Real Analysis Notes

Nikhil R.

Contents

1 Results from Set Theory

1.1 Orderings on Sets

Definition. A **partial ordering** on a nonempty set is a relation R (often written $\leq$) that satisfies:

- $x \leq x$ for all (**reflexivity**)
- If $x \leq y$ and $y \leq x$, then $x = y$ (**antisymmetry**)
- If $x \leq y$ and $y \leq z$, then $x \leq z$ (**transitivity**)

Definition. If $(X, \leq)$ is a partially ordered set (poset), then:

- An element $x \in X$ is an **upper bound** of $A \subseteq X$ if $a \leq x$ for all $a \in A$. If $x \in A$ then it is called the **greatest element** (either may not exist).
- An element $x \in X$ is a **lower bound** of $A \subseteq X$ if $x \leq a$ for all $a \in A$. If $x \in A$ then it is called the **least element** (either may not exist).
- A **maximal element** of X is an element $x \in X$ such there $\nexists\, y \in X$ s.t. $y > x$ (the minimal element is defined analogously). These are not unique in general.
- An element $x_0 \in X$ is the **supremum** of A ($\sup A$), or least upper bound, if it is an upper bound for A and $x_0 \leq x$ for any other upper bound x of A. This is unique, but may not exist.
- An element $x_0 \in X$ is the **infimum** of A ($\inf A$), or greatest lower bound, if it is a lower bound for A and $x \leq x_0$ for any other lower bound x of A. This is unique, but may not exist.

Definition. If the relation also satisfies that for all $x, y \in X$, either $x \leq y$ or $y \leq x$, it is a **linear ordering** or **total ordering**. In this case, if they exist, the maximal element, greatest element, and supremum are the same and the minimal element, least element, and infimum are the same.

Definition. A set $(X, \leq)$ is **well-ordered** if it is linearly ordered and every non-empty subset of X has a minimal (least) element.

Example. $(\mathbb{N}, \leq)$ is well-ordered, but $(\mathbb{Z}, \leq)$ is not, as $\mathbb{Z}$ itself has no minimal element.

1.2 Zorn's Lemma and Equivalents

Theorem (Fixed Point). Let $(X, \leq)$ is a poset where every linearly ordered subset has a supremum in X. Then every function $f : X \to X$ obeying $x \leq f(x)$ for all $x \in X$ has a fixed point. Namely, $\exists\, x^* \in X$ s.t. $f(x^*) = x^*$.

Proof. By hypothesis, the empty set is trivially ordered so its sup exists in X. Let $a = \sup \emptyset \in X$, which is the smallest element in X since every element in X is an upper bound for $\emptyset$. Construct a family $\mathcal{A}$ of subsets $A \subseteq X$ such that (a) $a \in A$, (b) $f[A] \subseteq A$, and (c) if $L \subseteq A$ is linearly ordered, then $\sup L \in A$. Let $A_* = \bigcap_{A \in \mathcal{A}} A$. This A_* also satisfies (a), (b), and (c). One can show that A_* must be linearly ordered. So by hypothesis, let $x_* = \sup A_*$. By (c), $x_* \in A_*$. By (b), $f(x_*) \in A_*$, so $f(x_*) \leq \sup A_* = x_*$. By hypothesis, $x_* \leq f(x_*)$. By antisymmetry, $x_* = f(x_*)$. □

Theorem (Supremum implies Maximal). If $(X, \leq)$ is a poset such that every linearly ordered subset has a supremum, then X contains a maximal element.

Proof. Suppose for a contradiction that X has no maximal element. Then for every $x \in X$, the set $A_x = \{y \in X : x < y\}$ is non-empty. By the Axiom of Choice (stated below), $\prod_{x \in X} A_x \neq \emptyset$. Let f be a choice function from this product, so $f(x) \in A_x$, which means $x < f(x)$ for all x. By the fixed point theorem, there must be $x_* \in X$ such that $f(x_*) = x_*$. This implies $x_* < f(x_*) = x_*$, a contradiction. □

Theorem. The following are equivalent statements:

1. **(a) Axiom of Choice.** If $(X_\alpha)_{\alpha \in A}$ is a nonempty collection of nonempty sets, then their Cartesian product $\prod_{\alpha \in A} X_\alpha \neq \emptyset$. Equivalently, there exists a function (called a choice function) $f : A \to \bigcup_{\alpha \in A} X_\alpha$ such that for every $\alpha \in A$, we have $f(\alpha) \in X_\alpha$.

2. **(b) Hausdorff Maximal Principle.** Every partially ordered set contains a maximal linearly ordered set. (A maximal linearly ordered set is a subset E that is linearly ordered, and no subset properly containing E is also linearly ordered.)

3. **(c) Kuratowski-Zorn Lemma.** If X is a partially ordered set and every linearly ordered subset of X has an upper bound, then X has a maximal element.

4. **(d) Well Ordering Principle.** Every nonempty set X can be well-ordered.

Proof. (a) ⇒ (b) Let X be a poset, and let $\mathcal{L}$ be the set of all linearly ordered subsets of X, ordered by inclusion $\subseteq$. Let $\mathcal{M} \subseteq \mathcal{L}$ be a linearly ordered subset (a chain of linearly ordered subsets of X). Define $\mathcal{S} = \bigcup_{M \in \mathcal{M}} M$. If $s_1, s_2 \in \mathcal{S}$, then $s_1 \in M_1$, $s_2 \in M_2$ for some $M_1, M_2 \in \mathcal{M}$. Since $\mathcal{M}$ is linearly ordered, either $M_1 \subseteq M_2$ or $M_2 \subseteq M_1$, so s_1 and s_2 are comparable. Hence $\mathcal{S}$ is linearly ordered, i.e. $\mathcal{S} \in \mathcal{L}$. By definition of union, $\mathcal{S}$ is the least upper bound (supremum) of $\mathcal{M}$ in $(\mathcal{L}, \subseteq)$. Thus every linearly ordered subset of $\mathcal{L}$ has a supremum. By the "Supremum implies Maximal" theorem (which follows from (a)), $\mathcal{L}$ has a maximal element $L \in \mathcal{L}$, which is the maximal linearly ordered subset of X.

(b) ⇒ (c) Let X be a poset where every linearly ordered subset has an upper bound. By (b), there exists a maximal linearly ordered set $L \subseteq X$. By hypothesis, L has an upper bound, call it $a \in X$. Since L is maximal, a must be in L (otherwise $L \cup \{a\}$ would be a larger linearly ordered set). This element a is then a maximal element of X.

(c) ⇒ (d) Let X be a nonempty set. Let $\mathcal{W}$ be the collection of all well-orderings $(E, \leq_E)$ where $E \subseteq X$. Define a partial order on $\mathcal{W}$ such that $(E_1, \leq_1)$ precedes $(E_2, \leq_2)$ if $E_1 \subseteq E_2$, the orderings agree in E_1, and $x \leq_2 y$ for all $x \in E_1$ and $y \in E_2 \setminus E_1$. Any linearly ordered chain in $\mathcal{W}$ has an upper bound (the union of the sets with the union of their orderings). By (c), $\mathcal{W}$ must have a maximal element $(E, \leq)$. This E must be X. If E were a proper subset, we could take $x_0 \in X \setminus E$, form $E' = E \cup \{x_0\}$, and extend $\leq$ to $\leq'$ by declaring $x \leq' x_0$ for all $x \in E$, contradicting the maximality of $(E, \leq)$. Thus X can be well-ordered.

(d) $\Rightarrow$ (a): Let $(X_\alpha)_{\alpha \in A}$ be a nonempty collection of nonempty sets. Let $X = \bigcup_{\alpha \in A} X_\alpha$. By (d), X can be well-ordered. Since each X_α is a non-empty subset of the well-ordered set X, each X_α has a minimal (least) element. We define a choice function f by setting $f(\alpha)$ to be this minimal element of X_α. Then $f \in \prod_{\alpha \in A} X_\alpha$, proving the product is non-empty. $\square$

1.3 Cardinality

Definition. If X and Y are nonempty sets, we define the **cardinality** relations by:

- $\operatorname{card}(X) = \operatorname{card}(Y)$ means there exists a **bijective** function $f : X \to Y$.
- $\operatorname{card}(X) \leq \operatorname{card}(Y)$ means there exists an **injective** function $f : X \to Y$.
- $\operatorname{card}(X) \geq \operatorname{card}(Y)$ means there exists a **surjective** function $f : X \to Y$.

Proposition. $\operatorname{card}(X) \leq \operatorname{card}(Y) \iff \operatorname{card}(Y) \geq \operatorname{card}(X)$.

Proof. ($\Rightarrow$) Assume $\operatorname{card}(X) \leq \operatorname{card}(Y)$. This means there is an injection $f : X \to Y$. f provides a bijection $f : X \to f[X]$ where $f[X] \subseteq Y$. Pick any $x_0 \in X$. We can define a surjection $g : Y \to X$ such that $g(y) = \begin{cases} f^{-1}(y) & \text{if } y \in f[X] \\ x_0 & \text{if } y \in Y \setminus f[X] \end{cases}$. This function g is surjective because its range includes all of X. Thus, $\operatorname{card}(Y) \geq \operatorname{card}(X)$.

($\Leftarrow$) Assume $\operatorname{card}(Y) \geq \operatorname{card}(X)$. This means there is a surjection $g : Y \to X$. For each $x \in X$, consider the non-empty set $g^{-1}[\{x\}] \subseteq Y$ (the set of elements in Y that map to x). These sets are disjoint for different x's. By the axiom of choice, the product $\prod_{x \in X} g^{-1}[\{x\}]$ is non-empty. Let f be a choice function from this product, so $f(x) \in g^{-1}[\{x\}]$ for each x. This function $f : X \to Y$ is an injection, because if $x_1 \neq x_2$, then $f(x_1) \in g^{-1}[\{x_1\}]$ and $f(x_2) \in g^{-1}[\{x_2\}]$, where the two sets are disjoint, so $f(x_1) \neq f(x_2)$. Thus, $\operatorname{card}(X) \leq \operatorname{card}(Y)$. $\square$

Proposition. For any two sets X and Y, either $\operatorname{card}(X) \leq \operatorname{card}(Y)$ or $\operatorname{card}(Y) \leq \operatorname{card}(X)$.

Proof. Let $\mathcal{I}$ be the set of all injective functions $f : A \to B$ where $A \subseteq X$ and $B \subseteq Y$. $\mathcal{I}$ is a poset by function extension; namely $f \preceq g$ if $\operatorname{dom}(f) \subseteq \operatorname{dom}(g)$ and $f(x) = g(x)\ \forall\ x \in \operatorname{dom}(f)$. Let $\mathcal{C} \subseteq \mathcal{I}$ be linearly ordered. Define $f^* = \bigcup_{f \in \mathcal{C}} f$. This is an upper bound for $\mathcal{C}$. Thus every linearly ordered subset of $\mathcal{I}$ has an upper bound meaning by Zorn's lemma, there exists a maximal element $f \in \mathcal{I}$. Then let $\operatorname{dom}(f) = A \subseteq X$ and $\operatorname{cod}(f) = B \subseteq Y$. We claim that either $A = X$ or $B = Y$. Since f is injective, if both $A \neq X$ and $B \neq Y$, one could find $x_0 \in X \setminus A$ and $y_0 \in Y \setminus B$ and extend f to $\tilde{f}$ with $\tilde{f}(x_0) = y_0$, contradicting $f's$ maximality. If $A = X$, then f is an injection from X into Y, so $\operatorname{card}(X) \leq \operatorname{card}(Y)$. If $B = Y$, in fact, $f[X] = Y$ in this case, since if it weren't we me extend f by picking some unused element in A (which must exist, otherwise we go back to the other case). Thus, we have a bijection from a subset of A to Y, meaning $\text{card}(Y) \leq \text{card}(X)$. $\square$

Lemma (Banach). If there are injections $f : A \to B$ and $g : B \to A$, then there are partitions of $A = A_1 \sqcup A_2$ and $B = B_1 \sqcup B_2$ such that $f[A_1] = B_1$ and $g[B_2] = A_2$.

Proof. Define a map $\Phi : \mathcal{P}(A) \to \mathcal{P}(A)$ by $X \mapsto A \setminus g[B \setminus f[X]]$. Note that Φ is monotone: if $X \subseteq Y$, then $f[X] \subseteq f[Y]$, so $B \setminus f[Y] \subseteq B \setminus f[X]$, hence $g(B \setminus f[Y]) \subseteq g(B \setminus f[X])$, and hence $\Phi(X) = A \setminus g(B \setminus f[X]) \subseteq A \setminus g(B \setminus f[Y]) = \Phi(Y)$. Construct an increasing chain of subsets of A: $X_0 := \emptyset$ and $X_{n+1} := \Phi(X_n)$ for $n \geq 0$. Let $\Omega := \bigcup_{n=0}^{\infty} X_n$. Since is injective, note that Φ preserves unions: for any family $(X_t)_{t \in T} \subseteq \mathcal{P}(A)$, we have $\Phi(\bigcup_{t \in T} X_t) = \bigcup_{t \in T} \Phi(X_t)$. Thus, $\Phi(\Omega) = \Phi(\bigcup_n X_n) = \bigcup_n \Phi(X_n) = \bigcup_n X_{n+1} = \Omega$, meaning Ω is a fixed point of Φ. Define the partitions: $A_1 := \Omega$, $A_2 := A \setminus A_1$, $B_1 := f[A_1]$, and $B_2 := B \setminus B_1$. From $\Phi(A_1) = A_1$, we have $A_1 = A \setminus g[B \setminus f[A_1]] = A \setminus g[B_2]$, hence $A_2 = g[B_2]$. $\square$

Theorem (Cantor-Bernstein-Schröder). If $\text{card}(X) \leq \text{card}(Y)$ and $\text{card}(Y) \leq \text{card}(X)$, then $\text{card}(X) = \text{card}(Y)$. Alternatively, if there is an injection $f : X \to Y$ and an injection $g : Y \to X$, then there must exist a bijection $h : X \to Y$.

Proof. Assume there are injections $f : X \to Y$ and $g : B \to A$. We use the Banach's Lemma to find partitions of A and B such that $f[A_1] = B_1$ and $g[B_2] = A_2$. The bijection h is then constructed as:

$$h(x) = \begin{cases} f(x) & \text{if } x \in A_1 \\ g^{-1}(x) & \text{if } x \in A_2 \end{cases}$$

$\square$

Example. $\text{card}([0,1]) = \text{card}((0,1))$

Proof. $f(x) = x$ is an injection from $(0,1)$ to $[0,1]$, so $\text{card}((0,1)) \leq \text{card}([0,1])$. $g(x) = \frac{x}{4} + \frac{1}{2}$ is an injection from $[0,1]$ to $(0,1)$ (it maps $[0,1]$ to $[\frac{1}{2}, \frac{3}{4}]$). Thus, by the Cantor-Bernstein-Schröder theorem, their cardinalities are equal. $\square$

Theorem (Cantor). For any set X, we have $\text{card}(X) < \text{card}(\mathcal{P}(X))$.

Proof. $\text{card}(X) \leq \text{card}(\mathcal{P}(X))$ because the map $f(x) = \{x\}$ is an injection from X to $\mathcal{P}(X)$. Now we must show there is no surjection. Suppose that $g : X \to \mathcal{P}(X)$ is a surjection. Consider the set $Y = \{x \in X : x \notin g(x)\}$. Since $Y \in \mathcal{P}(X)$, and g is surjective, there must be some $x_0 \in X$ such that $g(x_0) = Y$. Now, if $x_0 \in Y$, then by definition of , $x_0 \notin g(x_0) = Y$, a contradiction. If $x_0 \notin Y$, then $x_0 \in g(x_0) = Y$, also a contradiction. Hence both possibilities are impossible, meaning that no surjection $g : X \to \mathcal{P}(X)$ exists. $\square$

1.4 Countable and Uncountable Sets

Definition. We say a set X is **countable** if $\text{card}(X) \leq \text{card}(\mathbb{N})$.

- This includes all finite sets.

- If X is countable but not finite, it is **countably infinite**. In this case, $\text{card}(X) = \text{card}(\mathbb{N})$.

Proposition. 1. If X and Y are countable, so is $X \times Y$.

2. A countable union of countable sets is countable.

Proof. (1) This follows from $\text{card}(\mathbb{N} \times \mathbb{N}) = \text{card}(\mathbb{N})$, which can be shown by "snaking" through the grid of pairs: $(1,1),(1,2),(2,1),(1,3),(2,2),(3,1),\ldots$.

(2) Let A be a countable set and X_α be countable for each $\alpha \in A$. By axiom of choice, there is a surjection $f_\alpha : \mathbb{N} \to X_\alpha \ \forall \ \alpha \in A$. We can define a surjection $f : \mathbb{N} \times \mathbb{N} \to \bigcup_{\alpha \in A} X_\alpha$ by $f(n, \alpha) = f_\alpha(n)$. Since $\mathbb{N} \times \mathbb{N}$ is countable, the union is also countable. $\square$

Proposition. $\mathbb{Z}$ and $\mathbb{Q}$ are countable.

Proof. $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-n : n \in \mathbb{N}\}$ is a finite (countable) union of countable sets. $\mathbb{Q} = \bigcup_{n \in \mathbb{N}} \{\frac{m}{n} : m \in \mathbb{Z}\}$ is a countable union of countable sets (since each set $\{\frac{m}{n} : m \in \mathbb{Z}\}$ is in bijection with $\mathbb{Z}$). It follows that $\mathbb{Q} \times \mathbb{Q}$ is also countable. $\square$

Definition. We say a set is **uncountable** if it is not countable (it follows trivially that only infinite sets are uncountable).

Proposition. The set $\{0,1\}^{\mathbb{N}_0}$ (the set of all infinite sequences of 0s and 1s) is uncountable.

Cantor's Diagonal Argument. Suppose for contradiction that the set is countable. Then we can list all the sequences $\alpha_0, \alpha_1, \alpha_2, \ldots$ as rows of a matrix. Form a new sequence by reversing all the diagonal elements. This new element is not equal to any of the other elements, meaning $\{0,1\}^{\mathbb{N}}$ is not countable. $\square$

Definition. A set X is said to have **cardinality continuum** if $\text{card}(X) = \text{card}(\mathbb{R}) = \mathfrak{c}$.

Theorem. $\text{card}(\mathcal{P}(\mathbb{N})) = \text{card}(\{0,1\}^{\mathbb{N}}) = \text{card}(\mathbb{R})$.

Proof.

1. $\text{card}(\mathcal{P}(\mathbb{N})) = \text{card}(\{0,1\}^{\mathbb{N}})$. We define a function $f : \mathcal{P}(\mathbb{N}) \to \{0,1\}^{\mathbb{N}}$ by mapping as the selector function $f(A) = \chi_A$ where $\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$. Clearly f is bijective.

2. $\text{card}(\mathcal{P}(\mathbb{N}_0)) \leq \text{card}(\mathbb{R})$. Consider the map $f : \mathcal{P}(\mathbb{N}_0) \to \mathbb{R}$ by setting $f(A) = \sum_{n \in A} \frac{2}{3^{n+1}}$. Noting that $\sum_{n>m} \frac{2}{3^{n+1}} = \sum_{k=0}^{\infty} \frac{2}{3^{k+m+2}} = \frac{2}{3^{m+2}} \sum_{k=0}^{\infty} \left(\frac{1}{3}\right)^k = \frac{1}{3^{m+1}} < \frac{2}{3^{m+1}}$, this function is injective since if we miss a term we can never make up for it.

3. $\text{card}(\mathbb{R}) \leq \text{card}(\mathcal{P}(\mathbb{N}_0))$. We know $\text{card}(\mathbb{N}_0) = \text{card}(\mathbb{Q})$, so $\text{card}(\mathcal{P}(\mathbb{N}_0)) = \text{card}(\mathcal{P}(\mathbb{Q}))$. Define $g : \mathbb{R} \to \mathcal{P}(\mathbb{Q})$ by $g(x) = \{r \in \mathbb{Q} : r < x\}$. This Dedekind cut map g is injective.

4. By the Cantor-Bernstein theorem, $\text{card}(\mathcal{P}(\mathbb{N})) = \text{card}(\{0,1\}^{\mathbb{N}}) = \text{card}(\mathbb{R})$.

$\square$

Proposition.

- If $\text{card}(X) \geq \mathfrak{c}$, then X is uncountable.
- $\text{card}(\mathbb{R}\times\mathbb{R}) = \text{card}(\mathbb{R})$; equivalently, $\text{card}(\{0,1\}^{\mathbb{N}_0}\times\{0,1\}^{\mathbb{N}_0}) = \text{card}(\{0,1\}^{\mathbb{N}_0})$. *Sketch:* encode a pair of real binary expansions by interleaving their digits.
- $\text{card}(\mathbb{R}^k) = \text{card}(\mathbb{R})$ for any $k \in \mathbb{N}$. *Sketch:* follows by induction on above.
- If $\text{card}(A) \leq \mathfrak{c}$ and $\text{card}(X_\alpha) \leq \mathfrak{c}$ for all $\alpha \in A$, then $\text{card}(\bigcup_{\alpha\in A} X_\alpha) \leq \mathfrak{c}$. *Sketch:* The union can be injected into $A \times \mathbb{R}$, and applying the above.

Example. $\text{card}(\mathbb{R}) = \text{card}([0,1])$

Proof. The function $f(x) = \frac{x}{|x|+1}$ is an injection from $\mathbb{R} \to (-1,1)$, but we can achieve an injection between $(-1,1)$ and $[0,1]$. The other direction is trivial. By Cantor-Bernstein theorem, we achieve the desired result. $\square$

Example. if $X = \{[n, n+1) : n \in \mathbb{N}\}$ then $\text{card}(X) = \text{card}(\mathbb{N})$.

Proof. The function $f([n, n+1)) = n$ is a bijection. $\square$

Example. Any infinite set of pairwise disjoint closed intervals is countably infinite.

Proof. Define $f : X \to \mathbb{Q}$ by $f(I) = q$, where q is any rational number contained in the interval $I \in X$. Since $\mathbb{Q}$ is dense, q exists. Since the intervals are disjoint, f is injective. Thus, $\text{card}(X) \leq \text{card}(\mathbb{Q}) = \text{card}(\mathbb{N})$. $\square$

2 Construction of $\mathbb{N}$, $\mathbb{Q}$, and $\mathbb{Z}$

2.1 Construction of $\mathbb{N}$

Definition (Peano Axioms). $\mathbb{N}$ satisfies the following **Peano Axioms**. Let A be a set.

1. $\exists\, 0 \in A$
2. $\exists$ a well-defined **successor function** $S : A \to A$
3. S is an injection.
4. $(\forall\, n \in A)\ S(n) \neq 0$
5. $(\forall\, B \subseteq A)$ if $0 \in B$ and B is closed under S (i.e. $n \in B \implies S(n) \in B$), then $B = A$. This is the **axiom of induction**.

Definition. If (A, S) and (B, T) follow the above, then there is a bijection $f : A \to B$ given by $f(0_A) = 0_B$ and $f(S(a)) = T(f(a))$. We may then define:

- **Addition**: $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $\begin{cases} a + 0 = a \\ a + S(b) = S(a + b) \end{cases}$
- **Multiplication**: $\times : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $\begin{cases} a \times 0 = 0 \\ a \times S(b) = a + (a \times b) \end{cases}$

2.2 Principles on $\mathbb{N}$

Theorem. The following three principles on $\mathbb{N}$ are equivalent:

1. **Induction Principle.** if $A \subseteq \mathbb{N}_0$ with $0 \in A$ and $n \in A \implies n+1 \in A$, then $A = \mathbb{N}_0$
2. **Well-Ordering Principle.** Any non-empty subset $A \subseteq \mathbb{N}_0$ has a minimum number
3. **Maximum Principle.** If $A \subseteq \mathbb{N}_0$ is bounded from above, then A has a maximum

Proof. ($i \implies ii$): Suppose $A \subseteq \mathbb{N}_0$ without a minimum. Then it may not contain 0, 1, ..., by induction, so $A = \emptyset$.

($ii \implies i$): Suppose A satisfies the induction principle. Assume well-ordering. Let $B := \mathbb{N}_0 \setminus A \neq \emptyset$. So B has a smallest element, but it can't be $0 \in A$. By induction on A, $B = \emptyset$, so $A = \mathbb{N}_0$.

($i \implies iii$): Assume well-ordering. Let $A \subseteq \mathbb{N}_0$ be bounded above, with upper bound M. Consider $B := \{M - a : a \in A\}$ which is well-ordered. Let $b \in A$ be s.t. $M - b$ is the minimum of B. Then, $(\forall\, a \in A)\ M - b \leq M - a \implies a \leq b$. Thus A has a maximum.

($iii \implies ii$): Assume maximum principle. Let $A \neq \emptyset \subseteq \mathbb{N}_0$. Let B be the set of A's lower bounds. B is bounded above and contains 0. Let b_0 be the maximum of B. Then $(\forall\, a \in A)$, $(\forall\, b \in B)$ $b \leq b_0 \leq a$. If $b_0 \notin A$, then $b_0 + 1$ is in B, which is a contradiction. Thus $b_0 \in A$ is the smallest element of A. $\square$

2.3 Construction of $\mathbb{Z}$ and $\mathbb{Q}$

Definition. Recall the notion of equivalence relations, and that an equivalence class partitions a set.

- Define $\sim$ on $\mathbb{N} \times \mathbb{N}$ by $(a, b) \sim (c, d) \iff a + d = b + c$. Define $\mathbb{Z} := \{(a, b) : a, b \in \mathbb{N}\}/\sim$. Here we interpret (a, b) as $a - b$.

- Define $\sim$ on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ by $(a, b) \sim (c, d) \iff ad = bc$. Define $\mathbb{Q} := \{(a, b) : a, b \in \mathbb{Z}; b \neq 0\}/\sim$. Here we interpret (a, b) as a/b.

Notice $\mathbb{Q}$ does not have a "first" element. However, we do get Associativity, commutativity, closure for $+, \times$; Identity and inverses for $+, \times$; and distributivity.

Definition. We then define the arithmetic on $\mathbb{Q}$ as:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \quad , \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

3 The Real Field

3.1 Fields

Definition. A nonempty set $\mathbb{F}$ is a **field** if there are well-defined operations $+, \times : \mathbb{F}^2 \to \mathbb{F}$ and the set satisfies:

1. Closure under $+, \times$
2. Associativity, commutativity under $+, \times$
3. Identity and inverses under $+, \times$
4. Distributive property

Example. $\mathbb{Q}$ and $\mathbb{R}$ are fields but $\mathbb{N}$ and $\mathbb{Z}$ are not (no division).

Example. if p is prime then $\mathbb{Z}_p := \{0, 1, \ldots p-1\}$ is a field.

Proof. The only two interesting things to consider here are inverses, since everything else is inherited from $\mathbb{Z}$. Suppose $a \in \mathbb{Z}_p$. Then, $(p-a) \in \mathbb{Z}_p$ and $a+(p-a) = 0$. For multiplication, we must show $\exists\ b \in \mathbb{Z}_p$ s.t. $ab = 1$. We have $\mathrm{GCD}(a, p) = 1$, meaning $ax + py = 1$ for some $x, y \in \mathbb{Z}$, so $ax \equiv 1 \pmod{p}$. □

3.2 Orderings

Definition. Let S be ordered, $E \subseteq S$.

- $\beta \in S$ is an **upper bound** if $(\forall\ e \in E)\ e \le \beta$.
- $\beta \in S$ is a **lower bound** if $(\forall\ e \in E)\ \beta \le e$.
- If $\beta \in E$ is an upper/lower bound, then it is a **maximal element** or **minimal element** of E.
- β is a **supremum** (least upper bound) if it is an upper bound s.t. $\beta \le$ (any other upper bound of E).
- β is an **infimum** (greatest lower bound) if it is a lower bound s.t. (any other lower bound of E) $\le \beta$.

Example. if p is prime then the set $E := \{e : e \in \mathbb{Q}_{>0}, e^2 < p\}$ has no supremum in $\mathbb{Q}$ (i.e. $\sqrt{p} \notin \mathbb{Q}$).

Proof. By contradiction. Suppose $\exists\ a, b \in \mathbb{N}$ s.t. $\sqrt{p} = \frac{a}{b}$. Then $p = \frac{a^2}{b^2}$. Rewrite a, b in their prime decomposition and simplify to get $(a')^2 = p(b')^2$. The parity of the power of p cannot be equal on the LHS and RHS, a contradiction. □

Definition. An ordered set S has the **least upper bound property** if $(\forall\ E \subseteq S)$ if E has an upper bound, then E has a supremum.

Example. An ordered finite set, $\mathbb{N}$, $\mathbb{Z}$ have the property, but not $\mathbb{Q}$.

Theorem. Let S be ordered with the least upper bound property. If $B \neq \emptyset \subseteq S$ is bounded below then $\inf(B)$ exists.

Proof. In particular $\sup(L) = \inf(B)$ where L is the set of lower bounds of B. Every $x \in B$ is an upper bound for L. Thus L is bounded above and $\gamma := \sup(L)$ exists in S. If $\gamma > \beta$ then β is not an upper bound of L. γ is a lower bound for B and an upper bound for L. Thus $\gamma \in L$. If $\beta \in L \implies \beta \leq \gamma$. Thus $\gamma = \inf(B)$. $\square$

Definition. An **ordered field** is a field $\mathbb{F}$ for which there is an order defined and:

1. $(\forall\ x, y, z \in \mathbb{F})\ y < z \implies x + y < x + z$
2. $(\forall\ x, y \in \mathbb{F})\ x, y > 0 \implies x \cdot y > 0$

3.3 Constructing $\mathbb{R}$ from $\mathbb{Q}$

Theorem. There exists an ordered field $\mathbb{R} \supseteq \mathbb{Q}$ that has the least upper bound property.

Definition. A **Dedekind Cut** $\alpha \subseteq \mathbb{Q}$ is a set s.t.:

1. $\alpha \neq \emptyset$ and $\alpha \neq \mathbb{Q}$
2. If $p \in \alpha$, $q \in \mathbb{Q}$, and $q < p$, then $q \in \alpha$
3. If $p \in \alpha$ then $\exists\ r \in \alpha$ s.t. $p < r$ (there is no largest number)

Definition. Define $\mathbb{R} := \{\alpha : \alpha \text{ is a Dedekind cut}\}$ with $\alpha < \beta$ iff $\alpha \subset \beta$.

- *Claim:* This is a total order.

  Proof. Subset relations are a partial order. Assume $\alpha \not\subseteq \beta$ and $\alpha \neq \beta$. We show $\beta \subset \alpha$. Let $p \in \alpha, p \notin \beta$. For any $q \in \beta$, by (iii), $q < p$, so by (ii), $q \in \alpha$. Thus $\beta \subseteq \alpha$, and since they are not equal, $\beta \subset \alpha$. $\square$

- *Claim:* $\mathbb{R}$ has the least upper bound property.

  Proof. Let $A \neq \emptyset \subseteq \mathbb{R}$, with upper bound $\beta \in \mathbb{R}$. Define $\gamma := \bigcup_{\alpha \in A} \alpha$. We claim $\gamma = \sup(A) \in \mathbb{R}$.

  - Since $A \neq \emptyset$, take $\alpha_0 \in A$. Thus $\gamma \neq \emptyset$. Since $(\forall\ \alpha \in A)$, $\alpha \subset \beta$, we have $\gamma \subseteq \beta$, meaning $\gamma \neq \mathbb{Q}$.
  - If $p \in \gamma$ and $q < p$, then $p \in \alpha_1$ for some $\alpha_1 \in A$. Since α_1 is a cut, $q \in \alpha_1$, meaning $q \in \gamma$.
  - If $p \in \gamma$, then $\exists\ \alpha_2 \in A$ s.t. $p \in \alpha_2$. α_2 is a Dedekind cut, so $\exists\ r \in \alpha_2$ s.t. $r > p$. Then $r \in \gamma$.

- Thus γ is a Dedekind cut and an upper bound for A. If $\delta < \gamma$, take $s \in \gamma$ s.t. $s \notin \delta$. Then $\exists\ \alpha \in A$ s.t. $s \in \alpha$. Thus $\delta < \alpha$, meaning δ is not an upper bound of A.

$\square$

Definition. We can define operations on $\mathbb{R}$:

- **Addition:** $(\forall\ \alpha, \beta \in \mathbb{R})\ \alpha + \beta := \{r + s : r \in \alpha, s \in \beta\}$.
- **Multiplication:** $(\forall\ \alpha, \beta \in \mathbb{R}_+)$, where $\mathbb{R}_+ := \{\alpha \in \mathbb{R} : \alpha > 0^*\}$, we can define $\alpha \cdot \beta$.

3.4 Theorems on $\mathbb{R}$

Theorem (Nested Interval Property). If $I_n := [a_n, b_n]$ and $(\forall\ n \in \mathbb{N})$, $I_n \supseteq I_{n+1}$, then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$.

Proof. Take $A = \{a_n : n \in \mathbb{N}\}$. A is bounded above and $x = \sup(A)$. This satisfies $(\forall\ n \in \mathbb{N})$, $a_n \leq x \leq b_n$. $\square$

Theorem (Archimedean Property). $(\forall\ x, y \in \mathbb{R}, x > 0)$, $\exists\ n \in \mathbb{N}$ s.t. $nx > y$.

Proof. Let $A := \{nx : n \in \mathbb{N}\}$. Suppose $\forall\ n \in \mathbb{N}$, $nx \leq y$. Then y is an upper bound and $\alpha = \sup(A)$ exists. $\alpha - x < \alpha$ is not an upper bound of A. Then $\exists\ m \in \mathbb{N}$ s.t. $mx > \alpha - x$. Thus $\alpha < (m+1)x$, so α is not an upper bound, a contradiction. $\square$

Theorem (Density of Rationals). $(\forall\ x, y \in \mathbb{R}, x < y)$, $\exists\ p \in \mathbb{Q}$ s.t. $x < p < y$.

Proof. $y - x > 0$, so we may choose n s.t. $n(y - x) > 1$. Then $ny > 1 + nx > nx$. We may choose an integer m s.t. $nx < m < ny$. Then $x < \frac{m}{n} < y$. $\square$

Theorem. $(\forall\ x \in \mathbb{R}_{>0}, n \in \mathbb{N})$, $\exists\ !\ y$ s.t. $y^n = x$.

4 Sequences and Series

Definition. A **sequence** is a map $f : \mathbb{N} \to S$, $n \mapsto f(n) = p_n$. We note it by $\{p_n\}_{n=m}^{\infty}$. We can study sequences in any metric space, but we must define addition in order to discuss series.

4.1 Convergent Sequences

Definition. A sequence $\{p_n\}$ in a metric space (X, d) **converges** if there is a $p \in X$ s.t. $(\forall\ \epsilon > 0)$ $(\exists\ N \in \mathbb{N})$ s.t. $(\forall\ n > N)$ $d(p_n, p) < \epsilon$. This is noted $p_n \to p$.

- If the sequence does not converge then it **diverges**.
- A convergent sequence in a discrete metric space is one that is eventually constant.
- A sequence is **bounded** if its range is bounded in (X, d).

Proposition. We will mainly consider $\mathbb{R}$, but the following results hold true in any metric space.

- $\{p_n\}$ converges to $p \iff$ every $B(p, r)$ contains all but finitely many elements of $\{p_n\}$.
- If $\{p_n\}$ converges to p and p', then $p' = p$.
- If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.
- If $E \subseteq X$ we say p is a **limit point** of E if $\exists\ \{p_n\} \subseteq E \setminus \{p\}$ s.t. $p_n \to p$.

4.2 Limits on $\mathbb{R}, \mathbb{C}, \mathbb{R}^k$

Theorem. Suppose $\{s_n\} \to s$ and $\{t_n\} \to t$.

- $\{s_n + t_n\} \to s + t$
- $\{cs_n\} \to cs$ for $c \in \mathbb{C}$
- $\{s_n t_n\} \to st$
- $\{s_n/t_n\} \to s/t$ if $s_n, t \neq 0$
- A sequence $\{p_n\}$ in $\mathbb{R}^k$, where $p_n = (x_{1,n}, ..., x_{k,n})$, converges to $p = (x_1, ..., x_k)$ iff for each component i, $(x_{i,n}) \to x_i$.

Proposition (Order Limits). If $\{a_n\}, \{b_n\} \subseteq \mathbb{R}$ converge to a and b, then

- $a_n \geq 0$ $(\forall\ n \in \mathbb{N}) \implies a \geq 0$
- $a_n \leq b_n$ $(\forall\ n \in \mathbb{N}) \implies a \leq b$
- $c \leq b_n$ $(\forall\ n \in \mathbb{N}) \implies c \leq b$

Theorem (Squeeze Theorem). Suppose $\{x_n\}, \{y_n\}, \{z_n\} \subseteq \mathbb{R}$ s.t. $x_n \leq y_n \leq z_n$ ($\forall\ n \in \mathbb{N}$). Then, if $x_n \to L$ and $z_n \to L$, then $y_n \to L$.

Theorem (Monotone Convergence Theorem). Every bounded, monotonic sequence in $\mathbb{R}$ converges.

Proof. Let $\{a_n\}$ be a non-decreasing sequence that is bounded above. Let $S = \{a_n : n \in \mathbb{N}\}$ be the set of values in the sequence. Since S is non-empty and bounded above, the completeness of $\mathbb{R}$ guarantees that it has a least upper bound, let's call it $L = \sup(S)$. We claim that $a_n \to L$. □

Theorem (Bolzano-Weierstrass). Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.

Proof. Let $\{x_n\}$ be a bounded sequence and let $S = \{x_n : n \in \mathbb{N}\}$ be its range. If S is finite, then one of the values appears infinitely often, so take the constant subsequence. Alternatively, keep splitting the interval containing the lower and upper bound of S in half, and wlog one half must contain infinitely many points of the sequence. Construct the subsequence by picking points during each cut. □

Proposition.

- $p > 0 \implies \lim_{n\to\infty} \frac{1}{n^p} = 0$.
- $p > 0 \implies \lim_{n\to\infty} p^{1/n} = 1$.
- $\lim_{n\to\infty} n^{1/n} = 1$.
- $p > 0 \implies (\forall\ a \in \mathbb{R}), \lim_{n\to\infty} \frac{n^a}{(1+p)^n} = 0$.
- $|x| < 1 \implies \lim_{n\to\infty} x^n = 0$.

4.3 Means and Inequalities on $\mathbb{R}$

Definition. Let $a_1, ..., a_n \in \mathbb{R}_{>0}$. We have the following means:

- **Quadratic:** $QM_n := \left(\frac{a_1^2+...+a_n^2}{n}\right)^{1/2}$
- **Arithmetic:** $AM_n := \frac{a_1+...+a_n}{n}$
- **Geometric:** $GM_n := (a_1 \cdot ... \cdot a_n)^{1/n}$
- **Harmonic:** $HM_n := \left(\frac{1}{n}\left(\frac{1}{a_1} + ... + \frac{1}{a_n}\right)\right)^{-1}$

We have that $QM \geq AM \geq GM \geq HM$, and equality only happens when the a_i's are all the same.

Theorem (Bernoulli's Inequality). $x > -1 \implies (1+x)^n \geq 1 + nx$.

Proof. Let $a_1 = ... = a_{n-1} = 1, a_n = 1 + nx$. Then $AM = \frac{(n-1)+(1+nx)}{n} = 1 + x$. $GM = (1 \cdot ... \cdot 1 \cdot (1+nx))^{1/n} = (1+nx)^{1/n}$. By AM-GM, $1+x \geq (1+nx)^{1/n} \implies (1+x)^n \geq 1+nx$. $\square$

Theorem (Cauchy-Schwartz Inequality). $(\forall\, a_i, b_i \in \mathbb{R})$, $\left(\sum_{j=1}^n a_j b_j\right)^2 \leq \left(\sum_{j=1}^n a_j^2\right)\left(\sum_{j=1}^n b_j^2\right)$.

Proof. The quadratic $\sum(a_i x + b_i)^2 = (\sum a_i^2)x^2 + 2(\sum a_i b_i)x + \sum b_i^2 \geq 0$ since it is a sum of squares. Thus its discriminant must be ≤ 0. $\square$

Theorem (Minkowski's Inequality). $\left(\sum_{j=1}^n |a_j + b_j|^2\right)^{1/2} \leq \left(\sum_{j=1}^n |a_j|^2\right)^{1/2} + \left(\sum_{j=1}^n |b_j|^2\right)^{1/2}$.

Proof. Let $S_n^2 := \sum_{j=1}^n |a_j + b_j|^2 \leq \sum |a_j + b_j||a_j| + \sum |a_j + b_j||b_j|$. By Cauchy-Schwartz, this is $\leq (\sum |a_j + b_j|^2)^{1/2}((\sum |a_j|^2)^{1/2} + (\sum |b_j|^2)^{1/2})$. So $S_n^2 \leq S_n((\sum |a_j|^2)^{1/2} + (\sum |b_j|^2)^{1/2})$. $\square$

Theorem (Holder's Inequality). $(\forall\ x_i, y_i \in \mathbb{R})$, if p, q s.t. $\frac{1}{p} + \frac{1}{q} = 1$, then $\sum |x_i y_i| \leq (\sum |x_i|^p)^{1/p}(\sum |y_i|^q)^{1/q}$.

Theorem (Stolz-Cesàro). Let $(x_n), (y_n)$ be sequences in $\mathbb{R}$ s.t. y_n strictly increases to ∞. If $\lim_{n\to\infty} \frac{x_n - x_{n-1}}{y_n - y_{n-1}} = a$, then $\lim_{n\to\infty} \frac{x_n}{y_n} = a$.

Proof. wlog, assume $y_n > 0$ for all n. Since $\lim \frac{x_n - x_{n-1}}{y_n - y_{n-1}} = a$, for any $\epsilon > 0$, $\exists\ M$ such that for all $n > M$, we have $a - \frac{\epsilon}{2} < \frac{x_n - x_{n-1}}{y_n - y_{n-1}} < a + \frac{\epsilon}{2}$. This implies $(a - \frac{\epsilon}{2})(y_n - y_{n-1}) < x_n - x_{n-1} < (a + \frac{\epsilon}{2})(y_n - y_{n-1})$. Summing this inequality from $k = M$ to n gives $(a - \frac{\epsilon}{2})(y_n - y_{M-1}) < x_n - x_{M-1} < (a + \frac{\epsilon}{2})(y_n - y_{M-1})$. Then we can show $|\frac{x_n}{y_n} - a|$ becomes arbitrarily small. By the triangle inequality, $|\frac{x_n}{y_n} - a| \leq |\frac{x_{M-1}}{y_n}| + (1 - \frac{y_{M-1}}{y_n})|\frac{x_n - x_{M-1}}{y_n - y_{M-1}} - a| + |a|\frac{y_{M-1}}{y_n}$. As $y_n \to \infty$, we can make the right side less than ϵ. $\square$

4.4 Euler's Number

Let $a_n = (1 + \frac{1}{n})^n$ and $b_n = (1 + \frac{1}{n})^{n+1}$.

Proposition. (a_n) is strictly increasing, (b_n) is strictly decreasing, and $a_n < b_n$ for all n.

Proof.

1. By AM-GM, with $x_1 = 1$ and n terms of $1 + \frac{1}{n}$, we get $GM = ((1 + \frac{1}{n})^n)^{1/(n+1)} < AM = \frac{1 + n(1+1/n)}{n+1} = 1 + \frac{1}{n+1}$. Raising both sides to the power of $n+1$ gives $(1 + \frac{1}{n})^n < (1 + \frac{1}{n+1})^{n+1}$, so $a_n < a_{n+1}$.

2. By HM-GM, with $x_1 = 1$ and n terms of $\frac{n}{n-1} = 1 + \frac{1}{n-1}$, we get $HM = \frac{n+1}{1 + n(\frac{n-1}{n})} = 1 + \frac{1}{n}$. $GM = ((1 + \frac{1}{n-1})^n)^{1/(n+1)}$. Since $HM < GM$, $1 + \frac{1}{n} < (1 + \frac{1}{n-1})^{n/(n+1)}$, which implies $(1 + \frac{1}{n})^{n+1} < (1 + \frac{1}{n-1})^n$, so $b_n < b_{n-1}$.

$\square$

Definition (e). Since $a_1 < a_n < b_n < b_1$, by the Monotone Convergence Theorem, $\lim a_n$ and $\lim b_n$ exist and are equal, since $\lim b_n = \lim (1 + \frac{1}{n}) a_n = \lim a_n$. We call this limit e.

Proposition. If $\lim_{n\to\infty} a_n = \pm\infty$, then $\lim_{n\to\infty}(1+\frac{1}{a_n})^{a_n} = e$.

Proof. Assume $a_n \to \infty$. Let $b_n = \lfloor a_n \rfloor$. Then $b_n \leq a_n < b_n + 1$. So $(1+\frac{1}{b_n+1})^{b_n} < (1+\frac{1}{a_n})^{a_n} < (1+\frac{1}{b_n})^{b_n+1}$. The left and right terms both go to e, so by the Squeeze Theorem, the result holds. $\square$

4.5 Cauchy and Subsequences

Definition. A **subsequence** is formed from a sequence $\{p_n\}$ by taking indices $\{n_k\}$ s.t. $n_{k+1} > n_k$.

- If p_n converges to p, then every subsequence does as well.
- If p is a limit point of the range of the sequence, there is a subsequence that converges to p.
- If X is compact and $\{p_n\} \subseteq X$, then some subsequence of $\{p_n\}$ converges.

Definition. A sequence $\{p_n\}$ is **Cauchy** if $(\forall\, \epsilon > 0)\ \exists\, N \in \mathbb{N}$ s.t. $(\forall\, m, n > N)\ d(p_m, p_n) < \epsilon$.

- Every convergent sequence is Cauchy.
- A **complete metric space** is a space where every Cauchy sequence converges.

Example. Discrete spaces, $\mathbb{R}^k$, and compact metric spaces are complete.

Proof. (for Compact spaces): Let $\{p_n\}$ be a Cauchy sequence in a compact space X. Since X is compact, the infinite set of points $\{p_n\}$ must have a limit point $p \in X$. This means there exists a subsequence $\{p_{n_k}\}$ that converges to p. Since $\{p_n\}$ is Cauchy, for any $\epsilon > 0$, $\exists\, N$ s.t. if $m, n > N$, $d(p_m, p_n) < \epsilon/2$. Since $p_{n_k} \to p$, we can find a K s.t. for $k > K$, $d(p_{n_k}, p) < \epsilon/2$. If we choose k large enough such that $n_k > N$, then for any $n > N$, $d(p_n, p) \leq d(p_n, p_{n_k}) + d(p_{n_k}, p) < \epsilon/2 + \epsilon/2 = \epsilon$. Thus, $\{p_n\}$ converges to p. $\square$

4.6 Upper/Lower Limits of a Sequence

Definition. If (s_n) is a sequence in $\mathbb{R}$, then

- **Upper limit:** $\limsup_{n\to\infty}(s_n) = \inf_{k\geq 1} \sup_{n\geq k}(s_n)$
- **Lower limit:** $\liminf_{n\to\infty}(s_n) = \sup_{k\geq 1} \inf_{n\geq k}(s_n)$

Proposition. If $(s_n) \subseteq \mathbb{R}$, then the upper and lower limits always exist. Further we have that $\liminf_{n\to\infty} S_n = \lim_{n\to\infty} b_k \leq \lim_{n\to\infty} a_k = \limsup_{n\to\infty} S_n$.

Proof. Let $a_k := \sup_{n\geq k} s_n$. Then the sequence $\{a_k\}$ is non-increasing $(a_{k+1} \leq a_k)$, so its limit exists by the Monotone Convergence Theorem. Similarly, if $b_k := \inf_{n\geq k} s_n$, the sequence $\{b_k\}$ is non-decreasing, so its limit also exists. $\square$

Definition. An **accumulation point** of a sequence is a point for which there is a subsequence that converges to it. If E is the set of all accumulation points of (s_n), then $\limsup s_n = \sup(E)$ and $\liminf s_n = \inf(E)$.

Proposition. A sequence $(s_n) \subseteq \mathbb{R}$ converges if and only if its upper and lower limits are equal. In that case, the sequence converges to this common value.

Proof. ($\Longrightarrow$): If $s_n \to L$, then every subsequence converges to L. The set of accumulation points is just $E = \{L\}$, so $\sup(E) = \inf(E) = L$.

($\Longleftarrow$): Assume $\liminf s_n = \limsup s_n = L$. By definition, we know that $\inf_{n \geq k} s_n \leq s_k \leq \sup_{n \geq k} s_n$. The term on the left converges to $\liminf s_n = L$, and the term on the right converges to $\limsup s_n = L$. By the Squeeze Theorem, the sequence s_k must also converge to L. □

4.7 Series

Definition. A **series** is the sum of elements in a sequence. The series converges if the sequence of partial sums converges.

Theorem (Cauchy Condensation Test). Suppose $(b_n) \subseteq \mathbb{R}$ is decreasing and non-negative. Then $\sum b_n$ converges $\iff$ $\sum 2^k b_{2^k}$ converges.

Proof. Let $s_n = \sum_{i=1}^{n} b_i$ and $t_k = \sum_{i=1}^{k} 2^i b_{2^i}$. We can show that for appropriate n and k, $s_n \leq t_k$ and $s_n \geq \frac{1}{2} t_k$, so both sequences of partial sums are either bounded or unbounded together. □

Example. The p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges iff $p > 1$.

Proof. Consider $\sum_k 2^k \frac{1}{(2^k)^p} = \sum_k (2^{1-p})^k$, which is a geometric series that converges when $2^{1-p} < 1$, i.e., $p > 1$. □

Theorem. $\sum_{n=0}^{\infty} \frac{1}{n!} = e$.

Proof. The sequence of partial sums s_n is increasing and bounded, since $2 < s_n < 3$, so it converges. Let $t_n = (1 + \frac{1}{n})^n = \sum_{k=0}^{n} \frac{1}{k!}(1 - \frac{1}{n}) \dots (1 - \frac{k-1}{n})$. Thus $t_n \leq s_n$. Taking limits, we get $e \leq \sum \frac{1}{n!}$. For a fixed m, if $n \geq m$, $t_n \geq \sum_{k=0}^{m} \frac{1}{k!}(1 - \frac{1}{n}) \dots (1 - \frac{k-1}{n})$. Letting $n \to \infty$, we get $e \geq s_m$. Thus $e = \sum \frac{1}{n!}$.

- The error estimate is $0 < e - s_n < \frac{1}{n \cdot n!}$.
- It follows that **e is irrational**.

  Proof. If $e = p/q$, then $0 < q!(e - s_q) < \frac{1}{q}$. But $q!(e - s_q)$ must be an integer, a contradiction. □

□

Theorem (Algebraic Limits). If $\sum a_k = A$ and $\sum b_k = B$, then $\sum(\alpha a_k + \beta b_k) = \alpha A + \beta B$.

Theorem (Cauchy Criterion for Series). $\sum a_k$ converges $\iff$ $(\forall\, \epsilon > 0)(\exists\, N \in \mathbb{N})$ s.t. $n > m \geq N \implies |\sum_{k=m+1}^{n} a_k| < \epsilon$. As a result, if $\sum a_n$ converges, then $\lim a_n = 0$.

Theorem (Convergence Tests).

- **Comparison Test:** If $0 \leq a_k \leq b_k$ for all k, then if $\sum b_k$ converges, $\sum a_k$ converges. If $\sum a_k$ diverges, $\sum b_k$ diverges.
- **Root Test:** Define $\alpha := \limsup |a_n|^{1/n}$. The series $\sum a_n$ converges if $\alpha < 1$ and diverges if $\alpha > 1$, and is inconclusive if $\alpha = 1$.
- **Ratio Test:** Define $\alpha := \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$. The series $\sum a_n$ converges if $\alpha < 1$, diverges if $\alpha > 1$, and is inconclusive if $\alpha = 1$.

Definition.

- A series $\sum a_n$ **converges absolutely** if $\sum |a_n|$ converges. This implies convergence.
- A series $\sum a_n$ **converges conditionally** if it converges but does not converge absolutely.

Theorem (Abel Summation Formula). Given sequences $(a_n), (b_n) \subseteq \mathbb{R}$ let $A_n = \sum_{k=0}^{n} a_k$ with $A_{-1} = 0$. For $0 \leq p \leq q$, $\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$.

Proof.

$$\begin{aligned}
\sum_{n=p}^{q} a_n b_n &= \sum_{n=p}^{q} (A_n - A_{n-1}) b_n \\
&= \sum_{n=p}^{q} A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1} \\
&= \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p
\end{aligned}$$

□

Theorem (Dirichlet's Test). If the partial sums of (a_n) are bounded, and (b_n) is a decreasing sequence converging to 0, then $\sum a_n b_n$ converges.

Proof. Choose $M \geq 0$ as an upper bound for the partial sums $|A_n|$. Since $b_n \to 0$, for any $\epsilon > 0$, $\exists\, N$ such that for $n \geq N$, $b_n < \frac{\epsilon}{2M}$. By Abel's formula, for $q > p \geq N$, we have $|\sum_{n=p}^{q} a_n b_n| = |\sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p| \leq M \sum_{n=p}^{q-1} |b_n - b_{n+1}| + M|b_q| + M|b_{p-1}|$. Since b_n is non-increasing and non-negative, this simplifies to $M(b_p - b_q) + M b_q + M b_{p-1} \approx 2M b_p < 2M \frac{\epsilon}{2M} = \epsilon$. Thus the sequence of partial sums is Cauchy and converges. □

Theorem (Alternating Series Test). If (b_n) is a decreasing sequence converging to 0, then $\sum(-1)^n b_n$ converges.

Definition. Given $\sum a_n$ and $\sum b_n$, define their series product product as $\sum c_n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.

Theorem. If $\sum a_n$ converges absolutely to A and $\sum b_n$ converges to B, then $\sum c_n$ converges to AB.

Proof. Define the partial sums as A_n, B_n, C_n and let $\beta_n := B_n - B$. Then we can express the partial sum C_n as

$$\begin{aligned} C_n = \sum_{k=0}^{n} c_k &= \sum_{k=0}^{n} a_k B_{n-k} \\ &= \sum_{k=0}^{n} a_k (B + \beta_{n-k}) \\ &= \left(\sum_{k=0}^{n} a_k \right) B + \sum_{k=0}^{n} a_k \beta_{n-k} \\ &= A_n B + \sum_{k=0}^{n} a_k \beta_{n-k} \end{aligned}$$

We know that $A_n B \to AB$. Since $\sum a_n$ is absolutely convergent and $\beta_n \to 0$, it can be shown that the second term, $\sum_{k=0}^{n} a_k \beta_{n-k}$, converges to 0. Thus, $\lim_{n\to\infty} C_n = AB$. $\square$

Definition. Suppose $f : \mathbb{N} \to \mathbb{N}$ is a bijection s.t. $n \mapsto n'$. Then the series $\sum_{n'=0}^{\infty} a_{n'}$ is a rearrangement of the series $\sum_{n=0}^{\infty} a_n$. If S_n are the partial sums of the original series and $S_{n'}$ are the partial sums of the rearranged series, it is not necessarily true that the sequences of partial sums $\{S_n\}$ and $\{S_{n'}\}$, converge to the same limit.

Theorem (Rearrangement Theorems).

- **Riemann Rearrangement Theorem:** If $\sum a_n$ converges conditionally, then for any $-\infty \le \alpha \le \beta \le \infty$, there exists a rearrangement $\sum a_{n'}$ such that the partial sums $s_{n'}$ satisfy $\liminf s_{n'} = \alpha$ and $\limsup s_{n'} = \beta$. (Namely, we can play tricks with the rearrangement causing them to converge wherever we want to.)
- **Absolute Convergence Theorem:** If $\sum a_n$ converges absolutely, then every rearrangement converges to the same sum.

5 Basic Topology

5.1 Metric Spaces

Definition. A **metric** on a set X is a function $\rho : X \times X \to [0, \infty)$ such that for all $x, y, z \in X$:

1. $\rho(x, y) = 0$ iff $x = y$
2. $\rho(x, y) = \rho(y, x)$ (symmetry)
3. $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ (triangle inequality)

The function $\rho(x, y)$ represents the **distance** from x to y. A set X equipped with a metric ρ is called a **metric space**, denoted (X, ρ).

Example.

- The set $\mathbb{R}$ with $\rho(x, y) = |x - y|$ is a metric space.
- For any set X, the **discrete metric** is $\rho(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$.
- For $x, y \in \mathbb{R}^d$, we can define several metrics:
  - the "taxicab"/L_2 metric: $\rho_1(x, y) = \sum_{j=1}^{d} |x_j - y_j|$
  - the Euclidean/L_2 metric: $\rho_2(x, y) = \left(\sum_{j=1}^{d} |x_j - y_j|^2\right)^{1/2}$
  - the "max"/L_∞ metric: $\rho_\infty(x, y) = \max_{1 \leq j \leq d} |x_j - y_j|$
- For $x, y \in \mathbb{R}^{\mathbb{N}}$ (infinite sequences), $\rho(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min(1, |x_n - y_n|)$ is a metric.
- If ρ is a metric on X and $A \subseteq X$, then the restriction $\rho|_{A \times A}$ is a metric on A.
- For a product of metric spaces $X = X_1 \times ... \times X_d$, we can define in the same sense as we did for $\mathbb{R}^d$
  - $d_1(x, y) = \sum_{j=1}^{d} \rho_j(x_j, y_j)$
  - $d_2(x, y) = \left(\sum_{j=1}^{d} \rho_j(x_j, y_j)^2\right)^{1/2}$
  - $d_\infty(x, y) = \max_{1 \leq j \leq d} \rho_j(x_j, y_j)$

Definition. An **open ball** in a metric space (X, ρ) with center $x \in X$ and radius $r > 0$ is the set: $B(x, r) = \{y \in X : \rho(x, y) < r\}$.

5.2 Open and Closed Sets

Definition. If X is a metric space, we say a set $E \subseteq X$ is **open** if for every $x \in E$, there exists some $r > 0$ such that $B(x, r) \subseteq E$. We say a set $E \subseteq X$ is **closed** if its complement $X \setminus E$ is open. It follows that the entire space X and the empty set $\emptyset$ are both open and closed.

Proposition. Every open ball $B(x, r)$ is an open set.

Proof. If $y \in B(x, r)$, let $s = \rho(x, y) < r$. Then $B(y, r - s) \subseteq B(x, r)$. If $z \in B(y, r - s)$, then $\rho(x, z) \leq \rho(x, y) + \rho(y, z) < s + (r - s) = r$, so $z \in B(x, r)$. $\square$

Proposition. The union of any family (finite or infinite) of open sets is open. The intersection of a finite family of open sets is open.

Proof. The union case is trivial since each element lands in at least one of the sets in the family, and that set is open. For the intersection case, if $x \in \bigcap_{j=1}^n U_j$, then for each j, there is an $r_j > 0$ with $B(x, r_j) \subseteq U_j$. Let $r = \min\{r_1, ..., r_n\}$. Then $r > 0$ and $B(x, r) \subseteq \bigcap_{j=1}^n U_j$. Note that if we ask for a countably infinite intersection of open sets and we try to extend this notion by taking the inf of the set r_i this could end up being 0. $\square$

Corollary. The intersection of any family (finite or infinite) of closed sets is closed. The union of a finite family of closed sets is closed.

Proof. Intersections: Let $\{F_\alpha\}_{\alpha \in I}$ be a family of closed subsets of X. For each α, the complement $U_\alpha = X \setminus F_\alpha$ is open. By DeMorgan's laws, $X \setminus \bigcap_{\alpha \in I} F_\alpha = \bigcup_{\alpha \in I}(X \setminus F_\alpha) = \bigcup_{\alpha \in I} U_\alpha$, which is a union of open sets, hence open. Therefore, $\bigcap_{\alpha \in I} F_\alpha$ is closed.

Unions: Let $F_1, \ldots, F_n$ be closed. For each j, the complement $U_j = X \setminus F_j$ is open. Again by DeMorgan's laws, $X \setminus \bigcup_{j=1}^n F_j = \bigcap_{j=1}^n (X \setminus F_j) = \bigcap_{j=1}^n U_j$, and the finite intersection of open sets is open. Hence, $\bigcup_{j=1}^n F_j$ is closed. $\square$

Remark. We may instead define openness topologically. Given X, we take $\tau \subseteq \mathcal{P}(X)$ which is the collection of open sets in X, where τ contains the whole space and the emptyset as elements, and is closed under unions and finite intersections.

5.3 Interiors and Closures

Definition. Let $E \subseteq X$.

- The **interior** of E, denoted $\text{int}(E)$, is the union of all open sets contained in E. It is the largest open set contained in E.
- The **closure** of E, denoted $\text{cl}(E)$ or $\bar{E}$, is the intersection of all closed sets containing E. It is the smallest closed set containing E.
- The **boundary** of E, denoted ∂E, is the set difference $\text{cl}(E) \setminus \text{int}(E)$.

Proposition. In particular, we note that $(\text{int}(A))^c = \text{cl}(A^c)$ and $(\text{cl}(A))^c = \text{int}(A^c)$.

Proof. note that $\mathrm{int}(A) \subseteq A \subseteq \mathrm{cl}(A)$, so $A^c \subseteq (\mathrm{int}(A))^c$. $(\mathrm{int}(A))$ is open, so the complement is closed, but we know $\mathrm{cl}(A^c)$ is the smallest closed set containing A^c. Thus $\mathrm{cl}(A^c) \subseteq (\mathrm{int}(A))^c$. Also, $A^c \subseteq \mathrm{cl}(A^c)$ meaning $A \supseteq (\mathrm{cl}(A^c))^c$. But the right hand side is open and is contained in A meaning $(\mathrm{cl}(A^c))^c \subseteq (\mathrm{int}(A))$. Taking the complement we also get that $\mathrm{cl}(A^c) \supseteq (\mathrm{int}(A))^c$. Thus $(\mathrm{int}(A))^c = \mathrm{cl}(A^c)$. To prove the other direction, we simply consider the complement. □

Example.

- (a, b) is open. $[a, b]$ is closed, since its complement $(-\infty, a) \cup (b, +\infty)$ is open. $\mathbb{Z}$ is closed, since its complement $\mathbb{Z}^c = \bigcup_{n\in\mathbb{Z}}(n, n+1)$ is a union of open sets, and therefore open.
- $X = \mathbb{R}$, $E = (a, b)$. $\mathrm{cl}(E) = [a, b]$, $\mathrm{int}(E) = (a, b)$. $\partial E = \{a, b\}$.
- $X = \mathbb{R}$, $E = \mathbb{Z}$. $\mathrm{cl}(\mathbb{Z}) = \mathbb{Z}$, $\mathrm{int}(\mathbb{Z}) = \emptyset$. $\partial\mathbb{Z} = \mathbb{Z}$.
- $X = \mathbb{R}$, $E = \mathbb{Q}$. $\mathrm{cl}(\mathbb{Q}) = \mathbb{R}$, $\mathrm{int}(\mathbb{Q}) = \emptyset$. $\partial\mathbb{Q} = \mathbb{R}$.

Proposition. Every open set in $\mathbb{R}$ is a countable disjoint union of open intervals.

Proof. Suppose $U \subseteq \mathbb{R}$ is open. For $x \in U$ (open), let J_x be the union of all open intervals I such that $x \in I \subseteq U$. A union of two intervals mutually containing a point is still an interval, so J_x is the largest open interval containing x in U. For any $x, y \in U$, either $J_x = J_y$ or $J_x \cap J_y = \emptyset$, because otherwise they share a point meaning taking the union would result in a larger interval, which is a contradiction to saying J_x is the largest. Then the set $\mathcal{F} = \{J_x : x \in U\}$ is a collection of disjoint open intervals whose union is U. We may map each $J \in \mathcal{F}$ to a unique rational number $f(J) \in J$. This map $f : \mathcal{F} \to \mathbb{Q}$ is injective, since the intervals are disjoint. Since $\mathbb{Q}$ is countable, $\mathcal{F}$ must also be countable. □

5.4 Dense and Separable Sets

Definition. A set $E \subseteq X$ is **dense** in X if it intersects every non-empty open set U. Equivalently, E is dense if $E \cap B(x, r) \neq \emptyset$ for every $x \in X$ and $r > 0$. Equivalently, E is dense if $\mathrm{cl}(E) = X$.

Example.

- $\mathbb{Q}$ is dense in $\mathbb{R}$
- $\mathbb{Q}^d$ is dense in $\mathbb{R}^d$
- if $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, the set of fractional parts then $\{n\alpha - \lfloor n\alpha \rfloor : n \in \mathbb{Z}\}$ is dense in $[0, 1]$
- $\left\{\frac{k}{2^n} : n \in \mathbb{N},\ 0 < k < 2^n\right\}$ is dense in $(0, 1)$

Definition. A set $E \subseteq X$ is **nowhere dense** if the interior of its closure is empty ($\mathrm{int}(\mathrm{cl}\ E) = \emptyset$).

Example.

- In $\mathbb{R}$, any single point set $\{x\}$ is nowhere dense
- In $\mathbb{R}$, $\mathbb{Z}$ is nowhere dense but $\mathbb{Q}$ is not nowhere dense

Definition. A metric space (X, ρ) is **separable** if it has a countable dense subset.

Example.

- $\mathbb{R}$ is separable because $\mathbb{Q}$ is countable and dense
- $\mathbb{R}^d$ is separable because $\mathbb{Q}^d$ is countable and dense

5.5 Convergence

Definition. A sequence $(x_n)_{n\in\mathbb{N}}$ in (X, ρ) **converges** to $x \in X$ if for every $\epsilon > 0$, there is an integer $N_\epsilon \in \mathbb{N}$ such that $n \geq N_\epsilon$ implies $\rho(x_n, x) < \epsilon$. We write this as $\lim_{n\to\infty} x_n = x$ or $x_n \to x$. If a sequence does not converge, it **diverges**.

Definition. The **diameter** of a set $E \subseteq X$ is $\operatorname{diam}(E) = \sup\{\rho(x, y) : x, y \in E\}$. We a set E is **bounded** if $\operatorname{diam}(E) < \infty$.

Example.

- A single-point set $\{x\}$ is bounded, $\operatorname{diam}(\{x\}) = 0$.
- An open ball $B(x, r)$ is bounded. If $x_1, x_2 \in B(x, r)$, then $\rho(x_1, x_2) \leq \rho(x_1, x) + \rho(x, x_2) < 2r$. Thus, $\operatorname{diam}(B(x, r)) \leq 2r$.
- In $\mathbb{R}$, for $-\infty < a < b < \infty$, $\operatorname{diam}((a, b)) = b - a$.

Theorem. Let (x_n) be a sequence in a metric space (X, ρ).

1. $x_n \to x$ if and only if every open set V containing x contains x_n for all but finitely many n.
2. If $x_n \to x$ and $x_n \to x'$, then $x = x'$ (limits are unique).
3. If (x_n) converges, then (x_n) is bounded.

Proposition. For $E \subseteq X$ and $x \in X$, the following are equivalent:

1. (a) $x \in \operatorname{cl}(E)$
2. (b) $B(x, r) \cap E \neq \emptyset$ for all $r > 0$
3. (c) There is a sequence (x_n) in E that converges to x

Proof. (a) $\Rightarrow$ (b): Assume $x \in \operatorname{cl}(E)$. If $B(x, r) \cap E = \emptyset$ for some $r > 0$, then $B(x, r)^c$ is a closed set containing E, so it must contain its closure as well. But $x \notin B(x, r)^c$, which is a contradiction.

(b) $\Rightarrow$ (c): For each $n \in \mathbb{N}$, (b) implies we can find $x_n \in B(x, 1/n) \cap E$. Then $\rho(x_n, x) < 1/n$, so $\lim_{n\to\infty} \rho(x_n, x) = 0$, meaning $x_n \to x$.

(c) $\Rightarrow$ (b): If a sequence (x_n) in E converges to x, then for any $r > 0$, there exists an N such that $n \geq N \implies x_n \in B(x, r)$. This means $B(x, r) \cap E \neq \emptyset$.

(b) $\Rightarrow$ (a): Assume $B(x, r) \cap E \neq \emptyset$ for all $r > 0$. If $x \notin \operatorname{cl}(E)$, then x is in the open set $(\operatorname{cl}(E))^c$. This means there is some $r > 0$ such that $B(x, r) \subseteq (\operatorname{cl}(E))^c \subseteq E^c$, since $(\operatorname{cl}(E))^c = \operatorname{int}(E^c) \subseteq E^c$, which implies $B(x, r) \cap E = \emptyset$, a contradiction. $\square$

5.6 Accumulation and Isolated Points

Definition. An **accumulation point** (equivalently, limit point or cluster point) of E is a point $x \in X$ such that for every open set $U \ni x$, there is a point other than x in U. Namely, $E \setminus \{x\} \cap U \neq \emptyset$. Equivalently, x has points arbitrarily close to it. A point $x \in E$ is an **isolated point** of E if it is not an accumulation point of E. (Dense sets may indeed contain isolated points.)

Proposition. The following are equivalent:

1. x is an accumulation point of E
2. There exists a sequence $(x_n) \subseteq E \setminus \{x\}$ such that $x_n \to x$.

Proof. If x is an accumulation point, for each $n \in \mathbb{N}$ we can choose $x_n \in E \cap B(x, 1/n)$, giving a sequence (x_n) converging to x. Conversely, if such a sequence exists, then every neighborhood of x contains some $x_n \neq x$, so x is an accumulation point. □

Proposition. $\mathrm{cl}(E) = E \cup \mathrm{acc}(E)$. Thus, E is closed if and only if $\mathrm{acc}(E) \subseteq E$. Thus, E is dense if every point of X is either in E itself or is an accumulation point of E.

Proof. ($\supseteq$) We know $E \subseteq \mathrm{cl}(E)$. If $x \in \mathrm{acc}(E)$, then every $B(x, r)$ contains a point of E, so $B(x, r) \cap E \neq \emptyset$. By the previous proposition, $x \in \mathrm{cl}(E)$. Thus $E \cup \mathrm{acc}(E) \subseteq \mathrm{cl}(E)$.

($\subseteq$) If $x \in \mathrm{cl}(E)$ and $x \notin E$, we must show $x \in \mathrm{acc}(E)$. Since $x \in \mathrm{cl}(E)$, $B(x, r) \cap E \neq \emptyset$ for all $r > 0$. Since $x \notin E$, this intersection must be with $E \setminus \{x\}$. Thus $(E \setminus \{x\}) \cap B(x, r) \neq \emptyset$ for all $r > 0$, which means $x \in \mathrm{acc}(E)$. So $\mathrm{cl}(E) \subseteq E \cup \mathrm{acc}(E)$. □

Corollary. A set $E \subseteq X$ is closed if and only if for every sequence $(x_n)_{n \in \mathbb{N}} \subseteq E$ that converges to a point $x \in X$, the limit x is also in E. Equivalently, E is closed if it contains all of its accumulation points.

Proof. If E is closed, then it contains its accumulation points (precisely the points to which a sequence in E) converges. Alternatively, if E contains all of its accumulation points then $\mathrm{cl}(E) = E \cup \mathrm{acc}(E) = E$. □

Definition. A set $E \subseteq X$ is **perfect** if $E = \mathrm{acc}(E)$, where $\mathrm{acc}(E)$ is the set of all accumulation points of E. Clearly, perfect sets are closed.

Proposition. A nonempty perfect set P must be infinite.

Proof. Since P is perfect, every point $x \in P$ is a limit point. This means for any $r > 0$, $B(x, r) \cap P$ must be infinite. If $B(x_0, r_0) \cap P$ were finite, say $\{x_1, ..., x_n\}$, we could find a small $r = \min_{1 \leq i \leq n} \rho(x_0, x_i) > 0$. Then $B(x_0, r)$ would contain no points of P other than possibly x_0 itself, making x_0 an isolated point and contradicting that P is perfect. □

Proposition. The set of subsequential limits of a sequence (x_n) in a metric space (X, ρ) forms a closed subset of X.

Proof. Let E^* be the set of subsequential limits of (x_n). Let q be an accumulation point of E^*. We can find a point $x_{n_{i-1}} \in E^*$ such that $\rho(x_{n_{i-1}}, q) < \delta 2^{-i-1}$. Since $x_{n_{i-1}}$ is a subsequential limit, there is an index $n_i > n_{i-1}$ such that $\rho(x_{n_i}, x_{n_{-1}}) < \delta 2^{-i-1}$. By the triangle inequality: $\rho(q, x_{n_i}) \leq \rho(q, x) + \rho(x, x_{n_i}) < \delta 2^{-i-1} + \delta 2^{-i-1} = \delta 2^{-i}$. We have found a subsequence $(x_{n_i})_{i \in \mathbb{N}}$ that converges to q, meaning $q \in E^*$. Thus, $\text{acc}(E^*) \subseteq E^*$, meaning E^* is closed. $\square$

5.7 Cauchy Sequences and Completeness

Definition. A sequence (x_n) in (X, ρ) is a **Cauchy sequence** if for every $\epsilon > 0$, there exists an $N_\epsilon \in \mathbb{N}$ such that for all $m, n \geq N_\epsilon$, we have $\rho(x_m, x_n) < \epsilon$. It then follows that:

- Every convergent sequence is a Cauchy sequence
- Every Cauchy sequence is bounded
- If a Cauchy sequence has a subsequence (x_{n_k}) that converges to x, then the entire sequence converges to x

Definition. A subset E of a metric space (X, ρ) is **complete** if every Cauchy sequence in E converges to a limit that is also in E.

Proposition. A closed subset of a complete metric space is complete. A complete subset of an arbitrary metric space is closed.

Proof. For the first part, let X be complete and $E \subseteq X$ be closed. Let (x_n) be a Cauchy sequence in E. Since X is complete, (x_n) has a limit $x \in X$. Since E is closed, by the sequential property, x must be in E. Thus E is complete. For the second part, let $E \subseteq X$ be complete and $x \in \text{cl}(E)$. There exists a sequence $(x_n) \subseteq E$ converging to x. Every convergent sequence is Cauchy, so (x_n) is a Cauchy sequence in E. Since E is complete, the limit x must be in E. Thus $\text{cl}(E) \subseteq E$, so E is closed. $\square$

Theorem (Cantor's Intersection Theorem). A metric space (X, ρ) is complete if and only if for every decreasing sequence $F_1 \supseteq F_2 \supseteq F_3 \supseteq \ldots$ of non-empty closed sets in X with $\text{diam}(F_n) \to 0$, the intersection $\bigcap_{n=1}^{\infty} F_n$ contains exactly one point.

Proof. ($\implies$) Assume X is complete. Let (F_n) be such a sequence. Choose one $x_n \in F_n$ for each n. For any $\epsilon > 0$, $\exists$ N s.t. $n \geq N \implies \text{diam}(F_n) < \epsilon$. If $n, m \geq N$, wlog $n \geq m$. Then $F_n \subseteq F_m$. Both $x_n, x_m \in F_m$. So, $\rho(x_n, x_m) \leq \text{diam}(F_m) < \epsilon$. This shows (x_n) is a Cauchy sequence. Since X is complete, $x_n \to x_0$ for some $x_0 \in X$. Since x_0 is the limit of the sequence $(x_k)_{k \geq n}$, and F_n is closed, x_0 must be in F_n. This holds for all n, so $x_0 \in \bigcap F_n$. If $y \in \bigcap F_n$ and $y \neq x_0$, then $0 < \rho(x_0, y) \leq \text{diam}(F_n)$ for all n. This contradicts $\text{diam}(F_n) \to 0$. Thus $\bigcap F_n = \{x_0\}$.

($\impliedby$) Assume the intersection property holds. Let (x_n) be a Cauchy sequence. Let $F_n = \text{cl}(\{x_m : m \geq n\})$. F_n is closed and $F_1 \supseteq F_2 \supseteq \ldots$. Since (x_n) is Cauchy, $\text{diam}(\{x_m : m \geq n\}) \to 0$ as $n \to \infty$. Using the fact that $\text{diam}(E) = \text{diam}(\text{cl}(E))$, we have $\text{diam}(F_n) \to$

0. By assumption, $\bigcap F_n = \{x_0\}$ for some $x_0 \in X$. Since $x_0 \in F_n$ for all n, x_0 is a limit of (x_n). Thus X is complete. $\square$

5.8 Compactness

Definition. We note that openness and closedness are properties influenced by the ambient space. We now seek a property that is independent of the ambient space. Suppose (X, ρ) is a metric space. We say that for $E \subseteq X$, the family $(V_\alpha)_{\alpha \in A}$ is a **cover** of E if $E \subseteq \cup_{\alpha \in A} V_\alpha$. If the family of sets is open, then we call it an **open cover** of E. We say a subset $K \subseteq X$ is **compact** if every open cover of K contains a finite subcover.

Example. $K = \{1/n : n \in \mathbb{N}\} \cup \{0\}$ is compact in $\mathbb{R}$.

Proof. Let (V_α) be an open cover of K. One set, say V_{α_0}, must contain 0. Since V_{α_0} is open and $1/n \to 0$, V_{α_0} must contain all but finitely many of the points $1/n$. Let n_0 be such that $n \geq n_0 \implies 1/n \in V_{\alpha_0}$. For the remaining finite set $\{1/1, 1/2, ..., 1/(n_0 - 1)\}$, we pick one open set V_{α_j} to cover each $1/j$. The finite collection $\{V_{\alpha_0}, V_{\alpha_1}, ..., V_{\alpha_{n_0-1}}\}$ is a finite subcover for K. $\square$

Theorem. Compact subsets of metric spaces are closed and bounded.

Proof. Let K be compact. We show K^c is open. Let $x \in K^c$. For each $y \in K$, let $r_y = \frac{1}{2}\rho(x, y) > 0$. Consider the two open balls $W_y = B(y, r_y)$ and $V_y = B(x, r_y)$. By the triangle inequality, $V_y \cap W_y = \emptyset$. The family $\{W_y\}_{y \in K}$ is an open cover of K. Since K is compact, there is a finite subcover: $K \subseteq \bigcup_{j=1}^n W_{y_j} = W$. Let $V = \bigcap_{j=1}^n V_{y_j}$. V is open (finite intersection) and $x \in V$. V is disjoint from W, because $V \subseteq V_{y_j}$ for all j. This gives $x \in V \subseteq W^c \subseteq K^c$. So x is an interior point of K^c, and K^c is open. K is bounded since we can take the open cover to be $\{B(x, r) : x \in K, r > 0\}$ and take a finite subcover since it is compact. $\square$

Theorem. Closed subsets of compact sets are compact.

Proof. Suppose $F \subseteq K \subseteq X$, where F is closed and K is compact. Let $(V_\alpha)_{\alpha \in A}$ be an open cover of F. Then $K \subseteq (\bigcup_{\alpha \in A} V_\alpha) \cup F^c$. This is an open cover of K (since F^c is open). Since K is compact, there is a finite subcover of K, $K \subseteq V_{\alpha_1} \cup ... \cup V_{\alpha_n} \cup F^c$. But $F \subseteq K$ and $F \cap F^c = \emptyset$, meaning $F \subseteq V_{\alpha_1} \cup ... \cup V_{\alpha_n}$. $\square$

Theorem. If $(K_\alpha)_{\alpha \in A}$ is a collection of closed sets of a compact set K, such that finite intersections are nonempty, then there exists an element belonging to each K_α. Equivalently, if $A_f \subseteq A$ with $|A_f| < \infty$ and $\cap_{\alpha \in A_f} K_\alpha \neq \emptyset$ then $\cap_{\alpha \in A} K_\alpha \neq \emptyset$.

Proof. Let (K_α) be a family of closed subsets of K with the finite intersection property. Assume for contradiction that $\bigcap_\alpha K_\alpha = \emptyset$. Then (K_α^c) is an open cover of K since $\cup_{\alpha \in A} K_\alpha^c = (\cap_{\alpha \in A} K_\alpha)^c$. By compactness, finitely many suffice: $K \subseteq K_{\alpha_1}^c \cup ... \cup K_{\alpha_n}^c$. Taking complements: $K_{\alpha_1} \cap ... \cap K_{\alpha_n} = \emptyset$, contradicting the finite intersection property. Hence, $\bigcap_\alpha K_\alpha \neq \emptyset$. $\square$

Definition. A set $E \subseteq X$ is **totally bounded** if for every for every $\epsilon > 0$, there exist finitely many points $x_1, ..., x_n \in X$ such that $E \subseteq \bigcup_{i=1}^n B(x_i, \epsilon)$.

- If E is totally bounded, then so is $\mathrm{cl}(E)$.
- Every totally bounded set is bounded. The converse is not true.

Theorem. If E is a subset of a metric space (X, ρ), the following are equivalent:

1. (a) E is complete and totally bounded.
2. (b) E satisfies the Bolzano-Weierstrass property: every sequence in E has a subsequence that converges to a point of E.
3. (c) E satisfies the Heine-Borel property (it is compact): every open cover of E has a finite subcover.

Proof. (a) $\Rightarrow$ (b): Let (x_n) be a sequence in E. We need to find a convergent subsequence in E. Since E is totally bounded, we can cover E with finitely many balls of radius $1/2$. One of these balls, B_1, must contain x_n for infinitely many n. Let this infinite set of indices be $\mathbb{N}_1$. The set $E \cap B_1$ is also totally bounded. We can cover it with finitely many balls of radius $1/4$. One of them, B_2, must contain x_n for infinitely many $n \in \mathbb{N}_1$. Let this new set of indices be $\mathbb{N}_2$. We continue this process inductively, obtaining a sequence of balls B_j of radius 2^{-j} and a decreasing sequence of infinite index sets $\mathbb{N}_1 \supseteq \mathbb{N}_2 \supseteq \ldots \mathbb{N}_j \supseteq \ldots$ such that $x_n \in B_j$ for all $n \in \mathbb{N}_j$. We can now pick a subsequence $n_1 < n_2 < n_3 < \ldots$ where $n_j \in \mathbb{N}_j$. The resulting subsequence (x_{n_j}) is Cauchy, so by completeness, it converges to something in E.

(b) $\Rightarrow$ (a) We prove this by the contrapositive. First, assume E is not complete. Then there exists a Cauchy sequence (x_n) in E that does not converge to a limit in E. If a subsequence (x_{n_k}) converged to some $x \in E$, then the entire Cauchy sequence (x_n) would also have to converge to x, which is a contradiction. Therefore, this Cauchy sequence (x_n) has no subsequence that converges in E, so (b) fails. Second, assume E is not totally bounded. There must exist some $\epsilon > 0$ for which E cannot be covered by finitely many balls of radius ϵ. We can construct a sequence inductively: pick $x_1 \in E$ and $x_{n+1} \in E \setminus \bigcup_{j=1}^n B(x_j, \epsilon)$. By this construction, $\rho(x_n, x_m) \geq \epsilon$ for all $n \neq m$. This sequence has no convergent subsequence, so (b) fails.

(a) $\Rightarrow$ (c) Assume E is complete and totally bounded. Then E satisfies the Bolzano-Weierstrass property by above. We claim that for any open cover $(V_\alpha)_{\alpha \in A}$ of E, there exists an $\epsilon > 0$ (called a Lebesgue number) such that any ball B of radius ϵ that intersects E is fully contained in at least one V_α. In this case, since E is totally bounded, we can cover E using finitely many balls of radius ϵ, each of which is contained in a V_α. To prove the claim, assume otherwise. Then for each $n \in \mathbb{N}$, there is a ball of radius 2^{-n} such that $B_n \cap E \neq \emptyset$ but the ball is not contained in any single V_α. For each n, pick a point $x_n \in B_n \cap E$. By Bolzano-Weierstrass, this sequence (x_n) has a subsequence that converges to some $x \in E$, so wlog, take this sequence to be convergent. Since (V_α) is a cover, $x \in V_\alpha$ for some α. V_α is open, so there is some $\delta > 0$ such that $B(x, \delta) \subseteq V_\alpha$. Now, choose n large enough so that $2^{-n} < \delta/3$ and $\rho(x_n, x) < \delta/3$. Let $y \in B_n$. Then $\rho(y, x_n) < 2^{-n} < \delta/3$. By triangle inequality, $\rho(y, x) \leq \rho(y, x_n) + \rho(x_n, x) < \delta$. This shows $y \in B(x, \delta)$, so $B_n \subseteq B(x, \delta) \subseteq V_\alpha$, a contradiction.

(c) $\Rightarrow$ (b) Let E be compact and let (x_n) be a sequence in E. Assume for contradiction that (x_n) has no convergent subsequence in E. This means no point $x \in E$ is a subsequential limit. Therefore, for each $x \in E$, there must be an open ball B_x centered at x that contains x_n for only finitely many values of n. The collection of balls $(B_x)_{x \in E}$ is an open cover of E but may not have a finite subcover since it would be a finite union of finite sets, whereas (x_n) is an infinite sequence (cannot have repeating terms since we assumed no subsequence converged either). $\square$

Theorem (Heine–Borel). A set $K \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Proof. ($\Rightarrow$) Assume K is compact. Compact subsets of metric spaces are always closed. If K is not bounded, then we can pick points $x_n \in K$ with $\|x_n\| > n$. Then (x_n) has no convergent subsequence, contradicting sequential compactness.

($\Leftarrow$) Assume K is closed and bounded. We also know K is complete. By the Bolzano–Weierstrass theorem, every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence. Closure ensures the limit lies in K. Hence K satisfies the Bolzano-Weirestrass property, which is equivalent to compactness. Alternatively, since every closed subset of $\mathbb{R}^n$ is complete, we may show that bounded subsets are totally bounded. Every bounded set is contained in some cube $Q = [-R, R]^n$, so we must show Q is totally bounded. Given $\epsilon > 0$, pick pick an integer $k > \frac{R\sqrt{n}}{\epsilon}$ and divide $[-R, R]$ into k equal pieces. Each of these cubes has a diameter of $\frac{\sqrt{n}2R}{k} < 2\epsilon$. $\square$

Theorem. A non-empty perfect set $P \subseteq \mathbb{R}^k$ is uncountable.

Proof. recall $E \subseteq X$ is perfect if $E = \operatorname{acc}(E)$, and that nonempty perfect sets are infinite.

Now suppose for contradiction that P is countable, say $P = \{x_1, x_2, ...\}$. We will construct a nested sequence of non-empty compact sets whose intersection will end up empty, violating the finite intersection property for compact sets.

Intuitively, at each step we start with some open set V_n that still meets P, and we shrink it to a smaller open set $V_{n+1} \subset V_n$ that still touches P but avoids the point x_n. Because every point of P is a limit point, there are infinitely many other points of P inside V_n besides x_n. So we can choose another point $p \in V_n \cap P$ with $p \neq x_n$, and take a small ball around p that stays inside V_n and excludes x_n. Call that ball V_{n+1}. Then $V_{n+1} \subset V_n$, $x_n \notin \overline{V_{n+1}}$, and $V_{n+1} \cap P \neq \emptyset$.

This gives us a nested sequence of open sets $V_1 \supset V_2 \supset V_3 \supset \ldots$ each intersecting P and avoiding one more of the points $x_1, x_2, \ldots$.

Define $K_n = \overline{V_n} \cap P$. Each K_n is closed (intersection of closed sets) and bounded (since K_1 is bounded), so each K_n is compact. By construction $K_n \neq \emptyset$ and $K_{n+1} \subseteq K_n$.

But as we exclude a point with each new V_n, and since $K_n \subseteq P$, this means the intersection would have to be empty, which is a contradiction. Therefore our assumption that P is countable must be false. $\square$

5.9 Connected and Separated Sets

Definition. Two sets $A, B \subseteq X$ is **separated** if $\text{cl}(A) \cap B = A \cap \text{cl}(B) = \emptyset$. A set $E \subseteq X$ is **connected** if one of the following holds:

1. It cannot be written as a union of two non-empty, disjoint separated sets in E
2. It cannot be written as a union of two non-empty, disjoint closed sets in E
3. It cannot be written as a union of two non-empty, disjoint open sets in E

Proof. We show that the three definitions are equivalent:

(1) $\Rightarrow$ (2). Assume E is connected by (1). Suppose for contradiction that $E = F_1 \cup F_2$ with F_1 and F_2 non-empty, disjoint, and closed in E. For closed sets, $\overline{F_1} = F_1$ and $\overline{F_2} = F_2$, so we get a contradition.

(2) $\Rightarrow$ (3). Assume E is connected by (2). Suppose for contradiction that $E = U \cup V$ with U and V non-empty, disjoint, and open in E. The complements $U^c = E \setminus U$ and $V^c = E \setminus V$ are closed in E. Observe that $U = E \setminus V^c$ and $V = E \setminus U^c$, so U and V are complements of closed sets. If $E = U \cup V$, then $U^c \cap V^c = \emptyset$, and U^c, V^c are non-empty closed sets whose union is E, so we get a contradiction.

(3) $\Rightarrow$ (1) Assume E is connected by (3). Suppose for contradiction that $E = A \cup B$, where A and B are non-empty, disjoint, and separated in E. For separated sets, $A \cap \overline{B} = \emptyset$ and $B \cap \overline{A} = \emptyset$. Then A and B are open in E because $A = E \setminus \overline{B}$ and $B = E \setminus \overline{A}$, so we get a contradiction. $\square$

Theorem. A subset $E \subseteq \mathbb{R}$ is connected if and only if it is an interval: for all $x, y \in E$, if $x < z < y$, then $z \in E$.

Proof. ($\Rightarrow$) Assume E is not an interval. Then there exist $x, y \in E$ and $z \in (x, y)$ such that $z \notin E$. Let $A_z = E \cap (-\infty, z)$ and $B_z = E \cap (z, \infty)$. A_z is non-empty (contains x) and B_z is non-empty (contains y). $E = A_z \cup B_z$. $\text{cl}(A_z) \subseteq (-\infty, z]$, so $\text{cl}(A_z) \cap B_z = \emptyset$. $\text{cl}(B_z) \subseteq [z, \infty)$, so $A_z \cap \text{cl}(B_z) = \emptyset$. Thus, A_z and B_z are non-empty and separated, so E is not connected.

($\Leftarrow$) Assume E is not connected. Then $E = A \cup B$ for non-empty, disjoint, open sets A, B, which we know are intervals in $\mathbb{R}$. Thus E is not an interval since the union of two disjoint intervals is not an interval. $\square$

5.10 The Cantor Set

Our goal is to show that there exists a perfect set in $\mathbb{R}$ that contains no interior points.

Definition (Cantor Set). We will define the **Cantor Set** in the following way:

- $C_0 = [0, 1]$
- $C_1 = [0, 1/3] \cup [2/3, 1]$ (Remove the open middle third)
- $C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$ (Remove middle thirds of C_1)

- Continue this inductively. C_n is the union of 2^n disjoint closed intervals, each of length 3^{-n}.

- The Cantor Set is the intersection $\mathcal{C} = \bigcap_{n=0}^{\infty} C_n$

Proposition. $\mathcal{C}$ is compact, has no interior points, and is perfect.

Proof. $\mathcal{C}$ is compact. Each C_n is a finite union of compact sets, so C_n is compact. $\mathcal{C}$ is the intersection of a nested family of non-empty compact sets, so $\mathcal{C}$ is compact and non-empty.

$\mathcal{C}$ has no interior ($\text{int}(\mathcal{C}) = \emptyset$). This is because $\mathcal{C}$ contains no open interval. Any interval (α, β) must contain a "middle third" segment $(\frac{3k+1}{3^m}, \frac{3k+2}{3^m})$ for some m, and these segments are explicitly removed from $\mathcal{C}$.

$\mathcal{C}$ is perfect. $\mathcal{C}$ is closed (because it's compact). We must show $\mathcal{C}$ has no isolated points. Let $x \in \mathcal{C}$. For any $n \in \mathbb{N}$, x is in one of the 2^n intervals of C_n. Call this interval I_n and let x_n be endpoint of I_n such that $x_n \neq x$. By construction, endpoints are never removed, so $x_n \in \mathcal{C}$ for all n. Further, $|x - x_n| \leq \text{length}(I_n) = 3^{-n}$. Since $3^{-n} \to 0$, we have $x_n \to x$. We have found a sequence in $\mathcal{C}$ converging to x, which was arbitrary. $\square$

Proposition. We can describe the sets $C_n = \left\{\sum_{n=1}^{\infty} \frac{\varepsilon_j}{3^j} : \varepsilon_j \in \{0,1,2\} \text{ and } \varepsilon_j \neq 1 \text{ for } 1 \leq j \leq n\right\}$. Qualitatively, at stage n, we have removed the middle third from each interval n times. This leaves 2^n intervals, and each interval corresponds exactly to a choice of the first n ternary digits being 0 or 2 (never 1). Within each interval, the later ternary digits are free because the middle thirds corresponding to those positions haven't been removed yet, and with an infinite series we can still hit every element in these intervals. Then, we can describe $\mathcal{C}$, the intersection of all C_n's as

$$\mathcal{C} = \left\{\sum_{j=1}^{\infty} \frac{\epsilon_j}{3^j} : \epsilon_j \in \{0,2\}\right\}$$

It turns out that every number in $\mathcal{C}$ is uniquely determined by the choices for this sequence because it is impossible to make up an otherwise lost choice. Thus there is a bijection between $\{0,2\}^{\mathbb{N}}$ and $\mathcal{C}$. Namely,

$$\phi(z) = \frac{2}{3} \sum_{j=0}^{\infty} \frac{z_j}{3^j} \quad \text{for} \quad z = (z_j)_{j \in \mathbb{N}}, \quad z_j \in \{0,1\}$$

Thus, since $\{0,2\} \cong \{0,1\}$, we conclude that $\text{card}(\mathcal{C}) = \mathfrak{c}$.

6 Continuity

6.1 Limits of Functions

Definition. Let (X, ρ_X) and (Y, ρ_Y) be metric spaces. Suppose $E \subseteq X$, $f : E \to Y$, and p is a **limit point** of E. We write $\lim_{x\to p} f(x) = q$ if there is a point $q \in Y$ satisfying $(\forall\, \epsilon > 0)\ \exists\, \delta_\epsilon > 0$ such that $\rho_Y(f(x), q) < \epsilon$ for all points $x \in E$ for which $0 < \rho_X(x, p) < \delta$.

Example. If $X = Y = \mathbb{R}$ equipped with the standard metric, then we say for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in E$, if $0 < |x - p| < \delta$, then $|f(x) - q| < \epsilon$.

Theorem. Let (X, ρ_X), (Y, ρ_Y) be metric spaces, $E \subseteq X$, $f : E \to Y$, and p be a limit point of E. Then the following definitions are equivalent:

1. (a) (Cauchy) $\lim_{x\to p} f(x) = q$
2. (b) (Heine) $\lim_{n\to\infty} f(p_n) = q$ for every sequence $(p_n)_{n\in\mathbb{N}}$ in E such that $p_n \neq p$ and $\lim_{n\to\infty} p_n = p$.

Proof. (a) $\Rightarrow$ (b). Suppose (a) holds. Choose a sequence (p_n) as in (B). Let $\epsilon > 0$ be given. By (a), there exists $\delta > 0$ such that $\rho_Y(f(x), q) < \epsilon$ if $x \in E$ and $0 < \rho_X(x, p) < \delta$. Since $p_n \to p$, there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $0 < \rho_X(p_n, p) < \delta$. Thus, for $n \geq N$, we have $\rho_Y(f(p_n), q) < \epsilon$, showing that (b) holds.

(b) $\Rightarrow$ (a). We show the contrapositive; suppose (a) is false. Then there exists some $\epsilon > 0$ such that for every $\delta > 0$, there exists a point $x \in E$ for which $\rho_Y(f(x), q) \geq \epsilon$ but $0 < \rho_X(x, p) < \delta$. Taking $\delta_n = 1/n$ for each $n \in \mathbb{N}$, we can find a sequence $(p_n)_{n\in\mathbb{N}}$ in E satisfying $0 < \rho_X(p_n, p) < 1/n$ (so $p_n \to p$) but $\rho_Y(f(p_n), q) \geq \epsilon$. $\square$

Theorem (Algebraic Limits). Suppose (X, ρ_X) is a metric space, $E \subseteq X$, and p is a limit point of E. Let $f, g : E \to \mathbb{R}$ be functions such that $\lim_{x\to p} f(x) = A$ and $\lim_{x\to p} g(x) = B$. Then,

1. $\lim_{x\to p}(f + g)(x) = A + B$
2. $\lim_{x\to p}(f \cdot g)(x) = A \cdot B$
3. $\lim_{x\to p}(\frac{f}{g})(x) = \frac{A}{B}$, provided $B \neq 0$ and $g(x) \neq 0$ for $x \in E$.

Proof. we use the equivalence of the Cauchy and Heine definitions for limits, and the algebraic limit theorem for sequences. $\square$

6.2 Continuity

Definition. Suppose (X, ρ_X) and (Y, ρ_Y) are metric spaces, $E \subseteq X$, $p \in E$, and $f : E \to Y$. We say the function f is **continuous at point** p if for every $\epsilon > 0$ there exists $\delta > 0$ such that $\rho_Y(f(x), f(p)) < \epsilon$ for all points $x \in E$ for which $\rho_X(x, p) < \delta$. If f is continuous at every point of E, then f is said to be **continuous on** E.

Proposition. If p is an **isolated point** of E (i.e., not a limit point), then f is automatically continuous at p. Namely, we can pick $\delta > 0$ small enough so that the only point $x \in E$ with $\rho_X(x, p) < \delta$ is $x = p$. In this case, $\rho_Y(f(x), f(p)) = 0 < \epsilon$ is trivially satisfied. If p is a **limit point** of E, then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.

Example (Dirichlet Function). Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \begin{cases} 1, x \in \mathbb{Q} \\ 0, x \notin \mathbb{Q} \end{cases}$. Then f is not continuous at $x = 0$. Namely, consider the sequence $(a_n)_{n\in\mathbb{N}}$ where $a_n = \sqrt{2}/n$. Then $\lim_{n\to\infty} f(a_n) = 0$, but $f(0) = 1$.

Theorem (Compositions). Suppose (X, ρ_X), (Y, ρ_Y), and (Z, ρ_Z) are metric spaces, $E \subseteq X$, $f : E \to Y$, and $g : f[E] \to Z$. Define $h : E \to Z$ by $h(x) = g(f(x))$. If f is continuous at $p \in E$ and g is continuous at $f(p)$, then h is continuous at p.

Proof. Let $\epsilon > 0$ be given. Since g is continuous at $f(p)$, there is $\eta > 0$ such that $\rho_Z(g(y), g(f(p))) < \epsilon$ whenever $\rho_Y(y, f(p)) < \eta$ and $y \in f[E]$. Since f is continuous at p, there is $\delta > 0$ such that $\rho_Y(f(x), f(p)) < \eta$ whenever $\rho_X(x, p) < \delta$ and $x \in E$. It follows that $\rho_Z(h(x), h(p)) = \rho_Z(g(f(x)), g(f(p))) < \epsilon$ whenever $\rho_X(x, p) < \delta$ and $x \in E$. $\square$

Theorem. A mapping f of a metric space (X, ρ_X) into a metric space (Y, ρ_Y) is continuous on X if and only if $f^{-1}[V]$ is open in X for every open set V in Y.

Proof. ($\Rightarrow$) Suppose f is continuous on X and $V \subseteq Y$ is open. We must show $f^{-1}[V]$ is open. Let $p \in f^{-1}[V]$. This means $f(p) \in V$. Since V is open, there is some $\epsilon > 0$ such that $B_{\rho_Y}(f(p), \epsilon) \subseteq V$. Since f is continuous at p, there is $\delta > 0$ such that if $\rho_X(x, p) < \delta$, then $\rho_Y(f(x), f(p)) < \epsilon$. This means $x \in B_{\rho_X}(p, \delta)$ implies $f(x) \in B_{\rho_Y}(f(p), \epsilon) \subseteq V$. Thus $B_{\rho_X}(p, \delta) \subseteq f^{-1}[V]$. Since p was arbitrary, $f^{-1}[V]$ is open.

($\Leftarrow$) Suppose $f^{-1}[V]$ is open in X for any open $V \subseteq Y$. Fix $p \in X$ and $\epsilon > 0$. Consider the set $V = B_{\rho_Y}(f(p), \epsilon)$, which is an open set in Y. By hypothesis, $f^{-1}[V]$ is open in X. Since $p \in f^{-1}[V]$, there must be some $\delta > 0$ such that $B_{\rho_X}(p, \delta) \subseteq f^{-1}[V]$. This means if $\rho_X(x, p) < \delta$, then $x \in f^{-1}[V]$, which implies $f(x) \in V = B_{\rho_Y}(f(p), \epsilon)$. $\square$

Corollary. A mapping $f : X \to Y$ is continuous iff $f^{-1}[C]$ is closed in X for any closed set $C \subseteq Y$.

Proof. This follows from the theorem and the fact that $f^{-1}[E^c] = (f^{-1}[E])^c$. $\square$

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. $A = \{x \in \mathbb{R} : f(x) > a\}$ is open.

Proof. The set (a, ∞) is open in $\mathbb{R}$. We can write $A = \{x \in \mathbb{R} : f(x) \in (a, \infty)\} = f^{-1}[(a, \infty)]$. Since f is continuous, A is the pre-image of an open set, and is therefore open by the theorem. $\square$

Example. Let $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$. Then A is open in $\mathbb{R}^2$ with the Euclidean metric.

Proof. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x, y) = x^2 + y^2$. Each function $(x, y) \mapsto x^2$ and $(x, y) \mapsto y^2$ is continuous, and so their sum f is continuous. Observe that $A = \{(x, y) \in \mathbb{R}^2 : f(x, y) < 1\} = f^{-1}[B(0, 1)]$. □

Theorem. Let $f, g : X \to \mathbb{R}$ be two continuous functions on a metric space (X, ρ_X). Then $f + g$ and $f \cdot g$ are continuous. Furthermore, f/g is continuous, assuming $g(x) \neq 0$ for all $x \in X$.

Example. $f(x) = x^a$ for $a \in \mathbb{R}$ is continuous on $(0, \infty)$.

Example. every polynomial $p(x) = a_n x^n + ... + a_0$ is continuous on $\mathbb{R}$.

Example. if f, g are continuous, then $\max\{f, g\} = \frac{f+g+|f-g|}{2}$ and $\min\{f, g\} = \frac{f+g-|f-g|}{2}$ are continuous.

Theorem. Suppose f is a continuous mapping of a **compact** metric space (X, ρ_X) into a metric space (Y, ρ_Y). Then $f[X]$ is **compact** in Y.

Proof. Let $(V_\alpha)_{\alpha \in A}$ be an open cover of $f[X]$. Since f is continuous, each set $f^{-1}[V_\alpha]$ is open in X. The collection $\{f^{-1}[V_\alpha]\}_{\alpha \in A}$ forms an open cover of X (since $X \subseteq \bigcup_{\alpha \in A} f^{-1}[V_\alpha]$). Since X is compact, this cover has a finite subcover: $X \subseteq \bigcup_{j=1}^{n} f^{-1}[V_{\alpha_j}]$. Applying f to both sides (and using $f[f^{-1}[E]] \subseteq E$), we have that $f[X] \subseteq f[\bigcup_{j=1}^{n} f^{-1}[V_{\alpha_j}]] \subseteq \bigcup_{j=1}^{n} V_{\alpha_j}$. Thus, we have found a finite subcover for $f[X]$, so $f[X]$ is compact. □

Definition. A function $f : E \to \mathbb{R}$ is said to be **bounded** if there is $M > 0$ such that $|f(x)| \leq M$ ($\forall\, x \in E$).

Corollary. If $f : X \to \mathbb{R}$ is continuous on a compact metric space X, then $f[X]$ is closed and bounded in $\mathbb{R}$.

Theorem (Extreme Value Theorem). Suppose $f : X \to \mathbb{R}$ is continuous on a compact metric space X. Let $M = \sup_{p \in X} f(p)$ and $m = \inf_{p \in X} f(p)$ Then there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

Proof. The image $f[X]$ is compact in $\mathbb{R}$, so it is closed and bounded. Since $f[X]$ is closed, its supremum (M) and infimum (m) must be members of the set $f[X]$. □

Theorem. Suppose $f : X \to Y$ is continuous and bijective, where X is compact. Then the inverse mapping $f^{-1} : Y \to X$ is continuous.

Proof. To show f^{-1} is continuous, we use the topological characterization. We must show that for any open set V in X, the set $(f^{-1})^{-1}[V]$ is open in Y. Note that $(f^{-1})^{-1}[V] = f[V]$. So, we must prove $f[V]$ is open in Y for every open $V \subseteq X$. Let $V \subseteq X$ be open. Its complement V^c is closed in X. Since X is compact, the closed subset V^c is also compact. Since f is continuous, the image $f[V^c]$ is compact in Y. Compact subsets of a metric space are closed, so $f[V^c]$ is closed in Y. In particular, f being a bijection means that $f[V] = (f[V^c])^c$.

Since $f[V^c]$ is closed, its complement $f[V]$ must be open. $\square$

Theorem. If $f : X \to Y$ is a continuous mapping and E is a connected subset of X, then $f[E]$ is **connected** in Y.

Proof. Assume $f[E]$ is not connected. Then $f[E] = A \cup B$ where A and B are non-empty separated sets in Y. Let $G = E \cap f^{-1}[A]$ and $H = E \cap f^{-1}[B]$. Then $E = G \cup H$, and neither G nor H is empty. We must show G and H are separated (i.e., $\mathrm{cl}(G) \cap H = \emptyset$ and $G \cap \mathrm{cl}(H) = \emptyset$), which will contradict the fact that E is connected. $f[H] \subseteq B$. $f[G] \subseteq A \subseteq \mathrm{cl}(A)$. Since f is continuous, $f[\mathrm{cl}(G)] \subseteq \mathrm{cl}(f[G]) \subseteq \mathrm{cl}(A)$. Since A and B are separated, $\mathrm{cl}(A) \cap B = \emptyset$. Consider $f[H \cap \mathrm{cl}(G)] \subseteq f[H] \cap f[\mathrm{cl}(G)] \subseteq B \cap \mathrm{cl}(A) = \emptyset$. Since $f[H \cap \mathrm{cl}(G)] = \emptyset$, we must have $H \cap \mathrm{cl}(G) = \emptyset$. A symmetric argument shows $\mathrm{cl}(H) \cap G = \emptyset$. Thus G and H are separated, contradicting that E is connected. $\square$

Theorem (Intermediate Value/Darboux Property). Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function on the interval $[a, b]$. If $f(a) < f(b)$ and c is a number such that $f(a) < c < f(b)$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.

Proof. The interval $[a, b]$ is a connected set in $\mathbb{R}$. Since f is continuous, the image $f[[a, b]]$ is connected in $\mathbb{R}$. The connected subsets of $\mathbb{R}$ are intervals. Thus $f[[a, b]]$ is an interval. Since $f(a)$ and $f(b)$ are in this interval, any value c between them must also be in the interval. Therefore, $c \in f[[a, b]]$, which means there exists $x \in [a, b]$ such that $f(x) = c$. Since $c \neq f(a)$ and $c \neq f(b)$, we must have $x \in (a, b)$. $\square$

6.3 Uniform and Lipschitz Continuity

Definition. Let (X, ρ_X) and (Y, ρ_Y) be two metric spaces, and let $f : X \to Y$. We say that f is **uniformly continuous on** X if for every $\epsilon > 0$ there exists $\delta > 0$ such that $\rho_Y(f(x), f(y)) < \epsilon$ for all $x, y \in X$ for which $\rho_X(x, y) < \delta$. The key point here is that δ only depends on ϵ and not on $x \in X$. Thus uniform continuity is a stronger condition that implies continuity on the whole set.

Theorem. Let f be a continuous mapping of a compact metric space (X, ρ_X) into a metric space (Y, ρ_Y). Then f is uniformly continuous on X.

Proof. Let $\epsilon > 0$ be given. Since f is continuous at every point, for each $p \in X$, there is a $\delta_p > 0$ such that if $q \in B(p, \delta_p)$, then $\rho_Y(f(p), f(q)) < \epsilon/2$. Consider the collection of open balls $\{B(p, \delta_p/2)\}_{p \in X}$. This is an open cover of X. Since X is compact, this cover has a finite subcover. That is, there exist $p_1, ..., p_n \in X$ such that $X \subseteq \bigcup_{k=1}^n B(p_k, \delta_{p_k}/2)$. Now, let $\delta = \frac{1}{2}\min(\delta_{p_1}, ..., \delta_{p_n})$. If $p, q \in X$ are any two points such that $\rho_X(p, q) < \delta$, with say, $p \in B(p_m, \delta_{p_m}/2)$ for some m, we find the distance from q to p_m. By triangle inequality, $\rho_X(q, p_m) \leq \rho_X(q, p) + \rho_X(p, p_m) < \delta + \frac{\delta_{pm}}{2} < \delta_{p_m}$. This means both p and q are in the ball $B(p_m, \delta_{p_m})$. Hence by continuity of f, $\rho_Y(f(p), f(p_m)) < \epsilon/2$ and $\rho_Y(f(q), f(p_m)) < \epsilon/2$ and by triangle inequality, $\rho_Y(f(p), f(q)) \leq \epsilon$. $\square$

Example. Show $f(x) = \frac{1}{x}$ is not uniformly continuous on $(0, 1)$.

Proof. Assume f is uniformly continuous. Take $\epsilon = 1$. Then there is $\delta > 0$ such that for all x, y, $|x-y| < \delta \Rightarrow |f(x)-f(y)| < 1$. Let $x = 1/n$, $y = 1/(n+1)$. Then $|x-y| = 1/(n(n+1))$, which can be made $< \delta$ for large n. But $|f(x) - f(y)| = |n - (n+1)| = 1$, contradicting $1 < 1$. So f is not uniformly continuous. ☐

Definition. Let (X, ρ) be a metric space. A function $\phi : X \to X$ is a **Lipschitz mapping** if there exists a constant $C_\phi > 0$ such that $\rho(\phi(x), \phi(y)) \le C_\phi \rho(x, y)$. Every Lipschitz mapping is uniformly continuous.

Example. Distance from a set. Let (X, ρ) be a metric space and $\emptyset \neq E \subseteq X$. We will define the distance for $x \in X$ from the set E as $\rho_E(x) = \inf\{\rho(x, z) : z \in E\}$. Then $\rho(x, E) = 0$ iff $x \in \mathrm{cl}(E)$ and $f(x) = \rho(x, E)$ is Lipschitz with constant 1. Namely, $|\rho(x, E) - \rho(y, E)| \le \rho(x, y)$.

6.4 The Banach Contraction Principle

Definition. We say $\phi : X \to X$ is a **contraction** contraction map if it is Lipschitz with constant $c < 1$: $\rho(\phi(x), \phi(y)) \le c\rho(x, y)$.

Theorem (Banach Contraction Principle). If (X, ρ) is a complete metric space and $\phi : X \to X$ is a **contraction**, then there exists a unique $x \in X$ such that $\phi(x) = x$. This x is called a **fixed point**.

Proof. Uniqueness. Suppose x and y are two fixed points. Then by contraction, $0 \le \rho(x, y) = \rho(\phi(x), \phi(y)) \le c\rho(x, y)$ is impossible for $c < 1$ unless $\rho(x, y) = 0$.

Existence. Pick $x_0 \in X$. Define $x_{n+1} = \phi(x_n)$. Then, $\rho(x_{n+1}, x_n) \le c^n \rho(x_1, x_0)$. If $m > n$, then $\rho(x_n, x_m) \le \frac{c^n}{1-c}\rho(x_1, x_0) \to 0$. Thus x_n is Cauchy and converges to some $x \in X$ by completeness. Since contractions are continuous, $\phi(x) = \lim_{n\to\infty} \phi(x_n) = \lim_{n\to\infty} x_{n+1} = x$. ☐

Example. $x = \sqrt{2 + x}$. The sequence $\sqrt{2}, \sqrt{2 + \sqrt{2}}, \sqrt{2 + \sqrt{2 + \sqrt{2}}}, \ldots$ is generated by $x_{n+1} = \sqrt{x_n + 2}$. Note that $|\sqrt{x+2} - \sqrt{y+2}| = \frac{|x-y|}{\sqrt{x+2}+\sqrt{y+2}} \le \frac{1}{2\sqrt{2}}|x - y|$. Hence $f(x) = \sqrt{x+2}$ is a contraction with constant $1/(2\sqrt{2})$; so the fixed point exists and is unique.

6.5 Discontinuities

Definition. If a function f is not continuous at a point x in its domain, we say f has a **discontinuity** at x. In particular for $f : (a, b) \to \mathbb{R}$, we define the right and left hand limits as $f(x+) = q$ if $f(t_n) \to q$ for every sequence (t_n) in (x, b) with $t_n \to x$ and $f(x-) = q$ if $f(t_n) \to q$ for every sequence (t_n) in (a, x) with $t_n \to x$. Then, the limit $\lim_{t\to x} f(t)$ exists iff $f(x+) = f(x-)$. If f is discontinuous at x, we further categorize it. Let $f : (a, b) \to \mathbb{R}$ be discontinuous at x. We classify this as discontinuities if both $f(x+)$ and $f(x-)$ exist. They are either jumps if $f(x+) \neq f(x-)$ or removable if $f(x+) = f(x-) \neq f(x)$. Otherwise, at least one of $f(x+)$ or $f(x-)$ fails to exist.

Definition. We also define one-sided continuities as continuous from the left if $f(x-) = f(x)$ and continuous from the right if $f(x+) = f(x)$. Thus a function is continuous from left and right if and only if it is continuous.

Definition. We say a function $f : (a, b) \to \mathbb{R}$ is **monotonically increasing** if $x < y$ implies $f(x) \leq f(y)$.

Theorem. If f is monotonic on (a, b), then $f(x-)$ and $f(x+)$ exist at every $x \in (a, b)$, with $\sup_{a<t<x} f(t) = f(x-) \leq f(x) \leq f(x+) = \inf_{x<t<b} f(t)$. In particular, if $a < x < y < b$, then $f(x+) \leq f(y-)$.

Corollary. Monotonic functions have no discontinuities of the second kind.

Theorem. For monotonic f, the set of discontinuities is **at most countable**.

Proof. Let E be the set of points of discontinuity. Each $x \in E$ satisfies $f(x-) < f(x+)$. Choose a rational $r(x)$ with $f(x-) < r(x) < f(x+)$. This defines an injective function $r : E \to \mathbb{Q}$. Hence E is at most countable. $\square$

7 Differentiation

Definition. Let $f : [a, b] \to \mathbb{R}$. For any $x \in [a, b]$, we form the quotient function for $a < t < b$, $t \neq x$. $\phi(t) = \frac{f(t)-f(x)}{t-x}$ The **derivative** $f'(x)$ is defined as $f'(x) = \lim_{t \to x} \phi(t)$ provided the limit exists. If f' is defined at a point x, f is said to be **differentiable** at x. If f' is defined at every point of a set E, f is differentiable on E. (At endpoints a and b, the derivative is defined via right-hand and left-hand limits, respectively.)

Theorem. If f is differentiable at x, then f is **continuous** at x.

Proof. $\lim_{t \to x}(f(t) - f(x)) = \lim_{t \to x} \frac{f(t)-f(x)}{t-x}(t - x) = f'(x) \cdot 0 = 0$. □

Remark. The converse is false.

Proposition. Let f, g be differentiable at x. Then,

1. Sum: $(f + g)'(x) = f'(x) + g'(x)$.
2. Product: $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.
3. Quotient: $(\frac{f}{g})'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{(g(x))^2}$, provided $g(x) \neq 0$.

7.1 The Chain Rule

Theorem (Chain Rule). Suppose f is continuous on $[a, b]$ and differentiable at x, and g is defined on an interval containing the range of f and is differentiable at $f(x)$. If $h(t) = g(f(t))$, then h is differentiable at x and:

$$h'(x) = g'(f(x))f'(x)$$

Sketch. Define error terms $u(t)$ and $v(s)$ such that $f(t) - f(x) = (t - x)(f'(x) + u(t))$ and $g(s) - g(y) = (s - y)(g'(y) + v(s))$. Substitute $s = f(t)$ and $y = f(x)$ to show the difference quotient for h converges to the product of derivatives. □

7.2 Local Extrema

Definition. Let $f : X \to \mathbb{R}$. We say f has a **local maximum** at p if $f(q) \leq f(p)$ for all q in a neighborhood $(p - \delta, p + \delta)$. Analogously, we say f has a **local minimum** at p if $f(q) \geq f(p)$ for all q in a neighborhood.

Theorem. If f has a local extremum at $x \in (a, b)$ and $f'(x)$ exists, then $f'(x) = 0$.

Sketch. Examine the sign of the difference quotient $\frac{f(t)-f(x)}{t-x}$. For a local max, the quotient is ≥ 0 for $t < x$ (limit ≥ 0) and ≤ 0 for $t > x$ (limit ≤ 0). Thus, the limit must be 0. □

7.3 The Mean-Value Theorem

Theorem (Mean Value Theorem). If f, g are continuous on $[a, b]$ and differentiable in (a, b), then there exists $x \in (a, b)$ such that $(f(b) - f(a))g'(x) = (g(b) - g(a))f'(x)$. In particular, setting $g(x) = x$ we obtain that there exists $x \in (a, b)$ such that $f(b) - f(a) = f'(x)(b - a)$.

Sketch. Construct the auxiliary function $h(t) = (f(b) - f(a))g(t) - (g(b) - g(a))f(t)$. Note that $h(a) = h(b)$. Since h is continuous on a compact set, it attains a maximum and minimum. If extrema are at endpoints, h is constant ($h' = 0$). If an extremum is strictly inside, the derivative is 0 there. □

Corollary. If f is differentiable on (a, b), then $f'(x) \geq 0 \implies f$ is monotonically increasing, $f'(x) = 0 \implies f$ is constant, and $f'(x) \leq 0 \implies f$ is monotonically decreasing.

Theorem. If f is differentiable on (a, b) and $f'(a) < \lambda < f'(b)$, there exists $x \in (a, b)$ such that $f'(x) = \lambda$, even if the derivative is not continuous.

7.4 L'Hôpital's Rule

Theorem (L'Hôpital's Rule). Suppose f, g are differentiable in (a, b), $g'(x) \neq 0$, and $\lim_{x\to a} \frac{f'(x)}{g'(x)} = A$. If $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, or $\lim_{x\to a} g(x) = +\infty$ (regardless of f), then $\lim_{x\to a} \frac{f(x)}{g(x)} = A$.

Proof. This is a general version of Stolz' theorem. It relies on the Generalized MVT to relate the ratio of functions to the ratio of derivatives. □

7.5 Inverse Functions

Theorem. If $f : [a, b] \to \mathbb{R}$ is continuous and strictly increasing, its inverse g (defined on the range) is also continuous and strictly increasing.

Theorem. If f is continuous, strictly increasing, and differentiable with $f'(x) > 0$, then the inverse g is differentiable and $g'(y) = \frac{1}{f'(g(y))}$ for $y \in f((a, b))$.

Remark. It is possible to see this by applying the chain rule, but this makes an assumption that g is differentiable.

7.6 Higher Order Derivatives

Definition. If f' is differentiable, its derivative is the **second derivative**, denoted f''. We denote the n-th derivative as $f^{(n)}$. For $f^{(n)}(x)$ to exist, $f^{(n-1)}$ must exist in a neighborhood of x and be differentiable at x.

7.7 Convex and Concave Functions

Definition. A function $f : (a, b) \to \mathbb{R}$ is **convex** if for all $x, y \in (a, b)$ and $\alpha, \beta \in [0, 1]$ with $\alpha + \beta = 1$:

$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y)$$

Geometrically, the graph of the function lies below the secant line connecting any two points. We call the function **concave** if the inequality is reversed ($\geq$).

Proposition.

1. **Slope Inequality:** If f is convex and $a < s < t < u < b$, the slopes of the secant lines are non-decreasing $\frac{f(t)-f(s)}{t-s} \leq \frac{f(u)-f(s)}{u-s} \leq \frac{f(u)-f(t)}{u-t}$.

2. **Jensen's Inequality:** For convex f, $f(\sum \lambda_i x_i) \leq \sum \lambda_i f(x_i)$ where $\sum \lambda_i = 1$.

3. **Continuity:** Convex functions are continuous on open intervals.

Theorem. Let f is continuous on $[a, b]$ and twice differentiable on (a, b).

- If $f''(x) > 0$ on (a, b), then f is **strictly convex**.
- If $f''(x) < 0$ on (a, b), then f is **strictly concave**.

Proof. Uses the Mean-Value Theorem twice on an auxiliary function representing the difference between the secant line and the curve. □

8 Power Series and Applications

8.1 Power Series

Definition. A **power series** is a series of the form $\sum_{n=0}^{\infty} c_n x^n$, where (c_n) is a sequence of real coefficients and $x \in \mathbb{R}$.

Definition. The **radius of convergence**, R, is defined by $\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|}$ and $R = \frac{1}{\alpha}$. If $\alpha = 0$, then $R = +\infty$. The endpoints must be checked individually.

Theorem. The series $\sum c_n x^n$ converges if $|x| < R$ and diverges if $|x| > R$.

Proof. We use the ratio test with $a_n = c_n x^n$ so $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x|}{R}$. □

Example. $\sum_{n=0}^{\infty} n^n x^n$. $\alpha = \lim \sqrt[n]{n^n} = \lim n = \infty$, so $R = 0$.

Example. $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. $\alpha = \lim \sqrt[n]{1/n!} = 0$, so $R = +\infty$. This is the series for e^x.

Example. $\sum_{n=0}^{\infty} x^n$. $R = 1$. It diverges for $|x| = 1$. The sum is $\frac{1}{1-x}$ for $|x| < 1$.

Example. $\sum_{n=1}^{\infty} \frac{x^n}{n}$. $R = 1$. At $x = 1$ it diverges (harmonic series); at $x = -1$ it converges (alternating harmonic).

8.2 The Exponential

Definition. We define the **exponential function** as the power series:

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

for any $z \in \mathbb{C}$. The series converges absolutely for all z (by the ratio test), so $E(z)$ is well-defined. We usually denote $E(x)$ by e^x.

Proposition.

- **Multiplication:** $E(z)E(w) = E(z + w)$ for all $z, w \in \mathbb{C}$.

  Proof. Uses the Cauchy product of absolutely convergent series and the Binomial Theorem. □

- **Inverses:** $E(z)E(-z) = E(0) = 1$. Thus $E(z) \neq 0$ and $E(-z) = 1/E(z)$.
- **Positivity:** $E(x) > 0$ for all $x \in \mathbb{R}$. Since $E(x) > 0$ for $x > 0$ (sum of positive terms) and $E(-x) = 1/E(x)$, it is positive everywhere on $\mathbb{R}$.

Proposition.

- **Limits:** $\lim_{x\to+\infty} E(x) = +\infty$ and $\lim_{x\to-\infty} E(x) = 0$.
- **Monotonicity:** $E(x)$ is strictly increasing on $\mathbb{R}$.
- **Derivative:** $E'(x) = E(x)$ (i.e., $(e^x)' = e^x$).

  Proof. Analyzes the limit of the difference quotient $\frac{E(h)-1}{h}$ using the series definition to show it approaches 1. □

- **Growth Rate:** e^x grows faster than any polynomial. For any $n \in \mathbb{N}$, $\lim_{x\to+\infty} x^n e^{-x} = 0$.

8.3 The Logarithm

Definition. Since $E : \mathbb{R} \to (0, \infty)$ is strictly increasing and differentiable, it has an inverse $L : (0, \infty) \to \mathbb{R}$.

$$E(L(y)) = y \quad \text{for } y > 0$$

We write $L(x) = \log(x)$ (or $\ln(x)$).

Proposition.

- **Derivative:** Differentiating $E(L(x)) = x$ gives $L'(x) = \frac{1}{x}$ for $x > 0$.
- **Product Rule:** $\log(uv) = \log(u) + \log(v)$.
- **Limits:** $\lim_{x\to\infty} \log(x) = +\infty$ and $\lim_{x\to 0} \log(x) = -\infty$.
- **Limit Definition:** $\lim_{n\to\infty} n \log(1 + \frac{x}{n}) = x$.
- **Growth Rate:** $\lim_{x\to\infty} x^{-\alpha} \log(x) = 0$ for any $\alpha > 0$ (log grows slower than any power).

8.4 Applications of the Exponential and Logarithm

Definition. For $x > 0$ and $\alpha \in \mathbb{R}$, we define $x^\alpha = E(\alpha L(x)) = e^{\alpha \log(x)}$. Then, the derivative is $(x^\alpha)' = \alpha x^{\alpha-1}$.

Definition (Euler-Mascheroni Constant). The sequence $a_n = \sum_{k=1}^{n-1} \frac{1}{k} - \log(n)$ is increasing and bounded, converging to a limit γ, known as the **Euler-Mascheroni constant** ($\gamma \approx 0.5772$). This relates the divergence of the harmonic series to the natural logarithm.

Proof. Uses the inequality $\frac{1}{n+1} < \log(\frac{n+1}{n}) < \frac{1}{n}$ derived from e bounds. □

Proposition. For $x > 0$, $\frac{x}{x+1} < \log(1 + x) < x$ and $\frac{2x}{x+2} \leq \log(1 + x)$.

Proof. Analyze the auxiliary function's derivatives and values at $x = 0$. □

Theorem (Weighted Mean Inequality). If $x_i > 0, \alpha_i > 0$ with $\sum_{j=1}^{k} \alpha_j = 1$ then $x_1^{\alpha_1} \cdot ... \cdot x_k^{\alpha_k} \leq \alpha_1 x_1 + ... + \alpha_k x_k$.

Proof. Let $f(x) = \log(x)$. Since $f''(x) = -1/x^2 < 0$, f is **concave**. By the definition of concavity (Jensen's inequality), the log of the weighted geometric mean is less than or equal to the log of the weighted arithmetic mean. $\square$

Corollary (Young's Inequality). If $p, q > 0$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$ (conjugate exponents) and $x, y > 0$, then $xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q$.

Proof. Apply the previous weighted mean result with $\alpha_1 = 1/p, \alpha_2 = 1/q$ and inputs $x_1 = x^p, x_2 = y^q$. $\square$

8.5 Trigonometric Functions

Definition. We can define cosine and sine using the complex exponential function $E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$.

$$C(x) = \frac{E(ix) + E(-ix)}{2}, \quad S(x) = \frac{E(ix) - E(-ix)}{2i}$$

Proposition.

1. $C(x)$ and $S(x)$ are real-valued for real x, since $E(x^*) = E(x)^*$.
2. **Euler's Formula:** $E(ix) = C(x) + iS(x)$.
3. $|E(ix)| = 1$, which implies $C(x)^2 + S(x)^2 = 1$.
4. $C(0) = 1$, and $S(0) = 0$.
5. $C'(x) = -S(x)$ and $S'(x) = C(x)$.

Theorem. We may define $\pi = 2x_0$, where x_0 is the smallest positive number such that $C(x_0) = 0$. This indeed exists. $E(z)$ is periodic with period $2\pi i$. $C(x)$ and $S(x)$ are periodic with period 2π. If $t \in (0, 2\pi)$, then $E(it) \neq 0$. If $z \in \mathbb{C}$ with $|z| = 1$, then there is a unique $t \in (0, 2\pi)$ such that $e^{it} = z$.

Definition.

$$S(x) = \sin x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$$

$$C(x) = \cos x = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!}$$

8.6 Taylor's Theorem and Expansion

Theorem (Taylor's Theorem). Let $f : [a, b] \to \mathbb{R}$ satisfying for some $n \in \mathbb{N}$, $f^{(n-1)}$ is continuous on $[a, b]$ and $f^{(n)}$ exists on (a, b). For distinct $\alpha, \beta \in [a, b]$, let

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k$$

Then there exists $x \in (\alpha, \beta)$ such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n$$

Definition. Further, for any $x, x_0 \in [a, b]$ and $p > 0$, there exists $\theta \in (0, 1)$ such that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + r_n(x)$$

where the remainder term $r_n(x)$ is given by the **Schlömlich-Roche form**:

$$r_n(x) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{n!p}(1 - \theta)^{n+1-p}(x - x_0)^{n+1}$$

We may also get the **Lagrange Remainder** ($p = n + 1$):

$$r_n(x) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{(n + 1)!}(x - x_0)^{n+1}$$

and the **Cauchy Remainder** ($p = 1$):

$$r_n(x) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{n!}(1 - \theta)^n(x - x_0)^{n+1}$$

8.7 Applications

Example. For $|x| < 1$:

$$\log(1 + x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$$

Proof. Use Taylor expansion at $x_0 = 0$. The derivatives are $(\log(1 + x))^{(n)} = (-1)^{n+1} \frac{(n-1)!}{(1+x)^n}$.

- For $0 \leq x < 1$, use Lagrange remainder to show $r_n \to 0$.
- For $-1 < x < 0$, use Cauchy remainder to show $r_n \to 0$.

$\square$

Example (Newton's Binomial Formula). For $\alpha \in \mathbb{R} \setminus \mathbb{N}$ and $|x| < 1$:

$$(1+x)^\alpha = 1 + \sum_{n=1}^{\infty} \binom{\alpha}{n} x^n = 1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} x^n$$

The convergence of the remainder to 0 is shown using Lagrange remainder for $0 < x < 1$ and Cauchy remainder for $-1 < x < 0$.

Remark. Functions exist without power series. As an example, the function $f(x) = e^{-1/x^2}$ for $x \neq 0$ and $f(0) = 0$ is infinitely differentiable, and $f^{(n)}(0) = 0$ for all n. However, the Taylor series at $x = 0$ is identically zero, while $f(x) \neq 0$ for $x \neq 0$. Thus, $f(x)$ is not equal to its Taylor series.

Theorem (Bernoulli's Inequality). For $x > -1, x \neq 0$:

1. $(1+x)^\alpha > 1 + \alpha x$ if $\alpha > 1$ or $\alpha < 0$.
2. $(1+x)^\alpha < 1 + \alpha x$ if $0 < \alpha < 1$.

Proof. Using Taylor's expansion with Lagrange remainder, $(1+x)^\alpha = 1+\alpha x+\frac{\alpha(\alpha-1)(1+\theta x)^{\alpha-2}}{2}x^2$ and we analyze the sign of the remainder term. Alternatively, we analyze the auxiliary function $f(x) = (1+x)^\alpha - (1+\alpha x)$ and its derivative. $\square$

9 Riemann Integration

9.1 Partitions and Riemann Sums

Definition. Let $[a, b]$ be an interval. A **partition** P of $[a, b]$ is a finite set of points $a = x_0 \leq x_1 \leq ... \leq x_{n-1} \leq x_n = b$.

Definition. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. For a partition P, we define $m_i = \inf_{x\in[x_{i-1},x_i]} f(x)$, $M_i = \sup_{x\in[x_{i-1},x_i]} f(x)$, and $\Delta x_i = x_i - x_{i-1}$. Then we define the **upper and lower Riemann sums** as $U(P, f) = \sum_{i=1}^{n} M_i\Delta x_i$ and $L(P, f) = \sum_{i=1}^{n} m_i\Delta x_i$. In particular, we always have $L(P, f) \leq U(P, f)$.

9.2 The Riemann Integral

Definition. We define the upper and lower Riemann integrals as

- **Lower Riemann Integral:** $\underline{\int_a^b} f(x)dx = \sup_P L(P, f)$
- **Upper Riemann Integral:** $\overline{\int_a^b} f(x)dx = \inf_P U(P, f)$

We then say a function f is **Riemann integrable** ($f \in \mathcal{R}([a, b])$) if the upper and lower integrals are equal. The common value is denoted by $\int_a^b f(x)dx$. In particular, we always have the inequality $m(b - a) \leq L(P, f) \leq \int f \leq U(P, f) \leq M(b - a)$.

Example. The Dirichlet function $f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$ on $[0, 1]$ is **not** integrable.

Proof. For any partition, every subinterval contains rationals ($M_i = 1$) and irrationals ($m_i = 0$). Thus, $U(P, f) = 1$ and $L(P, f) = 0$ for all P. Thus, $\overline{\int} f = 1 \neq 0 = \underline{\int} f$. □

9.3 Refinements and Properties

Definition. We say a partition P^* is a **refinement** of a partition P if $P^* \supseteq P$. For partitions P_1, P_2 their **common refinement** is $P^* = P_1 \cup P_2$.

Proposition. If P^* is a refinement of P, then $L(P, f) \leq L(P^*, f)$ and $U(P^*, f) \leq U(P, f)$.

Proof. Suppose P^* contains one extra point x^* in $[x_{i-1}, x_i]$. Let $w_1 = \inf_{[x_{i-1},x^*]} f$ and $w_2 = \inf_{[x^*,x_i]} f$. Clearly $w_1 \geq m_i$ and $w_2 \geq m_i$. The change in the lower sum is $w_1(x^* - x_{i-1}) + w_2(x_i - x^*) - m_i(x_i - x_{i-1}) \geq 0$. We then repeat by induction for the remaining points. The proof for the upper sum is similar. □

Theorem. $\underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx$.

Proof. For any two partitions P_1, P_2, let P^* be their common refinement. Then, by the proposition above, $L(P_1, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P_2, f)$. Thus, $\sup L(P_1, f) \leq \inf U(P_2, f)$. □

9.4 Classes of Integrable Functions

Theorem. $f \in \mathcal{R}([a,b])$ if and only if for every $\epsilon > 0$, there exists a partition P such that $U(P,f) - L(P,f) < \epsilon$.

Proof. We must show that this is equivalent to our original definition. ($\Rightarrow$) If f is integrable, then $\underline{\int} = \overline{\int} = I$. There exist P_1, P_2 such that $L(P_1) > I - \epsilon/2$ and $U(P_2) < I + \epsilon/2$. Let $P = P_1 \cup P_2$. Then $U(P) - L(P) < \epsilon$. ($\Leftarrow$) We have $0 \le \overline{\int} f - \underline{\int} f \le U(P) - L(P) < \epsilon$. Since ϵ is arbitrary, the integrals must be equal. $\square$

Corollary (Riemann Sums). If $f \in \mathcal{R}([a,b])$, then for any partition $P = \{x_0, \dots, x_n\}$ and any choice of points $t_i \in [x_{i-1}, x_i]$, we have $\left|\sum_{i=1}^n f(t_i)\Delta x_i - \int_a^b f\right| < \epsilon$.

Proof. Since f is integrable, given $\epsilon > 0$ there exists a partition P with $U(P,f) - L(P,f) < \epsilon$. Also, $[x_{i-1}, x_i]$, $m_i \le f(t_i) \le M_i$, so $L(P,f) = \sum m_i \Delta x_i \le \sum f(t_i)\Delta x_i \le \sum M_i \Delta x_i = U(P,f)$. Also $L(P,f) \le \int_a^b f \le U(P,f)$. Thus $\left|\sum f(t_i)\Delta x_i - \int_a^b f\right| \le \max\{U(P,f) - \int f,\ \int f - L(P,f)\}$. But this is $\le U(P,f) - L(P,f) < \epsilon$. $\square$

We may now show that certain classes of functions are always Riemann integrable.

Theorem. If f is continuous on $[a,b]$, then $f \in \mathcal{R}([a,b])$.

Proof. Since f is continuous on a compact set, it is uniformly continuous. For any $\epsilon > 0$, choose δ such that $|x-y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b-a}$. Choose a partition with $\Delta x_i < \delta$. Then $M_i - m_i < \frac{\epsilon}{b-a}$. $U(P,f) - L(P,f) = \sum (M_i - m_i)\Delta x_i < \frac{\epsilon}{b-a}\sum \Delta x_i = \epsilon$. $\square$

Theorem. If f is **monotonic** on $[a,b]$, then $f \in \mathcal{R}([a,b])$.

Proof. Assume f is increasing. Choose equal partition $\Delta x_i = \frac{b-a}{n}$. Then $M_i - m_i = f(x_i) - f(x_{i-1})$. $U - L = \sum (f(x_i) - f(x_{i-1}))\frac{b-a}{n} = \frac{b-a}{n}(f(b) - f(a))$ as a telescoping sum. We then take n large to make this expression $< \epsilon$. $\square$

Theorem. If f is bounded and has **finitely many discontinuities**, then $f \in \mathcal{R}([a,b])$.

Proof. Cover the discontinuities with intervals of total length small enough (ϵ). On the rest of the interval, f is uniformly continuous. Construct a partition that handles the discontinuity intervals and the continuous parts separately to bound $U - L$. $\square$

9.5 Algebra of Integrals

Proposition (Algebraic Properties of Integrals).

- **Linearity:** $\int (f_1 + f_2) = \int f_1 + \int f_2$ and $\int cf = c\int f$.
- **Order:** If $f_1 \le f_2$, then $\int f_1 \le \int f_2$.
- **Additivity:** $\int_a^b f = \int_a^c f + \int_c^b f$.

- **Products:** If $f, g \in \mathcal{R}$, then $fg \in \mathcal{R}$.
- **Absolute Value:** $|\int f| \leq \int |f|$.

Theorem (Change of Variables). Let $\phi : [A, B] \to [a, b]$ be strictly increasing and continuous, with $\phi' \in \mathcal{R}$. If $f \in \mathcal{R}([a, b])$, then:

$$\int_a^b f(x)dx = \int_A^B f(\phi(t))\phi'(t)dt$$

Proof. Relates partitions of $[a, b]$ to $[A, B]$ via $y_i = \phi(x_i)$. Using Mean Value Theorem, $\Delta y_i = \phi'(t_i)\Delta x_i$. The Riemann sums converge to the integrals. □

Theorem (Fundamental Theorem of Calculus). Let $f \in \mathcal{R}([a, b])$ with $F(x) = \int_a^x f(y)dy$. We note that F is continuous on $[a, b]$ since it has a derivative. If f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$. In particular, if $f \in \mathcal{R}$ and there exists a differentiable function F such that $F' = f$, then:

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

Proof. $|\frac{F(t)-F(x_0)}{t-x_0} - f(x_0)| = |\frac{1}{t-x_0}\int_{x_0}^t (f(y)-f(x_0))dy|$. Using continuity of f, $|f(y)-f(x_0)| < \epsilon$, so the whole expression is $< \epsilon$. For any partition, $F(b) - F(a) = \sum(F(x_i) - F(x_{i-1}))$. By MVT, this equals $\sum f(t_i)\Delta x_i$, which is a Riemann sum. As mesh goes to 0, this converges to the integral. □

Theorem (Integration by Parts). If F, G are differentiable with integrable derivatives f, g:

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

Proof. Apply FTC to $H(x) = F(x)G(x)$, noting that $H' = Fg + fG$. □