^{(r)}(x)|.$$

**Definition 4.5.1** The derivative growth rate of  $f$  is defined as

$$\alpha = \limsup_{r \rightarrow \infty} \left( \frac{M_r}{r!} \right)^{1/r},$$

We can see that if  $R$  is the radius of convergence of the Taylor series of  $f$  at  $x_0$ , then  $R \geq 1/\alpha$ .

**Lemma 4.5.2** If  $\alpha\sigma < 1$ , then the Taylor series of  $f$  at  $x$  converges uniformly to  $f$  on  $(x - \sigma, x + \sigma)$ .

*Proof.* Choose  $\delta > 0$  such that  $(\alpha + \delta)\sigma < 1$ . By Taylor's theorem, exists some  $\theta$  between  $x$  and  $x + h$  such that

$$f(x + h) - \sum_{k=0}^{r-1} \frac{f^{(k)}(x)}{k!} h^k = \frac{f^{(r)}(\theta)}{r!}.$$

For  $r$  large and  $|h| \leq \sigma$ , we have

$$\left| f(x + h) - \sum_{k=0}^{r-1} \frac{f^{(k)}(x)}{k!} h^k \right| \leq \frac{M_r}{r!} \sigma^r \leq ((\alpha + \delta)\sigma)^r$$

approaches 0 as  $r \rightarrow \infty$ . Therefore, the Taylor series converges uniformly to  $f$  on  $(x - \sigma, x + \sigma)$ .  $\square$

**Lemma 4.5.3** If  $f$  is expressed as a convergent power series  $f(x + h) = \sum_{n=0}^{\infty} c_n h^n$  with radius of convergence  $R > \sigma$ , then  $f$  has bounded derivative growth rate on  $(x - \sigma, x + \sigma)$ .

*Proof.* We use the fact that

$$\limsup_{r \rightarrow \infty} \left( \sum_{k=r}^{\infty} \binom{k}{r} \lambda^l \right)^{\frac{1}{r}} < \infty$$

for all  $0 < \lambda < 1$ . By definition of radius of convergence, exists some  $0 < \lambda < 1$  such that for

---

all sufficiently large  $n$ ,  $|c_n\sigma^n|^{1/n} < \lambda$ . For  $|h| \leq \sigma$ , we have

$$\begin{aligned}|f^{(r)}(x+h)| &= \left| \sum_{n=r}^{\infty} c_n \frac{n!}{(n-r)!} h^{n-r} \right| \\ &\leq \frac{r!}{\sigma^r} \sum_{n=r}^{\infty} \binom{n}{r} |c_n \sigma^n| \\ &\leq \frac{r!}{\sigma^r} \sum_{n=r}^{\infty} \binom{n}{r} \lambda^n\end{aligned}$$

for all sufficiently large  $r$ . Therefore,

$$\alpha = \left( \frac{M_r}{r!} \right)^{1/r} \leq \frac{1}{\sigma} \left( \sum_{n=r}^{\infty} \binom{n}{r} \lambda^n \right)^{1/r},$$

which is bounded as  $r \rightarrow \infty$ .  $\square$

**Theorem 4.5.4** A smooth function  $f : (a, b) \rightarrow \mathbb{R}$  is analytic if and only if it has locally bounded derivative growth rate on  $(a, b)$ .

*Proof.* " $\Leftarrow$ ": Assume that  $f : (a, b) \rightarrow \mathbb{R}$  is smooth and has locally bounded derivative growth rate. For any  $x \in (a, b)$ , exists some neighborhood  $V$  of  $x$  such that  $f$  has bounded derivative growth rate on  $V$ , say  $\alpha$ . Choose  $\sigma > 0$  such that  $(x - \sigma, x + \sigma) \subseteq V$  and  $\alpha\sigma < 1$ . By lemma 4.5.2, the Taylor series of  $f$  at  $x$  converges uniformly to  $f$  on  $(x - \sigma, x + \sigma)$ , hence  $f$  is analytic at  $x$ .

" $\Rightarrow$ ": Assume that  $f : (a, b) \rightarrow \mathbb{R}$  is analytic. For any  $x \in (a, b)$ , exists a power series  $\sum_{n=0}^{\infty} c_n h^n$  with radius of convergence  $R > 0$  such that for all  $h$  with  $|h| < R$ ,

$$f(x+h) = \sum_{n=0}^{\infty} c_n h^n.$$

By lemma 4.5.3,  $f$  has bounded derivative growth rate on  $(x - \sigma, x + \sigma)$  for any  $0 < \sigma < R$ . Therefore,  $f$  has locally bounded derivative growth rate on  $(a, b)$ .  $\square$

**Corollary 4.5.5** A smooth function is analytic if its derivatives are uniformly bounded.

**Example 4.5.1** The trigonometric functions  $\sin x$  and  $\cos x$  are analytic on  $\mathbb{R}$ .

**Corollary 4.5.6** If  $f(x) = \sum_{n=0}^{\infty} c_n x^n$  and the power series has radius of convergence  $R > 0$ , then  $f$  is analytic on  $(-R, R)$ .

*Proof.* By lemma 4.5.3,  $f$  has bounded derivative growth rate on  $(-\sigma, \sigma)$  for any  $0 < \sigma < R$ . Therefore,  $f$  is analytic on  $(-R, R)$ .  $\square$

---

**Remark** The above corollary shows that if a function can be expressed as a convergent power series at  $x = 0$ , then it can also be expressed as a convergent power series at any other point in its interval of convergence.

## 4.6 Nowhere Differentiable Continuous Functions

**Theorem 4.6.1** There exists a continuous  $f : \mathbb{R} \rightarrow \mathbb{R}$  that does not have derivative at any point.

*Proof.* Define  $\sigma_0 : \mathbb{R} \rightarrow \mathbb{R}$  by

$$\sigma_0(x) = \begin{cases} x - 2n & \text{if } x \in [2n, 2n+1] \text{ for some } n \in \mathbb{Z}, \\ 2n+2-x & \text{if } x \in [2n+1, 2n+2] \text{ for some } n \in \mathbb{Z}. \end{cases}$$

Then  $\sigma_0$  is a continuous function with period 2. Define

$$\sigma_k(x) = \left(\frac{3}{4}\right)^k \sigma_0(4^k x)$$

for all  $k \in \mathbb{N}$ . By construction,  $\sigma_k$  is continuous with period  $2 \cdot 4^{-k}$ . Define

$$f(x) = \sum_{k=0}^{\infty} \sigma_k(x).$$

Since  $\sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k$  converges, by the Weierstrass M-test,  $f$  is continuous. We now show that  $f$  is nowhere differentiable. Fix  $x \in \mathbb{R}$ . Set  $\delta_k = 1/(2 \cdot 4^{-k})$ , then

$$\frac{f(x \pm \delta_k) - f(x)}{\delta_k} = \sum_{j=0}^{\infty} \frac{\sigma_j(x \pm \delta_k) - \sigma_j(x)}{\delta_k}.$$

Note that for  $j > k$ ,  $\sigma_j$  has period  $2 \cdot 4^{-j}$ , hence  $\sigma_j(x \pm \delta_k) = \sigma_j(x)$ , so

$$\frac{f(x \pm \delta_k) - f(x)}{\delta_k} = \frac{\sigma_k(x \pm \delta_k) - \sigma_k(x)}{\delta_k} + \sum_{n=0}^{k-1} \frac{\sigma_n(x \pm \delta_k) - \sigma_n(x)}{\delta_k}.$$

Note that  $\sigma_k$  is monotone on either  $[x, x + \delta_k]$  or  $[x - \delta_k, x]$ . Therefore, either

$$\left| \frac{\sigma_k(x + \delta_k) - \sigma_k(x)}{\delta_k} \right| = 3^n \quad \text{or} \quad \left| \frac{\sigma_k(x - \delta_k) - \sigma_k(x)}{\delta_k} \right| = 3^n.$$

WLOG assume the former, then

$$\left| \frac{f(x + \delta_k) - f(x)}{\delta_k} \right| \geq 3^k - \sum_{n=0}^{k-1} \left| \frac{\sigma_n(x + \delta_k) - \sigma_n(x)}{\delta_k} \right| \geq 3^k - \sum_{n=0}^{k-1} 3^n = \frac{3^k + 1}{2},$$

---

which approaches infinity as  $k \rightarrow \infty$ . Therefore,  $f$  does not have derivative at  $x$ . Since  $x$  is arbitrary,  $f$  is nowhere differentiable.  $\square$

Recall that a subset  $D \subseteq M$  of a metric space is dense if for any nonempty open set  $U \subseteq M$ ,  $U \cap D \neq \emptyset$ . In general, the intersection of two dense sets need not be dense. However, if  $U, V$  are both open and dense, then  $U \cap V$  is also open and dense.

If  $G$  is a countable intersection of open dense subsets of a metric space  $M$ , then we say that  $G$  is thick (or residual/comeager) in  $M$ . A set is thin (or meager) if its complement is thick. Note that a subset of  $H$  of  $M$  is thin iff it can be expressed as a countable union of nowhere dense closed subsets of  $M$ .

**Theorem 4.6.2** (Baire Category Theorem) If  $M$  is a complete metric space, then any thick subset of  $M$  is dense in  $M$ . Moreover, a complete metric space is not thin, i.e. it cannot be expressed as a countable union of nowhere dense closed subsets.

*Proof.* Let  $G = \bigcap_{n=1}^{\infty} U_n$  be a thick subset of  $M$ , where each  $U_n$  is open and dense in  $M$ . Given  $x_0 \in M$  and  $\epsilon > 0$ , pick points  $x_1, x_2, \dots$  and radii  $r_1, r_2, \dots$  converging to 0 such that for all  $n \geq 1$ ,

$$\begin{aligned} B(x_1, 2r_1) &\subseteq B(x_0, \epsilon) \cap U_1, \\ B(x_2, 2r_2) &\subseteq B(x_1, r_1) \cap U_1 \cap U_2, \\ &\vdots \\ B(x_n, 2r_n) &\subseteq B(x_{n-1}, r_{n-1}) \cap U_1 \cap \cdots \cap U_{n-1}. \end{aligned}$$

Then  $(x_n)$  is a Cauchy sequence in  $M$ , hence converges to some  $x \in M$ . For each  $n \geq 1$ , since  $x \in B(x_n, r_n) \subseteq U_n$ , we have  $x \in G$ . This shows that  $B(x_0, \epsilon) \cap G \neq \emptyset$ , hence  $G$  is dense in  $M$ .

For the second part, suppose that  $M = \bigcup_{n=1}^{\infty} F_n$ , where each  $F_n$  is closed and nowhere dense in  $M$ , then  $\emptyset = M^c = \bigcap_{n=1}^{\infty} F_n^c$  is thick in  $M$ , a contradiction.  $\square$

**Corollary 4.6.3** No subset of a complete metric space  $M$  is both thick and thin.

*Proof.* If  $S \subseteq M$  is both thick and thin, then  $S^c$  is also both thick and thin. Therefore, the intersection  $S \cap S^c = \emptyset$  is thick in  $M$ , a contradiction.  $\square$

If  $G \subseteq M$  is thick, we say that a property holds generically in  $M$  if it holds for all points in  $G$ .

**Theorem 4.6.4** A generic continuous function  $f \in C^0([a, b], \mathbb{R})$  is nowhere differentiable. In fact, it does not have a left or right derivative at any point, and is not monotone on any

---

interval.

*Proof.* For  $n \in \mathbb{N}$ , define

$$\begin{aligned} R_n &= \left\{ f \in C^0 \mid \forall x \in \left[a, b - \frac{1}{n}\right], \exists h > 0 \text{ such that } |f(x+h) - f(x)| > nh \right\}, \\ L_n &= \left\{ f \in C^0 \mid \forall x \in \left[a + \frac{1}{n}, b\right], \exists h > 0 \text{ such that } |f(x-h) - f(x)| > nh \right\}, \\ G_n &= \left\{ f \in C^0 \mid f \text{ is not monotone on any interval of length } \frac{1}{n} \right\}. \end{aligned}$$

We will show that each  $R_n, L_n, G_n$  is open and dense in  $C^0([a, b], \mathbb{R})$ .

Fix  $n \in \mathbb{N}$  and  $\epsilon > 0$ . For any  $f \in C^0$ , by Weierstrass approximation theorem, exists some polynomial  $P$  such that  $\|f - P\| < \epsilon/2$ . Consider a sawtooth function  $\sigma$  which has period less than  $1/n$ ,  $0 \leq \sigma < \epsilon/2$ , and

$$\min_x |\text{slope}_x(\sigma)| > n + \max_n |\text{slope}_x(P)|.$$

Define  $g := \sigma + P$ , then the slope of  $g$  alternate signs with period less than  $1/2n$ . At any  $x \in [a, b - 1/n]$ ,  $g$  has a right difference either greater than  $n$  or less than  $-n$ . Therefore,  $g \in R_n$ . Similarly,  $g \in L_n$  and  $g \in G_n$ . Since  $\|f - g\| < \epsilon$ , this shows that  $R_n, L_n, G_n$  are dense in  $C^0$ .

Let  $f \in R_n$ . By definition, for any  $x \in [a, b - 1/n]$ , exists some  $h > 0$  such that

$$\left| \frac{f(x+h) - f(x)}{h} \right| > n.$$

Since  $f$  is continuous, exists a neighborhood  $\mathcal{U}_x$  of  $x$  in  $[a, b]$  and a constant  $\nu = \nu(x) > 0$  such that for all  $y \in \mathcal{U}_x$ ,

$$\left| \frac{f(y+h) - f(y)}{h} \right| > n + \nu.$$

The collection  $\{\mathcal{U}_x \mid x \in [a, b - 1/n]\}$  is an open cover of the compact set  $[a, b - 1/n]$ , hence exists a finite subcover  $\{\mathcal{U}_{x_1}, \dots, \mathcal{U}_{x_m}\}$ . Let  $\nu_0 = \min\{\nu(x_1), \dots, \nu(x_m)\} > 0$ . For any  $g \in C^0$  with  $\|f - g\| < \nu_0 h/2$ , for all  $y \in [a, b - 1/n]$ ,

$$\begin{aligned} \left| \frac{g(y+h) - g(y)}{h} \right| &\geq \left| \frac{f(y+h) - f(y)}{h} \right| - \left| \frac{g(y+h) - f(y+h)}{h} \right| - \left| \frac{f(y) - g(y)}{h} \right| \\ &> n + \nu_0 - \frac{\nu_0}{2} - \frac{\nu_0}{2} = n. \end{aligned}$$

Therefore,  $g \in R_n$ , hence  $R_n$  is open. Similarly,  $L_n$  is open.

Finally, we will show that  $G_n^c$  is closed, and thus  $G_n$  is open. Let  $(f_k)$  be a sequence in  $G_n^c$  such that  $f_k \rightrightarrows f$ . Each  $f_k$  is monotone on some interval  $I_k$  of length  $1/n$ . There exists a subsequence of these intervals that converges to a limit interval  $I$  of length  $1/n$ . Since  $f_k \rightrightarrows f$ ,

---

$f$  is also monotone on  $I$ . Therefore,  $f \in G_n^c$ , hence  $G_n^c$  is closed.

Therefore,  $\bigcap_{n=1}^{\infty} (R_n \cap L_n \cap G_n)$  is thick in  $C^0([a, b], \mathbb{R})$ . Fix  $f$  in this intersection, then for any  $x \in [a, b]$ , exists sequences  $(h_n^+)$  and  $(h_n^-)$  such that

$$\left| \frac{f(x+h) - f(x)}{h} \right| > n \quad \text{and} \quad \left| \frac{f(x-h) - f(x)}{h} \right| > n.$$

Since  $f$  is bounded on  $[a, b]$ , both sequences  $(h_n^+)$  and  $(h_n^-)$  must converge to 0. Therefore,  $f$  does not have left or right derivative at  $x$ . Since  $x$  is arbitrary,  $f$  does not have left or right derivative at any point. Moreover, since  $f \in G_n$  for all  $n$ ,  $f$  is not monotone on any interval.  $\square$

## 4.7 Function Approximation and Artificial Neural Networks

Consider  $K \subseteq \mathbb{R}^d$ , where  $d$  is a large number. In many applications, we would like to approximate a continuous function  $f \in C^0(K, \mathbb{R})$  efficiently.

Let  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  be continuous. A single-layer neural network with activation function  $\sigma$  is a function

$$h(x) = \sum_{i=1}^n a_i \sigma(\langle w_i, x \rangle + b_i),$$

where  $a_i, b_i \in \mathbb{R}$  and  $w_i \in \mathbb{R}^d$  for all  $1 \leq i \leq n$ . The space of single-layer neural networks with  $n$  neurons is denoted by

$$\mathcal{F}_n^\sigma = \left\{ h(x) = \sum_{i=1}^n a_i \sigma(\langle w_i, x \rangle + b_i), a_i, b_i \in \mathbb{R}, w_i \in \mathbb{R}^d \right\}.$$

The set of all single-layer neural networks is  $\mathcal{F}^\sigma = \bigcup_{n=1}^{\infty} \mathcal{F}_n^\sigma$ .

**Lemma 4.7.1** For any non-polynomial  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ , exists some smooth function  $\varphi$  with compact support such that  $\varphi * \sigma$  is not a polynomial.

*Proof.* The space  $C_c^\infty(\mathbb{R})$  of smooth functions with compact support can be made complete under the norm

$$\rho(\varphi_1, \varphi_2) = \sum_{n=0}^{\infty} 2^{-n} \frac{\|\varphi_1^{(n)} - \varphi_2^{(n)}\|_\infty}{1 + \|\varphi_1^{(n)} - \varphi_2^{(n)}\|_\infty}.$$

Suppose that for all  $\varphi \in C_c^\infty(\mathbb{R})$ ,  $\varphi * \sigma$  is a polynomial. Then for each  $m \in \mathbb{N}$ , define

$$V_m = \{\varphi \in C_c^\infty([a, b]) \mid \deg(\varphi * \sigma) \leq m\}.$$

Then  $C_c^\infty([a, b]) = \bigcup_{m=0}^{\infty} V_m$ . We will show that each  $V_m$  is closed and nowhere dense in

---

$C_c^\infty([a, b])$ , contradicting the Baire Category Theorem.

Suppose  $V_m = C_c^\infty([a, b])$  for some  $m$ . Note that the choice of  $m$  only depends on the length of the interval  $[a, b]$ , hence for any  $\varphi \in C_c^\infty(\mathbb{R})$  with support  $[A, B]$ , we can cover  $[A, B]$  with finitely many intervals  $\{[a_i, b_i]\}_{1 \leq i \leq k}$ , where each interval has length  $b_i - a_i = b - a$ , and partition  $\varphi$  into  $\{\varphi_i\}_{1 \leq i \leq k}$  such that  $\varphi_i$  has support  $[a_i, b_i]$  and  $\varphi = \sum_{i=1}^k \varphi_i$ . Since each  $\varphi_i \in V_m$ ,  $\varphi * \sigma = \sum_{i=1}^k \varphi_i * \sigma$  is also a polynomial of degree at most  $m$ . Differentiating  $m + 1$  times, we have

$$\varphi^{(m+1)} * \sigma = 0.$$

Since  $\varphi$  is arbitrary, this implies that  $\sigma$  is a polynomial of degree at most  $m$ , a contradiction. Therefore, each  $V_m$  is a proper subspace of  $C_c^\infty([a, b])$ .

Now since each  $V_m$  is a proper vector subspace of  $C_c^\infty([a, b])$ , hence nowhere dense. Closedness of  $V_m$  follows from the fact that if  $\varphi_k \in V_m$  and  $\varphi_k \rightrightarrows \varphi$ , then  $\varphi_k * \sigma \rightrightarrows \varphi * \sigma$ , hence  $\varphi * \sigma$  is also a polynomial of degree at most  $m$ . Therefore,  $V_m$  is closed.  $\square$

**Theorem 4.7.2** (Universal Approximation Theorem) Let  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$  be continuous and  $K \subseteq \mathbb{R}^d$  be compact, then  $\mathcal{F}^\sigma$  is dense in  $C^0(K, \mathbb{R})$  if and only if  $\sigma$  is not a polynomial.

*Proof.* If  $\sigma$  is a polynomial of degree  $m$ , then any  $h \in \mathcal{F}_n^\sigma$  is a polynomial of degree at most  $m$ , hence  $\mathcal{F}^\sigma$  cannot approximate non-polynomial functions.

Assume that  $\sigma$  is not a polynomial. We first assume that  $\sigma$  is smooth. Fix  $w \in \mathbb{R}^d$  and  $b \in \mathbb{R}$ . Note that  $\sigma(\langle w, x \rangle + b) \in C^0(K, \mathbb{R})$ , and thus

$$\frac{\sigma(\langle w + he_i, x \rangle + b) - \sigma(\langle w, x \rangle + b)}{h} \in \mathcal{F}^\sigma$$

for any  $h \neq 0$ . Taking  $h \rightarrow 0$ , we have

$$\frac{\partial}{\partial w_i} \sigma(\langle w, x \rangle + b) = x_i \sigma'(\langle w, x \rangle + b) \in \overline{\mathcal{F}^\sigma},$$

Similarly, we see that

$$x_1^{\alpha_1} \cdots x_d^{\alpha_d} \sigma^{(r)}(\langle w, x \rangle + b) \in \overline{\mathcal{F}^\sigma},$$

where  $\alpha_1 + \cdots + \alpha_d = r$ . This holds for all  $w \in \mathbb{R}^d$  and  $b \in \mathbb{R}$ . Since  $\sigma$  is not a polynomial, for all  $r \geq 0$ ,  $\exists b \in \mathbb{R}$  such that  $\sigma^{(r)}(b) \neq 0$ . Hence  $\overline{\mathcal{F}^\sigma}$  contains all polynomials on  $K$ . By the Stone-Weierstrass theorem,  $\overline{\mathcal{F}^\sigma} = C^0(K, \mathbb{R})$ .

Suppose now that  $\sigma$  is only continuous. Define

$$C_c^\infty(\mathbb{R}) = \{\text{smooth function with compact support}\}.$$

---

There exists a function  $\varphi \in C_c^\infty(\mathbb{R})$  such that

$$(\varphi * \sigma)(t) = \int_{-\infty}^{\infty} \varphi(t-s)\sigma(s)ds$$

is not a polynomial and smooth. Let  $\epsilon > 0$  and  $f \in C^0(K, \mathbb{R})$ . Let  $\tilde{\sigma} = \varphi * \sigma$ , then we can approximate  $f$  by some  $\tilde{h} \in \mathcal{F}^{\tilde{\sigma}}$  such that  $\|f - h\| < \epsilon/2$ .

On the other hand, we can approximate  $\tilde{\sigma}$  by a Riemann sum

$$\begin{aligned} \tilde{\sigma}(\langle w_i, x \rangle + b_i) &= \int_{-\infty}^{\infty} \varphi(s)\sigma(\langle w_i, x \rangle + b_i - s)ds \\ &\approx \sum_{j=0}^k (s_{j+1} - s_j) \varphi(s_j) \sigma(\langle w_i, x \rangle + b_i - s_j), \end{aligned}$$

which is in  $\mathcal{F}^\sigma$ . Therefore, we can approximate  $\tilde{h}$  by some  $h \in \mathcal{F}^\sigma$  such that  $\|\tilde{h} - h\| < \epsilon/2$ . Hence,  $\|f - h\| < \epsilon$ . This shows that  $\mathcal{F}^\sigma$  is dense in  $C^0(K, \mathbb{R})$ .  $\square$

## CHAPTER 5

# Fourier Analysis

The one-dimensional wave equation is given by

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad u(0, t) = u(\pi, t) = 0,$$

where  $u(x, t)$  is the displacement of a string at position  $x$  and time  $t$ , and  $c$  is the wave speed.

To solve this equation, Fourier introduced the idea of separation of variables. Assume that  $u(x, t) = \varphi(x)\psi(t)$ , then

$$\varphi(x)\psi''(t) = c^2\varphi''(x)\psi(t) \rightsquigarrow \frac{\psi''(t)}{c^2\psi(t)} = \frac{\varphi''(x)}{\varphi(x)} = -\lambda,$$

where  $\lambda$  is some constant independent of  $x$  and  $t$ . Therefore, we arrive at a system of ODEs:

$$\begin{cases} \psi''(t) - c^2\lambda\psi(t) = 0, \\ \varphi''(x) - \lambda\varphi(x) = 0. \end{cases}$$

If  $\lambda \geq 0$ , the solutions to  $\varphi$  is

$$\psi(t) = A' \cosh(c\sqrt{\lambda}t) + B' \sinh(c\sqrt{\lambda}t), \quad \varphi(x) = A \cosh(\sqrt{\lambda}x) + B \sinh(\sqrt{\lambda}x),$$

and if  $\lambda < 0$ , then

$$\psi(t) = A' \cos(c\sqrt{-\lambda}t) + B' \sin(c\sqrt{-\lambda}t), \quad \varphi(x) = A \cos(\sqrt{-\lambda}x) + B \sin(\sqrt{-\lambda}x).$$

If we write  $\lambda = -m^2$ , we find that

$$u_m(x, t) = (A_m \cos(mt) + B_m \sin(mt)) \sin(mx)$$

is a solution to the wave equation. Since the wave equation is linear, any linear combination

---

of these solutions is also a solution. Therefore, we have the more general solution

$$u(x, t) = \sum_{m=1}^{\infty} (A_m \cos(mt) + B_m \sin(mt)) \sin(mx).$$

Suppose further that we're subject to the initial condition  $u(x, 0) = f(x)$  for some continuous function  $f$  with  $f(0) = f(\pi) = 0$ , then

$$f(x) = \sum_{m=1}^{\infty} A_m \sin(mx).$$

This prompts the question whether any (continuous) function on  $[0, \pi]$  can be expressed as a sine series with suitable coefficients  $A_m$ . Note that

$$\int_0^\pi \sin(mx) \sin(nx) dx = \frac{\pi}{2} \delta_{mn},$$

hence

$$\int_0^\pi f(x) \sin(nx) dx = A_n \cdot \frac{\pi}{2}.$$

We extend the domain to  $[-\pi, \pi]$  and consider complex values, then  $f$  can be written as

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx},$$

where the coefficients are given by

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

The coefficient  $a_n$  is called the  $n$ -th Fourier coefficient of  $f$ , denoted by  $\hat{f}(n)$ . The  $n$ -th partial Fourier sum of  $f$  is defined as

$$S_n(f)(x) = \sum_{k=-n}^n \hat{f}(k) e^{ikx}.$$

**Proposition 5.0.1** There exists a continuous function  $f : [-\pi, \pi] \rightarrow \mathbb{R}$  with  $f(-\pi) = f(\pi)$  such that the Fourier series of  $f$  diverges at some point.

## 5.1 Uniqueness of Fourier Series

We want to know that if  $\hat{f}(n) = \hat{g}(n)$  for all  $n \in \mathbb{Z}$ , then does it imply that  $f = g$  almost everywhere? This is equivalent to asking if  $\hat{f}(n) = 0$  for all  $n \in \mathbb{Z}$  implies that  $f = 0$  almost everywhere.

---

**Theorem 5.1.1** Suppose that  $f$  is  $2\pi$ -periodic and is Riemann integrable on  $[-\pi, \pi]$  such that  $\widehat{f}(n) = 0$  for all  $n \in \mathbb{Z}$ , then  $f(x) = 0$  whenever  $f$  is continuous at  $x$ .

*Proof.* We first assume that  $f$  is real-valued, and WLOG assume that  $f$  is continuous at 0 but  $f(0) \neq 0$ . Since  $f$  is continuous at 0, exists some  $\delta \in (0, \pi/2]$  such that for all  $x \in (-\delta, \delta)$ ,  $|f(x)| > |f(0)|/2$ . Consider

$$p(y) = \epsilon + \cos y$$

where  $\epsilon > 0$  is chosen such that  $|p(y)| < 1 - \epsilon/2$  whenever  $\delta \leq |y| \leq \pi$ . We then choose  $0 < \eta < \delta$  such that  $p(y) \geq 1 + \epsilon/2$  whenever  $|y| < \eta$ . Take  $p_k(x) = (p(x))^k$ , and let  $M > 0$  be such that  $|f(x)| \leq M$  for all  $x \in [-\pi, \pi]$ . Each  $p_k$  is a trigonometric polynomial, thus

$$\int_{-\pi}^{\pi} f(x)p_k(x)dx = 0$$

for all integers  $k$ . Note that

$$\begin{aligned} \left| \int_{\delta \leq |x| \leq \pi} f(x)p_k(x)dx \right| &\leq \int_{\delta \leq |x| \leq \pi} |f(x)||p_k(x)|dx \leq 2\pi M(1 - \epsilon/2)^k, \\ \int_{\eta \leq |x| < \delta} f(x)p_k(x)dx &\geq 0, \\ \int_{|x| < \eta} f(x)p_k(x)dx &\geq \frac{f(0)}{2} \left(1 + \frac{\epsilon}{2}\right)^k \cdot 2\eta \rightarrow \infty \end{aligned}$$

as  $k \rightarrow \infty$ . Therefore,

$$\int_{-\pi}^{\pi} f(x)p_k(x)dx \rightarrow \infty$$

as  $k \rightarrow \infty$ , a contradiction.

In general, we can write  $f = u + iv$ , where  $u, v$  are real-valued functions. Note that

$$u = \frac{f + \overline{f}}{2}, \quad v = \frac{f - \overline{f}}{2i},$$

and  $\widehat{f}(n) = \overline{\widehat{f}(-n)}$ , thus

$$\widehat{u}(n) = \frac{\widehat{f}(n) + \widehat{\overline{f}}(n)}{2} = \frac{\widehat{f}(n) + \overline{\widehat{f}(-n)}}{2} = 0, \quad \widehat{v}(n) = \frac{\widehat{f}(n) - \widehat{\overline{f}}(n)}{2i} = \frac{\widehat{f}(n) - \overline{\widehat{f}(-n)}}{2i} = 0,$$

thus  $u(x) = v(x) = 0$  whenever  $u, v$  are continuous at  $x$ . Therefore,  $f(x) = 0$  whenever  $f$  is continuous at  $x$ .  $\square$

**Corollary 5.1.2** Let  $f$  be a  $2\pi$ -periodic continuous function, then

- (a) If  $\widehat{f}(n) = 0$  for all  $n \in \mathbb{Z}$ , then  $f \equiv 0$ .
- (b) If  $\sum_{n=-\infty}^{\infty} |\widehat{f}(n)| < \infty$ , then the partial sum  $S_N(f)(x) = \sum_{n=-N}^N \widehat{f}(n)e^{inx}$  converges to

---

$f$  uniformly on  $[-\pi, \pi]$ .

*Proof.* (a) This follows directly from the previous theorem since  $f$  is continuous everywhere.

(b) By the Weierstrass M-test, the series  $S_N(f)$  converges absolutely and uniformly to

$$g(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx}.$$

By uniform convergence,  $g$  is continuous and  $2\pi$ -periodic. We can verify that  $\hat{g}(n) = \hat{f}(n)$  for all  $n \in \mathbb{Z}$ . Therefore, by applying part (a) to  $f - g$ , we obtain that  $f \equiv g$ , thus  $S_N(f)$  converges uniformly to  $f$ .  $\square$

**Proposition 5.1.1** If  $f$  is a  $2\pi$ -periodic twice-differentiable function, then

$$\hat{f}(n) = O(n^{-2}).$$

*Proof.* By integration by parts, we have

$$\hat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)e^{-inx} dx = \frac{1}{2\pi in} \int_{-\pi}^{\pi} f'(x)e^{-inx} dx = \frac{1}{(2\pi in)^2} \int_{-\pi}^{\pi} f''(x)e^{-inx} dx.$$

Since  $f''$  is continuous on  $[-\pi, \pi]$ , it is bounded, hence

$$|\hat{f}(n)| \leq \frac{1}{(2\pi n)^2} \int_{-\pi}^{\pi} |f''(x)| dx = O(n^{-2}).$$

$\square$

## 5.2 Convolution

Given a  $2\pi$ -periodic function  $f$ , the partial sum

$$S_N(f) = \sum_{-N}^N f(n)e^{inx}$$

can be written as a convolution. Define the Dirichlet kernel  $D_n : [-\pi, \pi] \rightarrow \mathbb{R}$  by

$$D_n(x) = \sum_{k=-n}^n e^{ikx} = \frac{\sin((n + \frac{1}{2})x)}{\sin(x/2)}.$$

---

Then

$$\begin{aligned}
S_N(x) &= \sum_{n=-N}^N \widehat{f}(n) e^{inx} \\
&= \sum_{n=-N}^N \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \right) e^{inx} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left( \sum_{n=-N}^N e^{in(x-t)} \right) dt \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t) dt.
\end{aligned}$$

**Proposition 5.2.1** Let  $f, g, h$  be  $2\pi$ -periodic integrable functions. Then

- $f * (g + h) = f * g + f * h$ ,
- For all  $c \in \mathbb{C}$ ,  $(cf) * g = c(f * g) = f * (cg)$ ,
- $f * g = g * f$ ,
- $(f * g) * h = f * (g * h)$ .

**Lemma 5.2.1** Suppose that  $f$  is Riemann integrable on  $[-\pi, \pi]$  and  $|f| \leq M$  for some  $M > 0$ . Then exists a sequence  $(f_k)$  of  $2\pi$ -periodic continuous functions such that for all  $k$ ,  $|f_k| \leq M$  and

$$\lim_{k \rightarrow \infty} \int_{-\pi}^{\pi} |f(t) - f_k(t)| dt = 0.$$

*Proof.* Although  $f$  can be complex-valued, we can approximate its real and imaginary parts separately. Therefore, we may assume that  $f$  is real-valued.

Let  $\epsilon > 0$ . Choose a partition  $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$  of  $[-\pi, \pi]$  such that the upper and lower Riemann sums satisfy

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \frac{\epsilon}{2}.$$

Set  $\tilde{f}(x) = \sup_{t \in [x_{i-1}, x_i]} f(t)$  for  $x \in [x_{i-1}, x_i]$ . Then  $|\tilde{f}| \leq M$  and

$$\int_{-\pi}^{\pi} |\tilde{f}(t) - f(t)| dt = U(f, \mathcal{P}) - \int_{-\pi}^{\pi} f(t) dt < \frac{\epsilon}{2}.$$

Next, we approximate  $\tilde{f}$  by a  $2\pi$ -periodic continuous function. Pick  $\delta = \epsilon/8nM$ , and define  $\tilde{\tilde{f}}(x) = \tilde{f}(x)$  for  $x$  not in the  $\delta$ -neighborhood of any  $x_i$ , and define  $\tilde{\tilde{f}}$  to be linear on each  $\delta$ -neighborhood of  $x_i$ . Specially, on the endpoints  $x_0 = -\pi$  and  $x_n = \pi$ , we set  $\tilde{\tilde{f}}(x_0) = \tilde{\tilde{f}}(x_n) = 0$ , and define  $\tilde{\tilde{f}}$  to be linear on the  $\delta$ -neighborhood of  $x_0$  and  $x_n$ . Then  $\tilde{\tilde{f}}$  is continuous and  $2\pi$ -

---

periodic. Moreover,  $|\tilde{f}| \leq M$ . Note that

$$\begin{aligned} \int_{-\pi}^{\pi} |\tilde{f}(t) - \tilde{f}(t)| dt &\leq \sum_{i=1}^n \int_{x_i-\delta}^{x_i+\delta} |\tilde{f}(t) - \tilde{f}(t)| dt \\ &\leq \sum_{i=1}^n \int_{x_i-\delta}^{x_i+\delta} 2M dt = 4Mn\delta = \frac{\epsilon}{2}. \end{aligned}$$

Therefore,

$$\int_{-\pi}^{\pi} |\tilde{f}(t) - f(t)| dt < \epsilon.$$

By choosing  $\epsilon = 1/k$  for  $k \in \mathbb{N}$ , we obtain a sequence of continuous functions  $(f_k) = (\tilde{f})$  that converges to  $f$  in  $L^1$  norm.  $\square$

**Proposition 5.2.2** If  $f, g$  are  $2\pi$ -periodic integrable functions, then  $f * g$  is continuous.

*Proof.* First assume that both  $f$  and  $g$  are continuous, then

$$\begin{aligned} |(f * g)(x) - (f * g)(y)| &= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)(g(x-t) - g(y-t)) dt \right| \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)| |g(x-t) - g(y-t)| dt \\ &\leq \|f\|_{\infty} \cdot \|g(x - \cdot) - g(y - \cdot)\|_{\infty}. \end{aligned}$$

Since  $g$  is continuous on  $[-\pi, \pi]$ , it is uniformly continuous, thus we can make the last term arbitrarily small by choosing  $x$  sufficiently close to  $y$ . Therefore,  $f * g$  is continuous.

For general integrable  $f$  and  $g$ , we can approximate them by continuous functions. By Lemma 5.2.1, exists sequences of continuous functions  $(f_k)$  and  $(g_k)$  such that  $f_k \rightarrow f$  and  $g_k \rightarrow g$  in  $L^1$  norm. Note that

$$\begin{aligned} |(f * g)(x) - (f_k * g_k)(x)| &= |f * (g - g_k)(x) + (f - f_k) * g_k(x)| \\ &\leq |f * (g - g_k)(x)| + |(f - f_k) * g_k(x)| \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)| |g(t) - g_k(t)| dt + \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t) - f_k(t)| |g_k(t)| dt \\ &\leq \frac{1}{2\pi} \|f\|_1 \int_{-\pi}^{\pi} |g(t) - g_k(t)| dt + \frac{1}{2\pi} \|g_k\|_1 \int_{-\pi}^{\pi} |f(t) - f_k(t)| dt, \end{aligned}$$

which approaches 0 as  $k \rightarrow \infty$ . Since each  $f_k * g_k$  is continuous, and  $f * g$  is the uniform limit of continuous functions,  $f * g$  is also continuous.  $\square$

**Proposition 5.2.3** If  $f, g$  are  $2\pi$ -periodic integrable functions, then for all  $n \in \mathbb{Z}$ ,

$$\widehat{f * g}(n) = \widehat{f}(n) \cdot \widehat{g}(n).$$

---

*Proof.* By definition,

$$\begin{aligned}\widehat{f * g}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (f * g)(x) e^{-inx} dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) g(x-t) dt \right) e^{-inx} dx \\ &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} f(t) \left( \int_{-\pi}^{\pi} g(x-t) e^{-inx} dx \right) dt.\end{aligned}$$

Making the substitution  $y = x - t$ , we have

$$\begin{aligned}\widehat{f * g}(n) &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} f(t) \left( \int_{-\pi}^{\pi} g(y) e^{-in(y+t)} dy \right) dt \\ &= \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} f(t) e^{-int} dt \cdot \int_{-\pi}^{\pi} g(y) e^{-iny} dy \\ &= \widehat{f}(n) \cdot \widehat{g}(n).\end{aligned}$$

□

**Definition 5.2.2** A sequence of  $2\pi$ -periodic integrable functions  $(K_n)$  is said to be an approximation to the identity if

(i) For each  $n$ ,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(x) dx = 1,$$

(ii) Exists some  $M > 0$  such that for all  $n$ ,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |K_n(x)| dx \leq M,$$

(iii) For all  $\delta \in (0, \pi)$ ,

$$\lim_{n \rightarrow \infty} \int_{\delta \leq |x| \leq \pi} |K_n(x)| dx = 0.$$

**Proposition 5.2.4** Let  $(K_n)$  be an approximation to the identity, and let  $f$  be a  $2\pi$ -periodic integrable function. Then, whenever  $f$  is continuous at  $x$ ,

$$\lim_{n \rightarrow \infty} (f * K_n)(x) = f(x).$$

Moreover, the convergence is uniform.

*Proof.* Let  $\epsilon > 0$ . Since  $f$  is continuous at  $x$ , exists some  $\delta \in (0, \pi)$  such that for all  $|y| < \delta$ ,  $|f(x-y) - f(x)| < \epsilon/(2M)$ , where  $M$  is as in condition (ii) of the definition of approximation

to the identity. Then

$$\begin{aligned}
& |(f * K_n)(x) - f(x)| \\
&= \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(y)(f(x-y) - f(x))dy \right| \\
&\leq \frac{1}{2\pi} \int_{|y|<\delta} |K_n(y)||f(x-y) - f(x)|dy + \frac{1}{2\pi} \int_{\delta \leq |y| \leq \pi} |K_n(y)||f(x-y) - f(x)|dy \\
&\leq \frac{\epsilon}{2M} \cdot \frac{1}{2\pi} \int_{|y|<\delta} |K_n(y)|dy + 2\|f\|_1 \cdot \frac{1}{2\pi} \int_{\delta \leq |y| \leq \pi} |K_n(y)|dy \\
&\leq \frac{\epsilon}{2} + 2\|f\|_1 \cdot \frac{1}{2\pi} \int_{\delta \leq |y| \leq \pi} |K_n(y)|dy.
\end{aligned}$$

By condition (iii) of the definition of approximation to the identity, integrating outside of a  $\delta$ -neighborhood of 0 can be made arbitrarily small for sufficiently large  $n$ . Therefore, exists some  $N$  such that for all  $n \geq N$ ,

$$|(f * K_n)(x) - f(x)| < \epsilon.$$

The convergence is uniform since  $\delta$  does not depend on  $x$ .  $\square$

Note that the Dirichlet kernel ( $D_n$ ) does not form an approximation to the identity since

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(x)|dx \rightarrow \infty$$

as  $n \rightarrow \infty$ , thus in general, we cannot expect pointwise convergence of the Fourier series, even for continuous functions.

### 5.3 Summability and Convergence

Let  $\sum_{k=0}^{\infty} c_k$  be a series of complex numbers. Define the  $n$ -th partial sum  $s_n = \sum_{k=0}^n c_k$ . We say that the series converges to  $s$  if  $\lim_{n \rightarrow \infty} s_n = s$ . Define the  $n$ -th Cesáro mean of the series to be

$$\sigma_n = \frac{1}{n} \sum_{k=0}^{n-1} s_k.$$

**Definition 5.3.1** If  $\lim_{n \rightarrow \infty} \sigma_n = \sigma$ , we say that the series is Cesáro summable to  $\sigma$ .

We can see that if a series converges, then it is also Cesáro summable to the same limit.

---

Define the Fejér kernel  $F_n : [-\pi, \pi] \rightarrow \mathbb{R}$  by

$$F_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(x).$$

**Lemma 5.3.2** For  $x \neq 0$ , we have

$$F_n(x) = \frac{1}{n} \left( \frac{\sin(\frac{nx}{2})}{\sin(\frac{x}{2})} \right)^2.$$

Moreover,  $(F_n)$  is an approximation to the identity.

*Proof.* Note that

$$\begin{aligned} F_n(x) &= \frac{1}{n} \sum_{k=0}^{n-1} \frac{\sin((k + \frac{1}{2})x)}{\sin(x/2)} \\ &= \frac{1}{n \sin(x/2)} \cdot \operatorname{Im} \left( \sum_{k=0}^{n-1} e^{i(k+\frac{1}{2})x} \right) \\ &= \frac{1}{n \sin(x/2)} \cdot \operatorname{Im} \left( e^{ix/2} \cdot \frac{1 - e^{inx}}{1 - e^{ix}} \right) \\ &= \frac{1}{n} \cdot \frac{\sin(nx/2)}{\sin(x/2)} \cdot \frac{\sin(nx/2)}{\sin(x/2)} \\ &= \frac{1}{n} \left( \frac{\sin(\frac{nx}{2})}{\sin(\frac{x}{2})} \right)^2. \end{aligned}$$

We now verify that  $(F_n)$  is an approximation to the identity. For each  $n$ ,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(x) dx = \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} D_k(x) dx = 1.$$

Moreover,

$$F_n(x) \geq 0$$

for all  $x$ , hence

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

Finally, for all  $\delta \in (0, \pi)$ ,

$$\begin{aligned} \int_{\delta \leq |x| \leq \pi} F_n(x) dx &\leq \int_{\delta}^{\pi} \frac{1}{n} \left( \frac{1}{\sin(x/2)} \right)^2 dx \\ &\leq \frac{1}{n} \cdot \frac{\pi - \delta}{\sin^2(\delta/2)} \rightarrow 0 \end{aligned}$$

as  $n \rightarrow \infty$ . Therefore,  $(F_n)$  is an approximation to the identity.  $\square$

---

**Corollary 5.3.3** If  $f$  is a  $2\pi$ -periodic integrable function, then its Fourier series is Cesáro summable to  $f(x)$  whenever  $f$  is continuous at  $x$ . Moreover, if  $f$  is continuous on  $[-\pi, \pi]$ , then the Cesáro means converge to  $f$  uniformly on  $[-\pi, \pi]$ .

*Proof.* Note that the  $n$ -th Cesáro mean of the Fourier series of  $f$  is given by

$$\sigma_n(f)(x) = \frac{1}{n} \sum_{k=0}^{n-1} S_k(f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) F_n(x-t) dt = (f * F_n)(x).$$

By the previous proposition, whenever  $f$  is continuous at  $x$ ,

$$\lim_{n \rightarrow \infty} \sigma_n(f)(x) = f(x).$$

Moreover, if  $f$  is continuous on  $[-\pi, \pi]$ , then the convergence is uniform.  $\square$

**Corollary 5.3.4**  $2\pi$ -periodic continuous functions can be uniformly approximated by trigonometric polynomials.

*Proof.* Let  $f$  be a  $2\pi$ -periodic continuous function and  $\epsilon > 0$ . By the previous corollary, the Cesáro means  $\sigma_n$  converge to  $f$  uniformly. Therefore, exists some  $N$  such that for all  $n \geq N$ ,

$$\|f - \sigma_n\|_\infty < \epsilon.$$

Since each  $\sigma_n$  is a trigonometric polynomial, we are done.  $\square$

A series of complex numbers  $\sum_{k=0}^{\infty} c_k$  is said to be Abel summable to  $s$  if the power series

$$A(r) := \sum_{k=0}^{\infty} c_k r^k$$

converges for all  $r \in [0, 1)$  and

$$\lim_{r \rightarrow 1^-} A(r) = s.$$

We can see that if a series converges, then it is also Abel summable to the same limit.

**Proposition 5.3.1** If a series is Cesáro summable to  $s$ , then it is also Abel summable to  $s$ .

*Proof.* Let  $\sum_{k=0}^{\infty} c_k$  be a series that is Cesáro summable to  $s$ . Define the partial sums  $s_n = \sum_{k=0}^n c_k$  and the Cesáro means  $\sigma_n = \frac{1}{n} \sum_{k=0}^{n-1} s_k$ . By assumption,  $\lim_{n \rightarrow \infty} \sigma_n = s$ .

---

For  $r \in [0, 1)$ , we have

$$\begin{aligned} A(r) &= \sum_{k=0}^{\infty} c_k r^k = (1-r) \sum_{k=0}^{\infty} s_k r^k \\ &= (1-r) \sum_{k=0}^{\infty} \left( \frac{1}{k} \sum_{j=0}^{k-1} s_j + (s_k - \sigma_k) \right) r^k \\ &= (1-r) \sum_{k=0}^{\infty} \sigma_k r^k + (1-r) \sum_{k=0}^{\infty} (s_k - \sigma_k) r^k. \end{aligned}$$

As  $r \rightarrow 1^-$ , the first term approaches  $s$  by the definition of Cesáro summability. The second term approaches 0 since  $s_k - \sigma_k$  is bounded. Therefore,  $\lim_{r \rightarrow 1^-} A(r) = s$ , and the series is Abel summable to  $s$ .  $\square$

**Remark** Consider the series

$$\sum_{n=0}^{\infty} (-1)^n (n+1).$$

The Abel sum of this series is given by

$$A(r) = \sum_{n=0}^{\infty} (-1)^n (n+1) r^n = \frac{1}{(1+r)^2},$$

thus the series is Abel summable to  $1/4$ . However, the Cesáro means of the series diverge. This shows that the converse of the previous proposition is not true in general.

For Fourier series, we define the Abel means of the Fourier series of a  $2\pi$ -periodic integrable function  $f$  to be

$$A_r(f)(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n) r^{|n|} e^{inx}, \quad r \in [0, 1).$$

This is the Abel mean with  $c_0 = \widehat{f}(0)$  and  $c_n = \widehat{f}(n)e^{inx} + \widehat{f}(-n)e^{-inx}$  for  $n \geq 1$ . Since

$$|\widehat{f}(n)e^{inx}| = |\widehat{f}(n)| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)| dt = \frac{1}{2\pi} \|f\|_1,$$

by the Weierstrass M-test, the series converges absolutely and uniformly for all  $r \in [0, 1)$ . Note

---

that

$$\begin{aligned}
A_r(f)(x) &= \sum_{n=-\infty}^{\infty} \widehat{f}(n) r^{|n|} e^{inx} \\
&= \sum_{n=-\infty}^{\infty} \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \right) r^{|n|} e^{inx} \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left( \sum_{n=-\infty}^{\infty} r^{|n|} e^{in(x-t)} \right) dt,
\end{aligned}$$

thus we can define the Poisson kernel  $P_r : [-\pi, \pi] \rightarrow \mathbb{R}$  by

$$P_r(x) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{inx} = \frac{1 - r^2}{1 - 2r \cos x + r^2}.$$

**Lemma 5.3.5** For  $r \in [0, 1)$ ,

$$P_r(x) = \frac{1 - r^2}{1 - 2r \cos x + r^2}.$$

Moreover, the Poisson kernels  $(P_r)_{r \in [0, 1)}$  form an approximation to the identity as  $r \rightarrow 1^-$ .

*Proof.* Write  $\omega = re^{ix}$ . Then

$$\begin{aligned}
P_r(x) &= \sum_{n=-\infty}^{\infty} r^{|n|} e^{inx} \\
&= \sum_{n=0}^{\infty} \omega^n + \sum_{n=1}^{\infty} \bar{\omega}^n \\
&= \frac{1}{1 - \omega} + \frac{1}{1 - \bar{\omega}} - 1 \\
&= \frac{1 - \omega + 1 - \bar{\omega} - (1 - \omega)(1 - \bar{\omega})}{(1 - \omega)(1 - \bar{\omega})} \\
&= \frac{1 - |\omega|^2}{|1 - \omega|^2} = \frac{1 - r^2}{1 - 2r \cos x + r^2}.
\end{aligned}$$

We now know that  $P_r$  is nonnegative. Integrating termwise, we obtain

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

If  $1/2 \leq r < 1$  and  $\delta \leq |x| \leq \pi$ , then exists  $c_\delta > 0$  such that

$$1 - 2r \cos x + r^2 = (1 - r)^2 + 2r(1 - \cos x) \geq c_\delta,$$

thus

$$P_r(x) \leq \frac{1 - r^2}{c_\delta} \leq \frac{2(1 - r)}{c_\delta}.$$

---

Therefore,

$$\int_{\delta \leq |x| \leq \pi} P_r(x) dx \leq \frac{4\pi(1-r)}{c_\delta} \rightarrow 0$$

as  $r \rightarrow 1^-$ . This shows that  $(P_r)$  form an approximation to the identity.  $\square$

We say that a sequence  $(f_n)$  of  $2\pi$ -periodic integrable function converges to  $f$  in  $L^2$  norm if

$$\int_{-\pi}^{\pi} |f_n(x) - f(x)|^2 dx \rightarrow 0$$

as  $n \rightarrow \infty$ . It concerns the global behavior of the functions instead of pointwise behavior, and is closely related to orthogonality.

**Theorem 5.3.6** Suppose that  $f$  is a  $2\pi$ -periodic integrable function. Then  $S_N(f)$  converges to  $f$  in  $L^2$  norm.

## 5.4 $L^2$ Convergence

The set of all double-sided sequences of complex numbers with finite 2-norm

$$\ell^2(\mathbb{Z}) = \left\{ (a_n)_{n \in \mathbb{Z}} \mid \sum_{n=-\infty}^{\infty} |a_n|^2 < \infty \right\}$$

is a vector space over  $\mathbb{C}$ . For two sequences  $A = (a_n)$  and  $B = (b_n)$  in  $\ell^2(\mathbb{Z})$ , we define their inner product to be

$$\langle A, B \rangle = \sum_{n=-\infty}^{\infty} a_n \overline{b_n}.$$

This inner product induces a norm  $\|A\|_2 = \sqrt{\langle A, A \rangle}$ . With this norm,  $\ell^2(\mathbb{Z})$  is a complete metric space. Such a complete inner product space is called a Hilbert space.

Let  $\mathcal{R}$  denote the space of all complex-valued Riemann integrable functions on  $[-\pi, \pi]$ . For  $f, g \in \mathcal{R}$ , we define their inner product to be

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \overline{g(x)} dx.$$

Note that this is technically not an inner product since  $\langle f, f \rangle = 0$  only implies that  $f$  is continuous and zero almost everywhere. However, if we identify two functions whose difference is continuous and zero almost everywhere, then this becomes a valid inner product. This inner product induces a norm  $\|f\|_2 = \sqrt{\langle f, f \rangle}$ .

Let  $e_n(x) = e^{inx}$  for  $n \in \mathbb{Z}$ . Then  $\{e_n\}_{n \in \mathbb{Z}}$  is an orthonormal set in  $\mathcal{R}$  since

$$\langle e_m, e_n \rangle = \delta_{mn}.$$

---

If  $f \in \mathcal{R}$ , then its Fourier coefficients are given by

$$\widehat{f}(n) = \langle f, e_n \rangle.$$

Since we can write

$$S_N(f)(x) = \sum_{n=-N}^N \widehat{f}(n) e^{inx} = \sum_{n=-N}^N \langle f, e_n \rangle e_n(x),$$

we see that  $f - S_N(f)$  is orthogonal to each  $e_n$  for  $|n| \leq N$ . Therefore, by the Pythagorean theorem,

$$\|f\|_2^2 = \|f - S_N(f)\|_2^2 + \|S_N(f)\|_2^2.$$

**Lemma 5.4.1** If  $f \in \mathcal{R}$ , then

$$\|f - S_N(f)\| \leq \|f - \sum_{n=-N}^N a_n e_n\|$$

for all choices of complex numbers  $a_n$ .

*Proof.* Note that

$$(f - S_N(f)) \perp \left( \sum_{n=-N}^N (\langle f, e_n \rangle) - e_n \right),$$

hence by the Pythagorean theorem,

$$\|f - \sum_{n=-N}^N a_n e_n\|^2 = \|f - S_N(f)\|^2 + \|S_N(f) - \sum_{n=-N}^N a_n e_n\|^2 \geq \|f - S_N(f)\|^2.$$

□

**Proposition 5.4.1** If  $f$  is a  $2\pi$ -periodic integrable function, then

$$\lim_{N \rightarrow \infty} \|f - S_N(f)\|_2 = 0.$$

*Proof.* First assume that  $f$  is continuous, then we can approximate  $f$  uniformly by trigonometric polynomials. Given  $\epsilon > 0$ , exists some trigonometric polynomial  $P(x)$  such that  $\|f - P\|_\infty < \epsilon$ , then we also have  $\|f - P\|_2 < \epsilon$ . Let  $M$  be the degree of  $P$ , then for all  $N \geq M$ ,

$$\|f - S_N(f)\|_2 \leq \|f - P\|_2 < \epsilon.$$

For general integrable  $f$ , we can approximate it by a  $2\pi$ -periodic continuous function such that

$$\sup_{x \in [-\pi, \pi]} |g(x)| \leq \sup_{x \in [-\pi, \pi]} |f(x)| =: K$$

and

$$\int_{-\pi}^{\pi} |f(x) - g(x)| dx < \frac{\pi\epsilon^2}{4K}.$$

Then

$$\begin{aligned} \|f - g\|_2^2 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - g(x)|^2 dx \\ &\leq \frac{2K}{2\pi} \int_{-\pi}^{\pi} |f(x) - g(x)| dx < \frac{\epsilon^2}{4}. \end{aligned}$$

Therefore,  $\|f - g\|_2 < \epsilon/2$ . Since  $g$  is continuous, exists some trigonometric polynomial  $P(x)$  such that  $\|g - P\|_2 < \epsilon/2$ , then

$$\|f - S_N(f)\|_2 \leq \|f - P\|_2 \leq \|f - g\|_2 + \|g - P\|_2 < \epsilon.$$

□

**Corollary 5.4.2** (Parseval's Identity) If  $f$  is a  $2\pi$ -periodic integrable function, then

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2.$$

*Proof.* Note that

$$(f - S_N(f)) \perp S_N(f),$$

hence by the Pythagorean theorem,

$$\|f\|_2^2 = \|f - S_N(f)\|_2^2 + \|S_N(f)\|_2^2.$$

As  $N \rightarrow \infty$ ,  $\|f - S_N(f)\|_2^2 \rightarrow 0$ , and

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2.$$

□

Parseval's identity implies that the mapping  $f \mapsto (\widehat{f}(n))_{n \in \mathbb{Z}}$  is an isometry from  $\mathcal{R}$  to  $\ell^2(\mathbb{Z})$ . Since  $\mathcal{R}$  is not complete, this mapping cannot be surjective. Therefore, exists some sequence  $(a_n) \in \ell^2(\mathbb{Z})$  that is not the sequence of Fourier coefficients of any Riemann integrable function.

**Theorem 5.4.3** (Riemann-Lebesgue lemma) If  $f$  is a  $2\pi$ -periodic integrable function, then

$$\lim_{|n| \rightarrow \infty} \widehat{f}(n) = 0.$$

---

*Proof.* By Parseval's identity,

$$\sum_{n=-\infty}^{\infty} |\widehat{f}(n)|^2 = \|f\|_2^2 < \infty,$$

thus  $|\widehat{f}(n)|^2 \rightarrow 0$  as  $|n| \rightarrow \infty$ .  $\square$

More generally, we have

$$\lim_{|n| \rightarrow \infty} \int_{-\pi}^{\pi} f(x) \sin(nx) dx = \lim_{|n| \rightarrow \infty} \int_{-\pi}^{\pi} f(x) \cos(nx) dx = 0.$$

**Proposition 5.4.2** If  $f, g \in \mathcal{R}$ , then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \overline{g(x)} dx = \sum_{n=-\infty}^{\infty} \widehat{f}(n) \cdot \overline{\widehat{g}(n)}.$$

*Proof.* The polarization identity states that

$$\langle f, g \rangle = \frac{1}{4} (\|f + g\|_2^2 - \|f - g\|_2^2 + i\|f + ig\|_2^2 - i\|f - ig\|_2^2).$$

Therefore by Parseval's identity,

$$\begin{aligned} \langle f, g \rangle &= \frac{1}{4} \left( \sum_{n=-\infty}^{\infty} |\widehat{f}(n) + \widehat{g}(n)|^2 - \sum_{n=-\infty}^{\infty} |\widehat{f}(n) - \widehat{g}(n)|^2 \right) \\ &\quad + \frac{i}{4} \left( \sum_{n=-\infty}^{\infty} |\widehat{f}(n) + i\widehat{g}(n)|^2 - \sum_{n=-\infty}^{\infty} |\widehat{f}(n) - i\widehat{g}(n)|^2 \right) \\ &= \sum_{n=-\infty}^{\infty} \widehat{f}(n) \cdot \overline{\widehat{g}(n)}. \end{aligned}$$

$\square$

**Theorem 5.4.4** Let  $f \in \mathcal{R}$ . Suppose that  $f$  is Lipschitz at  $x_0$ , i.e., exists some  $L > 0$  and  $\delta > 0$  such that

$$|f(x_0 + h) - f(x_0)| \leq L|h|$$

for all  $|h| < \delta$ . Then

$$\lim_{N \rightarrow \infty} S_N(f)(x_0) = f(x_0).$$

*Proof.* Define

$$F(t) = \begin{cases} \frac{f(x_0-t)-f(x_0)}{t}, & 0 < |t| < \pi, \\ 0, & t = 0. \end{cases}$$

---

Since  $f$  is Lipschitz at  $x_0$ ,  $F$  is bounded by  $L$  on  $[-\pi, \pi]$ . Also, the discontinuity set of  $F$  is contained in that of  $f$  and possibly at  $x_0$ , thus  $F$  is Riemann integrable on  $[-\pi, \pi]$ . Recall that  $S_N(f)(x_0) = (f * D_N)(x_0)$  and  $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t)dt = 1$ , we have

$$\begin{aligned} S_N(f)(x_0) - f(x_0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t)(f(x_0 - t) - f(x_0))dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} tF(t)D_N(t)dt. \end{aligned}$$

Write

$$tD_N(t) = \frac{\sin((N + \frac{1}{2})t)}{\sin(t/2)} \cdot t = \frac{t}{\sin(t/2)} \cdot (\sin(Nt)\cos(t/2) + \cos(Nt)\sin(t/2)).$$

Note that  $\frac{t/2}{\sin(t/2)}$  is continuous on  $[-\pi, \pi]$ , thus the functions

$$F(t) \cdot \frac{t}{\sin(t/2)} \cdot \cos(t/2) \quad \text{and} \quad F(t) \cdot t$$

are Riemann integrable on  $[-\pi, \pi]$ . By the Riemann-Lebesgue lemma,

$$\lim_{N \rightarrow \infty} \int_{-\pi}^{\pi} tF(t)D_N(t)dt = 0,$$

hence proved.  $\square$

**Corollary 5.4.5** (Localization principle) Let  $f, g \in \mathcal{R}$ . Suppose that there exists some  $\delta > 0$  such that  $f(x) = g(x)$  for all  $x \in (x_0 - \delta, x_0 + \delta)$ . Then

$$\lim_{N \rightarrow \infty} (S_N(f)(x_0) - S_N(g)(x_0)) = 0.$$

*Proof.* Follows from applying the previous theorem to the function  $f - g$ .  $\square$

## 5.5 Continuous function with divergent Fourier series

For each  $N \in \mathbb{N}$ , define

$$f_N(x) = \sum_{n=-N}^N \frac{e^{inx}}{n}, \quad g_N(x) = \sum_{n=-N}^{-1} \frac{e^{inx}}{n}.$$

Note that there exists  $c > 0$  such that  $|g_N(0)| \geq c \log N$  for all  $N$ .

**Lemma 5.5.1** Suppose that the Abel means  $A_r = \sum_{n=1}^{\infty} c_n r^n$  of a series  $\sum_{n=1}^{\infty} c_n$  are bounded as  $r \rightarrow 1^-$ . If  $c_n = O(n^{-1})$ , then the partial sums  $S_N = \sum_{n=1}^N c_n$  are also bounded.

---

*Proof.* Fix  $N$  and take  $r = 1 - 1/N$ . Choose  $M > 0$  such that  $n|c_n| \leq M$  for all  $n$ . Note that

$$1 - r^n = (1 - r)(1 + r + r^2 + \cdots + r^{n-1}) \leq n(1 - r),$$

thus

$$\begin{aligned} |S_N - A_r| &\leq \sum_{n=1}^N |c_n|(1 - r^n) + \sum_{n=N+1}^{\infty} |c_n|r^n \\ &\leq M \sum_{n=1}^N (1 - r) + \frac{M}{N} \sum_{n=N+1}^{\infty} r^n \\ &\leq M + \frac{M}{N} \cdot \frac{r^{N+1}}{1 - r} \leq M + M = 2M. \end{aligned}$$

Since  $A_r$  is bounded as  $r \rightarrow 1^-$ ,  $S_N$  is also bounded.  $\square$

From the lemma, we can deduce that the sums  $f_n(x)$  are uniformly bounded in both  $n$  and  $x$  since the Abel means

$$\sum_{n=1}^{\infty} \frac{r^n e^{inx}}{n} = -\log(1 - re^{ix})$$

are bounded as  $r \rightarrow 1^-$ . Therefore, exists some  $M > 0$  such that for all  $N$  and  $x$ ,  $|f_N(x)| \leq M$ .

Define

$$P_N(x) = e^{i2Nx} f_N(x), \quad Q_N(x) = e^{i2Nx} g_N(x).$$

Note that  $P_N$  has degree  $3N$  and  $Q_N$  has degree  $2N - 1$ . Then the partial sums of  $P_N$  are

$$S_M(P_N) = \begin{cases} P_N, & M \geq 3N, \\ Q_N, & M = 2N, \\ 0, & M < N. \end{cases}$$

Let  $(\alpha_k)$  be a convergent series of positive real numbers such that  $\sum_k \alpha_k < \infty$ . Also, let  $(N_k)$  be a increasing sequence of integers such that

- (i)  $N_{k+1} > 3N_k$ ,
- (ii)  $\alpha_k \log N_k \rightarrow \infty$  as  $k \rightarrow \infty$ .

Set

$$h(x) = \sum_{k=1}^{\infty} \alpha_k P_{N_k}(x).$$

Since  $|P_{N_k}(x)| = |f_{N_k}(x)| \leq M$  for all  $k$  and  $x$ , the series converges uniformly, thus  $h$  is continuous.

---

**Claim**

$$|S_{2N_m}(h)(0)| \geq c'\alpha_m \log N_m$$

for all  $N$  large enough, where  $c' > 0$  is some constant independent of  $m$ .

*Proof.* The terms corresponding to  $N_k$  for  $k < m$  contribute  $O(1)$  only since  $|P_{N_k}(0)|$  are uniformly bounded and  $\sum \alpha_k < \infty$ . The term that corresponds to  $N_m$  has absolute value at least  $\alpha_m |g_{N_m}(0)| \geq c\alpha_m \log N_m$ . The terms corresponding to  $N_k$  for  $k > m$  contribute 0 since their degrees are greater than  $2N_m$ . Therefore,

$$|S_{2N_m}(h)(0)| \geq c\alpha_m \log N_m - O(1).$$

Since  $\alpha_m \log N_m \rightarrow \infty$  as  $m \rightarrow \infty$ , the claim follows.  $\square$

The claim implies that the partial sums of the Fourier series of  $h$  at 0 are unbounded, thus they do not converge at 0.

## 5.6 Applications

Let  $\Gamma$  be a simple closed curve in  $\mathbb{R}^2$  with length  $\ell$ . Under fixed  $\ell$ , we want to find the curve that encloses the largest area.

**Definition 5.6.1** A parametrized curve is a continuous map  $\gamma : [a, b] \rightarrow \mathbb{R}^2$ . The image of  $\gamma$  is called a curve and is denoted by  $\Gamma$ . The curve  $\Gamma$  is simple and closed if  $\gamma(a) = \gamma(b)$  and  $\gamma(t_1) \neq \gamma(t_2)$  if  $t_1 \neq t_2$  unless  $t_1 = a$  and  $t_2 = b$ .

For a simple closed curve, we may extend  $\gamma$  to be periodic with period  $b - a$ . Furthermore, we will assume that  $\gamma$  is  $C^1$ , and  $\gamma'(t) \neq 0$  for all  $t \in [a, b]$ .

Note that any  $C^1$  bijective map  $t : [c, d] \rightarrow [a, b]$  gives rise to a new parametrization of the curve  $\Gamma$  by  $\gamma \circ t$ . Simplicity and closedness of  $\Gamma$  is independent of the parametrization. We say that  $\gamma$  and  $\gamma \circ t$  are equivalent if  $t'(s) > 0$  for all  $s$ .

**Definition 5.6.2** If  $\Gamma$  is a curve parametrized by  $\gamma(t) = (x(t), y(t))$ , we define the length of  $\Gamma$  by

$$\ell = \ell(\Gamma) = \int_a^b |\gamma'(t)| dt = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt.$$

---

Suppose that  $\tilde{\gamma} = \gamma \circ t$  and  $\gamma$  are equivalent, then

$$\begin{aligned}\int_a^b |\gamma'(t)| dt &= \int_c^d |\tilde{\gamma}'(s)| t'(s) ds \\ &= \int_c^d |\tilde{\gamma}'(s)| ds.\end{aligned}$$

**Definition 5.6.3** We say that  $\gamma$  is an arclength reparametrization if  $|\gamma'(t)| = 1$  for all  $t$ .

**Definition 5.6.4** The area enclosed by a simple closed curve  $\Gamma$  parametrized by  $\gamma(t) = (x(t), y(t))$  is given by

$$A(\Gamma) = \frac{1}{2} \left| \int_{\Gamma} (x dy - y dx) \right| = \frac{1}{2} \left| \int_a^b x(t) y'(t) - y(t) x'(t) dt \right|.$$

**Theorem 5.6.5** (Isoperimetric inequality) Suppose that  $\Gamma$  is a simple closed  $C^1$ -curve in  $\mathbb{R}^2$  with length  $\ell$ . Let  $A$  be the area enclosed by  $\Gamma$ . Then

$$A \leq \frac{\ell^2}{4\pi},$$

with equality if and only if  $\Gamma$  is a circle.

*Proof.* WLOG rescale the curve such that  $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$  is an arclength parametrization of  $\Gamma$ , which has length  $2\pi$ . Write  $\gamma(t) = (x(t), y(t))$ . Since  $\gamma$  is an arclength parametrization, we have

$$x'(t)^2 + y'(t)^2 = 1$$

for all  $t$ . Expanding  $x$  and  $y$  in Fourier series,

$$x(t) = \sum_{n=-\infty}^{\infty} a_n e^{int}, \quad y(t) = \sum_{n=-\infty}^{\infty} b_n e^{int}.$$

By differentiating termwise (justified by integration by parts), we have

$$x'(t) = \sum_{n=-\infty}^{\infty} i n a_n e^{int}, \quad y'(t) = \sum_{n=-\infty}^{\infty} i n b_n e^{int}.$$

Applying Parseval's identity to  $x'$  and  $y'$ , we obtain

$$\begin{aligned}1 &= \frac{1}{2\pi} \int_0^{2\pi} (x'(t)^2 + y'(t)^2) dt \\ &= \sum_{n=-\infty}^{\infty} n^2 (|a_n|^2 + |b_n|^2).\end{aligned}$$

---

The area enclosed by  $\Gamma$  is given by

$$\begin{aligned}
A &= \frac{1}{2} \left| \int_0^{2\pi} x(t)y'(t) - y(t)x'(t) dt \right| \\
&= \pi \left| \sum_{n=-\infty}^{\infty} n(a_n \bar{b}_n - b_n \bar{a}_n) \right| \\
&\leq \pi \sum_{n=1}^{\infty} 2n|a_n||b_n| \\
&= \pi \sum_{n=1}^{\infty} n(|a_n|^2 + |b_n|^2) \\
&\leq \pi \sum_{n=-\infty}^{\infty} n^2(|a_n|^2 + |b_n|^2) = \pi.
\end{aligned}$$

Therefore, the area enclosed by  $\Gamma$  is at most  $\pi$ . Equality holds only if  $a_n = b_n = 0$  for all  $|n| \geq 2$ . This means that  $x$  and  $y$  are of the form

$$x(t) = a_0 + a_1 e^{it} + a_{-1} e^{-it}, \quad y(t) = b_0 + b_1 e^{it} + b_{-1} e^{-it}.$$

Since  $\gamma$  is real-valued, we have  $a_{-1} = \bar{a}_1$  and  $b_{-1} = \bar{b}_1$ . Thus  $\gamma$  parametrizes a circle. Clearly if  $\Gamma$  is a circle, then  $A = \pi$ , hence proved.  $\square$

The second application looks at a number theory problem. Given  $\alpha \in \mathbb{R}$ , consider the fractional part of  $n\alpha$  for integers  $n \geq 1$ . We want to know how these fractional parts are distributed in the interval  $[0, 1)$ .

To formally define this, let  $[x]$  denote the largest integer not exceeding  $x$ . The fractional part of  $x$  is defined to be  $\{x\} = x - [x]$ .

Given  $\alpha \in \mathbb{R}$  consider the sequence

$$a_n = \{n\alpha\}, \quad n = 1, 2, 3, \dots$$

If  $\alpha$  is rational, then there are only finitely many distinct values in the sequence  $(a_n)$ . However, if  $\alpha$  is irrational, then the sequence  $(a_n)$  has infinitely many distinct values.

**Proposition 5.6.1** If  $\alpha$  is irrational, then the sequence  $(\{n\alpha\})_{n=1}^{\infty}$  is dense in  $[0, 1)$ .

*Proof.* Divide  $[0, 1)$  into  $n$  subintervals of equal length. Consider  $n + 1$  distinct elements

$$0, \{\alpha\}, \{2\alpha\}, \dots, \{n\alpha\}.$$

By the pigeonhole principle, exists some distinct  $0 \leq i, j \leq n$  such that  $\{i\alpha\}$  and  $\{j\alpha\}$  lie in the same subinterval. WLOG assume that  $\{i\alpha\} > \{j\alpha\}$  and  $i > j$  (the case  $i < j$  is similar).

Then

$$\{(i-j)\alpha\} = \{i\alpha\} - \{j\alpha\} < \frac{1}{n}.$$

Since  $\alpha$  is irrational,  $i-j \geq 1$ . Thus, we have found some positive integer  $m = i-j$  such that  $\{m\alpha\} < 1/n$ . For any open interval  $(a, b) \subset [0, 1)$ , choose  $n$  large enough such that  $1/n < b-a$ . Then exists some integer  $m$  such that  $\{m\alpha\} < 1/n < b-a$ . Let  $k$  be the smallest integer such that  $k\{m\alpha\} \geq a$ , then

$$(k-1)\{m\alpha\} < a \leq k\{m\alpha\} < a + \{m\alpha\} < b.$$

Therefore,  $\{(km)\alpha\} = k\{m\alpha\} \in (a, b)$ , hence proved.  $\square$

**Lemma 5.6.6** If  $f$  is a 1-periodic continuous function and if  $\alpha$  is irrational, then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\{n\alpha\}) = \int_0^1 f(x) dx.$$

*Proof.* Consider the special case that  $f(x) = e^{2\pi i kx}$  for some integer  $k$ . When  $k=0$ , the result is trivial. When  $k \neq 0$ , since  $\alpha$  is irrational,  $e^{2\pi i k\alpha} \neq 1$ . Therefore,

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N f(\{n\alpha\}) &= \frac{1}{N} \sum_{n=1}^N e^{2\pi i k n \alpha} \\ &= \frac{1}{N} \cdot \frac{e^{2\pi i k \alpha} (1 - e^{2\pi i k N \alpha})}{1 - e^{2\pi i k \alpha}} \rightarrow 0 \end{aligned}$$

as  $N \rightarrow \infty$ . By linearity, the lemma holds for all trigonometric polynomials with period 1. For general continuous  $f$ , we can approximate it uniformly by trigonometric polynomials. Given  $\epsilon > 0$ , exists some trigonometric polynomial  $P(x)$  such that  $\|f - P\|_\infty < \epsilon/3$ . For all  $N$  large enough,

$$\left| \frac{1}{N} \sum_{n=1}^N P(\{n\alpha\}) - \int_0^1 P(x) dx \right| < \epsilon/3.$$

Therefore, for all  $N$  large enough,

$$\begin{aligned} \left| \frac{1}{N} \sum_{n=1}^N f(\{n\alpha\}) - \int_0^1 f(x) dx \right| &\leq \left| \frac{1}{N} \sum_{n=1}^N (f(\{n\alpha\}) - P(\{n\alpha\})) \right| \\ &\quad + \left| \frac{1}{N} \sum_{n=1}^N P(\{n\alpha\}) - \int_0^1 P(x) dx \right| \\ &\quad + \left| \int_0^1 (P(x) - f(x)) dx \right| \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

Hence proved.  $\square$

---

**Definition 5.6.7** A sequence of numbers  $(\xi_n)$  in  $[0, 1]$  is equidistributed if for all  $(a, b) \subseteq [0, 1]$ ,

$$\lim_{N \rightarrow \infty} \frac{\#\{1 \leq n \leq N \mid \xi_n \in (a, b)\}}{N} = b - a.$$

**Theorem 5.6.8** (Weyl's equidistribution theorem) If  $\alpha$  is irrational, then the sequence  $(\{n\alpha\})_{n=1}^{\infty}$  is equidistributed in  $[0, 1]$ .

*Proof.* Consider  $\chi_{(a,b)}$  and extend it to  $\mathbb{R}$  to be periodic with period 1. It suffices to show that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \chi_{(a,b)}(\{n\alpha\}) = \int_0^1 \chi_{(a,b)}(x) dx = b - a.$$

Approximate  $\chi_{(a,b)}$  by continuous functions

$$f_{\epsilon}^{-}(x) = \begin{cases} 0, & x \leq a - \epsilon, \\ \frac{x-a}{\epsilon} + 1, & a - \epsilon < x < a, \\ 1, & a \leq x \leq b, \\ \frac{b-x}{\epsilon} + 1, & b < x < b + \epsilon, \\ 0, & x \geq b + \epsilon, \end{cases}, \quad f_{\epsilon}^{+}(x) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{\epsilon}, & a < x < a + \epsilon, \\ 1, & a + \epsilon \leq x \leq b - \epsilon, \\ \frac{b-x}{\epsilon}, & b - \epsilon < x < b, \\ 0, & x \geq b, \end{cases}$$

then

$$f_{\epsilon}^{-}(x) \leq \chi_{(a,b)}(x) \leq f_{\epsilon}^{+}(x).$$

We have

$$\int_0^1 f_{\epsilon}^{-}(x) dx \leq \int_0^1 \chi_{(a,b)}(x) dx \leq \int_0^1 f_{\epsilon}^{+}(x) dx,$$

and since both  $f_{\epsilon}^{-}$  and  $f_{\epsilon}^{+}$  differ from  $\chi_{(a,b)}$  on an interval of length  $2\epsilon$  only, letting  $\epsilon \rightarrow 0$  gives the desired equality.  $\square$

**Corollary 5.6.9** For any Riemann integrable function  $f$  with period 1 and any irrational  $\alpha$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(\{n\alpha\}) = \int_0^1 f(x) dx.$$

*Proof.* Let  $\epsilon > 0$ . Choose a partition  $P = \{x_0, x_1, \dots, x_n\}$  of  $[0, 1]$  such that  $U(f, P) - L(f, P) < \epsilon$ . Define

$$f_U = \sup_{x_{j-1} \leq x \leq x_j} f(y) \quad \text{if } x \in [x_{j-1}, x_j),$$

$$f_L = \inf_{x_{j-1} \leq x \leq x_j} f(y) \quad \text{if } x \in [x_{j-1}, x_j).$$

---

Then  $f_U$  and  $f_L$  are step functions such that

$$f_L(x) \leq f(x) \leq f_U(x)$$

for all  $x \in [0, 1]$ , and

$$L(f, P) = \int_0^1 f_L(x)dx \leq \int_0^1 f(x)dx \leq \int_0^1 f_U(x)dx = U(f, P).$$

Since  $f_L$  and  $f_U$  are linear combinations of characteristic functions of intervals, by Weyl's equidistribution theorem,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f_L(\{n\alpha\}) = \int_0^1 f_L(x)dx, \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f_U(\{n\alpha\}) = \int_0^1 f_U(x)dx.$$

Therefore, for all  $N$  large enough,

$$\begin{aligned} \left| \frac{1}{N} \sum_{n=1}^N f(\{n\alpha\}) - \int_0^1 f(x)dx \right| &\leq \left| \frac{1}{N} \sum_{n=1}^N (f(\{n\alpha\}) - f_L(\{n\alpha\})) \right| \\ &+ \left| \frac{1}{N} \sum_{n=1}^N f_L(\{n\alpha\}) - \int_0^1 f_L(x)dx \right| \\ &+ \left| \int_0^1 (f_L(x) - f(x))dx \right|. \end{aligned}$$

Letting  $N \rightarrow \infty$  gives the desired result.  $\square$

Consider the function  $\rho : [0, 2\pi) \rightarrow [0, 2\pi)$  defined by

$$\rho(\theta) = \theta + 2\pi\alpha \mod 2\pi.$$

The iterations of  $\rho$  are given by  $\rho^n(\theta) = \theta + 2\pi n\alpha \mod 2\pi$ . To each Riemann integrable  $f$  on the circle, we can associate the corresponding effects of  $\rho$ , and obtain a sequence of functions

$$f, f(\rho), f(\rho^2), \dots$$

By Weyl's equidistribution theorem, if  $\alpha \notin \mathbb{Q}$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\rho^n(\theta)) = \frac{1}{2\pi} \int_0^{2\pi} f(x)dx,$$

for all  $\theta \in [0, 2\pi)$ . Intuitively, the summation on the left can be thought of as a "time average", while the integral on the right is the "space average".

**Theorem 5.6.10** (Weyl's equidistribution criterion) A sequence  $(\xi_n)$  in  $[0, 1)$  is equidis-

---

tributed if and only if for all integers  $k \neq 0$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i k \xi_n} = 0.$$

*Proof.* "  $\implies$  ":

"  $\Leftarrow$  ": By the criterion, we know that it also holds for all trigonometric polynomials. Since we can approximate continuous functions uniformly by trigonometric polynomials, and characteristic functions of intervals can be approximated by continuous functions, the result follows.  $\square$

We can also construct nowhere differentiable continuous functions using Fourier series.

We first introduce other ways to sum the Fourier series:

- Ordinary partial sum:  $S_N(f) = f * D_N$ ,
- Cesàro sum:  $\sigma_N(f) = f * F_N$ , where  $F_N$  is the Fejér kernel,
- Delayed means:  $\Delta_N(f) = 2\sigma_{2N}(f) - \sigma_N(f) = f * (2F_{2N} - F_N)$ .

$2F_{2N} - F_N$  is called the de la Vallée Poussin kernel. Note that

$$\widehat{D}_N(n) = \begin{cases} 1, & |n| \leq N, \\ 0, & |n| > N, \end{cases}$$

and

$$\widehat{F}_N(n) = \frac{D_0(n) + D_1(n) + \cdots + D_{N-1}(n)}{N} = \begin{cases} 1 - \frac{|n|}{N}, & |n| \leq N, \\ 0, & |n| > N. \end{cases}$$

If  $f(x) \sim \sum_n a_n e^{inx}$ , then

- $S_N$  arises by multiplying  $a_n e^{inx}$  by 1 if  $|n| \leq N$  and 0 otherwise,
- $\sigma_N$  arises by multiplying  $a_n e^{inx}$  by  $1 - \frac{|n|}{N}$  if  $|n| \leq N$  and 0 otherwise,
- $\Delta_N$  arises by multiplying  $a_n e^{inx}$  by 1 if  $|n| \leq N$ , by  $2 - \frac{|n|}{N}$  if  $N < |n| \leq 2N$ , and 0 otherwise.

**Theorem 5.6.11** Let  $0 < \alpha < 1$ . The function

$$f(x) = f_\alpha(x) = \sum_{n=0}^{\infty} 2^{-n\alpha} e^{i2^n x}$$

is continuous but nowhere differentiable.

---

*Proof.* Note that

$$S_N(f) = \Delta_{N'}(f),$$

where  $N'$  is the largest power of 2 not exceeding  $N$ . Suppose that  $f$  is differentiable at some  $x_0$ .

**Lemma 5.6.12** Let  $g$  be differentiable at  $x_0$ . Then

$$(\sigma_N(g))'(x_0) = O(\log N).$$

*Proof.* We have

$$(\sigma_N(g))'(x_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F'_N(x_0 - t)g(t)dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} F'_N(t)g(x_0 - t)dt.$$

Note that  $\int_{-\pi}^{\pi} F'_N(t)dt = 0$ . Thus,

$$(\sigma_N(g))'(x_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F'_N(t)(g(x_0 - t) - g(x_0))dt.$$

Since  $g$  is differentiable at  $x_0$ , exists some  $c > 0$  such that

$$|(\sigma_N(g))'(x_0)| \leq c \int_{-\pi}^{\pi} |F'_N(t)| |t| dt.$$

It remains to estimate  $|F'_N(t)|$ .

Exists some  $A > 0$  such that

$$|F'_N(t)| \leq A \cdot N^2$$

since  $F_N$  is a trigonometric polynomial of degree  $N$  with coefficients  $\leq 1$ , hence  $F'_N$  is a trigonometric polynomial of degree  $N$  with coefficients  $\leq N$ .

On the other hand, write

$$F_N(t) = \frac{1}{N} \cdot \frac{\sin^2(Nt/2)}{\sin^2(t/2)}.$$

Differentiating gives

$$F'_N(t) = \frac{\sin(Nt/2) \cdot \cos(Nt/2)}{\sin^2(t/2)} - \frac{1}{N} \frac{\cos(t/2) \cdot \sin^2(Nt/2)}{\sin^3(t/2)}$$

Thus, exists some  $B > 0$  such that

$$|F'_N(t)| \leq \frac{B}{t^2}$$

for all  $t \in [-\pi, \pi] \setminus \{0\}$ .

From the estimates above, we conclude that

$$\begin{aligned}
|(\sigma_N(g))'(x_0)| &\leq c \int_{|t| \geq 1/N} |F'_N(t)| |t| dt + c \int_{|t| < 1/N} |F'_N(t)| |t| dt \\
&\leq c \int_{|t| \geq 1/N} \frac{B}{t^2} |t| dt + c \int_{|t| < 1/N} AN^2 |t| dt \\
&= cB \int_{1/N}^\pi \frac{1}{t} dt + cAN^2 \cdot \frac{1}{N^2} \\
&= cB \log(\pi N) + cA = O(\log N).
\end{aligned}$$

□

**Lemma 5.6.13** If  $2N = 2^n$ , then

$$\Delta_{2N}(f)(x) - \Delta_N(f)(x) = 2^{-n\alpha} e^{i2^nx}.$$

*Proof.* Note that

$$\begin{aligned}
\widehat{\Delta_{2N}}(k) - \widehat{\Delta_N}(k) &= \begin{cases} 0, & |k| \leq N, \\ 1 - \left(2 - \frac{|k|}{N}\right) = \frac{|k|}{N} - 1, & N < |k| \leq 2N, \\ 0, & |k| > 2N. \end{cases} \\
&= \begin{cases} 0, & |k| < 2N, \\ 1, & |k| = 2N, \\ 0, & |k| > 2N. \end{cases}
\end{aligned}$$

Therefore,

$$\Delta_{2N}(f)(x) - \Delta_N(f)(x) = 2^{-n\alpha} e^{i2^nx}.$$

□

By lemma 1, we have

$$(\Delta_{2N}(f))'(x_0) - (\Delta_N(f))'(x_0) = O(\log N).$$

On the other hand, by lemma 2,

$$|(\Delta_{2N}(f))'(x_0) - (\Delta_N(f))'(x_0)| = 2^{n(1-\alpha)} \geq cN^{1-\alpha}$$

for some constant  $c > 0$ . Since  $1 - \alpha > 0$ , this is a contradiction. Therefore,  $f$  is nowhere differentiable. □

**Remark** Note that  $f$  is complex-valued, so even if  $f$  is nowhere differentiable, it does not imply that  $\operatorname{Re}(f)$  and  $\operatorname{Im}(f)$  are also nowhere differentiable. However, we can modify the proof

---

to show that both  $\operatorname{Re}(f)$  and  $\operatorname{Im}(f)$  are nowhere differentiable as well.

Assume that  $g = \operatorname{Re}(f)$  is differentiable at some  $x_0$ . Then

$$(\sigma_N(g))'(x_0) = O(\log N).$$

On the other hand,

$$\Delta_{2N}(g)(x) - \Delta_N(g)(x) = 2^{-n\alpha} \cos(2^n x).$$

Thus,

$$\left| 2^{n(1-\alpha)} \sin(2^n x_0) \right| = |(\Delta_{2N}(g))'(x_0) - (\Delta_N(g))'(x_0)| = O(\log N).$$

Since  $1 - \alpha > 0$ , this implies that  $\sin(2^n x_0) \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore,  $2^n x_0$  is arbitrarily close to an integer multiple of  $\pi$  for all large  $n$ . This is possible only if  $x_0 = 0$ , which contradicts the assumption that  $g$  is differentiable at  $x_0$ . Hence,  $g$  is nowhere differentiable. A similar argument shows that  $\operatorname{Im}(f)$  is also nowhere differentiable.

## 5.7 Heat equation on the circle

Consider the heat equation on the circle  $S^1$ :

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, & x \in [0, 1), t > 0, \\ u(x, 0) = f(x), & x \in [0, 1], \end{cases}$$

By separating variables, we look for solutions of the form  $u(x, t) = X(x)T(t)$ . Substituting into the heat equation gives

$$X(x)T'(t) = X''(x)T(t) \implies \frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = -\lambda,$$

for some constant  $\lambda$ . Since  $X$  is periodic with period 1, we have  $\lambda = (2\pi n)^2$  for some integer  $n$ . Therefore, the general solution is given by

$$u(x, t) = \sum_{n=-\infty}^{\infty} a_n e^{-4\pi^2 n^2 t} e^{2\pi i n x},$$

where

$$a_n = \widehat{f}(n) = \int_0^1 f(x) e^{-2\pi i n x} dx.$$

If  $f$  is Riemann integrable, then  $a_n \rightarrow 0$  and  $e^{-4\pi^2 n^2 t}$  decays to 0 very rapidly as  $|n| \rightarrow \infty$ . Therefore, the series converges pointwise for all  $t > 0$ .

### Proposition 5.7.1

$$\lim_{t \rightarrow 0^+} \int_0^1 |u(x, t) - f(x)|^2 dx = 0.$$

---

*Proof.* Define the heat kernel

$$H_t(x) = \sum_{n=-\infty}^{\infty} e^{-4\pi^2 n^2 t} e^{2\pi i n x}.$$

Then the solution can be written as

$$u(x, t) = (f * H_t)(x) = \int_0^1 f(y) H_t(x - y) dy.$$

By Parseval's identity,

$$\begin{aligned} \int_0^1 |u(x, t) - f(x)|^2 dx &= \sum_{n=-\infty}^{\infty} |\widehat{u}(n, t) - \widehat{f}(n)|^2 \\ &= \sum_{n=-\infty}^{\infty} |e^{-4\pi^2 n^2 t} - 1|^2 |\widehat{f}(n)|^2. \end{aligned}$$

Let  $\epsilon > 0$ . Note that since  $|e^{-4\pi^2 n^2 t} - 1| = 1 - e^{-4\pi^2 n^2 t} \leq 1$  for all  $n$  and  $\sum_n |\widehat{f}(n)|^2 < \infty$ , exists some  $N$  such that

$$\sum_{|n|>N} |\widehat{f}(n)|^2 |e^{-4\pi^2 n^2 t} - 1|^2 < \epsilon/2.$$

For this fixed  $N$ , exists some  $\delta > 0$  such that if  $0 < t < \delta$ , then

$$\sum_{|n|\leq N} |\widehat{f}(n)|^2 |e^{-4\pi^2 n^2 t} - 1|^2 < \epsilon/2.$$

Therefore, for all  $0 < t < \delta$ ,

$$\int_0^1 |u(x, t) - f(x)|^2 dx < \epsilon,$$

hence proved.  $\square$

In  $d$  dimensions, the heat equation becomes

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_d^2} =: \Delta u.$$

$\Delta$  is known as the Laplace operator or Laplacian. We will focus on  $d = 2$ .

We can imagine that after an extended period of time, the heat exchange between different parts of the object will reach an equilibrium, i.e., the temperature distribution will no longer change with time. In this case,  $\frac{\partial u}{\partial t} = 0$ , so  $\Delta u = 0$ . A twice differentiable function  $u$  satisfying  $\Delta u = 0$  is called a harmonic function.

---

Consider

$$\begin{aligned}\mathbb{D} &= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}, \\ \partial\mathbb{D} &= \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.\end{aligned}$$

The Dirichlet problem for the Laplace equation on the unit disc is to find a function  $u : \bar{\mathbb{D}} \rightarrow \mathbb{R}$  such that  $\Delta u = 0$  in  $\mathbb{D}$  and  $u = f$  on  $\partial\mathbb{D}$ , where  $f : \partial\mathbb{D} \rightarrow \mathbb{R}$  is a given continuous function.

In order to perform separation of variables, we use polar coordinates

$$\mathbb{D} = \{(r, \theta) \mid 0 \leq r < 1, 0 \leq \theta < 2\pi\}.$$

It's not hard to deduce that

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

Assume that  $u(r, \theta) = A(r)B(\theta)$ . Then

$$\frac{A''(r)}{A(r)} + \frac{1}{r} \frac{A'(r)}{A(r)} + \frac{1}{r^2} \frac{B''(\theta)}{B(\theta)} = 0 \implies \frac{r^2 A''(r)}{A(r)} + r \frac{A'(r)}{A(r)} = -\frac{B''(\theta)}{B(\theta)} = \lambda,$$

hence

$$\begin{aligned}r^2 A''(r) + r A'(r) - \lambda A(r) &= 0, \\ B''(\theta) + \lambda B(\theta) &= 0.\end{aligned}$$

Since  $B$  is periodic with period  $2\pi$ , we have  $\lambda = n^2$  for some integer  $n$ . Write  $B(\theta) = ce^{in\theta} + de^{-in\theta}$  for some constants  $c, d \in \mathbb{C}$ . If  $n \neq 0$ , the simple solutions of  $A$  are  $r^n$  and  $r^{-n}$ ; if  $n = 0$ , the simple solutions of  $A$  are 1 and  $\log r$ . Since  $u$  must be bounded at  $r = 0$ , we discard  $r^{-n}$  and  $\log r$ . Therefore, the general solution is given by

$$u(r, \theta) = \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta},$$

where

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-in\theta} d\theta.$$

Define the Poisson kernel

$$P_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}.$$

**Proposition 5.7.2** Let  $f$  be a  $2\pi$ -periodic Riemann integrable function. Define

$$u(r, \theta) = (f * P_r)(\theta),$$

on  $\mathbb{D}$ , then

- 
- (a)  $u$  has two continuous derivatives in  $\mathbb{D}$  and satisfies  $\Delta u = 0$ ,
- (b) If  $\theta$  is any continuity point of  $f$ , then

$$\lim_{r \rightarrow 1^-} u(r, \theta) = f(\theta),$$

If  $f$  is continuous, then the convergence is uniform.

- (c) If  $f$  is continuous, then  $u$  is the unique solution to  $\Delta u = 0$  in  $\mathbb{D}$  that satisfies conditions (a) and (b).

*Proof.* (a) Since  $f$  is Riemann integrable, its Fourier coefficients  $a_n \rightarrow 0$  as  $|n| \rightarrow \infty$ . Therefore, for all  $0 \leq r < 1$ , the series

$$u(r, \theta) = \sum_{n=-\infty}^{\infty} a_n r^{|n|} e^{in\theta}$$

converges absolutely and uniformly in  $\theta$ . Differentiating term-by-term shows that  $u$  has two continuous derivatives in  $\mathbb{D}$  and satisfies  $\Delta u = 0$ .

- (b) Note that

$$u(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} f(\phi) P_r(\theta - \phi) d\phi.$$

Since  $P_r(\theta)$  is an approximate identity as  $r \rightarrow 1^-$ , the result follows.

- (c) Suppose that  $v$  is another solution satisfying (a) and (b). Write

$$v(r, \theta) = \sum_{n=-\infty}^{\infty} a_n(r) e^{in\theta},$$

where

$$a_n(r) = \frac{1}{2\pi} \int_0^{2\pi} v(r, \theta) e^{-in\theta} d\theta.$$

Since  $\Delta v = 0$ , we have

$$\frac{1}{2\pi} \int_0^{2\pi} \Delta v(r, \theta) e^{-in\theta} d\theta = 0.$$

Plugging in the expression of  $\Delta v$  gives

$$a_n''(r) + \frac{1}{r} a_n'(r) - \frac{n^2}{r^2} a_n(r) = 0.$$

Solving this ODE gives  $a_n(r) = c_n r^{|n|} + d_n r^{-|n|}$  for some constants  $c_n, d_n \in \mathbb{C}$  if  $n \neq 0$ . Since  $v$  is bounded,  $a_n(r)$  must also be bounded, so  $d_n = 0$  if  $n \neq 0$ . Similarly, we have  $a_0(r) = c_0$  for some constant  $c_0 \in \mathbb{C}$ . Therefore,

$$v(r, \theta) = \sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta}.$$

---

By condition (b), we have

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-in\theta} d\theta = a_n.$$

By uniqueness of Fourier series for continuous functions,  $v = u$ .  $\square$

## CHAPTER 6

# Fourier Transform

In this chapter, we extend the notion of Fourier series to non-periodic functions defined on the whole real line. The Fourier transform is a powerful tool in both pure and applied mathematics, with applications in signal processing, differential equations, and quantum mechanics, among others.

Consider a 1-periodic function  $f$  defined on  $\mathbb{R}$ . Its Fourier series is given by

$$f(x) \sim \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x},$$

where

$$a_n = \int_0^1 f(x) e^{-2\pi i n x} dx.$$

For general  $f$ , we can extend the frequency variable  $n$  from integers to real numbers, and define the Fourier transform of  $f$  as

$$\widehat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i \xi x} dx, \quad \xi \in \mathbb{R}.$$

In order for the integral to make sense, it needs to be at least improperly integrable.

**Definition 6.0.1** The Schwartz space  $\mathcal{S}(\mathbb{R})$  is the set of all infinitely differentiable functions  $f : \mathbb{R} \rightarrow \mathbb{C}$  such that for all non-negative integers  $m, n$ ,

$$\sup_{x \in \mathbb{R}} |x^m f^{(n)}(x)| < \infty.$$

**Example 6.0.1**  $f(x) = e^{-x^2}$  is in the Schwartz space since for all  $m, n$ ,

$$f^{(n)}(x) = P_n(x) e^{-x^2},$$

where  $P_n$  is a polynomial of degree  $n$ .

---

We can easily see that  $\mathcal{S}(\mathbb{R})$  is a vector space over  $\mathbb{C}$ , and closed under differentiation and multiplication by polynomials. The Fourier transform is well-defined on  $\mathcal{S}(\mathbb{R})$  since for all  $f \in \mathcal{S}(\mathbb{R})$  and  $\xi \in \mathbb{R}$ ,

$$|\widehat{f}(\xi)| \leq \int_{-\infty}^{\infty} |f(x)| dx < \infty.$$

**Proposition 6.0.1** Let  $f \in \mathcal{S}(\mathbb{R})$ ,  $h \in \mathbb{R}$ ,  $\delta > 0$ . Then

- (a) The Fourier transform of  $f(x + h)$  is  $\widehat{f}(\xi)e^{2\pi i \xi h}$ ,
- (b) The Fourier transform of  $f(x)e^{-2\pi i x h}$  is  $\widehat{f}(\xi + h)$ ,
- (c) The Fourier transform of  $f(\delta x)$  is  $\frac{1}{\delta} \widehat{f}\left(\frac{\xi}{\delta}\right)$ ,
- (d) The Fourier transform of  $f'(x)$  is  $2\pi i \xi \widehat{f}(\xi)$ ,
- (e) The Fourier transform of  $-2\pi i x f(x)$  is  $\frac{d}{d\xi} \widehat{f}(\xi)$ .

*Proof.* Let  $g(x) = -2\pi i x f(x)$ . Then

$$\begin{aligned} \frac{\widehat{f}(\xi + h) - \widehat{f}(\xi)}{h} &= \int_{-\infty}^{\infty} f(x) \frac{e^{-2\pi i (\xi+h)x} - e^{-2\pi i \xi x}}{h} dx \\ &= \int_{-\infty}^{\infty} f(x) e^{-2\pi i \xi x} \left( \frac{e^{-2\pi i h x} - 1}{h} \right) dx. \end{aligned}$$

Since  $f(x)$  and  $x f(x)$  decrease rapidly and

$$\lim_{h \rightarrow 0} \frac{e^{-2\pi i h x} - 1}{h} = -2\pi i x,$$

by the dominated convergence theorem, we have

$$\frac{d}{d\xi} \widehat{f}(\xi) = \int_{-\infty}^{\infty} -2\pi i x f(x) e^{-2\pi i \xi x} dx = \widehat{g}(\xi).$$

□

- (f)  $\widehat{f} \in \mathcal{S}(\mathbb{R})$ .

*Proof.* If  $f \in \mathcal{S}(\mathbb{R})$ , then

$$\widehat{f^{(n)}}(\xi) \leq \int_{-\infty}^{\infty} |f^{(n)}(x)| dx < \infty,$$

so  $\widehat{f}$  is bounded. Also, by (d) and (e),

$$\xi^k \left( \frac{d}{d\xi} \right)^l \widehat{f}(\xi) = \frac{1}{(2\pi i)^k} \widehat{\frac{d^k}{dx^k} ((-2\pi i x)^l f(x))},$$

---

which is the Fourier transform of a function in  $\mathcal{S}(\mathbb{R})$ , hence bounded. Therefore,  $\widehat{f} \in \mathcal{S}(\mathbb{R})$ .  $\square$

## 6.1 Inverse Fourier Transform

**Proposition 6.1.1** Let  $f(x) = e^{-\pi x^2}$ . Then  $\widehat{f}(\xi) = e^{-\pi \xi^2}$ .

*Proof.* Note that  $f'(x) = -2\pi x f(x)$ . Define  $F(\xi) = \widehat{f}(\xi)$ , then

$$\begin{aligned} F'(\xi) &= \frac{d}{d\xi} \int_{-\infty}^{\infty} f(x) e^{-2\pi i \xi x} dx \\ &= \int_{-\infty}^{\infty} f(x) (-2\pi i x) e^{-2\pi i \xi x} dx \\ &= i \int_{-\infty}^{\infty} f'(x) e^{-2\pi i \xi x} dx \\ &= i \cdot (2\pi i \xi) F(\xi) = -2\pi \xi F(\xi). \end{aligned}$$

Solving this differential equation, we get

$$F(\xi) = C e^{-\pi \xi^2}.$$

Since  $F(0) = \int_{-\infty}^{\infty} e^{-\pi x^2} dx = 1$ , we have  $C = 1$ . Therefore,

$$\widehat{f}(\xi) = e^{-\pi \xi^2}.$$

$\square$

**Corollary 6.1.1** For  $\delta > 0$  and  $x \in \mathbb{R}$ , define  $K_\delta(x) = \delta^{-1/2} e^{-\pi x^2/\delta}$ . Then  $\widehat{K}_\delta(\xi) = e^{-\pi \delta \xi^2}$ .

*Proof.* By Proposition 1(c) and Proposition 2,

$$\widehat{K}_\delta(\xi) = \frac{1}{\delta^{1/2}} \widehat{f}\left(\frac{\xi}{\delta}\right) = \frac{1}{\delta^{1/2}} e^{-\pi(\xi/\delta)^2} = e^{-\pi \delta \xi^2}.$$

$\square$

**Definition 6.1.2** If a family of functions  $\{f_\delta\}_{\delta>0}$  satisfy the following conditions, we call it an approximate to the identity on  $\mathbb{R}$ :

(a)

$$\int_{-\infty}^{\infty} f_\delta(x) dx = 1,$$

---

(b) exists some  $M > 0$  such that for all  $\delta > 0$ ,

$$\int_{-\infty}^{\infty} |f_{\delta}(x)| dx \leq M,$$

(c) For all  $\eta > 0$ ,

$$\lim_{\delta \rightarrow 0^+} \int_{|x| \geq \eta} f_{\delta}(x) dx = 0.$$

It's easy to see that  $\{K_{\delta}\}_{\delta>0}$  is an approximate to the identity on  $\mathbb{R}$ .

**Definition 6.1.3** If  $f, g \in \mathcal{S}(\mathbb{R})$ , define their convolution as

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

**Proposition 6.1.2** If  $f \in \mathcal{S}(\mathbb{R})$  and  $\{g_{\delta}\}_{\delta>0}$  is an approximate to the identity on  $\mathbb{R}$ , then

$$\lim_{\delta \rightarrow 0^+} (f * g_{\delta})(x) = f(x),$$

and the convergence is uniform on  $\mathbb{R}$ .

*Proof.* Note that

$$(f * g_{\delta})(x) - f(x) = \int_{-\infty}^{\infty} (f(x-y) - f(x))g_{\delta}(y) dy.$$

Given  $\epsilon > 0$ , since  $f$  is uniformly continuous on  $\mathbb{R}$ , exists some  $\eta > 0$  such that if  $|y| < \eta$ , then  $|f(x-y) - f(x)| < \epsilon/(2M)$ , where  $M$  is as in Definition 3(b). Therefore,

$$\begin{aligned} |(f * g_{\delta})(x) - f(x)| &\leq \int_{|y|<\eta} |f(x-y) - f(x)||g_{\delta}(y)| dy + \int_{|y|\geq\eta} |f(x-y) - f(x)||g_{\delta}(y)| dy \\ &< \frac{\epsilon}{2M} \int_{-\infty}^{\infty} |g_{\delta}(y)| dy + 2 \sup_{x \in \mathbb{R}} |f(x)| \int_{|y|\geq\eta} |g_{\delta}(y)| dy. \end{aligned}$$

By Definition 3(b), the first term is less than  $\epsilon/2$ . By Definition 3(c), exists some  $\delta_0 > 0$  such that if  $0 < \delta < \delta_0$ , then the second term is also less than  $\epsilon/2$ . Therefore, for all  $0 < \delta < \delta_0$ ,

$$|(f * g_{\delta})(x) - f(x)| < \epsilon,$$

hence proved. □

**Lemma 6.1.4** If  $f, g \in \mathcal{S}(\mathbb{R})$ , then

$$\int_{-\infty}^{\infty} f(x)\widehat{g}(x) dx = \int_{-\infty}^{\infty} \widehat{f}(y)g(y) dy.$$

---

*Proof.*

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\widehat{g}(x)dx &= \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} g(y)e^{-2\pi ixy} dy dx \\
(Fubini) \quad &= \int_{-\infty}^{\infty} g(y) \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy} dx dy \\
&= \int_{-\infty}^{\infty} \widehat{f}(y)g(y)dy.
\end{aligned}$$

□

**Theorem 6.1.5** (Fourier inversion) If  $f \in \mathcal{S}(\mathbb{R})$ , then

$$f(x) = \int_{-\infty}^{\infty} \widehat{f}(\xi)e^{2\pi i\xi x} d\xi.$$

*Proof.* We first show that

$$f(0) = \int_{-\infty}^{\infty} \widehat{f}(\xi)d\xi.$$

Let  $G_\delta(x) = e^{-\pi\delta x^2}$ . Note that  $\widehat{G}_\delta(\xi) = K_\delta(\xi)$ . By the previous lemma,

$$\begin{aligned}
\int_{-\infty}^{\infty} \widehat{f}(\xi)G_\delta(\xi)d\xi &= \int_{-\infty}^{\infty} f(x)\widehat{G}_\delta(x)dx \\
&= \int_{-\infty}^{\infty} f(x)K_\delta(x)dx \\
&= (f * K_\delta)(0).
\end{aligned}$$

By Proposition 6.1.2,

$$\lim_{\delta \rightarrow 0^+} (f * K_\delta)(0) = f(0).$$

On the other hand, by the dominated convergence theorem,

$$\lim_{\delta \rightarrow 0^+} \int_{-\infty}^{\infty} \widehat{f}(\xi)G_\delta(\xi)d\xi = \int_{-\infty}^{\infty} \widehat{f}(\xi)d\xi.$$

For general  $x \in \mathbb{R}$ , define  $F(y) = f(x + y)$ , then

$$f(x) = F(0) = \int_{-\infty}^{\infty} \widehat{F}(\xi)d\xi = \int_{-\infty}^{\infty} \widehat{f}(\xi)e^{2\pi i\xi x} d\xi.$$

□

---

Define  $\mathcal{F} : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$  and  $\mathcal{F}^{-1} : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$  by

$$\begin{aligned}\mathcal{F}(f)(\xi) &= \widehat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i \xi x} dx, \\ \mathcal{F}^{-1}(f)(x) &= \int_{-\infty}^{\infty} f(\xi)e^{2\pi i \xi x} d\xi.\end{aligned}$$

By the inverse formula, we know that  $\mathcal{F}^{-1} \circ \mathcal{F} = \text{Id}$ . Also, since  $\mathcal{F}(f)(y) = \mathcal{F}^{-1}(f)(-y)$ , we have  $\mathcal{F} \circ \mathcal{F}^{-1} = \text{Id}$ . Therefore,  $\mathcal{F}$  is a bijection on  $\mathcal{S}(\mathbb{R})$  with inverse  $\mathcal{F}^{-1}$ .

## 6.2 Convolutions

**Proposition 6.2.1** Let  $f, g \in \mathcal{S}(\mathbb{R})$ . Then

- (a)  $f * g \in \mathcal{S}(\mathbb{R})$ ,
- (b)  $f * g = g * f$ ,
- (c)  $\widehat{f * g}(\xi) = \widehat{f}(\xi) \cdot \widehat{g}(\xi)$ .

*Proof.*

- (a) We use the fact that for all  $l \geq 0$ ,  $\exists A_l > 0$  such that

$$\sup_{x \in \mathbb{R}} |x|^l |g(x - y)| \leq A_l (1 + |y|)^l.$$

Then, for all non-negative integers  $l$ ,

$$\sup_{x \in \mathbb{R}} |x|^l |(f * g)(x)| \leq \int_{-\infty}^{\infty} |f(y)| \sup_{x \in \mathbb{R}} |x|^l |g(x - y)| dy \leq A_l \int_{-\infty}^{\infty} |f(y)| (1 + |y|)^l dy < \infty.$$

We can verify that for all  $k \geq 1$ ,

$$(f * g)^{(k)}(x) = (f * g^{(k)})(x),$$

so by a similar argument, we can show that for all non-negative integers  $m, n$ ,

$$\sup_{x \in \mathbb{R}} |x|^m |(f * g)^{(n)}(x)| < \infty.$$

Therefore,  $f * g \in \mathcal{S}(\mathbb{R})$ .

- (b) By changing variable  $z = x - y$ , we have

$$(f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x - y) dy = \int_{-\infty}^{\infty} f(x - z)g(z) dz = (g * f)(x).$$

---

(c) Note that

$$\begin{aligned}
\widehat{f * g}(\xi) &= \int_{-\infty}^{\infty} (f * g)(x) e^{-2\pi i \xi x} dx \\
(\text{Fubini}) \quad &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y) g(x-y) e^{-2\pi i \xi x} dy dx \\
&= \int_{-\infty}^{\infty} f(y) e^{-2\pi i \xi y} \int_{-\infty}^{\infty} g(x-y) e^{-2\pi i \xi (x-y)} dx dy \\
&= \widehat{f}(\xi) \cdot \widehat{g}(\xi).
\end{aligned}$$

□

Define an inner product on  $\mathcal{S}(\mathbb{R})$  by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx, \quad \|f\| = \sqrt{\langle f, f \rangle}.$$

**Theorem 6.2.1** (Plancherel Theorem) For all  $f \in \mathcal{S}(\mathbb{R})$ ,

$$\langle f \rangle = \langle \widehat{f} \rangle.$$

This implies that  $\mathcal{F}$  is an isometry on  $\mathcal{S}(\mathbb{R})$ .

*Proof.* For  $f \in \mathcal{S}(\mathbb{R})$ , define  $f^*(x) = \overline{f(-x)}$ . Note that

$$\widehat{f^*}(\xi) = \int_{-\infty}^{\infty} \overline{f(-x)} e^{-2\pi i \xi x} dx = \overline{\widehat{f}(\xi)}.$$

Define  $h = f * f^*$ . By Proposition 6.2.1(c),

$$\widehat{h}(\xi) = \widehat{f}(\xi) \cdot \widehat{f^*}(\xi) = \widehat{f}(\xi) \cdot \overline{\widehat{f}(\xi)} = |\widehat{f}(\xi)|^2.$$

On the other hand,

$$\begin{aligned}
h(0) &= \int_{-\infty}^{\infty} f(y) f^*(-y) dy \\
&= \int_{-\infty}^{\infty} f(y) \overline{f(y)} dy = \|f\|^2.
\end{aligned}$$

By the Fourier inversion formula,

$$h(0) = \int_{-\infty}^{\infty} \widehat{h}(\xi) d\xi = \int_{-\infty}^{\infty} |\widehat{f}(\xi)|^2 d\xi = \|\widehat{f}\|^2.$$

Therefore,  $\|f\| = \|\widehat{f}\|$ , hence proved. □

**Remark**  $\mathcal{S}(\mathbb{R})$  is not complete with respect to the norm induced by the inner product.

### 6.3 Partial Differential Equations

We want to find functions  $u : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  satisfying the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2},$$

with initial condition  $u(x, 0) = f(x)$ . Fourier transform provides a powerful method to solve this problem by turning derivatives into polynomials. Taking Fourier transform with respect to  $x$  on both sides gives

$$\frac{\partial \hat{u}}{\partial t}(\xi, t) = -4\pi^2 \xi^2 \hat{u}(\xi, t).$$

This is an ordinary differential equation in  $t$  for each fixed  $\xi$ . Solving it gives

$$\hat{u}(\xi, t) = \hat{f}(\xi) e^{-4\pi^2 \xi^2 t}.$$

This  $\hat{f}(\xi)$  is a constant determined by the initial condition. Note that  $e^{-4\pi^2 \xi^2 t}$  is the Fourier transform of

$$\mathcal{H}_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t},$$

the heat kernel on  $\mathbb{R}$ . Therefore, we should have

$$u(x, t) = (f * \mathcal{H}_t)(x).$$

**Theorem 6.3.1** Given  $f \in \mathcal{S}(\mathbb{R})$ , define  $u(x, t) = (f * \mathcal{H}_t)(x)$  for  $x \in \mathbb{R}$  and  $t > 0$ . Then

- (a)  $u \in C^2$  and satisfies the heat equation for all  $x \in \mathbb{R}$  and  $t > 0$ ,
- (b)  $u(\cdot, t) \rightharpoonup f$  as  $t \rightarrow 0^+$ ,
- (c)  $\int_{-\infty}^{\infty} |u(x, t) - f(x)|^2 dx \rightarrow 0$  as  $t \rightarrow 0^+$ .

*Proof.*

- (a) Recall that

$$\hat{u}(\xi, t) = \hat{f}(\xi) e^{-4\pi^2 \xi^2 t}.$$

By Fourier inversion,

$$u(x, t) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{-4\pi^2 \xi^2 t} e^{2\pi i \xi x} d\xi.$$

Since  $\hat{f} \in \mathcal{S}(\mathbb{R})$ , we can differentiate under the integral sign to show that  $u$  has two continuous derivatives in  $x$  and satisfies the heat equation.

- (b) Since  $\{\mathcal{H}_t\}_{t>0}$  is an approximate to the identity on  $\mathbb{R}$ , by Proposition 6.1.2,  $u(\cdot, t) \rightharpoonup f$  as  $t \rightarrow 0^+$ .

(c) By Plancherel theorem,

$$\int_{-\infty}^{\infty} |u(x, t) - f(x)|^2 dx = \int_{-\infty}^{\infty} |\widehat{u}(\xi, t) - \widehat{f}(\xi)|^2 d\xi = \int_{-\infty}^{\infty} |\widehat{f}(\xi)|^2 |e^{-4\pi^2 \xi^2 t} - 1|^2 d\xi.$$

Given  $\epsilon > 0$ , exists some  $N > 0$  such that

$$\int_{|\xi|>N} |\widehat{f}(\xi)|^2 d\xi < \epsilon/2.$$

Since  $e^{-4\pi^2 \xi^2 t} \rightarrow 1$  as  $t \rightarrow 0^+$  for all  $\xi \in \mathbb{R}$ , by the dominated convergence theorem, exists some  $\delta > 0$  such that if  $0 < t < \delta$ , then

$$\int_{|\xi|\leq N} |\widehat{f}(\xi)|^2 |e^{-4\pi^2 \xi^2 t} - 1|^2 d\xi < \epsilon/2.$$

Therefore, for all  $0 < t < \delta$ ,

$$\int_{-\infty}^{\infty} |u(x, t) - f(x)|^2 dx < \epsilon,$$

hence proved.  $\square$

**Proposition 6.3.1** Let  $f \in \mathcal{S}(\mathbb{R})$  and  $u(x, t) = (f * \mathcal{H}_t)(x)$ . Then  $u(\cdot, t) \in \mathcal{S}(\mathbb{R})$  uniformly in  $t$ , in the sense that for all  $T > 0$ ,

$$\sup_{x \in \mathbb{R}, t \in (0, T)} |x|^k \left| \frac{\partial^l}{\partial x^l} u(x, t) \right| < \infty,$$

for all non-negative integers  $k, l$ .

*Proof.* Since  $f \in \mathcal{S}(\mathbb{R})$ , when  $|y| \leq |x|/2$ , for all  $N \geq 1$ , exists  $C_N > 0$  such that

$$|f(x - y)| \leq \frac{C_N}{1 + |x|^N}.$$

When  $|y| \geq |x|/2$ , we have

$$\mathcal{H}_t(y) \leq \frac{C}{\sqrt{t}} e^{-cx^2/t},$$

for some constant  $C, c > 0$ . Therefore,

$$\begin{aligned} |u(x, t)| &\leq \int_{|y| \leq |x|/2} \frac{C_N}{1 + |x|^N} \mathcal{H}_t(y) dy + C \int_{|y| \geq |x|/2} |f(x - y)| \frac{e^{-cx^2/t}}{\sqrt{t}} dy \\ &\leq \frac{C_N}{1 + |x|^N} + \frac{C'}{\sqrt{t}} e^{-cx^2/t}, \end{aligned}$$

---

for some constant  $C' > 0$ . Multiplying both sides by  $|x|^k$  and choosing  $N$  large enough shows that  $\sup_{x \in \mathbb{R}, t \in (0, T)} |x|^k |u(x, t)| < \infty$ . A similar argument works for the derivatives of  $u$ .  $\square$

**Theorem 6.3.2** Suppose that  $u : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  satisfies the following:

- (a)  $u$  is continuous on  $\mathbb{R} \times [0, \infty)$ ,
- (b)  $u$  satisfies the heat equation on  $\mathbb{R} \times (0, \infty)$ ,
- (c)  $u(x, 0) = 0$  for all  $x \in \mathbb{R}$ ,
- (d)  $u(\cdot, t) \in \mathcal{S}(\mathbb{R})$  uniformly in  $t$ .

Then  $u \equiv 0$ .

*Proof.* Define the energy function

$$E(t) = \int_{-\infty}^{\infty} |u(x, t)|^2 dx,$$

which is nonnegative and  $E(0) = 0$  by condition (c). Note that

$$\begin{aligned} E'(t) &= 2 \int_{-\infty}^{\infty} u(x, t) \frac{\partial u}{\partial t}(x, t) dx \\ (\text{integration by parts}) \quad &= -2 \int_{-\infty}^{\infty} \left( \frac{\partial u}{\partial x}(x, t) \right)^2 dx \leq 0, \end{aligned}$$

where the boundary terms vanish since  $u(\cdot, t) \in \mathcal{S}(\mathbb{R})$  uniformly in  $t$  by condition (d). Therefore,  $E(t)$  is non-increasing and nonnegative with  $E(0) = 0$ , so  $E(t) \equiv 0$ . This implies that  $u(x, t) \equiv 0$  for all  $x \in \mathbb{R}$  and  $t \geq 0$ .  $\square$

**Remark** The above theorem shows that the solution to the heat equation with initial condition in  $\mathcal{S}(\mathbb{R})$  is unique within the class of functions satisfying condition (d).

## 6.4 Laplace equation in the upper half plane

We can also use Fourier transform to solve the Dirichlet problem for the Laplace equation in the upper half plane  $\mathbb{H} = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ . Given a function  $f \in \mathcal{S}(\mathbb{R})$ , we want to find a function  $u : \mathbb{H} \rightarrow \mathbb{R}$  such that  $\Delta u = 0$  in  $\mathbb{H}$  and  $u(x, 0) = f(x)$  for all  $x \in \mathbb{R}$ .

Define the Poisson kernel on  $\mathbb{H}$  as

$$P_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2}, \quad x \in \mathbb{R}, y > 0.$$

---

Although  $P_y(x)$  is not in  $\mathcal{S}(\mathbb{R})$ , it is still integrable on  $\mathbb{R}$  with

$$\int_{-\infty}^{\infty} P_y(x) dx = 1.$$

Recall that the Fourier transform of the derivative is

$$\widehat{f}'(\xi) = 2\pi i \xi \widehat{f}(\xi).$$

Taking Fourier transform with respect to  $x$  on both sides of  $\Delta u = 0$  gives

$$\frac{\partial^2 \widehat{u}}{\partial y^2}(\xi, y) - 4\pi^2 \xi^2 \widehat{u}(\xi, y) = 0.$$

Solving this ODE gives

$$\widehat{u}(\xi, y) = A(\xi) e^{-2\pi|\xi|y} + B(\xi) e^{2\pi|\xi|y}.$$

Since  $u$  is bounded in  $\mathbb{H}$ , we must have  $B(\xi) = 0$ . By the boundary condition, we have

$$A(\xi) = \widehat{f}(\xi).$$

Therefore,

$$\widehat{u}(\xi, y) = \widehat{f}(\xi) e^{-2\pi|\xi|y}.$$

**Lemma 6.4.1**

$$\int_{-\infty}^{\infty} e^{-2\pi|\xi|y} e^{2\pi i \xi x} d\xi = P_y(x).$$

*Proof.* Note that

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-2\pi|\xi|y} e^{2\pi i \xi x} d\xi &= \int_0^{\infty} e^{-2\pi \xi y} e^{2\pi i \xi x} d\xi + \int_{-\infty}^0 e^{2\pi \xi y} e^{2\pi i \xi x} d\xi \\ &= \int_0^{\infty} e^{-2\pi \xi(y-ix)} d\xi + \int_0^{\infty} e^{-2\pi \xi(y+ix)} d\xi \\ &= \frac{1}{2\pi(y-ix)} + \frac{1}{2\pi(y+ix)} = \frac{y}{\pi(x^2+y^2)}. \end{aligned}$$

□

By the lemma and Fourier inversion, we have

$$u(x, y) = (f * P_y)(x).$$

**Lemma 6.4.2**  $(P_y)$  is an approximate to the identity on  $\mathbb{R}$ .

*Proof.*

---

(a) Follows from the plugging  $\xi = 0$  into the equation

$$\int_{-\infty}^{\infty} P_y(x) e^{-2\pi i \xi x} dx = e^{-2\pi |\xi| y}.$$

(b) Since  $P_y(x) > 0$  for all  $x \in \mathbb{R}$  and  $y > 0$ ,

$$\int_{-\infty}^{\infty} |P_y(x)| dx = \int_{-\infty}^{\infty} P_y(x) dx = 1.$$

(c) Given  $\eta > 0$ ,

$$\int_{|x| \geq \eta} \frac{y}{x^2 + y^2} dx = \tan^{-1} \left( \frac{x}{y} \right) \Big|_{-\infty}^{-\eta} + \tan^{-1} \left( \frac{x}{y} \right) \Big|_{\eta}^{\infty} = \pi - 2 \tan^{-1} \left( \frac{\eta}{y} \right).$$

As  $y \rightarrow 0^+$ ,  $\tan^{-1}(\eta/y) \rightarrow \pi/2$ , so

$$\lim_{y \rightarrow 0^+} \int_{|x| \geq \eta} P_y(x) dx = 0.$$

□

**Theorem 6.4.3** Let  $f \in \mathcal{S}(\mathbb{R})$  and define  $u(x, y) = (f * P_y)(x)$  for  $x \in \mathbb{R}$  and  $y > 0$ . Then

(a)  $u$  is harmonic in  $\mathbb{H}$ ,

(b)  $u(\cdot, y) \rightrightarrows f$  as  $y \rightarrow 0^+$ .

(c)  $\int_{-\infty}^{\infty} |u(x, y) - f(x)|^2 dx \rightarrow 0$  as  $y \rightarrow 0^+$ .

(d) If we further define  $u(x, 0) = f(x)$  for all  $x \in \mathbb{R}$ , then  $u$  is continuous on  $\mathbb{R} \times [0, \infty)$  and  $u(x, y) \rightarrow 0$  as  $|x| + y \rightarrow \infty$ .

*Proof.*

(a) Since

$$\widehat{u}(\xi, y) = \widehat{f}(\xi) e^{-2\pi |\xi| y},$$

by Fourier inversion, we have

$$u(x, y) = \int_{-\infty}^{\infty} \widehat{f}(\xi) e^{-2\pi |\xi| y} e^{2\pi i \xi x} d\xi.$$

We can differentiate under the integral sign to show that  $u$  is harmonic in  $\mathbb{H}$ .

(b) Since  $\{P_y\}_{y>0}$  is an approximate to the identity on  $\mathbb{R}$ , by Proposition 6.1.2,  $u(\cdot, y) \rightrightarrows f$  as  $y \rightarrow 0^+$ .

---

(c) By Plancherel theorem,

$$\int_{-\infty}^{\infty} |u(x, y) - f(x)|^2 dx = \int_{-\infty}^{\infty} |\widehat{f}(\xi)|^2 |e^{-2\pi|\xi|y} - 1|^2 d\xi.$$

Given  $\epsilon > 0$ , exists some  $N > 0$  such that

$$\int_{|\xi|>N} |\widehat{f}(\xi)|^2 d\xi < \epsilon/2.$$

Since  $e^{-2\pi|\xi|y} \rightarrow 1$  as  $y \rightarrow 0^+$  for all  $\xi \in \mathbb{R}$ , by the dominated convergence theorem, exists some  $\delta > 0$  such that if  $0 < y < \delta$ , then

$$\int_{|\xi|\leq N} |\widehat{f}(\xi)|^2 |e^{-2\pi|\xi|y} - 1|^2 d\xi < \epsilon/2.$$

Therefore, for all  $0 < y < \delta$ ,

$$\int_{-\infty}^{\infty} |u(x, y) - f(x)|^2 dx < \epsilon,$$

hence proved.

(d) Continuity at  $y = 0$  follows from (b). To show that  $u(x, y) \rightarrow 0$  as  $|x| + y \rightarrow \infty$ , note that

$$|(f * P_y)(x)| \leq C \cdot \min \left\{ \frac{1}{1+x^2} + \frac{y}{x^2+y^2}, \frac{1}{y} \right\} \rightarrow 0,$$

hence proved.

□

**Lemma 6.4.4** (Mean-value property) Let  $\Omega$  be an open subset of  $\mathbb{R}^2$ . Suppose that  $u \in C^2$  with  $\Delta u = 0$  is  $\Omega$ . If  $\overline{B_\ell(x, y), R} \subset \Omega$ , then

$$u(x, y) = \frac{1}{2\pi} \int_0^{2\pi} u(x + r \cos \theta, y + r \sin \theta) d\theta$$

for all  $0 \leq r \leq R$ .

*Proof.* Note that

$$\Delta u = 0 \implies \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} = 0.$$

---

Define  $F(r) = \frac{1}{2\pi} \int_0^{2\pi} u(r, \theta) d\theta$ . Then

$$\begin{aligned} r \frac{\partial}{\partial r} \left( r \frac{\partial F}{\partial r} \right) &= \frac{1}{2\pi} \int_0^{2\pi} r \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) d\theta \\ &= -\frac{1}{2\pi} \int_0^{2\pi} \frac{\partial^2 u}{\partial \theta^2} d\theta \\ &= -\frac{1}{2\pi} \frac{\partial u}{\partial \theta} \Big|_0^{2\pi} = 0. \end{aligned}$$

Therefore,  $r \frac{\partial F}{\partial r} = C$  for some constant  $C$ . Since  $F$  is bounded near  $r = 0$ , we must have  $C = 0$ . Therefore,  $F$  is constant, i.e.,

$$u(x, y) = F(0) = F(r) = \frac{1}{2\pi} \int_0^{2\pi} u(x + r \cos \theta, y + r \sin \theta) d\theta,$$

hence proved.  $\square$

**Theorem 6.4.5** Suppose that  $u$  is continuous on  $\overline{\mathbb{H}}$  and harmonic in  $\mathbb{H}$  with  $u(x, 0) = 0$  for all  $x \in \mathbb{R}$ . If  $u(x, y) \rightarrow 0$  as  $|x| + y \rightarrow \infty$ , then  $u \equiv 0$ .

*Proof.* Suppose that  $u$  is not identically zero. Then exists some point  $(x_0, y_0) \in \mathbb{H}$  such that  $u(x_0, y_0) \neq 0$ . Without loss of generality, assume that  $u(x_0, y_0) > 0$ . Since  $u$  vanishes at infinity, exists some  $R > 0$  such that if  $\sqrt{x^2 + y^2} \geq R$ , then  $|u(x, y)| < u(x_0, y_0)/2$ . Since  $u$  is continuous, it attains a maximum  $M > 0$  on the compact set  $\overline{B_{(0,0),R}} \cap \overline{\mathbb{H}}$  at some point  $(x_1, y_1)$ . In fact, the maximum is attained in the whole  $\overline{\mathbb{H}}$  by choice of  $R$ . Note that  $y_1 > 0$  since  $u(x, 0) = 0$  for all  $x \in \mathbb{R}$ . By the mean-value property, for all  $0 < r < y_1$ ,

$$M = u(x_1, y_1) = \frac{1}{2\pi} \int_0^{2\pi} u(x_1 + r \cos \theta, y_1 + r \sin \theta) d\theta.$$

Since  $u(x_1 + r \cos \theta, y_1 + r \sin \theta) \leq M$  for all  $\theta$ , we must have

$$u(x_1 + r \cos \theta, y_1 + r \sin \theta) = M$$

for all  $\theta$ . Taking  $r \rightarrow y_1^-$  shows that  $u \equiv M$  on the line segment  $\{(x, 0) \mid x_1 - y_1 \leq x \leq x_1 + y_1\}$ , contradicting the boundary condition. Therefore,  $u \equiv 0$ .  $\square$

## 6.5 Poisson summation formula

Given a function  $f \in \mathcal{S}(\mathbb{R})$ , we can construct a 1-periodic function by

$$F(x) = \sum_{n=-\infty}^{\infty} f(x + n).$$

---

Since  $f$  decays rapidly, the series converges absolutely and uniformly on  $\mathbb{R}$ . Therefore,  $F \in C^\infty(\mathbb{R})$  and  $F(x+1) = F(x)$  for all  $x \in \mathbb{R}$ .  $F$  is called the periodization of  $f$ . Another way to periodize  $f$  is to use the Fourier series

$$\widetilde{F}(x) = \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{2\pi i n x},$$

where  $\widehat{f}(n)$  is the Fourier transform of  $f$  evaluated at integer  $n$ . Since  $\widehat{f} \in \mathcal{S}(\mathbb{R})$ , the series converges absolutely and uniformly on  $\mathbb{R}$ . Therefore,  $\widetilde{F} \in C^\infty(\mathbb{R})$  and  $\widetilde{F}(x+1) = \widetilde{F}(x)$  for all  $x \in \mathbb{R}$ .

It turns out that these two periodizations are the same.

**Theorem 6.5.1** (Poisson summation formula) For all  $f \in \mathcal{S}(\mathbb{R})$  and  $x \in \mathbb{R}$ ,

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{2\pi i n x}.$$

*Proof.* Since both sides are 1-periodic in  $x$ , it suffices to show that they are equal for  $x \in [0, 1)$ . Note that the Fourier coefficients of  $F$  are

$$\int_0^1 F(x) e^{-2\pi i m x} dx = \int_{-\infty}^{\infty} f(x) e^{-2\pi i m x} dx = \widehat{f}(m),$$

for all integers  $m$ . Therefore, the Fourier series of  $F$  is

$$\sum_{n=-\infty}^{\infty} \widehat{f}(n) e^{2\pi i n x} = \widetilde{F}(x).$$

Since both  $F$  and  $\widetilde{F}$  are continuous functions on  $\mathbb{R}$  with the same Fourier coefficients, they must be equal. Hence proved.  $\square$

**Remark**  $f \in \mathcal{S}(\mathbb{R})$  is not necessary for the Poisson summation formula to hold. A sufficient condition is that both  $f$  and  $\widehat{f}$  are continuous on  $\mathbb{R}$  and the series  $\sum_{n=-\infty}^{\infty} f(x+n)$  converges absolutely for all  $x \in \mathbb{R}$ .

**Example 6.5.1** Recall the heat kernel on  $\mathbb{R}$ ,

$$\mathcal{H}_t(x) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t}.$$

Its Fourier transform is

$$\widehat{\mathcal{H}}_t(\xi) = e^{-4\pi^2 \xi^2 t}.$$

---

By the Poisson summation formula,

$$\sum_{n=-\infty}^{\infty} \mathcal{H}_t(x+n) = \sum_{n=-\infty}^{\infty} e^{-4\pi^2 n^2 t} e^{2\pi i n x},$$

which is precisely the heat kernel  $H_t$  on the circle  $S^1 = \mathbb{R}/\mathbb{Z}$ .

**Corollary 6.5.2**  $(H_t)$  is an approximate to the identity on  $S^1$ .

*Proof.*

$$\begin{aligned} \int_0^1 H_t(x) dx &= \int_0^1 \sum_{n=-\infty}^{\infty} \mathcal{H}_t(x+n) dx \\ (\text{Fubini}) \quad &= \sum_{n=-\infty}^{\infty} \int_n^{n+1} \mathcal{H}_t(x) dx = \int_{-\infty}^{\infty} \mathcal{H}_t(x) dx = 1. \end{aligned}$$

Also,  $H_t(x) > 0$  for all  $x \in \mathbb{R}$  and  $t > 0$  since  $\mathcal{H}_t(x) > 0$  for all  $x \in \mathbb{R}$ .

Note that

$$H_t(x) = \mathcal{H}_t(x) + \sum_{n \neq 0} \mathcal{H}_t(x+n).$$

When  $|x| < 1/2$ ,

$$\begin{aligned} H_t(x) - \mathcal{H}_t(x) &= \frac{1}{\sqrt{4\pi t}} \sum_{n \neq 0} e^{-(x+n)^2/4t} \\ &\leq C_1 t^{-1/2} \sum_{n=1}^{\infty} e^{-C_2 n^2/t} \end{aligned}$$

for some constants  $C_1, C_2 > 0$ . Note that  $n^2/t \geq n^2$  and  $n^2/t \geq 1/t$  for all  $n \geq 1$  and  $t \in (0, 1]$ , so  $n^2/t \geq (n^2 + 1/t)/2$ . Therefore,

$$\begin{aligned} |E_t(x)| &\leq C_1 t^{-1/2} \sum_{n=1}^{\infty} e^{-C_2(n^2 + 1/t)/2} \\ &= C_1 t^{-1/2} e^{-C_2/2t} \sum_{n=1}^{\infty} e^{-C_2 n^2/2} \\ &\leq C_3 t^{-1/2} e^{-C_2/2t}, \end{aligned}$$

Finally, given  $\eta > 0$ ,

$$\int_{|x| \geq \eta} H_t(x) dx = \int_{|x| \geq \eta} \mathcal{H}_t(x) dx + \int_{|x| \geq \eta} E_t(x) dx \rightarrow 0$$

as  $t \rightarrow 0^+$ , where the first term tends to 0 since  $(\mathcal{H}_t)$  is an approximate to the identity on  $\mathbb{R}$  and the second term tends to 0 since  $E_t(x)$  decays to 0 exponentially fast as  $t \rightarrow 0^+$ . Hence

proved.  $\square$

Similarly, we can use the Poisson summation formula to periodize the Poisson kernel on  $\mathbb{R}$  to obtain the Poisson kernel on the unit disk.

**Corollary 6.5.3**

$$P_r(2\pi x) = \sum_{n=-\infty}^{\infty} \mathcal{P}_y(x+n) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{2\pi i n x},$$

where  $r = e^{-2\pi y}$ .

## 6.6 Heisenberg uncertainty principle

**Theorem 6.6.1** Suppose that  $f \in \mathcal{S}(\mathbb{R})$  satisfies

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = 1,$$

then

$$\left( \int_{-\infty}^{\infty} x^2 |f(x)|^2 dx \right) \left( \int_{-\infty}^{\infty} \xi^2 |\widehat{f}(\xi)|^2 d\xi \right) \geq \frac{1}{16\pi^2},$$

with equality if and only if

$$f(x) = Ae^{-Bx^2}$$

where  $B > 0$  and  $|A|^2 = \sqrt{2B/\pi}$ . In fact, we have

$$\left( \int_{-\infty}^{\infty} (x - x_0)^2 |f(x)|^2 dx \right) \left( \int_{-\infty}^{\infty} (\xi - \xi_0)^2 |\widehat{f}(\xi)|^2 d\xi \right) \geq \frac{1}{16\pi^2},$$

for all  $x_0, \xi_0 \in \mathbb{R}$ .

*Proof.* The second inequality follows from the first by replacing  $f(x)$  by  $f(x + x_0)e^{-2\pi i \xi_0 x}$  and changing variables. To prove the first inequality, note that by integration by parts,

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} |f(x)|^2 dx = x|f(x)|^2 \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} x \frac{d}{dx} |f(x)|^2 dx \\ &= - \int_{-\infty}^{\infty} x \frac{d}{dx} (f(x) \overline{f(x)}) dx \\ &= - \int_{-\infty}^{\infty} x (f'(x) \overline{f(x)} + f(x) \overline{f'(x)}) dx \\ \implies 1 &\leq 2 \int_{-\infty}^{\infty} |x| |f(x)| |f'(x)| dx \\ (\text{Cauchy-Schwarz}) \quad &\leq 2 \left( \int_{-\infty}^{\infty} x^2 |f(x)|^2 dx \right)^{1/2} \left( \int_{-\infty}^{\infty} |f'(x)|^2 dx \right)^{1/2}. \end{aligned}$$

---

By Plancherel theorem,

$$\int_{-\infty}^{\infty} |f'(x)|^2 dx = \int_{-\infty}^{\infty} 4\pi^2 \xi^2 |\widehat{f}(\xi)|^2 d\xi.$$

Therefore,

$$\left( \int_{-\infty}^{\infty} x^2 |f(x)|^2 dx \right) \left( \int_{-\infty}^{\infty} \xi^2 |\widehat{f}(\xi)|^2 d\xi \right) \geq \frac{1}{16\pi^2}.$$

For the equality case, note that equality holds in the Cauchy-Schwarz inequality if and only if there exists some constant  $\alpha$  such that

$$f'(x) = \alpha x f(x) \implies f(x) = A e^{\frac{\alpha x^2}{2}}.$$

Since  $\int_{-\infty}^{\infty} |f(x)|^2 dx = 1$ , we must have  $\alpha < 0$ . Letting  $\alpha = -2B$  for some  $B > 0$  gives the desired form of  $f$ . Finally, computing the integral shows that  $|A|^2 = \sqrt{2B/\pi}$ .  $\square$

Physically, the Heisenberg uncertainty principle states that it is impossible to simultaneously measure the position and momentum of a particle with arbitrary precision. If  $f$  represents the wave function of a particle, then  $|f(x)|^2$  represents the probability density of finding the particle at position  $x$ , while

$$\int_{-\infty}^{\infty} (x - \bar{x})^2 |f(x)|^2 dx$$

represents the variance in position, where

$$\bar{x} = \int_{-\infty}^{\infty} x |f(x)|^2 dx.$$

Similarly,  $|\widehat{f}(\xi)|^2$  represents the probability density of finding the particle with momentum  $p = 2\pi\xi$ , while

$$\int_{-\infty}^{\infty} (\xi - \bar{\xi})^2 |\widehat{f}(\xi)|^2 d\xi$$

represents the variance in momentum, where

$$\bar{\xi} = \int_{-\infty}^{\infty} \xi |\widehat{f}(\xi)|^2 d\xi.$$

Therefore, the Heisenberg uncertainty principle states that the product of the variance in position and the variance in momentum is bounded below by a positive constant, making it impossible to simultaneously measure both quantities with arbitrary precision.

## 6.7 Generalizations of Fourier transform

In Fourier series, we considered Riemann integrable functions on the interval  $[0, 1]$ . From the Parseval's identity, we know that there exists sequences in  $\ell^2(\mathbb{Z})$  that do not correspond to any Riemann integrable functions on  $[0, 1]$ . To understand the actual correspondence, we need to understand the completion of  $\mathcal{R}$  under  $\|\cdot\|_{L^2}$ .

In Fourier transform on  $\mathbb{R}$ , we considered functions in  $\mathcal{S}(\mathbb{R})$ . We can extend the Fourier transform to a more general class of functions, even if  $\int_{-\infty}^{\infty} |f(x)|dx = \infty$ . If we can find a "function"  $\widehat{f}$  such that

$$\int_{-\infty}^{\infty} f(x)g(x)dx = \int_{-\infty}^{\infty} \widehat{f}(\xi)\widehat{g}(\xi)d\xi,$$

for all  $g \in \mathcal{S}(\mathbb{R})$ , then we can define the Fourier transform of  $f$  to be  $\widehat{f}$ .

Periodic functions can be seen as functions on the circle  $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ .  $\mathbb{R}$  is an abelian group under addition. The circle  $\mathbb{S}^1$  is also an abelian group under addition modulo 1. Moreover, they are locally compact Hausdorff topological groups under the usual topology.

**Definition 6.7.1** A **locally compact abelian group** (LCA group) is an abelian group  $G$  equipped with a topology such that

- (a)  $G$  is a topological group, i.e., the group operations (addition and inversion) are continuous,
- (b)  $G$  is Hausdorff,
- (c)  $G$  is locally compact, i.e., for every  $x \in G$ , there exists an open neighborhood  $U$  of  $x$  such that  $\overline{U}$  is compact.

A character on an LCA group  $G$  is a continuous group homomorphism  $\chi : G \rightarrow \mathbb{S}^1$ . The set of all characters on  $G$  is called the **dual group** of  $G$ , denoted by  $\widehat{G}$ .

**Example 6.7.1**  $\widehat{\mathbb{S}^1} \cong \mathbb{Z}$ ,  $\widehat{\mathbb{Z}} \cong \mathbb{S}^1$ ,  $\widehat{\mathbb{R}} \cong \mathbb{R}$ , and  $\widehat{\mathbb{Z}/n\mathbb{Z}} \cong \mathbb{Z}/n\mathbb{Z}$ .

Define the topology on  $\widehat{G}$  such that a sequence of characters  $\{\chi_n\}$  converges to a character  $\chi$  if and only if  $\chi_n(x) \Rightarrow \chi(x)$  on all compact subsets of  $G$ . With this topology,  $\widehat{G}$  is also an LCA group.

**Theorem 6.7.2** (Pontryagin duality theorem) For any LCA group  $G$ ,  $\widehat{\widehat{G}} \cong G$ .

The Fourier transform can be generalized to LCA groups using the dual group.

**Definition 6.7.3** Let  $G$  be an LCA group with a Haar measure  $\mu$ . For  $f \in L^1(G)$ , the

---

Fourier transform of  $f$  is the function  $\widehat{f} : \widehat{G} \rightarrow \mathbb{C}$  defined by

$$\widehat{f}(\chi) = \int_G f(x)\overline{\chi(x)}d\mu(x),$$

for all  $\chi \in \widehat{G}$ .

The integral  $\int_G$  satisfies the following properties for all continuous compactly supported functions  $f$ :

- (a)  $\int_G f(x \cdot y)d\mu(x) = \int_G f(x)d\mu(x)$  for all  $y \in G$  (left invariance),
- (b)  $\int_G f(x)d\mu(x) \geq 0$  if  $f(x) \geq 0$  for all  $x \in G$ ,
- (c)  $\int_G f(x)d\mu(x)$  is linear in  $f$ .

In fact, such an integral is unique up to a multiplicative constant.

# CHAPTER A

## A.1 Separation Axioms

**Definition A.1.1** Let  $X$  be a topological space.

- $X$  is  $T_0$  if for all distinct  $x, y \in X$ ,  $\exists$  a neighborhood of one of them that does not contain the other.
- $X$  is  $T_1$  if for all distinct  $x, y \in X$ ,  $\exists$  a neighborhood of  $x$  that does not contain  $y$ , and  $\exists$  a neighborhood of  $y$  that does not contain  $x$ .
- $X$  is  $T_2$  (Hausdorff) if for all distinct  $x, y \in X$ ,  $\exists$  a neighborhood of  $x$  and a neighborhood of  $y$  that are disjoint.

**Theorem A.1.2** A topological space  $X$  is  $T_1$  iff every one-point set is closed.

**Definition A.1.3** Let  $X$  be a topological space. We say that  $X$  is normal if for every pair of disjoint closed sets  $A, B \subseteq X$ ,  $\exists$  disjoint open sets  $\mathcal{U}, \mathcal{V}$  such that  $A \subseteq \mathcal{U}$  and  $B \subseteq \mathcal{V}$ .

**Definition A.1.4** Let  $X$  be a topological space. We say that  $X$  is completely regular if for every closed set  $A \subseteq X$  and every point  $x \in X \setminus A$ ,  $\exists$  a continuous function  $f : X \rightarrow [0, 1]$  such that  $f(x) = 1$  and  $f|_A = 0$ .

**Definition A.1.5** Let  $X$  be a topological space.

- $X$  is  $T_3$  if  $X$  is  $T_1$  and regular.
- $X$  is  $T_{3\frac{1}{2}}$  if  $X$  is  $T_1$  and completely regular.
- $X$  is  $T_4$  if  $X$  is  $T_1$  and normal.

---

**Example A.1.1** We'll show that

$$T_4 \subsetneq T_{3\frac{1}{2}} \subsetneq T_3 \subsetneq T_2 \subsetneq T_1 \subsetneq T_0,$$

by providing examples that separate each pair of consecutive classes.

- $T_1 \subsetneq T_0$ : Let  $X = \{a, b\}$ , and let  $\mathcal{T} = \{\emptyset, X, \{a\}\}$ . Then  $(X, \mathcal{T})$  is  $T_0$  but not  $T_1$ . This construction is called the Sierpinski space.
- $T_2 \subsetneq T_1$ : Let  $X = \mathbb{R}$  with the cofinite topology, i.e.  $\mathcal{T} = \{\emptyset\} \cup \{U \subseteq \mathbb{R} \mid U^c \text{ is finite}\}$ . Then  $(X, \mathcal{T})$  is  $T_1$  but not  $T_2$ .
- $T_3 \subsetneq T_2$ : Let  $X = \mathbb{R}$ . Consider the  $K$ -topology on  $\mathbb{R}$ , i.e.

$$\mathcal{T} = \{\emptyset\} \cup \left\{ U \subseteq \mathbb{R} \mid U = V \setminus K, V \text{ is open in the usual topology}, K \subseteq \left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\} \right\}.$$

Then  $(X, \mathcal{T})$  is  $T_2$  but not  $T_3$ . This is because  $\left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\}$  is closed in  $\mathcal{T}$ .

- $T_{3\frac{1}{2}} \subsetneq T_3$ : See [Muukres, Topology, Second Edition, P.214].

For  $T_4 \subsetneq T_{3\frac{1}{2}}$ , we first prove the following proposition:

**Proposition A.1.1** For  $k = 0, 1, 2, 3, 3\frac{1}{2}$ ,  $T_k$  is closed under subspace topology, while  $T_4$  is not.

*Proof.* Consider the Sorgenfrey line  $\mathbb{R}_l = (\mathbb{R}, \mathcal{T}_S)$ , where  $\mathcal{T}_S$  is generated by the basis

$$\{[a, b) \mid a < b, a, b \in \mathbb{R}\}.$$

Clearly,  $\mathbb{R}_l$  is  $T_4$ . Let  $\Delta = \{(x, x) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}_l \times \mathbb{R}_l$ . We claim that  $\Delta$  is not normal in the subspace topology.  $\square$

## A.2 Euclid's Theorem

**Theorem A.2.1** (Euclid's Theorem) There are infinitely many prime numbers.

*Proof.* Consider a topology on  $\mathbb{Z}$ . For each  $a \in \mathbb{Z}$  and  $b \in \mathbb{N}$ , let

$$N_{a,b} = \{a + nb \mid n \in \mathbb{Z}\}.$$

We say a set  $\mathcal{U} \subseteq \mathbb{Z}$  is open if for every  $a \in \mathcal{U}$ ,  $\exists b \in \mathbb{N}$  such that  $N_{a,b} \subseteq \mathcal{U}$ .

It is easy to verify that the union of any collection of open sets is open. The intersection of two open sets is also open, since if  $a \in \mathcal{U}_1 \cap \mathcal{U}_2$ , then  $\exists b_1, b_2 \in \mathbb{N}$  such that  $N_{a,b_1} \subseteq \mathcal{U}_1$  and  $N_{a,b_2} \subseteq \mathcal{U}_2$ , hence  $N_{a,\text{lcm}(b_1, b_2)} \subseteq \mathcal{U}_1 \cap \mathcal{U}_2$ . Thus this is indeed a topology.

---

Note that  $N_{a,b}$  is also closed, since

$$N_{a,b} = \mathbb{Z} \setminus \bigcup_{i=1}^{b-1} N_{a+i,b}.$$

Since any number  $n \neq \pm 1$  has a prime divisor, we have

$$\mathbb{Z} \setminus \{-1, 1\} = \bigcup_{p \text{ is prime}} N_{0,p}.$$

If there are only finitely many primes  $p_1, p_2, \dots, p_k$ , then  $\mathbb{Z} \setminus \{-1, 1\}$  is a finite union of closed sets, hence closed. This implies that  $\{-1, 1\}$  is open, which is a contradiction since any nonempty open set in this topology is infinite.  $\square$

### A.3 Quotient Topology

**Definition A.3.1** Let  $X, Y$  be topological spaces. A surjective map  $p : X \rightarrow Y$  is called a quotient map if a set  $U \subseteq Y$  is open iff  $p^{-1}(U)$  is open in  $X$ .

**Definition A.3.2** A subset  $C \subseteq X$  is called saturated w.r.t. a map  $p : X \rightarrow Y$  if  $C = p^{-1}(S)$  for some  $S \subseteq Y$ .

**Definition A.3.3** A mapping  $p : X \rightarrow Y$  is called open (resp. closed) if for every open (resp. closed) set  $U \subseteq X$ ,  $p(U)$  is open (resp. closed) in  $Y$ .

**Proposition A.3.1** The following are equivalent:

- (a)  $p$  is a quotient map.
- (b)  $p$  is surjective, continuous, and maps saturated open sets to open sets.
- (c)  $p$  is surjective, continuous, and is either open or closed.

**Definition A.3.4** Let  $X$  be a topological space,  $A$  be a set, and  $p : X \rightarrow A$  be a surjective map. Then there is a unique topology  $\mathcal{T}$  on  $A$  such that  $p : (X, \mathcal{T}_X) \rightarrow (A, \mathcal{T})$  is a quotient map, namely

$$\mathcal{T} = \{U \subseteq A \mid p^{-1}(U) \text{ is open in } X\}.$$

This topology is called the quotient topology induced by  $p$ .

---

**Example A.3.1** Let  $p : \mathbb{R} \rightarrow \{a, b, c\}$  be defined by

$$p(x) = \begin{cases} a, & x > 0; \\ b, & x < 0; \\ c, & x = 0. \end{cases}$$

Then the quotient topology on  $\{a, b, c\}$  is  $\{\emptyset, \{a\}, \{b\}, \{a, b\}, \{a, b, c\}\}$ .

**Definition A.3.5** Let  $X$  be a topological space, and let  $X^*$  be a partition of  $X$ . Let  $\pi : X \rightarrow X^*$  be the natural projection that maps each point to its partition. Then under the quotient topology on  $X^*$  induced by  $\pi$ ,  $X^*$  is called the quotient space or the identification space.

**Example A.3.2** Let  $X$  be the closed unit disk  $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ . Let

$$X^* = \{\{(x, y) \mid x^2 + y^2 < 1\}\} \cup \{\{(x, y) \mid x^2 + y^2 = 1\}\}.$$

Then  $X^*$  with the quotient topology is homeomorphic to the sphere  $S^2$ .

**Theorem A.3.6** Let  $p : X \rightarrow Y$  be a quotient map. Let  $Z$  be a topological space, and  $g : X \rightarrow Z$  be a map that is constant on each  $p^{-1}(\{y\})$ ,  $y \in Y$ . Then,  $g$  induces a map  $f : Y \rightarrow Z$  such that  $f \circ p = g$ . Furthermore,  $f$  is continuous iff  $g$  is continuous, and  $f$  is a quotient map iff  $g$  is a quotient map.

**Corollary A.3.7** Let  $g : X \rightarrow Z$  be a surjective continuous map. Let  $X^* = \{g^{-1}(\{z\}) \mid z \in Z\}$ . Equip  $X^*$  with the quotient topology.

- (a) The map  $g$  induces a bijection  $f : X^* \rightarrow Z$ , which is a homeomorphism iff  $g$  is a quotient map.
- (b) If  $Z$  is Hausdorff, then so is  $X^*$ .

## A.4 Paracompactness

**Definition A.4.1** Consider a collection of subset of space  $X$ , say  $\{U_\alpha\}_{\alpha \in A}$ .

1. The collection  $\{U_\alpha\}_{\alpha \in A}$  is said to be locally finite if for every  $x \in X$ ,  $\exists$  a neighborhood  $V$  of  $x$  such that  $V$  intersects only finitely many  $U_\alpha$ 's.
2. The collection  $\{U_\alpha\}_{\alpha \in A}$  is said to be point finite if for every  $x \in X$ ,  $x$  belongs to only finitely many  $U_\alpha$ 's.
3. Suppose  $\{U_\alpha\}_{\alpha \in A}$  is an open cover of  $X$ . A refinement of  $\{U_\alpha\}_{\alpha \in A}$  is another open cover  $\{V_\beta\}_{\beta \in B}$  such that for every  $\beta \in B$ ,  $\exists \alpha \in A$  such that  $V_\beta \subseteq U_\alpha$ , denoted by  $\{V_\beta\}_{\beta \in B} \prec \{U_\alpha\}_{\alpha \in A}$ .

---

**Proposition A.4.1** Let  $\{U_\alpha\}_{\alpha \in A}$  be a collection of subsets of  $X$ . If  $\{U_\alpha\}_{\alpha \in A}$  is locally finite, then

- (i) The collection  $\{\overline{U_\alpha}\}_{\alpha \in A}$  is also locally finite.
- (ii) For any  $B \subseteq A$ ,  $\bigcup_{\beta \in B} \overline{U_\beta} = \overline{\bigcup_{\beta \in B} U_\beta}$  is closed.

*Proof.* Given  $x \in X$ , let  $V$  be a neighborhood of  $x$  that intersects only finitely many  $U_\alpha$ 's, say  $U_{\alpha_1}, U_{\alpha_2}, \dots, U_{\alpha_n}$ . Then  $V$  also intersects only finitely many  $\overline{U_\alpha}$ 's, namely  $\overline{U_{\alpha_1}}, \overline{U_{\alpha_2}}, \dots, \overline{U_{\alpha_n}}$ . Thus (i) holds.

Let  $B \subseteq A$  be given. Since  $\{U_\beta\}_{\beta \in B} \subseteq \{U_\alpha\}_{\alpha \in A}$ ,  $\{U_\beta\}_{\beta \in B}$  is also locally finite. By (i),  $\{\overline{U_\beta}\}_{\beta \in B}$  is locally finite. For any  $x \in X \setminus \bigcup_{\beta \in B} \overline{U_\beta}$ , there is a neighborhood  $V$  of  $x$  that intersects only finitely many  $\overline{U_\beta}$ 's, say  $\overline{U_{\beta_1}}, \overline{U_{\beta_2}}, \dots, \overline{U_{\beta_m}}$ . Note that

$$U' = V \setminus \bigcup_{i=1}^m \overline{U_{\beta_i}}$$

is an open set, hence a neighborhood of  $x$  that does not intersect any  $\overline{U_\beta}$ . Thus  $X \setminus \bigcup_{\beta \in B} \overline{U_\beta}$  is open, i.e.  $\bigcup_{\beta \in B} \overline{U_\beta}$  is closed.  $\square$

**Lemma A.4.2** (Pasting Lemma) Let  $\{U_\alpha\}_{\alpha \in A}$  be a cover of  $X$  and  $f : X \rightarrow Y$  be a map. If either

- (i) each  $U_\alpha$  is open, or
- (ii) each  $U_\alpha$  is closed and  $\{U_\alpha\}_{\alpha \in A}$  is locally finite,

then  $f$  is continuous iff  $f|_{U_\alpha}$  is continuous for each  $\alpha \in A$ .

*Proof.* If  $f$  is continuous, then obviously  $f|_{U_\alpha}$  is continuous for each  $\alpha \in A$ .

Conversely, suppose  $f|_{U_\alpha}$  is continuous for each  $\alpha \in A$ . If (i) holds, then for any open set  $V \subseteq Y$ ,

$$f^{-1}(V) = \bigcup_{\alpha \in A} (f|_{U_\alpha})^{-1}(V) \cap U_\alpha$$

is open, hence  $f$  is continuous.

If (ii) holds, then for any closed set  $V \subseteq Y$ ,  $\square$

**Definition A.4.3** Let  $\{U_\alpha\}_{\alpha \in A}$  and  $\{V_\beta\}_{\beta \in B}$  be two covers of  $X$  such that  $\{V_\beta\}_{\beta \in B} \prec \{U_\alpha\}_{\alpha \in A}$ . We say that  $\{V_\beta\}_{\beta \in B}$  is precise if  $B = A$  and  $V_\alpha \subseteq U_\alpha$  for each  $\alpha \in A$ .

**Proposition A.4.2** Suppose  $\{A_\alpha\}_{\alpha \in A}$  has a locally point finite refinement  $\{B_\beta\}_{\beta \in B}$ . Then  $\{A_\alpha\}_{\alpha \in A}$  has a precise locally finite refinement  $\{C_\alpha\}_{\alpha \in A}$ .

---

*Proof.* As a refinement, for each  $\beta \in B$ , the set  $I_\beta = \{\alpha \in A \mid B_\beta \subseteq A_\alpha\}$  is nonempty. By the axiom of choice, there exists a choice function  $\phi : B \rightarrow A$  such that  $\phi(\beta) \in I_\beta$  for each  $\beta \in B$ . Then for each  $\alpha \in A$ , let  $C_\alpha = \bigcup_{\beta \in \phi^{-1}(\{\alpha\})} B_\beta$ . It is easy to verify that  $\{C_\alpha\}_{\alpha \in A}$  is a precise refinement of  $\{A_\alpha\}_{\alpha \in A}$ .  $\square$

**Definition A.4.4** A topological space  $X$  is said to be paracompact if every open cover of  $X$  has a locally finite open refinement.

**Remark** Every compact space is paracompact. The converse is not true, e.g.  $\mathbb{R}$  is paracompact but not compact.

**Theorem A.4.5** A paracompact  $T_2$ (Hausdorff) space is  $T_4$  (normal).

*Proof.* Let  $X$  be a paracompact  $T_2$  space. We first show that  $X$  is regular. Given a closed set  $F \subseteq X$  and a point  $x \in X \setminus F$ , for each  $y \in F$ , by Hausdorffness,  $\exists$  a neighborhood  $U_y$  of  $y$  such that  $x \notin \overline{U_y}$ . Then  $\{U_y\}_{y \in F} \cup \{X \setminus F\}$  is an open cover of  $X$ . Since  $X$  is paracompact,  $\exists$  a locally finite open refinement  $\{V_\alpha\}_{\alpha \in A}$ . Let  $W$  be the union of all  $V_\alpha$ 's that intersect  $F$ . Then  $W$  is open, contains  $F$ , and does not contain  $x$ . Let  $U = X \setminus \overline{W}$ . Then  $U$  is open, contains  $x$ , and is disjoint from  $W$ . Thus  $X$  is regular.  $\square$

**Definition A.4.6** Let  $X, Y$  be spaces, and  $f : X \rightarrow Y$  be a map. The support of  $f$  is the set  $\text{supp}(f) = \overline{\{x \in X \mid f(x) \neq 0\}}$ .

**Definition A.4.7** Let  $X$  be a space. A family of continuous mappings  $\{k_\alpha\}_{\alpha \in A}$  from  $X$  to  $[0, 1]$  is called a partition of unity if

- (i)  $\{\text{supp}(k_\alpha)\}_{\alpha \in A}$  is locally finite.
- (ii) for each  $x \in X$ ,  $\sum_{\alpha \in A} k_\alpha(x) = 1$ .

**Theorem A.4.8** If  $X$  is paracompact and  $T_2$ , then every open cover  $\mathcal{U}_{\alpha \in A}$  has a partition of unity  $\{k_\alpha\}_{\alpha \in A}$  subordinate to it, i.e.  $\text{supp}(k_\alpha) \subseteq U_\alpha$  for each  $\alpha \in A$ .

**Proposition A.4.3** Let  $X$  be a paracompact  $T_2$  space, and  $G, g : X \rightarrow \mathbb{R}$  are upper semi-continuous and lower semicontinuous functions respectively such that  $G(x) < g(x)$  for each  $x \in X$ . Then  $\exists$  a continuous function  $f : X \rightarrow \mathbb{R}$  such that  $G(x) < f(x) < g(x)$  for each  $x \in X$ .

## A.5 More on Cantor Set

**Lemma A.5.1** If  $M$  is a nonempty compact metric space, then for every  $\epsilon > 0$ ,  $M$  can be expressed as a finite union of nonempty compact sets of diameter less than  $\epsilon$ .

---

*Proof.* For each  $x \in M$ , let  $U_x = B(x, \epsilon/2)$ . Then  $\{U_x\}_{x \in M}$  is an open cover of  $M$ . Take its finite subcovering, then take the closure of each set.  $\square$

**Theorem A.5.2** (Cantor Surjection Theorem) Given any compact metric space  $M$ , there is a continuous surjection  $f : C \rightarrow M$ , where  $C$  is the Cantor set.

*Proof.* Let  $W(n)$  be the set of binary strings with 0 and 2 of length  $n$ . By Lemma A.5.1, we can divide  $M$  into a finite number of nonempty compact sets of diameter at most 1. Say there are  $m \leq 2^{n_1}$  such sets, indexed by  $W(n_1)$ . Denote these sets by  $S_\alpha$ ,  $\alpha \in W(n_1)$ . Then, divide  $S_\alpha$  into finitely many pieces of diameter at most  $1/2$ , say there are at most  $2^{n_2}$  pieces, for each  $\alpha \in W(n_1)$ . Label the pieces of  $S_{\alpha_1}$  by  $S_{\alpha_1 \alpha_2}$ ,  $\alpha_2 \in W(n_2)$ . Continuing this process, we obtain a sequence of integers  $n_1, n_2, \dots$ , and a family of nonempty compact sets  $S_{\alpha_1 \alpha_2 \dots \alpha_k}$  of diameter at most  $2^{-k+1}$ . For each infinite binary string  $w = \alpha_1 \alpha_2 \dots$  where  $\alpha_i \in W(n_i)$ , define

$$S_w = \bigcap_{k=1}^{\infty} S_{\alpha_1 \alpha_2 \dots \alpha_k}.$$

Since  $S_{\alpha_1 \alpha_2 \dots \alpha_k}$  is a decreasing sequence of nonempty compact sets and the diameter approaches 0, by 2.3.1,  $S_w$  is a singleton. Define  $f : C \rightarrow M$  by  $f(c_w) =$  the unique element in  $S_w$ , and  $c_w$  is the element in  $C$  denoted by the string  $w$ , i.e.

$$c_w = \sum_{i=1}^{\infty} \frac{w_i}{3^i}.$$

It is easy to verify that  $f$  is sequentially continuous, hence proved.  $\square$

**Definition A.5.3** We say  $M$  is a Cantor space if  $M$  is nonempty, compact, perfect, and totally disconnected.

**Theorem A.5.4** (Moore-Kline Theorem) Every Cantor space is homeomorphic to the Cantor set.

*Proof.* Let  $M$  be a Cantor space. By the Cantor Surjection Theorem,  $\exists$  a continuous surjection  $f : C \rightarrow M$ . We can divide a closed set  $S \subseteq M$  into finitely many disjoint clopen sets: for  $x \in S$ , by totally disconnectedness,  $\exists$  a clopen set  $U$  such that  $x \in U_x \subseteq B_{\epsilon/2}(x)$ . Then  $\{U_x\}_{x \in S}$  is an open cover of  $S$ . Take a finite subcovering  $U_{x_1}, U_{x_2}, \dots, U_{x_n}$ , then

$$\left\{ U_{x_1}, U_{x_2} \setminus U_{x_1}, U_{x_3} \setminus (U_{x_1} \cup U_{x_2}), \dots, U_{x_n} \setminus \bigcup_{i=1}^{n-1} U_{x_i} \right\}$$

forms a disjoint clopen cover of  $S$ . Also, we can increase the number of nonempty compact sets by just one: let  $U$  be a clopen subset of  $M$ . Since  $M$  is perfect, so is  $U$ , thus containing at least two points. Pick two distinct points  $x, y \in U$ , then  $\exists$  a clopen set  $V$  such that  $x \in V \subseteq U$

---

and  $y \notin V$ . Then  $\{V, U \setminus V\}$  is a disjoint clopen cover of  $U$ . By repeating this process, we can divide  $M$  into  $2^k$  disjoint clopen sets for any  $k \in \mathbb{N}$ . Hence, we may remove the non-injective part of  $f$  step by step, and finally obtain a continuous bijection from a closed subset of  $C$  to  $M$ . Since  $C$  is compact and  $M$  is Hausdorff, this bijection is a homeomorphism.  $\square$

## A.6 Local Compactness and Topological Groups

Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a smooth function such that for all  $x \in \mathbb{R}$ , there exists  $n \in \mathbb{N}$  such that  $f^{(n)}(x) = 0$ . We want to show that  $f$  is a polynomial.

**Definition A.6.1** Let  $X$  be a topological space.  $X$  is locally compact if for every  $x \in X$ ,  $\exists$  a neighborhood  $U$  of  $x$  such that  $\overline{U}$  is compact.

**Theorem A.6.2** Let  $X$  be Hausdorff. The followings are equivalent:

1.  $X$  is locally compact.
2. For every  $x \in X$ ,  $\exists$  a neighborhood basis  $\mathcal{B}$  at  $x$  such that for every  $U \in \mathcal{B}$ ,  $\overline{U}$  is compact.
3. For all compact subsets  $K \subseteq X$  and open set  $U$  containing  $K$ ,  $\exists$  an open set  $V$  such that  $K \subseteq V \subseteq \overline{V} \subseteq U$  and  $\overline{V}$  is compact.
4. For every  $x \in X$  and every neighborhood  $U$  of  $x$ ,  $\exists$  a neighborhood  $V$  of  $x$  such that  $\overline{V}$  is compact and  $\overline{V} \subseteq U$ .

**Theorem A.6.3** (Baire Category Theorem) Let  $(M, d)$  be a complete metric space. Then the countable intersection of dense open sets is still dense.

*Proof.* Let  $(D_n)_{n \in \mathbb{N}}$  be a sequence of dense open sets in  $M$ . Let  $B_1$  be a open nonempty set in  $M$ . Since  $D_1$  is dense,  $B_1 \cap D_1$  is nonempty. Let  $x_1 \in B_1 \cap D_1$ . Then  $\exists r_1 \in (0, 1)$  such that  $\overline{B(x_1, r_1)} \subseteq B_1 \cap D_1$ . Since  $D_2$  is dense,  $B(x_1, r_1) \cap D_2$  is nonempty. Let  $x_2 \in B(x_1, r_1) \cap D_2$ . Then  $\exists r_2 \in (0, 1/2)$  such that  $\overline{B(x_2, r_2)} \subseteq B(x_1, r_1) \cap D_2$ . Continuing this process, we obtain a sequence of points  $(x_n)_{n \in \mathbb{N}}$  and a sequence of positive numbers  $(r_n)_{n \in \mathbb{N}}$  such that for each  $n \in \mathbb{N}$ ,

$$x_n \in D_n, \quad r_n < \frac{1}{n}, \quad \overline{B(x_n, r_n)} \subseteq B(x_{n-1}, r_{n-1}) \cap D_n.$$

Note that  $(x_n)_{n \in \mathbb{N}}$  is a Cauchy sequence, hence converges to some  $x \in M$ . Since  $\overline{B(x_n, r_n)}$  is closed and contains  $x_m$  for all  $m \geq n$ , it also contains  $x$ . Thus  $x \in D_n$  for all  $n \in \mathbb{N}$ . Also,  $x \in B_1$  since  $\overline{B(x_1, r_1)} \subseteq B_1$ . Therefore,  $x \in B_1 \cap \bigcap_{n=1}^{\infty} D_n$ , which shows that  $\bigcap_{n=1}^{\infty} D_n$  is dense.  $\square$

**Definition A.6.4** A topological space  $X$  is called a Baire space if the countable intersection of dense open sets is still dense.

---

**Theorem A.6.5** Every locally compact Hausdorff space is a Baire space.

*Proof.* Let  $(D_n)_{n \in \mathbb{N}}$  be a sequence of dense open sets in  $X$ . Let  $B_1$  be a nonempty open set in  $X$ . Since  $D_1$  is dense,  $B_1 \cap D_1$  is nonempty. Pick  $x_1 \in B_1 \cap D_1$ . Since  $X$  is locally compact,  $\exists$  a neighborhood  $U_1$  of  $x_1$  such that  $\overline{U_1}$  is compact and  $\overline{U_1} \subseteq B_1 \cap D_1$ . Since  $D_2$  is dense,  $U_1 \cap D_2$  is nonempty. Pick  $x_2 \in U_1 \cap D_2$ . Again,  $\exists$  a neighborhood  $U_2$  of  $x_2$  such that  $\overline{U_2}$  is compact and  $\overline{U_2} \subseteq U_1 \cap D_2$ . Continuing this process, we obtain a sequence of points  $(x_n)_{n \in \mathbb{N}}$  and a sequence of neighborhoods  $(U_n)_{n \in \mathbb{N}}$  such that for each  $n \in \mathbb{N}$ ,

$$x_n \in D_n, \quad \overline{U_n} \text{ is compact}, \quad \overline{U_n} \subseteq U_{n-1} \cap D_n.$$

Note that  $\{\overline{U_n}\}_{n \in \mathbb{N}}$  is a decreasing sequence of nonempty compact sets, hence by 2.3.1,  $\bigcap_{n=1}^{\infty} \overline{U_n} \neq \emptyset$ . Let  $x$  be an element in this intersection. Then  $x \in D_n$  for all  $n \in \mathbb{N}$ . Also,  $x \in B_1$  since  $\overline{U_1} \subseteq B_1$ . Therefore,  $x \in B_1 \cap \bigcap_{n=1}^{\infty} D_n$ , which shows that  $\bigcap_{n=1}^{\infty} D_n$  is dense.  $\square$

**Definition A.6.6** A space  $X$  is called  $\sigma$ -compact if  $X$  is a countable union of compact subsets.

**Definition A.6.7** A group  $(G, \cdot)$  is called a topological group if  $G$  is a topological space and the group operations (multiplication and inversion) are continuous.

**Proposition A.6.1** Let  $G$  be a topological group and  $a \in G$ , then the maps  $x \mapsto a \cdot x$ ,  $x \mapsto x \cdot a$ , and  $x \mapsto x^{-1}$  are homeomorphisms.

**Definition A.6.8** Let  $X$  be a topological space and  $x \in X$ . A neighborhood basis at  $x$  is a collection  $\mathcal{B}$  of neighborhoods of  $x$  such that for every neighborhood  $U$  of  $x$ ,  $\exists V \in \mathcal{B}$  such that  $V \subseteq U$ .

**Proposition A.6.2** Let  $G$  be a topological group and  $N$  be a neighborhood basis at the identity  $e$ . Then the collections

$$\{Ux \mid x \in G, U \in N\} \quad \text{and} \quad \{xU \mid x \in G, U \in N\}$$

are bases for the topology of  $G$ .

**Proposition A.6.3** Let  $W$  be an open set in  $G$  and  $x \in W$ . Then since  $e \in x^{-1}W$ , there exists  $U \in N$  such that  $U \subseteq x^{-1}W$ . Thus  $xU \subseteq W$ , which shows that  $\{xU \mid x \in G, U \in N\}$  is a basis. Similarly,  $\{Ux \mid x \in G, U \in N\}$  is also a basis.

**Proposition A.6.4** Let  $G$  be a topological group, then the collection of open sets forms an ideal, i.e. for any open set  $U$  and any subset  $A \subseteq G$ ,  $AU$  and  $UA$  are open.

**Theorem A.6.9** Let  $G, H$  be locally compact Hausdorff topological groups, and let  $f : G \rightarrow H$  be a continuous epimorphism, then  $f$  is an open map whenever  $G$  is  $\sigma$ -compact.

---

## A.7 Oct 14, 2025

**Theorem A.7.1** (Arzela's Theorem) Let  $(f_n)$  be a sequence of Riemann integrable functions on  $[a, b]$  that converges to a Riemann integrable function  $f$ . If exists  $M > 0$  such that for all  $n \in \mathbb{N}$  and  $x \in [a, b]$ ,  $|f_n(x)| \leq M$ , then

$$\lim_{n \rightarrow \infty} \int_a^b |f_n(x) - f(x)| dx = 0.$$

In particular,

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx.$$

**Lemma A.7.2** (Dini's Theorem) Let  $(f_n)$  be a monotone sequence of continuous functions on  $[a, b]$ . If  $(f_n)$  converges pointwise to a continuous function  $f$ , then  $(f_n)$  converges to  $f$  uniformly.

**Lemma A.7.3** For any bounded nonnegative  $f : [a, b] \rightarrow \mathbb{R}$  and any  $\epsilon > 0$ , there exists a continuous function  $g : [a, b] \rightarrow \mathbb{R}$  such that  $0 \leq g(x) \leq f(x)$  for all  $x \in [a, b]$  and

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx + \epsilon.$$

**Lemma A.7.4** Let  $(f_n)$  be a decreasing sequence of bounded functions on  $[a, b]$ . If  $(f_n)$  converges pointwise to the zero function, then

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = 0.$$

*Proof.* By Lemma A.7.3,  $\forall \epsilon > 0$ ,  $\exists$  a continuous function  $g_n : [a, b] \rightarrow \mathbb{R}$  such that  $0 \leq g_n(x) \leq f_n(x)$  for all  $x \in [a, b]$  and

$$\int_a^b f_n(x) dx \leq \int_a^b g_n(x) dx + \frac{\epsilon}{2^n}.$$

Define  $h_n = \min\{g_1, g_2, \dots, g_n\}$ . Then  $0 \leq h_n \leq g_n \leq f_n$  and  $(h_n)$  is a decreasing sequence of continuous functions that converges pointwise to the zero function. By Dini's Theorem,  $(h_n)$  converges to the zero function uniformly, hence

$$\lim_{n \rightarrow \infty} \int_a^b h_n(x) dx = 0.$$

We claim that  $\int_a^b f_n(x) dx \leq \int_a^b h_n(x) dx + \epsilon(1 - 2^{-n})$ , which concludes the proof. We prove this by induction on  $n$ . The case  $n = 1$  is trivial. Suppose the claim holds for some  $n \in \mathbb{N}$ .

Then

$$\begin{aligned}
\int_a^b f_{n+1}(x) dx &\leq \int_a^b g_{n+1}(x) dx + \frac{\epsilon}{2^{n+1}} \\
&\leq \int_a^b \max\{h_n(x), g_{n+1}(x)\} dx + \frac{\epsilon}{2^{n+1}} \\
&= \int_a^b h_n(x) dx + \int_a^b \max\{0, g_{n+1}(x) - h_n(x)\} dx + \frac{\epsilon}{2^{n+1}} \\
&\leq \int_a^b h_n(x) dx + \int_a^b g_{n+1}(x) dx + \frac{\epsilon}{2^{n+1}} \\
&\leq \int_a^b h_n(x) dx + \int_a^b g_n(x) dx + \frac{\epsilon}{2^n} + \frac{\epsilon}{2^{n+1}} \\
&\leq \int_a^b h_n(x) dx + \epsilon(1 - 2^{-n}) + \frac{\epsilon}{2^{n+1}} \\
&= \int_a^b h_n(x) dx + \epsilon(1 - 2^{-(n+1)}).
\end{aligned}$$

□

**Theorem A.7.5** Let  $(f_n)$  be a sequence of Riemann integrable functions defined on  $[a, b]$  which converges to a Riemann integrable function  $f$ . If there exists a Riemann integrable function  $g$  such that for all  $n \in \mathbb{N}$  and  $x \in [a, b]$ ,  $|f_n(x)| \leq g(x)$ , then

$$\lim_{n \rightarrow \infty} \int_a^b |f_n(x) - f(x)| dx = 0.$$

*Proof.* WLOG assume  $f_n \geq 0$  and  $f \equiv 0$  by considering the positive parts and negative parts separately. Set  $g_{n,m} \equiv \max\{f_n, \dots, f_m\}$  for  $m \geq n$ . Then  $0 \leq g_{n,m} \leq g$ . Note that the sequence  $\int_a^b g_{n,m}(x) dx$  is increasing and bounded above by  $\int_a^b g(x) dx$ . Therefore, for  $\epsilon > 0$ ,  $\exists m_n > n$  such that  $m_{n+1} > m_n$  and

$$0 \leq \int_a^b (g_{n,k}(x) - g_{n,m_n}(x)) dx \leq \frac{\epsilon}{2^n}, \quad \forall k > m_n.$$

Define  $u_n = g_{n,m_n}$ . Then  $(u_n)$  is a decreasing sequence of bounded functions that converges pointwise to the zero function. We get

$$0 \leq f_n \leq u_n \leq \min\{u_1, \dots, u_n\} + \sum_{i=1}^{n-1} (\max\{u_i, \dots, u_n\} - u_i).$$

Since

$$\max\{u_i, \dots, u_n\} - u_i = \max\{f_i, \dots, f_{m_n}\} - \max\{f_i, \dots, f_{m_i}\} \leq g_{i,m_n} - g_{i,m_i},$$

---

we have

$$\begin{aligned} & \int_a^b (\max \{u_i, \dots, u_n\} - u_i) dx \leq \frac{\epsilon}{2^i}, \quad \forall i < n \\ \implies & 0 \leq \int_a^b f_n(x) dx \leq \int_a^b \min \{u_1(x), \dots, u_n(x)\} dx + \epsilon(1 - 2^{-(n-1)}). \end{aligned}$$

□

**Theorem A.7.6** (Fatou's Lemma) Let  $(f_n)$  be a sequence of nonnegative Riemann integrable functions on  $[a, b]$  that converges pointwise to a Riemann integrable function  $f$ . Then

$$\int_a^b f(x) dx \leq \liminf_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

*Proof.* Since  $f - f_n \leq f$ , the positive part is also bounded by  $f$ . By the previous theorem, we have

$$\int_a^b (f(x) - f_n(x))_+ dx \rightarrow 0.$$

Note that

$$f_n = (f_n - f)_+ + f - (f - f_n)_+.$$

Thus

$$\int_a^b f_n(x) dx = \int_a^b (f_n(x) - f(x))_+ dx + \int_a^b f(x) dx - \int_a^b (f(x) - f_n(x))_+ dx.$$

Taking  $\liminf$  on both sides, we get the desired result. □

## A.8 Applications of Arzela-Ascoli Theorem

Recall that a set of functions  $\mathcal{F} \subseteq C(X, \mathbb{R})$  is compact iff  $\mathcal{F}$  is closed, uniformly bounded, and equicontinuous. Every sequence in  $\mathcal{F}$  has a uniformly convergent subsequence iff  $\mathcal{F}$  is uniformly bounded and equicontinuous.

**Definition A.8.1** Let  $U \subseteq \mathbb{C}$  be open. A family of holomorphic functions  $\mathcal{F} \subseteq H(U)$  is normal if every sequence in  $\mathcal{F}$  has a subsequence that converges uniformly on compact subsets of  $U$  to a holomorphic function.

**Theorem A.8.2** (Montel's Theorem) Let  $D \subseteq \mathbb{C}$  be open. If  $\mathcal{F} \subseteq H(D)$  is locally uniformly bounded, i.e.  $\mathcal{F}|_K := \{f|_K \mid f \in \mathcal{F}\}$  is uniformly bounded for every compact subset  $K \subseteq D$ , then  $\mathcal{F}$  is normal.

*Proof.* Given  $K \subseteq U$  compact, we first show that  $\mathcal{F}|_K$  is equicontinuous.

---

Since  $K$  is compact and  $U$  is open,  $r_0 = d(K, \partial U) > 0$ , then for any  $r < r_0$ ,

$$\begin{aligned} |f'(z_0)| &= \left| \frac{1}{2\pi i} \int_{|z-z_0|=r} \frac{f(z)}{(z-z_0)^2} dz \right| \\ &\leq \frac{1}{2\pi} \cdot \frac{2\pi r M}{r^2} = \frac{M}{r}, \end{aligned}$$

where  $M$  is a uniform bound of  $\mathcal{F}|_K$ . Therefore,  $\mathcal{F}|_K$  is uniformly Lipschitz continuous with Lipschitz constant  $M/r$  and hence equicontinuous.

By Arzela-Ascoli Theorem, every sequence in  $\mathcal{F}|_K$  has a uniformly convergent subsequence. Write  $U = \bigcup_{n=1}^{\infty} K_n$  where  $K_n$ 's are compact and  $K_n \subseteq \text{int } K_{n+1}$ , where

$$K_n := \left\{ x \in U \mid d(X, \partial U) \leq \frac{1}{n}, |x| \leq n \right\}.$$

Given a sequence  $(f_n)$  in  $\mathcal{F}$  and  $K_i$ , by the above argument,  $\exists$  a subsequence that converges uniformly on  $K_i$ . By a diagonalization argument, we can find a subsequence that converges uniformly on each  $K_i$ , hence converges uniformly on compact subsets of  $U$ .  $\square$

**Definition A.8.3** A topological vector space is a vector space  $V$  over  $\mathbb{R}$  or  $\mathbb{C}$  equipped with a topology such that the vector addition  $V \times V \rightarrow V$ ,  $(x, y) \mapsto x + y$  and scalar multiplication  $\mathbb{F} \times V \rightarrow V$ ,  $(a, x) \mapsto ax$  are continuous.

**Definition A.8.4** A locally convex topological vector space is a topological vector space  $V$  such that the origin has a convex and balanced neighborhood basis. A set  $S \subseteq V$  is balanced if for every  $c \in \mathbb{F}$  with  $|c| \leq 1$ ,  $cS \subseteq S$ .

**Theorem A.8.5** (Shauder-Tychonoff Theorem) Every compact and convex subset of a locally convex Hausdorff topological vector space has the fixed point property, i.e. every continuous map from the set to itself has a fixed point.

An IVP for ODEs is a problem of the form

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(0) = 0 \end{cases}, \quad x \in \mathbb{R}^n, f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

**Theorem A.8.6** Let  $f(t, u)$  be defined on  $0 \leq t \leq a$  and  $u \in \mathbb{R}^n$  such that  $\|u\| \leq r$ . Assume that  $f$  is continuous and  $\|f(t, u)\| \leq r/a$ , then the IVP has a solution in  $C[0, a]$ .

*Proof.* Define  $(Ax)(t) = \int_0^t f(s, x(s)) ds$  and

$$D = \text{Dom}(A) := \left\{ x \in C^u[a, b] \mid \|x\|_{\infty} := \sup_{0 \leq t \leq a} \|x(t)\| \leq r \right\}.$$

□

## A.9 Nowhere Analytic Smooth Functions

Take an open set  $\mathcal{O} \subseteq \mathbb{R}$ .

**Definition A.9.1** (Fréchet Metric) Choose an exhaustion  $\mathcal{O} = \bigcup_{j \geq 0} K_j$ . For any  $f, g \in C^\infty(\mathcal{O})$ , define

$$d(f, g) = \sum_{j, m \geq 0} \frac{1}{2^{j+m+2}} \cdot \frac{\rho_{j,m}(f - g)}{1 + \rho_{j,m}(f - g)},$$

where

$$\rho_{j,m}(h) = \sup_{x \in K_j} |h^{(m)}(x)|.$$

## A.10 Riemann-Lebesgue Lemma

**Proposition A.10.1** If  $f$  is Riemann integrable on  $[a, b]$ , then

$$\lim_{n \rightarrow \infty} \int_a^b f(x) \sin(nx) dx = 0, \quad \lim_{n \rightarrow \infty} \int_a^b f(x) \cos(nx) dx = 0.$$

*Proof.* First, assume  $f$  is a step function. Then the result follows from direct computation. For general Riemann integrable  $f$ , given  $\epsilon > 0$ ,  $\exists$  a step function  $g$  such that  $\int_a^b |f(x) - g(x)| dx < \epsilon$ . Then

$$\begin{aligned} \left| \int_a^b f(x) \sin(nx) dx \right| &\leq \left| \int_a^b (f(x) - g(x)) \sin(nx) dx \right| + \left| \int_a^b g(x) \sin(nx) dx \right| \\ &\leq \int_a^b |f(x) - g(x)| dx + \left| \int_a^b g(x) \sin(nx) dx \right| \\ &< \epsilon + \left| \int_a^b g(x) \sin(nx) dx \right|. \end{aligned}$$

Since  $\epsilon$  is arbitrary and the second term tends to 0 as  $n \rightarrow \infty$ , we are done. □

**Corollary A.10.1**  $|\widehat{f}(n)| \rightarrow 0$  as  $n \rightarrow \infty$  for any Riemann integrable function  $f$  on  $[-\pi, \pi]$ , where

$$\widehat{f}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx.$$

**Proposition A.10.2** If  $f$  is  $L^1$  on  $\mathbb{R}$ , then

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(t) e^{-i\xi t} dt = 0.$$

---

*Proof.* Suppose that  $f$  is continuous and compact supported. Note that

$$\int_{-\infty}^{\infty} f(t)e^{-i\xi t} dt = \int_{-\infty}^{\infty} f\left(t + \frac{\pi}{\xi}\right) e^{-i\xi t} e^{-i\pi} dt = - \int_{-\infty}^{\infty} f\left(t + \frac{\pi}{\xi}\right) e^{-i\xi t} dt.$$

Thus,

$$\begin{aligned} \left| \int_{-\infty}^{\infty} f(t)e^{-i\xi t} dt \right| &= \frac{1}{2} \left| \int_{-\infty}^{\infty} \left( f(t) - f\left(t + \frac{\pi}{\xi}\right) \right) e^{-i\xi t} dt \right| \\ &\leq \frac{1}{2} \int_{-\infty}^{\infty} \left| f(t) - f\left(t + \frac{\pi}{\xi}\right) \right| dt. \end{aligned}$$

Since  $f$  is uniformly continuous, the right hand side tends to 0 as  $\xi \rightarrow \infty$ . For general  $f \in L^1(\mathbb{R})$ , given  $\epsilon > 0$ ,  $\exists$  a continuous compact supported function  $g$  such that  $\int_{-\infty}^{\infty} |f(t) - g(t)| dt < \epsilon$ . Then

$$\begin{aligned} \left| \int_{-\infty}^{\infty} f(t)e^{-i\xi t} dt \right| &\leq \left| \int_{-\infty}^{\infty} (f(t) - g(t))e^{-i\xi t} dt \right| + \left| \int_{-\infty}^{\infty} g(t)e^{-i\xi t} dt \right| \\ &\leq \int_{-\infty}^{\infty} |f(t) - g(t)| dt + \left| \int_{-\infty}^{\infty} g(t)e^{-i\xi t} dt \right| \\ &< \epsilon + \left| \int_{-\infty}^{\infty} g(t)e^{-i\xi t} dt \right|. \end{aligned}$$

Since  $\epsilon$  is arbitrary and the second term tends to 0 as  $\xi \rightarrow \infty$ , we are done.  $\square$

**Definition A.10.2** A Banach space is a complete normed space. For Banach spaces  $B_1$  and  $B_2$ , the norm of a (bounded) linear operator  $T: B_1 \rightarrow B_2$  is defined by

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_{B_2}}{\|x\|_{B_1}} = \sup_{\|x\|_{B_1}=1} \|Tx\|_{B_2} = \sup_{\|x\|_{B_1} \leq 1} \|Tx\|_{B_2}.$$

An operator  $T: B_1 \rightarrow B_2$  is called bounded if  $\|T\| < \infty$ .

**Lemma A.10.3** A linear operator  $T: B_1 \rightarrow B_2$  between Banach spaces is bounded iff it is continuous.

*Proof.* If  $T$  is bounded, then for any  $\epsilon > 0$ , let  $\delta = \epsilon/\|T\|$ . Then for any  $x, y \in B_1$  with  $\|x - y\|_{B_1} < \delta$ , we have

$$\|Tx - Ty\|_{B_2} = \|T(x - y)\|_{B_2} \leq \|T\| \cdot \|x - y\|_{B_1} < \|T\| \cdot \delta = \epsilon,$$

which shows that  $T$  is continuous.

Conversely, if  $T$  is continuous, then it is continuous at 0. Thus,  $\exists \delta > 0$  such that for all  $x \in B_1$  with  $\|x\|_{B_1} < \delta$ ,  $\|Tx\|_{B_2} < 1$ . For any  $x \in B_1$  with  $\|x\|_{B_1} \neq 0$ , let  $y = (\delta/(2\|x\|_{B_1}))x$ ,

---

then  $\|y\|_{B_1} < \delta$ , hence  $\|Ty\|_{B_2} < 1$ . Therefore,

$$\|Tx\|_{B_2} = \frac{2\|x\|_{B_1}}{\delta} \|Ty\|_{B_2} < \frac{2\|x\|_{B_1}}{\delta}.$$

This shows that  $T$  is bounded with  $\|T\| \leq 2/\delta$ .  $\square$

**Theorem A.10.4** Given a family  $\mathcal{A}$  of bounded linear operators from a Banach space  $B_1$  to a Banach space  $B_2$ , if  $\sup_{T \in \mathcal{A}} \|T\| = \infty$ , then there exists  $x \in B_1$  such that  $\sup_{T \in \mathcal{A}} \|Tx\|_{B_2} = \infty$ .

*Proof.* Suppose not, then for every  $x \in B_1$ ,  $\sup_{T \in \mathcal{A}} \|Tx\|_{B_2} < \infty$ . For each  $n \in \mathbb{N}$ , define

$$F_n = \left\{ x \in B_1 \mid \sup_{T \in \mathcal{A}} \|Tx\|_{B_2} \leq n \right\}.$$

Note that  $B_1 = \bigcup_{n=1}^{\infty} F_n$ . For a sequence  $(x_k)$  in  $F_n$  that converges to some  $x \in B_1$ , we also have  $x \in F_n$ , hence  $F_n$  is closed. By Baire Category Theorem,  $\exists n_0 \in \mathbb{N}$  such that  $\text{int}(F_{n_0}) \neq \emptyset$ . Thus,  $\exists x_0 \in B_1$  and  $r > 0$  such that  $B(x_0, r) \subseteq F_{n_0}$ . For any  $x \in B_1$  with  $\|x\|_{B_1} = 1$ , we have  $x_0 + rx \in F_{n_0}$ , hence for all  $T \in \mathcal{A}$ ,

$$\|Tx\|_{B_2} = \frac{1}{r} \|T(x_0 + rx) - Tx_0\|_{B_2} \leq \frac{1}{r} \left( \sup_{T \in \mathcal{A}} \|T(x_0 + rx)\|_{B_2} + \sup_{T \in \mathcal{A}} \|Tx_0\|_{B_2} \right) \leq \frac{2n_0}{r}.$$

This shows that  $\sup_{T \in \mathcal{A}} \|T\| < \infty$ , a contradiction.  $\square$