Real Variables

April 4, 2025

Chapter 1

Number Systems

Natural numbers: $\mathbb{N} = \{1, 2, 3, \ldots\}$

Integers: $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$

Rational numbers: $\mathbb{Q} = \left\{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\right\}$

Remark. Note for real numbers, $\mathbb{Q}$ has holes in it.
Example. $\nexists p \in \mathbb{Q}\ s.t\ p^2 = 2$

Proof. Assume $\exists p \in \mathbb{Q}$ s.t $p^2 = 2$. Then $p = \frac{a}{b}$ where $a, b \in \mathbb{Z}$ and $b \neq 0$. So, $p^2 = \frac{a^2}{b^2} = 2 \Rightarrow a^2 = 2b^2$. So, a^2 is even $\Rightarrow a$ is even. So, $a = 2k$ for some $k \in \mathbb{Z}$. So, $4k^2 = 2b^2 \Rightarrow b^2 = 2k^2$. So, b^2 is even $\Rightarrow b$ is even. So, $b = 2l$ for some $l \in \mathbb{Z}$. So, a and b are both even, which contradicts the fact that a and b are coprime. So, $\nexists p \in \mathbb{Q}$ s.t $p^2 = 2$. Q.E.D.

Definition 1.1 (Order)**.** An order on a set S is a relation $<$ such that:

(a) If $a, b \in S$, then exactly one of $a < b$, $a = b$, or $b < a$ is true.

(b) If $a, b, c \in S$ and $a < b$ and $b < c$, then $a < c$.

Definition 1.2 (Ordered Set)**.** An ordered set S is a set with an order $<$.

Definition 1.3. Let S be an ordered set. A set $E \subset S$ is bounded above if $\exists \beta \in S$ s.t $\forall x \in E : x \leq \beta$.
Similarly, a set S is bounded below if $\exists \beta \in S$ s.t $\forall x \in E : x \geq \beta$.

Definition 1.4 (LUB, GLB)**.** Let S be an ordered set and $E \subset S$, $E \neq \emptyset$, with E bounded above. If $\exists \alpha$ s.t. α is an upper bound for E and $\forall \gamma < \alpha$: γ is not an upper bound for E, then such α is called least upper bound (LUB), or *Supremum*. Similarly, if $\exists \alpha$ s.t. α is a lower bound for E and $\forall \gamma > \alpha$: γ is not a lower bound for E. Then such α is called greatest lower bound (GLB), or *Infimum.*

Definition 1.5 (LUB property). An ordered set S has the least upper bound (LUB) property if $\forall E \subset S$ if $E \neq \emptyset$ and E bounded above implies $\exists \sup E \in S$; i.e., Every bounded subset of S has the least upper bound(LUB).

Example.

- *$\mathbb{Z}$ has the LUB property.*
- *$\mathbb{Q}$ does not have the LUB property.*

Theorem 1.1. Let S be an ordered set. Then S has the LUB property if and only if S has the GLB property.

Proof. ($\Rightarrow$)Suppose S has the LUB property. Let $B \subset S$ be non-empty and bounded below. Let L be the set of all lower bounds of B. Then L is non-empty and bounded above. Let $\alpha = \sup L$. We claim that $\alpha = \inf B$. ($\Leftarrow$)Suppose S has the GLB property. Let $E \subset S$ be non-empty and bounded above. Let U be the set of all upper bounds of E. Then U is non-empty and bounded below. Let $\beta = \inf U$. We claim that $\beta = \sup E$. Q.E.D.

Definition 1.6 (Fields). Let F be a set with two operations, addition and multiplication. Then F is a field if the following axioms are satisfied:

(a) $a + b = b + a$ and $a \cdot b = b \cdot a$ for all $a, b \in F$ (Commutative laws).

(b) $(a+b)+c = a+(b+c)$ and $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all $a, b, c \in F$ (Associative laws).

(c) $a \cdot (b + c) = a \cdot b + a \cdot c$ for all $a, b, c \in F$ (Distributive law).

(d) $\exists 0 \in F$ s.t. $a + 0 = a$ for all $a \in F$.

(e) $\exists (-a) \in F$ s.t. $a + (-a) = 0$ for all $a \in F$.

(f) $\forall x, y \in F : xy \in E$.

(g) $\forall x, y \in F : xy = yx$.

(h) $\exists 1 \in F$ s.t. $a \cdot 1 = a$ for all $a \in F$.

(i) If $a \neq 0$, then $\exists a^{-1} \in F$ s.t. $a \cdot a^{-1} = 1$.

(j) $\forall x, y, z \in F : x(y + z) = xy + xz$

Example.

(a) $\mathbb{Q}$ is a field, while $\mathbb{Z}$ is not a field.

(b) $F_p = \{0, 1, \ldots, p - 1\}$ with mod p arithmetic is a field.

Read Text book: 114,115,116,118

Definition 1.7 (Ordered Field)**.** An ordered field F is a field that is an ordered set such that the following properties are satisfied:

(a) If $a, b, c \in F$ and $a < b$, then $a + c < b + c$.

(b) If $a, b \in F$ and $0 < a$ and $0 < b$, then $0 < ab$.

Remark. We say x is positive if $x > 0$ and x is negative if $x < 0$.

Example. *$\mathbb{Q}$ is an ordered field.*

Theorem 1.2. $\exists$ an ordered field $\mathbb{R}$ which has the LUB property and contains $\mathbb{Q}$ as a subfield.

Theorem 1.3.

(a) Arithmetic properties of $\mathbb{R}$: If $x, y \in \mathbb{R}$ and $x > 0$ then $\exists n \in \mathbb{N}$ such that $nx > y$.

(b) $\mathbb{Q}$ is dense in $\mathbb{R}$: If $x, y \in \mathbb{R}$ and $x < y$, then $\exists p \in \mathbb{Q}$ such that $x < p < y$.

(c) $x, y \in \mathbb{R}$ then $\exists \alpha \in \mathbb{R} \setminus \mathbb{Q}$ such that $x < \alpha < y$.

Proof. (a) Let $A = \{nx \mid n \in \mathbb{N}\}$. Suppose $\forall nx \in A : nx \le y$. Then y is an upper bound for A. So, A has a least upper bound α. Since $\alpha - x < \alpha$ as $x > 0$, $\alpha - x$ is not an upper bound for A. Thus, $\exists m \in \mathbb{N} : mx > \alpha - x$, so $\alpha < (m+1)x \in A$, contradicting the fact that α is a supremum of A. Therefore, $\exists n \in \mathbb{N}$ such that $nx > y$.

(b) Since $y - x > 0$, by (a), $\exists n \in \mathbb{N}$ such that $n(y - x) > 1$. $ny - nx > 1$ and therefore, $1 + nx < ny$. Let $m \in \mathbb{Z}$ such that $(m - 1) \le nx < m$. Such m exists by the extended version of (a). This implies there exists $m \in \mathbb{N}$ such that $nx < m \le nx + 1 < ny$. Therefore, $x < \frac{m}{n} < y$.

(c) $\exists k \in \mathbb{Q}$ such that $k^2 = 2$; i.e., $\exists \sqrt{2} \in \mathbb{R}$. $0 < \sqrt{2} < 2$ because if $\sqrt{2} \ge 2$ then $2 = \sqrt{2} \cdot \sqrt{2} \ge 2 \cdot 2 = 4$, which is a contradiction. By (b), $\exists p \in \mathbb{Q}$ such that $x < p < y$ and $\exists q \in \mathbb{Q}$ such that $x < p < q < y$. Let $\alpha = p + \frac{\sqrt{2}}{2}(q - p)$. Then $x < p < \alpha < q < y$ and $\alpha \notin \mathbb{Q}$ since otherwise $\sqrt{2} = 2 \cdot \frac{\alpha - p}{q - p}$ would be rational

Q.E.D.

Note. (a) can be improved to:

$$\forall x, y \in \mathbb{R}, x > 0, \exists n \in \mathbb{Z} \text{ such that } (n-1)x \leq y < nx. \tag{1.1}$$

Proof. Case 1: $y \geq 0$. Let $A = \{m \in \mathbb{N} : y < mx\} \subset \mathbb{N}$. By (a), $A \neq \emptyset$. Every non-empty subset of $\mathbb{N}$ has a smallest element. Let $n =$ smallest element of A. Then the inequality holds true. Case 2: Let $y < 0$, then there exists $n \in \mathbb{N}$ such that $(n-1)x \leq -y < nx$, which implies that (by changing sign for all terms) $-nx < y \leq -(n-1)x$. Hence, the statement holds. Q.E.D.

Lemma. Let $a, b \in \mathbb{R}$ such that $0 < a < b$, then $0 < b^n - a^n \leq nb^{n-1}(b-a)$ for some $n \in \mathbb{N}$.

Proof.

$$b^n - a^n = (b-a)(\underbrace{b^{n-1} + ab^{n-2} + \ldots + a^{n-2}b + a^{n-1}}_{n \text{ terms}})$$

$$< (b-a)nb^{n-1}$$

Q.E.D.

Theorem 1.4. $\forall x \in \mathbb{R}, x > 0, n \in \mathbb{N} : \exists!(\text{unique}) y > 0 : y^n = x$ (we write $y = x^{1/n} = \sqrt{x}^n$, the n^{th} root of x).

Proof. Uniqueness: For any $y_1, y_2 \in \mathbb{R}$, if $0 < y_1 < y_2$, then $0 < y_1^n < y_2^n$, hence y_1^n and y_2^n cannot both be equal to x.
Existence: Let $E_x = \{t \in \mathbb{R}_{>0} \mid t^n < x\}$. If $E \neq \varnothing$, E is bounded above, hence (by the least-upper-bound property) there exists a $\sup E$. Choose $y = \sup E$. Consider two cases.

(a) If $x \leq 1$, then $t_0 = \dfrac{x}{2}$ and thereby $t_0^n = \dfrac{x^n}{2^n} < x^n \leq x$ (by assumption that $x \leq 1$).

(b) If $x > 1$, then let $t_0 = 1$, leading to $t_0^n = 1 < x$.

In either case, $t_0 \in E$, and hence E is not empty. 1(a) (E is bounded above) Let $\beta = x + 1$. Then, $\beta^n = (x+1)^n > x + 1 > x$. Then, for any $t \in E$, we have that $t^n < x < \beta^n$, hence $t < \beta$, making t an upper bound of E.

(a) Assuming that $y^n < x$, we find $0 < h < 1$ such that $(y+h)^n < x$, which leads to $y + h \in E$, something that contradicts with the fact that $y = \sup E$. This is equivalent to finding an $0 < h < 1$ such that $(y+h)^n - y^n < x - y^n$. By the lemma , we have $0 < (y+h)^n - y^n < n(y+1)^{n-1}h$ for any $0 < h < 1$. Choose h so that $\frac{(x-y)^n}{n(y+1)^{n-1}}$. Then $0 < h < 1$ still holds and $hn(y+1)^{n-1} < x - y^n$, leading to $(y+h)^n < x$, and therefore $y + h \in E$. However, this contradicts the fact that $y = \sup E$ as $y + h > y$.

(b) Assuming that $y^n > x$, we find $k > 0$ such that $(y-k)^n > x$, which leads to a contradiction since otherwise $y - k$ would be an upper bound for E that's smaller than y, which is $\sup E$. By the lemma , $y^n - (y-k)^n \leq ny^{n-1}k < y^n - x$ for any $h < \dfrac{y^n - x}{ny^{n-1}}$. Therefore, $-(y-k)^n < -x$, or $x < (y-k)^n$. Thus, $y - k$ is also an upper bound of E and $y - k < y = \sup E$, which is a contradiction.

Since $y^n < x$ and $y^n > x$ are both contradictions, $y^n = x$. Q.E.D.

Definition 1.8 (Cut/Dedekind Cut)**.** The set $\mathbb{R}$ elements are (Dedekind) cuts, which are sets $\alpha \subset \mathbb{Q}$ such that

- $\forall p \in \alpha, q \in \mathbb{Q} : q < p \Rightarrow q \in \alpha$
- No greatest element in α

Example. $\alpha = \{p \in \mathbb{Q} \mid p < 0\}$, $\alpha = \{p \in \mathbb{Q} \mid p \leq 0 \vee p^2 < 2\}$

Definition 1.9 (Order of cuts)**.** For $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta := \alpha \subset \beta$

Proof (test)**.** Let γ be set of cuts A, and show that γ is a cut and that $\gamma = \sup A$. Q.E.D.

Theorem 1.5. There exists an ordered field $\mathbb{R}$ such that $\mathbb{Q} \subset \mathbb{R}$ and $\mathbb{R}$ has the LUB property.

Proof. Let $\mathbb{R}$ be the set of all cuts with:

order $a < b := a \subset b$.

addition $\alpha + \beta = \{p + q \mid p \in \alpha, q \in \beta\}$.

multiplication $\alpha \cdot \beta = \{p \cdot q \mid p \in \alpha, q \in \beta\}$.

Q.E.D.

Complex Numbers

Definition 1.10 (Complex Field)**.** The underlying set is $\mathbb{C} = \{(a,b) | a \in \mathbb{R}, b \in \mathbb{R}\}$
Addition is defined as $(a,b) + (c,d) = (a+c, b+d)$
Multiplication is defined as $(a,b) \cdot (c,d) = (ac - bd, ad + bc)$
Zero element is $(0,0)$
One element is $(1,0)$

Theorem 1.6. $\mathbb{C}$ is a field.

Proof. Verify the 11 field axioms. For just a few axioms:
(M3):
$x = (a,b), y = (c,d), z = (e,f)$. $x(yz) = (a,b)(ce - df, cf + de) = (a(ce-df) - b(cf+de), a(cf+de) + b(ce-df)) = (ac-bd, ad+bc)(e,f) = (xy)z$
(M4):
$(a,b)(1,0) = (a \cdot 1 - b \cdot 0, a \cdot 0 + b \cdot 1) = (a,b)$
(M5):
$x \neq 0$ means $x = (a,b)$ with $a \neq 0$ or $b \neq 0$. That is, $a^2 + b^2 > 0$. Let $\frac{1}{x} = (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$. Then $x \cdot \frac{1}{x} = (a,b)(\frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}) = (\frac{a^2+b^2}{a^2+b^2}, \frac{-ab+ba}{a^2+b^2}) = (1,0)$. Q.E.D.

Identification of $\mathbb{R}$ as a subfield of $\mathbb{C}$. Identify $(a,0) \in \mathbb{C}$ with $a \in \mathbb{R}$. Then $(a,0) + (b,0) = (a+b, 0)$, $(a,0)(b,0) = (ab,0)$, so we can represent them by $a + b = a + b$, $a \cdot b = a \cdot b$. Write $i = (0,1)$. $i^2 = (0,1)(0,1) = (-1,0)$. So, $i^2 = -1$. $(a,b) \leftrightarrow a+bi$. Usually write $z = a+bi$ for $z \in \mathbb{C}$. $\mathrm{Re}(z) = a$,$\mathrm{Im}(z) = b$.

Definition 1.11. Complex conjugate of $z = a+bi$ is defined as $a-bi$ and denoted by $\overline{z}$

Note.

(a) $\overline{z+w} = \overline{z} + \overline{w}$

(b) $\overline{zw} = \overline{z} \cdot \overline{w}$

(c) $z + \overline{z} = 2 \cdot \text{Re}(z)$

(d) $z - \overline{z} = 2i \cdot \text{Im}(z)$

(e) $z\overline{z} = (a+bi)(a-bi) = a^2 + b^2 \geq 0$, with $=$ if any only if $z = 0$

(f) $\frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{a-bi}{a^2+b^2}$

Definition 1.12. $|z| = \sqrt{z\overline{z}} = \sqrt{a^2+b^2}$

In particular, if $z = a \in \mathbb{R}$ then $|z| = \sqrt{a^2} = |a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$

Theorem 1.7. For $z, w \in \mathbb{C}$,

(a) $|z| \geq 0$ with $=$ iff $z = 0$

(b) $|z| = |\overline{z}|$

(c) $|zw| = |z| \cdot |w|$

(d) $|\text{Re}(z) \leq |z|, |\text{Im}(z)| \leq |z|$

Proof. Let $z = a + bi$. Then $|\text{Re}(z)| = |a| \leq \sqrt{a^2+b^2} = |z|$
Q.E.D.

(e) $|z+w| \leq |z| + |w|$ (Triangle inequality)

Proof.

$$\begin{aligned}
|z+w|^2 &= (z+w)(\overline{z+w}) \\
&= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} \\
&= |z|^2 + z\overline{w} + w\overline{z} + |w|^2 \\
&= |z|^2 + 2\text{Re}(z\overline{w}) + |w|^2 \\
&\leq (|z| + |w|)^2
\end{aligned}$$

Q.E.D.

Theorem (Cauchy-Schwarz inequality)**.** If $a_1, a_n, b_1, b_n \in \mathbb{C}$, then

$$|\sum_{j=1}^{n} a_j \overline{b_j}| \leq (\sum_{j=1}^{n} |a_j|^2)^{\frac{1}{2}} (\sum_{j=1}^{n} |b_j|^2)^{\frac{1}{2}}.$$

Interpretation: $(\vec{a}, \vec{b}) = \sum_{j=1}^{n} a_j \overline{b_j}$ defined on inner product on $\mathbb{C}^n$ and $|(\vec{a}, \vec{b})| \leq \sqrt{(\vec{a} \cdot \vec{a})(\vec{b} \cdot \vec{b})}$. (Note that $\vec{a} \cdot \vec{b} = \overline{\vec{b} \cdot \vec{a}}$)

Proof. Let $A = \sum |a_j|^2$, $B = \sum |b_j|^2$, $C = \sum a_j \overline{b_j}$ We can assume 1. $B \neq 0$ because $B = 0$ is $0 \leq 0$, 2. $C \neq 0$ because $C = 0$, LHS is 0.
For any $\lambda \in \mathbb{C}$, $0 \leq \sum_{j=1}^{n} |a_j + \lambda b_j|^2 = \sum_{j=1}^{n} (a_j + \lambda b_j)(\overline{a_j} + \overline{\lambda b_j}) = \sum_{j=1}^{n} |a_j|^2 + \lambda \sum_{j=1}^{n} b_j \overline{a_j} + \overline{\lambda} \sum_{j=1}^{n} a_j \cdot \overline{b_j} + |\lambda|^2 \sum_{j=1}^{n} |b_j|^2$. Let $\lambda = tC$ for $t \in \mathbb{R}$.
Then $0 \leq A + \lambda \overline{C} + \overline{\lambda} C + |\lambda|^2 B = A + 2|C|^2 t + B|C|^2 t^2 = p(t)$. $p(t)$ is a quadratic function in terms of t and it must be non-negative. Regardless of the value of t. Therefore, the discriminant of $p(t) = (2|C|^2)^2 - 4AB|C|^2 = 4|C|^2(|C|^2 - AB) \leq 0$. Since $|C| \geq 0$, $|C|^2 \leq AB$. Q.E.D.

Definition 1.13 (Euclidean k-space)**.** For $k \in N$, $\mathbb{R}^k := \{\vec{x} = (x_1, x_2, \ldots, x_k) : x_1, x_2, \ldots, x_k \in \mathbb{R}\}$ with the following properties:

Addition
$\vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_k + y_k)$

Scalar multiplication
$\lambda \vec{x} = (\lambda x_1, \lambda x_2, \ldots, \lambda x_k)$

Inner(dot) product
$(\vec{x}, \vec{y}) = \sum_{j=1}^{k} x_j y_j$, which is bilinear: $(\alpha \vec{x} + \beta \vec{y}) \cdot \vec{z} = \alpha \vec{x} \cdot \vec{z} + \beta \vec{y} \cdot \vec{z}$.

Norm
$|\vec{x}| = \sqrt{(\vec{x}, \vec{x})} = \sum_{j=1}^{k} |x_j|^{2^{1/2}}$

Remark. Addition and Scalar multiplication make $\mathbb{R}^k$ into a vector space.

Theorem 1.8. Let $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^k$. Then

(a) $|\vec{x}| \geq 0$

(b) $|\vec{x}| = 0 \leftrightarrow \vec{x} = \vec{0}$

(c) $|\alpha \vec{x}| = |\alpha||\vec{x}|$

(d) $|\vec{x} \cdot \vec{y}| \leq |\vec{x}||\vec{y}|$ (special case of Cauchy-Schwarz inequality)

(e) $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$ (Triangle inequality)

Proof. $|\vec{x} + \vec{y}|^2 = (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) = |\vec{x}|^2 + 2\vec{x} \cdot \vec{y} + |\vec{y}|^2 \leq |\vec{x}|^2 + 2|\vec{x} \cdot \vec{y}| + |\vec{y}|^2 \leq (|\vec{x}| + |\vec{y}|)^2$ Q.E.D.

(f) $|\vec{x} - \vec{y}| \leq |\vec{x} - \vec{z}| + |\vec{z} - \vec{y}|$

Proof. $|\vec{x} - \vec{y}| = |(\vec{x} - \vec{z}) + (\vec{z} - \vec{y})| \leq |\vec{x} - \vec{z}| + |\vec{z} - \vec{y}|$ Q.E.D.

Chapter 2

Basic Topology

Definition 2.1. Sets A and B have the same cardinality, if $\exists f : A \to B$ that is 1-1 and onto (i.e., bijective).

Theorem 2.1. Let $A \sim B$ be a relation between two sets having the same cardinality. Then is an equivalence relation. That is,

(a) $A \sim A$ (Reflexive)

(b) $A \sim B \Rightarrow B \sim A$ (Symmetry)

(c) $A \sim B \& B \sim C \Rightarrow A \sim C$ (Transitivity)

Definition 2.2. Let $\mathbb{N} = \{1, 2, 3, \ldots\}$. Let $J_n = \{1, 2, \ldots, n\}$ for $n \in \mathbb{N}$.

- A set A is finite if $A \sim J_n$ for some $n \in \mathbb{N}$(or if $A = \emptyset$).
- A set A is countably infinite if $A \sim \mathbb{N}$.
- A set A is countable if A is finite or countably infinite.

Example. *$\mathbb{Z}$ is a countably infinite. For $n \in \mathbb{Z} = \{0, +1, -1, +2, -2, \ldots\}$,*

Let $f(n) = \begin{cases} 1 & \textit{if } n = 0 \\ 2n & \textit{if } n > 0 \\ -2n + 1 & \textit{if } n < 0 \end{cases}$

Then f is bijective and therefore $|\mathbb{Z}| = |\mathbb{N}|$

Theorem 2.8. A subset of a countably infinite set is countable.

Proof. Let A be some countably infinite set and S be a infinite subset of A.
As A is a countably infinite set, we can remove duplicates and arrange A so that $A = \{a_1, a_2, a_3, \ldots\}$. Let n_1 be the smallest positive integer such that $x_{n_1} \in S$. Let n_k be the smallest positive integer greater than n_{k-1} such that $x_{n_k} \in E$ for $k = 2, 3, \ldots$. Let $f(k) = x_{n_k}$ for $k = 1, 2, 3, \ldots$. Then this is a bijection from $\mathbb{N}$ to S. Q.E.D.

Remark. Roughly speaking, countable sets represent the **smallest infinity**, as no uncountable set can be a subset of a countable set.

Theorem 2.12. Let $E_1, E_2, \ldots$ be countably infinite sets. Then $S = \cup_{n=1}^{\infty} E_n$ is countably infinite.

Proof. Write $E_1 = \{x_{11}, x_{12}, x_{13}, x_{14}, \ldots\}$
$E_2 = \{x_{21}, x_{22}, x_{23}, x_{24}, \ldots\}$
Form an array:

$$\begin{Bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & \cdots \\ x_{21} & x_{22} & x_{23} & x_{24} & \cdots \\ x_{31} & x_{32} & x_{33} & x_{34} & \cdots \\ x_{41} & x_{42} & x_{43} & x_{44} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{Bmatrix}.$$

This matrix might have duplicates. Let T be a subset of $\mathbb{N}$ such that $t \in T$ if and only if t is the smallest positive integer such that $x_t \in E_1 \cup E_2 \cup \ldots \cup E_n$.
Then a set $\{x_t | t \in T \text{ and } \exists_{i \in \mathbb{N}} : x_t \in E_i\}$ is S. Clearly, $|S| = |T|$, or $S \sim T$, and T is a subset of a countably infinite set, $\mathbb{N}$. Therefore, T is also countable, implying S is also countable. As S is infinite, S is countably infinite. Q.E.D.

Corollary 2.13. If A is countable and $n \in \mathbb{N}$, then $\{(a_1, a_2, \ldots, a_n) | a_1, a_2, \ldots, a_n \in A\}$ is countable.

Theorem 2.14. Let $A = \{(b_1, b_2, b_3 \ldots) | b_i \in \{0, 1\}\}$. I.e., A is a set of all infinite binary strings. Then A is uncountable.

Proof (Contor's Diaganalization argument,1891)**.** Let $E \subset A$ be countably infinite. $E = \{s^{(1)}, s^{(2)}, s^{(3)}, \dots \ |s^{(i)} \in A\}$. It suffices to find some $s \in A \setminus E$, for this shows every countably infinite subset of A is proper construction of s. Write

$$s^{(1)} = b_1^1 b_2^1 \dots \tag{2.1}$$

$$s^{(2)} = b_1^2 b_2^2 b_3^2 \dots \tag{2.2}$$

$$s^{(3)} = b_1^3 b_2^3 b_3^3 \dots \tag{2.3}$$

$$\vdots$$

On diagonal, flip each bit, i.e., $0 \to 1$ and $1 \to 0$ and represent the flipped bit of b_i^i by $\tilde{b_i^i}$. Let $s = \tilde{b_1^1}\tilde{b_2^2}\tilde{b_3^3}\dots$. Then $s \in A$ and $s \notin E$ as s differs from each $s^{(i)}$ in the i-th bit. Therefore, A is uncountable. Q.E.D.

Corollary 2.15. The set $\mathcal{P}(\mathbb{N})$ of subsets of $\mathbb{N}$ is uncountable.

Proof. We can create $f : \mathcal{P}(\mathbb{N}) \to A$ be a bijection, where A is the set of all infinite binary strings, by

$$f(S)_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \notin S \end{cases} \tag{2.4}$$

For example, if $f(\{\text{odd natural numbers}\}) = (1, 0, 1, 0, 1, 0, 1, 0 \dots)$. This f is a bijection, and therefore A is uncountable.

Q.E.D.

Theorem 2.16. $\mathbb{R}$ is uncountable.

Proof. This is a rough sketch of the proof:

(a) It's enough to show that $[0, 1]$ is uncountable.

(b) Consider binary decimal representation of $x \in [0, 1]$. For example, $x = 0.101001001\ldots$. Given x, choose maximal $b_1 \in \{0, 1\}$ such that $\frac{b_1}{2} \leq x$. Then choose $b_2 \in \{0, 1\}$ such that $\frac{b_1}{2} + \frac{b_2}{2} \leq x$. Continue this process to get $b_1, b_2, b_3, \ldots$. Then $x = \sup\left\{\sum_{i=1}^{n} \frac{b_i}{2^i}\right\}$. Consider any dyadic rational of the form $\frac{m}{2^n}$. Let it be $\frac{3}{2^4}$. Then this maps $\frac{3}{2^4} \to 0, 0, 1, 1, 0, 0, 0, \ldots$ and never produce $0, 0, 1, 0, 1, 1, 1, 1, \ldots$, which also represents $\frac{3}{2^4}$. Let A_1 be a subset of $A = \{\text{infinite binary strings}\}$ such that A_1 does not contain any strings ending in $1, 1, 1, 1, \ldots$. Then the decimal representation defines a bijection $f : [0) \to A \setminus A_1$.

(c) A_1 is countable because $A = (A \setminus A_1) \cup A_1$, which is uncountable.

This shows that $[0, 1]$ is uncountable, and therefore $\mathbb{R}$ is uncountable.
Q.E.D.

Definition 2.3 (Metric Spaces)**.** A set X is a metric space with metric $d : X \times X \to \mathbb{R}$ if

(a) $d(p, q) > 0$ if $p \neq q$ and $d(p, q) = 0$ if $p = q$, $\forall p, q \in X$

(b) $\forall_{p,q \in X} : d(p, q) = d(q, p)$

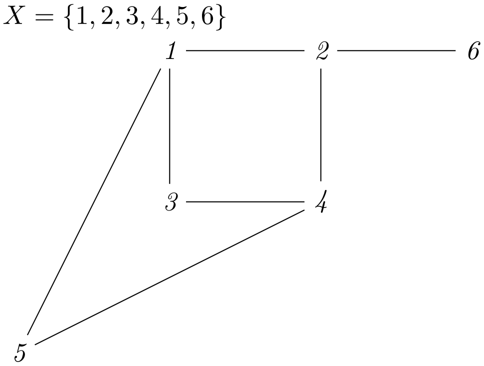
(c) $\forall_{p,q,r \in X} : d(p, q) \leq d(p, r) + d(r, q)$ (Triangle Inequality)

Remark. A metric space does not need to be an ordered set or to have addition or multiplication defined on it.

Example (Metric Spaces)**.** *(a)* $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{R}^k$ *are metric spaces with* $d(p, q) = |p - q|$. *Note the meaning of* $|x|$ *depends on the context.*

(b) Every subset of a metric space is a metric space.

(c) $X = \{1, 2, 3, 4, 5, 6\}$



Definition 2.4 (Neighborhood)**.** A neighborhood in X is a set $N_r(p) := \{q : d(q,p) < r\}$, where $p \in X, r > 0$.

> **Remark.** If $r_1 \le r_2$, then $N_{r_1}(p) \subset N_{r_2}(p)$.

Example.

$\mathbb{R}^1$ *intervals,* $N_r(x) = \{y \in R^1 : |x-y| < r\}$

$\mathbb{R}^2$ *disks* $N_r(x) = \{y \in R^2 : |x-y| < r\}$

$\mathbb{R}^3$ *balls,* $N_r(x) = \{y \in R^3 : |x-y| < r\}$

Given example (c), $N_1(2) = \{2\} = N_{\frac{1}{2}}(2)$, $N_2(2) = \{1,2,4,6\}$, $N_3(2) = \{1,2,3,4,5,6\} = X$.

Definition 2.5. Let $E \subset X$. $p \in E$ is an interior point of E if $\exists r > 0$ such that $N_r(p) \subset E$.

Example.

$X = \mathbb{R}^2, E = \{x \in \mathbb{R} : |x| \le 1\}$

$X = \mathbb{N}, E \subset X$.

Definition 2.6. $E \subset X$ is an open set if $\forall_{x \in E}$ is an interior point of E.

> **Theorem 2.19.** Every neighborhood is an open set.
>
> > **Proof.** Let $g \in N_r(p)$. Then we must find $s > 0$, such that $N_s(g) \subset N_r(p)$. We know $d(p,q) < r$. Choose s such that $0 < s < r - d(p,q)$. Let $x \in N_s(q)$, then $d(q,x) < s < r - d(p,q)$. By triangle inequality, $d(p,x) \le d(p,q) + d(q,x) < d(p,q) + r - d(p,q)$, so $x \in N_r(p)$, so $N_s(q) \subset N_r(p)$. Q.E.D.

Definition 2.7. Let $E \subset X$ and $p \in X$. p is a limit point of E if $\forall_{r>0} \exists_{q \in E}$ such that $q \ne p$ and $q \in N_r(p)$

Example. $E = \{\frac{1}{n} : n \in \mathbb{N}\} \subset \mathbb{R}$ *has exactly one limit point,* 0. *note* $0 \notin E$.

> **Theorem 2.20.** If p is a limit point of $E \subset S$, then every neighborhood of p contains infinitely many points of E.
>
> > **Proof.** Let $N_r(p)$ be a neighborhood of p. Then $N_r(p)$ contains at least one point $q_1 \in E$ such that $q_1 \ne p$. Let $r_1 = d(p,q_1)$. Then $N_{r_1}(p)$ contains some $q_2 \in E$ such that $q_2 \ne p$. Let $r_2 = d(p,q_2)$. Then $N_{r_2}(p)$ contains some $q_3 \in E$ such that $q_3 \ne p$. Continue this process to get $q_1, q_2, q_3, \ldots$. Q.E.D.

Corollary 2.21. If $E \subset X$ is finite then E has no limit points.

Definition 2.8 (Closed Set)**.** A set $E \subset X$ is closed if every limit point of E is in E.

Theorem 2.23. $E \subset X$ is open iff $E^c = \{x \in X : x \notin E\}$ is closed.

Proof.

- E is open $\Rightarrow E^c$ is closed.
  Let p be a limit point of E^c. Then every neighborhood of p contains some $q \in E^c$ such that $q \neq p$. If $p \in E$, then because E is open, p is an interior point, i.e., there exists some neighborhood of p that is a subset of E, which does not contain any points of E^c. This implies $p \notin E$ and therefore $p \in E^c$.

- E^c is closed $\Rightarrow E$ is open.
  Let $p \in E$. Then $p \notin E^c$, so p is not a limit point of E^c. Therefore, there exists some neighborhood of p that contains no points of E^c, i.e., all points of the neighborhood are in E. p Thus, Every $p \in E$ is an interior point of E, and hence E is open.

Q.E.D.

Theorem 2.24 (De Morgan's Laws)**.**

(a) $(\bigcup_\alpha E_\alpha)^c = \bigcap_\alpha E_\alpha^c$

(b) $(\bigcap_\alpha E_\alpha)^c = \bigcup_\alpha E_\alpha^c$

Theorem 2.24.

(a) For all collection $\{G_\alpha\}$ of open sets : $\bigcup_\alpha G_\alpha$ is open.

(b) For all collection $\{F_\alpha\}$ of closed sets : $\bigcap_\alpha F_\alpha$ is closed.

(c) For all finite collection $\{G_1, G_2, \ldots, G_n\}$ of open sets : $\bigcap_{i=1}^n G_i$ is open.

(d) For all finite collection $\{F_1, F_2, \ldots, F_n\}$ of closed sets : $\bigcup_{i=1}^n F_i$ is closed.

Proof. (a) Let $x \in \bigcup_\alpha G_\alpha$. Then $x \in G_{\alpha_0}$ for some α_0. So there exists a neighborhood N of x such that $N \subset G_{\alpha_0} \subset \bigcup_\alpha G_\alpha$.

(b) it's suffice to prove that $(\bigcap_\alpha F_\alpha)^c$ is open. But $(\bigcap_\alpha F_\alpha)^c = \bigcup_\alpha F_\alpha^c$ is open by (a).

(c) Let $x \in \bigcap_{i=1}^n G_i$. Then $x \in G_i$ for $i = 1, 2, \ldots, n$. So there exists a $r_i > 0$ such that $N_{r_i}(x) \subset G_i$. Let $r = \min\{r_1, r_2, \ldots, r_n\}$. Then $N_r(x) \subset N_{r_i} \subset G_i$ for $i = 1, 2, \ldots, n$ and therefore $N_r(x) \subset \bigcap_{i=1}^n G_i$.

(d) $(\bigcup_{i=1}^n F_i)^c = \bigcap_{i=1}^n F_i^c$ is open by (c).

Q.E.D.

Definition 2.9 (Closure). Let $E \subset X$. Let E' be a set of limit points of E in X. The set $\overline{E} = E \cup E'$ is the closure of E.

Theorem 2.27.

(a) $\overline{E}$ is closed.

(b) $E = \overline{E} \Leftrightarrow E$ is closed.

(c) If $F \subset X$ is closed and $E \subset F$, then $\overline{E} \subset F$. (i.e., $\overline{E}$ is the smallest closed set containing E, and $\overline{E} = \bigcap_{F:\text{closed set with } F \supset E} F$.)

Proof. (a) Let p be a limit point of $\overline{E}$. It suffices to show $p \in E'$ since this implies that $p \in E' \subset E \cup E' = \overline{E}$. Let $r > 0$. $\exists_{q \in \overline{E}, q \neq p} : q \in N_{\frac{r}{2}}(p)$, i.e., $d(p,q) < \frac{r}{2}$. Since $q \in E \cup E'$, $\exists_{s \in E}$ such that $d(q,s) < \frac{r}{2}$ (if $q \in E$, take $s = q$). But $d(p,s) \leq d(p,q) + d(q,s) < \frac{r}{2} + \frac{r}{2} = r$.

(b) $(\Rightarrow)$ by (a)

$(\Leftarrow)$ Suppose E is closed. Then $E' \subset E$, so $\overline{E} = E \cup E' = E$.

(c) Suppose F is closed. Then $F' \supset E'$ and also $F \supset F'$. So $F = \overline{F} = F \cup F' \supset E \cup E' = \overline{E}$

Q.E.D.

Theorem 2.28. Let E be a nonempty set of real numbers which is bounded above. Let $y = \sup\{E\}$. Then $y \in \overline{E}$. Hence, $y \in E$ if E is closed.

Example. *Let $X = \mathbb{R}, d(p,q) = |p-q|$. Let $E \subset \mathbb{R}$ be nonempty and bounded above, and let $y = \sup E$. Then $y \in \overline{E}$.*

Proof. Suppose for contradiction $y \notin \overline{E}$. Then y is neither a point in E nor a limit point of E, so $\exists$ some interval $N_r(y) = (y-r, y+r)$ such that $(y-r, y+r) \cap E = \emptyset$. However, then $y-r$ in an upper bound for E since y is a least upper bound, which is a contradiction. Therefore, $y \in \overline{E}$. Q.E.D.

Definition 2.10 (Relative Openness)**.** Suppose X is a metric space, so $Y \in X$ is a metric space with the same metric. Let $E \subset Y$. Then E is open relative to Y if E is an open set in the metric space Y

Example. *$X = R^2 \supset \mathbb{R} = y, E = (0,1) \subset Y$. Then E is open relative to Y, but E is neither open nor closed in X.*

Theorem 2.30. A set $E \subset Y \subset X$ is open relative to $Y \Leftrightarrow \exists_{\text{open set } G \subset X} :$ $E = G \cap Y$

Proof. ($\Rightarrow$) Suppose $E \subset Y$ is open relative to Y. Given $p \in E$, $\exists_{r_p > 0} : {N_{r_p}}^Y(p) \subset E$, where ${N_r}^Y(p) = \{q \in Y : d(p,q) < r\}$. Then $E \subset \bigcup_{p \in E} {N_{r_p}}^Y(p)$ and $\bigcup_{p \in E} {N_{r_p}}^Y(p) \subset E$. Therefore, $E = \bigcup_{p \in E} {N_{r_p}}^Y(p)$.
Let $G = \bigcup_{p \in E} {N_{r_p}}^X(p)$. This time, we are considering p's neighborhood in X, so each ${N_{r_p}}^X$ is open. Thus G is a union of open sets in X, and therefore open.
$\forall_{p \in E} : p \in N_{r_p}(p)^X$, so $E \subset G \cap Y$.
Let $p \in G \cap Y$. Then $p \in G$ and $p \in Y$. So $p \in {N_{r_p}}^X(p)$ for some $r_p > 0$. But $p \in Y$, so $p \in {N_{r_p}}^Y(p)$. Therefore, $p \in E$. This implies $G \cap Y \subset E$, and therefore $E = G \cap Y$.

($\Leftarrow$) Suppose $G \subset X$ is open and $E = G \cap Y$. Then $\forall_{p \in E} : \exists_{r_p > 0} :$ ${N_{r_p}}^X(p) \subset G$, so ${N_{r_p}}^Y(p) = {N_{r_p}}^X(p) \cap Y \subset G \cap Y = E$.

Q.E.D.

Note: Midterm 1 material ends here.

Definition 2.11 (Open Cover)**.** An open cover of $E \subset X$ is a collection $\{G_\alpha\}$ of open subsets of X s.t $E \subset \bigcup_\alpha G_\alpha$.

Definition 2.12 (Compact)**.** A set $K \subset X$ is compact if every open cover has a finite subcover; i.e., $\exists_{\alpha_1, \alpha_2, \ldots \alpha_n} :$ s.t $K \subset G_{\alpha_1} \cup G_{\alpha_2} \cup \ldots \cup G_{\alpha_n}$

Example.

- *If E is finite, then E is compact.*
- *$(0,1) \subset \mathbb{R}$ is not compact. Bad cover: $(\frac{1}{n}, 1), n > 2$*
- *$[0, \infty] \subset \mathbb{R}$ is not compact. Bad cover: $(-1, n)$ for $n \in \mathbb{N}$.*
- *$E \subset \mathbb{R}^k$ is compact if and only if E is closed and bounded.*

Theorem 2.34. If K is compact then K is closed.

Proof. Suppose K is compact. It suffices to prove that K^c is open. Let $p \in K^c$. We need to produce $r > 0$ s.t. $N_r(p) \subset K^c$. For $q \in K$, let $W_q = N_{r_q}(q)$, where $r_q = \frac{1}{2}d(p,q) > 0$. $\forall_{x \in N_{r_q}(p)} : x \in W_q \Rightarrow d(x,p) + d(x,q) < 2r_q = d(p,q)$. However, X is a metric space and $p, q, x \in X$, so $d(p,q) \leq d(p,x) + d(x,q)$, leading to $d(p,q) \leq d(p,x) + d(x,q) < d(p,q)$, which is a contradiction. Hence, $\forall_{x \in N_{r_q}} : x \notin W_q$. $N_{r_q}(p) \subset {W_q}^c$ for $\forall_{q \in K}$. Note that $\{W_q\}_{q \in K}$ is an open cover of K. K compact $\Rightarrow \exists_{\text{finite number of open sets } W_{q_1}, W_{q_2}, \dots W_{q_n}}$ s.t. $K \subset \bigcup_{i=1}^{n} W_{q_i}$. Let $r = \min\{r_{q_1}, r_{q_2}, \dots r_{q_n}\} > 0$.

$$\therefore N_r(p) \subset \left(\bigcap_{i \in \{1,2,\dots n\}} N_{r_p}(p) \right) \subset \left(\bigcap_{i \in \{1,2,\dots n\}} {W_{q_i}}^c \right) = \left(\bigcup_{i \in \{1,2,\dots \mathbb{N}\}} W_{q_i} \right)^c \subset K^c$$

Q.E.D.

Theorem 2.35. If $K \subset X$ is compact then K is bounded; i.e., $\exists_{M<\infty}$ s.t. $\forall_{p,q \in K}$: $d(p,q) \leq M$

Proof. Fix $p \in K$. An open cover of K is $\{N_n(p)\}_{n \in \mathbb{N}}$. In fact, this is an open cover of X. K compact $\Rightarrow \exists_{\text{finite subcover} N_{n_1}(p), N_{n_2}(p) \dots N_{n_m}(p)}$. Let $R = \max\{n_1, n_1, \dots n_m\}$. $K \subset N_R(p)$. Let $M = 2R$. $\forall_{q,r \in K}$: $d(q,r) \leq d(q,p) + d(p,r) < R + R = 2R = M$. Q.E.D.

Theorem 2.35. If F is closed, K is compact, and $F \subset K$ then F is compact.

Proof. Suppose $F \subset K$. Let $\{V_\alpha\}$ be an open cover of F. It suffice to produce a finite subcover:
Consider $\{V_\alpha\}$ together with F^c. This gives an open cover of X, hence of K, so $\exists_{\text{subcover of } K}$. Drop F^c from this finite subcover. The result is a finite subcover of $\{V_\alpha\}$, which covers F Q.E.D.

Corollary 2.36. If F is closed and K is compact then $F \cap K$ is compact.

Theorem 2.33. Suppose $K \subset Y \subset X$. Then K is compact relative to X iff K is compact relative to Y.

Note. This is not true for open sets. For instance, let $K = Y = [0, 1] \subset X = \mathbb{R}$. Y is open and closed relative to Y, but Y is not open relative to X

Proof.

(⇒) Suppose K is compact relative to X. Let $\{V_\alpha\}$ be an open cover of K relative to Y. For any subset of a metric space, there always exists an open cover relative to the whole space, as the whole metric space is trivially an open cover of itself. Then $\{V_\alpha\}$ is an open cover of K relative to X. Since K is compact relative to X, $\exists_{\text{finite subcover}}$.

(⇐) Suppose K is compact relative to Y. Let $\{V_\alpha\}$ be an open cover of K relative to X. Then $\{V_\alpha \cap Y\}$ is an open cover of K relative to Y. Since K is compact relative to Y, $\exists_{\text{finite subcover}}$.

Q.E.D.

Theorem 2.36. Suppose $\{K_\alpha\}$ is a collection of compact sets such that $\bigcap_{i\in\{1,2,\ldots,n\}} K_{\alpha_i} \neq \emptyset$ for any $n < \infty, \alpha_i$. Then, $\lim\limits_{n\to\infty} \bigcap_{i\in\{1,2,\ldots,n\}} K_{\alpha_i} \neq \emptyset$, or equivalently, $\bigcap_\alpha K_\alpha \neq \emptyset$.

Example. *Let $G_j = (0, \frac{1}{j}) \subset \mathbb{R}$. Then $\{G_j\}$ is a collection of open sets, but none of them are compact. (compact sets are closed) Then $\{G_j\}$ satisfies non-empty finite intersection property but $\bigcap_{j\in\mathbb{N}} G_j = \emptyset$.*

Proof. Suppose for contradiction $\bigcap_{i\in\{1,2,\ldots,n\}} K_{\alpha_i} \neq \emptyset$ for any $n < \infty, \alpha_i$ and $\bigcap_\alpha K_\alpha = \emptyset$. For any α_0, $K_{\alpha_0} \cap \left(\bigcap_{\alpha\neq\alpha_0} K_\alpha\right) = \emptyset$. Hence, $K_{\alpha_0} \subset \left(\bigcap_{\alpha\neq\alpha_0} K_\alpha\right)^c = \bigcup_{\alpha\neq\alpha_0} (K_\alpha)^c$ and $\{(K_\alpha)^c\}_{\alpha\neq\alpha_0}$ is an open cover of K_{α_0}, so $\exists$ a finite subcover of $K_{\alpha_0} \subset \bigcup_{i=1}^n {K_{\alpha_i}}^c$, which implies $K_{\alpha_0} \cap (\bigcap_{i=1}^n K_{\alpha_i}) = \emptyset$, contradiction. Q.E.D.

Corollary 2.37. If $\{K_1, K_2, \ldots\}$ are non-empty compact sets with $\forall_n$: $K_n \supset K_{n+1}$, then $\bigcap_{n=1}^\infty K_n \neq \emptyset$.

Proof. If $n_1 < n_N$ then $\bigcap_{i=1}^N K_{n_i} = K_{n_N} \neq \emptyset$ Q.E.D.

Theorem 2.37. If K is compact and $E \subset K$ is infinite, then E has a limit point in K.

Proof. Contrapositive of the statement is : *if $E \subset K$ has no limit point in K, then E is finite.*
Suppose every point $q \in K$ is not a limit point of E. Then

$$\exists_{V_q = N_{r_q}(q)} : V_q \cap E = \begin{cases} \emptyset & \text{if } q \notin E \\ \{q\} & \text{if } q \in E \end{cases}.$$

$\{V_q\}_{q \in K}$ is an open cover of K, so $\exists_{\text{finite subcover } V_{q_1} \cup V_{q_2} \cup \cdots \cup V_{q_n}}$. Then $E = E \cap K \subset (\bigcup_{i=1}^{n} V_{q_i} \cap E) \subset \{q_1, q_2, \ldots q_n\}$, so E is finite.

Q.E.D.

Theorem 2.38. Let $I_n = [a_n, b_n] \subset \mathbb{R}$ be such that $\forall_n : I_n \supset I_{n+1}$. Then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$.

Proof. Since $I_n \supset I_{n+1}$, $\forall_{n,m} : a_n \leq a_{n+m} \leq b_{n+m} \leq b_n$. Let $E = \{a_1, a_2, \ldots\}$. Then $E \neq \emptyset$, every b_k is an upper bound for E, so $\exists x = \sup E$ and $a_k \leq x \leq b_k$ for all k. Therefore, $x \in I_k$ for all k, so $x \in \bigcap_{n=1}^{\infty} I_n$. Q.E.D.

Theorem 2.39. Let $\{I_n\}$ be a sequence of k-cells such that $i_n \supset I_{n+1}$;i.e., $I_n = \{\mathbf{x} = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k : a_{nj} \leq x_j \leq b_{nj}, a_{nj} \leq a_{n+1,j} \leq b_{n+1,j} \leq b_{nj}$ for $j = 1, 2, \ldots, k\}$. Then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$.

Proof. Apply previous theorem to each component. Q.E.D.

Note. k-cell is a higher dimensional analog of a rectangle or rectangular solid, which is a Cartesian product of k *closed intervals on the real line.*
Formally, Given real numbers a_i and b_i such that $a_i < b_i$ for every integer i from 1 to k,

$$I = \{x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k : a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \ldots, k\}$$

Theorem 2.40. Let $I \subset \mathbb{R}^k$ be a k-cell. Then I is compact.

Proof. Let $I = I_0 = \{x = (x_1, x_2, \dots, x_k), a_j \le x_j \le b_j\}$.
Let $\Delta = \left\{\Sigma_i^k (b_j - a_j)^2\right\}^{1/2}$. Then $|\mathbf{x} - \mathbf{y}| \le \Delta$ for $\mathbf{x}, \mathbf{y} \in I$.
Suppose for contradiction $\{G_\alpha\}$ is an open cover of I that has no finite subcover.
Let $c_j = \frac{1}{2}(a_j + b_j)$ for $j = 1, 2, \dots, k$. Using $[a_j, c_j], [c_j, b_j]$, we get 2^k k-cells Q_i with $I = \bigcup_{i=1}^{2^k} Q_i$. At least one Q_i, call it I_1, has no finite subcover. Otherwise, every Q_i has a finite subcover, and I would have a finite subcover, namely the union of the finite subcovers of each Q_i.
Repeat this step to construct $I_0 = I, I_1, I_2, \dots$. Then the sequence $\{I_n\}$ constructed by this process satisfies the following properties:

(a) $I_0 = I \supset I_1 \supset I_2 \supset I_3 \supset \dots$

(b) $\forall_n : I_n$ has no finite subcover from $\{G_\alpha\}$

(c) if $x, y \in I_n$ then $|x - y| \le 2^{-n}\Delta$, where $\Delta =$ diagonal of $I = \left(\sum_{j=1}^k (b_j - a_j)^2\right)^{1/2}$.

By theorem 2.38 and (a), $\exists_{x^* \in \bigcap_{n=1}^\infty I_n}$. Since $x^* \in I$, $x^* \in G_{\alpha_0}$ for some α_0, so $\exists r > 0$ such that $N_r(x^*) \subset G_{\alpha_0}$. But by (c), $I_n \subset N_{2^{-n}\Delta}(x^*)$. As soon as n is large enough that $2^{-n}\Delta < r$, we have $I_n \subset N_{2^{-n}\Delta}(x^*) \subset G_{\alpha_0}$, which contradicts (b). Q.E.D.

Note. Reverse triangle inequality
$\forall_{a,b,c \in X} : d(a,b) \ge d(a,c) - d(c,b)$ because $d(a,c) \le d(a,b) + d(b,c)$.

Theorem 2.41. For $E \subset \mathbb{R}^k$, the following are equivalent:

(a) E is closed and bounded.

(b) E is compact.

(c) Every infinite subset of E has a limit point in E.

Proof.

$(a) \Rightarrow (b)$ Because E is bounded, i.e., $\exists_M$ s.t. $\forall_{x,y \in E} : |x - y| \leq M$, there exists a k-cell I such that $E \subset I$. Since every k-cell is compact, this implies E is a closed subset of a compact set. Hence, E is also compact.

$(b) \Rightarrow (c)$ by theorem 2.37

$(c) \Rightarrow (a)$ To see that E is bounded, suppose it were not. Then E has an infinite subset $S = \{x_1, x_2, x_3, \ldots\}$ with $\forall_n : |x_n| \geq n$. S has no limit point in $\mathbb{R}^k$ Let $S = \{(x_1, x_2, x_3, \ldots) \in E : |x_n - x_0| < \frac{1}{n}\}$. Then S is an infinite set because if S is finite, there exists a point $\mathbf{x} \in S$ such that $|\mathbf{x}| \geq |\mathbf{x}'|$ for $\mathbf{x}' \in S$. However, there exists $n \in \mathbb{N}$ such that $n > |\mathbf{x}|$ and by definition of S, there exists $x_n \in S$ such that $|x_n| \geq n > |\mathbf{x}|$, which is a contradiction. Thus, S is infinite. This S however, cannot have a limit point in E. By triangle inequality, for any $y \in R^k$, $|x_n| \leq |x_n - y| + |y|$, and from archimedean property, $\exists_{m \in \mathbb{N}}$ s.t. $m > |x_n - y| + |y|$, which implies for any $y \in R^k, r > 0$, $\exists_{m \in \mathbb{N}} : |x - y| < r < m$. However, by the definition of S, there are at most m such elements in S. Since a limit point y of E must contain an infinite number of points of E such that $d(x, y) < r$ for any $r > 0$, y cannot be a limit point, which contradicts the assumption that any infinite subset of E contains a limit point in E. Therefore, E must be bounded.
To see that E is closed, suppose it were not closed. Then $\exists_{x_0 \in} : E' \setminus E$. If T has no limit point in E except $x_0 \notin E$, it contradicts (c) because T is infinite and there must be a limit point of T in E.
Therefore, we can show that E is closed by showing that T has no limit point in E except x_0. Form an infinite sequence $(x_1, x_2, x_3, \ldots), x_n \in E$ with $|x_n - x_0| < \frac{1}{n}$. Let $y \in E$, $y \neq x_0$. We'll show that y cannot be a limit point of T. $|y - x_n| \geq |y - x_0| + |x_0 - x_n| > |y - x_0| - |x_0 - x_n| > |y - x_0| - \frac{1}{n}$. Choose $n \geq \frac{2}{|y - x_0|}$, so $\frac{1}{n} \leq \frac{|y - x_0|}{2}$. Then $|y - x_n| \geq |y - x_0| - \frac{1}{2}|y - x_0| = \frac{1}{2}|y - x_0|$. So only finitely many x_n can lie in $N_{\frac{1}{2}|y - x_0|}(y)$. So y cannot be a limit point of S. Therefore, E is closed.

Q.E.D.

Remark. (b) and (c) are equivalent in any metric space, but (a) does not, in general, imply (b) or (c) in a metric space other than $\mathbb{R}^k$.

Example. *Failure of Heine-Borel theorem in general metric spaces.*

some infinite set , discrete metric $d(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$. *Then E is bounded and closed but not compact.*

Theorem 2.42. [Weirstrass's theorem] Every bounded infinite subset $E \subset \mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.

Proof. Choose a k-cell $I \supset E$. Since I is compact, by theorem 2.41, E has a limit point in I. Q.E.D.

Example. *Let*

$$E_0 = [0,1] \tag{2.5}$$

$$E_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1] \tag{2.6}$$

$$E_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{1}{3}] \cup [\frac{2}{3},\frac{7}{9}] \cup [\frac{8}{9},1] \tag{2.7}$$

$$\vdots \tag{2.8}$$

This gives $E_0 \supset E_1 \supset E_2 \supset E_3 \supset \ldots$, *where each* E_n *is the union of* 2^n *closed intervals of length* $\frac{1}{3^n}$.

Definition 2.13 (Perfect Sets)**.** A set P is perfect if there is no isolated point in P; i.e.,

$$P = P'.$$

Theorem 2.43. Let P be a non-empty perfect set in R^k. Then P is uncountable.

Proof. Suppose for contradiction P is countable. Since P is non-empty, there exists some $p_1 \in P$. p_1 is then also a limit point of P. Let $p_2 \in P (\neq p_1)$ be a point in $V_1 = N_{r_1}(p_1)$ for some r_1 such that $d(p_1,p_2) > r_1/2$. Let $r_2 = r_1 - d(p_1,p_2)$, $V_2 = N_{r_2}p_2$. Then $\forall_{x \in V_2} : d(p_1,x) \leq d(p_1,p_2) + d(p_2,x) < d(p_1,p_2) + r_2 = r_1$. Hence, $V_2 \subset V_1$. $\overline{V_2} \subset V_1$ as well. Also, note that $d(p_1,p_2) > r_1/2$, so $r_2 = r_1 - d(p_1,p_2) < r_1/2 < d(p_1,p_2)$. So $p_1 \notin \overline{V_2}$. Repeat this process, and let $K_n = \overline{V_n} \cap P$. $K_n \subset \overline{V_n}$. Since $\overline{V_n}$ is closed and bounded, it's compact. $\overline{V_n} \cap P$ is a closed subset of $\overline{V_n}$, so K_n is also compact. However, for any p_n, $p_n \notin K_{n+1}$, so $\bigcap_{1 \in \infty} K_n \cap P = \emptyset$. Since $K_n \subset P$, this implies $\bigcap_{1 \in \infty} K_n = \emptyset$, but each K_n is not empty, $K_n \supset K_{n+1}$, and K_n is compact. Thus, $\bigcap_{1 \in \infty} K_n \cap P$ can't be empty, so this is a contradiction. Q.E.D.

Definition 2.14 (Cantor Set)**.** The cantor set $P := \bigcap_n^\infty E_n$.

Proposition. P is compact, non-empty and contains no open intervals (a, b) and uncountable.

Proof. **Compactness** P is compact because $P \subset E_0 = [0, 1]$ and E_0 is compact.

Non-emptiness P is non-empty because $P \subset E_0$ and E_0 is non-empty.

No open intervals P contains no open intervals (a, b) because any (a, b) contains some $(\frac{3k+1}{3^n}, \frac{3k+2}{3^n})$ and these are all removed.

Uncountability P is uncountable because P is a perfect set. Equivalently, P consists of points in $[0, 1]$ whose ternary, i.e., base 3, representation contains only 0's and 2's.

Note. ternary representation: $0.a_1a_2a_3 \ldots = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$ where $a_n \in \{0, 1, 2\}$.

Q.E.D.

Example (Cantor Set)**.** *Let* $E = [0, 1]$, $E_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. $E_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$, *etc. Keep removing open middle third. This gives* $E_0 \supset E_1 \supset E_2 \ldots$. *Each* E_n *is the union of* 2^n *closed intervals of length* $\frac{1}{3^n}$.

Definition 2.15 (Separated Sets)**.** **Separated Sets** $A, B \subset X$ are separated if $\overline{A} \cap B = A \cap \overline{B} = \emptyset$.

Connected Sets $E \subset X$ is connected if there is no non-empty separated sets $A, B \subset E$.

Example (Separated Sets)**.** *In* $\mathbb{R}^1$, $[0, 1)$ *and* $(1, 2]$ *are separated so* $[0, 1) \cup (1, 2]$ *is not connected. Every interval is connected (open, closed, semi-open).*

Theorem 2.47. $E \subset \mathbb{R}^1$ is connected if and only if E is an interval; i.e., $\forall_{x,y \in E, x<y}$ s.t. $\forall_{z \in (x,y)} : z \in E$

Proof. Let $x, y \in E$. Q.E.D.

Theorem 2.48. A metric space X is connected if and only if the only nonempty subset of X which is both open and closed is X itself. $|\limsup_{n\to\infty} |\gamma_n|| \leq 0 + \varepsilon\alpha$. Since ε is arbitrary, this implies $|\limsup_{n\to\infty} |\gamma_n|| = 0$, so $\lim\limits_{n\to\infty} |\gamma_n| = 0$.

Chapter 3

Sequence and Series

3.1 Sequences

Definition 3.1. In a metric space (X, d), a sequence $\{p_n\}$ converges to p if $\forall_{\varepsilon>0}\exists_N$ s.t. $n \geq N \Rightarrow d(p_n, p) < \varepsilon$.
We write $\lim\limits_{n\to\infty} p_n = p$ or $p_n \to p$.

If $\{p_n\}$ does not converge to any p then it is said to diverge.

Theorem 3.3. If s_n and t_n are sequences in $\mathbb{C}$ with $s_n \to s$ and $t_n \to t$, then the following hold:

(a) $s_n + t_n \to s + t$

(b) $cs_n \to cs$, $c + s_n \to c + s$ for any $c \in \mathbb{C}$

(c) $s_n t_n \to st$

(d) $\frac{1}{s_n} \to \frac{1}{s}$ if $s \neq 0$

Lemma (Squeeze Lemma)**.** In $\mathbb{R}$, if $\forall_{n\in\mathbb{N}} : 0 \leq x_n \leq s_n$ and $\lim\limits_{n\to\infty} s_n \to 0$, then $\lim\limits_{n\to\infty} x_n = 0$.

Proof. Let $\varepsilon > 0$. Choose N such that $n \geq N \Rightarrow 0 \leq s_n < \varepsilon$. Then $0 \leq x_n \leq s_n < \varepsilon$ for $n \geq N$, so $x_n \to 0$. Q.E.D.

Theorem 3.20. (a) If $p > 0$ then $\frac{1}{n^p} \to 0$.

Proof. Let $\varepsilon > 0$. Choose N such that $\frac{1}{N^p} < \varepsilon$; i.e., $N > \frac{1}{\varepsilon^{\frac{1}{p}}}$. Then for $n \geq N$, $\frac{1}{n^p} \leq \frac{1}{N^p} < \varepsilon$. Q.E.D.

(b) If $p > 0$ then $\sqrt[n]{p} \to 1$.

Proof. $p = 1$ is obvious.
Suppose $p > 1$. Let $x_n = \sqrt[n]{p} - 1 > 0$. Want to show $x_n \to 0$.
Since $(x_n+1)^n$, we have $p = (x_n+1)^n = \sum_{k=0}^{n} \binom{n}{k} x_n^k > \binom{n}{1} x_n' = nx_n$.
Therefore, $x_n \leq \frac{p}{n}$, so $x_n \to 0$ by the Squeeze Lemma.
Suppose $p \in (0,1)$. Let $q = \frac{1}{p} > 1$. Then $\sqrt[n]{q} \to 1$ by the previous case. By 3.3, $\sqrt[n]{p} = \frac{1}{\sqrt[n]{q}} \to 1$. Q.E.D.

(c) $\sqrt[n]{n} \to 1$

Proof. Let $x_n = \sqrt[n]{n} - 1 > 0$, for $n \geq 2$. $n = (x_n + 1)^n = \sum_{k=0}^{n} \binom{n}{k} (x_n)^k > \binom{n}{2} {x_n}^2 = \frac{n(n-1)}{2} x_n^2$. Therefore, $x_n \leq \sqrt{\frac{2}{n-1}}$.
Q.E.D.

(d) If $p > 0$ and $\alpha \in \mathbb{R}$, then $\frac{n^\alpha}{(1+p)^n} \to 0$; i.e., Exponentials beat powers.

Proof. We want an upper bound on $\frac{n^\alpha}{(1+p)^n}$, so seek a lower bound on $(1+p)^n$.
$(1+p)^n = \sum_{k=0}^{n} \binom{n}{k} p^k > \binom{n}{k} p^k$ for $k \leq n$
$(1+p)^n > \binom{n}{k} p^k = \frac{n(n-1)\cdots(n-k+1)}{k!} p^k$. Then for $k \leq \frac{n}{2}$, $\binom{n}{k} p^k > (\frac{n}{2})^k \frac{p^k}{k!}$. Therefore, $\frac{n^\alpha}{(1+p)^n} < (\frac{2}{p})^k k! \frac{1}{n^{k-\alpha}}$. Let $k_0 \in \mathbb{Z}$ s.t. $k > \alpha$. Then for $n \geq 2k_0$, RHS $\to 0$ by (a).

If $|x| < 1$ then $x^n \to 0$.

Proof. $|x^n - 0| = |x|^n$, so $x^n \to 0 \Leftrightarrow |x|^n \to 0$ and $|x|^n = \frac{n_0}{(\frac{1}{|x|})^n} \to 0$ by (d) with $\alpha = 0$ and $1 + p = \frac{1}{|x|} > 1$, so $p = \frac{1}{|x|} - 1 > 0$. Q.E.D.

Q.E.D.

Theorem 3.2. (a) $p_n \to p \Leftrightarrow \forall_{r>0} : N_r(p)$ contains all but finitely many p_n.

Proof. $\forall_{n \geq N} : p_n \in N_r(p)$ Q.E.D.

(b) If $p_n \to p$ and $p_n \to p'$ then $p = p'$.

Proof. $d(p, p') \le d(p_n, p) + d(p_n, p')$ for all n. Fix ε. Choose N such that $d(p_n, p) < \frac{\varepsilon}{2}$ and $d(p_n, p') < \frac{\varepsilon}{2}$ for $n \ge N'$. Then $d(p, p') < \varepsilon$. Then for $n \ge \max\{N, N'\}$, $d(p, p') < \varepsilon$. This is true for all $\varepsilon > 0$, so $d(p, p') = 0$. Q.E.D.

(c) If $\{p_n\}$ converges, then p_n is bounded, in a sense that $\exists_{M>0, q \in X}$ s.t. $d(p_n, q) \le M$ for all n.

Proof. If $p_n \to p$, then $\exists N$ s.t. $d(p_n, p) < 1$ for all $n \ge N$. Thus, $\forall_{n \ge 1} : d(p_n, p) \le \max\{1, d(p, p_1), \dots, d(p, p_{N-1})\} = M$ Q.E.D.

(d) If $E \subset X$ has a limit point p, then $\exists_{p_n \in E}$ s.t. $p_n \to p$.

Proof. We need to choose $p_n \in E$ s.t. $d(p, p_n) < \frac{1}{n}$. Let $\varepsilon > 0$. Then $d(p, p_n) < \varepsilon$ if $n > \frac{1}{\varepsilon}$ Q.E.D.

Definition 3.2. Given $p_n, n_1 < n_2 < n_3 < \dots$, we say $p_{n_i} = (p_{n_1}, p_{n_2}, \dots)$ is a subsequence of p_n.

Lemma. $p_n \to p \Leftrightarrow$ every subsequence of $\{p_n\}$ converges to p

Proof. Look at assignment 6 Q.E.D.

Theorem 3.6. (a) $\{p_n\}$ in X, X compact, then $\exists$ convergent subsequence.

Proof. Let $E =$ range of$\{p_n\}$. If E is finite, then $\exists p \in X$ and $n_1 < n_2 < \dots$ s.t. $p_n = p$ for $\forall i$. This subsequence converges to p. If E is infinite then by Theorem 2.37, E has a limit point $p \in X$; i.e., every neighborhood of p contains infinitely many points of E. Choose n_1 s.t. $d(p, p_{n_1}) < 1$.

Q.E.D.

(b) $\{p_n\}$ in $\mathbb{R}^k$, bounded, then $\exists$ convergent subsequence.

Proof. Choose a k-cell I that contains $\{p_n\}$. I is compact. Apply (a).

Q.E.D.

Definition 3.3 (Cauchy Sequence). $\{p_n\}$ is a Cauchy sequence in (X, d) if $\forall \varepsilon$:

$\exists_{N \in \mathbb{N}}$ s.t. $d(p_m, p_n) < \varepsilon \forall m, n \geq N$.

Definition 3.4. For $E \subset X$, $E \neq \emptyset$, we define diam $E = \sup\{d(p,q) : p, q \in E\}$. diam $E = \infty$ if the set is not bounded above.

Example. *For a sequence p_n in X, let $E_n = \{p_N, p_{N+1}, \ldots\}$. Then $\{p_n\}$ is a Cauchy sequence iff $\lim_{N\to\infty}$ diam $E_N = 0$.*

Theorem 3.11. (a) If $p_n \to p$ then $\{p_n\}$ is a Cauchy sequence.

(b) If X is a compact metric space and $\{p_n\}$ in X is a Cauchy sequence, then $\exists_{p \in X}$ s.t. $p_n \to p$.

(c) In R^K every Cauchy sequence converges.

Remark. If a Cauchy sequence has a convergent subsequence in a metric space, then the full sequence itself converges to the same point the subsequence converges to.

Proof. Let $\varepsilon > 0$. Choose N s.t., $d(p_n, p) < \varepsilon/2$ if $n \geq N$. Then for $m, n \geq N$, $d(p_m, p_n) \leq d(p_n, p) + d(p, p_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Suppose $\{p_n\}$ is Cauchy. Let $E_N = \{p_N, P_{N+1}, \ldots\}$. Then $\overline{E_N}$ is closed, hence compact. Also $\overline{E_N} \supset \overline{E_{N+1}}$ and $\lim_{N\to\infty}$ diam $\overline{E_N} = 0$ (use Theorem 3.10(a) to see diam $\overline{E_N}$ = diam E_N) By theorem 3.10(b), $\exists! p \in \bigcap_{N=1}^{\infty} \overline{E_N}$. Claim: $p_n \to p$.
Proof of the claim: Let $\varepsilon > 0$. Choose N_0 s.t.diam $\overline{E_{N_0}} < \varepsilon$, so $d(p,q) < \varepsilon \forall g \in \overline{E_{N_0}}$, and hence $\forall g \in N_0$; i.e., $d(p, p_n) < \varepsilon$ if $n \geq N_0$.

Let $\varepsilon > 0$. Choose N s.t., $d(p_n, p) < \varepsilon/2$ if $n \geq N$. Then for $m, n \geq N$, $d(p_m, p_n) \leq d(p_n, p) + d(p, p_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Suppose $\{p_n\}$ in $\mathbb{R}^k$ is Cauchy. Cauchy sequences are bounded in any metric space. Therefore, $\exists$ k-cell I, which is compact, containing $\{p_n\}$. Then (b) applies Q.E.D.

Note. The converse of Theorem 3.11(a) does not hold in general.
Example. *$X = \mathbb{Q}$ has a Cauchy sequence with no limit in $\mathbb{Q}$. (see assignment 6). Converse does hold if X is compact.*

Theorem 3.12. (a) diam $\overline{E}$ = diam E

(b) If $K_n \subset X$, $K_n \neq \emptyset$, K compact, $K_n \supset K_{n+1} \forall n$ and if $\lim_{n\to\infty}$ diam $K_n = 0$, then $\bigcap_{n=1}^{\infty} K_n$ is a single point.

Proof. (a) $E \subset \overline{E} \Rightarrow \text{diam } E \leq \text{diam } \overline{E}$. For the opposite inequality, let $\varepsilon > 0, p, q \in \overline{E},$. Choose $p', q' \in E$ s.t. $d(p,p') < \varepsilon, d(q,q') < \varepsilon$. Then $d(p,q) \leq d(p,p') + d(p',q') + d(q',q) < \varepsilon + \varepsilon = 2\varepsilon$. diam $\overline{E} \leq$ diam $E + 2\varepsilon$. Since ε is arbitrary, diam $\overline{E} \leq$ diam E.

(b) Let $K = \bigcap_{n=1}^{\infty} K_n$. By Theorem 2.36, $K \neq \emptyset$. Since $K \subset K_n \forall n$, diam $k \leq$ diam $K_n \forall n$, so diam $K = 0$. Therefore, $d(p,q) = 0 \forall p, q \in K$, so K is a simple point.

Q.E.D.

Definition 3.5 (Complete Metric Space)**.** A metric space (X,d) is complete if every Cauchy sequence in X converges to a point in X.

Example. *(a) X compact $\Rightarrow X$ complete.*

(b) $\mathbb{R}^k$ is complete, so is $\mathbb{C}$.

(c) $\mathbb{Q}$ is not complete. (see assignment 6)

(d) See assignment 6 for reference to completion of a metric space.

We've seen that convergent sequences are bounded. $p_n = (-1)^n$ shows the converse if false. However the converse does hold for monotonic sequences.

Definition 3.6 (Monotone)**.**
- A sequence $\{s_n\}$ in $\mathbb{R}$ is monotonically increasing if $s_n \leq s_{n+1} \forall n$.
- A sequence $\{s_n\}$ in $\mathbb{R}$ is monotonically decreasing if $s_n \geq s_{n+1} \forall n$.

Theorem 3.14. A monotone sequence in $\mathbb{R}$ converges if and only if it is bounded.

Proof. $\Rightarrow$ all convergent sequences are bounded in any metric space.

$\Leftarrow$ **Increasing case** Let $\{s_n\}$ be monotonically increasing and $s_n \leq M \forall n$. Let $s = \sup\{s_n : n \in \mathbb{N}\}$. Then $s_n \leq s \forall n$. Let $\varepsilon > 0$. $\exists N$ s.t. $s - \varepsilon < s_N \leq s$. But then $s - \varepsilon < s_N \leq s_{N+1} \leq s_{N+2} \ldots \leq s$, so $|s - s_n| < \varepsilon \forall n \geq N$, and therefore $s_n \to s$.

Q.E.D.

Definition 3.7 (Infinite Limits)**.** We say

- $s_n \to \infty$ if $\forall_{M \in \mathbb{R}} : \exists_N$ s.t. $s_n \geq M \forall_{n \in N}$.
- $s_n \to -\infty$ if $\forall_{M \in \mathbb{R}} : \exists_N$ s.t. $s_n \leq M \forall_{n \in N}$.

Definition 3.8. Let $\{s_n\}$ be a sequence in $\mathbb{R}$. We define $\limsup_{n\to\infty} s_n = \overline{\lim\limits_{n\to\infty}} s_n = \inf_{n\geq 1}\{\sup_{m\geq n}\{s_m\}\} = \lim\limits_{n\to\infty} \sup_{m\geq n}\{s_m\}$.
$\liminf_{n\to\infty} s_n = \underline{\lim\limits_{n\to\infty}} s_n = \sup_{n\geq 1}\{\inf_{m\geq n}\{s_m\}\} = \lim\limits_{n\to\infty} \inf_{m\geq n}\{s_m\}$.

> **Note.** Alternate definition; see ass 7 for equivalence

> **Remark.** (a) If $a_n \leq b_n \forall n$ and $a_n \to a$ and $b_n \to b$, then $a \leq b$.
>
> (b) $\liminf_{n\to\infty} s_n \leq \limsup_{n\to\infty} s_n$

Example. *(a)* $s_n = (-1)^n(1+\frac{1}{n^2})$ $1 \leq \sup_{m\geq n} s_m \leq 1+\frac{1}{n^2}$, *so* $\limsup_{n\to\infty} s_n = 1$. *Similarly,* $\liminf_{n\to\infty} s_n = -1$

(b) If $\{s_n\}$ *has no upper bound, then* $\sup_{m\geq n} s_m = \infty$ *and in this case we say* $\limsup_{n\to\infty} s_n = \infty$*; e.g.,*

$$s_n = \begin{cases} n & n \text{ odd} \\ -n & n \text{ even} \end{cases}$$

has $\limsup_{n\to\infty} s_n = \infty$, $\liminf_{n\to\infty} s_n = -\infty$

> **Lemma.** $\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = L \Leftrightarrow s_n \to L$.
>
> **Proof** (L finite).
>
> $\Rightarrow$ This follows from $\inf_{m\geq n} s_m \leq s_n \leq \sup_{m\geq n} s_m$. $\lim\limits_{n\to\infty} \inf_{m\geq n} s_m = \liminf_{n\to\infty} s_n$, and $\lim\limits_{n\to\infty} \sup_{m\geq n} s_m = \limsup_{n\to\infty} s_n$. Therefore, $\lim\limits_{n\to\infty} s_n = L$.
>
> $\Leftarrow$ If $s_n \to L$, then $\forall_{\varepsilon>0} : \exists_N$ s.t. $s_m \in [L-\varepsilon, L+\varepsilon] \forall m \geq N$. Therefore, $\forall_{n\geq N} : L-\varepsilon \leq \inf_{m\geq N} s_m \leq \inf_{m\geq n} s_m \leq \sup_{m\geq n} s_m \leq \sup_{m\geq N} s_m \leq L+\varepsilon$. Let $n\to\infty$: $L-\varepsilon \leq \liminf_{n\to\infty} s_n \leq \limsup_{n\to\infty} s_n \leq L+\varepsilon$. Since ε is arbitrary, so $L \leq \liminf_{n\to\infty} s_n \leq \limsup_{n\to\infty} s_n \leq L$.
>
> Q.E.D.

3.2 Series

Definition 3.9 (Series). Let $\{a_n\}$ be a sequence in $\mathbb{C}$. Form a new sequence $\{s_n\}$, the sequence of partial sums, by $s_n = a_1 + a_2 + \ldots + a_n = \sum_{k=1}^{n} a_k$. If $s_n \to s$, we say **the series** $\sum_{k=1}^{\infty} a_k$ **converges** and that $\sum_{k=1}^{\infty} a_k = s$. If $\{s_n\}$ diverges then we say $\sum_{k=1}^{\infty} a_k$ diverges.

Theorem 3.15. $\sum_{n\in\mathbb{N}} a_n$ converges if and only if $\forall_{\varepsilon>0} : \exists N$ s.t. $\forall n \geq m \geq N$: $|\sum_{k=m}^{n} a_k| < \varepsilon$.

Proof. $\sum_n a_n$ converges $\Leftrightarrow$ $\{s_n\}$ converges $\Leftrightarrow$ $\{s_n\}$ is a Cauchy sequence ($\because$ $\mathbb{C}$ is compact). Use $s_n - s_{m-1} = \sum_{k=m}^{n} a_k$. Q.E.D.

Corollary 3.16. If $\sum_n a_n$ converges then $a_n \to 0$.

Proof. Take $m = n$ in Theorem 3.22. $\sum_n a_n$ converges $\Rightarrow \forall_{\varepsilon>0}$: $\exists_N$ s.t. $|a_n| < \varepsilon$ if $n \geq N$. Q.E.D.

Remark. n-th term test for divergence: If $a_n \not\to 0$ then $\sum_n a_n$ diverges.
Example. $\sum_{n=1}^{\infty} \frac{n}{n+1}$ *diverges because* $\frac{n}{n+1} \to 1 \neq 0$.

Converse to Corollary 3.16 is false! E.g., $\sum_n \frac{1}{n}$ diverges but $\frac{1}{n} \to 0$.

Theorem 3.24. If $a_n \geq 0$, then $\sum_n a_n$ converges if and only if $\{s_n\}$ is bounded.

Proof. $\{s_n\}$ is monotone increasing, so by Theorem 3.14, it converges if and only if it is bounded. Q.E.D.

Theorem 3.25. [Comparison Test]

(a) If $|a_n| \leq c_n \forall n \geq N_0$ and $\sum_n c_n$ converges, then $\sum_n a_n$ converges.

Proof. Suppose $|a_n| \leq c_n \forall n \geq N_0$ and $\sum_n c_n$ converges. Let $\varepsilon > 0$. By theorem 3.22, $\exists N$ s.t. $\sum_{k=m}^{n} c_k < \varepsilon$ if $n \geq m \geq N$. Can take $N \geq N_0$. Then $|N \geq N_0|$. $|\sum_{k=m}^{n} a_k| \leq \sum_{k=m}^{n} |a_k| \leq \sum_{k=m}^{n} c_k < \varepsilon$ if $n \geq m \geq N$. By theorem 3.22 again, $\sum_n a_n$ converges. Q.E.D.

(b) If $a_n \geq d_n \geq 0 \forall n \geq N_0$ and if $\sum_n d_n$ diverges, then $\sum_n a_n$ diverges.

Proof. This follows from (a): if $\sum_n a_n$ converges then $\sum_n d_n$ converges. Thus it's contrapositive, (b) is true. Q.E.D.

Theorem 3.26. [Geometric Series] $\sum_{n=0}^{\infty} x^n = \begin{cases} \frac{1}{1-x} & \text{if } -1 < x < 1 \\ \text{diverges} & \text{otherwise} \end{cases}$

Proof. Let $S_n = 1 + x + x^2 + \cdots + x^n$, $xS_n = x + x^2 + \cdots x^n + x^{n+1}$. Then

$$S_n - xS_n = 1 - x^{n+1} \Rightarrow S_n = \frac{1 - x^{n+1}}{1 - x}$$

If $|x| < 1 (\Leftrightarrow -1 < x < 1)$, then $x^{n+1} \to 0$ and $S_n \to \frac{1}{1-x}$. If $|x| \geq 1$, then x^{n+1} does not converge to 0, so $\sum_{n=0}^{\infty} x^n$ diverges. Q.E.D.

Theorem 3.27. Suppose $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$. Then $\sum_{n=1}^{\infty} a_n$ converges $\Leftrightarrow \sum_{k=1}^{\infty} 2^k a_{2^k}$ converges.

Proof. ($\Leftarrow$) We show that if $\sum_n a_n$ diverges, then $\sum_k 2^k a_{2^k}$ diverges. For this, note that $a_1 + a_2 + \ldots + a_n \leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \cdots (a_{2^k} + a_{2^k+1} + \cdots + a_{2^{k+1}-1})$ if $2^{k+1} > n$. $a_1 + a_2 \cdots + a_n \leq a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k}$. LHS unbounded as $n \to \infty$, so RHS is also unbounded as $k \to \infty$.

($\Rightarrow$) $a_1 + a_2 + a_3 + \cdots + a_n \geq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \cdots + (a_{2^{k-1}+1} + a_{2^{k-1}+2} + \cdots + a_{2^k})$ if $2^k \leq n$. $a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \cdots + (a_{2^{k-1}+1} + a_{2^{k-1}+2} + \cdots + a_{2^k}) \geq a_1 + a_2 + 2a_4 + 4a_8 + \cdots + 2^{k-1} a_{2^k} \geq \frac{1}{2}(a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k})$. If $\sum_n a_n$ converges, then LHS is bounded for all n so RHS is bounded for all k. Hence RHS converges since it is monotone.

Q.E.D.

Theorem 3.28. [p-series] $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \leq 1$.

Proof. For $p \leq 0$, $\frac{1}{n^p} \not\to 0$, so series diverges. For $p > 0$, $\frac{1}{n^p}$ is decreasing, so $\sum_{n=1} \frac{1}{n^p}$ converges iff $\sum_k 2^k \cdot \frac{1}{(2^k)^p}$ converges. But $\sum_k 2^k \cdot \frac{1}{(2^k)^p} = \sum_k (\frac{1}{2^{p-1}})^k$ converges iff $\frac{1}{2^{p-1}} < 1 (\Leftrightarrow p - 1 > 0)$, which is equivalent to $p > 1$. Q.E.D.

Theorem 3.29. $\sum_{n=3}^{\infty} \frac{1}{n(\log n)^p}$ converges if $p > 1$ and diverges if $p \leq 1$. (log is to base e.)

Proof. If $p \leq 0$, then $\frac{1}{n(\log n)^p} \geq \frac{1}{n}$, so $\sum_n \frac{1}{n(\log n)^p}$ diverges by the comparison test. If $p > 0$ then $\frac{1}{n(\log n)^p}$ decreases since $\log n$ increases. By theorem 3.27, $\sum_n \frac{1}{n(\log n)^p}$ converges $\Leftrightarrow$ $\sum_k 2^k \cdot \frac{1}{2^k(\log 2^k)^p}$ converges $\Leftrightarrow \sum_k \frac{1}{(\log 2)^p} \cdot \frac{1}{k^p}$ converges $\Leftrightarrow p > 1$ Q.E.D.

Definition 3.10 (e). $e := \sum_{n=0}^{\infty} \frac{1}{n!}$.

Remark. Convergence $\frac{1}{n!} = \frac{1}{(n(n-1)(n-2)\cdots 3\cdot 2\cdot 1)} \leq \frac{1}{2\cdot 2\cdot 2\cdot 2\cdot \cdots \cdot 2\cdot 1} = \frac{1}{2^{n-1}}$.

Therefore, $S_n = \sum_{k=0}^{n} \frac{1}{k!} \leq 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} < 1 + \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} = 1 + \frac{1}{1-1/2} = 3$. Then S_n is a monotonically increasing sequence that's also bounded. Hence, $e \leq 3$

Rate of Convergence

$0 < e - S_n = \sum_{k=n+1}^{\infty} \frac{1}{k!} < \sum_{k=n+1}^{\infty} \frac{1}{(n+1)^{k-n+1}} \cdot \frac{1}{(n+1)!}$

$= \frac{1}{(n+1)!} \sum_{k=n+1}^{\infty} \frac{1}{(n+1)^{k-(n+1)}}$

$= \frac{1}{(n+1)!} \cdot \frac{1}{1-1/(n+1)} = \frac{1}{(n!)\cdot n}$.

Theorem 3.32. $e \notin \mathbb{Q}$.

Proof. For contradiction, suppose $e = \frac{p}{q}, p, q \in \mathbb{N}$. As $0 < e - S_q < \frac{1}{q\cdot q!}$, $0 < q! \cdot e - q! \cdot S_q < \frac{1}{q}$. Since $S_q = \sum_{k=0}^{q} \frac{1}{k!}$, $q! \cdot e$ and $S_q \cdot q!$ are both integers. However, then $q! \cdot e - q! \cdot S_q$ is an integer between 0 and $\frac{1}{q} < 1$, which is a contradiction. Q.E.D.

Theorem 3.31. $e = \lim\limits_{n\to\infty} (1 + \frac{1}{n})^n$.

Proof. Let $t_n = (1+\frac{1}{n})^n$. Then $t_n = \sum_{k=0}^{n} \binom{n}{k} \cdot \frac{1}{n^k} = \sum_{k=0}^{n} \frac{1}{k!} \cdot (\frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}) \leq S_n$. So $\limsup_{n\to\infty} t_n \leq \limsup_{n\to\infty} S_n = \lim\limits_{n\to\infty} S_n = e$. On the other hand, for fixed m and $n \geq m$, $t_n \geq \sum_{k=0}^{m} \binom{n}{k} \cdot \frac{1}{n^k} = \sum_{k=0}^{m} \frac{1}{k!}(1 - \frac{1}{n})(1 - \frac{2}{n}) \cdot \cdots \cdot (1 - \frac{k-1}{n})$. Let $n \to \infty$ with m fixed. $\liminf_{n\to\infty} t_n \geq \sum_{k=0}^{m} \frac{1}{k!} \cdot 1 = S_m$. This is true for any m. Now let $m \to \infty$. $\liminf_{n\to\infty} t_n \geq \limsup_{m\to\infty} s_m = e$. $e \leq \liminf_{n\to\infty} t_n \leq \limsup_{n\to\infty} t_n \leq e$. Therefore, $\lim\limits_{n\to\infty} t_n$ exists and equals e. Q.E.D.

Theorem 3.33. [Root test] Let $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. Then,

$$\sum a_n \begin{cases} \text{converges} & \text{if } \alpha < 1 \\ \text{diverges} & \text{if } \alpha > 1 \\ \text{inconclusive} & \text{if } \alpha = 1 \end{cases}$$

Proof (Just outline). $\alpha < \beta < 1$ Eventually $|a_n| \leq \beta^n$, thus convergence.

$\alpha > 1$ $|a_n| > 1$ for infinitely many n, thus divergence.

$\alpha = 1$ $\frac{1}{n}$ diverges, $\frac{1}{n^2}$ converges.

Q.E.D.

Theorem 3.34. [Ratio test] The series $\sum_{n=1}^{\infty} a_n$ with $a_n \neq 0$ converges if $\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$ and diverges if $\exists_{N\in\mathbb{N}}$ s.t. $\forall_{n\geq N} : |\frac{a_{n+1}}{a_n}| \geq 1$. Otherwise, inconclusive.

Proof. (see textbook).

$$\textbf{Convergence} \ \sum a_n \begin{cases} \text{converges} & \text{if } \lim\limits_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1 \\ \text{diverges} & \text{if } \lim\limits_{n\to\infty} |\frac{a_{n+1}}{a_n}| > 1 \\ \text{diverges} & \text{if } \liminf_{n\to\infty} |\frac{a_{n+1}}{a_n}| > 1 \\ \text{inconclusive} & \text{otherwise. e.g., } \lim\limits_{n\to\infty} |\frac{a_{n+1}}{a_n}| = 1 \end{cases}$$

Note. Note that we cannot replace $\liminf$ with $\limsup$ in the third case. For inconclusive case, check $\sum 1/n \to \infty$ and $\sum 1/n^2 \to \pi^2/6$

Q.E.D.

Example. *Let* $s_n = \sum_{n=0}^{\infty} a_n = \frac{1}{2} + 1 + \frac{1}{8} + \frac{1}{4} + \frac{1}{32} + \frac{1}{16} + \frac{1}{128} + \frac{1}{64} + \cdots$. *First, note that* $a_{2k} = \frac{1}{2} \cdot \frac{1}{4^k}, a_{2k+1} = 2a_{2k} = \frac{1}{4^k}$ *for* $k \geq 0$.

Ratio test *Then the ratio* $\frac{a_{n+1}}{a_n}$ *is the sequence* $2, \frac{1}{8}, 2, \frac{1}{8}, 2, \frac{1}{8}, \ldots$. *Therefore,* $\liminf_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{8}, \limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = 2$. *The ratio test is inconclusive for* s_n.

Root test $a_n = \begin{cases} \frac{1}{2} \cdot \frac{1}{2^n} & n \text{ even} \\ \frac{2}{2^n} & n \text{ odd} \end{cases}$, *so* $(\frac{1}{2})^{1/n} \cdot \frac{1}{2} \leq \sqrt[n]{a_n} \leq \frac{2^{1/n}}{2}$, *thus* $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = \frac{1}{2} < 1$. *Therefore,* $s_n = \sum_{n=0}^{\infty} a_n$ *converges.*

This is an example where the ratio test is inconclusive but the root test is conclusive.

Theorem 3.47. If $\sum a_n = A$ and $\sum b_n = B$, then $\sum a_n + b_n = A + B$ and $\sum c \cdot a_n = cA$.

3.3 Power Series

Definition 3.11 (Power Series)**.** For $z \in \mathbb{C}$ and a complex sequence $\{c_n\}$, $\sum_{n=0}^{\infty} c_n z^n$ is a power series.

Remark. As $z^0 = 1$ for all $z \in \mathbb{C}$, by convention we write $\sum_{n=0}^{\infty} c_n z^n = c_0 + \sum_{n=1}^{\infty} c_n z^n$.

Theorem 3.39. Let $R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}}$, where

$$R = \begin{cases} 0 & \text{if } \limsup_{n\to\infty} \sqrt[n]{|c_n|} = \infty \\ \infty & \text{if } \limsup_{n\to\infty} \sqrt[n]{|c_n|} = 0 \end{cases}.$$

Then $\sum c_n z^n \begin{cases} \text{converges} & \text{if } |z| < R \\ \text{diverges} & \text{if } |z| > R \\ \text{inconclusive} & \text{if } |z| = R \end{cases}$. Note $R = 0$ implies the series diverges for $z \neq 0$, and $R = \infty$ implies the series converges for any $z \in \mathbb{C}$.

Proof. $\limsup_{n\to\infty} \sqrt[n]{|c_n z^n|} = |z| \cdot \limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{|z|}{R}$. By root test, the series converges if $\frac{|z|}{R} < 1$ and diverges if $\frac{|z|}{R} > 1$.

Note. In practice, often use the ratio test to find R.

Q.E.D.

Example. *(a) $\sum n! \cdot z^n$ has $R = 0$.*

By ratio test $\forall z \neq 0$: $\left|\frac{(n+1)! \cdot z^{n+1}}{n! \cdot z^n}\right| = |z| \cdot (n+1) \to \infty$. *Hence, the series diverges .*

By root test *Note $n \neq \frac{1}{2}(\frac{2}{3})^2(\frac{3}{4})^3 \cdots (\frac{n-1}{n})^{n-1} n^n$ for $n \geq 2$. Then $n \neq \frac{n^n}{(1+1)^1(1+\frac{1}{2})^2(1+\frac{1}{n-1})^{n-1}}$. In the proof of Theorem 3.31, we saw $(1+\frac{1}{j}) \leq e$. So $n! \geq \frac{n^n}{e^{n-1}} = e \cdot (\frac{n}{e})^n$. $\sqrt[n]{n!} \geq e^{1/n} \cdot \frac{n}{e} \to \infty$ as $n \to \infty$. Therefore, $R = \frac{1}{\infty} = 0$.*

> **Note.** Cf. Stirling's formula: $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

Definition 3.12 (Absolute Convergence, Conditional Convergence)**.**

(a) $\sum a_n$ converges absolutely if $\sum |a_n|$ converges.

(b) $\sum a_n$ converges conditionally if $\sum a_n$ converges but $\sum |a_n|$ diverges.

> **Remark.** All other convergence tests seen so far are actually tests for absolute convergence.

Example.

- *$\sum_{n=1}^{\infty} (-1)^{n-1}\frac{1}{n}$ converges conditionally by the alternating series test of assignment 7's practice problem 1. See also Theorem 3.43. (MUST TAKE A LOOK!)*
- *Given $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$ and $a_n \to 0$, then $\sum (-1)^n a_n$ converges.*
- *$\sum_{n=0}^{\infty} n!2^n$ has $R = 0$*
- *$\sum_{n=0}^{\infty} \frac{z^n}{n^n}$ has $R = \infty$ since $R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}} = \frac{1}{\limsup_{n\to\infty} \frac{1}{n}} = 1/0 = \infty$, or use ratio test, $|\frac{z^{n+1}/(n+1)^{n+1}}{z^n/n^n}| = |z| = \frac{n^n}{(n+1)^{n+1}} = |z|\frac{1}{n+1}\frac{1}{(1+\frac{1}{n})^n}$. Since $(1+\frac{1}{n})^n \to \frac{1}{e}$ as $n \to \infty$, $|z|\frac{1}{n+1}\frac{1}{(1+\frac{1}{n})^n} \to 0$ as $n \to \infty \forall_{z\in\mathbb{C}}$ so $R = \infty$.*

> **Theorem 3.45.** If $\sum a_n$ converges absolutely, then $\sum a_n$ converges.

> **Theorem 3.54.** Suppose $\sum a_n$ converges conditionally. Let $-\infty \leq \alpha \leq \beta \leq +\infty$. Then $\exists$bijection $f : \mathbb{N} \to \mathbb{N}$ such that with $a'_n = a_{f(n)}$ and $S'_n = \sum_{k=1}^{n} a'_k$, $\liminf_{n\to\infty} s'_n = \alpha$ and $\limsup_{n\to\infty} s'_n = \beta$. In other words, there exists a rearrangement of $\sum a_n$, say $\sum a'_n$, such that $\liminf_{n\to\infty} \sum a'_n = \alpha, \limsup_{n\to\infty} \sum a'_n = \beta$.
>
> **Proof.** Take a look at the textbook Q.E.D.

> **Theorem 3.55.** If $\sum a_n$ converges absolutely, then every rearrangement of $\sum a_n$ converges to the same sum.
>
> **Proof.** Take a look at the textbook Q.E.D.

3.4 Products of Series

Motivation Consider z^N in $(\sum_{n=0}^{\infty} a_n z^n)(\sum_{n=0}^{\infty} b_n z^n)$. Since $(\sum_{n=0}^{\infty} a_n z^n)(\sum_{n=0}^{\infty} b_n z^n) = (a_0 + a_1 z + a_2 z^2 + \cdots)(b_0 + b_1 z + b_2 z^2 + \cdots) = (a_0 b_0) + (a_0 b_1 + a_1 b_0) z + (a_0 b_2 + a_1 b_1 + a_2 b_0) z^2 + \cdots$, z^N has coefficient $\sum_{k=0}^{N} a_k b_{N-k}$.

Definition 3.13. The product of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ is $\sum_{n=0}^{\infty} c_n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.

> **Note.** This is a discrete convolution.

Question If $\sum a_n = A$ and $\sum b_n = B$ both converge, does $\sum c_n$ converge and if so, does it converge to AB?

Answer $\sum c_n$ converges if $\sum a_n$ and $\sum b_n$ converge absolutely. (Theorem 3.50). Moreover, if $\sum c_n$ does converge, then it must converge to AB (Theorem 3.51). Maybe no otherwise (ref: Example 3.49).

Theorem 3.50. Suppose $\sum a_n$ converges absolutely to A and $\sum b_n$ converges to B. Then $\sum c_n$ converges to AB.

Proof. Let $A_n = \sum_{k=0}^{n} a_k$, $B_n = \sum_{k=0}^{n} b_k$, $C_n = \sum_{k=0}^{n} c_k$. Then $A_n \to A, B_n \to B$. By definition, $C_n = \sum_{k=0}^{n} \sum_{j=0}^{k} a_j b_{k-j} = \sum_{j=0}^{n} \sum_{k=j}^{n} a_j b_{k-j} = \sum_{j=0}^{n} a_j \sum_{k=j}^{n} b_{k-j} = \sum_{j=0}^{n} a_j B_{n-j} = \sum_{j=0}^{n} a_j B + \sum_{j=0}^{n} a_j (B_{n-j} - B)$. Let β_{n-j}, where $\beta_k = B_k - B$. Then $C_n = A_n B + \sum_{j=0}^{n} a_j \beta_{n-j}$. Let $\gamma_n = \sum_{j=0}^{n} a_j \beta_{n-j}$. Note that $A_n B \to AB, \beta_k \to 0$ as $n \to \infty$. Let $\alpha = \sum_{k=0}^{\infty} |a_k| < \infty$ ($\because a_n$ converges absolutely by assumption). Rewrite γ_n as $\gamma_n = \sum_{j=0}^{n} a_{n-j} \beta_j$. We know $\beta_j \to 0$ as $j \to \infty$. Let $\varepsilon > 0$. Choose N s.t. $|\beta_j| < \varepsilon$ if $j \geq N$. Then for $n \geq N+1$, $|\gamma_n| \leq |\sum_{j=0}^{N} a_{n-j} \beta_j| + |\sum_{j=N+1}^{n} a_{n-j} \beta_j|$. Note $|\sum_{j=N+1}^{n} a_{n-j} \beta_j| \leq \varepsilon \sum_{j=N+1}^{n} |a_{n-j}| \leq \varepsilon \alpha$. Let $n \to \infty$ with N fixed. Then $a_{n-j} \to 0$ for $0 \leq j \leq N$ since $|a_n| \to 0$. Q.E.D.

Theorem 3.51. If the series $\Sigma a_n, \Sigma b_n, \Sigma c_n$ converge to A, B, C respectively and $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, then $C = AB$.

Midterm 2 covers up to this point (TB p.36-82, A.5-8)

Chapter 4

Continuity

Assume general metric spaces X, Y and $f : X \to Y$.

Definition 4.1. Suppose X, Y are metric spaces, $E \subset X$, $f : E \to Y$, $p \in E'$, where E': set of limit points in metric space X. We say $\lim_{x\to p} f(x) = q$, or $f(x) \to q$ as $x \to p$, if $\forall_{\varepsilon>0} : \exists_{\delta>0}$ s.t. $(0 < d_X(x,p) < \delta$ and $x \in E) \Rightarrow d_Y(f(x), q) < \varepsilon$.

> **Note.** We don't say anything about $x = p$, $f(p)$ may not even be defined.

> **Theorem 4.2.** $\lim_{x\to p} f(x) = q \Leftrightarrow \forall_{\{p_n\} \text{ in } E}$: if $p_n = p$ or $p_n \to p$, then
>
> $$\lim_{n\to\infty} f(p_n) = q,$$
>
> where the RHS is the limit of Definition 3.1.
>
> > **Note.** This implies uniqueness of q in Definition 4.1.

Proof. $\Rightarrow$ Suppose $\lim_{x\to p} f(x) = q$. Let $\varepsilon > 0$. Choose $\delta > 0$ s.t. $d_Y(f(x), q)\varepsilon$ if $0 < d_X(x, p) < \delta$. Let $\{p_n\}$ be a sequence in E such that $p_n \to p$ and $p_n \neq p$. Then $\exists_N$ s.t. $0 < d_X(p_n, p) < \delta$ if $n \geq N$; i.e., $f(p_n) \to q$.

$\Leftarrow$ Consider the contrapositive of ($\Leftarrow$): $\neg(\lim_{x\to p} f(x) = q) \Rightarrow \neg(\forall_{\{p_n\} \text{ in } E} : \lim_{n\to\infty} f(p_n) = q)$. Suppose $\neg(\lim_{x\to p} f(x) = q)$. Then $\exists_{\varepsilon>0}$ s.t. $\forall_{\delta>0}$: $\exists_{x\in N_\delta^E(p)}$ s.t. $x \neq p$ and $d_Y(f(x), q) \geq \varepsilon$. Take $\delta = \delta_n = \frac{1}{n}$ and let p_n be an x as above for δ_n. Then $p_n \to p$, but $d_Y(f(p_n), q) \geq \varepsilon \forall n$, so $f(p_n) \not\to q$.

Q.E.D.

Theorem 4.4. When $Y = \mathbb{C}$, limit as defined in Definition 4.1 respects sums, products and quotients.

Proof. By Theorem 4.2, it suffices to show that the theorem holds for sequences. Q.E.D.

Definition 4.2. Suppose X, Y are metric spaces, $p \in E \subset X$, $f : E \to Y$. Then f is continuous at p if $\forall_{\varepsilon>0} : \exists_{\delta>0}$ s.t. $d_X(x, p) < \delta \Rightarrow d_Y(f(x), f(p)) < \varepsilon$; i.e., $f(N_\delta^E(p)) \subset N_\varepsilon^y(f(p))$. We say f is continuous if f is continuous at p for all $p \in E$.

Note. If p is an isolated point; i.e., $\exists_{\delta>0}$ s.t. $N_\delta^E(p) = \{p\}$, then every $f : E \to Y$ is continuous at p.

Theorem 4.6. Suppose $E \subset X, p \in E \cap E', f : E \to Y$. Then f is continuous at p if and only if $\lim_{x\to p} f(x) = f(p)$.

Proof. By Definition 4.1 and Definition 4.2 with $q = f(p)$. Q.E.D.

Theorem 4.7. For $E \subset X, f : E \to Y, g : f(E) \to Z$, let $h = g \circ f : E \to Z$. If f is continuous at $p \in E$ and if g is continuous at $f(p) \in Y$, then h is continuous at p.

Proof. Choose $\eta > 0$ such that $d_Y(f(p), y) < \eta \Rightarrow d_Z(g(f(p)), g(y)) < \varepsilon$ (continuity of g at $f(p)$). Choose $\delta > 0$ s.t. $d_E(x, p) < \delta \Rightarrow d_Y(f(x), f(p)) < \eta$ (continuity of f at p). Then $d_E(x, p) < \delta \Rightarrow d_Z(g(f(x)), g(f(p))) = d_Z(h(x), h(p)) < \varepsilon$. Q.E.D.

Theorem 4.8. [Topological Characterization of Continuity] $f : X \to Y$ is continuous $\Leftrightarrow$ $f^{-1}(V)$ is open for every open $V \subset Y$.

Proof. ($\Rightarrow$) Suppose f is continuous. Let $V \subset Y$ be open. Then $f^{-1}(V)$ is open. Let $p \in f^{-1}(V)$. We need to show $\exists_{\delta>0}$ s.t. $N_\delta^X(p) \subset f^{-1}(V)$. Since V is open, $\exists_{\varepsilon>0}$ s.t. $N_\varepsilon^Y(f(p)) \subset V$. Since f is continuous, $\exists_{\delta>0}$ s.t. $f(N_\delta^X(p)) \subset N_\varepsilon^Y(f(p)) \subset V$.

($\Leftarrow$) Suppose $f^{-1}(V)$ is open for every open $V \subset Y$. Let $p \in X$ and $\varepsilon > 0$. Then $N_\varepsilon^Y(f(p))$ is open, so $f^{-1}(N_\varepsilon^Y(f(p)))$ is open. Take $V = N_\varepsilon^Y(f(p))$, which is open. Since $f^{-1}(V)$ is open and $p \in f^{-1}(V)$, there exists $\delta > 0$ such that $N_\delta^X(p) \subset f^{-1}(V)$. Then $f(N_\delta^X(p)) \subset V = N_\varepsilon^Y(f(p))$; i.e., f is continuous at p.

Q.E.D.

Remark.

(a)

(b) Continuity is determined by the open sets, not the metric. For instance, if metrics l_1, l_2, l_∞ have the same open sets in $\mathbb{R}^k$, hence the same continuous functions.

$$l_1(x,y) = \sum_{i=1}^{k} |x_i - y_i|$$

$$l_2(x,y) = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$

$$l_\infty(x,y) = \max_{1\le i\le k} |x_i - y_i|$$

(c) f with open $U \subset X \Rightarrow f(U)$ is open are called open maps. Continuous maps need not be open(e.g., $f(x) =$ some constant, $f(x) = x^2$), and open maps need not be continuous(e.g., floor function: $\lfloor\cdot\rfloor : \mathbb{R} \to \mathbb{Z}$).

Corollary 4.9. $f : X \to Y$ is continuous if and only if $f^{-1}(F)$ is closed for every closed $F \subset Y$.

Proof. Let $V \subset Y$ be open and $F = V^c$. Then the above condition (RHS) is the same as $f^{-1}(V) = f^{-1}(F^c) = (f^{-1}(F))^c$ is open. Q.E.D.

Theorem 4.9. Let $f : X \to \mathbb{C}, g : X \to \mathbb{C}$ be continuous. Then $f + g, f \cdot g, f/g$(at points p where $g(p) \neq 0$) are also continuous.

Theorem 4.10. Given $f_i : X \to \mathbb{R}(i = 1, 2, \ldots, k)$, define $f : X \to \mathbb{R}^k$ by $f(x) = (f_1(x), \ldots, f_k(x))$. Then

(a) f is continuous if and only if each f_i is continuous.

(b) if $f, g : X \to \mathbb{R}^k$ are continuous, then so are $f + g : X \to \mathbb{R}^k, f \cdot g : X \to \mathbb{R}^1$

Example. *(a) For $i = 1, \ldots, k$, define $\varphi_i : \mathbb{R}^k \to \mathbb{R}$ by $\varphi_i(x) = x_i$, where $x = (x_1, x_2, \ldots, x_k)$. Then $|\varphi_i(x) - \varphi_i(y)| = |x_i - y_i| \leq \left(\sum_{j=1}^{k} |x_j - y_j|^2\right)^{1/2} = |x - y|$, so φ_i is continuous(take $\delta = \varepsilon$. If $|x - y| < \delta = \varepsilon$, then $|\varphi_i(x) - \varphi_i(y)| < \varepsilon$).*

(b) The functions $\mathbb{R}^k \to \mathbb{R}$ defined by $x \mapsto x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k} (n_i \in \{0, 1, 2, \ldots\})$ is continuous on $\mathbb{R}^k$ and so is any polynomial $P(x) = \sum C_{n_1, n_2, n_3, \ldots, n_k} x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k}$, where $C_{n_1, n_2, n_3, \ldots, n_k}$ is a constant (function) in $\mathbb{C}$.

(c) Rational functions $f(x) = \frac{P(x)}{Q(x)}$ are continuous at points where $Q(x) \neq 0$.

(d) The function $\mathbb{R}^k \to \mathbb{R}$ defined by $x \mapsto |x|$ is continuous.

Proof. $|x| = |y + (x - y)| \leq |y| + |x - y|$, so $|x| - |y| \leq |x - y|$. Similarly, $|y| - |x| \leq |y - x|$, so $||x| - |y|| \leq |x - y|$. Thus by taking $\delta = \varepsilon$, $|x - y| < \delta \Rightarrow ||x| - |y|| < \varepsilon$. Q.E.D.

(e) Suppose $f : X \to \mathbb{R}^k$ is continuous. Then $p \mapsto |f(p)|$ is continuous.

Proof. $p \mapsto |f(p)| = (y \mapsto |y|) \circ (p \mapsto f(p))$. Since both $(y \mapsto |y|)$, $(p \mapsto f(p))$ are continuous, $p \mapsto |f(p)|$ is continuous by Theorem 4.7. Q.E.D.

Note. A function is said to be continuous on the *domain*, not on the *range*.

Theorem 4.14. Let $f : X \to Y$ be continuous and X be compact. Then $f(X)$ is compact.

Proof. Let $\{V_\alpha\}$ be an open cover of $f(X)$. We need to find a finite subcover of $f(X)$. By Theorem 4.8, each set $O_\alpha = f^{-1}(V_\alpha)$ is open and $\bigcup_\alpha O_\alpha = \bigcup_\alpha f^{-1}(V_\alpha) = f^{-1}(\bigcup_\alpha V_\alpha) = f^{-1}(f(X)) = X$. Hence, $\{O_\alpha\}$ is an open cover of X, so there exists a finite subcover $X = O_{\alpha_1} \cup \cdots \cup O_{\alpha_n}$. However, then $f(x) = \bigcup_{i=1}^n f(O_{\alpha_i}) = \bigcup_{i=1}^n f(f^{-1}(V_{\alpha_i})) \subset \bigcup_{i=1}^n V_{\alpha_i}$. Therefore, $\{V_{\alpha_i}\}_{i=1}^n$ is a finite subcover of $f(X)$. Q.E.D.

Definition 4.3 (4.13). $f : E \to \mathbb{R}^k$ is bounded if $\exists_{M>0}$ s.t. $|f(x)| \leq M\, \forall x \in E$.

Theorem 4.15. If X is compact and $f : X \to \mathbb{R}^k$, then $f(X)$ is closed and bounded (so f is bounded).

Proof. $f(X)$ is compact by Theorem 4.14, and since $f(X) \subset \mathbb{R}^k$, it is closed and bounded. Q.E.D.

Theorem 4.16. If X is compact and $f : X \to \mathbb{R}^1$ is continuous, then $\exists_{p,q\in X}$ s.t. $f(p) \leq f(x) \leq f(q)$ for all $x \in X$.

Proof. By Theorem 4.15, $f(X)$ is closed and bounded. By Theorem 2.28, $M \in f(X)$ and similarly $m \in f(X)$. Q.E.D.

Example. *Let $X = (0,1)$, not compact, let $f(x) = \frac{1}{x}$, continuous. However, $\nexists_{p\in X}$ s.t. $\forall_{x\in X} : f(p) \leq f(x)$ and $\nexists_{q\in X}$ s.t. $\forall_{x\in X} : f(x) \leq f(q)$.*

Theorem 4.17. Suppose $f : X \to Y$ is one-to-one, onto, continuous, where X is compact. Define $f^{-1} : Y \to X$ by $f^{-1}(f(x)) = x$. Then f^{-1} is continuous.

Proof. By Theorem 4.8, it suffices to prove that if $V \subset X$ is open then $(f^{-1})^{-1}(v)(= f(v))$ is open. However, $V^c \subset X$ is closed, hence V^c is compact by Theorem 4.14 and $(f(V^c))^c = f(V)$ is open. Q.E.D.

Example (Compactness is needed in Theorem 4.17)**.** *Let $X = [0, 2\pi), Y = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Define $f : X \to Y$ by $f(\theta) = (\cos\theta, \sin\theta)$. This f is 1-1, onto, and continuous, but f^{-1} is not continuous as X is not compact.*

Proof.

(1) $[0,1) \subset X$ is open but $(f^{-1})^{-1}([0,1)) = f([0,1))$ is not open because $(1,0)$ is not an interior point of Y.

(2) In Y, as $(x,y) \to (1,0)$ from above, $f((x,y)) \to 0$. As $(x,y) \to (1,0)$ from below, $\lim f^{-1}(x,y)$ does not exist in X. (Wants to be $2\pi \notin X$), so f^{-1} is not continuous at $(1,0) \in Y$.

Q.E.D.

Definition 4.18. Let X, Y be metric spaces and $f : X \to Y$. f is uniformly continuous on X if $\forall_{\varepsilon>0} : \exists_{\delta>0}$ s.t. for all $p, q \in X$ with $d_x(p, q) < \delta$, we have $d_Y(f(p), f(q)) < \varepsilon$.

Remark. The point is for any ε, there is some δ that works for every $p, q \in X$ such that $d(p, q)$.

Example. *(a) $X = (0, 1), Y = \mathbb{R}, f(x) = \frac{1}{x}$. f is continuous on X but is not uniformly continuous.*

Proof. For $x \in (0, \frac{1}{2})$, $|f(x) - f(2x)| = |\frac{1}{x} - \frac{1}{2x}| = \frac{1}{2x} \to \infty$ as $x \to 0$. Then for $\varepsilon = 1$, given any $\delta \in (0, \frac{1}{2})$, we can pick $x < \delta$ s.t. $d_(X)(x, 2x) = x < \delta$, but $d_Y(f(x), f(2x)) = \frac{1}{2x} > \frac{1}{2\delta} > 1$. Q.E.D.

(b) $X = [0, 5], Y = \mathbb{R}, f(x) = x^2$ is uniformly continuous.

Proof. For $0 \leq x_1 \leq x_2 \leq 5$ and $\varepsilon > 0$, $|f(x_1) - f(x_2)| = x_2^2 - x_1^2 = (x_2 - x_1)(x_2 + x_1) \leq 10 \cdot (x_2 - x_1)$, which is less than ε if $|x_2 - x_1| < \frac{\varepsilon}{10} = \delta$ Q.E.D.

Theorem 4.19. Suppose X is a compact metric space, Y is a metric space, and $f : X \to Y$ is continuous. Then f is uniformly continuous.

Proof. Fix $\varepsilon > 0$. For $p \in X$ there exists $\delta = \delta_p(\varepsilon)$ s.t. $d_X(p, q) < \delta_p \Rightarrow d_Y(f(p), f(q)) < \frac{\varepsilon}{2}$. We need to remove the p-dependence of δ_p. Let $J_p = N_{\frac{1}{2}\delta_p}(p)$. Then $\{J_p\}_{p \in X}$ is an open cover of X. Then there exists subcover $X = J_{p_1} \cup J_{p_2} \cdots \cup J_{p_n}$ (equality works as X is the whole metric space, so $X \subset J \Rightarrow X = J$). Let $\delta = \min\{\frac{1}{2}\delta_{p_1}, \frac{1}{2}\delta_{p_2}, \ldots, \frac{1}{2}\delta_{p_n}\}$. Suppose p, q with $d_X(p, q) < \delta$. Choose $m \in \{1, 2, \ldots n\}$ s.t. $p \in J_{p_m}$. Then $d_X(p, p_n) < \frac{1}{2}\delta_{p_m}$. $d_X(q, p_m) \leq d_X(q, p) + d_X(p, p_m) < \delta + \frac{1}{2}\delta p_m \leq \delta p_m$. $\therefore d_Y(f(q), f(p)) \leq d_Y(f(q), f(p_m)) + d_Y(f(q_m), f(p)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Q.E.D.

Theorem 4.22. If X, Y are metric spaces, $f : X \to Y$ is continuous, and $E \subset X$ is connected, then $f(E)$ is connected.

Proof (By Contradiction)**.** Suppose for contradiction E is connected and there exists $A, B \subset Y$ s.t. $f(E) = A \cap B$, $f(E) \neq \emptyset$, $\overline{A} \cup B = A \cap \overline{B} = \emptyset$. Let $G = f^{-1}(A) \cap E, H = f^{-1}(B) \cap E$. Then $E = G \cup H$, G, H are nonempty. If $G \cap \overline{H} = \overline{G} \cap H = \emptyset$, it leads to a contradiction. First, $G \subset f^{-1}(A) \subset (\because A \subset \overline{A}) f^{-1}(\overline{A})$, where $f^{-1}(\overline{A})$ is closed by the corollary to Theorem 4.8, so $\overline{G} \subset f^{-1}(\overline{A})$. Second, $f(H) = B, \overline{A} \cap B = \emptyset$. Therefore, $\overline{G} \cap H = \emptyset$. WLOG, $G \cap \overline{H} = \emptyset$ as well. Hence a contradiction. Q.E.D.

Theorem 4.23. [Intermediate Value Theorem] Suppose $f : [a, b] \to \mathbb{R}$ is continuous. $\forall_{c \in (\min\{f(a), f(b)\}, \max\{f(a), f(b)\})} : \exists_{x_0 \in (a,b)}$ s.t. $f(x_0) = c$.

Proof. $[a, b]$ is connected by Theorem 2.47. Hence, by Theorem 4.22, $f([a, b])$ is connected and therefore contains all points between $f(a)$ and $f(b)$. In particular, $c \in f((a, b))$ Q.E.D.

Example. *(a) there exists a continuous function called (Peano/space-filling curve) from* $[0, 1]$ *onto the closed unit square* $S = [0, 1] \times [0, 1] \subset \mathbb{R}^2$.

Proof. Omitted. See Rudin's problem 7.14 for an explicit example(covered in MATH-321). Q.E.D.

(b) But no such function can be one-to-one.

Proof. Suppose $f : [0, 1] \to S$ is 1-1, onto, continuous. Since $[0, 1]$ is compact, f^{-1} is continuous by Theorem 4.17. Let $E = [0, \frac{1}{2}) \cup (\frac{1}{2}, 1]$. Then, $f(E) = S \setminus \{f(\frac{1}{2})\}$ is S minus one point, which is connected (pf omitted). But then, $f^{-1}(f(E)) = E$ must be connected by Theorem 4.22. E is not connected, so this is a contradiction. Q.E.D.

Theorem 4.29. Let f be a monotonically increasing function on (a, b). Then $f(x+)$ and $f(x-)$ exist at every $x \in (a, b)$; i.e.,

$$\sup_{a<t<x} f(t) = f(x^-) \leq f(x) \leq f(x^+) = \inf_{x<t<b} f(t).$$

Moreover, if $a < x < y < b$, then

$$f(x^+) \leq f(y^-).$$

Analogous results hold for monotonically decreasing functions.

Example (18)**.** *Every rational* x *can be written in the form* $x = m/n$, *where*

$n > 0$ and m and n are integers without any common divisors. When $x = 0$, we take $n = 1$. Consider the function f defined on $\mathbb{R}^1$ by

$$f(x) = \begin{cases} 0 & (x \in \mathbb{R} \setminus \mathbb{Q}) \\ \dfrac{1}{n} & (x = \dfrac{m}{n}) \end{cases} .$$

Prove that f is continuous at every irrational point, and that f has a simple discontinuity at every rational point, and that f has a simple discontinuity at every rational point.

Chapter 5

Differentiation

We consider $f : [a, b] \to \mathbb{R}$.

Definition. For $f : [a, b] \to \mathbb{R}$ and $x \in [a, b]$, let $f'(x) = \lim_{t\to x} \frac{f(t)-f(x)}{t-x}$ if limit exists. Equivalently, $f(t) = f(x) + (t - x)[f'(x) + u(x, t)]$ with $\lim_{t\to x} u(x, t) = 0$.

Example. *(a)* $f(x) = c$ *for all* $x \Rightarrow f'(x) = \lim_{t\to x} \frac{c-c}{t-x} = 0$.

(b) $f(x) = x$ *for all* $x \Rightarrow f'(x) = \lim_{t\to x} \frac{t-x}{t-x}$

(c) $f(x) = e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}$. *Write* $t = x + h$, *so* $t \to x \Leftrightarrow h \to 0$. $\frac{e^{x+h}-e^x}{(x+h)-x} = e^x \frac{e^h-1}{h} = e^x \frac{e^h-1}{h} + e^x - e^x = e^x + e^x \frac{e^h-1-h}{h}$. *Let* $u(h) = \frac{e^h-1-h}{h}$. *Then* $u(h) = \frac{1}{h} \sum_{n=2}^{\infty} \frac{h^n}{n!}$, *so* $|u(h)| = |\frac{1}{h} \sum_{n=2}^{\infty} \frac{h^n}{n!}| \leq |h| \sum_{n=2}^{\infty} \frac{1}{n!} = (e - 2)|h|$ *(note for* $n \geq 2, |h^{n-1}| \leq |h|$ *if* $|h| \leq 1$*). Hence,* $u(h) \to \infty$ *as* $h \to 0$ *and therefore* $f'(x) = e^x$.

> **Remark.** $e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}$ is well defined. $e^1 = \sum_{n=0}^{\infty} \frac{1}{n!} = e$. Regarding it as a power series, its radius of convergence is $R = \infty$. Also, $e^{x+y} = e^x e^y$ using definition 3.48 of product series (Rudin's p.178-180).

> **Note.** $f'(x)$: Lagrange's notation, $\frac{df}{dx}$: Leibnitz notation

> **Theorem 5.2.** Suppose $f : [a, b] \to \mathbb{R}$ and $f'(x)$ exists. Then f is continuous at x.

Proof. The existence of $f'(x) \Leftrightarrow f(t) = f(x) + (t-x)[f'(x) + u(x,t)]$ with $\lim_{t\to x} u(x,t) = 0$. Let $t \to x$. $\lim_{t\to x} f(x) + (t-x)[f'(x) + u(x,t)] = f(x) + 0[f'(x) + 0] = f(x)$, so $\lim_{t\to x} f(t) = f(x)$; i.e., f is continuous at x. Q.E.D.

Remark. The converse is false; e.g., $f(x) = |x|$ is continuous for all x, but $f'(0)$ does not exist.

Theorem 5.3. If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are both differentiable at x then so are $f+g, fg, \frac{f}{g}($ if $g(x) \neq 0)$, and $(f+g)'(x) = f'(x) + g'(x), (fg)'(x) = f'(x)g(x) + f(x)g'(x), (\frac{f}{g})'(x) = \frac{f'(x)g(x)-g'(x)f(x)}{g(x)^2}$.

Proof (Only the quotient rule). $h(t)-h(x) = \frac{f(t)}{g(t)} - \frac{f(x)}{g(x)} = \frac{1}{g(t)g(x)}[(f(t)g(x) - f(x)g(x)) - (f(x)g(t) - f(x)g(x))]$.
Then $\frac{h(t)-h(x)}{t-x} = \frac{1}{g(t)g(x)}\left[\frac{f(t)-f(x)}{t-x}g(x) - f(x)\frac{g(t)-g(x)}{t-x}\right]$. Let $t \to x$. $h'(x) = \frac{1}{g(x)^2}\left[f'(x)g(x) - f(x)g'(x)]\right]$. Q.E.D.

Remark. By induction, $(f_1 \cdots f_n)' = f_1' f_2 \cdots f_n + f_1 f_2' f_3 \cdots f_n + \cdot f_1 f_2 \cdots f_n'$.

Example. *For* $n = 2, 3$, $\dfrac{d}{dx}x^n = nx^{n-1}$ *and we know this already for* $n = 0, 1$. *For* $n = -1, -2$, *let* $m = -n > 0$. *Then* $\dfrac{d}{dx}x^n = \dfrac{d}{dx}\dfrac{1}{x^m} = \dfrac{(\frac{d}{dx}1)x^m - (\frac{d}{dx}x^m)1}{(x^m)^2} = \dfrac{0x^m - mx^{m-1}1}{x^{2m}} = -mx^{m-1} = nx^{n-1}$. *Hence,* $\forall_{n\in\mathbb{Z}} : \dfrac{d}{dx}^{\,n} = nx^{n-1}$.

Theorem 5.5. [Chain Rule] Suppose $f : [a,b] \to \mathbb{R}$, $f'(x)$ exists for some $x \in [a,b], f([a,b]) \subset I$, where I is some interval in $\mathbb{R}$. Suppose $g : I \to \mathbb{R}$ and $g'(f(x))$ exists. Then $g \circ f$ is differentiable at x and $(g \circ f)(x)' = g'(f(x))f'(x)$

Proof. Let $h(t) = (g \circ f)(t) = g(f(t))$ for $t \in [a, b]$. Fix $x \in [a, b]$ where $f'(x)$ exists. We know:

(a) $f(t) - f(x) = (t - x)[f'(x) + u(t)]$ with $\lim_{t \to x} u(t) = 0$.

(b) With $y = f(x)$, $g(s) - g(y) = (s - y)(g'(y) + v(s))$ with $\lim_{s \to y} v(s) = 0$

As $\dfrac{h(t) - h(x)}{t - x} = \dfrac{g(f(t)) - g(f(x))}{t - x}$. By (2), $\dfrac{g(f(t)) - g(f(x))}{t - x}[g' f(x) + v(f(t))]$. Let $t \to x$. Then RHS $\to f'(x)[g'(f(x)) + 0]$ since $f(t) \to f(x)$ by continuity of f at x. Therefore, $h'(x) = f'(x)g'(f(x))$. Q.E.D.

Note. Suppose you produce $f(t)$ meters of wire by time t; i.e., rate of wire production is $f'(t)\ m/x$. Also suppose you get \$ $g(l)$ for l meters of wire; rate of profit is $g'(l)$ $\$/m$. Then the rate of earning by time t is $g'(f(t))f'(t)$ $\$/m$.

Example. *(a)* $\dfrac{d}{dx} e^{x^2} = 2xe^{x^2}$

(b) $f(x) = \begin{cases} x \sin \dfrac{1}{x} & \textit{if } x \neq 0 \\ 0 & \textit{if } x = 0 \end{cases}$

Remark. (a) f is continuous on $\mathbb{R}$, including at $x = 0$.

Proof. $|f(x)| \leq |x|$, so by the Squeeze theorem, $\lim_{x \to 0} f(x) = 0 = f(0)$. Q.E.D.

(b) f is differentiable on $x \neq 0$, but not differentiable at $x = 0$. For $x \neq 0$, $f'(x) = \sin \dfrac{1}{x} + x(\cos \dfrac{1}{x})(\dfrac{-1}{x^2}) = \sin \dfrac{1}{x} - \dfrac{1}{x} \cos \dfrac{1}{x}$. For $x = 0$, $\dfrac{f(t) - f(0)}{t - 0} = \dfrac{t \sin \frac{1}{t}}{t} = \sin \dfrac{1}{t}$, which does not converge. Therefore, f not differentiable at $x = 0$.

(c) Let $f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \textit{if } x \neq 0 \\ 0 & x = 0 \end{cases}$.

(a) f is continuous in $\mathbb{R}$ including at $x = 0$ ($\because |f(x)| \leq |x^2|$).

(b) f is differentiable in $\mathbb{R}$ including at $x = 0$.

Proof. For $x \neq 0$, $f'(x) = 2x \sin \frac{1}{x} + x^2 (\cos \frac{1}{x})(\frac{-1}{x^2}) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$. For $x = 0$, $\frac{f(t) - f(0)}{t - 0} = \frac{t^2 \sin \frac{1}{t}}{t} = t \sin \frac{1}{t} \to 0$ as $t \to 0$. Hence, $f'(0) = 0$. HOWEVER! f' is not continuous at $x = 0$, because $\lim_{x \to 0} f'(x)$ does not exist. Q.E.D.

Definition 5.1. Let X be a metric space, $f : X \to \mathbb{R}$. f has a *local* max at $x \in X$ if $\exists_{\delta > 0}$ s.t. $f(y) \geq f(x)$ for all $y \in N_\delta(x)$.

Theorem 5.8. Let $f : [a, b] \to \mathbb{R}$. If f has a local min or a local max at $x \in (a, b)$ and if $f'(x)$ exists, then $f'(x) = 0$.

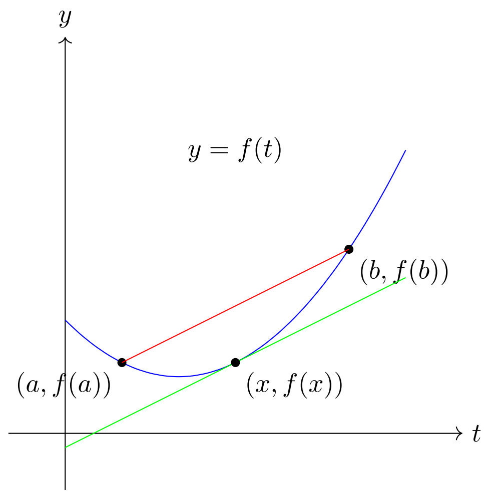
Proof (local min). Suppose f has a local min at x and $f'(x)$ exists. Choose $\delta > 0$ s.t. $N_\delta(x) \subset (a, b)$ and $f(t) \geq f(x)$ if $t \in (x - \delta, x + \delta)$. For $x < t < x + \delta$, $\frac{f(t) - f(x)}{t - x} \geq 0 (\because f(t) \geq f(x), t > x)$, so $f'(x) \geq 0$. For $x - \delta < t < x$, $\frac{f(t) - f(x)}{t - x} \leq 0 (\because f(t) - f(x) \geq 0, t < x)$, so $f'(x) \leq 0$. Hence, $f'(x) = 0$. Q.E.D.

Remark. Note that the hypothesis of the theorem requires *open* interval and existence $f'(x)$. If these conditions are not met, then $f'(x) = 0$ doesn't have to be the case.

Example. *(a) $f(x) = |x|$ has a local min at $x = 0$ but $f'(0)$ does not exist.*

(b) $f : [0, 1] \to \mathbb{R}$ defined by $f(x) = x$ has a local max at $x = 1$ and local min at $x = 0$, but $f'(0) = f'(1) = 1$.

Theorem 5.10. [Mean-Value Theorem] If $f : [a, b] \to \mathbb{R}$ is continuous and differentiable on (a, b), then $\exists_{x \in (a, b)}$ s.t. $f(b) - f(a) = f'(x)(b - a)$.



Proof. Let $L : y = f(a) + m(t-a)$, where $m = \dfrac{y - f(a)}{t-a} = \dfrac{f(b)-f(a)}{b-a}$. Subtract L from the curve $y = f(t)$. Let $h(t) = f(t) - [f(a) + m(t-a)]$. Then $h(a) = h(b) = 0$. $h'(t) = f'(t) - m = f'(t) - \dfrac{f(b)-f(a)}{b-a}$. Therefore, it suffices to find x s.t. $h'(x) = 0$. h is continuous and $[a,b]$ is compact, so $h([a,b])$ is also compact. Hence, h attains its global max$(= \sup\{h([a,b])\})$ and global min$(= \inf\{h([a,b])\})$ on $[a,b]$. If $h(t) = 0$ for all $t \in [a,b]$ then $h'(t) = 0$ for all $t \in [a,b]$ so any $x \in (a,b)$ will do. Otherwise, h attains its global max or global min at some $x \in (a,b)$. By Theorem 5.8, $h'(x) = 0$. Q.E.D.

Theorem 5.11. If f is differentiable on (a,b) then

(a) $f'(x) \geq 0$ for all $x \in (a,b)$ implies f is monotone increasing.

(b) $f'(x) \leq 0$ for all $x \in (a,b)$ implies f is monotone decreasing.

(c) $f'(x) = 0$ for all $x \in (a,b)$ implies f is constant.

Proof ((a) only). Suppose $f'(x) \geq 0$ for all $x \in (a,b)$. For $a < x_1 < x_2 < b$, $\exists_{x \in (x_1, x_2)}$ s.t. $f(x_2) - f(x_1) = f'(x)(x_2 - x_1)$ by Theorem 5.10. As $f'(x) \geq 0, x_2 \geq x_1$, $f(x_2) - f(x_1) \geq 0$. Q.E.D.

Definition 5.2. $f^{(0)} = f, f^{(1)} = f', f^{(2)} = f'', f^{(3)} = f'''$ and so on.

Theorem 5.15. [Taylor's Theorem] Suppose $f : [a,b] \to \mathbb{R}$, $n \in \mathbb{N}$, and

$f^{(n+1)}(x)$ exists for all $x \in (a,b)$. Let $x, x_0 \in [a,b]$. Then $\exists_{c\in(x,x_0)}$ s.t.

$$f(x) = \underbrace{\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}_{P_n(x),n^{\text{th}} \text{ Taylor polynomial of } f \text{ at } x_0} + \underbrace{\frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}}_{\text{Taylor Remainder}}.$$

Proof. If $n = 0$, the mean-value theorem guarantees existence of c. For general n, $A \in \mathbb{R}$ by $R_n(x) = f(x) - P_n(x) = \frac{A}{(n+1)!}(x-x_0)^{n+1}$, where A depends on f, n, x, x_0. Claim: $A = f^{(n+1)}(c)$ for some c between x and x_0.

Define $g(t) = f(t) - P_n(t) - \frac{A}{(n+1)!}(t-x_0)^{n+1}$ for $t \in [a,b]$. Then $g(x_0) = 0$. $g(x) = f(x) - P_n(x) - \frac{A}{(n+1)!}(x-x_0)^{n+1} = 0$ by the definition of A. Also for $j = 1, \ldots, n$, then $P_n^{(j)}(x_0) = f^{(j)}(x_0)$, $\frac{\mathrm{d}^j}{\mathrm{d}x^j}(x-x_0)^{n+1} \mid_{x=x_0} = 0$. Hence, $g^{(j)}(x_0) = f^{(j)}(x_0) - P_n^{(j)}(x_0) - 0 = 0$. $g^{(n+1)}(t) = f^{(n+1)}(t) - 0 - A$. We need to find c s.t. $g^{(n+1)}(c) = 0$.

$g(x) = g(x_0) = 0 \Rightarrow \exists_{c_1 \in (\min\{x_0,x\}, \max\{x_0,x\})}$ s.t. $g'(c_1) = 0$.

$g'(x_0) = g'(c_1) = 0 \Rightarrow \exists_{c_2 \in (\min\{x_0,x,c_1\}, \max\{x_0,x,c_1\})}$ s.t. $g''(c_2) = 0$.

$\vdots$

Finally, $\exists_{c_{n+1}=c}$ s.t. $g^{n+1}(c) = 0$ and hence $f^{(n+1)}(c) = A$. Q.E.D.

Example. *(not in Rudin) Does* $\sum_{n=1}^{\infty} \left(\sqrt{1+\frac{1}{n^2}} - 1\right)$ *converge or diverge?*

Method 1:

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

so the series converges by the comparison test since $\sum \frac{1}{n^2}$ *converges.*

Method 2: Using Taylor's theorem. Let $f(x) = \sqrt{1+x} = (1+x)^{1/2}$. *Then*

$$\begin{aligned}
f'(x) &= \frac{1}{2}(1+x)^{-1/2} \\
f''(x) &= \frac{1}{2}(-\frac{1}{2})(1+x)^{-3/2} = -\frac{1}{4}(1+x)^{-3/2} \\
f(x) &= P_1(x) + R_1(x) \\
&= f(0) + \frac{f'(0)}{1!}(x-0)^1 + \frac{f''(c)(x-0)^2}{2!} \\
&= 1 + \frac{1}{2}x + R_1(x).
\end{aligned}$$

$|R_1(x)| \leq (\frac{1}{2} \cdot \frac{1}{2} \cdot 1)\frac{1}{2!}x^2 = \frac{1}{8}x^2$ *for* $x \in [0,1]$. *Therefore,* $\sqrt{1+\frac{1}{n^2}} - 1 = f(\frac{1}{n^2}) - 1 = \frac{1}{2}(\frac{1}{n^2}) + R_1(\frac{1}{n^2}) \leq \frac{1}{2n^2} + \frac{1}{8}\frac{1}{n^4}$. *Since* $\sum\left(\frac{1}{2n^2} + \frac{1}{8n^4}\right)$ *converges,* $\sum\left(\sqrt{1+\frac{1}{n^2}} - 1\right)$ *converges by comparison test.*

Example. *Let* $f(x) = \sin x, x_0 = 0$.

Taylor series for $f(x)$. $f'(x) = \cos x, f''(x) = -\sin x, f'''(x) = -\cos x, \ldots$, *so* $f^{(k)}(x) = \begin{cases} (-1)^m \sin x & (k = 2m) \\ (-1)^m \cos x & (k = 2m+1) \end{cases}$. *Hence* $n \geq 0, f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k + \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}$, *where* c *between* 0 *and* x. *Remainder estimate:* $\left|\frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}\right| \leq \frac{|x|^{n+1}}{(n+1)!} \to 0$ *as* $n \to \infty$ *because* $\frac{|x|^{n+1}}{(n+1)!}$ *is the* $(n+1)^{th}$ *term in convergent series* $e^{|x|} = \sum_{n=0}^{\infty} \frac{|x|^n}{n!}$.

$\sin x = \sum_{n=0}^{\infty} \frac{f^{(k)}(0)}{k!}x^k = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1}$.

Taylor approximation. *Find* $\sin 0.2$ *to within an error* $\pm 10^{-6}$. *Use* $\sin 0.2 = \frac{2}{10} - \frac{1}{3!}(\frac{2}{10})^3 + \frac{1}{5!}(\frac{2}{10})^5 - \cdots$.

Method 1: Alternating Series Test. *If* $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$, *and* $a_n \to 0$, *then* $\sum_{k=1}^{\infty}(-1)^{k-1}a_k = s$ *converges and* $|s - s_n| \leq a_{n+1}$. *Above series satisfies the hypotheses, so truncation error is* $\leq$ *first omitted term. We look for when* $\frac{1}{(2k+1)!}(\frac{2}{10})^{2k+1} \leq 10^{-6}$; *i.e.,*

$(2k+1)! \cdot \frac{10^{2k+1}}{2^{2k+1}} \geq 10^6$.

If $k = 1$, $3! \cdot \frac{10^3}{2^3} < 10^6$.

If $k=2$, $5! \cdot \dfrac{10^5}{2^5} < 10^6$.

If $k=3$, $7! \cdot \dfrac{10^7}{2^7} \geq 10^6$, *so* $k=3$ *works. Therefore,* $\sin 0.2 = 0.2 - \dfrac{1}{3!}(0.2)^3 + \dfrac{1}{5!}(0.2)^5 \pm 10^{-6} = 0.198669 \pm 10^{-6}$.

Method 2: General Case. *If alternating series test does not apply, estimate remainder using the worst* c *for* $\left|\dfrac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}\right|$. *In our example,* $\left|\dfrac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}\right| \leq \dfrac{1}{(n+1)!}(0.2)^{n+1}$, *so seek* n *s.t.* $\dfrac{1}{(n+1)!}\left(\dfrac{2}{10}\right)^{n+1} \leq 10^{-6}$. *First* n *that works is* $n=6$, *same as before.*

Chapter 6

Riemann-Stieltjes Integral

Definition (Partition)**.** A partition P of $[a,b]$ is $\{x_0, x_1, x_2, \ldots, x_n\}$ for some $n \geq 1$, with $a = x_0 \leq x_1 \ldots \leq x_{n-1} \leq x_n = b$.

Notation. $\Delta x_i \quad = x_i - x_{i-1}$ for $i = 1, \ldots, n$

$f \quad : [a,b] \to \mathbb{R}$ be bounded, which is not necessarily continuous

$M_i \quad = \sup\{f(x) : x_{i-1} \leq x \leq x_i\}$, $m_i = \inf\{f(x) : x_{i-1} \leq x \leq x_i\}$

$U(P,f)$
$= \sum_{i=1}^{n} M_i \Delta x_i$

$L(P,f)$
$= \sum_{i=1}^{n} m_i \Delta x_i$.

> **Note.** $L(P,f) \leq U(p,f)$ always.

Definition (Riemann Integral)**. Upper Riemann Integral** : $\overline{\int_a^b} f(x)dx = \inf_P\{U(P,f)\} = \inf\{U(P,f) : P \text{ is a partition of } [a,b]\}$.

Lower Riemann Integral : $\underline{\int_a^b} f(x)dx = \sup_P\{L(P,f)\} = \sup\{L(P,f) : P \text{ is a partition of } [a,b]\}$.

Riemann Integrable : f is Riemann integrable on $[a,b]$ if $\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$.
If f is Riemann integrable on $[a,b]$, we write $f \in \mathscr{R}[a,b]$ and

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

Note. Since f is bounded, $m = \inf\{f(x) : a \leq x \leq b\}$ and $M = \sup\{f(x) : a \leq x \leq b\}$ are both finite. Hence, for any P, $m \leq m_i \leq M_i \leq M$ and $\forall_i : m(b-a) \leq L(P,f) \leq U(P,f) \leq M(b-a)$.

Notation. Let $\alpha : [a,b] \to \mathbb{R}$ is a monotone increasing function. Then $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$.

Definition 6.2. Given P, let $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$. (Note: $\Delta\alpha_i \geq 0$).
For bounded f, let $U(P,f,\alpha) = \sum_{i=1}^n M_i \Delta\alpha_i$, $L(P,f,\alpha) = \sum_{i=1}^n m_i \Delta\alpha_i$.

Upper Riemann-Stieltjes Integral $\overline{\int_a^b} f(x)d\alpha = \overline{\int_a^b} f(x)d\alpha(x) = \inf_P\{U(P,f,\alpha)\} = \inf\{U(P,f,\alpha) : P \text{ is a partition of } [a,b]\}$.

Lower Riemann-Stieltjes Integral $\underline{\int_a^b} f(x)d\alpha = \underline{\int_a^b} f(x)d\alpha(x) = \sup_P\{L(P,f,\alpha)\} = \sup\{L(P,f,\alpha) : P \text{ is a partition of } [a,b]\}$.

If $\overline{\int_a^b} f(x)d\alpha = \underline{\int_a^b} f(x)d\alpha$, then $f \in R[a,b,\alpha]$ and $\int_a^b f(x)d\alpha = \overline{\int_a^b} f(x)d\alpha = \underline{\int_a^b} f(x)d\alpha$.
If $\alpha(x) = x$, then equivalent to $\int_a^b f(x)dx$.

Definition 6.3. (a) Partition P^* is called a refinement of P if $P \subset P^*$.

(b) Partition P^* is called the common refinement of P_1 and P_2 if $P^* = P_1 \cup P_2$.

Theorem 6.4. If P^* is a refinement of P then $L(P,f,\alpha) \leq L(P^*,f,\alpha) \leq U(P^*,f,\alpha) \leq U(P,f,\alpha)$.

Proof. It's enough to consider p^* with one extra point: $x_{i-1} \leq x^* \leq x_i$.
Sketch for L:

$$\begin{aligned}
&L(P^*,f,\alpha) - L(p,f,\alpha) \\
&= m^*[\alpha(x^*) - \alpha(x_{i-1})] + m_i[\alpha(x_i)\alpha(x^*)] - m_i[\alpha(x^*) - \alpha(x_{i-1})] - m_i[\alpha(x_i) - \alpha(x^*)] \\
&= (m^* - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (m_i - m_i)[\alpha(x_i) - \alpha(x^*)]
\end{aligned}$$

Q.E.D.

Notation. When f, α are fixed, we write $L(P) = L(P,f,\alpha), U(P) = U(P,f,\alpha)$

Theorem 6.5. $\underline{\int_a^b} f \mathrm{d}\alpha \le \overline{\int_a^b} f \mathrm{d}\alpha$.

Proof. For partitions P_1, P_2, let $P^* = P_1 \cup P_2$. By Theorem 6.4, $L(P_1) \le L(P^*) \le U(P^*) \le U(P_2)$. In particular, $\sup_{P_1}\{L(P_1)\} \le U(P_2)$ for all P_2. Hence, $\sup_{P_1}\{L(P_1)\} \le \inf_{P_2}\{U(P_2)\}$. Q.E.D.

Theorem 6.6. $f \in \mathscr{R}_\alpha[a,b] \Leftrightarrow \forall_{\varepsilon>0} : \exists P_\varepsilon$ s.t. $U(P_\varepsilon) - L(P_\varepsilon) < \varepsilon$

Proof. Let $\varepsilon > 0$.

($\Rightarrow$) By hypothesis, $\sup_P\{L(P)\} = \underline{\int_a^b} f \mathrm{d}\alpha = \overline{\int_a^b} f \mathrm{d}\alpha = \inf_P\{U(P)\}$.
$\exists P_1, P_2$ s.t. $L(P_1) > \underline{\int_a^b} f \mathrm{d}\alpha - \varepsilon/2$ and $U(P_2) < \overline{\int_a^b} f \mathrm{d}\alpha + \varepsilon/2$. Then $U(P_2) - L(P_1) < \varepsilon$. Let $P_\varepsilon = P^* = P_1 \cup P_2$. By Theorem 6.4, $L(P_1) \le L(P^*) \le U(P^*) \le U(P_2)$, so $U(P_\varepsilon) - L(P_\varepsilon) \le U(P_2) - L(P_1) < \varepsilon$.

($\Leftarrow$) $0 \le \overline{\int_a^b} f \mathrm{d}\alpha - \underline{\int_a^b} f \mathrm{d}\alpha \le U(P_\varepsilon) - L(P_\varepsilon) < \varepsilon$. Since ε is arbitrary, $\overline{\int_a^b} f \mathrm{d}\alpha = \underline{\int_a^b} f \mathrm{d}\alpha$.

Q.E.D.

Remark. very important

Theorem 6.7. Let $\varepsilon_0 > 0$ be fixed. Suppose there exists a partition $P = \{x_0 = a, \dots, x_n = b\}$ s.t. $U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon_0$. Let s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$. Then,

(a) For any refinement of P, denoted by P^*, $U(P^*, f, \alpha) - L(P^*, f, \alpha) < \varepsilon_0$ also holds true

(b) $\sum_{i=1}^n |f(s_i) - f(t_i)| \Delta\alpha_i < \varepsilon_0$

(c) If $f \in \mathscr{R}_\alpha$, then $\left|\sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f \mathrm{d}\alpha\right| < \varepsilon_0$

Theorem 6.8. If f is continuous on $[a,b]$ then $f \in \mathscr{R}_\alpha[a,b]$.

Proof. For any P, $U(P) - L(P) = \sum_{i=1}^n (M_i - m_i)\Delta\alpha_i$. Since $[a, b]$ is compact, f is uniformly continuous on $[a, b]$ (Theorem 4.19), so $\forall \eta > 0 : \exists \delta > 0$ s.t. $|x - t| < \delta \Rightarrow |f(x) - f(t)| < \eta$. Given $\varepsilon > 0$, choose η s.t. $\eta[\alpha(b) - \alpha(a)] < \varepsilon$ and choose P with $\Delta x_i < \delta = \delta(\eta)$ for all i. For such P, $M_i - m_i \le \eta$. Then $U(P) - L(P) \le \sum_{i=1}^n \eta\Delta\alpha_i = \eta[\alpha(b) - \alpha(a)] < \varepsilon$. Therefore, $f \in \mathscr{R}_\alpha[a, b]$. Q.E.D.

Theorem 6.9. If f is monotone increasing or decreasing on $[a, b]$ and α is continuous on $[a, b]$ then $f \in \mathscr{R}_\alpha[a, b]$.

Proof. By definition, $U(P) - L(P) = \sum_{i=1}^n (M_i - m_i)\Delta\alpha_i$. Given $n \in \mathbb{N}$, let P s.t. $\Delta\alpha_i = \dfrac{\alpha(b) - \alpha(a)}{n}$ for all i. Such P exists by the intermediate value theorem (Theorem 4.23) as α is continuous. Then, $U(P) - L(P) = \dfrac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^n M_i - m_i$. Suppose f is increasing, so $M_i - m_i = f(x_i) - f(x_{i-1})$. Then $U(P) - L(P) = \dfrac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^n f(x_i) - f(x_{i-1}) = \dfrac{\alpha(b) - \alpha(a)}{n}[f(b) - f(a)]$. Given $\varepsilon > 0$, we can choose n (hence P) s.t. $U(P) - L(P) < \varepsilon$. Therefore, $f \in \mathscr{R}_\alpha[a, b]$ by Theorem 6.6. Q.E.D.

Note. We always assume α is monotone.

Theorem 6.10. If f is bounded on $[a, b]$ and has only finitely many discontinuities, and α is continuous at each point where f is not, then $f \in \mathscr{R}_\alpha[a, b]$.

Proof. We apply Theorem 6.6. Use $U(P)-L(P) = \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i$. Let $\varepsilon > 0$ and $E = \{e_1, \ldots, e_k\}$ be the set of points where f is discontinuous. α is assumed to be continuous at each e_i, which implies $\exists(u_j, v_j)$ s.t. $u_j < e_j < v_j$ and $\alpha(v_j) - \alpha(u_j) < \varepsilon$. (Relax inequality to include equality if $e_1 = a$, $e_k = b$).
Let $K = [a,b] \cap \left(\bigcup_{j=1}^{k}(u_j, v_j)\right)^c$. K is compact. f is continuous on K, so f is uniformly continuous on K by Theorem 4.19. Hence, $\exists\delta > 0$ s.t. for $s, t \in K$, $|s-t| < \delta \Rightarrow |f(s) - f(t)| < \varepsilon$.
Form P to consist of $\{u_1, v_1, \ldots, u_k, v_k\}$ and additional points in K so that $\Delta x_i < \delta$. If x_i is in K, then $M_i - m_i < \varepsilon$. Otherwise, $x_i = u_j$ or $x_i = v_j$ for some j, so $\Delta\alpha_i \le \varepsilon$. Then

$$
\begin{aligned}
0 \le U(P) - L(P) &= \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i \\
&\le \underbrace{k \cdot 2M\varepsilon}_{\text{from intervals in } \bigcup_{j=1}^{k}(u_j, v_j)} + \underbrace{\varepsilon[\alpha(b) - \alpha(a)]}_{\text{from intervals in } K}.
\end{aligned}
$$

As RHS is as small as we want by taking ε small enough, $f \in \mathscr{R}_\alpha[a,b]$.
Q.E.D.

Remark. (a) Theorem 6.10 implies part of A1.2 but do the problem from first principles. Do not apply Theorem 6.10 directly.

(b) A1.4 shows what can happen if f, α are discontinuous at the same point.

Theorem 6.11. If $f \in \mathscr{R}_\alpha[a,b], m \le f(x) \le M$ for all $x \in [a,b]$, and $\varphi : [m,M] \to \mathbb{R}$ is continuous, then $\varphi \circ f \in \mathscr{R}_\alpha[a,b]$.

Proof. Let $\varepsilon > 0$. As φ continuous on $[m, M]$, φ is uniformly continuous on $[m, M]$ by Theorem 4.19. That is, $\exists \delta < \varepsilon$ s.t. $|\varphi(s) - \varphi(t)| < \varepsilon$ if $|s - t| < \delta$ for $s, t \in [m, M]$.
Since $f \in \mathscr{R}_\alpha$, there exists a partition $P = \{x_0, x_1, \dots, x_n\}$ of $[a, b]$ such that $U(P) - L(P) < \delta^2$.
Let $A = \{i \in \{1, 2, \dots, n\} : M_i - m_i < \delta\}, B = \{i \in \{1, 2, \dots, n\} : M_i - m_i \geq \delta\}$. Note $A \cup B = \{1, 2, \dots, n\}$.
Let $M_i^* = \sup\{\varphi(f(x)) : x_{i-1} \leq x \leq x_i\}$ and $m_i^* = \inf\{\varphi(f(x)) : x_{i-1} \leq x \leq x_i\}$. Suppose $i \in A$. Then $M_i - m_i < \delta$. By definition of δ, this implies $|M_i^* - m_i^*| \leq \varepsilon$.
Suppose $i \in B$. By definition of P,

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i \in B} (M_i - m_i)\Delta\alpha_i < \delta^2.$$

As $M_i - m_i \geq \delta$,

$$\delta \sum_{i \in B} \Delta\alpha_i \leq \sum_{i \in B} (M_i - m_i)\Delta\alpha_i < \delta^2.$$

Hence, $\sum_{i \in B} \Delta\alpha_i < \delta$. Then,

$$\begin{aligned}
&\sum_{i \in B} (M_i^* - m_i^*)\Delta\alpha_i \\
&\leq 2 \cdot \underbrace{\sup\{|\varphi|\}}_{=\sup\{|\varphi(t)| : m \leq t \leq M\}} \cdot \left(\sum_{i \in B} \Delta\alpha_i\right) \\
&< 2 \cdot \sup\{|\varphi|\} \cdot \delta \\
&< 2 \cdot \sup\{|\varphi|\} \cdot \varepsilon.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&U(P, \varphi \circ f, \alpha) - L(P, \varphi \circ f, \alpha) = \sum_{i=1}^{n} (M_i^* - m_i^*)\Delta\alpha_i \\
&= \sum_{i \in A} (M_i^* - m_i^*)\Delta\alpha_i + \sum_{i \in B} (M_i^* - m_i^*)\Delta\alpha_i \\
&< \varepsilon[(\alpha(b) - \alpha(a)) + 2 \cdot \sup\{|\varphi|\}]
\end{aligned}$$

Q.E.D.

Example. *$f \in \mathscr{R}_\alpha[a, b] \Rightarrow f^2 \in \mathscr{R}_\alpha[a, b], |f| \in \mathscr{R}_\alpha[a, b]$ where $\varphi(t) = t^2$ and*

$\varphi(t) = |t|$ *respectively.*

Note. $\varphi \in \mathscr{R}_\alpha[m, M]$ does not imply $\varphi \circ f \in \mathscr{R}_\alpha[a, b]$. See A2.

Theorem 6.12. (Linearity and related properties)

(a) If $f, f_1, f_2 \in \mathscr{R}_\alpha[a, b]$, then $f_1 + f_2 \in \mathscr{R}_\alpha[a, b]$, $cf \in \mathscr{R}_\alpha[a, b]$, and

$$\int_a^b (f_1 + f_2) \mathrm{d}\alpha = \int_a^b f_1 \mathrm{d}\alpha + \int_a^b f_2 \mathrm{d}\alpha$$

$$\int_a^b cf \mathrm{d}\alpha = c \int_a^b f \mathrm{d}\alpha.$$

Proof. TEXTBOOK Q.E.D.

(b) $f_1, f_2 \in \mathscr{R}_\alpha$ and $f_1(x) \leq f_2(x)$ for all $x \in [a, b]$, then $\int_a^b f_1 \mathrm{d}\alpha \leq \int_a^b f_2 \mathrm{d}\alpha$.

Proof. $L(P, f_1) \leq L(P, f_2) \leq \sup_P L(P, f_2) = \int_a^b f_2 \mathrm{d}\alpha$. $\int_a^b f_1 \mathrm{d}\alpha = \sup_P L(P, f_1) \leq \int_a^b f_2 \mathrm{d}\alpha$. Q.E.D.

(c) If $f \in \mathscr{R}_\alpha[a, b], c \in [a, b]$ then $f \in \mathscr{R}_\alpha[a, c]$ and $f \in \mathscr{R}_\alpha[a, b]$ and $\int_a^b f \mathrm{d}\alpha = \int_a^c f \mathrm{d}\alpha + \int_c^b f \mathrm{d}\alpha$.

(d) If $f \in \mathscr{R}_\alpha$ and $|f(x)| \leq M$ for all $x \in [a, b]$, then $|\int_a^b f \mathrm{d}\alpha| \leq M(\alpha(b) - \alpha(a))$.

Proof. Let $P = \{a, b\}$. Then $-M[\alpha(b) - \alpha(a)] \leq m_1 \Delta\alpha_1 = L(P) \leq \int_a^b f \mathrm{d}\alpha \leq U(P) = M_1 \Delta\alpha_1 \leq M[\alpha(b) - \alpha(a)]$. Q.E.D.

(e) If $f \in \mathscr{R}_{\alpha_1}$ and $f \in \mathscr{R}_{\alpha_2}$, then $f \in \mathscr{R}_{\alpha_1 + \alpha_2}$ and

$$\int_a^b f \mathrm{d}(\alpha_1 + \alpha_2) = \int_a^b f \mathrm{d}\alpha_1 + \int_a^b f \mathrm{d}\alpha_2. \qquad (*)$$

If $f \in \mathscr{R}_\alpha$ and $c \geq 0$ then $f \in \mathscr{R}_\alpha$ and $\int_a^b f \mathrm{d}c\alpha = c \int_a^b f \mathrm{d}\alpha$.

Proof (*). Let $\varepsilon > 0$. Choose P_1, P_2 s.t. $U(P_j, f, \alpha_j) - L(P_j, f, \alpha_j) < \frac{\varepsilon}{2}$, where $j = 1, 2$. Let $P^* = P_1 \cup P_2$. By Theorem 6.4,

$$U(P^*, f, \alpha_j) - L(P^*, f, \alpha_j) < \frac{\varepsilon}{2}. \qquad (**)$$

Since $(\Delta\alpha_1)_i + (\Delta\alpha_2)_i = (\Delta(\alpha_1 + \alpha_2))_i$,

$$\begin{aligned}
&U(P^*, f, (\alpha_1 + \alpha_2)) - L(P^*, f, (\alpha_1 + \alpha_2)) \\
&= \sum_{i=1}^{n} (M_i - m_i)(\Delta(\alpha_1 + \alpha_2))_i \\
&= \sum_{i=1}^{n} (M_i - m_i)[(\Delta\alpha_1)_i + (\Delta\alpha_2)_i] \\
&= \sum_{i=1}^{n} (M_i - m_i)(\Delta\alpha_1)_i + \sum_{i=1}^{n} (M_i - m_i)(\Delta\alpha_2)_i \\
&= U(P^*, f, \alpha_1) - L(P^*, f, \alpha_1) + U(P^*, f, \alpha_2) - L(P^*, f, \alpha_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

By Theorem 6.6, $f \in \mathscr{R}_{\alpha_1+\alpha_2}$. Also $\int_a^b f\mathrm{d}(\alpha_1 + \alpha_2) \leq U(P^*, f, \alpha_1 + \alpha_2) = U(P^*, f, \alpha_1) + U(P^*, f, \alpha_2) < \int_a^b f\mathrm{d}\alpha_1 + \frac{\varepsilon}{2} + \int_a^b f\mathrm{d}\alpha_2 + \frac{\varepsilon}{2}$ by (**). Similarly, $\int_a^b f\mathrm{d}(\alpha_1 + \alpha_2) \geq L(P^*, f, \alpha_1 + \alpha_2) = L(P^*, f, \alpha_1) + L(P^*, f, \alpha_2) > \int_a^b f\mathrm{d}\alpha_1 - \frac{\varepsilon}{2} + \int_a^b f\mathrm{d}\alpha_2 - \frac{\varepsilon}{2}$ by (**). As ε is arbitrary, (*) holds. Q.E.D.

Theorem 6.13. (a) $f, g \in \mathscr{R}_\alpha \Rightarrow fg \in \mathscr{R}_\alpha$

Proof. By Theorem 6.11 with $\varphi(t) = t^2$, $h \in \mathscr{R}_\alpha \Rightarrow h^2 \in \mathscr{R}_\alpha$. By Theorem 6.12(a), $f, g \in \mathscr{R}_\alpha \Rightarrow f + g \in \mathscr{R}_\alpha$, so $(f \pm g)^2 \in \mathscr{R}_\alpha$. Since $(f + g)^2 - (f - g)^2 = 4fg$, by Theorem 6.12(a), $fg \in \mathscr{R}_\alpha$. Q.E.D.

(b) If $f \in \mathscr{R}_\alpha$, then $|f| \in \mathscr{R}_\alpha$, and $\left|\int_a^b f\mathrm{d}\alpha\right| \leq \int_a^b |f|\mathrm{d}\alpha$.

Proof. By Theorem 6.11, $|f| \in \mathscr{R}_\alpha$ (take $\varphi(t) = |t|$). Let

$$c = \operatorname{sgn}\left(\int_a^b f \mathrm{d}\alpha\right) = \begin{cases} +1 & \text{if } \int_a^b f \mathrm{d}\alpha > 0 \\ 0 & \text{if } \int_a^b f \mathrm{d}\alpha = 0 \\ -1 & \text{if } \int_a^b f \mathrm{d}\alpha < 0 \end{cases}.$$

As $cf \le |f|$, $\left|\int_a^b f \mathrm{d}\alpha\right| = c\int_a^b f \mathrm{d}\alpha = \int_a^b cf \mathrm{d}\alpha \le \int_a^b |f| \mathrm{d}\alpha$. Q.E.D.

Definition 6.14. [Unit Step Function]

$$I(x) = \begin{cases} 0 & x \le 0 \\ 1 & x > 0 \end{cases}$$

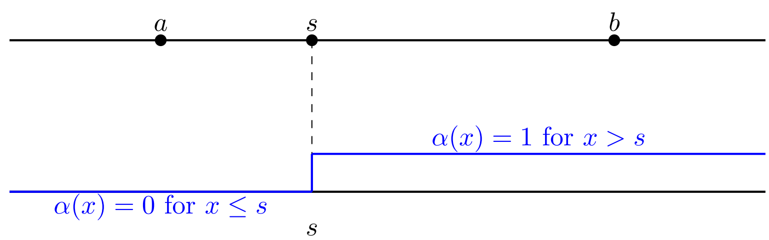
Theorem 6.15. Suppose f is bounded on $[a, b]$ and continuous at $s \in (a, b)$. Let

$$\alpha(x) = \begin{cases} 0 & x \le s \\ 1 & x > s \end{cases}.$$

Then $\int_a^b f \mathrm{d}\alpha$ exists, and

$$\int_a^b f \mathrm{d}\alpha = f(s).$$

Proof. Let $P = \{a, s, x, b\}$ with $x \in (s, b)$. $U(P) = \sum_{i=1}^3 M_i \Delta\alpha_i = 0 + M_2\Delta\alpha_1 = M_2$. $L(P) = \sum_{i=1}^3 m_i \Delta\alpha_i = 0 + m_2\Delta\alpha_1 = m_2$. Hence, $m_2 \le \underline{\int_a^b} f \mathrm{d}\alpha \le \overline{\int_a^b} f \mathrm{d}\alpha \le M_2$. Let x approach s. By continuity of f at s, M_2 approaches $f(s)$ from above and m_2 approaches $f(s)$ from below. Hence, $\underline{\int_a^b} f \mathrm{d}\alpha = f(s) = \overline{\int_a^b} f \mathrm{d}\alpha = \int_a^b f \mathrm{d}\alpha = f(s)$.



Q.E.D.

Remark. (a) By Theorem 4.29, if α is monotone-increasing, then $\alpha(x^+)$ and $\alpha(x^-)$ exist for all $x \in (a,b)$, and $\alpha(x^-) \le \alpha(x) \le \alpha(x^+)$.

(b) In Theorem 6.15, α is left-continuous at s.

Exercise

Prove the same conclusion for $\alpha(x) = \begin{cases} 0 & x < s \\ 1 & x \ge s \end{cases}$.

(c) This α plays the role of the Dirac delta function:

$$\int_a^b f(x)\delta(x-s)dx = f(s)$$

where $\delta(x-s) = \begin{cases} 0 & x \ne s \\ \infty & x = s \end{cases}$ is the Dirac delta function. Technically, there is no such function as $\delta(x-s)$, but it is a useful concept in physics. Note δ is kind of like α' in Theorem 6.15. (See A2).

(d) Theorem 6.15 has $\alpha(x) = I(x-s)$.

Theorem 6.16. Let $c_n \ge 0, \sum_{n=1}^{\infty} c_n < \infty, s_n \in (a,b)$, where $s_i \ne s_j$ if $i \ne j$.
Let $\alpha(x) = \sum_{n=1}^{\infty} c_n I(x-s_n)$, a monotone increasing function. Let f be continuous on $[a,b]$. Then $\int_a^b f \mathrm{d}\alpha = \sum_{n=1}^{\infty} c_n f(s_n)$.

Note.

(a) $\alpha(x)$ converges for any $x \in (a,b)$ by the comparison test, with

$$0 \le \sum_{i=1}^{\infty} c_n I(x-s_n) \le \sum_{n=1}^{\infty} c_n < \infty$$

(b) $\sum_{n=1}^{\infty} c_n f(s_n)$ converges by the comparison test, with

$$|c_n f(x_n)| \le M \cdot c_n$$

Proof. Let $R_N = \int_a^b f\mathrm{d}\alpha - \sum_{i=1}^N c_n f(s_n)$.
Claim: $R_N \to 0$ as $N \to \infty$; i.e., given $\varepsilon > 0$, $\exists N_0$ s.t. $|R_N| < \varepsilon$ for all $N \geq N_0$.
Let

$$\alpha_1(x) = \sum_{n=1}^{N} c_n I(x - s_n)$$

$$\alpha_2(x) = \sum_{n=N+1}^{\infty} c_n I(x - s_n).$$

By Theorem 6.12 and Theorem 6.15,

$$\int_a^b f\mathrm{d}\alpha_1 = \sum_{n=1}^{N} c_n \int_a^b f(x)\mathrm{d}I(x - s_n) = \sum_{n=1}^{N} c_n f(s_n)$$

$$\int_a^b f\mathrm{d}\alpha_2 = \sum_{n=N+1}^{\infty} c_n f(s_n).$$

By Theorem 6.12,

$$\int_a^b f\mathrm{d}\alpha = \int_a^b f\mathrm{d}\alpha_1 + \int_a^b f\mathrm{d}\alpha_2.$$

Then $R_N = \int_a^b f\mathrm{d}\alpha_2$. Choose N_0 s.t. $\sum_{n=N_0+1}^{\infty} c_n < \varepsilon$. By Theorem 6.12(d),

$$|R_N| \leq M \cdot [\alpha_2(b) - \underbrace{\alpha_2(a)}_{=0}] = M \cdot \sum_{n=N+1}^{\infty} c_n < M \cdot \varepsilon$$

for all $N \geq N_0$. Q.E.D.

Theorem 6.17. Suppose

(a) $|f(x)| \leq M$ for all $x \in [a, b]$,

(b) α is differentiable on $[a, b]$ and increasing on $[a, b]$.

(c) $\alpha'(x) \in \mathscr{R}[a, b]$

Then

$$f \in \mathscr{R}_\alpha[a, b] \Leftrightarrow f\alpha' \in R[a, b] \qquad (*)$$

If (*) holds, then

$$\int_a^b f \mathrm{d}\alpha = \int_a^b f\alpha' \mathrm{d}x \tag{**}$$

Proof. (i) It suffices to show that $\overline{\int_a^b} f \mathrm{d}\alpha = \overline{\int_a^b} f\alpha' \mathrm{d}x$ and $\underline{\int_a^b} f \mathrm{d}\alpha = \underline{\int_a^b} f\alpha' \mathrm{d}x$.

(ii) Let $\varepsilon > 0$. Since $\alpha' \in \mathscr{R}$, there exists P s.t. $U(P,\alpha') - L(P,\alpha') < \varepsilon$ by Theorem 6.6.
Note this and the rest of the proof hold for any refinement of P. $U(P,\alpha') - L(P,\alpha') = \sum_{i=1}^{n}(A_i - a_i)\Delta x_i$, where $A_i = \sup\{\alpha'(x) : x \in [x_{i-1}, x_i]\}$ and $a_i = \inf\{\alpha'(x) : x_{i-1} \le x \le x_i\}$.
By the mean value theorem,

$$\exists t_1 \in [x_{i-1}, x_i] \text{ s.t. } \Delta\alpha_i = \alpha'(t_i) \cdot \Delta x_i.$$

For any $s_i \in [x_{i-1}, x_i]$, $|\alpha'(s_i) - \alpha'(t_i)| \le A_i - a_i$, so

$$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)| \, \Delta x_i < \varepsilon.$$

[cf. Theorem 6.7(b)]

(iii) For any $s_i \in [x_{i-1}, x_i]$,

$$\left| \sum_{i=1}^{n} f(s_i) \cdot \underbrace{\Delta\alpha_i}_{\alpha'(t_i)\Delta x_i} - \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i \right| = \left| \sum_{i=1}^{n} f(s_i)[\alpha'(t_i) - \alpha'(s_i)]\Delta x_i \right| < M\varepsilon.$$

Therefore,

$$\sum_{i=1}^{n} f(s_i)\Delta\alpha_i \le \sum_{i=1}^{n} f(s_i)\alpha'(s_i)\Delta x_i + M\varepsilon \le U(P, f\alpha') + M\varepsilon.$$

Taking supremum over s_i's, $U(P,f,\alpha) \le U(P,f\alpha') + M\varepsilon$, so

$$\begin{aligned} \overline{\int_a^b} f \mathrm{d}\alpha &\le U(P,f,\alpha) \\ &\le U(P,f\alpha') + M\varepsilon. \end{aligned}$$

As $\inf_P U(P,f,\alpha) \le \inf_P U(P,f\alpha')$,

$$\overline{\int_a^b} f \mathrm{d}\alpha \le \overline{\int_a^b} f\alpha' \mathrm{d}x + M\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\overline{\int_a^b} f \mathrm{d}\alpha \le \overline{\int_a^b} f\alpha' \mathrm{d}x.$$

Similarly, $\underline{\int_a^b} f \mathrm{d}\alpha \ge \underline{\int_a^b} f\alpha' \mathrm{d}x$.

Q.E.D.

Theorem 6.19. [Change of Variables]
Suppose $\varphi : [A,B] \to [a,b]$ is strictly increasing, continuous, and onto. Suppose α is increasing on $[a,b]$ and $f \in \mathscr{R}_\alpha[a,b]$. Let $g = f \circ \varphi : [A,B] \to \mathbb{R}, \beta = \alpha \circ \varphi : [A,B] \to \mathbb{R}$. Then $g \in \mathscr{R}_\beta[A,B]$ and $\int_A^B g \mathrm{d}\beta = \int_a^b f \mathrm{d}\alpha$.

Note. This is the change of variables formula for Riemann-Stieltjes integrals. It generalizes the calculus formula. $\int_a^b f(x)\mathrm{d}x = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(y))\varphi'(y)\mathrm{d}y$. Here $\alpha(x) = x$, so $\beta = \varphi$ and $\mathrm{d}\beta = \varphi'(y)\mathrm{d}y$

Proof. Partition P of $[a,b]$ and Q of $[A,B]$ are in one-to-one correspondence via $x_i = \varphi(y_i)$. $g([y_{i-1}, y_i]) = f([x_{i-1}, x_i])$ and $\alpha(x_i) = (\alpha \circ \varphi)(y_i) = \beta(y_i)$, so $U(P,f,\alpha) = U(Q,g,\beta)$, and $L(P,f,\alpha) = L(Q,g,\beta)$. Let $\varepsilon > 0$. Since $f \in \mathscr{R}_\alpha[a,b]$, there exists P s.t. $U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon$, so $U(Q,g,\beta) - L(Q,g,\beta) < \varepsilon$, and $g \in \mathscr{R}_\beta[A,B]$. Also, $\int_A^B g\mathrm{d}\beta = \inf_Q U(Q,g,\beta) = \inf_P (P,f,\alpha) = \int_a^b f\mathrm{d}\alpha$. Q.E.D.

Example. *$\int_a^b \sin x^2 \mathrm{d}x$ for $0 \le a < b$. Here, $f(x) = \sin x^2, \alpha(x) = x$. Let $x^2 = y$, so $x = \varphi(y) = \sqrt{y}, \varphi^{-1}(y) = y^2$. Then $\varphi : \underbrace{[a^2, b^2]}_{=[A,B]} \to [a,b]$ is continuous, strictly increasing, and onto. $g(y) = (f \circ \varphi)(y) = \sin y$. $\beta(y) = (\alpha \circ \varphi)(y) = \sqrt{y}$. Theorem 6.19 gives*

$$\int_a^b \sin x^2 \mathrm{d}\mathrm{d}x = \int_{a^2}^{b^2} \sin y \mathrm{d}\beta = \int_{a^2}^{b^2} \sin y \frac{1}{2\sqrt{y}} \mathrm{d}y,$$

where the last equality follows from the Theorem 6.17 as $\beta' = \dfrac{1}{2\sqrt{y}} \in \mathscr{R}[a^2, b^2]$. Hence,

$$\int_a^b \sin x^2 \mathrm{d}x = \int_{a^2}^{b^2} \frac{\sin y}{2\sqrt{y}} \mathrm{d}y.$$

Theorem 6.20. Let $f \in \mathscr{R}[a,b]$ and for $x \in [a,b]$, define $F(x) = \int_a^x f(t)\mathrm{d}t$. Then F is continuous on $[a,b]$, and if f is continuous at $x_0 \in [a,b]$, then $F'(x_0)$ exists and $F'(x_0) = f(x_0)$.

Proof. **Continuity of F** : Choose M s.t. $|f(t)| \leq M$ for all $t \in [a,b]$. For $a \leq x < y \leq b$, $|F(y) - F(x)| = \left|\int_a^y f(t)\mathrm{d}t - \int_a^x f(t)\mathrm{d}t\right| = \left|\int_x^y f(t)\mathrm{d}t\right| \leq M\,|y-x|$. Let $\varepsilon > 0$. If $|y-x| < \delta = \dfrac{\varepsilon}{M}$, then $|F(y) - F(x)| < M \cdot \dfrac{\varepsilon}{M} = \varepsilon$. Therefore, F is continuous on $[a,b]$.

Differentiability of F and $F'(x_0) = f(x_0)$: Let $h > 0$.

$$\frac{[F(x_0+h) - F(x_0)}{h} - f(x_0) = \frac{1}{h}\left[\int_{x_0}^{x_0+h} f(t)\mathrm{d}t\right] - f(x_0) = \frac{1}{h}\int_{x_0}^{x_0+h}[f(t) - f(x_0)]\mathrm{d}t.$$

As f is continuous at x_0, $\forall \varepsilon > 0 : \exists \delta > 0$ s.t. $|t - x_0| < \delta \Rightarrow |f(t) - f(x_0)| < \varepsilon$. Thus, if $h < \delta$,

$$\left|\frac{[F(x_0+h) - F(x_0)}{h} - f(x_0)\right| = \left|\frac{1}{h}\int_{x_0}^{x_0+h} \underbrace{[f(t) - f(x_0)]}_{\in(-\varepsilon,\varepsilon)}\mathrm{d}t\right| \leq \frac{1}{h}\cdot h\varepsilon = \varepsilon$$

Q.E.D.

Theorem 6.21. [Fundamental Theorem of Calculus] If $f \in \mathscr{R}[a,b]$ and if $\exists$ differentiable function F on $[a,b]$ s.t. $F' = f$, then $\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$.

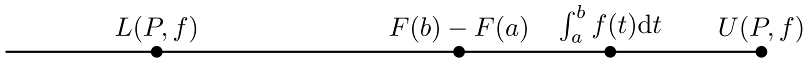
Proof. For any partition P, $F(b) - F(a) = \sum_{i=1}^{n} [F(x_i) - F(x_{i-1})]$, and by the mean value theorem 5.10,

$$\forall i \in \{1, 2, \ldots, n\} : \exists t_i \in [x_{i-1}, x_i] \text{ s.t. } F(x_i) - F(x_{i-1}) = F'(t_i)\Delta x_i = f(t_i)\Delta x_i.$$

Therefore,

$$F(b) - F(a) = \sum_{i=1}^{n} F'(t_i)\Delta x_i = \left[\sum_{i=1}^{n} \underbrace{f(t_i)}_{\in[m_i, M_i]} \Delta x_i\right] \in [L(P, f), U(P, f)].$$

Also, $\int_a^b f(t)\mathrm{d}t \in [L(P, f), U(P, f)]$.



Let $\varepsilon > 0$. Choose P s.t. $U(P, f) - L(P, f) < \varepsilon$. Then

$$\left|[F(b) - F(a)] - \int_a^b f(t)\mathrm{d}t\right| < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\left|[F(b) - F(a)] - \int_a^b f(t)\mathrm{d}t\right| = 0.$ Q.E.D.

Theorem 6.22. [Integration by Parts] If F, G are differentiable on $[a, b]$ and $F' = f, G' = g \in \mathscr{R}[a, b]$, then

$$\int_a^b F(x)g(x)\mathrm{d}x = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\mathrm{d}x.$$

Moreover, if F, G are (monotone) increasing, on $[a, b]$, then

$$\int_a^b F\mathrm{d}G = FG|_a^b - \int_a^b G\mathrm{d}F.$$

Proof. Let $H(x) = F(x)G(x)$. Then $H' = F'G + FG' = fG + Fg \in \mathscr{R}[a,b]$, so

$$\int_a^b H'(x)\mathrm{d}x = \int_a^b f(x)G(x)\mathrm{d}x + \int_a^b F(x)g(x)\mathrm{d}x.$$

By Theorem 6.21 to H,

$$\begin{aligned}\int_a^b H'(x)\mathrm{d}x = H(b) - H(a) &= F(b)G(b) - F(a)G(a)\\ &= \int_a^b f(x)G(x)\mathrm{d}x + \int_a^b F(x)g(x)\mathrm{d}x\end{aligned}$$

$$\therefore \int_a^b F(x)g(x)\mathrm{d}x = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\mathrm{d}x.$$

Q.E.D.

Remark. See problem 6.17 for a version with α.

Definition (Limits of Integration)**.** If $f \in \mathscr{R}[a,b]$ for all $b > 0$, then we define $\int_a^\infty f(x)\mathrm{d}x = \lim\limits_{b\to\infty} \int_a^b f(x)\mathrm{d}x$ if the limit exists (in $(-\infty,\infty)$) and we say $\int_a^\infty f(x)\mathrm{d}x$ converges. If $\int_a^\infty |f(x)|\,\mathrm{d}x$ converges, then we say the integral converges absolutely.

Example (1)**.** *Prove that* $\int_0^\infty \sin t^2 \mathrm{d}t$ *converges but not absolutely. [*$\int_0^x \sin t^2\mathrm{d}t$ *is a Fresnel integral and* $\int_0^\infty \sin t^2\mathrm{d}t = \sqrt{\frac{\pi}{8}}$*, as can be shown by contour integration.]*

Proof. Let $I_x = \int_0^x \sin t^2\mathrm{d}t$ for $x > 0$.

Proof of Convergence :

(a) Claim: I_n is a Cauchy Sequence ($n \in \mathbb{N}$), so it has a limit in $\mathbb{R}$.

Proof. For $0 < x < y$, $\int_x^y \sin t^2 \mathrm{d}t = \int_{x^2}^{y^2} \sin u \cdot \frac{1}{2\sqrt{u}} \mathrm{d}u$. Using Theorem 6.22,

$$\int_{x^2}^{y^2} \underbrace{\frac{1}{2\sqrt{u}}}_{F(u)} \cdot \underbrace{\sin u \mathrm{d}u}_{\mathrm{d}G(u)} = FG\Big|_{x^2}^{y^2} - \int_{x^2}^{y^2} G\mathrm{d}F$$

$$= \frac{\cos y^2}{2y} + \frac{\cos x^2}{2x} - \int_{x^2}^{y^2} \cos u \frac{1}{4u^{3/2}} \mathrm{d}u$$

$$\therefore |I_x - I_y| \leq \frac{1}{2y} + \frac{1}{2x} + \underbrace{\int_{x^2}^{y^2} \frac{1}{4u^{3/2}} \mathrm{d}u}_{-\frac{1}{2u^{1/2}}\Big|_{x^2}^{y^2} = -\frac{1}{2y} + \frac{1}{2x}}$$

$$= \frac{1}{x}.$$

In particular, for $n \geq m$, $|I_n - I_m| \leq \frac{1}{m}$, so (I_n) is a Cauchy sequence. Therefore, $\exists I \in \mathbb{R}$ s.t. $I_n \to I$ as $n \to \infty$. Q.E.D.

(b) Let $\varepsilon > 0$. Choose $N_0 > \frac{1}{\varepsilon}$, so that $N \geq N_0 \Rightarrow |I_N - I| < \varepsilon$. Let $b > N_0$. Choose N s.t. $b \in [N, N+1)$. As $N_0 \leq N$

$$I_b = \int_0^b \sin t^2 \mathrm{d}t = I_N + \int_N^b \sin t^2 \mathrm{d}t \leq I_N + \frac{1}{N}.$$

Hence,

$$|I - I_b| \leq |I - I_N| + \frac{1}{N} < \frac{1}{N} + \frac{1}{N} = \frac{2}{N} \leq \frac{2}{N_0} < 2\varepsilon.$$

$$\therefore \int_0^\infty \sin t^2 \mathrm{d}t = I.$$

Failure of Absolute Convergence :

Proof. For $n \geq 0$, let $[A_n, B_n] = [2n\pi, (2n+1)\pi]$, and let $a_n^2 = A_n, b_n^2 = B_n$.
Then $\int_0^{b_N} \left|\sin t^2\right| \mathrm{d}t \geq \sum_{n=0}^{N} \int_{a_n}^{b_n} \sin t^2 \mathrm{d}t$. Now

$$\int_{a_n}^{b_n} \sin t^2 \mathrm{d}t = \int_{A_n}^{B_n} \sin u \cdot \frac{1}{2\sqrt{u}} \mathrm{d}u \geq \frac{1}{2\sqrt{B_n}} \cdot \int_{A_n}^{B_n} \sin u \mathrm{d}u = \frac{1}{\sqrt{B_n}}.$$

$$\therefore \int_0^{b_N} \left|\sin t^2\right| \mathrm{d}t \geq \sum_{n=0}^{N} \frac{1}{\sqrt{(2n+1)\pi}} = \frac{1}{\sqrt{\pi}} \cdot \underbrace{\sum_{n=0}^{N} \frac{1}{\sqrt{2n+1}}}_{\to\infty \text{ as } N\to\infty}.$$

Hence,

$$\int_0^{\infty} \left|\sin t^2\right| \mathrm{d}t = \infty.$$

Q.E.D.

Q.E.D.

Note. Material for MT 1 ends here, including A1,A2,A3.

Chapter 7

Sequences and Series of Functions

Example. *Bad behaviour of limits*

(1) *For $m, n \in \mathbb{N}$, let $p_{m,n} = \frac{m}{n}$. Then*

$$\lim_{m\to\infty} p_{m,n} = \infty$$

and

$$\lim_{n\to\infty} p_{m,n} = 0.$$

Hence,

$$\lim_{n\to\infty} \underbrace{\lim_{m\to\infty} p_{m,n}}_{=\infty} = \infty \neq \lim_{m\to\infty} \underbrace{\lim_{n\to\infty} p_{m,n}}_{=0}.$$

This shows the order of limits matters.

(2) *Let*

$$f_n(x) = \begin{cases} 1 & x \geq 0 \\ 1 + nx & -\frac{1}{n} < x < 0 \\ 0 & x \leq -\frac{1}{n} \end{cases}.$$

Then $f_n(x)$ is a continuous function. Then $\lim_{n\to\infty} f_n(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$, which is not continuous at 0. This shows that the limit of continuous functions need not be continuous. Moreover,

$$\lim_{n\to\infty} \underbrace{\lim_{x\to 0} f_n(x)}_{=1} = 1,$$

while
$$\lim_{x\to 0}\underbrace{\lim_{n\to\infty} f_n(x)}_{=f(x)}$$
does not exist. This shows again that the order of limits matters.

(3) *For $x \in [0,1]$, let*
$$f_n(x) = \begin{cases} 1 & n!\cdot x \in \mathbb{Z} \\ 0 & n!\cdot x \notin \mathbb{Z} \end{cases}.$$
Each $f_n \in \mathscr{R}[0,1]$ by Theorem 6.10. However,
$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 1 & x \in Q\cap[0,1] \\ 0 & x \notin Q\cap[0,1] \end{cases}.$$
Hence, $f(x)$ is nowhere continuous, and $f \notin \mathscr{R}[0,1]$.

(4) *Let*
$$f_n(x) = \begin{cases} 0 & |x| \geq \frac{1}{n} \\ n^2x+n & -\frac{1}{n} < x < 0 \\ -n^2x+n & 0 < x < \frac{1}{n} \\ 0 & x = 0 \end{cases}.$$
Ten $f(x) = \lim\limits_{n\to\infty} f_n(x) = 0$ for all $x \in \mathbb{R}$. Moreover,
$$\forall n : \int_{-1}^{1} f_n(x)\mathrm{d}x = 1,$$
$$\int_{-1}^{1} f(x)\mathrm{d}x = 0.$$
Hence, $\lim\limits_{n\to\infty}\int_{-1}^{1} f_n(x)\mathrm{d}x = 1 \neq \int_{-1}^{1}\left(\lim\limits_{n\to\infty} f_n(x)\right)\mathrm{d}x$

(5) *Let*
$$f_n(x) = \frac{\sin nx}{\sqrt{n}} \text{ for } n \in \mathbb{N}, x \in \mathbb{R}.$$
Let
$$f(x) = \lim_{n\to\infty} f_n(x) = 0,$$
so
$$\forall x \in \mathbb{R} : f'(x) = 0.$$
However,
$$f_n'(x) = \frac{n\cos x}{\sqrt{n}} = \sqrt{n}\cos nx.$$
Therefore, $f_n'(\pi) = \sqrt{n}(-1)^n$ diverges as $n \to \infty$. Hence,
$$\underbrace{f'(\pi) = \left(\lim_{n\to\infty} f_n\right)'(\pi)}_{=0} \neq \underbrace{\lim_{n\to\infty} f_n'(\pi)}_{DNE}.$$

These examples show bad behaviour under interchange of limits, which suggests the need of a stronger notion of convergence.

Definition 7.7. Let E be any set and $f_n : E \to \mathbb{R}($ or $\mathbb{C})$ for $n \in \mathbb{N}$. Then f_n converges uniformly to f on E if

$$\forall \varepsilon > 0 : \exists N \text{ s.t. } \forall n \geq N : \forall x \in E : |f_n(x) - f(x)| < \varepsilon.$$

Example. *(a) Consider the example (2)*

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

$$f_n(x) - f(x) = \begin{cases} 1 + nx & -\frac{1}{n} < x < 0 \\ 0 & \textit{otherwise} \end{cases}.$$

In particular, $f_n(-\frac{1}{2n}) - f(-\frac{1}{2n}) = \frac{1}{2}$, *so we cannot choose* N *s.t.* $n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon = \frac{1}{4}$ *for all* x.

$$\therefore f_n \textit{ does not converge uniformly to } f \textit{ on } \mathbb{R}.$$

(b) Consider the example (5).

$$f(x) = \lim_{n\to\infty} f_n(x) = 0,$$

$$\forall x \in \mathbb{R} : f'(x) = 0.$$

Then

$$\left| f_n(x) - \underbrace{f(x)}_{=0} \right| = \left| \frac{\sin nx}{\sqrt{n}} \right| \leq \frac{1}{\sqrt{n}},$$

so $f_n \to f$ *uniformly on* $\mathbb{R}$. *[Note: uniform convergence is not enough for* $\lim_{n\to\infty} f_n' = (\lim_{n\to\infty} f_n)'$.*]*

Theorem 7.8. [Cauchy Criteria for Uniform Convergence]
f_n converges uniformly to f on E if and only if

$$\forall \varepsilon > 0 : \exists N \text{ s.t. } m, n \geq \mathbb{N} \Rightarrow \forall x \in E : |f_n(x) - f_m(x)| < \varepsilon.$$

That is, we can choose such N independent of x.

Proof. ($\Rightarrow$) Suppose f_n converges uniformly to f on E. Then

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

For $\varepsilon > 0$, choose N s.t. $\forall n \geq N : \forall x \in E : |f_n(x) - f(x)| < \frac{\varepsilon}{2}$. Then

$$\forall m, n \geq N : \forall x \in E : |f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \varepsilon.$$

($\Leftarrow$) Let $x \in E$. $\{f_n(x)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{C}$, so has a limit $f(x) = \lim\limits_{n\to\infty} f_n(x)$. To check uniformity, let $\varepsilon > 0$. We know that $\exists N$ s.t. $|f_n(x) - f_m(x)| < \varepsilon$ if $n, m \geq N$ for all $x \in E$. Let $m \to \infty$. Then $|f_n(x) - f(x)| \leq \varepsilon$ if $n \geq N$ for all $x \in E$.

Q.E.D.

Definition. $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on E if $s_n(x) = \sum_{i=1}^{n} f_i(x)$ is a uniformly convergence sequence of functions.

Theorem 7.10. [Weierstass M-test] If $|f_n(x)| \leq M_n$ for all $n \geq N_0$ and all $x \in E$ and if $\sum_{n=N_0}^{\infty} M_n < \infty$, then $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on E.

Proof. Let $s_n(x) = \sum_{i=1}^{n} f_i(x)$. For $n > m \geq N_0$,

$$\forall x \in E : |s_n(x) - s_m(x)| = \left| \sum_{i=m+1}^{n} f_i(x) \right| \leq \sum_{i=m+1}^{n} M_i.$$

Let $\varepsilon > 0$. Choose $N \geq N_0$ s.t. $\sum_{i=N+1}^{\infty} M_i < \varepsilon$. Then $|s_n(x) - s_m(x)| < \varepsilon$ if $n > m \geq N$ for all $x \in E$. Hence, s_n converges uniformly on E by Theorem 7.8. Q.E.D.

Theorem 7.11. Let $E \subset X$ and $f_n : E \to \mathbb{R}($ or $\mathbb{C})$, $n \in \mathbb{N}$. Suppose $f_n \to f$ uniformly on E. Let $x \in E'$ and suppose $\lim\limits_{t\to x} f_n(t) = A_n$ exists for each n. Then $A_n \to A$ for some A and $\lim\limits_{t\to x} f(t) = A$; i.e.,

$$\lim_{t\to x} \underbrace{\lim_{n\to\infty} f_n(t)}_{f(t)} = \lim_{n\to\infty} \underbrace{\lim_{t\to x} f_n(t)}_{A_n} = A.$$

Proof. $A_n \to A$ for some A:

It suffices to show that $\{A_n\}$ is a Cauchy sequence. Since $f_n \to f$ uniformly on E, for any $\varepsilon > 0$, we can choose N s.t. $m, n \geq N \Rightarrow |f_m(t) - f_n(t)| < \varepsilon$ for all $t \in E$. Let $t \to x$. $|A_m - A_n| < \varepsilon$ if $m, n \geq N$. Therefore, $\{A_n\}$ is a Cauchy sequence and converges to some A by completeness of $\mathbb{R}$(or $\mathbb{C}$).

$f(t) \to A$ as $t \to x$:

For $t \in E$ and $n \in \mathbb{N}$,

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| \qquad (*)$$

Let $\varepsilon > 0$. Since $f_n \to f$ uniformly $\exists N_1$ s.t. $|f(t) - f_n(t)| < \frac{\varepsilon}{3}$ if $n \geq N_1$ for all $t \in E$.
Since $A_n \to A$, $\exists N_2$ s.t. $|A_n - A| < \frac{\varepsilon}{3}$ if $n \geq N_2$.
Let $N = \max\{N_1, N_2\}$ and use $n = N$ in (*). Then

$$\forall t \in E : |f(t) - A| \leq \frac{\varepsilon}{3} + |f_N(t) - A_N| + \frac{\varepsilon}{3}.$$

Since $\lim_{t \to x} f_N(t) = A_N$, there exists $\delta > 0$ s.t. $t \in N_\delta^E(x) \setminus \{x\} \Rightarrow |f_N(t) - A_N| < \frac{\varepsilon}{3}$. Then

$$t \in N_\delta^E(x) \setminus \{x\} \Rightarrow |f(t) - A| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

Hence, $\lim_{t \to x} f(t) = A$.

Q.E.D.

Corollary 7.12. If f_n is continuous on E and $f_n \to f$ uniformly on E, then f is continuous on E.

Proof. Every function is continuous at an isolated point, so only need to consider $x \in E' \cap E$.

$$\begin{aligned} f(x) &:= \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{t \to x} f_n(t) && (\because f_n \text{ continuous}) \\ &= \lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{t \to x} f(t) && (\because \text{Theorem 7.11}). \end{aligned}$$

Q.E.D.

Remark. VERY IMPORTANT

Theorem 7.13. Suppose K is compact and

(a) f_n is continuous on K for all $n \in \mathbb{N}$

(b) $f_n \to f$ pointwise on K and f is continuous on K

(c) $f_n(x) \geq f_{n+1}(x)\ \forall x \in K, \forall n \in \mathbb{N}$.

Then $f_n \to f$ uniformly on K.

Proof. Let $g_n = f_n - f$. Then

(a) g_n is continuous on K for all $n \in \mathbb{N}$

(b) $g_n \to 0$ pointwise on K

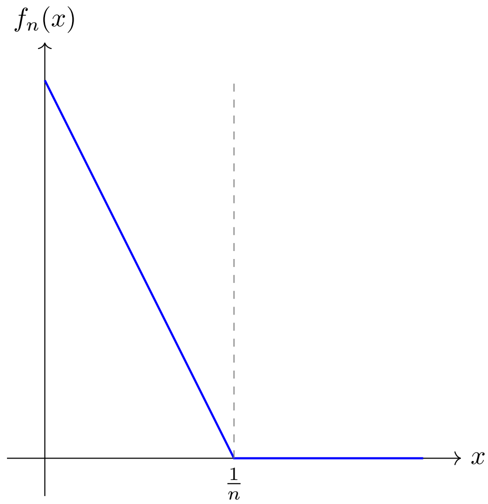
(c) $g_n(x) \geq g_{n+1}(x)\ \forall x \in K, \forall n \in \mathbb{N}$.

Goal: Prove $g_n \to 0$ uniformly on K; i.e., given $\varepsilon > 0$, $\exists N$ s.t. $\forall n \geq N : \forall x \in K : |g_n(x)| < \varepsilon$.
For this, it suffices if $\exists N$ s.t. $g_N(x) < \varepsilon$ for all $x \in K$.
Let $K_n = g_n^{-1}([\varepsilon, \infty))$.
Then the goal becomes: find N s.t. $K_N = \emptyset$, since this implies $(K_N)^c = g_N^{-1}([0, \varepsilon)) = K$; i.e., $\forall x \in K : g_N(x) < \varepsilon$.
Since g_n is continuous, K_n is closed by Theorem 4.8.
Since $K_n \subset K$, K_n is compact. Also $K_{n+1} \subset K_n$ because $g_{n+1}(x) \geq \varepsilon \Rightarrow g_n(x) \geq \varepsilon$.
Let $x \in K$. Since $g_n(x) \to 0$, $\exists N_x$ s.t. $\underbrace{x \notin K_{N_x} = g_{N_x}^{-1}([\varepsilon, \infty))}_{\because g_n(x) < \varepsilon \text{ for large } n}$ for all $n \geq N_x$. Hence, $x \notin \bigcap_{n=1}^{\infty} K_n$. x is arbitrary, so $\bigcap_{n=1}^{\infty} K_n = \emptyset$. By corollary to Theorem 2.36, this implies $\exists N$ s.t. $K_N = \emptyset$. Q.E.D.

Example. *(a) Let $K = (0, 1]$ (not compact). Let*

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

$f_n(x) \to 0$ for all $x \in K$ so (a), (b), (c) all hold, but f_n does not converge uniformly to zero function on K as K not compact.

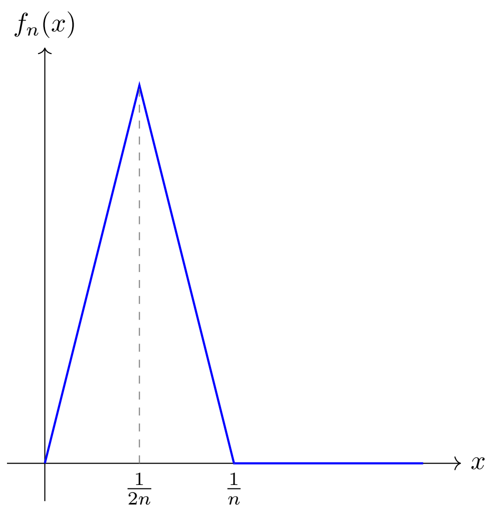


$K = (0, 1]$

(b) Let $K = [0, 1]$ *(compact). Let*

$$f_n(x) = \begin{cases} 2nx & 0 \le x \le \frac{1}{2n} \\ -2nx + 2 & \frac{1}{2n} \le x \le \frac{1}{n} \\ 0 & \frac{1}{n} < x \le 1 \end{cases}.$$

f_n *continuous,* $f_n(x) \to 0$ *for all* $x \in K$ *but not uniformly. Here, (a) and (b) hold, but (c) does not.*



$K = [0, 1]$

Definition 7.14. For a metric space X, let

$$\mathscr{C}(X) = \{f : X \to \mathbb{C} \text{ s.t. } f \text{ is continuous}\}.$$

The supremum norm of $f \in \mathscr{C}(X)$ is defined by

$$\|f\| = \sup_{x \in X}\{|f(x)|\}.$$

Notation. When $X = [a, b]$, often write $\|f\|_\infty$ instead of $\|f\|$ since $\lim\limits_{p\to\infty} \|f\|_p = \|f\|_\infty$ (C.f A3-Q2).

Proposition.

$$d(f, g) = \|f - g\| \text{ defines a metric on } \mathscr{C}(X).$$

$d(f, g) = 0 \Leftrightarrow \sup_{x\in X}\{|f(x) - g(x)|\} = 0$ and hence $f(x) = g(x)$ $\forall x \in X$; i.e., $f = g$.

(b) $d(f, g) = d(g, f)$

(c) $d(f, g) \leq d(f, h)+d(h, g)$ since $\forall x \in X : |f(x) - g(x)| \leq |f(x) - h(x)|+ |h(x) - g(x)| \leq \|f-h\|+\|h-g\|$. Hence, $\|f-g\| \leq \|f-h\|+\|h-g\|$.

As a consequence, $f_n \to f$ uniformly on X if and only if $f_n \to f$ in the metric space $(\mathscr{C}(X), \|\cdot\|)$. Proof:

$$\begin{aligned}
\text{LHS} &\Leftrightarrow \forall\varepsilon > 0 : \exists N \text{ s.t. } |f_n(x) - f(x)| < \varepsilon \text{ if } n \geq N \text{ for all } x \in X.\\
&\Leftrightarrow \forall\varepsilon > 0 : \exists N \text{ s.t. } \|f_n - f\| < \varepsilon \text{ if } n \geq N\\
&\Leftrightarrow f_n \to f \text{ in } (\mathscr{C}(X), \|\cdot\|).
\end{aligned}$$

Theorem 7.15. $\mathscr{C}$ is a complete metric space.

Proof. Let $\{f_n\}$ be a Cauchy sequence in $\mathscr{C}(X)$; i.e.,

$$\forall \varepsilon > 0 : \exists N \text{ s.t. } n, m \geq N \Rightarrow \|f_m - f_n\| < \varepsilon.$$

By Cauchy criterion (Theorem 7.8), $f_n \to f$ uniformly for some f.
Now it is sufficient to check that $f \in \mathscr{C}(X)$.
By Cor 7.12, f is continuous. Also, f is bounded since $\exists N_0$ s.t. $|f(x) - f_{N_0}(x)| < 1$ for all $x \in X$. Then

$$|f(x)| \leq |f_{N_0}(x)| + |f(x) - f_{N_0}(x)| \leq \underbrace{M_0}_{\text{bound for } f_{N_0}} +1 \text{ for all } x \in X.$$

$\therefore f \in \mathscr{C}(X)$. Q.E.D.

Theorem 7.16. Suppose $f_n \in \mathscr{R}_\alpha[a, b]$ for $n \in \mathbb{N}$ and $f_n \to f$ uniformly on $[a, b]$. Then $f \in \mathscr{R}_\alpha[a, b]$ and

$$\int_a^b f \mathrm{d}\alpha = \lim_{n \to \infty} \int_a^b f_n \mathrm{d}\alpha.$$

Proof. First, we prove $f \in \mathscr{R}_\alpha[a,b]$; i.e.,

$$\overline{\int_a^b} f \mathrm{d}\alpha = \underline{\int_a^b} f \mathrm{d}\alpha.$$

Let $\varepsilon > 0$. Since $f_n \to f$ uniformly on $[a,b]$, $\exists N$ s.t.

$$n \geq N \Rightarrow \forall x \in [a,b] : f_n(x) - \varepsilon < f(x) < f_n(x) + \varepsilon.$$

Hence,

$$\underline{\int_a^b} (f_n - \varepsilon) \mathrm{d}\alpha \leq \underline{\int_a^b} f \mathrm{d}\alpha \leq \overline{\int_a^b} f \mathrm{d}\alpha \leq \overline{\int_a^b} (f_n + \varepsilon) \mathrm{d}\alpha.$$

Therefore,

$$\overline{\int_a^b} f \mathrm{d}\alpha - \underline{\int_a^b} f \mathrm{d}\alpha \leq \int_a^b 2\varepsilon \mathrm{d}\alpha = 2\varepsilon \left[\alpha(b) - \alpha(a)\right].$$

As ε arbitrary, this implies

$$\overline{\int_a^b} f \mathrm{d}\alpha = \underline{\int_a^b} f \mathrm{d}\alpha,$$

so $f \in \mathscr{R}_\alpha[a,b]$.
Since

$$\int_a^b f \mathrm{d}\alpha - \int_a^b f_n \mathrm{d}\alpha = \int_a^b f - f_n \mathrm{d}\alpha \leq \int_a^b \underbrace{|f - f_n|}_{<\varepsilon \text{ for all } x \text{ if } n \geq N} \mathrm{d}\alpha \leq \varepsilon \int_a^b \mathrm{d}\alpha = \varepsilon \left[\alpha(b) - \alpha(a)\right],$$

We have

$$\lim_{n\to\infty} \int_a^b f_n \mathrm{d}\alpha = \int_a^b f \mathrm{d}\alpha.$$

Q.E.D.

Corollary. If $f_n \in \mathscr{R}_\alpha[a,b]$ and $f(x) = \sum_{n=1}^{\infty} f_n(x)$ converges uniformly on $[a,b]$ then

$$\int_a^b f \mathrm{d}\alpha = \sum_{n=1}^{\infty} \int_a^b f_n \mathrm{d}\alpha.$$

Proof. Let $s_n(x) = \sum_{i=1}^{n} f_i(x)$. Then $s_n \to f$ uniformly so

$$\int_a^b f \mathrm{d}\alpha = \lim_{n\to\infty} \int_a^b s_n \mathrm{d}\alpha = \lim_{n\to\infty} \sum_{i=1}^{n} \int_a^b f_i \mathrm{d}\alpha = \sum_{i=1}^{\infty} \int_a^b f_i \mathrm{d}\alpha.$$

Q.E.D.

Note. Recall the example $f_n(x) = \frac{\sin nx}{\sqrt{n}}$, $f_n \to 0$ uniformly on $\mathbb{R}$, but $f_n'(x)$ does not converge.

Notation. For a, b, $\int_b^a f \mathrm{d}\alpha = -\int_a^b f \mathrm{d}\alpha$.

Theorem 7.17. Suppose

(a) f_n is differentiable on $[a, b]$.

(b) $\exists x_0 \in [a, b]$ s.t. $f_n(x_0)$ converges, say to L_0, as $n \to \infty$.

(c) f_n' converges uniformly on $[a, b]$.

Then $\exists f$ s.t. $f_n \to f$ uniformly on $[a, b]$ and $\lim_{n\to\infty} f_n'(x) = f'(x)$ for all $x \in [a, b]$.

Remark.

1. The hypothesis (b) is needed. E.g., $f_n(x) = n$ obeys hypotheses (a),(c), but the conclusion of the theorem fails as it fails (b). We can assume $f_n(x) \to 0$ by replacing f_n by $f_n - L_0$.
2. We add a hypothesis (d) to make the proof simpler: f_n' is continuous on $[a, b]$ for all $n \in \mathbb{N}$

Proof (With hypothesis (d) added).

1. By (c), $\exists g$ s.t. $f_n' \to g$ uniformly on $[a,b]$, and hence on any subintervals of $[a,b]$. By (d), g is continuous.

2. By Theorem 7.16 (applied to $[x, x_0]$ or $[x_0, x]$ in $[a,b]$),

$$(\pm)\int_{x_0}^{x} f_n'(t)\mathrm{d}t \to \int_{x_0}^{x} g(t)\mathrm{d}t \underbrace{=}_{\text{by defn. of } f(x)} f(x).$$

By Theorem 6.21,

$$\int_{x_0}^{x} f_n'(t)\mathrm{d}t = f_n(x) - f_n(x_0)$$

$$\int_{x_0}^{x} g(t)\mathrm{d}t = f(x) - f(x_0).$$

Hence,

$$[f_n(x) - f_n(x_0)] \to f(x)$$

$$g(x) = f'(x).$$

Note we assume $f_n(x_0) \to 0$.
Then for all $x \in [a,b]$,

$$f_n(x) - f_n(x_0) \to f(x) - 0 = f(x).$$

Therefore,

$$f_n'(x) \to g(x) = f'(x)..$$

3. It remains to prove that $f_n \to f$ uniformly on $[a,b]$: Let $\varepsilon > 0$.
By (b), $\exists N_0$ s.t. $n \geq N_0 \Rightarrow |f_n(x_0) - L_0| = |f_n(x_0)| < \frac{\varepsilon}{2}$.
By (c), $\exists N_1$ s.t. $n \geq N_1 \Rightarrow |f_n'(x) - g(x)| < \frac{\varepsilon}{2(b-a)}$ for all $x \in [a,b]$.
Let $N = \max\{N_0, N_1\}$. Then N is independent of x.
For $n \geq N$,

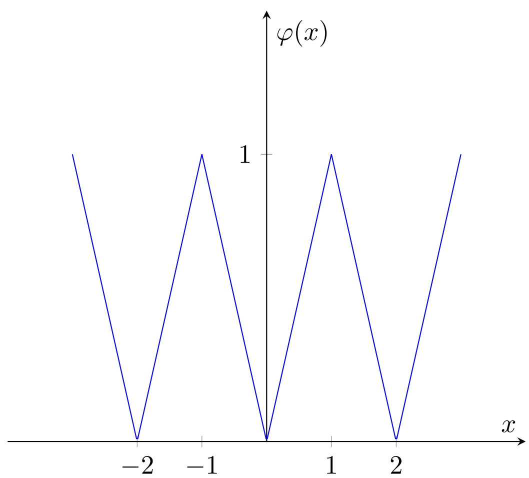
$$\begin{aligned} \forall x \in [a,b] : |f(x) - f_n(x)| &= \left| \int_{x_0}^{x} g(t)\mathrm{d}t - \left(\int_{x_0}^{x} f_n'(t)\mathrm{d}t + f_n(x_0) \right) \right| \\ &\leq \left| \int_{x_0}^{x} [g(t) - f_n'(t)]\,\mathrm{d}t \right| + |f_n(x_0)| \\ &< \left| \int_{x_0}^{x} \frac{\varepsilon}{2(b-a)}\mathrm{d}t \right| + \left| \frac{\varepsilon}{2} \right| = \int_{x_0}^{x} \frac{\varepsilon}{2(b-a)}\mathrm{d}t + \frac{\varepsilon}{2} \\ &= \frac{\varepsilon}{2(b-a)} \cdot (b-a) + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Q.E.D.

Theorem 7.18. There exists a continuous function $f : \mathbb{R} \to \mathbb{R}$ s.t. for all $x \in \mathbb{R}$, $f'(x)$ does not exist.

Proof (Theorem 7.18)**.** Let

$$\varphi : \mathbb{R} \to \mathbb{R} \text{ by } \varphi(x) = \begin{cases} |x| & |x| \leq 1 \\ \varphi(x) = \varphi(x-2) & \text{otherwise} \end{cases}.$$



Then φ is continuous on $\mathbb{R}$, and in fact, $|\varphi(x) - \varphi(y)| \leq |x - y|$ for all $x, y \in \mathbb{R}$ (Lipschitz bound).
Let

$$f_n(x) = \sum_{k=0}^{n} \left(\frac{3}{4}\right)^n \varphi(4^n x)$$

$$f(x) = \lim_{n \to \infty} f_n(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x).$$

Then the series $f(x)$ converges uniformly on $\mathbb{R}$ by the Weierstrass M-test (Theorem 7.10) since $\left|\left(\frac{3}{4}\right)^n \varphi(4^n x)\right| \leq \left(\frac{3}{4}\right)^n$ and $\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n < \infty$.

Also, f is continuous on $\mathbb{R}$ by Corollary 7.12.
Claim: $f'(x)$ does not exist for any $x \in \mathbb{R}$.

It suffices to show that for $x \in \mathbb{R}$, there exists $\delta_m \to 0$ s.t.

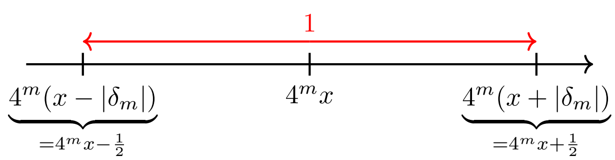
$$\left| \underbrace{\frac{f(x+\delta_m) - f(x)}{\delta_m}}_{=\sum_{n=0}^{\infty} \frac{3}{4}^n \gamma_n, \gamma_n = \frac{\varphi(4^n(x+\delta_m)) - \varphi(4^n x)}{\delta_m}} \right| \to \infty \text{ as } m \to \infty.$$

Note

$$4^m\left(x+\delta_m\right) = 4^m x - \frac{1}{2},$$

$$4^m\left(x+1-\delta_m\right) = 4^m x + \frac{1}{2}.$$

Choose $\delta_m = \pm\frac{1}{2} \cdot 4^{-m}$ with sign chosen (depending on x) so that there is no integer between $4^m x$ and $4^m(x+\delta_m)$.



Then $\gamma_n = 0$ if $n \neq m$ and $\gamma_m = \frac{\varphi(4^m(x+4^{-m})) - \varphi(4^m x)}{4^{-m}}$. In particular,

$$|\varphi(4^m(x+\delta_m)) - \varphi(4^m x)| = |4^m \delta_m| = \frac{1}{2}.$$

If $n > m$, then

$$\varphi(4^n(x+\delta_m)) - \varphi(4^m x) = \varphi(4^n x \pm \underbrace{4^n \cdot \frac{1}{2} \cdot \frac{1}{4^m}}_{\text{even integer}}) - \varphi(4^m x) = 0.$$

Hence, $\gamma_n = 0$ for $n > m$.
Since $\varphi(x) - \varphi(y) \leq |x-y|$, for $n \leq m$,

$$\gamma_n \leq \frac{1}{|\delta_m|} \cdot |4^m \delta_m| = 4^m.$$

In fact, $|\gamma_m| = \frac{1}{|\delta_m|} \cdot \frac{1}{2} = 4^m$.
Therefore,

$$\left| \frac{f(x+\delta_m) - f(x)}{\delta_m} \right| = \left| \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \gamma_n \right| = \left| \sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_n \right|$$

$$\geq \left(\frac{3}{4}\right)^m |\gamma_m| - \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n |\gamma_n|$$

$$\geq \left(\frac{3}{4}\right)^m \cdot 4^m - \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n 4^n$$

$$= 3^m - \frac{3^m - 1}{3 - 1} = \frac{1}{2}(3^m + 1) \to \infty \text{ as } m \to \infty.$$

Q.E.D.

Definition (Equicontinuity)**.** A family $\mathscr{F}$ of functions on E is equicontinuous if $\forall \varepsilon > 0 : \exists \delta > 0$ s.t. $f \in \mathscr{F}; x, y \in E; d(x, y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Remark. (a) If $\mathscr{F}$ is equicontinuous, then every $f \in \mathscr{F}$ is uniformly continuous on E.

(b) Any finite set of uniformly continuous functions is equicontinuous.

Example. *Let $\mathscr{F} = \{f_1, f_2, \ldots\}$, where $f_n(x) = \frac{\sin nx}{\sqrt{n}}$, $x \in E = [0, 1]$. We have seen $f_n \to 0$ uniformly on E.*
Claim $\mathscr{F}$ is equicontinuous.
Let $\varepsilon > 0$. Choose N s.t. $\frac{2}{\sqrt{n}} < \varepsilon$ for all $n \geq N$.
Since $\{f_1, f_2, \ldots f_{N-1}\}$ is a finite set of uniformly continuous functions, it is equicontinuous.
Hence, $\exists \delta > 0$ s.t. $n < N$ and $|x - y| < \delta \Rightarrow |f_n(x) - f_n(y)| < \varepsilon$.
Therefore, $n \in \mathbb{N}$ and $|x - y| < \delta \Rightarrow |f_n(x) - f_n(y) <| \varepsilon$.

Problem (7.16)**.** Let $\{f_n\}$ be an equicontinuous sequence of functions $f_n : K \to \mathbb{C}$, where K is compact.
Suppose $\lim_{n\to\infty} f_n(x) = f(x)$ exists for all $x \in K$. Then $f_n \to f$ uniformly on K.

Proof. We use $\frac{\varepsilon}{3}$ argument.
Let $\varepsilon > 0$. $\exists \delta > 0$ s.t. $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$ for all $x, y \in K$ s.t. $d(x, y) < \delta$ and $n \in \mathbb{N}$.
Since K is compact, the open cover $\{N_\delta(x)\}_{x \in K}$ has a finite subcover $\{N_\delta(x_1), \ldots, N_\delta(x_k)\}$.
Thus, given any $x \in K$, $\exists j$ s.t. $d(x, x_j) < \delta$. Then,

$$|f_n(x) - f_m(x)| \leq \underbrace{|f_n(x) - f_n(x_j)|}_{< \frac{\varepsilon}{3}} + |f_n(x_j) - f_m(x_j)| + \underbrace{|f_m(x_j) - f_m(x)|}_{< \frac{\varepsilon}{3}}.$$

For each $i = 1, \ldots, k$, we know $\{f_n(x_i)\}_n$ is a convergent sequence, so it's a Cauchy sequence. Hence, $\exists N_i$ s.t. $m, n \geq N_i \Rightarrow |f_m(x_i) - f_n(x_i)| < \frac{\varepsilon}{3}$.

Let $N = \max\{N_1, N_2, \dots, N_k\}$. Then

$$m, n \geq N \Rightarrow |f_m(x) - f_n(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Therefore, $f_n \to f$ uniformly on K. Q.E.D.

Theorem 7.24. If $f_n : K \to \mathbb{C}$ are continuous and K is compact and $f_n \to f$ uniformly on K, then $\{f_n\}$ is equicontinuous.

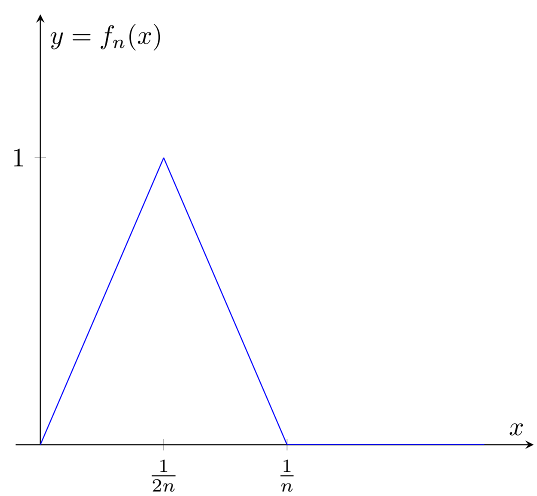
Proof. (Using $\frac{\varepsilon}{3}$ argument)
Let $\varepsilon > 0$. Since $f_n \to f$ uniformly, $\exists N$ s.t. $m, n \geq N \Rightarrow |f_m(x) - f_n(x)| < \frac{\varepsilon}{3}$ for all $x \in K$. Since K is compact, each f_n is uniformly continuous.
Hence, $\{f_1, \dots, f_N\}$ is equicontinuous, so $\exists \delta > 0$ s.t. $|f_i(x) - f_i(y)| < \frac{\varepsilon}{3}$ if $d(x, y) < \delta$ for all $i \in \{1, 2, \dots, N\}$.
For $n \geq N$,

$$|f_n(x) - f_n(y)| \leq \underbrace{|f_n(x) - f_N(x)|}_{<\frac{\varepsilon}{3}} + \underbrace{|f_N(x) - f_N(y)|}_{<\frac{\varepsilon}{3} \text{ if } d(x,y)<\delta} + \underbrace{|f_N(y) - f_n(y)|}_{<\frac{\varepsilon}{3}} < \varepsilon \text{ if } d(x, y) < \delta.$$

Q.E.D.

Example. *Let*

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



$\{f_n\}$ *obeys* $|f_n(x)| \leq 1$ *for all* $n \in \mathbb{N}, x \in [0, 1]$. *It is not equicontinuous because* $\left|f_n(\frac{1}{2n}) - f_n(0)\right| = 1 - 0 = 1$ *for all* n. *Also, no subsequence of* $\{f_n\}$ *can* *converge uniformly since* $f_n(\frac{1}{2n}) = 1$ *whereas* $\lim\limits_{n\to\infty} f_n(x) = 0$ *for all* $x \in [0,1]$.

Definition (7.19). (a) $\{f_n\}$ is pointwise bounded on E if $\exists \varphi : E \to \mathbb{R}$ s.t. $|f_n(x)| < \varphi(x)$ for all $n \in \mathbb{N}, x \in E$.
Note: $<$ can be replaced with $\leq$ but Rudin uses strict inequality

(b) $\{f_n\}$ is uniformly bounded on E if $\exists M > 0$ s.t. $|f_n(x)| < M$ for all $n \in \mathbb{N}, x \in E$.

Example. *Let*

$$f_n(x) = \frac{1}{x} + \frac{1}{n}, n \in \mathbb{N}, x \in E = (0,1]..$$

Then $f_n(x)$ *is pointwise bounded on* E *since* $|f_n(x)| < \frac{1}{x} + 2$ *for all* $n \in \mathbb{N}, x \in E$. *However,* $\{f_n\}$ *is not uniformly bounded on* E *since* $f_n(x) > \frac{1}{x}$ *and* $\frac{1}{x} \to \infty$ *as* $x \to 0^+$.

Theorem 7.23. [Selection Theorem] Suppose E is a countable set and $f_n : E \to \mathbb{C}$ is pointwise bounded. Then there exists $\{f_{n_k}\}$ of $\{f_n\}$ which is pointwise convergent on E; i.e.,

$$\lim_{k\to\infty} \{f_{n_k}(x)\} \text{ exists for all } x \in E.$$

Proof. (Using diagonal argument)
Let $E = \{x_1, x_2, x_3, \dots\}$.
Since $\{f_n(x_1)\}_n$ is bounded, $\exists$ subsequence $\{f_{1,k}\}$ s.t. $\lim_{k\to\infty} f_{1,k}(x_1)$ exists by Weierstrass theorem 2.42.
Successive subsequences can be constructed as follows:

$$
\begin{aligned}
&S_1 : f_{1,1}, f_{1,2}, f_{1,3}, f_{1,4}, \dots \text{ converges on } x_1 \\
&S_2 : f_{2,1}, f_{2,2}, f_{2,3}, f_{2,4}, \dots \text{ converges on } x_1, x_2 \\
&S_3 : f_{3,1}, f_{3,2}, f_{3,3}, f_{3,4}, \dots \text{ converges on } x_1, x_2, x_3 \\
&\quad \vdots \\
&\quad .
\end{aligned}
$$

Note we can make S_n converges not just on x_n but also for $\{x_1, x_2, \dots, x_{n-1}\}$ by taking a subsequence of S_{n-1}.
This is because $f_{n-1,k}$ is also a pointwise bounded sequence itself.
Form a diagonal subsequence $S : f_{1,1}, f_{2,2}, f_{3,3}, f_{4,4}, \dots$. Then S is eventually a subsequence of each S_j, so it converges on $x_1, x_2, x_3, \dots, x_j$.
This is true for all j, so $\{f_{n,n}(x_i)\}_n$ converges for each $x_i \in E$.
Q.E.D.

Lemma. If K is compact then K has a countable dense subset $E \subset K$; i.e., $\overline{E} = K$, or $\forall p \in K : \forall r > 0 : \exists x \in E$ s.t. $d(p, x) < r$. We also say K is separable. See Problem 2.25.

Proof. For $n \in \mathbb{N}$, $\{N_{\frac{1}{n}}(p)\}_{p\in K}$ is an open cover of K, so $\exists$ finite subcover $\{N_{\frac{1}{n}}(p)\}_{p\in E_n}$, $E_n \subset K$, E_n finite. Let $E = \bigcup_{n=1}^{\infty} E_n$, a countable set.
Let $p \in K$ and $r > 0$. Choose n_0 s.t. $\frac{1}{n_0} < r$.
By definition of E_{n_0}, there exists $x_0 \in E_{n_0}$ s.t. $p \in N_{\frac{1}{n_0}}(x_0)$.
Then $d(p, x_0) < \frac{1}{n_0} < r$, so E is dense in K. Q.E.D.

Theorem 7.25. Suppose K is compact and that $\mathscr{F} = \{f_n\} \subset \mathscr{C}(K)$ is equicontinuous and pointwise bounded on K. Then

(a) $\{f_n\}$ is uniformly bounded

(b) $\{f_n\}$ has a uniformly convergent subsequence; i.e., a subsequence that converges uniformly in $(\mathscr{C}(K), \|\cdot\|)$

Remark.

1. A more topological statement: A5.3
2. Converse: [(b)$\Rightarrow$ equicontinuity and uniform boundedness] see A5.4
3. need for compactness of K: see A5.5
4. good theorem/proof to master..

Proof. (a) Goal: Find M s.t. $|f_n(x)| \le M$ for all $n \in \mathbb{N}, x \in K$.
Since $\mathscr{F}$ is equicontinuous,

$$\forall \varepsilon > 0 : \exists \delta > 0 \text{ s.t. } \forall n \in \mathbb{N} : d(x,y) < \delta \Rightarrow |f_n(x) - f_n(y)| < \varepsilon.$$

As K is compact, $\exists \{p_1, p_2, \ldots, p_r\} \in K$ s.t. $\{N_\delta(p_i)\}_{i \in \{1,\ldots,r\}}$ covers K.
For each i, $\{f_n(p_i)\}_n$ is bounded as $\mathscr{F}$ is pointwise bounded, so there exists M_i s.t. $|f_n(p_i)| \le M_i$ for all $n \in \mathbb{N}$.
Let $M_0 = \max\{M_1, \ldots, M_r\}$.
Given $x \in K$, choose $p \in \{p_1, p_2, \ldots, p_r\}$ such that $x \in N_\delta(p)$.
Then $|f_n(x)| \le |f_n(p)| + |f_n(x) - f_n(p)| \le M_0 + \varepsilon$ for any $\varepsilon > 0$.
Let $M = M_0 + 1$. Then $|f_n(x)| \le M$ for all $n \in \mathbb{N}, x \in K$.

(b) Goal: Construct a uniformly convergent subsequence of $\{f_n\}$.
Step 1. By the lemma, K has a countable dense subset $E \subset K$.
By Theorem 7.23, $\exists \{f_{n_i}\}$ s.t. $\lim\limits_{i\to\infty} f_{n_i}(x)$ exists for all $x \in E$.
Write $g_i = f_{n_i}$. We can show g_i converges uniformly on K via the Uniform Cauchy Criterion Theorem 7.8.
Step 2. Let $\varepsilon > 0$. By equicontinuity, $\exists \delta > 0$ s.t. $d(x,y) < \delta \Rightarrow |g_i(x) - g_i(y)| < \frac{\varepsilon}{3}$ for all i.
$\{N_\delta(p)\}_{p \in E}$ covers K since E is dense. There exists a finite subcover $\{N_\delta(x_1), \ldots, N_\delta(x_m)\}$ with each $x_s \in E$. Given $x \in K$ there exists x_s s.t. $d(x, x_s) < \delta$.
Step 3. (using $\frac{\varepsilon}{3}$ argument)

$$g_i(x) - g_j(x) \le \underbrace{|g_i(x) - g_i(x_s)|}_{<\frac{\varepsilon}{3} \ \ (\because d(x,x_s)<\delta)} + |g_i(x_s) - g_j(x_s)| + \underbrace{|g_j(x_s) - g_j(x)|}_{<\frac{\varepsilon}{3}}.$$

Since $\lim\limits_{i\to\infty} \{g_i(x)\}$ converges for any $x \in K$, $\{g_i(x_s)\}_i$ is a Cauchy sequence.
Hence for $s = 1, \ldots, m$, we can choose N_s s.t. $i, j \ge N_s \Rightarrow |g_i(x_s) - g_j(x_s)| < \frac{\varepsilon}{3}$.
Let $N = \max\{N_1, \ldots, N_m\}$. Then for $i, j \ge N$(independent of x),

$$|g_i(x) - g_j(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}.$$

Hence, $\{g_i\}$ converges uniformly on K by Theorem 7.8.

Q.E.D.

Theorem 7.26. [Weierstrass's Theorem] Let $f : [a,b] \to \mathbb{R}($ or $\mathbb{C})$ be continuous.

> There exists a polynomial sequence $\{P_n\}$ s.t. $P_n \to f$ uniformly on $[a, b]$.

Exercise

By a linear change of variable, can assume $[a, b] = [0, 1]$.
We'll discuss Bernstein's proof (1912), which needs some background in probability theory.
Fix $p \in [0, 1]$. Perform a sequence of n independent of Bernoulli trials:
Success with probability p, failure with probability $1 - p$.
Let S_n be the number of successes observed. Then

$$p_m = P(S_n = m) = \binom{n}{m} p^m (1-p)^{n-m}.$$

> **Note.**
>
> $$\sum_{m=0}^{n} p_m = \sum_{m=0}^{n} \binom{n}{m} p^m (1-p)^{n-m} = [p + (1-p)]^n = 1.$$

Random variable: a function $X : \{0, 1, \ldots, n\} \to \mathbb{R}$; e.g.,

- $X = S_n$ is the identity function $S_n(m) = m$.
- $X(m) = \begin{cases} 0 & m < \frac{n}{2} \\ 1 & m \geq \frac{n}{2} \end{cases}$ is the indicator random variable for the event that at least half of trials are success.

Expectation of X: $EX = E(X) = \sum_{m=0}^{n} X(m) p_m$.
E.g.,

$$\begin{aligned} ES_n &= \sum_{m=0}^{n} m \binom{n}{m} p^m (1-p)^{n-m} \\ &= np \sum_{m=1}^{n} \frac{(n-1)!}{(m-1)!(n-m)!} p^{m-1} (1-p)^{n-m} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} \quad (k := m-1) \\ &= np. \end{aligned}$$

Hence, $ES_n = np$.
Variance of X: $VarX = Var(X) = E[(X - EX)^2] = E(X^2) - (EX)^2$.
E.g., $VarS_n = ES_n^2 - (ES_n)^2 = np(1-p)$

Standard deviation of X: $\sigma_X = \sqrt{VarX}$.
E.g., $\sigma_{S_n} = \sqrt{np(1-p)}$.
Now let $X_n = \frac{1}{n}S_n$ = proportion of successes.
Then

- $EX_n = \frac{1}{n}ES_n = \frac{1}{n}np = p$
- $VarX_n = \frac{1}{n^2} \cdot VarS_n = \frac{p(1-p)}{n}$
- $\sigma_{X_n} = \sqrt{\frac{p(1-p)}{n}}$

Proposition (Chebyshev's Inequality)**.** For all $\delta > 0$, $p \in [0,1]$, $n \in \mathbb{N}$,

$$P(|X_n - p| > \delta) \leq \frac{1}{\delta^2}\frac{p(1-p)}{n}.$$

Proof.

$$\text{LHS} = \sum_{m:|\frac{m}{n}-p|>\delta} p_m \cdot 1 \leq \sum_{m=0}^{n} p_m \cdot \left(\frac{\frac{m}{n}-p}{\delta}\right)^2$$

$$= \frac{1}{\delta^2}\sum_{m=0}^{n}(X_n(m)-p)^2 p_m = \frac{1}{\delta^2}E\left[(X_n-p)^2\right]$$

$$= \frac{1}{\delta^2}E\left[(X_n-EX_n)^2\right] = \frac{1}{\delta^2}\text{Var}X_n = \frac{1}{\delta^2}\frac{p(1-p)}{n}$$

Q.E.D.

Proof (Theorem 7.26)**.** Choose $p = x \in [0,1]$ in above, so now

$$p_m = \binom{n}{m}x^m(1-x)^{n-m}.$$

Let $P_n(x) = E\left[f(X_n)\right] = \sum_{m=0}^{n} f(\frac{m}{n})p_m$.
Then

$$f(x) - P_n(x) = \sum_{m=0}^{n}\left[f(x) - f(\frac{m}{n})\right]p_m.$$

Remark. C.f. A4-Q3
For any $\delta > 0$

$$|f(x) - P_n(x)| \leq \sum_{m:|\frac{m}{n}-x|\leq\delta} \left|f(x) - f(\frac{m}{n})\right| p_m + \sum_{m:|\frac{m}{n}-x|>\delta} 2 \underbrace{M}_{\sup|f(x)|} p_m.$$

Let $\varepsilon > 0$. Choose $\delta > 0$ s.t.

$$|x-y| \leq \delta \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{2}.$$

Note this is possible since f is continuous on $[0,1]$, which is a compact set, and thus uniformly continuous on $[0,1]$.
Since

$$\sum_{m:|\frac{m}{n}-x|\leq\delta} \left|f(x) - f(\frac{m}{n})\right| p_m \leq \frac{\varepsilon}{2} \sum_{m=0}^{n} p_m = \frac{\varepsilon}{2}$$

$$\sum_{m:|\frac{m}{n}-x|>\delta} 2Mp_m \leq 2M\frac{1}{\delta^2}\frac{\overbrace{x(1-x)}^{\leq\frac{1}{4}}}{n} \leq \frac{M}{2\delta^2 n} \quad (\because \text{Chebyshev's}).$$

Choose N s.t. $n \geq N \Rightarrow \frac{M}{2\delta^2}\frac{1}{n} < \frac{\varepsilon}{2}$. Then for $n \geq N$, $|f(x) - p_n(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Q.E.D.

Definition 7.28.

Algebra: Let $\mathscr{A}$ be a set of functions $f : E \to \mathbb{C}$. Then $\mathscr{A}$ is an *algebra* if $\forall f, g \in \mathscr{A}\ \forall c \in \mathbb{C} : f + g \in \mathscr{A},\ fg \in \mathscr{A},\ cf \in \mathscr{A}$.
Example. *Let* $E = [0,1]$, $\mathscr{A} = \mathscr{P} =$ *polynomials, or* $\mathscr{A} = \mathscr{C}(E)$.

Uniform Closure: The *uniform closure* $\mathscr{B}$ of $\mathscr{A}$ is $\{f : E \to \mathbb{C} \mid \exists\{f_n\}$ in $\mathscr{A}$ s.t. $f_n \to f$ uniformly $\}$.

Note. We say an algebra $\mathscr{A}$ is *uniformly closed* if $\mathscr{A} = \mathscr{B}$.

Example.
- $\mathscr{C}[0,1]$ *is the uniform closure of* $\mathscr{P}$. *(by Theorem 7.26)*
- $\mathscr{P}$ *is uniformly closed on* $\mathbb{R}$ *(See A6)*
- $\mathscr{P}$ *is not uniformly closed on* $[0,1]$
- $\mathscr{C}[0,1]$ *is uniformly closed*

Note. If $\mathscr{A}$ is an algebra of bounded functions then it has a metric

$$d(f,g) = \|f-g\| = \sup_{x\in E} |f(x)-g(x)|.$$

In this case, uniform convergence is merely convergence in this metric, and the uniform closure is the closure in this metric; i.e., $\mathscr{B} = \overline{\mathscr{A}}$.

Theorem 7.29. The uniform closure $\mathscr{B} = \overline{\mathscr{A}}$ of an algebra $\mathscr{A}$ of bounded functions is a uniformly closed algebra.

Proof. Suppose $f, g \in \mathscr{B}$ and $c \in \mathbb{C}$. Then $\exists f_n, g_n \in \mathscr{A}$ s.t. $f_n \to f$ uniformly and $g_n \to g$ uniformly:

$$\|f - f_n\| \to 0, \quad \|g - g_n\| \to 0.$$

Then $f_n + g_n \to f + g$ uniformly (Prob. 7.2 of Rudin's) so $f+g \in \mathscr{B}$, and $f_n g_n \to fg$ uniformly (Prob. 7.2,7.3 of Rudin's; need for bounded functions here), so $fg \in \mathscr{B}$, $cf_n \to cf$ uniformly so $cf \in \mathscr{B}$. Therefore, $\mathscr{B}$ is an algebra. $\mathscr{B}$ is uniformly closed because it consists of $\mathscr{A}$ and all limit points of $\mathscr{A}$;i.e., $B = \overline{\mathscr{A}}$. Q.E.D.

Definition 7.30.

Separation of Points : A set $\mathscr{A}$ (not necessarily algebra) of functions $f : E \to \mathbb{C}$ *separates points* on E if for all $x_1, x_2 \in E$ with $x_1 \neq x_2$, there exists $f \in \mathscr{A}$ such that $f(x_1) \neq f(x_2)$

Non-Vanishment : A set $\mathscr{A}$ *vanishes at no point* if $\forall x \in E : \exists f \in \mathscr{A}$ s.t. $f(x) \neq 0$.

Example.

- *The set of polynomials on* $[-1,1]$ *separates points and vanishes at no point.*
- *The set of even polynomials on* $[-1,1]$ *does not separate points.*
- *The set of odd polynomials on* $[-1,1]$ *vanishes at* $x=0$.

Theorem 7.31. If an algebra $\mathscr{A}$ of functions on E separates points and vanishes at no point, then given any $x_1, x_2 \in E$ with $(x_1 \neq x_2)$ and constants c_1, c_2, there exists $f \in \mathscr{A}$ such that $f(x_1) = c_1$ and $f(x_2) = c_2$.

Proof. By hypothesis, $\exists g, h, k \in \mathscr{A}$ such that

$$g(x_1) \neq g(x_2) \qquad (\because A \text{ separates points })$$

$$h(x_1) \neq 0 \qquad (\because A \text{ does not vanish at } x_1)$$

$$k(x_2) \neq 0 \qquad (\because A \text{ does not vanish at } x_2).$$

Q.E.D.

Let $u(x) = [g(x) - g(x_1)]k(x)$. Then $u \in \mathscr{A}, u(x_1) = 0,\ u(x_2) \neq 0$.
Let $v(x) = [g(x) - g(x_2)]h(x)$. Then $v \in \mathscr{A}, v(x_1) \neq 0,\ v(x_2) = 0$.
Let

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

Then $f \in \mathscr{A},\ f(x_1) = c_1 \cdot 1 + c_2 \cdot 0 = c_1,\ f(x_2) = c_1 \cdot 0 + c_2 \cdot 1 = c_2$.

Theorem 7.32. [Stone-Weierstrass Theorem] The uniform closure of any algebra of real continuous functions on a compact set K, which separates points and vanishes at no point, is the set of all real-valued continuous functions on K; i.e., $\mathscr{C}(K)$.

For the proof of Theorem 7.32, we need the following lemmas

Lemma (1). Let $\mathscr{A}$ be an algebra of real-valued continuous functions on a compact set K. If $f \in \overline{\mathscr{A}}$ then $|f| \in \overline{\mathscr{A}}$.

Proof. Note that $\overline{\mathscr{A}}$ is an algebra since it is a uniform closure of a bounded algebra, where the boundedness is due to the continuity of functions in $\mathscr{A}$ on a compact set K.
Let $f \in \overline{\mathscr{A}}$. Let $M = \sup_{x \in K} |f(x)|$. Then $M < \infty$.
Let $\varepsilon > 0$. By Theorem 7.26, there exists a polynomial $\tilde{P}$ such that $\sup_{|y| \leq M} \left| \tilde{P}(y) - |y| \right| < \frac{\varepsilon}{2}$. Let $P(y) = \tilde{P}(y) - \tilde{P}(0) = \sum_{j=1}^{n} c_j y^j$. Then

$$\begin{aligned} |P(y) - |y|| &= \left| \tilde{P}(y) - \tilde{P}(0) - |y| \right| \\ &\leq \left| \tilde{P}(y) - |y| \right| + \left| \tilde{P}(0) - |0| \right| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ for all } y \text{ such that } |y| \leq M. \end{aligned}$$

Also, $P(f) = \sum_{j=1}^{n} c_j f^j \in \overline{\mathscr{A}}$ since $\overline{A}$ is an algebra, and $\sup_{x \in K} |(P(f))(x) - |f(x)|| < \varepsilon$ as $|f(x)| \leq M$.
Thus, $|f|$ can be approximated arbitrarily by an element of $\overline{\mathscr{A}}$, so $|f| \in \overline{\mathscr{A}}$.

[Note we don't know if constants are in $\overline{\mathscr{A}}$, so we omitted the $j = 0$ term] Q.E.D.

Lemma (2)**.** Let $\mathscr{A}$ as in Lemma 1.
If $f_1, f_2, \ldots f_n \in \overline{\mathscr{A}}$, then $\max\{f_1, f_2, \ldots, f_n\} \in \overline{\mathscr{A}}$ and $\min\{f_1, f_2, \ldots, f_n\} \in \overline{\mathscr{A}}$.

Note.

$$\max\{f, g\}(x) := \begin{cases} f(x) & f(x) \geq g(x) \\ g(x) & f(x) < g(x) \end{cases}$$

$$\min\{f, g\}(x) := \begin{cases} g(x) & f(x) \geq g(x) \\ f(x) & f(x) < g(x) \end{cases}.$$

Proof. It suffices to consider $n = 2$, and for this, use

$$\max\{f_1, f_2\} = \frac{1}{2}(f_1 + f_2) + \frac{1}{2}|f_1 - f_2| \qquad (*)$$

$$\min\{f_1, f_2\} = \frac{1}{2}(f_1 + f_2) - \frac{1}{2}|f_1 - f_2|,$$

and apply Lemma 1.
(*) holds because if $f_1(x) \leq f_2(x)$, then $(\text{LHS}(*))(x) = f_2(x)$, and $(\text{RHS}(*))(x) = \frac{1}{2}(f_1(x) + f_2(x)) + \frac{1}{2}(f_2(x) - f_1(x)) = f_2(x)$.
Q.E.D.

Proof (Theorem 7.32)**.** It suffices to prove that if $\mathscr{A}$ is an algebra of real continuous functions on a compact set K that separates points and vanishes at no points, then $\forall f \in \mathscr{C}(K) : \forall \varepsilon > 0 : \exists h \in \mathscr{B} = \overline{\mathscr{A}}$ such that $\|f - h\| < \varepsilon$. This suffices because if there exists such an h, then we can choose $g \in \mathscr{A}$ such that $\|g - h\| < \varepsilon$, so $\|f - g\| \leq \|f - h\| + \|h - g\| < \varepsilon + \varepsilon = 2\varepsilon$.

Claim (1)**.** Let $f : K \to \mathbb{R}$ be continuous, $\varepsilon > 0$. Then

$\exists g_x \in \mathscr{B}$ such that $g_x(x) = f(x)$ and $g_x(t) - f(t) > -\varepsilon$ for all $t \in K$.

Proof (Claim 1)**.** Given $y \in K$ with $y \neq x$, by Theorem 7.31, there exists $h_y \in \mathscr{A}$ such that $h_y(x) = f(x)$, $h_y(y) = f(y)$. Since $h_y - f$ is continuous and $h_y(y) - f(y) = 0$, there exists open $J_y \subset K$ with $y \in J_y$ such that $h_y(t) - f(t) > -\varepsilon$ for all $t \in J_y$.
Since K is compact, $K \subset J_{y_1} \cup \cdots \cup J_{y_n}$ for some $y_1, y_2, \ldots, y_n \in K$, where some y_j may be x.
Let $g_x = \max\{h_{y_1}, h_{y_2}, \ldots, h_{y_n}\}$. Then $g(x) \in \mathscr{B}$ by Lemma , and $g_x(x) = \max\{h_{y_1}(x), \ldots, h_{y_n}(x)\} = f(x)$.
Also, $g_x(t) - f(t) \geq h_{y_i}(t) - f(t) > -\varepsilon$, where i is chosen such that $t \in J_{y_i}$.
Q.E.D.

Claim (2)**.** Let $f : K \to \mathbb{R}$ be continuous, $\varepsilon > 0$. Then

$$\exists h \in \mathscr{B} \text{ such that } \sup_{x \in K} |h(x) - f(x)| < \varepsilon.$$

Proof (Claim 2)**.** Since $g_x - f$ is continuous and $g_x(x) - f(x) = 0$, there exists open $V_x \subset K, V_x \ni x, g_x(t) - f(t) < \varepsilon$ for all $t \in V_x$. Since K is compact, $K \subset V_{x_1} \cup V_{x_2} \cup \cdots \cup V_{x_m}$ for some $x_1, x_2, \ldots, x_m \in K$. Let $h = \min\{g_{x_1, g_{x_2}}, \ldots, g_{x_n}\} \in \mathscr{B}$. Then by Claim 1,

$$\forall t \in K : h(t) - f(t) \underbrace{=}_{\text{choose } i \text{ s.t. } g_{x_i}(t) \text{ gives the minimum}} g_{x_i}(t) - f(t) > -\varepsilon.$$

We also have $h(t) - f(t) \leq g_{x_j}(t) - f(t) < \varepsilon$ (Choose j s.t. $t \in V_{x_j}$). Therefore, $-\varepsilon < h(t) - f(t) < \varepsilon$ for all $t \in K$, and hence $\sup_{t \in K} |h(t) - f(t)| < \varepsilon$. Q.E.D.

Q.E.D.

Theorem 7.33. [Complex Stone-Weierstrass Theorem] Let $\mathscr{A}$ be a self-adjoint algebra of complex continuous functions on a compact set K that separates points and vanishes at no point. Then the uniform closure of $\mathscr{A}$ is the set of complex continuous functions on K; i.e.,

$$\overline{\mathscr{A}} = \mathscr{C}_{\mathbb{C}}(K).$$

Note. Complex case of Theorem 7.32 requires additional hypothesis that $\mathscr{A}$ is self-adjoint; i.e., if $f \in \mathscr{A}$ then $\overline{f} \in \mathscr{A}$, where $\overline{f}(x) = \overline{f(x)} =$ conjugate of $f(x)$.

Proof. Let $\mathscr{A}_{\mathbb{R}}$ be the algebra of real continuous functions in $\mathscr{A}$.

(i) If $f = u + iv \in \mathscr{A}$, then $u, v \in \mathscr{A}_{\mathbb{R}}$.
Proof. Since $u = \frac{1}{2}(f + \overline{f}) \in \mathscr{A}$ by self-adjointness of $\mathscr{A}$ and u is real-valued. Similarly, $v = \frac{1}{2i}(f - \overline{f}) \in \mathscr{A}_{\mathbb{R}}$.

(ii) $\mathscr{A}_{\mathbb{R}}$ separates points.
Proof. Let $x_1, x_2 \in K$ with $x_1 \neq x_2$. By Theorem 7.31, $\exists f \in \mathscr{A}$ such that $f(x_1) = 1$ and $f(x_2) = 0$.
Write $f = u + iv$. Then $u(x_1) = 1, u(x_2) = 0, u \in \mathscr{A}_{\mathbb{R}}$

(iii) $\mathscr{A}_{\mathbb{R}}$ vanishes at no point.
Proof. Let $x \in K$. Choose $g \in \mathscr{A}$ s.t. $g(x) \neq 0$.
Write $g(x) = re^{i\theta}(r > 0)$.
Let $f(t) = e^{-i\theta}g(t)$.
Then $\operatorname{Re} f \in \mathscr{A}$ obeys $(\operatorname{Re} f)(x) = \operatorname{Re} e^{-i\theta}g(x) = \operatorname{Re} e^{-i\theta}re^{i\theta} = r \neq 0$.

Therefore, by Theorem 7.32, $\overline{\mathscr{A}_{\mathbb{R}}} = \mathscr{C}(K)$. Let $f = u + iv \in \mathscr{C}_{\mathbb{C}}(K)$.
Let $\varepsilon > 0$.
Since $u, v \in \mathscr{C}_{\mathbb{R}}(K) = \overline{\mathscr{A}_{\mathbb{R}}}$, there exists $\tilde{u}, \tilde{v} \in \mathscr{A}_{\mathbb{R}}$ s.t. $\|u - \tilde{u}\| < \frac{\varepsilon}{2}$ and $\|v - \tilde{v}\| < \frac{\varepsilon}{2}$, so

$$\|f - \underbrace{(\tilde{u} + i\tilde{v})}_{\in \mathscr{A}}\| \leq \|u - \tilde{u}\| + \|v - \tilde{v}\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Q.E.D.

Chapter 8

Some Special Functions

Power series, $e^x, \log x, \sin x, \cos x$, Fourier series (we're omitting Gamma function)

8.1 Power Series

Recall $\sum_{n=0}^{\infty} c_n x^n$ has a radius of convergence $R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|c_n|}}$ such that the series absolutely converges for $|x| < R$, diverges for $|x| > R$ and anything possible for $|x| = R$.

Remark. To determine R, we often use the ratio test instead:

$$\left|\frac{c_{n+1}x^{n+1}}{c_n x^n}\right| = |x|\cdot\left|\frac{c_{n+1}}{c_n}\right| \underbrace{\rightarrow}_{\text{if limit exists}} |x|\cdot L \Rightarrow \text{ absolute convergence if } |x| < \frac{1}{L}, \text{ diverges if } |x| > \frac{1}{L},$$

where $R = \frac{1}{L} = \frac{1}{\lim\limits_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right|}$.
In general, by Theorem 3.37, we also have

$$\frac{1}{\liminf_{n\to\infty}\left|\sqrt[n]{c_n}\right|} \le R \le \frac{1}{\liminf_{n\to\infty}\frac{c_{n+1}}{c_n}}.$$

Theorem 8.1. Suppose $\sum_{n=0}^{\infty} c_n x^n$ has a radius of convergence $R > 0$ and define $f(x) = \sum_{n=0}^{\infty} c_n x^n$ for $|x| < R$. Such a function $f(x)$ is called an *analytic function* .
If $R < \infty$, then the series converges uniformly on $[-R+\varepsilon, R-\varepsilon]$ for all $\varepsilon > 0$.
If $R = \infty$, then the series converges uniformly on $[-M, M]$ for all $M < \infty$.

The function f is continuous and differentiable on $(-R, R)$ with $f'(x) = \sum_n c_n n x^{n-1}$

Note. a function of the form $f(x) = \sum_{n=0}^{\infty} c_n x^n$ is called an *analytic function* .

Remark. Uniform convergence may not hold on $(-R, R)$. C.f. A7-Q3.

Proof.

Uniform convergence: For $|x| \leq R - \varepsilon$, $|c_n x^n| \leq |c_n| (R-\varepsilon)^n$. $\sum_{n=0}^{\infty} |c_n| (R-\varepsilon)^n < \infty$ by absolute convergence on $(-R, R)$. Hence, by Weierstass M-test, $\sum_{n=0}^{\infty} c_n x^n$ converges uniformly on $[-R+\varepsilon, R-\varepsilon]$.

Derivative: The radius of convergence of $\sum_{n=1}^{\infty} n c_n x^{n-1}$ is

$$\frac{1}{\limsup_{n\to\infty} \sqrt[n]{n\,|c_n|}} = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{n}\,\sqrt[n]{|c_n|}}$$

$$= R = \text{ radius of convergence of } \sum_{n=0}^{\infty} c_n x^n \ (\because \lim_{n\to\infty} \sqrt[n]{n} = 1).$$

Let $S_n(x) = \sum_{m=0}^{n} c_m x^m$. Then $S_n'(x) = \sum_{m=1}^{n} c_m m x^{m-1}$. By the first part of the proof, $S_n'(x) \to \sum_{m=1}^{\infty} c_m m x^{m-1}$ uniformly on $[-R+\varepsilon, R-\varepsilon]$. Since also $S_n(x) \to f(x)$, by Theorem 7.17, f' exists on $[-R+\varepsilon, R-\varepsilon]$, and $f'(x) = \sum_{m=1}^{\infty} c_m m x^{m-1}$. Since ε is arbitrary, $f'(x) = \sum_{m=1}^{\infty} c_m m x^{m-1}$ for all $x \in (-R, R)$. In particular, f is also continuous.

Q.E.D.

Corollary 8.2. If $f(x) = \sum_{n=0}^{\infty} c_n x^n$ converges for $|x| < R$, then $f^{(k)}(x)$ exists for all $k \in \mathbb{N}$, and

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

Consequently, $c_k = f^{(k)}(0)$ and $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$.

Note. C.f. Taylor's theorem.

Proof. By Theorem 8.1, $f'(x) = \sum_{n=1}^{\infty} c_n n x^{n-1}$, $f''(x) = (f')'(x) = \sum_{n=2}^{\infty} c_n n(n-1)x^{n-2}, \ldots$.
Set $x = 0$ in (*) to get $f^{(k)}(0) \underbrace{=}_{\text{only } n=k \text{ term survives}} c_k k(k-1)\cdot\cdots\cdot 1 = c_k \cdot k!$. Q.E.D.

Example. *Let* $f(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases}$. *By Rudin's problem 8.1,* $f^{(n)}(0) = 0$ *for all* $n = 0, 1, 2, \ldots$, *so* $f(x) \neq \sum_{n=0}^{\infty} \frac{f^{(n)}}{n!} x^n$ *except for* $x = 0$.

Remark (Bump Function). *Bump functions* are infinitely differentiable functions with compact support. E.g.,

$$f(x) = \begin{cases} e^{-\frac{1}{1-x^2}} & x \in (-1, 1) \\ 0 & |x| \geq 1 \end{cases}.$$

Theorem 8.2. [Abel's Theorem] Suppose $\sum_{n=0}^{\infty} c_n$ converges (perhaps conditionally). Let $f(x) = \sum_{n=0}^{\infty} c_n x^n$. Then $f(x)$ converges for $|x| < 1$ and $\lim_{x \to 1^-} f(x) = f(1) = \sum_{n=0}^{\infty} c_n$.

Remark. Interesting case is at $R = 1$, since $R > 1$ implies continuity of f for $|x| < R$.

Proof. By the root test, $\limsup_{n\to\infty} \sqrt[n]{|c_n|} \le 1$, so $\sum_{n=0}^{\infty} c_n x^n$ has $R \ge 1$. Let $S_n = \sum_{m=0}^{n} c_m$ and $S = \sum_{m=0}^{\infty} c_m = \lim_{n\to\infty} S_n$.
Set $S_{-1} = 0$. Then $c_n = s_n - s_{n-1}$ for $n \ge 0$.
Let $\varepsilon > 0$. We need to show $\exists \delta > 0$ such that $1 - \delta < x < 1 \Rightarrow |f(x) - S| < \varepsilon$.
Start with partial sum for $f(x)$. For $|x| < 1$,

$$\sum_{m=0}^{n} c_m x^m = \sum_{m=0}^{n} (S_m - S_{m-1}) x^m = \sum_{m=0}^{n} S_m x^m - \sum_{m=0}^{n} S_{m-1} x^m.$$

Let $k = m - 1$ so that $m = k + 1$. Then

$$\sum_{m=0}^{n} S_{m-1} x^m = \underbrace{S_{0-1}}_{0} \cdot 1 + \sum_{m=1}^{n} S_{m-1} x^m = x \sum_{k=0}^{n-1} S_k x^k$$

$$\sum_{m=0}^{n} c_m x^m = (1-x) \sum_{m=0}^{n} S_m x^m + \underbrace{S_n x^{n+1}}_{\to 0 \text{ as } n\to\infty \text{ since } S_n \text{ is bounded and } |x|<1}.$$

Let $n \to \infty$. Then

$$f(x) = (1-x) \sum_{n=0}^{\infty} S_n x^n.$$

$$|f(x) - S| = \left| (1-x) \sum_{n=0}^{\infty} S_n x^n - S(1-x) \cdot \frac{1}{1-x} \right| = \left| (1-x) \sum_{n=0}^{\infty} S_n x^n - S(1-x) \sum_{n=0}^{\infty} x^n \right|$$

$$= (1-x) \left| \sum_{n=0}^{\infty} (S_n - S) x^n \right| \le (1-x) \sum_{n=0}^{\infty} |S_n - S| \cdot |x|^n.$$

Choose N s.t. $n \ge N \Rightarrow |S_n - S| < \frac{\varepsilon}{2}$. For $x \in (0,1)$,

$$|f(x) - S| \le (1-x) \sum_{n=0}^{N} |S_n - S| \, x^n + (1-x) \sum_{n=N+1}^{\infty} |S_n - S| \, x^n$$

$$< (1-x) \sum_{n=0}^{N} |S_n - S| \, x^n + (1-x) \left(\frac{\varepsilon}{2} \cdot \frac{1}{1-x} \right) = (1-x) \left(\sum_{n=0}^{N} |S_n - S| \, x^n \right) + \frac{\varepsilon}{2}.$$

Since $(1-x) \sum_{n=0}^{N} |S_n - S| \, x^n$ is a polynomial in x, so it is continuous and equals 0 at $x = 1$.
Hence, $(1-x) \sum_{n=0}^{N} |S_n - S| \, x^n < \frac{\varepsilon}{2}$ if $|x - 1| < \delta$ for some $\delta > 0$.
Therefore, for $1 - \delta < x < 1$, $|f(x) - S| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Q.E.D.

Note. For an application of Abel's theorem, see Rudin's p. 175. For the case $\sum_{n=0}^{\infty} c_n = \infty$, see A7.

Theorem 8.3. If $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}|$ converges, then

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij},$$

where both sides converge.

Proof. Rudin has a too clever proof ... A7 involves a more straightforward proof. Q.E.D.

Theorem 8.4. Suppose $f(x) = \sum_{n=0}^{\infty} c_n x^n$ (Taylor series of f at $x = 0$, a.k.a Maclauren series) has a radius of convergence $R > 0$. Let $|a| < R$. Then $f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$ for (at least) $|x - a| < R - |a|$.

Proof. Note

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

We want to interchange the order of summation.
By Theorem 8.3, interchange of summations is justified if

$$\sum_{n=0}^{\infty} \sum_{m=0}^{n} |c_n| \binom{n}{m} |x-a|^m |a|^{n-m} < \infty.$$

Note $\sum_{n=0}^{\infty} |c_n| \left(|x-a| + |a|\right)^n$ does converge as we assume $|x-a| + |a| < R$.
Therefore, for $|x-a| < R - |a|$,

$$\begin{aligned} f(x) &= \sum_{m} \frac{1}{m!}(x-a)^m \sum_{n=m}^{\infty} c_n [n \cdot (n-1)(n-2) \cdot \dots \cdot (n-m+1)] a^{n-m} \\ &= \sum_{m} \left(\sum_{n} c_n \binom{n}{m} a^{n-m} \right) (x-a)^m \\ &= \sum_{m} \left(\sum_{n} \frac{f^{(n)}(a)}{n!} n(n-1) \cdots (n-m+1) a^{n-m} \right) (x-a)^m \end{aligned}$$

.

Q.E.D.

Example. *Let* $f(x) = \sum_{n=0}^{\infty} x^n$ *for* $|x| < 1$. *Then* $f(x) = \frac{1}{1-x}$ *for* $|x| < 1$.
Taylor series of $\frac{1}{1-x}$ *at* $x = -\frac{1}{2}$: *For* $|x| < 1$, $f^{(n)}(x) = \frac{n!}{(1-x)^{n+1}}$, *so* $f^{(n)}(-\frac{1}{2}) = \frac{n!}{(\frac{3}{2})^{n+1}}$.
By Theorem 8.4,

$$f(x) = \sum_{n=0}^{\infty} \frac{n!}{(\frac{3}{2})^{n+1} n!} \left(x + \frac{1}{2}\right)^n = \sum_{n=0}^{\infty} (\frac{2}{3})^n \left(x + \frac{1}{2}\right)^n$$

for $\left|x + \frac{1}{2}\right| < 1 - \left|-\frac{1}{2}\right| = \frac{1}{2}$.
In fact, the series converges even when $\left|\frac{2}{3}\left(x + \frac{1}{2}\right)\right| < 1$; *i.e.,* $\left|x + \frac{1}{2}\right| < \frac{3}{2}$.
C.f. Analytic continuation.
Another way to get Taylor series at $x = -\frac{1}{2}$:

For $|x| < 1$,

$$f(x) = \frac{1}{1-x} = \frac{1}{\left(1+\frac{1}{2}\right) - \left(x+\frac{1}{2}\right)} = \frac{2}{3}\frac{1}{1-\frac{2}{3}\left(x+\frac{1}{2}\right)} = \frac{2}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n\left(x+\frac{1}{2}\right)^n.$$

Theorem 8.5. [Principle of Permanence of Form] Suppose $\sum_n a_n x^n$ and $\sum_n b_n x^n$ have radii of convergence at least R. Suppose $D \subset (-R, R)$ has a limit point in $(-R, R)$.
If $\sum_n a_n x^n = \sum_n b_n x^n$ for all $x \in D$, then $a_n = b_n$ for all n, and hence $\sum_n a_n x^n = \sum_n b_n x^n$ for all $|x| < R$.

Proof. Let $c_n = a_n - b_n$ and $f(x) = \sum_{n=0}^{\infty} c_n x^n$. Then $f(x) = 0$ for all $x \in D$.
Let $E = \{x \in (-R, R) : f(x) = 0\}$. Then $D \subset E$.
We want to show $E = (-R, R)$. Let $A = E' \cap (-R, R)$. Then $A \neq \emptyset$ because D has a limit point in $(-R, R)$.
Also, relative to $(-R, R)$, A is closed as the set of all limit points is always closed (C.f. Problem 2.6).
Let $B = (-R, R) \setminus A$. Then $A \cup B = (-R, R)$, $A \cap B = \emptyset$, and B is open.

Claim. A is also open.

Proof. Let $x_0 \in A = E' \cap (-R, R)$. Then $\exists\{d_n\}$ such that $f(x) = \sum_{n=0}^{\infty} d_n (x - x_0)^n$ for $|x - x_0| < R - |x_0|$. If we show that $d_n = 0$ for all $n \geq 0$, it proves $(x_0 - r, x_0 + r) \subset A$, and hence that A is open. Suppose for contradiction there exists $k \geq 0$ such that $d_k \neq 0$ and $f(x) = \sum_{n=k}^{\infty} d_n (x - x_0)^n$. Then $f(x) = (x - x_0)^k \sum_{n=k}^{\infty} d_n (x - x_0)^{n-k}$. Let $m = n - k$. Then

$$f(x) = (x - x_0)^k \sum_{m=0}^{\infty} d_{m+k} (x - x_0)^m.$$

Say $g(x) = \sum_{m=0}^{\infty} d_{m+k} (x - x_0)^m$. Then $g(x_0) = d_k \neq 0$, and by continuity of g, $\exists \delta > 0$ such that $g(x) \neq 0$ if $|x - x_0| < \delta$. However, $f(x) = (x - x_0)^k g(x) \neq 0$ if $0 < |x - x_0| < \delta$, so x_0 is an isolated zero of f and $x_0 \in E'$. This is a contradiction, so $d_n = 0$ for all $n \geq 0$. Q.E.D.

Given the claim, in $(-R, R)$, A and B are both open and closed. Since $(-R, R)$ is connected (C.f. MATH-320's A5.3(c)), one of

(i) $A = \emptyset, B = (-R, R)$

(ii) $A = (-R, R), B = \emptyset$

must hold.
As A is non-empty, $A = (-R, R) = E'$.
Therefore, $\forall x \in (-R, R) : \exists\{x_n\}$ in E such that $x_n \to x$. Since f is continuous on $(-R, R)$, $f(x) = \lim_{n \to \infty} f(x_n) = 0$.
Therefore, for all $x \in (-R, R)$, $x \in E$, so $(-R, R) \subset E$.

Q.E.D.

Remark. CUT-OFF for MT2 ends here, including A4,5,6,7 and up to Rudin's p.143-178.

8.2 Exponential Function

Recall

$$e := \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Also, $e = \lim\limits_{n\to\infty} \left(1+\frac{1}{n}\right)^n$, $e < 3, e \in \mathbb{Q}$.

Definition. For $z \in \mathbb{C}$, $E(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$ and it has a radius of convergence $R = \infty$. Note $E(1) = e$ is given by the definition of e.

Theorem 8.6. For $p \in \mathbb{Q}$, $E(p) = e^p$.

Proof. (i) Note

$$E(w)E(z) = \sum_{m=0}^{\infty} \frac{w^m}{m!} \sum_{n=0}^{\infty} \frac{z^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z^k}{k!} \frac{w^{n-k}}{(n-k)!} \quad (\text{ by Theorem 3.50})$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} = \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = E(z+w).$$

In particular,

$$E(2) = E(1+1) = E(1)^2 = e^2$$

$$E(3) = E(1+2) = E(1)E(2) = e \cdot e^2 = e^3$$

$$\vdots E(n) = e^n \text{ for } n = 0,1,2,3,\ldots,(E(0) = \frac{1}{0!} = 1).$$

(ii) Since $E(z)E^{-z} = E(z-z) = 1$, $E(-z) = \frac{1}{E(z)}$. Hence, $E(-n) = \frac{1}{E(n)} = \frac{1}{e^n} = e^{-n}$.
Thus, $E(n) = e^n$ for $n \in \mathbb{Z}$.

(iii) Let $p = \frac{m}{n}$, $m \in \mathbb{Z}, n \in \mathbb{N}$. Then

$$e^m = E(m) = E(np) = E(p+p+\cdots+p) = E(p)^n.$$

Hence, $E(\frac{m}{n}) = E(p) = e^{\frac{m}{n}} = e^p$.

Q.E.D.

Definition. For $x \in \mathbb{R}$, let $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = E(x)$. This defines $x \mapsto e^x$ as an analytic function on $\mathbb{R}$, which must be infinitely differentiable and

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = \sum_{n=1}^{\infty} \frac{1}{n!} n x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} = e^x.$$

We then have $\frac{\mathrm{d}^k}{\mathrm{d}x^k} e^x = e^x$ for all $x = 0, 1, 2, \ldots$.
Since $e^x e^{-x} = 1$ and since $e^x > 0$ for $x \geq 0$, $e^{-x} > 0$, so $e^x > 0$ for all $x \in \mathbb{R}$.
From the first two derivatives, $x \mapsto e^x$ is strictly increasing and strictly convex.
For $n \geq 0$ and $x > 0$, $e^x > \frac{x^{n+1}}{(n+1)!}$, so $\frac{e^x}{x^n} > \frac{x}{(n+1)!} \to \infty$ as $x \to \infty$.
In particular, with $n = 0$, $e^x \to \infty$ as $x \to \infty$, and $e^{-x} = \frac{1}{e^x} \to 0$ as $x \to \infty$.

Definition. For $y > 0$, define $L(y) = \log y (= \ln y)$ by $E(L(y)) = y$. Equivalently, $L(E(x)) = x$ for all $x \in \mathbb{R}$, or $e^{\log y} = y$ for all $y > 0$ and $\log e^x = x$.
Since $x \mapsto e^x$ is strictly increasing and differentiable, its inverse $y \mapsto \log y$ is also strictly increasing and differentiable.

Theorem. $L'(y) = \frac{1}{y}$ for $y > 0$, $L(uv) = L(u) + L(v)$ for $u, v > 0$.

Proof. Apply the chain rule to $L(E(x)) = x$ (C.f. Rudin's Problem 5.2).
Then $L'(E(x))E'(x) = 1$, so $L'(y) = L'(E(x)) = \frac{1}{E(x)} = \frac{1}{y}$.
Also, since there exist unique x, y such that $E(x) = u, E(y) = v$, $L(uv) = L(E(x)E(y)) = L(E(x+y)) = x + y = L(u) + L(v)$.
Q.E.D.

Corollary. For $y > 0$, $L(y) = \int_1^y \frac{1}{t} \mathrm{d}t$.

Proof. By Theorem 6.21, RHS $= L(t)\Big|_1^y = L(y) - L(1) = L(y)$.
Q.E.D.

Theorem. For $p \in \mathbb{Q}$ and $x > 0$, $L(x^p) = pL(x)$, so $x^p = e^{p \log x}$.

Proof. C.f. Rudin's p.181. Q.E.D.

Definition. For $\alpha \in \mathbb{R}$, $x > 0$, define $x^\alpha = e^{\alpha \log x}$.

Theorem 8.7. For $x > 0$, $\alpha \in \mathbb{R}$, $\frac{\mathrm{d}}{\mathrm{d}x} x^\alpha = \alpha x^{\alpha - 1}$. Hence, x^α has an antiderivative of $\begin{cases} \frac{x^{\alpha+1}}{\alpha+1} + C & \alpha \neq -1 \\ \log x + C & \alpha = -1 \end{cases}$.

Proof. $\frac{\mathrm{d}}{\mathrm{d}x}x^\alpha = \frac{\mathrm{d}}{\mathrm{d}x}e^{\alpha\log x} = e^{\alpha\log x}\frac{\alpha}{x} = \frac{x^\alpha\alpha}{x} = \alpha x^{\alpha-1}$. Q.E.D.

Theorem 8.8. (a) $\lim_{x\to\infty}\log x = \infty$

(b) $\lim_{x\to 0^+}\log x = -\infty$

(c) $\lim_{x\to\infty}\frac{\log x}{x^\alpha} = 0$ if $\alpha > 0$

.

Proof. (a) exercise

(b) exercise

(c) For $a > 0$, $x > 1$,

$$\log x = \int_1^x \frac{1}{t}\mathrm{d}t \leq \int_1^x t^a\frac{1}{t}\mathrm{d}t = \frac{1}{a}(x^a - 1) < \frac{1}{a}x^a.$$

Choose $a \in (0, \alpha)$. Then

$$\frac{1}{x^\alpha}\log x \leq \frac{1}{x^\alpha}\frac{1}{a}x^a = \frac{1}{a}\frac{1}{x^{\alpha-a}} \to 0 \text{ as } x \to \infty.$$

Q.E.D.

For sine,cosine, pi, read Rudin's p. 182-184.

Theorem 8.8. [Fundamental Theorem of Algebra]
Let $n \in \mathbb{N}$, $a_0, a_1, a_2, \ldots, a_n \in \mathbb{C}$, $a_n \neq 0$,
$P(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{n-1}z^{n-1} + a_n z^n$ for $z \in \mathbb{C}$.
Then there exists $z_0 \in \mathbb{C}$ such that $P(z_0) = 0$.

Proof. Can assume $a_n = 1$. Let $\mu = \inf_{z\in\mathbb{C}}|P(z)|$. We want to show that $\mu = 0$ and is attained at some $z_0 \in \mathbb{C}$.

Claim (1). There exists $z_0 \in \mathbb{C}$ such that $|P(z_0)| = \mu$.

Proof. Suppose $|z| = R$. Then

$$\begin{aligned} |P(z)| &= |z|^n \cdot \left|1 + \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n}\right| \\ &\geq |z|^n \left(1 - \frac{|a_{n-1}|}{|z|} - \cdots - \frac{|a_0|}{|z|^n}\right) \\ &= R^n \left(1 - \frac{|a_{n-1}|}{R} - \cdots - \frac{|a_0|}{R^n}\right). \end{aligned}$$

Hence, $|P(z)| \to \infty$ as $R \to \infty$, so there exists R_0 that depends on μ such that $|P(z)| \geq \mu$ if $|z| \geq R_0$.
Since $|P|$ is continuous and $|z| \leq R_0$ is compact, by Theorem 4.16, $\mu = \inf_{|z| \leq R_0} |P(z)| = P(z_0)$ for some z_0 with $|z_0| \leq R_0$. Q.E.D.

Claim (2). $\mu = 0$.

Proof. Suppose for contradiction $\mu > 0$, so $\mu = P(z_0) \neq 0$. Let

$$Q(z) = \frac{P(z_0 + z)}{P(z_0)}.$$

As we assume $P(z_0) = \mu > 0$, $Q(z)$ is well defined.
As $Q(0) = 1$, we can write $Q(z) = 1 + b_k z^k + \cdots + b_n z^n$, where b_k is the first non-zero coefficient.
By definition of z_0, for all $z \in \mathbb{C}$,

$$|P(z_0 + z)| \geq P(z_0) = \mu.$$

Hence,

$$|Q(z)| \geq |Q(0)| = 1.$$

Note

$$|Q(z)| \leq \left|1 + b_k z^k\right| + \sum_{m=k+1}^{n} |b_m| \, |z|^m.$$

We need to choose z s.t. $\left|1 + b_k z^k\right| < 1$.
Write $b_k z^k = |b_k| \frac{b_k}{|b_k|} z^k$, $\frac{b_k}{|b_k|} = e^{it}$. $z = re^{i\theta}$ to be chosen..
Then

$$b_k z^k = |b_k| \, e^{it} r^k e^{ik\theta} = |b_k| \, r^k e^{i(t+k\theta)}.$$

Let θ so that $t + k\theta = \pi$; i.e., $\theta = \frac{\pi - t}{k}$. Let $\varepsilon = 1 + b_k z^k = 1 - |b_k| \, r^k$.
Then $\varepsilon < 1$, and for a small enough r, we have

$$\sum_{m=k+1}^{n} |b_m| \, r^m < 1 - \varepsilon.$$

Therefore, there exists $z = re^{i\theta} \in \mathbb{C}$ such that

$$|Q(z)| = \left|Q(re^{i\theta})\right| \leq 1 - |b_k|^{r^k} + \sum_{m=k+1}^{n} |b_m| \, r^m = \varepsilon + \sum_{m=k+1}^{n} |b_m| \, r^m < 1.$$

This contradicts $|Q(z)| \geq 1$ for all $z \in \mathbb{C}$, so $\mu = 0$. Q.E.D.

Q.E.D.

8.3 Fourier Series

Definition (Fourier Series)**.** Given $a < b$ and integrable $f, g : [a, b] \to \mathbb{C}$, we write

$$< f, g >= \int_a^b f(x) \overline{g(x)} \mathrm{d}x.$$

Remark. $< f,g >= \overline{< g,f >}$, $\|f\|_2^2 =< f,f >= \int_a^b |f(x)|^2 \,\mathrm{d}x$.

Note. (a) $d(f,g) = \|f-g\|_2$ defines a metric on $\mathscr{C}([a,b])$ but not on $\mathscr{R}[a,b]$, as we can have $f \in \mathscr{R}[a,b], f \neq 0, \int_a^b |f(x)|^2 \,\mathrm{d}x = 0$.

(b) Sometimes when $[a,b] = [-\pi,\pi]$, we prefer $< f,g >= \frac{1}{2\pi} \cdot \int_{-\pi}^{\pi} f(x)\overline{g(x)}\mathrm{d}x$.

Definition (Orthogonal Family of Functions)**.** Functions $\varphi_n : [a,b] \to \mathbb{C}$ are *orthogonal* if $< \varphi_n, \varphi_m >= 0$ for $n \neq m$.

They are *orthonormal* if $< \varphi_m, \varphi_n >= \delta_{mn} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$.

Example. *(a)* $[a,b] = [-\pi,\pi]$, $\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$ *for* $n \in \mathbb{Z}$ *obey* $< \varphi_m, \varphi_n >= \int_{-\pi}^{\pi} \frac{1}{2\pi} e^{-i(m-n)x}\mathrm{d}x = \frac{1}{2\pi}\int_{-\pi}^{\pi} [\cos(m-n)x + i\sin(m-n)x]\,\mathrm{d}x = \delta_{mn}$*. Hence* $\{\varphi_n\}_{n\in\mathbb{Z}}$ *is an orthonormal set.*

(b) $\varphi_1, \varphi_2, \varphi_3, \ldots$ *given by* $\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(nx), \ldots$ *are orthonormal on* $[-\pi,\pi]$*.*

(c) Legendre polynomials $\{P_n(x)\}$ *are orthogonal with* $\|P_n\|_2 = \sqrt{\frac{2}{2n+1}}$ *on* $[-1,1]$*:*

$$P_n(x) = \sum_{k=0}^{n} \binom{n}{k}\binom{n+k}{n}\left(\frac{x-1}{2}\right)^k \text{ for } n = 0,1,2,\ldots.$$

Definition (Fourier Coefficients)**.** Let $\{\varphi_n\}_{n\in\mathbb{Z}}$ be an orthonormal set on $[a,b]$. Suppose $f(x) = \sum_{m=0}^{N} c_m\varphi_m(x)$ with $\{\varphi_n\}_{n\in\mathbb{Z}}$ orthonormal on $[a,b]$. Then

$$< f,\varphi_n > =< \sum_{m=0}^{N} c_m\varphi_m, \varphi_n >= \sum_{m=0}^{N} c_m < \varphi_m,\varphi_n > = \sum_{m=0}^{N} c_m\delta_{mn} = c_n.$$

I.e., $c_n =< f,\varphi_n >= \int_a^b f(x)\overline{\varphi_n(x)}\mathrm{d}x$. E.g., if $f(x) = \sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N} c_n\sqrt{2\pi}\cdot\left(\frac{1}{\sqrt{2\pi}}e^{inx}\right)$, then $\sqrt{2\pi}c_n =< f,\varphi_n >= \int_{-\pi}^{\pi} f(x)\frac{1}{\sqrt{2\pi}}e^{-inx}\mathrm{d}x$; i.e., $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\mathrm{d}x$.

Question: When does the Fourier series converge and when does it equal f?

Theorem 8.11. Suppose $f \in \mathscr{R}[a,b]$ and $\{\varphi_m\}_{m=1,2,\ldots}$ is orthonormal. Let $c_n =< f,\varphi_n >, s_n = \sum_{m=1}^{n} c_m\varphi_m, t_n = \sum_{m=1}^{n} a_m\varphi_m$ for some $a_m \in \mathbb{C}$. Then

$$\|f - s_n\|_2 \leq \|f - t_n\|_2,$$

where the equality holds if and only if $a_m = c_m$ for all $m = 1,2,\ldots$.

Proof. $\|f-t_n\|_2^2 = < f-t_n, f-t_n > = < f,f > - < t_n, f > - < f, t_n > + \underbrace{< t_n, t_n >}_{=\|t_n\|_2^2}$. Note

$$\|t_n\|_2^2 = < t_n, t_n > = \sum_{k,m=1}^{n} a_k \overline{a_m} \underbrace{< \varphi_k, \varphi_m >}_{=\delta_{k,m}} = \sum_{m=1}^{n} a_m \overline{a_m} \qquad (*)$$

$$< f, t_n > = \sum_{m=1}^{n} \overline{a_m} \underbrace{< f, \varphi_m >}_{=c_m} = \sum_{m=1}^{n} \overline{a_m} c_m.$$

$$\|f-t_n\|_2^2 = \|f\|_2^2 + \sum_{m=1}^{n} \underbrace{a_m\bar{a}_m - \bar{a}_m c_m - a_m \bar{c}_m + c_m \bar{c}_m}_{=|a_m-c_m|^2} - \underbrace{\sum_{m=1}^{n} c_m \bar{c}_m}_{=\|s_n\|_2^2 \text{ by } (*)} .$$

Thus, $\|f-t_n\|_2^2 = \|f\|_2^2 - \|s_n\|^2 + \sum_{m=1}^{n} |a_m - c_m|^2$. With $a_m = c_m$, this gives $\|f-s_n\|_2^2 = \|f\|_2^2 - \|s_n\|_2^2 + 0$. Hence,

$$\|f-t_n\|_2^2 = \|f-s_n\|_2^2 + \underbrace{\sum_{m=1}^{n} |a_m - c_m|^2}_{\geq 0},$$

so

$$\|f-s_n\|_2^2 \leq \|f-t_n\|_2^2,$$

with equality if and only if $a_m = c_m$ for all $m = 1, 2, \ldots$. Q.E.D.

Note. We have proven that $\|f\|_2^2 = \|s_n\|_2^2 + \|f-s_n\|_2^2$. Let V_n be a subspace of linear combinations of $\{\varphi_m\}_{m=1,2,\ldots,n}$; i.e., $V_n = \{\sum_{m=1}^{n} a_m \varphi_m : a_m \in \mathbb{C}\}$. Then $s_n = L^2$ projection of f onto V_n. Also, $\|s_n\|_2^2 = \sum_{m=1}^{n} |c_m|^2 \leq \|f\|_2^2$. If we have an infinite set $\{\varphi_m\}_{m\in\mathbb{N}}$ then we can let $n \to \infty$, and we get $\sum_{m=1}^{\infty} |c_m|^2 \leq \|f\|_2^2$. This is called *Bessel's inequality.* In particular, $\lim_{n\to\infty} c_n = 0$.

Example. *If $f \in \mathscr{R}[-\pi, \pi]$, then*

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} f(x) \cos(nx) \mathrm{d}x = 0.$$

This is an instance of Riemann-Lebesgue lemma.

From here on, we restrict to $f : [-\pi, \pi] \to \mathbb{C}$ with $f \in \mathscr{R}[-\pi, \pi]$ and take $\varphi_n(x) = e^{inx}, n \in \mathbb{Z}$. We extend f by 2π-periodicity to $\mathbb{R}$; i.e., $f(x+2\pi) = f(x)$ for all $x \in \mathbb{R}$.

In order to make $\{\varphi_n\}$ orthonormal, we change the definition of $< f, g >$ to

$$< f, g >= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\mathrm{d}x,$$

and

$$\|f\|_2^2 =< f, f >= \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 \,\mathrm{d}x.$$

Then, $< \varphi_m \varphi_n >= \delta_{mn}$, so Theorem 8.11 still holds.

Definition (Fourier Series).

- For $f \in \mathscr{R}[-\pi, \pi]$, the *Fourier series* of f is

$$s(x) = \sum_{n=-\infty}^{\infty} c_n \varphi_n(x)$$

with $c_n =< f, \varphi_n >$.

> **Remark.** At this point, we don't claim that the series converges or equals $f(x)$

- The $N^{\ th}$ *partial Fourier series* of f is

$$s_N(x) = \sum_{n=-N}^{N} c_n e^{inx}.$$

We have seen that $\sum_{n=-N}^{N} |c_n|^2 = \|s_N\|_2^2 \leq \|f\|_2^2$.

Lemma. For $x \in \mathbb{R}$, $S_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_N(t)\mathrm{d}t$ with $D_N(t) = \sum_{n=-N}^{N} e^{int} = \frac{\sin{(N+\frac{1}{2})t}}{\sin{\frac{t}{2}}}$.

> **Remark.** D_N is called the Dirichlet kernel.

> **Note.** (a)
>
> $$\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(t)\mathrm{d}t =< D_N, \underbrace{1}_{\varphi_0} >= \sum_{n=-N}^{N} \underbrace{< \varphi_n, \varphi_0 >}_{=\delta_{n0}} = 1.$$
>
> (b)
>
> $$D_N(t) = \frac{1}{\sin{(\frac{t}{2})}} \left[\sin{(\frac{t}{2})}\cos{(Nt)} + \cos{(\frac{t}{2})}\sin{(Nt)}\right] = \cos{(Nt)} + \cot{(\frac{t}{2})}\sin{(Nt)}.$$

Proof. D_N :

$$D_N(t) = \sum_{n=-N}^{N} e^{int} = e^{-iNt} \sum_{k=0}^{2N} (e^{it})^k$$

$$= e^{-iNt} \cdot \frac{e^{it(2N+1)} - 1}{e^{it} - 1} \cdot \frac{e^{-\frac{it}{2}}}{e^{-\frac{it}{2}}} = \frac{e^{i(N+\frac{1}{2})t} - e^{-i(N+\frac{1}{2})t}}{e^{i\frac{t}{2}} - e^{-i\frac{t}{2}}} = \frac{\sin{(N+\frac{1}{2})t}}{\sin\frac{t}{2}}.$$

S_N :

$$S_N(x) = \sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N} < f, \varphi_n > e^{inx}$$

$$= \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{e^{int}} \mathrm{d}t e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=-N}^{N} e^{in(x-t)} \mathrm{d}t$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t) \mathrm{d}t$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_N(t) \mathrm{d}t.$$

(Last equality left as an exercise)

Q.E.D.

Theorem 8.14. Let $x \in \mathbb{R}$ and $f \in \mathscr{R}[-\pi,\pi]$ with $f(x+2\pi) = f(x)$. Suppose there exist δ, M such that

$$|f(x+t) - f(x)| \leq M\,|t| \text{ for } |t| \leq \delta.$$

Then $\lim_{N\to\infty} S_N(x) = f(x)$.

Remark. Such f is said to be *Lipschitz continuous* at x.

Proof.

$$f(x)-S_N(x)=\underbrace{f(x)\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi}D_N(t)\mathrm{d}t}_{=1}-\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x-t)D_N(t)\mathrm{d}t$$
$$=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(x)-f(x-t)\right]D_N(t)\mathrm{d}t$$
$$=\underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(x)-f(x-t)\right]\cos\left(Nt\right)\mathrm{d}t}_{*}+\underbrace{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[f(x)-f(x-t)\right]\cot\left(\frac{t}{2}\right)\sin\left(Nt\right)\mathrm{d}t}_{**}.$$

Since $f(x)-f(x-t)\in\mathscr{R}(dt)$, $*\to 0$ as $N\to\infty$ by Riemann-Lebesgue lemma.
For $**$, note that $\left[f(x)-f(x-t)\right]\cot\left(\frac{t}{2}\right)=\frac{f(x)-f(x-t)}{t}\cdot 2\frac{\frac{t}{2}}{\sin\left(\frac{t}{2}\right)}\cos\left(\frac{t}{2}\right)$.
Since $\frac{f(x)-f(x-t)}{t}$ is bounded by hypothesis and $\cos\left(\frac{t}{2}\right)$ is continuous, $**\to 0$ as $N\to\infty$ by Riemann-Lebesgue lemma. Q.E.D.

Corollary 8.15.
- If $f(x)=0$ for all $x\in(x_0-\varepsilon,x_0+\varepsilon)$ then $S_N(f;x)\to 0$ for all such x because f is Lipschitz continuous at x.
- If $f(x)=g(x)$ for all $x\in(x_0-\varepsilon,x_0+\varepsilon)$, then by linearity, $S_N(f;x)-S_N(g;x)\to 0$ as $N\to\infty$ for all such x.

Example. *Suppose*

$$f(x)=\begin{cases}0 & x\in(0,\pi)\\ \frac{3}{\pi}x & x\in(\pi,\frac{4}{3}\pi)\\ 1 & x\in(\frac{4}{3}\pi,\frac{5}{3}\pi)\\ -\frac{3}{\pi}(x-2\pi) & x\in(\frac{5}{3}\pi,2\pi)\\ f(x)=f(x+2\pi) & \textit{otherwise}\end{cases}$$

$$g(x)=\begin{cases}0 & x\in(0,\pi)\\ \frac{2}{\pi}x & x\in(\pi,\frac{3}{2}\pi)\\ -\frac{2}{\pi}(x-2\pi) & x\in(\frac{3}{2}\pi,2\pi)\\ g(x)=g(x+2\pi) & \textit{otherwise}\end{cases}.$$

Then f,g have different Fourier series, but both converge to zero on $[0,\pi]$: Localisation Principle. Contract this to the behaviour of power series in Theorem 8.5.

Theorem 8.15. If $f:\mathbb{R}\to\mathbb{C}$ is continuous and 2π-periodic, then for all $\varepsilon>0$, there exists a trigonometric polynomial $P(x)=\sum_{n=-N}^{N}c_ne^{inx}$ such that $\|f-P\|=\sup_{x\in\mathbb{R}}|f(x)-P(x)|<\varepsilon$.

Proof. Let $T = \{z \in \mathbb{C} \mid |z| = 1\}$, the unit circle in $\mathbb{C}$.
T is compact. Define $F : T \to \mathbb{C}$ by $F(z) = f(t)$, where $z = e^{it}$. This is well defined by the 2π-periodicity of f.
Let $\mathscr{A}$ = Algebra of trigonometric polynomials $\sum_{n=-N}^{N} a_n z^n$, $z \in T, a_n \in \mathbb{C}, N = 0, 1, 2, \ldots$. Then $\mathscr{A}$ vanishes at no point since it contains 1.
Also, $\mathscr{A}$ separates points and self-adjoint since

$$\overline{\left(\sum_{n=-N}^{N} a_n z^n\right)} = \sum_{n=-N}^{N} \overline{a_n z^n} = \sum_{n=-N}^{N} \overline{a_n} z^{-n}$$

$$\underbrace{=}_{m=-n} \sum_{m=-N}^{N} \overline{a_{-m}} z^m \in \mathscr{A}.$$

By Stone-Weierstrass, $\overline{\mathscr{A}} = \mathscr{C}(T)$, so given $\varepsilon > 0$, there exists $\mathscr{P} \in \mathscr{A}$ such that $|F(z) - \mathscr{P}(z)| < \varepsilon$ for all $z \in T$. Write $\mathscr{P}(z) = \sum_{n=-N}^{N} a_n z^n$ and let $P(x) = \sum_{n=-N}^{N} c_n e^{inx} = \mathscr{P}(e^{ix})$. Then $|f(x) - P(x)| = \left|F(e^{ix}) - \mathscr{P}(e^{ix})\right| < \varepsilon$ for all $x \in \mathbb{R}$. Q.E.D.

Remark. (a) $\mathscr{P}$ in Theorem 8.15 need not involve the Fourier coefficients. In fact, there exists a continuous f whose Fourier series does not converge pointwise.

(b) For a continuous f, Rudin's Problem 8.15 gives an explicit sequence of trigonometric polynomials that converge uniformly to f (Cesaro summation of Fourier partial sums).

(c) (theorem not in Rudin's) If f is 2π-periodic and continuously differentiable then its Fourier series converge uniformly to f (proof omitted).

(d) Theorem (Kolmogorov): There exists a Lebesgue integrable function f on $[-\pi, \pi]$ whose Fourier series diverges everywhere.

Notation.

- For $\{c_n\}_{n\in\mathbb{Z}}$ and $\{\gamma_n\}_{n\in\mathbb{Z}}$, let $(c, \gamma) = \sum_{n=-\infty}^{\infty} c_n \overline{\gamma}_n$ when the series converges.
- $< f, g >= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\mathrm{d}x$, $\|f\|_2^2 =< f, f >$.
- $\varphi_n(x) = e^{inx}, S_N(f) = \sum_{n=-N}^{N} < f, \varphi_n > \varphi_n$.

Problem (6.10)**.** Cauchy-Schwarz inequality: $|< f, g >| \leq \|f\|_2 \|g\|_2$.
Problem (6.11)**.** Minkowski inequality: $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$.
(so $\|f - g\|_2 \leq \|f - h\|_2 + \|h - g\|_2$)
Problem (6.12)**.** For $f \in \mathscr{R}[-\pi, \pi], \varepsilon > 0$, there exists a continuous function (piecewise-linear) h s.t. $\|f - h\| < \varepsilon$. If $f(-\pi) = f(\pi)$ then we can choose h s.t. $h(-\pi) = h(\pi)$.

Theorem 8.16. For $f, g \in \mathscr{R}[-\pi, \pi]$ that are 2π-periodic, let $c_n =< f, \varphi_n >, \gamma_n =< g, \varphi_n >$. Then $\lim\limits_{N\to\infty} \|f - s_N(f)\|_2 = 0$ (convergence of $s_N(f)$ to f in L^2-norm), and $< f, g >= (c, \gamma)$ (Parseval's relation). In particular, $\|f\|_2^2 = (c, c)$; i.e., $\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 \, dx = \sum_{n=-\infty}^{\infty} |c_n|^2$. This is called Bessel's equality.

Proof (8.16). **L^2 convergence:**

Let $\varepsilon > 0$.
Choose a continuous h (Problem 6.12) s.t. $\|f - h\|_2 < \frac{\varepsilon}{3}$.
Then

$$\|s_N(f) - f\|_2 \leq \underbrace{\|s_N(f) - s_N(h)\|_2}_{(*)} + \underbrace{\|s_N(h) - h\|_2}_{(**)} + \underbrace{\|h - f\|_2}_{<\frac{\varepsilon}{3}}.$$

Note $(*) = \|s_N(f - h)\|_2 \leq \|f - h\|_2$ from the proof of Bessel's inequality, so

$$(*) < \frac{\varepsilon}{3}.$$

Since h is continuous and by Theorem 8.15, there exists a trigonometric polynomial P s.t. $\|P - h\|_\infty < \frac{\varepsilon}{3}$ and hence

$$\|P - h\|_2 \leq \|P - h\|_\infty < \frac{\varepsilon}{3}.$$

Say $\deg(P) = N_0$. By Theorem 8.11, for $N \geq N_0$,

$$\|s_N(h) - h\|_2 \leq \|P - h\|_2 < \frac{\varepsilon}{3}.$$

Therefore, for $N \geq N_0$, we have

$$\|s_N(f) - f\|_2 < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Hence, $\lim\limits_{N\to\infty} \|s_N(f) - f\|_2 = 0$.

> **Remark.** Good proof to master.

> **Note.** $\|f\|_2 \leq \|f\|_\infty$ for all $f \in \mathscr{R}[-\pi, \pi]$ since $\|f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^2 \, dx \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \|f\|_\infty^2 dx = \|f\|_\infty^2$.

Parseval relation:

Note $< s_N(f), g >= \sum_{n=-N}^{N} c_n \underbrace{< \varphi_n, g >}_{\overline{\gamma}_n}$.

$$\left| < f, g > - \sum_{n=-N}^{N} c_n \overline{\gamma}_n \right| = |< f, g > - < s_N(f), g >| = |< f - s_N(f), g >| \leq \|f - s_N(f)\|_2 \|g\|_2 \to 0$$

as $N \to \infty$, where the last inequality follows from Cauchy-Schwarz inequality (Problem 6.10).

Note. We showed the symmetric sum $\sum_{n=-N}^{N} c_n \overline{\gamma}_n$ converges to $< f, g >$, but the series also converges absolutely. Since

$$\sum_{n=-N}^{N} |c_n \overline{\gamma}_n| \leq \left(\sum_{n=-N}^{N} |c_n|^2 \right)^{\frac{1}{2}} \cdot \left(\sum_{n=-N}^{N} |\overline{\gamma}_n|^2 \right)^{\frac{1}{2}}$$
$$\leq \|f\|_2 \cdot \|g\|_2, ,$$

The series $\sum_{n=-\infty}^{\infty} c_n \overline{\gamma}_n$ converges absolutely and its sum is independent of how we make partial sums.

Q.E.D.

Chapter 9

Functions of Several variables

Theorem 9.23. [Bananach's Fixed Point Theorem/Contraction Mapping Theorem]
Let (X, d) be a complete metric space. Suppose $\exists c < 1$ such that the map $\varphi : X \to X$ obeys $d(\varphi(x), \varphi(y)) \leq c \cdot d(x, y)$ for all $x, y \in X$ (φ is a *contraction*). Then $\exists! x \in X$ such that $\varphi(x) = x$.

Proof. **Uniqueness:** If $\varphi(x) = x$ and $\varphi(y) = y$, then $d(x, y) = d(\varphi(x), \varphi(y)) \leq c \cdot d(x, y)$, so $d(x, y) = 0$ and $x = y$.

Existence: Given $x_0 \in X$, let $x_1 = \varphi(x_0), x_2 = \varphi(x_1) = (\varphi \circ \varphi)(x_0)$, $x_{n+1} = \varphi(x_n) = \underbrace{(\varphi \circ \varphi \circ \cdots \circ \varphi)}_{n+1}(x_0)$. Then $d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq c \cdot d(x_n, x_{n-1})$, so by induction, $d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$ and hence for $n > m$,

$$d(x_n, x_m) \leq \sum_{i=m+1}^{n} d(x_i, x_{i-1}) \leq \sum_{i=m+1}^{n} c^{i-1} d(x_1, x_0) \leq \sum_{i=m+1}^{\infty} c^{i-1} d(x_1, x_0) = \frac{c^m}{1-c} d(x_1, x_0).$$

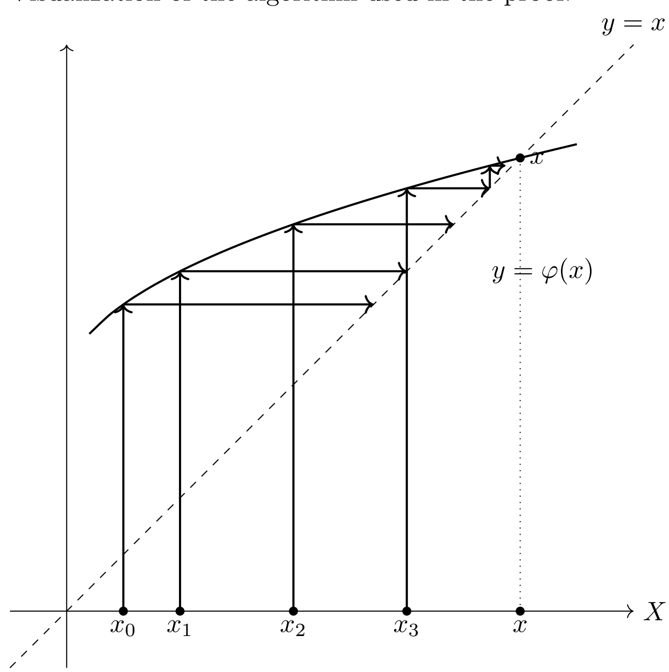
Hence, $\{x_n\}$ is a Cauchy sequence. As X is assumed to be complete, $\exists x \in X$ such that $x_n \to x$. Since φ is continuous $\varphi(x) = \lim_{n\to\infty} \varphi(x_n) = \lim_{n\to\infty} x_{n+1} = x$.

Q.E.D.

Remark. (a) φ is uniformly continuous (take $\delta = \varepsilon$).

(b) The proof gives an exponentially convergent algorithm to find the fixed point.

(c) Visualization of the algorithm used in the proof:



Reading assignment: Rudin's theorems 9.1-9.9.

9.1 Diffrentiation of functions of $f : \mathbb{R}^n \to \mathbb{R}^m$

Recall for $n = m = 1$,

$$f'(x) = \lim_{h \to \infty} \frac{f(x+h) - f(x)}{h}.$$

Equivalently,

$$\lim_{h \to 0} \left| \frac{f(x+h) - f(x) - f'(x)h}{h} \right| = 0.$$

This again is equivalent to

$$f(x+h) = f(x) + \underbrace{f'(x)h}_{\text{best linear approximation to } f(x+h)-f(x)} + \underbrace{r(h)}_{r(h)=o(h)},$$

with $\lim\limits_{h \to 0} \frac{r(h)}{h} = 0$.

Definition 9.11. Let $m, n \in \mathbb{N}$, $E \subset \mathbb{R}^n$ an open set, $f : E \to \mathbb{R}^m, x \in E$. Then f is *differentiable* at x and $f'(x) = A$ if

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Ah\|}{\|h\|} = 0,$$

where $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ depends on x (but not h).
That is, if $f(x+h) = f(x)+Ah+r(h)$ with $\lim_{h \to 0} \frac{\|r(h)\|}{\|h\|} = 0$, then f is differentiable at x and $f'(x) = A$.
Note $L(\mathbb{R}^n, \mathbb{R}^m)$ is the set of all linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Notation. $r(h) = o(h)$ means $\lim_{h \to 0} \frac{\|r(h)\|}{\|h\|} = 0$.

Theorem 9.12. A in the definition of differentiability is unique if it exists.

Proof. Suppose

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + A_1\mathbf{h} + r_1(\mathbf{h})$$

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + A_2\mathbf{h} + r_2(\mathbf{h})$$

.

with $r_1\mathbf{h} = o(\mathbf{h}), r_2\mathbf{h} = o(\mathbf{h})$. Let $B = A_1 - A_2$. Then $B\mathbf{h} = r_2(\mathbf{h}) - r_1(\mathbf{h})$, so for a fixed $\mathbf{h} \neq \mathbf{0}$ and $t > 0$, we have

$$\frac{|B\mathbf{h}|}{|\mathbf{h}|} = \frac{|B(t\mathbf{h})|}{|t\mathbf{h}|} \leq \frac{|r_2(t\mathbf{h})|}{t\mathbf{h}} + \frac{|r_1(t\mathbf{h})|}{t\mathbf{h}}.$$

Since $\frac{|B\mathbf{h}|}{|\mathbf{h}|}$ is independent of t, so the fact $\frac{|B\mathbf{h}|}{|\mathbf{h}|} \to 0$ as $t \to 0$ implies $B\mathbf{h} = \mathbf{0}$ for all $\mathbf{h} \in \mathbb{R}^n$; i.e., $B = A_1 - A_2 = 0$. Q.E.D.

Note. (a) If $f'(\mathbf{x})$ exists for all $\mathbf{x} \in E$ then we can regard f' as a function $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$.

(b) Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear. Then for all $\mathbf{x}, \mathbf{h} \in \mathbb{R}^n$,

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \underbrace{f(\mathbf{h})}_{=A\mathbf{h} \text{ with } A=f} + 0.$$

Hence, $f'(\mathbf{x}) = A = f$ for all $x \in \mathbb{R}^n$.
Note that $\mathbf{x}$ does not appear on the right hand side of the above equality. In fact, $f'(\mathbf{x}) \in L(\mathbb{R}^n, \mathbb{R}^m)$, while $f(\mathbf{x})$ is a vector in $\mathbb{R}^m$. For $n = m = 1$, we can identify a linear map

$$f : \mathbb{R} \to \mathbb{R}, f(x) = ax,$$

where $a \in \mathbb{R}$. Then $f'(x) = a$ is consistent with the above.

(c) If $f'(x)$ exists then f is continuous at x since $\lim_{h \to 0} f(x + h) = f(x) + 0 + 0 = f(x)$.

Definition. For a linear map $A : \mathbb{R}^n \to \mathbb{R}^m$, we define the *norm* of A as

$$\|A\| = \sup_{\|h\|=1} \|Ah\|.$$

Note. $|Ax| = \left|A \frac{x}{|x|}\right| |x| \leq \|A\| \, |x|$ for all $x \in \mathbb{R}^n, x \neq 0$; i.e., $\|A\|$ is the best constant such that $\|Ax\| \leq \|A\| \|x\|$.

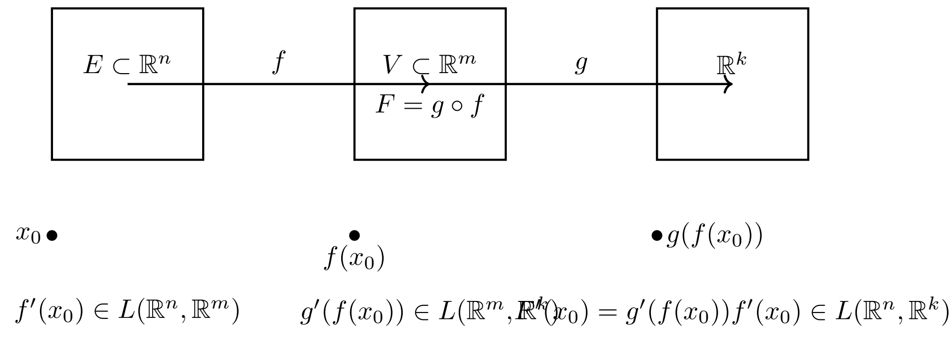
Note. The following facts are immediate from Theorem 9.7 of Rudin's.

(a) $\|A\| < \infty$

(b) For $A : \mathbb{R}^n \to \mathbb{R}^m, B : \mathbb{R}^m \to \mathbb{R}^k$, $\|BA\| \leq \|B\| \cdot \|A\|$

(c) $d(A, B) = \|A - B\|$ defines a metric on $L(\mathbb{R}^n, \mathbb{R}^m)$.

Theorem 9.5. [Chain Rule] Suppose $E \subset \mathbb{R}^n$ is open, $f : E \to \mathbb{R}^m, V \subset \mathbb{R}^m$ is open, $f(E) \subset V$, $g : V \to \mathbb{R}^k$, $x_0 \in E$, $f'(x)$ exists, $g'(f(x_0))$ exists. Let $F = g \circ f : E \to \mathbb{R}^k$. Then $F'(x_0)$ exists and $F'(x) = g'(f(x_0))f'(x_0) \in L(\mathbb{R}^n, \mathbb{R}^k)$, where $g'(f(x_0)) \in L(\mathbb{R}^m, \mathbb{R}^k)$, $f'(x_0) \in L(\mathbb{R}^n, \mathbb{R}^m)$.



Remark. Similar to the proof of Theorem 5.5.

$$F(x_0 + h) - F(x_0) = g(f(x_0 + h)) - g(f(x_0)) = g'(f(x_0)) + \underbrace{v(k)}_{o(k)}$$

$$= g'(f(x_0))\,[f'(x_0)h + u(h)] + v(k) = g'(f(x_0))f'(x_0)h + r(h),$$

where $r(h) = g'(f(x_0))u(h) + v(k)$.
We want to prove that $r(h) = o(h)$; i.e., $\lim\limits_{h \to 0} \frac{|r(h)|}{|h|}$. Note

$$\frac{|g'(f(x_0))u(h)|}{|h|} \le \|g'(f(x_0))\| \cdot \frac{|u(h)|}{|h|} \to 0 \text{ as } h \to 0 \text{ since } u(h) = o(h).$$

$$\frac{|v(k)|}{|h|} = \frac{|v(k)|}{|k|} \cdot \frac{|k|}{|h|} \le \frac{|v(k)|}{|k|} \left[\frac{|f'(x_0)h|}{|h|} + \frac{|u(h)|}{|h|} \right]$$

$$\le \frac{|v(k)|}{|k|} \cdot \left[\|f'(x_0)\| \cdot \frac{|h|}{|\,|h|} + \frac{|u(h)|}{|h|} \right] \to 0 \text{ as } h \to 0.$$

Thus, $r(h) = o(h)$. Q.E.D.

Notation. Standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ are denoted by $e_1, \ldots, e_n$ and $u_1, \ldots, u_m$ respectively, where

$$e_i = (0, \ldots, 0, 1, 0, \ldots, 0),$$

$$u_j = (0, \ldots, 0, 1, 0, \ldots, 0),$$

where the 1 is in the i-th position.
For $f : \mathbb{R}^n \to \mathbb{R}^m$, write

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

so that

$$f = (f_1, \dots, f_m),$$

where $f_j : \mathbb{R}^n \to \mathbb{R}$, $f_j(x) = u_j \cdot f$.

Definition 9.16. [Partial Derivatives] For $i = 1, 2, \dots n$, $j = 1, 2, \dots m$,

$$\frac{\partial f_i}{\partial x_j}(x) = (D_j f_i)(x) = \lim_{t \to 0} \frac{f_i(x + te_j) - f_i(x)}{t} \text{ if the limit exists.}$$

Theorem 9.17. If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $x \in \mathbb{R}^n$, then $(D_i f_j)(x)$ exists for all $i = 1, 2, \dots, n$, $j = 1, 2, \dots m$, and

$$[f'(x)] = \begin{bmatrix} (D_1 f_1)(x) & \dots & (D_n f_1)(x) \\ \vdots & & \vdots \\ (D_1 f_m)(x) & \dots & (D_n f_m)(x) \end{bmatrix}.$$

In other words, $[f'(x)]_{ij} = (D_i f_j)(x) = \frac{\partial f_j}{\partial x_i}(x)$.

Proof. We want to prove that $A = f'(x)$ has matrix elements:

$$u_i \cdot Ae_j = (D_j f_i)(x) = \frac{\partial f_i}{\partial x_j}(x).$$

That is,

$$f_i(+te_j) = f_i(x) + A_{ij}t + o(t).$$

Let

$$\varepsilon(t) = f_i(x + te_j) - f_i(x) - A_{ij}t = u_i \cdot [f(x + te_j) - f(x) - A(te_j)].$$

By Cauchy-Schwartz inequality,

$$|\varepsilon(t)| \leq |u_i| \cdot |f(x + te_j) - f(x) - A(te_j)|.$$

Since $|u_i| = 1$ and

$$|f(x + te_j) - f(x) - A(te_j)| = o(te_j) = o(t) \ \ (\because \text{ by def. of } f'(x)),$$

we have $\varepsilon(t) = o(t)$.

Note. Existence of all partial derivatives does not imply differentiability.

Q.E.D.

Example. *(see also p.215 of Rudin's)*

(a) *(check desmos3d for the graph of)*

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

We have

$$\left.\frac{\partial f}{\partial x}\right|_{(0,0)} = \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t\to 0} \frac{0-0}{t} = 0,$$

$$\left.\frac{\partial f}{\partial y}\right|_{(0,0)} = \lim_{t\to 0} \frac{f(0,t) - f(0,0)}{t} = \lim_{t\to 0} \frac{0-0}{t} = 0.$$

However, f is not continuous at $(0,0)$:

$$\lim_{t\to 0} f(t,at) = \lim_{t\to 0} \frac{at^2}{t^2 + (at)^2} = \lim_{t\to 0} \frac{a}{1+a^2} = \frac{a}{1+a^2} \neq f(0,0) \text{ if } a \neq 0.$$

Definition (Gradient). Let $E \subset \mathbb{R}^n$ be open, $f : E \to \mathbb{R}^1$. If the partial derivatives $\frac{\partial f}{\partial x_i}$ exist for all $i = 1, 2, \dots n$, then the *gradient* of f at $x \in E$ is given by

$$\nabla f(x) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x) e_j = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Hence, $\nabla f : E \to \mathbb{R}^n$.
If $f'(x)$ exists then the gradient is the matrix representation of $f'(x)$.
In particular, if $h = (h_1, h_2, \dots, h_n) \in R^n$, then

$$f'(x)h = \nabla f(x) \cdot h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i.$$

Also, for $u \in \mathbb{R}^n$ with $|u| = 1$,

$$\lim_{t\to 0} \frac{f(x+tu) - f(x)}{t} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(x+tu) = f'(x)u = \nabla f(x) \cdot u.$$

This is the *directional derivative* of f at x in the direction of u, and is denoted by $(D_u f)(x)$. Note that $(D_u f)(x) = \nabla f(x) \cdot u$ is maximal when $u \parallel \nabla f$. Hence, ∇f points in the direction of maximal increase of f at x.

Example. *Let $z = f(x,y) = x^2 + y^2$. Then*

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x, 2y).$$

The maximal rate of increase of f at (x,y) is in direction of $\nabla f(x,y)$.

Definition (Convexity). A $E \subset \mathbb{R}^n$ is *convex* if for all $x, y \in E$ and $\lambda \in [0,1]$, $(1-\lambda)x + \lambda y \in E$.

Example.

Theorem 9.19. Let $E \subset \mathbb{R}^n$ be open and convex, and let $f : E \to \mathbb{R}^m$. Suppose f is differentiable on E and $\|f'\| \leq M$ for all $x \in E$. Then for all $a, b \in E$, we have

$$|f(b) - f(a)| \leq M\,|b - a|\,.$$

Proof. Note

$$f(b) - f(a) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f((1-t)a + tb)\mathrm{d}t$$

by convexity of E, component-wise integration (Theorem 6.23) and Theorem 6.21. We now have

$$f(b) - f(a) = \int_0^1 f'((1-t)a + tb)(b-a)\mathrm{d}t.$$

Since $f'((1-t)a + tb) \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $(b-a) \in \mathbb{R}^n$, we have

$$\begin{aligned} |f(b) - f(a)| &= \left| \int_0^1 f'((1-t)a + tb)(b-a)\mathrm{d}t \right| \\ &\leq \int_0^1 |f'((1-t)a + tb)(b-a)|\,\mathrm{d}t \leq \int_0^1 M\,|b-a|\,\mathrm{d}t \leq M\,|b-a|\,. \end{aligned}$$

Q.E.D.

Corollary 9.20. If in addition, $f'(x) = 0$ all $x \in E$, then f is constant on E.

Problem (9.9)**.** Corollary 9.20 remains for open connected E.

Definition. $f : E \to \mathbb{R}^m$ is continuously differentiable if $f'(x)$ exists for all $x \in E$ and if $f' : E \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous; i.e., if

$$\forall x \in E : \forall \varepsilon > 0 : \exists \delta > 0 \text{ such that } |x - y| < \delta, y \in E \Rightarrow \|f'(x) - f'(y)\| < \varepsilon.$$

Notation. If $f : E \to \mathbb{R}^m$ is continuously differentiable, then we write $f \in \mathcal{C}'(E)$.

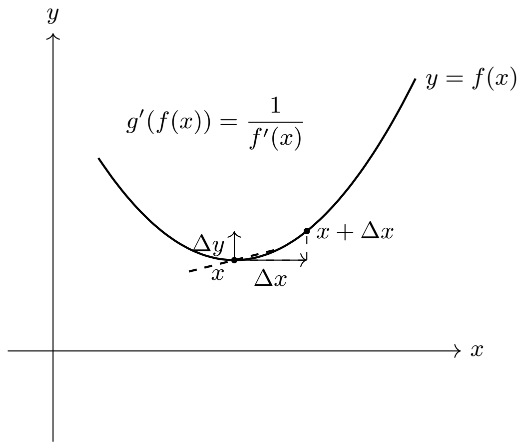
Theorem 9.21. Suppose $f : E \to \mathbb{R}^m$, where E is an open subset of R^n. Then $f \in \mathcal{C}'(E) \Leftrightarrow \frac{\partial f_i}{\partial x_j}$ exists for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$ and are continuous on E.

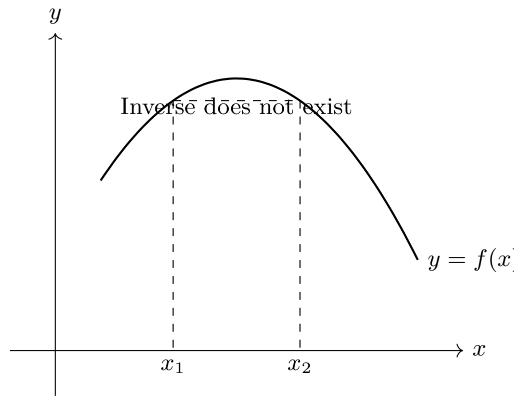
Proof. check out Rudin's proof. Q.E.D.

> **Note.** We've seen an example where all $\frac{\partial f_i}{\partial x_j}$ exist, but $f \notin \mathcal{C}'(E)$ when the partial derivatives are not continuous.

9.2 Inverse Function Theorem

For 1-dimensional case: (check out Rudin's Problem 5.2) If $f'(x) > 0$ for all $x \in (a, b)$ then f is strictly increasing on (a, b) and thus has an inverse function g. This g is differentiable and $g'(f(x)) = \frac{1}{f'(x)}$ for all $x \in (a, b)$.





Theorem 9.24. Suppose $f : E \to \mathbb{R}^n$, where E is an open subset of $\mathbb{R}^n$, is in $\mathcal{C}'(E)$, $a \in E$, $f'(a)$ invertible (c.f., $f'(a) \neq 0$ for $n = 1$). Let $b = f(a)$. Then

(a) there exists an open U containing a and an open V containing b such that $f : U \to V$ is a bijection. [Thus $g = f^{-1}$ exists on V and $g : V \to U$ obeys $g'(f(x)) = x$ for all $x \in E$]

(b) $g \in \mathcal{C}'(V)$.

Remark. (a) By the chain rule, $g'(f(x))f'(x) = I$; i.e., $g'(f(x)) = [f'(x)]^{-1}$ for $x \in U$.

(b) write

$$y_1 = f_1(x_1, \dots, x_n)$$
$$y_2 = f_2(x_1, \dots, x_n)$$
$$\vdots$$
$$y_n = f_n(x_1, \dots, x_n).$$

If $f'(a)$ is invertible, then for y near $b = f(a)$, there exists a unique solution $(x_1, \dots, x_n) \in C'(E)$ to the system of equations such that

$$x_1 = g_1(y_1, \dots, y_n)$$
$$x_2 = g_2(y_1, \dots, y_n)$$
$$\vdots$$
$$x_n = g_n(y_1, \dots, y_n).$$

(c) Invertibility of $f'(a)$ is equivalent to the Jacobian determinant of f at a being non-zero; i.e.,

$$J_f(a) := \begin{vmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \cdots & \frac{\partial f_n}{\partial x_n}(a) \end{vmatrix} \neq 0.$$

Proof (Theroem 9.24).

(a) **Step 1. Existence of open U containing a such that f is one-to-one on U:** Write $A = f'(a)$. Given $y \in \mathbb{R}^n$, let $\varphi : E \to \mathbb{R}^n$ by $\varphi(x) = x + A^{-1}(y - f(x))$. Then $\varphi(x) = x \Leftrightarrow y = f(x)$ $[A^{-1}z = 0 \Leftrightarrow z = 0]$. Therefore, uniqueness of fixed points of φ would imply uniqueness of x such that $y = f(x)$, which is the one-to-one point of our goal.

Note. If φ is a contraction then fixed points are unique (if there is one) since if $\varphi(x_1) = x_1$ and $\varphi(x_2) = x_2$, then

$$|x_1 - x_2| = |\varphi(x_1) - \varphi(x_2)| \le c\,|x_1 - x_2|$$

with $c < 1$ so $x_1 = x_2$.

Now we have $\varphi'(x) = I - A^{-1}f'(x) = A^{-1}(A - f'(x)) = A^{-1}(f'(a) - f'(x))$. Hence,

$$\|\varphi'(x)\| \le \|A^{-1}\| \cdot \underbrace{\|f'(a) - f'(x)\|}_{\le \frac{1}{2\|A^{-1}\|} \text{ if } |x-a| < \delta \text{ for some } \delta > 0}.$$

Thus, if $x \in N_\delta(a)$, then $\|\varphi'(x)\| \le \frac{1}{2}$, and $N_\delta(a)$ is the desired open set U.

Step 2. Openness of V: Let $V = f(U)$. Let $y_0 \in V$, say $y = f(x_0)$, $x_0 \in U$. V is then open if and only if for all $y_0 \in V$, there exists $\varepsilon > 0$ such that $N_\varepsilon(y_0) \subset V$; i.e., $\forall y \in N_\varepsilon(y_0)$: $\exists x \in U$ such that $f(x) = y$.
Given $y \in \mathbb{R}^n$, as before, define $\varphi_y : E \to \mathbb{R}^n$

$$\varphi_y(x) = x + A^{-1}(y - f(x)).$$

Choose $r > 0$ s.t. $\overline{B} = \overline{N_r(x)} \subset U$, which is possible since U is open and $x_0 \in U$. Next, we show that $\varphi_y : \overline{B} \to \overline{B}$ if $y \in N_\varepsilon(y_0)$ with ε sufficiently small. In fact, for $x \in \overline{B}$, we have

$$|\varphi_y(x) - x_0| \le |\varphi_y(x) - \varphi_y(x_0)| + |\varphi_y(x_0) - x_0| \le \frac{1}{2}\,|x - x_0| \le \frac{1}{2}r = \left|A^{-1}(y - f(x_0))\right| = \left|A^{-1}(y -\right.$$

Hence,

$$|\varphi_y(x_0) - x_0| = \left|A^{-1}(y - y_0)\right| \le \|A^{-1}\|\,|y - y_0|\,.$$

Choose $\varepsilon = \frac{1}{2\|A^{-1}\|}r$ for $y \in N_\varepsilon(y_0)$. Then

$$\|A^{-1}\,|y - y_0|\,\| \le \|A^{-1}\|\varepsilon = \frac{1}{2}r.$$

Altogether, we have

$$|\varphi_y(x) - x_0| \le \frac{1}{2}r + \frac{1}{2}r = r.$$

Therefore,

$$\varphi_y(x) \in \overline{B} \text{ for all } x \in \overline{B} \text{ if } y \in N_\varepsilon(y_0).$$

hence, $\varphi_y : \overline{B} \to \overline{B}$, and since φ_y is a contraction and $\overline{B}$ is complete, by Theorem 9.23, φ_y has a unique fixed point $x_* \in \overline{B}$, for which $\varphi(y)(x_*) = x_*$. This x_* obeys $f(x_*) = y$, so $y \in V$. Thus, $N_\varepsilon(y_0) \subset V$. This proves (i). of the theorem.

$g = f^{-1} : V \to U$ **is in** $\mathcal{C}^1(V)$**:** Let $y \in V = f(U)$, $y = f(x), x \in U$. Choose k small enough that $y+k \in V$, say $f(x_K) = y+k$, $x_k \in U$. Let $S = f'(x), T = S^{-1}$.

> **Note.** $(f'(x))^{-1}$ exists by Theorem 9.8 since $f'(x)$ is close to $f'(a)$ and $f'(a)$ is invertible,

We want to show $\frac{|g(\mathbf{y}+\mathbf{k})-g(\mathbf{y})-T\mathbf{k}|}{|\mathbf{k}|} \to 0$ as $k \to 0$. We have

$$\begin{aligned} g(\mathbf{y}+\mathbf{k}) - g(\mathbf{y}) - Tk &= \mathbf{x_k} - \mathbf{x} - Tk \\ &= -T\left[\underbrace{f(\mathbf{x_k}) - f(\mathbf{x})}_{=k} - S(\mathbf{x_k} - \mathbf{x})\right]. \end{aligned}$$

Hence,

$$|g(\mathbf{y}+\mathbf{k}) - g(\mathbf{y}) - Tk| \le \|T\| \left|f(\mathbf{x_k}) - f(\mathbf{x}) - S(\mathbf{x_k} - \mathbf{x})\right|.$$

> **Claim.**
> $$\frac{1}{|k|} \le \frac{2\|A^{-1}\|}{\mathbf{x_k} - \mathbf{x}}.$$
> I.e.,
> $$|\mathbf{x_k} - \mathbf{x}| \le 2\|A^{-1}\| \, |k|.$$

> **Proof.** $\mathbf{x_k} - \mathbf{x} = \varphi_{\mathbf{y}}(\mathbf{x_k}) - \varphi_{\mathbf{y}}(\mathbf{x}) + A^{-1}k$, so
> $$|x_k - x| \le |\varphi_y(x_k) - \varphi_y(x)| + \|A^{-1}\,|k|\,\| \le \frac{1}{2}\,|x_k - x| + \|A^{-1}\,|k|\,\| \le 2\|A^{-1}\|\,|k|.$$
> Q.E.D.

Since we have

$$\frac{|g(\mathbf{y}+\mathbf{k}) - g(\mathbf{y}) - Tk|}{|k|} \le 2\|A^{-1}\| \cdot \|T\| \cdot \frac{|f(\mathbf{x_k}) - f(\mathbf{x}) - S(\mathbf{x_k} - \mathbf{x})|}{|\mathbf{x_k} - \mathbf{x}|},$$

as $k \to 0$, by the claim, $|\mathbf{x_k} - \mathbf{x}| \to 0$, so the right hand side of the inequality goes to 0 as $k \to 0$ by definition of $S = f'(\mathbf{x})$. This shows that $g'(\mathbf{y}) = T$ exists for all $\mathbf{y} \in V = f(U)$. $f : U \to V = (U)$.

(b) Note $g'(\mathbf{y})$ exists for all $\mathbf{y} \in V$. Since $f(g(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{y} \in V$, the chain rule gives $f'(g(\mathbf{y}))g'(\mathbf{y}) = I$; i.e., $g'(\mathbf{y}) = (f'(g(\mathbf{y})))^{-1}$. Since g is continuous on V because of its differentiability, f' is differentiable on U by hypothesis and the inversion is continuous by Theorem **??** for $L(\mathbb{R}^n, \mathbb{R}^n)$. We conclude $g' : V \to L(\mathbb{R}^n, \mathbb{R}^n)$ is continuous.

> **Note.** In coordinates, we have $g(f(x)) = x$, $g'(f(x))f'(x) = I$. The latter implies
>
> $$\sum_{k=1}^{n} \frac{\partial g_i}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_j} = \delta_{ij} \text{ for } i, j = 1, 2, \ldots, n.$$

Q.E.D.

Notation.
For $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$, define $A_{\mathbf{x}} \in L(\mathbb{R}^n, \mathbb{R}^n)$ by $A_{\mathbf{x}}\mathbf{h} = A(\mathbf{h}, \mathbf{0})$ for $\mathbf{h} \in \mathbb{R}^n$. $A_{\mathbf{y}} \in L(\mathbb{R}^m, \mathbb{R}^n)$ by $A_{\mathbf{y}}\mathbf{k} = A(0, \mathbf{k})$ for $\mathbf{k} \in \mathbb{R}^m$.
E.g., if $A = f'(\mathbf{a}, \mathbf{b})$ for $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$, then

$$[f'(\mathbf{a}, \mathbf{b})] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & & & \vdots & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} & \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_m} \end{bmatrix} (\mathbf{a}, \mathbf{b})$$

$$[A_{\mathbf{x}}] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} (\mathbf{a}, \mathbf{b})$$

$$[A_{\mathbf{y}}] = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & & \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_m} \end{bmatrix} (\mathbf{a}, \mathbf{b})$$

$$A = \begin{bmatrix} A_{\mathbf{x}} & A_{\mathbf{y}} \end{bmatrix}.$$

Example (29)**.** *Let $f = (f_1, f_2) : \mathbb{R}^{2+3} \to \mathbb{R}^2$ defined by*

$$f_1(x_1, x_2, y_1, y_2, y_3) = 2e^{x_1} + x_2 y_1 - 4y_2 + 3,$$

$$f_2(x_1, x_2, y_1, y_2, y_3) = x_2 \cos(x_1) - 6x_1 + 2y_1 - y_3.$$

For $\mathbf{a} = (0, 1), \mathbf{b} = (3, 2, 7)$, we have $f(\mathbf{a}, \mathbf{b}) = 0$. Problem: if $\mathbf{y}$ is near $(3, 2, 7)$, can we find $\mathbf{x} = x(\mathbf{y})$ s.t. $f(\mathbf{x}, \mathbf{y}) = \mathbf{0}$?; i.e., can we solve the implicit equation $f(\mathbf{x}, \mathbf{y}) = 0$ for $\mathbf{x}$ in terms of $\mathbf{y}$.

Remark. This motivates the Implicit Function Theorem.

Theorem 9.28. [Implicit Function Theorem] Let $f : E \to \mathbb{R}$ be $\mathcal{C}'(E)$ where $E \subset \mathbb{R}^{n+m}$ is open.
Suppose $f(\mathbf{a}, \mathbf{b})$ for some $(\mathbf{a}, \mathbf{b}) \in \mathbb{R}^{n+m}$ ($\mathbf{a} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$) and that for $A = f'(\mathbf{a}, \mathbf{b})$, $A_{\mathbf{x}}$ is invertible. Then there exists $U \subset \mathbb{R}^{n+m}$, $W \subset \mathbb{R}^m$ with $(\mathbf{a}, \mathbf{b}) \in U$, $\mathbf{b} \in W$ s.t.

- $\forall \mathbf{y} \in W : \exists! \mathbf{x} \in \mathbb{R}^n$ such that $(\mathbf{x}, \mathbf{y}) \in U$ and $f(\mathbf{x}, \mathbf{y}) = 0$.
- Write $\mathbf{x} = g(\mathbf{y})$.
  Then $g : W \to \mathbb{R}^n$ in $\mathcal{C}'(W)$, $g(\mathbf{b}) = \mathbf{a}$, $f(g(\mathbf{y}), \mathbf{y}) = 0$ for all $\mathbf{y} \in W$ and $g'(\mathbf{b}) = -A_{\mathbf{x}}^{-1} A_{\mathbf{y}}$, where $g'(\mathbf{b}) \in L(\mathbb{R}^m, \mathbb{R}^n)$, $A_{\mathbf{x}} \in L(\mathbb{R}^n, \mathbb{R}^n)$, $A_{\mathbf{y}} \in L(\mathbb{R}^m, \mathbb{R}^n)$.