NCKU 112.2
General Analysis

Eric Liu

Contents

Chapter 1

General Topology

1.1 Equivalent Characterizations of Basic Notions

Abstract

This section give a compact and comprehensive development of some of the most basic notions in the study of topology. In this section, $(X, \mathscr{T})$ is a topological space.

Given a collection $\mathcal{B} \subseteq \mathscr{T}$ of open sets, we say $\mathcal{B}$ is a

(a) **basis** if for each $O \in \mathscr{T}$ there exists a subcollection $\mathcal{B}_0$ such that $O \subseteq \mathcal{B}_0$.

(b) **subbase** if $\mathscr{T}$ is the collection of unions of finite intersections of $\mathcal{B}$. In a more formal language, $\mathcal{B}$ has to satisfy $\mathscr{T} = \{\bigcup A : A \subseteq \mathcal{A}\}$ where $\mathcal{A} = \{\bigcap \mathcal{B}_0 : \text{ card } \mathcal{B} \in \mathbb{N} \text{ and } \mathcal{B}_0 \subseteq \mathcal{B}\}$

Theorem 1.1.1. (Equivalent Definition of Basis) The following statements are equivalent.

(a) $\mathcal{B}$ is a basis.

(b) For all $O \in \mathscr{T}$ and $x \in O$, there exists $B \in \mathcal{B}$ such that $x \in B \subseteq O$.

Proof. Check straightforward. ∎

Theorem 1.1.2. (Equivalent Definition of Subbase) The following statements are equivalent.

(a) $\mathcal{B}$ is a subbase of $\mathscr{T}$.

(b) $\mathcal{B}$ cover X and $\mathscr{T}$ is the smallest topology containing $\mathcal{B}$.

Proof. Check straightforward. ■

Immediately, with the equivalent definitions, one should check

(a) Given any cover $\mathcal{B}$ of X, there always exists a unique topology $\mathscr{T}$ containing $\mathcal{B}$ as a subbase. We say $\mathscr{T}$ is the **topology generated by** $\mathcal{B}$.

(b) The set $\mathcal{A} \triangleq \{\bigcap S : S \subseteq \mathcal{B},\ \text{card } S \in \mathbb{N}\}$ of finite intersections of cover $\mathcal{B}$ is a basis of the topology generated by $\mathcal{B}$.

(c) Not every cover $\mathcal{B}$ of X has some topology $\mathscr{T}$ containing $\mathcal{B}$ as a basis. Consider $\mathcal{B} \triangleq \{(-\infty, a) : a \in \mathbb{R}\} \cup \{(b, \infty) : b \in \mathbb{R}\}$. Even the smallest topology containing $\mathcal{B}$, i.e. the standard topology, does not have $\mathcal{B}$ as a basis.

(d) However, cover $\mathcal{B}$ is the basis of the topology $\mathscr{T}$ generated by itself if for all $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, there exists B_3 such that $x \in B_3 \subseteq B_1 \cap B_2$.

(e) If $\mathcal{B}$ is a basis of $\mathscr{T}$, then $\mathcal{B}$ is also a subbase of $\mathscr{T}$.

(f) Basis is not necessarily closed under finite intersection. Consider the basis $\{(a, a+\frac{1}{n}) : a \in \mathbb{R}, n \in \mathbb{N}\}$ for $\mathbb{R}$'s standard topology.

Note that in (a), to check the generated $\mathscr{T}$ is indeed a topology, one may need to utilize the identity

$$\Big(\bigcup_{i\in I} A_i\Big) \cap \Big(\bigcup_{j\in J} B_j\Big) = \bigcup_{i\in I, j\in J} A_i \cap B_j.$$

Now, given an arbitrary subset $E \subseteq X$, we

(a) say $x \in X$ is a **limit point of** E if every open O containing x contain a point $y \in E$ such that $y \neq x$.

(b) say $x \in E$ is an **interior point of** E if there exists $O \in \mathscr{T}$ such that $x \in O \subseteq E$.

(c) define the **interior** E° **of** E to be the union of all open sets contained by E.

(d) say $E \subseteq X$ is a **closed set** if $E^c \in \mathscr{T}$.

(e) define the **closure** $\overline{E}$ **of** E by $\overline{E} \triangleq E \cup E'$ where E' is the set of limit points of E.

(f) say E is **dense** in X if $\overline{E} = X$.

(g) define the **boundary** ∂E **of** E by $\partial E \triangleq \overline{E} \setminus E^\circ$

Theorem 1.1.3. (Equivalent Definitions of Interior) The following sets are equivalent

(a) E°

(b) The largest open set contained by E.

(c) The set of interior points of E.

Proof. Check straightforward. ■

Theorem 1.1.4. (Equivalent Definitions of Closed) The following statements are equivalent.

(a) E is closed.

(b) the set of limit points of E is contained by E.

(c) $\overline{E} = E$.

Proof. The proof of (a) $\implies$ (b) $\implies$ (c) are straight forward. The proof of (c) $\implies$ (a) follows from first noting no $x \in E^c$ is a limit point of E. Then shows $E^c = \bigcup_{x \notin E} O_x$ where O_x is an open set containing x and disjoint with E. ■

Theorem 1.1.5. (Equivalent Definitions of Closure) The following sets are equivalent.

(a) $\overline{E}$

(b) $((E^c)^\circ)^c$

(c) The smallest closed set containing E.

(d) $\{x \in X : \text{ every open } O \text{ containing } x \text{ intersect with } E\ \}$

Proof. (a) = (d) is obvious. To verify (a) = (c), check $(\overline{E})' \subseteq E'$ and check $E' \subseteq F' \subseteq F$ for each closed F containing E. Lastly, to verify (b) = (c), check $(\overline{E})^c = (E^c)^\circ$ using the largest open set and the smallest closed set characterization of interior and closure. ■

Theorem 1.1.6. (Equivalent Definitions of Dense) The following statements are equivalent.

(a) E is dense in X.

(b) Every non-empty open set intersect with E.

(c) $(E^c)^\circ = \varnothing$

Proof. (a) = (c) follows from $\overline{E} = ((E^c)^\circ)^c$, and (a) = (b) follows from $\overline{E} = \{x \in X : \text{ every open } O \text{ containing } x \text{ intersect with } E\}$. ■

Theorem 1.1.7. (Equivalent Definitions of Boundary) The following sets are equivalent.

(a) ∂E

(b) $\overline{E} \cap \overline{E^c}$

(c) $\{x \in X :$ every open O containing x intersect with both E and E^c $\}$

Proof. (a) $=$ (b) follows from $(E^\circ)^c = \overline{E^c}$ and (b) $=$ (c) follows from $\overline{E} = \{x \in X :$ every open O containing x intersect with $E\}$. ■

We now develop the theory of continuity by first giving a pointwise definition. Given another topological space $(Y, \mathscr{S})$ and a function $f : X \to Y$, we say f is **continuous at** $x \in X$ if for all open O containing $f(x)$, there exists open E containing x such that $f(E) \subseteq O$. We say f is a **continuous (or $(\mathscr{T}, \mathscr{S})$-continuous, if necessary) function** if f is continuous at all $x \in X$.

It is easy to see the composition of two continuous function must be continuous. However, one should notice that the composition of a continuous function and a discontinuous function can be continuous. Just let one of them be a constant function.

Theorem 1.1.8. (Equivalent Definitions of Continuous function) The following are equivalent

(a) f is continuous.

(b) $f^{-1}(O) \in \mathscr{T}$ for all $O \in \mathscr{S}$.

(c) $f^{-1}(F)$ is closed for all closed F in Y.

(d) For all $B \subseteq Y$, $f^{-1}(B^\circ) \subseteq (f^{-1}(B))^\circ$

(e) For all $A \subseteq X$, $f(\overline{A}) \subseteq \overline{f(A)}$.

(f) For all $B \subseteq Y, \overline{f^{-1}(B)} \subseteq f^{-1}(\overline{B})$

(g) For all subbase $\mathcal{B}$ of Y, $f^{-1}(B) \in \mathscr{T}$ for all $B \in \mathcal{B}$.

Proof. It is straightforward to check (a) $\implies$ (b) $\implies$ (c). To verify (c) $\implies$ (a), check $x \in (f^{-1}(O^c))^c$ and $f((f^{-1}(O^c))^c) \subseteq)$ for each $x \in X$ and $O \in \mathscr{S}$ containing $f(x)$. Respectively, to verify (b) $\implies$ (d), (c) $\implies$ (e) and (c) $\implies$ (f), check $f^{-1}(B^\circ) \subseteq f^{-1}(B), A \subseteq f^{-1}(\overline{f(A)})$ and $f^{-1}(B) \subseteq f^{-1}(\overline{B})$. Note that (e) $\implies$ (f) follows from noting $A = f^{-1}(B)$. Check (d) $\implies$ (b) and (f) $\implies$ (c) straightforwardly, and we have proved the equivalency of statements from (a) to (f). Lastly, check (b) $\iff$ (g) straightforwardly. ■

One may wonder: Why isn't "For all $A \subseteq X$, $f(A)^\circ \subseteq f(A^\circ)$" a characterization of f being continuous? Consider a function that maps some topological space with a subset that has an empty interior into the topological space Y having only a single point.

1.2 Equivalent Definition of Subspace and Product

Abstract

Theorem 1.2.1. (Equivalent Definition of Finer/Coarser Topologies) Given another topology $\mathscr{T}'$ on X, the following are equivalent:

(a) $\mathscr{T} \subseteq \mathscr{T}'$.

(b) $\mathbf{id} : (X, \mathscr{T}') \to (X, \mathscr{T})$ is continuous.

(c) Given any basis $\mathcal{B}$ of $\mathscr{T}$ and any basis $\mathcal{B}'$ of $\mathscr{T}'$, for all $x \in X$ and basic open $B \in \mathcal{B}$ containing x, there exists a basic open $B' \in \mathcal{B}'$ such that $x \in B' \subseteq B$.

(d) There exists subbases $\mathcal{B}, \mathcal{B}'$ of $\mathscr{T}, \mathscr{T}'$ such that $\mathcal{B} \subseteq \mathcal{B}'$.

Proof. (d) $\iff$ (a) $\iff$ (b) are straightforward. From (a) to (c), one finds B' by noting $B \in \mathscr{T}'$ and utilizing the definition of basis. From (c) to (a), one shows $O \in \mathscr{T}$ belongs to $\mathscr{T}'$ by taking $O = \bigcup_{x \in O} B'_x$. ■

Now, given a collection $(X_\alpha)_{\alpha \in J}$ of topological spaces, we define the **product topology** on $X \triangleq \prod_{\alpha \in J} X_\alpha$ to be the smallest topology such that for all $\alpha \in J$, the projection $\pi_\alpha : X \to X_\alpha$ that maps $(y_\alpha)_{\alpha \in J}$ to y_α is continuous.

Theorem 1.2.2. (Equivalent Definition of Product Topology) Let $\mathcal{B}_\alpha$ each be the subbase of X_α. The following topologies are equivalent.

(a) The product topology on X.

(b) The topology on X generated by the basis $\{\prod U_\alpha : U_\alpha \neq X_\alpha$ for finitely many α and $U_\alpha \in \mathcal{B}_\alpha\}$.

(c) The smallest topology on X satisfying the statement: For all topological spaces $(Z, \mathscr{T}_Z)$, a function $f : Z \to X$ is continuous if and only if for all $\alpha \in J$, the function $f \circ \pi_\alpha : Z \to X_\alpha$ is continuous.

Proof. By definition, we know X has the subbase $\bigcup_{\alpha \in J} \mathcal{G}_\alpha$ where $\mathcal{G}_\alpha \triangleq \{\pi_\alpha^{-1}(U) : U \in \mathscr{T}_\alpha\}$. This gives us (a) = (b). For (a) = (c), we have to prove that: Given a topology on X, the following two statements are equivalent:

(i) All projections π_α are continuous.

(ii) For all topological spaces $(Z, \mathscr{T}_Z)$ and functions $f : Z \to X$, the function f is continuous if and only if for all $\alpha \in J$, the function $f \circ \pi_\alpha : Z \to X_\alpha$ is continuous.

Taking $Z \triangleq X$ and $f \triangleq \mathbf{id}$ proves (ii) $\implies$ (i). For (i) $\implies$ (ii), if f is continuous, then $f \circ \pi_\alpha$ is clearly continuous for all α. Conversely, if $f \circ \pi_\alpha$ is continuous for all α, use the pointwise definition of continuity and the basis given in (b) to show that f is continuous. ■

Immediately, one should check:

(a) In X, a sequence (p_n) converges to p if and only if $\pi_\alpha(p_n)$ converges to $\pi_\alpha(p)$ for all $\alpha \in J$. This generalize what happen in $\mathbb{R}^n$.

(b) If $f : X \times Y \to Z$ is continuous, then for all $x \in X$, the function $f(x, \cdot) : Y \to Z$ defined by $f(x, \cdot)(y) \triangleq f(x, y)$ is continuous. The converse is not true, in the sense that f can be discontinuous even if $f(x, \cdot)$ and $f(\cdot, y)$ are continuous for all x and y. Elementary counterexample can be constructed in $\mathbb{R}^2$.

(c) The third characterization of product topology shows that the product is independent of expression. For example, given an enumeration $(X_\alpha)_{\alpha \leq \gamma}$ of topological spaces, $X_1 \times \prod_{1<\alpha\leq\gamma} X_\alpha$ is homeomorphic to $\prod_{1\leq\alpha\leq\gamma} X_\alpha$.

We will later introduce an inferior alternative to assigning topology onto the Cartesian product of topological spaces. Now, given a topological space $(X, \mathscr{T})$ and a subset $E \subseteq X$, we define the **subspace topology** $\mathscr{T}_E$ on E by $\mathscr{T}_E \triangleq \{O \cap E : O \in \mathscr{T}\}$. Immediately, one can check that $\mathscr{T}_E$ is indeed a topology, and:

(a) Given a subset $F \subseteq E$, viewing F as a subspace of E or X makes no difference.

(b) The collection of closed sets in $(E, \mathscr{T}_E)$ is $\{F \cap E : F \text{ is a closed set in } (X, \mathscr{T})\}$.

(c) If (U_α) is an open cover of X and $(\mathcal{B}_\alpha)$ are basis of (U_α), then $\bigcup_\alpha \mathcal{B}_\alpha$ form a basis of X.

(d) For all $F \subseteq X$, $\text{cl}_E(F \cap E) \subseteq \text{cl}_X(F) \cap E$. The equality holds when $F \subseteq E$.

(e) Given a function $f : X \to Y$ and a subset F of Y containing $f(X)$, $f : X \to Y$ is continuous at p if and only if $f : X \to F$ is continuous at p.

(f) Given a function $f : X \to Y$ and $p \in E \subseteq X$, if f is continuous at p, then $f|_E : E \to Y$ is continuous at p. The converse is true only when E is open in X.

(g) Given a finite collection of closed subspace E_j such that $X = \bigcup E_j$, if $f|_{E_j} : E_j \to Y$ are all continuous, then $f : X \to Y$ is continuous. (**Paste Lemma**)

Theorem 1.2.3. (Equivalent Definition of Subspace Topology) Given a basis $\mathcal{B}$ and a subbase $\mathcal{B}'$ of $\mathscr{T}$, the following sets are equivalent:

(a) $\mathscr{T}_E$.

(b) The topology on E generated by the basis $\mathcal{B}_E \triangleq \{B \cap E : B \in \mathcal{B}\}$.

(c) The topology on E generated by the subbase $\mathcal{B}'_E \triangleq \{B' \cap E : B' \in \mathcal{B}'\}$.

(d) The smallest topology on E such that the inclusion map $\iota : E \to X$ is continuous.

Proof. Check straightforward. ■

At this point, one should check the compatibility between the definitions of subspace topology and product topology. Given a collection $(X_\alpha)_{\alpha \in J}$ of topological spaces and a subspace $(A_\alpha)_{\alpha \in J}$ of each X_α, one can view $A \triangleq \prod A_\alpha$ either as a subspace of the product $X \triangleq \prod X_\alpha$ or the product of subspaces $(A_\alpha)_{\alpha \in J}$. The two topologies are identical, and the proof goes as follows:

(a) Show that the product topology has the subbase $\{\pi_{\alpha,A}^{-1}(U_\alpha) : \alpha \in J, U_\alpha \in \mathscr{T}_{A_\alpha}\}$, where $\pi_{\alpha,A} : A \to A_\alpha$ is the projection mapping.

(b) Show that the subspace topology has the subbase $\{\pi_{\alpha,X}^{-1}(U_\alpha) \cap A : \alpha \in J, U_\alpha \in \mathscr{T}_{X_\alpha}\}$, where $\pi_{\alpha,X} : X \to X_\alpha$ is the projection mapping.

(c) Show that $\{\pi_{\alpha,A}^{-1}(U_\alpha) \subseteq A : \alpha \in J, U_\alpha \in \mathscr{T}_{A_\alpha}\} = \{\pi_{\alpha,X}^{-1}(U_\alpha) \cap A : \alpha \in J, U_\alpha \in \mathscr{T}_{X_\alpha}\}$.

1.3 Connected

Abstract

Given a topological space $(X, \mathscr{T})$, we say nonempty $E \subseteq X$ is

(a) **connected** if E can not be written as $E = A \cup B$ so that $\overline{A} \cap B = A \cap \overline{B} = \varnothing$ and $A \neq \varnothing \neq B$.

(b) **path-connected** if for each $p, q \in E$, there exists a continuous function $f : [0, 1] \to E$ such that $f(0) = p, f(1) = q$.

These two notions are sometimes referred as **topological properties**, since they are invariant under continuous function, the "morphism" between topological space. Put more precisely, If $E \subseteq X$ satisfy a topological property and $f : X \to Y$ is continuous, then $f(E)$ also satisfy the topological property.

Theorem 1.3.1. (Equivalent Definitions of Connected) Given a subset $E \subseteq X$, the following statements are equivalent

(a) E is connected in $(X, \mathscr{T})$.

(b) E is connected in $(E, \mathscr{T}_E)$.

(c) The only clopen sets in $(E, \mathscr{T}_E)$ are E and $\varnothing$.

(d) In $(E, \mathscr{T}_E)$, the only set that has empty boundary are E and $\varnothing$.

(e) All continuous function from $(E, \mathscr{T}_E)$ to $\{0, 1\}$ with discrete topology is constant.

Proof. For (a) $\iff$ (b), use the identity $\forall A \subseteq E$, $\text{cl}_X(A) \cap E = \text{cl}_E(A)$. Check straightforward for (b) $\iff$ (c) and (d) $\iff$ (c) $\iff$ (e). ■

Three things to note here

(a) If $E \subseteq X$ is connected, E can not be covered by any two disjoint open sets intersecting with E in $(X, \mathscr{T}_X)$. The converse is not true. Consider finite subset of an infinite set with cofinite topology.

(b) Union of collection $(A_\alpha)_{\alpha \in J}$ of connected sets with non-empty intersection is connected. Prove this by a proof of contradiction. Path-connectedness has the same property, and the proof is much easier.

(c) If E is connected and $E \subseteq F \subseteq \mathrm{cl}(E)$, then F is connected. Use the fifth equivalent definitions for connected to prove this. Path-connectedness doesn't have the same property this time. Consider the "fattened" Topologist's sine curve $\{(x, y) \in \mathbb{R}^2 : \left|y - \sin \frac{1}{x}\right| < x\}$. It's closure can be easily proved to be $\{(x, y) \in \mathbb{R}^2 : \left|y - \sin \frac{1}{x}\right| \leq x\} \cup \{(0, y) \in \mathbb{R}^2 : y \in [-1, 1]\}$.

(d) Path-connectedness is strictly stronger than connectedness. This can be proved using a proof of contradiction and supremum. Two famous counterexamples of the converse are Topologist's Sine Curve and Long line.

1.4 Compact

Abstract

We now give definitions to three notions so important that they drive us to study Topology in the first place. Given a topological space $(X, \mathscr{T})$, we say nonempty $E \subseteq X$ is

(a) **compact** if every open cover has a finite subcover.

These three properties are often called **topological properties**, since they are invariant under continuous function, the "morphism" between topological space. Put more precisely, If $E \subseteq X$ satisfy a topological property and $f : X \to Y$ is continuous, then $f(E)$ also satisfy the topological property.

Immediately, one should again check the "natural" behaviors of subspace topology: Whether a set E is connected, path-connected, or compact is independent of the choices of ambient space. In other words, given $E \subseteq X$, E is connected, path-connected or compact in $(X, \mathscr{T})$ if and only if E is connected, path-connected or compact in $(E, \mathscr{T}_E)$.

Theorem 1.4.1. (Equivalent Definitions of Compact) The following statements are equivalent

(a) E is compact in $(X, \mathscr{T})$.

(b) E is compact in $(E, \mathscr{T}_E)$.

(c) Given subbase $\mathcal{B}$ of $(X, \mathscr{T})$, every cover of E consisting of the elements of $\mathcal{B}$ has a finite subcover.

(d) Every infinite subset M of E has a complete limit point in E, that is, a point $x \in E$ such that all open set O containing x satisfy $|O \cap M| = |M|$.

(e) Every collection of closed sets of $(E, \mathscr{T}_E)$ that has finite intersection property has nonempty intersection.

(f) For all topological space Y, the projection $\pi_Y : E \times Y \to Y$ is a closed mapping.

Proof. For (b) $\iff$ (e), use proofs by contradiction. (b) $\iff$ (a) $\implies$ (c) are clear. We now prove

$$(c) \implies (b)$$

Fix $\mathcal{B}$. Assume E is not compact. Then the collection $\mathbb{S}$ of all open covers that have no finite subcover is non-empty. Let $\mathcal{C}$ be a maximal element of $\mathbb{S}$. It is clear that $\mathcal{C} \cap \mathcal{B}$ is not a cover of E by premise. Let $x \in E \setminus \bigcup(\mathcal{C} \cap \mathcal{B})$. Let U be an element of $\mathcal{C} \setminus \mathcal{B}$ containing x. Because $\mathcal{B}$ is a subbase, there exists finite $B_1, \ldots, B_n \in \mathcal{B}$ such that $x \in B_1 \cap \cdots \cap B_n \subseteq U$. Because $\mathcal{C}$ is a maximal element of $\mathbb{S}$, for all j, the collection $\mathcal{C} \cup \{B_j\}$ does not belong to $\mathbb{S}$. This implies that for each $j \in \{1, \ldots, n\}$, there exists a finite sub-collection $\mathcal{C}_j \subseteq \mathcal{C}$ such that $\mathcal{C}_j \cup \{B_j\}$ covers E. Let $\mathcal{C}_F \triangleq \bigcup_{j=1}^n \mathcal{C}_j$. Because $\mathcal{C}_j \cup \{B_j\}$ are covers of E, $\mathcal{C}_F \cup \{B_1, \ldots, B_n\}$ is a cover of E. This implies $\mathcal{C}_F \cup \{U\} \subseteq \mathcal{C}$ is a finite subcover. CaC (done)

We now prove

$$\text{(a)} \implies \text{(d)}$$

Assume there exists infinite $M \subseteq E$ that has no complete limit point. Because of our assumption, for each $x \in E$, there exists an open set O_x containing x such that $|M \cap O_x| < |M|$. Because $(O_x)_{x \in E}$ is an open cover of E, there exists a finite sub-cover $(O_x)_{x \in I}$. Note that M is infinite, so we can deduce

$$|M| = \left| \bigcup_{x \in I} M \cap O_x \right| \leq \sum_{x \in I} |M \cap O_x| < |M| \text{ CaC } \text{ (done)}$$

We now prove

$$\text{(d)} \implies \text{(a)}$$

Assume E is not compact. Let O be an open cover of E that has no finite subcover with smallest cardinality c. Well-order O by $O \triangleq \{O_\alpha\}_{\alpha < c}$. Use transfinite recursion to build $M \triangleq \{x_\alpha : \alpha < c\}$ where $x_\alpha \in E \setminus \bigcup_{\beta < \alpha} O_\beta$. Such x_α always exists; otherwise, there exists an open cover of E that has no finite subcover with cardinality smaller than c. To cause a contradiction, it remains to show

$$M \text{ has no complete limit point in } E$$

Because O is an open cover of E, for all x, there exists some O_α containing x. Observe using the definition of M

$$|O_\alpha \cap M| \leq |\{x_\gamma : \gamma \leq \alpha\}| \leq |\alpha| < c = |M| \text{ CaC } \text{ (done)}$$

Before we prove (a) $\implies$ (f), we first prove the Generalized Tube Lemma. That is,

> Given a product space $X \times Y$, compact $A \subseteq X$, compact $B \subseteq Y$, and $N \subseteq X \times Y$ open containing $A \times B$, there exists $U \subseteq X$ open, $V \subseteq Y$ open such that $A \times B \subseteq U \times V \subseteq N$.

First note that for all $(a,b) \in A \times B$, there exists $U_{(a,b)} \subseteq X$ open and $V_{(a,b)} \subseteq Y$ open such that $(a,b) \in U_{(a,b)} \times V_{(a,b)} \subseteq N$. Because A is compact and for all b, the collection $(U_{(a,b)})_{a \in A}$ is an open cover of A, there exists a finite subset $A_b \subseteq A$ for all b such that $A \subseteq \bigcup_{a \in A_b} U_{(a,b)}$. Now, let $U_b \triangleq \bigcup_{a \in A_b} U_{(a,b)}$ and $V_b \triangleq \bigcap_{a \in A_b} V_{(a,b)}$. It is clear that U_b, V_b are open, and it is straightforward to check $A \times \{b\} \subseteq U_b \times V_b \subseteq N$. Again, because B is compact and $(V_b)_{b \in B}$ is an open cover of B, there exists a finite subset $B_0 \subseteq B$ such that $B \subseteq \bigcup_{b \in B_0} V_b$. Let $V \triangleq \bigcup_{b \in B_0} V_b$ and $U \triangleq \bigcap_{b \in B_0} U_b$. It is straightforward to check U, V suffice. (done)

We now prove

$$\text{(a)} \implies \text{(f)}$$

Given $A \subseteq X \times Y$ closed, we are required to prove $\pi_Y(A)$ is closed. WOLG, assume $\pi_Y(A) \neq Y$. Fix $y \in Y \setminus \pi_Y(A)$. Because X and $\{y\}$ are compact and $X \times \{y\}$ is a subset of the open set A^c, by the Generalized Tube Lemma, there exists open $V \subseteq Y$ such that $X \times \{y\} \subseteq X \times V \subseteq A^c$. It is straightforward to check $V \cap \pi_Y(A) = \varnothing$. (done)

Lastly, we prove

$$\text{(f)} \implies \text{(a)}$$

Assume X is not compact. Let $(O_\alpha)_{\alpha \in J}$ be an open cover of X with no finite subcover. Consider the following construction:

(a) $\mathcal{U} \triangleq \{\bigcup_{\alpha \in I} O_\alpha : I \text{ is a finite subset of } J\}$ is an open cover of X with no finite subcover,

(b) $\mathcal{U}$ is closed under finite union,

(c) $\mathcal{F} \triangleq \{U^c : U \in \mathcal{U}\}$ is a collection of non-empty closed sets that has the finite intersection property.

(d) If we let $Y \triangleq X \cup \{p\}$ where $p \notin X$, then $\mathscr{T}_Y \triangleq \mathcal{P}(X) \cup \{\{p\} \cup A : \exists F \in \mathcal{F}, F \subseteq A \subseteq X\}$ is a topology on Y, where $\mathcal{P}(X)$ is the collection of all subsets of X.

(e) Let $C \triangleq \text{cl}_{X \times Y}\{(x,x) \in X \times Y : x \in X\}$.

(f) Fix $x \in X$. Because $\mathcal{U}$ is an open cover of X, there exists $U \in \mathcal{U}$ containing x. Note that $\{p\} \cup U^c$ is open in Y. This implies $U \times (\{p\} \cup U^c)$ is an open subset of $X \times Y$ containing (x,p). We have proved $C \subseteq X \times X$.

(g) It is clear that X is not closed in Y. Now observe that π_Y maps the closed set C to the open set $X \subseteq Y$. CaC (done)

$\blacksquare$

Notably, one can easily check that closed subspace of compact space must be compact, and the converse is true when the ambient space is Hausdorff.

Theorem 1.4.2. (Compact Subspace of Hausdorff Space is Closed) If E is a compact subspace of Hausdorff space X, then E is closed in X.

Proof. Fix $x \in E^c$. Because X is Hausdorff, we can associate each $y \in E$ an open set U_y containing y and an open set $U_{x,y}$ containing x such that $U_y, U_{x,y}$ are disjoint. Now, because E is compact and (U_y) is an open cover of E, we know there exists a finite sub-cover

$$E \subseteq \bigcup_{n=1}^{N} U_{y_n}$$

It then follows that open $\bigcap_{n=1}^{N} U_{x,y_n}$ contain x and disjoint with E. ■

Corollary 1.4.3. (Homeomorphism between Compact Space and Hausdorff Space) Suppose

(a) X is compact.

(b) Y is Hausdorff.

(c) $f : X \to Y$ is a continuous bijective function.

Then

$$f \text{ is a homeomorphism between } X \text{ and } Y$$

Proof. Because closed subset of compact set is compact and continuous function send compact set to compact set, we see for each closed $E \subseteq X$, $f(E) \subseteq Y$ is compact. The result then follows from $f(E) \subseteq Y$ being closed since Y is Hausdorff. ■

1.5 Countability Axioms

Abstract

This section introduce countability axioms.

Given a topological space X and some $p \in X$, by an **open-neighborhood basis $\mathcal{B}_p$ at p**, we mean a collection of open-neighborhood of p such that each open set containing p contain some element of $\mathcal{B}_p$. If we say X is **first countable at** p, we mean there exists some countable open-neighborhood basis at p, and if we say X is **first countable**, we mean X is first countable at all $p \in X$. Recall that in general topology, we have the following propositions

(a) If there exists some sequence x_n in A such that $x_n \to x$, then $x \in \overline{A}$.

(b) Given a function $f : X \to Y$, if f is continuous, then f preserve convergence, i.e., $x_n \to x \implies f(x_n) \to f(x)$

In the specific case of $\mathbb{R}^n$, the converses of these propositions also hold true, a property attributed to $\mathbb{R}^n$ being first countable. This can be verified through the construction of countable bases at x. However, the converse of this is generally not true, consider

Example 1 (Space of real eventually 0 sequence)

$$X \triangleq \{(x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}} : \text{ All but finitely many of } x_n \text{ are } 0 \ \}$$

We define $S \subseteq X$ is open $\overset{\triangle}{\iff}$ for all n, $\{(x_1, \dots, x_n) \in \mathbb{R}^n : (x_1, \dots, x_n, 0, 0, \dots) \in S\}$ is open in $\mathbb{R}^n$. Consider

$$B_n \triangleq \left\{ x \in X : \Big(\sum_{j=1}^{n} x_j^2 \Big)^{\frac{1}{2}} < \frac{1}{n} \text{ and } x_{n+k} = 0 \text{ for all } k \in \mathbb{N} \right\}$$

and

$$S \triangleq X \setminus \bigcup_{n \in \mathbb{N}} B_n$$

Suppose $x \in B_N$ for some N. By definition, we know

$$\Big(\sum_{j=1}^{N} x_j^2 \Big)^{\frac{1}{2}} < \frac{1}{N} \text{ and } x_{N+k} = 0 \text{ for all } k$$

Let U be an open-neighborhood around x. By definition, we know there exists some

$$y = (x_1, \dots, x_N, \epsilon_1, \epsilon_2, \dots, \epsilon_M, 0, 0, \dots) \in U$$

such that $y \in S$ for large enough M.

In other words, the converse of propositions (a), (b) hold true whenever X is first countable at x.

Theorem 1.5.1. (Sequential Compactness and Countable Compactness is equivalent in first countable space)

Stronger and perhaps more interesting, we say X is **second countable** if X has a countable basis.

Theorem 1.5.2. (Basis Property of Second Countable Space) Given a second countable space X, if $\mathcal{B}$ is a basis of X, then there exists some countable $\mathcal{B}'$ basis of X such that $\mathcal{B}' \subseteq \mathcal{B}$.

Proof. Let (C_n) be a countable basis of X. Define

$$\mathcal{B}_{(n,m)} \triangleq \{B \in \mathcal{B} : C_n \subseteq B \subseteq C_m\}$$

If $\mathcal{B}_{(n,m)}$ is nonempty, pick one $B_{(n,m)} \in \mathcal{B}_{n,m}$. Clearly, the collection of $B_{(n,m)}$ is countable. To see that this collection is a basis, fix x and open U containing x. Because (C_k) and $\mathcal{B}$ are both basis, there exists some n, m and B such that $x \in C_n \subseteq B \subseteq C_m \subseteq U$. This implies that $\mathcal{B}_{(n,m)}$ is nonempty and there exists some $B_{(n,m)}$ such that $x \in B_{n,m} \subseteq C_m \subseteq U$. ■

Theorem 1.5.3. (Basic Property of Second Countable Space) Open cover of second countable space always have a countable subcover.

Proof. Let X be a space with basis (B_n) and some open cover $\mathcal{S}$. For each n, let S_n be an element of $\mathcal{S}$ containing B_n if possible. To see that (S_n) form a cover, fix x and some $S \in \mathcal{S}$ containing x. Because (B_n) is a basis, there exists some B_m such that $x \in B_m \subseteq S$. This implies the existence of S_m, which contain x. ■

1.6 Locally Compact Hausdorff Space

Abstract

We say a topological space is **locally compact** if for each point p there exists some compact subspace containing some open-neighborhood of p.

Theorem 1.6.1. (Locally Compact Hausdorff Space admits a pre-compact basis) Suppose X is a locally compact Hausdorff space. There exists a basis of X whose elements all have compact closure.

Proof. If we define

$$\mathcal{B}_p \triangleq \{U : U \text{ is an open-neighborhood around } p \text{ with compact closure.}\}$$

Because compact subspace of Hausdorff space is closed and closed subspace of compact space is compact, one can deduce $\mathcal{B}_p$ is a local basis. ∎

1.7 Quotient Topology

Abstract

This is a short section introducing quotient topology.

Let $\sim$ be an equivalence relation on some topological space X, and let $\pi : X \to Y \triangleq X \backslash \sim$ be the projection map, we can define the **quotient topology** on Y by

$$U \subseteq Y \text{ is open } \stackrel{\triangle}{\iff} \pi^{-1}(U) \subseteq X \text{ is open}$$

It is easily checked that

(a) Quotient topology is indeed a topology.

(b) Quotient topology is the largest (finest) topology on Y such that $\pi : X \to Y$ is continuous.

(c) A function $f : Y \to Z$ is continuous if and only if $f \circ \pi$ is continuous. (Universal Property)

Because the quotient map π is continuous, we know quotient preserve topological properties like connected and compact, yet notably, quotient does not preserve separation axioms. Consider the following examples.

Example 2 (Quotient does NOT preserve Second Countable)

$$X \triangleq \mathbb{R} \text{ and } x \sim y \stackrel{\triangle}{\iff} x = y \text{ or } x, y \in \mathbb{Z}$$

Let $Y \triangleq X \backslash \sim$. We show that Y is not even first countable at $[0]$. Let $U_n \subseteq Y$ be an arbitrary sequence of open-neighborhood of $[0]$. It is easily checked that for each $n \in \mathbb{N}$, there exists $\epsilon_{n,k}$ such that

$$\pi\Big[\bigcup_{k \in \mathbb{Z}} (k - \epsilon_{n,k}, k + \epsilon_{n,k})\Big] \subseteq U_n$$

Define $\delta_k \triangleq \frac{\epsilon_{k,k}}{2}$ and

$$V = \pi\Big[\bigcup_{k \in \mathbb{Z}} (k - \delta_k, k + \delta_k)\Big]$$

It is easily checked that V is an open neighborhood of $[0]$ contained in no U_n.

Example 3 (Quotient does NOT preserve Hausdorff)

$$X = \mathbb{R} \text{ and } Y = \{(-\infty, 0), [0, \infty)\}$$

However, with the criterion of π being an open mapping, we can draw some useful conclusions. For example, if $\pi : X \to Y$ is an open mapping and $\mathcal{B}$ is a basis of X, then $\pi(\mathcal{B})$ is a basis of Y.

Theorem 1.7.1. (Hausdorff and Quotient) If $\pi : X \to Y$ is an open mapping, and we define

$$R_\pi \triangleq \{(x, y) \in X^2 : \pi(x) = \pi(y)\}$$

Then

$$R_\pi \text{ is closed } \iff Y \text{ is Hausdorff}$$

Proof. Suppose R_π is closed. Fix some x, y such that $\pi(x) \neq \pi(y)$. Because R_π is closed, we know there exists open neighborhood U_x, U_y such that $U_x \times U_y \subseteq (R_\pi)^c$. It is clear that $\pi(U_x), \pi(U_y)$ are respectively open neighborhood of $\pi(x)$ and $\pi(y)$. To see $\pi(U_x)$ and $\pi(U_y)$ are disjoint, assume that $\pi(a) \in \pi(U_x) \cap \pi(U_y)$. Let $a_x \in U_x$ and $a_y \in U_y$ satisfy $\pi(a_x) = \pi(a) = \pi(a_y)$, which is impossible because $(a_x, a_y) \in (R_\pi)^c$. CaC

Suppose Y is Hausdorff. Fix some x, y such that $\pi(x) \neq \pi(y)$. Let U_x, U_y be open neighborhoods of $\pi(x), \pi(y)$ separating them. Observe that $(x, y) \in \pi^{-1}(U_x) \times \pi^{-1}(U_y) \subseteq (R_\pi)^c$ ∎

Notably, quotient topology give us a famously weird homeomorphism.

Example 4 (Weird Quotient)

$$\mathbb{R} \setminus \mathbb{Z} \simeq S^1 \triangleq \{e^{ix} \in \mathbb{C} : x \in \mathbb{R}\}$$

Clearly, we can well define a map $F : \mathbb{R} \setminus \mathbb{Z} \to S^1$ by

$$F(\pi(x)) \triangleq e^{i2\pi x}$$

It is straightforward to check F is a continuous bijection and $\mathbb{R} \setminus \mathbb{Z}$ is compact. It then follows F is a homeomorphism.

1.8 Topological Group

Abstract

This section introduce the notion of topological group and prove that quotient group of a topological group when equipped with the quotient topology is again a topological group.

By a **topological group**, we mean a topological space M equipped with a group structure such that addition $M^2 \to M$ and inversion $M \to M$ are both continuous. Equivalently, one can simply require

$$M^2 \to M; (g,h) \mapsto gh^{-1}$$

to be continuous.

Theorem 1.8.1. (Quotient Group of Topological Group)

Chapter 2

Metric Space

2.1 Completion

2.2 Bounded and Totally Bounded

2.3 Compactness

2.4 Limit Interchange

In this section, we

(a) discuss the condition in which we can change the limit order of double sequence in general metric space. (Theorem 2.4.1 and Theorem 2.4.2)

(b) prove that the space of functions is complete if and only if the codomain is complete. (Theorem 2.4.3)

(c) prove that the uniform limit of a sequence of convergent sequences in a complete metric space converge. (Theorem 2.4.5)

Remark on structure of the Theory: The proof of (Theorem 2.4.5: convergent sequences in complete metric space is closed under uniform convergence) relies on (Theorem 2.4.1: exchange limit order), while that of (Theorem 2.4.3: Space of functions (X^Y, d_∞) is complete iff Y is complete) does not.

(Theorem 2.4.1: exchange limit order) will later be used to prove the Uniform Limit Theorem (Theorem 2.5.2) which is a "pointwise" Theorem, and justify abundant of limit exchange, e.g. (Theorem 2.5.1: exchange limit order for functions)

An important consequence of (Corollary 2.4.4: space of bounded functions into complete space is complete) is that $\big(L(\mathbb{R}^n, \mathbb{R}^m), \|\cdot\|_{\text{op}}\big)$ is complete. This will be later shown with extra tools.

Theorem 2.4.1. (Change Order of Limit Operations: Part 1) Given a double sequence $a_{n,k}$ whose codomain is (Y, d). Suppose

(a) $a_{n,k} \to a_{\bullet}$, uniformly as $n \to \infty$

(b) $a_{n,k} \to A_n$ pointwise as $k \to \infty$.

(c) $A_n \to A$

We have

$$\lim_{k\to\infty} a_{\bullet,k} = A$$

In other words, we can switch the order of limit operations

$$\lim_{k\to\infty}\lim_{n\to\infty} a_{n,k} = \lim_{n\to\infty}\lim_{k\to\infty} a_{n,k}$$

Proof. We wish to prove

$$a_{\bullet,k} \to A \text{ as } k \to \infty$$

Fix ϵ. Because $a_{n,k} \to a_{\bullet,k}$ uniformly and $A_n \to A$ as $n \to \infty$, we know there exists m such that

$$d(A_m, A) < \frac{\epsilon}{3} \text{ and } \forall k \in \mathbb{N}, d(a_{m,k}, a_{\bullet,k}) < \frac{\epsilon}{3} \tag{2.1}$$

Then because $a_{m,k} \to A_m$ as $k \to \infty$, we know there exists K such that

$$\forall k > K, d(a_{m,k}, A_m) < \frac{\epsilon}{3} \tag{2.2}$$

We now claim

$$\forall k > K, d(a_{\bullet,k}, A) < \epsilon$$

The claim is true since by Equation 2.1 and Equation 2.2, we have

$$\forall k > K, d(a_{\bullet,k}, A) \leq d(a_{\bullet,k}, a_{m,k}) + d(a_{m,k}, A_m) + d(A_m, A) < \epsilon \text{ (done)}$$

■

Theorem 2.4.2. (Change Order of Limit Operations: Part 2) Given a double sequence $a_{n,k}$ whose codomain is (Y, d). Suppose

(a) $a_{n,k} \to a_{\bullet,k}$ uniformly as $n \to \infty$

(b) $a_{n,k} \to A_n$ pointwise as $k \to \infty$

(c) $a_{\bullet,k} \to A$ as $k \to \infty$

We have

$$A_n \to A$$

Proof. Fix ϵ. We wish to find N such that

$$\forall n > N, d(A_n, A) < \epsilon$$

Because $a_{n,k} \to a_{\bullet,k}$ uniformly as $n \to \infty$, we can let N satisfy

$$\forall n > N, \forall k \in \mathbb{N}, d(a_{n,k}, a_{\bullet,k}) < \frac{\epsilon}{3} \tag{2.3}$$

We claim

$$\text{such } N \text{ works}$$

Arbitrarily pick $n > N$. Because $a_{\bullet,k} \to A$, and because $a_{n,k} \to A_n$, we know there exists j such that

$$d(a_{\bullet,j}, A) < \frac{\epsilon}{3} \text{ and } d(a_{n,j}, A_n) < \frac{\epsilon}{3} \tag{2.4}$$

From Equation 2.3 and Equation 2.4, we now have

$$d(A_n, A) \leq d(A_n, a_{n,j}) + d(a_{n,j}, a_{\bullet,j}) + d(a_{\bullet,j}, A) < \epsilon \text{ (done)}$$

■

In summary of Theorem 2.4.1 and Theorem 2.4.2, given a double sequence $a_{n,k}$ converging both side

(a) $a_{n,k} \to a_{\bullet,k}$ pointwise as $n \to \infty$

(b) $a_{n,k} \to a_{n,\bullet}$ pointwise as $k \to \infty$

As long as

(a) one side of convergence is uniform

(b) between two sequence $\{a_{\bullet,k}\}_{k\in\mathbb{N}}$ and $\{a_{n,\bullet}\}_{n\in\mathbb{N}}$, one of them converge, say, to A

Then the other sequence also converge, and the limit is also A.

It is at this point, we shall introduce two other terminologies. Suppose f_n is a sequence of functions from an arbitrary set X to a metric space Y. We say f_n is **pointwise Cauchy** if for all fixed $x \in X$, the sequence $f_n(x)$ is Cauchy. We say f_n is **uniformly Cauchy** if for all ϵ, there exists $N \in \mathbb{N}$ such that

$$\forall n, m > N, \forall x \in X, d\big(f_n(x), f_m(x)\big) < \epsilon$$

In last Section (Section 2.6), we define the **uniform metric** d_∞ on X^Y by

$$d_\infty(f, g) = \sup_{x\in X} d\big(f(x), g(x)\big)$$

and say that $f_n \to f$ uniformly if and only if $f_n \to f$ in (X^Y, d_∞). Similar to this clear fact, we have

$$f_n \text{ is uniformly Cauchy} \iff f_n \text{ is Cauchy in } (X^Y, d_\infty)$$

It should be very easy to verify that if f_n uniformly converge, then f_n is uniformly Cauchy, and just like sequences in metric space, the converse hold true if and only if the space $\big(X^Y, d_\infty\big)$ is complete. In Theorem 2.4.3, we give a necessary and sufficient condition for $\big(X^Y, d_\infty\big)$ to be complete.

Theorem 2.4.3. (Space of functions $\big(X^Y, d_\infty\big)$ is Complete iff Y is Complete) Given an arbitrary set X and a metric space (Y, d), we have

$$\text{the extended metric space } \big(X^Y, d_\infty\big) \text{ is complete} \iff Y \text{ is complete}$$

Proof. $(\longleftarrow)$

Suppose f_n is uniformly Cauchy. We wish

$$\text{to construct a } f : X \to Y \text{ such that } f_n \to f \text{ uniformly}$$

Because f_n is uniformly Cauchy, we know that for all $x \in X$, the sequence $f_n(x)$ is Cauchy in (Y, d). Then because Y is complete, we can define $f : X \to Y$ by

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

We claim

$$\text{such } f \text{ works, i.e. } f_n \to f \text{ uniformly}$$

Fix ϵ. We wish

$$\text{to find } N \in \mathbb{N} \text{ such that for all } n > N \text{ and } x \in X \text{ we have } d\big(f_n(x), f(x)\big) < \epsilon$$

Because f_n is uniformly Cauchy, we know there exists N such that

$$\forall n, m > N, \forall x \in X, d\big(f_n(x), f_m(x)\big) < \frac{\epsilon}{2} \tag{2.5}$$

We claim

$$\text{such } N \text{ works}$$

Assume there exists $n > N$ and $x \in X$ such that $d\big(f_n(x), f(x)\big) \geq \epsilon$. Because $f_k(x) \to f(x)$ as $k \to \infty$, we know

$$\exists m \in \mathbb{N}, d\big(f_m(x), f(x)\big) < \frac{\epsilon}{2} \tag{2.6}$$

Then from Equation 2.5 and Equation 2.6, we can deduce

$$\epsilon \leq d\big(f_n(x), f(x)\big) \leq d\big(f(x), f_m(x)\big) + d\big(f_n(x), f_m(x)\big) < \epsilon \text{ CaC } \text{ (done)}$$

$(\longrightarrow)$

Let K be the set of constant functions in X^Y. We first prove

$$K \text{ is closed}$$

Arbitrarily pick $f \in K^c$. We wish

$$\text{to find } \epsilon \in \mathbb{R}^+ \text{ such that } B_\epsilon(f) \in K^c$$

Because f is not a constant function, we know there exists $x_1, x_2 \in X$ such that

$$d\big(f(x_1), f(x_2)\big) > 0$$

We claim that

$$\epsilon = \frac{d\big(f(x_1), f(x_2)\big)}{3} \text{ works}$$

Arbitrarily pick $g \in B_\epsilon(f)$. We wish

$$\text{to show } g \in K^c$$

Notice the triangle inequality

$$3\epsilon = d\big(f(x_1), f(x_2)\big) \leq d\big(f(x_1), g(x_1)\big) + d\big(g(x_1), g(x_2)\big) + d\big(g(x_2), f(x_2)\big) \tag{2.7}$$

Also, because $g \in B_\epsilon(f)$, we have

$$\forall x \in X, d\big(f(x), g(x)\big) < \epsilon \tag{2.8}$$

Then by Equation 2.7 and Equation 2.8, we see

$$d\big(g(x_1), g(x_2)\big) > \epsilon$$

This then implies g is not a constant function. (done)

Now, Because by premise (X^Y, d_∞) is complete, and we have proved K is closed in (X^Y, d_∞), we know K is complete. Then, we resolve the whole problem into proving

$$Y \text{ is isometric to } K$$

Define $\sigma : Y \to K$ by

$$y \mapsto \tilde{y} \text{ where } \forall x \in X, \tilde{y}(x) = y$$

It is easy to verify σ is an isometry. (done) ■

Corollary 2.4.4. (Space of Bounded functions $\big(B(X,Y), d_\infty\big)$ is Complete iff Y is Complete)

$$\big(B(X,Y), d_\infty\big) \text{ is complete } \iff Y \text{ is complete}$$

Proof. $(\longleftarrow)$

By Theorem 2.4.3, the space (X^Y, d_∞) is complete. Then because $B(X,Y)$ is closed in (X^Y, d_∞), we know $B(X,Y)$ is complete.

$(\longrightarrow)$

Notice that the set of constant function K is a subset of the galaxy $B(X,Y)$. The whole proof in Theorem 2.4.3 works in here too. ■

Remember in the beginning of this section we say we will prove convergent sequences in Y is closed under uniform convergence if Y is complete. The proof of this result relies on Theorem 2.4.3.

Theorem 2.4.5. (Convergent Sequences are Closed under Uniform Convergence if Codomain $\big(Y, d\big)$ is Complete) Given a complete metric space $\big(Y, d\big)$, let $\mathcal{C}_{\mathbb{N}}^{Y}$ be the set of convergent sequences in Y.

$$Y \text{ is complete} \implies \mathcal{C}_{\mathbb{N}}^{Y} \text{ is closed under uniform convergent}$$

Proof. Let $a_{n,k} \to a_{\bullet,k}$ uniformly as $n \to \infty$ where for all $n, k \in \mathbb{N}, a_{n,k} \in Y$ and let $A_n = \lim_{k\to\infty} a_{n,k}$ for all $n \in \mathbb{N}$.

$$\text{to prove } a_{\bullet,k} \text{ converge}$$

By Theorem 2.4.2, we can reduce the problem to

$$\text{proving } A_n \text{ converge}$$

Then because Y is complete, we can then reduce the problem into proving

$$A_n \text{ is Cauchy}$$

Fix ϵ. We wish to find N such that

$$\forall n, m > N, d(A_n, A_m) < \epsilon$$

Because $a_{n,k} \to a_{\bullet,k}$ uniformly, we can find N such that

$$\forall n, m > N, d_\infty(\{a_{n,k}\}_{k\in\mathbb{N}}, \{a_{m,k}\}_{k\in\mathbb{N}}) < \frac{\epsilon}{3} \tag{2.9}$$

We claim

$$\text{such } N \text{ works}$$

Arbitrarily pick $n, m > N$. We wish to prove

$$d(A_n, A_m) < \epsilon$$

Because $a_{n,k} \to A_n$ and $a_{m,k} \to A_m$ as $k \to \infty$, we can find j such that

$$d(a_{n,j}, A_n) < \frac{\epsilon}{3} \text{ and } d(a_{m,j}, A_m) < \frac{\epsilon}{3} \tag{2.10}$$

Then from Equation 2.9 and Equation 2.10, we can deduce

$$d(A_n, A_m) \le d(A_n, a_{n,j}) + d(a_{n,j}, a_{m,j}) + d(a_{m,j}, A_m) < \epsilon \text{ (done)}$$

■

2.5 Closed under Uniform Convergence

Given $(E, d_E), (Y, d_Y)$ and a sequence of functions $f_n : E \to Y$, converging uniformly to some $f : E \to Y$ such that each f_n has the property

(a) Boundedness

(b) Unboundedness

(c) Continuity

(d) Uniform continuity

(e) K-Lipschitz continuity

on E, then f also has the same property. These fact will later be proved in **??** and **??**. are again all closed under uniform convergence, where the proof for continuity is closed under uniform convergence use Theorem 2.4.1 as a lemma.

The reason we require the co-domain Y of sequence to be complete is explained in the last paragraph of Section 2.6. An example of such beautiful closure is lost if the codmain (Y, d) is not complete is $Y = \mathbb{R}^*$ and $a_{n,k} = \frac{1}{n} + \frac{1}{k}$.

Theorem 2.5.1. (Change Order of Limit Operation in Complete Metric Space) Given a sequence of function $f_n : E \to (Y, d)$ and a function $f : E \to (Y, d)$ such that

(a) $f_n \to f$ uniformly on E

(b) $\lim_{t \to x} f_n(t)$ exists for all $n \in \mathbb{N}$

(c) (Y, d) is complete

We have

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

Proof. Fix a sequence t_k in E that converge to x. We reduced the problem into proving

$$\lim_{n\to\infty} \lim_{k\to\infty} f_n(t_k) = \lim_{k\to\infty} \lim_{n\to\infty} f_n(t_k)$$

Set

$$a_{n,k} \triangleq f_n(t_k) \tag{2.11}$$

We then reduced the problem into proving

$$\lim_{n\to\infty} \lim_{k\to\infty} a_{n,k} = \lim_{k\to\infty} \lim_{n\to\infty} a_{n,k}$$

Set

$$A_n \triangleq \lim_{t\to x} f_n(t) \text{ and } a_{\bullet,k} \triangleq \lim_{n\to\infty} f_n(t_k)$$

We now prove

A_n converge

Fix ϵ. We wish

to find N such that $d(A_n, A_m) \leq \epsilon$ for all $n, m > N$

Because $a_{n,k}$ uniformly converge (to $a_{\bullet,k}$) as $n \to \infty$ by our setting, we know there exists N such that

$$\forall n, m > N, \forall k \in \mathbb{N}, d(a_{n,k}, a_{m,k}) < \frac{\epsilon}{3}$$

We claim

such N works

Fix $n, m > N$. Because $a_{n,k} \to A_n$ and $a_{m,k} \to A_m$, we know there exists $j \in \mathbb{N}$ such that

$$d(a_{n,j}, A_n) < \frac{\epsilon}{3} \text{ and } d(a_{m,j}, A_m) < \frac{\epsilon}{3}$$

We now have

$$\begin{aligned} d(A_n, A_m) &\leq d(A_n, a_{n,j}) + d(a_{n,j}, a_{m,j}) + d(a_{m,j}, A_m) \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \text{ (done)} \end{aligned}$$

Now, because $a_{n,k} \to a_{\bullet,k}$ uniformly, by Theorem 2.4.2, we have

$$\lim_{n\to\infty} \lim_{k\to\infty} a_{n,k} = \lim_{n\to\infty} A_n = \lim_{k\to\infty} a_{\bullet,k} = \lim_{k\to\infty} \lim_{n\to\infty} a_{n,k} \text{ (done)}$$

■

The end goal for this section is to prove that the following properties

(a) continuity

(b) uniform continuity

(c) K-Lipschitz Continuity

Theorem 2.5.2. (Uniform Limit Theorem) Given a sequence of function f_n from a topological space (X, τ) to a metric space (Y, d), suppose

(a) $f_n \to f$ uniformly as $n \to \infty$

(b) f_n is continuous for all $n \in \mathbb{N}$

Then f is also continuous.

Proof. Fix $x \in X$, and let $x_k \to x$. We wish to prove

$$f(x_k) \to f(x)$$

Because $f_n \to f$ uniformly as $n \to \infty$, we know

$$\Big\{f_n(x_k)\Big\}_{k\in\mathbb{N}} \to \Big\{f(x_k)\Big\}_{k\in\mathbb{N}} \text{ uniformly as } n \to \infty \tag{2.12}$$

Also, because for each $n \in \mathbb{N}$, the function f_n is continuous at x, we know

$$\forall n \in \mathbb{N}, f_n(x_k) \to f_n(x) \text{ as } k \to \infty \tag{2.13}$$

Then because $f_n \to f$ pointwise, we know

$$f_n(x) \to f(x) \tag{2.14}$$

Now, because Equation 2.12, Equation 2.13 and Equation 2.14, by Theorem 2.4.1, we have

$$\lim_{k\to\infty} f(x_k) = \lim_{k\to\infty}\lim_{n\to\infty} f_n(x_k) = \lim_{n\to\infty}\lim_{k\to\infty} f_n(x_k) = \lim_{n\to\infty} f_n(x) = f(x) \text{ (done)}$$

∎

Suppose X is a compact Hausdroff space, with Theorem **??**, we can now say that the set $\mathcal{C}(X)$ of complex-valued continuous functions on X

Theorem 2.5.3. (Uniformly Continuous functions are Closed under Uniform Convergence) Given a sequence of functions f_n from a metric space (X, d_X) to metric space (Y, d_Y), suppose

(a) $f_n \to f$ uniformly

(b) f_n is uniformly continuous for all $n \in \mathbb{N}$

Then f is also uniformly continuous

Proof. Fix ϵ. We wish

$$\text{to find } \delta \text{ such that } \forall x, y \in X, d_X(x, y) < \delta \implies d_Y\big(f(x), f(y)\big) < \epsilon$$

Because $f_n \to f$ uniformly, we know there exists $m \in \mathbb{N}$ such that

$$\forall x \in X, d_Y\big(f_m(x), f(x)\big) < \frac{\epsilon}{3} \tag{2.15}$$

Because f_m is uniformly continuous, we know

$$\exists \delta, \forall x, y \in X, d_X(x,y) < \delta \implies d_Y\big(f_m(x), f_m(y)\big) < \frac{\epsilon}{3} \tag{2.16}$$

We claim

$$\text{such } \delta \text{ works}$$

Let $x, y \in X$ satisfy $d_X(x,y) < \delta$. We wish

$$\text{to prove } d_Y\big(f(x), f(y)\big) < \epsilon$$

From Equation 2.15 and Equation 2.16, we have

$$d_Y\big(f(x), f(y)\big) \leq d_Y\big(f(x), f_m(x)\big) + d_Y\big(f_m(x), f_m(y)\big) + d_Y\big(f_m(y), f(y)\big) = \epsilon \text{ (done)}$$

■

Theorem 2.5.4. (K-Lipschitz functions are Closed under Uniform Convergence) Given a sequence of functions f_n from metric space (X, d_X) to metric space (Y, d_Y), suppose

(a) $f_n \to f$ uniformly as $n \to \infty$

(b) f_n is K-Lipschtize continuous for all $n \in \mathbb{N}$

Then f is also K-Lipschtize continuous.

Proof. Arbitrarily pick $x, y \in X$, to show f is K-Lipschtize continuous, we wish

$$\text{to show } d_Y\big(f(x), f(y)\big) \leq K d_X(x,y)$$

Fix ϵ. We reduce the problem into proving

$$d_Y\big(f(x), f(y)\big) < K d_X(x,y) + \epsilon$$

Because $f_n \to f$ uniformly as $n \to \infty$, we know there exists m such that

$$\forall z \in X, d_Y\big(f(z), f_m(z)\big) < \frac{\epsilon}{2} \tag{2.17}$$

Because f_m is K-Lispchitz continuous, we know

$$d_Y\big(f_m(x), f_m(y)\big) \leq K d_X(x,y) \tag{2.18}$$

Now, from Equation 2.18 and Equation 2.17, we now see

$$d_Y\big(f(x), f(y)\big) \leq d_Y\big(f(x), f_m(x)\big) + d_Y\big(f_m(x), f_m(y)\big) + d_Y\big(f_m(y), f(y)\big) < K d_X(x,y) + \epsilon$$

■

An example of sequences of Lipschitz continuous functions with unbounded Lipschitz constant can uniformly converge to a non-Lipschitz continuous function is given below

Example 5 (Lipschitz functions with Unbounded Lipschitz constant Uniformly Converge to a non-Lipschitz function)

$$X = [0,1] \text{ and } f_n(x) = \sqrt{x + \frac{1}{n}}$$

2.6 Modes of Convergence

This section is the starting point for us to study spaces of function. At first, we will define two modes of convergence for sequence of function and point out some basic properties and the difference between two modes of convergence.

Given an arbitrary set X and a metric space Y, we say a sequence of functions f_n from X to Y **pointwise converge** to f if for all ϵ and x in X, there exists N such that

$$\forall n > N, f_n(x) \in B_\epsilon\big(f(x)\big)$$

In other words, for each fixed x in X, we have $f_n(x) \to f(x)$.

We say f_n **uniformly converge** to f if for all ϵ there exists N such that

$$\forall x \in X, \forall n > N, f_n(x) \in B_\epsilon\big(f(x)\big)$$

The difference between pointwise convergence and uniform convergence is that if we require $f_n(x)$ to be ϵ-close to $f(x)$ for all $n > N$, then

(•) N depend on both ϵ and x if $f_n \to f$ pointwise

(•) N depend on only ϵ if $f_n \to f$ uniformly

A few properties of sequence of functions similar to that of sequences in metric space is obvious. If $f_n \to f$ pointwise, then all sub-sequences $f_{n_k} \to f$ pointwise. If $f_n \to f$ uniformly, then all sub-sequences $f_{n_k} \to f$ uniformly. Suppose $Z \subseteq X$. It is clear that if $f_n \to f$ uniformly (resp: pointwise) the restricts $f_n|_Z \to f|_Z$ uniformly (resp: pointwise). Also, if $f_n \to f$ uniformly, then $f_n \to f$ pointwise.

Suppose we have a family $\mathcal{F}$ of functions $f : X \to (Y, d)$. If we define

$$d_\infty(f, g) = \sup_{x\in X} d\big(f(x), g(x)\big)$$

instead of a metric, d_∞ become an extended metric. If f is bounded and g is unbouneded, we have $d_\infty(f, g) = \infty$. If f, g are both bounded, then $d_\infty(f, g) \in \mathbb{R}^+$. Because of such, for d_∞ to be a metric, one but not the only condition is for $\mathcal{F}$ to be space of bounded functions.

Now, regardless of d_∞ is an extended metric or not, we have

$$f_n \to f \text{ uniformly} \iff d_\infty(f_n, f) \to 0$$

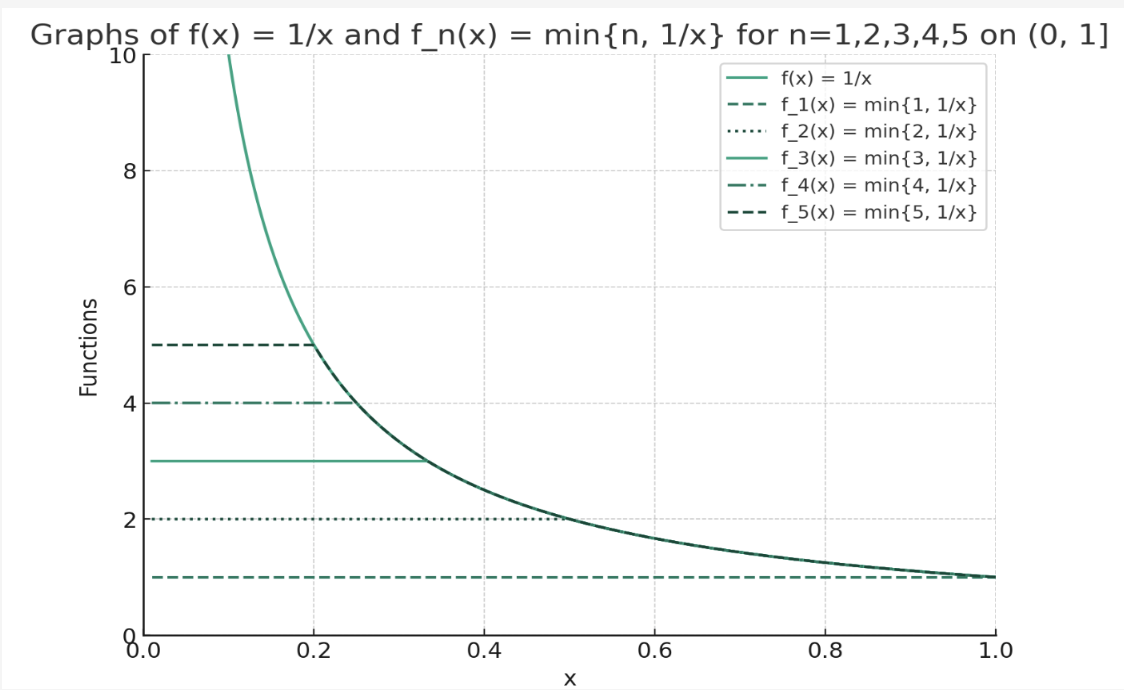
With this in mind, it shall be clear that the uniform limit of bounded (resp: unbounded) functions is always bounded (resp: unbounded).

Examples for bounded (resp: unbounded) function f_n pointwise converge to unbounded (resp: bounded) function f are as follows.

Example 6 (Bounded functions pointwise converge to unbounded function)

$$X = (0, 1], f_n(x) = \min\{n, \frac{1}{x}\}$$

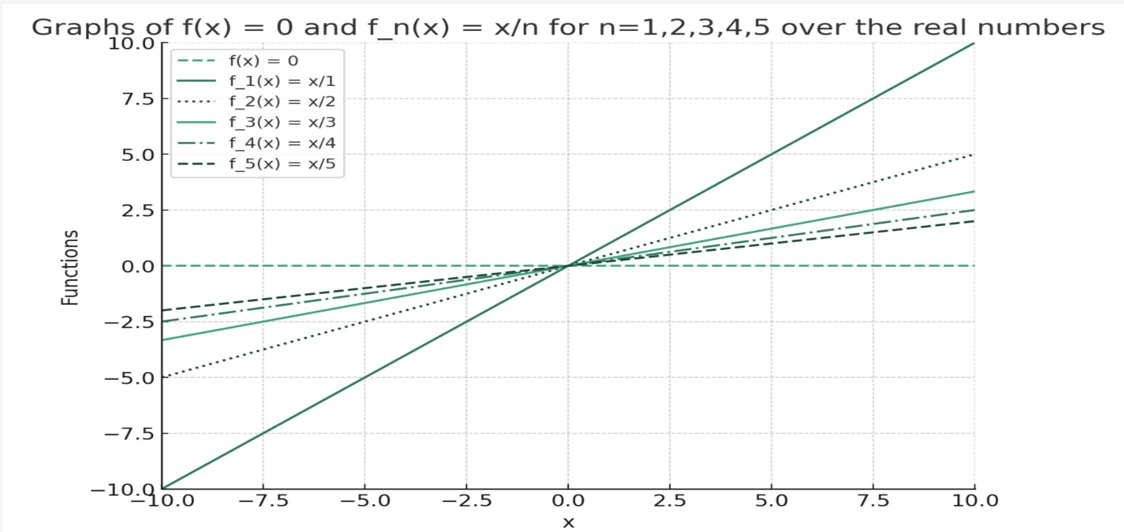
It is clear that $\forall n \in \mathbb{N}, f_n(x) \in [0, n]$, and the limit $f : X \to \mathbb{R}$ is $f(x) = \frac{1}{x}$



Example 7 (Unbounded functions pointwise converge to bounded function)

$$X = \mathbb{R}, f_n(x) = \frac{1}{n}x$$

The limit function is $f(x) = 0$



As pointed out earlier, if $f : X \to (Y, d)$ is bounded and $g : X \to (Y, d)$ is unbounded, then $d_\infty(f, g) = \infty$. This means that if Y is unbounded, the uniform metric d_∞ is extended on X^Y. For this, it is necessary to develop some basic fact concerning extended metric space.

Suppose (X, d) is an extended metric space. If we define $\sim$ on X by $x \sim y \iff d(x, y) < \infty$, then $\sim$ is an equivalence relation. We say each equivalence class is a **galaxy** of (X, d). Suppose T is the collection of the galaxies of (X, d). For each $\mathcal{T} \in T$, the space $(\mathcal{T}, d)$ is just a metric space.

It is easy to see that the way we induce topology from metric space is still valid if the metric is extended. That is

$$\tau = \{Z \in X : \forall z \in Z, \exists \epsilon, B_\epsilon(z) \subseteq Z\}$$

is still a topology, even though d is an extended metric on X.

We can verify that a set Y in X is open if and only if for all $\mathcal{T} \in T$, the set $Y \cap \mathcal{T}$ is open, and the set Y in X is closed if and only if all convergent sequences y_n in Y

converge to points in Y.

Now, suppose we are given an arbitrary set X and a complete metric space $(\overline{Y}, d)$, and on $X^{\overline{Y}}$, we define the uniform metric d_∞. We say a set $\mathcal{F} \subseteq X^{\overline{Y}}$ of functions is **closed under uniform convergence** if for all uniform convergent sequence $f_n \subseteq \mathcal{F}$, the limit function f is also in $\mathcal{F}$. There are justified reasons for us to give the premise that $\overline{Y}$ is complete prior to the definition of the term **closed under uniform convergence**. One reason is that by Theorem 2.4.3, if Y is not complete, then the extended metric space (X^Y, d_∞) is also not complete, which implies the possibility a Cauchy sequence f_n in X^Y converge to a function $f \in X^{\overline{Y}} \setminus X^Y$ where $\overline{Y}$ is the completion of Y. For instance, if we let $Y = \mathbb{R} \setminus \{1\}$ where $X = \mathbb{R}$, and let $f_n(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ 1 + \frac{1}{n} & \text{if } x = 0 \end{cases} \in Y$, we see that the set $\mathcal{F} = \{f_n : n \in \mathbb{N}\}$ is "closed under uniform convergence" in the context of X^Y, but when in fact f_n uniformly converge to $f(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$ which is not in $\mathcal{F}$. This awkward usage of words can be solved if we define the term **closed under uniform convergence** after the premise that Y is complete.

Now, given a set of functions $\mathcal{F} \subseteq X^{\overline{Y}}$, one can verify that

$$\begin{aligned} \mathcal{F} \text{ is closed under uniform convergence } &\iff (\mathcal{F}, d_\infty) \text{ is complete} \\ &\iff \mathcal{F} \text{ is closed with respect to } (X^{\overline{Y}}, d_\infty) \end{aligned}$$

Let $\mathcal{G}$ be a galaxy of $(X^{\overline{Y}}, d_\infty)$. With multiple ways, we can verify that $\mathcal{G}$ is closed with respect to $(X^{\overline{Y}}, d_\infty)$. Then, acknowledging the space of bounded functions $B(X, \overline{Y})$ is a galaxy of $X^{\overline{Y}}$, we see that $B(X, \overline{Y})$ is closed under uniform convergence. The statement that $B(X, \overline{Y})$ is closed under uniform convergence, although already "proved" before as we pointed out the limit of uniform convergent sequence of bounded functions must be bounded, is now in fact actually proved in the sense the term "closed under uniform convergence" is formally given a satisfying definition.

2.7 Arzelà–Ascoli Theorem

In this section, we will give a complete proof of Arzelà–Ascoli Theorem for functions from arbitrary compact topological space to arbitrary metric space. Note that in Baby Rudin, Arzelà–Ascoli Theorem are given for functions from compact metric space to metric space. Because Arzelà–Ascoli Theorem are concerned with family of equicontinuos functions, it is crucial for us to give a definition to equicontinuity for functions from topological space to metric space, for the sake of our generalization.

Let X,Y be metric space. Let Z be topological space. Let $\mathcal{F}_X$ be family of functions from X to Y, and let $\mathcal{F}_Z$ be family of functions from Z to Y. We say $\mathcal{F}_Z$ is **pointwise equicontinuous** if

$$\text{For all } \epsilon \text{ and for all } x\text{, there exists a neighborhood } U_x \text{ such that}$$
$$d_Y(f(x),f(y)) < \epsilon \text{ for all } y \in U_x$$

We say $\mathcal{F}_X$ is **equicontinuous** if

$$\text{For all } \epsilon\text{, there exists } \delta \text{ such that } d_Y(f(x),f(y)) < \epsilon \text{ for all } \delta\text{-close } x,y \in X$$
$$\text{and all } f \in \mathcal{F}.$$

It is easy to verify that if $\mathcal{F}_X$ is equicontinuous, then $\mathcal{F}_X$ is pointwise equicontinuous. The converse don't always hold true. Say, $\mathcal{F} = \{n+x^2\}_{n\in\mathbb{N}}$, the set $\{n+x^2\}_{n\in\mathbb{N}}$ is clearly pointwise equicontinuous on $\mathbb{R}$, and is not equicontinuous on $\mathbb{R}$, since no function $n+x^2$ is uniform continuous on $\mathbb{R}$. However, the same set $\mathcal{F} = \{n+x^2\}$ is equicontinous on compact domain $[a,b]$. This is a general result, as we shall prove below.

Theorem 2.7.1. (Pointwise Equicontinous is Uniform on Compact Domain) Given two metric space $(X,d_X),(Y,d_Y)$, and a family $\mathcal{F}$ of functions from X to Y such that

(a) X is compact

(b) $\mathcal{F}$ is pointwise equicontinuous

Then

$$\mathcal{F} \text{ is equicontinuous}$$

Proof. Fix ϵ. We wish to

$$\text{find } \delta \text{ such that } d_X(x,y) < \delta \implies d_Y(f(x),f(y)) \leq \epsilon \text{ for all } f \in \mathcal{F}$$

Because $\mathcal{F}$ is pointwise equicontinuous, we know for each $x \in X$, there exists δ_x such that

$$\forall y \in B_{\delta_x}(x), d_Y\big(f(x),f(y)\big) < \frac{\epsilon}{2} \text{ for all } f \in \mathcal{F} \tag{2.19}$$

It is clear that $\{B_{\frac{\delta_x}{2}}(x) : x \in X\}$ form an open cover of X. Then because X is compact, we know

$$\text{there exists a finite open sub-cover: } \{B_{\frac{\delta_x}{2}}(x) : x \in X_{\text{finite}}\}$$

We claim

$$\delta = \min_{x \in X_{\text{finite}}} \frac{\delta_x}{2} \text{ works}$$

Fix $y, z \in X : d_X(y, z) < \delta$. We have to prove

$$d_Y\big(f(y), f(z)\big) < \epsilon$$

We know y must lie in some $B_{\frac{\delta_x}{2}}(x)$ for some $x \in X_{\text{finite}}$. Because $d_X(y, z) < \frac{\delta_x}{2}$, we see that z must lie in $B_{\delta_x}(x)$. We now know y, z are both in $B_{\delta_x}(x)$. Then from (2.19), we can now deduce

$$d_Y\big(f(y), f(z)\big) \leq d_Y\big(f(y), f(x)\big) + d_Y\big(f(x), f(z)\big) < \epsilon \text{ (done)}$$

■

The proof above should be a great example why in the discussion of metric space, instead of using sequential definition of compactness, which leads to the beautiful Bolzano-Weierstrass Theorem, some people prefer the open-cover definitions.

Now, we give proof for the Arzelà–Ascoli Theorem.

Theorem 2.7.2. (Arzelà–Ascoli Theorem) Given a compact topological space (X, τ), a metric space (Y, d_Y), and a family $\mathcal{F} \subseteq C\big(X, Y\big)$ of continuous function

$\mathcal{F}$ is pointwise equicontinuous and $\{f(x) : f \in \mathcal{F}\}$ has compact closure in Y for all $x \in X$
$\implies \mathcal{F}$ has a compact closure in $C(X, Y)$

Proof. Fix a sequence f_n in $\mathcal{F}$. We wish to show

$$f_n \text{ has a sub-sequence } f_{n_k} \text{ uniformly converge to some } f : X \to Y$$

First, we prove

$$\text{there exists a countable set } P \text{ such that } P \text{ works like a dense set}$$

Because $\mathcal{F}$ is pointwise equicontinuous, we know for all $x \in X$

$$\exists U_{x,n}, \forall y \in U_{x,n}, \forall f \in \mathcal{F}, d_Y\big(f(x), f(y)\big) < \frac{1}{n} \text{ for each fixed } n \in \mathbb{N}$$

Now, because X is compact, for each $n \in \mathbb{N}$, there exists a finite subset $P_n \subseteq X$ such that $\left\{U_{x,n} : x \in P_n\right\}$ is a cover of X. Let $P = \bigcup_{n\in\mathbb{N}} P_n$. (done)

Now, we wish to

$$\text{construct a sub-sequence } f_{n_k} \text{ pointwise converge on } P$$

Express $P = \{p_k\}_{k\in\mathbb{N}}$. By premise (pointwise image has compact closure), we know there exists a compact set that contain $\{f_n(p_1)\}_{n\in\mathbb{N}}$, so by Bolzano-Weierstrass Theorem, there exists a sub-sequence

$$\left\{f_{g_1(k)}(p_1)\right\}_{k\in\mathbb{N}} \text{ converge to some point in } Y$$

Now, again by premise and Bolzano-Weierstrass Theorem, there exists a sub-sequence

$$\left\{f_{g_2\circ g_1(k)}(p_2)\right\}_{k\in\mathbb{N}} \text{ converge to some point in } Y$$

Repeatedly doing such, we have

$$\begin{array}{cccc}
f_{g_1(1)}(p_1) & f_{g_2\circ g_1(1)}(p_2) & f_{g_3\circ g_2\circ g_1(1)}(p_3) & \cdots \\
f_{g_1(2)}(p_1) & f_{g_2\circ g_1(2)}(p_2) & f_{g_3\circ g_2\circ g_1(2)}(p_3) & \cdots \\
f_{g_1(3)}(p_1) & f_{g_2\circ g_1(3)}(p_2) & f_{g_3\circ g_2\circ g_1(3)}(p_3) & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
\downarrow & \downarrow & \downarrow & \\
y_1 & y_2 & y_3 & \cdots
\end{array}$$

Now, let

$$n_k = g_k \circ \cdots \circ g_1(k)$$

Then

$$n_k \text{ is eventually a sub-sequence of } g_m \circ \cdots \circ g_1(k) \text{ for all } m$$

This then implies

$$f_{n_k}(p_m) \to y_m \text{ for all } p_m \in P \text{ (done)}$$

Next, we show

$$\text{To prove } f_{n_k} \text{ uniformly converge on } X\text{, it suffice to prove } f_{n_k} \text{ is uniformly Cauchy on } X.$$

By premise (pointwise image has compact closure), if f_{n_k} is uniformly Cauchy, then we know f_{n_k} pointwise converge to some f.

Fix ϵ. We reduced the problem into

$$\text{finding } N \text{ such that for all } k > N, \text{ we have } d_Y\big(f_{n_k}(x), f(x)\big) \leq \epsilon \text{ for all } x \in X$$

Because f_{n_k} is uniformly Cauchy, we know there exists N such that for all $m, k > M$ $d_Y\big(f_{n_k}(x), f_{n_m}(x)\big) \leq \frac{\epsilon}{2}$ for all $x \in X$. We claim

$$\text{such } N \text{ works}$$

Let $k > N$. Assume $d_Y\big(f_{n_k}(x), f(x)\big) > \epsilon$. We see that

$$d_Y\big(f(x), f_{n_m}(x)\big) \geq d_Y\big(f(x), f_{n_k}(x)\big) - d_Y\big(f_{n_k}(x), f_{n_m}(x)\big) > \frac{\epsilon}{2} \text{ for all } m > N \text{ CaC } \text{ (done)}$$

Lastly, we wish to prove

$$f_{n_k} \text{ is uniformly Cauchy}$$

Fix ϵ. We wish

$$\text{to find } N \text{ such that } \forall j, k > N, \forall x \in X, d_Y\big(f_{n_j}(x), f_{n_k}(x)\big) \leq \epsilon$$

Fix $m > \frac{3}{\epsilon}$. Express $P_m = \{p_1^m, \dots, p_u^m\}$. Because $f_{n_k}(p_t^m)$ converge for each $t \in \{1, \dots, u\}$, we know

$$\forall t, \exists N_t, d_Y\big(f_{n_j}(p_t^m), f_{n_k}(p_t^m)\big) < \frac{\epsilon}{3} \text{ for all } j, k > N_t$$

We claim

$$N = \max_t N_t \text{ works}$$

Fix $j, k > N$ and $x \in X$. We have to show

$$d_Y\big(f_{n_j}(x), f_{n_k}(x)\big) \leq \epsilon$$

Because $\{U_{p_t^m, m}\}$ form an open cover of X, we know there exists t such that $x \in U_{p_t^m, m}$. We can now deduce

$$d_Y\big(f_{n_j}(x), f_{n_k}(x)\big) \leq d_Y\big(f_{n_j}(x), f_{n_j}(p_t^m)\big) + d_Y\big(f_{n_j}(p_t^m), f_{n_k}(p_t^m)\big) + d_Y\big(f_{n_k}(p_t^m), f_{n_k}(x)\big) < \epsilon$$

(done)

$\blacksquare$

2.8 Banach Fixed Point Theorem

This section give a complete statement and proof of Banach Fixed Point Theorem. The setting is

(a) a metric space $\Big(X, d_X\Big)$

(b) a subset $E \subseteq X$

(c) another metric space $\Big(Y, d_Y\Big)$

(d) a function $f : E \to Y$

(e) another function $g : E \to X$

We say f is a **contraction** on E if there exists $r \in [0,1)$ such that

$$d_Y(f(x), f(y)) \leq r d_X(x,y) \qquad (x, y \in E)$$

or equivalently

$$\sup_{x \neq y \in E} \frac{d_Y(f(x), f(y))}{d_X(x,y)} < 1$$

Note that the restriction of a contraction is again a contraction. We say g admits a **fixed point** x if we have

$$g(x) = x$$

Theorem 2.8.1. (Banach Fixed Point Theorem) If g is a contraction that maps E into X, then

$$g \text{ admits at most one fixed point}$$

Moreover, if E is complete and $g(E) \subseteq E$, then

$$\text{the fixed point exists}$$

And if we use the notation g^n to denote $g \circ g^{n-1}$, then for all $x \in E$,

$$\text{the fixed point can be written in the form } \lim_{n \to \infty} g^n(x)$$

Proof. We first

prove the uniqueness of the fixed point

Suppose x, y are both fixed by g. We have

$$d(g(x), g(y)) = d(x, y)$$

Because g is a contraction mapping, this implies $d(x, y) = 0$. (done)

Suppose E is complete and $g(E) \subseteq E$. We now

prove the existence of the fixed point

Fix $x \in E$. Because we have already prove the uniqueness of the fixed point, we only have to prove

$$\lim_{n\to\infty} g^n(x) \text{ exists and } \lim_{n\to\infty} g^n(x) \text{ is a fixed point of } g$$

Because E is complete, to prove $\lim_{n\to\infty} g^n(x)$ exists, we only have to prove

$$\{g^n(x)\}_{n\in\mathbb{N}} \text{ is Cauchy}$$

Observe

$$\begin{aligned}
d(g^n(x), g^{n+k}(x)) &\leq \sum_{i=0}^{k-1} d(g^{n+i}(x), g^{n+i+1}(x)) \\
&\leq d(x, g(x)) \sum_{i=0}^{k-1} r^{n+i} \\
&\leq \frac{r^n}{1-r} d(x, g(x)) \to 0 \text{ as } n \to \infty \quad \text{(done)}
\end{aligned}$$

Note that contraction is Lipschitz thus continuous, and note that $\lim_{n\to\infty} g^n(x) \in E$. This allow us to carry the below limit process

$$g\big(\lim_{n\to\infty} g^n(x)\big) = \lim_{n\to\infty} g(g^n(x)) = \lim_{n\to\infty} g^{n+1}(x) = \lim_{n\to\infty} g^n(x) \text{ (done)}$$

■

Banach Fixed Point Theorem is one of the most important Theorem in Mathematics. It will be used to prove

(a) Inverse Function Theorem

(b) Picard-Lindelof Theorem

(c) Nash-Embedding Theorem

Chapter 3

Algebraic Topology

3.1 Fundamental Group

3.2 Invariance of Domain

Theorem 3.2.1. (Invariance of Domain) If U is an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^n$ is one-to-one and continuous, then

$$f(U) \text{ is open and } f \text{ is a homeomorphism between } U \text{ and } f(U)$$

Theorem 3.2.2. (Invariance of Dimension) If U is a non-empty open subset of $\mathbb{R}^n$ and V is a non-empty subset of $\mathbb{R}^m$ homeomorphic to U, then $n = m$.

Chapter 4

Differential Calculus

4.1 Operator Norm

Abstract

This section introduces the concept of the operator norm and proves some fundamental results related operator norm and finite-dimensional normed spaces. For example, we establish results such as a linear operator being bounded if and only if it is continuous and the equivalence of all norms on finite-dimensional vector spaces.

In this section, and particularly in functional analysis, we say a function T between two metric space is a **bounded operator** if T always map bounded set to bounded set. In particular, if T is a linear transformation between two normed space, we say T is a **bounded linear operator**. Now, suppose $\mathcal{X},\mathcal{Y}$ are two normed space over $\mathbb{R}$ or $\mathbb{C}$. In space $L(\mathcal{X},\mathcal{Y})$, alternatively, we can define

$$T \text{ is bounded} \quad \overset{\triangle}{\iff} \exists M \in \mathbb{R}, \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|$$

The proof of equivalency is simple. For $(\longrightarrow)$, observe

$$\|Tx\| = \|x\| \cdot \|T\frac{x}{\|x\|}\| \leq \Big(\sup\{\|Ty\| : \|y\| = 1\}\Big)\|x\|$$

For $(\longleftarrow)$, observe

$$\|Tx - Ty\| = \|T(x-y)\| \leq M\|x-y\|$$

We first show that a linear transformation is continuous if and only if it is bounded.

Theorem 4.1.1. (Liner Operator is Bounded if and only if it is Continuous) Given two normed space $\mathcal{X}, \mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$ and $T \in L(\mathcal{X}, \mathcal{Y})$, we have

$$T \text{ is a bounded operator } \iff T \text{ is continuous on } \mathcal{X}$$

Proof. If T is bounded, we see that T is Lipschitz.

$$\|Tx - Ty\| \leq M\|x - y\|$$

Now, suppose T is linear and continuous at 0. Let ϵ satisfy

$$\sup_{\|y\| \leq \epsilon} \|Ty\| \leq 1$$

Observe that for all $x \in \mathcal{X}$, we have

$$\|Tx\| = \frac{\|x\|}{\epsilon}\Big\|T\frac{\epsilon x}{\|x\|}\Big\| \leq \frac{\|x\|}{\epsilon}$$

■

Here, we introduce a new terminology, which shall later show its value. Given a set X, we say two metrics d_1, d_2 on X are **equivalent**, and write $d_1 \sim d_2$, if we have

$$\exists m, M \in \mathbb{R}^+, \forall x, y \in X, md_1(x, y) \leq d_2(x, y) \leq Md_1(x, y)$$

Now, given a fixed vector space V, naturally, we say two norms $\|\cdot\|_1, \|\cdot\|_2$ on V are **equivalent** if

$$\exists m, M \in \mathbb{R}^+, \forall x \in X, m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

We say two metric d_1, d_2 on X are **topologically equivalent** if the topology they induce on X are identical.

A few properties can be immediately spotted.

(a) Our definition of $\sim$ between metrics of a fixed X is an equivalence relation.

(b) Our definition of $\sim$ between norms on a fixed V is an equivalence relation.

(c) Equivalent norms induce equivalent metrics.

(d) Equivalent metrics are topologically equivalent.

We now prove if V is finite-dimensional, then all norms on V are equivalent. This property will later show its value, as used to prove linear map of finite-dimensional domain is always continuous

Theorem 4.1.2. (All Norms on Finite-dimensional space are Equivalent) Suppose V is a finite-dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$. Then

$$\text{all norms on } V \text{ are equivalent}$$

Proof. Let $\{e_1, \ldots, e_n\}$ be a basis of V. Define ∞-norm $\|\cdot\|_\infty$ on V by

$$\left\|\sum \alpha_i e_i\right\|_\infty \triangleq \max |\alpha_i|$$

It is easily checked that $\|\cdot\|_\infty$ is indeed a norm. Fix a norm $\|\cdot\|$ on V. We reduce the problem into

$$\text{finding } m, M \in \mathbb{R}^+ \text{ such that } m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$$

We first claim

$$M = \sum \|e_i\| \text{ suffices}$$

Compute

$$\|x\| = \left\|\sum \alpha_i e_i\right\| \leq \sum |\alpha_i| \, \|e_i\| \leq \|x\|_\infty \sum \|e_i\| = M\|x\|_\infty \text{ (done)}$$

Note that reverse triangle inequality give us

$$\Big|\|x\| - \|y\|\Big| \leq \|x-y\| \leq M\|x-y\|_\infty \tag{4.1}$$

Then we can check that

(a) $\|\cdot\| : \Big(V, \|\cdot\|_\infty\Big) \to \mathbb{R}$ is Lipschitz continuous because of Equation 4.1.

(b) $S \triangleq \{y \in V : \|y\|_\infty = 1\}$ is sequentially compact in $\|\cdot\|$ and non-empty.

Now, by EVT, we know $\min_{y \in S} \|y\|$ exists. Note that $\min_{y \in S} \|y\| > 0$, since $0 \notin S$. We claim

$$m = \min_{y \in S} \|y\| \text{ suffices}$$

Fix $x \in V$ and compute

$$m\|x\|_\infty = \|x\|_\infty (\min_{y \in S} \|y\|) \leq \|x\|_\infty \cdot \left\|\frac{x}{\|x\|_\infty}\right\| = \|x\| \text{ (done)} \quad \text{(done)}$$

∎

Theorem 4.1.3. (Linear map of Finite-dimensional Domain is always Continuous) Given a finite-dimensional normed space $\mathcal{X}$ over $\mathbb{R}$ or $\mathbb{C}$, an arbitrary normed space $\mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$ and a linear transformation $T : \mathcal{X} \to \mathcal{Y}$, we have

$$T \text{ is continuous}$$

Proof. Fix $x \in \mathcal{X}, \epsilon$. We wish

$$\text{to find } \delta \text{ such that } \forall h \in \mathcal{X} : \|h\| \leq \delta, \|T(x+h) - Tx\| \leq \epsilon$$

Let $\{e_1, \dots, e_n\}$ be a basis of $\mathcal{X}$. Note that $\|\sum \alpha_i e_i\|_1 \triangleq \sum |\alpha_i|$ is a norm. Because $\mathcal{X}$ is finite-dimensional, we know $\|\cdot\|$ and $\|\cdot\|_1$ are equivalent. Then, we can fix $M \in \mathbb{R}^+$ such that

$$\|x\|_1 \leq M\|x\| \quad (x \in V)$$

We claim

$$\delta = \frac{\epsilon}{M(\max \|Te_i\|)} \text{ suffices}$$

Fix $\|h\| \leq \delta$ and express $h = \sum \alpha_i e_i$. Compute using linearity of T

$$\begin{aligned}
\|T(x+h) - Tx\| &= \|\sum \alpha_i T e_i\| \\
&\leq \sum |\alpha_i| \, \|Te_i\| \\
&\leq \|h\|_1 (\max \|Te_i\|) \\
&\leq M\|h\|(\max \|Te_i\|) = \epsilon \text{ (done)}
\end{aligned}$$

■

We now see that, because Linear transformation is bounded if and only if it is continuous and Linear map of finite-dimensional domain is always continuous, if $\mathcal{X}$ is finite-dimensional, then all linear map of domain $\mathcal{X}$ are bounded. A counter example to the generalization of this statement is followed.

Example 8 (Differentiation is an Unbounded Linear Operator)

$$\mathcal{X} = \Big(\mathbb{R}[x]|_{[0,1]}, \|\cdot\|_\infty\Big), D(P) \triangleq P'$$

Note that $\{x^n\}_{n\in\mathbb{N}}$ is bounded in $\mathcal{X}$ and $\{D(x^n)\}_{n\in\mathbb{N}}$ is not.

Now, suppose $\mathcal{X}, \mathcal{Y}$ are two fixed normed spaces over $\mathbb{R}$ or $\mathbb{C}$. We can easily check that the set $BL(\mathcal{X}, \mathcal{Y})$ of bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$ form a vector space over whichever field $\mathcal{Y}$ is over.

Naturally, our definition of boundedness of linear operator derive us a norm on $BL(\mathcal{X},\mathcal{Y})$, as followed

$$\|T\|_{\text{op}} \triangleq \inf\{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|\} \tag{4.2}$$

Before we show that our definition is indeed a norm, we first give some equivalent definitions and prove their equivalency.

Theorem 4.1.4. **(Equivalent Definitions of Operator Norm)** Given two fixed normed space $\mathcal{X},\mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$, a bounded linear operator $T : \mathcal{X} \to \mathcal{Y}$, and define $\|T\|_{\text{op}}$ as in Equation 4.2, we have

$$\|T\|_{\text{op}} = \sup_{x\in\mathcal{X},x\neq 0} \frac{\|Tx\|}{\|x\|}$$

Proof. Define $J \triangleq \{M \in \mathbb{R}^+ : \forall x \in \mathcal{X}, \|Tx\| \leq M\|x\|\}$ and observe

$$J = \{M \in \mathbb{R}^+ : M \geq \frac{\|Tx\|}{\|x\|}, \forall x \neq 0 \in \mathcal{X}\}$$

This let us conclude

$$\sup_{x\in\mathcal{X},x\neq 0} \frac{\|Tx\|}{\|x\|} = \min J = \|T\|_{\text{op}}$$

∎

It is now easy to see

$$\|T\|_{\text{op}} = \sup_{x\in\mathcal{X},x\neq 0} \frac{\|Tx\|}{\|x\|} \tag{4.3}$$
$$= \sup_{x\in\mathcal{X},\|x\|=1} \|Tx\| \tag{4.4}$$

It is not all in vain to introduce the equivalent definitions. See that the verification of $\|\cdot\|_{\text{op}}$ being a norm on $BL(\mathcal{X},\mathcal{Y})$ become simple by utilizing the equivalent definitions.

(a) For positive-definiteness, fix non-trivial T and fix $x \in \mathcal{X} \setminus N(T)$. Use Equation 4.3 to show $\|T\|_{\text{op}} \geq \frac{\|Tx\|}{\|x\|} > 0$.

(b) For absolute homogenity, use Equation 4.4 and $\|Tcx\| = |c| \cdot \|Tx\|$.

(c) For triangle inequality, use Equation 4.4 and $\|(T_1 + T_2)x\| \leq \|T_1x\| + \|T_2x\|$.

Naturally, and very very importantly, Equation 4.3 give us

$$\|Tx\| \leq \|T\|_{\text{op}} \cdot \|x\| \qquad (x \in \mathcal{X})$$

This inequality will later be the best tool to help analyze the derivatives of functions between Euclidean spaces, and perhaps better, it immediately give us

$$\frac{\|T_1T_2x\|}{\|x\|} \leq \frac{\|T_1\|_{\text{op}} \cdot \|T_2\|_{\text{op}} \cdot \|x\|}{\|x\|} = \|T_1\|_{\text{op}} \cdot \|T_2\|_{\text{op}}$$

Then Equation 4.3 give us

$$\|T_1T_2\|_{\text{op}} \leq \|T_1\|_{\text{op}} \cdot \|T_2\|_{\text{op}}$$

4.2 Directional Derivative and Gradient

Abstract

This short section introduce the idea of directional derivative and gradient. It shall be noted that, although both gradient and directional derivative are defined for real-valued function in this section, the notion of directional derivative can be easily generalized to function between Euclidean space; while the notion of gradient, as the way we define it, is only for real-valued function.

Given two normed space $\mathcal{X}, \mathcal{Y}$, suppose f maps an open neighborhood O around x in $\mathcal{X}$ into $\mathcal{Y}$. We say f is **differentiable at** x if there exists a bounded linear transformation $A_x : \mathcal{X} \to \mathcal{Y}$ (from now, A_x will be denoted df_x) such that

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - df_x(h)\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} = 0 \tag{4.5}$$

Immediately, we should check that the linear approximation is unique. Suppose df_x and df'_x both satisfy Equation 4.5. We are required to show $(df_x - df'_x)h = 0$ for all $\|h\|_{\mathcal{X}} = 1$. Fix $h \in \mathcal{X}$ such that $\|h\|_{\mathcal{X}} = 1$. Note that

$$\frac{(df_x - df'_x)th}{t} \text{ is a constant in } t \text{ for } t \neq 0$$

This then reduced the problem into showing

$$\frac{(df_x - df'_x)th}{t\|h\|_{\mathcal{X}}} \to 0 \text{ as } t \to 0 \tag{4.6}$$

Observe

$$(df_x - df'_x)th = \Big(f(x+th) - f(x) - df'_x(th)\Big) - \Big(f(x+th) - f(x) - df_x(th)\Big)$$

which implies

$$\|(df_x - df'_x)th\|_{\mathcal{Y}} \leq \|f(x+th) - f(x) - df'_x(th)\|_{\mathcal{Y}} + \|f(x+th) - f(x) - df_x(th)\|_{\mathcal{Y}}$$

and thus implies Equation 4.6.

It shall be quite clear that a function f differentiable at x must be continuous at x, by noting the nominator of Equation 4.5 must tend to 0. For clarity, we here specify the notation. By $\mathbb{R}$, we mean a field equipped with the usual norm $\|x\| \triangleq |x|$. By $\mathbb{R}^n$ we mean the set of functions from $\{1,\dots,n\}$ to $\mathbb{R}$ equipped with the usual vector addition, scalar multiplication, dot product and induced norm.

Definition 4.2.1. (Definition of Directional Derivative of Scalar function) Given a normal vector $v \in \mathbb{R}^n$ and a function f that maps an open-neighborhood E around $x \in \mathbb{R}^n$ into $\mathbb{R}$, by the **directional derivative** $\partial_v f(x)$ of f with respect to v at x, we mean

$$\partial_v f(x) \triangleq \lim_{t\to 0} \frac{f(x+tv)-f(x)}{t} \text{ if exists}$$

Something to note in our definition for directional derivative (Definition 4.2.1)

(a) The limit on the right hand side is done in $\Big(\mathbb{R}, |\cdot|\Big)$

(b) If f is differentiable at x, then we have

$$\partial_{av+bw} f(x) = df_x(av+bw) = adf_x(v) + bdf_x(w) = a\partial_v f(x) + b\partial_w f(x) \tag{4.7}$$

With what we observed, one can immediately see that if a function f is differentiable at x, then f has directional derivative with respect to any direction at x. The converse is not true. It is possible that f has directional derivative with respect to all directions, and yet f is still not differentaible. Consider

Example 9 (Discontinuous function such that all directional derivatives exist)

$$f : \mathbb{R}^2 \to \mathbb{R}, f(x,y) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{if } x^2+y^2 \neq 0 \\ 0 & \text{if } x = y = 0 \end{cases}$$

Definition 4.2.2. (Definition of Gradient of $\mathbb{R}^n \to \mathbb{R}$ function) Given a point $x \in \mathbb{R}^n$ with open neighborhood E, a function $f : E \to \mathbb{R}$ differentiable at x, we define the **gradient** $\nabla f(x) \in \mathbb{R}^n$ of f at x to be the unique vector that satisfy

$$\nabla f(x) \cdot v = df_x(v) \text{ for all } v \in \mathbb{R}^n$$

We should immediately discuss whether our definition of gradient is well-defined. The proof of existence and uniqueness follows from generating an orthogonal basis $\{v_1,\dots,v_n\}$ and noting $\nabla f(x)$ must equal to $\sum_{i=1}^n df_x(v_i)v_i$. A few things one must know about gradient is as followed

(a) $\nabla f(x)$ is only defined when f is differentiable at x.

(b) gradient $\nabla f(x)$ "points toward" the direction at which $f : \mathbb{R}^n \to \mathbb{R}$ grow the fastest. Suppose v is normal. See

$$\nabla f(x) \cdot v = df_x(v) = \partial_v f(x)$$

Using Cauchy-Schwarz Inequality, we see that $\partial_v f(x)$ is of largest value when $v = \frac{\nabla f(x)}{|\nabla f(x)|}$. If $v = \frac{\nabla f(x)}{|\nabla f(x)|}$, then

$$\partial_v f(x) = |\nabla f(x)|$$

(c) It is possible $\nabla f(x) = 0$. This is true if and only if df_x maps $\mathbb{R}^n$ into 0. This fact echos with the fact gradient points toward the fastest growing direction. See (b).

4.3 MVT

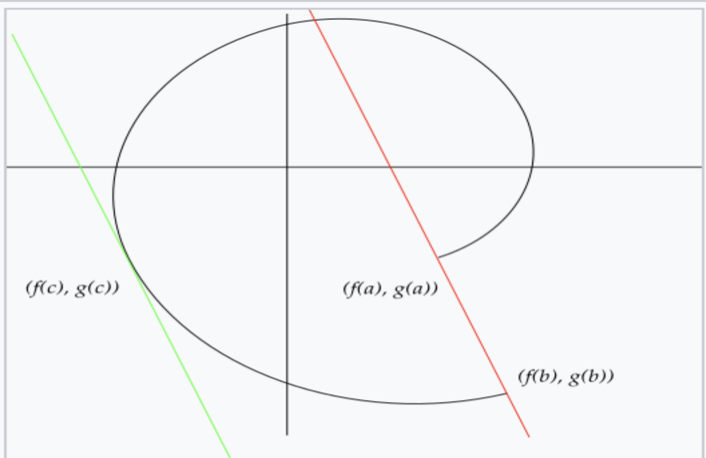
Abstract

This section introduce Cauchy's MVT and MVT, which will be heavily used later to prove multiple important and common results, e.g., L'Hospital Rule, Differentiability Theorem, The criterion for commutation of partial derivative and part 2 of FTC. For how important MVT is in the development of Theory of Calculus, it is worth pointing out that, fundamentally, MVT relies on the order of the real numbers, as usage of property of function differentiable at extremum in the proof of Cauchy's MVT suggest.

Theorem 4.3.1. (Local Extremum of Differentiable Function must have zero derivative) If $f : (x - \epsilon, x + \epsilon) \to \mathbb{R}$ is differentiable at x and attain maximum or minimum at x, then $f'(x) = 0$.

Proof. WOLG, suppose f attain maximum at x. We see that the value of $\frac{f(x+h)-f(x)}{h}$ is non-negative when $h < 0$ and non-positive when $h > 0$. This suggest that $f'(x) = 0$. ∎

Below is a graph to help visualize Cauchy's MVT.



Geometrical meaning of Cauchy's theorem

We now prove Cauchy's MVT.

Theorem 4.3.2. (Cauchy's MVT) Given a function $f : [a, b] \to \mathbb{R}$ such that

(a) f, g are differentiable on (a, b)

(b) f, g are continuous on $[a, b]$

There exists $x \in (a,b)$ such that

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

Proof. Define h on (a,b) by

$$h(x) \triangleq \big[f(b)-f(a)\big]g'(x) - \big[g(b)-g(a)\big]f'(x)$$

We reduced our problem into

finding $x \in (a,b)$ such that $h(x)=0$

Because f,g are both differentiable on (a,b), we know there exists an anti-derivative H on (a,b), which have the form

$$H(x) = \big[f(b)-f(a)\big]g(x) - \big[g(b)-g(a)\big]f(x)$$

Theorem 4.3.1 now allow us to reduce the problem into

finding a local extremum of H on (a,b)

Because f,g are both continuous on $[a,b]$, we know H is continuous on $[a,b]$. Then by EVT, we know

$$\exists x \in [a,b], H(x) = \max_{t\in[a,b]} H(t) \text{ and } \exists y \in [a,b], H(y) = \min_{t\in[a,b]} H(t)$$

If any of such x,y is in (a,b), we are done. If not, says that x,y both are on end points a or b. Compute that

$$H(a) = f(b)g(a) - g(b)f(a) = H(b)$$

We see H is constant on $[a,b]$. Then all points in (a,b) are extremums. (done) ∎

Corollary 4.3.3. (Lagrange's MVT) Given a function $f:[a,b]\to\mathbb{R}$ such that

(a) f is differentiable on (a,b)

(b) f is continuous on $[a,b]$

Then there exists $x \in (a,b)$ such that

$$f'(x) = \frac{f(b)-f(a)}{b-a}$$

Proof. The proof is done by applying Cauchy's MVT, where $g(x) \triangleq x$. ∎

There are two hypotheses in Lagrange's MVT

(a) f is differentiable on (a, b)

(b) f is continuous on $[a, b]$

They are all necessary. The necessity of differentiability on (a, b) is clear as shown by the canonical example using absolute value. The necessity of continuity on $[a, b]$ can be shown by the example

$$f(x) = \begin{cases} 1 & \text{if } a < x \leq b \\ 0 & \text{if } x = a \end{cases}$$

4.4 Differentiability Theorem

Abstract

This section prove

(a) The matrix representation of derivative for function between Euclidean spaces

(b) Differentiability Theorem

Note that the proof of Differentiability Theorem use MVT and the fact that all norms on $\mathbb{R}^k$ are equivalent where $k = nm$, and utilize the Frobenius norm.

Given an orthonormal basis $\{q_1, \dots, q_m\}$ of $\mathbb{R}^m$ and a function $f : \mathbb{R}^n \to \mathbb{R}^m$, we let $f_j(x)$ be a real number

$$f_j(x) = f(x) \cdot q_j$$

It shall be clear that

$$f(x) = \sum_{j=1}^{m} f_j(x) q_j$$

which explain why we require $\{q_1, \dots, q_m\}$ to be orthonormal in the first place. For brevity of the statement of the next theorem (Theorem 4.4.1), we introduce another notation. If we are provided a normal basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$, we denote $\partial_{e_i} f_j(x)$ by $\partial_i f_j(x)$

Theorem 4.4.1. (Derivative is Jacobian) Suppose $\alpha = \{e_1, \dots, e_n\}$ is a basis of $\mathbb{R}^n$, and $\beta = \{q_1, \dots, q_m\}$ is an orthonormal basis of $\mathbb{R}^m$. Suppose f maps an open neighborhood O around $x \in \mathbb{R}^n$ to $\mathbb{R}^m$. Then

$$f \text{ is differentiable at } x \implies \begin{cases} \partial_i f_j(x) \text{ exists for all } i, j \\ [df_x]^{\beta}_{\alpha} = \begin{bmatrix} \partial_1 f_1(x) & \cdots & \partial_n f_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 f_m(x) & \cdots & \partial_n f_m(x) \end{bmatrix} \end{cases}$$

Proof. Suppose $e_1, \dots, e_n$ are all normal. Fix i, j. We wish to show

$$\partial_i f_j(x) \text{ exists}$$

Because f is differentiable at x, by definition of df_x, we have

$$\lim_{t\to 0} \frac{|f(x+te_i)-f(x)-df_x(te_i)|}{|te_i|}=0$$

Set $R_i:\mathbb{R}\to\mathbb{R}^m$ by $R_i(t)\triangleq f(x+te_i)-f(x)-df_x(te_i)$. We have

$$\lim_{t\to 0}\frac{|R_i(t)|}{|t|}=0 \tag{4.8}$$

Compute

$$\begin{aligned} f_j(x+te_i)-f_j(x) &= \big(f(x+te_i)-f(x)\big)\cdot q_j \\ &= \big(R_i(t)+df_x(te_i)\big)\cdot q_j \\ &= R_i(t)\cdot q_j + t df_x(e_i)\cdot q_j \end{aligned}$$

This then give us

$$\frac{f_j(x+te_i)-f_j(x)}{t}=\frac{R_i(t)\cdot q_j}{t}+df_x(e_i)\cdot q_j$$

and

$$df_x(e_i)\cdot q_j-\frac{|R_i(t)\cdot q_j|}{|t|}\le\frac{f_j(x+te_i)-f_j(x)}{t}\le df_x(e_i)\cdot q_j+\frac{|R_i(t)\cdot q_j|}{|t|}$$

By Cauchy-Schwarz Inequality, we now have

$$df_x(e_i)\cdot q_j-\frac{|R_i(t)|}{|t|}\le\frac{f_j(x+te_i)-f_j(x)}{t}\le df_x(e_i)\cdot q_j+\frac{|R_i(t)|}{|t|}$$

Now applying Squeeze Theorem and Equation 4.8, we have

$$\partial_i f_j(x)=\lim_{t\to 0}\frac{f_j(x+te_i)-f_j(x)}{t}=df_x(e_i)\cdot q_j \text{ (done)}$$

Using the fact β is orthonormal, we now have

$$df_x(e_i)=\sum_{j=1}^m\Big(df_x(e_i)\cdot q_j\Big)q_j=\sum_{j=1}^m \partial_i f_j(x)q_j$$

and suggest the matrix representation. ■

Note that the converse is not always true. It is possible that a function f has all partial derivatives with respect to a given basis, or even all directions, and yet f is still discontinuous. We have given an example already in Directional Derivative and Gradient. Consider a less trivial one.

Example 10 (Non-differentiable Continuous Funciton with Partial Derivative)

$$f : \mathbb{R}^2 \to \mathbb{R} \text{ and } f(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq 0 \\ 0 & \text{if } (x,y) = 0 \end{cases}$$

We have

$$\partial_x f(0) = \partial_y f(0) = 0$$

By Theorem 4.4.1 (Derivative is Jacobian), if f is differentiable at 0, then df_0 must be trivial. Yet

$$\frac{|f(h,h) - f(0) - df_0(h,h)|}{|(h,h)|} = \frac{h}{2\,|h|} \not\to 0$$

Note that f is continuous at 0, by observing

$$x^2 + y^2 - 2\,|xy| = \big(\,|x| - |y|\,\big)^2 \geq 0 \implies \frac{x^2+y^2}{2} \geq |xy|$$

which implies

$$|f| \leq \frac{\sqrt{x^2+y^2}}{2}$$

We now introduce a property of function between normed space that are stronger than differentiability. Given two normed space $\mathcal{X}, \mathcal{Y}$, and an open $E \subseteq \mathcal{X}$, we say $f : E \to \mathcal{Y}$ is **continuously differentiable** on $\mathcal{Y}$ if the map $D : (E, \|\cdot\|_{\mathcal{X}}) \to \Big(BL(\mathcal{X}, \mathcal{Y}), \|\cdot\|_{\text{op}}\Big)$ defined by

$$D(x) = df_x$$

is continuous. Note that the definition of the term "continuously differentiable" coincide when $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ and df_x is just $h \mapsto f'(x)h$. We now give proof to the Differentiability

Theorem, which links between the continuity of total derivative and the continuity of partial derivatives.

Theorem 4.4.2. (Differentiability Theorem) Suppose $\alpha = \{e_1, \dots, e_n\}$ is an orthonormal basis of $\mathbb{R}^n$, and $\beta = \{q_1, \dots, q_m\}$ is an orthonormal basis of $\mathbb{R}^m$. Suppose f maps an open set $E \subseteq \mathbb{R}^n$ to $\mathbb{R}^m$. Then

$$f \text{ is continuously differentiable on } E \iff \partial_i f_j \text{ exists and is continuous on } E \text{ for all } i, j$$

Proof. $(\longrightarrow)$

Fix i, j. Because f is differentiable on E, we know $\partial_i f_j$ exists on E by Theorem 4.4.1. Fix $x \in E$. We only have to show

$$\partial_i f_j \text{ is continuous at } x$$

Fix ϵ. We wish

$$\text{to find } \delta \text{ such that } |\partial_i f_j(y) - \partial_i f_j(x)| \leq \epsilon \text{ for all } |y - x| < \delta$$

Because f is continuously differentaible at x, we know there exists δ such that

$$\|df_y - df_x\|_{\text{op}} < \epsilon \text{ for all } |y - x| \leq \delta$$

We claim

$$\text{such } \delta \text{ suffices}$$

By the the matrix representation, we know

$$\partial_i f_j(y) - \partial_i f_j(x) = (df_y - df_x)e_i \cdot q_j$$

Then by Cauchy-Inequality, we have

$$\begin{aligned}|\partial_i f_j(y) - \partial_i f_j(x)| &\leq |(df_y - df_x)e_i| \\ &\leq \|df_y - df_x\|_{\text{op}} < \epsilon \text{ (done)}\end{aligned}$$

$(\longleftarrow)$

We first show

$$f \text{ is differentiable on } E$$

We first prove

$$\forall j \in \{1, \dots, m\}, f_j : \mathbb{R}^n \to \mathbb{R} \text{ is differentiable on } E \implies f \text{ is differentiable on } E$$

Fix $x \in E$. We wish to prove

$$f \text{ is differentiable at } x$$

Define $A : E \to \mathbb{R}^m$ by

$$A(h) \triangleq \sum_{j=1}^{m} (df_j)_x(h) q_j$$

We claim

$$A \text{ suffices to be the } df_x$$

Using the fact q_j are orthonormal, we have

$$f(x+h) - f(x) - A(h) = \sum_{j=1}^{m} \big(f_j(x+h) - f_j(x) - (df_j)_x(h) \big) q_j$$

This give us

$$\begin{aligned}
\lim_{h \to 0} \frac{|f(x+h) - f(x) - A(h)|}{|h|} &= \lim_{h \to 0} \frac{\left| \sum_{j=1}^{m} \big(f_j(x+h) - f_j(x) - (df_j)_x(h) \big) q_j \right|}{|h|} \\
&\leq \lim_{h \to 0} \frac{\sum_{j=1}^{m} |f_j(x+h) - f_j(x) - (df_j)_x(h)|}{|h|} = 0 \text{ (done)}
\end{aligned}$$

Fix $j \in \{1, \dots, m\}$. We can now reduce the problem into

$$f_j : \mathbb{R}^n \to \mathbb{R} \text{ is differentiable on } E$$

Fix $x \in E$. We wish to prove

$$f_j \text{ is differentiable at } x$$

Express $h = \sum_{i=1}^{n} h_i e_i$. Define $B : E \to \mathbb{R}$ by

$$B(h) = \sum_{i=1}^{n} \partial_i f_j(x) h_i$$

We claim

$$B \text{ suffices to be } (df_j)_x$$

By continuity of each $\partial_i f_j$ on E, we can let δ satisfy

$$|\partial_i f_j(y) - \partial_i f_j(x)| < \frac{\epsilon}{n} \text{ for all } y \in B_\delta(x)$$

We claim

$$\frac{|f_j(y)-f_j(x)-B(y-x)|}{|y-x|}\leq \epsilon \text{ for all } y\in B_\delta(x)$$

Express $y-x=\sum_{k=1}^n h_k e_k$. Define $v_0,\dots,v_n\in\mathbb{R}^n$ by

$$v_0\triangleq 0 \text{ and } v_k\triangleq \sum_{i=1}^k h_i e_i \text{ for all } k\in\{1,\dots,n\}$$

Now observe

$$\begin{aligned}
\frac{|f_j(y)-f_j(x)-B(y-x)|}{|y-x|} &= \frac{|f_j(x+v_n)-f_j(x)-B(\sum_{k=1}^n h_k e_k)|}{|y-x|}\\
&= \frac{\left|\left(\sum_{k=1}^n f_j(x+v_k)-f_j(x+v_{k-1})\right)-\sum_{k=1}^n \partial_k f_j(x)h_k\right|}{|y-x|}\\
&= \frac{|\sum_{k=1}^n f_j(x+v_k)-f_j(x+v_{k-1})-\partial_k f_j(x)h_k|}{|y-x|}\\
&= \frac{|\sum_{k=1}^n f_j(x+v_{k-1}+h_k e_k)-f_j(x+v_{k-1})-\partial_k f_j(x)h_k|}{|y-x|}\\
(\text{ For some } e_k\in(0,1) \text{ by MVT}) \quad &= \frac{|\sum_{k=1}^n \partial_k f_j(x+v_{k-1}+t_k e_k)h_k-\partial_k f_j(x)h_k|}{|y-x|}\\
&\leq \frac{\sum_{k=1}^n \left|\left(\partial_k f_j(x+v_{k-1}+t_k e_k)-\partial_k f_j(x)\right)h_k\right|}{|y-x|}\\
&< \frac{\sum_{k=1}^n \frac{\epsilon}{n}|h_k|}{|y-x|}\leq \epsilon \text{ (done)}
\end{aligned}$$

We now prove

$$f \text{ is continuously differentiable on } E$$

Fix ϵ and $x\in E$. We are required

$$\text{to find } \delta \text{ such that } \|df_y-df_x\|_{\text{op}}\leq \epsilon \text{ for all } y\in B_\delta(x)$$

Note that one can define a norm $\|\cdot\|_F$ called "Forbenius Norm" on $BL(\mathbb{R}^n,\mathbb{R}^n)$ by

$$\|A\|_F\triangleq\sqrt{\sum_{i=1}^n\sum_{j=1}^n a_{i,j}^2} \text{ where } [A]_\alpha^\beta=\begin{bmatrix} a_{1,1} & \cdots & a_{n,1}\\ \vdots & \ddots & \vdots\\ a_{1,n} & \cdots & a_{n,n}\end{bmatrix}$$

Because all norms on finite-dimensional real vector spaces are equivalent, we know there exists M such that for all $x \in E$, we have

$$\|df_x\|_{\text{op}} \leq M\|df_x\|_F$$

Because the partial derivatives are all continuous by definition, we can let δ satisfy

$$\big(\partial_i f_j(x+h)\big)^2 - \big(\partial_i f_j(x)\big)^2 < \frac{\epsilon^2}{M^2 n^2} \text{ for all } h \in B_\delta(0)$$

We claim

such δ suffices

Let $|y-x| < \delta$. We see

$$\|df_y - df_x\|_{\text{op}} \leq M\|df_y - df_x\|_F < M\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\epsilon^2}{M^2 n^2}} = \epsilon \text{ (done)}$$

■

4.5 Product Rule and Chain Rule

Abstract

This short section prove Product rule in the setting of gradient of real-valued function and Chain rule for functions between normed spaces, which is heavily used in next section on smooth function and Differential Geometry.

Although distinct versions of product rule exists in the context of vector calculus, here we prove the almost most simple kind.

Theorem 4.5.1. (Product Rule) Given two function $f, g : \mathbb{R}^d \to \mathbb{R}$ differentiable at x, we have

$$\nabla(fg)(x) = g(x)\nabla f(x) + f(x)\nabla g(x)$$

Proof. Note that

$$\lim_{h\to 0} f(x+h)\Big[\frac{g(x+h) - g(x) - \nabla g(x)\cdot h}{|h|}\Big] = 0$$
$$\text{and } \lim_{h\to 0} g(x)\Big[\frac{f(x+h) - f(x) - \nabla f(x)\cdot h}{|h|}\Big] = 0$$

Adding these two equations together, we have

$$\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x) - [g(x)\nabla f(x) + f(x)\nabla g(x)]\cdot h}{|h|} = 0$$

∎

We now prove the Chain Rule for function between normed space.

Theorem 4.5.2. (Chain Rule) Given three normed space $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, a point $x \in \mathcal{X}$, a function g that map an open set $U \subseteq \mathcal{Y}$ containing $f(x)$ into $\mathcal{Z}$, a function f that map an open-neighborhood around x into U such that

(a) f is differentiable at x

(b) g is differentiable at $f(x)$

we have

$$d(g\circ f)_x = dg_{f(x)} \circ df_x$$

Proof. For brevity, we use $F \triangleq g \circ f$. We wish to prove

$$\lim_{h \to 0} \frac{\|F(x+h) - F(x) - dg_{f(x)} df_x(h)\|_{\mathcal{Z}}}{\|h\|_{\mathcal{X}}} = 0$$

Fix $k \triangleq f(x+h) - f(x)$. Observe

$$F(x+h) - F(x) - dg_{f(x)} df_x(h) = \Big(g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k) \Big) + dg_{f(x)}\big(k - df_x(h)\big)$$

This now implies

$$\frac{\|F(x+h) - F(x) - dg_{f(x)} df_x(h)\|_{\mathcal{Z}}}{\|h\|_{\mathcal{X}}} \text{ is smaller than}$$

$$\frac{\|g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k)\|_{Z} + \|dg_{f(x)}\big(k - df_x(h)\big)\|_{\mathcal{Z}}}{\|h\|_{\mathcal{X}}}$$

This let us reduce the problem into proving

$$\lim_{h \to 0} \frac{\|g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k)\|_{Z}}{\|h\|_{\mathcal{X}}} = 0$$

$$\text{and } \lim_{h \to 0} \frac{\|dg_{f(x)}\big(k - df_x(h)\big)\|_{\mathcal{Z}}}{\|h\|_{\mathcal{X}}} = 0$$

We first prove

$$\lim_{h \to 0} \frac{\|g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k)\|_{Z}}{\|h\|_{\mathcal{X}}} = 0$$

Note that if $\|k\|_{\mathcal{Y}} = 0$, we have

$$\frac{\|g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k)\|_{Z}}{\|h\|_{\mathcal{X}}} = 0$$

Now, observe that

$$\frac{\|g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k)\|_{Z}}{\|h\|_{\mathcal{X}}} = \frac{\|g\big(f(x)+k\big) - g\big(f(x)\big) - dg_{f(x)}(k)\|_{Z}}{\|k\|_{\mathcal{Y}}} \cdot \frac{\|k\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}}$$

Because $h \to 0 \implies k \to 0$, we can now reduce the problem into proving

$$\limsup_{h \to 0} \frac{\|k\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} \text{ exists}$$

Observe

$$\begin{aligned}
\frac{\|k\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} &= \frac{\|f(x+h) - f(x) - df_x(h) + df_x(h)\|_{\mathcal{X}}}{\|h\|_{\mathcal{X}}} \\
&\leq \frac{\|f(x+h) - f(x) - df_x(h)\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} + \frac{\|df_x(h)\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} \\
&\leq \frac{\|f(x+h) - f(x) - df_x(h)\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} + \|df_x\|_{\text{op}} \text{ (done)}
\end{aligned}$$

We now prove

$$\lim_{h \to 0} \frac{\|dg_{f(x)}\big(k - df_x(h)\big)\|_{\mathcal{Z}}}{\|h\|_{\mathcal{X}}} = 0$$

Note that if $f(x+h) - f(x) - df_x(h) = 0$, then $\|dg_{f(x)}\big(k - df_x(h)\big)\|_{\mathcal{Z}} = 0$. Now, observe

$$\begin{aligned}
\frac{\|dg_{f(x)}\big(k - df_x(h)\big)\|_{\mathcal{Z}}}{\|h\|_{\mathcal{X}}} &= \frac{\|dg_{f(x)}\big(f(x+h) - f(x) - df_x(h)\big)\|_{\mathcal{Z}}}{\|f(x+h) - f(x) - df_x(h)\|_{\mathcal{Y}}} \cdot \frac{\|f(x+h) - f(x) - df_x(h)\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} \\
&\leq \|dg_{f(x)}\|_{\text{op}} \cdot \frac{\|f(x+h) - f(x) - df_x(h)\|_{\mathcal{Y}}}{\|h\|_{\mathcal{X}}} \to 0 \text{ (done)}
\end{aligned}$$

$\blacksquare$

4.6 Smoothness

Abstract

This is a short section introducing the idea of smooth functions between Euclidean spaces, which heavily rely on Chain Rule. One should note that our **notation** $\partial_{21}f$ here means $\partial_2(\partial_1 f)$.

Theorem 4.6.1. (Structure of Mixed Partial Derivative) Given an open set $E \subseteq \mathbb{R}^2$, a point $p \in E$, a basis $\{e_1, e_2\}$ of $\mathbb{R}^2$ and a function $f : E \to \mathbb{R}$ such that

(a) $\partial_1 f$ exists on E

(b) $\partial_2 f$ exists on E

(c) $\partial_{21} f$ exists on E and is continuous at p

We have

$$\partial_{12} f(p) = \partial_{21} f(p)$$

Proof. Express elements of E in the basis $\{e_1, e_2\}$, and express $p = (a, b)$. We are required to prove

$$\lim_{h \to 0} \frac{\partial_2 f(a + h, b) - \partial_2 f(a, b)}{h} = \partial_{21} f(a, b)$$

Let $D \triangleq \{q - (a, b) \in \mathbb{R}^2 : q \in E \setminus \{p\}\}$ and define $\Delta(h, k) : D \to \mathbb{R}$ by

$$\Delta(h, k) \triangleq f(a + h, b + k) - f(a, b + k) - f(a + h, b) + f(a, b)$$

Note that because $\partial_2 f$ exists on E, for all $h \neq 0$, we have

$$\lim_{k \to 0} \frac{\Delta(h, k)}{hk} = \frac{\partial_2 f(a + h, b) - \partial_2 f(a, b)}{h}$$

This let us reduce the problem into proving

$$\lim_{h \to 0} \lim_{k \to 0} \frac{\Delta(h, k)}{hk} = \partial_{21} f(a, b)$$

We first show

> For all $(h, k) \in D : hk \neq 0$, $\frac{\Delta(h,k)}{hk} = \partial_{21} f(x, y)$ for some $(x, y) \in D : |x - a| < h$ and $|y - b| < k$.

Fix $(h,k) \in D : hk \neq 0$. Define $u : (b-\epsilon, b+\epsilon) \to \mathbb{R}$ by

$$u(t) \triangleq f(t, b+k) - f(t,b)$$

Chain Rule allow us to compute

$$u'(t) = \partial_1 f(t, b+k) - \partial_1 f(t,b)$$

We then can deduce

$$\begin{aligned}\Delta(h,k) &= u(a+h) - u(a) \\ &= hu'(x) \text{ for some } x \in (a, a+h) \text{ by MVT} \\ &= h\big(\partial_1 f(x, b+k) - \partial_1 f(x,b)\big)\end{aligned}$$

Now, because $\partial_{21} f$ exists on E, we can deduce

$$\begin{aligned}\Delta(h,k) &= h(\partial_1 f(x, b+k) - \partial_1 f(x,b)) \\ &= hk\partial_{21} f(x,y) \text{ for some } y \in (b, b+k) \text{ by MVT (done)}\end{aligned}$$

Fix ϵ. We wish

$$\text{to find some } \delta \text{ such that for all } h : 0 < |h| < \delta, \left| \lim_{k\to 0} \frac{\Delta(h,k)}{hk} - \partial_{21} f(x,y) \right| \leq \epsilon$$

Because $\partial_{21} f$ is continuous at p, by olive lemma, we know there exists δ such that

$$\left| \frac{\Delta(h,k)}{hk} - \partial_{21} f(a,b) \right| < \frac{\epsilon}{2} \text{ for all } h,k \in (-\delta, \delta) \setminus \{0\}$$

We claim

$$\text{such } \delta \text{ works}$$

Fix $h : 0 < |h| < \delta$. Note that $\lim_{k\to 0} \frac{\Delta(h,k)}{hk} = \frac{\partial_2 f(a+h,b) - \partial_2 f(a,b)}{h}$ exists, so we can find small enough k' such that

$$0 < |k'| < \delta \text{ and } \left| \lim_{k\to 0} \frac{\Delta(h,k)}{hk} - \frac{\Delta(h,k')}{hk'} \right| < \frac{\epsilon}{2}$$

Now observe

$$\left| \lim_{k\to 0} \frac{\Delta(h,k)}{hk} - \partial_{21} f(x,y) \right| \leq \left| \lim_{k\to 0} \frac{\Delta(h,k)}{hk} - \frac{\Delta(h,k')}{hk'} \right| + \left| \frac{\Delta(h,k')}{hk'} - \partial_{21} f(a,b) \right| \leq \epsilon \text{ (done)}$$

$\blacksquare$

Corollary 4.6.2. (Clairaut's Theorem on equality of mixed partial) Given a basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$, an open set $E \subseteq \mathbb{R}^n$, a function $f : E \to \mathbb{R}$ such that

$$\partial_{ij} f \text{ exist and is continuous on } E \text{ for all } i, j \in \{1, \dots, n\}$$

We have

$$\partial_{ij} f = \partial_{ji} f \text{ on } E \text{ for all } i, j \in \{1, \dots, n\}$$

Given a function $f : E \subseteq \mathbb{R}^n \to \mathbb{R}$, we now see that, by Differentiability Theorem, if $\partial_x f, \partial_{xy} f$ exist on E and the latter is continuous on E, then f is continuously differentiable on E.

Example 11 (A C^1 but not C^2 function)

$$f(x, y) \triangleq \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x, y) \neq 0 \\ 0 & \text{if } (x, y) = 0 \end{cases}$$

All second order partial derivatives exist, yet f_{xy}, f_{yx} does not commute at 0.

Now, if we define a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ on some open subset U to be **smooth** if for all $j \in \{1, \dots, m\}$, all partial derivatives of f exists, we then see that the partial derivatives of smooth function always commute. Notably, one can check that if $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, and $g : V \subseteq \mathbb{R}^m \to \mathbb{R}^d$ are both smooth and $f(U) \subseteq V$, then $g \circ f : U \to \mathbb{R}^d$ is also smooth, using long tedious induction with Product rule and Chain rule.

4.7 Holomorphic Functions

Abstract

This is a short section introducing the idea of holomorphic and prove some basic properties of holomorphic functions, i.e., Cauchy Riemann Criteria and Product and Quotient Rule for Holomorphic Function.

Given a complex-valued function f defined on some open subset of $\mathbb{C}$ containing z, we say f is **holomorphic at** z if there exists some complex number denoted by $f'(z)$ such that

$$\frac{f(z+h)-f(z)-f'(z)h}{h} \to 0 \text{ as } h \to 0; h \in \mathbb{C}$$

Immediately, one can see that a holomorphic function is differentiable when viewed as a function between $\mathbb{R}^2$ with derivative

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \text{ where } f'(z) = a + bi \tag{4.9}$$

With the form of derivative in mind, one may conjecture that holomorphic is a 'stricter' condition than merely differentiable when regarded as function between $\mathbb{R}^2$. This is exactly true. Consider the following example.

Example 12 (A not holomorphic function)

$$f(z) \triangleq \overline{z}$$

This is a linear function with matrix representation

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

which doesn't fit the necessary form in Equation 4.9.

Theorem 4.7.1. (Cauchy Riemann Criteria) Given a complex-valued function f defined on some open subset U of $\mathbb{C}$ containing z, if we write

$$f(x+yi) = u(x,y) + iv(x,y)$$

where $u, v : U \to \mathbb{R}$, then the following two statements are equivalent.

(a) f is holomorphic at z.

(b) u, v are differentiable at z and

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = \frac{-\partial v}{\partial x} \text{ at } z$$

Proof. (a) to (b) is an immediate result of Chain Rule and matrix representation of $f'(z)$. Suppose (b) is true. Let $h = h_1 + ih_2$ and $z = x + yi$. Because u, v are differentiable at (x, y), by the matrix representation of derivative, we have

$$\begin{aligned}
\frac{f(z+h)-f(z)}{h} &= \frac{u(x+h_1, y+h_2) - u(x,y)}{h_1 + ih_2} + i\frac{v(x+h_1, y+h_2) - v(x,y)}{h_1 + ih_2} \\
&= \frac{h_1 u_x + h_2 u_y + ih_1 v_x + ih_2 v_y + o(|h|)}{h_1 + ih_2} \\
&= \frac{(h_1 + ih_2)u_x + i(h_1 + ih_2)v_x + o(|h|)}{h_1 + ih_2} \\
&= u_x + iv_x + \frac{o(|h|)}{h_1 + ih_2} \to u_x + iv_x
\end{aligned}$$

∎

Theorem 4.7.2. (Product and Quotient Rule for Holomorphic Function) Given two function f, g holomorphic at z, we have

$$(fg)'(z) = f'(z)g(z) + f(z)g'(z)$$

and if $g(z) \neq 0$, we also have

$$\Big(\frac{f}{g}\Big)'(z) = \frac{f'(z)g(z) - f(z)g'(z)}{(g(z))^2}$$

Proof. Observe

$$\begin{aligned}
&\frac{f(z+h)g(z+h) - f(z)g(z)}{h} \\
=&f(z+h)\Big[\frac{g(z+h)-g(z)}{h}\Big] + g(z)\Big[\frac{f(z+h)-f(z)}{h}\Big] \to f'(z)g(z) + f(z)g'(z)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\frac{f(z+h)}{g(z+h)} - \frac{f(z)}{g(z)}}{h} &= \frac{f(z+h)g(z) - f(z)g(z+h)}{g(z+h)g(z)h} \\
&= \frac{1}{g(z+h)g(z)}\Big[g(z)\Big(\frac{f(z+h)-f(z)}{h}\Big) - f(z)\Big(\frac{g(z+h)-g(z)}{h}\Big)\Big] \\
&\to \frac{f'(z)g(z) - f(z)g'(z)}{(g(z))^2}
\end{aligned}$$

∎

4.8 Uniform Convergence and Differentiation

Abstract

This is a section discussing the relationship between uniform convergence and differentiation, which heavily rely on the usage of MVT, and is used to prove Analytic function is smooth.

Before stating Theorem 4.8.1, let's see three examples why we don't (can't) use the hypothesis: $f_n \to f$ uniformly in our statement of Theorem 4.8.1

Example 13 (Differentiable functions are NOT closed under uniform convergence)

$$X = [-1, 1] \text{ and } f(x) = |x|$$

By Weierstrass approximation Theorem, there is a sequence of polynomials (differentiable) that uniformly converge to f, which is not differentiable at 0.

Example 14 (Derivative won't necessarily converge to the right place)

$$X = \mathbb{R} \text{ and } f_n(x) = \frac{\sin nx}{\sqrt{n}}$$

Compute

$$f'(x) = 0 \text{ and } f_n'(x) = \sqrt{n} \cos nx$$

Example 15 (Derivative won't necessarily converge to the right place)

$$X = \mathbb{R} \text{ and } f_n(x) = \frac{x}{1 + nx^2}$$

Compute

$$f = \tilde{0} \text{ and } f_n'(0) = 1$$

Informally speaking, these examples together with the fact Riemann integral are closed under uniform convergence should give you some ideas that differentiation and integration although are operations inverse to each other, are NOT symmetric. There is a certain hierarchy on continuous functions on a fixed compact interval. Thus, we have the next Theorem in its form. Note that in application, the next Theorem only require us to prove f_n' uniformly converge, and doesn't require us to prove to where does it converge.

Theorem 4.8.1. (Uniform Convergence and Differentiation) Given a bounded interval $[a,b]$ and some sequence of function $f_n : [a,b] \to \mathbb{R}$ such that

(a) f_n' uniformly converge on (a,b)

(b) f_n are continuous on $[a,b]$

(c) $f_n(x_0) \to L$ for some $x_0 \in [a,b]$

Then

(a) f_n uniformly converge on $[a,b]$

(b) and

$$\Big(\lim_{n\to\infty} f_n\Big)'(x_0) = \lim_{n\to\infty} f_n'(x_0) \text{ on } (a,b)$$

Proof. We first prove

$$f_n \text{ uniformly converge on } [a,b] \tag{4.10}$$

Fix ϵ. We wish

$$\text{to find } N \text{ such that } \|f_n - f_m\|_\infty \le \epsilon \text{ for all } n,m > N$$

Because $f_n(x_0)$ converge, and f_n' uniformly converge, we know there exists N such that

$$\begin{cases} |f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2} \\ \|f_n' - f_m'\|_\infty < \frac{\epsilon}{2(b-a)} \end{cases} \quad \text{for all } n,m > N \tag{4.11}$$

We claim

$$\text{such } N \text{ works}$$

Fix $x \in [a,b]$ and $n,m > N$. We first show

$$\big|(f_n - f_m)(x) - (f_n - f_m)(x_0)\big| \le \frac{\epsilon}{2}$$

Because $(f_n - f_m)' = f_n' - f_m'$, by MVT and Equation 4.11, we can deduce

$$\begin{aligned} \big|(f_n - f_m)(x) - (f_n - f_m)(x_0)\big| &= \Big|\big[(f_n - f_m)'(t)\big](x - x_0)\Big| \text{ for some } t \text{ between } x, x_0 \\ &< \frac{\epsilon}{2(b-a)} \cdot |x - x_0| \\ &\le \frac{\epsilon}{2(b-a)} \cdot (b-a) = \frac{\epsilon}{2} \quad \big(\because x, x_0 \in [a,b]\big) \text{ (done)} \end{aligned}$$

Now, by Equation 4.11, we have

$$\begin{aligned}\big|(f_n - f_m)(x)\big| &\leq \big|(f_n - f_m)(x) - (f_n - f_m)(x_0)\big| + \big|(f_n - f_m)(x_0)\big| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ (done)}\end{aligned}$$

Let $f : [a, b] \to \mathbb{R}$ be the limit of f_n. It remains to prove

$$f'(x) = \lim_{n\to\infty} f'_n(x) \text{ on } (a, b) \tag{4.12}$$

Fix $x \in (a, b)$ and define $\varphi, \varphi_n : [a, b] \setminus x \to \mathbb{R}$ by

$$\varphi(t) \triangleq \frac{f(t) - f(x)}{t - x} \text{ and } \varphi_n(t) \triangleq \frac{f_n(t) - f_n(x)}{t - x}$$

It is clear that $\varphi_n \to \varphi$ pointwise on $[a, b] \setminus x$. We now show

$$\varphi_n \to \varphi \text{ uniformly on } [a, b] \setminus x$$

Fix ϵ. We have

$$\text{to find } N \text{ such that } |\varphi_n(t) - \varphi_m(t)| \leq \epsilon \text{ for all } n, m > N \text{ and } t \in [a, b] \setminus x$$

Because f'_n uniformly converge on (a, b), we know there exists N such that

$$\|f'_n - f'_m\|_\infty \leq \epsilon \text{ for all } n, m > N \tag{4.13}$$

We claim

$$\text{such } N \text{ works}$$

Fix $n, m > N$ and $t \in [a, b] \setminus x$. We wish to prove

$$|\varphi_n(t) - \varphi_m(t)| \leq \epsilon$$

Because $(f_n - f_m)' = f'_n - f'_m$, by MVT and Equation 4.13, we can deduce

$$\begin{aligned}|\varphi_n(t) - \varphi_m(t)| &\leq \left|\frac{f_n(t) - f_n(x)}{t - x} - \frac{f_m(t) - f_m(x)}{t - x}\right| \\ &= \left|\frac{(f_n - f_m)(t) - (f_n - f_m)(x)}{t - x}\right| \\ &= \big|(f'_n - f'_m)(t_0)\big| \text{ for some } t_0 \text{ between } t, x \\ &\leq \epsilon \text{ (done)}\end{aligned}$$

Note that

$$\lim_{n\to\infty}\lim_{t\to x}\varphi_n(t)=\lim_{n\to\infty}f'(x)\text{ exists}$$

We can now exchange the limit and see that the derivative of f at x exists.

$$\begin{aligned}f'(x)=\lim_{t\to x}\varphi(t)&=\lim_{t\to x}\lim_{n\to\infty}\varphi_n(t)\\&=\lim_{n\to\infty}\lim_{t\to x}\varphi_n(t)=\lim_{n\to\infty}f_n'(x)\text{ (done)}\end{aligned}$$

■

Theorem 4.8.2. (Uniform Convergence and Holomorphic) Given a real number r and some sequence of function $f_n:\overline{D_r(z_0)}\to\mathbb{C}$ such that

(a) f_n' uniformly converge on $D_r(z_0)$

(b) f_n are continuous on $\overline{D_r(z_0)}$

(c) $f_n(v)\to L$ for some $v\in\overline{D_r(z_0)}$

Then

(a) f_n uniformly converge on $\overline{D_r(z_0)}$

(b) and

$$\Big(\lim_{n\to\infty}f_n\Big)'(z)=\lim_{n\to\infty}f_n'(z)\text{ on }D_r(z_0)$$

Proof. We first prove

$$f_n\text{ uniformly converge on }\overline{D_r(z_0)}\tag{4.14}$$

Fix ϵ. We wish

$$\text{to find }N\text{ such that }\|f_n-f_m\|_\infty\leq\epsilon\text{ for all }n,m>N$$

Because $f_n(v)$ converge, and f_n' uniformly converge, we know there exists N such that

$$\begin{cases}|f_n(v)-f_m(v)|<\frac{\epsilon}{2}\\\|f_n'-f_m'\|_\infty<\frac{\epsilon}{8r}\end{cases}\quad\text{for all }n,m>N\tag{4.15}$$

We claim

$$\text{such }N\text{ works}$$

Fix $z \in \overline{D_r(z_0)}$ and $n, m > N$. We first show

$$\big|(f_n - f_m)(z) - (f_n - f_m)(z_0)\big| \leq \frac{\epsilon}{2}$$

Denote $f_n - f_m : \overline{D_r(z_0)} \to \mathbb{C}$ by g. Because

$$|g(z) - g(z_0)| \leq \Big|\text{Re}\Big(g(z) - g(z_0)\Big)\Big| + \Big|\text{Im}\Big(g(z) - g(z_0)\Big)\Big|$$

WOLG, we only have to prove

$$\Big|\text{Re}\Big(g(z) - g(z_0)\Big)\Big| \leq \frac{\epsilon}{4}$$

Because $\overline{D_r(z_0)}$ is convex, we can define $h : [0,1] \to \mathbb{R}$ by

$$h(t) \triangleq \text{Re}\Big(g(tz + (1-t)z_0)\Big)$$

By Chain Rule and matrix representation of derivative, we see that for all $t \in (0,1)$

$$\begin{aligned} h'(t) = ac - bd \text{ where } z_0 - z = a + bi \\ \text{and } g'(tz + (1-t)(z_0)) = c + di \end{aligned}$$

Because $|a + bi| \leq r$ and $|c + di| \leq \frac{\epsilon}{8r}$ by Equation 4.15, if we use MVT, we see that

$$\begin{aligned} \Big|\text{Re}\Big(g(z) - g(z_0)\Big)\Big| = |h(1) - h(0)| &= |h(t)| \text{ for some } t \in (0,1) \\ &= |ac| + |bd| \leq \frac{\epsilon}{4} \text{ (done)} \end{aligned}$$

Now, by Equation 4.15, we have

$$\begin{aligned} \big|(f_n - f_m)(z)\big| &\leq \big|(f_n - f_m)(z) - (f_n - f_m)(v)\big| + \big|(f_n - f_m)(v)\big| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ (done)} \end{aligned}$$

Let $f : \overline{D_r(z_0)} \to \mathbb{C}$ be the limit of f_n. It remains to prove

$$f'(z) = \lim_{n\to\infty} f_n'(z) \text{ on } D_r(z_0) \tag{4.16}$$

Fix $z \in D_r(z_0)$ and define $\varphi, \varphi_n : \overline{D_r(z_0)} \setminus z \to \mathbb{R}$ by

$$\varphi(u) \triangleq \frac{f(u) - f(z)}{u - z} \text{ and } \varphi_n(u) \triangleq \frac{f_n(u) - f_n(z)}{u - z}$$

It is clear that $\varphi_n \to \varphi$ pointwise on $\overline{D_r(z_0)} \setminus z$. We now show

$$\varphi_n \to \varphi \text{ uniformly on } \overline{D_r(z_0)} \setminus z$$

Fix ϵ. We have

$$\text{to find } N \text{ such that } |\varphi_n(t) - \varphi_m(t)| \leq \epsilon \text{ for all } n, m > N \text{ and } t \in \overline{D_r(z_0)} \setminus z$$

Because f_n' uniformly converge on $D_r(z_0)$, we know there exists N such that

$$\|f_n' - f_m'\|_\infty \leq \frac{\epsilon}{4} \text{ for all } n, m > N \tag{4.17}$$

We claim

$$\text{such } N \text{ works}$$

Fix $n, m > N$ and $u \in \overline{D_r(z_0)}$. We wish to prove

$$|\varphi_n(u) - \varphi_m(u)| \leq \epsilon$$

Denote $f_n - f_m : \overline{D_r(z_0)} \to \mathbb{C}$ by g. Because

$$|\varphi_n(u) - \varphi_m(u)| = \left|\frac{g(u) - g(z)}{u - z}\right| \leq \frac{\left|\text{Re}\Big(g(u) - g(z)\Big)\right|}{|u - z|} + \frac{\left|\text{Re}\Big(g(u) - g(z)\Big)\right|}{|u - z|}$$

WOLG, we only have to prove

$$\frac{\left|\text{Re}\Big(g(u) - g(z)\Big)\right|}{|u - z|} \leq \frac{\epsilon}{2}$$

Again, define $h : [0, 1] \to \mathbb{R}$ by

$$h(t) \triangleq \text{Re}\Big(g(tu + (1 - t)z_0)\Big)$$

Then by Chain Rule and matrix representation of derivative, we see that for all $t \in (0, 1)$

$$\begin{aligned} h'(t) = ac - bd \text{ where } u - z &= a + bi \\ \text{and } g'(tu + (1 - t)(z)) &= c + di \end{aligned}$$

Now, by MVT and Equation 4.17, we can deduce

$$\begin{aligned} \frac{\left|\text{Re}\Big(g(u) - g(z)\Big)\right|}{|u - z|} = \frac{|h(1) - h(0)|}{|u - z|} &= \frac{|h'(t)|}{|a + bi|} \text{ for some } t \in (0, 1) \\ &= \frac{|ac| + |bd|}{|a + bi|} \leq |c| + |d| \leq \frac{\epsilon}{2} \text{ (done)} \end{aligned}$$

Note that

$$\lim_{n\to\infty}\lim_{t\to x}\varphi_n(t) = \lim_{n\to\infty} f'(x) \text{ exists}$$

We can now exchange the limit and see that the derivative of f at x exists.

$$\begin{aligned} f'(x) = \lim_{t\to x}\varphi(t) &= \lim_{t\to x}\lim_{n\to\infty}\varphi_n(t) \\ &= \lim_{n\to\infty}\lim_{t\to x}\varphi_n(t) = \lim_{n\to\infty} f_n'(x) \text{ (done)} \end{aligned}$$

■

4.9 Basic Technique on Sequence and Series

Abstract

This section prove some basic result on sequence and series, which will be heavily used in next section on analytic functions and Chapter: Beauty. Although written in an almost glossary form, we present the Theorems in a structural order based on the necessity of notion of absolute convergence and limit superior. Note that in this section, z, v, w always represent complex numbers, and a, b, c always represent real numbers.

Theorem 4.9.1. (Weierstrass M-test) Given sequences $f_n : X \to \mathbb{C}$, and suppose

$$\forall n \in \mathbb{N}, \forall x \in X, |f_n(x)| \leq M_n$$

Then

$$\sum_{n=1}^{\infty} M_n \text{ converge } \implies \sum_{n=1}^{\infty} f_n \text{ uniformly converge}$$

Proof. The proof follows from noting

$$\forall x \in X, \left| \sum_{k=m}^{n} f_k(x) \right| \leq \sum_{k=m}^{n} |f_k(x)| \leq \sum_{k=m}^{n} M_k$$

∎

Note that in our proof of Weierstrass M-test, we reduce the proof for uniform convergence into uniform Cauchy, which is a technique we shall also use later in Abel's test for uniform convergence. We now prove summation by part, which is a result hold in all fields, and is the essence of the proof of Dirichlet's test and Abel's test for uniform convergence.

Theorem 4.9.2. (Summation by Part)

$$\begin{aligned} f_n g_n - f_m g_m &= \sum_{k=m}^{n-1} f_k \Delta g_k + g_k \Delta f_k + \Delta f_k \Delta g_k \\ &= \sum_{k=m}^{n-1} f_k \Delta g_k + g_{k+1} \Delta f_k \end{aligned}$$

Proof. The proof follows induction which is based on

$$f_m g_m + f_m \Delta g_m + g_m \Delta f_m + \Delta f_m \Delta g_m = f_{m+1} g_{m+1}$$

∎

Theorem 4.9.3. (Dirichlet's Test) Suppose

(a) $a_n \to 0$ monotonically.

(b) $\sum_{n=1}^N z_n$ is bounded.

We have

$$\sum a_n z_n \text{ converge}$$

Proof. Define $Z_n \triangleq \sum_{n=1}^N z_n$ and let M bound $|Z_n|$. Using summation by part by letting $f_k = a_k$ and $g_k = Z_{k-1}$, we have

$$\begin{aligned}
\left|\sum_{k=m}^n a_k z_k\right| &= \left|a_{n+1}Z_n - a_m Z_{m-1} - \sum_{k=m}^n Z_k(a_{k+1}-a_k)\right| \\
&\leq |a_{n+1}Z_n| + |a_m Z_{m-1}| + \left|\sum_{k=m}^n Z_k(a_{k+1}-a_k)\right| \\
(\because a_n \text{ is monotone }) \qquad &\leq M\Big(|a_{n+1}| + |a_m| + |a_{n+1}-a_m|\Big)
\end{aligned}$$

∎

Theorem 4.9.4. (Abel's Test for Uniform Convergence) Suppose $g_n : X \to \mathbb{R}$ is a uniformly bounded pointwise monotone sequence. Then given a sequence $f_n : X \to \mathbb{R}$,

$$\sum f_n \text{ uniformly converge } \implies \sum f_n g_n \text{ uniformly converge}$$

Proof. Define $R_n \triangleq \sum_{k=n}^\infty f_k$. Let M uniformly bound g_n. Because $R_n \to 0$ uniformly, we can let N satisfy

$$\forall n \geq N, \forall x \in X, |R_n(x)| < \frac{\epsilon}{6M}$$

Then for all $n, m \geq N$, using summation by part, we have

$$\begin{aligned}
\left|\sum_{k=m}^n f_k g_k\right| &= \left|\sum_{k=m}^n g_k \Delta R_k\right| \\
&\leq |R_{n+1}g_{n+1}| + |R_{m+1}g_{m+1}| + \sum_{k=m}^n |R_{k+1}\Delta g_k| \\
(\because g_n \text{ is pointwise monotone }) \qquad &\leq |R_{n+1}g_{n+1}| + |R_{m+1}g_{m+1}| + \frac{\epsilon}{6M}|g_{n+1}-g_m| \leq \epsilon
\end{aligned}$$

∎

Although the proofs of Dirichlet's test and Abel's test for uniform convergence are quite similar, one should note that the "ways" summation by part is applied are slightly different, as one use $R_n \triangleq \sum_{k=n}^{\infty} f_k$ instead of $\sum_{k=1}^{n} f_k$, like $Z_n \triangleq \sum_{j=1}^{n} z_j$. As corollaries of Dirichlet's test, one have the famous **alternating series test** and Abel's test for complex series.

Theorem 4.9.5. (Abel's Test for Complex Series) Suppose

(a) $\sum z_n$ converge.

(b) b_n is a bounded monotone sequence.

We have

$$\sum z_n b_n \text{ converge}$$

Proof. Denote $B \triangleq \lim_{n\to\infty} b_n$. By Dirichlet's Test, we know $\sum z_n(b_n - B)$ converge. The proof now follows form noting

$$\sum z_n b_n = \sum z_n(b_n - B) + B \sum z_n$$

∎

We now introduce the idea of absolute convergence, which we shall use throughout the remaining of the section. By a **permutation** $\sigma : E \to E$ on some set E, we merely mean σ is a bijective function. We say $\sum z_n$ **absolutely converge** if $\sum |z_n|$ converge, and say $\sum z_n$ **unconditionally converge** if for all permutation $\sigma : \mathbb{N} \to \mathbb{N}$, the series $\sum z_{\sigma(n)}$ converge and converge to the same value.

Theorem 4.9.6. (Absolutely Convergent Series Unconditionally Converge)

$$\sum z_n \text{ absolutely converge } \implies \sum z_n \text{ unconditionally converge}$$

Proof. The fact $\sum z_n$ converge follows from noting

$$\left| \sum_{k=n}^{m} z_k \right| \leq \sum_{k=n}^{m} |z_k| \leq \sum_{k=n}^{\infty} |z_k|$$

Now, fix ϵ and permutation σ. Let N_1 and N_2 satisfy

$$\sum_{n=N_1}^{\infty} |z_n| < \frac{\epsilon}{2} \text{ and } \left| \sum_{n=N}^{\infty} z_n \right| < \frac{\epsilon}{2} \text{ for all } N > N_2$$

Let $M \triangleq \max\{N_1, N_2\}$. Observe that for all $N > \max_{1\leq r\leq M} \sigma^{-1}(r)$, we have

$$\left| \sum z_n - \sum_{n=1}^{N} z_{\sigma(n)} \right| \leq \left| \sum_{n=M+1}^{\infty} z_n \right| + \sum_{n=M+1}^{\infty} |z_n| < \epsilon$$

■

Theorem 4.9.7. (Riemann Rearrangement Theorem) If $\sum a_n$ converge but not absolutely, then for each $L \in \overline{\mathbb{R}}$, there exists a permutation σ such that

$$\sum a_{\sigma(n)} = L$$

Proof. Define a_n^+ and a_n^- by

$$a_n^+ \triangleq \max\{a_n, 0\} \text{ and } a_n^- \triangleq \min\{a_n, 0\}$$

Because

$$\sum (a_n^+ + a_n^-) \text{ converge but } \sum (a_n^+ - a_n^-) = \infty$$

We know

$$\sum a_n^+ = \sum (-a_n^-) = \infty$$

WOLG, (why?), fix $L \in \mathbb{R}$ and suppose $a_n \neq 0$ for all n. Let $A = B = L$, and let two increasing sequence $\sigma^+, \sigma^- : \mathbb{N} \to \mathbb{N}$ satisfy

$$\sigma^+(k+1) = \min\{n \in \mathbb{N} : a_n > 0 \text{ and } n > \sigma^+(k)\}$$

and similar for σ^-. Now, recursively define p_k, q_k by

$$p_1 \text{ is the smallest number such that } \sum_{n=1}^{p_1} a_{\sigma^+(n)} \geq A \tag{4.18}$$

$$q_1 \text{ is the smallest number such that } \sum_{n=1}^{p_1} a_{\sigma^+(n)} + \sum_{n=1}^{q_1} a_{\sigma^-(n)} \leq B \tag{4.19}$$

$$p_{k+1} \text{ is the smallest number such that } \sum_{n=1}^{p_{k+1}} a_{\sigma^+(n)} + \sum_{n=1}^{q_k} a_{\sigma^-(n)} \geq A \tag{4.20}$$

$$q_{k+1} \text{ is the smallest number such that } \sum_{n=1}^{p_{k+1}} a_{\sigma^+(n)} + \sum_{n=1}^{q_{k+1}} a_{\sigma^-(n)} \leq B \tag{4.21}$$

We then define σ by

$$\sigma^+(1), \dots, \sigma^+(p_1), \sigma^-(1), \dots, \sigma^-(q_1), \sigma^+(p_1+1), \dots, \sigma^+(p_2), \sigma^-(q_1+1), \dots, \sigma^-(q_2), \dots$$

It then follows from

$$\left|\sum_{n=1}^{p} a_{\sigma^+}(n) + \sum_{n=1}^{q_k} a_{\sigma^-}(n) - L\right| \leq \min\{a_{\sigma^+(p_{k+1})}, \left|a_{\sigma^-(q_k)}\right|\} \text{ for all } p_k \leq p \leq p_{k+1}$$

and $a_n \to 0$ that $\sum a_{\sigma(n)} = L$. ■

Note that the method we deploy in the proof of Riemann rearrangement Theorem can be used to control the sequence to have arbitrary large set of subsequential limits by modifying the number of A, B in Equation (4.1), (4.2), (4.3) and (4.4).

Using Riemann rearrangement Theorem and equation

$$\max_{1\leq r\leq d} |x_n| \leq |\mathbf{x}| \leq \sum_{r=1}^{d} |x_r|$$

we can now generalize and strengthen Theorem 4.9.6 to

$$\begin{aligned}\sum \mathbf{x}_n \text{ absolutely converge } &\iff \sum_n x_{n,r} \text{ absolutely converge for all } r\\ &\iff \sum_n x_{n,r} \text{ unconditionally converge for all } r\\ &\iff \sum \mathbf{x}_n \text{ unconditionally converge}\end{aligned}$$

With this in mind, we can now well state the Fubini's Theorem for Double Series.

Theorem 4.9.8. (Fubini's Theorem for Double Series) If

$$\sum_n \sum_k |z_{n,k}| \text{ converge}$$

Then

$$\sum_{n,k} |z_{n,k}| \text{ converge and } \sum_{n,k} z_{n,k} = \sum_n \sum_k z_{n,k} = \sum_k \sum_n z_{n,k}$$

Proof. The fact $\sum z_{n,k}$ absolutely converge follow from

$$\sum_{n=1}^{N} \sum_{k=1}^{N} |z_{n,k}| \leq \sum_n \sum_k |z_{n,k}| \text{ for all } N$$

WOLG, it remains to prove

$$\sum_{n,k} z_{n,k} = \sum_n \sum_k z_{n,k}$$

Because $\sum_n \sum_k |z_{n,k}|$ converge, we can reduce the problem into proving the same statement for nonnegative series $a_{n,k}$. (why?)

$$\sum_n \sum_k |a_{n,k}| \text{ converge } \implies \sum_{n,k} a_{n,k} = \sum_n \sum_k a_{n,k}$$

Because

$$\sum_{n=1}^N \sum_{k=1}^N a_{n,k} \leq \sum_{n=1}^N \sum_k a_{n,k} \leq \sum_n \sum_k a_{n,k} \text{ for all } N$$

we see

$$\sum_{n,k} a_{n,k} \leq \sum_n \sum_k a_{n,k}$$

It remains to prove

$$\sum_{n,k} a_{n,k} \geq \sum_n \sum_k a_{n,k}$$

Fix N and ϵ. We reduce the problem into proving

$$\sum_{n,k} a_{n,k} \geq \sum_{n=1}^N \sum_k a_{n,k} - \epsilon$$

Let K satisfy

$$\text{For all } 1 \leq n \leq N, \ \sum_{k=K+1}^{\infty} a_{n,k} < \frac{\epsilon}{N}$$

It then follows

$$\sum_{n,k} a_{n,k} \geq \sum_{n=1}^N \sum_{k=1}^K a_{n,k} \geq \sum_{n=1}^N \sum_k a_{n,k} - \epsilon \text{ (done)}$$

■

Example 16 (Counter-Example for Fubini's Theorem for Double Series)

$$a_{n,k} \triangleq \begin{cases} 1 & \text{if } n = k \\ -1 & \text{if } n = k+1 \\ 0 & \text{if otherwise} \end{cases}$$

$$\sum |a_{n,k}| = \infty \text{ and } \sum_n \sum_k a_{n,k} = 1 \text{ and } \sum_k \sum_n a_{n,k} = 0$$

Theorem 4.9.9. (Merten's Theorem for Cauchy Product) Suppose

(a) $\sum_{n=0}^{\infty} z_n$ converge absolutely

(b) $\sum_{n=0}^{\infty} z_n = Z$

(c) $\sum_{n=0}^{\infty} v_n = V$

(d) $w_n = \sum_{k=0}^{n} z_k v_{n-k}$

Then we have

$$\sum_{n=0}^{\infty} w_n = ZV$$

Proof. We prove

$$\left| V \sum_{n=0}^{N} z_n - \sum_{n=0}^{N} w_n \right| \to 0 \text{ as } N \to \infty$$

Compute

$$\begin{aligned} V \sum_{n=0}^{N} z_n - \sum_{n=0}^{N} w_n &= \sum_{n=0}^{N} z_n (V - \sum_{k=0}^{N-n} v_k) \\ &= \sum_{n=0}^{N} z_n \sum_{k=N-n+1}^{\infty} v_k \end{aligned}$$

Because $\sum_{k=n}^{\infty} v_k \to 0$ as $n \to \infty$, we know there exists M such that

$$\left| \sum_{k=n}^{\infty} v_k \right| < M \text{ for all } n$$

Let N_0 satisfy

$$\sum_{n=N_0+1}^{\infty} |z_n| < \frac{\epsilon}{2M}$$

Let $N_1 > N_0$ satisfy

$$\left| \sum_{k=N-N_0+1}^{\infty} v_k \right| < \frac{\epsilon}{2(N_0+1)\sum_n |z_n|} \text{ for all } N > N_1$$

Now observe that for all $N > N_1$

$$\left| \sum_{n=0}^{N} z_n \Big(\sum_{k=N-n+1}^{\infty} v_k \Big) \right| \leq \sum_{n=0}^{N_0} |z_n| \left| \sum_{k=N-n+1}^{\infty} v_k \right| + \sum_{n=N_0+1}^{N} |z_n| \left| \sum_{k=N-n+1}^{\infty} v_k \right| < \epsilon \text{ (done)}$$

■

We first define the **limit superior** by

$$\limsup_{n\to\infty} a_n \triangleq \lim_{n\to\infty} (\sup_{k\geq n} a_k)$$

Note that $\limsup_{n\to\infty} a_n$ must exists because $(\sup_{k\geq n} a_k)_n$ is a decreasing sequence.

Theorem 4.9.10. (Equivalent Definition for Limit Superior) If we let E be the set of subsequential limits of a_n

$$E \triangleq \{L \in \overline{\mathbb{R}} : L = \lim_{k\to\infty} a_{n_k} \text{ for some } n_k\}$$

The set E is non-empty and

$$\max E = \limsup_{n\to\infty} a_n$$

Proof. Let $n_1 \triangleq 1$. Recursively, because

$$\sup_{j\geq n_k} a_k \geq \limsup_{n\to\infty} a_n > \limsup_{n\to\infty} a_n - \frac{1}{k} \text{ for each } k$$

We can let n_{k+1} be the smallest number such that

$$a_{n_{k+1}} > \limsup_{n\to\infty} a_n - \frac{1}{k}$$

It is straightforward to check $a_{n_k} \to \limsup_{n\to\infty} a_n$ as $k \to \infty$. Note that no subsequence can converge to $\limsup_{n\to\infty} a_n + \epsilon$ because there exists N such that $\sup_{k\geq N} a_k < \limsup_{n\to\infty} a_n + \epsilon$. ■

We can now state the **limit comparison test** as follows. Given a positive sequence b_n,

$$\limsup_{n\to\infty} \frac{|z_n|}{b_n} \in \mathbb{R} \text{ and } \sum b_n \text{ converge } \implies \sum z_n \text{ absolutely converge}$$
$$\liminf_{n\to\infty} \frac{b_n}{|z_n|} > 0 \text{ and } \sum z_n \text{ diverge } \implies \sum b_n \text{ diverge}$$

Theorem 4.9.11. (Geometric Series)

$$|z| < 1 \implies \sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$$

Proof. The proof follows from noting

$$(1-z)\sum_{n=0}^{N} z^n = 1 - z^{N+1} \to 1 \text{ as } N \to \infty$$

■

Theorem 4.9.12. (Ratio and Root Test)

$$\limsup_{n\to\infty} \sqrt[n]{|z_n|} < 1 \text{ or } \limsup_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| < 1 \implies \sum z_n \text{ absolutely converge}$$
$$\liminf_{n\to\infty} \sqrt[n]{|z_n|} > 1 \text{ or } \liminf_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| > 1 \implies \sum z_n \text{ diverge}$$

Proof. The convergent part follows from comparison to an appropriate geometric series and the diverge part follows from noting $|z_n|$ does not converge to 0. ■

Theorem 4.9.13. (Root Test is Stronger Than Ratio Test)

$$\liminf_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| \leq \liminf_{n\to\infty} \sqrt[n]{|z_n|} \leq \limsup_{n\to\infty} \sqrt[n]{|z_n|} \leq \limsup_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right|$$

Proof. Fix ϵ and WOLG suppose $\liminf_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| > 0$. We prove

$$\liminf_{n\to\infty} \sqrt[n]{|z_n|} \geq \liminf_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| - \epsilon$$

Let $\alpha \in \mathbb{R}$ satisfy

$$\liminf_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right| - \epsilon < \alpha < \liminf_{n\to\infty} \left|\frac{z_{n+1}}{z_n}\right|$$

Let N satisfy

$$\text{For all } n \geq N, \left| \frac{z_{n+1}}{z_n} \right| > \alpha$$

We then see

$$\sqrt[N+n]{|z_{N+n}|} \geq \sqrt[N+n]{|z_N| \, \alpha^n} = \alpha \Big(\frac{|z_N|^{\frac{1}{N+n}}}{\alpha^{\frac{N}{N+n}}} \Big) \to \alpha \text{ as } n \to \infty \text{ (done)}$$

The proof for the other side is similar. ∎

Theorem 4.9.14. (Root Test Trick) For all $k \in \mathbb{N}$

$$\limsup_{n\to\infty} |z_{n+k}|^{\frac{1}{n}} = \limsup_{n\to\infty} |z_n|^{\frac{1}{n}}$$

Proof. This is a direct corollary of equivalent definition for limit superior. ∎

> Lastly, we prove Cauchy's condensation Test, whose existence is almost solely for investigating p-Series.

Theorem 4.9.15. (Cauchy's Condensation Test) Suppose $a_n \searrow 0$. We have

$$\sum_{n=0}^{\infty} 2^n a_{2^n} \text{ converge } \iff \sum_{n=1}^{\infty} a_n \text{ converge}$$

Proof. Observe that for all $N \in \mathbb{N}$

$$\sum_{n=0}^{N} 2^n a_{2^n} \geq \sum_{n=0}^{N} \sum_{k=1}^{2^n} a_{2^n+k-1} = \sum_{n=1}^{2^{N+1}-1} a_n$$

and

$$2 \sum_{n=1}^{2^N-1} a_n = 2 \sum_{n=1}^{N} \sum_{k=0}^{2^{n-1}-1} a_{2^{n-1}+k} \geq 2 \sum_{n=1}^{N} 2^{n-1} a_{2^n} = \sum_{n=1}^{N} 2^n a_{2^n}$$

∎

Theorem 4.9.16. (p-Series)

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converge } \iff p > 1$$

Proof. Observe that

$$\sum_{n=0}^{\infty} 2^n \frac{1}{(2^n)^p} = \sum_{n=0}^{\infty} (2^{1-p})^n$$

The result then follows from Cauchy's Condensation Test and geometric series. ∎

4.10 Analytic Functions

Abstract

This section introduces the concept of analytic functions and proves some of their basic properties, including the Identity Theorem. We will rely on the tools developed in the previous section on sequences and series. Note that throughout this section, z will always denote a complex number.

In this section, by a **power series**, we mean a pair (z_0, c_n) where $z_0 \in \mathbb{C}$ is called the **center** of power series, and $c_n \in \mathbb{C}$ are the coefficients sequence. By **radius of convergence**, we mean a unique $R \in \mathbb{R}_0^+ \cup \infty$ such that

$$\sum_{n=0}^{\infty} c_n(z-z_0)^n \begin{cases} \text{converge absolutely} & \text{if } |z-z_0| < R \\ \text{diverge} & \text{if } |z-z_0| > R \end{cases}$$

Such R always exist (and is unique, the uniqueness can be checked without computing the actual value of R) and is exactly

$$R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{c_n}} \tag{4.22}$$

This result is called **Cauchy-Hadamard Theorem** and is proved by applying Root Test to $\sum c_n(z-z_0)^n$. Note that Cauchy-Hadamard Theorem does not tell us whether a power series converges at points of boundary of disk of convergence. It require extra works to determine if the power series converge at boundary.

Theorem 4.10.1. (Abel's Test for Power Series) Suppose $a_n \to 0$ monotonically and $\sum a_n z^n$ has radius of convergence R .

$$\text{The power series } \sum a_n z^n \text{ at least converge on } \overline{D_R(0)} \setminus \{R\}$$

Proof. Note that

$$\sum \frac{a_n}{R^n} z^n \text{ has radius of convergence } R$$

Fix $z \in \overline{D_R(0)} \setminus \{R\}$. Note that

$$\left| \sum_{n=0} \frac{z^n}{R^n} \right| = \left| \frac{1-(\frac{z}{R})^{N+1}}{1-\frac{z}{R}} \right| \leq \frac{2}{|1-\frac{z}{R}|} \text{ for all } N$$

It then follows from Dirichlet's Test that $\sum a_n (\frac{z}{R})^n$ converge. ■

Example 17 (Discussion of Convergence on Boundary)

$$f_q(z) = \sum_{n=0}^{\infty} n^q z^n \text{ provided } q \in \mathbb{R}$$

It is clear that f_q has convergence radius 1 for all $q \in \mathbb{R}$. For boundary, we have

$$\begin{cases} q < -1 \implies f_q \text{ converge on } S^1 \\ -1 \leq q < 0 \implies f_q \text{ converge on } S^1 \setminus \{1\} \\ 0 \leq q \implies f_q \text{ diverge on } S^1 \end{cases}$$

Note that

(a) At $z = 1$, the discussion is just p-series.

(b) $n^q \searrow 0$ if and only if $q < 0$; and if $n^q \searrow 0$, then the series converge by Abel's test for power series.

(c) If $q \geq 0$, $n^q z^n$ does not converge to 0 on $S^1 \setminus \{1\}$

Notice that the fact $\sum c_n(z-z_0)^n$ absolutely converge in $D_R(z_0)$ implies the convergence is uniform on all $\overline{D_{R-\epsilon}(z_0)}$ by M-Test. However, on $D_R(z_0)$, the convergence is not always uniform.

Example 18 (Failure of Uniform Convergence on $D_R(z_0)$)

$$f(z) = \sum_{n=0}^{\infty} z^n$$

Note $R = 1$. Use Geometric series formula to show $f(z) = \frac{1}{1-z}$ on $D_1(0)$. It is then clear that f is unbounded on $D_1(0)$ while all partial sums $\sum_{k=0}^{n} z^k$ is bounded on $D_1(0)$.

We now introduce some terminologies. We say a complex function f is **analytic at** $z_0 \in \mathbb{C}$ if f there exists a power series (z_0, c_n) whose convergence radius is greater than 0 and f agrees with $\sum_{n=0}^{\infty} c_n(z-z_0)^n$ on $D_R(z_0)$ for some R (of course, such R must not be strictly greater than the radius of convergence of (a, c_n)). It shall be quite clear that if f, g are both analytic at $z \in \mathbb{C}$ with radius $R_f \leq R_g$, then by Merten's Theorem for Cauchy product, $f + g$ and fg are analytic at z with radius at least R_f. We now investigate deeper into analytic functions.

Theorem 4.10.2. (Term by Term Differentiation) Given a power series (z_0, c_n) of convergence radius $R > 0$, if we define $f : D_R(z_0) \to \mathbb{C}$ by

$$f(z) \triangleq \sum_{n=0}^{\infty} c_n (z - z_0)^n$$

Then f is holomorphic on $D_R(z_0)$ and its derivative at z_0 is also a power series with radius of convergence R

$$f'(z) = \sum_{n=0}^{\infty} (n+1) c_{n+1} (z - z_0)^n$$

Proof. Because $(n+1)^{\frac{1}{n}} \to 1$, we can use Theorem 4.9.14 to deduce

$$\limsup_{n\to\infty} \left((n+1) \left| c_{n+1} \right| \right)^{\frac{1}{n}} = \limsup_{n\to\infty} \left| c_{n+1} \right|^{\frac{1}{n}} = \limsup_{n\to\infty} \left| c_n \right|^{\frac{1}{n}}$$

which implies that the power series $\sum_{n=0}^{\infty} (n+1) c_{n+1} (z - z_0)^n$ is of radius of convergence R. We now prove

$$f'(z) = \sum_{n=0}^{\infty} (n+1) c_{n+1} (z - z_0)^n \text{ on } D_R(z_0)$$

Define $f_m : D_R(z_0) \to \mathbb{C}$ by

$$f_m(z) \triangleq \sum_{n=0}^{m} c_n (z - z_0)^n$$

Observe

(a) $f_m \to f$ pointwise on $D_R(a)$

(b) $f_m'(z) = \sum_{n=0}^{m-1} (n+1) c_{n+1} (z - z_0)^n$ for all m

Fix $z \in D_R(z_0)$. Proposition (b) allow us to reduce the problem into proving

$$f'(z) = \lim_{m\to\infty} f_m'(z) \text{ on } D_R(a) \tag{4.23}$$

Let $z \in D_r(z_0)$ where $r < R$. With proposition (a) in mind, to show Equation 4.23, by Theorem 4.8.2, we only have to prove f_m' uniformly converge on $D_r(z_0)$, which follows from M-Test and the fact that $\sum_{n=0}^{\infty} (n+1) c_{n+1} (z - z_0)^n$ absolutely converge on $D_R(z_0)$. (done) ∎

Suppose

$$f(z) \triangleq \sum_{n=0}^{\infty} c_n (z-z_0)^n$$

Now by repeatedly applying Theorem 4.10.2, we see

$$f^{(k)}(z) = \sum_{n=0}^{\infty} (n+k)\cdots(n+1)c_{n+k}(z-z_0)^n \text{ for all } k \in \mathbb{Z}_0^+ \tag{4.24}$$

This then give us

$$c_k = \frac{f^{(k)}(z_0)}{k!} \text{ for all } k \in \mathbb{Z}_0^+ \tag{4.25}$$

and

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + \frac{f''(z_0)}{2!}(z-z_0)^2 + \cdots \text{ on } D_R(z_0) \tag{4.26}$$

Equation 4.26 is often called the **Taylor expansion of f at z_0**. Notably, Equation 4.25 tell us that if f is constant 0, then $c_n = 0$ for all n.

Example 19 (Smooth but not Analytic Function)

$$f(x) = \begin{cases} e^{\frac{-1}{x^2}} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

Use induction to show that

$$f^{(k)}(x) = P_k(\frac{1}{x})e^{-(\frac{1}{x})^2} \quad \exists P_k \in \mathbb{R}[x^{-1}], \forall k \in \mathbb{Z}_0^+, \forall x \in \mathbb{R}^*$$

and again use induction to show that

$$f^{(k)}(0) = 0 \quad \forall k \in \mathbb{Z}_0^+$$

The trick to show $f^{(k)}(0) = 0$ is let $u = \frac{1}{x}$.

Now, with Theorem 4.10.2, we see that f is not analytic at 0.

Example 20 (Bump Function)

$$f(x) = \begin{cases} e^{\frac{-1}{1-x^2}} & \text{if } |x| < 1 \\ 0 & \text{otherwise} \end{cases}$$

Use the same trick (but more advanced) to show f is smooth, and note that f is not analytic at ± 1.

Now, it comes an interesting question. Given a complex-valued function f analytic at z_0 with radius R, and suppose $z_1 \in D_R(z_0)$.

(a) Is f also analytic at z_1?

(b) What do we know about the radius of convergence of f at z_1?

(c) Suppose f is indeed analytic at z_1. It is trivial to see that the power series $(z_0, c_{0;n})$ and $(z_1, c_{1;n})$ must agree in the intersection of their convergence disks, and because f is given, we by Theorem 4.10.2 and Equation 4.25, have already known the value of $c_{1;n}$. Can we verify that the power series $(z_0, c_{0;n})$ and $(z_1, c_{1;n})$ do indeed agree with each other on the common convergence interval?

Taylor's Theorem for power series give satisfying answers to these problems.

Theorem 4.10.3. (Taylor's Theorem for Power Series) Given a function f analytic at z_0 with radius R, and suppose $z_1 \in D_R(z_0)$. Then

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_1)}{k!}(z - z_1)^k \text{ on } D_{R-|z_1-z_0|}(z_1)$$

Proof. WOLG, let $z_0 = 0$. Suppose z satisfy $|z - z_1| < R - |z_1|$. By Equation 4.25, we can compute

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} z^k \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} (z - z_1 + z_1)^k \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \sum_{n=0}^{k} \binom{k}{n} (z - z_1)^n z_1^{k-n} \\ &= \sum_{k=0}^{\infty} \sum_{n=0}^{k} \frac{f^{(k)}(0)}{k!} \binom{k}{n} (z - z_1)^n z_1^{k-n} \end{aligned}$$

Note that

$$\begin{aligned}
\sum_{k=0}^{\infty}\left|\sum_{n=0}^{\infty}\frac{f^{(k)}(0)}{k!}\binom{k}{n}(z-z_1)^n z_1^{k-n}\right| &\le \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\left|\frac{f^{(k)}(0)}{k!}\right|\binom{k}{n}|z-z_1|^n\cdot|z_1|^{k-n} \\
&= \sum_{k=0}^{\infty}\left|\frac{f^{(k)}(0)}{k!}\right|\sum_{n=0}^{\infty}\binom{k}{n}|z-z_1|^n\cdot|z_1|^{k-n} \\
&= \sum_{k=0}^{\infty}\left|\frac{f^{(k)}(0)}{k!}\right|\big(|z-z_1|+|z_1|\big)^k
\end{aligned}$$

is a convergent series, by Cauchy-Hadamard Theorem and $|z-z_1|+|z_1|<R$; thus, we can use Fubini's Theorem for double series to deduce

$$\begin{aligned}
\sum_{k=0}^{\infty}\sum_{n=0}^{k}\frac{f^{(k)}(0)}{k!}\binom{k}{n}(z-z_1)^n z_1^{k-n} &= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\frac{f^{(k)}(0)}{k!}\binom{k}{n}(z-z_1)^n z_1^{k-n} \\
&= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}\binom{k}{n}(z-z_1)^n z_1^{k-n} \\
&= \sum_{n=0}^{\infty}\sum_{k=n}^{\infty}\frac{f^{(k)}(0)}{k!}\binom{k}{n}(z-z_1)^n z_1^{k-n} \\
&= \sum_{n=0}^{\infty}\Big[\sum_{k=n}^{\infty}\frac{f^{(k)}(0)}{k!}\binom{k}{n}z_1^{k-n}\Big](z-z_1)^n
\end{aligned}$$

We have reduced the problem into proving

$$\sum_{k=n}^{\infty}\frac{f^{(k)}(0)}{k!}\binom{k}{n}z_1^{k-n}=\frac{f^{(n)}(z_1)}{n!}$$

Because z_1 is in $D_R(0)$, by Equation 4.24 and Equation 4.25, we can compute

$$\begin{aligned}
f^{(n)}(z_1) &= \sum_{k=0}^{\infty}(k+n)\cdots(k+1)\cdot\frac{f^{(n+k)}(0)}{(n+k)!}z_1^k \\
&= \sum_{k=n}^{\infty}(k)\cdots(k-n+1)\cdot\frac{f^{(k)}(0)}{k!}\cdot z_1^{k-n} \\
&= \sum_{k=n}^{\infty}\frac{f^{(k)}(0)}{(k-n)!}z_1^{k-n}
\end{aligned}$$

We now have

$$\frac{f^{(n)}(z_1)}{n!} = \sum_{k=n}^{\infty} \frac{f^{(k)}(0)}{n!(k-n)!} z_1^{k-n} = \sum_{k=n}^{\infty} \frac{f^{(k)}(0)}{k!} \binom{k}{n} z_1^{k-n} \text{ (done)}$$

■

Lastly, to close this section, we prove the Identity Theorem. By a **region** $D \subseteq \mathbb{C}$, we mean a connected and open subset of $\mathbb{C}$

Theorem 4.10.4. (Identity Theorem) Given two analytic complex-valued function $f, g : D \to \mathbb{C}$ defined on some region $D \subseteq \mathbb{C}$, if f, g agree on some subset $S \subseteq D$ such that S has a limit point in D, then f, g agree on the whole region D.

Proof. Define

$$T \triangleq \{z \in D : f^{(k)}(z) = g^{(k)}(z) \text{ for all } k \geq 0\}$$

Since D is connected, we can reduce the problem into proving T is non-empty, open and closed in D. Let c be a limit point of S in D. We first show

$$c \in T$$

Assume $c \notin T$. Let m be the smallest integer such that $f^{(m)}(c) \neq g^{(m)}(c)$. We can write the Taylor expansion of $f - g$ at c by

$$\begin{aligned}(f-g)(z) &= (z-c)^m \Big[\frac{(f-g)^{(m)}(c)}{m!} + \frac{(f-g)^{(m+1)}(c)}{(m+1)!}(z-c) + \cdots \Big] \\ &\triangleq (z-c)^m h(z)\end{aligned}$$

Clearly, $h(c) \neq 0$. Now, because h is continuous at c (h is a well-defined power series at c with radius greater than 0), we see h is non-zero on some $B_\epsilon(c)$, which is impossible, since $(f-g) \equiv 0$ on $S \setminus \{c\}$ implies $h = 0$ on $S \setminus \{c\}$. CaC (done)

Fix $z \in T$. Because f, g are analytic at z and $f^{(k)}(z) = g^{(k)}(z)$ for all k, we see $f - g$ is constant 0 on some open disk $B_\epsilon(z)$. We have proved that T is open. To see T is closed in D, one simply observe that

$$T = \bigcap_{k \geq 0} \{z \in D : (f-g)^{(k)}(z) = 0\}$$

and $(f-g)^{(k)}$ is continuous on D. ■

4.11 Abel's Theorem and its application

In this section, we use the notation $\mathbb{S}_M(R)$ to denote **stolz region**

$$\mathbb{S}_M(R) \triangleq \{z \in \mathbb{C} : \frac{|R-z|}{R-|z|} \in (0, M)\}$$

Theorem 4.11.1. (Abel's Theorem for Power Series) Given a complex Maclaurin series $f(z) = \sum_{n=0}^{\infty} c_n z^n$ of convergence radius R such that

$$\sum_{n=0}^{\infty} c_n R^n \text{ converge}$$

Then for all $M > 1$, we have

$$f|_{\mathbb{S}_M(R)}(z) \to \sum_{n=0}^{\infty} c_n R^n = f(R) \text{ as } z \to R$$

Proof. We first

$$\text{prove when } R = 1$$

Fix ϵ. We wish

$$\text{to find } \delta \text{ such that } \left| \sum_{n=0}^{\infty} c_n z^n - c_n \right| < \epsilon \text{ for all } z \in \mathbb{S}_M(1) \cap D_\delta(1)$$

To use summation by part, we first fix

$$s_n \triangleq \sum_{k=0}^{n} c_k \text{ and } s \triangleq \lim_{n\to\infty} s_n$$

Now Use summation by part

$$\begin{aligned}
\sum_{n=0}^{k} c_n z^n &= \sum_{n=0}^{k} (s_n - s_{n-1}) z^n \\
&= \sum_{n=0}^{k} s_n z^n - \sum_{n=0}^{k-1} s_n z^{n+1} \\
&= s_k z^k + (1-z) \sum_{n=0}^{k-1} s_n z^n
\end{aligned}$$

Note that

$$(1-z)\sum_{n=0}^{\infty} z^n = 1 \quad (|z|<1)$$

This give us

$$\begin{aligned}\lim_{z\to 1^-}\Big(\sum_{n=0}^{\infty} c_n z^n - \sum_{n=0}^{\infty} c_n\Big) &= \lim_{z\to 1^-}\Big(\lim_{k\to\infty} s_k z^k + (1-z)\sum_{n=0}^{k-1} s_n z^n - s\Big)\\ &= \lim_{z\to 1^-}(1-z)\sum_{n=0}^{\infty}(s_n - s)z^n \quad (\because \forall z\in\mathbb{C}:|z|<1, \lim_{k\to\infty} s_k z^k = 0)\end{aligned}$$

We reduce the problem into

$$\text{finding } \delta \text{ such that } \left|(1-z)\sum_{n=0}^{\infty}(s_n-s)z^n\right| \le \epsilon \text{ for all } z\in\mathbb{S}_M(1)\cap D_\delta(1)$$

Because $s_n\to s$, we know there exists N such that $|s_n - s| < \frac{1}{2M}$ for all $n>N$. We claim

$$\delta = \frac{\epsilon}{2\sum_{n=0}^{N}|s_n - s|} \text{ suffices}$$

Note that $\sum_{n=0}^{\infty}(s_n - s)z^n$ absolutely converges by direct comparison test. Then we can deduce

$$\begin{aligned}\left|(1-z)\sum_{n=0}^{\infty}(s_n-s)z^n\right| &= |1-z|\cdot\left|\sum_{n=0}^{N}(s_n-s)z^n + \sum_{n=N+1}^{\infty}(s_n-s)z^n\right|\\ &\le |1-z|\Big(\sum_{n=0}^{N}|s_n - s| + \frac{\epsilon}{2M}\sum_{n=N+1}^{\infty}|z|^n\Big)\\ &\le \frac{\epsilon}{2}+\frac{\epsilon}{2M}\cdot\frac{|1-z|}{1-|z|}\cdot|z|^{N+1}\le\epsilon \quad (\because |z|<1 \text{ and } \frac{|1-z|}{1-|z|}<M) \text{ (done)}\end{aligned}$$

We now prove

$$\text{when } R\in\mathbb{R}^+$$

Fix ϵ. We wish

$$\text{to find } \delta \text{ such that } \left|\sum_{n=0}^{\infty} c_n z^n - c_n R^n\right| < \epsilon \text{ for all } z\in\mathbb{S}_M(R)\cap D_\delta(R)$$

Fix

$$a_n = c_n R^n \text{ and } g(z) \triangleq \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} c_n R^n z^n \quad (|z| < 1)$$

By premise and our result, we know

$$g(1) \text{ exists and there exists } \delta' \text{ such that } |g(z) - g(1)| < \epsilon \text{ for all } z \in \mathbb{S}_M(1) \cap D_{\delta'}(1)$$

We claim

$$\delta = R\delta' \text{ suffices}$$

First note that

$$\frac{|R - z|}{R - |z|} \in (0, M) \implies \frac{\left|1 - \frac{z}{R}\right|}{1 - \left|\frac{z}{R}\right|} \in (0, M)$$

This tell us

$$z \in \mathbb{S}_M(R) \implies \frac{z}{R} \in \mathbb{S}_M(1)$$

Fix $z \in \mathbb{S}_M(R) \cap D_\delta(R)$. We now have

$$\frac{z}{R} \in \mathbb{S}_M(1) \cap D_{\delta'}(1)$$

This then let us conclude

$$\left| \sum_{n=0}^{\infty} c_n z^n - c_n R^n \right| = \left| g\left(\frac{z}{R}\right) - g(1) \right| < \epsilon \text{ (done)}$$

■

Example 21 (Identity of $\ln$ derived from Abel's Theorem)

$$\ln(1 + x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} \text{ for all } x \in (-1, 1]$$

Check that both side satisfy $y' = \frac{1}{1+x}$, and $y(0) = 0$. This tell us that two sides equal on $(-1, 1)$. Now using Abel's Theorem and the continuity of ln, we have

$$\ln 2 = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$$

Example 22 ($1-1+1-1+\cdots=\frac{1}{2}$)

$$1-1+1-1+\cdots=\frac{1}{2} \text{ is WRONG!!!}$$

When people say: "$1-1+1-1+\cdots=\frac{1}{2}$", they mean the sum of the series in the sense of Abel. Compute the Macularin series of $\frac{1}{1+r}$

$$\frac{1}{1+r}=\sum_{n=0}^{\infty}(-1)^n r^n$$

Check both side do equal on $(-1,1)$ by direct computation. Apply Abel's Theorem to see the magic.

4.12 L'Hospital Rule

Abstract

This section state and prove the L'Hospital Rule, and provide examples to show the necessity of each hypotheses of L'Hospital Rule. Note that although L'Hospital Rule is not really directly used in most results in Theory of Calculus, it is used in the proof of Taylor's Theorem.

Theorem 4.12.1. (L'Hospital Rule) Let $I \subseteq \mathbb{R}$ be an open interval containing c and let $f, g : I \to \mathbb{R}$ be two function continuous on I and differentiable on I everywhere except possibly at c, where

$$g'(x) \neq 0 \text{ for all } x \in I \setminus \{c\}$$

If $\frac{f}{g}$ is indeterminate form, i.e.

$$\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = L \text{ where } L \in \{0, \infty, -\infty\}$$

and

$$\lim_{x\to c} \frac{f'(x)}{g'(x)} \in \mathbb{R}$$

Then we have

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)} \tag{4.27}$$

Proof. Suppose $I = (a, b)$. Note that since $g'(x) \neq 0$ on (c, b), by MVT, we know there exists at most one $x \in (c, b)$ such that $g(x) = 0$. With similar argument for (a, c), we see that

$$g(x) \neq 0 \text{ on } (c - \epsilon, c + \epsilon) \setminus \{c\} \text{ for some } \epsilon$$

We now see that the expression $\lim_{x\to c} \frac{f(x)}{g(x)}$ is at least well-defined, and WOLG, we can suppose $g(x) \neq 0$ on $I \setminus \{c\}$. Define $m, M : I \setminus \{c\} \to \mathbb{R}$ by

$$m(x) \triangleq \inf \frac{f'(t)}{g'(t)} \text{ and } M(x) \triangleq \sup \frac{f'(t)}{g'(t)} \text{ where } t \text{ ranges over values between } x \text{ and } c$$

Because the value $\frac{f'(t)}{g'(t)}$ converge at c, we can deduce

$$\lim_{x\to c} m(x) = \lim_{x\to c} M(x) = \lim_{x\to c} \frac{f'(x)}{g'(x)} \tag{4.28}$$

We now prove

$$\text{the case when } \lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$$

Fix $x \in I \setminus \{c\}$, and WOLG suppose $x < c$. By Cauchy's MVT, we know that

$$m(x) \le \frac{f(x) - f(y)}{g(x) - g(y)} \le M(x) \text{ for all } y \in (x, c)$$

Note that $g(x) \ne g(y)$ by MVT, since $g' \ne 0$ on $I \setminus \{c\}$. It now follows from

$$\lim_{y\to c^-} \frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f(x)}{g(x)}$$

that

$$m(x) \le \frac{f(x)}{g(x)} \le M(x)$$

The proof of Equation 4.27 then follows from Equation 4.28. (done)

We now prove

$$\text{the case when } \lim_{x\to c} f(x) = \lim_{x\to c} g(x) = \infty$$

Again, fix $x \in I \setminus \{c\}$, and WOLG suppose $x < c$. By Cauchy's MVT, we know that

$$m(x) \le \frac{f(y) - f(x)}{g(y) - g(x)} \le M(x) \text{ for all } y \in (x, c)$$

Note that $g(x) \ne g(y)$ by MVT, since $g' \ne 0$ on $I \setminus \{c\}$. It now follows

$$\frac{f(y) - f(x)}{g(y) - g(x)} = \frac{\frac{f(y)}{g(y)} - \frac{f(x)}{g(y)}}{1 - \frac{g(x)}{g(y)}} \to \lim_{x\to c} \frac{f(x)}{g(x)} \text{ as } y \to c^-$$

The proof of Equation 4.27 then follows form Equation 4.28. (done) ■

We now introduce some examples to demonstrate the necessity of our hypotheses.

Chapter 5

Measure Theory

5.1 Sigma-Algebra

Abstract

In this section, we first discuss properties of σ-algebra and some of its substructure for better understanding of a slightly generalized version of Carathéodory's extension theorem. Note that in this section, the terms 'ring', 'field', or 'algebra' do not refer to algebraic structures like the integer ring.

Given a set X and an non-empty set R of subsets of X, we say R is a **semi-ring**, if for each $A, B \in R$, we have

(a) $A \cap B \in R$ (closed under finite intersection)

(b) $A \setminus B = \bigsqcup_{i=1}^{n} K_i$ for some disjoint $K_1, \dots, K_n \in R$. (relative complements can be written as finite disjoint union)

and we say R is a **ring**, if for each $A, B \in R$, we have

(a) $A \cup B \in R$ (closed under finite union)

(b) $A \setminus B \in R$ (closed under relative complement)

One should check

(a) Semi-ring always contain the empty set.

(b) Since $A \cap B = A \setminus (A \setminus B)$, closure under relative complement implies closure under finite intersection. Thus, a ring, or any collection closed under relative complement, is always a semi-ring.

(c) Note that $A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)$. This implies we can replace closure under finite union with closure under finite disjoint union as definition for ring.

(d) Given a family S of subsets of X, there exists smallest ring $R(S)$ containing S. Such ring $R(S)$ is called **the ring generated by** S.

Theorem 5.1.1. (Ring Generated by Semi-Ring) If S is a semi-ring, then

$$\begin{aligned} R(S) &= \{A : A \text{ is the union of some finite pair-wise disjoint sub-family } S' \text{ of } S\} \\ &= \{A : A \text{ is the union of some finite sub-family } S' \text{ of } S\ \} \end{aligned}$$

Proof. Let $R \triangleq \{A : A \text{ is the union of some finite pair-wise disjoint sub-family } S' \text{ of } S\}$. $R' \triangleq \{A : A \text{ is the union of some finite sub-family } S' \text{ of } S\ \}$. We first prove $R = R(S)$.

Clearly, the problem can be reduced into proving R is a ring. Because R is clearly closed under finite disjoint union, by property (c) of the ring, we can reduce the problem into proving R is closed under relative complement.

We first show R is closed under finite intersection. Given $\bigsqcup E_i, \bigsqcup F_j \in R$, we see

$$\Big(\bigsqcup_i E_i\Big) \cap \Big(\bigsqcup_j F_j\Big) = \bigsqcup_{i,j} E_i \cap F_j \in R \text{ (done)}$$

Now observe

$$\begin{aligned} \Big(\bigsqcup_i E_i\Big) \setminus \Big(\bigsqcup_j F_j\Big) &= \bigcap_j \Big(\bigsqcup_i (E_i \setminus F_j)\Big) \\ &= \bigcap_j A_j \text{ for some } A_j \in R \quad \text{(done)} \end{aligned}$$

We now prove $R = R'$. It is clear that $R \subseteq R'$. We only have to prove $R' \subseteq R$. This is trivially true, since any finite sub-family S' of S is a finite sub-family of R and R is closed under finite union. (done) ∎

We now give definition to the most important structure in this section. Given a family Σ of subsets of X, we say Σ is a **σ-algebra** (or sometimes **σ-field**) on X and say (X, Σ) is a **measurable space**, if Σ is ring and for each sequence $(A_n)_{n \in \mathbb{N}}$ of elements of Σ we have

(a) $X \in \Sigma$

(b) $\bigcup_{n\in\mathbb{N}} A_n \in \Sigma$ (Closed under countable union)

Similarly, one should check

(a) Because $\bigcap_{n=1}^{\infty} A_n = A_1 \setminus (\bigcup_{n=2}^{\infty}(A_1 \setminus A_n))$, we see σ-algebra is closed under countable intersection.

(b) Because $\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty}(A_n \setminus (\bigcup_{k=n+1}^{\infty} A_k))$, we can replace closure under countable union with closure under countable disjoint union.

(c) Because $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}}(\bigcup_{k=1}^{n} A_k)$, we can replace closure under countable union with the condition for all $(A_n) \subseteq \Sigma$ such that $\forall n, A_n \subseteq A_{n+1}$, we have $\bigcup_n A_n \in \Sigma$.

(d) Given a family S of subsets of X, there exists smallest sigma-algebra $\sigma(S)$ containing S. Such σ-algebra $\sigma(S)$ is called **the sigma-algebra generated by** S.

Now, given a family S of subsets of X, there in fact exists an explicit expression of $\sigma(S)$, albeit infamous. Let ω_1 be the smallest uncountable ordinal, and let $\mathbf{\Sigma}_1^0 \triangleq S$. For each ordinal $\alpha < \omega_1$, we recursively define $\mathbf{\Pi}_\alpha^0 \triangleq \{X \setminus A : A \in \mathbf{\Sigma}_\alpha^0\}$ and $\mathbf{\Sigma}_\alpha^0 \triangleq \{\bigcup_{n\in\mathbb{N}} A_n : (A_n) \subseteq \bigcup_{1\leq\gamma<\alpha} \mathbf{\Pi}_\gamma^0\}$. One can use transfinite induction to check that $\sigma(S) = \bigcup_{\alpha<\omega_1} \mathbf{\Sigma}_\alpha^0$.

5.2 Carathéodory's Extension Theorem

Abstract

In this section, we introduce a general process to construct measure on X. The process involve first inducing an outer measure from a pre-measure on some weaker structure S, and then restricting the outer measure onto a subfamily that contain exactly all the subset that "sharply cut" all the other subsets of X. The subfamily, as we shall prove, is a sigma-algebra. This extending-restricting process is known mostly by the name of Carathéodory Extension Theorem, and the rigorous definition of "sharply cut" is known by the name of Carathéodory criterion. The selected 'some weaker structure' S, which we begin out extension from, is a semi-ring. Although as long as S contain the empty set, the process works, in the sense that one can generate a measure with pre-measure $\mu : S \to [0, \infty]$, to have the generated measure agree with μ on S, some necessary condition needs to be satisfied, and the axioms of semi-ring, not equivalent to some other popular choices that also suffice, e.g., ring or quasi-semi-ring, suffices to be a set of necessary conditions.

Note that in this section, if we write $\mu(A)$ without specifying whether A is in the domain of μ, we mean that the statement always hold true as long as A is in the domain of μ, and note that the difference between the term **measure space** and measurable space lies in that the latter is not equipped with a measure yet.

Given a collection S of subsets of X containing the empty set, we say $\mu : S \to [0, \infty]$ is a **pre-measure** (or **content**) on (X, S) if

(a) $\mu(\varnothing) = 0$ (null empty set)

(b) $A \subseteq B \implies \mu(A) \leq \mu(B)$ (monotone)

(c) $\mu\big(\bigsqcup_{n\in\mathbb{N}} E_n\big) = \sum_{n\in\mathbb{N}} \mu(E_n)$ (countably additive, or σ-additive)

and say μ is a **measure** if S is a σ-algebra on X. Note that if S is closed under relative complement, e.g., S is a ring or any stronger structure, then monotone is implied by countable additive. Now, with the hints given below, one can check straightforward that if S is a semi-ring, we have

(a) $\mu(A_1 \sqcup \cdots \sqcup A_n) = \mu(A_1) + \cdots + \mu(A_n)$ (finitely additive)

(b) $\mu\big(\bigcup_{n\in J} A_n\big) \leq \sum_{n\in J} \mu(A_n)$, for each finite or countable J.

(c) $A_n \nearrow A \implies \mu(A_n) \nearrow \mu(A)$.

(d) $A_n \searrow A$ and $\mu(A_1) < \infty$ and S is a σ-algebra $\implies \mu(A_n) \searrow \mu(A)$

Hints: Properties (b) and (c) are proved by letting $B_n \triangleq A_n \setminus (A_{n-1} \cup \cdots \cup A_1)$; and property (d) is proved by letting $B_n \triangleq A_1 \setminus A_n, \forall n \in \mathbb{N}$ and the equation

$$\begin{aligned}
\mu(A_1) - \lim_{n\to\infty} \mu(A_n) &= \lim_{n\to\infty} \mu(A_1) - \mu(A_n) \\
&= \lim_{n\to\infty} \mu(B_n) \\
\text{(property (c) is used here)} \quad &= \mu(\bigcup_{n=1}^{\infty} B_n) = \mu(A_1 \setminus \bigcap_{n=1}^{\infty} A_n) = \mu(A_1) - \mu(\bigcap_{n=1}^{\infty} A_n)
\end{aligned}$$

Note that our proof for property (d) require A_1 to be of finite measure in last step.

Now, suppose S is a semi-ring and $\mu : S \to [0,\infty]$ is a pre-measure on (X, S). This section address the question: Is there a unique extension of μ onto $\sigma(S)$? The answer is indeed affirmative: The extension always exists, and if μ is σ-finite, the extension is unique. While extensions of pre-measures can be constructed from structures weaker than a semi-ring, we focus on semi-rings here, as they are the starting points in common applications, e.g., Lebesgue-Stieltjes measure.

We first extend μ from S onto $R(S)$. Since each element of $R(S)$ is a union of some finite pair-wise disjoint sub-family of S, as we proved before, for each $A = \bigsqcup_{j=1}^{n_a} A_j \in R(S)$, we can assign $\mu(A) \triangleq \sum_{j=1}^{n_a} \mu(A_j)$. Such assignment is well-defined, as one can check using

$$\bigsqcup_{j=1}^{n_a} A_j = \bigsqcup_{k=1}^{n_b} B_k \implies A = \bigsqcup_{j,k} A_j \cap B_k \text{ where } A_j \cap B_k \in S$$

At this point, one should check μ remains countably additive after the extension onto $R(S)$ by dissecting $A = \bigsqcup A_n \in R(S)$ into a countable disjoint union of element in S.

We now give definition to **outer measure**, and shows that given any pre-measure μ on some semi-ring S of subsets of X, there exists outer measure μ^* on X such that μ^* and μ agrees on S.

Definition 5.2.1. (Definition of outer measure) Given a set X, by an **outer measure**, we mean a function $\nu : 2^X \to [0,\infty]$ such that

(a) $\nu(\varnothing) = 0$ (null empty set)

(b) $A \subseteq \bigcup_{n\in\mathbb{N}} A_n \implies \nu(A) \leq \sum \nu(A_n)$ (countably subadditive)

Equivalently, one can replace countably subadditive with the following two axioms

(a) $A \subseteq B \implies \nu(A) \leq \nu(B)$ (monotone)

(b) $\nu(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum \nu(A_n)$

Theorem 5.2.2. (Pre-measure on semi-ring induces outer measure) Given a pre-measure μ on some semi-ring S of subsets of X, if we define $\mu^* : 2^X \to [0,\infty]$ by

$$\mu^*(E) \triangleq \inf \Big\{ \sum_n \mu(T_n) : E \subseteq \bigcup_n T_n \text{ and } T_1, T_2, \dots \in S \Big\} \text{ where } \inf \varnothing = \infty$$

Then

$$\mu^* \text{ is an outer measure agreeing with } \mu \text{ on } R(S)$$

Proof. It is clear $\mu^*(\varnothing) = 0$ and $A \subseteq B \implies \mu^*(A) \leq \mu^*(B)$. It remains to prove that for arbitrary B_n we have

$$\mu^*\Big(\bigcup_n B_n\Big) \leq \sum_n \mu^*(B_n)$$

If $\sum_n \mu^*(B_n) = \infty$, the proof is trivial. We from now suppose $\sum_n \mu^*(B_n) < \infty$. Fix ϵ. We prove

$$\mu^*\Big(\bigcup_n B_n\Big) \leq \sum_n \mu^*(B_n) + \epsilon$$

Because $\mu^*(B_n) < \infty$ for each $n \in \mathbb{N}$, we know for each $n \in \mathbb{N}$ there exists an countable cover $(T_{n,k})_{k\in\mathbb{N}} \subseteq R$ of B_n such that

$$\sum_k \mu(T_{n,k}) \leq \mu^*(B_n) + \epsilon(2^{-n})$$

It is clear that $\{T_{n,k} : n, k \in \mathbb{N}\}$ is a countable cover of $\bigcup_n B_n$, we now see

$$\begin{aligned} \mu^*\Big(\bigcup_n B_n\Big) &\leq \sum_{n,k} \mu(T_{n,k}) = \sum_n \sum_k \mu(T_{n,k}) \quad (\because \text{Fubini's Theorem for Double Series}) \\ &\leq \sum_n \mu^*(B_n) + \epsilon(2^{-n}) = \sum_n \mu^*(B_n) + \epsilon \text{ (done)} \end{aligned}$$

At this point, one should check that even if we first extend μ onto $R(S)$ before doing the same procedure, we would still have the same outer measure. Now, because for each $T \in R(S)$, we clearly have $\mu^*(T) \leq \mu(T)$, to finish the proof it only remains to see for each cover $T_n \subseteq R(S)$ of T, we have

$$\mu(T) \leq \sum_n \mu(T \cap T_n) \leq \sum_n \mu(T_n)$$

∎

So far, we have proved that given a semi-ring S and a pre-measure $\mu : S \to [0, \infty]$, there exist some outer measure agree with μ on $R(S)$. One may wish to ask if such outer measure, an extension of the pre-measure μ, is unique? The answer is negative even in the most trivial case, as the example below shows. In fact, the outer measure induced in Theorem 5.2.2 is called the **maximal outer extension**, in the sense that if ν is an outer measure agreeing with μ on $R(S)$, then $\nu(E) \leq \mu^*(E), \forall E \subseteq X$, as one can check straightforwardly. Also, one can check that it make no difference if we first extend μ from S onto $R(S)$ or not, before we extend μ to the maximal outer measure μ^*.

Example 23 (Non-uniqueness of outer extension)

$$X \triangleq \{1, 2\} \text{ and } R \triangleq \{\varnothing, X\}$$

Define the pre-measure $\mu : R \to [0, \infty]$ and an outer measure $\nu : \mathcal{P}(X) \to [0, \infty]$ agreeing with μ on S by

$$\mu(A) \triangleq \begin{cases} 0 & \text{if } A = \varnothing \\ 1 & \text{if } A = X \end{cases} \text{ and } \nu(A) \triangleq \begin{cases} \mu(A) & \text{if } A \in R \\ \frac{1}{2} & \text{if } A \notin R \end{cases}$$

One can check that the maximal outer extension μ^* disagree with ν on $\mathcal{P}(X) \setminus R$.

It is important to note that the final step, Theorem 5.2.3, inducing measure from outer measure, is a general Theorem and operate independently of Theorem 5.2.2. This distinction is crucial because many constructions of measures, such as the Hausdorff measure, begin by defining an outer measure explicitly, rather than inducing it from a weaker structure like a semi-ring.

Theorem 5.2.3. (Outer measure induce measure) Given an outer measure μ^* on X, if we let

$$\mathcal{A} \triangleq \{A \subseteq X : \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A) \text{ for all } E \subseteq X \}$$

then $\mathcal{A}$ is a sigma-algebra on X and $\mu^*|_{\mathcal{A}} : \mathcal{A} \to [0, \infty]$ is a measure.

Proof. Because of the following facts

(a) $A \setminus B = A \cap B^c$

(b) $A \cup B = (A^c \cap B^c)^c$

(c) property (c) of sigma-algebra

we can reduce the problem into proving the following propositions

(i) $\mathcal{A}$ is closed under complement.

(ii) $X \in \mathcal{A}$.

(iii) $\mathcal{A}$ is closed under finite intersection.

(iv) $\mu^*|_{\mathcal{A}}$ is countably additive. (Thus at least form a pre-measure)

(v) $\mathcal{A}$ is closed under countable disjoint union.

We will prove the propositions sequentially, as the proof of each subsequent proposition may rely on the proofs of the preceding ones. The first two are straightforward to check. We now prove $\mathcal{A}$ is closed under finite intersection. Fix $A, B \in \mathcal{A}$ and $E \subseteq X$, we wish to show

$$\mu^*(E) = \mu^*(E \cap A \cap B) + \mu^*\big(E \setminus (A \cap B)\big)$$

Because $B \in \mathcal{A}$, we can "sharply cut $E \setminus (A \cap B)$ by B"; that is

$$\mu^*(E \setminus (A \cap B)) = \mu^*\big((E \cap B) \setminus A\big) + \mu^*(E \setminus B) \tag{5.1}$$

Equation 5.1 together with $A, B \in \mathcal{A}$ then give us

$$\begin{aligned}\mu^*(E \cap A \cap B) + \mu^*\big(E \setminus (A \cap B)\big) &= \mu^*(E \cap A \cap B) + \mu^*\big((E \cap B) \setminus A\big) + \mu^*(E \setminus B) \\ &= \mu^*(E \cap B) + \mu^*(E \setminus B) = \mu^*(E) \text{ (done)}\end{aligned}$$

We now prove the claim: For each pairwise disjoint sequence $(A_n) \subseteq \mathcal{A}$ and $E \subseteq X$, we have the equality

$$\mu^*\Big(E \cap \bigsqcup_n A_n\Big) = \sum_n \mu^*(E \cap A_n)$$

The countably subadditivty of μ^* trivially implies the inequality

$$\mu^*\Big(E \cap \bigsqcup_n A_n\Big) \leq \sum_n \mu^*(E \cap A_n)$$

Using induction and the fact $(A_n) \subseteq \mathcal{A}$, we see that

$$\mu^*(E \cap \bigsqcup_n A_n) = \sum_{n=1}^{N} \mu^*(E \cap A_n) + \mu^*\Big(E \cap \bigsqcup_{n=N+1}^{\infty} A_n\Big) \text{ for all } N \in \mathbb{N}$$

Then since μ^* has codomain $[0, \infty]$, we see

$$\mu^*(E \cap \bigsqcup_n A_n) \geq \sum_{n=1}^{N} \mu^*(E \cap A_n) \text{ for all } N \in \mathbb{N}$$

This implies the desired inequality $\mu^*(E \cap \bigsqcup_n A_n) \geq \sum_{n=1}^{\infty} \mu^*(E \cap A_n)$. (done)

Using $E \triangleq \bigsqcup_n A_n$, one see our claim implies that $\mu^*|_{\mathcal{A}}$ is countably additive. Lastly, we prove $\mathcal{A}$ is closed under countable disjoint union. Fix a pairwise disjoint sequence $(A_n) \subseteq \mathcal{A}$ and $E \subseteq X$. We wish to prove

$$\mu^*(E) \geq \mu^*(E \cap \bigsqcup_n A_n) + \mu^*(E \setminus \bigsqcup_n A_n)$$

Using induction and the fact $(A_n) \subseteq \mathcal{A}$, we see that

$$\mu^*(E \cap \bigsqcup_{n=1}^{N} A_n) = \sum_{n=1}^{N} \mu^*(E \cap A_n) \text{ for all } N \in \mathbb{N}$$

Then our claim give us

$$\mu^*(E \cap \bigsqcup_{n=1}^{N} A_n) \to \mu^*(E \cap \bigsqcup_n A_n) \text{ as } N \to \infty \tag{5.2}$$

Now, because of the identity $F \cup G = (F^c \cap G^c)^c$, proposition (i) and (iii) have shown $\mathcal{A}$ is closed under finite union. This implies $\bigsqcup_{n=1}^{N} A_n \in \mathcal{A}$ for all $N \in \mathbb{N}$, which, together with monotone of μ^*, give

$$\begin{aligned}\mu^*(E \cap \bigsqcup_{n=1}^{N} A_n) + \mu^*(E \setminus \bigsqcup_n A_n) &\leq \mu^*(E \cap \bigsqcup_{n=1}^{N} A_n) + \mu^*(E \setminus \bigsqcup_{n=1}^{N} A_n) \\ &\leq \mu^*(E) \text{ for all } N \in \mathbb{N} \end{aligned} \tag{5.3}$$

Equation 5.2 and Equation 5.3 gives the desired inequality. (done) ∎

Theorem 5.2.2 together with Theorem 5.2.3 shows that for each pre-measure μ on semi-ring S, we can induce a measure $\mu^*|_{\mathcal{A}}$ agreeing with μ on S. Although this result is correct, it doesn't show $S \subseteq \mathcal{A}$, which is necessary to refer to $\mu^*|_{\mathcal{A}}$ as an extension. However, it is straightforward to verify that $S \subseteq \mathcal{A}$ using the definition of a semi-ring and the property that $\mu^*(T) = \mu(T)$ for all $T \in S$.

Moving forward, here are some additional concepts we will utilize in subsequent sections: Given a measurable space (X, Σ, μ), we define a measurable set $N \in \Sigma$ as a **null set** if $\mu(N) = 0$. Moreover, we say that μ is a **complete measure** if every subset of a null set is also measurable. It is important to note that every measure induced by an outer measure is complete, as one can readily verify.

Lastly, we wish to ask: Given a sigma-algebra Σ containing S and contained by $\mathcal{A}$, under what condition, is Carathéodory the only extension of μ onto Σ ?

This question turns out to have direct connection with the notion named 'σ-finite'. Given a pre-measure space (X, S, μ), we say μ is **σ-finite** if there exists a countable cover $(A_n) \subseteq S$ of E such that $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$. It is clear that

(a) μ is σ-finite only if S form a cover of X.

(b) If $\mu : S \to [0, \infty]$ is σ-finite and ν is a pre-measure defined on a class larger than S, such that ν agree with μ on S, then ν is also σ-finite.

Theorem 5.2.4. (Uniqueness of Extension) Suppose

(a) (X, S, μ) is a pre-measure space, and S is a semi-ring.

(b) $\mathcal{A}$ is the induced sigma-algebra in Theorem 5.2.3

(c) $\Sigma \subseteq \mathcal{A}$ is a sigma-algebra containing S

(d) $\nu : \Sigma \to [0, \infty]$ is a measure agreeing with μ on S

We have

$$\nu(A) \leq \mu^*(A) \text{ for all } A \in \Sigma \tag{5.4}$$

and, if μ is σ-finite, we have

$$\nu(A) = \mu^*(A) \text{ for all } A \in \Sigma \tag{5.5}$$

Proof. The inequality 5.4 follows from the greatest-lower-bound definition of induced outer-measure, property (b) of measure and monotone of measure.

From now on, we suppose μ is σ-finite. Before we prove Equation 5.5, we first prove the claim: for each $A \in \Sigma$, there exists a pairwise disjoint sequence $(D_n) \subseteq R(S) \subseteq \Sigma$ such that $A \subseteq \bigsqcup_n D_n$ and $\nu(D_n) = \mu^*(D_n) < \infty$ for all $n \in \mathbb{N}$.

Because μ is σ-finite, there exists a sequence $(A_n) \subseteq S$ such that $A \subseteq \bigcup_n A_n$ and $\mu(A_n) < \infty, \forall n \in \mathbb{N}$. Define $D_1 \triangleq A_1$ and $D_n \triangleq A_n \setminus (A_1 \cup \cdots \cup A_{n-1})$ for all $n > 2$. Noting the structure of $R(S)$, it is clear that D_n is a pairwise disjoint sequence in $R(S)$. It is also clear that $A \subseteq \bigcup A_n = \bigsqcup D_n$. Fix $n \in \mathbb{N}$. It remains to prove

$$\nu(D_n) = \mu^*(D_n) < \infty$$

The inequality $\mu^*(D_n) < \infty$ follows from $\mu^*(D_n) \leq \mu(A_n) < \infty$, and the equation $\nu(D_n) = \mu^*(D_n)$ follows from $R(S) \subseteq \Sigma$ and $R(S) \subseteq \mathcal{A}$. (done)

Note that for all $n \in \mathbb{N}$, $A \cap D_n \in \Sigma \subseteq \mathcal{A}$. Now, since

$$\nu(A) = \sum_{n=1}^{\infty} \nu(A \cap D_n) \text{ and } \mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A \cap D_n)$$

To prove Equation 5.5, it only remains to prove $\nu(A \cap D_n) = \mu^*(A \cap D_n), \forall n \in \mathbb{N}$.

Because ν is a measure and $A \cap D_n \in \mathcal{A}$, we have the equations set

$$\begin{cases} \nu(A \cap D_n) = \nu(D_n) - \nu(D_n \setminus A) \\ \mu^*(A \cap D_n) = \mu^*(D_n) - \mu^*(D_n \setminus A) \end{cases} \tag{5.6}$$

The proof then follows from the facts

(a) $\nu(D_n) = \mu^*(D_n) < \infty$

(b) $\nu(A \cap D_n) \leq \mu^*(A \cap D_n)$ and $\nu(D_n \setminus A) \leq \mu^*(D_n \setminus A)$ (done)

Note that fact (b) can be checked straightforwardly. ■

5.3 Equivalent Definition of Lebesgue Measure

Abstract

It is clear that the collection S of half-open interval $\prod[a_j,b_j)$ form a semi-ring. If we define a volume function $\mu:S\to[0,\infty]$ on S by

$$\mu\Big(\prod_{j=1}^d[a_j,b_j)\Big)\triangleq\prod_{j=1}^d(b_j-a_j)$$

we see that the empty set is indeed null and μ is indeed monotone. To check that μ form a pre-measure, it remains to prove

$$\prod_{j=1}^d[a_j,b_j)=\bigsqcup_{n=1}^\infty\prod_{j=1}^d[a_{j,n},b_{j,n})\implies\mu\Big(\prod_{j=1}^d[a_j,b_j)\Big)=\sum_{n=1}^\infty\mu\Big(\prod_{j=1}^d[a_{j,n},b_{j,n})\Big)$$

To check

$$\mu\Big(\prod_{j=1}^d[a_j,b_j)\Big)\geq\sum_{n=1}^\infty\mu\Big(\prod_{j=1}^d[a_{j,n},b_{j,n})\Big)\tag{5.7}$$

One fix arbitrary N and cut $\prod_{j=1}^d[a_j,b_j)$ into finite amount of grids to see

$$\mu\Big(\prod_{j=1}^d[a_j,b_j)\Big)\geq\sum_{n=1}^N\mu\Big(\prod_{j=1}^d[a_{j,n},b_{j,n})\Big)$$

To check

$$\mu\Big(\prod_{j=1}^d[a_j,b_j)\Big)\leq\sum_{n=1}^\infty\mu\Big(\prod_{j=1}^d[a_{j,n},b_{j,n})\Big)\tag{5.8}$$

One fix ϵ and have each ϵ_n satisfy

$$\mu\Big(\prod_{j=1}^d[a_{j,n}-\epsilon_n,b_{j,n}+\epsilon_n)\Big)\leq(1+\epsilon)\mu\Big(\prod_{j=1}^d[a_{j,n},b_{j,n})\Big)$$

Then because

$$\Big\{\prod_{j=1}^d (a_{j,n}-\epsilon_n, b_{j,n}+\epsilon_n) \subseteq \mathbb{R}^d : n \in \mathbb{N}\Big\} \text{ form an open cover for compact } \prod_{j=1}^d [a_j, b_j]$$

there exists finite subset $I \subseteq \mathbb{N}$ such that

$$\prod_{j=1}^d [a_j, b_j] \subseteq \bigcup_{n\in I}\prod_{j=1}^d [a_{j,n}-\epsilon_n, b_{j,n}+\epsilon_n)$$

This then give us

$$\mu\Big(\prod_{j=1}^d [a_j, b_j)\Big) \leq \sum_{n\in I}\mu\Big(\prod_{j=1}^d [a_{j,n}-\epsilon_n, b_{j,n}+\epsilon_n)\Big) \leq (1+\epsilon)\sum_{n=1}^\infty \mu\Big(\prod_{j=1}^d [a_{j,n}, b_{j,n})\Big)$$

which give us Equation 5.8. Having proved that the volume function $\mu : S \to [0,\infty]$ is a pre-measure, we can induce a outer measure on $\mathbb{R}^d$ by

$$|E|_e \triangleq \inf\Big\{\mu(T_n) : E \subseteq \bigcup_n T_n \text{ and } T_1, T_2, \dots \in S\Big\}$$

and restrict the outer measure into a measure by letting the collection $\mathcal{L}(\mathbb{R}^d)$ of Lebesgue measurable set to be

$$\mathcal{L}(\mathbb{R}^d) \triangleq \{A \subseteq \mathbb{R}^d : \forall E \subseteq \mathbb{R}^d, |E|_e = |E\cap A|_e + |E \cap A^c|_e\}$$

Notably, if we define the class K_n of half-open dyadic cubes by

$$K_n \triangleq \Big\{\prod_{j=1}^d [\frac{m_j}{2^n}, \frac{m_j+1}{2^n}) \subseteq \mathbb{R}^d : m_j \in \mathbb{Z} \text{ for all } j\Big\}$$

We see that for each p in some open set U, there exists some small enough half-open dyadic cube $Q \in K_n$ (for some n) such that $p \in Q \subseteq U$. Then, since the collection $\bigcup K_n$ of half-open dyadic cubes is countable, we see U is in the sigma-algebra $\mathcal{L}(\mathbb{R}^d)$.

Theorem 5.3.1. (Equivalent Definition of Lebesgue Measurability) The following statements are equivalent.

(a) $E \in \mathcal{L}(\mathbb{R}^d)$.

(b) For all ϵ, there exists some open O containing E such that $|O \setminus E|_e < \epsilon$.

(c) For all ϵ, there exists some closed F contained by E such that $|E \setminus F|_e < \epsilon$.

(d) $E = H \setminus Z$ for some null Z and some $H \in G_\delta$.

(e) $E = H \cup Z$ for some null Z and some $H \in F_\sigma$.

Proof. Since $\mathbb{R}^d$ is σ-finite, to prove from (a) to (b), we can WOLG suppose $|E| < \infty$; the proof then follows from the definition of $|E|_e$, the trick of enlarging the cover to appropriate size and that open set is measurable. The proof for (b) $\implies$ (d) $\implies$ (a) is straightforward. The proof for (b) $\implies$ (c) is observation of $O \setminus E^c = E \setminus O^c$, and the proof for (c) $\implies$ (e) $\implies$ (a) is again straightforward. ■

It is worth pointing out that every open set can be expressed as a countable disjoint union of half-open dyadic cubes by an algorithmic construction. This implies that if we write

$$U_1 \triangleq \bigsqcup_n Q_{n,1} \text{ and } U_2 \triangleq \bigsqcup_n Q_{n,2}$$

we have

$$|U_1 \times U_2| = \left| \bigsqcup_{n,k} Q_{n,1} \times Q_{k,2} \right| = \sum_{n,k} |Q_{n,1}| \, |Q_{k,2}| = \sum_n |Q_{n,1}| \sum_k |Q_{k,2}| = |U_1| \, |U_2|$$

which in tern tell us that the product of measurable set is indeed measurable. (proved using an open set approximation from outside and the fact $\mathbb{R}^d$ is σ-finite)

Although most of the sets in Euclidean space we encounter in daily mathematics is measurable, It is worth pointing out there do exist a non Lebesgue measurable set, called the Vitali Set.

Theorem 5.3.2. (Lipschitz Continuity Preserve Lebesgue Measurability) If $E \in \mathcal{L}(\mathbb{R}^d)$ and $f : \mathbb{R}^d \to \mathbb{R}^m$ is K-Lipschitz, where $d \leq m$, then $f(E) \in \mathcal{L}(\mathbb{R}^m)$.

Proof. Write $E = H \cup Z$ where $H \in F_\sigma$ and Z is null. Because $f(E) = f(H) \cup f(Z)$, to show $f(E)$ is measurable, we can reduce the problem into proving f maps F_σ into F_σ and f maps null set to null set.

By splitting $\mathbb{R}^d$ into a countable union of compact set

$$\mathbb{R}^d = \bigcup_{n \in \mathbb{N}} \overline{B_n(\mathbf{0})}$$

We see that each closed set in $\mathbb{R}^d$ can be expressed as a countable union of compact sets. Then because continuous function preserve compactness, we see f must map closed set into class F_σ and thus map class F_σ into F_σ. (done)

Fix ϵ. Because $f:\mathbb{R}^d\to\mathbb{R}^d$ is K-Lipschitz, we know all its components $f_1,\dots,f_d:\mathbb{R}^d\to\mathbb{R}$ are also K-Lipschitz. Because Z is measurable and open set can be expressed as a disjoint countable union of dyadic half-open cubes, we can let $T_n\subseteq S$ be a countable cover of Z consisting of dyadic half-open cubes such that

$$\sum_n |T_n| < \frac{\epsilon}{K^m d^{\frac{m}{2}}}$$

Now, note that for each half-open cube T_n, if we write

$$T_n = \prod_{j=1}^d [a_j, a_j+h)$$

We clearly have

$$\operatorname{diam} T_n = h\sqrt{d} \text{ and } |T_n| = h^d$$

This give us the relationship between diameter and volume of a cube by

$$\operatorname{diam} T_n = |T_n|^{\frac{1}{d}} d^{\frac{1}{2}}$$

Because $f_k:\mathbb{R}^d\to\mathbb{R}$ are all K-Lipschitz, we know $f_k(T_n)$ can be contained by an interval of length $K(\operatorname{diam} T_n) = K\,|T_n|^{\frac{1}{d}} d^{\frac{1}{2}}$. This then tell us that $f(T_n)$ can be contained by closed cube of side length $K(\operatorname{diam} T_n)$, and give us the estimation

$$|f(T_n)|_e \le K^m (\operatorname{diam} T_n)^m = K^m d^{\frac{m}{2}} |T_n|^{\frac{m}{d}} \le K^m d^{\frac{m}{2}} |T_n|$$

where the last inequality hold true because $d\le m$ and WOLG we can suppose $|T_n|\le 1$. We now see that

$$|f(Z)|_e \le \sum_n |f(T_n)|_e \le K^m d^{\frac{m}{2}} \sum_n |T_n| < \epsilon \text{ (done)}$$

■

Although Lipschitz function from $\mathbb{R}^d$ to $\mathbb{R}^m$ preserve Lebesgue measurability when $d\le m$, the proposition does not hold true when $d>m$. Simple counter example can be constructed with projection $\pi:\mathbb{R}^2\to\mathbb{R}$ and the set $V\times\{0\}$ where V is the Vitali set. Also, it is worth pointing out that there do exists continuous function between $\mathbb{R}$, called the Cantor-Lebesgue function, that maps measurable set to non-measurable set.

Theorem 5.3.3. (Volume of Parallelepiped) Given $A \in M_d(\mathbb{R})$ and $E \triangleq [0,1]^d$, we have

$$|A(E)| = |\det(A)|$$

Proof. ∎

Ii is worth pointing out that our development of Theory of Lebesgue measure mostly depend on the abstract Carathéodory's Extension Theorem, which is indeed quite unorthodox. If one wish to instead develop the Theory of Lebesgue measurability starting at the second definition of Lebesgue measurability, one may follows the following steps

(i) Prove that the outer measure is countably subadditive using $\epsilon 2^{-k}$ trick.

(ii) Prove that null set is measurable.

(iii) Prove that compact set is measurable.

(iv) Prove that the class of measurable set is closed under countable union. (This step is independent of the second step)

(v) Prove that closed set is measurable by expressing the closed set is a countable union of compact sets.

(vi) Prove that complement E^c of measurable set E is measurable by expressing E^c as a countable union of closed set and a null set.

At this point, one can easily show that the collection of measurable sets form a sigma-algebra without usage of outer measure. It remains to show that the measure is countably additive.

(i) Prove that disjoint compact sets have positive distance.

(ii) Prove that $d(E_1, E_2) > 0 \implies |E_1 \cup E_2|_e = |E_1|_e + |E_2|_e$.

(iii) Prove the third equivalent definition of Lebesgue measurability.

(iv) Prove that if measurable E_n are bounded and disjoint then $|\bigcup E_n| = \sum |E_n|$ using the first three steps.

(v) Generalize the forth step to E_n that may not be bounded.

Chapter 6

Riemann Calculus

6.1 Riemann-Stieltjes Integral

‘

6.2 Riemann-Stieltjes on Computation

Theorem 6.2.1. (Change of Variable) Given two functions $g, \beta : [A, B] \to \mathbb{R}$, a function $\varphi : [A, B] \to [a, b]$ and two functions $f, \alpha : [a, b] \to \mathbb{R}$ such that

(a) $g = f \circ \varphi$ for all $x \in [a, b]$

(b) $\beta = \alpha \circ \varphi$ for all $x \in [a, b]$

(c) α, β increase respectively on $[a, b]$ and $[A, B]$

(d) $\varphi : [A, B] \to [a, b]$ is a homeomorphism

(e) $\int_a^b f d\alpha$ exist

Then

$$\int_A^B g d\beta = \int_a^b f d\alpha \text{ (This implies } \int_A^B g d\beta \text{ exists)}$$

Proof. Fix ϵ. We only wish

$$\text{to find a partition } Q \text{ of } [A, B] \text{ such that } U(Q, g, \beta) - L(Q, g, \beta) < \epsilon$$
$$\text{and such that } \int_a^b f d\alpha \in \Big[L(Q, g, \beta), U(Q, g, \beta)\Big]$$

Because $\int_a^b f d\alpha$ exists, we know

$$\text{there exists a partition } P \text{ of } [a, b] \text{ such that } U(P, f, \alpha) - L(P, f, \alpha) < \epsilon \tag{6.1}$$

where, of course, $\int_a^b f d\alpha \in \Big[L(P, f, \alpha), U(P, f, \alpha)\Big]$.

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$. Because φ is a homeomorphism, we can let φ be strictly increasing WOLG.

Define a partition Q on $[A, B]$ by

$$Q = \varphi^{-1}[P] = \{A = \varphi^{-1}(x_0), \varphi^{-1}(x_1), \ldots, \varphi^{-1}(x_n) = B\}$$

Now, because $\beta = \alpha \circ \varphi$ and $g = f \circ \varphi$ for all $x \in [a, b]$ by premise, and because φ is a homeomorphism, we have

$$\begin{aligned}
U(Q,g,\beta) &= \sum_{k=1}^{n} \Big[\sup_{t\in[\varphi^{-1}(x_{k-1}),\varphi^{-1}(x_k)]} g(t) \Big] \Big[\beta(\varphi^{-1}(x_k)) - \beta(\varphi^{-1}(x_{k-1})) \Big] \\
&= \sum_{k=1}^{n} \Big[\sup_{t\in[\varphi^{-1}(x_{k-1}),\varphi^{-1}(x_k)]} f\circ\varphi(t) \Big] \Big[\alpha\circ\varphi\big(\varphi^{-1}(x_k)\big) - \alpha\circ\varphi\big(\varphi^{-1}(x_{k-1})\big) \Big] \\
&= \sum_{k=1}^{n} \Big[\sup_{t\in[x_{k-1},x_k]} f(t) \Big] \big(\alpha(x_k) - \alpha(x_{k-1}) \big) = U(P,f,\alpha) \qquad (6.2)
\end{aligned}$$

Similarly, we can deduce $L(Q,g,\beta) = L(P,f,\alpha)$. Now, from Equation 6.2 and by definition of P (Equation 6.1), we see

$$U(Q,g,\beta) - L(Q,g,\beta) = U(P,f,\alpha) - L(P,f,\alpha) < \epsilon$$

$$\text{and } \int_a^b f d\alpha \in \Big[L(P,f,\alpha), U(P,f,\alpha) \Big] = \Big[L(Q,g,\beta), U(Q,g,\beta) \Big] \text{ (done)}$$

■

Theorem 6.2.2. (Reduction of Riemann-Stieltjes Integral: Part 1) Given two functions $f,\alpha : [a,b] \to \mathbb{R}$ such that

(a) α increase on $[a,b]$

(b) α is differentiable on (a,b)

(c) $\lim_{x\to b^-} \frac{\alpha(x)-\alpha(b)}{x-b}$ exists and $\lim_{x\to a^+} \frac{\alpha(x)-\alpha(a)}{x-a}$ exists

(d) α' is properly Riemann-Integrable on $[a,b]$

(e) f is bounded on $[a,b]$

Then

$$\int_a^b f d\alpha \text{ exists} \iff \int_a^b f(x)\alpha'(x)dx \text{ exists and they equal to each other if exists}$$

Proof. We wish to prove

$$\overline{\int_a^b} f d\alpha = \overline{\int_a^b} f(x)\alpha'(x)dx$$

Fix ϵ. We reduce the problem into proving

$$\left| \overline{\int_a^b} f d\alpha - \overline{\int_a^b} f(x)\alpha'(x)dx \right| < \epsilon$$

Then, because for all partition P of $[a,b]$, we have

$$\left|\overline{\int_a^b} f d\alpha - \overline{\int_a^b} f(x)\alpha'(x)dx\right|$$

$$\leq \left|\overline{\int_a^b} f d\alpha - U(P,f,\alpha)\right| - \left|U(P,f,\alpha) - U(P,f\alpha')\right| - \left|U(P,f\alpha') - \overline{\int_a^b} f(x)\alpha'(x)dx\right|$$

We only wish

$$\text{to find } P \text{ such that} \left|\overline{\int_a^b} f d\alpha - U(P,f,\alpha)\right| < \frac{\epsilon}{3}$$

$$\text{and } \left|U(P,f,\alpha) - U(P,f\alpha')\right| < \frac{\epsilon}{3} \text{ and } \left|\overline{\int_a^b} f(x)\alpha'(x)dx - U(P,f\alpha')\right| < \frac{\epsilon}{3}$$

Because f is bounded on $[a,b]$, we can let $M = \sup_{x\in[a,b]} |f(x)|$. Because $\int_a^b \alpha'(x)dx$ exists, we can let P satisfy

$$U(P,\alpha') - L(P,\alpha') < \frac{\epsilon}{4M} \tag{6.3}$$

By definition of Riemann Upper sum, we can further refine P to let P satisfy

$$\left|\overline{\int_a^b} f d\alpha - U(P,f,\alpha)\right| < \frac{\epsilon}{3} \text{ and } \left|\overline{\int_a^b} f(x)\alpha'(x)dx - U(P,f\alpha')\right| < \frac{\epsilon}{3}$$

It is clear that the statement concerning P (Equation 6.3) remain valid after refinement of P. Fix such P. We now have reduced the problem into proving

$$|U(P,f,\alpha) - U(P,f\alpha')| < \frac{\epsilon}{3}$$

Express P in the form $P = \{a = x_0, x_1, \dots, x_n = b\}$. By MVT (Theorem 4.3.3), we know for all $k \in \{1,\dots,n\}$ there exists $t_k \in [x_{k-1}, x_k]$ such that

$$\Delta\alpha_k = \alpha'(t_k)\Delta x_k \tag{6.4}$$

Then, because $U(P,\alpha') - L(P,\alpha)' < \frac{\epsilon}{3M}$ (Equation 6.3), we now see

$$\sum_{k=1}^n |\alpha'(s_k) - \alpha'(t_k)| \Delta x_k < \frac{\epsilon}{3M} \text{ if } s_k \in [x_{k-1}, x_k] \text{ for all } k \in \{1,\dots,n\} \tag{6.5}$$

Then from Equation 6.4, definition of M and Equation 6.5, we have

$$\begin{aligned}\left|\sum_{k=1}^{n} f(s_k)\Delta\alpha_k - \sum_{k=1}^{n} f(s_k)\alpha'(s_k)\Delta x_k\right| &= \left|\sum_{k=1}^{n} f(s_k)\big(\alpha'(s_k)-\alpha'(t_k)\big)\Delta x_k\right| \\ &\leq \sum_{k=1}^{n} |f(s_k)|\cdot|\alpha'(s_k)-\alpha'(t_k)|\,\Delta x_k \\ &\leq M\sum_{k=1}^{n} |\alpha'(s_k)-\alpha'(t_k)|\,\Delta x_k \\ &< \frac{\epsilon}{4}\end{aligned}$$

Then because $\sum_{k=1}^{m} f(s_k)\alpha'(s_k)\Delta x_k \leq U(P, f\alpha')$, we now have

$$\sum_{k=1}^{n} f(s_k)\Delta\alpha_k < U(P, f\alpha') + \frac{\epsilon}{4} \tag{6.6}$$

Because Equation 6.6 hold true for all choices of s_k, we have

$$U(P, f, \alpha) < U(P, f\alpha') + \frac{\epsilon}{3}$$

Similarly, we can deduce

$$U(P, f\alpha') < U(P, f, \alpha) + \frac{\epsilon}{3} \text{ (done)}$$

■

Theorem 6.2.3. (Substitution Law) Given a function $\varphi : [a, b] \to [A, B]$ and a function $f : [A, B] \to \mathbb{R}$ such that

(a) φ is a homoeomorphism.

(b) φ is differentiable on (a, b)

(c) $\int_a^b \varphi'(x)dx$ exists.

(d) f is integrable on $[A, B]$

We have

$$\int_a^b f\big(\varphi(x)\big)\varphi'(x)dx = \int_A^B f(u)du$$

Proof. Because $f \circ \varphi$ and φ' is integrable on $[a, b]$, by reduction of Riemann-Stieljes Integral (Tehroem 6.2.2), we know

$$\int_a^b \big(f \circ \varphi\big)(x)\varphi'(x)dx = \int_a^b \big(f \circ \varphi\big)(x)d\varphi$$

Let $\alpha(x) = x$. Let $\beta = \alpha \circ \varphi$. Define $g = f \circ \varphi$. By Change of Variable (Theorem 6.2.1), we now have

$$\int_a^b \big(f \circ \varphi\big)(x)d\varphi = \int_a^b g(x)d\beta = \int_A^B f(x)dx$$

$\blacksquare$

6.3 FTC

Theorem 6.3.1. (Fundamental Theorem of Calculus: Part 1) Suppose a function $f : [a, \infty) \to \mathbb{R}$ satisfy

$$f \text{ is } \underline{\text{proper-Riemann integrable}} \text{ on } [a, b] \text{ for all } b > a$$

If we set $F : [a, \infty) \to \mathbb{R}$

$$F(x) = \int_a^x f(t)dt$$

Then

(a) F is continuous on $[a, \infty)$

(b) F is differentiable at $x_0 \in [a, \infty)$ where $F'(x_0) = f(x_0)$ if f is continuous at x_0

Proof. Fix ϵ and $[a, b]$. We only wish

$$\text{to prove } F \text{ is continuous on } [a, b]$$

To prove F is continuous on $[a, b]$, we only wish

$$\text{to find } \delta \text{ such that } \forall [x, y] \subseteq [a, b], |x - y| < \delta \implies |F(x) - F(y)| < \epsilon$$

Because f is proper-Riemann-Integrable on $[a, b]$, we know f is bounded on $[a, b]$. Let M be an upper bound of $|f|$ on $[a, b]$. We claim

$$\delta = \frac{\epsilon}{M} \text{ works}$$

Because $y - x < \delta = \frac{\epsilon}{M}$, we have

$$\begin{aligned} |F(x) - F(y)| &= \left| \int_x^y f(t)dt \right| \\ &\leq \int_x^y |f(t)| \, dt \\ &\leq (y - x) < \epsilon \text{ (done)} \end{aligned}$$

Now, to prove $F'(x_0) = f(x_0)$, we wish

$$\text{to prove } \lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0)$$

Fix ϵ. We wish

$$\text{to find } \delta \text{ such that } |x - x_0| < \delta \implies \left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| < \epsilon$$

Because f is continuous at x_0, we know

$$\exists \delta, |x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon \tag{6.7}$$

We claim

such δ in Equation 6.7 works

WOLG, let $x > x_0$. Deduce

$$\begin{aligned}
\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| &= \left|\frac{\int_{x_0}^x f(t)dt}{x - x_0} - f(x_0)\right| \\
&= \left|\frac{\int_{x_0}^x [f(t) - f(x_0)]\,dt}{x - x_0}\right| \\
&\leq \frac{\int_{x_0}^x |f(t) - f(x_0)|\,dt}{|x - x_0|} \\
&\leq \frac{\int_{x_0}^x \epsilon dt}{|x - x_0|} = \epsilon \text{ (done)}
\end{aligned}$$

■

Theorem 6.3.2. (Fundamental Theorem of Calculus: Part 2, Leibniz Rule) Suppose two functions $f, F : [a, \infty) \to \mathbb{R}$ satisfy

(a) f is proper Riemann-Integrable on $[a, b]$ for all $b > a$

(b) $F'(x) = f(x)$ for all $x \in (a, \infty)$

(c) F is continuous on $[a, \infty)$

Then for all $b > a$,

$$\int_a^b f(x)dx = F(b) - F(a)$$

Proof. Fix ϵ. We wish

$$\text{to show that } \left|\Big(F(b) - F(a)\Big) - \int_a^b f(x)dx\right| < \epsilon$$

Because f is proper Riemann-Integrable on $[a,b]$, we know there exists a partition $P = \{a = x_0, x_1, \dots, x_n = b\}$ of $[a,b]$ such that

$$U(P,f) - L(P,f) < \epsilon \tag{6.8}$$

Because $f = F'$ on (a,b), for each $k \in \{1,\dots,n\}$, by MVT , we know

$$\exists t_k \in (x_{k-1}, x_k), \frac{F(x_k) - F(x_{k-1})}{x_k - x_{k-1}} = f(t_k)$$

This let us deduce

$$F(b) - F(a) = \sum_{k=1}^{n} F(x_k) - F(x_{k-1}) = \sum_{k=1}^{n} f(t_k)\Delta x_k$$

Now, we have

$$\int_a^b f(x)dx \text{ and } F(b) - F(a) \text{ are both in } \big[L(P,f), U(P,f)\big]$$

Then by Equation 6.8, we can deduce

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| < \epsilon \text{ (done)}$$

■

Theorem 6.3.3. (Integral By Part) Given four function $f, g, F, G : [a,b] \to \mathbb{R}$ such that

(a) $F'(x) = f(x)$ and $G'(x) = g(x)$ for all $x \in (a,b)$

(b) f, g are properly Riemann-Integrable on $[a,b]$

(c) F, G are continuous on $[a,b]$

We have

$$\int_a^b F(x)g(x)dx = FG\Big|_a^b - \int_a^b f(x)G(x)dx \tag{6.9}$$

Proof. To prove Equation 6.9, we only with

$$\text{to prove } \int_a^b F(x)g(x)dx + \int_a^b f(x)G(x)dx = FG\Big|_a^b$$

We can reduce the problem

$$\text{into proving } \int_a^b (Fg + fG)dx = FG\Big|_a^b$$

Notice that by Chain Rule,

$$\big(FG\big)'(x) = F(x)g(x) + f(x)G(x) \text{ for all } x \in (a, b)$$

Then the result follows from Part 2 of Fundamental Theorem of Calculus (Theorem 6.3.2).
(done) ■

Example 24 (Discontinuous Derivative)

$$f(x) = \begin{cases} \frac{x^2}{\sin x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

6.4 Bounded Variation

This section prove some key properties of functions of bounded variations. These properties are worthy of discuss, as they make the set $BV\big([a,b]\big)$ of function of bounded variation on $[a,b]$ a natural candidate for the class of Riemann-Stieltjes integrator. The key properties include

(a) Functions of bounded variation must be continuous almost everywhere. (Corollary 6.4.9)

(b) Functions of bounded variation can be expressed as difference of two increasing functions. (Theorem 6.4.5)

(c) Functions of bounded variation can only have jump discontinuity. (Corollary 6.4.6)

Definition 6.4.1. (Definition of variation and function of bounded variation) Given a compact interval $[a,b]$, by a **partition** P of $[a,b]$, we mean a finite set $\{a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\}$ contain by $[a,b]$ and containing a,b. If f is a complex-valued function defined on $[a,b]$, we say the **total variation** $V_f(a,b)$ of f on $[a,b]$ is

$$\sup_P \sum_{k=1}^n |f(x_k) - f(x_{k-1})|$$

We say f is of **bounded variation** on $[a,b]$ if $V_f(a,b) < \infty$, and we denote the set of real-valued function from $[a,b]$ of bounded variation on $[a,b]$ by $BV\big([a,b]\big)$.

It is straightforward to check

(a) $f \in BV\big([a,b]\big)$ must be bounded on $[a,b]$

(b) real-valued f monotone on $[a,b]$ is of bounded variation on $[a,b]$

(c) $BV\big([a,b]\big)$ form a vector space over $\mathbb{R}$

In fact, $BV\big([a,b]\big)$ also form a commutative algebra, as below proved.

Theorem 6.4.2. (Bounded variation is closed under multiplication) Given two real-valued (or more generally complex-valued) f,g defined on $[a,b]$

$$V_{fg}(a,b) \le AV_f(a,b) + BV_f(a,b)$$

where

$$A = \sup_{[a,b]} |g| \text{ and } B = \sup_{[a,b]} |f|$$

Proof. For every partition P, we have

$$\begin{aligned}
\sum_{k=1}^{n} |fg(x_k) - fg(x_{k-1})| &= \sum_{k=1}^{n} |(f(x_k) - f(x_{k-1}))g(x_k) + f(x_{k-1})g(x_k) - fg(x_{k-1})| \\
&\leq \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \, |g(x_k)| + \sum_{k=1}^{n} |f(x_{k-1})| \, |g(x_k) - g(x_{k-1})| \\
&\leq AV_f(a,b) + BV_f(a,b)
\end{aligned}$$

■

Note that the proof above only consider when g, f are both bounded on $[a,b]$. If not, the statement hold trivially. For the brevity of the proof of the next Theorem, if we are given a partition $P = \{a = x_0 < \cdots < x_n = b\}$ of $[a,b]$, we denote $\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$ by $\sum(P)$.

Theorem 6.4.3. (Additive property of total variation) Given a real-valued function f defined $[a,b]$, and $c \in (a,b)$

$$V_f(a,b) = V_f(a,c) + V_f(c,b)$$

Proof. We first prove

$$V_f(a,b) \geq V_f(a,c) + V_f(c,b)$$

Note that it is possible $V_f(a,c) = \infty$. We only prove when $V_f(a,c) < \infty$, since the proof for the statement when $V_f(a,c) = \infty$ is similar. Fix ϵ. We reduce the problem into proving

$$V_f(a,b) \geq V_f(a,c) + V_f(c,b) - \epsilon$$

Let Y, Z respectively be the set of all partitions of $[a,c]$ and $[a,b]$. Fix $P_y \in Y$ and $P_z \in Z$ such that

$$\sum(P_y) > V_f(a,c) - \frac{\epsilon}{2} \text{ and } \sum(P_z) > V_f(c,b) - \frac{\epsilon}{2}$$

It is clear that $P_y \cup P_z$ is a partition of $[a,b]$. Observe

$$V_f(a,b) \geq \sum(P_y \cup P_z) = \sum(P_y) + \sum(P_z) > V_f(a,c) + V_f(c,b) - \epsilon \text{ (done)}$$

It remains to prove

$$V_f(a,b) \leq V_f(a,c) + V_f(c,b)$$

Fix a partition P of $[a,b]$. We are required to prove

$$\sum(P) < V_f(a,c) + V_f(c,b)$$

Let $P_c = P \cup \{c\}$. Observe

$$\sum(P) \le \sum(P_c) \le V_f(a,c) + V_f(c,b) \text{ (done)}$$

■

Corollary 6.4.4. (Additive property of total variation) Given a real-valued function f defined $[a,b]$, and $c \in (a,b)$

$$f \in BV\big([a,b]\big) \iff f \in BV\big([a,c]\big) \text{ and } f \in BV\big([c,b]\big)$$

Perhaps, the best property of bounded variation are the following.

Theorem 6.4.5. (Function of bounded variation can be expressed as difference of two increasing functions) Given real-valued f defined on $[a,b]$

$$f \in BV\big([a,b]\big) \iff \exists \text{ increasing } g,h : [a,b] \to \mathbb{R}, f = g - h$$

Proof. From right to left is trivial. From left to right, we claim

$$g(x) \triangleq V_f(a,x) \text{ and } h(x) \triangleq V_f(a,x) - f(x) \text{ suffices}$$

It is clear that $V_f(a,x)$ is increasing. Fix $a \le x \le y \le b$. We prove

$$h(x) \le h(y)$$

Use Theorem 6.4.3

$$\begin{aligned} h(y) - h(x) &\ge V_f(a,y) - f(y) - V_f(a,x) + f(x) \\ &= \big(V_f(a,y) - V_f(a,x)\big) - \big(f(y) - f(x)\big) \\ &= V_f(x,y) - \big(f(y) - f(x)\big) \ge 0 \text{ (done)} \end{aligned}$$

■

Corollary 6.4.6. (Function of bounded variation can only have jump discontinuity) Given $f : [a,b] \to \mathbb{R}$, if $f \in BV\big([a,b]\big)$, then f can only have jump discontinuity.

Corollary 6.4.7. (Function of bounded variation can be expressed as difference of two strictly increasing functions) Given real-valued f defined on $[a,b]$

$$f \in BV\big([a,b]\big) \iff \exists \text{ strictly increasing } g,h : [a,b] \to \mathbb{R}, f = g - h$$

Proof. From right to left is again trivial. If $g(x) \triangleq V_f(a,x)$ and $h(x) \triangleq V_f(a,x) - f(x)$ is not strictly increasing, define $g' \triangleq g + (x-a)$ and $h' \triangleq h + (x-a)$. Such g', h' suffice. ■

The reason we showed functions of bounded variation can be expressed as difference of two increasing functions is the following (Theorem 6.4.9). Theorem 6.4.9 make function of bounded variation continuous almost everywhere. This make functions of bounded variation a natural candidate of the class of Riemann-Stieltjes integrators, in the perspective of Lebesgue-Riemann Criterion. (Theorem **??**)

Theorem 6.4.8. (Function monotone on $[a, b]$ must be continuous almost everywhere on $[a, b]$) If f is monotone on $[a, b]$, then the set of discontinuities of f is countable.

Proof. WOLG, suppose f is increasing on $[a, b]$. Because f is increasing on $[a, b]$, we know that every discontinuities of f on $[a, b]$ is a jump discontinuity. Define

$$S_n \triangleq \{x \in [a, b] : f(x+) - f(x-) > \frac{1}{n}\}$$

Then the set of discontinuity of f is exactly $\bigcup_n S_n$. Note that each S_n must be finite, otherwise $f(b) = \infty$. This conclude the proof. ■

Corollary 6.4.9. (Function of bounded variation on $[a, b]$ must be continuous almost everywhere on $[a, b]$) If f is of bounded variation on $[a, b]$, then the set of discontinuities of f on $[a, b]$ is countable.

If we only consider real-valued continuous function of bounded variation on $[a, b]$, the structure is even richer.

Theorem 6.4.10. (Continuous function of bounded variation) Given $f : [a, b] \to \mathbb{R}$ such that $f \in BV\big([a, b]\big)$. If we define $V : [a, b] \to \mathbb{R}$ by

$$V(x) \triangleq V_f(a, x)$$

Then for each $x \in [a, b]$

$$V \text{ is continuous at } x \iff f \text{ is continuous at } x$$

Proof. $(\longrightarrow)$

Assume f is discontinuous at x. WOLG, suppose that there exists $x_n \searrow x$ such that

$$\lim_{n\to\infty} f(x_n) \in \mathbb{R} \text{ and } c \triangleq \left|\lim_{n\to\infty} f(x_n) - f(x)\right| > 0$$

Now, observe

$$|V(x_n) - V(x)| = |V_f(x, x_n)| \geq |f(x_n) - f(x)|$$

This give us

$$\begin{aligned}\liminf_{n\to\infty} |V(x_n) - V(x)| &\geq \liminf_{n\to\infty} |f(x_n) - f(x)| \\ &= \lim_{n\to\infty} |f(x_n) - f(x)| = \left|\lim_{n\to\infty} f(x_n) - f(x)\right| = c > 0 \text{ CaC}\end{aligned}$$

$(\longleftarrow)$

Fix ϵ. Let $P = \{x = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[x, b]$ such that

$$V_f(x, b) - \frac{\epsilon}{2} < \sum(P)$$

Let δ satisfy

$$|f(y) - f(x)| < \frac{\epsilon}{4n} \text{ for all } y \in [x, x + \delta]$$

We claim

$$\text{such } \delta \text{ satisfy } |V(y) - V(x)| \leq \epsilon \text{ for all } y \in [x, x + \delta]$$

Let $y \triangleq x + \delta$. Because V is increasing on $[a, b]$, we only have to prove

$$|V(y) - V(x)| \leq \epsilon$$

Theorem 6.4.3 allow us to reduce the problem into proving

$$V_f(x, y) \leq \epsilon$$

Denote $P \cup \{y\}$ by P' and express $P' = \{x = x'_0 < \cdots < x'_r\}$. Express $y = x'_m$. Note that $m < r \leq n + 1$ and observe

$$\begin{aligned}V_f(x, b) - \frac{\epsilon}{2} < \sum(P) \leq \sum(P') \\ &= \sum_{k=1}^{m} |f(x'_k) - f(x'_{k-1})| + \sum_{k=m+1}^{n} |f(x'_k) - f(x'_{k-1})| \\ &\leq \frac{m\epsilon}{2n} + V_f(y, b) \leq \frac{\epsilon}{2} + V_f(y, b)\end{aligned}$$

Theorem 6.4.3 now give us

$$V_f(x, y) = V_f(x, b) - V_f(y, b) \leq \epsilon \text{ (done)}$$

Proof for $V(x-) = V(x)$ is similar, and when $x = a$ or b, some trivial modifications are needed. ■

Give very careful attention to the statement of Theorem 6.4.10. Note that we require to the domain of f to be $[a,b]$. If the domain of f contain a or b as interior point, the statement isn't always true.

Corollary 6.4.11. (Continuous function of bounded variation can be expressed as difference of two continuous strictly increasing functions) Given continuous real-valued f defined on $[a,b]$

$$f \in BV\big([a,b]\big) \iff \exists \text{ continuous strictly increasing } g,h:[a,b]\to\mathbb{R}, f=g-h$$

Proof. From right to left is again trivial. If $g(x) \triangleq V_f(a,x)$ and $h(x) \triangleq V_f(a,x)-f(x)$ is not strictly increasing, define $g' \triangleq g+(x-a)$ and $h' \triangleq h+(x-a)$. Such g',h' suffice. ■

6.5 Uniform Convergence and Riemann Integration

Theorem 6.5.1. (Riemann-Integration and Uniform Convergence) Given a function $\alpha : [a,b] \to \mathbb{R}$ and a sequence of functions $f_n : [a,b] \to \mathbb{R}$ such that

(a) α increase on $[a,b]$

(b) $\int_a^b f_n d\alpha$ exists for all $n \in \mathbb{N}$

(c) $f_n \to f$ uniformly on $[a,b]$

Then

$$\lim_{n\to\infty} \int_a^b f_n d\alpha \text{ exists and } \int_a^b f d\alpha = \lim_{n\to\infty} \int_a^b f_n d\alpha$$

Proof. We first prove

$$\int_a^b f d\alpha \text{ exists}$$

Fix ϵ. We wish to prove

$$\overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha < \epsilon$$

Let $\epsilon_n = \|f_n - f\|_\infty$. Because $f_n \to f$ uniformly, we know

$$\text{there exists } n \in \mathbb{N} \text{ such that } \epsilon_n = \|f_n - f\|_\infty < \frac{\epsilon}{2\big[\alpha(b) - \alpha(a)\big]}$$

Because α increase, by definition of ϵ_n, we see

$$\int_a^b (f_n - \epsilon_n) d\alpha \leq \underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha \leq \int_a^b (f_n + \epsilon_n) d\alpha$$

Because $\epsilon_n < \frac{\epsilon}{2\big[\alpha(b)-\alpha(a)\big]}$, we now see

$$\begin{aligned} \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha &\leq \int_a^b (f_n + \epsilon_n) d\alpha - \int_a^b (f_n - \epsilon_n) d\alpha \\ &= \int_a^b (2\epsilon_n) d\alpha < 2\epsilon_n \cdot \big[\alpha(b) - \alpha(a)\big] = \epsilon \text{ (done)} \end{aligned}$$

We now prove

$$\int_a^b f_n d\alpha \to \int_a^b f d\alpha \text{ as } n \to \infty$$

Fix ϵ. We wish

$$\text{to find } N \text{ such that } \forall n > N, \left| \int_a^b f_n d\alpha - \int_a^b f d\alpha \right| < \epsilon$$

Recall the definition $\epsilon_n = \|f_n - f\|_\infty$. Because $\epsilon_n \to 0$, we know

$$\text{there exists } N \text{ such that } \forall n > N, \epsilon_n < \frac{\epsilon}{\alpha(b) - \alpha(a)} \tag{6.10}$$

We claim

$$\text{such } N \text{ works}$$

Fix $n > N$. From Equation 6.10, we see

$$\begin{aligned}
\left| \int_a^b f_n d\alpha - \int_a^b f d\alpha \right| &= \left| \int_a^b (f_n - f) d\alpha \right| \\
&\leq \int_a^b |f_n - f| \, d\alpha \\
&\leq \int_a^b \epsilon_n d\alpha = \epsilon_n \big[\alpha(b) - \alpha(a)\big] < \epsilon \text{ (done)}
\end{aligned}$$

■

As Rudin remarked, a much shorter (and much more intuitive) proof can be given, if we require f' to be continuous on $[a, b]$.

Theorem 6.5.2. (Uniform Convergence and Differentiation: Weaker Version) Given a sequence of function $f_n : [a, b] \to \mathbb{R}$ such that

(a) f_n' uniformly converge on $[a, b]$

(b) $f_n(x_0) \to L$ for some $x_0 \in [a, b]$

(c) f_n are of class C^1 on $[a, b]$

Then

(a) f_n uniformly converge on $[a, b]$

(b) and

$$\frac{d}{dx}\Big(\lim_{n\to\infty} f_n(x)\Big)\Big|_{x=x_0} = \lim_{n\to\infty} f_n'(x_0) \text{ on } (a,b)$$

Proof. We claim

$$f(x) = \lim_{n\to\infty}\int_{x_0}^{x} f_n'(t)dt + L \text{ works}$$

Note that $\lim_{n\to\infty}\int_{x_0}^{x} f_n'(t)dt$ exists because f_n' uniformly converge (Theorem 6.5.1).

Because f_n' uniformly converge and are continuous on $[a,b]$, by ULT, we know

$$\int_{x_0}^{x} \lim_{n\to\infty} f_n'(t)dt + L \text{ exists}$$

and know

$$f(x) = \int_{x_0}^{x} \lim_{n\to\infty} f_n'(t)dt + L$$

By FTC, we see

$$f'(x) = \lim_{n\to\infty} f_n'(x) \text{ on } (a,b)$$

Such convergence is uniform by premise. To finish the proof, we now only have to prove

$$f_n \to f \text{ uniformly on } [a,b]$$

Fix ϵ. We wish

$$\text{to find } N \text{ such that } |f_n(x) - f(x)| \leq \epsilon \text{ for all } n > N \text{ and } x \in [a,b]$$

Because $f_n' \to f'$ uniformly, and $f_n(x_0) \to L = f(x_0)$ (Check $L = f(x_0)$), we know there exists N such that

$$\begin{cases} \|f_n' - f'\|_\infty < \frac{\epsilon}{2(b-a)} \\ |f_n(x_0) - f(x_0)| < \frac{\epsilon}{2} \end{cases} \quad \text{for all } n > N$$

We claim

$$\text{such } N \text{ works}$$

Fix $n > N$ and $x \in [a,b]$. Observe

$$\begin{aligned}
|f(x) - f_n(x)| &= \left| \int_{x_0}^{x} \big(f'(t) - f_n'(t)\big) dt + f(x_0) - f_n(x_0) \right| \\
&\leq \int_{x_0}^{x} |f'(t) - f_n'(t)|\, dt + |f(x_0) - f_n(x_0)| \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ (done)}
\end{aligned}$$

■

6.6 MVT for Definite Integral

Theorem 6.6.1. (First Mean Value Theorem for Definite Integral) Given a function $f : [a, b] \to \mathbb{R}$ such that

(a) f is continuous on (a, b)

There exists $\xi \in (a, b)$ such that

$$\int_a^b f(x)dx = f(\xi) \cdot (b - a)$$

Proof. We wish

$$\text{to find } \xi \in (a, b) \text{ such that } f(\xi) = \frac{\int_a^b f(x)dx}{b - a}$$

Define $\tilde{f} : [a, b] \to \mathbb{R}$ on $[a, b]$ by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in (a, b) \\ \lim_{t \to a} f(t) & \text{if } x = a \\ \lim_{t \to b} f(t) & \text{if } x = b \end{cases} \tag{6.11}$$

Then, because $\int_a^b f(x)dx = \int_a^b \tilde{f}(x)dx$, we reduce our problem into

$$\text{finding } \xi \in (a, b) \text{ such that } \tilde{f}(\xi) = \frac{\int_a^b \tilde{f}(x)dx}{b - a}$$

Because $\tilde{f}$ is continuous on $[a, b]$ by definition Equation 6.11, by EVT, we know we there exists $\alpha, \beta \in [a, b]$ such that

$$\tilde{f}(\alpha) = \inf_{x \in [a,b]} \tilde{f}(x) \text{ and } \tilde{f}(\beta) = \sup_{x \in [a,b]} \tilde{f}(x) \tag{6.12}$$

WOLG, suppose $\alpha \leq \beta$. Deduce

$$\tilde{f}(\alpha) = \inf_{x \in [a,b]} \tilde{f}(x) \leq \frac{\int_a^b \tilde{f}(x)dx}{b - a} \leq \sup_{x \in [a,b]} \tilde{f}(x) = \tilde{f}(\beta)$$

by IVT, we then know there exists $\xi \in [\alpha, \beta]$ such that

$$\exists \xi \in [\alpha, \beta], \tilde{f}(\xi) = \frac{\int_a^b \tilde{f}(x)dx}{b - a} \tag{6.13}$$

If $a < \alpha$ and $\beta < b$, our proof is done.

If not, notice that if $\tilde{f}(\alpha) = \tilde{f}(\beta)$, then by definition of α, β (Equation 6.12), the proof is trivial since $\tilde{f}$ is a constant, so we only have to consider when $\tilde{f}(\alpha) < \tilde{f}(\beta)$, and we wish to show

$$\xi \text{ can not happen at } a \text{ nor } b$$

Assume $\xi = a$, WOLG. Because $\xi \in [\alpha, \beta]$, we know $\alpha = a$. Because $\tilde{f}(\beta) > \tilde{f}(\alpha)$, we can find δ such that

$$\inf_{x \in [\beta - \delta, \beta]} \tilde{f}(x) \geq \frac{\tilde{f}(\alpha) + 2\tilde{f}(\beta)}{3} \tag{6.14}$$

We then from Equation 6.13 see that

$$\int_a^b \tilde{f}(x)dx = \tilde{f}(\xi)(b-a) = \tilde{f}(\alpha)(b-a) \tag{6.15}$$

Also, we see from definition of α (Equation 6.12) and Equation 6.14 that

$$\int_a^b \tilde{f}(x)dx = \int_a^{\beta-\delta} \tilde{f}(x)dx + \int_{\beta-\delta}^{\beta} \tilde{f}(x)dx + \int_\beta^b \tilde{f}(x)dx \tag{6.16}$$

$$\geq (b-\delta-a)\tilde{f}(\alpha) + \delta \cdot \left(\frac{\tilde{f}(\alpha) + 2\tilde{f}(\beta)}{3}\right) \tag{6.17}$$

$$> (b-\delta-a)\tilde{f}(\alpha) + \delta \cdot \left(\frac{\tilde{f}(\alpha) + \tilde{f}(\beta)}{2}\right) \tag{6.18}$$

$$= \tilde{f}(\alpha)\left(b-a-\frac{\delta}{2}\right) + \tilde{f}(\beta) \cdot \left(\frac{\delta}{2}\right) \tag{6.19}$$

Now, from Equation 6.15 and Equation 6.19, we can deduce

$$\tilde{f}(\alpha)(b-a) > \tilde{f}(\alpha)(b-a-\frac{\delta}{2}) + \tilde{f}(\beta) \cdot (\frac{\delta}{2})$$

Then we can deduce

$$\tilde{f}(\alpha) \cdot (\frac{\delta}{2}) > \tilde{f}(\beta) \cdot (\frac{\delta}{2}) \text{ CaC (done)}$$

■

Theorem 6.6.2. (Second Mean Value Theorem for Definite Integral) Given functions $G, \varphi : [a,b] \to \mathbb{R}$ such that

(a) G is monotonic

(b) φ is Riemann-Integrable

Let $G(a^+) = \lim_{t\to a^+} G(t)$ and $G(b^-) = \lim_{t\to b^-} G(t)$. Then there exists $\xi \in (a,b)$ such that

$$\int_a^b G(t)\varphi(t)dt = G(a^+)\int_a^\xi \varphi(t)dt + G(b^-)\int_\xi^b \varphi(t)dt$$

Proof. Define f on $[a,b]$ by

$$f(x) = G(a^+)\int_a^x \varphi(t)dt + G(b^-)\int_x^b \varphi(t)dt$$

We then reduce the problem into

$$\text{finding } \xi \in (a,b) \text{ such that } \int_a^b G(t)\varphi(t)dt = f(\xi)$$

By Theorem 6.3.1, we know f is continuous on $[a,b]$. Then by IVT, we can reduce the problem into

$$\text{finding an interval } [c,d] \subseteq (a,b) \text{ such that } \int_a^b G(t)\varphi(t) \text{ is between } f(c) \text{ and } f(d)$$

Observe that

$$f(a) = G(b^-)\int_a^b \varphi(t)dt \text{ and } f(b) = G(a^+)\int_a^b \varphi(t)dt$$

■

6.7 Taylor's Theorem

Abstract

This section prove Taylor's Theorem in both single and multi variables, and give some explicit formulas for the remainder terms. In this section, k is a fixed integer greater than 1, $I \subseteq \mathbb{R}$ is an open interval containing a and $U \subseteq \mathbb{R}^n$ is an open set containing $\mathbf{a}$

Theorem 6.7.1. (Single Variable Taylor's Theorem) If real function $f : I \to \mathbb{R}$ is k time differentiable at a and if we write

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + R_k(x)$$

then the approximation error $R_k(x)$ satisfy

$$R_k(x) = o(|x-a|^k)$$

Proof. Denote $P_k : I \to \mathbb{R}$

$$P_k(x) \triangleq f(a) + f'(a)(x-a) + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k$$

The proof then follows from applying L'Hospital Rule to have

$$\lim_{x \to a} \frac{f(x) - P_k(x)}{(x-a)^k} = 0$$

∎

Theorem 6.7.2. (Schlömilch form of the Remainders) If $f : I \to \mathbb{R}$ is $k+1$ times differentiable on I, then for all positive integer p, the approximation error $R_k : I \to \mathbb{R}$ can be written in the Schlömilch form

$$R_k(x) = \frac{f^{(k+1)}(\xi)}{k!}(x-\xi)^{k+1-p}\frac{(x-a)^p}{p} \text{ for some } \xi \text{ between } x, a$$

Corollary 6.7.3. (Lagrange and Cauchy form of the Remainders) If $f : I \to \mathbb{R}$ is $k+1$ times differentiable on I, then for all positive integer p, the approximation error $R_k : I \to \mathbb{R}$ can be written in the Lagrange form

$$R_k(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-a)^{k+1} \text{ for some } \xi \text{ between } x, a$$

or the Cauchy form

$$R_k(x) = \frac{f^{(k+1)}(\xi)}{k!}(x-\xi)^k(x-a) \text{ for some } \xi \text{ between } x, a$$

Notably, the remainder can also be written in form of integral.

Theorem 6.7.4. (Integral form of the Remainder) If $f : I \to \mathbb{R}$ is C^{k+1} on I, then the approximation error $R_k : I \to \mathbb{R}$ can be written in the integral form

$$R_k(x) = \int_a^x \frac{f^{(k+1)}(t)}{k!}(x-t)^k dt$$

We now consider when f is multi-variables. We first introduce the multi-index notation. Given $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ and $\mathbf{x} \in \mathbb{R}^n$, we write

$$\begin{aligned} |\alpha| &\triangleq \alpha_1 + \cdots + \alpha_n \\ \alpha! &\triangleq \alpha_1! \cdots \alpha_n! \\ \mathbf{x}^\alpha &\triangleq x_1^{\alpha_1} \cdots x_n^{\alpha_n} \end{aligned}$$

Suppose $f : U \to \mathbb{R}$ is C^{k+1} on U. One shall note that the notation

$$D^\alpha f \triangleq \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} \cdots (\partial x_n)^{\alpha_n}} \text{ for all } |\alpha| \leq k+1$$

is well-defined.

Theorem 6.7.5. (Multi Variables Taylor's Theorem) If real-valued function $f : U \to \mathbb{R}$ is C^{k+1} on U, then we can write

$$f(\mathbf{x}) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(\mathbf{a})}{\alpha!}(\mathbf{x}-\mathbf{a})^\alpha + \sum_{|\beta|=k+1} R_\beta(\mathbf{x})(\mathbf{x}-\mathbf{a})^\beta$$

where

$$R_\beta(\mathbf{x}) \triangleq \frac{|\beta|}{\beta!} \int_0^1 (1-t)^{|\beta|-1} D^\beta f(\mathbf{a} + t(\mathbf{x}-\mathbf{a}))dt$$

Proof. ■

Chapter 7

Complex Analysis

Chapter 8

Lebesgue Calculus

8.1 Basic Property of Measurable Functions

In this section, we give two equivalent general definition (Theorem 8.1.1) of measurable function and show

(a) Measurable $f : X \to \mathbb{R}$ or $\mathbb{C}$ is closed under addition and multiplication. (Theorem 8.1.3)

(b) The (superior) limit of a sequence of measurable ($\mathbb{R} \cup \{\pm\infty\}$ or $\mathbb{C} \cup \{\infty\}$)-valued functions on general (X, Σ_X) is measurable. (Theorem 8.1.4)

(c) Positive and negative part of a measurable $f : X \to [-\infty, \infty]$ are measurable. (Corollary 8.1.6)

If we are given a function $f : (X, \Sigma_X) \to (Y, \Sigma_Y)$ such that for all $E \in \Sigma_Y$, we have $f^{-1}(E) \in \Sigma_X$, then we say f is a **measurable function**. Immediately, one can check that the composition of two measurable functions must be measurable. We now introduce an equivalent definition of measurable function, which makes checking if a function is measurable much easier.

Theorem 8.1.1. (Equivalent Definition of measurable function) If a function $f : (X, \Sigma_X) \to (Y, \sigma(T))$ satisfy

$$f^{-1}(E) \in \Sigma_X \text{ for all } E \in T$$

then

$$f \text{ is measurable}$$

Proof. Define

$$\mathcal{A} \triangleq \{E \subseteq Y : f^{-1}(E) \in \Sigma_X\}$$

Check that $\mathcal{A}$ is a σ-algebra on Y. By premise, $T \subseteq \mathcal{A}$. It then follows from definition that $\sigma(T) \subseteq \mathcal{A}$. This conclude that f is $(\Sigma_X, \sigma(T))$-measurable. ■

There are two important consequences of Theorem 8.1.1

(a) If X, Y are both Borel, then a continuous function $f : X \to Y$ must also be a measurable function.

(b) If there exists T such that $\Sigma_Y = \sigma(T)$, then we only have to check that $f^{-1}(E) \in \Sigma_X$ for all $E \in T$ to show f is measurable. In particular, if $Y = \mathbb{R}^n$, then T can be just the set of all open boxes.

These two consequences give the following Lemma, which later prove Theorem 8.1.3, Theorem that make checking if $\mathbb{C}$ or $\mathbb{R}$-valued function is measurable easier.

Lemma 8.1.2. (Computational Lemma) Given two measurable function $u, v : (X, \Sigma_X) \to \mathbb{R}$ and a continuous function $\Phi : \mathbb{R}^2 \to (Y, \tau_Y)$, define $h : (X, \Sigma_X) \to (Y, \mathcal{B}_Y)$ by

$$h(x) \triangleq \Phi(u(x), v(x))$$

We can deduce

$$h : (X, \Sigma_X) \to (Y, \mathcal{B}_Y) \text{ is measurable}$$

Proof. Define $f : X \to \mathbb{R}^2$ by $f(x) \triangleq (u, v)(x)$. Because $h = \Phi \circ f$, we can reduce the problem into proving

$$f \text{ is measurable}$$

Fix an open-rectangle $I_1 \times I_2 \subseteq \mathbb{R}^2$. Because $\mathcal{B}_{\mathbb{R}^2}$ can be generated by the set of all open rectangles, we can reduce the problem into proving

$$f^{-1}(I_1 \times I_2) \in \Sigma_X$$

Because u, v are measurable, we know

$$f^{-1}(I_1 \times I_2) = u^{-1}(I_1) \cap v^{-1}(I_2) \in \Sigma_X \text{ (done)}$$

■

Theorem 8.1.3. (Basic tools to show a real/complex-valued function is measurable) Given a measurable set X, and two real-valued function $u, v : X \to \mathbb{R}$

(a) $u + iv : X \to \mathbb{C}$ is measurable if and only if u, v are measurable.

(b) If $f, g : X \to \mathbb{R}$ or $\mathbb{C}$ are measurable, so are $f + g$, fg and $|f|$.

(c) If $f : X \to \mathbb{R}$ or $\mathbb{C}$ is measurable and $g : X \to \mathbb{R}$ or $\mathbb{C}$ isn't, then $f + g$ is not measurable.

(d) If $f : X \to \mathbb{C}$ is measurable, then there exists $\alpha : X \to \mathbb{C}$ such that $|\alpha| = 1$ and $f = \alpha\,|f|$

Proof. (a) follows from Lemma 8.1.2 and noting $u = \text{Re}(u + iv), v = \text{Im}(u + iv)$, since Re,Im : $\mathbb{C} \to \mathbb{R}$ are continuous.

(b) follows from the fact $\mathbb{R}, \mathbb{C}$ are topological field and Lemma 8.1.2.

It is clear that the set of all function from X to $\mathbb{R}$ or $\mathbb{C}$ form a group under addition. (b) shows that the set of measurable functions from a subgroup, thus giving (c).

It remains to prove (d). Define

$$E \triangleq \{x \in X : f(x) = 0\} \text{ and } \varphi(z) \triangleq \frac{z}{|z|}$$

We claim

$$\alpha \triangleq \varphi \circ (f + \mathbf{1}_E) \text{ suffices}$$

Because f is measurable, we know E is measurable. This implies that $\mathbf{1}_E : X \to \mathbb{C}$ is measurable. It follows that $f + \mathbf{1}_E$ is $(X, \mathcal{B}_{\mathbb{C}})$-measurable. Note that $f + \mathbf{1}_E$ is never 0 on X.

Now by Theorem **??**, we see $f + \mathbf{1}_E$ is $(X, \mathcal{B}_{\mathbb{C}^*})$-measurable. It then follows from the fact $\varphi : \mathbb{C}^* \to \mathbb{C}$ is continuous that α is $(X, \mathcal{B}_{\mathbb{C}})$ measurable.

Observe that α maps E into $\{1\}$, and when $x \notin E$, we have $\alpha(x) = \frac{f(x)}{|f(x)|}$. (done) ∎

Note that Theorem 8.1.3 does not consider function whose range include ∞. This will be later addressed using approximation of simple functions.

Theorem 8.1.4. (Superior limit of measurable $f_n : X \to [-\infty, \infty]$ is measurable) Given a sequence $f_n : X \to [-\infty, \infty]$ of measurable functions

$$g \triangleq \sup f_n \text{ and } f \triangleq \limsup_{n\to\infty} f_n \text{ are both measurable}$$

Proof. It is straightforward to check

$$g^{-1}(\alpha, \infty] = \bigcup_n f_n^{-1}(\alpha, \infty]$$

It is straightforward to check

$$\mathcal{B}_{[-\infty,\infty]} = \sigma\Big(\{(\alpha,\infty] \subseteq [-\infty,\infty] : \alpha \in \mathbb{R}\}\Big)$$

These two facts and Theorem 8.1.1 shows g is measurable. The same arguments shows that $\inf g_k$ is measurable if g_k are all measurable.

It is straight forward to check

$$f = \inf_{n\geq 1} \sup_{k\geq n} f_k$$

It then follows f is measurable. ■

Corollary 8.1.5. (Pointwise limit of measurable $f_n : X \to [-\infty,\infty]$ is measurable) If the sequence $f_n : X \to [-\infty,\infty]$ pointwise converge to $f : X \to [-\infty,\infty]$, then f is measurable.

Corollary 8.1.6. (Positive and Negative parts of a measurable function are measurable) If we are given measurable $f : X \to [-\infty,\infty]$, and we define the **positive and negative part $f^+, f^- : X \to [0,\infty]$ of f** by

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

Then

$$f^+ \text{ and } f^- \text{ are measurable}$$

Proof. Define

$$h_n \triangleq \begin{cases} f & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even} \end{cases}$$

then we have

$$\limsup_{n\to\infty} h_n = f^+ \text{ and } \liminf_{n\to\infty} h_n = -f^-$$

Now, because 0 is a measurable function, by Theorem 8.1.4, f^+ is measurable. Moreover, because -1 is a measurable function, by Theorem 8.1.3, f^- is measurable. ■

8.2 Egorov's and Lusin's Theorem

8.3 Abstract integration

We say a function is a **simple function** if its range is of finite cardinality. Simple function is the cornerstone of the development of Lebesgue Integral Theory, as we shall see.

Suppose s is a simple function defined on some measurable $E \subseteq (X, \Sigma_X)$ with range $\{c_1, \cdots, c_n\} \subseteq [0, \infty]$. Define

$$E_j \triangleq \{x \in X : s(x) = c_j\}$$

It is clear that $\{E_j\}$ is a finite disjoint decomposition of E, and s is measurable if and only if all E_j are measurable.

We can write

$$s = \sum_{j=1}^{n} c_i \mathbf{1}_{E_j}$$

Then if s is measurable, we can define

$$\int_E s d\mu \triangleq \sum_{j=1}^{n} c_j \mu(E_j)$$

and if s is defined on some larger domain while $s|_E$ is measurable, we can define

$$\int_E s d\mu \triangleq \sum_{j=1}^{n} c_j \mu(E_j \cap E)$$

Note that if s is defined on some measurable F containing E and s is measurable on F, then s is surely measurable on E. It is straightforward to verify our definition so far is consistent.

We now expand our definitions to the class of all measurable functions range in $[0, \infty]$.

Given a function f range in $[0, \infty]$ and measurable on E, we define

$$\int_E f d\mu \triangleq \sup_{s \leq f \text{ on } E} \int_E s d\mu$$

If f is defined on some large measurable domain F, we see that

$$\int_E f d\mu = \int_F f \mathbf{1}_E d\mu$$

It is at this point one should verify our definition is so far consistent. That is, for each simple s, we have

$$\int_E s d\mu = \sup_{s' \leq s \text{ on } E} \int_E s' d\mu$$

The trick is to decompose each E'_j into $\bigcup E'_j \cap E_i$.

We now expand our definitions to the class of all measurable functions range in $[-\infty, \infty]$, but before such, we have to first introduce the idea of **Lebesgue integrable**. For either s or f, range in either $[0, \infty)$ or $[0, \infty]$, it is always possible that $\int_E s$ or $f d\mu = \infty$.

If we have $\int_E f d\mu = \infty$, we say f is **not Lebesgue integrable**, and if $\int_E f d\mu < \infty$, we say f is **Lebesgue integrable on** E. Because $\mathbb{R}$ is complete in order, we know f is either Lebesgue integrable or not Lebesgue integrable on E.

Given a function range in $[-\infty, \infty]$ and measurable on E, if both f^+, f^- are Lebesgue integrable on E, we say f is Lebesgue integrable on E, and write

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu$$

If either f^+ or f^- are not Lebesgue integrable on E, we say f is not Lebesgue integrable on E.

Notice that

$$f \text{ is Lebesgue integrable on } E \iff \int_E |f| \, d\mu < \infty$$

It is clear that our definition is again so far consistent.

We now

8.4 Basic property of abstract integration

This section prove some basic properties of Lebesgue integral over general measure space (X, Σ_X, μ). From now when we use the notation X, it shall be understood X is equipped with an σ-algebra Σ_X and a measure μ. We will prove

(a) Lebesgue Monotone Convergence Theorem (Theorem 8.4.1)

(b) Fatou's Lemma (Theorem 8.4.3)

(c) Reverse Fatou's Lemma (Theorem 8.4.4)

(d) Dominated Convergence Theorem (Theorem 8.4.5)

Theorem 8.4.1. (Lebesgue Monotone Convergence Theorem) Given a sequence of measurable $f_n : X \to [0, \infty]$ such that $\{f_n(x)\}_{n\in\mathbb{N}}$ is an increasing sequence for each $x \in X$, then

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

Proof. f is measurable by Corollary 8.1.5. Because $f_n \nearrow f$ on X, we know

$$\lim_{n\to\infty} \int_X f_n d\mu = \sup_n \int_X f_n d\mu \leq \int_X f d\mu$$

It remains to prove

$$\int_X f d\mu \leq \lim_{n\to\infty} \int_X f_n d\mu$$

Fix simple $0 \leq s \leq f$ on X. We reduce the problem into proving

$$\int_X s d\mu \leq \lim_{n\to\infty} \int_X f_n d\mu$$

Fix $c \in (0, 1)$. We reduce the problem into proving

$$c \int_X s d\mu \leq \lim_{n\to\infty} \int_X f_n d\mu$$

Define

$$E_n \triangleq \{x \in X : f_n(x) \geq cs(x)\}$$

E_n are measurable because $f_n - cs$ are measurable. Now because f_n are non-negative on X, we have

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

Taking limit, we see

$$\lim_{n\to\infty}\int_X f_n d\mu \geq \lim_{n\to\infty} c\int_{E_n} sd\mu$$

It is straightforward to check E_n is increasing and $\bigcup E_n = X$. Then if we decompose $s=\sum_j c_j \mathbf{1}_{F_j}$, by Theorem **??**, we can take limit

$$\lim_{n\to\infty}\mu(F_j\cap E_n)=\mu(F_j)$$

It then follows that

$$\lim_{n\to\infty}\int_X f_n d\mu \geq \lim_{n\to\infty} c\int_{E_n} sd\mu = c\int_X sd\mu \text{ (done)}$$

■

It is worth pointing out in our proof for Lebesgue Monotone Convergence Theorem, instead of proving $\int_X sd\mu \leq \lim_{n\to\infty}\int_X f_n d\mu$, we proved $c\int_X sd\mu \leq \lim_{n\to\infty}\int_X f_n d\mu$. Multiplying $\int_X sd\mu$ with $c\in(0,1)$ is not just a random limit technique. Our action play a much more profound role. Consider the Example.

Example 25 (Why we take $c\int_X sd\mu$?)

$$X=[0,1] \text{ and } f_n = 1-\frac{1}{n}$$

We can take $s=f$, and see $E_n=\varnothing$ for all n, which renders our proceeding proof invalid.

Corollary 8.4.2. (Monotone Convergence Theorem for general functions) Given a sequence of measurable $f_n : X\to[0,\infty]$ such that

(a) $\{f_n(x)\}_{n\in\mathbb{N}}$ is an increasing sequence on N^c

(b) $f:X\to[0,\infty]$ is the limit of f_n on N^c

(c) $\mu(N)=0$

(d) μ is complete

We have

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

Proof. Let

$$g(x) \triangleq \begin{cases} f(x) & \text{if } x \in N^c \\ 0 & \text{if } x \in N \end{cases}$$

Note that

$$\int_X f d\mu = \int_{N^c} f d\mu = \lim_{n\to\infty} \int_{N^c} f d\mu = \lim_{n\to\infty} \int_X f d\mu$$

■

Theorem 8.4.3. (Fatou's Lemma) Given measurable $f_n : X \to [0,\infty]$

$$\int_X \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu$$

Proof. Since $\inf_{k\geq n} f_k \leq f_n$ for each n, x, we see

$$\int_X \inf_{k\geq n} f_k d\mu \leq \int_X f_n d\mu \text{ for all } n$$

Because $\inf_{k\geq n} f_k \nearrow \liminf_{m\to\infty} f_m$ as $n \to \infty$, by (Theorem 8.4.1: Lebesgue Monotone Convergence Theorem), we can take limit

$$\int_X \liminf_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_X \inf_{k\geq n} f_k d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu$$

■

Example 26 (Fatou's Lemma strict inequality)

$$f_{2k}(x) = \begin{cases} 0 & \text{if } x \in [0, \frac{1}{2}] \\ 1 & \text{if } x \in (\frac{1}{2}, 1] \end{cases} \quad \text{and} \quad f_{2k+1}(x) = \begin{cases} 1 & \text{if } x \in [0, \frac{1}{2}) \\ 0 & \text{if } x \in [\frac{1}{2}, 1] \end{cases}$$

From now, we use $L^1(\mu)$ to denote the set of all function defined and μ-Lebesgue-integrable on X, and we say a sequence of function $f_n : X \to [0,\infty]$ is **dominated** by g, if g is a $[0,\infty]$-valued function defined on X such that

$$\sup_n |f_n(x)| \leq g(x) \text{ for all } x \in X$$

Theorem 8.4.4. (Reverse Fatou's Lemma) Given measurable $f_n : X \to [0,\infty]$ dominated by some $g \in L^1(\mu)$

$$\limsup_{n\to\infty} \int_X f_n d\mu \leq \int_X \limsup_{n\to\infty} f_n d\mu$$

Proof. By (Theorem 8.4.3: Fatou Lemma)

$$\int_X \liminf_{n\to\infty}(g-f_n)d\mu \le \liminf_{n\to\infty}\int_X (g-f_n)d\mu$$

Multiplying both side with -1, we have

$$\limsup_{n\to\infty}\int_X (f_n-g)d\mu \le \int_X \limsup_{n\to\infty}(f_n-g)d\mu$$

Then adding both side the constant $\int_X gd\mu$, we reach to the conclusion. ■

Note that in our proof above, when we "pull" the negative sign out from $\liminf_{n\to\infty}$, it changed to $\limsup_{n\to\infty}$. This is a standard technique, which can be justified using the sub-sequence definition of limit superior.

Theorem 8.4.5. (Dominate Convergence Theorem) Given a sequence $f_n : X \to \mathbb{C}\cup\{\infty\}$ of measurable function such that

$$f \triangleq \lim_{n\to\infty} f_n \text{ exists on } X$$

If there exists $g \in L^1(\mu)$ dominating f_n, then

$$\lim_{n\to\infty}\int_X |f_n-f|\,d\mu = 0 \text{ and } \int_X fd\mu = \lim_{n\to\infty}\int_X f_n d\mu \text{ exists in } \mathbb{C}\cup\{\infty\}$$

Proof. Because f is measurable by Corollary 8.1.5 and $|f|\le g$ on X, $f\in L^1(\mu)$.

We first prove

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

We relax the problem into proving

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

Note that $|f_n-f|\le 2g$. We can now apply (Theorem 8.4.3: Fatou lemma) to $2g-|f_n-f|$ and see

$$\begin{aligned}\int_X 2gd\mu &= \int_X \lim_{n\to\infty}(2g-|f_n-f|)d\mu \\ &\le \liminf_{n\to\infty}\int_X (2g-|f_n-f|)d\mu \\ &= \int_X 2gd\mu + \liminf_{n\to\infty}\Big(-\int_X |f_n-f|\,d\mu\Big) \\ &= \int_X 2gd\mu - \limsup_{n\to\infty}\int_X |f_n-f|\,d\mu\end{aligned}$$

Then because $g \in L^1(\mu)$, we can subtract it and obtain

$$\limsup_{n\to\infty} \int_X |f_n - f| \, d\mu = 0 \text{ (done)}$$

It then follows that

$$\limsup_{n\to\infty} \left| \int_X (f_n - f) d\mu \right| \le \limsup_{n\to\infty} \int_X |f_n - f| \, d\mu = 0$$

which implies

$$\lim_{n\to\infty} \int_X (f_n - f) d\mu = 0$$

and because $f \in L^1(\mu)$, we have

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

■

Example 27 (Counterexample for Dominate Convergence Theorem)

$$f_n(x) = \begin{cases} \frac{1}{n} & \text{if } |x| < n \\ 0 & \text{if } |x| \ge n \end{cases}$$

8.5 Equivalent Definitions of Lebesgue Measurable Functions and Integral

Chapter 9

Harmonic Analysis

9.1 Weierstrass approximation Theorem: $[a, b] \to \mathbb{R}$

Theorem 9.1.1. (Bernoulli's Inequality) Given $r, x \in \mathbb{R}$, suppose

(a) $r \geq 1$

(b) $x \geq -1$

Then

$$(1 + x)^r \geq 1 + rx$$

Proof. Fix $r \geq 1$. We wish

$$\text{to prove } (1 + x)^r \geq 1 + rx \text{ for all } x \geq -1$$

Define $f : [-1, \infty) \to \mathbb{R}$ by

$$f(x) = (1 + x)^r - (1 + rx) \tag{9.1}$$

We reduced the problem into

$$\text{proving } f(x) \geq 0 \text{ for all } x \geq -1$$

Because $r \geq 1$ by premise, by definition of $f(x)$ (Equation 9.1), we see that

$$f(0) = 0, \text{ and } f(-1) = r - 1 \geq 0$$

Notice that by definition of f (Equation 9.1), $f(x)$ is clearly differentiable on $(-1, \infty)$.

Then, by MVT (Theorem 4.3.3), to prove $f(x) \geq 0$ on $(-1, \infty)$, we only wish

$$\text{to prove } f'(x) \geq 0 \text{ for all } x > 0 \text{ and } f'(x) \leq 0 \text{ for all } x \in (-1, 0)$$

Compute f'

$$f'(x) = r(1+x)^{r-1} - r$$
$$= r\Big((1+x)^{r-1} - 1\Big)$$

Because $r \geq 1$, we can deduce

$$x > 0 \implies (1+x)^{r-1} \geq 1 \implies f'(x) = r\Big((1+x)^{r-1} - 1\Big) \geq 0$$

and deduce

$$x \in (-1,0) \implies 1+x \in (0,1) \implies (1+x)^{r-1} \leq 1 \implies f'(x) = r\Big((1+x)^{r-1} - 1\Big) \leq 0$$

(done) ∎

In this section, notation $\mathcal{C}\big([a,b]\big)$ means the set of **real-valued continuous function on** $[a,b]$.

Theorem 9.1.2. (Weierstrass approximation Theorem: $[a,b] \to \mathbb{R}$**)** Let $\mathbb{R}[x]\big|_{[a,b]}$ be the space of polynomials on $[a,b]$ with real coefficient. We have

$$\mathbb{R}[x]\big|_{[a,b]} \text{ is dense in } \Big(\mathcal{C}\big([a,b]\big), \|\cdot\|_\infty\Big)$$

Proof. WOLG, we can let $[a,b] = [0,1]$. The reason we can assume such is explained at last. Now, let $f : [0,1] \to \mathbb{R}$ be a continuous function. Fix ϵ. We only wish

$$\text{to find } P \in \mathbb{R}[x]\big|_{[0,1]} \text{ such that } \|f - P\|_\infty < \epsilon$$

Define $\tilde{f} \in \mathcal{C}\big([0,1]\big)$ by

$$\tilde{f}(x) = f(x) - f(0) - x\big[f(1) - f(0)\big] \tag{9.2}$$

It is easy to check $\tilde{f}$ is continuous. We first prove that

$$\big(\tilde{f}(x) - f(x)\big) \in \mathbb{R}[x]\big|_{[0,1]}$$

By definition of $\tilde{f}$ (Equation 9.2), we see

$$\tilde{f}(x) - f(x) = (f(0) - f(1))x - f(0) \in \mathbb{R}[x]\big|_{[0,1]} \text{ (done)}$$

This reduce our problem into

$$\text{finding } P \in \mathbb{R}[x]\big|_{[0,1]} \text{ such that } \|\tilde{f} - P\|_\infty < \epsilon$$

Notice that by definition of $\tilde{f}$ (Equation 9.2), we have

$$\tilde{f}(0) = 0 = \tilde{f}(1)$$

Then, we can expand the definition of $\tilde{f}$ by

$$\tilde{f}(x) = \begin{cases} \tilde{f}(x) & \text{if } x \in [0,1] \\ 0 & \text{if } x \notin [0,1] \end{cases} \tag{9.3}$$

This makes $\tilde{f}$ uniformly continuous on $\mathbb{R}$, since $\tilde{f}$ is uniformly continuous on $[0,1]$ and $[0,1]^c$. Now, for each $n \in \mathbb{N}$, define $Q_n \in \mathbb{R}[x]$ by

$$Q_n = c_n(1-x^2)^n \text{ where } c_n \text{ is chosen to satisfy } \int_{-1}^{1} Q_n(x)dx = 1 \tag{9.4}$$

Define $P_n : [0,1] \to \mathbb{R}$ by

$$P_n(x) = \int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt$$

We now prove

$$P_n \in \mathbb{R}[x]\big|_{[0,1]}$$

Because $\tilde{f}(x) = 0$ for all $x \notin (0,1)$ by definition of $\tilde{f}$ (Equation 9.3), we see that

$$P_n(x) = \int_{-x}^{1-x} \tilde{f}(x+t)Q_n(t)dt \text{ for all } x \in [0,1] \tag{9.5}$$

Fix $x \in [0,1]$. Now, by change of variable, we see

$$P_n(x) = \int_{-x}^{1-x} \tilde{f}(x+t)Q_n(t)dt = \int_0^1 \tilde{f}(u)Q_n(u-x)du$$

Because Q_n is a polynomial by definition (Equation 9.4), we can express $Q_n(u-x)$ by

$$Q_n(u-x) = \sum_{k=0}^{m} a_k x^k \text{ for some } \{a_0, \dots, a_m\} \text{ depending on } u$$

Then we see

$$P_n(x) = \int_0^1 \tilde{f}(u)Q_n(u-x)du = \sum_{k=0}^{m} x^k \Big(\int_0^1 \tilde{f}(u)a_k du \Big)$$

This shows that $P_n \in \mathbb{R}[x]\big|_{[0,1]}$
(done)

Now, because $\tilde{f}$ is uniformly continuous on $\mathbb{R}$, we can fix $\delta < 1$ such that

$$\forall x, y \in \mathbb{R}, |x-y| < \delta \implies \left|\tilde{f}(x) - \tilde{f}(y)\right| < \frac{\epsilon}{2} \tag{9.6}$$

By definition of $\tilde{f}$ (Equation 9.3), we know $\tilde{f}$ is a bounded function. Then we can set M by

$$M = \sup_{x \in \mathbb{R}} |f(x)|$$

Let n satisfy

$$4M\sqrt{n}(1-\delta^2)^n < \frac{\epsilon}{2} \tag{9.7}$$

Such n exists, because $\delta < 1 \implies \sqrt{n}(1-\delta^2)^n \to 0$. We claim

$$P_n \text{ satisfy } \|\tilde{f} - P_n\|_\infty < \epsilon$$

We first prove

$$c_n < \sqrt{n}$$

By Bernoulli's Inequality (Theorem 9.1.1). Compute

$$\begin{aligned} 1 = \int_{-1}^{1} Q_n(x)dx &= c_n \int_{-1}^{1} (1-x^2)^n dx \\ &= 2c_n \int_0^1 (1-x^2)^n dx \\ &\geq 2c_n \int_0^{\frac{1}{\sqrt{n}}} (1-x^2)^n dx \\ &\geq 2c_n \int_0^{\frac{1}{\sqrt{n}}} 1 - nx^2 dx = c_n\left(\frac{4}{3\sqrt{n}}\right) > c_n\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

This implies

$$\sqrt{n} > c_n \text{ (done)}$$

Because $\sqrt{n} > c_n$, by definition of Q_n (Equation 9.4), we have

$$Q_n(x) < \sqrt{n}(1-x^2)^n \leq \sqrt{n}(1-\delta^2)^n \text{ for all } x \text{ such that } \delta \leq |x| \leq 1$$

Fix $x \in [0,1]$. Finally, because

(a) $\int_{-1}^{1} Q_n(x)dx = 1$ by definition of Q_n (Equation 9.4)

(b) $Q_n(x) = c_n(1-x^2)^n \geq 0$ for all $x \in [-1,1]$

(c) $\left|\tilde{f}(x+t) - \tilde{f}(x)\right| < \frac{\epsilon}{2}$ for all t such that $|t| < \delta$, by definition of δ (Equation 9.7)

(d) $Q_n(x) \leq \sqrt{n}(1-\delta^2)^n$ for all x such that $\delta \leq |x| \leq 1$

(e) $4M\sqrt{n}(1-\delta^2)^n < \frac{\epsilon}{2}$ by definition of n (Equation 9.7)

we have

$$\begin{aligned}
\left|P_n(x) - \tilde{f}(x)\right| &= \left|\int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt - \tilde{f}(x)\right| \\
&= \left|\int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt - \tilde{f}(x)\int_{-1}^{1} Q_n(t)dt\right| \\
&= \left|\int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt - \int_{-1}^{1} \tilde{f}(x)Q_n(t)dt\right| \\
&= \left|\int_{-1}^{1} \big[\tilde{f}(x+t) - \tilde{f}(x)\big]Q_n(t)dt\right| \\
&\leq \int_{-1}^{1} \left|\big[\tilde{f}(x+t) - \tilde{f}(x)\big]Q_n(t)\right| dt \\
&= \int_{-1}^{1} \left|\tilde{f}(x+t) - \tilde{f}(x)\right| Q_n(t)dt \\
&\leq \int_{-1}^{-\delta} 2MQ_n(t)dt + \int_{-\delta}^{\delta} \left|\tilde{f}(x+t) - \tilde{f}(x)\right| Q_n(t)dt + \int_{\delta}^{1} 2MQ_n(t)dt \\
&\leq 2M\Big(\int_{-1}^{-\delta} Q_n(t)dt + \int_{\delta}^{1} Q_n(t)dt\Big) + \int_{-\delta}^{\delta} (\frac{\epsilon}{2})Q_n(t)dt \\
&\leq 4M(1-\delta)\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} < \epsilon
\end{aligned}$$

Because x is arbitrarily picked from $[0,1]$, we now have $\|P_n - \tilde{f}\|_\infty < \epsilon$ (done)

Lastly, we show

our result can be transplanted to arbitrary $\mathcal{C}\big([a,b]\big)$

Let $[a,b]$ be arbitrary. Fix ϵ and $f \in \mathcal{C}\big([a,b]\big)$. We wish

to find $P \in \mathbb{R}[x]\big|_{[a,b]}$ such that $\|f - P\|_\infty \leq \epsilon$

Define $g : [0,1] \to \mathbb{R}$ by

$$g(x) \triangleq f(a + (b-a)x) \tag{9.8}$$

We know there exists $P_n : [0,1] \to \mathbb{R}$ such that

$$\|P_n - g\|_\infty < \epsilon$$

Define $H_n : [a,b] \to \mathbb{R}$ by

$$H_n(x) = P_n\left(\frac{x-a}{b-a}\right)$$

Because P_n is a real polynomial on $[0,1]$, we know H_n is a real polynomial on $[a,b]$. We now claim

$$\text{such } H_n \text{ works}$$

Fix $x \in [a,b]$. Observe

$$\begin{aligned} \big|f(x) - H_n(x)\big| &= \left|f(x) - P_n\left(\frac{x-a}{b-a}\right)\right| \\ &= \left|g\left(\frac{x-a}{b-a}\right) - P_n\left(\frac{x-a}{b-a}\right)\right| < \epsilon \text{ (done)} \end{aligned}$$

■

It is at now, we will show that every real-valued continuous functions on $[a,b]$ can be approximated by polynomials with rational coefficient. This fact enable our computer to more easily approximate real-valued continuous function on $[a,b]$.

Note that since $\mathcal{C}\big([a,b]\big)$ is a separable metric space, we can show that $\mathcal{C}\big([a,b]\big)$ has cardinality of at most continuum $\mathfrak{c}$.

Theorem 9.1.3. (The space $\mathbb{Q}[x]\big|_{[a,b]}$ is dense in $\Big(\mathcal{C}\big([a,b]\big), \|\cdot\|_\infty\Big)$, thus $\mathcal{C}\big([a,b]\big)$ is separable)

$$\Big(\mathcal{C}\big([a,b]\big), \|\cdot\|_\infty\Big) \text{ is separable}$$

Proof. Because $\mathbb{Q}[x]\big|_{[a,b]}$ is countable, to show $\mathcal{C}\big([a,b]\big)$ is separable, we only wish to show

$$\mathbb{Q}[x]\big|_{[a,b]} \text{ is dense in } \mathcal{C}\big([a,b]\big)$$

Because $\mathbb{R}[x]\big|_{[a,b]}$ is dense in $\mathcal{C}\big([a,b]\big)$, we reduce our problem into proving

$$\mathbb{Q}[x]\big|_{[a,b]} \text{ is dense in } \mathbb{R}[x]\big|_{[a,b]}$$

Fix ϵ and $P \in \mathbb{R}[x]\big|_{[a,b]}$. We must

$$\text{find } Q \in \mathbb{Q}[x]\big|_{[a,b]} \text{ such that } \|Q - P\|_\infty \leq \epsilon$$

Express $P(x) = \sum_{k=0}^{n} r_k x^k$. Let $M > \max\{|a|, |b|\}$. Because $\mathbb{Q}$ is dense in $\mathbb{R}$, we know there exists $c_k \in \mathbb{Q}$ such that $|c_k - r_k| < \frac{\epsilon}{(n+1)M^n}$. We claim

$$Q(x) = \sum_{k=0}^{n} c_k x^k \text{ works}$$

Fix $x \in [a,b]$. See

$$\begin{aligned}
|P(x) - Q(x)| &= \left| \sum_{k=0}^{n} (c_k - r_k) x^k \right| \\
&\leq \sum_{k=0}^{n} |c_k - r_k| \cdot |x|^k \\
&\leq \sum_{k=0}^{n} |c_k - r_k| \cdot M^k \\
&\leq (M^n) \sum_{k=0}^{n} |c_k - r_k| \\
&< M^n (n+1) \Big(\frac{\epsilon}{(n+1)M^n} \Big) = \epsilon \text{ (done)}
\end{aligned}$$

$\blacksquare$

9.2 The Stone-Weierstrass Theorem

Recall that a **vector space over a field** $\mathbb{F}$ is a set V equipped with **vector addition** $+ : V \times V \to V$ and **scalar multiplication** such that

(a) $(V, +)$ is an abelian group.

(b) Scalar multiplication is compatible with field multiplication: $\Big((ab)v = a(bv)\Big)$

(c) Scalar multiplication is distributive: $\Big((a+b)v = av + bv$ and $a(v+w) = av + aw\Big)$

There are many ways to define the term **algebra over a field** $\mathbb{F}$. One can exhaust all the laws an algebra should obey. In short, an **algebra over a field** $\mathbb{F}$ (or $\mathbb{F}$**-algebra**) is a vector space V equipped with a vector multiplication such that

(i) Multiplication is <u>distributive</u> to addition.

(ii) Scalar and vector multiplications are compatible: $(av) \cdot (bw) = ab(v \cdot w)$

Given an arbitrary set E and a field $\mathbb{F}$, let A be the set of all functions from E to $\mathbb{F}$. The following is a list of some algebra

(a) $(\mathbb{R}^3, \text{cross product})$ over $\mathbb{R}$

(b) $(\mathbb{C}, \text{complex multiplication})$ over $\mathbb{C}$

(c) $(\mathbb{Q}[x], \text{function multiplication})$ over $\mathbb{Q}$

(d) $(\text{Functions from } E \text{ to } \mathbb{F}, \text{function multiplication})$ over $\mathbb{F}$

(e) $(\text{Continuous functions from } (E, \tau) \text{ to } \mathbb{C}\ , \text{function multiplication})$ over $\mathbb{C}$

(f) (Linear transformation from V to V, composition) over $\mathbb{F}$ where V is over $\mathbb{F}$

(g) $\big(M_n(\mathbb{F}), \text{matrix multiplication}\big)$ over $\mathbb{F}$

Note that $B = (\text{continuous functions from } \mathbb{C} \text{ to } \mathbb{C}, \text{composition})$ over $\mathbb{C}$ is not an algebra, even though B is both a vector space and a ring. ($\because$ scalar multiplication and multiplication are not compatible).

It is at here we shall introduce some general terminologies. Given an arbitrary set E, a field $\mathbb{F}$ and a point $x \in E$, we say a family $\mathcal{F}$ of functions from E to $\mathbb{F}$ **vanish at** x if for all $f \in \mathcal{F}$, we have $f(x) = 0$. We say $\mathcal{F}$ **separate points** in E if for all $x_2 \neq x_1 \in E$, there exists $f \in \mathcal{F}$ such that $f(x_2) \neq f(x_1)$.

Chapter 10

Calculus in Euclidean Space

10.1 Inverse Function Theorem

Interestingly, if $f : \Big(\mathbb{R}, \|\cdot\|_2\Big) \to \Big(\mathbb{R}^n, \|\cdot\|_2\Big)$ is a curve in $\mathbb{R}^n$

$$f(t) = (f_1(t), \cdots, f_n(t))$$

and we define

$$f'(t) \triangleq (f_1'(t), \cdots, f_n'(t))$$

We have

$$|f'(t)| = \|df_t\|_{\text{op}}$$

This give us the following expected result (Corollary 10.1.2).

Theorem 10.1.1. (Basic Property of Derivative) Suppose f maps a convex open set $E \subseteq \Big(\mathbb{R}^n, \|\cdot\|_2\Big)$ into $\Big(\mathbb{R}^m, \|\cdot\|_2\Big)$, f is differentiable on E, and there exists $M \in \mathbb{R}$ such that

$$\|df_x\|_{\text{op}} \leq M \qquad (x \in E)$$

Then for all $a, b \in E$, we have

$$|f(b) - f(a)| \leq M\,|b - a|$$

Proof. Define $\gamma : [0, 1] \to E$ by

$$\gamma(t) \triangleq a + (b - a)t$$

Now, note that

$$\begin{aligned}
|f(b)-f(a)| &= |(f\circ\gamma)(1)-(f\circ\gamma)(0)| \\
&= \left|\int_0^1 (f\circ\gamma)'(t)dt\right| \\
&\leq \int_0^1 |(f\circ\gamma)'(t)|\,dt \\
&= \int_0^1 \|d(f\circ\gamma)_t\|_{\text{op}}dt \\
&\leq \int_0^1 \|df_{\gamma(t)}\|_{\text{op}}\cdot\|d\gamma_t\|_{\text{op}}dt \\
&\leq \int_0^1 M\cdot|b-a|\,dt = M\,|b-a|
\end{aligned}$$

■

Corollary 10.1.2. (Basic Property of Derivative) Suppose f maps a convex open set $E\subseteq\Big(\mathbb{R}^n,\|\cdot\|_2\Big)$ into $\Big(\mathbb{R}^m,\|\cdot\|_2\Big)$, f is differentiable on E, and $df_x=0$ for all $x\in E$, then

$$f\text{ stay constant on }E$$

In this section, we will give a local statement and the proof for Inverse Function Theorem in $\mathbb{R}^n$ (Theorem 10.1.4). Let $L(\mathbb{R}^n)$ be the set of linear transformation that maps $\mathbb{R}^n$ into itself, and let Ω be the set of all invertibles in $L(\mathbb{R}^n)$. We will first prove that Ω is open (Theorem 10.1.3).

Theorem 10.1.3. (Ω is Open) Suppose $A\in\Omega$. If we define $\epsilon\triangleq\frac{1}{\|A^{-1}\|_{\text{op}}}$, then

$$B_\epsilon(A)\overset{\text{def}}{=}\{T\in L(\mathbb{R}^n):\|T-A\|_{\text{op}}<\epsilon\}\subseteq\Omega$$

Proof. Fix $T\in B_\epsilon(A)$ and $x\neq 0\in\mathbb{R}^n$. We are required to show

$$|Tx|>0$$

If $T=A$, then the proof is trivial. We therefore suppose $T\neq A$. Define

$$\beta\triangleq\|T-A\|_{\text{op}}$$

Note that $T\neq A\in B_\epsilon(A)$ implies $0<\beta<\epsilon$. We claim

$$(\epsilon-\beta)\,|x|\leq|Tx|$$

Observe

$$\epsilon |x| = \epsilon |A^{-1}Ax| \leq |Ax| \tag{10.1}$$

Observe

$$|Ax| \leq |(A-T)x| + |Tx| \leq \beta |x| + |Tx| \tag{10.2}$$

Equation 10.1 and Equation 10.2 implies

$$\epsilon |x| \leq \beta |x| + |Tx|$$

which implies

$$(\epsilon - \beta) |x| \leq |Tx| \quad \text{(done)}$$

■

Theorem 10.1.3 is essential for proving Inverse Function Theorem (Theorem 10.1.4). Think about what happen if f is linear. If f is linear, then f^{-1} is linear, and we will have

$$df^{-1} = f^{-1} = (f)^{-1} = (df)^{-1}$$

Because derivative is unique, it is reasonable to guess that if f is not linear and f^{-1} is differentiable, we would have

$$df^{-1} = (df)^{-1}$$

Now, if we wish df^{-1} to exists everywhere on $f(U)$, we must guarantee that df is invertible on U, and this is when Theorem 10.1.3 kick in. Note that in our proof of Inverse Function Theorem (Theorem 10.1.4), our selection of U guarantee df_x is invertible for all $x \in U$, by Theorem 10.1.3.

The rest of the proof boiled down to a fixed point argument (Theorem 2.8.1) to show f is one-to-one in U and $f(U)$ is open.

Theorem 10.1.4. (Inverse Function Theorem) Given a function f that maps an open neighborhood $E \subseteq \mathbb{R}^n$ around a into $\mathbb{R}^n$ such that

(a) f is differentiable on E

(b) df_a is invertible

(c) f is continuously differentiable at a

Then there exists open $U \subseteq E$ containing a such that

(a) f is one-to-one in U

(b) $f(U)$ is open

(c) The inverse of $f|_U$ is differentaible at $f(a)$.

Proof. Fix

$$\lambda \triangleq \frac{1}{2\|(df_a)^{-1}\|_{\text{op}}} \tag{10.3}$$

Because f is continuously differentaible at a, we know there exists δ such that

$$\|df_x - df_a\|_{\text{op}} < \lambda \qquad (x \in B_\delta(a)) \tag{10.4}$$

We claim

$$U \triangleq B_\delta(a) \text{ suffices}$$

For each $y \in \mathbb{R}^n$, define $\varphi_y : U \to \mathbb{R}^n$ by

$$\varphi_y(x) \triangleq x + (df_a)^{-1}(y - f(x))$$

Before anything, we first prove

$$\text{for all } y \in \mathbb{R}^n, \varphi_y : U \to \mathbb{R}^n \text{ is a contraction of } U$$

Fix $y \in \mathbb{R}^n$, and $x_1, x_2 \in U$. We claim

$$|\varphi_y(x_1) - \varphi_y(x_2)| \leq \frac{1}{2}|x_1 - x_2| \tag{10.5}$$

Because U is convex, Theorem 10.1.1 allow us to reduce the problem into proving

$$\|d(\varphi_y)_x\|_{\text{op}} \leq \frac{1}{2} \text{ for all } x \in U$$

Fix $x \in U$. Using Chain Rule (Theorem 4.5.2) and the fact that the derivative of a bounded linear transformation is itself, we can compute $d(\varphi_y)_x$

$$\begin{aligned} d(\varphi_y)_x &= I + (df_a)^{-1}(-df_x) \\ &= (df_a)^{-1}(df_a - df_x) \end{aligned}$$

This together with Equation 10.3 and Equation 10.4 give us

$$\|d(\varphi_y)_x\|_{\text{op}} \leq \|(df_a)^{-1}\|_{\text{op}}\|df_a - df_x\|_{\text{op}} < \frac{1}{2} \text{ (done)}$$

We now prove

$$f \text{ is one-to-one in } U$$

Fix y in $f(U)$. We wish to show

$$\text{there exists at most one } x \in U \text{ such that } f(x) = y$$

Because $f(x) = y \iff x$ is a fixed point of φ_y, we can reduce the problem into

$$\varphi_y \text{ has at most one fixed point}$$

Because φ_y is a contraction of U, Banach Fixed Point Theorem (Theorem 2.8.1) tell us φ_y has at most one fixed point. (done)

We now prove

$$f(U) \text{ is open in } \mathbb{R}^n$$

Fix $y_0 \in f(U)$. Let $x_0 = f^{-1}(y_0)$. Because U is open, we know there exists r such that

$$\overline{B_r(x_0)} \subseteq U$$

We claim

$$B_{\lambda r}(y_0) \subseteq f(U)$$

Fix $y \in B_{\lambda r}(y_0)$. We are required to prove

$$y \in f(U)$$

Because

$$y = f(x) \iff x \text{ is a fixed point of } \varphi_y$$

We then can use Banach Fixed Point Theorem (Theorem 2.8.1) to reduce the problem into proving

$$\varphi_y \text{ is a contraction that maps some complete subset of } U \text{ into itself}$$

We claim

$$\overline{B_r(x_0)} \text{ suffices}$$

We have already known φ_y is a contraction on U, and it is clear that $\overline{B_r(x_0)}$ is complete. We reduce the problem into proving

$$\varphi_y\big(\overline{B_r(x_0)}\big) \subseteq \overline{B_r(x_0)}$$

Using

(a) definition of φ_y

(b) $|y-y_0|<\lambda r$

(c) $\|(df_a)^{-1}\|_{\text{op}}=\frac{1}{2\lambda}$

We can deduce

$$\begin{aligned}|\varphi_y(x_0)-x_0| &= \big|(df_a)^{-1}(y-f(x_0))\big|\\ &\leq \|(df_a)^{-1}\|_{\text{op}}\,|y-y_0|<\frac{r}{2}\end{aligned}$$

Fix $x\in\overline{B_r(x_0)}$. We can now deduce

$$\begin{aligned}|\varphi_y(x)-x_0| &\leq |\varphi_y(x_0)-\varphi_y(x)|+|x_0-\varphi_y(x_0)|\\ &\leq \frac{1}{2}\,|x_0-x|+\frac{r}{2}\leq r \text{ (done)}\end{aligned}$$

We now prove

$$df_x \text{ is invertible for all } x\in U$$

Fix $x\in U$.

$$\|df_x-df_a\|_{\text{op}}\cdot\|(df_a)^{-1}\|_{\text{op}}<\frac{1}{2}$$

Theorem 10.1.3 now implies df_x is invertible. (done)

Lastly, it remains to prove

$$f^{-1}:f(U)\to U \text{ is differentiable on } f(U)$$

Fix $y\in f(U)$, and $x\triangleq f^{-1}(y)$. We are required to prove

$$\lim_{k\to 0}\frac{\big|f^{-1}(y+k)-x-(df_x)^{-1}k\big|}{|k|}=0$$

Fix $h(k)\triangleq f^{-1}(y+k)-f(x)$. In other words, $h\in\mathbb{R}^n$ is fixed to be the unique vector such that

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

We now see

$$\begin{aligned}f^{-1}(y+k)-x-(df_x)^{-1}k &= h-(df_x)^{-1}k\\ &= -(df_x^{-1})(f(x+h)-f(x)-df_xh)\end{aligned}$$

and see

$$\left|f^{-1}(y+k)-x-(df_x)^{-1}k\right| \leq \|(df_x)^{-1}\|_{\text{op}}\,|f(x+h)-f(x)-df_xh|$$

which give us

$$\frac{\left|f^{-1}(y+k)-x-(df_x)^{-1}k\right|}{|k|} \leq \|(df_x)^{-1}\|_{\text{op}}\frac{|f(x+h)-f(x)-df_xh|}{|h|}\cdot\frac{|h|}{|k|}$$

This allow us to reduce the problem into proving

$$\limsup_{k\to 0}\frac{|h|}{|k|}\in\mathbb{R}$$

We claim

$$\frac{|h|}{|k|}\leq\lambda^{-1}\text{ for all }k\text{ such that }y+k\in f(U)$$

Compute

$$\varphi_y(x+h)-\varphi_y(x)=h-(df_a)^{-1}k$$

Equation 10.5 let us deduce

$$\left|h-(df_a)^{-1}k\right|=|\varphi_y(x+h)-\varphi_y(x)|\leq\frac{|h|}{2}$$

This with triangle inequality implies

$$\|(df_a)^{-1}\|_{\text{op}}\,|k|\geq\left|(df_a)^{-1}k\right|\geq\frac{|h|}{2}\text{ (done)}$$

■

The following is a technical recap of our proof for the Inverse Function Theorem (Theorem 10.1.4).

1: Let $\lambda\triangleq\frac{1}{2\|(df_a)^{-1}\|_{\text{op}}}$

2: Claim $B_\delta(a)$ suffices to be U, where $\|df_x-df_a\|_{\text{op}}<\lambda$

3: For each $y\in\mathbb{R}^n$, define $\varphi_y:U\to\mathbb{R}^n$ by $\varphi_y(x)\triangleq x+(df_a)^{-1}(y)$.

4: Prove that φ_y is a contraction of U by taking derivative and utilize step 1 and 2.

5: Prove that φ_y fix $x\iff f(x)=y$

6: Prove f is one-to-one in U using step 4,5.

7: Prove $f(U)$ is open by proving $B_{\lambda r}(y_0)\subseteq f(U)$, while $\overline{B_r(x_0)}\subseteq U$. The proof use step 4,5, some computation and ultimately claim that φ_y admits a fixed point as φ_y maps $\overline{B_r(x_0)}$ into itself.

8: Prove df_x is invertible in U by Theorem 10.1.3

9: Algebraically prove $df^{-1} = (df)^{-1}$, using $\left|h - (df_a)^{-1}k\right| = |\varphi_y(x+h) - \varphi_y(x)| \leq \frac{|h|}{2}$

Theorem 10.1.5. (Inversion is Continuous)

$$\text{The mapping } A \to A^{-1} \text{ is continuous on } \Omega$$

Proof. Fix $A \in \Omega$ and let $T \in \Omega$. We are required to prove

$$\lim_{T \to A} \|T^{-1} - A^{-1}\|_{\text{op}} = 0$$

We know

$$T^{-1} - A^{-1} = T^{-1}(A - T)A^{-1}$$

This implies

$$\|T^{-1} - A^{-1}\|_{\text{op}} \leq \|T^{-1}\|_{\text{op}} \|A - T\|_{\text{op}} \|A^{-1}\|_{\text{op}}$$

This allow us to reduce the problem into proving

$$\limsup_{T \to A} \|T^{-1}\|_{\text{op}} \in \mathbb{R}$$

Fix $\epsilon \triangleq \frac{1}{\|A^{-1}\|_{\text{op}}}$, $T \in B_\epsilon(A)$ and $\beta \triangleq \|T - A\|_{\text{op}} < \epsilon$. We claim

$$\|T^{-1}\|_{\text{op}} \leq (\epsilon - \beta)^{-1}$$

Following the proof of Theorem 10.1.3, we have

$$(\epsilon - \beta)\,|x| \leq |Tx| \text{ for all } x \in \mathbb{R}^n$$

This implies

$$\frac{\left|T^{-1}x\right|}{|x|} \leq (\epsilon - \beta)^{-1} \text{ for all } x \neq 0 \in \mathbb{R}^n \quad \text{(done)}$$

■

Corollary 10.1.6. (Continuously Differentaible Version of Inverse Function Theorem) Given a function f that maps open $E \subseteq \mathbb{R}^n$ containing a into $\mathbb{R}^n$ such that

(a) f is differentiable on E

(b) df_a is invertible

(c) f is continuously differentiable on E

Then there exists open $U \subseteq E$ containing a such that

$$f(U) \text{ is open and } f|_U : U \to f(U) \text{ is a diffeomorphism}$$

10.2 Implicit Function Theorem

Some notations first. If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^k)$. We define $A|_{\mathbb{R}^n} : L(\mathbb{R}^n, \mathbb{R}^k)$ by

$$A|_{\mathbb{R}^n}(x) \triangleq A(x, 0)$$

Theorem 10.2.1. (Implicit Function Theorem) Suppose a function f that maps open $E \subseteq \mathbb{R}^n \times \mathbb{R}^m$ containing (a, b) into $\mathbb{R}^n$ satisfy

(a) $f(a, b) = 0$

(b) $(df_{(a,b)})|_{\mathbb{R}^n}$ is invertible

(c) f is continuously differentiable on E

Then there exists open $U \subseteq E$ containing (a, b) and open $W \subseteq \mathbb{R}^m$ containing b such that there exists a unique function g from W to $\mathbb{R}^n$ such that

(a) $\big(g(y), y\big) \in U$ for all $y \in W$

(b) $f(g(y), y) = 0$ for all $y \in W$

Moreover, g satisfy

(a) g is continuously differentiable on W

(b) g satisfy $dg_b = -(df_{(a,b)}|_{\mathbb{R}^n})^{-1} \circ df_{(a,b)}|_{\mathbb{R}^m}$

Proof. Define $F : E \to \mathbb{R}^n \times \mathbb{R}^m$ by

$$F(x, y) \triangleq \Big(f(x, y), y\Big)$$

Because f is continuously differentiable on E, using Differentaibility Theorem (Theorem 4.4.2), we can deduce

$$F \text{ is continuously differentiable on } E$$

Again using Differentiability Theorem (Theorem 4.4.1), we can write down $dF_{(a,b)}$ in the matrix form with respect to standard basis

$$[dF_{(a,b)}] = \begin{bmatrix} df_{(a,b)}|_{\mathbb{R}^n} & O \\ df_{(a,b)}|_{\mathbb{R}^m} & I \end{bmatrix}$$

Now, because $df_{(a,b)}|_{\mathbb{R}^n}$ is invertible, we know $dF_{(a,b)}$ is invertible.

We can now apply Inverse Function Theorem (Theorem 10.1.4) to $F : E \to \mathbb{R}^n \times \mathbb{R}^m$. This give us

(a) an open $U \subseteq E \subseteq \mathbb{R}^n \times \mathbb{R}^m$ containing (a, b)

(b) open $V \triangleq F(U) \subseteq \mathbb{R}^n \times \mathbb{R}^m$ containing $(0, b)$

(c) $F|_U : U \to V$ is a diffeomorphism.

Define W by

$$W \triangleq \{y \in \mathbb{R}^m : (0, y) \in V\}$$

We claim

$$\text{such } U, W \text{ suffices}$$

Note that it is easy to check W is open, utilizing V is open and the same ϵ.

Now, because $F|_U : U \to V$ is bijective, we know for each $y \in W$, there exists unique $(x, y) \in U$ such that

$$F(x, y) = (0, y)$$

We can now well define a function $g : W \to \mathbb{R}^n$ such that

$$\Big(g(y), y\Big) \in U \text{ and } f(g(y), y) = 0 \text{ for all } y \in W$$

It remains to show

(a) g is continuously differnetiable on W

(b) $dg_b = -(df_{(a,b)})|_{\mathbb{R}^n}^{-1}(df_{(a,b)})|_{\mathbb{R}^m}$

Fix $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. We wish to prove

$$\partial_i g_j \text{ exists and is continuous on } W$$

Express

$$g(y_1, \dots, y_m) = \Big(g_1(y_1, \dots, y_m), \dots, g_n(y_1, \dots, y_m)\Big)$$

Express

$$F^{-1}(z_1, \dots, z_{n+m}) = \Big(F_1^{-1}(z_1, \dots, z_{n+m}), \dots, F_{n+m}^{-1}(z_1, \dots, z_{n+m})\Big)$$

Because F^{-1} is continuously differentiable on V, we reduce the problem into proving

$$\partial_i g_j(y) = \partial_{n+i} F_j^{-1}(0, y) \text{ for all } y \in W$$

Because $F^{-1}(0,y) = (g(y),y)$ for all $y \in W$, we know

$$F_j^{-1}(0,\dots,0,y_1,\dots,y_m) = g_j(y_1,\dots,y_m) \text{ for all } y \in W \tag{10.6}$$

Fix arbitrary $y = (y_1,\dots,y_m) \in W$. Because W is open, we can see from Equation 10.6 that

$$\begin{aligned}\partial_i g_j(y) &= \lim_{t\to 0} \frac{g_j(y_1,\dots,y_i+t,\dots,y_m)}{t} \\ &= \lim_{t\to 0} \frac{F_j^{-1}(0,\dots,0,y_1,\dots,y_i+t,\dots,y_m)}{t} \\ &= \partial_{n+i} F_j^{-1}(y) \text{ (done)}\end{aligned}$$

Define $\Phi : W \to U$ by

$$\Phi(y) = \Big(g(y), y\Big)$$

By definition of g, we have

$$f \circ \Phi = 0 \text{ on } W$$

This by Chain Rule give us

$$df_{\Phi(y)} \circ d\Phi_y = 0 \text{ on } W$$

In particular

$$df_{(a,b)} \circ d\Phi_b = 0$$

Now, compute

$$d\Phi_b = \begin{bmatrix} dg_b \\ I \end{bmatrix} \text{ and } df_{(a,b)} = \begin{bmatrix} df_{(a,b)}|_{\mathbb{R}^n} & df_{(a,b)}|_{\mathbb{R}^m} \end{bmatrix}$$

This then give us

$$df_{(a,b)}|_{\mathbb{R}^n} \circ dg_b + df_{(a,b)}|_{\mathbb{R}^m} = 0$$

and of course

$$dg_b = -\big(df_{(a,b)}|_{\mathbb{R}^n}\big)^{-1} \circ df_{(a,b)}|_{\mathbb{R}^m} \text{ (done) (done)}$$

■

Example 28 (Unit Circle Example)

$$f(x,y) \triangleq x^2 + y^2 - 1 \text{ and } (a,b) \triangleq (1,1)$$

We have

$$g(y) = \sqrt{2 - y^2} \text{ on } y \in (1-\epsilon, 1+\epsilon)$$

Compute

$$df_{(a,b)} = \begin{bmatrix} 2 & 2 \end{bmatrix} \text{ and } dg_1 = \begin{bmatrix} -1 \end{bmatrix}$$

This established

$$dg_a = -(df_{(a,b)}|_{\mathbb{R}^1})^{-1} \circ df_{(a,b)}|_{\mathbb{R}^1}$$

Example 29 (Implicit Function Theorem Implies Inverse Function Theorem)

Given continuously differentaible $h : E \overset{\text{open}}{\subseteq} \mathbb{R}^n \ni a \to \mathbb{R}^n$ such that dh_a is invertible

Define $f : E \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$f(x,y) \triangleq h(x) - y$$

It is easily checked that $f(x, h(x)) = 0$ and the rest of the condition is satisfied. Now by Implicit Function Theorem (Theorem 10.2.1), we see that there exists $g : W \overset{\text{open}}{\subseteq} \mathbb{R}^n \to E$ such that

$$f(g(y), y) = 0 \text{ for all } y \in W$$

In other words,

$$h(g(y)) = y \text{ for all } y \in W$$

10.3 Feynman's Trick

In this section

Theorem 10.3.1. (Feynman's Trick) Given a real-valued function $f(x,t)$ defined on $[a,b]\times[c,d]$, and an real-valued function α of bounded variation on $[a,b]$, such that

(a) $\partial_2 f(x,t)$ exists on $[a,b]\times[c,d]$

(b) For all $t\in[c,d]$, the integral $\int_a^b f(x,t)d\alpha(x)$ exists.

(c) For all $s\in[c,d]$ and ϵ, there exists δ such that

$$|\partial_2 f(x,t)-\partial_2 f(x,s)|<\epsilon \text{ for all } x\in[a,b] \text{ and all } t\in(s-\delta,s+\delta)$$

Then we have

$$\frac{d}{dt}\int_a^b f(x,t)d\alpha(x)=\int_a^b \frac{\partial}{\partial t}f(x,t)d\alpha(x)$$

In other words, if we define $g(t)\triangleq\int_a^b f(x,t)d\alpha(x)$, then we have

$$g'(t)=\int_a^b \partial_2 f(x,t)d\alpha(x)$$

Proof. Fix $s\in[c,d]$. We are required to prove

$$g'(s)=\int_a^b \partial_2 f(x,s)dx$$

Note that, for all $t\neq s\in[c,d]$, we have

$$\frac{g(t)-g(s)}{t-s}=\int_a^b \frac{f(x,t)-f(x,s)}{t-s}d\alpha(x)$$

This allow us to reduce the problem into proving

$$\frac{f(x,t)-f(x,s)}{t-s}\to\partial_2 f(x,s) \text{ uniformly for all } x\in[a,b] \text{ as } t\to s$$

By MVT (Corollary 4.3.3), we know for all $t\neq s\in[c,d]$, there exists u_t between t and s such that

$$\frac{f(x,t)-f(x,s)}{t-s}=\partial_2 f(x,u_t)$$

The proof now follows from (c). (done) ■

Example 30 (Introductive application of Feynman's Trick)

$$\text{What is the value of } \int_0^1 \frac{x-1}{\ln x} dx \ ?$$

Define $f(x,t) \triangleq \frac{x^t-1}{\ln x}$ on $[0,1]\times[0,1]$. Observe $\partial_2 f(x,t) = x^t$, and observe

$$\int_0^1 f(x,0)dx = 0 \text{ and } \int_0^1 \partial_2 f(x,t)dx = \frac{1}{t}$$

We can con compute

$$\begin{aligned}\int_0^1 \frac{x-1}{\ln x}dx &= \int_0^1 f(x,1)dx \\ &= \int_0^1 f(x,0)dx + \int_0^1 \Big(\int_0^1 \partial_2 f(x,t)dx\Big)dt = 0\end{aligned}$$

Example 31 (Dirichlet's Integral)

$$\text{What is the value of } \int_0^\infty \frac{\sin t}{t} dt \ ?$$

Define the Laplace transformation

$$f(s,t) \triangleq e^{-st}\frac{\sin t}{t} \text{ on } \mathbb{R}_0^+ \times \mathbb{R}^+$$

Observe

$$\partial_1 f(s,t) = -e^{-st}\sin t$$

Now compute

$$\begin{aligned}\int_0^\infty -e^{-st}\sin t dt &= \frac{1}{-2i}\int_0^\infty e^{-st}(e^{it}-e^{-it})dt \\ &= \frac{1}{-2i}\Big(\frac{e^{t(i-s)}}{i-s} - \frac{e^{t(-i-s)}}{-i-s}\Big)\Big|_{t=0}^\infty \\ &= \frac{-1}{1+s^2} = \frac{d}{ds}(-\arctan s)\end{aligned}$$

It is clear that

$$\lim_{s\to\infty}\int_0^\infty f(s,t)dt = 0$$

We now have

$$\begin{aligned}\int_0^\infty \frac{\sin t}{t} dt &= \int_0^\infty f(0,t)dt \\ &= \lim_{s\to\infty} \int_0^\infty f(s,t)dt - \int_0^\infty \int_0^\infty \partial_1 f(s,t)dtds \\ &= \int_0^\infty \frac{1}{1+s^2} ds = \frac{\pi}{2}\end{aligned}$$

10.4 Appendix: Linear Algebra

This section contains

(a) definition and basic properties of the term **norm**

(b) definition and basic properties of the term **inner product**

(c) definition and basic properties of the term **positive semi-definite Hermitian form**

(d) full statement and proof of **Cauchy Schwarz Inequality** for both inner product space and positive semi-definite Hermitian form

(e) statement and proof of **SVD** (singular value decomposition).

(Norm Axiom Part)

Recall that by a **normed space** V, we mean a vector space over a sub-field $\mathbb{F}$ of $\mathbb{C}$ equipped with $\|\cdot\| : V \to \mathbb{R}_0^+$ satisfying the following axioms:

(a) $\|x\| = 0 \implies x = 0$ (positive-definiteness)

(b) $\|sx\| = |s| \cdot \|x\|$ for all $s \in \mathbb{F}$ and $x \in V$ (absolute-homogenity)

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality)

Observe

$$\|0\| = \|0 + x\| \leq \|0\| + \|x\| \text{ for all } x \in V$$

This shows that $\|x\| \geq 0$ for all $x \in V$. Also observe

$$\|0\| = \|0(x)\| = |0| \cdot \|x\| = 0$$

We can now rewrite the normed space axioms into

(a) $\|x\| = 0 \iff x = 0$ (positive-definiteness)

(b) $\|sx\| = |s| \cdot \|x\|$ for all $s \in \mathbb{F}$ and $x \in V$ (absolute-homogeneity)

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (triangle inequality)

(d) $\|x\| \geq 0$ for all $x \in V$ (non-negativity)

(Inner Product Axiom Part)

Recall that by an **inner product space** V, we mean a vector space over $\mathbb{R}$ or $\mathbb{C}$ equipped with $\langle \cdot, \cdot \rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$ satisfying the following <u>axioms</u>

(a) $\langle x, x \rangle > 0$ for all $x \neq 0$ (Positive-definiteness)

(b) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (Conjugate symmetry)

(c) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ and $\langle cx, z \rangle = c\langle x, z \rangle$ (Linearity in the first argument)

Note that conjugate symmetry let us deduce

$$\langle x, x \rangle = \overline{\langle x, x \rangle} \implies \langle x, x \rangle \in \mathbb{R}$$

Also, one can easily use linearity in first argument to deduce

$$\langle 0, 0 \rangle = 2\langle 0, 0 \rangle \implies \langle 0, 0 \rangle = 0$$

This now let us rewrite the inner product space over $\mathbb{C}$ axioms into

(a) $\langle x, x \rangle \geq 0$ for all $x \in V$ (non-negativity)

(b) $\langle x, x \rangle = 0 \iff x = 0$ (positive-definiteness)

(c) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (conjugate symmetry)

(d) $\langle cx + y, z \rangle = c\langle x, z \rangle + \langle y, z \rangle$ and $\langle x, cy + z \rangle = \overline{c}\langle x, y \rangle + \langle x, z \rangle$ (Linearity)

Note that using $c = 1$ and $y = 0$, ($\because \langle 0, z \rangle = 0\langle x, z \rangle = 0$) one can check that the latter expression of linearity implies the first expression.

If the scalar field is $\mathbb{R}$, then conjugate symmetry is just symmetry and we also have linearity in the second argument.

This now let us rewrite the inner product space over $\mathbb{R}$ axioms into

(a) $\langle x, x \rangle \geq 0$ for all $x \in V$ (non-negativity)

(b) $\langle x, x \rangle = 0 \iff x = 0$ (positive-definiteness)

(c) $\langle x, y \rangle = \langle y, x \rangle$ (symmetry)

(d) Linearity in both arguments

If we do not require $\langle\cdot,\cdot\rangle$ to be positive-definite, but only non-negative, i.e. $\langle x,x\rangle \geq 0$ for all $x \in V$, then we have a **positive semi-definite Hermitian form**. Formally speaking, a positive semi-definite Hermitian form $\langle\cdot,\cdot\rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$ satisfy the following axioms

(a) $\langle x,x\rangle \geq 0$ for all $x \in V$ (non-negativity)

(b) $\langle x,y\rangle = \overline{\langle y,x\rangle}$ (conjugate symmetry)

(c) $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$ and $\langle cx,z\rangle = c\langle x,z\rangle$ (Linearity in the first argument)

Example 32 (Example of Positive semi-definite Hermitian form)

arbitrary V over $\mathbb{R}$ or $\mathbb{C}$ $\quad \langle x,y\rangle \triangleq 0$ for all x,y

(Norm Induce Part)

Given a vector space V over $\mathbb{R}$ or $\mathbb{C}$, one can check that if V is equipped with an inner product $\langle\cdot,\cdot\rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$, then we can induce a norm on V by

$$\|x\| \triangleq \sqrt{\langle x,x\rangle} \qquad (x \in V)$$

Note that

$$\|x\| = 0 \iff \langle x,x\rangle = 0$$

This implies that if $\langle\cdot,\cdot\rangle$ is an inner product (satisfy positive-definiteness), then $\|\cdot\|$ is also positive-definite. And if $\langle\cdot,\cdot\rangle$ is not positive-definite, then there exists $x \neq 0 \in V$ such that $\|x\| = 0$, which make $\|\cdot\|$ a **semi-norm**.

Absolute homogeneity follows from the linearity of inner product.

To check triangle inequality, we first have to prove Cauchy-Schwarz inequality.

Theorem 10.4.1. (Basic Property of Positive semi-definite Hermitian form) Given a positive semi-definite Hermitian form $\langle\cdot,\cdot\rangle : V^2 \to \mathbb{R}$ or $\mathbb{C}$ and $x,y \in V$, we have

$$\langle x,x\rangle = 0 \implies \langle x,y\rangle = 0$$

Proof. Assume $\langle x,y\rangle \neq 0$. Fix $t > \frac{\|y\|^2}{2|\langle x,y\rangle|^2}$. Compute

$$\begin{aligned}\|y - t\langle y,x\rangle x\|^2 &= \|y\|^2 + \|(-t)\langle y,x\rangle x\|^2 + \langle -t\langle y,x\rangle x, y\rangle + \langle y, -t\langle y,x\rangle x\rangle \\ &= \|y\|^2 + t^2 \left|\langle x,y\rangle\right|^2 \|x\|^2 - t\langle y,x\rangle\langle x,y\rangle - t\langle x,y\rangle\langle y,x\rangle \\ &= \|y\|^2 - 2t\left|\langle x,y\rangle\right|^2 < 0 \text{ CaC}\end{aligned}$$

∎

Theorem 10.4.2. (Cauchy-Schwarz Inequality) Given a positive semi-definite Hermitian form $\langle\cdot,\cdot\rangle : V^2 \to \mathbb{C}$ on vector space V over $\mathbb{C}$, we have

(a) $|\langle x,y\rangle| \leq \|x\| \cdot \|y\| \quad (x,y \in V)$

(b) the equality hold true if x, y are linearly dependent

(c) the equality hold true if and only if x, y are linearly dependent (provided $\langle\cdot,\cdot\rangle$ is an inner product)

Proof. We first prove

$$|\langle x,y\rangle| \leq \|x\| \cdot \|y\| \qquad (x,y \in V)$$

Fix $x, y \in V$. Theorem 10.4.1 tell us $\|x\| = 0 \implies \langle x,y\rangle = 0$. Then we can reduce the problem into proving

$$\frac{|\langle x,y\rangle|^2}{\|x\|^2} \leq \|y\|^2$$

Set $z \triangleq y - \frac{\langle y,x\rangle}{\|x\|^2}x$. We then have

$$\langle z,x\rangle = \langle y - \frac{\langle y,x\rangle}{\|x\|^2}x, x\rangle = \langle y,x\rangle - \frac{\langle y,x\rangle}{\|x\|^2}\langle x,x\rangle = 0$$

Then from $y = z + \frac{\langle y,x\rangle}{\|x\|^2}x$, we can now deduce

$$\begin{aligned}\langle y,y\rangle &= \langle z + \frac{\langle y,x\rangle}{\|x\|^2}x, z + \frac{\langle y,x\rangle}{\|x\|^2}x\rangle \\ &= \langle z,z\rangle + \left|\frac{\langle y,x\rangle}{\langle x,x\rangle}\right|^2 \langle x,x\rangle \\ &= \langle z,z\rangle + \frac{|\langle x,y\rangle|^2}{\langle x,x\rangle}\end{aligned}$$

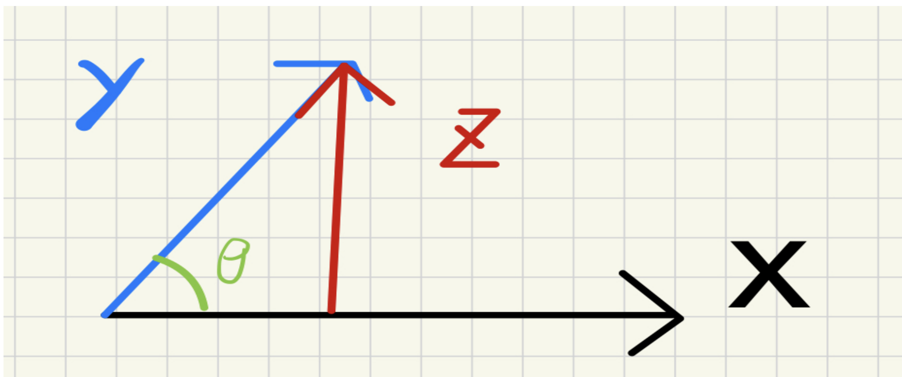
Because $\langle z, z\rangle \geq 0$, we now have

$$\langle y, y\rangle = \langle z, z\rangle + \frac{|\langle x, y\rangle|^2}{\langle x, x\rangle} \geq \frac{|\langle x, y\rangle|^2}{\langle x, x\rangle} \text{ (done)}$$

The equality hold true if and only if $\langle z, z\rangle = 0$. This explains the other two statements regarding the equality. ■

The proof is clearly geometrical. If one wish to remember the proof, one should see the trick we use is exactly

$$z \triangleq y - |y|(\cos\theta)\hat{x} \text{ is the projection of } y \text{ onto } x^{\perp}$$



Then all we do rest is just expanding $|y|^2 = |z + \tilde{x}|^2$, where $\tilde{x} = y - z = |y|(\cos\theta)\hat{x}$, which give the answer and is easy to compute since $z \cdot \tilde{x} = 0$.

Now, with Cauchy-Schwarz Inequality, we can check the triangle inequality

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \langle x, x\rangle + \langle y, y\rangle + 2 \text{ Re } \langle x, y\rangle \\
&\leq \|x\|^2 + \|y\|^2 + 2\,|\langle x, y\rangle| \\
&\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = \Big(\|x\| + \|y\|\Big)^2
\end{aligned}$$

(Euclidean Space Abstract Part)

By a **concrete Euclidean Space**, we mean some space of n-tuple $(x_1, \ldots, x_n)$ over $\mathbb{R}$, equipped with inner product $\langle \cdot, \cdot \rangle_E$ defined by

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle_E = \sqrt{\sum_{k=1}^{n} (y_k - x_k)^2}$$

By an **Euclidean Space**, we simply mean a finite dimensional vector space V over $\mathbb{R}$, equipped with an inner product $\langle \cdot, \cdot \rangle$ such that there exists a concrete Euclidean space E and an isomorphism $\varphi : V \to E$ such that

$$\langle x, y \rangle = \langle \varphi(x), \varphi(y) \rangle_E \qquad (x, y \in V)$$

Note that if you define $\langle \cdot, \cdot \rangle$ on the space of n-tuples $(x_1, \dots, x_n)$ over $\mathbb{R}$ by

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = 2\sqrt{\sum_{k=1}^{n} (y_k - x_k)^2}$$

Then, the space of n-tuple is clearly not a concrete Euclidean space, and clearly an Euclidean space.

(SVD)

Chapter 11

Differential Geometry

11.1 Smooth Manifold

Abstract

This section prove that smooth manifolds always admit smooth partition of unity and some of its corollaries which are heavily used throughout the preceding sections.

Given a topological space $(X, \mathscr{T})$ and fix n, if $: U \to \mathbb{R}^n$ is a homeomorphism between some open subspace U of X and some open subspace of $\mathbb{R}^n$, we say (Φ, U) is a **chart** on X, and if there exists a collection $A = \{(\Phi_i, U_i)\}_{i \in I}$ of chart cover the whole X, we say A is an **atlas** and X is **locally Euclidean**. We now give definition to the term **topological manifold**, which we will use throughout this chapter.

Definition 11.1.1. (Definition of Topological Manifold) We say a Topological space X is a **topological manifold** if

(a) X is locally Euclidean.

(b) X is Hausdorff.

(c) X is second countable.

Immediately, one should check that all three conditions are necessary: There are locally Euclidean space that is not Hausdorff, Bug-Eyed Line for example. There also are locally Euclidean space that is not second countable, Long Line for example. Also, one can check that any open subspace or finite products of topological manifold is still topological manifold. To have a better understanding why we require more than locally Euclidean in our definition of topological manifold, we first introduce some more topological notions. Given a collection of subsets (E_α) of some topological space X, we say (E_α) is **locally finite** if for each $p \in X$, there exists some neighborhood of p intersecting with only finitely many of E_α. Given an open cover (E_α) of topological space M, we say a family of continuous function $\psi_\alpha : M \to \mathbb{R}$ is a **partition of unity subordinate to (or dominated by)** (E_α) if

(i) $0 \leq \psi_\alpha(x) \leq 1$ for all $\alpha \in A, x \in M$.

(ii) $\operatorname{supp} \psi_\alpha \subseteq X_\alpha$ for all α.

(iii) The collection $\{\operatorname{supp} \psi_\alpha \subseteq M : \alpha \in A\}$ is locally finite.

(iv) $\sum_{\alpha \in A} \psi_\alpha(x) = 1$ for all $x \in M$. (Note that the sum is finite because $\{\operatorname{supp} \psi_\alpha\}$ is locally finite)

Given a cover (E_α) of topological space X, we say another cover (F_β) of X is a **refinement of** E if for each β there exists some α such that $F_\beta \subseteq E_\alpha$. Suppose we have some open cover $(U_\alpha)_{\alpha \in A}$ of some topological space X A topological space X if said to be **paracompact** if every open cover has a locally finite open refinement. It shall be clear that each compact space is paracompact.

Theorem 11.1.2. (Topological Manifold are Paracompact) Every locally compact, Hausdorff second-countable space M is paracompact. Moreover, for each open cover $\mathcal{S}$ and basis $\mathcal{B}$, there exists a countable locally finite open refinement of $\mathcal{S}$ consisting of elements of $\mathcal{B}$.

Proof. We first show that

> M admits an exhaustion by compact sets, i.e., there exists a sequence K_n of compact sets such that $M = \bigcup K_n$ and $K_n \subseteq K_{n+1}^\circ$.

Since M is locally compact and Hausdorff, we know there exists some basis of X consisting of precompact open sets, and since M is second countable, we can WOLG write this basis as (U_n). Let $K_1 \triangleq \overline{U_1}$. Now, since (U_k) cover the whole M, for each n we can let m_n be some integer greater than n and $K_n \subseteq U_1 \cup \cdots \cup U_{m_n}$. Defining $K_{n+1} \triangleq \overline{U_1} \cup \cdots \cup \overline{U_{m_n}}$, we see K_{n+1} is compact because it is a finite union of compact subspace. We also see $K_n \subseteq U_1 \cup \cdots \cup U_{m_n} \subseteq K_{n+1}^\circ$. (done)

Now, for each $n \in \mathbb{Z}_0^+$, define

$$V_n \triangleq K_{n+1} \setminus K_n^\circ \text{ and } W_n \triangleq K_{n+2}^\circ \setminus K_{n-1} \text{ where } K_0 = K_{-1} = \varnothing$$

Note that W_n are open and $V_n \subseteq W_n$. We then can associate each n and $x \in V_n$ with some $S_x^n \in \mathcal{S}$ and $B_x^n \in \mathcal{B}$ such that $x \in B_x^n \subseteq S_x^n \cap W_n$. Now, because V_n is compact (closed in K_{n+1}), we know there exists a finite subcollection $\{B^n_{x_1},\dots,B^n_{x_{n_k}}\}$ covering V_n and contained by W_n. Define

$$\mathcal{S}' \triangleq \bigcup_{n\in\mathbb{Z}_0^+}\{B^n_{x_1},\dots,B^n_{x_{n_k}}\}$$

We then see that $\mathcal{S}'$ is a countable refinement of $\mathcal{S}$ (The fact $\mathcal{S}'$ is a cover follows from V_n covering the whole M). To see that $\mathcal{S}'$ is locally finite, observe that each $B^n_{x_j}$ is contained by W_n and if $|p-q|>2$, then $W_p\cap W_q=\varnothing$. ∎

An atlas A is said to be a **smooth atlas**, if for each two charts $(\Phi_i,U_i),(\Phi_j,U_j)$ in A,

The function $\Phi_i\circ\Phi_j^{-1}:\Phi_j(U_i\cap U_j)\to\Phi_i(U_i\cap U_j)$ is a smooth diffeomorphism

It is easily checked that for each two chart Φ_i,Φ_j, the **transition map** $\Phi_i\circ\Phi_j^{-1}$ is a homeomorphism. Now, if the union of two smooth atlas A_1,A_2 is again smooth, we say A_1,A_2 are **compatible**. With some effort, one can check that compatibility is an equivalence relation on the collection of all possible atlas on X. Thus, it make sense for us to define the **smooth structure**, i.e., a maximal smooth atlas on X. Now, by a **smooth manifold**, we merely mean a manifold equipped with a maximal smooth atlas, and given a function $F:M\to N$ that maps a smooth manifold M into another smooth manifold N, we say F is **smooth at** p if F is continuous at p and there exists some charts $(U,\varphi),(V,\psi)$ respectively at $p,F(p)$ such that

(i) $U\subseteq F^{-1}(V)$.

(ii) $\psi\circ F\circ\varphi^{-1}:\varphi(U)\subseteq\mathbb{R}^m\to\psi(V)\subseteq\mathbb{R}^n$ is smooth at $\varphi(p)$.

Note that F is required to be continuous at p in the first place to guarantee that for all (V,ψ) at $F(p)$ there exists some (U,φ) at p such that $F(U)\subseteq V$. Immediately, one can check that if F is smooth at p, then

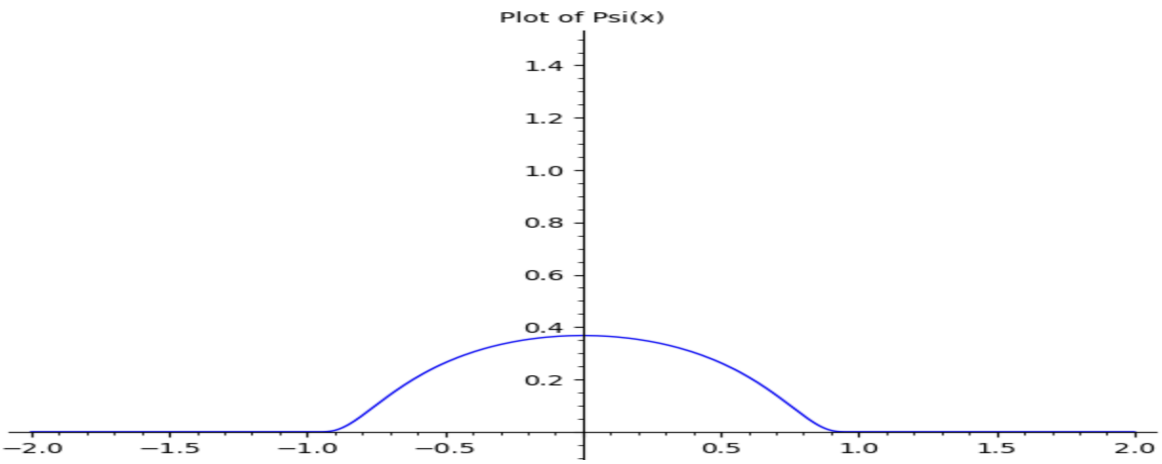
(a) For all charts $(U,\varphi),(V,\psi)$ at $p,F(p)$, the function $\psi\circ F\circ\varphi^{-1}:\varphi(F^{-1}(V)\cap U)\subseteq\mathbb{R}^m\to\psi(V)\subseteq\mathbb{R}^n$ is smooth at p.

(b) If $G:N\to R$ is another map smooth at $F(p)$, then $G\circ F:M\to R$ is smooth at p.

When F is a $\mathbb{R}^n$-valued function on M, we define the smoothness of F by considering $\mathbb{R}^n$ as a manifold with the standard atlas $\{(\mathbb{R}^n,\mathbf{id})\}$. With these definitions specified, we are almost ready to prove that smooth manifolds always admits smooth partition of unity, but before we actually give a proof, we first need to know how to manipulate smooth function between Euclidean Spaces. The simplest smooth function $\Psi : \mathbb{R} \to \mathbb{R}$ is perhaps

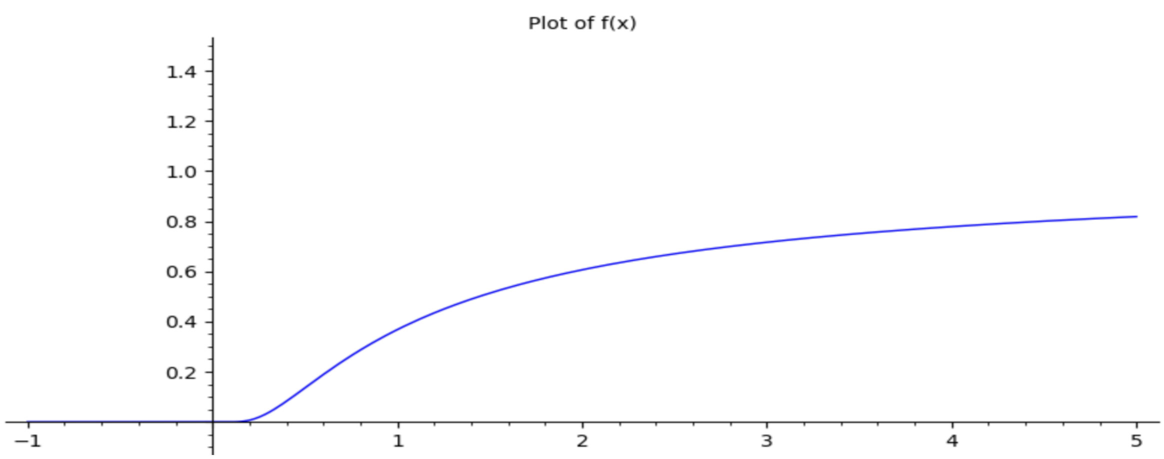
$$\Psi(x) \triangleq \begin{cases} e^{\frac{-1}{1-x^2}} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1 \end{cases}$$

Although the Ψ we just defined is smooth, it is not particularly useful in construction of partition of unity.



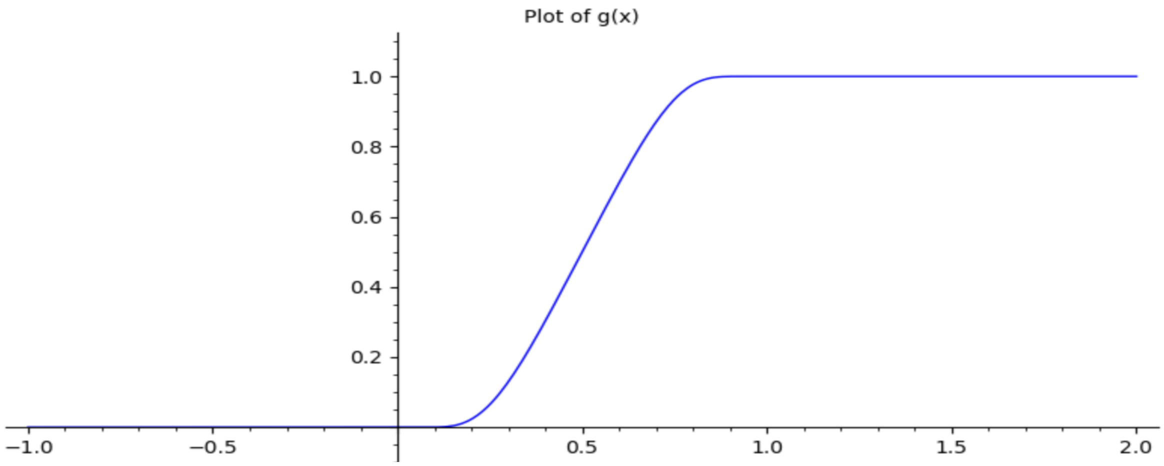
Consider the smooth function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) \triangleq \begin{cases} e^{\frac{-1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$



and $g : \mathbb{R} \to \mathbb{R}$ defined by

$$g(x) \triangleq \frac{f(x)}{f(x) + f(1-x)}$$

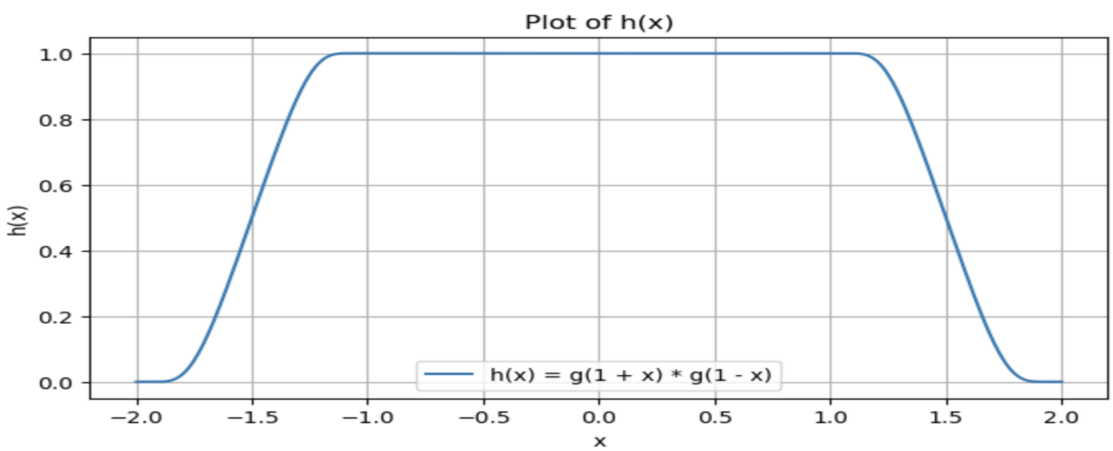


which is a smooth increasing function such that

$$g(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x \geq 1 \end{cases} \quad \text{and } g(x) \in (0,1) \text{ if } x \in (0,1)$$

Lastly, if we define $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) \triangleq g(2+x)g(2-x)$$



We see that h is a smooth function and

$$h(x) = \begin{cases} 1 & \text{if } |x| \leq 1 \\ 0 & \text{if } |x| \geq 2 \end{cases} \quad \text{and } h(x) \in (0,1) \text{ if } 1 \leq |x| \leq 2$$

A n-dimensional generalization of this smooth function h is $H : \mathbb{R}^n \to \mathbb{R}$

$$H(\mathbf{x}) \triangleq h\Big(\frac{|\mathbf{x}|}{r}\Big)$$

which is a non-negative smooth function such that

$$H(\mathbf{x}) > 0 \text{ if and only if } |x| < 2r$$

Theorem 11.1.3. (Smooth manifolds always admit smooth partition of unity) Given a smooth manifold M and some open cover (U_α) of M, there exists some smooth partition of unity (ψ_α) subordinate to (U_α).

Proof. For each U_α and each chart (V, φ) contained by U_α, if we let $\mathcal{B}_{\alpha,V}$ be the collection of the pre-images of the open balls $B_r(x)$ in $\varphi(V)$ that satisfy $\exists r' > r : B_{r'}(x) \subseteq \varphi(V)$, we know that $\mathcal{B}_{\alpha,V}$ form a basis of V. It follows that there exists a basis $\mathcal{B}_\alpha$ of U_α of the form $\mathcal{B}_\alpha \triangleq \bigcup_{(V,\varphi)\subseteq U_\alpha} \mathcal{B}_{\alpha,V}$ and a basis $\mathcal{B}$ of M of the form $\mathcal{B} \triangleq \bigcup_\alpha \mathcal{B}_\alpha$. By Theorem 11.1.2, we know there exists a countable locally finite open refinement (B_n) of U_α consisting of element of $\mathcal{B}$.

Now, for each n, there exists corresponding chart (V, φ) and r, r' such that

$$\varphi(B_n) = B_r(x) \subseteq B_{r'}(x) \subseteq \varphi(V)$$

This with some tedious effort guarantee that

$$\overline{B_n} = \varphi^{-1}(\overline{B_r(x)})$$

We then can well-define a function $f_n : M \to \mathbb{R}$ by

$$f_n \triangleq \begin{cases} H_n \circ \varphi & \text{on } V \\ 0 & \text{on } M \setminus \overline{B_n} \end{cases}$$

where $H_n : \mathbb{R}^n \to \mathbb{R}$ is a smooth function that is positive on $B_r(x)$ and zero else where. (One should at this point check that f_n is well-defined)

Since V and $M \setminus \overline{B_n}$ are both open and f_n are both smooth on them, we now see f_n is smooth on the whole M. Now, since (B_n) is a locally finite cover of M and f_n can only be positive on B_n, we see that the function $f : M \to \mathbb{R}$

$$f(x) \triangleq \sum_{n\in\mathbb{N}} f_n(x) \text{ is well defined}$$

and that f is positive on whole M because each f_n is positive on B_n and (B_n) cover the whole M. Note that f is also smooth on M since for each x, there exists an open-neighborhood around x on which f is just a finite sum of f_n.

We now can define for each n a smooth function $g_n : M \to \mathbb{R}$ by

$$g_n(x) \triangleq \frac{f_n(x)}{f(x)}$$

Lastly, because (B_n) is an open refinement of U_α, we can associate each n with an $\alpha(n)$ satisfying $B_n \subseteq U_{\alpha(n)}$. Then, because g_n are non-negative and $\sum_{n\in\mathbb{N}} g_n \equiv 1$, we can define for each α a function $\psi_\alpha : M \to \mathbb{R}$ by

$$\psi_\alpha \triangleq \sum_{n:\alpha(n)=\alpha} g_n$$

Because (B_n) is locally finite and g_n is only positive on B_n, we see that ψ_α is indeed smooth. Some tedious effort shows that ψ_α is indeed a partition of unity.

■

Corollary 11.1.4. (Existence of Smooth Bump Function) Given some U open in M and some K closed in M and contained by U, there exists a smooth function $\psi : M \to [0,1]$ such that

(a) $\psi \equiv 1$ on K.

(b) $\operatorname{supp} \psi \subseteq U$

We first develop some important tools we shall heavily used later. Let K be closed in M. We say $f : K \to \mathbb{R}$ is smooth if there exists some open $U \subseteq M$ containing K and smooth $\tilde{f} : U \to \mathbb{R}$ such that $\tilde{f} \equiv f$ on K.

Corollary 11.1.5. (Extension of Smooth Function) Suppose K is closed in M and $f : K \to \mathbb{R}$ is smooth. Let U be open in M and containing K, and let $g \in C^\infty(U)$ satisfy $g|_K = f$. There exists a smooth function $\tilde{f} : M \to \mathbb{R}$ such that $\tilde{f}|_K = f$ and $\operatorname{supp} \tilde{f} \subseteq U$.

Proof. Corollary 11.1.4 give us a smooth function $\psi \in C^\infty(M)$ such that $\psi \equiv 1$ on K and $\operatorname{supp} \psi \subseteq U$. We simply define

$$\tilde{f}(p) \triangleq \begin{cases} \psi(p)g(p) & \text{if } p \in U \\ 0 & \text{if } p \notin U \end{cases}$$

It is straightforward to check $\tilde{f}|_K = f$. To see $\operatorname{supp} \tilde{f} \subseteq U$, check $\operatorname{supp} \tilde{f} \subseteq \operatorname{supp} \psi$. It is clear that $\tilde{f}$ is smooth at all $p \in U$. It remains to show $\tilde{f}$ is smooth at all $p \notin U$. Fix $p \notin U$. Because $p \notin \operatorname{supp} \psi$, we know that $\tilde{f} \equiv 0$ on some open neighborhood around p. ■

11.2 Tangent Space

Abstract

In this section and from now on, $C^\infty(M)$ denote the space of all smooth real-valued function defined on M.

Given a point p in M, we define the **tangent space** T_pM **of** M **at** p to be the space of linear functional $D : C^\infty(M) \to \mathbb{R}$ that satisfy the product rule at p

$$D(fg) = D(f)g(p) + f(p)D(g) \text{ for all } f, g \in C^\infty(M)$$

It is clear that tangent space form a vector space when endowed with pointwise scalar multiplication and addition. Note that

$$D(1) = D(1^2) = D(1) + D(1)$$

This implies that for all $D \in T_pM$, if $f \in C^\infty(M)$ is constant, then $D(f) = 0$.

Theorem 11.2.1. (Dimension of $T_{\mathbf{p}}\mathbb{R}^n$ is n) For all $\mathbf{p} \in \mathbb{R}^n$, the vector space $T_{\mathbf{p}}\mathbb{R}^n$ is n-dimensional.

Proof. Define a function $\varphi : \mathbb{R}^n \to T_{\mathbf{p}}(\mathbb{R}^n)$ by

$$\mathbf{x} \mapsto D_{\mathbf{x}} \text{ where } D_{\mathbf{x}}(f) = \mathbf{x} \cdot \nabla f(\mathbf{p})$$

It is straightforward to check that φ is well defined and a linear transformation. It remains to prove φ is bijective. Let $D_{\mathbf{y}} = 0$, to show φ is one-to-one, we are required to show $\mathbf{y} = 0$. For each $j \in \{1, \dots, n\}$, define a smooth function $g_j : \mathbb{R}^n \to \mathbb{R}$ by

$$g_j(\mathbf{x}) \triangleq \mathbf{x}^j \tag{11.1}$$

We then see that for all $\mathbf{y} \in \mathbb{R}^n$

$$D_{\mathbf{y}}(g_j) = \mathbf{y} \cdot \mathbf{e}_j = \mathbf{y}^j$$

Then if $D_{\mathbf{y}} = 0$, we must have $\mathbf{y} = 0$. (done)

We now show φ is onto. Fix $w \in T_p(\mathbb{R}^n)$ and define $\mathbf{v} \in \mathbb{R}^n$ by

$$\mathbf{v}^j \triangleq w(g_j)$$

where g_j is defined in Equation 11.1. We claim $w = D_{\mathbf{v}}$. Fix $f \in C^\infty(\mathbb{R}^n)$. By Multi Variables Taylor Theorem, we know

$$f(\mathbf{x}) = f(\mathbf{p}) + \sum_{i=1}^n \frac{\partial f}{\partial \mathbf{x}^i}(\mathbf{p})(\mathbf{x}^i - \mathbf{p}^i)$$
$$+ \sum_{i,j=1}^n (\mathbf{x}^i - \mathbf{p}^i)(\mathbf{x}^j - \mathbf{p}^j) \int_0^1 (1-t)\frac{\partial f}{\partial \mathbf{x}^i \partial \mathbf{x}^j}(\mathbf{p} + t(\mathbf{x}-\mathbf{p}))dt \tag{11.2}$$

Now, putting both side into w, by product rule we can cancel term 11.2, so

$$w(f) = \sum_{i=1}^n \frac{\partial f}{\partial \mathbf{x}^i}(\mathbf{p})w(g_i) = \sum_{i=1}^n \frac{\partial f}{\partial \mathbf{x}^i}(\mathbf{p})\mathbf{v}^i = D_{\mathbf{v}}(f) \text{ (done)}$$

■

As pointed out by the proof of Theorem 11.2.1, for each $D_{\mathbf{x}} \in T_{\mathbf{p}}\mathbb{R}^n$, the value of $D_{\mathbf{x}}(f)$ is determined locally at $\mathbf{p}$ by f. The same behavior happens in the general case.

Theorem 11.2.2. (Tangent Vector care only about local behavior) Suppose $v \in T_pM$. If $f, g \in C^\infty(M)$ agree on some open neighborhood around p, then $v(f) = v(g)$.

Proof. Let $h \triangleq f - g \in C^\infty(M)$. We know $h \equiv 0$ on some open neighborhood around p. This implies

$$\text{supp}(h) \subseteq M \setminus \{p\} \subseteq M$$

By Lemma 11.1.4, there exists some bump function $\psi : M \to [0,1]$ such that $\psi \equiv 1$ on $\text{supp}(h)$ and $\text{supp}(\psi) \subseteq M \setminus \{p\}$. Since $\text{supp}(\psi) \subseteq M \setminus \{p\}$, we know $\psi(p) = h(p) = 0$, and since $\psi \equiv 1$ on $\text{supp}(h)$, we also know $\psi h = h$. This let us deduce

$$v(h) = v(\psi h) = v(\psi)h(p) + \psi(p)v(h) = 0$$

■

We now define the derivative $F_{*,p} : T_pM \to T_{F(p)}N$ of a smooth map $F : M \to N$ at $p \in M$ between smooth manifold.

$$(F_{*,p}(w))(f) \triangleq w(f \circ F)$$

It is straightforward to check that given another smooth manifold R and smooth map $G : N \to R$ we have

(a) $F_{*,p}(w) \in T_{F(p)}N$.

(b) $F_{*,p}$ is linear.

(c) $(G \circ F)_{*,p} = G_{*,F(p)} \circ F_{*,p} : T_pM \to T_{G\circ F(p)}R$.

(d) If $\mathbf{id} : M \to M$ is the identity function on M, then $\mathbf{id}_{*,p}$ is the identity map on T_pM.

(e) If $F : M \to N$ is a diffeomorphism, then $F_{*,p} : T_pM \to T_{F(p)}(N)$ is an isomorphism of vector space and $(F_{*,p})^{-1} = (F^{-1})_{*,F(p)}$

Theorem 11.2.3. (Tangent Space of points in open restriction) Given some open subset $U \subseteq M$ and inclusion map $\iota : U \to M$, for all $p \in U$, the map $\iota_{*,p} : T_pU \to T_pM$ is a vector space isomorphism.

Proof. We first prove that $\iota_{*,p}$ is one-to-one. Fix $v \in T_pU$ such that $\iota_{*,p}(v) = 0$ and $f \in C^\infty(U)$. We are required to show $v(f) = 0$. Let (V, φ) be a chart contained in U and containing p, and let ϵ satisfy

$$\mathrm{cl}_{\mathbb{R}^n}(B_\epsilon(\varphi(p))) \subseteq \varphi(V)$$

Let $A \triangleq \varphi^{-1}(B_\epsilon(\varphi(p)))$. Since $\mathrm{cl}_{\mathbb{R}^n}(B_\epsilon(\varphi(p)))$ is compact, and compact subspace is closed in Hausdorff space, we now see

$$\mathrm{cl}_M(A) \subseteq \varphi^{-1}(\mathrm{cl}_{\mathbb{R}^n}(B_\epsilon(\varphi(p)))) \subseteq U$$

By Corollary 11.1.5, we know there exists $\tilde{f} \in C^\infty(M)$ such that $\tilde{f} \equiv f$ on $\mathrm{cl}_M(A)$. We can now deduce

$$v(f) = v(\tilde{f}|_U) = v(\tilde{f} \circ \iota) = \iota_{*,p}(v)(\tilde{f}) = 0 \text{ (done)}$$

We now prove $\iota_{*,p}$ is onto. Fix $w \in T_pM$, and again the same $A \triangleq \varphi^{-1}(B_\epsilon(\varphi(p)))$. Define $v \in T_pU$ by

$$v(f) \triangleq w(\tilde{f})$$

where $\tilde{f} \in C^\infty(M)$ satisfy $\tilde{f} \equiv f$ on $\mathrm{cl}_M(B)$. Because for all $g \in C^\infty(M)$, $g \equiv g \circ \iota$ on $\mathrm{cl}_M(B)$, we know

$$\iota_{*,p}(v)(g) = v(g \circ \iota) = w(g) \text{ (done)}$$

∎

Theorem 11.2.4. (Dimension of T_pM is the same as M) If M is an n-dimensional smooth manifold, then T_pM is an n-dimensional vector space.

Proof. Let (U, φ) be a chart containing p. Theorem 11.2.3 tell us that T_pU and T_pM are isomorphic, and since $\varphi : U \to \varphi(U)$ is a diffeomorphism between smooth manifolds $U, \varphi(U)$, we know T_pU is isomorphic to $T_{\varphi(p)}\varphi(U)$ which is isomorphic to $T_p\mathbb{R}^n$ again by Theorem 11.2.3. ∎

With all the tools we have gathered, we can now see that if we are given a chart (U, φ) containing p, and we define for each $i \in \{1, \ldots, n\}$ the functional $\frac{\partial}{\partial \mathbf{x}^i}|_p : C^\infty(M) \to \mathbb{R}$ by

$$\frac{\partial}{\partial \mathbf{x}^i}\Big|_p(f) \triangleq \frac{\partial(f \circ \varphi^{-1})}{\partial \mathbf{x}^i}(\varphi(p))$$

then since

$$\varphi_{*,p}\Big(\frac{\partial}{\partial \mathbf{x}^i}\Big|_p\Big)(g) = \frac{\partial}{\partial \mathbf{x}^i}\Big|_p(g \circ \varphi) = \frac{\partial g}{\partial \mathbf{x}^i}(\varphi(p))$$

We see

$$\Big\{\varphi_{*,p}\Big(\frac{\partial}{\partial \mathbf{x}^i}\Big|_p\Big) : 1 \leq i \leq n\Big\} \text{ form a basis for } T_{\varphi(p)}\varphi(U)$$

and thus

$$\Big\{\frac{\partial}{\partial \mathbf{x}^i}\Big|_p : 1 \leq i \leq n\Big\} \text{ form a basis for } T_pM$$

Example 33 (Loring Tu Example)

$$F : \mathbb{R}^2 \to \mathbb{R}^3; F(x, y) \triangleq (x, y, xy) = (u, v, w)$$

We have

$$F_{*,(x,y)}\big(\frac{\partial}{\partial x}\big) = \frac{\partial}{\partial u} + y\frac{\partial}{\partial w}$$

11.3 Lie Group

By a **Lie group**, we mean a smooth manifold equipped with a group structure such that the inversion and group addition are both smooth map, or equivalently, that

$$M^2 \to M; (g, h) \mapsto gh^{-1}$$

is smooth.

Chapter 12

Beauty

12.1 Fundamental Theorem of Algebra

Theorem 12.1.1. (Fundamental Theorem of Algebra)

12.2 Euler's Formula

Abstract

This section give a precise definition to exponential function and trigonometric functions, prove Euler's Formula and prove some of their basic properties seen in the real case. In this section, v, v are complex numbers, and x, y are real numbers.

Suppose that we define

$$\exp(z) \triangleq \sum_{n=0}^{\infty} \frac{z^n}{n!}$$
$$\sin(z) \triangleq \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1}$$
$$\cos(z) \triangleq \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n}$$

Some properties we are familiar with is now easily seen using basic techniques on sequence and series.

(a) Because $\limsup_{n\to\infty} |c_n|^{\frac{1}{n}} \leq \limsup_{n\to\infty} \left|\frac{c_{n+1}}{c_n}\right|$, by Cauchy-Hadamard Theorem, $\exp(z)$, $\sin(z)$ and $\cos(z)$ are defined on the whole complex plane.

(b) $\exp(z) \cdot \exp(v) = \exp(z+v)$ by Merten's Theorem of Cauchy product.

(c) $\frac{d}{dz}\exp(z) = \exp(z)$, $\frac{d}{dz}\sin(z) = \cos(z)$ and $\frac{d}{dz}\cos(z) = -\sin(z)$ on the whole complex plane, since we can differentiate them term by term.

(d) For all $x \in \mathbb{R}$, $\exp(x), \sin(x)$ and $\cos(x)$ lie in $\mathbb{R}$.

(e) $\exp(x)$ strictly increase on $\mathbb{R}$ and $\exp(0) = 1$.

(f) $\exp(x) \nearrow \infty$ as $x \to \infty$.

(g) Because $\exp(-x) = \frac{\exp(0)}{\exp(x)} = \frac{1}{\exp(x)}$, by proposition (f), $\exp(x) \searrow 0$ as $x \to -\infty$.

(h) For all $x \in \mathbb{R}$, $\exp(x) \in \mathbb{R}^+$ by proposition (g).

(i) $\exp(x)$ is convex on $\mathbb{R}$ $\quad (\because (e^x)'' = e^x > 0)$

In particular, we have **Euler's Formula**.

Theorem 12.2.1. (Euler's Formula) For all $z \in \mathbb{C}$, we have

$$\exp(iz) = \cos(z) + i\sin(z)$$

Proof. Define

$$I(n) \triangleq \begin{cases} 1 & \text{if } n \equiv 0 \pmod 4 \\ i & \text{if } n \equiv 1 \pmod 4 \\ -1 & \text{if } n \equiv 2 \pmod 4 \\ -i & \text{if } n \equiv 3 \pmod 4 \end{cases}$$

Compute

$$\begin{aligned} \exp(iz) &= \sum_{n=0}^{\infty} \frac{(iz)^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{I(n) z^n}{n!} \qquad (12.1) \\ &= \sum_{n=0}^{\infty} \frac{I(2n)}{(2n)!} z^{2n} + \frac{I(2n+1)}{(2n+1)!} z^{2n+1} \qquad \left(\because \text{ this is a sub-sequence of } (12.1)\right) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} + i \cdot \frac{(-1)^n}{(2n+1)!} z^{2n+1} \end{aligned}$$

Now, we can conclude

$$\begin{aligned} \cos(z) + i\sin(z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} + i \cdot \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} + \sum_{n=0}^{\infty} i \cdot \frac{(-1)^n}{(2n+1)!} z^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} + i \cdot \frac{(-1)^n}{(2n+1)!} z^{2n+1} = \exp(iz) \end{aligned}$$

∎

12.3 Equivalent Definitions of Exponential Functions

Abstract

This section give several equivalent definitions of the exponential function.

Theorem 12.3.1. (First Characterization) For all $z \in \mathbb{C}$, we have

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n = e^z$$

Proof. Fix z and ϵ. Let N satisfy

$$\sum_{n=N+1}^{\infty}\frac{|z|^n}{n!} < \frac{\epsilon}{3}$$

It is clear that

$$\lim_{n\to\infty}\sum_{k=0}^{N}\binom{n}{k}\left(\frac{z}{n}\right)^k = \sum_{k=0}^{N}\frac{z^k}{k!}$$

We then can let N_1 satisfy

$$\left|\sum_{k=0}^{N}\binom{n}{k}\left(\frac{z}{n}\right)^k - \sum_{k=0}^{N}\frac{z^k}{k!}\right| < \frac{\epsilon}{3} \text{ for all } n > N_1$$

Now, because for all k smaller than n we have

$$\binom{n}{k}\frac{1}{n^k} = \frac{1}{k!}\prod_{j=1}^{k-1}(1-\frac{j}{n}) \le \frac{1}{k!}$$

we have for all $n > N_1$

$$\begin{aligned}
\left|\left(1+\frac{z}{n}\right)^n - \sum_{k=0}^{\infty}\frac{z^k}{k!}\right| &\le \left|\sum_{k=0}^{n}\binom{n}{k}\left(\frac{z}{n}\right)^k - \sum_{k=0}^{\infty}\frac{z^k}{k!}\right| \\
&\le \left|\sum_{k=0}^{N}\binom{n}{k}\left(\frac{z}{n}\right)^k - \sum_{k=0}^{N}\frac{z^k}{k!}\right| + \left|\sum_{k=N+1}^{n}\binom{n}{k}\left(\frac{z}{n}\right)^k\right| + \left|\sum_{k=N+1}^{\infty}\frac{z^k}{k!}\right| \\
&\le \left|\sum_{k=0}^{N}\binom{n}{k}\left(\frac{z}{n}\right)^k - \sum_{k=0}^{N}\frac{z^k}{k!}\right| + \sum_{k=N+1}^{n}\binom{n}{k}\frac{|z|^k}{n^k} + \sum_{k=N+1}^{\infty}\frac{|z|^k}{k!} < \epsilon
\end{aligned}$$

∎

$$\ln(x) \triangleq \int_1^x \frac{1}{t} dt$$

By FTC (Theorem 6.3.1), it is easy to see that

$$\frac{d}{dx} \ln(x) = \frac{1}{x} \qquad (x \in \mathbb{R}^+)$$

To see

$$\ln(xy) = \ln(x) + \ln(y)$$

Fix $y \in \mathbb{R}^+$ and set

$$f(x) \triangleq \ln(x) \text{ and } g(x) \triangleq \ln(xy)$$

Conclude $f'(x) = g'(x)$, and use FTC (Theorem 6.3.2) to conclude $f - g$ is some fixed constant k. Now, see that

$$g(1) = f(1) + k \implies k = \ln(y)$$

Then, we have

$$\ln(xy) = g(x) = f(x) + k = \ln(x) + \ln(y)$$

Using induction, it is now easy to see

$$\ln(x^n) = n \ln(x) \qquad (n \in \mathbb{Z}_0^+)$$

Theorem 12.3.2. (Second Characterization)

12.4 Equivalent Definitions of Trigonometric Functions

12.5 Equivalent Definitions of Gamma and Beta Functions

12.6 Prime Number Theorem

Chapter 13

and the Beast

13.1 Topologist's Sine Curve

13.2 Long Line

13.3 Bug-Eyed Line

Abstract

This section introduce Bug-Eyed Line, which is a second-countable locally Euclidean space that is not Hausdorff.

Let a, b be distinct real numbers and let $\mathbb{R}'$ be the quotient space of $\mathbb{R} \times \{a, b\}$ where

$$(x, a) \sim (x, b) \text{ for all } x \neq 0$$

The space $\mathbb{R}'$ is commonly referred as the **Bug-Eyed Line**. Now by definition of quotient topology, given a subset E of $\mathbb{R}'$, E is open if and only if

$$\{x \in \mathbb{R} : [(x, a)] \in E\} \text{ and } \{x \in \mathbb{R} : [(x, b)] \in E\} \text{ are both open in } \mathbb{R}$$

With this in mind, it is clear that $\mathbb{R}'$ is not Hausdorff, since $[(0, a)]$ and $[(0, b)]$ can not be distinguished, and it is straightforward to check $\mathbb{R}'$ is locally Euclidean.

Now, suppose we write some subset A of $\mathbb{R} \times \{a, b\}$ in the form

$$\{(x, a) : x \in A_1\} \cup \{(x, b) : x \in A_2\}$$

We can deduce

$$\pi^{-1}(\pi(A)) = \Big\{(x, a) : x \in A_1 \cup (A_2 \setminus \{0\})\Big\} \cup \Big\{(x, b) : x \in A_2 \cup (A_1 \setminus \{0\})\Big\}$$

It is then clear that π is an open mapping, thus $\mathbb{R}'$ is second countable.

13.4 Weierstrass Function

13.5 Fabius Function

13.6 Vitali Set

Abstract

This section construct the Vitali Set in $\mathbb{R}^d$ for reference.

Fix $p > m > 0 \in \mathbb{R}$. Define an equivalence relation in $[0,m]^d$ by

$$\mathbf{x} \sim \mathbf{y} \overset{\triangle}{\iff} \mathbf{x} - \mathbf{y} \in \mathbb{Q}^d$$

Using axiom of choice, we can let $V \subseteq [0,m]^d$ contain exactly one element in each equivalence class. Enumerate $([0,p] \cap \mathbb{Q})^d$ by $\mathbf{z}_n$. For each $n \in \mathbb{N}$, define

$$V + \mathbf{z}_n \triangleq \{\mathbf{x} + \mathbf{z}_n \in \mathbb{R}^d : \mathbf{x} \in V\}$$

It is straightforward to check from definition of V that

(a) $V + \mathbf{z}_n$ are pairwise disjoint.

(b) $[0,m]^d \subseteq \bigsqcup_n (V + \mathbf{z}_n) \subseteq [0, m+p]^d$.

Now, if V is really measurable, we see that

$$0 < m^d \leq \sum_n |V| \leq (m+p)^d \tag{13.1}$$

since

$$\left| \bigsqcup_n V + \mathbf{z}_n \right| = \sum_n |V + \mathbf{z}_n| = \sum_n |V|$$

Observing that Equation 13.1 is impossible implies V is not measurable.

13.7 Cantor Set

Abstract

This section construct the classical ternary Cantor set and some of its variant, and prove they are uncountable and perfect.

By the term **Classical Ternary Cantor Set** $\mathcal{C}$, one usually mean

$$\mathcal{C} \triangleq \bigcap_{n\in\mathbb{N}} \mathcal{C}_n$$

where $\mathcal{C}_0 \triangleq [0,1]$, and $\mathcal{C}_{k+1}$ is the result of deleting the open middle third of each connected component – which are clearly compact – of $\mathcal{C}_k$. Immediately, one can see that

(a) $\mathcal{C}_n$ has 2^n amount of connected components, which are all compact with the length of $\frac{1}{3^n}$.

(b) $\mathcal{C}$ has zero measure, since $|\mathcal{C}_{n+1}| = |\mathcal{C}_n| - \frac{2^n}{3^{n+1}}$.

(c) $\mathcal{C}$ is totally disconnected, since if $a, b \in \mathcal{C}$ are connected then $[a,b] \subseteq \mathcal{C}$, which is impossible by proposition (a).

(d) $\mathcal{C}$ is perfect, since for each ϵ and $x \in \mathcal{C}$, there exists large enough n such that the length of each connected component of $\mathcal{C}_n$ is smaller than ϵ, and the end point of the connected component – in which x lies – that isn't x would belong to $\mathcal{C}$ and be ϵ-close to x.

(e) The endpoints $\{x \in \mathcal{C} : x \text{ is the endpoint of some connected component of some } \mathcal{C}_n\}$ are countable and dense in $\mathcal{C}$.

With some tedious effort, one can see that $x \in \mathcal{C}$ if and only if

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \text{ for some } a_n \in \{0, 2\}$$

With base 3 representation, one can use a diagonal argument to show $\mathcal{C}$ is uncountable. Another approach to show $\mathcal{C}$ is uncountable is to show non-singleton perfect sets in $\mathbb{R}^n$ are uncountable.

Theorem 13.7.1. (Non-Singleton Perfect Set in $\mathbb{R}^n$ is Uncountable) Given a perfect set $E \subseteq \mathbb{R}^n$, if E contain more than one element, then E must be uncountable.

Proof. ■

Notably, a variant of the Cantor set includes the **Fat Cantor Set**, which is constructed similarly to the classical ternary Cantor set, except that the removed open middle intervals at $n$stage is each of length δ^n, where $0 < \delta < 3$. Note that the construction cannot be done if $\delta > 3$ and that the Fat Cantor Sets all have positive measure, are perfect and totally disconnected.

13.8 Cantor-Lebesgue Function

Abstract

This section construct the Cantor-Lebesgue Function for reference.

Consider the Classical Ternary Cantor Set $\mathcal{C}$. Let

$$\mathcal{D}_n \triangleq [0,1] \setminus \mathcal{C}_n \text{ for all } n$$

For example,

$$\mathcal{D}_1 = (\frac{1}{3}, \frac{2}{3}) \text{ and } \mathcal{D}_2 = (\frac{1}{9}, \frac{2}{9}) \cup (\frac{1}{3}, \frac{1}{2}) \cup (\frac{7}{9}, \frac{8}{9})$$

Because $\mathcal{C}_n$ has 2^n amount of connected components, we know $\mathcal{D}_n$ has $2^n - 1$ amount of connected components. Order these connected components by $\{I_j^n : 1 \leq j \leq 2^n - 1\}$. We now define a sequence of function $f_n : [0,1] \to [0,1]$ by letting

$$f_n(x) \triangleq \begin{cases} 0 & \text{if } x = 0 \\ \frac{j}{2^n} & \text{if } x \in \overline{I_j^n} \\ 1 & \text{if } x = 1 \end{cases}$$

where f_n is linear on $\overline{\mathcal{D}_n}$.

For each $x \in [0,1]$ there exists a (not always unique) base 3 representation

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \text{ for some } a_n \in \{0,1,2\}$$

The Cantor-Lebesgue Function $f : [0,1] \to [0,1]$ is defined as follows. Given $x = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \in [0,1]$, if there exists some smallest n such that $a_n = 1$, define

$$b_k \triangleq \begin{cases} 0 & \text{if } k < n \text{ and } a_k = 0 \\ 1 & \text{if } k < n \text{ and } a_k = 2 \\ 1 & \text{if } k = n \\ 0 & \text{if } k > n \end{cases}$$

It $1 \notin \{a_n : n \in \mathbb{N}\}$, then we simply define

$$b_n \triangleq \begin{cases} 0 & \text{if } a_n = 0 \\ 1 & \text{if } a_n = 2 \end{cases}$$

We then define

$$f(x) \triangleq \sum_{n=1}^{\infty} \frac{b_n}{2^n}$$

Note that if the base 3 representation of $x \in (0,1)$ has the trailing 0 or 2, then the representation must not be unique, yet the procedure described above does give the same value $f(x)$ while the base 2 representation can be different. Some tedious effort can now be applied to show that

(a) $f(\mathcal{C}) = [0,1]$ where $\mathcal{C}$ is the classical ternary set.

(b) $f : [0,1] \to [0,1]$ is increasing on $[0,1]$.

13.9 Volterra's Function

13.10 Peano Space-filling Curve