NCKU 112.2
Advanced Calculus 2

Eric Liu

Contents

Chapter 1

General Topology

1.1 Equivalent axiomazations

In the first lecture of Topology, we learned that a topological space (X, τ) is a space X with a topology $\tau \subseteq \mathcal{P}(X)$ such that τ satisfy the following three axioms.

$$\begin{cases} X, \varnothing \in \tau \\ \forall O, Y \in \tau, O \cap Y \in \tau \text{ (Closed under finite intersection)} \\ \forall T \subseteq \mathcal{P}(\tau), \bigcup T \in \tau \text{ (Closed under arbitrary union)} \end{cases} \tag{1.1}$$

It is only after the explicit listing of the above three of open sets, we then start defining "closed sets", "neighborhoods", "continuous functions", "compact sets" or "connected sets" based on open sets.

Although this approach, axiomatization via open sets, is mathematically sufficient, in history, there are other axiomatization proved to be equivalent to the traditional axiomaization via open sets.

In this section, we will give other three axiomatizations of topology, via neighborhood systems, via nets and via filters, and show they are equivalent with each other.

In this note, a *neighborhood system* is a function $\mathcal{N} : X \to \mathcal{P}(\mathcal{P}(X))$. We shall first show that there exists an explicit one-to-one correspondence between collection of topologies on X and collection of neighborhood systems that satisfy the following axioms.

Axiom 1.1.1. (Axioms of neighborhood systems)

$$\begin{cases} \forall x \in X, \mathcal{N}(x) \neq \varnothing \text{ (Non empty)} \\ \forall x \in X, \forall M \in \mathcal{N}(x), x \in M \text{ (Around)} \\ \forall x \in X, \forall M \subseteq X, \exists N \in \mathcal{N}(x), N \subseteq M \implies M \in \mathcal{N}(x) \text{(Closed under superset)} \\ \forall x \in X, \forall M, N \in \mathcal{N}(x), M \cap N \in \mathcal{N}(x) \text{(Closed under finite intersection)} \\ \forall x \in X, \forall M \in \mathcal{N}(x), \exists N \in \mathcal{N}(x), N \subseteq M \text{ and } \forall y \in N, M \in \mathcal{N}(y) \text{(Open neighborhood)} \end{cases}$$

Suppose C is the collection of topologies on X and D is the collection of neighborhood systems that satisfy Axioms 1.1.1. Now we wish to prove that the function $f : C \to D$

$$\tau \mapsto \mathcal{N}_\tau \text{ where } \mathcal{N}_\tau(x) = \{A \in \mathcal{P}(X) : \exists O \in \tau, x \in O \subseteq A\} \tag{1.2}$$

is bijective. In order to prove this, we first have to prove that f is well-defined, meaning for each topology τ, the function $\mathcal{N}_\tau$ is indeed a neighborhood system that satisfy the above axioms. This is easy, which we will omit here. It remains that we have to prove two statements: f is one-to-one, and f is onto.

Theorem 1.1.2. (f in Equation 1.2 is one-to-one) As titled.

Proof. Suppose τ and τ' are two different topologies on X. WOLG, suppose $O \in \tau \setminus \tau'$. By definition of f (Equation 1.2), we know

$$\forall x \in O, O \in \mathcal{N}_\tau(x)$$

This means, to prove $\mathcal{N}_\tau \neq \mathcal{N}_{\tau'}$, we only have to prove the existence of some $x \in O$ such that $O \notin \mathcal{N}_\tau(x)$. Assume $\forall x \in O, O \in \mathcal{N}_{\tau'}(x)$. By definition of f (Equation 1.2), we can deduce $\forall x \in O, \exists A_x \in \tau', x \in A_x \subseteq O$. It is easy to verify that $O = \bigcup_{x \in O} A_x$. Then by Axiom 1.1 of open sets , we $O = \bigcup_{x \in O} A_x \in \tau'$ CaC to that $O \in \tau \setminus \tau'$. ∎

To prove f is onto is a little bit complicated.

Theorem 1.1.3. (f in Equation 1.2 is onto) As titled.

Proof. Define $g : D \to C$ by

$$\mathcal{N} \mapsto \sigma_\mathcal{N} \text{ where } \sigma_\mathcal{N} = \{O \in \mathcal{P}(X) : \forall x \in O, O \in \mathcal{N}(x)\} \tag{1.3}$$

It is easy to check for each neighborhood system $\mathcal{N} : X \to \mathcal{P}(\mathcal{P}(X))$ satisfying Axioms 1.1.1, the collection $\sigma_\mathcal{N}$ is a topology.

Then, to prove f is onto, it suffice to prove

$$\forall \mathcal{N} \in D, \mathcal{N} = f(\sigma_{\mathcal{N}})$$

In other words, we wish to prove

$$\forall \mathcal{N} \in D, \forall x \in X, \mathcal{N}(x) = \big(f(\sigma_{\mathcal{N}})\big)(x)$$

By definition of f (Equation 1.2), we know

$$\big(f(\sigma_{\mathcal{N}})\big)(x) = \{A \in \mathcal{P}(X) : \exists O \in \sigma_{\mathcal{N}}, x \in O \subseteq A\} \tag{1.4}$$

Suppose $A \in \big(f(\sigma_{\mathcal{N}})\big)(x)$. Let $O \in \sigma_{\mathcal{N}}$ satisfy $x \in O \subseteq A$. By definition of g (Equation 1.3), we know $O \in \mathcal{N}(x)$. Then because Axiom 1.1.1 says that $\mathcal{N}(x)$ is closed under superset, we now see $A \in \mathcal{N}(x)$. We have proved $\big(f(\sigma_{\mathcal{N}})\big)(x) \subseteq \mathcal{N}(x)$.

Suppose $M \in \mathcal{N}(x)$. By the fifth axiom listed in Axiom 1.1.1, we know there exists some $N \in \mathcal{N}(x)$ such that $N \subseteq M$ and $\forall y \in N, M \in \mathcal{N}(y)$. Notice that the union of such N is still a subset of M and M is a neighborhood of each point in the union. This means that there exists a maximum N that satisfy

$$\begin{cases} N \subseteq M \\ \forall y \in N, M \in \mathcal{N}(y) \end{cases}$$

We now prove $N \in \sigma_{\mathcal{N}}$. Arbitrarily pick $z \in N$, by definition of N above, we can deduce $M \in \mathcal{N}(z)$. Then by the fifth axiom listed in Axiom 1.1.1, we know there exists some $N' \in \mathcal{N}(z)$ satisfying $N' \subseteq M$ and $\forall y \in N', M \in \mathcal{N}(y)$. Because N by definition is the maximum of such neighborhood, we see $N' \subseteq N$. Then because Axiom 1.1.1 says that $\mathcal{N}(z)$ is closed under superset, we see $N \in \mathcal{N}(z)$.

We have shown $\forall z \in N, N \in \mathcal{N}(z)$. Then by definition of g (Equation 1.3), we see $N \in \sigma_{\mathcal{N}}$ (done) .

Then because $N \subseteq M$ by definition of N and because $N \in \sigma_{\mathcal{N}}$, we see $M \in \big(f(\sigma_{\mathcal{N}})\big)(x)$, according to Equation 1.4. We have proved $\mathcal{N}(x) \subseteq \big(f(\sigma_{\mathcal{N}})\big)(x)$ (done) ■

Before embarking on the axiomatization via nets, we first have to settle the terminologies. Recall that a set D is **directed** if there exists an ordering $\leq$ on D such that $\leq$ is reflexive, transitive, and we have $\forall i, j \in D, \exists k \in D, i \leq k$ and $j \leq k$. By a **net**, we mean a function w whose domain is directed. We say a subset D' of a directed set is **cofinal** if $\forall d \in D, \exists d' \in D', d \leq d'$. By a **subnet** of $w : D \to X$, we mean a net

$v : E \to X$ such that there exists $h : E \to D$ such that

$$\begin{cases} \forall e, e' \in E, e \leq e' \implies h(e) \leq h(e')(\text{Monotone}) \\ h[E] \text{ is cofinal} \\ v = w \circ h \end{cases}$$

One can check that when w is a sequence x_n and v is the sub-sequence x_{n_k}, the corresponding h is just n_k.

By a **tail** T_d of a directed set D, we mean $T_d = \{e \in D : d \leq e\}$. We say $w : D \to X$ **is eventually in** $A \subseteq X$ if $\exists d \in D, w[T_d] \subseteq A$. We say $w : D \to X$ **is frequently in** $A \subseteq X$ if $\forall d \in D, \exists e \in T_d, w(e) \in A$. Given a topological space (X, τ), we say w **converge** to a point x, if for all neighborhood O around x there exists $d \in D$ such that $w[T_d] \subseteq O$. Notice that if we wish to prove $w \to x$ we only have to verify for all open neighborhoods O around x. Also notice that w can converge to multiple points. A trivial example is when two point are topologically indistinguishable.

1.2 Equivalent definitions

1.3 Product topology

1.4 ABOVE ARE FIXED AND SHOULD BE FOCUSED BEFORE ARRANGEMENT OF THE FOLLOWING

1.5 Manifold

1.6 Quotient topology

1.7 Countability axioms

1.8 Separation axioms

1.9 Path connected

1.10 Basic notion on compact

1.11 Baire space

1.12 Homotopy

1.13 Simply connected

1.14 Fundamental group

Chapter 2

Metric Space and some Linear Algebra

2.1 Pseudo Metric

2.2 Completion

2.3 Bounded and Totally Bounded

2.4 Compactness

2.5 Holder Continuity

2.6 Limit Interchange

Given an arbitrary set X and a complete metric space $(\overline{Y}, d)$, in Section 2.8, we have proved that the set of functions with the following properties

(a) boundedness

(b) unboundedness

are respectively closed under uniform convergence. In next section (Section 2.7), we will prove that the following three properties

(a) continuity

(b) uniform continuity

(c) K-Lipschitz continuity

are again all closed under uniform convergence, where the proof for continuity is closed under uniform convergence use Theorem 2.6.1 as a lemma.

Here, we prove

(a) convergent of sequences

in, of course, complete metric space, is also closed under uniform convergence.

The reason we require the codomain $\overline{Y}$ of sequence to be complete is explained in the last paragraph of Section 2.8. 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 2.6.1. (Change Order of Limit Operations: Part 1) Given a double sequence $a_{n,k}$ whose codomain is (Y, d). Suppose

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

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

(c) $A_n \to A$

Then we can deduce

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

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 2.6.2. (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 $k \to \infty$

Then we can deduce

$$A_n \text{ converge and } A_n \to A$$

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 2.6.1 and Theorem 2.6.2, 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 2.8), 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 2.6.3, we give a necessary and sufficient condition for $\big(X^Y, d_\infty\big)$ to be complete.

Theorem 2.6.3. (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 2.6.4. (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 2.6.3, 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 2.6.3 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 2.6.3.

Now, before we actually prove convergence sequences are closed under uniform convergence if codomain (Y,d) is complete (Theorem 2.6.6), we will state and prove Weierstrass M-test (Theorem 2.6.5), which concerns the uniform convergence of series of complex functions.

Theorem 2.6.5. (Weierstrass M-test) Given sequences $f_n : X \to \mathbb{C}$, and suppose

$$\forall n \in \mathbb{N}, \forall x \in X, |f_n(x)| \leq M_n \tag{2.9}$$

Then

$$\sum_{n=1}^{\infty} M_n \text{ converge } \implies \sum_{n=1}^{\infty} f_n \text{ uniformly converge}$$

Proof. Because $(\mathbb{C}, \|\cdot\|_2)$ is complete, by Corollary 2.6.4, we only wish to prove

$$\left\{\sum_{k=1}^{n} f_k\right\}_{n\in\mathbb{N}} \text{ is uniformly Cauchy}$$

Fix ϵ. We wish

$$\text{to find } N \text{ such that } \forall n,m > N, \forall x \in X, \left|\sum_{k=n}^{m} f_k(x)\right| < \epsilon$$

Because $\sum_{n=1}^{\infty} M_n$ converge, we know there exists N such that

$$\forall n,m > N, \sum_{k=n}^{m} M_k < \epsilon$$

We claim

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

By Premise 2.9, we have

$$\forall n,m > N, \forall x \in X, \left|\sum_{k=n}^{m} f_k(x)\right| \leq \sum_{k=n}^{m} |f_k(x)| \leq \sum_{k=n}^{m} M_k < \epsilon$$

■

Theorem 2.6.6. (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 2.6.2, 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

$$\forall n,m > N, d_{\infty}(\{a_{n,k}\}_{k\in\mathbb{N}}, \{a_{m,k}\}_{k\in\mathbb{N}}) < \frac{\epsilon}{3} \tag{2.10}$$

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

$$d(a_{n,j}, A_n) < \frac{\epsilon}{3} \text{ and } d(a_{m,j}, A_m) < \frac{\epsilon}{3} \tag{2.11}$$

Then from Equation 2.10 and Equation 2.11, we can deduce

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

$\blacksquare$

2.7 Closed under Uniform Convergence

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

(a) continuity

(b) uniform continuity

(c) K-Lipschitz Continuity

Theorem 2.7.1. (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 2.6.1, 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 2.7.2. (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 2.7.3. **(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

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

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

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

■

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

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

2.8 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

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

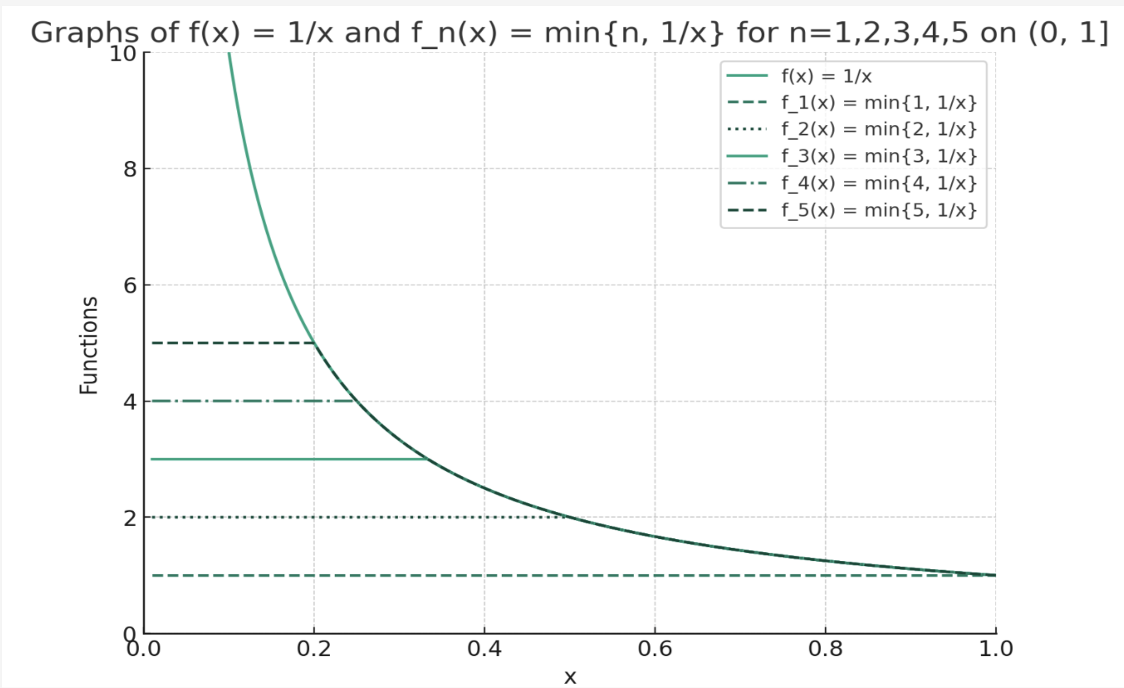
With this in mind, it shall be clear that the uniform limit of bounded (resp: unbounded) functions is always bounded (resp: unbounded).

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

Example 2 (Bounded functions pointwise converge to unbounded function)

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

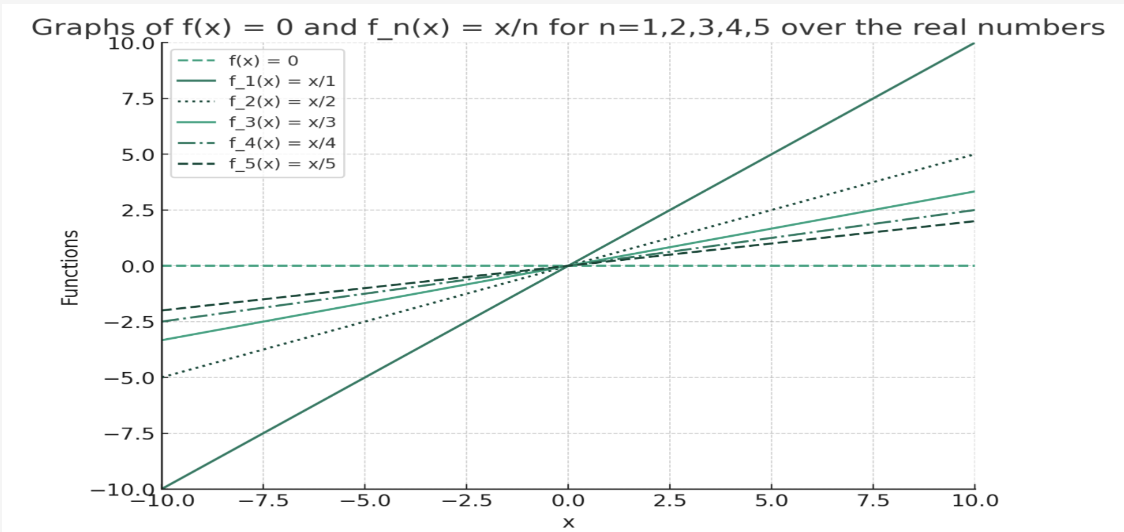
It is clear that $\forall n \in \mathbb{N}, f_n(x) \in [0, n]$, and the limit $f : X \to \mathbb{R}$ is $f(x) = \frac{1}{x}$



Example 3 (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 2.6.3, 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.

2.9 Arzelà–Ascoli Theorem

Given a topological space (X, τ) and a metric space (Y, d), let $\mathcal{F}$ be a collection of functions from X to Y. We say $\mathcal{F}$ is **equicontinuous** if for all ϵ and for all $x \in X$ there exists a neighborhood U_x around x such that

$$\forall y \in U_x, \forall f \in \mathcal{F}, d_Y\big(f(x), f(y)\big) < \epsilon$$

Theorem 2.9.1. (Arzelà–Ascoli theorem) Given a compact metric space (X, d_X), let $\mathcal{C}(X)$ denote the set of complex-valued continuous function on X. Let $\mathcal{F}$ be a subset of $\mathcal{C}(X)$.

$\mathcal{F}$ is compact in metric space $\big(\mathcal{C}(X), d_\infty\big) \iff \mathcal{F}$ is equicontinous and pointwise bounded

Proof. $(\longleftarrow)$

Let f_n be a sequence in $\mathcal{F}$. To prove $\mathcal{F}$ is compact in $\big(\mathcal{C}(X), d_\infty\big)$, we wish

to find a sub-sequence f_{n_k} that uniformly converge

Because X is compact, we know X is separable. Let $G = \{x_k\}_{k\in\mathbb{N}}$ be a countable dense subset of (X, d_X). Because $\mathcal{F}$ is pointwise bounded, we know

$$\Big\{f_n(x_k)\Big\}_{n\in\mathbb{N}} \text{ is bounded for all } k \in \mathbb{N} \tag{2.19}$$

Then by Bolzano-Weierstrass Theorem, we know there exists a convergent sub-sequence

$$\Big\{f_{g_1(k)}(x_1)\Big\}_{k\in\mathbb{N}}$$

Notice that $\Big\{f_{g_1(k)}(x_2)\Big\}_{k\in\mathbb{N}}$ is a sub-sequence of $\Big\{f_n(x_2)\Big\}_{n\in\mathbb{N}}$. Then because Equation 2.19, we know $\Big\{f_{g_1(k)}(x_2)\Big\}_{k\in\mathbb{N}}$ is bounded. Then again by Bolzano-Weierstrass Theorem, we know there exists a convergent sub-sequence

$$\Big\{f_{g_2\circ g_1(k)}(x_2)\Big\}_{k\in\mathbb{N}}$$

Again, notice that $\Big\{f_{g_2\circ g_1(k)}(x_3)\Big\}$ is a sub-sequence of $\Big\{f_n(x_3)\Big\}_{n\in\mathbb{N}}$. Then again because Equation 2.19, again by Bolzano-Weierstrass Theorem, we know there exists a convergent sub-sequence

$$\Big\{f_{g_3\circ g_2\circ g_1(k)}(x_3)\Big\}_{k\in\mathbb{N}}$$

Repeatedly do such, we see

$$\begin{array}{cccc}
f_{g_1(1)}(x_1) & f_{g_2\circ g_1(1)}(x_2) & f_{g_3\circ g_2\circ g_1(1)}(x_3) & \cdots \\
f_{g_1(2)}(x_1) & f_{g_2\circ g_1(2)}(x_2) & f_{g_3\circ g_2\circ g_1(2)}(x_3) & \cdots \\
f_{g_1(3)}(x_1) & f_{g_2\circ g_1(3)}(x_2) & f_{g_3\circ g_2\circ g_1(3)}(x_3) & \cdots \\
\vdots & \vdots & \vdots & \\
\lim_{k\to\infty} f_{g_1(k)}(x_1) & \lim_{k\to\infty} f_{g_2\circ g_1(k)}(x_2) & \lim_{k\to\infty} f_{g_3\circ g_2\circ g_1(k)}(x_3) \in Y &
\end{array}$$

Now, let

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

Then, we see

$$\forall m \in \mathbb{N}, n_k \text{ is eventually a sub-sequence of } g_m \circ \cdots \circ g_1(k)$$

Then we see

$$\forall j \in \mathbb{N}, f_{n_k}(x_j) \to \lim_{k\to\infty} f_{g_j\circ\cdots\circ g_1(k)}(x_k) \text{ as } k \to \infty$$

We claim

$$f_{n_k} \text{ uniformly converge}$$

Fix ϵ. We wish

$$\text{to find } K \in \mathbb{N} \text{ such that } \forall k,l > K, \forall x \in X, |f_{n_k}(x) - f_{n_l}(x)| < \epsilon$$

Because $\mathcal{F}$ is equicontinuous and $\{f_{n_k}\}_{k\in\mathbb{N}} \subseteq \mathcal{F}$, we know

$$\exists \delta, \forall k \in \mathbb{N}, \forall x,y \in X, d_X(x,y) < \delta \implies |f_{n_k}(x) - f_{n_k}(y)| < \frac{\epsilon}{3} \tag{2.20}$$

Fix such δ. Now, because

$$f_{n_j}(x_p) \text{ converge as } j \to \infty \text{ for each fixed } p \in \mathbb{N}$$

For each fixed $p \in \mathbb{N}$, we can let N_p satisfy

$$\forall k,l > N_p, |f_{n_k}(x_p) - f_{n_l}(x_p)| < \frac{\epsilon}{3} \tag{2.21}$$

Because $G = \{x_k\}_{k\in\mathbb{N}}$ is dense (See the statement above Equation 2.19, the statement contains the definition of G), we see that

$$\Big\{B_\delta(x_k) : k \in \mathbb{N}\Big\} \text{ is an open cover of } (X, d_X)$$

Recall by premise that (X, d_X) is compact. Then we know

$$\text{there exists a finite set } Q \text{ of natural number such that } \Big\{B_\delta(x_k) : k \in Q\Big\} \text{ is an open cover} \tag{2.22}$$

We claim

$$K = \max_{q \in Q} N_q \text{ where } N_q \text{ is defined by Equation 2.21 works.} \tag{2.23}$$

Fix $k, l > K$ and $x \in X$. We wish

$$\text{to prove } |f_{n_k}(x) - f_{n_l}(x)| < \epsilon$$

Because $\Big\{B_\delta(x_k) : k \in Q\Big\}$ is an open cover (We create this in Equation 2.22 from the compactness of X), we know there exists $i \in Q$ such that

$$d_X(x_i, x) < \delta$$

By Equation 2.20 we know

$$|f_{n_k}(x) - f_{n_k}(x_i)| < \frac{\epsilon}{3} \text{ and } |f_{n_l}(x) - f_{n_l}(x_i)| < \frac{\epsilon}{3} \tag{2.24}$$

From the definition of K (Equation 2.23 and Equation 2.21), we know

$$|f_{n_k}(x_i) - f_{n_l}(x_i)| < \frac{\epsilon}{3} \tag{2.25}$$

Then from Inequality 2.24 and Inequality 2.25, we see

$$|f_{n_k}(x) - f_{n_l}(x)| \leq |f_{n_k}(x) - f_{n_k}(x_i)| + |f_{n_k}(x_i) - f_{n_l}(x_i)| + |f_{n_l}(x_i) - f_{n_l}(x)| < \epsilon \text{ (done)}$$

$\blacksquare$

2.10 The Stone-Weierstrass Theorem

2.11 Contraction Mapping Theorem

Chapter 3

Calculus

3.1 Equivalent Definitions for Riemann Integral

In this section, we will give a principal Theorem (Theorem 3.1.1) that can serve as a lemma to prove the equivalency of multiple different definitions of Riemann integral on a compact interval. With this approach, we shall diminish the trouble of getting through miscellaneous minor definitions, where they are all equivalent, with only the difference of taking different tags and partitions of certain pattern, which solely serve as a pedagogical tool to give students a concrete idea of integration.

A caveat will be made clear here: this section concern only the proper Riemann Integral. That is, we only consider the integration of a function bounded on a compact interval. For a treatment of inproper integral, see Section **??**.

In this section, by a **partition** P of $[a,b]$, we mean a finite set of values $P=\{a=x_0\leq x_1\leq\cdots\leq x_{n_P-1}\leq x_{n_P}=b\}$. We say the partition P' is **finer** than P if $P\subseteq P'$. Given a partition P, we put

$$\begin{cases}M_i=\sup_{[x_{i-1},x_i]}f(x)\\ m_i=\inf_{[x_{i-1},x_i]}f(x)\end{cases}\quad\text{and}\quad\begin{cases}U(P,f)=\sum_{i=1}^{n_P}M_i\Delta x_i\\ L(P,f)=\sum_{i=1}^{n_P}m_i\Delta x_i\end{cases}\quad\text{where }\Delta x_i=x_i-x_{i-1}$$

We shall write n instead of n_P if no confusion will be made.

The word **norm of partition** $\|P\|$ is defined by $\max_{1\leq i\leq n}\Delta x_i$. We say $U(P,f)$ is an **upper sum** of f. We say the **upper integral** $\overline{\int_a^b} f\,dx$ of f on $[a,b]$ is $\inf_P U(P,f)$ where the infimum run through all partitions P of $[a,b]$. The **lower integral** $\underline{\int_a^b} f\,dx$ is defined similarly. We say a function f is **integrable** on $[a,b]$ if $\overline{\int_a^b} f dx = \underline{\int_a^b} f dx$.

Give close attention to the setting that P is finite. This is crucial for making the integration operation possible, since if P is countable and we define $U(P,f)$ by taking limits for sums, the order of addition can make a difference if the sum does not converge absolutely. This fact is backed by Riemann Rearrangement Theorem (Theorem 3.1.7), of which we will later give a proof.

Theorem 3.1.1. (Principal for Proving Equivalency of Definitions for Riemann Integral)

$$\int_a^b f dx \in \mathbb{R} \iff \forall \{P_k\} : \|P_k\| \to 0, U(P_k,f) - L(P_k,f) \to 0$$

Proof. From right to left is obvious. We prove only

$$\int_a^b f dx \in \mathbb{R} \implies \forall \{P_k\} : \|P_k\| \to 0, U(P_k,f) - L(P_k,f) \to 0$$

Fix ϵ. We wish to find a positive number $\beta \in \mathbb{R}^+$ such that $\forall P : \|P\| \leq \beta, U(P,f) - L(P,f) < \epsilon$. Because $\int_a^b f dx \in \mathbb{R}$, we can let W be a partition such that

$$U(W,f) - L(W,f) < \frac{\epsilon}{2}$$

Let $W = \{a = x_0^*, x_1^*, \dots, x_{n_W}^* = b\}$, and let $J = \{1, \dots, n_W\}$ be the set of indices of W. Suppose

$$L = \max_{1 \leq j \leq n_W - 1} \big(\sup_{[x_{j-1}^*, x_{j+1}^*]} f(x) - \inf_{[x_{j-1}^*, x_{j+1}^*]} f(x) \big) \tag{3.1}$$

Notice that if $L = 0$, then f must be constant and the proof become trivial, so we can assume $L > 0$. We claim that

$$L\beta n_W \leq \frac{\epsilon}{2} \text{ and } \beta < \min_{j \in J} \Delta x_j$$

suffice so that $\forall P : \|P\| \leq \beta, U(P,f) - L(P,f) < \epsilon$. Let $C = \min_{j \in J} \Delta x_j$. In other words, we now reduce the problem into proving

$$\|P\| \leq \min\{\frac{\epsilon}{2Ln_W}, C\} \implies U(P,f) - L(P,f) < \epsilon$$

Let $I = \{1, \dots, n_P\}$ be the set of indices for P. Suppose

$$P = \{a = x_0, x_1, \dots, x_{n_P} = b\}$$

We partition I into

$$\begin{cases} A = \Big\{ i \in I : \exists j \in J, [x_{i-1}, x_i] \subseteq [x^*_{j-1}, x^*_j] \Big\} \\ B = I \setminus A \end{cases}$$

We now have

$$U(P, f) - L(P, f) = \sum_{i \in A} \big(M_i - m_i \big) \Delta x_i + \sum_{i \in B} \big(M_i - m_i \big) \Delta x_i \tag{3.2}$$

Because for each $i \in A$, there is a unique corresponding $j \in J$ such that $[x_{i-1}, x_i] \subseteq [x^*_{j-1}, x^*_j]$, we have

$$\sum_{i \in A} \big(M_i - m_i \big) \Delta x_i \leq \sum_{j \in J} \big(M^W_j - m^W_j \big) \Delta x^*_j = U(W, f) - L(W, f) < \frac{\epsilon}{2} \tag{3.3}$$

Because $\|P\| \leq C = \min_{j \in J} \Delta x_j$, we know for each distinct $i \in B$, there exists a distinct $j \in J$ such that $[x_{i-1}, x_i] \subseteq [x^*_{j-1}, x^*_{j+1}]$, so by definition of L (Equation 3.1), we have

$$\sum_{i \in B} \big(M_i - m_i \big) \Delta x_i \leq L \sum_{i \in B} \Delta x_i \leq L n_W \|P\| \leq \frac{\epsilon}{2} \tag{3.4}$$

Combining Equation 3.2, Equation 3.3 and Equation 3.4, we now see

$$U(P, f) - L(P, f) < \epsilon \text{ (done)}$$

■

Recall that we say a series $\sum_{n=1}^{\infty} a_n$ **absolutely converge** if $\sum_{n=1}^{\infty} |a_n|$ converge. We can show that a series converges if it absolutely converges by proving it is Cauchy. In this section, by a **permutation on** $\mathbb{N}$, we mean a bijective function σ from $\mathbb{N}$ to $\mathbb{N}$. Another two important terminologies are the followings. We say that $\sum_{n=1}^{\infty} a_n$ **unconditionally converge** if for all permutation $\sigma : \mathbb{N} \to \mathbb{N}$, the series $\sum_{n=1}^{\infty} a_{\sigma(n)}$ converge to the same number. We say $\sum_{n=1}^{\infty} a_n$ **conditionally converge** if it converge but not unconditionally.

In our treatment, Riemann Rearrangement Theorem will be split into 4 parts. The summary is at Theorem 3.1.7. The first part (Theorem 3.1.2) states that the limit an absolutely convergent series remain the same under any permutations. The proof for the first part (Theorem 3.1.2) may seem technical, but the essence is quite easy to remember. Just "see" that

$$\left|\sum_{k=1}^{\infty} a_{\sigma(k)} - L\right| \leq \left|\sum_{i<M} a_i - L + \sum_{i\geq M} a_i\right|$$
$$\leq \left|\sum_{i<M} a_i - L\right| + \sum_{i\geq M} |a_i| \to 0 \text{ as } M \to 0$$

Theorem 3.1.2. (Riemann Rearrangement Theorem, Part 1)

$$\sum_{k=1}^{\infty} a_k \text{ absolutely converge } \implies \sum_{k=1}^{\infty} a_k \text{ unconditionally converge}$$

Proof. Suppose $\sum_{k=1}^{\infty} |x_k|$ converge. Let $\sum_{k=1}^{\infty} x_k = L$. Fix permutation $\sigma : \mathbb{N} \to \mathbb{N}$. We wish to prove

$$\sum_{k=1}^{\infty} x_{\sigma(k)} = L$$

Fix ϵ. We reduce the problem into

$$\text{finding } N \text{ such that } \forall n > N, \left|\sum_{k=1}^{n} x_{\sigma(k)} - L\right| < \epsilon$$

Because both $\sum_{k=1}^{\infty} x_k$ and $\sum_{k=1}^{\infty} |x_k|$ converge by premise. We know there exists M such that

$$\forall n > M, \left|\sum_{k=1}^{n} x_k - L\right| < \frac{\epsilon}{2} \text{ and } \sum_{k=n}^{\infty} |x_k| < \frac{\epsilon}{2} \tag{3.5}$$

Let

$$I = \sigma^{-1}\big(\{1, \dots, M\}\big) \text{ and } N = \max I$$

We claim

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

To prove our claim, fix $n > N$. We wish to show

$$\left|\sum_{k=1}^{n} x_{\sigma(k)} - L\right| < \epsilon$$

Let $I_n = \{1, \dots, n\}$. Observe that

$$\left|\sum_{k=1}^{n} x_{\sigma(k)} - L\right| = \left|\sum_{k \in I_n \setminus I} x_{\sigma(k)} + \sum_{k \in I} x_{\sigma(k)} - L\right| \tag{3.6}$$

Notice that $k \notin I \implies \sigma(k) > M$. Then by definition of M (Equation 3.5), we have

$$\left|\sum_{k \in I_n \setminus I} x_{\sigma(k)}\right| \leq \sum_{k \in I_n \setminus I} \left|x_{\sigma(k)}\right| \leq \sum_{j > M}^{\infty} |x_j| < \frac{\epsilon}{2} \tag{3.7}$$

Notice that $\sigma[I] = \{1, \dots, M\}$. Then also by definition of M (Equation 3.5), we have

$$\left|\sum_{k \in I} x_{\sigma(k)} - L\right| = \left|\sum_{j=1}^{M} x_j - L\right| < \frac{\epsilon}{2} \tag{3.8}$$

Then by inequalities Equation 3.6, Equation 3.7 and Equation 3.8, we now have

$$\begin{aligned}
\left|\sum_{k=1}^{n} a_{\sigma(k)} - L\right| &= \left|\sum_{k \in I_n \setminus I} x_{\sigma(k)} + \sum_{k \in I} x_{\sigma(k)} - L\right| \\
&\leq \left|\sum_{k \in I_n \setminus I} x_{\sigma(k)}\right| + \left|\sum_{k \in I} x_{\sigma(k)} - L\right| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \text{ (done)}
\end{aligned}$$

■

The second, third and forth parts (Theorem 3.1.4, Theorem 3.1.5 and Theorem 3.1.6), respectively states that if a series converge but not absolutely, then the limit value can be changed to any real number, infinite, negative infinite and even jumping-ly diverges.

The detail of the proof is very tedious and cumbersome, while the essence is easy to understand. The only two tools for proving Theorem 3.1.4, Theorem 3.1.5 and Corollary 3.1.6, is Lemma 3.1.3 and the fact $\sum a_k \to L \implies a_k \to 0$. If any part of the proof can be considered interesting, I believe it lies in that of Lemma 3.1.3, where one split the series $\sum a_k$ into two $\sum a_k^+, \sum a_k^-$, and shows that they must both diverge.

Lemma 3.1.3. (Intrinsic Structure of Series that Converge but not Absolutely) Let $f^+ : \mathbb{N} \to \mathbb{N}$ and $f^- : \mathbb{N} \to \mathbb{N}$ satisfy that $\{a_{f^+(n)}\}$ contain all and only positive terms of $\{a_n\}$ and $\{a_{f^-(n)}\}$ contain all and only negative terms. If $\sum_{k=1}^{\infty} a_k$ converge but not absolutely, then for each $\alpha \in \mathbb{R}^+$ and $n \in \mathbb{N}$, there exists $u_n > n$ and $t_n > n$ such that

$$\sum_{n \leq k \leq u_n} a_{f^+(k)} > \alpha \text{ and } \sum_{n \leq k \leq t_n} a_{f^-(k)} < -\alpha$$

Proof. Let $a_n^+ = \max\{0, a_n\}$ and $a_n^- = \min\{0, a_n\}$. It is easy to check $\forall n \in \mathbb{N}, a_n = a_n^+ + a_n^-$. Because $\sum_{k=1}^{\infty} a_k$ converge but not absolutely, we know

$$\begin{cases} \sum_{k=1}^{\infty} a_k^+ \to \infty \\ \sum_{k=1}^{\infty} a_k^- \to -\infty \end{cases} \tag{3.9}$$

This is true because if both of them converge then $\sum_{k=1}^{\infty} |a_k|$ converges and if only one of them converge them $\sum_{k=1}^{\infty} a_k$ diverges.

Because of Equation 3.9, we know

$$\begin{cases} \sum_{k=1}^{\infty} a_{f^+(k)} \to \infty \\ \sum_{k=1}^{\infty} a_{f^-(k)} \to -\infty \end{cases}$$

The result then follow, since

$$\forall n \in \mathbb{N}, \sum_{k \geq n} a_{f^+(k)} \nearrow \infty \text{ and } \sum_{k \geq n} a_{f^-(k)} \searrow -\infty$$

■

Theorem 3.1.4. (Riemann Rearrangement Theorem, Part 2) If $\sum_{k=1}^{\infty} a_k$ converge but not absolutely, then there exists permutations $\sigma_{\infty}, \sigma_{-\infty} : \mathbb{N} \to \mathbb{N}$ such that $\sum_{k=1}^{\infty} a_{\sigma_{\infty}(k)} \to \infty$ and $\sum_{k=1}^{\infty} a_{\sigma_{-\infty}(k)} \to -\infty$.

Proof. We wish

$$\text{to construct } \sigma_{\infty} : \mathbb{N} \to \mathbb{N} \text{ such that } \sum_{k=1}^{\infty} a_{\sigma_{\infty}(k)} \to \infty$$

Using Lemma 3.1.3, construct $\sigma_{\infty} : \mathbb{N} \to \mathbb{N}$ as follows. Let p_n be a sequence of natural number such that for each $n \in \mathbb{N}$, p_{n+1} is the smallest natural number such that

$$\sum_{k=p_n+1}^{p_{n+1}} a_{f^+(k)} > 3 \text{ and } p_1 = 0 \tag{3.10}$$

Similarly, let q_n be a sequence of natural number such that for each $n \in \mathbb{N}, q_{n+1}$ is the smallest natural number such that

$$\sum_{k=q_n+1}^{q_{n+1}} a_{f^-(k)} < -1 \text{ and } q_1 = 0 \tag{3.11}$$

Notice that the definition of p_n and q_n (Equation 3.10, Equation 3.11) are done recursively. Now, recursively define σ_∞ to follow the order

$$\begin{aligned} &f^+(p_1+1), \dots, f^+(p_2), f^-(q_1+1), \dots, f^-(q_2) \\ \longrightarrow &f^+(p_2+1), \dots, f^+(p_3), f^-(q_2+1), \dots, f^-(q_3), f^+(p_3+1), \dots \end{aligned}$$

If there exists $k \in \mathbb{N}$ such that $a_k = 0$, which is not in the range $f^+[\mathbb{N}] \cup f^-[\mathbb{N}]$, we can merge these zero term into our σ_∞ by putting them in terms of even order. This way, our σ_∞ then become bijecetive, a permutation.

We claim

$$\text{such } \sigma_\infty \text{ works}$$

Recall the definition of p_n (Equation 3.10) is that for each $n \in \mathbb{N}, p_{n+1}$ is the smallest natural number such that

$$\sum_{k=p_n+1}^{p_{n+1}} a_{f^+(k)} > 3$$

Also recall the similarly defined q_n. This tell us

$$\sum_{k=p_n+q_n+1}^{p_{n+1}+q_{n+1}} a_{\sigma_\infty(k)} \to 2 \text{ as } n \to \infty$$

where

$$\sum_{k=p_n+q_n+1}^{p_{n+1}+q_n} a_{\sigma_\infty(k)} \to 3 \text{ and } \sum_{k=p_{n+1}+q_n+1}^{p_{n+1}+q_{n+1}} a_{\sigma_\infty(k)} \to -1 \text{ as } n \to \infty$$

With this, it is easy to verify $\sum_{k=1}^{\infty} a_{\sigma_\infty(k)} \to \infty$ (done) . The construction of $\sigma_{-\infty}$ and the proof for its validity is done similarly. ∎

Theorem 3.1.5. (Riemann Rearrangement Theorem, Part 3) If $\sum_{k=1}^{\infty} a_k$ converges but not absolutely, then for all $[L, M] \subseteq \mathbb{R}$, there exists a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\liminf_{n\to\infty} \sum_{k=1}^{n} a_{\sigma(k)} = L$ and $\limsup_{n\to\infty} \sum_{k=1}^{n} a_{\sigma(k)} = M$.

Proof. We wish

to construct a working σ

The construction of σ is similar to that of σ_∞ in Theorem 3.1.4. WOLG, let $M > 0$. Let $p_1 = 0$, and let p_2 be the smallest natural number such that

$$\sum_{k=1}^{p_2} a_{f^+(k)} > M \text{ and } p_1 = 0$$

Next, define $q_1 = 0$ and let q_2 be the smallest natural number such that

$$\sum_{k=1}^{p_2} a_{f^+(k)} + \sum_{k=1}^{q_2} a_{f^-(k)} < L$$

Then, let p_3 be the smallest natural number such that

$$\sum_{k=1}^{p_3} a_{f^+(k)} + \sum_{k=1}^{q_2} a_{f^-(k)} > M$$

Recursively do such. We get two sequences $\{p_n\}, \{q_n\}$ of natural number such that for all $n \in \mathbb{N}, p_{n+1}$ is the smallest natural number such that

$$\sum_{k=1}^{p_{n+1}} a_{f^+(k)} + \sum_{k=1}^{q_n} a_{f^-(k)} > M$$

and for all $n \in \mathbb{N}, q_{n+1}$ is the smallest natural number such that

$$\sum_{k=1}^{p_{n+1}} a_{f^+(k)} + \sum_{k=1}^{q_{n+1}} a_{f^-(k)} < L$$

Them, recursively define σ to follow the order

$$\begin{aligned} &f^+(p_1+1), \dots, f^+(p_2), f^-(q_1+1), \dots, f^-(q_2) \\ \longrightarrow &f^+(p_2+1), \dots, f^+(p_3), f^-(q_2+1), \dots, f^-(q_3), f^+(p_3+1), \dots \end{aligned}$$

Again, merge in the zero terms like in Theorem 3.1.4. The proof for the claim such σ works is easy to verify knowing $a_{\sigma(k)} \to 0$ (done) ■

Corollary 3.1.6. (Riemann Rearrangement Theorem, Part 4) If $\sum_{k=1}^\infty a_k$ converges but not absolutely, then for all $L \in \mathbb{R}$, there exists a permutation σ such that $\sum_{k=1}^\infty a_{\sigma(k)} = L$

Theorem 3.1.7. (Summary of Riemann Rearrangement Theorem) If $\sum_{k=1}^{\infty} a_k$ converge, then

$$\sum_{k=1}^{\infty} a_k \text{ absolutely converges} \iff \sum_{k=1}^{\infty} a_k \text{ unconditionally converges}$$

Proof. $(\longrightarrow)$

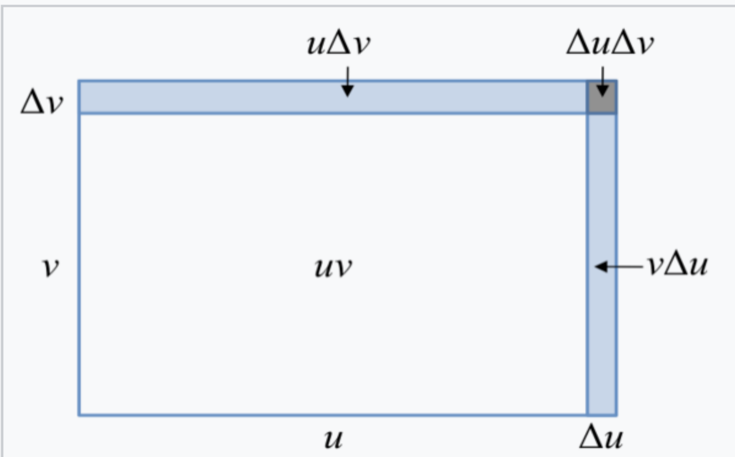
This is Theorem 3.1.2.

$(\longleftarrow)$

The fact that the contraposition of this statement is true is implied by any of Theorem 3.1.4, Theorem 3.1.5 and Corollary 3.1.6. ■

3.2 Product, Quotient and Chain Rule

This section concern mostly the computation of actual value of the derivative and integral of function. With this in mind, we first prove the product and quotient rules for derivative of $\mathbb{R}$ to $\mathbb{R}$ functions taught in most Calculus 1 classes. The proofs for the laws are easy, as it require no ingenious idea but ability to manipulate the limit symbol. However, without philosophical comments, we left an graph for geometric intuition for product rule. There are also graphs for geometric intuition for quotient rule on Internet, but we won't put it here as it require more than subtle work to understand the graph.



Geometric illustration of a proof of the product rule

Theorem 3.2.1. (Product Rule and Quotient Rule for Real to Real Function) Suppose f and g is differentiable at x, and $g'(x) \neq 0$. We have

(a) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$ (Product Rule)

(b) $(\frac{f}{g})'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{\big(g(x)\big)^2}$ (Quotient Rule)

Proof. Compute

$$\begin{aligned}
(fg)'(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \big(g(x+h)\big)\frac{f(x+h) - f(x)}{h} + \big(f(x)\big)\frac{g(x+h) - g(x)}{h} \\
&= g(x)f'(x) + f(x)g'(x)
\end{aligned}$$

Compute

$$
\begin{aligned}
(\frac{f}{g})'(x) &= \lim_{h\to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x+h)}{hg(x+h)g(x)} \\
&= \lim_{h\to 0} (\frac{1}{g(x+h)g(x)}) \cdot (\frac{f(x+h)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(x+h)}{h}) \\
&= \lim_{h\to 0} \Big(\frac{1}{g(x+h)g(x)}\Big) \cdot \Big(\big(g(x)\big)\frac{f(x+h) - f(x)}{h} + \big(f(x)\big)\frac{g(x) - g(x+h)}{h}\Big) \\
&= \Big(\frac{1}{\big(g(x)\big)^2}\Big) \cdot \Big(\big(g(x)f'(x)\big) - f(x)g'(x)\Big) = \frac{f'(x)g(x) - f(x)g'(x)}{\big(g(x)\big)^2}
\end{aligned}
$$

■

Even a year has past, I can still remember what happened in the first class of Vector Analysis last year. The professor asked: "What is derivative?". A lot of answers emerge, from extremely formal and abstract like $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ to those following geometric intuition like tangent line. Everyone gave a correct answer, but none of them philosophically satisfy the requirement of the question from the professor. Then, he stated: "Derivative is exactly linear approximation", and stated on black board the most general definition:

Definition 3.2.2. (Definition of Differential) Given two normed space V, W and an open subset $U \subseteq V$, we say a function $f : U \to W$ is **differentiable at** x if there exists a bounded linear operator $A : V \to W$ such that

$$\lim_{h\to 0} \frac{\|f(x+h) - f(x) - A(h)\|_W}{\|h\|_V} = 0$$

and we say the bounded linear map A is the **(total) derivative** of f at x.

If one put the key words "proof for chain rule" in Google search box, just like the situation in my classes, lots of rigorous proof emerge, but none of them is philosophical satisfying. For this reason, I shall give a proof of chain rule for real to real function based on the concept of linear approximation.

In Baby Rudin, derivative of a real to real function f is defined by

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

Immediately from this definition, we can derive a linear approximation P of f at a by setting

$$P(x) = f(a) + f'(a)(x-a)$$

Then, we see if we set $R(x) = f(x) - P(x)$ as the error (or the remainder) of the approximation, then trivially we have the behavior

$$R(x) \to 0 \text{ as } x \to a$$

what behavior of $R(x)$ that give P the name approximation is

$$\frac{R(x)}{x-a} \to 0 \text{ as } x \to a$$

The difference between the two behaviors is symbolically apparent, yet without geometric help, it may be difficult to precisely describe how insignificant the first behavior is compared to the second behavior. For this, observe that any function g that converge to $f(a)$ at a satisfy the first behavior, yet only a few satisfy the second. One can easily verify that the only linear $\mathbb{R}$ to $\mathbb{R}$ function that satisfy the second behavior is $P(x) = f(a) + f'(a)(x-a)$. Geometrically, this means that $R(x) = o(f'(x)dx)$ as $x \to a$.

Theorem 3.2.3. (Chain Rule for $\mathbb{R}$ to $\mathbb{R}$ function) Suppose g is differentiable at a and f is differentiable at $g(a)$. We have

$$(f \circ g)'(a) = f'(g(a))g'(a)$$

Proof. Define the remainders $R_{f(g(a))}(x)$ and $R_{g(a)}(x)$ by

$$\begin{cases} R_{f(g(a))}(x) = f(x) - f(g(a)) - f'(g(a))(x - g(a)) \\ R_{g(a)}(x) = g(x) - g(a) - g'(a)(x-a) \end{cases}$$

Compute

$$(f \circ g)'(a) = \lim_{x \to a} \frac{f(g(x)) - f(g(a))}{x-a} \tag{3.12}$$

$$= \lim_{x \to a} \frac{R_{f(g(a))}(g(x)) + f'(g(a))(g(x) - g(a))}{x-a} \tag{3.13}$$

Notice that because $x \to a \implies g(x) \to g(a)$, we have

$$\lim_{x \to a} \frac{R_{f(g(a))}(g(x))}{x-a} = \lim_{x \to a} \frac{R_{f(g(a))}(g(x))}{g(x) - g(a)} \cdot \frac{g(x) - g(a)}{x-a} = 0 \cdot g'(x) = 0 \tag{3.14}$$

Notice that the above deduction (Equation 3.14) is quite informal for two reasons: First, it may happen that $g(x) = g(a)$ locally. Second, for some reader it may require a mini proof to verify that $\frac{R_{f(g(a))}(g(x))}{g(x)-g(a)} \to 0$ as $x \to a$. These two obstacles for advanced readers should be insignificant.

Getting back to Equation 3.12, by Equation 3.14, we now see

$$(f \circ g)'(a) = \lim_{x \to a} \frac{f'(g(a))(g(x) - g(a))}{x - a} = f'(g(a))g'(a)$$

which finish the proof. ■

3.3 IVT, EVT and MVT

If one wish to understand most of the Theorems after this section, one must first know MVT, for what an important role it cast in the sections after. Logically prior to MVT is IVT. Yet, unlike MVT involve the intrinsic nature of field and limit structure of $\mathbb{R}$. IVT can be considered as purely topological in the sense that its proof can be stated almost in the language of topology. There are only two facts (the first are purely topological and the second is very close to purely topological) one need to know to prove IVT.

First, continuous functions map a connected sets to connected set. Second, a set in $\mathbb{R}$ is connected if and only if it is an interval.

Combining the above two facts, we have the following statement:

Theorem 3.3.1. (Continuous Real to Real Function Maps Interval to Interval) as titled.

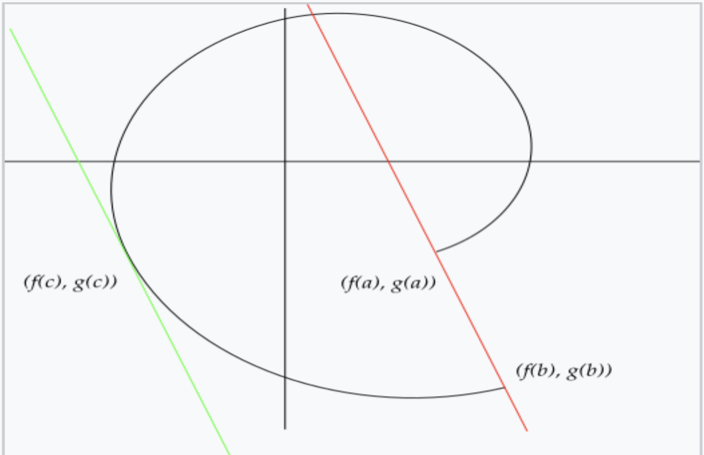
Proof. Consider the fact a continuous function map connected sets to connected sets and the fact a set in $\mathbb{R}$ is connected if and only if it is an interval. ∎

Then, given the necessary constraint (the interval considered is compact) to give the conclusion, we have the famous statement:

Theorem 3.3.2. (IVT) Given a continuous function $f : [a, b] \to \mathbb{R}$, for each y that lies between $f(a)$ and $f(b)$, there exists $x \in [a, b]$ such that $f(x) = y$.

Given the simplicity of the logical deduction, we shall not give a rigorous proof here. However, one can notice that the interval considered in IVT "must be" compact, otherwise the Theorem is invalid. This constraint is in some sense a showcase how the concept of compact really match the description of "smallness (bounded) and rigidness (closed)".

Compared to IVT, another famous MVT is richer in both the results and the proof. Clearly for a logical and economic purpose, we shall first prove the Cauchy MVT.



Geometrical meaning of Cauchy's theorem

Theorem 3.3.3. (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. We wish to find $x \in (a, b)$ such that

$$[f(b) - f(a)]g'(x) - [g(b) - g(a)]f'(x) = 0$$

Define h on (a, b) by

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

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 of h on (a, b) such that

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

We have $h = H'$ on (a, b). This let us reduce our 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 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 3.3.4. (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. Let $g(x) = x$ in Cauchy's MVT (Theorem 3.3.3), and we are done. ■

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 differentiablity 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

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

Theorem 3.3.5. (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

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in (a,b) \\ \lim_{t\to a} f(t) & \text{if } x = a \\ \lim_{t\to b} f(t) & \text{if } x = b \end{cases} \tag{3.15}$$

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 3.15, by EVT, we know we there exists $\alpha, \beta \in [a,b]$ such that

$$\tilde{f}(\alpha) = \inf_{x\in[a,b]} \tilde{f}(x) \text{ and } \tilde{f}(\beta) = \sup_{x\in[a,b]} \tilde{f}(x) \tag{3.16}$$

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

$$\exists \xi \in [\alpha,\beta], \tilde{f}(\xi) = \frac{\int_a^b \tilde{f}(x)dx}{b-a} \tag{3.17}$$

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 3.16), 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

$$\inf_{x\in[\beta-\delta,\beta]} \tilde{f}(x) \geq \frac{\tilde{f}(\alpha) + 2\tilde{f}(\beta)}{3} \tag{3.18}$$

We then from Equation 3.17 see that

$$\int_a^b \tilde{f}(x)dx = \tilde{f}(\xi)(b-a) = \tilde{f}(\alpha)(b-a) \tag{3.19}$$

Also, we see from definition of α (Equation 3.16) and Equation 3.18 that

$$\int_a^b \tilde{f}(x)dx = \int_a^{\beta-\delta} \tilde{f}(x)dx + \int_{\beta-\delta}^{\beta} \tilde{f}(x)dx + \int_{\beta}^{b} \tilde{f}(x)dx \tag{3.20}$$

$$\geq (b-\delta-a)\tilde{f}(\alpha) + \delta \cdot (\frac{\tilde{f}(\alpha) + 2\tilde{f}(\beta)}{3}) \tag{3.21}$$

$$> (b-\delta-a)\tilde{f}(\alpha) + \delta \cdot (\frac{\tilde{f}(\alpha) + \tilde{f}(\beta)}{2}) \tag{3.22}$$

$$= \tilde{f}(\alpha)\big(b-a-\frac{\delta}{2}) + \tilde{f}(\beta) \cdot (\frac{\delta}{2}) \tag{3.23}$$

Now, from Equation 3.19 and Equation 3.23, 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

$$\tilde{f}(\alpha) \cdot (\frac{\delta}{2}) > \tilde{f}(\beta) \cdot (\frac{\delta}{2}) \text{ CaC } \text{ (done)}$$

■

Theorem 3.3.6. (Second Mean Value Theorem for Definite Integral) Given functions $G, \phi : [a,b] \to \mathbb{R}$ 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

$$\int_a^b G(t)\phi(t)dt = G(a^+)\int_a^{\xi} \phi(t)dt + G(b^-)\int_{\xi}^{b} \phi(t)dt$$

Proof. Define f on $[a,b]$ by

$$f(x) = G(a^+)\int_a^{x} \phi(t)dt + G(b^-)\int_{x}^{b} \phi(t)dt$$

We then reduce the problem into

$$\text{finding } \xi \in (a,b) \text{ such that } \int_a^b G(t)\phi(t)dt = f(\xi)$$

By Theorem 3.5.1, we know f is continuous on $[a,b]$. Then by IVT, we can reduce the problem into

$$\text{finding an interval } [c,d] \subseteq (a,b) \text{ such that } \int_a^b G(t)\phi(t) \text{ is between } f(c) \text{ and } f(d)$$

Observe that

$$f(a) = G(b^-)\int_a^b \phi(t)dt \text{ and } f(b) = G(a^+)\int_a^b \phi(t)dt$$

$\blacksquare$

3.4 Computation for Riemann-Stieltjes Integral

Theorem 3.4.1. (Change of Variable) Given two functions $g, \beta : [A, B] \to \mathbb{R}$, a function $\phi : [A, B] \to [a, b]$ and two functions $f, \alpha : [a, b] \to \mathbb{R}$ such that

(a) $g = f \circ \phi$ for all $x \in [a, b]$

(b) $\beta = \alpha \circ \phi$ for all $x \in [a, b]$

(c) α, β increase respectively on $[a, b]$ and $[A, B]$

(d) $\phi : [A, B] \to [a, b]$ 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

$$\text{there exists a partition } P \text{ of } [a, b] \text{ such that } U(P, f, \alpha) - L(P, f, \alpha) < \epsilon \tag{3.24}$$

where, of course, $\int_a^b f d\alpha \in \Big[L(P, f, \alpha), U(P, f, \alpha)\Big]$.

Let $P = \{a = x_0, x_1, \dots, x_n = b\}$. Because ϕ is a homeomorphism, we can let ϕ be strictly increasing WOLG.

Define a partition Q on $[A, B]$ by

$$Q = \phi^{-1}[P] = \{A = \phi^{-1}(x_0), \phi^{-1}(x_1), \dots, \phi^{-1}(x_n) = B\}$$

Now, because $\beta = \alpha \circ \phi$ and $g = f \circ \phi$ for all $x \in [a, b]$ by premise, and because ϕ is a homeomorphism, we have

$$\begin{aligned}
U(Q,g,\beta) &= \sum_{k=1}^{n} \Big[\sup_{t\in[\phi^{-1}(x_{k-1}),\phi^{-1}(x_k)]} g(t) \Big] \big[\beta(\phi^{-1}(x_k)) - \beta(\phi^{-1}(x_{k-1}))\big] \\
&= \sum_{k=1}^{n} \Big[\sup_{t\in[\phi^{-1}(x_{k-1}),\phi^{-1}(x_k)]} f\circ\phi(t) \Big] \big[\alpha\circ\phi\big(\phi^{-1}(x_k)\big) - \alpha\circ\phi\big(\phi^{-1}(x_{k-1})\big)\big] \\
&= \sum_{k=1}^{n} \Big[\sup_{t\in[x_{k-1},x_k]} f(t) \Big] \big(\alpha(x_k) - \alpha(x_{k-1})\big) = U(P,f,\alpha) \qquad (3.25)
\end{aligned}$$

Similarly, we can deduce $L(Q,g,\beta) = L(P,f,\alpha)$. Now, from Equation 3.25 and by definition of P (Equation 3.24), 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 3.4.2. (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

$$\begin{aligned}&\left|\overline{\int_a^b} f d\alpha - \overline{\int_a^b} f(x)\alpha'(x)dx\right|\\&\leq \left|\overline{\int_a^b} f d\alpha - U(P,f,\alpha)\right| - \left|U(P,f,\alpha) - U(P,f\alpha')\right| - \left|U(P,f\alpha') - \overline{\int_a^b} f(x)\alpha'(x)dx\right|\end{aligned}$$

We only wish

$$\begin{aligned}\text{to find } P \text{ such that} &\left|\overline{\int_a^b} f d\alpha - U(P,f,\alpha)\right| < \frac{\epsilon}{3}\\&\text{and } \left|U(P,f,\alpha) - U(P,f\alpha')\right| < \frac{\epsilon}{3} \text{ and } \left|\overline{\int_a^b} f(x)\alpha'(x)dx - U(P,f\alpha')\right| < \frac{\epsilon}{3}\end{aligned}$$

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

$$U(P,\alpha') - L(P,\alpha') < \frac{\epsilon}{4M} \tag{3.26}$$

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 3.26) 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 3.3.4), we know for all $k \in \{1,\dots,n\}$ there exists $t_k \in [x_{k-1}, x_k]$ such that

$$\Delta\alpha_k = \alpha'(t_k)\Delta x_k \tag{3.27}$$

Then, because $U(P,\alpha') - L(P,\alpha)' < \frac{\epsilon}{3M}$ (Equation 3.26), we now see

$$\sum_{k=1}^n |\alpha'(s_k) - \alpha'(t_k)|\,\Delta x_k < \frac{\epsilon}{3M} \text{ if } s_k \in [x_{k-1}, x_k] \text{ for all } k \in \{1,\dots,n\} \tag{3.28}$$

Then from Equation 3.27, definition of M and Equation 3.28, 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

$$\sum_{k=1}^{n} f(s_k)\Delta\alpha_k < U(P, f\alpha') + \frac{\epsilon}{4} \tag{3.29}$$

Because Equation 3.29 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 3.4.3. (Substitution Law) Given a function $\phi : [a,b] \to [A,B]$ and a function $f : [A,B] \to \mathbb{R}$ such that

(a) ϕ is a homoeomorphism.

(b) ϕ is differentiable on (a,b)

(c) $\int_a^b \phi'(x)dx$ exists.

(d) f is integrable on $[A,B]$

We have

$$\int_a^b f\big(\phi(x)\big)\phi'(x)dx = \int_A^B f(u)du$$

Proof. Because $f \circ \phi$ and ϕ' is integrable on $[a, b]$, by reduction of Riemann-Stieljes Integral (Tehroem 3.4.2), we know

$$\int_a^b \big(f \circ \phi\big)(x)\phi'(x)dx = \int_a^b \big(f \circ \phi\big)(x)d\phi$$

Let $\alpha(x) = x$. Let $\beta = \alpha \circ \phi$. Define $g = f \circ \phi$. By Change of Variable (Theorem 3.4.1), we now have

$$\int_a^b \big(f \circ \phi\big)(x)d\phi = \int_a^b g(x)d\beta = \int_A^B f(x)dx$$

■

3.5 FTC

Theorem 3.5.1. (Fundamental Theorem of Calculus: Part 1) Given a function $f : [a, b] \to \mathbb{R}$ proper-Riemann-Integrable on $[a, b]$, and let f be continuous at $x_0 \in [a, b]$. If we define $F : [a, b] \to \mathbb{R}$ by

$$F(x) = \int_a^x f(t)dt$$

Then

$$F \text{ is continuous on } [a, b] \text{ and is differentiable at } x_0 \text{ where } F'(x_0) = f(x_0)$$

Proof. Fix ϵ. 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

$$\exists \delta, |x - x_0| < \delta \implies |f(x) - f(x_0)| < \epsilon \tag{3.30}$$

We claim

$$\text{such } \delta \text{ in Equation 3.30 works}$$

WOLG, let $x > x_0$. Deduce

$$\begin{aligned}
\left|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\right| &= \left|\frac{\int_{x_0}^{x} f(t)dt}{x-x_0} - f(x_0)\right| \\
&= \left|\frac{\int_{x_0}^{x} \left[f(t)-f(x_0)\right]dt}{x-x_0}\right| \\
&\leq \frac{\int_{x_0}^{x} |f(t)-f(x_0)|\, dt}{|x-x_0|} \\
&\leq \frac{\int_{x_0}^{x} \epsilon dt}{|x-x_0|} = \epsilon \text{ (done)}
\end{aligned}$$

■

Theorem 3.5.2. (Fundamental Theorem of Calculus: Part 2, Leibniz Rule) Given two functions $f, F : [a,b] \to \mathbb{R}$ such that

(a) f is proper Riemann-Integrable on $[a,b]$

(b) $F'(x) = f(x)$ for all $x \in (a,b)$

(c) F is continuous on $[a,b]$

Then

$$\int_a^b f(x)dx = F(b) - F(a)$$

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

$$U(P,f) - L(P,f) < \epsilon \tag{3.31}$$

Because $f = F'$ on $[a,b]$, for each $k \in \{1,\dots,n\}$, by MVT (Theorem 3.3.4), 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 3.31, we can deduce

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| < \epsilon \text{ (done)}$$

■

With the

Theorem 3.5.3. (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

$$\int_a^b F(x)g(x)dx = FG\Big|_a^b - \int_a^b f(x)G(x)dx \tag{3.32}$$

Proof. To prove Equation 3.32, 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

$$\text{into proving } \int_a^b \big(Fg + fG\big)dx = FG\Big|_a^b$$

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 3.5.2). (done) ■

3.6 The Stone-Weierstrass Theorem

Theorem 3.6.1. (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

$$f(x) = (1+x)^r - (1+rx) \tag{3.33}$$

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 3.33), we see that

$$f(0) = 0, \text{ and } f(-1) = r - 1 \geq 0$$

Notice that by definition of f (Equation 3.33), $f(x)$ is clearly differentiable on $(-1, \infty)$.

Then, by MVT (Theorem 3.3.4), 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'

$$\begin{aligned} f'(x) &= r(1+x)^{r-1} - r \\ &= r\Big((1+x)^{r-1} - 1\Big) \end{aligned}$$

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) ∎

Theorem 3.6.2. (The Stone-Weierstrass Theorem: Real Version) Let $\mathcal{C}\big([a,b]\big)$ be the set of real continuous functions on $[a,b]$, and let $\mathbb{R}[x]\big|_{[a,b]}$ be the set of polynomials in x with real coefficients on $[a,b]$. Then

$$\mathbb{R}[x]\big|_{[a,b]} \text{ is dense in } \Big(\mathcal{C}\big([a,b]\big), \|\cdot\|_\infty\Big)$$

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

$$\tilde{f}(x) = f(x) - f(0) - x\big[f(1)-f(0)\big] \tag{3.34}$$

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 3.34), 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 3.34), we have

$$\tilde{f}(0) = 0 = \tilde{f}(1)$$

Then, we can expand the definition of $\tilde{f}$ by

$$\tilde{f}(x) = \begin{cases} \tilde{f}(x) & \text{if } x \in [0,1] \\ 0 & \text{if } x \notin [0,1] \end{cases} \tag{3.35}$$

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

$$Q_n = c_n(1-x^2)^n \text{ where } c_n \text{ is chosen to satisfy } \int_{-1}^{1} Q_n(x)dx = 1 \tag{3.36}$$

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 3.35), we see that

$$P_n(x) = \int_{-x}^{1-x} \tilde{f}(x+t)Q_n(t)dt \text{ for all } x \in [0,1] \tag{3.37}$$

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 3.36), 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

$$\forall x, y \in \mathbb{R}, |x-y| < \delta \implies \left| \tilde{f}(x) - \tilde{f}(y) \right| < \frac{\epsilon}{2} \tag{3.38}$$

By definition of $\tilde{f}$ (Equation 3.35), 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

$$4M\sqrt{n}(1-\delta^2)^n < \frac{\epsilon}{2} \tag{3.39}$$

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 3.6.1). Compute

$$\begin{aligned}
1 = \int_{-1}^{1} Q_n(x)dx &= c_n \int_{-1}^{1} (1-x^2)^n dx \\
&= 2c_n \int_0^1 (1-x^2)^n dx \\
&\geq 2c_n \int_0^{\frac{1}{\sqrt{n}}} (1-x^2)^n dx \\
&\geq 2c_n \int_0^{\frac{1}{\sqrt{n}}} 1 - nx^2 dx = c_n(\frac{4}{3\sqrt{n}}) > c_n(\frac{1}{\sqrt{n}})
\end{aligned}$$

This implies

$$\sqrt{n} > c_n \text{ (done)}$$

Because $\sqrt{n} > c_n$, by definition of Q_n (Equation 3.36), 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 3.36)

(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 3.39)

(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 3.39) we have

$$
\begin{aligned}
\left|P_n(x) - \tilde{f}(x)\right| &= \left|\int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt - \tilde{f}(x)\right| \\
&= \left|\int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt - \tilde{f}(x)\int_{-1}^{1} Q_n(t)dt\right| \\
&= \left|\int_{-1}^{1} \tilde{f}(x+t)Q_n(t)dt - \int_{-1}^{1} \tilde{f}(x)Q_n(t)dt\right| \\
&= \left|\int_{-1}^{1} \left[\tilde{f}(x+t) - \tilde{f}(x)\right]Q_n(t)dt\right| \\
&\leq \int_{-1}^{1} \left|\left[\tilde{f}(x+t) - \tilde{f}(x)\right]Q_n(t)\right| dt \\
&= \int_{-1}^{1} \left|\tilde{f}(x+t) - \tilde{f}(x)\right| Q_n(t)dt \\
&\leq \int_{-1}^{-\delta} 2MQ_n(t)dt + \int_{-\delta}^{\delta} \left|\tilde{f}(x+t) - \tilde{f}(x)\right| Q_n(t)dt + \int_{\delta}^{1} 2MQ_n(t)dt \\
&\leq 2M\Big(\int_{-1}^{-\delta} Q_n(t)dt + \int_{\delta}^{1} Q_n(t)dt\Big) + \int_{-\delta}^{\delta} (\frac{\epsilon}{2})Q_n(t)dt \\
&\leq 4M(1-\delta)\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} < \epsilon
\end{aligned}
$$

Because x is arbitrarily picked from $[0,1]$, we now have $\|P_n - \tilde{f}\|_\infty < \epsilon$ (done)

3.7 Closed under Uniform Convergence

Theorem 3.7.1. (Riemann-Integrable Functions are Closed under 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

$$\lim_{n\to\infty} \int_a^n f_n d\alpha \text{ exists and } \int_a^b f d\alpha = \lim_{n\to\infty} \int_a^b f_n d\alpha$$

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

$$\text{there exists } N \text{ such that } \forall n > N, \epsilon_n < \frac{\epsilon}{\alpha(b) - \alpha(a)} \tag{3.40}$$

We claim

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

Fix $n > N$. From Equation 3.40, we see

$$\begin{aligned} \left| \int_a^b f_n d\alpha - \int_a^b f d\alpha \right| &= \left| \int_a^b (f_n - f) d\alpha \right| \\ &\leq \int_a^b |f_n - f| \, d\alpha \\ &\leq \int_a^b \epsilon_n d\alpha = \epsilon_n \big[\alpha(b) - \alpha(a)\big] < \epsilon \text{ (done)} \end{aligned}$$

■

Theorem 3.7.2. (Uniform convergence and differentiation) Given a sequence of function $f_n : I \to \mathbb{R}$ such that

(a) $I \subseteq \mathbb{R}$ is bounded

(b) $f_n(x_0)$ converge as $n \to \infty$ for some $x_0 \in I$

(c) f_n are differentiable on I°

(d) f_n' converge uniformly on I°

Then there exists a function $f : I^\circ \to \mathbb{R}$ such that

$$f_n \to f \text{ uniformly and } f'(x) = \lim_{n\to\infty} f_n'(x) \text{ for all } x \in I^\circ$$

3.8 Taylor's Theorem