Analysis (I)

by Poshen Chang

v.1

2025-10-21

Contents

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 $d : M \times M \to \mathbb{R}$. $d(x, y)$ 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 $M \neq \varnothing$, 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 $\varnothing \neq A \subseteq M$, 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 $f : M \to N$ 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 $f : M \to N$, $g : N \to P$ be continuous functions.

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

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

thus $g \circ f$ is continuous. $\qquad\square$

Example 1.2.1 The following are some examples of continuous functions:

- The identity map id $: M \to M$ is continuous.

- Every constant function $f : M \to N$ is continuous.

- Every function $f : M \to N$ is continuous if M is equipped with the discrete metric.

Definition 1.2.2 If $f : M \to N$ is a bijection such that f and $f^{-1} : N \to M$ 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 $f : [0, 2\pi) \to \mathbb{S}^1$ 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 $f : M \to N$ is continuous iff it satisfies the following: $\forall \epsilon$ 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. " $\implies$ ": Suppose that f fails to satisfy the ϵ-δ 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.

" $\impliedby$ ": Suppose that f satisfies the ϵ-δ condition at x. Let (x_n) be a sequence in M susch that (x_n) converges to x. Let $\epsilon > 0$. $\exists \delta > 0$ such that $d_M(x, y) < \delta \implies d_N(f(x), f(y)) < \epsilon$. 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. " $\implies$ ": 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, $x \in S$.

" $\impliedby$ ": 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 q 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 $\varnothing$ and M.

Proof. Clearly, $\varnothing$ 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 α. 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 $1 \le i \le n$, $\exists r_i > 0$ such that $d(x, y) < r_i \implies y \in \mathcal{U}_i$. Take $r = \min r_1, \ldots, r_n$, then for any y satisfying $d(x, y) < r$, we have $y \in \mathcal{U}_i$ for all $1 \le i \le n$, 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) $\varnothing, X \in \mathcal{T}$.

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 $\mathcal{T} = \{\varnothing, X\}$. Then $(X, \mathcal{T})$ is a topological space, which is know 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 $\varnothing, X$ 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 $S = \varnothing$, then $\overline{S} = \varnothing$, which is closed.

Suppose that $S \neq \varnothing$, and let $(x_n) \to x$ 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 $k \to \infty$. 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 $n \to \infty$, 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. $\qquad\square$

Corollary 1.3.5 $\overline{S}$ is the smallest closed set that conatins 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 $\overline{S}$. $\qquad\square$

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

Definition 1.3.6 Lex 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 Lex 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 $f : X \to Y$ continuous if for any open set $V \subseteq Y$, the preimage $f^{-1}(Y)$ is open in X.

Proposition 1.3.4 A function $f : X \to Y$ 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. " $\implies$ ": Let $\mathcal{V} \subseteq Y$ is open. We need to show $f^{-1}(Y)$ 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 $f(\mathcal{U}_x) \subseteq \mathcal{V} \rightsquigarrow \mathcal{U}_x \subseteq f^{-1}(\mathcal{V})$. Take the union of all such $\mathcal{U}_x$ over $x \in f^{-1}(\mathcal{V})$, then

$$\bigcup_{x \in f^{-1}(\mathcal{V})} \mathcal{U}_x \subseteq f^{-1}(\mathcal{V}) \quad \text{and} \quad \forall x \in f^{-1}(\mathcal{V}), \, x \in \mathcal{U}_x \subseteq f^{-1}(\mathcal{V})$$

$$\implies f^{-1}(\mathcal{V}) = \bigcup_{x \in f^{-1}(\mathcal{V})} \mathcal{U}_x$$

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

" $\impliedby$ ": 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 $f(f^{-1}(\mathcal{V})) \subseteq \mathcal{V}$. $\qquad\square$

Definition 1.3.10 A homeomorphism is a continuous bijection between topoloical spaces.

Corollary 1.3.11 A homeomorphism $f : X \to Y$ 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 paints $x_1, x_2 \in X$, $\exists$ neighborhoods $\mathcal{U}_1$ of x_1 and $\mathcal{U}_2$ of x_2 such that $\mathcal{U}_1 \cap \mathcal{U}_2 = \varnothing$.

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 $x_n \in \mathcal{U} \cap \mathcal{U}' = \varnothing$, contradiction. Therefore, any converging sequence has a unique limit.

$\square$

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 to prove that $\mathcal{U}$ belongs to the metric toplogy 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. $\square$

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

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, $\varnothing = \bigcup_{B \in \varnothing} B \in \mathcal{T}$ 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 $X_1, \ldots, X_n$ be topological spaces. Define a basis in $X_1 \times \cdots \times X_n$ by

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

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

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

$$f : X_1 \times \cdots \times X_n \to Y_1 \times \cdots \times Y_n, \quad f(x_1, \ldots, x_n) = (f_1(x_1), \ldots, f_n(x_n))$$

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 ∞-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_∞ are the same.

- The metric topology coincides with the product topology.

Proposition 1.5.5 Let M be a metric space, then the metric $d : M \times M \to \mathbb{R}$ 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

$$|d(x_1, y_1) - d(x_2, y_2)| \leq |d(x_1, y_1) - d(x_1, y_2)| + |d(x_1, y_2) - d(x_2, y_2)|$$
$$\leq d(y_1, y_2) + d(x_1, x_2).$$

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. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

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

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

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

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

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

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

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,

$$|d(x_n, y_n) - d(x_m, y_m)| \le |d(x_n, y_n) - d(x_m, y_n)| + |d(x_m, y_n) - d(x_m, y_m)|$$
$$\le d(x_n, x_m) + d(y_n, y_m),$$

and since $(x_n), (y_n)$ are Cauchy sequences, $\forall \epsilon > 0, \exists N \in \mathbb{N}$ such that $m, n \ge N$ 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

$$L = \lim_{n \to \infty} d(x_n, y_n), \quad L' = \lim_{n \to \infty} d(x_n', y_n').$$

By triangle inequality again, we have

$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq |d(x_n, y_n) - d(x'_n, y_n)| + |d(x'_n, y_n) - d(x'_n, y'_n)|$$
$$\leq d(x_n, x'_n) + d(y_n, y'_n).$$

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 $d(x_n, y_n) \to 0$, 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,

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

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

For each $x \in M$, define

$$\overline{x} = (x, x, x, \ldots) \in \mathcal{C},$$

then $\overline{x}$ is a Cauchy sequence, and $D(\overline{x}, \overline{y}) = d(x, y)$ for any $x, y \in M$. It is clear that the mapping $x \mapsto \overline{x}$ 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 $X_n \to Y$.

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

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

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 $X_n \to Y$. 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

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

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

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 ϵ'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 $[(y_0, y_0, y_0, \ldots)]$ is an upper bound for $\mathcal{A}$.

2. For $n \geq 1$, y_n is the smallest rational with denominator 2^n such that $[(y_n, y_n, y_n, \ldots)]$ 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

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

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

$$z_m + \frac{\epsilon}{2} < y_k - \frac{\epsilon}{2} < y_k - \frac{1}{2^k}$$

$$\implies Z < [(y_k - \frac{1}{2^k}, y_k - \frac{1}{2^k}, y_k - \frac{1}{2^k}, \ldots)],$$

but y_k is the smallest fraction with denominator 2^k such that $[(y_k, y_k, y_k, \ldots)]$ 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 $\varnothing$ and X.

Proposition 1.7.2 Let X, Y be topological spaces and let $f : X \to Y$ 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 $U \cap V \cap f(X) = \varnothing$ and $U \cup V \supset f(X)$. 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 $f : X \to \mathbb{R}$ 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. $\qquad\square$

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

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

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

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 $V \cap \overline{A} = \varnothing$, contradicting the fact that $\mathcal{U}, \mathcal{V}$ is a separation of $\overline{A}$. Therefore, $\overline{A}$ is connected. $\qquad\square$

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_α 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_α 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. $\qquad\square$

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 $f :$ $[0,1] \to X$ 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 $f : [0,1] \to X$ 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

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

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. "$\implies$": 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, $x_n \to x$.

"$\impliedby$": Let (x_n) be a sequence of distinct points in S such that $x_n \to x$. 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, $x \in S \cup S'$. $\qquad\square$

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 $M = \{x_1, x_2, \ldots\}$.

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 $Y_1 \supset Y_2 \supset \cdots$ such that $x_1, x_2, \ldots, x_n \notin Y_n$ and $r_n < 1/n$. Since the centers y_n form a Cauchy sequence, by completeness, $y_n \to y$ 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 $y \notin \{x_1, x_2, \ldots\}$, contradicting the fact that $M = \{x_1, x_2, \ldots\}$. Therefore, M is uncountable. $\qquad\square$

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. $\qquad\square$

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

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

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

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 $\mathcal{U} = \left\{ (\frac{1}{n}, 2) \mid n \in \mathbb{N} \right\}$, 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 λ 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. $\qquad\square$

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. " $\implies$ ": 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

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

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.

" $\Longleftarrow$ ": Let $\mathcal{U}$ be an open cover of K. By the Lebesgue number lemma, $\mathcal{U}$ has a positive Lebesgue number λ. 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 $x_3, x_4, \ldots$. 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 $x_{n_m} \to x$. 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 ϵ.

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 $f : X \to Y$ 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, $\exists U_1, \ldots, U_n \in \mathcal{U}$ such that

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

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 $f : X \to \mathbb{R}$ 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 $f : M \to N$ 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 $y_{n_k} \to x$ as well. Since f is continuous, $f(x_{n_k}) \to f(x)$ and $f(y_{n_k}) \to f(x)$, 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

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

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 $Z_a = V_1 \cap \cdots \cap V_m$, then $Z_a \times B \subseteq U_1 \cup \cdots \cup U_n$. 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 $A \subseteq Z_{a_1} \cup \cdots \cup Z_{a_k}$. Then

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

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. $\qquad\square$

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. $\qquad\square$

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 $b \in \widetilde{V}_b$. $\left\{ \widetilde{V}_b \mid b \in B \right\}$ 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 $B \subseteq \widetilde{V}$. $\quad\square$

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. $\quad\square$

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 $\iff$ $\bigcap_{C \in \mathcal{C}} C = \varnothing$.

3. The finite subcollection $\{U_1, \ldots, U_n\}$ of $\mathcal{U}$ covers X $\iff$ $\bigcap_{i=1}^{n}(X \setminus U_i) = \varnothing$.

Now, if X is compact, we have the contrapositive statment 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,

$$\lim_{n \to \infty} \operatorname{diam}(F_n) = 0,$$

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 $\operatorname{diam}(F) = 0$. 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 $\mathcal{C}_1, \ldots, \mathcal{C}_n$ of $\mathcal{C}$, $\exists \mathcal{B}_1, \ldots, \mathcal{B}_n \in \mathbb{B}$ 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 $\mathcal{C}_1, \ldots, \mathcal{C}_n \in \mathcal{B}_k$. Since $\mathcal{B}_k$ has the FIP, $\bigcap_{i=1}^n \mathcal{C}_i \neq \varnothing$. 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. $\qquad\square$

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, $B \neq \varnothing$. 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, $A \in \mathcal{D}$. $\qquad\square$

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_α is compact.

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

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

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

$\bigcap_{D \in \mathcal{D}} \overline{D} \neq \varnothing$. Given $\alpha \in I$, let $\pi_\alpha : X \to X_\alpha$ 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_α is compact, by Theorem 2.3.2, we have

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

Pick an element x_α 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_β be an open neighborhood of x_β in X_β. 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,

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

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. $\square$

CHAPTER 3

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 $C_0 \supseteq C_1 \supseteq C_2 \supseteq \cdots$. 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 3.0.1   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 3.0.1   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 3.0.2   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 ϵ are arbitrary, C is totally disconnected. $\qquad \square$

Corollary 3.0.3 Cantor set C is uncountable.

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

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 $f : C \to M$. 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 $f : C \to \{0, 2\}^{\mathbb{N}}$ 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 3.0.4 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 3.0.1 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 $\mathbb{R}$. $\qquad\square$

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

Definition 3.0.5 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 3.0.2 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 $Z_1, Z_2, \ldots$ 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 $\left\{(a_k^{(n)}, b_k^{(n)})\right\}_{k \in \mathbb{N}}$ such that

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

Then $\left\{(a_k^{(n)}, b_k^{(n)}) \mid n, k \in \mathbb{N}\right\}$ 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. $\qquad\square$

- 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 4

Integration Theory

4.1 Riemann Integrals on $\mathbb{R}$

Let $f : [a, b] \to \mathbb{R}$ 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 $P = \{x_0, x_1, \ldots, x_n\}$ and $T = \{t_1, t_2, \ldots, t_n\}$ are interlaced as

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

Assume that the points $x_0, x_1, \ldots, x_n$ 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 mesh(P) $< \delta$, we have

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

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

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

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 $f : [a, b] \to [-M, M]$ be a bounded function. For a partition $P = \{x_0, x_1, \ldots, x_n\}$ 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

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

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

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

Definition 4.1.1 A partition P' refines another partition P if $P \subseteq P'$.

Lemma 4.1.2 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. $\qquad \square$

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

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

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

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

$$< \sum_{i=1}^{n} \frac{\epsilon}{4(b-a)}(x_i - x_{i-1}) = \frac{\epsilon}{4(b-a)}(b - a) = \frac{\epsilon}{4}.$$

Similarly, choose $T' = \{t'_1, t'_2, \ldots, t'_n\}$ 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 $f : [a, b] \to [-M, M]$ 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
$$\delta = \frac{\epsilon}{12 n_1 M},$$

where n_1 is the number of points in P_1. Choose any partition P, T with 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

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

then

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

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

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

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,

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

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. $\qquad\square$

Example 4.1.1 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 mesh(P) $< \delta$, we have

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

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. $\qquad\square$

Proposition 4.1.1 A function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if for all $\epsilon > 0$, $\exists$ functions $g, h : [a, b] \to \mathbb{R}$ such that $g(x) \le f(x) \le h(x)$ 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 $g(x) \le f(x) \le h(x)$ for all $x \in [a, b]$ and

$$\int_{a}^{b} (h(x) - g(x))\, dx < \epsilon/3.$$

For all partition P, we have

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

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

$$U(f, P) - L(f, P) < U(h, P) - L(g, P)$$
$$< \int_{a}^{b} h(x)\, dx - \int_{a}^{b} g(x)\, dx + 2\epsilon/3$$
$$= \int_{a}^{b} (h(x) - g(x))\, dx + 2\epsilon/3 < \epsilon.$$

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

Definition 4.1.4 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 4.1.2 (Thomae's Function)   Define $[0,1] \to \mathbb{Q}$ 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_ϵ is Riemann integrable with integral 0 since S_ϵ 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 4.1.5 (Riemann's Integrability Criterion) $f : [a,b] \to \mathbb{R}$ is Riemann integrable iff f is bounded and continuous almost everywhere.

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

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

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

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

" $\implies$ ": 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

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

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 $M_i - m_i \geq \mathrm{osc}_x(\mathrm{f}) \geq 1/k$. 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.

" $\Longleftarrow$ ": 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 $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a, b]$ with $\mathrm{mesh}(\mathcal{P}) < \lambda$. 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

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

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. □

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 4.1.6 Every bounded piecewise continuous function is Riemann integrable.

Proof. The set of discontinuities of a piecewise continuous function is finite, hence has measure zero. □

Corollary 4.1.7 The characteristic function of a set $S \subseteq [a, b]$ is Riemann integrable iff the boundary ∂S has measure zero.

Proof. Note that χ_S is continuous at x iff $x \notin \partial S$. Therefore, the set of discontinuities of χ_S is exactly ∂S. □

Corollary 4.1.8 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. □

Corollary 4.1.9 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. □

Corollary 4.1.10 If $f : [a, b] \to [c, d]$ is Riemann integrable and $\phi : [c, d] \to \mathbb{R}$ 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, ϕ is bounded, thus $\phi \circ f$ is bounded as well. □

Corollary 4.1.11 If f is Riemann integrable then so is $|f|$.

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

Corollary 4.1.12 If $a < c < b$ and $f : [a, b] \to \mathbb{R}$ 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. $\qquad\square$

Corollary 4.1.13 A Riemann integrable function $f : [a, b] \to [0, M]$ has integral 0 iff $f(x) = 0$ almost everywhere.

Proof. "$\implies$": 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.

"$\impliedby$": 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. $\qquad\square$

Corollary 4.1.14 If $f : [a, b] \to \mathbb{R}$ is Riemann integrable and if $\psi : [c, d] \to [a, b]$ is a homeomorphism such that ψ^{-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 ψ^{-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, $\left\{ \left(\psi^{-1}(a_k), \psi^{-1}(b_k) \right) \right\}_{k \in \mathbb{N}}$ is a countable cover of $\psi^{-1}(\mathcal{D})$ by open intervals of total length at most $C\epsilon$. Since ϵ is arbitrary, $\psi^{-1}(\mathcal{D})$ has measure zero. Hence $f \circ \psi$ is Riemann integrable. $\qquad \square$

Remark If f is Riemann integrable and ψ 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.

4.2 Fundamental Theorem of Calculus

Let $f : [a, b] \to \mathbb{R}$ 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 $F : [a, b] \to \mathbb{R}$ such that $F'(x) = f(x)$ for all $x \in [a, b]$.

Corollary 4.2.1 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 4.2.2 There exists a continuous $H : [0, 1] \to \mathbb{R}$ 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,

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

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. $\quad\square$

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

4.3   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 $P = \{x_0, x_1, \ldots, x_m\}$ of $[a,b]$ and $Q = \{y_0, y_1, \ldots, y_n\}$ of $[c,d]$. Let G be the grid partition G consisting of the rectangles

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

Let S be a set of sample points $s_{ij} \in R_{ij}$ for all $1 \le i \le m$ and $1 \le j \le n$. The Riemann sum of a bounded function $f : R \to \mathbb{R}$ 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

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

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:

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

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 4.3.1 • 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 $f(x,y) \le g(x,y)$ for all $(x,y) \in R$, then

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

- If f is Riemann integrable on R, then

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

Definition 4.3.1 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 ϵ:

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

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

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

Denote the set of discontinuities of f by

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

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

Theorem 4.3.2 (Riemann Integrability Criterion in Higher Dimensions) A bounded function

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

Suppose that $f : R \to \mathbb{R}$ 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

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

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 4.3.3 (Fubini's Theorem) If $f : R \to \mathbb{R}$ 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 $P = \{x_0, x_1, \ldots, x_m\}$ is a partition of $[a, b]$ and $Q = \{y_0, y_1, \ldots, y_n\}$ 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 $y \in J_j$,

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

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

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

$\square$

Corollary 4.3.4 If $f : R \to \mathbb{R}$ is Riemann integrable, then the order of integration does not affect the value of the integral.

Corollary 4.3.5 (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 ∂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. $\square$

Definition 4.3.6 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

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

4.4   Change of Variables

Proposition 4.4.1 If $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation, then for every Riemann measurable set $S \subseteq \mathbb{R}^2$, $T(S)$ is Riemann measurable and

$$|T(S)| = |\det(\mathrm{T})| \cdot |S|.$$

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

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

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

This is always possible since χ_S is Riemann integrable and
$$\sum_{R \subseteq S} |R| = L(\chi_S, G), \qquad \sum_{R : R \cap S \neq \varnothing} |R| = U(\chi_S, G).$$

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

Similarly,
$$|\det(E)|(|S| + \epsilon) \geq |\det(E)| \sum_{R : R \cap S \neq \varnothing} |R|$$
$$= \sum_{R : R \cap S \neq \varnothing} |T(R)| = \sum_{R : R \cap S \neq \varnothing} \int \chi_{T(R)} \, dA \geq \overline{\int} \chi_{T(S)}.$$

Since ϵ is arbitrary, we have
$$\underline{\int} \chi_{T(S)} = |\det(E)||S| = \overline{\int} \chi_{T(S)},$$

hence proved. $\qquad \square$

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

Definition 4.4.1 A norm on a vector space V (over $\mathbb{R}$) is a function $\| \cdot \| : V \to [0, \infty)$ 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 $T : V \to W$ 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 $S(x, r)$ under this norm. Write $S(r) := S(0, r)$.

Lemma 4.4.2 Let $\mathcal{U}$ be open with $0 \in \mathcal{U}$, and let $\sigma \in (0, 1)$. Suppose that $\psi : \mathcal{U} \to \mathbb{R}^2$ 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 ρ 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 $K_v : S(r) \to \mathbb{R}^2$ 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 4.4.3 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 $h : \mathbb{R}^2 \to \mathbb{R}^2$ 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
$$\operatorname{diam}(h(S_k \cap Z)) \leq L \operatorname{diam}(S_k \cap Z) \leq L \operatorname{diam}(S_k)$$

w.r.t. the sup-norm. Therefore, $h(S_k \cap Z)$ is contained in an open square S_k' of side length $L \operatorname{diam}(S_k)$, 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 ϵ is arbitrary, $h(Z)$ has measure zero. $\qquad\square$

Theorem 4.4.4 Let f be Riemann integrable, and let $\varphi : R \to \mathbb{R}^2$ 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 φ is a C^1-diffeomorphism, φ^{-1} is Lipschitz, thus by Lemma 4.4.3, $\mathcal{D}$ 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 φ is a C^1-diffeomorphism, in particular it is a homeomorphism, hence $\varphi(\partial R) = \partial(\varphi(R))$. Since ∂R has measure zero, so does $\varphi(\partial R)$ by Lemma 4.4.3 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 φ 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 4.4.2, $\psi(R_{ij})$ is sandwiched between the squares $S(z_{ij}, r(1-\sigma))$

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

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

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

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

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

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

Since σ, τ 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. $\qquad \square$

CHAPTER 5

Function Spaces

5.1 Uniform Convergence and Completeness

A sequence of functions $f_n : [a, b] \to \mathbb{R}$ converges uniformly to a function $f : [a, b] \to \mathbb{R}$ 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 5.1.1 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

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

Hence proved. $\square$

Remark Even if $f_n \to f$ 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 $f_n \to 0$ pointwise, but not uniformly.

Let

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

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 5.1.2 Convergence with respect to the supremum norm is equivalent to uniform convergence.

Proof. $f_n \rightrightarrows f$ 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. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Theorem 5.1.1 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

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

Clearly $f_n \to f$ 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 $f_n \to f$ 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 5.1.2 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 5.1.1. $\square$

Proposition 5.1.3 (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

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

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

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

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 5.1.1, hence f is Riemann integrable. To prove the equality, note that

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

$\square$

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

Proof. Suppose that $f_n : [a, b] \to \mathbb{R}$ 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

$$\phi_n(t) = \begin{cases} \dfrac{f_n(t) - f_n(x)}{t - x}, & t \neq x, \\ f_n'(x), & t = x, \end{cases}$$

$$\phi(t) = \begin{cases} \dfrac{f(t) - f(x)}{t - x}, & t \neq x, \\ g(x), & t = x. \end{cases}$$

Note that each ϕ_n is continuous and $\phi_n \to \phi$ 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 ξ 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 ϕ_n is a Cauchy sequence in C_b as well. By

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

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 $f_n \to |x|$ uniformly, but $|x|$ is not differentiable at 0. The derivatives converge pointwise but not uniformly.

5.2   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

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

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

Definition 5.2.1   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 δ does not depend on α.

Theorem 5.2.2 (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 $(f_{n(1,k)}(d_1)) \to y_1$ as $k \to \infty$. 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 $k \to \infty$. 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

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

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 $\{d_1, d_2, \ldots, d_J\} \subseteq \mathcal{D}$ such that every $x \in [a, b]$ is within δ of some d_j. Since $g_m(d_j) \to y_j$ 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

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

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 5.2.1 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 5.2.3 Assume that $f_n : [a, b] \to \mathbb{R}$ 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

$$|f_n(x)| \leq |f_n(x_0)| + |f_n(x) - f_n(x_0)|$$
$$= |f_n(x_0)| + |f_n'(\xi)||x - x_0|$$
$$\leq K + M(b - a).$$

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

Theorem 5.2.4 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 $B(f_i, \epsilon/3)$, $i = 1, 2, \ldots, N$. 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 $|x - y| < \delta$,

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

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 $|x - y| < \delta$,

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

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. $\qquad \square$

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 ρ is a bounded metric on Y that induces the same topology as d_Y. In particular, (Y, ρ) 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 5.2.2 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 $d_Y(f_n(x), f(x)) \rightrightarrows 0$, 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 $f(x) := \lim_{n \to \infty} f_n(x)$ exists for each $x \in X$. Cauchyness of (f_n) implies uniform convergence and $f_n \to f$ with respect to d.

(c) If $f_n \in \mathcal{F}_b$ and $d(f_n, f) \to 0$, then for any constant function c,

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

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

(d) If $f_n \in C^0(X, Y)$ and $d(f_n, f) \to 0$, then for each fixed $x \in X$, $f_n(x) \to f(x)$. Given $\epsilon > 0$, choose N such that for all $n \ge N$ 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 $n \ge N$,

$$d_Y(f(x), f(y)) \le d_Y(f(x), f_n(x)) + d_Y(f_n(x), f_n(y)) + d_Y(f_n(y), f(y))$$
$$< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Hence f is continuous at x.

$\square$

Definition 5.2.5

- 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 δ 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 5.2.3 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 $x'_{n_k} \to x_0$ 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 5.2.4 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 5.2.6 A topological space X is σ-compact if it is a countable union of compact subsets.

Example 5.2.1 $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are σ-compact. Uncountable sets equipped with the discrete topology are not σ-compact.

Proposition 5.2.5 If X is σ-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). $\qquad\square$

Corollary 5.2.7 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 $a = x_0 < x_1 < x_2 < \cdots < x_m = b$ 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 $x \in \mathbb{R}$. $\qquad\square$

5.3 Uniform Approximation in C^0

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

Theorem 5.3.1 (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 $n \to \infty$.

Let $u = x + t$, then

$$P_n(x) = \int_{x-1}^{x+1} f(u)\beta_n(u - x)\,du$$

$$= \int_{0}^{1} f(u)\beta_n(u - x)\,du,$$

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

$$1 = \int_{-1}^{1} \beta_n(t)\,dt$$

$$\geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} b_n(1 - t^2)^n\,dt$$

$$= 2b_n \int_{0}^{1/\sqrt{n}} (1 - t^2)^n\,dt$$

$$\geq 2b_n \int_{0}^{1/\sqrt{n}} \left(1 - \frac{1}{n}\right)^n\,dt$$

$$= \frac{2b_n}{\sqrt{n}} \left(1 - \frac{1}{n}\right)^n,$$

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

$$\beta_n(t) = b_n(1 - t^2)^n < c\sqrt{n}(1 - \delta^2)^n \to 0 \text{ as } n \to \infty.$$

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 δ, 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 $[0,1]$. $\qquad\square$

We now extend this to functions defined on a compact Hausdorff space X.

Definition 5.3.2

- 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 5.3.3   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 $f(x_2) = c_2$. $\qquad\square$

Lemma 5.3.4   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

$$\begin{aligned}
|q(y) - |y|| &= |p(y) - p(0) - |y|| \\
&\leq |p(y) - |y|| + |p(0)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

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 $|f| \in \overline{\overline{\mathcal{A}}} = \overline{\mathcal{A}}$. $\qquad\square$

Lemma 5.3.5 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 $\max(f, g), \min(f, g) \in \overline{\mathcal{A}}$. $\qquad\square$

Theorem 5.3.6 (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 $\|f - g\| < \epsilon$. $\qquad\square$

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 $U_1 \cap U_2 = \varnothing$. Then $f^{-1}(U_1), f^{-1}(U_2)$ are open neighborhoods of x_1, x_2 respectively such that $f^{-1}(U_1) \cap f^{-1}(U_2) = \varnothing$, a contradiction.

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 $\mathcal{T} = \{\varnothing, X, \{a\}\}$. 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. $\mathcal{T} = \{\varnothing\} \cup \{U \subseteq \mathbb{R} \mid U^c \text{ is finite}\}$. 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.

$$\mathcal{T} = \{\varnothing\} \cup \left\{ U \subseteq \mathbb{R} \mid U = V \setminus K, V \text{ is open in the usual topology}, K \subseteq \left\{ \frac{1}{n} \mid n \in \mathbb{N} \right\} \right\}.$$

  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 Δ is not normal in the subspace topology. $\qquad\square$

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 $N_{a, \mathrm{lcm}(b_1, b_2)} \subseteq \mathcal{U}_1 \cap \mathcal{U}_2$. 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 $p_1, p_2, \ldots, p_k$, 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. $\qquad\square$

A.3   Quotient Topology

Definition A.3.1   Let X, Y be topological spaces. A surjective map $p : X \to Y$ 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 $p : X \to Y$ if $C = p^{-1}(S)$ for some $S \subseteq Y$.

Definition A.3.3   A mapping $p : X \to Y$ 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 $p : X \to A$ be a surjective map. Then there is a unique topology $\mathcal{T}$ on A such that $p : (X, \mathcal{T}_X) \to (A, \mathcal{T})$ 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 $p : \mathbb{R} \to \{a, b, c\}$ 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 $\{\varnothing, \{a\}, \{b\}, \{a, b\}, \{a, b, c\}\}$.

Definition A.3.5 Let X be a topological space, and let X^* be a partition of X. Let $\pi : X \to X^*$ be the natural projection that maps each point to its partition. Then under the quotient topology on X^* induced by π, 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 $p : X \to Y$ be a quotient map. Let Z be a topological space, and $g : X \to Z$ be a map that is constant on each $p^{-1}(\{y\})$, $y \in Y$. Then, g induces a map $f : Y \to Z$ 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 $g : X \to Z$ 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 $f : X^* \to Z$, 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_α'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_α'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_α's, say $U_{\alpha_1}, U_{\alpha_2}, \ldots, U_{\alpha_n}$. Then V also intersects only finitely many $\overline{U_\alpha}$'s, namely $\overline{U_{\alpha_1}}, \overline{U_{\alpha_2}}, \ldots, \overline{U_{\alpha_n}}$. 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 $\overline{U_{\beta_1}}, \overline{U_{\beta_2}}, \ldots, \overline{U_{\beta_m}}$. 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. $\qquad\square$

Lemma A.4.2 (Pasting Lemma)   Let $\{U_\alpha\}_{\alpha \in A}$ be a cover of X and $f : X \to Y$ be a map. If either

(i) each U_α is open, or

(ii) each U_α 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 $V \subseteq Y$, $\qquad\square$

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 $\phi : B \to A$ 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 $\{A_\alpha\}_{\alpha \in A}$. $\qquad \square$

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_α'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. $\qquad \square$

Definition A.4.6 Let X, Y be spaces, and $f : X \to Y$ be a map. The support of f is the set $\mathrm{supp}(f) = \overline{\{x \in X \mid f(x) \neq 0\}}$.

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) $\{supp(k_\alpha)\}_{\alpha \in A}$ 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. $supp(k_\alpha) \subseteq U_\alpha$ for each $\alpha \in A$.

Proposition A.4.3 Let X be a paracompact T_2 space, and $G, g : X \to \mathbb{R}$ are upper semicontinuous and lower semicontinuous functions respectively such that $G(x) < g(x)$ for each $x \in X$. Then $\exists$ a continuous function $f : X \to \mathbb{R}$ 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 ϵ.

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. □

Theorem A.5.2 (Cantor Surjection Theorem)   Given any compact metric space M, there is a continuous surjection $f : C \to M$, 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_α, $\alpha \in W(n_1)$. Then, divide S_α 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_{α_1} by $S_{\alpha_1 \alpha_2}$, $\alpha_2 \in W(n_2)$. Continuing this process, we obtain a sequence of integers $n_1, n_2, \ldots$, and a family of nonempty compact sets $S_{\alpha_1 \alpha_2 \ldots \alpha_k}$ of diameter at most 2^{-k+1}. For each infinite binary string $w = \alpha_1 \alpha_2 \ldots$ where $\alpha_i \in W(n_i)$, define

$$S_w = \bigcap_{k=1}^{\infty} S_{\alpha_1 \alpha_2 \ldots \alpha_k}.$$

Since $S_{\alpha_1 \alpha_2 \ldots \alpha_k}$ is a decreasing sequence of nonempty compact sets and the diameter approaches 0, by 2.3.1, S_w is a singleton. Define $f : C \to M$ 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. □

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 $f : C \to M$. 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 $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$, then

$$\left\{ U_{x_1}, U_{x_2} \setminus U_{x_1}, U_{x_3} \setminus (U_{x_1} \cup U_{x_2}), \ldots, U_{x_n} \setminus \bigcup_{i=1}^{n-1} U_{x_i} \right\}$$

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. $\qquad\square$

A.6 Local Compactness and Topological Groups

Let $f : \mathbb{R} \to \mathbb{R}$ 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. $\qquad\square$

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, $\bigcap_{n=1}^{\infty} \overline{U_n} \neq \varnothing$. 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 σ-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 $f : G \to H$ be a continuous epimorphism, then f is an open map whenever G is σ-compact.

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

$$\lim_{n \to \infty} \int_a^b |f_n(x) - f(x)| \, dx = 0.$$

In particular,

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

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 $f : [a, b] \to \mathbb{R}$ and any $\epsilon > 0$, there exists a continuous function $g : [a, b] \to \mathbb{R}$ 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

$$\lim_{n \to \infty} \int_a^b f_n(x) \, dx = 0.$$

Proof. By Lemma A.7.3, $\forall \epsilon > 0$, $\exists$ a continuous function $g_n : [a, b] \to \mathbb{R}$ 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 $h_n = \min\{g_1, g_2, \ldots, g_n\}$. 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

$$\lim_{n \to \infty} \int_a^b h_n(x) \, dx = 0.$$

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

$$\int_a^b f_{n+1}(x)\,dx \le \int_a^b g_{n+1}(x)\,dx + \frac{\epsilon}{2^{n+1}}$$

$$\le \int_a^b \max\{h_n(x), g_{n+1}(x)\}\,dx + \frac{\epsilon}{2^{n+1}}$$

$$= \int_a^b h_n(x)\,dx + \int_a^b \max\{0, g_{n+1}(x) - h_n(x)\}\,dx + \frac{\epsilon}{2^{n+1}}$$

$$\le \int_a^b h_n(x)\,dx + \int_a^b g_{n+1}(x)\,dx + \frac{\epsilon}{2^{n+1}}$$

$$\le \int_a^b h_n(x)\,dx + \int_a^b g_n(x)\,dx + \frac{\epsilon}{2^n} + \frac{\epsilon}{2^{n+1}}$$

$$\le \int_a^b h_n(x)\,dx + \epsilon(1 - 2^{-n}) + \frac{\epsilon}{2^{n+1}}$$

$$= \int_a^b h_n(x)\,dx + \epsilon(1 - 2^{-(n+1)}).$$

$\square$

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 $x \in [a, b]$, $|f_n(x)| \le g(x)$, then

$$\lim_{n \to \infty} \int_a^b |f_n(x) - f(x)|\,dx = 0.$$

Proof. WLOG assume $f_n \ge 0$ and $f \equiv 0$ by considering the positive parts and negative parts separately. Set $g_{n,m} \equiv \max\{f_n, \ldots, f_m\}$ for $m \ge n$. Then $0 \le g_{n,m} \le g$. 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

$$0 \le \int_a^b \left(g_{n,k}(x) - g_{n,m_n}(x)\right)\,dx \le \frac{\epsilon}{2^n}, \quad \forall k > m_n.$$

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

$$0 \le f_n \le u_n \le \min\{u_1, \ldots, u_n\} + \sum_{i=1}^{n-1} \left(\max\{u_i, \ldots, u_n\} - u_i\right).$$

Since

$$\max\{u_i, \ldots, u_n\} - u_i = \max\{f_i, \ldots, f_{m_n}\} - \max\{f_i, \ldots, f_{m_i}\} \le g_{i,m_n} - g_{i,m_i},$$

we have

$$\int_a^b \left(\max \{u_i, \ldots, u_n\} - u_i \right) dx \leq \frac{\epsilon}{2^i}, \quad \forall i < n$$

$$\implies 0 \leq \int_a^b f_n(x)\, dx \leq \int_a^b \min \{u_1(x), \ldots, u_n(x)\}\, dx + \epsilon(1 - 2^{-(n-1)}).$$

$\square$

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

$$\int_a^b f(x)\, dx \leq \liminf_{n \to \infty} \int_a^b f_n(x)\, dx.$$

Proof. Since $f - f_n \leq f$, the positive part is also bounded by f. By the previous theorem, we have

$$\int_a^b (f(x) - f_n(x))_+\, dx \to 0.$$

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. $\square$

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 $r < r_0$,

$$|f'(z_0)| = \left| \frac{1}{2\pi i} \int_{|z-z_0|=r} \frac{f(z)}{(z-z_0)^2} \, dz \right|$$
$$\leq \frac{1}{2\pi} \cdot \frac{2\pi r M}{r^2} = \frac{M}{r},$$

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 $V \times V \to V$, $(x, y) \mapsto x + y$ and scalar multiplication $\mathbb{F} \times V \to V$, $(a, x) \mapsto ax$ 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

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(0) = 0 \end{cases}, \quad x \in \mathbb{R}^n, f : \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^n.$$

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

$$D = \text{Dom}(A) := \left\{ x \in C^u[a, b] \mid \|x\|_\infty := \sup_{0 \leq t \leq a} \|x(t)\| \leq r \right\}.$$