MA522 Lecture Notes

Instructor: Betsy Stovall   Note taken by: Yujia Bao

Definition. Let (X, d) be a metric space, let $K \subseteq X$, then

- An open cover of K (in X) is a set $\mathcal{G} = \{G_\alpha\}_{\alpha \in \mathcal{A}}$ where each G_α is an open subset of X and $K \subseteq \cup_{\alpha \in \mathcal{A}} G_\alpha$.
- K is compact if every open cover of K contains a finite subcover of K, i.e. if for every open cover $\mathcal{G} = \{G_\alpha\}_{\alpha \in \mathcal{A}}$, there $\exists\, \alpha_1, \ldots, \alpha_N \in \mathcal{A}$, s.t. $K \subseteq \cup_{j=1}^{N} G_{\alpha_j}$.

Example. Show that $(0, 1]$ is not compact.

Proof. Let $\mathcal{G} = \{(1/n,\, 2) : n \in \mathbb{N}\}$. Then $\mathcal{G}$ is an open cover of $(0, 1]$. But if $\{n_1, \ldots, n_N\}$ is any finite set, $\cup_{j=1}^{N}(1/n,\, 2) = (1/\max n_j,\, 2) \nsupseteq (0, 1]$. $\square$

Example. Show that $\mathbb{R}$ is not compact.

Proof. Let $\mathcal{G} = \{(-n, n) : n \in \mathbb{N}\}$. $\square$

Theorem 1 (Heine-Borel)**.** *A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Example. Show that $\{0\} \cup \{1/n : n \in \mathbb{N}\}$ is compact.

Proof. Let $\mathcal{G} = \{G_\alpha\}_{\alpha \in \mathcal{A}}$ be any open cover. Then $0 \in G_{\alpha_0}$ for some $\alpha_0 \in \mathcal{A}$. Since G_{α_0} is open, $\exists \epsilon > 0$ such that $B_\epsilon(0) \subseteq G_{\alpha_0}$. For $N = \lceil \epsilon \rceil$, we have $1/n \in B_\epsilon(0) \subseteq G_{\alpha_0}$ for all $n > N$. Choose α_n such that $1/n \in G_{\alpha_n}$ for each $n \leq N$. Thus $\{G_{\alpha_j}\}_{j=0}^{N}$ is a finite subcover of the origin set. So the origin set is compact. $\square$

Definition. $U \subseteq X$ is precompact if $\bar{U}$ is compact. (Here $\bar{U}$ stands for the closure of U.)

Example. By Theorem 1, every Borel subset of $\mathbb{R}^n$ is precompact.

Question: Definition of Borel set.

Definition. $K \subseteq X$ is sequentially compact if every sequence $\{x_n\}$ contains a subsequence $\{x_{n_k}\}$ that converges to a limit in K.

Definition. $U \subseteq X$ is totally bounded if $\forall \epsilon > 0$, U is covered by a finite collection of ϵ-balls, i.e., $\exists\, x_1, \ldots, x_{N_\epsilon} \in X$, s.t. $U \subseteq \cup_{j=1}^{N_\epsilon} B_\epsilon(x_j)$.

Example. Every bounded subset of $\mathbb{R}^n$ is totally bounded.

Example. In discrete metric space,

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

Then, every set is both open and closed. For $\epsilon < 1$, the ϵ Ball becomes a single ball. Infinite sets are bounded but not totally bounded.

Definition. $K \subseteq X$ is complete if every Cauchy sequence in K converges to some limit in K.

Definition. x is an accumulation point of $E \subseteq X$ if $\forall \epsilon > 0$, $(B_\epsilon(x) \setminus \{x\}) \cap E \neq \emptyset$. This is equivalent to say $\forall \epsilon > 0, B_\epsilon(x) \cap E$ contains infinitly many points.

Theorem 2. *Let (X, d) be a metric space, then TFAE*

1. *K is compact.*
2. *K has the Bolzano-Weierstrass property (Every infinity subset of K has an accumulation point in K).*
3. *K is sequentially compact.*
4. *K is complete and totally bounded.*

Proof. $1 \Rightarrow 2$**:** Assume K is compact. Let $E \subseteq K$ be an infinite set. If there doesn't exist such E, then Bolzano-Weierstrass property holds trivially. Now suppose E has no accumulation point in K. That means for every x in K, $\exists$ neighbour U_x of x (i.e. $x \in U_x$ and U_x is open), that contains no points of E other than (possibly) x itself. Since $K = \cup_{x \in K}\{x\} \subseteq \cup_{x \in K} U_x$, $\{U_x : x \in K\}$ is an open cover of K. By compactness, $\exists x_1, \ldots, x_N$, s.t. $K \subseteq \cup_{j=1}^N U_{x_j}$. But $\cup_{j=1}^N U_{x_j}$ contains at most N points of E. So it cannot contain all points of E because E is an infinity set. Contradiction! Since $E \subseteq K$, $\cup_{j=1}^N U_{x_j}$ should be a cover of E.

$2 \Rightarrow 3$**:** Assume K has the Bolzano-Weierstrass property. Let $\{x_n\}$ be a sequence in K. We need to show that it has a convergent subsequence. Let $E = \{x_n : n \in \mathbb{N}\}$.

Case I: E is a finite set. By the pigeonhole principle, $\exists x \in E \subseteq K$, s.t. $x_n = x$ for infinitely many n. Here $\{x_n\}$ has a constant subsequence which only takes the value x. This is a subsequnce converge to $x \in K$.

Case II: E is an infinite set. By Bolazano-Weierstrass propert, E has an accumulation point $x \in K$. Thus every ball centered at x contains infinitely many x_ns. Choose n_1 such that $x_{n_1} \in B_1(x)$; Choose $n_2 > n_1$ such that $x_{n_2} \in B_{1/2}(x)$; and so on. Proceeding by induction, we may find $n_1 < n_2 < \cdots < n_k < \cdots$, s.t. $x_{n_k} \in B_{1/k}(x)$ for all k. Then $\{x_{n_k}\}$ is a subsequent of $\{x_n\}$ and $x_{n_k} \to x$.

$3 \Rightarrow 4$**:** Assume K is sequentially compact.

Completeness: Let $\{x_n\}$ be a Cauchy sequence in K. By sequentially compactness, there exists a convergent subsequent $\{x_{n_k}\}$ of $\{x_n\}$, say $x_{n_k} \to x \in K$. Then we claim that $x_n \to x$. Since $x_{n_k} \to x$, let $\epsilon > 0$, $\exists M$ s.t. $\forall k \geq M$, $d(x_{n_k}, x) < \epsilon$. Since $\{x_n\}$ is a

Cauchy sequence, $\exists N$ s.t. $\forall n, m \geq N$, $d(x_n, x_m) < \epsilon$. Now fix $k_0 \geq \max\{M, N\}$ and let $n \geq N$. Then

$$d(x_n, x) \leq d(x_n, x_{n_{k_0}}) + d(x_{n_{k_0}}, x).$$

Since $k_0 \geq M$, we have $d(x_{n_{k_0}}, x) < \epsilon$. Since $n_{k_0} \geq k_0 \geq N$, we have $d(x_n, x_{n_{k_0}}) < \epsilon$. Then $d(x_n, x) < 2\epsilon$. So $\{x_n\}$ does converge.

Totally boundedness: Suppose not. Then $\exists \epsilon > 0$ s.t. K cannot be covered by a finite union of ϵ-balls. Thus, we may (inductively) construct a sequence $\{x_n\}$ in K such that $\forall n \geq 2$, $x_n \notin \cup_{j=1}^{n-1} B_\epsilon(x_j)$. Let $\{x_{n_k}\}$ be any subsequence of $\{x_n\}$. Pick any k_1, k_2 with $k_1 < k_2$. Then $x_{k_2} \notin B_\epsilon(x_{k_1})$ which means $d(x_{n_{k_1}}, x_{n_{k_2}}) \geq \epsilon$. So $\{x_{n_k}\}$ is not Cauchy and hence not convergent. This contradicts with sequential compactness of K.

$4 \Rightarrow 3$**:** Assume K is complete and totally bounded. Let $\{x_n\}$ be a sequence in K. We want to find a convergent subsequence. By totally boundedness, K is covered by a finite number of 1-balls, $K \subseteq \cup_{j=1}^N B_1(y_j)$. By the pigenhole principle, there must exist an $B_1(y_j)$ which contains x_n for infinitely many n. Denote that y_j as z_1. So there exists a subsequence $\{x_{n_k^1}\}$ contained in $B_1(z_1)$. By the same argument, $\exists z_2$ s.t. $B_{1/2}(z_2)$ contains a subsequent $\{x_{n_k^2}\}$ of $\{x_{n_k^1}\}$, and so on. So for each $m \in \mathbb{N}$, we find z_m and a subsequent $\{x_{n_k^m}\}$ of $\{x_{n_k^{m-1}}\}$ s.t. $\{x_{n_k^m}\}$ is contained in $B_{1/m}(z_m)$. Now we define $x_{n_k} = x_{n_k^k}$ (diagonalization).

Claim 1: $\{x_{n_k}\}$ is a subsequence of $\{x_n\}$.

This is because $n_k = n_k^k \geq n_k^{k-1} > n_{k-1}^{k-1} = n_{k-1}$, where the first inequality comes from the fact that $\{x_{n_j^k}\}$ is a subsequence of $\{x_{n_j^{k-1}}\}$.

Claim 2: $\{x_{n_k}\}$ is a Cauchy sequence.

For $k_1, k_2 \geq M$, $x_{n_{k_1}}$ and $x_{n_{k_2}}$ are both terms in the sequence $\{x_{n_j^M}\}$. So both of them lie in $B_{1/M}(z_M)$ which means $d(x_{n_{k_1}}, x_{n_{k_2}}) \leq 2/M$.

By completeness of K, $\{x_{n_k}\}$ converges in K.

$4 \Rightarrow 2$**:** Assume K is complete and totally bounded. Let $E \subseteq K$ be an infinite subset. Since K is totally bounded, K can be covered by finitely many 1-balls. By pigeonhole principle, $\exists x_1 \in K$, s.t. $B_1(x_1) \cap E =: E_1$ is an infinite set. By induction, for each $n \in \mathbb{N}^+$, $\exists x_n \in K$, s.t. $B_{1/n}(x_n) \cap E_{n-1} =: E_n$ is an infinite set.

Claim 1: $\{x_n\}$ is a Cauchy sequence.

Notice that $\forall n, m$,

$$B_{1/n}(x_n) \cap B_{1/m}(x_m) \supseteq E_{\max\{m,n\}} \neq \emptyset.$$

This implies

$$d(x_n, x_m) < \frac{1}{n} + \frac{1}{m} \leq \frac{2}{\min\{n, m\}}.$$

Then as long as $n, m > 2/\epsilon$, we have $d(x_n, x_m) < \epsilon$. So $\{x_n\}$ is a Cauchy sequence.

By completeness of K, $\{x_n\}$ converges, say $x_n \to x_0$.

Claim 2: x_0 is an accumulation point for E.

Let $\epsilon > 0$. Choose N sufficiently large, such that $\forall n \geq N$, $d(x_0, x_n) < \epsilon/2$ and $1/n < \epsilon/2$. For any $y \in B_{1/n}(x_n)$, we have $d(y, x_n) < 1/n$. Then

$$d(x_0, y) \leq d(x_0, x_n) + d(x_n, y) < \frac{\epsilon}{2} + \frac{1}{n} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

So we have $y \in B_\epsilon(x_0)$. Then $B_\epsilon(x_0) \supseteq B_{1/n}(x_n) \supseteq E_n$. Since E_n is an infinite subset of E and ϵ is selected arbitrary, x_0 is an accumulation point of E.

$4, 3 \Rightarrow 1$**:** Assume K is complete, totally bounded and sequentially compact. Let $\mathcal{G} \coloneqq \{G_\alpha\}_{\alpha \in \mathcal{A}}$ be an open cover of K. Then $\forall x \in K$, $\exists \alpha \in \mathcal{A}$, s.t. $x \in G_\alpha$. Since G_α is open, $\exists r > 0$, s.t. $B_r(x) \subseteq G_\alpha$. Thus we define

$$\epsilon(x) \coloneqq \sup\{r > 0 \,:\, \exists \alpha \in \mathcal{A}, \text{ s.t. } B_r(x) \subseteq G_\alpha\}$$

By definition , $\epsilon(x) > 0$. If $\epsilon(x) = +\infty$ for some $x \in K$, then since K is bounded (because K is totally bounded), there must be one G_α containing all of K. Now we assume $\epsilon(x)$ is finite for all $x \in K$.

Claim: $\exists \epsilon_0 > 0$, s.t. $\epsilon(x) \geq \epsilon_0$, $\forall x \in K$.

Suppose such ϵ_0 doesn't exist. Then $\exists \{x_n\}$, s.t. $\epsilon(x_n) \to 0$. By sequential compactness of K, there exists a subsequence $x_{n_k} \to x_0 \in K$. As $\epsilon(x_0) > 0$, we can choose some $r > 0$ and α s.t. $B_r(x_0) \subseteq G_\alpha$. But for k sufficiently large, $\epsilon(x_{n_k}) < r/2$ and $d(x_0, x_{n_k}) < r/2$. Then

$$B_{r/2}(x_{n_k}) \subseteq B_r(x_0) \subseteq G_\alpha,$$

and this implies $r/2 \leq \epsilon(x_{n_k}) < r/2$ which leads to contradiction.

Since K is totally bounded, so $\exists x_1, \ldots, x_N$, s.t.

$$K \subseteq \bigcup_{j=1}^{N} B_{\epsilon_0/2}(x_j) \subseteq \bigcup_{j=1}^{N} B_{\epsilon(x_j)/2}(x_j)$$

Furthermore, as $\epsilon(x_j)/2 < \epsilon(x_j)$ (because $\epsilon(x_j) > 0$), $\exists r > 0$, s.t. $r > \epsilon(x_j)/2$ and $B_r(x_j) \subseteq G_{\alpha_j}$ for some $\alpha_j \in \mathcal{A}$. Finally, $\cup_{j=1}^{N} G_{\alpha_j}$ is a finite subcover of K. Thus K is compact.

□

Corollary. *A subset of a complete metric space is compact if and only if it is closed and totally bounded.*

Corollary. *A subset of a complete metric space is precompact if and only if it is totally bounded.*

Example. A closed subset of a compact set is compact.

Proof. Suppose $K \subseteq E$. Since E is compact, by Bolzano-Weierstrass property, every infinite subset of K has an accumulation point. Since K is closed, K contains all its accumulation points. So K also satisfies B-W property which means K is compact. □

Definition. A subset E of a metric space X is not connected if $\exists$ sets A, B s.t.

$$\begin{cases} E = A \cup B, \ A \neq \emptyset, \ B \neq \emptyset \\ \bar{A} \cap B = \emptyset, \ A \cap \bar{B} = \emptyset \end{cases}$$

We say A and B form a separation of E, or A and B separate E.

Definition. E is connected if it is not disconnected.

Theorem 3. *Let E be a subset of X. Then TFAE:*

1. *$E \subseteq X$ is not connected.*
2. *$\exists$ open sets $U, V \subseteq X$, s.t.*
$$E \subseteq U \cup V, \ E \cap U \neq \emptyset, \ E \cap V \neq \emptyset, \ U \cap V = \emptyset.$$
3. *$\exists A \subseteq E$, s.t.*
$$A \neq \emptyset, \ A \neq E, \ A = E \cap F = E \cap G,$$
where F is closed and G is open.

Proof. $2 \Rightarrow 3$**:** Assume 2 holds. We pick $A = E \cap U$. Then $A \neq \emptyset$. Let $G = U$. Then $A = E \cap G$. Let $F = V^c$. Then F is closed since V is open.

Claim 1: $A = E \cap F$.
Let $a \in A$. Then $a \in E$ and $a \in U \subseteq V^c = F$. So $a \in E \cap F$. Let $x \in E \cap F$. Then $x \in E$ and $x \notin V$. Since $E \subseteq U \cup V$, $x \in U$. So $x \in E \cap U = A$.

Claim 2: $A \neq E$.
This is true because $E \setminus A = E \setminus (E \cap F) = E \cap V$ and we know $E \cap V \neq \emptyset$ by 2.

$3 \Rightarrow 1$**:** Assume 3 holds. We define A as in 3 and let $B = E \setminus A$. Then by definition, $A \cup B = E$, $A \neq \emptyset$ and $B \neq \emptyset$ (since $A \neq E$). Note that

$$\bar{A} \cap B = \overline{E \cap F} \cap (E \cap F^c) \subseteq \bar{F} \cap F^c = F \cap F^c = \emptyset,$$

since F is closed. Also

$$A \cap \bar{B} = E \cap G \cap \overline{E \cap G^c} \subseteq G \cap \overline{G^c} = G \cap G^c = \emptyset,$$

since G^c is closed. Then E is not connected.

$1 \Rightarrow 2$**:** Assume 1 holds. Let $a \in A$. Then $a \notin \bar{B}$, since $A \cap \bar{B} = \emptyset$. So $\exists r(a) > 0$, s.t. $B_{r(a)}(a) \cap B = \emptyset$. Likewise, if $b \in B$, $\exists r(b) > 0$, s.t. $B_{r(b)}(b) \cap A = \emptyset$. Now define $U = \cup_{a \in A} B_{r(a)/2}(a)$ and $V = \cup_{b \in B} B_{r(b)/2}(b)$. Then U and V are open sets since they are unions of open sets. Also $E \subseteq A \cup B \subseteq U \cup V$ and $E \cap U \supseteq A \neq \emptyset$, $E \cap V \supseteq B \neq \emptyset$. Now we just need to show $U \cap V = \emptyset$. If it is not true, $\exists x \in U \cap V$. By our construction of U and V, $\exists a \in A$, $b \in B$, s.t. $x \in B_{r(a)/2}(a) \cap B_{r(b)/2}(b)$. So

$$d(a, b) \leq d(a, x) + d(b, x) < \frac{r(a)}{2} + \frac{r(b)}{2} \leq \max\{r(a), r(b)\}$$

Then if $r(a) \geq r(b)$, we have $b \in B_{r(a)}(a)$. If $r(b) \geq r(a)$, we have $a \in B_{r(b)}(b)$. Both of them contradicts with our definition of $r(a)$ or $r(b)$. So $U \cap V = \emptyset$.

$\square$

Example. $(-\infty, 0) \cup (0, +\infty)$ is not connected.

Example. $E \coloneqq \{(x, y) : y \in [-1, 1] \text{ with } x = 0 \text{ or } y = \sin(1/x) \text{ with } x > 0\}$ is connected.

Proposition. *If E is connected and $f : E \to Y$ is continuous, then $f(E)$ is connected.*

Proof. Assume $f(E)$ is not connected. Then $\exists U, V$ open with $f(E) \subseteq U \cup V$, $U \cap V = \emptyset$, $f(E) \cap U \neq \emptyset$, $f(E) \cap V \neq \emptyset$. Note that $f^{-1}(U)$ is open, nonempty. Likeliwise for $f^{-1}(V)$. Moreover $f^{-1}(U) \cap f^{-1}(V) = \emptyset$ (If $\exists x \in f^{-1}(U) \cap f^{-1}(V)$, then $f(x) \in U \cap V = \emptyset$. Contradiction!) and $f^{-1}(U) \cup f^{-1}(V) \supseteq E$. By Theorem 3, E is not connected. Contradiction! □

Definition. Let (X, D) be a metric space. (Assume $X \neq \emptyset$.) Say $\Phi : X \to X$ is a contraction if $\exists r < 1$, s.t. $\forall x, y \in X$,

$$d(\Phi(x),\ \Phi(y)) \leq r \cdot d(x, y)$$

Example. By mean value theorem, if $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $\exists r < 1$, s.t. $\forall x$, $|f'(x)| \leq r$. Then f is a contraction.

Proposition. *If Φ is a contraction, then Φ is continuous.*

Theorem 4 (Contraction Mapping Theorem). *If (X, d) is a nonempty complete metric space and $\Phi : X \to X$ is a contraction. Then Φ has a unique fixed point, i.e. $\exists! x_0 \in X$, s.t. $\Phi(x_0) = x_0$.*

Proof. Let Φ be a contraction with shrinking constant $r < 1$.

Uniqueness: Suppose x_1, x_2 are both fixed points of the contraction Φ. Then

$$d(x_1, x_2) = d(\Phi(x_1), \Phi(x_2)) \leq r \cdot d(x_1, x_2).$$

Thus $d(x_1, x_2) = 0$ and $x_1 = x_2$, since $r < 1$.

Existence: Let $x \in X$. Define a sequence $\{x_n\}$ inductively by setting $x_0 = x$ and $x_n = \Phi(x_{n-1})$ for $n \in \mathbb{N}^+$.

Claim 1: $\{x_n\}$ is a Cauchy sequence and hence convergent.
Suppose $m, n \in \mathbb{N}^+$, $m > n$. Then

$$d(x_m, x_n) = d(\Phi(x_{m-1}), \Phi(x_{n-1})) \leq r d(x_{m-1}, x_{n-1}).$$

By induction,

$$\begin{aligned}
d(x_m, x_n) &\leq r^n d(x_{m-n}, x_0) \\
&\leq r^n \left(d(x_{m-n}, x_{m-n-1}) + d(x_{m-n-1}, x_{m-n-2}) + \cdots + d(x_1, x_0)\right) \\
&\leq r^n \left(r^{m-n-1} d(x_1, x_0) + r^{m-n-2} d(x_1, x_0) + \cdots + d(x_1, x_0)\right) \\
&\leq r^n d(x_1, x_0) \sum_{j=0}^{\infty} r^j \\
&= \frac{r^n}{1 - r} d(x_1, x_0)
\end{aligned}$$

Given any $\epsilon > 0$, choose N such that $r^N d(x_1, x_0)/(1 - r) < \epsilon$. Then by the preceeding calculation, $\forall n, m \geq N$, $d(x_n, x_m) < \epsilon$.

Claim 2: $\lim_{n\to\infty} x_n$ is a fixed point of Φ.

Since Φ is a contraction, Φ is continuous. So we can exchange Φ with the limit operation. Then

$$\Phi(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} \Phi(x_n) = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n.$$

□

Corollary. *If* $\exists r < 1$, *s.t.* $|f'(x)| \le r$, $\forall x \in \mathbb{R}$. *Then* $\exists! x_0$, *s.t.* $f(x_0) = x_0$.

Example. $\Phi(x) = x + 1$ on $\mathbb{R}$ is not a contraction.

Example. $X = (-\infty, 0) \cup (0, \infty)$ is not complete. For $\Phi(x) = \frac{1}{2}x$, it is a contraction but the fixed point $x_0 = 0$ is not in X.

Definition. Let $E \subseteq X$.

- E is dense in X if $\bar{E} = X$.
- E is nowhere dense in X if $\bar{E}$ has empty interior.

Example. $\mathbb{Z} \subseteq \mathbb{R}$ is nowhere dense in $\mathbb{R}$. $\mathbb{Q} \cap (-2, 2)$ is neither dense nor nowhere dense in $\mathbb{R}$.

Definition. Let $E \subseteq X$, where X is some metric space.

- E is meager if E can be written as a countable union of nowhere dense sets.
- E is generic if E^c is meager.

Theorem 5 (Baire Category Theorem)**.** *A nonempty complete metric space cannot be written as a countable union of nowhere dense sets.*

This is equivalent to say that a complete metric space cannot be meager and is also equivalent to say that a subset of a complete metric space cannot be both meager and generic. In particular, generic subsets of complete metric space are nonempty.

Proof. Let (X, d) be a complete metric space and $\{F_n\}$ be a collection of nowhere dense subsets of X. Suppose for contradiction that $X = \cup_n F_n$. Then $X = \cup_n \bar{F}_n$ and $\bar{F}_n$ are also nowhere dense. Without loss of generality, we assume each F_n is closed.
Since $X \nsubseteq F_i$ for any i (Otherwise, $\overline{\text{interior}(F_i)} = X \neq \emptyset$.), $\exists x_1 \in F_1^c$ and $\exists r_1 > 0$, s.t. $\bar{B}_{r_1}(x_1) \subseteq F_1^c$. Since $\text{interior}(F_2) = \emptyset$, $B_{r_1}(x_1) \nsubseteq F_2$. Then $\exists x_2$, s.t. $x_2 \in F_2^c \cap B_{r_1}(x_1)$. Since F_2^c and $B_{r_1}(x_1)$ are both open, their intersection is open. $\exists r_2 > 0$, s.t. $\bar{B}_{r_2}(x_2) \subseteq F_2^c \cap B_{r_1}(x_1)$ and $r_2 \le r_1/2$. By induction, we obtain a sequence $\{x_n\}$ in X and $\{r_n\}$ in $(0, +\infty)$, s.t. for all n,

$$B_{r_n}(x_n) \subseteq F_n^c \cap B_{r_{n-1}}(x_{n-1}), \quad r_n \le \frac{1}{2} r_{n-1}.$$

So $r_n \to 0$.

Claim 1: $\{x_n\}$ is a Cauchy sequence.

For $n, m \ge M$, $x_n, x_m \in B_{r_M}(x_M)$. So $d(x_n, x_m) < r_M$ and $r_M \to 0$.

Claim 2: $x_\infty \coloneqq \lim_{x\to\infty} x_n \notin \cup_n F_n$.

Since $\{x_n\}_{n\geq M}$ is a sequence in $B_{r_M}(x_M)$, $x_\infty \in \overline{B_{r_M}(x_M)} \subseteq F_M^c$. As M was arbitrary,

$$x_\infty \in \bigcap_M F_M^c = \left(\bigcup_M F_M\right)^c.$$

Then we have $x_\infty \notin \cup_n F_n = X$. However, X is a complete metric. Contradiction. ☐

Proposition. *$\mathbb{R}^n$ cannot be written as a countable union of hyperplanes.*

Proof. Let P be any hyperplane. So $P = \{x \in \mathbb{R}^n : \langle x, a\rangle = d\}$ for fixed a, d, where $a \neq 0$.

Claim 1: Hyperplane P is closed.

P can also be defined by $f^{-1}(d)$ where $f(x) = \langle x, a\rangle$. Since d is closed, f is continuous, the preimage f^{-1} of a closed set d is also closed.

Claim 2: Hyperplane P has empty interior.

Let $x_0 \in P$. For any $r > 0$, the point $x_0 + (r/2|a|)\cdot a$ is inside $B_r(x_0)$. However,

$$\langle x_0 + \frac{ra}{2|a|}, a\rangle = d + \frac{r|a|}{2} \neq d.$$

So P has empty interior.

Then by definition, all hyperplanes are nowhere dense in $\mathbb{R}^n$. So $\mathbb{R}^n$ cannot be written as a countable union of hyperplanes. ☐

Proposition (Well-approximable numbers)**.** *Let*

$$\Lambda_n = \left\{x \in \mathbb{R} : \left|x - \frac{p}{q}\right| < \frac{1}{q^n} \text{ for infinitely many } \frac{p}{q} \in \mathbb{Q}\right\}.$$

Λ_n is generic. Thus by Theorem 5, $\exists$ well-approximable irrationals, since $\mathbb{Q}$ is meager (Every countable set is meager).

Proof. By definition

$$\Lambda_n^c = \left\{x \in \mathbb{R} : \left|x - \frac{p}{q}\right| \geq \frac{1}{q^n} \text{ for all but finitely many } \frac{p}{q} \in \mathbb{Q}\right\}.$$

If we can show that Λ_n^c is meager, then since R is a complete metric space, Λ_n is generic. Now define

$$F_q \coloneqq \left\{x \in \mathbb{R} : \forall p \in \mathbb{Z}, \left|x - \frac{p}{q}\right| \geq \frac{1}{q^n}\right\}, \quad E_q \coloneqq \bigcap_{q'\geq q} F_{q'}$$

Since $F_{q'}$ is closed for all q', E_q is also closed for all q. Then

$$\Lambda_n^c = \bigcup_{q\in\mathbb{N}} E_q = \bigcup_{q\in\mathbb{N}} \bigcap_{q'\geq q} \left\{x \in \mathbb{R} : \forall p \in \mathbb{Z}, \left|x - \frac{p}{q'}\right| \geq \frac{1}{(q')^n}\right\}.$$

Now we need to show that E_q is nowhere dense in $\mathbb{R}$. We know $\bar{E}_q = E_q$. But $E_q \cap \{p/q' : p, q' \in \mathbb{Z}, q' > q\} = \emptyset$. Since the latter set is a dense set in $\mathbb{R}$, then E_q^c countains a dense set which implies the interier of E_q is empty. ☐

Furthermore, we can express Λ_n in another way,

$$\begin{aligned}\Lambda_n &= (\Lambda_n^c)^c\\ &= \bigcap_{q\in\mathbb{N}}\bigcup_{q'\geq q}\left\{x\in\mathbb{R} : \exists p\in\mathbb{Z}, \left|x-\frac{p}{q'}\right| < \frac{1}{(q')^n}\right\}\\ &= \bigcap_{q\in\mathbb{N}}\bigcup_{q'\geq q}\bigcup_{p\in\mathbb{Z}}\left(\frac{p}{q'}-\frac{1}{(q')^n}, \frac{p}{q'}+\frac{1}{(q')^n}\right)\end{aligned}$$

The inner part is the union of intervals of width $2/(q')^n$ with spacing $1/q$. So heuristically,

$$\begin{aligned}\Pr\left(x\in\bigcup_{p\in\mathbb{Z}}\left(\frac{p}{q'}-\frac{1}{(q')^n}, \frac{p}{q'}+\frac{1}{(q')^n}\right)\right) &\leq \frac{2/(q')^n}{1/q} = \frac{2}{(q')^{n-1}},\\ \Pr\left(x\in\bigcup_{q'\geq q}\bigcup_{p\in\mathbb{Z}}\left(\frac{p}{q'}-\frac{1}{(q')^n}, \frac{p}{q'}+\frac{1}{(q')^n}\right)\right) &\leq \sum_{q'\geq q}^{\infty}\frac{2}{(q')^{n-1}}\\ &\leq \frac{2}{q^{n-\alpha}}\sum_{q'\geq q}^{\infty}\frac{(q')^{n-\alpha}}{(q')^{n-1}}\\ &\leq \frac{2}{q^{n-\alpha}}\sum_{q'\geq q}^{\infty}\frac{1}{(q')^{\alpha-1}}\\ &\leq \frac{C_n}{q^{n-\alpha}},\end{aligned}$$

as long as $n > 2$ (α is also greater than 2). So as $n\to+\infty$, the probability that $x\in\Lambda_n$ goes to zero.

Definition. Let X be a nonempty set and (Y, d_y) be a metric space. Let $\{f_n\}$ be a sequence of functions from X to Y, and let f be a function from X to Y.

- Say $f_n\to f$ pointwise if $\forall x\in X$ and $\forall\epsilon>0$, $\exists N=N(\epsilon,x)$, s.t. $\forall n\geq N$, $d_y(f_n(x), f(x)) < \epsilon$.
- Say $\{f_n\}$ is pointwise Cauchy if $\forall x\in X$, $\forall\epsilon>0$, $\exists N=N(\epsilon,x)$, s.t. $\forall n,m>N$, $d_y(f_n(x), f_m(x)) < \epsilon$.
- Say $f_n\to f$ uniformly if $\forall\epsilon>0$, $\exists N=N(\epsilon)$, s.t. $\forall n>N$, $\forall x\in X$, $d_y(f_n(x)-f(x)) < \epsilon$.
- Say $\{f_n\}$ is uniformly Cauchy if $\forall\epsilon>0$, $\exists N=N(\epsilon)$, s.t. $\forall n,m>N$, $\forall x\in X$, $d_y(f_n(x), f_m(x)) < \epsilon$. That is to say $\lim_{n\to\infty}\sup_{x\in X} d(f_n(x), f(x)) = 0$.

Note. Uniform convergence is much better than pointwise convergence.

Definition. $f : X\to Y$ is a bounded function if $f(X)$ is a bounded subset of Y. (i.e. if $f(x)$ is contained in some metric ball $B_{r_f}(y_f)$ in Y)

Example. The pointwise limit of a sequence of bounded functions need not be bounded. For example,

$$f_n(x) = \begin{cases} x & \text{for } |x|\leq N\\ N & \text{for } x\geq N\\ -N & \text{for } x\leq -N.\end{cases}$$

Each f_n is bounded in $[-N, N]$. However, its pointwise limit is $f(x) = x$, which is unbounded.

Definition. Let X, Y be two nonempty metric spaces. Let

$$\mathcal{B}(X,Y) \coloneqq \{f : X \to Y \ : \ f(X) \text{ is a bounded set}\}.$$

Proposition. *The uniform limit of a sequence of bounded function is bounded.*

Proof. Let $\{f_n\}$ be asequence in $\mathcal{B}(X,Y)$ and assume that $f_n \to f$ uniformly. By uniform convergence, $\exists n \in \mathbb{N}$, s.t. $\forall x \in X$, $d_Y(f(x), f_N(x)) < 1$. Since f_N is a bounded function, $\exists y_0, k$, s.t. $f_N(x) \in B_k(y_0)$ for every x in X. So $f(x) \in B_{k+1}(y_0)$ for every x in X. Thus f is bounded. □

Proposition. *$\mathcal{B}(X,Y)$ is a metric sapce with metric $d_{\mathcal{B}}(f,g) = \sup_{x\in X} d(f(x), g(x))$.*

1. *If f, g are bounded functions, then $d_{\mathcal{B}}(f,g)$ is finite.*
2. *$d_{\mathcal{B}}$ is a metric on $\mathcal{B}(X,Y)$.*
3. *Uniform convergent of a sequence in $\mathcal{B}(X,Y)$ is equivalent to metric convergence with respect to $d_{\mathcal{B}}$.*
4. *If Y is complete, then so is $\mathcal{B}(X,Y)$.*

Definition. Let X, Y be two nonempty metric spaces. Let

$$\mathcal{C}(X,Y) \coloneqq \{\text{continuous functions from } X \text{ to } Y\}.$$

Example. The pointwise limit of a sequence of continuous functions need not to be continuous. Let $f_n : [0,1] \to \mathbb{R}$, $f_n(x) = x^n$. But

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & \text{if } x \in [0,1) \\ 1, & \text{if } x = 1 \end{cases}$$

is not continuous.

Proposition (Honors HW)**.** *If $\{f_n\}$ is a sequence of functions on $\mathbb{R}$ and $f_n \to f$ pointwise, then the set of continuity points for f is generic.*

Theorem 6. *If $\{f_n\}$ is a sequence in $\mathcal{C}(X,Y)$ and $f_n \to f$ uniformly. Then $f \in \mathcal{C}(X,Y)$.*

Proof. Assume $f_n \to f$ uniformly. Let $x_0 \to X$ and $\epsilon > 0$. By uniform convergence, $\exists N \in \mathbb{N}$, s.t. $\forall x$, $d_Y(f(x), f_N(x)) < \epsilon$. Since f_N is continuous, $\exists \delta > 0$, s.t. $\forall x$ with $d(x, x_0) < \delta$, we have $d_Y(f_N(x), f_N(x_0)) < \epsilon$. Finally, by triangle inequality, $\forall x$ with $d_x(x, x_0) < \delta$,

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(x_0)) + d_Y(f_N(x_0), f(x_0)) < 3\epsilon.$$

□

Definition. Let X, Y be two nonempty metric spaces. Define

$$\mathcal{C}^0(X,Y) \coloneqq \mathcal{C}(X,Y) \cap \mathcal{B}(X,Y).$$

Then $\mathcal{C}^0(X,Y)$ is a metric space with metric $d_{\mathcal{C}^0}(f,g) \coloneqq d_{\mathcal{B}}(f,g)$.

Definition. For $Y = \mathbb{R}$, X being any metric space, let $\mathcal{C}^0(X) \coloneqq \mathcal{C}^0(X, \mathbb{R})$ and define the norm (Question: why it is a norm?)

$$\|f\|_{\mathcal{C}^0(X)} = \sup_{x \in X} |f(x)|.$$

Proposition. *Let X, Y be two metric spaces. $(\mathcal{C}^0(X, Y), d_{\mathcal{C}^0})$ is a metric space which is complete if Y is complete.*

Proof. Let $\{f_n\}$ be a Cauchy sequence in $\mathcal{C}^0(X, Y)$, then $\{f_n\}$ is uniformly Cauchy in $\mathcal{C}^0(X, Y)$. Thus $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, s.t. $\forall n, m \geq N$, $d_{\mathcal{C}^0}(f_n, f_m) < \epsilon$. Notice that $d_{\mathcal{C}^0}(f_n, f_m) = \sup_{x \in X} d_Y(f_n(x), f_m(x))$. So in paticular, for any $x \in X$, $d_Y(f_n(x), f_m(x)) < \epsilon$. Now fix $x \in X$. Then $\{f_n(x)\}$ is a Cauchy in Y. As Y is complete, $\forall x \in X$, $\exists f(x) \coloneqq \lim_{n \to \infty} f_n(x)$.

Claim: $f_n \to f$ uniformly.

Let $\epsilon > 0$. By uniform Cauchyness, we may choose $N \in \mathbb{N}$, s.t. $\forall n, m \geq N$, $\forall x \in X$, $d_Y(f_n(x), f_m(x)) < \epsilon$. Now fix $n \geq N$, $x \in X$. Choose m_x s.t. $m_x \geq N$ and $d_Y(f_{m_x}(x), f(x)) < \epsilon$. Then

$$d_Y(f_n(x), f(x)) \leq d_Y(f_n(x), f_{m_x}(x)) + d_Y(f_{m_x}(x), f(x)) < 2\epsilon.$$

Note, it is okay that m_x depends on x, since it doesn't appear on either side of the inequality. Since $d_Y(f_n(x), f(x)) < 2\epsilon$ for all $x \in X$,

$$d_{\mathcal{C}^0}(f_n(x), f) \leq 2\epsilon < 3\epsilon.$$

Thus, $f_n \to f$ in $\mathcal{C}^0$.

□

Theorem 7. *There exists a nowhere differentiable (not differentiable at any point) continuous function $f \in \mathcal{C}^0([0,1])$.*

Proof. Since $\mathbb{R}$ is complete, $\mathcal{C}^0([0,1])$ is a complete metric space and thus it is not meager. It sufficies to prove that

$$F \coloneqq \{f \in \mathcal{C}^0([0,1]) \,:\, \exists x_0, \text{ s.t. } f'(x_0) \text{ exists}\}$$

is meager, i.e. a countable union with nowhere dense sets. It then suffices to prove that F is contained in a countable union of nowhere dense sets.

Claim 1: $F \subseteq \cup_{n=1}^{\infty} F_n$, where

$$F_n \coloneqq \{f \in \mathcal{C}^0([0,1]) \,:\, \exists x_0 \in [0,1], \text{ s.t. } |f(x) - f(x_0)| \leq n|x - x_0|, \forall x \in [0,1]\}.$$

If $f \in F$, then $\exists x_0$, s.t.

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. Thus, there exists $\delta > 0$, s.t. $\forall x \in [0,1]$ with $|x - x_0| < \delta$,

$$|f(x) - f(x_0)| \leq (|f'(x_0)| + 1)|x - x_0|.$$

For $x \in [0,1]$ with $|x - x_1| \geq \delta$, we have

$$|f(x) - f(x_0)| \leq |f(x)| + |f(x_0)| \leq 2\|f\|_{\mathcal{C}^0} \cdot \frac{|x - x_0|}{|x - x_0|} \leq \frac{2\|f\|_{\mathcal{C}^0}}{\delta} \cdot |x - x_0|.$$

Finally, for any $n \geq |f'(x_0)| + 1 + 2\|f\|_{\mathcal{C}^0}/\delta$, we have $f \in F_n$.

Claim 2: F_N is closed.

Claim 3: F_N is nowhere dense.

Since F_N is closed, it suffices to prove that F_N has an empty interior, i.e. that $\forall f \in F_N$, $\forall \epsilon > 0$, $\exists g \in \mathcal{C}^0([0,1])$, s.t. $\|f-g\|_{\mathcal{C}^0([0,1])} < \epsilon$ and $g \notin F_N$. Let $f \in F_N$ and $\epsilon > 0$. The idea is to find g, piecewise linear, such that the slopes of linear parts of g has absolute value $> N$.

Since f is continuous and $[0,1]$ is compact, f is uniformly continuous. So there exists $\delta > 0$, s.t. $|x-y| \leq \delta$, $|f(x)-f(y)| \leq \epsilon$. Choose $n \in \mathbb{N}$, s.t. $1/n < \delta$ and $2\epsilon/(1/n) > 1000N$. Set $x_j = j/n$, $0 \leq j \leq n$. Define

$$g(x_j) \coloneqq f(x_j) + (-1)^j \epsilon,$$

and make g linear in $x \in (x_j, x_{j+1})$, for $j = 0, \ldots, n$. Note, if $x \in [x_j, x_{j+1}]$, then $x = (1-\theta)x_j + \theta x_{j+1}$ for some $0 \leq \theta \leq 1$, and this implies $g(x) = (1-\theta)g(x_j) + \theta g(x_{j+1})$.

Subclaim 1: $\|g-f\|_{\mathcal{C}^0([0,1])} < 3\epsilon$.

Suffices to show that $\forall j$ and $\forall x \in [x_j, x_{j+1}]$, $|g(x)-f(x)| < 3\epsilon$. Write $x = (1-\theta)x_j + \theta x_{j+1}$, with $0 \leq \theta \leq 1$. Then

$$\begin{aligned}
|g(x)-f(x)| &= |(1-\theta)(g(x_j)-f(x_j)) + (1-\theta)(f(x_j)-f(x)) \\
&\qquad + \theta(g(x_{j+1})-f(x_j)) + \theta(f(x_j)-f(x))| \\
&\leq (1-\theta)|g(x_j)-f(x_j)| + (1-\theta)|f(x_j)-f(x)| \\
&\qquad + \theta|g(x_{j+1})-f(x_j)| + \theta|f(x_j)-f(x)| \\
&\leq (1-\theta)\epsilon + (1-\theta)\epsilon + \theta\epsilon + \theta\epsilon \\
&= 2\epsilon < 3\epsilon
\end{aligned}$$

Claim 2: The slopes of g have absolute value greater than N:

Suffices to prove

$$\frac{|g(x_{j+1})-g(x_j)|}{1/n} > N.$$

Indeed,

$$|g(x_{j+1})-g(x_j)| \geq 2\epsilon - |f(x_{j+1})-f(x_j)| \geq \epsilon,$$

because $|x_{j+1}-x_j| \leq \delta$. Since $n\epsilon > 500N > N$, done.

Finally, observe that $g \notin F_N$ since any x_0 belongs to some $[x_j, x_{j+1}]$ and $|g(x)-g(x_0)| > N|x-x_0|$ for $x_0 \neq x \in [x_j, x_{j+1}]$.

□

Uniform Convergence and Integration

Definition. Let $f : [a,b] \to \mathbb{R}$ be a bounded function.

- A partition of $[a,b]$ is a finite set

$$P = \{a = x_0 < x_1 < \cdots < x_N = b\}.$$

- Define intervals $I_j \coloneqq [x_{j-1}, x_j]$ for $j = 1, \ldots, N$, with lengths $\Delta x_j \coloneqq x_j - x_{j-1}$.
- Upper Riemann sums:
$$U(f,p) \coloneqq \sum_{j=1}^{N} M_j(f,P)\Delta x_j,$$
where $M_j(f,p) = \sup_{x \in I_j} f(x)$.
- Lower Riemann sums:
$$L(f,p) \coloneqq \sum_{j=1}^{N} m_j(f,P)\Delta x_j,$$
where $m_j(f,p) = \inf_{x \in I_j} f(x)$.

Theorem 8. *f is Riemann integrable if and only if $\forall \epsilon > 0$, there exists an partition P, s.t. $U(f,P) - L(f,P) < \epsilon$. In this case,*

$$\int_a^b f(x)\,\mathrm{d}x = \inf_P U(f,P) = \sup_P L(f,P).$$

Example. Pointwise limit of Riemann integrable functions need not be Riemann integrable. Let $f_n(x)$ be a function from $[0,1]$ to $\mathbb{R}$ defined as following

$$f_n(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \cap [0,1] \text{ with denominator of x at most } n \\ 0, & \text{otherwise} \end{cases}$$

$f_n(x)$ is Riemann integrable since it is piecewise linear. Consider the limit

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \cap [0,1] \\ 0, & \text{if } x \in [0,1] \setminus \mathbb{Q} \end{cases}$$

For any partition P, $U(f,P) = 1$ and $L(f,P) = 0$. We see the limit $f(x)$ is not Riemann integrable.

Theorem 9. *Let $\{f_n\}$ be a sequence of Riemann integrable functions on $[a,b]$ and assume $f_n \to f$ uniformly. Then f is Riemann integrable and*

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

Note. Under the assumption of pointwise convergence, the above formula can fail even if the limit f is Riemann integrable.
Under the assumption of uniform convergence, this is a total mess if $[a,b]$ is replaced by $[a,\infty)$ and $\int_a^b$ is replaced by $\int_a^\infty$.

Uniform Convergence and Differentiation

Example. Consider the sequence

$$f_n(x) = \sqrt{\frac{1}{n} + x^2}, \quad x \in \mathbb{R}.$$

Claim: $f_n \to f$ uniformly, where $f = |x|$, on $\mathbb{R}$.

Let $\epsilon > 0$. Let $N = \lceil 1/\epsilon \rceil$. Then $\forall n > N$ and $\forall x$, we have

$$|f_n(x) - |x|| < \frac{1}{n} < \epsilon.$$

Furthermore, each f_n is differentiable and

$$f_n'(x) = \frac{x}{\sqrt{\frac{1}{n} + x^2}}.$$

However, $f_n' \to g$ pointwise, where

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

Also we observe that the limiting function f is not differentiable.

Definition. Let I be an interval with nonempty interior. Let

$$\mathcal{C}^k(I) \coloneqq \{f : I \to \mathbb{R} \,:\, f \text{ is } k\text{-‘times’ differentiable and } f^{(j)} \in \mathcal{C}^0(I),\, 0 \leq j \leq k\}.$$

Define the norm

$$\|f\|_{\mathcal{C}^k(I)} \coloneqq \sum_{j=0}^{k} \|f^{(j)}\|_{\mathcal{C}^0(I)}.$$

Proposition (HW)**.** *$\exists A = A(I)$, s.t. $\|f\|_{C^k(I)} \leq A(\|f\|_{C^0(I)} + \|f^{(k)}\|_{C^0(I)})$. Further, this A can be independent of I.*

Theorem 10. *$\mathcal{C}^k(I)$ is a complete metric space.*

Proof. Prove by induction. First, we know that $\mathcal{C}^0(I)$ is complete. We want to deduce completeness of $\mathcal{C}^{k+1}(I)$ from completeness of $\mathcal{C}^k(I)$.
Now we assume $\mathcal{C}^k(I)$ is complete and let $\{f_n\}$ be a Cauchy sequence in $\mathcal{C}^{k+1}(I)$. Notice $\forall n, m$,

$$\|f_n - f_m\|_{\mathcal{C}^k(I)} + \|f_n^{(k+1)} - f_m^{(k+1)}\|_{\mathcal{C}^0(I)} = \|f_n - f_m\|_{\mathcal{C}^{k+1}(I)},$$

so $\{f_n\}$ is a Cauchy sequence in $C^k(I)$. Then by hypothesis, $\exists f \in \mathcal{C}^k(I)$, s.t. $\|f_n - f\|_{\mathcal{C}^k(I)} \to 0$ as $n \to \infty$. Since $\{f_n^{(k+1)}\}$ is Cauchy in $\mathcal{C}^0(I)$, $\exists g \in \mathcal{C}^0(I)$, s.t. $f_n^{(k+1)} \to g$ uniformly on I. If we can show that $f \in \mathcal{C}^{k+1}(I)$ and $f^{(k+1)} = g$, then

$$\begin{aligned} \|f_n - f\|_{\mathcal{C}^{k+1}(I)} &= \|f_n - f\|_{\mathcal{C}^k(I)} + \|f_n^{(k+1)} - f^{(k+1)}\|_{\mathcal{C}^0(I)} \\ &= \|f_n - f\|_{\mathcal{C}^k(I)} + \|f_n^{(k+1)} - g\|_{\mathcal{C}^0(I)} \end{aligned}$$

goes to 0 as $n \to \infty$. For this, it suffices to prove that $f^{(k)}$ is differentiable and $(f^{(k)})' = g$. Fix $x_0 \in I$. Then $\forall x \in I$, by the fundamental theorem of calculus, we have

$$f_n^{(k)}(x) = f_n^{(k)}(x_0) + \int_{x_0}^{x} (f_n^{(k)})'(y)\,\mathrm{d}y.$$

Since $f_n \to f$ in $\mathcal{C}^k$, $\lim_{n\to\infty} f_n^{(k)}(x) = f^{(k)}(x)$ for all $x \in I$. Since $f_n^{(k+1)} = (f_n^{(k)})' \to g$ uniformly, we know

$$\lim_{n\to\infty} \int_{x_0}^{x} (f_n^{(k)})'(y)\,\mathrm{d}y = \int_{x_0}^{x} \lim_{n\to\infty} (f_n^{(k)})'(y)\,\mathrm{d}y = \int_{x_0}^{x} g(y)\,\mathrm{d}y.$$

Finally, by linearity of limits,

$$f^{(k)}(x) = f^{(k)}(x_0) + \int_{x_0}^{x} g(y)\,\mathrm{d}y.$$

Since g is continuous, the fundamental theorem of calculus says $\int_{x_0}^{x} g(y)\,\mathrm{d}y$ is differentiable with derivative $g(x)$. Then we can conclude that $f^{(k)}$ is differentiable and $f^{(k)} = g$. $\square$

Proposition (HW)**.** *Let $\{f_n\}$ be a sequence in $\mathcal{C}^k(I)$. Assume $\{f_n^{(k)}\}$ is Cauchy in $\mathcal{C}^0(I)$ and $\exists x_0 \in I$, s.t. $\forall j = 0, \ldots, k-1$, $\{f_n^{(j)}(x_0)\}$ is a Cauchy sequence. Then $\{f_n\}$ is convergent in $\mathcal{C}^k(I)$.*

Theorem 11 (7.17)**.** *Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a, b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a, b]$. If $\{f_n'\}$ converges uniforly on $[a, b]$, then $\{f_n\}$ converges uniformly on $[a, b]$, to a function f, and*

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad (a \le x \le b).$$

Example. Let

$$f(x) = \begin{cases} x^2 \sin 1/x, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

It has derivative

$$f'(x) = \begin{cases} 2x \sin 1/x - \cos 1/x, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

and $f'(x)$ is discontinuous. So there exists some functions which has derivative at every point, but the derivative is not continuous.

Recall. Let (X, d) be a complete metric space. A subset $K \subseteq X$ is compact if and only if K is closed and totally bounded.

Corollary. *$\mathcal{F} \subseteq \mathcal{C}^0(X)$ is compact if and only if $\mathcal{F}$ is closed and totally bounded.*

But what does totally bounded mean for $\mathcal{C}^0(X)$? We want a simpler characterization of totally boundedness.

Definition. Let $\mathcal{F} \subseteq \mathcal{C}(X)$.

- $\mathcal{F}$ is pointwise bounded if $\forall x \in X$, $\exists M_x$, s.t. $\forall f \in \mathcal{F}$, $|f(x)| \le M_x$ (i.e. $\forall x \in X$, $\{f(x) : f \in \mathcal{F}\}$ is a bounded set).

- $\mathcal{F}$ is equicontinuous if $\forall \epsilon > 0$, $\forall x \in X$, $\exists \delta = \delta(x, \epsilon) > 0$, s.t. $\forall f \in \mathcal{F}$ and $\forall y \in \mathcal{B}_\delta(x)$, $|f(x) - f(y)| < \epsilon$.
- $\mathcal{F}$ is uniformly equicontinuous if $\forall \epsilon > 0$, $\exists \delta = \delta(\epsilon) > 0$, s.t. $\forall f \in \mathcal{F}$ and $\forall x, y \in X$ with $d(x, y) < \delta$, $|f(x) - f(y)| < \epsilon$.

Example. $\mathcal{F} \coloneqq \{f \in \mathcal{C}^0([0,1]) : f$ if differentiable on $(0,1)$ and $|f'(x)| \le 1, \forall x \in (0,1)\}$ is uniformly equicontinuous.

Proof. $\forall f \in \mathcal{F}$, $x, y \in [0,1]$, $|f(x) - f(y)| \le |x - y|$ by mean value theorem. So $\forall \epsilon > 0$, pick $\delta = \epsilon$. Then $|f(x) - f(y)| < \epsilon$ for all $f \in \mathcal{F}$ and for all $x, y \in [0,1]$ with $|x - y| < \delta$. □

Example. $f_n(x) = x^n$ defined on $[0,1]$. Then $\mathcal{F} \coloneqq \{f_n : n \in \mathbb{N}\}$ is a pointwise bounded, but not a equicontinuous subset of $\mathcal{C}^0([0,1])$.

Proof. For any $x \in [0,1]$, $\{f(x) : f \in \mathcal{F}\} \subseteq [0,1]$. So $\mathcal{F}$ is pointwise bounded. Fix $x = 1$ and $\epsilon = 0.5$. For any $\delta > 0$, there exists N sufficiently large such that for all $n > N$, $f_n(x - \delta/2) < 0.5$. Thus, $f_n(x)$ is not equilcontinous. □

Proposition. *If K is compact, then $\mathcal{C}^0(K) = \mathcal{C}(K)$. ($\mathcal{C}^0$ means bounded continuous functions, while $\mathcal{C}$ means continuous functions)*

Example. Show by example that a pointwise bounded subset of $\mathcal{C}^0(K)$ need not be uniformly pointwise bounded (i.e. a bounded subset of $\mathcal{C}^0(K)$).

$$f_n(x) = \begin{cases} n^2 x, & 0 \le x \le 1/n \\ 2n - n^2 x, & 1/n < x \le 2/n \\ 0 & 2/n < x \le 1 \end{cases}$$

Proposition. *If K is compact, then $\mathcal{F} \subseteq \mathcal{C}^0(K)$ is equicontinuous if and only if $\mathcal{F}$ is uniformly equicontinuous.*

Proof. $\Leftarrow$: is immediate.
$\Rightarrow$: Assume K is compact and $\mathcal{F} \subseteq \mathcal{C}^0(K)$ is equicontinuous. Let $\epsilon > 0$. Then $\forall x \in K$, $\exists \delta_x > 0$, s.t. $\forall y \in B_{\delta_x}(x)$, $\forall f \in \mathcal{F}$, $|f(x) - f(y)| < \epsilon/2$. Since $K \subseteq \cup_{x \in K} B_{\delta_x}(x)$, $\exists x_1, \ldots, x_N$, s.t. $K \subseteq \cup_{j=1}^N B_{\delta_{x_j}}(x_j)$.

Claim: $\exists \delta > 0$, s.t. $\forall y, z \in K$, if $d(y, z) < \delta$, then $y, z \in B_{\delta_{x_j}}(x_j)$.

Suppose not. Then there exists sequences $\{y_n\}$ and $\{z_n\}$ such that $d(y_n, z_n) \to 0$, but y_n and z_n never belong to the same $B_{\delta_{x_j}}(x_j)$. Each z_n lives in some $\mathcal{B}_{\delta_{x_j}}(x_j)$ and since there are only finitely many such balls, there must be a ball that contains infinitely many z_n. Passing to a subsequence, we may assume $\exists j_0$, s.t. $z_n \in B_{\delta_{x_{j_1}}}(x_{j_1})$ for all n. Similarly, passing to a further subsequence, we may assume $y_n \in B_{\delta_{x_{j_2}}}(x_{j_2})$ for all n. Since $\{z_n\}$ is in K, which is compact, passing to a further subsequence, we may assume $z_n \to z$. Since $d(z_n, y_n) \to 0$, $y_n \to z$. Notice $\exists j$, s.t. $z \in B_{\delta_{x_j}}(x_j)$ (z is not necessarily in $B_{\delta_{x_{j_1}}}(x_{j_1})$). For n sufficiently large, y_n and z_n are both in $B_{\delta_{x_j}}(x_j)$. Contradiction.

Then we have
$$|f(y) - f(z)| \le |f(y) - f(x_j)| + |f(x_j) - f(z)| < \epsilon$$
for any $f \in \mathcal{F}$. □

Theorem 12 (Arzela-Ascali Theorem)**.** *If K is a compact metric space, then $\mathcal{F} \subseteq \mathcal{C}^0(K)$ is totally bounded if and only if $\mathcal{F}$ is pointwise bounded and equicontinuous.*

Proof. $\Rightarrow$: Assume $\mathcal{F} \subseteq \mathcal{C}^0(K)$ is totally bounded.

Pointwise bounded: Since $\mathcal{F}$ is totally bounded, $\exists N$ and $f_1, \ldots, f_n \in \mathcal{F}$, s.t. $\mathcal{F} \subseteq \cup_{j=1}^N B_1(f_1)$. Then $\forall x$ and $\forall f \in \mathcal{F}$,

$$|f(x)| < \sum_{j=1}^{N} |f_j(x)| + 1 \leq \sum_{j=1}^{N} \|f_j\|_{\mathcal{C}^0(K)} + 1.$$

Equicontinuous: Let $\epsilon > 0$. Since $\mathcal{F}$ is totally bounded, $\exists N$ and $f_1, \ldots, f_n \in \mathcal{F}$, s.t. $\mathcal{F} \subseteq \cup_{j=1}^N B_{\epsilon/3}(f_1)$. Since each f_j is uniform continuous (being continuous on a compact set), $\exists \delta_j > 0$, s.t. $\forall x, y \in K$ with $d(x,y) < \delta_j$, we have $|f_j(x) - f_j(y)| < \epsilon/3$. Let $\delta := \min\{\delta_1, \ldots, \delta_N\}$. Let $f \in \mathcal{F}$. Then $\exists j$, s.t. $\|f - f_j\|_{\mathcal{C}^0} < \epsilon/3$. Finally, if $d(x,y) < \delta$,

$$|f(x) - f(y)| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_i(y) - f(y)| < \epsilon.$$

Thus $\mathcal{F}$ is uniform equicontinuous.

$\Leftarrow$: Assume $\mathcal{F}$ is pointwise bounded and equicontinouus. Let $\epsilon > 0$. By proposition, $\mathcal{F}$ is uniformly equicontinous, so $\exists \delta > 0$, s.t. $\forall f \in \mathcal{F}$, $\forall x, y \in K$ with $d(x,y) < \delta$, $|f(x) - f(y)| < \epsilon/4$. Since K is compact, K is totally bounded. So $\exists N$ and $x_1, \ldots, x_N \in K$, s.t. $K \subseteq \cup_{j=1}^N B_\delta(x_j)$. The idea is to discretize the functions in $\mathcal{F}$. Consider $P := \{(f(x_1), f(x_2), \ldots, f(x_N)) : f \in \mathcal{F}\} \subseteq \mathbb{R}^N$. By pointwise boundedness, for each j, $\exists M_j$, s.t. $\{f(x_j) : f \in \mathcal{F}\} \subseteq [-M_j, M_j]$. Therefore

$$P \subseteq [-M_1, M_1] \times [-M_2, M_2] \times \cdots \times [-M_N, M_N],$$

which is a bounded (hence totally bounded) subset of $\mathbb{R}^N$. So P is totally bounded. Then $\exists L$ and $y_1, \ldots, y_L \in P$, s.t. $P \subseteq \cup_{j=1}^L B_{\epsilon/4}(y_j)$. (Note: $B_{\epsilon/4}(y_j)$ is the ball in Euclidean space.) For each y_j, by definition, $\exists f_j \in \mathbb{F}$, s.t. $(y_j)_i = f_j(x_i)$, $i = 1, \ldots, N$. Thus, $\forall f \in \mathcal{F}$, $\exists j$, s.t. $1 \leq j \leq L$ and

$$|f(x_i) - f_j(x_i)| \leq |(f(x_1), \ldots, f(x_N)) - (f_j(x_1), \ldots, f_j(x_N))| < \epsilon/4, \quad i = 1, \ldots, N$$

For any $x \in K$, by totally boundedness of K, $\exists i \in 1, \ldots, N$, s.t. $x \in B_\delta(x_i)$. So

$$|f(x) - f_j(x)| \leq |f(x) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(x)| < 3\epsilon/4.$$

Taking the supremum over x, $\|f_j - f\|_{\mathcal{C}^0(K)} \leq 3\epsilon/4 < \epsilon$. Thus, $f \in B_\epsilon(f_j)$ and we have $\mathcal{F} \subseteq \cup_{j=1}^L B_\epsilon(f_j)$.

$\square$

Corollary. *Let K be a compact metric space.*

1. *$\mathcal{F} \subseteq \mathcal{C}^0(K)$ is compact if and only if $\mathcal{F}$ is closed, pointwise bounded and equicontinuous.*
2. *$\mathcal{F} \subseteq \mathcal{C}^0(K)$ is precompact if and only if $\mathcal{F}$ is pointwise bounded and equicontinuous.*
3. *Let $\{f_n\}$ be a pointwise bounded equicontinuous sequence in $\mathcal{C}^0(K)$. Then $\{f_n\}$ has a uniformly convergent subsequence.*

Proposition. *For $k \in \mathbb{N}$ and I compact, a subset of $\mathcal{C}^k(I)$ is totally bounded if and only each of its ith derivatives, $i = 0, 1, \dots, k$, is pointwise bounded and equicontinuous.*

Theorem 13 ((Baby) Stone-Weierstress Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. There exists a sequence $\{P_n\}$ of polynomials, s.t. $P_n \to f$ uniformly on $[a, b]$.*

Note that even if f has derivatives of all orders, it probably doesn't work if we just take P_n equal to the n-th Taylor polynomial. Also, unless f itself is a polynomial, the degress of P_n must tend to infinity.

Proof. First, we do some reductions:

1. We assume $[a, b] = [0, 1]$. If this is not true, we can approximate $g(x) \coloneqq f(\frac{x-a}{b-a})$ and replace $P_n(x)$ by $P_n((b-a)x + a)$.
2. We assume $f(0) = f(1) = 0$. If this is not the case, we approximate $g(x) = f(x) - f(0) - x(f(1) - f(0))$ and replace $P_n(x)$ by $P_n(x) + f(0) + x(f(1) - f(0))$.

Now, we extend f to all of $\mathbb{R}$ by setting $f(x) = 0$ off of $[0, 1]$. (Warning: We are only approximating f on $[0, 1]$, even though the domain of f is larger.) Then f is continous and thus uniformly continuous on $\mathbb{R}$. Define

$$Q_n(x) \coloneqq c_n(1 - x^2)^n,$$

where c_n is chosen s.t.

$$\int_{-1}^{1} Q_n(x)\,\mathrm{d}x = 1,$$

i.e. $c_n = \left(\int_{-1}^{1}(1 - x^2)^n\,\mathrm{d}x\right)^{-1}$. Let

$$P_n(x) \coloneqq \int_0^1 f(t)Q_n(x - t)\,\mathrm{d}t = f * Q_n(x) = \int_{x-1}^{x} f(x - t)Q_n(t)\,\mathrm{d}t.$$

We need to prove that P_n is a polynomial and $P_n \to f$ uniformly on $[0, 1]$.

Claim 1: P_n is a polynomial of degree $\leq 2n$.

Subproof. Expand the expression for $Q_n(x - t)$,

$$\begin{aligned}
Q_n(x - t) = c_n(1 - (x - t)^2)^n &= \sum_{j=0}^{n} c_n \binom{n}{j} (-1)^j (x - t)^{2j} \\
&= \sum_{j=0}^{n} \sum_{k=0}^{2j} c_n \binom{n}{j} \binom{2j}{k} (-1)^{j+k} x^{2j-k} t^k \\
&= \sum_{m=0}^{2n} \left(\sum_{m \leq 2j \leq 2n} c_n \binom{n}{j} \binom{2j}{2j - m} (-1)^{j-m} x^m t^{2j-m} \right)
\end{aligned}$$

So $P_n(x) = \sum_{m=0}^{2n} a_{n,m} x^m$, where

$$a_{n,m} = \sum_{m \leq 2j \leq 2n} c_n \binom{n}{j} \binom{2j}{2j - m} (-1)^{j-m} \int_0^1 f(t) t^{2j-m}\,\mathrm{d}t.$$

□

Subproof (An alternative approach). For $k \in \mathbb{N}$, P_n is k times differentiable and

$$P_n^{(k)}(x) = \int_0^1 f(t) Q_n^{(k)}(x-t)\,\mathrm{d}t.$$

In particular, since Q_n is a polynomial of degree $2n$,

$$p_n^{(2n+1)}(x) = \int_0^1 f(t) \cdot 0\,\mathrm{d}t = 0.$$

So P_n is a polynomial of degree at most $2n$. By induction, it suffices to prove the following proposition, which sometimes is of independent interest. □

Proposition. *Let $g \in \mathcal{C}^1(\mathbb{R})$ and let $f \in \mathcal{C}^0(\mathbb{R})$ with $f = 0$ out of $[0,1]$. Then the convolution*

$$h(x) \coloneqq \int_0^1 f(t) g(x-t)\,\mathrm{d}t$$

is also in $\mathcal{C}^1(\mathbb{R})$, where its derivative is

$$h'(x) = \int_0^1 f(t) g'(x-t)\,\mathrm{d}t.$$

Note: Since $f(x) = 0$ off of $[0,1]$, we can also replace the proper integral $\int_0^1$ by the inproper integral $\int_{-\infty}^{\infty}$.

Proof of Proposition. First, observe that h is bounded in $\mathcal{C}^1(\mathbb{R})$. This is because

$$|h(x)| = \left|\int_0^1 f(t) g(x-t)\,\mathrm{d}x\right| \le \int_0^1 |f(t) g(x-t)|\,\mathrm{d}t \le \|f\|_{\mathcal{C}^0} \|g\|_{\mathcal{C}^1}.$$

By a similar argument, we observe that

$$|h'(x)| \le \|f\|_{\mathcal{C}^0} \|g\|_{\mathcal{C}^1}.$$

In addition, we claim that $h'(x)$ is continuous. Let $x_n \to x$. Define

$$\varphi_n(t) \coloneqq f(t) g'(x_n - t) \quad \text{and} \quad \varphi(t) \coloneqq f(t) g'(x-t).$$

Without loss of generality, assume $x_n \in [x-1, x+1]$ for all n. Observe that every φ_n and φ lives on $[0,1]$. For any $t \in [0,1]$, $x_n - t \in [x-2, x+1]$. Since g' is continuous, it is uniform continuous on $[x-2, x+1]$. Hence $\forall \epsilon > 0$, $\exists N$, s.t. $\forall n \ge N$, $\forall t \in [0,1]$, $|g'(x_n - t) - g'(x-t)| < \epsilon$ (because $|x_n - x| <$ some δ). Thus, $\varphi_n \to \varphi$ uniformly on $[0,1]$. Then we have

$$\lim_{x_n \to x} h'(x_n) = \lim_{x_n \to x} \int_0^1 \varphi_n(t)\,\mathrm{d}t = \int_0^1 \varphi(t)\,\mathrm{d}t = h'(x).$$

□

Claim 2: $P_n \to f$ uniformly on $[0,1]$.

Subclaim 1: $c_n \leq \sqrt{n}, \forall n$.

Let $g(y) = (1-y)^n$. Then $g''(y) = n(n-1)(1-y)^{n-2} \geq 0$ for $y \in [0,1]$. By Taylor's theorem, for $y \in [0,1]$,

$$\begin{aligned} g(y) &= g(0) + g'(0)y + \frac{1}{2}g''(ty)y^2, \quad \text{for some } t \in [0,y] \subseteq [0,1] \\ &\geq g(0) + g'(0)y \\ &= 1 - ny. \end{aligned}$$

Apply the above result by setting y to x^2, then we have

$$\begin{aligned} \int_{-1}^{1} (1-x^2)^n \, \mathrm{d}x &\geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} (1-x^2)^n \, \mathrm{d}x \\ &\geq \int_{-1/\sqrt{n}}^{1/\sqrt{n}} 1 - nx^2 \, \mathrm{d}x = \frac{4}{3\sqrt{n}}. \end{aligned}$$

Hence

$$c_n \leq \frac{3\sqrt{n}}{4} \leq \sqrt{n}.$$

Subclaim 2: $\forall \delta > 0$, $Q_n \to 0$ uniformly on $\{x : \delta \leq |x| \leq 1\}$.

On $\delta \leq |x| \leq 1$,

$$Q_n(x) = c_n(1-x^2)^n \leq \sqrt{n}(1-\delta^2)^n \to 0,$$

since $1 - \delta^2 < 1$.

Now let's prove the claim. Let $x \in [0,1]$.

$$\begin{aligned} P_n(x) - f(x) &= \int_0^1 f(t)Q_n(x-t) \, \mathrm{d}t - f(x) \int_{-1}^{1} Q_n(s) \, \mathrm{d}s \\ &= \int_{x-1}^{x} f(x-s)Q_n(s) \, \mathrm{d}s - \int_{-1}^{1} f(x)Q_n(s) \, \mathrm{d}s \\ &= \int_{-1}^{1} f(x-s)Q_n(s) \, \mathrm{d}s - \int_{-1}^{1} f(x)Q_n(s) \, \mathrm{d}s. \end{aligned}$$

since $f = 0$ off of $[0,1]$ (so $f(x-s)$ vanishes for $s \notin [x-1, x]$). Thus

$$P_n(x) - f(x) = \int_{-1}^{1} (f(x-s) - f(x))Q_n(s) \, \mathrm{d}x = \int_{-\delta}^{\delta} + \int_{-1}^{-\delta} + \int_{\delta}^{1} =: I_1 + I_2 + I_3$$

For I_1, we have

$$\begin{aligned} I_1 &\leq \int_{-\delta}^{\delta} |f(x-s) - f(x)| \cdot |Q_n(s)| \, \mathrm{d}s \\ &\leq \max_{|s| \leq \delta} |f(x-s) - f(x)| \int_{-1}^{1} Q_n(s) \, \mathrm{d}s. \\ &= \max_{|s| \leq \delta} |f(x-s) - f(x)| \end{aligned}$$

For I_2 and I_3, we have

$$I_2 + I_3 \le 2 \cdot \max_{s \in \mathbb{R}} |f(x-s) - f(x)| \cdot \max_{\delta \le |s| \le 1} |Q_n(s)| \le 4 \cdot \|f\|_{\mathcal{C}^0} \cdot \max_{\delta \le |s| \le 1} |Q_n(s)|$$

Now, let $\epsilon > 0$. Since f is uniformly continuous, $\exists \delta > 0$, s.t. whenever $|s| \le \delta$, $|f(x - s) - f(x)| < \epsilon/2$, $\forall x \in \mathbb{R}$. So $I_1 < \epsilon/2$, $\forall n$. Also, by the subclaim, $\exists N$, s.t. $\forall n \ge N$, $\max_{\delta \le |s| \le 1} |Q_n(s)| < \epsilon/(4\|f\|_{\mathcal{C}^0})$. So $\forall n \ge N$ and $x \in [0, 1]$,

$$|P_n(x) - f(x)| \le |I_1| + |I_2 + I_3| < \epsilon.$$

□

Definition. Let E be a nonempty set.

$$\mathcal{F}(E) \coloneqq \mathcal{F}(E; \mathbb{R}) \coloneqq \{f : E \to \mathbb{R}\}.$$

Definition. A family $\mathcal{A} \subseteq \mathcal{F}(E)$ is an algebra if $\forall c \in \mathbb{R}$, and $f, g \in \mathcal{A}$,

$$cf,\ f + g,\ fg \in \mathcal{A}.$$

Definition. Let $\mathcal{A} \subseteq \mathcal{F}(E)$.

- $\mathcal{A}$ separates points in E if $\forall x_1, x_2 \in E$ with $x_1 \ne x_2$, $\exists f \in \mathcal{A}$, s.t. $f(x_1) \ne f(x_2)$.
- $\mathcal{A}$ is nonvanishing on E if $\forall x \in E$, $\exists f \in \mathcal{A}$, s.t. $f(x) \ne 0$.

Example. The set P of polynomials on $\mathbb{R}$ is an algebra, which separates points in $\mathbb{R}$ and vanishes nowhere.

Example. The set P_{odd} is not an algebra and is not nonvanishing (because it is always 0 at $x = 0$). The set P_{even} is an algebra, which is nonvaishing on $\mathbb{R}$ but it doesn't separate points.

Proposition. *Let $\mathcal{A} \subseteq \mathcal{F}(E)$ be an algebra that separates points and vanishes nowhere. Then $\forall x_1 \ne x_2 \in E$ and $\forall c_1, c_2 \in \mathbb{R}$, $\exists f \in \mathcal{A}$, s.t. $f(x_1) = c_1$ and $f(x_2) = c_2$.*

Proof. By definition, $\exists g, h, k \in \mathcal{A}$, s.t. $g(x_1) \ne g(x_2)$, $h(x_1) \ne 0$, $k(x_2) \ne 0$. Now let

$$\begin{aligned} u &= (g - g(x_1))k = gk - g(x_1)k \in \mathcal{A}, \\ v &= (g - g(x_2))h = gh - g(x_2)h \in \mathcal{A}. \end{aligned}$$

Then $u(x_1) = 0, u(x_2) \ne 0, v(x_1) \ne 0, v(x_2) = 0$. Finally, let

$$f = c_1 \frac{v}{v(x_1)} + c_2 \frac{u}{u(x_2)} \in \mathcal{A}.$$

□

Theorem 14 (Full Stone Weierstress)**.** *Let K be a compact set and $\mathcal{A} \subseteq \mathcal{C}^0(K)$ be an algebra that separtes points and vanishes nowhere. Then $\mathcal{A}$ is dense in $\mathcal{C}^0(K)$. In other words, $\forall f \in \mathcal{C}^0(K)$, there exists a sequence $\{f_n\}$ in $\mathcal{A}$, s.t. $f_n \to f$ uniformly.*

Proof. Let's $\mathcal{C} = \bar{\mathcal{A}}$. Claim $\mathcal{C}$ is an algebra (HW). We now need to show that $\mathcal{C} = \mathcal{C}^0(K)$.

Claim 1: If $f \in \mathcal{C}$, then so is $|f|$.

Let $a := \|f\|_{\mathcal{C}^0(K)}$. By baby Stone Weierstress (aka Weierstress preparation lemma), there exists a polynomial P on $\mathbb{R}$ such that $|P(y) - |y|| < \epsilon$ for any $y \in [-a, a]$. Define $g := P \circ f$. Then $g(x) = \sum_{n=0}^{N} a_n f(x)^n$, where the a_n's are the coeffients of P. Since $\mathcal{C}$ is an algebra, $g \in \mathcal{C}$. Furthermore, $\forall x \in K$, $f(x) \in [-a, a]$. So

$$|g(x) - |f(x)|| = |P(y) - |y|| < \epsilon.$$

Thus $\||f| - g\|_{\mathcal{C}^0(K)} \leq \epsilon$. Since ϵ was arbitrary, $|f| \in \bar{\mathcal{C}} = \mathcal{C}$.

Claim 2: If $f_1, \ldots, f_n \in \mathcal{C}$, then so are $\max\{f_1, \ldots, f_n\}$ and $\min\{f_1, \ldots, f_N\}$.

If $N = 2$, this follows from Claim 1 and

$$\max\{f, g\} = \frac{f+g}{2} + \frac{|f-g|}{2}, \quad \min\{f, g\} = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

For larger N, by induction, we have

$$\max\{f_1, \ldots, f_{N+1}\} = \max\{\max\{f_1, \ldots, f_N\}, f_{N+1}\},$$
$$\min\{f_1, \ldots, f_{N+1}\} = \min\{\min\{f_1, \ldots, f_N\}, f_{N+1}\}.$$

Claim 3: Let $f \in \mathcal{C}^0(K)$, $\epsilon > 0$ and $x_0 \in K$. Then $\exists g_{x_0} \in \mathcal{C}$, s.t. $g_{x_0}(x_0) = f(x_0)$ and $g_{x_0}(x) > f(x) - \epsilon$ for all $x \in K$ (approximate f from not too far below).

Let $x_1 \in K$. Then $\exists h_{x_1} \in \mathcal{C}$, s.t. $h_{x_1}(x_0) = f(x_0)$ and $h_{x_1}(x_1) = f(x_1)$. For $y \in K$, define

$$G_y := \{x \in K : h_y(x) > f(x) - \epsilon\}.$$

Then G_y is open since h_y is continuous. Also $y \in G_y$ since $h_y(y) = f(y)$. Thus, $\{G_y : y \in K\}$ is an open cover of K, so $\exists y_1, \ldots, y_N \in K$, s.t. $K \subseteq \cup_{n=1}^{N} G_{y_n}$. Now, let $g_{x_0} = \max\{h_{y_1}, \ldots, h_{y_N}\}$. By Claim 2, $g_{x_0} \in \mathcal{C}$. Furthermore, $g_{x_0}(x_0) = f(x_0)$. Finally, for $x \in G_{y_n}$, $g_{x_0}(x) \geq h_{y_n}(x) > f(x) - \epsilon$. Thus, g_{x_0} is the function we want.

Claim 4: $\forall f \in \mathcal{C}^0(K)$ and $\forall \epsilon > 0$, $\exists g \in \mathcal{C}$, s.t. $\forall x$, $f(x) - \epsilon < g(x) < f(x) + \epsilon$.

Let $x_0 \in K$. By Claim 3, $\exists g_{x_0} \in \mathcal{C}$, s.t. $g_{x_0} = f(x_0)$ and $g_{x_0}(x) > f(x) - \epsilon$, $\forall x \in K$. For $y \in K$, define

$$H_y := \{x \in K : g_y(x) < f(x) + \epsilon\}.$$

Then H_y is open since g_y is continuous. Again, $y \in H_y$ since $g_y(y) = f(y)$. So $\{H_y : y \in K\}$ is an open cover of K. Then $\exists y_1, \ldots, y_N \in K$, s.t. $K \subseteq \cup_{n=1}^{N} H_{y_n}$. Finally, define $g = \min\{g_{y_1}, \ldots, g_{y_n}\}$. Then $\forall x$, $g(x) > f(x) - \epsilon$. Also $\exists n$, s.t. $x \in H_{y_n}$. Then $g(x) \leq g_{y_n}(x) < f(x) + \epsilon$.

$\square$

Theorem 15 (Picard's theorem)**.** *Let $t_0 \in \mathbb{R}$ and $y_0 \in \mathbb{R}^k$. Let $a, b \in \mathbb{R}$ and define*

$$B := \{y \in \mathbb{R}^k : |y - y_0| \leq b\},$$

and

$$R := [t_0 - a, t_0 + a] \times B.$$

Let $F : R \to \mathbb{R}^k$ be a bounded, continuous function and let $M := \|F\|_{\mathcal{C}^0(R)}$. Assume that $\exists C \in \mathbb{R}$, s.t. $\forall t \in (t_0 - a, t_0 + a)$, $\forall u, y \in B$, $|F(t,u) - F(t,y)| \leq C|u-y|$. Then, $\exists!$ function $g : (t_0 - \tilde{a}, t_0 + \tilde{a}) \to B$, s.t. g is differentiable and g solves the initial value problem

$$\begin{cases} g(t_0) &= y_0 \\ g'(t) &= F(t, g(t)), \quad \forall t \in (t_0 - \tilde{a}, t_0 + \tilde{a}). \end{cases}$$

Here, $\tilde{a} = \min\{a, b/M\}$.

Warning: The C in the assumption is crucial for this theorem. Consider $k = 1$, $F(t, y) = y^{1/3}$ and the initial value problem

$$\begin{cases} g(0) &= 0 \\ g'(t) &= (g(t))^{1/3} \end{cases}$$

Here is one solution, $g(t) = 0$ for all t. Here is another solution,

$$g(t) = \begin{cases} ct^{3/2}, & t \geq 0, \\ 0, & t < 0, \end{cases}$$

where $c^2 = 8/27$. Actually, there are infinitely many distinct solutions.

Proof. Observe that g solves our initial value problem if and only if

- g is continuous;
- $|g(t) - y_0| \leq b$, $\forall t \in I$.
- $g(t) = y_0 + \int_{t_0}^t F(s, g(s))\,\mathrm{d}s$, $\forall t \in I := [t_0 - \tilde{a}, t_0 + \tilde{a}]$.

Here we are using the fact that differentiable functions are continuous and equal the integral of their derivative plus some constant. On the other hand, that condition of g implies continuity of the integrand and then we can use the fundamental theorem of calculus.

Consider the set

$$\mathcal{M} := \{g \in \mathcal{C}^0(I; \mathbb{R}^k) : g(t) \in B, \forall t \in I; g(t_0) = y_0\}.$$

Consider the map $\Phi : \mathcal{M} \to \mathcal{C}^0(I, \mathbb{R}^k)$,

$$[\Phi(g)](t) := y_0 + \int_{t_0}^t F(s, g(s))\,\mathrm{d}s.$$

By the fundamental theorem of calculus, $\Phi(g)$ is differentiable and hence it is continuous.

Now, we want to show that there exists an unique fixed point g of Φ in $\mathcal{M}$. The idea is to apply contraction mapping theorem. Observe that $\mathcal{M}$ is closed since it is the intersection of two closed sets (careful here). Also since $\mathcal{C}^0(I, \mathbb{R})$ is complete, $\mathcal{M}$ is complete. Since the function $g(t) = y_0, \forall t \in I$ is in $\mathcal{M}$, $\mathcal{M} \neq \emptyset$. Thus, by the contraction mapping theorem, any contraction on $\mathcal{M}$ has a unique fixed point p. Now we want to show Φ is a contraction on $\mathcal{M}$.

Compute

$$
\begin{aligned}
|\Phi(g)(t) - y_0| = \left| \int_{t_0}^{t} F(s, g(s))\, \mathrm{d}s \right| &\le \left| \int_{t_0}^{t} |F(s, g(s))|\, \mathrm{d}s \right| \\
&\le \left| \int_{t_0}^{t} M\, \mathrm{d}s \right| = |t - t_0| M \le \tilde{a} M \le b.
\end{aligned}
$$

Let $g_1, g_2 \in \mathcal{M}$.

$$
\begin{aligned}
|\Phi(g_1)(t) - \Phi(g_2)(t)| &= \left| \int_{t_0}^{t} F(s, g_1(s)) - F(s, g_2(s))\, \mathrm{d}s \right| \\
&\le \left| \int_{t_0}^{t} |F(s, g_1(s)) - F(s, g_2(s))|\, \mathrm{d}s \right| \\
&\le \left| \int_{t_0}^{t} C |g_1(s) - g_2(s)|\, \mathrm{d}s \right| \\
&\le |t - t_0| \cdot C \cdot \|g_1 - g_2\|_{\mathcal{C}^0(I, \mathbb{R}^k)} \\
&\le \tilde{a} C \|g_1 - g_2\|_{\mathcal{C}^0(I; \mathbb{R}^k)}
\end{aligned}
$$

So Φ is a contraction if and only if $\tilde{a} C < 1$. We have two approaches to fix this.

Approach 1. By contraction mapping theorem and above computation, we can uniquely solve the initial value problem for a shorter time, say on $I_0 := [t_0 - a_0, t_0 + a_0]$, where $a_0 := \min\{a, b/M, 1/(2C)\}$. Now define $t_{-1} := t_0 - a_0$ and $t_1 := t_0 + a_0$. Look at the new initial value problem on $[t_{-1} - a_0, t_1 + a_0]$. Define $y_{\pm 1} := g_0(t_{\pm 1})$. The new initial value problem can be written as

$$
\begin{cases}
g_{\pm 1}(t_{\pm 1}) = y_{\pm 1} \\
\quad g'_{\pm 1}(t) = F(t, g(t))
\end{cases}
$$

We a find a unique solution on

$$
I_{\pm 1} := [t_{\pm 1} - b_1, t_{\pm 1} + b_1],
$$

where $b_1 = \min\{a - a_0, \frac{b - a_0 M}{M}, 1/2C\}$. Furthermore, by uniqueness (in contraction mappinig theorem), $g_{\pm 1}$ equals g on $I_{\pm 1} \cap I_0$. Thus there exists an unique solution on

$$
I_1 \cup I_0 \cup I_{-1} = [t_0 - a_0 - b_1, t_0 + a_0 + b_1].
$$

Let $a_1 := a_0 + b_1$. Then $a_1 \ge \{a, b/M, 2/2C\}$. Iterate this process will finish the proof.

Approach 2. Find a better (but equivalent) metric on $\mathcal{M}$.

Definition. Two metrics d_1 and d_2 are equivalent if there exists positive constants c_1, c_2, s.t. $\forall f, g$,

$$
c_1 d_2(f, g) \le d_1(f, g) \le c_2 d_2(f, g).
$$

Proposition (HW). *Two equivalent metrics yield the same open and closed set, the same continuous functions, the same Cauchy and convergent sequences.*

Now define

$$d_C(f,g) := \sup_{t\in I} e^{-2C|t-t_0|}|g(t)-f(t)|.$$

Then

$$e^{-2C\tilde{a}}\|f-g\|_{\mathcal{C}^0(I,\mathbb{R}^k)} \le d_C(f,g) \le \|f-g\|_{\mathcal{C}^0(I,\mathbb{R}^k)}.$$

So the metrics are equivalent and $\mathcal{M}$ is complete with respect to d_C. Finally, for $t \in I$,

$$\begin{aligned}
e^{-2C|t-t_0|}\left|\Phi(g_1)(t)-\Phi(g_2)(t)\right| &= e^{-2C|t-t_0|}\left|\int_{t_0}^{t} F(s,g_1(s))-F(s,g_2(s))\,\mathrm{d}s\right| \\
&\le e^{-2C|t-t_0|}\left|\int_{t_0}^{t} Ce^{2C|s-t_0|}e^{-2C|s-t_0|}|g_1(s)-g_2(s)|\,\mathrm{d}s\right| \\
&\le e^{-2C|t-t_0|}\left|\int_{t_0}^{t} Ce^{2C|s-t_0|}d_C(g_1,g_2)\,\mathrm{d}s\right| \\
&\le e^{-2C|t-t_0|}\cdot C\cdot d_C(g_1,g_2)\cdot \frac{1}{2C}e^{2C|t-t_0|} \\
&= \frac{1}{2}d_C(g_1,g_2)
\end{aligned}$$

So Φ is a contraction on $\mathcal{M}$ with respect to d_C and hence has unique fixed point.

□

Analytic Functions

Definition. $\mathbb{C}$ is the complex field defined by

$$\mathbb{C} := \{x+iy : x,y \in \mathbb{R}\}.$$

Define $\overline{x+iy} := x-iy$. Then $(x+iy)(x-iy) = x^2+y^2$. Define $|x+iy| := \sqrt{x^2+y^2}$ and $d(z,w) = |z-w|$. Note that $\mathbb{C}$ is a complete metric space. Recall that for a field, if $x_1+iy_1, x_2+iy_2 \in \mathbb{C}$, then

- $(x_1+iy_1)+(x_2+iy_2) = (x_1+x_2)+i(y_1+y_2) \in \mathbb{C}$.
- $(x_1+iy_1)\cdot(x_2+iy_2) = (x_1y_1-x_2y_2)+i(x_1y_2+x_2y_1) \in \mathbb{C}$.
- $-(x+iy) = -x-iy \in \mathbb{C}$.
- If $x_2+iy_2 \neq 0$ (i.e. $x_2 \neq 0$ or $y_2 \neq 0$), then

$$\frac{(x_1+iy_1)}{(x_2+iy_2)} = \frac{(x_1+iy_1)}{(x_2+iy_2)}\cdot\frac{(x_2-iy_2)}{(x_2-iy_2)} = \frac{(x_1x_2+y_1y_2)}{x_2^2+y_2^2}+i\frac{(-x_1y_2+x_2y_1)}{x_2^2+y_2^2}.$$

Definition. If $\{c_n\}$ is a sequence of complex numbers, $\sum_{n=0}^{\infty} c_n$ converges if the sequence $\{s_n\}$ of partial sums $s_n := \sum_{n=0}^{N} c_n$ converges. Furthermore, by completeness, $\{s_n\}$ converges if and only if $\{s_n\}$ is Cauchy, in which case, we say $\{s_n\}$ satisfies the Cauchy criterion, that $\forall\epsilon$, $\exists N$, s.t. $\forall n > m \ge N$,

$$\left|\sum_{j=m+1}^{n} c_j\right| < \epsilon.$$

Definition. $\sum_{n=0}^{\infty} c_n$ converges absolutely if $\sum_{n=0}^{\infty} |c_n|$ converges.

Proposition. *If $\sum_{n=0}^{\infty} c_n$ converges absolutely, then $\sum_{n=0}^{\infty} c_n$ converges.*

Proof. Use the Cauchy criterion and triangle inequality. □

Definition. Let X be a nonempty set and $\{f_n\}$ be a sequence of functions mapping from X into $\mathbb{C}$. Say that $\sum_{n=1}^{\infty} f_n$ converges (uniformly) if the sequence $s_n := \sum_{j=0}^{n} f_n$ converges (uniformly).

Proposition. *$\sum_{n=0}^{\infty} f_n$ converges uniformly if and only if $\sum_{n=0}^{\infty} f_n$ satisfies a uniform Cauchy criterion, which means $\forall \epsilon > 0$, $\exists N$, s.t. $\forall n \geq m \geq N$ and $\forall x \in X$,*

$$\left| \sum_{j=m}^{n} f_j(x) \right| < \epsilon.$$

Note. Usually, it is much easier to show the convergence a sequence by showing that it is Cauchy.

Theorem 16 (Weierstress M test)**.** *Let $\{M_n\}$ be a sequence in $[0, \infty)$. Let f_n be a sequence of functions mapping from $X \neq \emptyset$ into $\mathbb{C}$. Assume that $\forall n \in \mathbb{N}$ and $x \in X$, $|f_n(x)| \leq M_n$. Then, if $\sum_{n=0}^{\infty} M_n$ converges, then $\sum_{n=0}^{\infty} f_n$ converges uniformly on X.*

Proof. Show $\sum_{n=0}^{\infty} f_n$ satisfies the uniform Cauchy criterion:

$$\left| \sum_{j=m}^{n} f_j(x) \right| \leq \sum_{j=m}^{n} |f_j(x)| \leq \sum_{j=m}^{n} M_j.$$

Since $\sum_{n=0}^{\infty} M_n$ is Cauchy, we are done. □

Theorem 17 (Root test)**.** *Let $\{c_n\}$ be a complex sequence in $\mathbb{C}$. Let $L := \limsup |c_n|^{1/n}$. If $L < 1$, then $\sum c_n$ converges absolutely. If $L > 1$, then $\sum c_n$ diverges (badly).*

Recall. Let $\{c_n\}$ be a sequence in $\mathbb{R}$. Then

$$\limsup s_n = \sup\{\text{subsequence limits of } \{s_n\}\} = \lim_{n\to\infty} \sup\{s_k : k \geq n\}.$$

Proof. If $L < 1$, then $\frac{L+1}{2} > L$. So $\exists N$, s.t. $\forall n \geq N$, $|c_n|^{1/n} < \frac{L+1}{2}$. Thus $|c_n| < \left(\frac{L+1}{2}\right)^n$ for $n \geq N$ and $\sum \left(\frac{L+1}{2}\right)^n$ converges since $\frac{L+1}{2} < 1$. By comparison, $\sum |c_n|$ converges.
If $L > 1$, define $L' := \min\{\frac{L+1}{2}, 2\}$. Since $L' < L$, $|c_n|^{1/n} > L'$ infinitely often. Thus $|c_n| > (L')^n$ infinitely often and since $L' > 1$, c_n cannot converge to zero. So $\sum c_n$ diverges. □

Definition. Let $\{c_n\}$ be a complex sequence. The radius of convergence of the power series $\sum_{n=0}^{\infty} c_n(z - z_0)^n$ is

$$R := \liminf |c_n|^{-1/n} = \left(\limsup |c_n|^{1/n}\right)^{-1}.$$

Theorem 18. *Let R denote the radius of convergence of the complex power series $\sum_{n=0}^{\infty} c_n(z-z_0)^n$. Then $\forall R' < R$, $\sum_{n=0}^{\infty} c_n(z - z_0)^n$ converges absolutely uniformly on $\{z \in \mathbb{C} : |z - z_0| \leq R'\}$ and $\sum_{n=0}^{\infty} c_n(z - z_0)^n$ diverges on $\{z \in \mathbb{C} : |z - z_0| > R\}$.*

Proof. If $|z - z_0| \leq R' < R$,

$$\limsup |c_n(z - z_0)^n|^{1/n} = \limsup |c_n|^{1/n} \cdot |z - z_0| \leq \frac{1}{R} \cdot R'.$$

So eventually, $|c_n| \cdot |z - z_0|^n < \alpha^n$ for some $\alpha \in (R'/R, 1)$. Outside $\{|z - z_0| \leq R\}$, root test shows divergence. $\square$

Lemma. *The series*

$$f^{(k)}(y) = \sum_{n=k}^{\infty} c_n n(n-1) \cdots (n-k+1)(x-a)^{n-k},$$

has the same radius of convergence as $f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$*.*

Proof. We want to show

$$\lim_{n\to\infty} (n(n-1)\cdots(n-k+1))^{1/n} = 1.$$

Since $\lim_{n\to\infty} (n(n-1)\cdots(n-k+1))^{1/n} \geq 1$, it sufficies to show that the $\limsup$ is ≤ 1. But

$$\lim_{n\to\infty} \sup (n(n-1)\cdots(n-k+1))^{1/n} \leq \lim_{n\to\infty} n^{k/n} = (\lim_{n\to\infty} n^{1/n})^k = 1.$$

$\square$

Theorem 19. *Let R be the radius of convergence of the power series $\sum_{n=0}^{\infty} c_n(x-a)^n$. Assume $R > 0$ and define $f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$ on $I := (a-R, a+R)$. Then f is indefinitely differentiable on I and we can differentiate it term by term:*

$$f^{(k)}(y) = \sum_{n=k}^{\infty} c_n n(n-1) \cdots (n-k+1)(x-a)^{n-k}.$$

Proof. Let $g_N(x) = \sum_{n=0}^{N} c_n (x-a)^n$. Let $0 < R' < R$. By the lemma and previous theorem, $\{g_N\}$ is a Cauchy sequence in $\mathcal{C}^k((a-R', a+R'))$. Thus, $\{g_N\}$ converges in $\mathcal{C}^k((a-R', a+R'))$. By uniqueness of limits, the limit must be f. Thus f is differentiable to order k on $(a-R', a+R')$ and

$$f^{(k)}(x) = \lim_{N\to\infty} g_N^{(k)}(x) = \lim_{N\to\infty} \sum_{n=0}^{N} c_n (n \cdots (n-k+1))(x-a)^{n-k}.$$

Since k is arbitrary and R' is arbitrary, we are done. $\square$

Corollary. *Under the hypotheses of the theorem, $f^{(k)}(a) = k! c_k$, i.e. $c_k = \frac{1}{k!} f^{(k)}(a)$.*

Theorem 20. *Let I denote the set of all points x at which $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges. Then*

$$f(x) := \sum_{n=0}^{\infty} c_n (x-a)^n$$

is continuous on I.

Proof. First, we do some reductions:

- By replacing $f(x)$ with $f(x+a)$, we can assume $a=0$.
- Also, without loss of generality, we just need to prove for $0 < R < \infty$. This is because $R=0$, then there is only one point at which the infinity sum converges. If $R=\infty$, then $I=\mathbb{R}$. So we can replace $f(x)$ with $f(Rx)$ and assume $R=1$.
- It suffices to prove continuity at each interval and replacing $f(x)$ with $f(-x)$, it sufficies to prove the following theorem.

□

Theorem 21. *If $\sum c_n$ converges, then $\sum c_n x^n$ converges $\forall |x| < 1$ and $\lim_{x\to 1^-} \sum c_n x^n = \sum c_n$.*

Proof. Define $s_{-1} := 0$, $s_n := \sum_{j=0}^{n} c_j$, $s := \sum_{j=0}^{\infty} c_j$. Define

$$f(x) := \sum_{n=0}^{\infty} c_n x^n, \quad x \in (-1, 1].$$

Then we can write

$$\sum_{n=0}^{m} c_n x^n = \sum_{n=0}^{m} (s_n - s_{n-1}) x^n = \sum_{j=0}^{m-1} s_j (x^j - x^{j+1}) + s_m x^m.$$

If $|x| < 1$, $s_m x^m \to 0$ as $x \to \infty$ since s_m is bounded. Thus, for $|x| < 1$,

$$f(x) = \sum_{j=0}^{\infty} s_j (x^j - x^{j+1}) = (1-x) \sum_{j=0}^{\infty} s_j x^j.$$

Since $\sum_{j=0}^{\infty} x^j = 1/(1-x)$, we can write

$$\sum_{n=0}^{\infty} c_n x^n - \sum_{n=0}^{\infty} c_n = f(x) - f(1) = (1-x) \sum_{j=0}^{\infty} s_j x^j - s(1-x) \sum_{j=0}^{\infty} x^j = (1-x) \sum_{j=0}^{\infty} (s_j - s) x^j.$$

Let $\epsilon > 0$. Choose N, s.t. $\forall j \geq N$, $|s_j - s| < \epsilon$. Thus,

$$\left| (1-x) \sum_{j=N}^{\infty} (s_i - s) x^j \right| < (1-x) \sum_{j=N}^{\infty} \epsilon |x|^j < \epsilon \cdot \frac{1-x}{1-|x|} < \epsilon,$$

if $0 \leq x \leq 1$. Since $\sum c_n$ converges, $\sum c_n x^n$ also converges. Furthermore, let $\delta := \epsilon / \sum_{j=0}^{N-1} |s - s_j|$. Then if $1 - \delta < x < 1$,

$$\left| (1-x) \sum_{j=0}^{N-1} (s - s_j) x^j \right| \leq (1-x) \sum_{j=0}^{N-1} |s - s_j| < \delta \cdot \sum_{j=0}^{N-1} |s - s_j| = \epsilon.$$

Combining with the previous result, we obtain

$$|f(x) - f(1)| = \left| (1-x) \sum_{j=0}^{\infty} (s_j - s) x^j \right| < 2\epsilon.$$

□

Definition. The Cauchy product of the two series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ is the series $\sum_{n=0}^{\infty} c_n$, where $c_n \coloneqq a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0$.

Theorem 22. *Assume $\sum a_n$ converges absolutely and $\sum b_n$ converges. Then their Cauchy product $\sum c_n$ converges. Furthermore, it converges to $\sum a_n \cdot \sum b_n$. If $\sum b_n$ converges absolutely as well, then $\sum c_n$ converges absolutely.*

Proof. Some notations: $A_N \coloneqq \sum_{n=0}^{N} a_n$, $B_N \coloneqq \sum_{n=0}^{N} b_n$, $C_N \coloneqq \sum_{n=0}^{N} c_n$, $A \coloneqq \sum a_n$, $B \coloneqq \sum b_n$, $\bar{A} = \sum |a_n|$, $M \coloneqq \sup B_N$.

$$\begin{aligned} C_N &= c_0 + \cdots + c_N \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_N + \cdots + a_N b_0) \\ &= a_0 B_N + a_1 B_{N-1} + \cdots + a_N B_0 \\ &= A_N B + a_0 (B_N - B) + \cdots + a_N (B_0 - B) \end{aligned}$$

Define $\Gamma_N \coloneqq a_0(B_N - B) + \cdots + a_N(B_0 - B)$. If we can show that $\lim_{N\to\infty} \Gamma_N = 0$, it follows that $\lim_{n\to\infty} C_n = AB$.

Let $\epsilon > 0$. By absolute convergence of $\sum a_n$, $\exists N$, s.t. $\forall n \geq m \geq N_1$, $\sum_{j=m}^{n} |a_j| < \epsilon/2M$. Since $B_n \to B$, $\exists N_2$, s.t. $\forall n > N_2$, $|B_n - B| < \epsilon/\bar{A}$. Finally, if $N \geq N_1 + N_2$,

$$\begin{aligned} |\Gamma_N| &\leq \sum_{j=0}^{N} |a_j| \cdot |B_{N-j} - B| \\ &= \sum_{j=0}^{N_1} |a_j| \cdot |B_{N-j} - B| + \sum_{j=N_1+1}^{N} |a_j| \cdot |B_{N-j} - B| \\ &< \sum_{j=0}^{N_1} |a_j| \cdot \frac{\epsilon}{\bar{A}} + \sum_{j=N_1+1}^{N} |a_j| \cdot 2M \\ &< \bar{A} \cdot \frac{\epsilon}{\bar{A}} + \frac{\epsilon}{2M} \cdot 2M \\ &= 2\epsilon \end{aligned}$$

Now we want to show that $\sum c_n$ converges absolutely if $\sum b_n$ converges absolutely. Note,

$$\begin{aligned} \sum_{n=0}^{N} |c_n| &= |c_0| + \cdots + |c_N| \\ &= |a_0 b_0| + |a_0 b_1 + a_1 b_0| + \cdots + |a_0 b_N + \cdots + a_N b_0| \\ &\leq |a_0| \cdot |b_0| + (|a_0| \cdot |b_1| + |a_1| \cdot |b_0|) + \cdots + (|a_0| \cdot |b_N| + \cdots + |a_N| \cdot |b_0|) \end{aligned}$$

Since $\sum |a_n|$ and $\sum |b_n|$ are both convergent, we are done. $\square$

Example. Consider the Cauchy product of the two absolutely convergent power series

$$\begin{aligned} &\sum_{n=0}^{\infty} a_n \quad \text{with} \quad a_n = \frac{1}{n!} x^n, \\ &\sum_{n=0}^{\infty} b_n \quad \text{with} \quad b_n = \frac{1}{n!} y^n. \end{aligned}$$

By definition, we have

$$\begin{aligned}
c_n &= a_0b_n + \cdots + a_nb_0 \\
&= \sum_{j=0}^{n} a_j b_{n-j} \\
&= \sum_{j=0}^{n} \frac{1}{j!}x^j \frac{1}{(n-j)!}y^{n-j} \\
&= \frac{1}{n!}\sum_{j=0}^{n} \frac{n!}{j!(n-j)!}x^j y^{n-j} \\
&= \frac{1}{n!}(x+y)^n
\end{aligned}$$

Upshot:

$$\left(\sum_{n=0}^{\infty} \frac{1}{n!}x^n\right)\left(\sum_{n=0}^{\infty} \frac{1}{n!}y^n\right) = \sum_{n=0}^{\infty} \frac{1}{n!}(x+y)^n,$$

i.e. $e^x e^y = e^{x+y}$.

Unordered Series

Definition. Let S be a set and $\{a_x\}_{s\in S}$ be a function from S into $\mathbb{R}$. We say that the unordered series $\sum_{s\in S} a_s$ converges to $b \in \mathbb{R}$ if $\forall \epsilon > 0$, $\exists$ a finite set $S_\epsilon \subseteq S$, s.t. $\forall$ finite set S' with $S_\epsilon \subseteq S' \subseteq S$,

$$\left|\sum_{s\in S'} a_s - b\right| < \epsilon.$$

Proposition. *An unordered series can have at most one sum.*

Theorem 23. *The following are equivalent:*

1. *The unordered series $\sum_{s\in S} a_s$ converges.*
2. *$\forall \epsilon > 0$, $\exists$ a finite set $S_\epsilon \subseteq S$, s.t. $\forall$ finite set $S' \subseteq S \setminus S_\epsilon$, $\sum_{s\in S'} |a_s| < \epsilon$.*
3. *$\sum_{s\in S} |a_s|$ converges absolutely.*
4. *$\sup\{\sum_{s\in S'} |a_s| : S' \subseteq S \text{ is finite}\} < \infty$.*

Proof. **1 $\Rightarrow$ 2:** Assume $\sum_{s\in S} a_s$ converges to b and let $\epsilon > 0$. Then $\exists$ a finite set $S_\epsilon \subseteq S$, s.t. $\forall$ finite set S' with $S_\epsilon \subseteq S' \subseteq S$, $\left|\sum_{s\in S'} a_s - b\right| < \epsilon$. Let $S'' \subseteq S \setminus S_\epsilon$ be a finite set. Now let

$S''_+ \coloneqq \{s \in S'' : a_s > 0\}$ and $S''_- \coloneqq \{s \in S'' : a_s < 0\}$. Then

$$\begin{aligned}\left|\sum_{s\in S''}|a_s|\right| &= \left|\sum_{s\in S''_+} a_s - \sum_{s\in S''_-} a_s\right| \\ &= \left|\sum_{s\in S''_+\cup S_\epsilon} a_s - \sum_{s\in S_\epsilon} a_s - \sum_{s\in S''_-\cup S_\epsilon} a_s + \sum_{s\in S\epsilon} a_s\right| \\ &\le \left|\sum_{s\in S''_+\cup S_\epsilon} a_s - b\right| + 2\left|\sum_{s\in S_\epsilon} a_s - b\right| + \left|\sum_{s\in S_\epsilon\cup S''_-} a_s - b\right| \\ &< 4\epsilon.\end{aligned}$$

$2 \Rightarrow 4$: $\sup\{\sum_{s\in S'}|a_s| : S' \subseteq S \text{ is finite}\} \le \epsilon + \sum_{s\in S_\epsilon}|a_s| < \infty$.

$4 \Rightarrow 3$: Let $B \coloneqq \sup\{\sum_{s\in S'}|a_s| : S' \subseteq S \text{ is finite}\}$. We want to show that $\sum_{s\in S}|a_s| = B$. Let $\epsilon > 0$. By the definition of sup, there exists a finite subset $S_\epsilon \subseteq S$ such that $\sum_{s\in S_\epsilon}|a_s| > B-\epsilon$. So if S' is a finite set with $S_\epsilon \subseteq S' \subseteq S$,

$$B - \epsilon < \sum_{s\in S_\epsilon}|a_s| \le \sum_{s\in S'}|a_s| \le B.$$

Thus, by definition, the series $\sum_{s\in S}|a_s|$ converges.

$3 \Rightarrow 2$: The argument is similar as the argument for showing $1 \Rightarrow 2$.

$3 \Rightarrow 1$: Suppose for contradiction that the unordered series $\sum_{s\in S} a_s$ does not converge. This means for all $b \in \mathbb{R}$, $\exists \epsilon > 0$, s.t. $\forall$ finite $S_\epsilon \subseteq S$, $\exists$ a finite set S' with $S_\epsilon \subseteq S' \subseteq S$, s.t.

$$\left|\sum_{s\in S'} a_s - b\right| > \epsilon.$$

This implies

$$\sum_{s\in S'} a_s - b > \epsilon, \quad \text{or} \quad \sum_{s\in S'} a_s - b < -\epsilon,$$

which further implies

$$b - \epsilon > \sum_{s\in S'} a_s > b + \epsilon.$$

Now, let's choose b to be the limit of $\sum_{s\in S}|a_s|$. We have

□

Lemma. *If $\sum_{s\in S}|b_s|$ converges and $|a_s| \le |b_s|$ for all s. Then $\sum_{s\in S}|a_s|$ converges.*

Proof. Let $\epsilon > 0$. If $\sum_{s\in S}|b_s|$ converges, by the theorem, there exists a finite set $S_\epsilon \subseteq S$, s.t. $\forall$ finite set $S' \subseteq S \setminus S_\epsilon$, $\sum_{s\in S'}|b_s| < \epsilon$. Then $\forall$ finite set $S' \subseteq S \setminus S_\epsilon$, $\sum_{s\in S'}|a_s| \le \sum_{s\in S'}|b_s| < \epsilon$. Then $\sum_{s\in S}|a_s|$ converges. □

Corollary. *The unordered series $\sum_{n\in\mathbb{N}} a_n$ converges if and only if $\sum_{n=1}^{\infty} a_n$ converges absolutely.*

Proposition (HW)**.** *Show directly (without the theorem), that if $\lambda \in \mathbb{R}$ and $\sum_{s\in S} a_s$, $\sum_{s\in S} b_s$ converges, then*

$$\sum_{s\in S} \lambda a_s = \lambda \sum_{s\in S} a_s$$

and

$$\sum_{s\in S} a_s + b_s = \sum_{s\in S} a_s + \sum_{s\in S} b_s$$

Proposition. *The unordered series $\sum_{(i,j)\in\mathbb{N}^2} a_{i,j}$ converges if and only if $\sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} |a_{i,j}|\right)$ converges. In this case,*

$$\sum_{(i,j)\in\mathbb{N}^2} a_{i,j} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{i,j} = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{i,j}.$$

Proof. Idea: $\Leftrightarrow$ follows directly from $\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} |a_{i,j}| = \lim_{N\to\infty}\lim_{M\to\infty}\sum_{j=1}^{N}\sum_{i=1}^{M} |a_{i,j}|$. For the identity, consider the positive and negative parts separately. $\square$

Definition. Let $G \subseteq \mathbb{R}$ be an open set and let $f : G \to \mathbb{R}$. Say that f is analytic on G if $\forall a \in G$, $\exists \epsilon > 0$, s.t. $f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x-a)^n$, on $(a-\epsilon, a+\epsilon)$ (i.e. f has a power series representation on $(a-\epsilon, a+\epsilon)$).

Theorem 24. *f is analytic on the open set $G \subseteq \mathbb{R}$ if and only if G can be written as a union of open intervals at which f has a power series representation.*

Note. Every open subset of $\mathbb{R}$ is a countable union of disjoint open intervals.

Theorem 25. *Suppose that $f(x) = \sum_{n=0}^{\infty} c_n x^n$ on $\{|x| < R\}$, some $R > 0$. If $|a| < R$, then f has a power series representation centered at a converging on $|x-a| < R - |a|$.*

Proof. Notice that

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a+a)^n = \sum_{n=0}^{\infty}\sum_{j=0}^{n} c_n \binom{n}{j} a^{n-j}(x-a)^j.$$

Observe that

$$\sum_{n=0}^{\infty}\sum_{j=0}^{n} \left| c_n \binom{n}{j} a^{n-j}(x-a)^j \right| = \sum_{n=0}^{\infty} |c_n|(|x-a| + |a|)^n.$$

Converges on $|x-a| + |a| < R$. So by proposition, we can switch the order of summation,

$$f(x) = \sum_{j=0}^{\infty} \left(\sum_{n=j}^{\infty} c_n \binom{n}{j} a^{n-j} \right) (x-a)^j, \quad |x-a| < R - |a|,$$

and $\sum_{n=j}^{\infty} c_n \binom{n}{j} a^{n-j}$ is the coefficient $c_j(a)$. In particular, the sum representing $c_j(a)$ converges absolutely $\forall |a| < R$ and

$$c_j(a) = \frac{1}{j!} f^{(j)}(a).$$

$\square$

Theorem 26. *If f and g are analytic on the open interval I, then so are $f+g$ and $f \cdot g$. In particular, if f and g both have power series representations centered at a on $(a-R, a+R)$, then so do $f+g$ and $f \cdot g$.*

Theorem 27. *If f is analytic on the open interval I, g is analytic on the open interval J, and $g(J) \subseteq I$. Then $f \circ g$ is analytic on J.*

Proof. Let $a \in J$. By translating and adding a constant to g, we can assume $a = g(a) = 0$. Now we expand

$$f(y) = \sum_{n=0}^{\infty} b_n y^n, \quad g(x) = \sum_{k=0}^{\infty} c_k x^k, \quad \text{on } |y| < \epsilon, |x| < \delta.$$

Define $\bar{g}(x) := \sum_{k=0}^{\infty} |c_k| x^k$. The $\bar{g}(x)$ is continuous on $|x| < \delta$. So by shrinking δ if needed, we may assume that $|\bar{g}(x)| < \epsilon$, $|x| < \delta$.

$$f \circ g(x) = \sum_{n=0}^{\infty} b_n (g(x))^n.$$

By previous theorem and induction, for each n,

$$(g(x))^n = \sum_{k=0}^{\infty} a_k^{(n)} x^k, \quad (\bar{g}(x))^n = \sum_{k=0}^{\infty} \bar{a}_k^{(n)} x^k, \quad \text{on } |x| < \delta.$$

Furthermore, $|a_k^{(n)}| \leq \bar{a}_k^{(n)}$. Therefore,

$$f \circ g(x) = \sum_{n=0}^{\infty} b_n \sum_{k=0}^{\infty} a_k^{(n)} x^k.$$

We want to switch the order of summation, but it requires absolute convergence. Now

$$\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |b_n a_k^{(n)} x^k| \leq \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |b_n| \cdot \bar{a}_k^{(n)} \cdot |x|^k = \sum_{n=0}^{\infty} |b_n| (\bar{g}(|x|))^n,$$

which converges since $\bar{g}(|x|) < \epsilon$. So

$$f \circ g(x) = \sum_{k=0}^{\infty} \left(\sum_{n=0}^{\infty} b_n a_k^{(n)} \right) x^k.$$

□

Corollary. *If f and g are analytic on the open interval I and $g \neq 0$ on I. Then f/g is analytic on I.*

Proof. Since $1/x$ is analytic on $\mathbb{R} \setminus \{0\}$, by the previous theorem, $1/g$ is analytic. And so is $f \cdot 1/g = f/g$. □

Theorem 28. *Let $E \subseteq \mathbb{R}$. Then $[x, y] \subseteq E$ for all $x, y \in E$ if and only if $E = \emptyset$, E is a singleton (i.e. E is a single point), or E is an interval.*

Proof. $\Leftarrow$: This direction is trivial.

$\Rightarrow$: Assume $\forall x, y \in E$, $[x, y] \subseteq E$. We may assume E has at least 2 points. Thus,

$$\alpha := \inf E < \sup E =: \beta.$$

Claim that $(\alpha, \beta) \subseteq E$. If so, we're done, since $E \subseteq [\alpha, \beta]$.
Let $x \in (\alpha, \beta)$. Then $\exists y \in E$, s.t. $y < x$ and $\exists z \in E$, s.t. $z > x$. By hypothesis, $x \in [y, z] \subseteq E$.

$\square$

Theorem 29. *$E \subseteq \mathbb{R}$ is connected if and only if E is empty, E is singleton, or E is an interval.*

Notice, by the last theorem, the R.H.S. is equivalent to $[x, y] \subseteq E$ for all $x, y \in E$.

Proof. $\Rightarrow$: It suffices to prove the contrapositive: If $E \neq \emptyset, E \neq \{x\}$ and E is not an interval, then $\exists x, y \in E$ and z with $x < z < y$ s.t. $z \notin E$. Then $E \cap (-\infty, z)$ and $E \cap (z, \infty)$ form a separation of E. So E is not connected.

$\Leftarrow$: Assume that $E = \emptyset$ or $E \neq \{x\}$ or E is an interval and E is not connected. Since a set that isn't connected must contain at least 2 points, E is an interval. Now we fix a separation $E = A \cup B$. By definition, $A, B \neq \emptyset$ and $\bar{A} \cap B = A \cap \bar{B} = \emptyset$. Fix $x \in A$ and $y \in B$. Since $A \cap B = \emptyset$, $x \neq y$. So we may assume $x < y$ (otherwise, just rename A and B). Define $z := \sup A \cap [x, y]$. Then, $z \geq x$ because $x \in A \cap [x, y]$ and $z \leq y$ because y is an upper bound of A. So $z \in [x, y] \subseteq E = A \cup B$. Furthermore $z \in \overline{A \cap [x, y]} \subseteq \bar{A}$. Since $\bar{A} \cap B = \emptyset$, $z \notin B$. Then $z \in A$. Since $A \cap \bar{B} = \emptyset$, $z \notin \bar{B}$. So $\exists r > 0$, s.t. $(z - r, z + r) \cap B = \emptyset$. But $y \geq z$ and $y \in B$, so $z + r \leq y$. So $z + r/2 \in [z, y] \cap B^c \subseteq [x, y] \cap A$. But z was an upper bound for $[x, y] \cap A$ and $z + r/2 > z$. Contradiction!

$\square$

Theorem 30. *Let I be an open interval and assume that f and g are analytic on I. Let $E := \{x \in I : f(x) = g(x)\}$. If E has an accumulation point in I, then $E = I$, i.e. $f \equiv g$ on I.*

Note, it is extremely important for I to be an open interval and the assumption that E has an accumulation point in I.

Proof. By replacing f with $f - g$, we may assume that $g \equiv 0$ on I. Since f is continuous, E is closed in I. Let E' be a set of accumulation points of E. Therefore $E' \cap I \subseteq E$. Then $E' \cap I$ is closed in I.

Claim: $E' \cap I$ is open.
Let $a \in E' \cap I$. Then for some $\epsilon > 0$, $f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x - a)^n$, on $|x - a| < \epsilon$. Since $a \in E' \cap I \subseteq E$, $f(a) = f^{(0)}(a) = 0$. Suppose $f(a) = f'(a) = \cdots = f^{(k)}(a) = 0$ and $f^{(k+1)}(a) \neq 0$ for some $k \geq 0$. By Taylor's theorem with remainder,

$$f(x) = \frac{1}{(k+1)!} f^{(k+1)}(a)(x - a)^{k+1} + \frac{1}{(k+2)!} f^{(k+2)}(t_{x,a})(x - a)^{k+2},$$

for some $t_{x,a}$ between a and x. By continuity of $f^{(k+2)}$, there exists $\delta > 0$, s.t. $\forall |x - a| < \delta$ and $|t - a| < \delta$,

$$\frac{1}{(k+2)!} \left| f^{(k+2)}(t) \right| \cdot |x - a| < \frac{1}{2} \cdot \frac{1}{(k+1)!} \cdot \left| f^{(k+1)}(a) \right|.$$

Thus for $0 < |x-a| < \delta$,

$$|f(x)| \geq \frac{1}{2} \cdot \frac{1}{(k+1)!} \cdot \left|f^{(k+1)}(a)\right| \cdot |x-a|^{k+1} \neq 0.$$

So $a \notin E'$. Contradiction. So $f^{(n)}(a) = 0$ for all n. So $f \equiv 0$ on $(a-\epsilon, a+\epsilon)$. So $(a-\epsilon, a+\epsilon) \subseteq E' \cap I$. Since a was arbitrary, $E' \cap I$ is open.

Then we conclude that $E' \cap I$ is closed and open. Since I is connected (I is an open interval), $E' \cap I = \emptyset$ or $E' \cap I = I$. In the latter case, $E = I$. □

The exponential function

Properties. Let $E(z) \coloneqq \sum_{n=0}^{\infty} \frac{1}{n!} z^n$, $z \in \mathbb{C}$. By root test, its radius of convergence is ∞ and it has the following properties:

1. $E(z) \cdot E(w) = E(z+w)$ for every $z, w \in \mathbb{C}$.
2. $E(z) \neq 0$ and $E(-z) = \frac{1}{E(z)}$ for every $z \in \mathbb{C}$.
3. $E(x) > 0$ for every $x \in \mathbb{R}$.
4. $E'(x) = E(x)$ for every $x \in \mathbb{R}$.

Proof.

1. By Cauchy products.
2. $E(z)E(-z) = E(0) = 1$.
3. For $x \geq 0$, $E(x) \geq 1$ by inspection. For $x < 0$, $E(x) = \frac{1}{E(-x)} > 0$.
4. Differentiate term by term.

□

Definition. $e \coloneqq E(1)$.

Proposition. $E(x) = e^x$, $\forall x \in \mathbb{R}$.

Proof. For $x = n \in \mathbb{N}$, this follows from property 1 and induction. $E(n) = E(1 + \cdots + 1) = (E(1))^n = e^n$. For $x = n \in \mathbb{Z}$, this follows from preceeding and property 2. For $x = \frac{n}{m}$ with $n \in \mathbb{Z}$, $m \in \mathbb{N}$, we know that

$$E\left(\frac{n}{m}\right)^m = E(n)$$

by property 1. So by property 2 and uniqueness of m-th roots, $E(n/m) = E(n)^{1/m} = e^{n/m}$. So the conclusion holds for all rationals. Finally, for $x \in \mathbb{R}$, e^x is (by definition, since $e > 1$) $\sup\{e^p : p \in \mathbb{Q} \text{ and } p \leq x\}$. So the general case of proposition follows from continuity of E. □

Theorem 31.

1. *e^x is continuous and differentiable on $\mathbb{R}$ and $\frac{\mathrm{d}}{\mathrm{d}x} e^x = e^x$.*
2. *e^x is positive and strictly increasing on $\mathbb{R}$.*
3. *$\lim_{x\to\infty} e^x = \infty$ and $\lim_{x\to-\infty} e^x = 0$.*

4. *For fixed n, $\lim_{x\to\infty} e^x x^{-n} = \infty$ and $\lim_{x\to\infty} e^{-x}x^n = 0$.*

Proof. 1. Proved by the 4th statement in the proposition.

2. Combine the 3th statement in the proposition and the fact that $E(x) = e^x$, $e^x > 0$. Also, since $E'(x) = E(x) > 0$, so e^x is strictly increasing.

3. It is clear that $\lim_{x\to\infty} e^x = \lim_{x\to\infty} E(x) = \infty$. Also by inspection,

$$\lim_{x\to-\infty} e^x = \lim_{x\to\infty} e^{-x} = \lim_{x\to\infty} \frac{1}{e^x} = 0.$$

4. For $x \geq 0$, $e^x \geq \frac{1}{(n+1)!}x^{n+1}$. So $e^x x^{-n} \geq \frac{1}{(n+1)!}x \to \infty$. On the other hand, $\lim_{x\to\infty} e^{-x}x^n = \lim_{x\to\infty} \frac{1}{e^x x^{-n}} = 0$.

$\square$

Lemma. *Let I and J be intervals. Let $f : I \to J$ be a continuous bijection onto J. Then f has a continuous inverse. Furthermore, if $y \in int(J)$ and f is differentiable at $f^{-1}(y)$ with $f'(f^{-1}(y)) \neq 0$. Then f^{-1} is differentiable at y and*

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

Note. The conclusion is not true if I is an open interval and J is a subset of an arbitrary metric space. For example, consider $I = \mathbb{R}$ and $J = 8$. Map $\mathbb{R}$ into a bounded open interval (s.t. arctan).

Remarks. If K is compact and $f : K \to J$ is a continuous bijection, then f^{-1} is continuous.

Proof. f is either strictly increasing or strictly decreasing. Without loss of generality, we may assume f is strictly increasing. Let $g := f^{-1}$. Then g is strictly increasing.

Claim 1: g is continuous.
Let $y \in J$. Then both $g(y^+) = \lim_{t\to y^+} g(t)$ and $g(y^-) = \lim_{t\to y^-} g(t)$ exists (modify as needed if y is an end point). Furthermore, we know $g(y^-) \leq g(y) \leq g(y^+)$. If $g(y^-) = g(y) = g(y^+)$, then g is continuous at g. Let's suppose $g(y^-) < g(y)$. Because g is increasing, then $g(t) \leq g(y^-)$ for all $t < y$ and $g(t) \geq g(y)$ for all $t \geq y$. So g omits in $(g(y^-), g(y)) \subseteq I$. But I is the domain of f, which is the range of g. Contradiction. So $g(y^-) = g(y)$. Similarly, $g(y^+) = g(y)$.

Clam 2: If $y \in \text{int}(J)$ and f is differentiable at $f^{-1}(y)$ with $f'(f^{-1}(y)) \neq 0$. Then f^{-1} is differentiable at y.
We want to evaluate

$$\lim_{h\to 0} \frac{g(y+h) - g(y)}{y+h-y} = \lim_{h\to 0} \frac{g(y+h) - g(y)}{f(g(y+h)) - f(g(y))}.$$

Define

$$\varphi(t) = \begin{cases} \dfrac{t - g(y)}{f(t) - f(g(y))}, & t \neq g(y), \\ \dfrac{1}{f'(g(y))}, & t = g(y). \end{cases}$$

Then $\frac{1}{\varphi}$ is continuous on I and $\frac{1}{\varphi(g(y))} = f'(g(y)) \neq 0$. So φ is continuous on a nerighbourhood of $g(y)$. So $\varphi \circ g$ is continuous at y, i.e.

$$\lim_{h\to 0} \frac{g(y+h) - g(y)}{f(g(y+h)) - f(g(y))} = \lim_{h\to 0} \varphi(g(y+h)) = \varphi(g(y)) = \frac{1}{f'(g(y))}.$$

□

Definition. Since $E : x \mapsto e^x$ is continuous, differentiable, strictly increasing and maps $\mathbb{R}$ onto $(0, \infty)$. Thus it has an inverse, which we call log, that is continuous, differentiable, strictly increasing and maps $(0, \infty)$ onto $\mathbb{R}$.

Properties.

1. $e^{\log y} = y$ for all $y > 0$ and $\log(e^x) = x$ for all $x \in \mathbb{R}$.
2. $\frac{\mathrm{d}}{\mathrm{d}x} \log y = \frac{1}{y}$, $\forall y > 0$.
3. $\log(1) = 0$.
4. $\log(y) = \int_1^y 1/s \,\mathrm{d}x$, $\forall y > 0$.
5. $\log(uv) = \log u + \log v$, $\forall u, v > 0$.
6. $\log(y) \to +\infty$ as $y \to +\infty$. $\log y \to -\infty$ as $y \to 0$.
7. For $u > 0$ and $\alpha \in \mathbb{R}$, $\log(u^\alpha) = \alpha \log(u)$ and $u^\alpha = e^{\alpha \log u}$.

Proof.

1. Done.
2. By the lemma, log is differentiable and $\frac{\mathrm{d}}{\mathrm{d}y} \log(y) = \frac{1}{E'(\log(y))} = \frac{1}{E(\log(y))} = \frac{1}{y}$.
3. Because $E(0) = 1$, done.
4. By property two and three, done.
5. $\log(uv) = \log(e^{\log u} \cdot e^{\log v}) = \log(e^{\log u + \log v}) = \log u + \log v$.
6. Done.
7. $\log(u^\alpha) = \log((e^{\log u})^\alpha) = \log(e^{\alpha \log u}) = \alpha \log u$.

□

Proposition. *For any $a > 0$, $x \mapsto a^x$ is differentiable on $\mathbb{R}$ and*

$$\frac{\mathrm{d}}{\mathrm{d}x} a^x = (\log a) \cdot a^x.$$

Proof. $a^x = e^{x \log a}$. By the chain rule, done. □

Proposition (HW). *If $\alpha \neq -1$, then $x \mapsto x^\alpha$ has $x \mapsto \frac{x^{\alpha+1}}{\alpha+1}$ as an antiderivative.*

Proposition. $\forall \epsilon > 0$, $0 = \lim_{x\to\infty} x^{-\epsilon} \log x = \lim_{x \to 0+} x^\epsilon \log x$.

This is saying log goes to infinity very slow.

Proof. Notice that

$$\lim_{x\to 0+} x^{\epsilon}\log x = \lim_{y\to\infty}\left(\frac{1}{y}\right)^{\epsilon}\log\left(\frac{1}{y}\right) = -\lim_{y\to\infty} y^{-\epsilon}\log y.$$

So it suffices to prove the first equation.

$$\lim_{x\to\infty} x^{-\epsilon}\log x = \lim_{t\to\infty}(e^t)^{-\epsilon}\log(e^t) = \lim_{t\to\infty} e^{-\epsilon t}t = 0.$$

□

Trigonometric Functions

Definition. Define

$$C(z) \coloneqq \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}z^{2n} \quad \text{and} \quad S(z) \coloneqq \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}z^{2n+1},$$

where $z \in \mathbb{C}$ (radius of convergence is $+\infty$).

Definition. Let $z \coloneqq x + iy$ where $x, y \in \mathbb{R}$. Recall that $\bar{z} = x - iy$. Define $\mathrm{Re}(z) \coloneqq x$ and $\mathrm{Im}(z) \coloneqq y$. Then

- $\mathrm{Re}(z) = (z + \bar{z})/2$,
- $\mathrm{Im}(z) = (z - \bar{z})/(2i)$,
- $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$,
- $\overline{z + w} = \bar{z} + \bar{w}$.

Note. For all $x \in \mathbb{R}$

$$\overline{E(ix)} = \overline{\sum_{n=0}^{\infty}\frac{1}{n!}(ix)^n} = \sum_{n=0}^{\infty}\overline{\frac{1}{n!}(ix)^n} = E(\overline{ix}) = E(-ix).$$

By this, we have

$$C(x) = \sum_{n=0}^{\infty}\frac{(ix)^{2n}}{(2n)!} = \mathrm{Re}\left(E(ix)\right) = \frac{1}{2}\left(E(ix) + E(-ix)\right).$$

$$S(x) = \frac{1}{i}\sum_{n=0}^{\infty}\frac{(ix)^{2n+1}}{(2n+1)!} = \mathrm{Im}\left(E(ix)\right) = \frac{1}{2i}(E(ix) - E(-ix)).$$

Properties. Trigonometric functions have the following properties:

1. $C(x)^2 + S(x)^2 = 1$, $\forall x \in \mathbb{R}$.
2. $C(0) = 1$ and $S(0) = 0$.

3. $C'(x) = -S(x)$, $S'(x) = C(x)$, $\forall x \in \mathbb{R}$.

Proof. 1. $C(x)^2 + S(x)^2 = |E(ix)|^2 = E(ix) \cdot \overline{E(ix)} = E(ix)E(-ix) = E(0) = 1$.

2. Done.

3. $C'(x) + iS'(x) = \frac{\mathrm{d}}{\mathrm{d}x}E(ix) = iE'(ix) = iE(ix) = iC(x) + i^2S(x)$. So we have $C'(x) = -S(x)$, $S'(x) = C(x)$.

□

Proposition. *C has a positive zero.*

Proof. Since $C(0) = 1 \neq 0$ and since $C(x) = C(-x)$, it suffices to show C has a zero. Suppose not. Since $C(0) > 0$ and C is continuous, $C(x) > 0$, $\forall x \in \mathbb{R}$. Then S is strictly increasing since $S' = C$. So for $0 < x < y$,

$$S(x)(y - x) = \int_x^y S(x)\,\mathrm{d}t < \int_x^y S(t)\,\mathrm{d}t = C(x) - C(y).$$

Since $C(x) \leq 1$ (because $C^2 + S^2 = 1$), and $C(y) > 0$, so $S(x) < \frac{1}{y-x}$, $\forall 0 < x < y$. Fix x and let $y \to \infty$, we see $S(x) = 0$. But S strictly increasing and $S(0) = 0$, so $S(x) > 0$. Contradiction! So C must has a zero. □

Definition. Define $\pi \coloneqq 2\inf\{x > 0 : C(x) = 0\}$. By continuity and $C(0) \neq 0$, $\pi > 0$ and the infimum is a minimum, i.e. $C(\pi/2) = 0$.

Theorem 32. *The following statements hold:*

1. *$C(k\pi) = (-1)^k$, $S(k\pi) = 0$, $C(k\pi + \pi/2) = 0$ and $S(k\pi + \pi/2) = (-1)^k$, $k \in \mathbb{Z}$.*

2. *$E(z + 2k\pi i) = E(z)$, $\forall k \in \mathbb{Z}$ and consequently C and S are periodic with period 2π.*

Proof. 1. Since $C(\pi/2) = 0$ and $C^2 + S^2 = 1$, we see $S(\pi/2) = \pm 1$. By definition of π and continuity of C and $C(0) > 0$, S is increasing on $[0, \pi/2]$ and since $S(0) = 0$, $S(\pi/2) = 1$. So $E(i\pi/2) = C(\pi/2) + iS(\pi/2) = 0 + i$. So $E(ik\pi/2) = i^k$ and all identities in 1 hold.

2. $E(z + 2\pi ik) = E(z)E(2\pi i)^k = E(z)$ by part 1.

□

Lemma. *If $E(it) = 1$ and $t \in [0, 2\pi)$, then $t = 0$.*

Proof. Suppose $t \in (0, 2\pi)$ and $E(it) = 1$. By the proof of part 1 of the previous theorem, $S \neq 0$ on $(0, \pi/2]$. Since $S(\pi/2 - t) = S(\pi/2 + t)$ (Exercise) so $S > 0$ on $(0, \pi)$. Since $S(t) = -S(-t) = -S(2\pi - t)$, $S < 0$ on $(\pi, 2\pi)$. So $t = \pi$. But $E(i\pi) = -1$. □

Theorem 33. *If $z \in \mathbb{C}$ with $|z| = 1$, then $\exists! t \in [0, 2\pi)$, s.t. $E(it) = z$.*

Proof. The above lemma proves the uniqueness of z. It remains to prove the existence. Let $z \in \mathbb{C}$ with $|z| = 1$. Write $z = x + iy$, $x, y \in \mathbb{R}$. Thus $x^2 + y^2 = 1$.

Case 1: $x \geq 0, y \geq 0$.

Then $0 \leq x \leq 1$ and since $\cos 0 = 1$ and $\cos \pi/2 = 0$, by the intermediate value theorem, $\exists t \in [0, \pi/2]$, s.t. $\cos t = x$. Furthermore, $y = \sqrt{1-x^2} = \sqrt{1-\cos^2 t} = \sqrt{\sin^2 t} = |\sin t|$. Since sin is increasing on $[0, \pi/2]$ (because $\sin' = \cos$ is nonnegative) and $\sin(0) = 0$, $\sin t \geq 0$ and $|\sin t| = \sin t$.

Case 2: $x \geq 0, y > 0$.

Then $\bar{z}$ is in the first quadrant. By Case 1, there exists $t \in (0, \pi/2]$, s.t. $e^{it} = \bar{z}$. Therefore, $z = e^{-it} = e^{i(2\pi - t)}$ and $2\pi - t \in [3\pi/2, 2\pi)$.

Case 3: $x < 0$.

Then by Case 1 and Case 2, $\exists t \in (-\pi/2, \pi/2)$, s.t. $-z = e^{it}$. Therefore $z = (-1)e^{it} = e^{i(\pi+t)}$ and $\pi + t \in (\pi/2, 3\pi/2)$.

□

Corollary. *The circumference of the unit circle is 2π.*

Proof.

$$\text{Circumference} = \int_0^{2\pi} |E'(it)| \, \mathrm{d}t = \int_0^{2\pi} |E(it)| \, \mathrm{d}t = \int_0^{2\pi} C^2(t) + S^2(t) \, \mathrm{d}t = 2\pi.$$

□

Corollary. *If $z \in \mathbb{C}$ with $z \neq 0$, then $\exists! t \in [0, 2\pi)$, s.t. $z = |z|e^{it}$.*

Theorem 34 (Algebraic Completeness of $\mathbb{C}$)**.** *Let $P(z) = a_0 + a_1 z + \cdots + a_n z^n$ be a complex polynomial with $a_n \neq 0$ and $n \geq 1$. Then there exists $z_0 \in \mathbb{C}$ such that $P(z_0) = 0$.*

Proof. Let $\mu = \inf_{z \in \mathbb{C}} |P(z)|$. We claim that μ is a minimum (it is achieved). Indeed, $|P(z)| \geq |a_n| \cdot |z|^n - \sum_{j=0}^{n-1} |a_j| \cdot |z|^j$. So $\exists R$, s.t. $\forall |z| > R$, $|P(z)| \geq \mu + 1$. Therefore by the continuity of $|P(z)|$, $\mu = \inf_{|z| \leq R} |P(z)| = \min_{|z| \leq R} |P(z)|$. Thus $\exists z_0 \in \mathbb{C}$, s.t. $|P(z_0)| = \mu$. If $\mu = 0$, we are done.

So now suppose $\mu > 0$. Define $Q(z) = \frac{P(z+z_0)}{P(z_0)}$. Then Q is a polynomial. $Q(0) = 1$ and $|Q(z)| \geq 1$ for all z (because $P(z_0)$ is the minimum). Thus

$$Q(z) = 1 + \sum_{j=k}^{n} b_j z^j, \quad \text{with } b_k \neq 0.$$

By the previous theorem, $\exists \theta \in [0, 2pi/k)$, s.t. $e^{ik\theta} = -\frac{|b_k|}{b_k}$. Thus for $r > 0$,

$$\begin{aligned}
\left|Q(re^{i\theta})\right| &= \left|1 + |b_k| r^k \cdot \frac{e^{ik\theta} b_k}{|b_k|} + \sum_{k+1}^{n} b_j r^j e^{ij\theta}\right| \\
&= \left|1 - |b_k| r^k + \sum_{k+1}^{n} b_j r^j e^{ij\theta}\right| \\
&\leq 1 - |b_k| r^k + \sum_{j=k+1}^{n} |b_j| r^j
\end{aligned}$$

Notice that $|b_j|r^j \leq \frac{1}{2}|b_k|r^k$ for sufficiently small r. So we further have

$$\left|Q(re^{i\theta})\right| \leq 1 - \frac{1}{2}|b_k|r^k < 1.$$

Contradiction! Since $|Q(x)| \geq 1$ for all x. Tracing back, we see $\mu = 0$. So there exists $z_0 \in \mathbb{C}$ such that $P(z_0) = 0$. $\square$

Corollary. *Let $P(z) = a_0 + a_1 z + \cdots + a_n z^n$ be a complex polynomial with $a_n \neq 0$ and $n \geq 1$. There exists $z_1, \ldots, z_n \in \mathbb{C}$, s.t.*

$$P(z) = a_n(z - z_1) \cdots (z - z_n).$$

Proof. By the theorem, there exists $z_n \in \mathbb{C}$ s.t. $P(z_n) = 0$. By long division algorithm, $P(z) = (z - z_n)Q(z) + \text{constant}$, where Q is a polynomial with degree $n - 1$. Evaluating both sides at $z = z_n$, we see that the constant is zero, i.e. $z - z_n | P$. Now repeat this procedure and we're done. $\square$

Banach Spaces

Definition. X is a real (or complex) vector space if $\forall x, y \in X$ and $\alpha, \beta \in \mathbb{R}$ (or $\mathbb{C}$), $\alpha x + \beta y \in X$ and some axioms hold.

Definition. A norm on X is a function $\|\cdot\|_X : X \to \mathbb{R}$ satisfying

- $\forall x \in X$, $\|x\|_X \geq 0$ and $\|x\|_X = 0$ if and only if $x = 0$.
- $\forall \alpha \in \mathbb{R}$ (or $\mathbb{C}$), $\forall x \in X$, $\|\alpha x\|_X = |\alpha| \cdot \|x\|_X$.
- $\forall x, y \in X$, $\|x + y\|_X \leq \|x\|_X + \|y\|_X$.

Definition. The normed vector space $(X, \|\cdot\|)$ is a Banach space if X is a complete metric space with respect to the distance $d(x, y) = \|x - y\|$.

Example. $\mathbb{R}^k$ with euclidean metric is a Banach space. For any interval $I \subseteq \mathbb{R}$, $\mathcal{C}^k(I)$ is a Banach space. For any metric space X, $\mathcal{C}^0(X)$ is a Banach space.

Definition. Define

$$l^\infty := l^\infty(\mathbb{N}) := \left\{\text{bounded sequences } \{x_n\}_{x\in\mathbb{N}} \text{ in } \mathbb{R}\right\}.$$

Define $\|\{x_n\}\|_{l^\infty} := \sup_n |x_n|$.

Definition. For $1 \leq p < \infty$,

$$l^p := l^p(\mathbb{N}) := \left\{\text{real sequences } \{x_n\}_{n\in\mathbb{N}} \text{ with } \|\{x_n\}\|_{l^p} < \infty\right\},$$

where

$$\|\{x_n\}\|_{l^p} := \left(\sum |x_n|^p\right)^{1/p}.$$

Note. In fact, $(l^\infty(\mathbb{N}), \|\cdot\|_{l^\infty}) = (\mathcal{C}^0(\mathbb{N}), \|\cdot\|_{\mathcal{C}^0(\mathbb{N})})$. So we have already seen that l^∞ is a Banach space.

Theorem 35 (Holders inequality). *Let $1 \le p, q \le \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ $(1/\infty = 0)$. Let $\{a_n\} \in l^p$ and $\{b_n\} \in l^q$. Then $\{a_n b_n\} \in l^1$ and*

$$\|\{a_n b_n\}\|_{l^1} = \sum_{n=1}^{\infty} |a_n b_n| \le \|\{a_n\}\|_{l^p} \cdot \|\{b_n\}\|_{l^q}.$$

Proof. If $\{a_n\} = \{0\}$ or $\{b_n\} = \{0\}$, the inequality holds trivially. Now suppose $p = \infty$ and $q = 1$ (the argument also works for $p = 1$ and $q = \infty$). We have

$$\sum_{n=1}^{\infty} |a_n b_n| \le \sum_{n=1}^{\infty} \|\{a_n\}\|_{l^\infty} \cdot |b_n| = \|\{a_n\}\|_{l^\infty} \cdot \|\{b_n\}\|_{l^1}.$$

Now consider $p, q \neq \infty$. Replacing $\{a_n\}$ with $\{a_n / \|\{a_n\}\|_{l^p}\}$ and $\{b_n\}$ with $\{b_n / \|\{b_n\}\|_{l^q}\}$ if needed, we may assume $\|\{a_n\}\|_{l^p} = \|\{b_n\}\|_{l^q} = 1$.

Claim: For $1 < p, q < \infty$ and $x, y \ge 0$, $xy \le x^p/p + y^q/q$.

Define
$$f_y(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy.$$
Then $f_y'(x) = x^{p-1} - y$ and $f_y''(x) = (p-1)x^{p-2} \ge 0$. So f_y has a global minimum at the zero of f_y', i.e. $x = y^{1/(p-1)}$. Remember that $1/q = (p-1)/p$, so

$$\begin{aligned} f_y(y^{1/(p-1)}) &= \frac{y^{p/(p-1)}}{p} + \frac{y^q}{q} - y^{p/(p-1)} \\ &= y^q \left(\frac{1}{p} + \frac{1}{q} - 1 \right) \\ &= 0. \end{aligned}$$

Then we see
$$\begin{aligned} \sum_{n=0}^{\infty} |a_n b_n| &\le \sum_{n=0}^{\infty} \frac{|a_n|^p}{p} + \frac{|b_n|^q}{q} \\ &= \frac{1}{p} \sum_{n=1}^{\infty} |a_n|^p + \frac{1}{q} \sum_{n=1}^{\infty} |b_n|^q \\ &= \frac{1}{p} + \frac{1}{q} \\ &= 1. \end{aligned}$$

□

Proposition. *Let $1 \le p, q \le \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for every real sequence $\{a_n\}$,*

$$\left(\sum_{n=1}^{\infty} |a_n|^p \right)^{1/p} = \sup_{\{b_n\} \in l^q,\ \|\{b_n\}\|_{l^q} = 1} \sum_{n=1}^{\infty} |a_n b_n|.$$

Proof. By Holders inequality, RHS $\leq$ LHS. If $\{a_n\} \in l^p$, take $b_n = |a_n|^{p-1}$. Then

$$\|\{b_n\}\|_{l^q} = \left(\sum_{n=1}^{\infty} \left(|a_n|^{p-1}\right)^q\right)^{1/q} = \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/q} = \|\{a_n\}\|_{l^p}^{p/q}.$$

So $\{b_n\} \in l^q$. Divide $\{b_n\}$ by $\|\{b_n\}\|_{l^q}$ to make $\|\{b_n\}\|_{l^q=1}$ and we have

$$\sum_{n=1}^{\infty} |a_n b_n| = \sum_{n=1}^{\infty} \frac{|a_n|^p}{\|\{a_n\}\|_{l^p}^{p/q}} = \frac{\|\{a_n\}\|_{l^p}^p}{\|\{a_n\}\|_{l^p}^{p/q}} = \|\{a_n\}\|_{l^p}.$$

If $\{a_n\} \notin l^p$, take

$$b_n^N = \begin{cases} |a_n|^{p-1}/\text{normalizing factor}, & n \leq N, \\ 0, & n > N. \end{cases}$$

Then

$$\lim_{N\to\infty} \sum_{n=1}^{N} |a_n b_n^N| = \lim_{N\to\infty} \left(\sum_{n=1}^{N} |a_n|^p\right)^{1/p} = \infty.$$

$\square$

Theorem 36 (Triangle inequality)**.** *For $\{a_n\},\ \{b_n\} \in l^p$, $\{a_n + b_n\} \in l^p$ and*

$$\|(a_n + b_n)\|_{l^p} \leq \|\{a_n\}\|_{l^p} + \|\{b_n\}\|_{l^p}.$$

Consequently, l^p is a normed vector space.

Proof. Choose q such that $\frac{1}{q} + \frac{1}{p} = 1$. Then by the above proposition, we see

$$\begin{aligned} \|\{a_n + b_n\}\|_{l^p} &= \sup_{\{c_n\}\in l^q, \|\{c_n\}\|_q = 1} \sum_{n=1}^{\infty} |c_n| \cdot |a_n + b_n| \\ &\leq \sup_{\|\{c_n\}\|_q = 1} \left(\sum_{n=1}^{\infty} |c_n a_n| + \sum_{n=1}^{\infty} |c_n b_n|\right) \\ &\leq \|\{a_n\}\|_{l^p} + \|\{b_n\}\|_{l^p}. \end{aligned}$$

$\square$

Definition. $\sum x_n$ converges absolutely if $\sum \|x_n\|$ converges. $\sum x_n$ converges if the sequence $s_n := x_1 + \cdots + x_n$ of partial sums converges.

Example. Define

$$l_K^\infty = \{\text{real sequences } \{x_n\}, \text{s.t. } x_n = 0, \text{for all except finitely many } n\},$$

where K stands for compact support. Let $x_n := (0, \ldots, 0, 2^{-n}, 0, \ldots)$ where 2^{-n} appears at the nth position. Then $\sum x_n$ converges absolutely but it doesn't converge in l_K^∞.

Theorem 37. *The normed vector space $(X, \|\cdot\|)$ is a Banach space if and only if every absolutely convergence series in X converges.*

Proof. **⇒:** If X is a Banach space and $\sum x_n$ converges absolutely, then by traingle inequality, for $n \geq m$,

$$\|s_n - s_{m-1}\| = \|a_m + \cdots + a_n\| \leq \sum_{j=m}^{n} \|a_j\|.$$

So the sequence of partial sums is Cauchy. Since X is complete, the sequnce of partial sums converge and thus $\sum x_n$ converges.

⇐: Assume that every absolute convergence series in X converges. Let $\{x_n\}$ be a Cauchy sequence in X. Recall that a Cauchy sequence with a convergence subsequence must converges (in any metric space). We need to prove that $\{x_n\}$ has a convergent subsequence. The idea is to use telescoping series.

By Cauchyness, there exists a subsequence $\{x_{n_k}\}$ s.t. $\|x_{n_{k+1}} - x_{n_k}\| \leq 2^{-(k+1)}$, $\forall k \in \mathbb{N}$. Then $\sum x_{n_{k+1}} - x_{n_k}$ converges absolutely, so it converges by hypothesis. Therefore,

$$\lim_{k\to\infty} x_{n_k} = x_{n_1} + \sum_{k=1}^{\infty}(x_{n_{k+1}} - x_{n_k}),$$

which implies $\{x_{n_k}\}$ is a convergent subsequence of $\{x_n\}$.

□

Theorem 38. *l^p is a Banach space for $1 \leq p \leq \infty$.*

Proof. We've already checked the case when $p = \infty$. Let $1 \leq p < \infty$ and assume $\sum \|\boldsymbol{x}_n\|_{l^p} \leq \infty$, where $\boldsymbol{x}_n$ is a sequence for each n. If we can show $\sum \boldsymbol{x}_n$ converges, then by the above theorem, l^p is a Banach space. For each $k \in \mathbb{N}$, the k-th coordinate satisfy

$$\sum_n |x_{n,k}| \leq \sum_n \left(\sum_k |x_{n,k}|^p\right)^{1/p} = \sum_n \|\boldsymbol{x}_n\|_{l^p} < \infty.$$

So we can define a sequence $\{y_k\}$ by $y_k = \sum_n x_{n,k}$. We need to show $\{y_k\} \in l^p$ and $\sum \boldsymbol{x}_n = \{y_k\}$ with convergence in l^p. Choose q with $\frac{1}{q} + \frac{1}{p} = 1$ and pick any $\{a_k\} \in l^q$ with $\|\{a_k\}\|_{l^q} = 1$. We see

$$\sum_k |a_k y_k| = \sum_k \left| a_k \sum_n x_{n,k} \right| \leq \sum_k \sum_n |a_k x_{n,k}| \leq \sum_n \sum_k |a_k x_{n,k}| \leq \sum_n \|\boldsymbol{x}_n\|_{l^p} \|\boldsymbol{a}_k\|_{l^q} = \sum_n \|\boldsymbol{x}_n\|_{l^p} < \infty.$$

Thus, $\|\{y_k\}\|_{l^p} < \infty$. Question. To show $\sum_n \boldsymbol{x}_n = \{y_n\}$ is similar. Show $\sum_{n=N}^{\infty} \boldsymbol{x}_n \to 0$ in l^p as $N \to \infty$ using a similar argument. □

Bounded linear operators

Definition. If X, Y are real vector spaces, a map $T : X \to Y$ is linear if $\forall \alpha_1, \alpha_2 \in \mathbb{R}$, $\forall x_1, x_2 \in X$, $T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)$.

Definition. Let $T : X^{\text{normed v.s.}} \to Y^{\text{normed v.s.}}$ be linear. We say T is a bounded linear operator (or T is bounded) if $\exists C$, s.t. $\forall x \in X$, $\|T(x)\|_Y \leq C \cdot \|x\|_X$.

Example. $\frac{\mathrm{d}}{\mathrm{d}x} : \mathcal{C}^k(I) \to \mathcal{C}^{k-1}(I)$ is a bounded linear operator.

Example. $X = \mathcal{C}^0([a,b])$ and $Y = \mathbb{R}$, $T(f) = \int_a^b f(x)\,\mathrm{d}x$ is a bounded linear operator.

For convention, we write $T(x)$ as Tx.

Theorem 39. *Let X, Y be normed vector spaces and $T : X \to Y$ being a linear operator. Then the following are equivalent:*

1. *T is a bounded linear opeartor;*
2. *T is uniformly continuous on X;*
3. *T is countinuous on X;*
4. *T is continuous at 0.*

Proof. $1 \Rightarrow 2:$ By 1, we have $\forall x, y$, $\|Tx - Ty\| = \|T(x-y)\| \le C\|x-y\|$. So T is Lipschitz, which implies 2.

$2 \Rightarrow 3:$ Done by definition.

$3 \Rightarrow 4:$ Done by definition.

$4 \Rightarrow 1:$ Assume T is continuous at 0. Then $\exists \delta > 0$, s.t. $\|x\| \le \delta$ implies $\|Tx\| \le 1$. Now if $x = 0$, $\|Tx\| = 0$. For $x \neq 0$, since $\|\delta x/\|x\|\| = \delta$, so $\|T(\delta x/\|x\|)\| \le 1$. By linearity and arithmetic, $\|Tx\| \le \|x\|/\delta$. So T is a bounded linear operator ($C = 1/\delta$).

$\square$

Definition. Define

$$\mathcal{L}(X,Y) \coloneqq \{\text{bounded linear operators from } X \text{ to } Y\}.$$

Then $\mathcal{L}(X,Y)$ is a vector space, since

$$\begin{aligned} \|(\alpha T + \beta S)x\|_Y &= \|\alpha(Tx) + \beta(Sx)\|_Y \\ &\le |\alpha| \cdot \|Tx\|_Y + |\beta| \cdot \|Sx\|_Y \\ &\le (|\alpha| C_1 + |\beta| C_2)\|x\|_X. \end{aligned}$$

Definition. Define the norm

$$\|T\|_{X\to Y} \coloneqq \sup_{0\neq x\in X} \frac{\|Tx\|_Y}{\|x\|_X}.$$

Proposition. *The following properties holds,*

1. $\|T\| = \sup_{\|x\|=1} \|Tx\|$;
2. $\|T\| = \min\{C : \|Tx\| \le C\|x\|,\ \forall x \in X\}$, *in particular,* $\forall x$, $\|Tx\| \le \|T\| \cdot \|x\|$;
3. $\|\cdot\|$ *is a norm;*
4. *If Y is a Banach space, so is $\mathcal{L}(X,Y)$;*

5. *If X, Y, Z are normed vector spaces and $T \in \mathcal{L}(X, Y)$, $S \in \mathcal{L}(Y, Z)$, then $S \circ T \in \mathcal{L}(X, Z)$ and $\|S \circ T\| \leq \|S\| \cdot \|T\|$.*

Proof. 1. By definition

$$\|T\| = \sup_{0 \neq x \in X} \frac{\|Tx\|}{\|x\|} = \sup_{0 \neq x \in X} \left\| T \frac{x}{\|x\|} \right\| = \sup_{\|y\|=1} \|Ty\|.$$

2. For any $C < \|T\|$, by definition of $\|T\|$, $\exists x \neq 0$, s.t. $\|Tx\|/\|x\| > C$. So $\|Tx\| > C\|x\|$. Also check that $\|T\| \geq \|Tx\|/\|x\|$, which implies $\|Tx\| \leq \|T\| \cdot \|x\|$.

3. Done.

4. Assume that $\{T_n\}$ is a Cauchy sequence. Then $\forall x$, $\{T_n x\}$ is a Cauchy sequence (because $\|T_n x - T_m x\| \leq \|T_n - T_m\| \cdot \|x\|$). Since Y is complete, $\{T_n x\}$ converges. Define $Tx \coloneqq \lim T_n x$. We need to show that $T \in L(X, Y)$. We see T is linear by linearity of limits and the T_n's. For the boundedness of T, we see $\|Tx\| = \lim \|T_n x\| \leq (\limsup \|T_n\|)\|x\|$. Note that $\{\|T_n\|\}$ is Cauchy in $\mathbb{R}$ since
$$\|T_m\| - \|T_n\| \leq \|T_m - T_n\|$$
and hence $\{\|T_n\|\}$ is convergent and bounded. So T is bounded.

5. Apply 2, we have
$$\|S \circ T(x)\| \leq \|S\| \cdot \|Tx\| \leq \|S\| \cdot \|T\| \cdot \|x\|.$$

□

Definition. Define the following norms on $\mathbb{R}^n$:

1. $\|x\|_{l_n^\infty} \coloneqq \max_{1 \leq i \leq n} |x_i|$
2. $\|x\|_{l_n^p} \coloneqq \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$, $1 \leq p \leq \infty$.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ takes the form $Tx = Ax$, where $A = (a_{ij})$ is a $m \times n$ matrix.

Proposition. *In the above notation,*

1. $\|T\|_{l_n^\infty \to l_m^\infty} = \max_{i=1,\ldots,m} \sum_{j=1}^n |a_{ij}|$;
2. $\|T\|_{l_n^1 \to l_m^1} = \max_{j=1,\ldots,n} \sum_{i=1}^m |a_{ij}|$.

Proof. 1. Define $C \coloneqq \max_{i=1,\ldots,m} \sum_{j=1}^n |a_{ij}|$. We need to first show $\|Tx\|_{l_m^\infty} \leq C\|x\|_{l_n^\infty}$, $\forall x$. This is true since

$$\max_{i=1,\ldots,m} |(Tx)_i| = \max_{i=1,\ldots,m} \left| \sum_j a_{ij} x_j \right| \leq \max_{i=1,\ldots,m} \sum_j |a_{ij}| \cdot \|x\|_{l_n^\infty}.$$

Now we need to show if $C' < C$, then $\exists x$, s.t. $\|Tx\|_{l_m^\infty} > C'\|x\|_{l_n^\infty}$. This is enough to find $x \neq 0$, s.t. $\|Tx\| = C\|x\|$. This is always possible in finite dimensional vector space (may not be possible in infinity dimensional vector space). Choose i to maximize $\sum_j |a_{ij}|$. Take

$$x_j = \operatorname{sign}(\alpha_{ij}) \coloneqq \begin{cases} 1 & \text{if } a_{ij} \geq 0 \\ -1 & \text{if } a_{ij} < 0. \end{cases}$$

Then $\|Tx\|_{l^\infty} = \sum_j |a_{ij}| = C\|x\|_{l^\infty}$.

2. Homework.

$\square$

The open mapping and closed graph theorem

Definition. Let X, Y be metric spaces. A function $f : X \to Y$ is an open map if $f(U)$ is open in Y whenever U is open.

Theorem 40 (Open Mapping Theorem)**.** *Let X and Y be Banach spaces. Then a surjective map $T \in \mathcal{L}(X, Y)$ is also an open map.*

Proof. Assume $T \in \mathcal{L}(X, Y)$ is surjective. Then, $Y = \cup_{n=1}^{\infty} T(B_n(0))$. By the Baire Category Theorem (Y is complete), for some n, $\overline{T(B_n(0))}$ has nonempty interior. Since $\overline{T(B_n(0))} = n\overline{T(B_1(0))}$ by linearity of T, $\overline{T(B_1(0))}$ has nonempty interior. Suppose $y_0 \in \operatorname{int}(\overline{T(B_1(0))})$. Then $\exists r > 0$, s.t. $B_r(y_0) \subseteq \overline{T(B_1(0))}$.

Claim 1: $B_{2r}(0) \subseteq B_r(y_0) - B_r(y_0) := \{y - y' : y, y' \in B_r(y_0)\}$

If $\|x\| < 2r$, $y_0 + x/2, y_0 - x/2 \in B_r(y_0)$. If $y, y' \in B_r(y_0)$, then $\|y - y'\| \le \|y - y_0\| + \|y_0 - y'\| < 2r$.

So $B_{2r}(0) \subseteq \overline{T(B_1(0))} - \overline{T(B_1(0))} = \overline{T(B_2(0))}$. So $B_r(0) \subseteq \overline{T(B_1(0))}$.

Claim 2: $B_{r/2}(0) \subseteq T(B_1(0))$.

Let $y_1 \in B_{r/2}(0)$. Then $\exists x_1 \in B_{1/2}(0)$, s.t. $\|y_1 - Tx_1\| < r/4$. Let $y_2 := y_1 - Tx_1$. In general, given $y_n \in B_{2^{-n}r}(0) \subseteq \overline{T(B_{2^{-n}}(0)}$, $\exists x_n \in B_{2^{-n}}(0)$, s.t. $\|y_n - Tx_n\| < 2^{-(n+1)}r$. Set $y_{n+1} = y_n - Tx_n$. Then $y_{n+1} \in B_{2^{-(n+1}r}(0)$. So we repeat. We obtain sequences $\{y_n\}$ in Y and $\{x_n\}$ in X, s.t. $\|x_n\| < 2^{-n}$, $\forall n$ and $\|y_n - Tx_n\| < 2^{-(n+1)}r$. Notice that

$$\|y_n - Tx_n\| = \|y_{n-1} - Tx_{n-1} - Tx_n\| = \cdots = \|y_1 - \sum_{j=1}^{n} Tx_j\| = \|y_1 - T\sum_{j=1}^{n} x_j\|.$$

Now $\sum x_n$ converges absolutely (since $\|x_n\| < 2^{-n}$). Since X is complete, $\exists x \in X$, s.t. $\sum x_n = x$. Moreover, $\|x\| \le \sum \|x_n\| < \sum 2^{-n} = 1$. So $x \in B_1(0)$. Finally,

$$y_1 = \lim_{n\to\infty} T\sum_{j=1}^{n} x_j = T(\lim_{n\to\infty} \sum_{j=1}^{n} x_j) = T(x) \in T(B_1(x)).$$

Let U be open and let $y \in T(U)$. Then $\exists x \in U$, s.t. $Tx = y$. Since U is open, $\exists \epsilon > 0$, s.t. $B_\epsilon(x) \subseteq U$. Notice that

$$T(B_\epsilon(0)) + Tx = T(B_\epsilon(0) + x) = T(B_\epsilon(x)) \subseteq T(U),$$

and $Tx = y$. Finally, by Claim 2,

$$B_{\epsilon r/2}(y) = B_{\epsilon r/2}(0) + y \subseteq \epsilon T(B_1(0)) + y \subseteq T(U).$$

$\square$

Corollary. *If X and Y are Banach spaces and $T \in L(X,Y)$ is a bijection. Then T^{-1} is also a bounded linear operator.*

Proof. It suffices to show that T^{-1} is continuous. Let $U \subseteq X$ be an open set. Then

$$\text{preImage}(T^{-1}) = T(U)$$

is open. So T^{-1} is continuous. $\square$

Theorem 41 (Closed Graph Theorem)**.** *Let X, Y be Banach spaces. The map $T : X \to Y$ is a bounded linear operator if and only if the graph*

$$\Gamma_T := \{(x, Tx) \in X \times Y : x \in X\}$$

is a closed linear subspace of $X \times Y$. Here $X \times Y$ is the Banach space with operations $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ and $\alpha(x, y) = (\alpha x, \alpha y)$, $\|(x, y)\| = \|x\|_X + \|y\|_Y$.

Proof. $\Rightarrow$: Homework.

$\Leftarrow$: Assume that Γ_T is a closed linear subspace.

Claim 1: T is linear.
Let $x, x' \in X$ and $\alpha, \alpha' \in \mathbb{R}$, then $(x, Tx), (x', Tx') \in \Gamma_T$. So $(\alpha x + \alpha' x', \alpha Tx + \alpha' Tx') \in \Gamma_T$. Thus by definition of Γ_T, $\alpha Tx + \alpha' Tx' = T(\alpha x + \alpha' x')$.

Claim 2: T is continuous.
We know Γ_T is a Banach space. We consider the projections $P_x : \Gamma_T \to X$ with $(x, y) \mapsto x$ and $P_y : \Gamma_T \to Y$ with $(x, y) \mapsto y$. Then P_x and P_y are bounded, since $\|y\|, \|x\| \leq \|(x, y)\|$. Furthermore, P_x is a bijection. So $P_x^{-1} : X \to \Gamma_T$, $x \mapsto (x, Tx)$, is a bounded linear operator (by the previous corollary). Finally, $T = P_y \circ P_x^{-1}$.

$\square$

Invertible linear operators and the van Neumann series

Theorem 42. *Let X and Y be Banach spaces and let $\Omega(X, Y)$ denote the set of bijections in $\mathcal{L}(X, Y)$. Then $\Omega(X, Y)$ is open subset of $\mathcal{L}(X, Y)$ and the inversion map $A \mapsto A^{-1}$ is a continuous bijection of $\Omega(X, Y)$ onto $\Omega(Y, X)$*

Proof. By corollary of the open mapping theorem, we know $A \in \Omega(X, Y)$ implies $A^{-1} \in \mathcal{L}(Y, X)$. Thus $A \mapsto A^{-1}$ is a bijection. Remains to prove that $\Omega(X, Y)$ is open and $A \mapsto A^{-1}$ is continuous.

Lemma: If $A \in \Omega(X, Y)$ and $\|B - A\|_{X \to Y} < 1/\|A^{-1}\|_{Y \to X}$. Then $B \in \Omega(X, Y)$ and

$$B^{-1} = B^{-1} A A^{-1} = (A^{-1} B)^{-1} A^{-1} = (I - (I - A^{-1} B))^{-1} A^{-1} = \sum_{n=0}^{\infty} (I - A^{-1} B)^n A^{-1},$$

with convergence (indeed absolute convergence) in $\mathcal{L}(Y, X)$ and

$$\|B^{-1} - A^{-1}\|_{Y \to X} \leq \frac{\|A^{-1}\|^2 \cdot \|A - B\|}{1 - \|A^{-1}\| \cdot \|A - B\|}$$

Proof of the lemma. Observe that

$$\|(I-A^{-1}B)^nA^{-1}\| \le \|I-A^{-1}B\|^n\cdot\|A^{-1}\| = \|A^{-1}(A-B)\|^n\cdot\|A^{-1}\| \le \|A^{-1}\|^{n+1}\cdot\|A-B\|^n.$$

By hypothesis, we know $\|A^{-1}\|\cdot\|A-B\| < 1$. So by geometric series test, $\sum_{n=0}^{\infty}(I-A^{-1}B)^nA^{-1}$ converges absolutely. And since $L(Y,X)$ is a Banach space, the series converges. Now we compute

$$\begin{aligned}
B\cdot\sum_{n=0}^{\infty}(I-A^{-1}B)^nA^{-1} &= \sum_{n=0}^{\infty}B(I-A^{-1}B)^nA^{-1}\\
&= \sum_{n=0}^{\infty}A(I-(I-A^{-1}B))(I-A^{-1}B)^nA^{-1}\\
&= \sum_{n=0}^{\infty}A(I-A^{-1}B)^nA^{-1} - \sum_{n=1}^{\infty}A(I-A^{-1}B)^nA^{-1}\\
&= A(I-A^{-1}B)^0A^{-1}\\
&= I.
\end{aligned}$$

For the other direction, we compute

$$\begin{aligned}
\left(\sum_{n=0}^{\infty}(I-A^{-1}B)^nA^{-1}\right)B &= \sum_{n=0}^{\infty}(I-A^{-1}B)^nA^{-1}B\\
&= \sum_{n=0}^{\infty}(I-A^{-1}B)^n(I-(I-A^{-1}B))\\
&= \sum_{n=0}^{\infty}(I-A^{-1}B)^n - \sum_{n=1}^{\infty}(I-A^{-1}B)^n\\
&= (I-A^{-1}B)^0\\
&= I.
\end{aligned}$$

Note the first equality is due to the continuity of B.

Aside: $S_N = \sum_{n=0}^{N}(I-A^{-1}B)$. We know $\lim S_N = S$. We want $B\cdot S = \lim Bs_N$.

$$\|BS - BS_N\| \le \|B\|\cdot\|S-S_N\| \to 0.$$

So $B\in\Omega(X,Y)$ and B^{-1} is indeed the given sum. Then finally, we have

$$\begin{aligned}
\|\sum_{n=1}^{\infty}(I-A^{-1}B)^nA^{-1}\| &\le \sum_{n=1}^{\infty}\|I-A^{-1}B\|^n\cdot\|A^{-1}\|\\
&\le \sum_{n=1}^{\infty}\|A^{-1}\|^{n+1}\cdot\|A-B\|^n\\
&= \frac{\|A^{-1}\|^2\cdot\|A-B\|}{1-\|A^{-1}\|\cdot\|A-B\|}.
\end{aligned}$$

□

□

Multivariable Calculus

Definition. Let $(V, \|\cdot\|_C)$, $(W, \|\cdot\|_W)$ be normed vector spaces. Let E be an open subset of V and $f : E \to W$. We say f is differentiable at $v \in E$ if $\exists L_v \in \mathcal{L}(V, W)$ s.t.

$$\lim_{h \in V, h \to 0} \frac{\|f(v+h) - f(v) - L_v(h)\|}{\|h\|} = 0.$$

We call L_v the total derivative of f at v and write

$$Df_v = Df(v) = L_v \in \mathcal{L}(V, W).$$

Proposition. *Total derivatives, if they exist, are unique.*

Proof. If L_v and L'_v both satisfies the equation above. Then by triangle inequality, we have

$$\lim_{h \to 0} \frac{\|L_v(h) - L'_v(h)\|}{\|h\|} = 0.$$

So $\forall \epsilon > 0$, $\exists \delta > 0$, s.t. $\|h\| < \delta$ implies

$$\frac{\|L_v(h) - L'_v(h)\|}{\|h\|} < \epsilon.$$

This implies (HW) $\|L_v - L'_v\| < \epsilon$. Since ϵ is arbitrary, $L_v = L'_v$. □

Proposition. *If f is differentiable at v, then f is continuous at v.*

Proof. If f is differentiable at v, look at

$$\|f(v+h) - f(v)\| = o(\|h\|) + \|L_v(h)\| \leq o(\|h\|) + \|L_v\| \cdot \|h\| \to 0.$$

□

Example. If $f \in \mathcal{L}(V, W)$, then f is differentiable at each point $v \in V$ and $Df_v = f$.

Proof. By linearity, $f(v+h) - f(v) - f(h) \equiv 0$ and we are done. □

Example. Let $V = \mathbb{R}^k$, $W = \mathbb{R}$. Let $f(x) = \|x\|_2^2 = \sum_{j=1}^k x_j^2$. Then $\forall v \in \mathbb{R}^k$, f is differentiable at v and $Df_v(h) = 2v \cdot h$. In matrix form, $Df_v = [2v_1, 2v'_2, \ldots, 2v_k]$.

Proof.

$$\lim_{h \to 0} \frac{\|f(v+h) - f(v) - 2v \cdot h\|}{\|h\|} = \lim_{h \to 0} \frac{\|h\|^2}{\|h\|} = 0.$$

□

Example. Let V be an arbitrary space but $V \neq \{0\}$, $W = \mathbb{R}$. Let $f(x) = \|x\|$ (any norm). Then f is not differentiable at 0.

Proof. Let $L_v \in \mathcal{L}(V, \mathbb{R})$. Let $e \in V$ with $\|e\| = 1$. By definition,

$$0 = \lim_{h \in \mathbb{R}, h \to 0} \frac{\|f(0+he) - f(0) - L_v(he)\|}{\|h\|} = \lim_{h \in \mathbb{R}, h \to 0} \frac{\||h| - hL_v(e)\|}{\|h\|} = \lim_{h \in \mathbb{R}, h \to 0} \left|1 - \frac{h}{|h|} L_v(E)\right|.$$

If $L_v(e) = 0$, this is false because the limit is 1. If $L_v(e) \neq 0$, this is false because the left and right limits disagree. □

Example. Let $V = \mathcal{C}^0([a,b])$, $W = \mathbb{R}$. Let $F(g) = \int_a^b (g(x))^2 \,\mathrm{d}x$ is differentiable at g, $\forall g \in \mathcal{C}^0([a,b])$ and $DF_g(f) = \int_a^b 2f(x)g(x)\,\mathrm{d}x$.

Proof.

$$\|F(g+h) - F(g) - DF_v(h)\| = \left\| \int_a^b h^2(x)\,\mathrm{d}x \right\| \leq (b-a)\|h\|^2$$

□

Example. Let X, Y be Banach spaces. Define $V = \mathcal{L}(X,Y)$, $W = \mathcal{L}(Y,X)$ and $E = \Omega(X,Y)$ (the set of all bijections from X to Y), $F(T) = T^{-1}$, $T \in \Omega(X,Y)$. Then F is differentiable at each point of E and

$$DF_T(h) = -T^{-1}hT^{-1}$$

Remember that $DF_T \in \mathcal{L}(V,W)$, $h \in \mathcal{L}(X,Y)$ and $DF_T(h) \in \mathcal{L}(Y,X)$.

Proof. Let $T \in \Omega(X,Y)$. For h small, we can write

$$F(T+h) = (T+h)^{-1} = (T(I+T^{-1}h))^{-1} = (I+T^{-1}h)^{-1}T^{-1} = \sum_{n=0}^{\infty}(-1)^n(T^{-1}h)^nT^{-1}.$$

Now

$$F(T+h) - F(T) + T^{-1}hT^{-1} = \sum_{n=2}^{\infty}(-1)^n(T^{-1}h)^nT^{-1} = \mathcal{O}(\|h\|^2),$$

since $F(t)$ cancels out the zeroth term and $T^{-1}hT^{-1}$ cancels out the first term. So

$$\|F(T+h) - F(T) + T^{-1}hT^{-1}\| \leq \frac{\|T^{-1}h\|^2 \cdot \|T^{-1}\|}{1 - \|T^{-1}h\|} \leq \frac{\|T^{-1}\|^3 \cdot \|h\|^2}{1 - \|T^{-1}\| \cdot \|h\|}.$$

□

Theorem 43 (Chain rule)**.** *Let V, W, X be normed vector spaces and assume that $f : V \to W$ is differentiable at $v \in V$ and $g : W \to X$ is differentiable at $f(v)$. Then $g \circ f$ is differentiable at v and $\mathrm{D}(g \circ f)_v = \mathrm{D}g_{f(v)} \circ \mathrm{D}f_v$ (in $\mathcal{L}(V,X)$).*

Special case: If V, W, X are finite dimensional vector spaces, then we can think of $\mathrm{D}g_{f(v)}$, $\mathrm{D}f_v$ as matrices and we can think of the composition $\mathrm{D}g_{f(v)} \circ \mathrm{D}f_v$ as matrix multiplication.

Proof. Notation:

$$\epsilon(h) = \|h\|^{-1}\left(f(v+h) - f(v) - \mathrm{D}f_v(h)\right),$$
$$\delta(h) = \|k\|^{-1}\left(g(f(v)+k) - g(f(v)) - \mathrm{D}g_{f(v)}(k)\right).$$

We know that

$$\lim_{h \to 0} \epsilon(h) = 0_W, \quad \lim_{k \to 0} \delta(k) = 0_X.$$

Let $k = f(v+h) - f(v)$, we have

$$\begin{aligned} & g(f(v+h)) - g(f(v)) - \mathrm{D}g_{f(v)} \circ \mathrm{D}f_v(h) \\ =& g\left(f(v) + (f(v+h) - f(v))\right) - g(f(v)) - \mathrm{D}g_{f(v)}(f(v+h) - f(v)) + \mathrm{D}g_{f(v)}(f(v+h) - f(v) - \mathrm{D}f_v(h)) \\ =& \|k\| \cdot \delta(k) + \mathrm{D}g_{f(v)}(\|h\| \cdot \epsilon(h)) =: T_1 + T_2. \end{aligned}$$

Notice that $\|T_2\| \leq \|\mathrm{D}g_{f(v)}\| \cdot \|h\| \cdot \|\epsilon(h)\|$, so

$$\lim_{h\to 0} \frac{\|T_2\|}{\|h\|} = 0.$$

On the other hand,

$$\|k\| = \|(\|h\|\epsilon(h) + \mathrm{D}f_v(h))\| \leq \|h\| \cdot (\|\epsilon(h)\| + \|\mathrm{D}f_v\|)\,.$$

By continuity of f at v, $\lim_{h\to 0} \|f(v+h) - f(v)\| = 0$. So $\lim_{h\to 0} \delta(f(v+h) - f(v)) = 0$. So

$$\lim_{h\to 0} \frac{\|T_1\|}{\|h\|} \leq \lim_{h\to 0}(\|\epsilon(h)\| + \|\mathrm{D}f_v\|) \cdot \|\delta(k)\| = 0.$$

$\square$

Henceforth, $V = \mathbb{R}^k$, $W = \mathbb{R}^m$, all norms are Euclidean.

Proposition. *If $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable at x_0, then all first order partial derivatives of f at x_0 exist and moreover,*

$$\mathrm{D}f(x_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{bmatrix}.$$

Proof. Notice that

$$\lim_{h\to 0, h\in\mathbb{R}} \frac{\|f(x_0 + he_j) - f(x_0) - Df(x_0)he_j\|}{\|h\|} = 0.$$

Also we have

$$\lim_{h\to 0, h\in\mathbb{R}} \frac{\|f(x_0 + he_j) - f(x_0) - Df(x_0)he_j\|}{\|h\|} = \left\| \frac{\partial f}{\partial x_j}(x_0) - Df(x_0)e_j \right\|.$$

Since $\frac{\partial f}{\partial x_j}(x_0)$ is the jth column of the Jacobian matrix and $Df(x_0)e_j$ is the jth column of total derivative. So we are done. $\square$

Theorem 44 (Single variable mean value theorem)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Then $\exists c \in (a, b)$, s.t.*

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

How can we generalize this result to $f : E \to \mathbb{R}^m$, where $E \subseteq \mathbb{R}^n$? First, we consider the case $n \geq 1$, $m = 1$.

Definition. The set $E \subseteq V^{\text{vector space}}$ is convex if $\forall x, y \in E$, and $0 \leq t \leq 1$, $(1-t)x + ty \in E$. In other words, the line segment connecting x and y is in E.

Theorem 45. *Let $E \subseteq \mathbb{R}^n$ be open and convex. Let $f : E \to \mathbb{R}$ be differentiable. Let $x, y \in E$. Then $\exists t \in [0, 1]$, s.t.*

$$f(y) - f(x) = \mathrm{D}f((1-t)x + ty)(y - x).$$

Proof. Define $g(t) = f((1-t)x + ty)$. Then g is continuous on $[0, 1]$. By the chain rule, g is differentiable on $(0, 1)$ and $g'(t) = \mathrm{D}f((1-t)x+ty) \cdot (y-x)$. Finally, by the mean value theorem, $\exists t \in [0, 1]$, s.t. $g'(t) = g(1) - g(0) = f(y) - f(x)$. We are done. $\square$

Now consider $m > 1$. The direct analogue of mean value theorem is false.

Example. Let $\gamma(t) = (\cos(t), \sin(t))$. Then, $\gamma(0) = \gamma(2\pi)$, i.e. $\gamma(0) - \gamma(2\pi) = 0$. But $\gamma'(t) = (-\sin(t), \cos(t))$ is never 0.

Theorem 46. *Let $E \subseteq \mathbb{R}^n$ be open, convex and $f : E \to \mathbb{R}^m$ be differentiable. If $\|\mathrm{D}f(x)\|_{\mathcal{L}(l_n^2, l_m^2)} \leq M$ (this norm is not the 2-norm of the matrix), $\forall x \in E$, then*

$$\|f(x) - f(y)\| \leq M\|x - y\|$$

for all $x, y \in E$.

Proof. Fix $x_0, y_0 \in E$. Define $v := f(y_0) - f(x_0)$, $g(x) := \langle v, f(x)\rangle$. By the previous theorem, $\exists t \in [0, 1]$, s.t.

$$\|v\| \cdot \|f(x_0) - f(y_0)\| = g(y_0) - g(x_0) = \langle v, \mathrm{D}f((1-t)x_0 + ty_0)(y_0 - x_0)\rangle \leq \|v\| \cdot M \cdot \|y_0 - x_0\|.$$

$\square$

Both theorems are false if E is not convex.

Example. Define

$$E := \{(x, y) \in \mathbb{R}^2 : |x| < 1, |y| < 1, y = 0 \Rightarrow x = 0\}.$$

Let

$$f(x, y) = \begin{cases} 0 & \text{if } x \leq 0, \\ 0 & \text{if } x > 0, y < 0, \\ e^{-1/x} & \text{if } x > 0, y > 0. \end{cases}$$

Theorem 47. *Let $E \subseteq \mathbb{R}$ be open connected and $f : E \to \mathbb{R}^m$ be differentiable with $\mathrm{D}f \equiv 0$ on E. Then f is a constant function on E.*

Proof. Let $x_0 \in E$. Let $F := \{x \in E : f(x) = f(x_0)\}$. Then F is closed (relative to E) and f is continuous. Let $y_0 \in F$. Then $\exists r > 0$, s.t. $B_r(y_0) \subseteq E$. Since $B_r(y_0)$ is convex, by the previous theorem, $f(x) - f(y_0) = 0$, $\forall x \in B_r(y_0)$. So $B_r(y_0) \in F$. Thus F is open. Since E is connected, F closed and open implies $F = E$ or $F = \emptyset$. But $x_0 \in F$, so $F \neq \emptyset$. Thus $F = E$. $\square$

Question: Is there any relationship between partial derivatives and differentiability of f?
Answer: No.

Example. Let

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0) \end{cases}$$

Then all partial derivatives of f of all orders exist. But f is not differentiable, since it is not continuous at 0.

Definition. Let $E^{\text{open}} \subseteq \mathbb{R}^n$ and $f : E \to \mathbb{R}^m$. Say f is continuously differentiable on E if f is differentiable on E and $\mathrm{D}f$ is continuous on E (i.e. as a map from E into $\mathcal{L}(l_n^2, l_m^2)$).

Definition. Say $f \in \mathcal{C}^1(E)$ if f is continuously differentiable on E and

$$\|f\|_{\mathcal{C}^1(E)} \coloneqq \|f\|_{\mathcal{C}^0(E)} + \|\mathrm{D}f\|_{\mathcal{C}^0(E)}.$$

Definition. Say $f \in \mathcal{C}^1_{loc}(E)$ if f is continuously differentiable on E. Warning: $\mathcal{C}^1_{loc}(E)$ is not a normed (not even a normable) space.

Theorem 48. *Let $E^{open,\neq\emptyset} \subseteq \mathbb{R}^n$ and $f : E \to \mathbb{R}^m$. Then f is continuously differentiable on E if and only if all first order partial derivatives of f exist and are continuous on E.*

Proof. **$\Rightarrow$:** Immediate follows from the proposition earlier.

$\Leftarrow$: We will show it on Monday's class. By the proposition, we just need to show that

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - \mathrm{D}f(x_0)h\|}{\|h\|} = 0, \quad \forall x_0 \in E.$$

Note, here $\mathrm{D}f(x_0)$ is the Jacobian matrix. Considering each component separately, it suffices to consider the case $m = 1$.

$\square$