Notes on Real Analysis

Pranav Konda

December 21, 2020

Contents

Chapter 1

Preface

These are my notes from an introductory Analysis class. I tried to be as detailed as I could, but most of these are taken off of lectures or lecture notes. If you do have any questions, you can contact me at pranavk123455@gmail.com, and I'll try to respond as soon as I can.

Chapter 2

Sets

✾ 2.1 Introduction

Sets have important operations, such as the **cartesian product** $A \times B = \{(a, b) : a \in A, b \in B\}$.

Proposition
$A \setminus B = A \cap B^c$.

Proof. Notice that $x \in A \setminus B \Leftrightarrow x \in A, x \notin B \Leftrightarrow x \in A, x \in B^c \Leftrightarrow x \in A \cup B^c$. □

Note that this makes sense intuitively, as the set minus operation is all elements $a \in A \notin B$, which would be $A \cap B^c$.

Theorem 2.1.1 (DeMorgan's Laws)
DeMorgan's Laws are the following:

1. $(A \cap B)^c = A^c \cup B^c$.
2. $(A \cup B)^c = A^c \cap B^c$.

Proof. Suppose $x \in (A \cap B)^c$. Then $x \notin A \cap B$. Since x cannot be in both A and B, suppose WLOG that $x \notin A$. Hence $x \in A^c \subseteq A^c \cup B^c$. Now suppose $x \in A^c \cup B^c$. THen $x \in A^c$ or $x \in B^c$, so suppose WLOG that $x \in A^c$. Hence $x \notin A$, so $x \notin A \cap B$. Thus $x \in (A \cap B)^c$.

Notice that $x \in (A \cup B)^c \Leftrightarrow x \notin (A \cup B) \Leftrightarrow x \notin A$ and $x \notin B \Leftrightarrow x \in A^c$ and $x \in B^c \Leftrightarrow x \in A^c \cap B^c$. □

Definition 2.1.2: A **function** $f : S \longrightarrow T$ is a map of each element $s \in S$ to *at most* one element $t \in T$. f can also be defined as a subset $F \subseteq S \times T$, such that each $s \in S$ occurs in at most one ordered pair in F. A function f is **onto** or **surjective** if $\operatorname{ran} f = T$. f is **one-to-one** or **injective** if $f(s_1) = f(s_2)$ implies $s_1 = s_2$. f is a **bijection** from S to T if it is both injective and surjective. This also implies f has an inverse.

✾ 2.2 Bounded Sets

Definition 2.2.1: A set S is **bounded above** if there exists a number $b \in \mathbb{R}$ such that $s \leq b$ for all $s \in S$. Similarly, S is **bounded below** if there is a number $\ell \in \mathbb{R}$ such that $s \geq \ell$ for all $s \in S$.

A number $b_0 \in \mathbb{R}$ is the **supremum** of S if it is the *least* upper bound of S. It is denoted as $b_0 = \sup S$. Similarly, a number ℓ_0 is the **infimum** of S if it is the *greatest* lower boundof S. It is denoted as $\ell_0 = \inf S$. Note that if S contains its supremum, then we say that it is also the **maximum** of S. Similarly if S contains its infimum, it is also the **minimum** of S.

Theorem 2.2.2 (Supremum Property)
Every nonempty set of real numbers that is bounded above has a least upper bound.

We take this as an axiom, but we can prove it later.

Proposition 2.2.3
Assume $\mu \in \mathbb{R}$ is an upper bound for a set $S \subset \mathbb{R}$. Then $\mu = \sup S \Leftrightarrow$ for every $\varepsilon > 0$, there exists an element $s \in S$ such that $s \in [\mu - \varepsilon, \mu]$.

Proof. hi □

<++> Consequences of the supremum property are the following:

Theorem 2.2.4 (Archimedean Property)
For any $b > 0$, there are positive integers $n, m \in \mathbb{B}$ satisfying $\frac{1}{m} < b < n$.

Theorem 2.2.5 (Nested Interval Property)
For each $n \in \mathbb{N}$, assume we have a closed interval $I_n = [a_n, b_n]$. Assume these intervals are *nested*: $I_{n+1} \subseteq I_n$. Then

$$\bigcap_{n=1}^{\infty} I_n \neq \emptyset.$$

✾ 2.3 Cardinality, Countability, and Density

We can use bijections to compare the cardinality of sets.

Definition 2.3.1: Two sets have the same **cardinality** if there is a bijection $f : A \longrightarrow B$ $(A \sim B)$.

Note that sets are either finite or infinite. But infinite sets do not necessarily have the same cardinality.

Definition 2.3.2: A set A is **countable** if $\mathbb{N} \sim A$. If A is an infinite set but $\mathbb{N} \nsim A$, then A is **uncountable**.

We can establish some properties of countable sets.

Theorem 2.3.3
If $S \subset T$ is infinite and T is countable, then S is also countable.

Theorem 2.3.4
If $A_1, A_2, \cdots, A_n$ are each countable sets, then $\bigcup_{k=1}^{n} A_k$ is a countable set. If A_n is a countable set for each $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n$ is countable.

Theorem 2.3.5
$\mathbb{Q}$ is countable.

Theorem 2.3.6 (Cantor)
The open interval $(0, 1)$ is uncountable.

Cantor's Diagonalization Argument. Suppose there exists a bijection from $\mathbb{N} \longrightarrow (0, 1)$. Then $1 \longrightarrow f(1) = 0.a_{11}a_{12}\cdots$, and so forth for all $n \in \mathbb{N}$. Now define $x \in (0, 1) = 0.b_1b_2b_3\cdots$, where

$$b_n = \begin{cases} 2 \text{ if } a_{nn} \neq 2 \\ 3 \text{ if } a_{nn} = 2 \end{cases}$$

Since no such $x = f(n)$ for any $n \in \mathbb{N}$, $(0, 1)$ must be uncountable. $\square$

An immediate corollary of this is that $\mathbb{R}$ must be uncountable, as $(0, 1) \subset \mathbb{R}$. Note that $\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}^c$. Since $\mathbb{Q}$ is countable, $\mathbb{Q}^c$ must be uncountable.

> **Theorem 2.3.7** (Density of $\mathbb{Q}$ in $\mathbb{R}$)
> For every two real numbers a and b with $a < b$, there is a rational number r such that $a < r < b$.

2.4 Open and Closed Sets (Topology in $\mathbb{R}$)

Definition 2.4.1: Given $a \in \mathbb{R}$ and $\varepsilon > 0$, the **ε-neighborhood** of a is the set $V_\varepsilon(a) = \{x \in \mathbb{R} : |x - a| < \varepsilon\}$.

Definition 2.4.2: A set $O \subset \mathbb{R}$ is **open** if for all points $p \in O$, there exists an ε-neighborhood $V_\varepsilon(p) \subset O$.

Proposition 2.4.3

If A and B are open, then $A \cup B$ is open.

Proposition 2.4.4

If A and B are open, then $A \cap B$ is open.

Theorem 2.4.5

If $A_1, A_2, \cdots$ are open, then

$$\bigcup_{\lambda \in S} A_\lambda$$

is open. The indexing set S need not be countable.

Theorem 2.4.6

If $A_1, \cdots, A_N$ are open, then

$$\bigcap_{n=1}^{n} A_n$$

is open.

Note that the above theorem does *not* work when there are an infinite number of sets involved.

Definition 2.4.7: A set $F \subset \mathbb{R}$ is **closed** if its complement F^c is open.

Theorem 2.4.8

The union of a finite collection of closed sets is closed, and the intersection of an arbitraty collection of closed sets is closed.

The above theorem can be proved using DeMorgan's laws to shift the expressions to **Theorem 1.4.5** and **Theorem 1.4.6**'s statements.

Definition 2.4.9: A point $x \in \mathbb{R}$ is a **limit point** of A if every ε-neighborhood of x intersects the set A at some point other than x.

Theorem 2.4.10

A set $F \subset \mathbb{R}$ is closed iff it contains its limit points.

Remark. Note that $\mathbb{R}$ and $\emptyset$ are both open and closed at the same time. These are the only two such sets in $\mathbb{R}$ that have this property.

2.5 The Cantor Set

Consider the following sets:

1. $C_1 = [0, 1]$.
2. $C_2 = C_1 \setminus (\frac{1}{3}, \frac{2}{3})$.
3. $C_3 = C_2 \setminus [(\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9})]$.
4. So on and so forth for all $n \in \mathbb{N}$.

Definition 2.5.1: The **Cantor set** is defined as $C = \bigcap_{n=1}^{\infty} C_n$.

Each closed interval in C_n has length $(\frac{1}{3})^{n-1}$, so the total length of C_n is $2^{n-1}(\frac{1}{3})^{n-1}$. Note that the length of C is 0.

Definition 2.5.2: A set $A \subset \mathbb{R}$ is **perfect** if it is closed and contains no isolated points (all points are limit points). The Cantor set is one example of a perfect set.

The Cantor set is not empty, but contains an uncountable amount of points.

Chapter 3

Sequences and Limits

3.1 Sequences and Convergence

Definition 3.1.1: A **sequence** is a function whose domain is $\mathbb{N}$.

Definition 3.1.2: A sequence $\{a_n\}$ **converges** to $a \in \mathbb{R}$ if for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that whenever $n > \mathbb{N}$, if follows that $|a_n - a| \leq \varepsilon$. We write $\lim a_n = a$, or simply $a_n \to a$.

The proof of a limit like this is very straightforward and formulaic.

Example 3.1.3: Prove that $\lim \frac{1}{n} = 0$.

Proof. Let $\varepsilon > 0$, and choose $N \geq \frac{1}{\varepsilon}$. Then if $n \geq N$, we have $|\frac{1}{n} - 0| = |\frac{1}{n}| \leq \frac{1}{N} \leq \varepsilon$. Thus $\lim \frac{1}{n} - 0$. $\square$

When computing and proving a limit, we can drop terms in the inequality $|a_N - a| \leq \varepsilon$, because the *inequality* must hold when calculating N. If we can find an expression greater than $|a_N - a|$ and prove it is less than ε for some choice of N, then we can still proceed. This can be done by dropping certain terms, for example.

Definition 3.1.4: A **divergent** sequence is one that does not converge.

Definition 3.1.5: A sequence $\{b_n\}$ is bounded if there exists an $M > 0$ such that $|b_n| \leq M$ for all $n \in \mathbb{N}$.

Theorem 3.1.6
Every convergent sequence is bounded.

3.2 Limit Theorems

Theorem 3.2.1 (Squeeze Theorem)
Let $\{a_n\}$ and $\{b_n\}$ both converge to an $L \in \mathbb{R}$. If $\{c_n\}$ is a sequence satisfying $\{a_n\} \leq \{c_n\} \leq \{b_n\}$ for all $n \in \mathbb{N}$, then $c_n \to L$.

Theorem 3.2.2 (Algebraic Limit Theorem)
Suppose $a_n \to a$, $b_n \to b$, and $\kappa \in \mathbb{R}$. Then the following hold:

1. $\lim(a_n + b_n) = a + b$.
2. $\lim \kappa a_n = \kappa a$.
3. $\lim(a_n b_n) = ab$.

4. If $b, b_n \neq 0$, then $\lim \frac{a_n}{b_n} = \frac{a}{b}$.

Theorem 3.2.3 (Order Limit Theorem)

Suppose $a_n \to a$ and $b_n \to b$. Then the following hold:

1. If $a_n \geq 0$ for all $n \in \mathbb{N}$, then $a \geq 0$.
2. If $a_n \leq b_n$ for all $n \in \mathbb{N}$, then $a \leq b$.
3. If there exists a $c \in \mathbb{R}$ for which $c \leq b_n$ for all $n \in \mathbb{N}$, then $c \leq b$. Similarly, if $a_n \leq c$ for all $n \in \mathbb{N}$, then $a \leq c$.

Definition 3.2.4: let $\{a_n\}$ be a sequence. Define $\overline{s}_N = \sup\{a_n | n \geq N\}$ and $\underline{s}_N = \inf\{a_n | n \geq N\}$. The **limit superior** of $\{a_n\}$ is

$$\limsup a_n = \lim_{N\to\infty} \overline{s}_N,$$

and the **limit inferior** of $\{a_n\}$ is

$$\liminf a_n = \lim_{N\to\infty} \underline{s}_N.$$

Example 3.2.5: Let $a_n = 1 + \frac{1}{n}$. Find $\limsup a_n$ and $\liminf a_n$.

Solution.

$$\limsup a_n = \lim_{N\to\infty} \sup\{a_n | n \geq N\} = \lim_{n\to\infty} \sup\left\{1 + \frac{1}{N}, 1 + \frac{1}{N+1}, \cdots\right\} = \lim_{N\to\infty} (1 + \frac{1}{N}) = \boxed{1}.$$

The infimum of the above set is 1, so $\liminf a_n = \boxed{1}$. ☐

Theorem 3.2.6

Let a_n be a sequence.

1. If $\limsup a_n$ is finite, then for any $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ so that $a_n \leq \varepsilon + \limsup a_n$ for all $n \geq N$.
2. If $\liminf a_n$ is finite, then for any $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ so that $a_n \geq \varepsilon + \liminf a_n$ for all $n \geq N$.

Theorem 3.2.7

A sequence a_n converges to $a \in \mathbb{R}$ iff $\limsup a_n = a = \liminf a_n$.

✾ 3.3 Cauchy Sequences

Definition 3.3.1: A sequence a_n is a **Cauchy sequence** if, given any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ so that $|a_n - a_m| \leq \varepsilon$ if $m \geq n \geq N$.

Proposition

If a_n is Cauchy, then a_n is bounded.

Theorem 3.3.2 (Cauchy Criterion)
Every convergence sequence a_n in $\mathbb{R}$ is Cauchy. (and vice versa)

The forwards direction of this proof is easy. The backwards direction of this proof is more difficult, because the definition of a Cauchy sequence does not contain a limit. In order to prove this, we must introduce subseqences:

Definition 3.3.3: Let a_n be a sequence of real numbers. Let $K \to n(k)$ be a not onto function from $\mathbb{N} \longrightarrow \mathbb{N}$ having the property that $n(k+1) > n(k)$ for all $k \in \mathbb{N}$. Then $\{a_{n(k)}\}_{k=1}^{\infty}$ is called a **subsequence** of a_n.

Remark 3.3.4. A subsequence must be a sequence $\sim \mathbb{N}$. Note that subsequences must preserve order ($n(k+1) > n(k)$)

Theorem 3.3.5 (Monotone Subsequence Theorem)
Every sequence has a subsequence which is monotone

Cauchy implying convergence does not happen in every domain.

Definition 3.3.6: Met M be a metric space. M is **complete** if every Cauchy sequence in M converges to a point $p \in M$.

$\mathbb{R}$ is complete, and this is called the Axiom of Completeness.

Theorem 3.3.7
A set $E \subset \mathbb{R}$ is closed if every Cauchy sequence in E converges to a point $p \in E$.

3.4 The Bolzano-Weierstrass Theorem

Theorem 3.4.1 (Monotone Convergence Theorem)
Every bounded monotone sequence converges.

Theorem 3.4.2
Let $a_n \to a$. THen every subsequence of a_n also converges to a.

We now arrive at the **Bolzano Weierstrass** theorem.

Theorem 3.4.3 (Bolzano Weierstrass)
Every bounded sequence has a convergent subsequence.

Chapter 4

Continuity

4.1 Continuous Functions

Definition 4.1.1: A function f is said to be **continuous** at $c \in \operatorname{dom} f$ if for every sequence x_n of points in $\operatorname{dom} f$ which converge to c we have

$$\lim f(x_n) = f(c).$$

We say f is continuous if it is continuous at every point $c \in \operatorname{dom} f$.

Example 4.1.2: Let's show that $f : \mathbb{R} \to \mathbb{R}, f(x) = 2x$ is continous on its domain. Choose any $c \in \mathbb{R}$, and suppose that $\{x_n\} \to c$. THen using the Algebraic Limit Theorem,

$$\lim f(x_n) = \lim 2x_n = 2 \lim x_n = 2c = f(c).$$

Thus f is continuous on $\mathbb{R}$.

Theorem 4.1.3

Let f and g be continuous functions, and define $D = \operatorname{dom} f \cap \operatorname{dom} g$. Then,

1. $f(x) + g(x)$ is continuous on D.
2. $f(x)$ is continuous on $\operatorname{dom} f$ for any $k \in \mathbb{R}$.
3. $f(x)g(x)$ is continuous on D.
4. $\dfrac{f(x)}{g(x)}$ is continuous at all $x \in D$ such that $g(x) \neq 0$.

Proposition 4.1.4 (Sequential Criterion for Discontinuity)

Let f be a function defined on A, and let $c \in A$. If there exist two sequences x_n and y_n in A with $x_n \neq c$ and $y_n \neq c$, where $\lim x_n = c = \lim y_n$ but $\lim f(x_n) \neq \lim f(y_n)$, then f is not continuous at c.

Theorem 4.1.5 (δ-ε Convergence)

A function $f(x)$ is continuous at a point $c \in \operatorname{dom} f$ iff for each $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x \in \operatorname{dom} f$, $|x - c| \leq \delta \Rightarrow |f(x) - f(c)| \leq \varepsilon$.

Let's look at an example of a proof using this:

Example 4.1.6: Prove $f(x) = 3x - 7$ is continuous on $\mathbb{R}$.

Proof. Let $\varepsilon > 0$, and choose $\delta > 0$ such that $\delta \leq \frac{\varepsilon}{3}$. Then if $|x - c| \leq \delta$, we find $|f(x) - f(c)| = |3x - 7 - 3c + 7| = 3|x - c| \leq 3\delta \leq \varepsilon$. Hence f is continuous for all $c \in \mathbb{R}$, so f is continuous on $\mathbb{R}$. □

Theorem 4.1.7 (Composition of Continuous Functions)
Let $F : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$, and assume the range of f is contained in B so that the function $(g \circ f)(x) : A \to \mathbb{R}$ is well defined. Then if f is continuous at $c \in A$, and if g is continuous at $f(c) \in B$, then $g \circ f$ is continuous at c.

Continuity at $c \in \operatorname{dom} f$ is the same thing as

$$\lim_{x \to c} f(x) = f(c),$$

where the limit is the familiar **functional limit** and $x \in \operatorname{dom} f$.

4.2 Continuity on Closed Bounded Intervals

Definition 4.2.1: A real valued function f is said to be **bounded** on $S \subset \operatorname{dom} f$ if there exists a real number $B > 0$ such that $|f(x)| \leq B$ for all $x \in S$.

Theorem 4.2.2
If f is continuous on $[a, b]$, then $f([a, b]) = \{f(x) : a \leq x \leq b\}$ is bounded.

Definition 4.2.3: $\sup_S f = \sup \{f(x) | x \in S\}$. $\inf_S f = \inf \{f(x) | x \in S\}$.

Theorem 4.2.4 (Extreme Value Theorem)
If f is continuous on $[a, b]$, then f obtains its supremum (maximum) and infimum (minimum); There exists $c, d \in [a, b]$ such that $f(c) = \sup_{[a,b]} f$ and $f(d) = \inf_{[a,b]} f$.

Theorem 4.2.5 (Intermediate Value Theorem)
If f is continuous on $[a, b]$ and $f(a) \neq f(b)$, then for any y between $f(a)$ and $f(b)$ there exists $c \in (a, b)$ such that $f(c) = y$.

Definition 4.2.6: A continuous function f is **uniformly continuous** on $S \subset \operatorname{dom} f$ if for each $\varepsilon > 0$ there is a $\delta > 0$ such that for all $x_1, x_2 \in S$, $|x_1 - x_2| \leq \delta \Rightarrow |f(x_1) - f(x_2)| \leq \varepsilon$.

Theorem 4.2.7
If f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.

Definition 4.2.8: A function f is said to be **Lipschitz continuous** on $S \subset \operatorname{dom} f$ if there is an $M > 0$ such that for all $x_1 \neq x_2 \in S$,

$$\frac{|f(x_1) - f(x_2)}{|x_1 - x_2|} \leq M.$$

Theorem 4.2.9
If f is a Lipschitz continuous function on S, then f is uniformly continuous on S.

Chapter 5

The Riemann Integral

5.1 Riemann Integration

Definition 5.1.1: A **partition** P of $[a, b]$ is a finite ordered set

$$P = \{a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b\}.$$

Suppose f is a not necessarily continuous bounded function on $[a, b]$. Given a partiton P, we can consider the ith subinterval $[x_{i-1}, x_i]$. Since f is bounded, the supremum property implies we can define

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\},$$
$$M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}.$$

Then we can define the following:

Definition 5.1.2: The **lower sum** with respect to P is

$$L_P(f) = \sum_{i=1}^{N} m_i(x_i - x_{i-1}).$$

Similarly, the **upper sum** with respect to P is

$$U_P(f) = \sum_{i=1}^{N} M_i(x_i - x_{i-1}).$$

Note that we have $U_P(f) - L_P(f) = \sum_{i=1}^{N}(M_i - m_i)(x_i - x_{i-1}) \geq 0$, as $M_i - m_i \geq 0$ and the difference between the x_i is positive. Then we can conclude that $U_P(f) \geq L_P(f)$.

Definition 5.1.3: A partition Q is called a **refinement** of a partition P if Q contains all the points of P, or $P \subseteq Q$.

Lemma

If $P \subseteq Q$, then $L_P(f) \leq L_Q(f)$, and $U_Q(f) \leq U_P(f)$.

Proof. Since P and Q are finite, we can assume that Q adds just a single new point to P. Suppose $[x_{k-1}, x_k]$ of P is refined to $[x_{k-1}, \overline{x}]$ and $[\overline{x}, x_k]$. Let $m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$. Ten we have $L_Q(f)$ being the sum of the $m_i \Delta x_i$ and the infimi of $[x_{k-1}, \overline{x}]$ and $[\overline{x}, x_k]$, which is greater than $L_P(f)$, which is the same sum but $[x_{k-1}, x_k]$ instead. □

Lemma

If P and Q are *any* two partitions of $[a, b]$, then $L_P(f) \leq U_P(f)$.

Proof. Let $P' = P \cup Q$. Then

$$L_P(f) \leq L_R(f) \leq U_R(f) \leq U_Q(f).$$

□

Definition 5.1.4: Let $\mathcal{P}$ be the collection of all possible partitions of $[a, b]$. Then the **upper integral** of f is defined as

$$U(f) = \inf_{\mathcal{P}}\{U_P(f) : P \in \mathcal{P}\}.$$

Similarly, the **lower integral** of f is defined as

$$L(f) = \sup_{\mathcal{P}}\{L_P(f) : P \in \mathcal{P}\}.$$

Proposition

For any bounded function $f : [a, b] \to \mathbb{R}$, $U(f) \geq L(f)$.

Proof. Suppose $U(f) < L(f)$. Then by the characterization of the supremum, there exists some P and Q such that $L_P(f) \in [L(f) - \frac{L(f)-U(f)}{3}, L(f)]$, and $U_Q(f) \in [U(f), U(f) + \frac{L(f)-U(f)}{3}]$. This implies $U_Q(f) < L_P(f)$, a contradiction. □

Definition 5.1.5: A bounded function $f : [a, b]$ is **Riemann integrable** if $U(f) = L(f)$. Then we define

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

Lemma 5.1.6 (Riemann Integrability Criterion)

A bounded function $f : [a, b]$ is integrable iff for every $\varepsilon > 0$ there exists a partition P_ε such that

$$|U_{P_\varepsilon}(f) - L_{P_\varepsilon}(f)| \leq \varepsilon.$$

Proof. Let $\varepsilon > 0$, and choose P_ε accordingly. Note that

$$|U(f) - L(f)| = U(f) - L(f) \leq U_{P_\varepsilon}(f) - L(f) \leq U_{P_\varepsilon}(f) - L_{P_\varepsilon} \leq \varepsilon.$$

Since ε was arbitrary, $U(f) = L(f)$.

Now suppose f is Riemann integrable, and let $\varepsilon > 0$. By the characterization of the supremum, there exists a P_1 such that $U_{P_1}(f) \in [U(f), U(f) + \frac{\varepsilon}{2}]$. Similarly there exists a P_2 such that $L_{P_2}(f) \in [L(f) - \frac{\varepsilon}{2}, L(f)]$. Let $P_\varepsilon = P_1 \cup P_2$. Then

$$U_{P_\varepsilon}(f) - L_{P_\varepsilon}(f) \leq U(f) + \frac{\varepsilon}{2} - (L(f) - \frac{\varepsilon}{2}) = \varepsilon,$$

as desired. □

Next we show that continuity on an interval implies integrability on that interval.

Theorem 5.1.7

If f is continuous on $[a, b]$, then f is Riemann integrable on $[a, b]$.

Proof. Since f is conitnuous and $[a, b]$ is a closed, bounded interval, f is uniformly continuous on $[a, b]$. Let $\varepsilon > 0$, and choose $\delta > 0$ such that $\forall x, y \in [a, b]$ with $|x - y| \leq \delta$,

$$|f(x) - f(y)| \leq \frac{\varepsilon}{b - a}.$$

Let P_ε be any partition of $[a, b]$ for which $\Delta x_i \leq \delta$ for all i. For each $[x_{i-1}, x_i]$, f must attain its maximum, M_i, and its minimum, m_i. By uniform continuity, we have $|M_i - m_i| \leq \frac{\varepsilon}{b-a}$. Then we have

$$U_{P_\varepsilon}(f) - L_{P-\varepsilon}(f) = \sum (M_i - m_i)\Delta x_i \leq \sum \frac{\varepsilon}{b - a}\Delta x_i = \frac{\varepsilon}{b - a}(b - a) = \varepsilon.$$

Thus f is Riemann integrable. □

Definition 5.1.8: Let's define some properties of the integral, assuming $f, g : [a, b] \to \mathbb{R}$ are integrable functions on $[a, b]$, and let $\alpha, \beta \in \mathbb{R}$.

1. **Linearity**: The function $\alpha f + \beta g$ is integrable, and
$$\int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x + \beta \int_a^b g(x)\,\mathrm{d}x.$$

2. **Monotonicity**: If $f(x) \leq g(x)$ for all $x \in [a, b]$ then
$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x.$$

3. The function $|f(x)|$ is integrable, and
$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x.$$

4. If $m \leq f(x) \leq M$ for all $x \in [a, b]$, then
$$m(b-a) \leq \int_a^b f(x)\,\mathrm{d}x \leq M(b-a).$$

This last theorem is pretty useful.

Theorem 5.1.9

Assume $f : [a, b] \to \mathbb{R}$ is bounded, and let $c \in (a, b)$. Then f is integrable on $[a, b]$ iff f is integrable on $[a, c]$ and $[c, b]$. Then we have
$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

The proof is kind of long and I'm lazy so rip.

5.2 Riemann-Stieltjes Integrals

Definition 5.2.1: Let α be a monotonically increasing function on $[a, b]$. Since $\alpha(a)$ and $\alpha(b)$ are finite, α is bounded on $[a, b]$. For each partition P of $[a, b]$ we write
$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$
For any real f bounded on $[a, b]$ we define
$$U_P(f, \alpha) = \sum_{i=1}^n M_i \Delta\alpha_i,$$
$$L_P(f, \alpha) = \sum_{i=1}^n m_i \Delta\alpha_i$$
as the upper and lower sums with respect to P. We define the upper and lower integral similarly, as the infimum and supremum of the upper and lower sums, respectively. When these are equal we

denote them by

$$\int_a^b f \, d\alpha = \int_a^b f(x) \, d\alpha(x).$$

This is known as the **Riemann-Stieltjes integral**, or simply the **Stieltjes integral** of f with respect to α. The Riemann integral is a special case of this integral, with $\alpha(x) = x$. But α need not even be continuous!

Sometimes we define the space of integrable functions as $\mathcal{R}$, and if f is integrable we say that $f \in \mathcal{R}$. Let's investigate the integrability of functions with regards to the Stieltjes integral:

Theorem 5.2.2

If f is monotonic on $[a, b]$, and if α is continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$. (In other words, f is Sjteltjes integrable).

Proof. Let $\varepsilon > 0$. For any $n \in \mathbb{Z}^+$, choose a partition

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n}$$

for $i = 1, \cdots, n$. It is always possible to choose such a partition as α is continuous on $[a, b]$. Assume, WLOG that f is monotonically increasing. Then if $M_i = f(x_i)$ and $m_i = f(x_{i-1})$, we have

$$U_P(f, \alpha) - L_P(f, \alpha) = \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} f(x_i) - f(x_{i-1}) = \frac{\alpha(b) - \alpha(a)}{n} f(b) - f(a) < \varepsilon$$

for large n, as desired. By the integrability criterion, $f \in \mathcal{R}$. □

5.3 Change of Variable

Theorem 5.3.1 (Change of Variable)

Let φ be a strictly increasing continuous function that maps $[A, B] \to [a, b]$. Let α be monotonically increasing on $[a, b]$ and $f \in \mathcal{R}(\alpha)$ on $[a, b]$. Define β and g on $[A, B]$ by $\beta(y) = \alpha(\varphi(y))$, and $g(y) = f(\varphi(y))$. Then $g \in \mathcal{R}(\beta)$ and

$$\int_A^B g \, d\beta = \int_a^b f \, d\alpha.$$

<++>

5.4 Integration With Discontinuities

Just because a function is discontinuous on an interval $[a, b]$ does *not* mean that is not Riemann integrable on that interval.

Theorem 5.4.1

If $f : [a, b] \to \mathbb{R}$ is monotone, then f is Riemann integrable.

Proof. Without loss of generality, suppose f is monotone increasing. If f is constant, it is continuous and thus integrable, so suppose $f(b) > f(a)$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that $\delta = \frac{\varepsilon}{f(b)-f(a)}$. Choose any partition P_ε of $[a, b]$ so that $\Delta x_i \leq \delta \ \forall 1 \leq i \leq n$. Then

$$U_{P_\varepsilon}(f) - L_{P_\varepsilon}(f) = \sum (M_i - m_i)\delta x_i \leq \sum (f(x_i) - f(x_{i-1}))\delta =$$

$$\delta(f(x_1) - f(a) + f(x_2) - f(x_1) + \cdots + f(b) - f(x_{n-1})) = \frac{\varepsilon}{f(b) - f(a)}(f(b) - f(a)) = \varepsilon.$$

□

This next theorem shows that a finitely discontinuous function on $[a, b]$ is still Riemann integrable.

Theorem 5.4.2

Any function $f : [a, b]$ with a finite number of discontinuities is integrable.

Proof. Let $\varepsilon > 0$, and choose $M > 0$ as a bound for f. We have two cases:

- Case f has a discontinuity at an endpoint: WLOG, assume f is discontinuous at a. Then f is continuous on $[a + \frac{\varepsilon}{4M}, b]$, so there exists a partition P such that $U_P(f) - L_P(f) \leq \frac{\varepsilon}{2}$. Let $P_\varepsilon = \{a\} \cup P$, and suppose $|P_\varepsilon| = n + 1$. Then we have

$$U_{P_\varepsilon}(f) - L_{P_\varepsilon}(f) = \sum (M_i - m_i)\Delta x_i = (M_1 - m_1)\Delta x_1 + \sum_2^n (M_i - m_i)\Delta x_i \leq 2M\left(a + \frac{\varepsilon}{4M} - a\right) + \frac{\varepsilon}{2},$$

which equals ε. Thus f is integrable on $[a, b]$. The proof is similar for discontinuity at b.

- f has discontinuity at some $c \in [a, b]$: If f is discontinuous at c, then it is integrable on $[a, c]$ and $[b, c]$, and is thus integrable on $[a, b]$.

Assume f is integrable if it has $k \geq 2$ points of discontinuity in $[a, b]$, denoted as $\{y_1, \cdots, y_n\}$. Consider f on $[a, \frac{y_1+y_2}{2}]$ and $[\frac{y_1+y_2}{2}, b]$. On these intervals, f has 1 and $k - 1$ points of discontinuity. Thus we can always create $P_\varepsilon = P_1 \cup P_2$, and keep on going. Thus f is integrable on $[a, b]$, and the result follows from induction. $\square$

Theorem 5.4.3

If $f : [a, b] \to \mathbb{R}$ is bounded, and f is integrable on $[c, b]$ for all $c \in (a, b)$, then f is integrable on $[a, b]$.

Let's look at an interesting example.

Example 5.4.4 (Thomae's Function)**:** Let's see if this function is integrable or not.
Let $f(x) : [0, 1] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \text{ or } 1 \\ \frac{1}{q} & \text{if } x \in \mathbb{Q} \cap (0, 1) \text{ with } x = \frac{p}{q} \text{ in reduced form} \\ 0 & \text{otherwise.} \end{cases}$$

This function is discontinuous at every rational, but continuous at every irrational, and is still Riemann integrable. If the set of discontinuities has **measure zero**, f is integrable. A simple definition of this is that if a set X is finite, or countable, then it has measure zero (which fits with our earlier theorem that the set of discontinuities must be *finite*).

5.5 Improper Integrals

Definition 5.5.1: Suppose that f is a continuous function on $(a, b]$. If the following limit exists, then

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{\sigma \searrow 0} \int_{a+\sigma}^b f(x)\,\mathrm{d}x$$

is called the **improper Riemann integral** of f on $[a, b]$.

An important technique that we can use is **comparison**. Recall that if a limit $L \geq 0$ is less than a limit L' on $[a, b]$, then if L' converges, so does L. By finding a function with an improper integral on $[a, b]$ that is greater than the function f in question, we can state that f has an improper integral on $[a, b]$ provided that $f(x) \geq 0\ \forall x \in [a, b]$.

Proposition 5.5.2

Suppose $f(x) = x^n$. Then

$$\int_0^1 f(x)\, \mathrm{d}x$$

exists provided that $n > -1$.

Proof. We can easily integrate

$$\int_\sigma^1 x^n\, \mathrm{d}x = \frac{x^{n+1}}{n+1}|_\sigma^1.$$

From this we know that

$$\lim_{\sigma \searrow 0} \int_\sigma^1 f(x)\, \mathrm{d}x$$

diverges if $n \leq -1$. □

Remark. When using comparison, note that it still holds under u and trig substitution, as well as integration by parts.

Definition 5.5.3: Suppose that f is a continuous function on $[a, \infty)$. If the following limit exists, then

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

is called the **improper Riemann integral** of f on $[a, \infty)$.

For polynomial f, the opposite of the result for $[0, 1]$ holds here: f has an improper integral on $[a, \infty)$ if $n < -1$.

Let's look at an example.

Example 5.5.4: Does

$$\int_0^1 \frac{\cos x}{\sqrt{1-x}}$$

have an improper Riemann integral?

Proof. Note that we have a u substitution, namely $u = 1 - x$. Then $\mathrm{d}u = \mathrm{d}x$ and we have

$$\int_1^0 \frac{\cos 1 - u}{\sqrt{u}}(-\,\mathrm{d}u) = \int_0^1 \frac{\cos 1 - u}{\sqrt{u}}\, \mathrm{d}u.$$

We have $0 \leq \dfrac{\cos 1 - u}{\sqrt{u}} \leq \dfrac{1}{\sqrt{u}}$, which we know has a convergent limit, so the improper integral exists. □

Chapter 6

Derivatives

6.1 Differentiable Function

Definition 6.1.1: A function f is said to be **differentiable** at a point $x \in \operatorname{dom} f$ if the limit of the difference quotient

$$\frac{f(x+h)-f(x)}{h}$$

exists as $h \to 0$. Then this limit is called the **derivative** of f at x, denoted by

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \frac{\mathrm{d}f}{\mathrm{d}x}.$$

Theorem 6.1.2

If $g : A \to \mathbb{R}$ is differentiable at $c \in A$, then g is continuous at c.

Proof. If $f'(c)$ exists, then

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

Then we have

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

Thus $\lim\limits_{h\to 0} f(c+h) = f(c)$, so f is continuous at c. □

Theorem 6.1.3 (Combinations of Differentiable Functions)

Let $f, g : A \to \mathbb{R}$ and assume both are differentiable at a point $c \in A$. Let $k \in \mathbb{R}$. Then we have:

1. $(f+g)'(c) = f'(c) + g'(c)$
2. $(kf)'(c) = kf'(c)$
3. $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$
4. $\left(\frac{f}{g}\right)'(c) = \frac{g(c)f'(c) - g'(c)f(c)}{|g(c)|^2}$

One last property of derivatives:

Theorem 6.1.4 (Chain Rule)

Let $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ satisfy $f(A) \subset B$ so that composition $g \circ f$ is well-defined. If f is differentiable at $c \in A$ and g is differentiable at $f(c) \in B$, then $g \circ f$ is differentiable at c with

$$(g \circ f)'(x) = g'(f(c)) \cdot f'(c).$$

All of these should be familiar from Calculus.

> **Remark.** Another way to compute the derivative at $x = c$ is
> $$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}.$$

✻ 6.2 The Mean Value Theorem

We state a few theorems first:

> **Theorem 6.2.1** (Interior Extremum Theorem)
> Suppose that f is continuous on (a, b), differentiable at some $c \in (a, b)$, and attains a local minimum or maximum at c. Then $f'(c) = 0$.

Proof. We present the proof for the maximum case. Since $c \in (a, b)$, we can construct a sequence $\{x_n\} \to c$ such that $x_n < c$ for all $n \in \mathbb{N}$. Then

$$f'(c) = \lim_{n \to \infty} \frac{f(x_n) - f(c)}{x_n - c},$$

where the numerator must be less than or equal to 0 and the denominator must be less than 0. Similarly, we can construct a sequence $\{y_n\} \to c$ so that $y_n > c$ for all $n \in \mathbb{N}$, where the numerator will be less than or equal to zero and the denominator must be greater than 0:

$$f'(c) = \lim_{n \to \infty} \frac{f(y_n) - f(c)}{y_n - c}.$$

Since the above two limits imply that $f'(c) \geq 0$ and $f'(c) \leq 0$ respectively, $f'(c) = 0$. $\square$

> **Theorem 6.2.2** (Darboux's Theorem)
> If f is differentiable on an interval $[a, b]$ and there is som α such that $f'(a) < \alpha < f'(b)$, then there exists some $c \in (a, b)$ such that $f'(c) = \alpha$.

Proof. Let $g(x) = f(x) - \alpha x$. Since $[a, b]$ is closed and bounded, and g is continuous, g attains its extrema. Note that $g'(a) = f'(a) - \alpha < 0$ and $g'(b) = f'(b) - \alpha > 0$, g attains an interior minimum at $x = c$. Then $g'(c) = 0 \longrightarrow f'(c) - \alpha = 0 \longrightarrow f'(c) = \alpha$, as desired. $\square$

> **Theorem 6.2.3** (Rolle's Theorem)
> Assume f is continuous on $[a, b]$, differentiable on (a, b), and $f(a) = f(b)$. Then there exists some $c \in (a, b)$ such that $f'(c) = 0$.

Proof. Suppose $f(x) = k$ for all $x \in [a, b]$, then for all $c \in (a, b)$,

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

If $f(x)$ is a nonconstant function, then wlog there exists a point $x \in (a, b)$ with $f(x) > f(a)$. Then since f is continuous on $[a, b]$., f must attain a maximum value, which cannot happen at $x = a, b$. Then f must achieve an interior maximum at some $c \in (a, b)$. By the Interior Extremum Theorem, $f'(c) = 0$. $\square$

Theorem 6.2.4 (Mean Value Theorem)
If f is continuous on $[a, b]$, and differentiable on (a, b), there exists a $c \in (a, b)$ such that

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

Proof. Note that $\frac{f(b)-f(a)}{b-a}$ computes the slope of the secant line through points a and b. If $f(a) = f(b)$, then the result immediately follows from Rolle's theorem. Suppose $f(a) \neq f(b)$. Construct $d(x)$ such that

$$d(x) = f(x) - \left(\frac{f(b) - f(a)}{b - a}(x - a) + f(a)\right).$$

Note $d(a) = f(a) - \frac{f(b)-f(a)}{b-a}(a - a) - f(a) = 0$, and similarly $d(b) = 0$. Since $d(x)$ is a differentiable function, there exists a $c \in (a, b)$ such that $d'(c) = 0$ by Rolle's Theorem. Since we have

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

we have

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

□

This is actually a special case of the **General Mean Value Theorem**:

Theorem 6.2.5 (General Mean Value Theorem)
Let f and g be continuous functions on $[a, b]$ which are differentiable on (a, b). Then there is a point $c \in (a, b)$ such that

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

Proof. □

We can prove the Mean Value Theorem by letting $g(x) = x$ in the above theorem.

✾ 6.3 The Fundamental Theorem of Calculus

This theorem relates integration to differentiation; it links both of the main 'components' of calculus.

Theorem 6.3.1 (Fundamental Theorem of Calculus)
We have two statements:

1. If $f \in C([a, b])$, and
$$g(x) = \int_a^x f(t)\, dt,$$
then g is differentiable at each $x \in (a, b)$ and $g'(x) = f(x)$.
2. If G is differentiable and $G'(x) = g(x)$ is continuous on $[a, b]$, then
$$G(b) - G(a) = \int_a^b g(t)\, dt.$$

Proof. We prove both statements.

1. Note that
$$\frac{g(x+h)-g(x)}{h} = \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t.$$
Since f is a continuous function, for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $|t-x| < \delta$ implies that $|f(t) - f(x)| \le \varepsilon$. As h appraoches 0, we have
$$\left|\frac{g(x+h)-g(x)}{h} - f(x)\right| = \left|\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t - \frac{1}{h}f(x)(x+h-x)\right| = \frac{1}{|h|}\varepsilon\left|\int_x^{x+h} 1\,\mathrm{d}t\right| = \varepsilon.$$
Thus $g'(x)$ is defined, and equals $f(x)$.

2. Define
$$f(x) = \int_a^x G'(t)\,\mathrm{d}t.$$
Note that $f(a) = 0$. From the first part of this theorem, $f'(x) = G'(x)$. Define $h(x) = f(x) - G(x)$, so that $h'(x) = 0 \Rightarrow h(x) = k$ for some constant k. Since $h(a) = f(a) - G(a) = -G(a)$, we have $-G(a) = f(x) - G(x)$ for all $x \in [a, b]$. Thus
$$f(x) = G(x) - G(a) \longrightarrow f(b) = \int_a^b G'(t)\,\mathrm{d}t = \int_a^b g(t)\,\mathrm{d}t = G(b) - G(a).$$

□

Corollary

All continuous functions have antiderivatives.

Even though we cannot always express these antiderivatives in terms of elementary functions, they still exist. (A rather well known example of this is $f(t) = e^{t^2}$)

Remark. Assuming f is continuous, then we have
$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^x f(t)\,\mathrm{d}t\right) = f(x),$$
and that
$$\int_a^x \left(\frac{\mathrm{d}}{\mathrm{d}t}f(t)\right)\mathrm{d}t = f(x) - f(a).$$
Differentiation and integration are inverse processes.

✻ 6.4 Derivatives of Inverse Functions

Suppose f is a monotone function. If $f(x_1) < f(x_2) \Leftrightarrow x_1 < x_2$, then f must be bijective, and is an **invertible** map between $\operatorname{Dom} f$ and $\operatorname{Ran} f$, or $\exists f^{-1} : \operatorname{Ran} f \to \operatorname{Dom} f$.

Theorem 6.4.1

Suppose f is a strictly monotone function on $[a, b]$. If $\operatorname{Ran} f$ is an interval, then f is continuous.

Proof. WLOG let f be strictly increasing. Choose any $c \in (a, b)$ and let $\varepsilon > 0$. Let $y_1 = \max f(c) - \varepsilon, f(a)$ and $y_2 = \min f(c) + \varepsilon, f(b)$. Note that
$$f(a) \le y_1 < f(c) < y_2 \le f(b).$$
Note that $\operatorname{Ran} f = [f(a), f(b)]$, we have $y_1, y_2 \in \operatorname{Ran} f$. Choose $x_1, x_2 \in [a, b]$ such that $f(x_1) = y_1$ and $f(x_2) = y_2$. Note that $x_1 < c < x_2$ because f is monotonic. Let $\delta = \min|x_1 - c|, |x_2 - c|$. Note that $|x - c| \le \delta$ implies $x_1 \le x \le x_2$. Then we may write
$$y_1 \le f(x) \le y_2 \Leftrightarrow f(c) - \varepsilon \le f(x) \le f(x) + \varepsilon \Leftrightarrow -\varepsilon \le f(x) \le \varepsilon.$$

Thus $|x - c| \leq \delta$ implies $|f(x) - f(c)| \leq \varepsilon$, and f is continuous for any $c \in (a, b)$. The proofs are similar for $x = a, b$. Thus f is continuous. ☐

Theorem 6.4.2

Suppose f is a strictly monotone function on $[a, b]$. If f is continuous, then so is f^{-1}.

Proof. WLOG, suppose f is increasing. Since f is strictly monotone, the inverse exists. We have $f^{-1} : [f(a), f(b)] \longrightarrow [a, b]$, so $\operatorname{Ran} f$ is an interval. Choose any $y_1 < y_2$ in $\operatorname{Ran} f$. Then $f^{-1}(y_1) < f^{-1}(y_2)$, so f^{-1} is a strictly monotone increasing, and must be continuous. ☐

The next theorem shows the derivative of the inverse in terms of the inverse. There's a common proof that's wrong using the chain rule which doesn't work because it assumes f^{-1} is differentiable.

Theorem 6.4.3

Suppose f is a strictly monotone function on $[a, b]$. If f is differentiable at x_0 and $f'(x_0) \neq 0$, then f^{-1} is differentiable at $y_0 = f(x_0)$, and

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Proof. Let $\{y_n\} \to y_0$, where $y_n = f^{-1}(x_n)$ for some $\{x_n\} \to x_0$. Then we have

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \lim_{n\to\infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} = \lim_{n\to\infty} \frac{1}{\frac{f(x_n) - f(x_0)}{x_n - x_0}} = \frac{1}{f'(x_0)}.$$

☐

Chapter 7

Sequences of Functions

7.1 Introduction

This chapter deals with sequences of functions and their properties.

Definition 7.1.1: A **sequence of functions** on a set $E \subset \mathbb{R}$ can be defined as

$$\{f_n(x)\} = \{f_1(x), f_2(x), \cdots\}.$$

7.2 Pointwise Convergence

Definition 7.2.1: For each $n \in \mathbb{N}$, let f_n be a function defined on a set $E \subset \mathbb{R}$. The sequence f_n of functions **converges pointwise** on E to a **limiting function** $f : E \to \mathbb{R}$ if, for all $x \in E$, the sequence of real numbers $f_n(x)$ converges to $f(x)$.

Another way of looking at the above is that a sequence of functions converges pointwise if for all f_n,

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

for all $x \in E$. If for every $\varepsilon > 0$ and $x \in E$ we have a choice of $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| \leq \varepsilon$ whenever $n \geq N$, then we may establish that $\{f_n(x)\}$ converges to $f(x)$. Let's look at an example.

Example 7.2.2: Prove that

$$f_n(x) = \frac{x^2 + nx}{n}$$

converges pointwise on $[0, \infty]$.

Proof. Let $\varepsilon > 0$. Choose any $N \in \mathbb{N}$ such that $N \geq \frac{1}{\varepsilon} \geq \frac{x^2}{\varepsilon}$. If $n \geq N$, we have

$$\left|\frac{x^2 + nx}{n} - x\frac{n}{n}\right| = \left|\frac{x^2}{n}\right| \leq \frac{x^2}{N} \leq \varepsilon.$$

Thus $f_n(x)$ converges pointwise on $[0, \infty]$ to $f(x) = x$. $\square$

If $f_n(x)$ converges to $f(x)$ pointwise on E, it is *not* necessarily continuous, integrable, or differentiable. Note that even if $f_n(x)$ converges pointwise to $f(x)$ on $[a, b]$ and f_n, f are all integrable, the following doesn't necessarily hold:

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

Example 7.2.3: Consider the Weierstrass Function on $E = [-1, 1]$ such that

$$f_n(x) = \sum_{k=0}^{n} \left(\frac{1}{2}\right)^k \cos\left(3^k x\right).$$

Each f_n is differentiable on E, but $f(x) = \lim f_n(x)$ is continuous and *nowhere* differentiable.

If $f_n(x) \to f(x)$ pointwise, and each $f_n(x)$ is a continuous function on E, then we can write

$$\lim_{x \to a} f_n(x) = f_n(a),$$

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

Note that the above does *not* show continuity at $x = a$ for the limit function $f(x)$. However, if the above convergence is *uuniform*, then we can swap the limits like so:

$$\lim_{x \to a} \lim_{x \to \infty} f_n(x) = f(a) \iff \lim_{x \to a} f(x) = f(a),$$

meaning f *is* continuous at $x = a$.

✻ 7.3 Uniform Convergence

Definition 7.3.1: A sequence of functions f_n defined on $E \subset \mathbb{R}$ is **uniformly convergent** on E to a limit function $f : E \to \mathbb{R}$ if, for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \varepsilon$$

for $n \geq N$ and $x \in E$.

Let's look at a quick example.

Example 7.3.2: Let

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

Prove that f_n converges to 0 uniformly on $\mathbb{R}$.

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $N \geq \frac{1}{\varepsilon}$. For any $n \geq N$, we have

$$\left| \frac{1}{n(1 + x^2)} - 0 \right| = \frac{1}{n(1 + x^2)} \leq \frac{1}{n} \leq \frac{1}{N} \leq \varepsilon.$$

Since N is independent of x, f_n converges to 0 uniformly on $\mathbb{R}$. ☐

For the next few propositions, assume that $f_n \to f$ and $g_n \to g$ are uniformly convergent sequences of functions on a set E.

Proposition 7.3.3

$f_n + g_n$ is a uniformly convergent sequence of functions.

Proof. We claim that $f_n + g_n$ converges uniformly to $f + g$. Given any $\varepsilon > 0$, there exists an N_f such that for all $n \geq N_f$, $|f_n(x) - f(x)| \leq \frac{\varepsilon}{2}$. Similarly we have N_g. If $N = \max N_f, N_g$, then for all $n \geq N$, we have

$$|f_n(x) + g_n(x) - (f(x) + g(x))| \leq |f_n(x) - f(x)| + |g_n(x) - g(x)| \leq \varepsilon,$$

for all $x \in E$. ☐

While addition of sequences seems to hold, multiplication does *not*, there are some extra conditions involved.

Proposition 7.3.4

$f_n g_n$ is uniformly convergent if there exists an $M > 0$ such that $|f_n| \le M$ and $|g_n| \le M$ for all $n \in \mathbb{N}$.

Proof. Let $\varepsilon > 0$. There exists an $N_f \in \mathbb{N}$ such that $|f_n(x) - f(x)| \le \frac{\varepsilon}{2M}$ when $n \ge N$. Similarly, we can define N_g. Let $N = \max N_f, N_g$. Then we have the following:

$$|f_n(x)g_n(x) - f(x)g(x)| \le |f_n(x)g_n(x) - f(x)g_n(x)| + |f(x)g_n(x) - f(x)g(x)| =$$

$$|f_n(x) - f(x)||g_n(x)| + |g_n(x) - g(x)||f(x)| \le \frac{\varepsilon}{2M}M + \frac{\varepsilon}{2M}M = \varepsilon.$$

Thus $f_n g_n$ is uniformly convergent. □

Theorem 7.3.5

Let f_n be a sequence of functions defined on $E \subset \mathbb{R}$ that converges uniformly on A to a function f. If each f_n is continuous at $c \in A$, then f is continuous at c.

Proof. Let $\varepsilon > 0$. As $f_n \to f$ uniformly on the set E, choose a fixed $N \in \mathbb{N}$ such that $|f_N(x) - f(x)| \le \frac{\varepsilon}{3}$, for all $x \in E$. Since f_N is continuous, for any $c \in E$, we can choose $\delta > 0$ such that $|x - c| \le \delta$ implies $|f_N(x) - f_N(c)|$ for all $x \in E$. Then we have

$$|f(x) - f(c)| = |f(x) - f_N(x)| + |f_N(x) - f_N(c)| + |f_N(x) - f(c)| \le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

as desired. Thus

$$\lim_{x \to c} f(x) = f(c),$$

and f is continuous at c for all $c \in E$. □

Finally, let's look at an example.

Example 7.3.6: Prove that $f_n(x) = x^n$ converges uniformly on $[0, c]$ for all $c \in (0, 1)$.

Proof. Let $\varepsilon > 0$. Since $c \in (0, 1)$, we know that $\{c^n\} \to 0$. Then there exists an $N \in \mathbb{N}$ such that for all $n \ge N$, $|c^n - 0| = c^n \le \varepsilon$. Then we have

$$|x^n - 0| = x^n \le c^n \le \varepsilon,$$

for all $x \in [0, c]$, and x^n converges to 0 uniformly on $[0, c]$. □

✾ 7.4 Limit Theorems

Just like with sequences, sequences of functions also have a Cauchy criterion.

Theorem 7.4.1 (Cauchy Criterion for Uniform Convergence)

A sequence of functions $f_n(x)$ converges uniformly on a set $E \subset \mathbb{R}$ iff for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$|f_n(x) - f_m(x)| \le \varepsilon$$

for all $m \ge n \ge N$ and $x \in E$.

Proof. Suppose $f_n \to f$ uniformly on E. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ so that for all $n \ge N$,

$$|f_n(x) - f_m(x)| \le \frac{\varepsilon}{2}$$

for all $x \in E$. Then if $m \geq n \geq N$, we have

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Now suppose that for any $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $m \geq n \geq N$,

$$|f_n(x) - f_m(x)| \leq \varepsilon$$

for all $x \in E$. Then for each $x \in E$, $\{f_n(x)\}$ is Cauchy, and $\{f_n(x)\}$ converges, and f_n converges to f pointwise. Then we have

$$|f_n(x) - f(x)| = \lim_{m \to \infty} |f_n(x) - f_m(x)| \leq \varepsilon$$

for all $n \geq N$ and $x \in E$. So f_n converges to f uniformly. ☐

Theorem 7.4.2

If $f_n \to f$ uniformly on $[a, b]$, and each f_n is integrable then

$$\lim_{n \to \infty} \int_a^b f_n = \int_a^b f,$$

and f is integrable.

Proof. We prove f is integrable first. Let $\varepsilon > 0$. Since $f_n \to f$ uniformly on $[a, b]$, there exists an $N \in \mathbb{N}$ such that

$$|f_N(x) - f(x)| \leq \frac{\varepsilon}{2(b-a)}$$

for all $x \in [a, b]$. Let $d(x) = f(x) - f_N(x)$. Note that $|d(x)| \leq \dfrac{\varepsilon}{2(b-a)}$. Let P be any partition of $[a, b]$. Then

$$U_P(d) - L_P(d) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \leq 2\left(\frac{\varepsilon}{2(b-a)}\right) \sum_{i=1}^{n} \Delta x_i = \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

Then $d(x)$ is integrable, and $f(x) = d(x) + f_N(x)$ is integrable on $[a, b]$.

Now that f is integrable, we prove that

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

Let $\varepsilon > 0$. Since $f_n \to f$ uniformly on $[a, b]$, there exists an $M \in \mathbb{N}$ such that for all $n \geq M$,

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{b-a}$$

for all $x \in [a, b]$. Then we have

$$\left| \int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x \right| = \left| \int_a^b f_n(x) - f(x)\,\mathrm{d}x \right| \leq \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x \leq \frac{\varepsilon}{b-a}(b-a) = \varepsilon,$$

as desired. ☐

<++>

✾ 7.5 The Supremum Norm

We have already seen an example of a **norm** before, namely the **Euclidean Norm**:

Definition 7.5.1: The **Euclidean Norm** of a vector $v \in \mathbb{R}^n$ is

$$||v|| = \sqrt{\langle x, x \rangle} = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}.$$

We can think of the norm as a notion of the length or size of something. Now for the full definition:

Definition 7.5.2: A **norm** is a function $|| \cdot || : \mathcal{V} \to \mathbb{R}$ satisfying the following:

1. **Positivity**: $||v|| \geq 0$ for all $v \in \mathcal{V}$.
2. **Nondegeneracy**: $||v|| = 0$ iff $v = 0$.
3. **Multiplicativity**: $||\lambda v|| = |\lambda| ||v||$ for all $v \in \mathcal{V}$ and scalar λ.
4. **Triangle Inequality**: $||v + w|| \leq ||v|| + ||w||$ for all $w, v \in \mathcal{V}$.

We now consider the **ℓ_p norm**. For $p = 1$, we can define the ℓ_1 norm as

$$||x||_1 = |x_1| + |x_2| + \cdots + |x_n|.$$

For $p = 2$, the ℓ_2 norm is

$$||x||_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2)^{\frac{1}{2}}.$$

As $p \to \infty$ we have $||x||_\infty = \sup \{|x_1|, |x_2|, \cdots, |x_n|\}$. This last one is especially important:

Definition 7.5.3: Let f be a bounded function on a set $E \subset \mathbb{R}$. Then the **supremum norm** of f on E is defined as

$$||f||_\infty = \sup_{x \in E} |f(x)|.$$

Let's look at some examples.

Example 7.5.4: Find the supremum norm for $f(x) = \sin x$ on $E = \mathbb{R}$.

Solution. We have

$$||f||_\infty = \sup \{|\sin x| : x \in \mathbb{R}\} = 1.$$

Example 7.5.5: Find the supremum norm of $f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{otherwise.} \end{cases}$ on the set $E = [0, 1]$.

Solution. It's ok if the function is not continuous, like this one. We have

$$||f||_\infty = 1.$$

Let's establish the notion of convergence.

Definition 7.5.6: Let $\{f_n\}$ be a sequence of functions defined on $E \subset \mathbb{R}^n$. We say $\{f_n\}$ converges to f in the supremum norm if for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$||f_n - f||_\infty \leq \varepsilon$$

for all $n \geq N$.

This is similar to the normal definition of convergence, but in the supremum norm instead.

Theorem 7.5.7

Let $\{f_n\}$ be a sequence of functions defined on $E \subset \mathbb{R}^n$. Then $\{f_n\}$ converges uniformly on E to a limit function f iff $\{f_n\}$ converges to f in the supremum norm on E.

Proof. Suppose $f_n \to f$ uniformly on E. Then given any $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| \leq \varepsilon$$

for all $n \geq \mathbb{N}$ and $x \in E$. Then ε is an upper bound of

$$\{|f_n(x) - f(x)| : x \in E\},$$

and it follows that $||f_n - f||_\infty \leq \varepsilon$ for all $n \geq N$.

Now suppose for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $||f_n - f||_\infty \leq \varepsilon$ for all $n \geq N$. Then $|f_n(x) - f(x)| \leq \varepsilon$ for all $n \geq N$ and $x \in E$. Thus $f_n \to f$ uniformly on E, and we are done. □

Now that we've established convergence, we can also establish the notion of a Cauchy sequence:

Definition 7.5.8: Let $\{f_n\}$ be a sequence of functions defined on $E \subset \mathbb{R}^n$. We say $\{f_n\}$ is a Cauchy sequence in the supremum norm if for all $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$||f_n - f_m||_\infty \leq \varepsilon$$

for all $m \geq n \geq N$.

Recall that a **complete** metric space is one such that all Cauchy sequences converge to a point in that space. If $f_n \in \mathcal{S}$ is Cauchy in the supremum norm, then there must be a limit function $f \in \mathcal{S}$ such that $f_n \to f$ if $\mathcal{S}$ is a complete metric space.

Theorem 7.5.9

$C([a, b])$ is complete in the supremum norm.

Proof. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that for $m \geq n \geq N$, $||f_n - f_m|| \leq \varepsilon$. Then

$$|f_n(x) - f_m(x)| \leq \varepsilon$$

for all $x \in [a, b]$. Then $\{f_n(x)\}$ is Cauchy for all x. By the Cauchy Criterion for uniform convergence, $\{f_n\}$ converges uniformly to $f = \lim\limits_{n\to\infty} f_n(x)$. Since all $f_n \in C([a, b])$ and $f_n \to f$ uniformly, $f \in C([a, b])$, and $C([a, b])$ is complete in the supremum norm. □

In other words, this theorem states that if $\{f_n\}$ is Cauchy with respect to the supremum norm, then there is an $f \in C([a, b])$ such that $f_n \to f$.

Remark 7.5.10. Just because a space is complete in the supremum norm does *not* make it complete in other norms. For example, $C([a, b])$ is not complete with respect to the L^1 norm

$$||f||_1 = \int_a^b |f(x)|\, dx.$$

✾ 7.6 Metric Spaces

We now formally consider **metric spaces**.

Definition 7.6.1: A **metric space** $(\mathcal{M}, \rho)$ is a set $\mathcal{M}$ and a function $\rho : \mathcal{M} \times \mathcal{M}$ that satisfies the following:

1. **Positivity**: $\rho(x, y) \geq 0$ for all $x, y \in \mathcal{M}$.
2. **Nondegeneracy**: $\rho(x, y) = 0$ iff $x = y$.
3. **Symmetry**: $\rho(x, y) = \rho(y, x)$ for all $x, y \in \mathcal{M}$.
4. **Triangle Inequality**: $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for all $x, y, z \in \mathcal{M}$.

For instance, we can define $\rho(x, y)$ as the Euclidean norm, then we can make a metric space $(\mathbb{R}, \rho)$.

Definition 7.6.2: The **taxicab metric** or the **rectilinear distance** between two points (x_1, y_1) and (x_2, y_2) is $|x_1 - x_2| + |y_1 - y_2|$. Note that the taxicab distance between two points is *never* less than the straight line distance (euclidean norm) between themm.

There are many paths with equal distance from one point to another, trivial by combo.

Definition 7.6.3: The **discrete metric** is defined as

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

As shown, there are many different metrics. We can define the notion of equivalency (assuming both metrics are functions on the same set):

Definition 7.6.4: Two metrics, say ρ and σ on the same set $\mathcal{M}$ are **equivalent** if for every $\varepsilon > 0$ and $x \in \mathcal{M}$ there exists a $\delta > 0$ such that for all $y \in \mathcal{M}$ we have

$$\rho(x, y) < \delta \text{ implies } \sigma(x, y) < \varepsilon,$$

and that

$$\sigma(x, y) < \delta \text{ implies } \rho(x, y) < \varepsilon.$$

Let's look at an example.

Example 7.6.5: Let ρ be the Euclidean metric and σ be the taxicab metric. Prove that under $\mathbb{R}^2$ that ρ and σ are equivalent.

Solution. Let $\varepsilon > 0$ and choose $\delta < \frac{\varepsilon}{2}$. If $\rho(p, q) < \delta$, then

$$|p_i - q_i| \leq \rho(x, y) < \delta.$$

Then

$$\sigma(p, q) = |p_1 - q_1| + |p_2 - q_2| < 2\delta < \varepsilon.$$

If $\sigma(p, q) < \delta$, then $|p_i - q_i| < \delta$ means

$$\rho(p, q) = (|p_1 - q_1|^2 + |p_2 - q_2|^2)^{\frac{1}{2}} < (2\delta^2)^{\frac{1}{2}} = \frac{\varepsilon}{\sqrt{2}} < \varepsilon.$$

Thus ρ and δ are equivalent. $\square$

Proposition 7.6.6

Let ρ and σ be two equivalent metrics on a set $\mathcal{M}$ and suppose that $\{x_n\}$ is a sequence in $\mathcal{M}$. Then $x_n \to x$ in ρ iff $x_n \to x$ in σ.

This is useful to check whether two metrics are equivalent.

Example 7.6.7: Let ρ be the Euclidean metric and σ be the discrete metric. Prove that ρ and σ are *not* equivalent on $[0, 1]$.

Solution. Consider $\{\frac{1}{n}\}$. In the Euclidean metric, $\frac{1}{n} \to 0$, so $\{\frac{1}{n}$ converges. But with the discrete metric, $\sigma(\frac{1}{n}, 0) = 1$ for all $n \in \mathbb{N}$, and $\frac{1}{n}$ does not converge to 0. Thus ρ and σ are not equivalent. □

✾ 7.7 Contraction Mapping

Contraction mappings are functions that are important to DP problems.

Definition 7.7.1: Let $(\mathcal{M}, \rho)$ be a metric space. A function $T : \mathcal{M} \to \mathcal{M}$ is a **contraction** if there is an α which satisfies $0 \le \alpha < 1$ such that

$$\rho(T(x), T(y)) \le \alpha\rho(x, y)$$

for all $x, y \in \mathcal{M}$.

A quick example:

Example 7.7.2: Assuming ρ is the Euclidean metric, we have $f(x) = mx + b$ for $m \in [0, 1)$ on $\mathbb{R}$. Then we have $|f(x) - f(y)| = m|x - y|$, so f is a contraction.

Note that while something may seem to be a contraction, the property that $T : \mathcal{M} \to \mathcal{M}$ *must* hold in order for T to be a contraction mapping.

Theorem 7.7.3 (Contraction Mapping Principle)

Let T be a contraction on a complete metric space $(\mathcal{M}, \rho)$. Then there exists a unique point $x \in \mathcal{M}$ such that $T(x) = x$. Furthermore, the sequence defined recursively by

$$x_{n+1} = T(x_n)$$

where $x_0 \in \mathcal{M}$ must converge to x as $n \to \infty$.

Proof. Since T is a contraction mapping, we have $|T^{n+1}(x_0) - T^n(x_0)| \le \alpha|T^n(x_0) - T^{n-1}(x_0)| \le \alpha^2|T^{n-1}(x_0) - T^{n-2}(x_0)| \le \cdots \le \alpha^n|T(x_0) - T^0(x_0)| = \alpha^n|x_1 - x_0|$.

Suppose $m > n$. Then

$$|T^m(x_0) - T^n(x_0)| \le |T^m(x_0) - T^{m-1}(x_0)| + |T^{m-1}(x_0) - T^{m-2}(x_0)| + \cdots + |T^{n+1}(x_0) - T^n(x_0)| \le$$

$$\sum_{k=n}^{m-1} \alpha^k|x_1 - x_0| = \alpha^n\frac{|x_1 - x_0}{1 - \alpha}.$$

This tends to 0 as $n \to \infty$, so $\{T^n(x_0)\}$ is Cauchy. Since $\mathcal{M}$ is a complete metric space, this means that the above is convergent in $\mathcal{M}$. As T must be continuous by the definition of a contraction, we have

$$x = \lim x_n + 1 = \lim T(x_n) = T(\lim x_n) = T(x),$$

meaning x is a fixed point. Assume, for the sake of contradiction, that $T(x) = x$, $T(y) = y$, and $x \neq y$ for some $x, y \in \mathcal{M}$. Then we have

$$\alpha|x - y| \geq |T(x) - T(y)| = |x - y| \to \alpha \geq 1,$$

a contradiction. Thus x is unique, and we are done. $\square$

<++>

Chapter 8

Series

8.1 Introduction

Recall that an **infinite series** is a sum

$$\sum_{j=1}^{\infty} a_j$$

for some a_j, and the nth **partial sum**

$$S_n = \sum_{j=1}^{n} a_j.$$

Definition 8.1.1: Let $\{a_j\}$ be a sequence of real numbers, and consider the partial sums S_n for all $n \in \mathbb{N}$. If the sequence of partial sums $\{S_n\}$ converges to an $S \in \mathbb{R}$ as $n \to \infty$, then the series

$$\sum_{j=1}^{\infty}$$

converges, and we define

$$S = \sum_{j=1}^{\infty}.$$

Let's look at an example.

Example 8.1.2: Consider $a_j = \left(\frac{1}{10}\right)^j$. This converges to $0.\overline{1} = \frac{1}{9}$ as $n \to \infty$.

Next we define the notion of a geometric series.

Definition 8.1.3: Let $\alpha \in \mathbb{R}$. Then the series $\sum_{j=k}^{\infty} \alpha^j$ is called a **geometric series**, which converges if $|\alpha| < 1$ to

$$\sum_{j=k}^{\infty} \alpha^j = \frac{\alpha^k}{1-\alpha}.$$

This is easily proved by looking at the partial sum S_n. Now we establish the notion of convergence for series.

Theorem 8.1.4

Let $\{a_j\}$ be a sequence of real numbers. Then we have the following:

1. $\sum_{j=1}^{\infty} a_j$ converges iff for all $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ so that

$$\left| \sum_{j=n}^{m} a_j \right| \leq \varepsilon$$

for all $m \geq n \geq N$.

2. $\sum_{j=1}^{\infty} a_j$ converges iff $\{a_j\} \to 0$.

3. If $\sum_{j=1}^{\infty} |a_j|$ converges, $\sum_{j=1}^{\infty} a_j$ converges as well.

4. If $\sum a_j$ and $\sum b_j$ converge, then $\sum (ca_j + db_j)$ converges to $c \sum a_j + d \sum b_j$ for all $c, d \in \mathbb{R}$.

✾ 8.2 Convergence Tests

Recall from calculus that there are various 'convergence tests' to check whether a series converges.

Theorem 8.2.1 (Comparison Test)

Suppose $\{a_j\}$, $\{b_j\}$, and $\{c_j\}$ are *nonnegative* sequences with $a_j \leq b_j \leq c_j$ for all $j \in \mathbb{N}$. Then we have the following:

1. If $\sum c_j$ is convergent, then $\sum b_j$ converges.

2. If $\sum a_j$ diverges, then $\sum b_j$ converges.

Proof. We offer the proof for (1). Since we have $0 \leq b_j \leq c_j$ for all j, and $\sum c_j$ converges, let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that if $m \geq n \geq N$,

$$\left| \sum_{j=n}^{m} c_j \right| \leq \varepsilon.$$

Then we have

$$\left| \sum_{j=n}^{m} b_j \right| \leq \left| \sum_{j=n}^{m} c_j \right| \leq \varepsilon,$$

so $\sum b_j$ is Cauchy, and thus converges. The proof is similar for (2). □

Theorem 8.2.2 (Limit Comparison Test)

Suppose $\{a_j\}$ and $\{b_j\}$ are *nonnegative* sequences with $b_j > 0$. If

$$\lim_{j \to \infty} \frac{a_j}{b_j} = L$$

for some $L \in (0, \infty)$, then both a_j and b_j either both diverge or both converge.

Proof. Since the limit equals some L, there exists an $N \in \mathbb{N}$ large enough so that for all $j \geq N$, we have

$$-\frac{L}{2} \leq \frac{a_j}{b_j} - L \leq \frac{L}{2} \iff \frac{L}{2} b_j \leq a_j \leq \frac{3L}{2} b_j.$$

If $\sum b_j$ converges, then $a_j \leq \frac{3L}{2} b_j$ also converges. If $\sum b_j$ diverges, then $\frac{L}{2} b_j \leq a_j$, so $\sum a_j$ also diverges. $\square$

Theorem 8.2.3 (Integral Test)
Suppose $f(x)$ is continuous, positive, and monotone decreasing on $[K, \infty)$, and $f(j) = a_j$ for all $j \in \mathbb{N}$. Then

$$\int_K^\infty f(x)\,\mathrm{d}x \text{ and } \sum_{j=K}^\infty a_j$$

either both converge or both diverge.

A rather 'famous' example of this is the p-test for series of the form $\sum \frac{1}{x^p}$, where the series is convergent iff $p > 1$.When $p = 1$, we get the famously divergent **Harmonic Series**.

Theorem 8.2.4 (Alternating Series Test)
Suppose a_j is a monotone decreasing sequence of nonnegative numbers. If $a_j \to 0$, then

$$\sum_{j=1}^\infty (-1)^j a_j$$

converges.

This proof is done by noting that by the Nested Interval Property, the intersection of the partial sums is not empty, and the $a_j \to 0$ (the 'widths' of the intervals) meaning that the intersection of the intervals is a single element S, and that $S_n \to S$, meaning the series as a whole is convergent.

We list the last two tests without discussion.

Theorem 8.2.5 (Root Test)
Let $L = \lim\limits_{j\to\infty} |a_j|^{\frac{1}{j}}$.

1. If $L < 1$, the series is **absolutely convergent**, meaning that $\sum |a_j|$ converging $\to \sum a_j$ converges.
2. If $L > 1$, the series is divergent.
3. If $L = 1$, the series is inconclusive.

Theorem 8.2.6 (Ratio Test)
Let $L = \lim\limits_{j\to\infty} \left|\frac{a_{j+1}}{a_j}\right|$. Then the conditions for convergence and divergence are the same as the Root Test.

Note that pointwise and uniform convergence refer to $f_n \to f$ on $E \subset \mathbb{R}$, while **conditionally convergent** ($\sum a_j$ cocnverges but $\sum |a_j|$ does not) and absolute convergence refer to series of real numbers.

✾ 8.3 The Weierstrass M-Test

Note that the partial sums of a series S_n form a sequence of functions.

Definition 8.3.1 (Convergence of a Series)**:** For each $j \in \mathbb{N}$, let $f_j(x)$ and f be functions defined on a set $E \subset \mathbb{R}^n$. The infinite series

$$\sum_{j=1}^{\infty} f_j(x)$$

converges pointwise / uniformly on E to $f(x)$ if the sequence $S_n(x)$ converges pointwise / uniformly to $f(x)$.

Let's look at an example

Example 8.3.2: Does $\sum_{j=1}^{\infty} x_j$ converge? In what manner?

Solution. For any $x \in E$, we have $\sum_{j=1}^{\infty} x^j = \frac{x}{1-x}$ as $|x| < 1\ \forall x$. The nth partial sum $S_n(x) = \frac{x - x^{n+1}}{1-x}$. On E, this convergence is not uniform:

$$S_n(x) - f(x)| = \frac{|x|^{n+1}}{1-x},$$

and as $x \to 1^-$ the denominator converges to 0. Consider any interval $[-a, a] \subset E$. Then $|S_n(x) - f(x)| = \frac{a^{n+1}}{1-a}$, which converges to 0. Since this is independent of x, we have $S_n \to f$ uniformly on $[-a, a]$.

We can also interchange the continuity and summation:

Theorem 8.3.3 (Interchange of continuity and $\sum$)

Let $f_j(x)$ be continuous functions on $E \subset \mathbb{R}$. Suppose $\sum_{j=1}^{\infty} f_j(x)$ converges uniformly on E to f. Then f is continuous on E.

Proof. Note that if f_n are continuous, then the partial sums S_n are continuous as well. $\{S_n\} \to f$ uniformly, as the series converges uniformly on E, meaning f is continuous by the Uniform Limit Theorem. □

Theorem 8.3.4 (Cauchy Criterion for Uniform Convergence of a Series)

A series $\sum_{j=1}^{\infty} f_j(x)$ converges unformly on $E \subset \mathbb{R}$ if and only i for evey $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $m \geq n \geq N$,

$$|f_{n+1}(x) + \cdots + f_m(x)| \leq \varepsilon$$

for all $x \in E$.

Proof. insert later □

Now for the big theorem of this section!

Theorem 8.3.5 (The Weierstrass M Test)
For each $j \in \mathbb{N}$, let $f_j(x)$ be a function defined on a set $E \subset \mathbb{R}$, and let $M_j \in \mathbb{R}$ such that $M_j > 0$ and for all $x \in E$,

$$|f_j(x)| \le M_j.$$

Then if $\sum_{j=1}^{\infty} M_j$ converges, $\sum_{j=1}^{\infty} f_j(x)$ converges uniformly on $E \subset \mathbb{R}$.

Proof. Let $\varepsilon > 0$. Since $\sum M_j$ is a convergent series, there exists an $N \in \mathbb{N}$ such that for $m \ge n \ge N$, we have

$$\left|\sum_{j=n+1}^{m} M_j\right| \le \varepsilon.$$

Since $|f_j(x)| \le M_j$ for all $x \in E$, for any such $m \ge n \ge N$, we have

$$\left|\sum_{j=n+1}^{m} f_j(x)\right| \le |f_{n+1}(x)| + \cdots + |f_m(x)| \le \sum_{j=n+1}^{m} M_j \le \varepsilon,$$

and we are done. □

Theorem 8.3.6 (Interchange of $\int$ and $\sum$)
Let $f_j(x)$ be integrable functions on $[a, b]$ and suppose $\sum f_j(x)$ converges uniformly on $[a, b]$ to a function f. Then f is integrable on $[a, b]$, and

$$\int_a^x \sum_{j=1}^{\infty} f_j(t)\,\mathrm{d}t = \int_a^x f(t)\,\mathrm{d}t = \sum_{j=1}^{\infty} \int_a^x f_j(t)\,\mathrm{d}t.$$

Proof. Since $f_j(x)$ is integrable for all j, each partial sum is also integrable. Since the partial sums converge uniformly to f on $[a, b]$, f is also integrable. Let $\varepsilon > 0$, and choose $N \in \mathbb{N}$ such that

$$\left|\sum_{j=1}^{n} f_j(x) - f(x)\right| \le \frac{\varepsilon}{b-a},$$

for all $x \in [a, b]$ and $n \ge N$. Then when $n \ge N$, we have

$$\left|\int_a^x f(t)\,\mathrm{d}t - \sum_{j=1}^{n} \int_a^x f(t)\,\mathrm{d}t\right| = \left|\int_a^x f(t) - \sum_{j=1}^{n} f_j(t)\,\mathrm{d}t\right| \le \int_a^x \left|f(t) - \sum_{j=1}^{n} f_j(t)\right| \mathrm{d}t \le \int_a^x \frac{\varepsilon}{b-a}\,\mathrm{d}t \le \varepsilon,$$

as desired. □

Theorem 8.3.7

Let $f_j(x)$ be differentiable functions defined on an interval $E \subset \mathbb{R}$, and assume that $\sum_{j=1}^{\infty} f_j'(x)$ converges uniformly to $g(x)$ on E. If there exists an $x_0 \in [a, b]$ where $\sum_{j=1}^{\infty} f_j(x_0)$ converges, then $\sum_{j=1}^{\infty} f_j(x)$ converges uniformly to a differentiable function $f(x)$ such that $f'(x) = g(x)$ on E. Then we have

$$f(x) = \sum_{j=1}^{\infty} f_j(x), \text{ and } f'(x) = \sum_{j=1}^{\infty} f_j'(x).$$

Note that this essentially states that if the above are satisfied, then $\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{j=1}^{\infty} f_j(x)\right) = \sum_{j=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_j(x)$.

✾ 8.4 Test Section

This is a test section on new keyboard. Supose we have a set S such that $\forall x \in S$ there exists a value $y \in S^c$ such that $x + y = 0$. If $1 \in S$, find S. *Solution.* We claim $S = \mathbb{R}^+$ or $\mathbb{R}^-$. Clearly, if $x \in S$, then $y = -x$ must be in S^c. Note that 0 cannot be in S, as then $0 \in S^c$, which is a contradiction. We also have $|S^c| > |S|$, as S^c must contain 0. Now we can also show that $S^c = \mathbb{R}_{\geq 0}$ or $\mathbb{R}_{\leq 0}$.

<++>