Real Analysis

MATH 104

Instructor: Peng Zhou

Kelvin Lee

UC Berkeley

Contents

Chapter 1

The Real Number Systems

1.1 Natural Numbers $\mathbb{N}$

Definition 1.1.1 (Peano Axioms (Peano Postulates))**.** The properties of the set of natural numbers, denoted $\mathbb{N}$, are as follows:

(i) 1 belongs to $\mathbb{N}$.

(ii) If n belongs to $\mathbb{N}$, then its successor $n+1$ belongs to $\mathbb{N}$.

(iii) 1 is not the successor of any element in $\mathbb{N}$.

(iv) If $n, m \in \mathbb{N}$ have the same successor, then $n = m$.

(v) A subset of $\mathbb{N}$ which contains 1, and which contains $n+1$ whenever it contains n, must equal to $\mathbb{N}$.

Remark. The last axiom is the basis of mathematical induction. Let $P_1, P_2, P_3, \ldots$ be a list of propositions that may or may not be true. The principle of mathematical induction asserts all the statements $P_1, P_2, \ldots$ are true provided

- P_1 is true. (Basis for induction)
- $P_n \implies P_{n+1}$. (Induction step)

1.2 Rational Numbers $\mathbb{Q}$

Definition 1.2.1 (Rational Numbers)**.** The set of **rational numbers**, denoted $\mathbb{Q}$, is defined by

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid n, m \in \mathbb{Z}, n \neq 0 \right\},$$

which supports addition, multiplication, subtraction, and division.

Remark. $\mathbb{Q}$ is a very nice algebraic system. However, there is no rational solution to equations like $x^2 = 2$.

Definition 1.2.2 (Algebraic Number)**.** A number is called an **algebraic number** if it satisfies a polynomial equation

$$c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0 = 0$$

where $c_0, \ldots, c_n$ are integers, $c_n \neq 0$ and $n \geqslant 1$.

Remark. Rational numbers are always algebraic numbers.

Theorem 1.2.3 (Rational Zeros Theorem)**.** Suppose $c_0, c_1, \ldots, c_n$ are integers and r is a rational number satisfying the polynomial equations

$$c_n x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 = 0$$

where $n \geqslant 1, c_n, c_0 \neq 0$. Let $r = \frac{c}{d}$ where $\gcd(c,d) = 1$. Then $c \mid c_0$ and $d \mid c_n$. In simpler terms, the only rational candidates for solutions to the equation have the form $\frac{c}{d}$ where c is a factor of c_0 and d is a factor of c_n.

Proof. Plug in $r = \frac{c}{d}$ to the equation, we get

$$c_n \left(\frac{c}{d}\right)^n + c_{n-1}\left(\frac{c}{d}\right)^{n-1} + \cdots + c_1\left(\frac{c}{d}\right) + c_0 = 0.$$

Then we multiply by d^n on both sides and get

$$c_n c^n + c_{n-1}c^{n-1}d + \cdots + c_1 c d^{n-1} + c_0 d^n = 0.$$

Solving for $c_0 d^n$, we obtain

$$c_0 d^n = -c\left(c_n c^n + c_{n-1}^{n-2} + \cdots + c_2 c d^{n-2} + c_1 d^{n-1}\right).$$

Then it follows that $c \mid c_0 d^n$. Since $\gcd(c,d) = 1$, c can only divide c_0.
Now let's instead solve for $c_n c^n$, then we have

$$c_n c^n = -d\left(c_{n-1}c^{n-1} + c_{n-2}c^{n-2}d + \cdots + c_1 c d^{n-2} + c_0 d^{n-1}\right).$$

Thus $d \mid c_n c^n$, which implies $d \mid c_n$ because $\gcd(c,d) = 1$. $\square$

Corollary 1.2.4. Consider

$$x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 = 0,$$

where $c_0, c_1, \ldots, c_{n-1}$ are integers and $c_0 \neq 0$. Any rational solution of this equation must be an integer that divides c_0.

Proof. Since the Rational Zeros Theorem states that d must divide c_n, which is 1 in this case, r is an integer and it divides c_0. $\square$

Example 1.2.5. $\sqrt{2}$ is not a rational number.

Proof. Using Corollary 5, if $r = \sqrt{2}$ is rational, then $\sqrt{2}$ must be an integer, which is a contradiction. $\square$

1.3 Real Numbers $\mathbb{R}$

1.3.1 The Completeness Axiom

Definition 1.3.1 (Maximum/minimum). Let S be a nonempty subset of $\mathbb{R}$.

(i) If S contains a largest element s_0(i.e., $s_0 \in S$, $s \leqslant s_0 \forall s \in S$), then s_0 is the **maximum** of S, denoted $s_0 = \max S$.

(i) If S contains a smallest element, then it is called the **minimum** of S, denoted as $\min S$.

Remark.

- If s_1, s_2 are both maximum of S, then $s_1 \geqslant s_2, s_2 \geqslant s_1$, which implies that $s_1 = s_2$. Thus the maximum is **unique** if it exists.
- However, the maximum may not exist (e.g. $S = \mathbb{R}$).
- If $S \subset \mathbb{R}$ is a finite subset, then $\max S$ exists.

Definition 1.3.2 (Upper/Lower bound). Let S be a nonempty subset of $\mathbb{R}$.

(i) If a real number M satisfies $s \leqslant M$ for all $s \in S$, then M is an **upper bound** of S and S is said to be *bounded above*.

(i) If a real number m satisfies $\leqslant s$ for all $s \in S$, then m is a **lower bound** of S and S is said to be *bounded below*.

(i) S is said to be *bounded* if it is bounded above and bounded below. Thus S is bounded if there exist real numbers m and M such that $S \subset [m, M]$.

Definition 1.3.3 (Supremum/Infimum). Let S be a nonempty subset of $\mathbb{R}$.

- If S is bounded above and S has a least upper bound, then it is called the **supremum** of S, denoted by $\sup S$.
- If S is bounded below and S has a greatest lower bound, then it is called the **infimum** of S, denoted by $\inf S$.

Remark. If S has a maximum, then $\max S = \sup S$. Similarly, if S has a minimum, then $\min S = \inf S$. Also note that $\sup S$ and $\inf S$ need not belong to S.

Example 1.3.4. Suppose we have $S = \{1 - \frac{1}{n} \mid n \in \mathbb{N}\}$. Then $\max S$ does not exist and $\sup S = 1$.

Proof. Suppose for contradiction that it exists. Then it must be of the form $1 - \frac{1}{n_0}$ for some $n_0 \in \mathbb{N}$. However,

$$1 - \frac{1}{n_0 + 1} > 1 - \frac{1}{n_0},$$

and $1 - \frac{1}{n_0+1} \in S$. Hence a contradiction. □

Theorem 1.3.5 (Completeness Axiom). Every nonempty subset $S \subset \mathbb{R}$ that is bounded above has a least upper bound. In other words, $\sup S$ exists and is a real number.

Corollary 1.3.6. Every nonempty subset $S \subset \mathbb{R}$ that is bounded below has a greatest lower bound $\inf S$.

Proof. Consider the set $-S = \{-s \mid s \in S\}$. Since S is bounded below there exists an $m \in \mathbb{R}$ such that $m \leqslant s$ for all $s \in S$. This implies $-m \geqslant -s$ for all $s \in S$, so $-m \geqslant u$ for all $u \in -S$. Thus, $-S$ is bounded above by $-m$. The Completeness Axiom applies to $-S$, so $\sup -S$ exists.
Now we show that $\inf S = -\sup -S$. Let $s_0 = \sup -S$, we need to prove

$$-s_0 \leqslant s \qquad \text{for all } s \in S,$$

and if $t \leqslant s$ for all $s \in S$, then $t \leqslant -s_0$. The first inequality will show that $-s_0$ is a lower bound while the second inequality will show that $-s_0$ is the greatest lower bound, i.e., $-s_0 = \inf S$. The proofs of the two claims are left as an exercise. □

Theorem 1.3.7 (Archimedean Property)**.** If $a, b > 0$, then $na > b$ for some positive integer n.

Proof. Suppose the property fails for some pair of $a, b > 0$. That is, for all $n \in \mathbb{N}$, we have $na \leqslant b$, meaning that b is an upper bound for the set $S = \{na \mid n \in \mathbb{N}\}$. Using the Completeness Axiom, we can let $s_0 = \sup S$. Since $a > 0$, we have $s_0 - a < s_0$, so $s_0 - a$ cannot be an upper bound for S. It follows that $s_0 - a < n_0 a$ for some $n_0 \in \mathbb{N}$, which then implies that $s_0 < (n_0 + 1)a$. Since $(n_0 + 1)a$ is in S, s_0 is not an upper bound for S, which is a contradiction. □

Theorem 1.3.8 (Denseness of $\mathbb{Q}$)**.** If $a, b \in \mathbb{R}$ and $a < b$, then there is a rational $r \in \mathbb{Q}$ such that $a < r < b$.

Proof. We need to show that $a < \frac{m}{n} < b$ for some integers m and n where $n \neq 0$. Equivalently, we want

$$an < m < bn.$$

Since $b - a > 0$, the Archimedean property shows that there exists an $n \in \mathbb{N}$ such that

$$n(b - a) > 1 \implies bn - an > 1.$$

Now we need to show that there is an integer m between an and bn. □

1.4 + and −

We adjoint $+$ and $-$ to $\mathbb{R}$ and extend our ordering to the set $\mathbb{R} \cup \{-, +\}$. Explicitly, we have $- \leqslant a \leqslant +$ for all $a \in \mathbb{R} \cup \{-, +\}$.

Remark. $+$ and $-$ are not real numbers. Theorems that apply to real numbers would not work.

We define

$$\sup S = + \qquad \text{if } S \text{ is not bounded above}$$

and

$$\inf S = - \qquad \text{if } S \text{ is not bounded below.}$$

1.5 Reading (Rudin's)

1.5.1 Ordered Sets

Definition 1.5.1 (Order). Let S be a set. An **order** on S is a relation, denoted by $<$, with the following two properties:

- If $x \in S$ and $y \in S$, then one and only one of the statements
$$s < y, \quad x = y, \quad , y < x$$
is true.
- If $x, y, z \in S$, if $x < y$ and $y < z$, then $x < z$.

Definition 1.5.2 (Ordered Set). An **ordered set** is a set S in which an order is defined.

For example, Q is an ordered set if $r < s$ is defined to mean that $s - r$ is a positive rational number.

1.5.2 Fields

Definition 1.5.3 (Field). A **field** is a set F with two operations: *addition* and *multiplication*, which satisfy the following **field axioms**:

(A) **Axioms for addition**

(A1) If $x, y \in F$, then $x + y \in F$.

(A2) (Commutativity) $\forall x, y \in F$, $x + y = y + x$.

(A3) (Associativity) $\forall x, y, z \in F$, $(x + y) + z = x + (y + z)$.

(A4) (Identity) $\forall x \in F$, $0 + x = x$.

(A5) (Inverse) $\forall x \in F$, there exists a corresponding $-x \in F$ such that
$$x + (-x) = 0.$$

(M) **Axioms for multiplication**

(M1) If $x, y \in F$, then $xy \in F$.

(M2) (Commutativity) $\forall x, y \in F$, $xy = yx$.

(M3) (Associativity) $\forall x, y, z \in F$, $(xy)z = x(yz)$.

(M4) (Identity) $\forall x \in F$, $1x = x$.

(M5) (Inverse) $\forall x \in F$, there exists a corresponding $\frac{1}{x} \in F$ such that
$$x\left(\frac{1}{x}\right) = 1.$$

(D) **The distributive law**

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

Definition 1.5.4 (Ordered Field). An **ordered field** is a field F which is also an *ordered set*, such that

(i) if $y < z$ and $x, y, z \in F$, $x + y < x + z$,

(i) if $x, y > 0$ and $x, y \in F$, $xy > 0$.

Chapter 2

Sequences

2.1 Limits of Sequences

Definition 2.1.1 (Sequence). A **sequence** is a function whose domain is a set of the form $\{n \in \mathbb{Z} \mid n \geqslant m\}$ where m is usually 1 or 0.

One may wonder why do we care about sequence, and the answer is that sequences are useful for *approximation*.

Definition 2.1.2. A sequence $\{s_n\}$ of real numbers is said to **converge** to the real number s if $\forall \epsilon > 0$, $\exists N > 0$ such that for all positive integers $n > N$, we have

$$|s_n - s| < \epsilon.$$

If $\{s_n\}$ converges to s, we write $\lim_{n\to} s_n = s$, or simply $s_n \to s$, where s is called the **limit** of the sequence. A sequence that does not converge to some real number is said to **diverge**.

2.2 Proofs of Limits

Example 2.2.1. Prove $\lim_{n\to} \frac{1}{n^2} = 0$.

Scratch. For any $\epsilon > 0$, we want

$$\left|\frac{1}{n^2} - 0\right| < \epsilon \iff \frac{1}{n^2} < \epsilon \iff \frac{1}{\epsilon} < n^2 \iff \frac{1}{\sqrt{\epsilon}} < n.$$

Thus, we can just take $N = \frac{1}{\sqrt{\epsilon}}$. □

Proof. Let $\epsilon > 0$ and $N = \frac{1}{\sqrt{\epsilon}}$. Then $n > N$ implies $n > \frac{1}{\sqrt{\epsilon}}$ which implies $n^2 > \frac{1}{\epsilon}$ and hence $\epsilon > \frac{1}{n^2}$. Thus $n > N$ implies $\left|\frac{1}{n^2} - 0\right| < \epsilon$. This proves our claim. □

Example 2.2.2. Prove $\lim_{n\to} \frac{3n+1}{7n-4} = \frac{3}{7}$.

Scratch. $\forall \epsilon > 0$, we need $\left|\frac{3n+1}{7n-4} - \frac{3}{7}\right| < \epsilon$, which implies that

$$\left|\frac{21n + 7 - 21n + 12}{7(4n - 4)}\right| < \epsilon \implies \left|\frac{19}{7(7n - 4)}\right| < \epsilon.$$

Since $7n - 4 > 0$, we can remove the absolute value sign and have

$$\frac{19}{7\epsilon} < 7n - 4 \implies \frac{19}{49\epsilon} + \frac{4}{7} < n.$$

Thus, we have $N = \frac{19}{49\epsilon} + \frac{4}{7}$. □

Proof. Let $\epsilon > 0$ and let $N = \frac{19}{49\epsilon} + \frac{4}{7}$. Then $n > N$ implies $n > \frac{19}{49\epsilon} + \frac{4}{7}$, hence $7n > \frac{19}{7\epsilon} + 4$, which gives us $\frac{19}{7(7n-4)} < \epsilon$, and thus $\left|\frac{3n+1}{7n-4} - \frac{3}{7}\right| < \epsilon$. Then we are done. □

Example 2.2.3. Prove $\lim_{n\to} 1 + \frac{1}{n}(-1)^n = 1$.

Scratch. $\forall\epsilon > 0$, we want n large enough, such that

$$|a_n - 1| < \epsilon \iff \left|1 + \frac{1}{n}(-1)^n - 1\right| < \epsilon \iff \left|\frac{1}{n}(-1)^n\right| < \epsilon \iff \frac{1}{n} < \epsilon \iff n > \frac{1}{\epsilon}.$$

Just take $\alpha = \frac{1}{\epsilon}$, then $n > N \to |a_n - 1| < \epsilon$ □

2.3 Limit Theorems for Sequences

Definition 2.3.1 (Bounded)**.** A sequence $\{s_n\}$ of real numbers is said to be **bounded** if the set $\{s_n \mid n \in \mathbb{N}\}$ is a bounded set, i.e., if there exists a constant M such that $|s_n| \leqslant M$ for all n.

Theorem 2.3.2. Convergent sequences are bounded.

Proof. Let $\{s_n\}$ be a convergent sequence and let $s = \lim_{n\to} s_n$. Let $\epsilon > 0$ be fixed. Then by convergence of the sequence, there exists an number $N \in \mathbb{N}$ such that

$$n > N \implies |s_n - s| < \epsilon.$$

By the triangle inequality we see that $n > N$ implies $|s_n| < |s| + \epsilon$. Define $M = \max\{|s| + \epsilon, |s_1|, \ldots, |s_N|\}$. Then $|s_n| \leqslant M$ for all $n \in \mathbb{N}$, so $\{s_n\}$ is a bounded sequence. □

Theorem 2.3.3. Let $\{s_n\}$ and $\{t_n\}$ be sequences in $\mathbb{R}$ such that $s_n \to s$ and $t_n \to t$. Let $k \in \mathbb{R}$ be a constant. Then

(i) $ks_n \to ks$.

(ii) $(s_n + t_n) \to s + t$.

(iii) $s_nt_n \to st$.

(iv) If $s_n \neq 0$ for all n, and if $s \neq 0$, then $\frac{1}{s_n} \to \frac{1}{s}$.

(v) If $s_n \neq 0$ and $s \neq 0$ for all n, then $\frac{t_n}{s_n} \to \frac{t}{s}$.

Proof of (i). Since the case where $k = 0$ is trivial, we assume $k \neq 0$. Let $\epsilon > 0$ and we want to show that $|ks_n - ks| < \epsilon$ for large n. Since $\lim_{n\to} = s$, there exists N such that

$$n > N \implies |s_n - s| < \frac{\epsilon}{|k|}.$$

Then

$$n > N \implies |ks_n - ks| < \epsilon.$$

□

Proof of (ii). Let $\epsilon > 0$. We need to show

$$|s_n + t_n - (s + t)| < \epsilon \quad \text{for large } n.$$

Using triangle inequality, we have $|s_n + t_n - (s+t)| \leqslant |s_n - s| + |t_n - t|$. Since $s_n \to s$, there exists N_1 such that

$$n > N_1 \implies |s_n - s| < \frac{\epsilon}{2}.$$

Similarly, there exists N_2 such that

$$n > N_2 \implies |t_n - t| < \frac{\epsilon}{2}.$$

Let $N = \max\{N_1, N_2\}$. Then clearly

$$n > N \implies |s_n + t_n - (s+t)| \leqslant |s_n - s| + |t_n - t| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Proof of (iii). We use the identity

$$s_n t_n - st = (s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s).$$

Given $\epsilon > 0$, there are integers N_1, N_2 such that

$$\begin{aligned} n > N_1 &\implies |s_n - s| < \sqrt{\epsilon} \\ n > N_2 &\implies |t_n - t| < \sqrt{\epsilon} \end{aligned}$$

If we take $N = \max\{N_1, N_2\}, n \geqslant N$ implies

$$|(s_n - s)(t_n - t)| < \epsilon$$

which implies that

$$\lim_{n\to\infty} (s_n - s)(t_n - t) = 0.$$

Applying (i) and (ii), we get

$$\lim_{n\to\infty} (s_n t_n - st) = 0.$$

□

Proof of (iv). Choosing m such that $|s_n - s| < \frac{1}{2}|s|$ if $n \geqslant m$, we see that

$$|s_n| > \frac{1}{2}|s| \quad (n \geqslant m).$$

Given $\epsilon > 0$, there is an integer $N > m$ such that $n > N$ implies

$$|s_n - s| < \frac{1}{2}|s|^2\epsilon.$$

Hence, for $n \geqslant N$

$$\left|\frac{1}{s_n} - \frac{1}{s}\right| = \left|\frac{s_n - s}{s_n s}\right| < \frac{2}{|s|^2}|s_n - s| < \epsilon.$$

□

Proof of (v). Using (iv), we have $\frac{1}{s_n} \to \frac{1}{s}$, and by (iii), we get

$$\lim_{n\to} \frac{t_n}{s_n} = \lim_{n\to} \frac{1}{s_n} \cdot t_n = \frac{1}{s} \cdot t = \frac{t}{s}.$$

□

Theorem 2.3.4.

(i) $\lim_{n\to} \frac{1}{n^p} = 0$ for $p > 0$.

(ii) $\lim_{n\to\infty} a^n = 0$ if $|a| < 1$.

(iii) $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$.

(iv) $\lim_{n\to} a^{\frac{1}{n}} = 1$ for $a > 0$.

Proof of (i). Let $\epsilon > 0$ and let $N = \left(\frac{1}{\epsilon}\right)^{\frac{1}{p}}$. Then $n > N$ implies $n^p > \frac{1}{\epsilon}$ and thus $\epsilon > \frac{1}{n^p}$. Since $\frac{1}{n^p} > 0$, this shows $n > N$ implies $\left|\frac{1}{n^p} - 0\right| < \epsilon$. □

Proof of (ii). The case for $a = 0$ is trivial. Suppose that $a \neq 0$. Since $|a| < 1$, we can write $|a| = \frac{1}{1+b}$ where $b > 0$. By the binomial theorem, we have $(1+b)^n \geqslant 1 + nb > nb$, then

$$|a^n - 0| = |a^n| = \frac{1}{(1+b)^n} < \frac{1}{nb}.$$

Consider $\epsilon > 0$ and let $N = \frac{1}{\epsilon b}$. Then $n > N$ implies $n > \frac{1}{\epsilon b}$ and thus $|a^n - 0| < \frac{1}{nb} < \epsilon$. □

Proof of (iii). Let $s_n = n^{\frac{1}{n}} - 1$. Then $s_n \geqslant 0$ and by the binomial theorem,

$$n = (1 + s_n)^n \geqslant \frac{n(n-1)}{2}s_n^2.$$

Hence,

$$0 \leqslant s_n \leqslant \sqrt{\frac{2}{n-1}} \implies s_n \to 0.$$

□

Proof of (iv). Suppose $a > 1$. Let $s_n = a^{\frac{1}{n}} - 1$. Then $s_n > 0$, and by the binomial theorem,

$$1 + ns_n \leqslant (1 + s_n)^n = a,$$

so that

$$0 < s_n \leqslant \frac{p-1}{n}.$$

Hence, $s_n \to 0$. The case for $a = 1$ is trivial, and if $0 < p < 1$, the result is obtained by taking reciprocals. □

2.3.1 Upper and lower limits

Definition 2.3.5. Let $\{s_n\}$ be a sequence of real numbers with the property that for every real M there is an integer N such that $n \geqslant N$ implies $s_n \geqslant M$. We then write

$$s_n \to +.$$

Similarly, if for every real M there is an integer N such that $n \geqslant N$ implies $s_n \leqslant M$, we write

$$s_n \to -.$$

2.4 Monotone Sequences and Cauchy Sequences

Definition 2.4.1 (Monotone sequence)**.** A sequence $\{s_n\}$ of real numbers is called an *increasing sequence* if $s_n \leqslant s_{n+1}$ for all n, and $\{s_n\}$ is called a *decreasing sequence* if $s_n \geqslant s_{n+1}$ for all n. If $\{s_n\}$ is increasing, then $s_n \leqslant s_m$ whenever $n < m$. A sequence that is increasing or decreasing will be called a **monotone sequence** or a **monotonic sequence**.

Theorem 2.4.2. All bounded monotone sequences converge.

Proof. Let $\{s_n\}$ be a bounded increasing sequence, Let $= \{s \mid n \in \mathbb{N}\}$ and let $u = \sup S$, Since S is bounded, u represents a real number. We show $s_n \to u$. Let $\epsilon > 0$. Since $u - \epsilon$ is not an upper bound for S, there exists N such that $s_N > u - \epsilon$.Since $\{s_n\}$ is increasing, $s_N \leqslant s_n$ for all $n \geqslant N$. Of course $s_n \leqslant u$ for all n, so $n > N$ implies $u - \epsilon < s_n \leqslant u$, which implies $|s_n - u| < \epsilon$. Hence $s_n \to u$. The proof for bounded decreasing sequences is left as an exercise. □

Theorem 2.4.3.

(i) If $\{s_n\}$ is an unbounded increasing sequence, then $s_n \to +$.

(ii) If $\{s_n\}$ is an unbounded decreasing sequence, then $s_n \to -$.

Corollary 2.4.4. If $\{s_n\}$ is a monotone sequence, then the sequence either converges, diverges to $+$, or $-$. Thus $\lim s_n$ is always meaningful for monotone sequences.

Proof. Simply apply the previous two theorems. □

Definition 2.4.5. Let $\{s_n\}$ be a sequence in $\mathbb{R}$. We define

$$\limsup s_n = \lim_{N\to} \sup\{s_n \mid n > N\}$$

and

$$\liminf s_n = \lim_{N\to} \inf\{s_n \mid n > N\}$$

Theorem 2.4.6. Let $\{s_n\}$ be a sequence in $\mathbb{R}$.

(i) If $\lim s_n$ is defined (real, or $\pm\infty$), then

$$\liminf s_n = \lim s_n = \limsup s_n.$$

(ii) If $\liminf s_n = \limsup s_n$, then $\lim s_n$ is defined and

$$\lim s_n = \liminf s_n = \limsup s_n.$$

Definition 2.4.7 (Cauchy sequence)**.** A sequeunce $\{s_n\}$ of real numbers i called a **Cauchy sequeunce** if for each $\epsilon > 0$ there exists a number N such that

$$m, n > N \implies |s_n - s_m| < \epsilon.$$

Lemma 2.4.8. Convergent sequences are Cauchy sequences.

Proof. Suppose $\lim s_n = s$. Since the terms s_n are close to s for large n, they must also be close to each other; indeed

$$|s_n - s_m| = |s_n - s + s - s_m| \leqslant |s_n - s| + |s - s_m|.$$

Let $\epsilon > 0$. Then there exists N such that

$$n > N \implies |s_n - s| < \frac{\epsilon}{2}.$$

Clearly we can also write

$$m > N \implies |s_m - s| < \frac{\epsilon}{2},$$

so

$$m, n > N \implies |s_n - s_m| \leqslant |s_n - s| + |s - s_m| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Thus $\{s_n\}$ is a Cauchy sequence. □

Lemma 2.4.9. Cauchy sequences are bounded.

Proof. Let $\epsilon = 1$. By definition, we have N in $\mathbb{N}$ such that

$$m, n > N \implies |s_n - s_m| < 1.$$

In particular, $|s_n - s_{N+1}| < 1$ for $n > N$, so $|s_n| < |s_{N+1}| + 1$ for $n > N$. If $M = \max\{|s_{N+1} + 1, |s_1|, |s_2|, \ldots, |s_N|\}$, then $|s_n| \leqslant M$ for all $n \in \mathbb{N}$. □

Theorem 2.4.10. A sequence is a convergent sequence if and only if it is a Cauchy sequence.

Proof. Earlier we have shown one direction in a lemma. We now only need to show that Cauchy sequences are convergent sequences. Consider a Cauchy sequence $\{s_n\}$ and it is bounded by previous lemma. We now need to show that

$$\liminf s_n = \limsup s_n.$$

Let $\epsilon > 0$. Since $\{s_n\}$ is a Cauchy sequence, there exists N so that

$$m, n > N \implies |s_n - s_m| < \epsilon.$$

In particular, $s_n < s_m + \epsilon$ for all $m, n > N$. This shows $s_m + \epsilon$ is an upper bound for $\{s_n \mid n > N\}$, so $v_N = \sup\{s_n \mid n > N\} \leqslant s_m + \epsilon$ for $m > N$. This, in turn, shows $v_N - \epsilon$ is a lower bound for $\{s_m \mid m > N\}$, so $v_N - \epsilon \leqslant \inf\{s_m \mid m > N\} = u_N$. Thus

$$\limsup s_n \leqslant v_N \leqslant u_N + \epsilon \leqslant \liminf s_n + \epsilon.$$

Since this holds for all $\epsilon > 0$, we have $\limsup s_n \leqslant \liminf s_n$. Since $\limsup s_n \geqslant \liminf s_n$ always holds, we are done. □

2.4.1 Subsequences

Definition 2.4.11 (Subsequence)**.** Suppose $\{s_n\}_{n\in\mathbb{N}}$ is a sequence. A **subsequence** of this sequence is a sequence of the form $\{t_k\}_{k\in\mathbb{N}}$

Theorem 2.4.12. Every sequence $\{s_n\}$ has a monotonic subsequence.

Proof. We say that the n-th term is *dominant* if $s_m < s_n$ for all $m > n$. There are two cases:

Case 1: Suppose there are infinitely many dominant terms, and let $\{s_{nk}\}$ be any subsequence consisting solely of dominant terms. Then $s_{nk+1} < s_{n_k}$ for all k, so $\{s_{n_k}\}$ is a decreasing sequence.

Case 2: Suppose there are only finitely many dominant terms. Select n_1 so that s_{n_1} is beyond all the dominant terms of the sequence. Then given $N \geqslant n_1$, there exists $m > N$ such that $s_m \geqslant s_N$. □

Theorem 2.4.13 (Bolzano-Weierstrass Theorem)**.** Every bounded sequence has a convergent subsequence.

Proof. Using previous theorem, we have a monotonic subsequence. Since monotonic bounded sequence are convergent, we are done. □

Alternative proof. Suppose that $\{s_n\}$ is bounded. Then there exists $M > 0$ such that $|s_n| < M$ for all $n \in \mathbb{N}$. Let $A_1 = \{n \in \mathbb{N} \mid s_n \in [0, M]\}, B_1 = \{n \in \mathbb{N} \mid s_n \in [-M, 0]\}$. Since $A_1 \cup B_1 = \mathbb{N}$ is an infinite set, hence at least one of A_1, B_1 is infinite. WLOG assume that A_1 is infinite. We then cut $[0, M]$ into two halves, and repeat the same procedure, then at least one of $[0, M/2]$ and $[M/2, M]$ contains infinitely many points of the sequence. Then, we get a nested sequence of closed intervals,

$$I_1 \supset I_2 \supset \cdots, \qquad |I_{n+1}| = \frac{1}{2}|I_n|.$$

One can pick subsequence $\{s_{nk}\}$ such that for all k, s_{nk} is in I_k, and $n_{k+1} > n_k$. Then this subsequence is Cauchy, hence is convergent. $\square$

Definition 2.4.14 (Subsequential limit). A **subsequential limit** is any real number or symbol $\pm$ that is the limit of some subsequence of $\{s_n\}$.

Example 2.4.15. Consider $\{s_n\}$ where $s_n = n^2(-1)^n$. The subsequence of even terms diverges to $+$ where as that of odd terms diverges to $-$. Hence, the set $\{-, +\}$ is the set of subsequential limits of $\{s_n\}$.

Example 2.4.16. Consider $\{r_n\}$, a list of all rational numbers. Every real number is a subsequential limit of $\{r_n\}$ as well as $\pm$. Thus, $\mathbb{R} \cup \{-, +\}$ is the set of subsequential limits of $\{r_n\}$.

Theorem 2.4.17. Let $\{s_n\}$ be any sequence. There exists a monotonic subsequence whose limit is $\limsup s_n$ and there exists a monotonic subsequence whose limit is $\liminf s_n$.

Proof. If $\{s_n\}$ is not bounded above, then a monotonic subsequence of $\{s_n\}$ has limit $\limsup s_n = +$. Similarly, if $\{s_n\}$ is not bounded below, a monotonic subsuquence has limit $\liminf s_n$.
Consider the case that it is bounded above. Let $t = \limsup s_n$, and consider $\epsilon > 0$. There exists N_0 so that for $N \geqslant N_0$,

$$\sup \{s_n \mid n > N\} < t + \epsilon.$$

In particular, $s_n < t + \epsilon$ for all $n > N_0$. We now claim

$$\{n \in \mathbb{N} \mid |s_n - t| < \epsilon\} \text{ is infinite.}$$

Otherwise, there exists $N_1 > N_0$ $\square$

Theorem 2.4.18. Let $\{s_n\}$ be any sequence in $\mathbb{R}$, and let S denote the set of subsequential limits of $\{s_n\}$.

(i) S is non-empty.

(ii) $\sup S = \limsup s_n$ and $\inf S = \liminf s_n$.

(iii) $\lim s_n$ exists if and only if S has exactly one element, namely $\lim s_n$.

Proof. (i) is an immediate consequence of the previous theorem.
To prove (ii), consider any limit t of a subsequence $\{s_{nk}\}$ of $\{s_n\}$. By the $\square$

2.5 lim sup's and lim inf's

Let $\{s_n\}$ be any sequence of real numbers, and let S be the set of subsequential limits of $\{s_n\}$. Recall the following definition:

$$\limsup s_n = \lim_{N\to} \sup s_n \mid n > N = \sup S$$

and

$$\liminf s_n = \lim_{N\to} \inf s_n \mid n > N = \inf S.$$

Claim.

$$\liminf s_n \leqslant \limsup s_n.$$

Proof. We know that

$$\sup_{n>N} s_n \geqslant \inf_{n>N} s_n.$$

Then take limit $N \to$. ☐

Claim. If $\{s_{n_k}\}$ is a subsequence, then

$$\limsup s_{n_k} \leqslant \limsup s_n.$$

Theorem 2.5.1. If $\{s_n\} \to s > 0$ and $\{t_n\}$ is any sequence, then

$$\limsup s_n t_n = s \cdot \limsup t_n.$$

Here we allow the conventions $s \cdot (\pm) = \pm$ for $s > 0$.

Proof. ☐

Question. If $\{s_{n_k} \cdot t_{n_k}\}$ converges, does that imply $\{t_{n_k}\}$ converges?

Answer. Yes. (Why?)

Theorem 2.5.2. Let $\{s_n\}$ be any sequence of nonzero real numbers. Then we have

$$\liminf \left|\frac{s_{n+1}}{s_n}\right| \leqslant \liminf |s_n|^{1/n} \leqslant \limsup |s_n|^{1/n} \leqslant \limsup \left|\frac{s_{n+1}}{s_n}\right|.$$

Question. If $\{s_n\}$ is a bounded positive sequence, is $\frac{s_{n+1}}{s_n}$ a bounded sequence?

Answer. No. Consider $0 < a, b < 1$, and take $a = \frac{1}{2}$ and $b = \frac{1}{n}$, then $\frac{a}{b} = \frac{n}{2}$.

Claim. If $\{s_n\}$ is bounded and monotone, then the ratio $\frac{s_{n+1}}{s_n}$ eventually converges to 1.

Proof. Since $\{s_n\}$ is bounded and monotone, it must converge to some limit s. Then

$$\lim \frac{s_{n+1}}{s_n} = \frac{\lim s_{n+1}}{s_n} = \frac{s}{s} = 1.$$

☐

Question. Is it possible to have s_n to be bounded, but $\frac{s_{n+1}}{s_n}$ unbounded?

Answer. Yes. Consider

$$s_n = \begin{cases} 1 & n \text{ is even;} \\ \frac{1}{n} & n \text{ is odd.} \end{cases}$$

Question. If $\{s_n\}$ is positive and bounded, is it possible that $\frac{s_{n+1}}{s_n} \to 0$?

Answer. Yes. Consider $s_n = \frac{1}{n!}$. Then

$$\lim \frac{s_{n+1}}{s_n} = \lim_{n \to} \frac{1}{n+1} = 0.$$

Chapter 3

Metric Spaces and Topology

3.1 Metric Spaces

Definition 3.1.1 (Metric Space). A set X, containing elements called **points**, is said to be a **metric space** if with any two points p and q of X there is associated a real number $d(p,q)$, called the *distance* from p to q, such that

(i) $d(p,q) > 0$ if $p \neq q$; $d(p,p) = 0$;

(ii) $d(p,q) = d(q,p)$;

(iii) $d(p,q) \leqslant d(p,r) + d(r,q)$, for any $r \in X$.

Any function with these three properties is called a **distance function**, or a **metric**.

Definition 3.1.2 (Induced Metric). Let (X,d) be a metric space, and let $S \subset X$. Then, $(S, d|_S)$ is a metric space, where $d|_S$ is the **induced metric**, which is the metric d when restricted to S.

3.1.1 Topological Definitions

Definition 3.1.3 (Topology). A **topology** on a set X is a collection $\mathcal{T}$ of subsets of X having the following properties:

(i) $\varnothing$ and X are in $\mathcal{T}$.

(ii) The union of the elements of any subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

(iii) The intersection of the elements of any finite subcollection of $\mathcal{T}$ is in $\mathcal{T}$.

A set X for which a topology $\mathcal{T}$ has been specified is called a **topological space**.

Definition 3.1.4 (Open). If X is a topological space with topology $\mathcal{T}$, we say that a subset $U \subset X$ is an **open set** of X if U belongs to the collection $\mathcal{T}$. Hence, a topological space is a set X together with a collection of open subsets of X, such that:

(i) $\varnothing$ and X are both open;

(ii) arbitrary unions of open sets are open;

(iii) finite intersections of open sets are open.

Definition 3.1.5 (Open/Closed Balls). Let (X, d) be a metric space. The **open ball** of *radius* ϵ at x is defined by:

$$\mathcal{B}_\epsilon(p) := \{x \in X \mid d(p, x) < \epsilon\}$$

and the **closed ball** is defined by:

$$\bar{\mathcal{B}}_\epsilon(p) := \{x \in X \mid d(p, x) \leqslant \epsilon\}.$$

Sometimes we also use the **neighborhood** of p to represent any open ball of any radius centered at p.

Definition 3.1.6 (Limit Point). A point $p \in E$ is a **limit point** if every open ball of p contains a point $q \neq p$ such that $q \in E$, i.e., for every $\delta > 0$,

$$\mathcal{B}_\delta^x(p) \cup E \neq \varnothing.$$

Definition 3.1.7 (Dense). $E \subset X$ is **dense** in X if every points of X is a limit point of E or a point of E, i.e., $\overline{E} = X$.

Definition 3.1.8 (Interior Point). Let (X, d) be a metric space, and $E \subset X$. A point $p \in E$ is called an **interior point** of E if there is a open ball $\mathcal{B}$ of p such that $\mathcal{B} \subset E$.

Definition 3.1.9 (Open Sets). A subset $U \subset X$ is **open** if and only if for any $p \in U$, there exists $\delta > 0$ such that the open ball

$$\mathcal{B}_\delta(p) = \{x \in X \mid d(p, x) < \delta\} \subset U.$$

In other words, U is open if every point of U is interior.

Definition 3.1.10 (Closed Sets). A subset $E \subset X$ is **closed** if every limit point of E is a point of E. Equivalently, E is closed if and only if for any point $x \in E^c$, there exists $\delta > 0$, such that $\mathcal{B}_\delta(x) \cap E = \varnothing$.

Theorem 3.1.11 (Open/Closed). A set E is open if and only if its complement E^c is closed. Similarly, it is closed if and only if its complement is open.

Definition 3.1.12 (Closure). Let X be a metric space, if $E \subset X$, the **closure** of E is the set $\overline{E} = E \cup E'$, where E' is the set of all limit points of E. In other words, the **closure** of E is the intersection of all closed sets containing E, i.e., it is the smallest closed set containing E.

Theorem 3.1.13. If X is a metric space and $E \subset X$, then

(i) the closure $\overline{E}$ is closed;

(ii) $E = \overline{E}$ if and only if E is closed;

(iii) $\overline{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$.

3.1.2 Compact Sets

Definition 3.1.14 (Open Cover). An **open cover** of a set E in a metric space X is a collection $\{U_i\}$ of open subsets of X such that $E \subset \bigcup_i U_i$.

Definition 3.1.15 (Compact Set). Let $K \subset S$. K is **compact** if every open cover of K contains a *finite* subcover. More explicitly, the requirement is that if $\{G_\alpha\}$ is an open cover of K, then there are finitely many indices $\alpha_1, \ldots, \alpha_n$ such that

$$K \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}.$$

Remark. Every finite set is compact. $\mathbb{R}$ is not compact.

Theorem 3.1.16. Compact subsets of metric spaces are closed.

Theorem 3.1.17. Closed subsets of compact sets are compact.

Corollary 3.1.18. If F is closed and K is compact, then $F \cup K$ is compact.

Theorem 3.1.19 (Heine-Borel Theorem). A subset $E \subset \mathbb{R}^k$ is compact if and only if it is closed and bounded.

Theorem 3.1.20. If $E \subset X$ is compact, then E is a closed and bounded subset of X.

Theorem 3.1.21 (Weierstrass). Every bounded infinite subset of $\mathbb{R}^k$ has a limit point in $\mathbb{R}^k$.

Definition 3.1.22 (Convergence of Metric Space). A sequence $\{s_n\}$ in a metric space (S, d) **converges** to $s \in S$ if $\lim_{n\to} d(s_n, s) = 0$. The sequence is a *Cauchy sequence* if for each $\epsilon > 0$, there exists an N such that

$$m, n > N \implies d(s_m, s_n) < \epsilon.$$

Lemma 3.1.23. If $\{s_n\}$ converges to s, then s_n is Cauchy.

Proof. For any $\epsilon > 0$, there exists $N > 0$ such that for all $n > N$

$$d(s_n, s) < \frac{\epsilon}{2}.$$

Thus, for all $n, m > N$, we have

$$\begin{aligned} d(s_n, s_m) &\leqslant d(s_n, s) + d(s_m, s) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

□

Definition 3.1.24 (Completeness). The metric space (S, d) is **complete** if every Cauchy sequence in S converges to some element in S.

Example 3.1.25 (Non-complete Metric Spaces).

1. $S = \mathbb{R}\backslash\{0\}$.
2. $S = \mathbb{Q}$.

Lemma 3.1.26. A sequence $\{\boldsymbol{x}^{(n)}\} \in \mathbb{R}^k$ converges iff for each $j = 1, 2, \ldots, k$, the sequence (x_j^n) converges in $\mathbb{R}$. A sequence $\{\boldsymbol{x}^{(n)}\}$ in $\mathbb{R}^k$ is a Cauchy sequence iff each sequence $\{x_j^{(n)}\}$ is a Cauchy sequence in $\mathbb{R}$.

Theorem 3.1.27. Euclidean k-space $\mathbb{R}^k$ is complete.

Theorem 3.1.28 (Bolzano-Weierstrass Theorem). Every bounded sequence in $\mathbb{R}^k$ has a convergent subsequence.

Theorem 3.1.29. Let $\{F_n\}$ be a decreasing sequence ($F_1 \supseteq F_2 \supseteq \cdots$) of closed bounded nonempty sets in $\mathbb{R}^k$. Then $F = \bigcap_{n=1} F_n$ is also closed, bounded and nonempty.

Definition 3.1.30 (Open Cover). Let $E \subset S$. An **open cover** of E is a collection $\{G_\alpha\}$of open subsets of S such that $E \subset \bigcup_\alpha G_\alpha$.

Remark. Every finite set is compact. $\mathbb{R}$ is not compact.

Theorem 3.1.31 (Heine-Borel Theorem). A subset E of $\mathbb{R}^k$ is compact iff it is closed and bounded.

Proof. Suppose $E \subset S$ is compact. Then pick some point $p \in S$ and consider $\{B_n(p) \mid n \in \mathbb{N}\}$, which covers S and thus covers E as well:

$$E \subset S = \bigcup_{n \in \mathbb{N}} B_n(p).$$

Since E is compact, there is a finite subcover such that

$$E \subset \bigcup_{i=1}^{M} B_{n_i}(p).$$

We can order the indices such that $n_1 < n_2 < \cdots, n_M$ then

$$E \subset B_{n_M}(p),$$

which implies that E is bounded. In particular, for any points $x, y \in E$,

$$d(x, y) \leqslant d(x, p) + d(y, p) \leqslant 2 \cdot n_M.$$

The remaining of the proof is left as an exercise. $\square$

Theorem 3.1.32. Every k-cell F in $\mathbb{R}^k$ is compact.

3.2 Connected Sets

Definition 3.2.1 (Separated)**.** Two subsets A, B of a metric space X are **separated** if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty, i.e

Definition 3.2.2 (Connected Sets)**.** A set $E \subset X$ is **connected** if E is not a union of two nonempty separated sets.

Theorem 3.2.3. A subset E of $\mathbb{R}$ is connected if and only if $x, y \in E$ and $z \in (x, y)$ implies $z \in E$.

Chapter 4

Series

4.1 Series

In this section we are interested in convergence of series, thus we use $\sum a_n$ to denote $\sum_{i=1} a_i$.

Definition 4.1.1 (Convergence/Divergence)**.** The n-th partial sum of a sequence $\{a_n\}$ is defined as $s_n = \sum_{i=1}^{n} a_i$. We say that $\sum a_n$ **converges** iff the sequence of partial sums $\{s_n\}$ converges to a real number. Otherwise, we say that the series **diverges**.

Definition 4.1.2 (Absolute Convergence)**.** The series $\sum a_n$ **converges absolutely** if $\sum |a_n|$ converges.

Definition 4.1.3 (Geometric Series)**.** A series of the form $\sum_{n=0} ar^n$ for constants a and r is a **geometric series**. For $r \neq 1$,

$$\sum_{k=0}^{n} ar^k = \frac{a(1-r^{n+1})}{1-r}.$$

For $|r| < 1$, since $\lim_{n\to} r^{n+1} = 0$, using the formula above gives

$$\sum_{k=0} ar^k = \frac{a}{1-r}.$$

If $a \neq 0$ and $|r| \geqslant 1$, then the sequence $\{ar^n\}$ does not converge to 0, so the series diverges.

Definition 4.1.4 (Cauchy Criterion)**.** A series $\sum a_n$ satisfies the **Cauchy criterion** if its sequence $\{s_n\}$ of partial sums is a Cauchy sequence, i.e., for each $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$n \geqslant m > N \implies \left| \sum_{i=m}^{n} a_i \right| < \epsilon.$$

Theorem 4.1.5. A series converges iff it satisfies the Cauchy criterion.

Corollary 4.1.6. If a series $\sum a_n$ converges, then $\lim a_n = 0$.

Proof. By Cauchy criterion, take $n = m$. Then for $\epsilon > 0$, there exists N such that $n > N$ implies $|a_n| < \epsilon$. Thus, $\lim a_n = 0$. $\square$

Remark. The converse is not true. Consider $\sum \frac{1}{n} = +\infty$.

Theorem 4.1.7 (Comparison Test)**.** Let $\sum a_n$ be a series where $a_n \geqslant 0$ for all n.

(i) If $\sum a_n$ converges and $|b_n| \leqslant a_n$ for all n, then $\sum b_n$ converges.

(ii) If $\sum a_n = +\infty$ and $b_n \geqslant a_n$ for all n, then $\sum b_n = +\infty$.

Proof of (i). For $n \geqslant m$, by the triangle inequality, we have

$$\left|\sum_{k=m}^{n} b_k\right| \leqslant \sum_{k=m}^{n} |b_k| \leqslant \sum_{k=m}^{n} a_k.$$

Since $\sum a_n$ converges, it satisfies the Cauchy criterion. It follows from the above that $\sum b_n$ also satisfies the Cauchy criterion, and so $\sum b_n$ converges. □

Proof of (ii). Let $\{s_n\}$ and $\{t_n\}$ be the sequences of partial sums for $\sum a_n$ and $\sum b_n$ respectively. Since $b_n \geqslant a_n$ for all n, we have $t_n \geqslant s_n$ for all n. Since $\lim s_n = +$, $\lim t_n = +$, and so $\sum b_n = +$. □

Theorem 4.1.8 (Ratio Test)**.** A series $\sum a_n$ of nonzero terms

1. converges absolutely if $\limsup \left|\frac{a_{n+1}}{a_n}\right| < 1$;
2. diverges if $\liminf \left|\frac{a_{n+1}}{a_n}\right| > 1$.
3. Otherwise $\liminf \left|\frac{a_{n+1}}{a_n}\right| \leqslant 1 \leqslant \limsup \left|\frac{a_{n+1}}{a_n}\right|$ and the test gives no information.

Theorem 4.1.9 (Root Test)**.** Let $\sum a_n$ be a series and let $\alpha = \limsup |a_n|^{\frac{1}{n}}$. The series $\sum a_n$

(i) converges absolutely if $\alpha < 1$;

(ii) diverges if $\alpha > 1$.

(iii) Otherwise, the test gives no information if $\alpha = 1$.

4.2 Alternating Series

Theorem 4.2.1. $\sum \frac{1}{n^p}$ converges iff $p > 1$.

Proof. If $p > 1$, then

$$\sum_{k=1}^{n} \frac{1}{k^p} \leqslant 1 + \int_1^n \frac{1}{x^p} dx = 1 + \frac{1}{p-1}\left(1 - \frac{1}{n^{p-1}}\right) < 1 + \frac{1}{p-1} = \frac{p}{p-1}.$$

Hence,

$$\sum_{n=1} \frac{1}{n^p} \leqslant \frac{p}{p-1} < +.$$

If $0 < p \leqslant 1$, then $\frac{1}{n} \leqslant \frac{1}{n^p}$ for all n. Since $\sum \frac{1}{n}$ diverges, $\sum \frac{1}{n^p}$ diverges as well by the Comparison Test. $\square$

Theorem 4.2.2 (Integral Tests). Suppose that $f(x) > 0$ and is decreasing on the infinite interval $[k, \infty)$ (for some $k \geqslant 1$) and that $a_n = f(n)$. Then the series $\sum_{n=1}^{\infty} a_n$ converges if and only if the improper integral $\int_1^{\infty} f(x)dx$ converges.

Theorem 4.2.3 (Alternating Series Theorem). If $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n \geqslant \cdots \geqslant 0$ and $\lim a_n = 0$, then the alternating series $\sum(-1)^{n+1}a_n$ converges. Moreover, the partial sums $s_n = \sum_{k=1}^{n}(-1)^{k+1}a_k$ satisfy $|s - s_n| \leqslant a_n$ for all n.

Proof. Define $s_n = \sum_{j=1}^{n} a_j$. The subsequence $\{s_{2n}\}$ is increasing because $s_{2n+2} - s_{2n} = -a_{2n+2} + a_{2n+1} \geqslant 0$, Similarly, the subsequence $\{s_{2n-1}\}$ is decreasing. $\square$

Chapter 5

Continuity

5.1 Limits of Functions

Definition 5.1.1 (ϵ-δ limit)**.** Let X, Y be metric spaces, and $E \subset X$, and p a limit point of E. We write the **limit**

$$\lim_{x \to p} f(x) = f(p)$$

if there exists $f(q) \in Y$ such that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$0 < d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \epsilon.$$

Theorem 5.1.2.

$$\lim_{x \to p} f(x) = q$$

if and only if

$$\lim_{n \to \infty} f(p_n) = q$$

for every sequence $\{p_n\}$ such that $p_n \neq p$ (for all n) and $p_n \to p$.

5.1.1 Continuous Functions

Definition 5.1.3 (Continuity)**.** Let X and Y be metric spaces. A function $f : X \to Y$ is **continuous** at $p \in X$ if for any $\epsilon > 0$, there exists $\delta > 0$ such that for every $x \in X$,

$$d_X(x, p) < \delta \implies d_Y(f(x), f(p) < \epsilon.$$

Or equivalently, for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$f(B_\delta(p)) \subset B_\epsilon(f(p)).$$

Theorem 5.1.4. If p is a limit point of E. Then f is continuous at p if and only if $\lim_{x \to p} f(x) = f(p)$.

Theorem 5.1.5 (Preimage of open subset is open)**.** Let X and Y be metric spaces. A function $f : X \to Y$ is continuous if and only if for every open subset $U \subset Y$, $f^{-1}(U)$ is open.

Theorem 5.1.6 (Composition of continuous functions is continuous)**.** If $f : X \to Y$ and $g : Y \to Z$ are continuous, then

$$g \circ f : X \to Z \text{ is continuous.}$$

Theorem 5.1.7. Let f, g be complex continuous functions on metric space X. Then $f + g$, fg, and $f|g$ are continuous on X.

5.2 Continuity and Compactness

Definition 5.2.1. A function $f : X \to Y$ is **bounded** if there exists $M \in \mathbb{R}$ such that $|f(x)| \leqslant M$ for all $x \in X$.

Theorem 5.2.2 (Compactness is preserved under continuity)**.** If f is a continuous mapping of a compact metric space X into a metric space Y. Then $f(X)$ is compact.

Theorem 5.2.3. Suppose f is a continuous real function on a compact metric space X, and

$$M = \sup{}_{p \in X} f(p), \quad m = \inf{}_{p \in X} f(p)$$

Then there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

5.3 Uniform Continuity

Definition 5.3.1 (Uniformly Continuous)**.** Let f be a mapping of a metric space X into a metric space Y. We say that f is uniformly continuous on X if for every $\epsilon > 0$ there exists $\delta > 0$ such that (15)

$$d_Y(f(p), f(q)) < \epsilon$$

for all p and q in X for which $d_X(p, q) < \delta$

Theorem 5.3.2. Let f be a continuous mapping of a compact metric space X into a metric space Y. Then f is **uniformly continuous** on X.

5.4 Continuity and Connectedness

Theorem 5.4.1 (Connectedness is preserved under continuity)**.** If f is a continuous mapping of metric space X to metric space Y and if E is a connected subset of X, then $f(E)$ is connected.

Theorem 5.4.2 (Intermediate Value Theorem)**.** Let f be a continuous real function on $[a, b]$. If $f(a) < f(b)$ and if $c \in (f(a), f(b))$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.

Proof. Since $[a, b]$ is connected, $f([a, b])$ is also connected subset of $\mathbb{R}$, which implies that $[f(a), f(b)] \subset f([a, b])$. $\square$

Chapter 6

Differentiation

6.1 The Derivative of a Real Function

Definition 6.1.1 (Derivative). Let $f : [a, b] \to \mathbb{R}$ be a real valued function. We say f is **differentiable** at a point $p \in [a, b]$ if the following limit exists:

$$f'(p) = \lim_{x \to p} \frac{f(x) - f(p)}{x - p} \qquad (x \in [a, b] \backslash \{p\})$$

f' is called the **derivative** of f.

Theorem 6.1.2. If f is differentiable at $p \in [a, b]$, then f is continuous at p.

Proof. We simply show that $\lim_{x \to p} f(x) = f(p)$, or $\lim_{x \to p}(f(x) - f(p)) = 0$. Since $f'(p)$ exists, we have

$$\begin{aligned}
\lim_{x \to p}(f(x) - f(p)) &= \lim_{x \to p} \left(\frac{f(x) - f(p)}{x - p} \cdot (x - p) \right) \\
&= \left(\lim_{x \to p} \frac{f(x) - f(p)}{x - p} \right) \cdot \left(\lim_{x \to p} x - p \right) \\
&= f'(p) \cdot 0 \\
&= 0.
\end{aligned}$$

□

Remark. It is not true that if f is differentiable at p, then f is continuous in a neighborhood of p. Consider

$$f(x) = \begin{cases} x^2 & x \in \mathbb{Q} \\ -x^2 & x \notin \mathbb{Q}. \end{cases}$$

f is both continuous and differentiable only at $x = 0$.

Remark. Consider

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

$f'(0)$ does not exist because
$$\lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} \sin\left(\frac{1}{x}\right)$$
does not exist.

Question. If $f : \mathbb{R} \to \mathbb{R}$ is a continuous function, and $f'(x)$ exists at all $x \in \mathbb{R}$. Is f' continuous?

Answer. No. Consider
$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x > 0 \\ 0 & x \leqslant 0. \end{cases}$$
Since $f'(0^+) = f'(0^-) = 0$, $f'(0) = 0$. For $x > 0$, $\lim_{x \to 0^+} f'(x) \neq 0$.

Theorem 6.1.3. Let $f, g : [a, b] \to \mathbb{R}$ and assume f, g are differentiable at p. Then

(i) $(f + g)'(p) = f'(p) + g'(p)$;

(ii) $(f \cdot g)'(p) = f'(p)g(p) + f(p)g'(p)$;

(iii) if $g(p) \neq 0$, then
$$(f/g)'(p) = \frac{f'g - fg'}{g^2}.$$

Proof of (ii).
$$\begin{aligned} \lim_{x \to p} \frac{f(x)g(x) - f(p)g(p)}{x - p} &= \lim_{x \to p} \frac{f(x)g(x) - f(x)g(p) + f(x)g(p) - f(p)g(p)}{x - p} \\ &= \lim_{x \to p} f(x) \cdot \frac{g(x) - g(p)}{x - p} + \lim_{x \to p} \frac{f(x) - f(p)}{x - p} \cdot g(p) \\ &= f(p)g'(p) + f'(p)g(p). \end{aligned}$$
□

Theorem 6.1.4 (Chain Rule)**.** Let $f : [a, b] \to \mathbb{R}$ be differentiable at $x_0 \in [a, b]$, and $g : I \to \mathbb{R}$ where $f([a, b]) \subset I$, and g is differentiable at $f(x_0)$. If
$$h(x) = g(f(x)) \qquad (x \in [a, b]),$$
then h is differentiable at x_0 and
$$h'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Let $y = f(x)$ and $y_0 = f(x_0)$.
$$\lim_{x \to x_0} \frac{h(x) - h(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{g(y) - g(y_0)}{x - x_0}.$$
Since $f'(x_0)$ exists, there exist functions u, v such that
$$f(x) = f(x_0) + (x - x_0)(f'(x_0) + u(x));$$

$$g(y) = g(y_0) + (y - y_0)(g'(y_0) + v(y)),$$

and $\lim_{x\to x_0} u(x) = 0$,$\lim_{y\to y_0} v(y) = 0$. Then

$$\begin{aligned} g(f(x)) - g(f(x_0)) &= (f(x) - f(x_0))(g'(f(x_0)) + v(f(x))) \\ &= (x - x_0)(f'(x_0) + u(x))(g'(f(x_0)) + v(f(x))). \end{aligned}$$

Hence,

$$\begin{aligned} \lim_{x\to x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} &= \lim_{x\to x_0} (f'(x_0) + u(x))(g'(f(x_0)) + v(f(x))) \\ &= f'(x_0)g'(f(x_0)). \end{aligned}$$

□

6.2 Mean Value Theorem

Definition 6.2.1 (Local Maximum). A point p is a **local maximum** of f if there exists a $\delta > 0$ such that $f(p) = \max f(\mathcal{B}_\delta(p))$. Likewise for local minimum.

Remark. If f is locally constant at p, then p is both a local maximum and local minimum.

Lemma 6.2.2. Let $f : [a, b] \to \mathbb{R}$. If f has a local maximum or local minimum at $p \in (a, b)$, and if $f'(p)$ exists, then $f'(p) = 0$.

Proof. Suppose f has a local maximum at p. Then there exists $\delta > 0$ such that $f(p) \geqslant f(x)$ for $x \in (p - \delta, p + \delta)$. The derivative is

$$f'(p) = \lim_{x\to p} \frac{f(x) - f(p)}{x - p}.$$

This limit is $\geqslant 0$ when $x \leqslant p$ and $\leqslant 0$ when xp. Since $f'(p)$ exists, then by squeeze theorem we must have $f'(p) = 0$. □

Remark. The conditions that $p \in (a, b)$ and $f'(p)$ exists are required since the endpoints a, b can be local maxima but the slopes there are not zero. In addition, there can be cases where p is a local maximum but $f'(p)$ does not exist, consider $f(x) = -|x|$.

Theorem 6.2.3 (Rolle's Theorem). Let $f : [a, b] \to \mathbb{R}$ be continuous, and suppose f is differentiable on (a, b), and $f(a) = f(b)$. Then there exists some $c \in (a, b)$ such that $f'(c) = 0$.

Remark. Note that $[a, b] \subset \mathbb{R}$ is compact, and so $f([a, b])$ is also compact.

Proof. Consider the following cases:

- if $f([a, b])$ is a single point, then f is a constant function, any $c \in (a, b)$ has $f'(c) = 0$.
- if $\max(f([a, b)] \neq f(a)$, then let $p \in (a, b)$ such that $f(p) = \max(f([a, b])$. Then by the above lemma, we have $f'(p) = 0$, where we let $c = p$.
- if $\min(f[a, b])) \neq f(a)$, then similar argument shows $f'(p) = 0$.

□

Theorem 6.2.4 (Generalized Mean Value Theorem). Let $f, g : [a, b] \to \mathbb{R}$ be continuous and differentiable in (a, b). Then there exists $c \in (a, b)$ such that

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

Proof. Take $h(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x)$. Then we have $h(a) = h(b)$. Hence, by Rolle's theorem, there exists $c \in (a, b)$ such that $h'(c) = 0$ as desired. $\square$

Theorem 6.2.5 (Mean Value Theorem). If $f : [a, b] \to \mathbb{R}$ is continuous and differentiable on (a, b), then there exists $c \in (a, b)$ such that

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

Proof. Use the generalized Mean Value Theorem by taking $g(x) = x$. $\square$

Corollary 6.2.6. Let f be differentiable on (a, b). Then for all $x \in (a, b)$,

(i) if $f'(x) \geqslant 0$, then f is strictly increasing;

(ii) if $f'(x) = 0$, then f is constant;

(iii) if $f'(x) \leqslant 0$, then f is strictly decreasing.

Proof of (i). Let $x < y$ be in (a, b). Then applying Mean Value Theorem to $[x, y]$, there exists some $c \in (x, y)$ such that

$$f'(c) = \frac{f(y) - f(x)}{y - x} \geqslant 0.$$

Hence, we have $f(y) \geqslant f(x)$. Similar arguments apply to the other two claims. $\square$

Corollary 6.2.7. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and differentiable everywhere on $\mathbb{R}$. Suppose there exists $M > 0$ such that $|f'(x)| \leqslant M$ for all $x \in \mathbb{R}$. Then f is uniformly continuous.

Proof. For any $\epsilon > 0$, take $\delta = \frac{\epsilon}{M}$. Then for any $x \neq y$, with $|x - y| < \delta$, there exists some $c \in (x, y)$ such that

$$f(y) - f(x) = (y - x)f'(c),$$

which implies

$$\begin{aligned} |f(y) - f(x)| &= |y - x| \cdot |f'(c)| \\ &< \delta \cdot M = \epsilon. \end{aligned}$$

$\square$

Theorem 6.2.8 (Intermediate Value Theorem for Derivatives). Let $f : [a,b] \to \mathbb{R}$ be differentiable such that $f'(a) < f'(b)$. Then for any $\lambda \in (f'(a), f'(b))$, there exists some $c \in (a,b)$ such that $f'(c) = \lambda$.

Remark. This is not an immediate application of the intermediate value theorem as the derivatives of continuous functions may not be continuous.

Proof. Let $g(x) = f(x) - \lambda x$. Our goal is to show that g has a root in (a,b). Since $g'(a) = f'(a) - \lambda < 0$, and $g'(b) = f'(b) - \lambda > 0$. Let $c \in [a,b]$ such that $c = \min g([a,b])$. Since $g'(a) < 0$ and $g'(b) > 0$, a, b are not global minimum, which implies that there exists some $c \in (a,b)$ that is a global minimum. Then using the previous lemma, we know that $g'(c) = f'(c) - \lambda = 0$ and so $f'(c) = \lambda$. □

6.3 L'Hospital's Rule

Theorem 6.3.1 (L'Hospital's Rule). Suppose $f, g : [a,b] \in \mathbb{R}$ are differentiable in (a,b) and $g'(x) \neq 0$ for all $x \in (a,b)$, where $-\infty \leqslant a < b \leqslant +\infty$. Suppose

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{+\infty, -\infty\}$$

and one of the following holds:

(i) $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$;

(ii) $\lim_{x\to a} |g(x)| = \lim_{x\to a} |f(x)| = +\infty$.

Then we have

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

Proof. TODO. □

Example 6.3.2.

$$\begin{aligned}
\lim_{x\to\infty} \left(1 + \frac{1}{x}\right)^x &= \lim_{x\to\infty} e^{x \log\left(1+\frac{1}{x}\right)} \\
&= e^{\lim_{x\to\infty} x \log\left(1+\frac{1}{x}\right)} \\
&= e.
\end{aligned}$$

6.4 Derivatives of Higher Order

Definition 6.4.1. If $f'(x)$ is differentiable at x_0, then the *second derivative* is defined as $f''(x_0) = (f')'(x_0)$. Similarly, if the $(n-1)$-th derivative $f^{(n-1)}$ exists and is differentiable at x_0, then the *n-th derivative* is defined as $f^{(n)}(x_0) = (f^{(n-1)})'(x_0)$.

Definition 6.4.2 (Smoothness). $f(x)$ is a *smooth* function on (a,b) if for any $x \in (a,b)$, $f^{(k)}(x)$ exists for all $k \in \mathbb{N}$. We also say that f is *infinitely differentiable*.

6.5 Taylor's Series

Definition 6.5.1 (Power Series). Given a sequence $\{c_n\}_{n\geqslant 0}$. A **power series** is defined by

$$\sum_{n=0}^{\infty} c_n(x-x_0)^n.$$

Proposition 6.5.2. Given a power series

$$f(z)=\sum_{n=0}^{\infty} c_n z^n.$$

Let $\alpha = \limsup \sqrt[n]{|c_n|}$ and $R=\frac{1}{\alpha}$. Then $f(z)$ converges for $|z|<R$ and diverges for $|z|>R$ (equality gives no info), where R is the **radius of convergence**.

Proof. Use root test for absolute convergence. If $|z|<R$, then $|c_n z^n|^{1/n}=|c_n|^{1/n}|z|$. Hence,

$$\lim_{n\to\infty}\sup |c_n z^n|^{1/n}=\alpha|z|<1.$$

Thus, $\sum_n |c_n z^n|$ is convergent, which implies that $\sum_n c_n z^n$ is convergent (absolute convergence implies convergence). If $|z|>R$, one can show that $|c_n z^n|$ does not converge to 0. □

Definition 6.5.3 (Taylor Series). Let f be a smooth function for which all higher derivatives exist at α. Then the **Taylor series** of f at α is defined as the power series

$$T_\alpha(x)=\sum_{k=0}^{\infty}\frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k.$$

Remark. The series may not converge. Even if it converges, the limit may not be $f(x)$.

Theorem 6.5.4 (Taylor's Theorem). Let $f:[a,b]\to\mathbb{R}$, $f^{(n-1)}$ exists and is continuous on $[a,b]$ and $f^{(n)}$ exists on (a,b). Let $\alpha,\beta\in[a,b]$ be distinct points and define

$$P_\alpha(x)=\sum_{k=0}^{n-1}\frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k.$$

Then for any $\beta\in(a,b)$, if $\beta\neq\alpha$, there exists $\gamma\in[\alpha,\beta]$ such that

$$f(\beta)=P_\alpha(\beta)+\frac{f^{(n)}(\gamma)}{n!}(\beta-\alpha)^n.$$

Intuition: Given a smooth function f, we can approximate $f(x)$ near α of different levels:

(i) 0-th order:

$$P_{\alpha,0}=f(\alpha).$$

(ii) 1-th order:

$$P_{\alpha,1}(x)=f(\alpha)+f'(\alpha)(x-\alpha).$$

(iii) 2-nd order
$$P_{\alpha,2}(x) = f(\alpha) + f'(\alpha)(x-\alpha) + \frac{f''(\alpha)}{2!}(x-\alpha)^2.$$

Taylor's theorem is all about the *error term* $f(x) - P_{\alpha,n-1}(x)$.

Remark. If $n = 1$, then $P_\alpha(x) = f(\alpha)$. The statement then becomes there exists $\gamma \in (\alpha, \beta)$ such that
$$f(\beta) - f(\alpha) = f'(\gamma)(\beta - \alpha),$$
which is the Mean Value Theorem. In general, the theorem shows that f can be approximated by a polynomial of degree $n-1$, and we can estimate the error, if we know bounds on $|f^{(n)}(x)|$.

Proof. Let $P(x) \doteq P_\alpha(x)$ for simplicity and let M be the number defined by
$$f(\beta) - P(\beta) = (\beta - \alpha)^n M.$$

Define
$$g(x) = f(x) - P(x) - M(x-\alpha)^n.$$
Then $g(\beta) = f(\beta) - P(\beta) - M(\beta - \alpha)^n = 0$ by the choice of M and $g(\alpha) = f(\alpha) - P(\alpha) - 0 = 0$. We want to show that $M = \frac{f^{(n)}(\gamma)}{n!}$ for some $\gamma \in (\alpha, \beta)$. By definition of g,
$$g^{(n)}(x) = f^{(n)}(x) - n!M \qquad (P(x) \text{ is degree } n-1 \text{ polynomial in } X).$$

Now our goal is to show that for any $x \in (a, b)$ there exists $\gamma \in (\alpha, \beta)$ such that $g^{(n)}(\gamma) = 0$.
Since we have $g(\alpha) = g(\beta) = 0$, by Rolle's there exists some $\gamma_1 \in (\alpha, \beta)$ such that $g'(\gamma_1) = 0$.

In addition, we have $g^{(k)}(\alpha) = 0$ for $k \in \{1, \ldots, n-1\}$. Since $g'(\alpha) = 0$ and $g'(\gamma_1) = 0$, by Rolle's there exists $\gamma \in (\alpha, \gamma_1)$ such that $g''(\gamma_2) = 0$. Then we repeat the argument and get $\gamma_n \in (\alpha, \gamma_{n-1})$ such that $g^{(n)}(\gamma_n) = 0$. Let $\gamma = \gamma_n$, then $g^{(n)}(\gamma) = 0$. □

Definition 6.5.5 (Analytic function)**.** If a smooth function $f(x)$ satisfies the condition that for any $x_0 \in (a, b)$ there exists $\gamma_0 > 0$ such that
$$f(x) = T_{x_0}(x), \qquad \forall |x - x_0| < \gamma_0,$$
then we say $f(x)$ is a (real) **analytic function**.

Remark. $\sin(x), \cos(x), e^x$, polynomials, and combinations of any of them are real analytic functions.

Chapter 7

The Riemann-Stieltjes Integral

7.1 Definition and Existence of the Integral

Definition 7.1.1 (Partition). A **partition** P of $[a, b] \subset \mathbb{R}$ is a finite set of points $\{x_i\}_{i=0}^{n}$ where $a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_{n-1} \leqslant x_n = b$, i.e.,

$$[a, b] = \bigcup_{i=0}^{n-1} [x_i, x_{i+1}]$$

Define

$$\Delta x_i = x_i - x_{i-1}, \qquad \forall i \in \mathbb{N}.$$

Let $f : [a, b] \to \mathbb{R}$ be real and bounded for the remaining of this section.

Definition 7.1.2 (Upper/lower Darboux sums). Given f and a partition P of $[a, b]$, the **upper** and **lower Darboux sums** are defined by

$$U(P, f) = \sum_{i=1}^{n} M_i \Delta x_i \quad \text{where } M_i = \sup \{f(x) \mid x \in [x_{i-1}, x_i]\},$$
$$L(P, f) = \sum_{i=1}^{n} m_i \Delta x_i \quad \text{where } m_i = \inf \{f(x) \mid x \in [x_{i-1}, x_i]\}.$$

Definition 7.1.3 (Upper/lower Darboux integrals). The **upper** and **lower Darboux integrals** are defined by

$$U(f) \doteq \overline{\int_a^b} f(x)dx = \inf U(P, f),$$
$$L(f) \doteq \underline{\int_a^b} f(x)dx = \sup L(P, f).$$

Definition 7.1.4 (Riemann Integral). If $U(f) = L(f)$, then the common value is denoted by

$$\int_a^b f dx, \quad \text{or} \quad \int_a^b f(x)dx,$$

which is the **Riemann integral** of f over $[a, b]$ and f is said to be *Riemann-integrable* on $[a, b]$ and we write $f \in \mathcal{R}$ (set of Riemann-integrable functions).

Since f is bounded, there exists $m, M \in \mathbb{R}$ such that $m \leqslant f(x) \leqslant M$ over $[a,b]$. Hence, for every P,

$$m(b-a) \leqslant L(P,f) \leqslant U(P,f) \leqslant M(b-a).$$

Remark. This shows that the upper and lower integrals are defined for every bounded function f.

Theorem 7.1.5. Suppose $f : [a,b] \to \mathbb{R}$ is bounded. Then $f \in \mathcal{R}$ if and only if for each $\epsilon > 0$ there exists a partition P of $[a,b]$ such that

$$U(P,f) - L(P,f) < \epsilon.$$

Let $\alpha : [a,b] \to \mathbb{R}$ be a monotonically increasing weight function. Define

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

Then define

$$U(P,f,\alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i,$$
$$L(P,f,\alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i,$$

and

$$U(f,\alpha) \doteq \overline{\int_a^b} f\,d\alpha = \inf U(P,f,\alpha),$$
$$L(f,\alpha) \doteq \underline{\int_a^b} f\,d\alpha = \sup L(P,f,\alpha),$$

Definition 7.1.6 (Riemann-Stieltjes integral)**.** If $U(f,\alpha) = L(f,\alpha)$, then the common value is denoted by

$$\int_a^b f\,d\alpha, \quad \text{or} \quad \int_a^b f(x)\,d\alpha(x),$$

which is the **Riemann-Stieltjes integral** of f with respect to α over $[a,b]$. f is also said to be *integrable with respect to* α, and write $f \in \mathcal{R}(\alpha)$ on $[a,b]$.

Remark. By taking $\alpha(x) = x$, the Riemann integral is seen to be a special case of the Riemann-Stieltjes integral.

Remark. Similarly as above, since f is bounded, we have the following inequalities:

$$m(\alpha(b) - \alpha(a)) \leqslant L(P,f,\alpha) \leqslant U(P,f,\alpha) \leqslant M(\alpha(b) - \alpha(a)).$$

Definition 7.1.7 (Refinement)**.** Let P, Q be two partitions of $[a,b]$, where

$$P = \{a = x_0 \leqslant x_1 \leqslant x_2 \leqslant \cdots \leqslant x_n = b\}$$
$$Q = \{a = y_0 \leqslant y_1 \leqslant y_2 \leqslant \cdots \leqslant y_m = b\}.$$

Q is a **refinement** of P if $Q \supset P$. Further, any two partitions P and Q have a **common refinement** $P \cup Q$.

Lemma 7.1.8. If Q is a refinement of P, then

$$L(P, f, \alpha) \leqslant L(Q, f, \alpha) \leqslant U(Q, f, \alpha) \leqslant U(P, f, \alpha).$$

In simpler terms, the refinement of partition improves the approximation.

Proof. It suffices to prove the case that Q has one more point than P. Let that point be x^* such that $x^* \in (x_{i-1}, x_i)$. Then let

$$\begin{aligned} w_1 &= \inf\{f(x) \mid x \in [x_{i-1}, x^*]\} \\ w_2 &= \inf\{f(x) \mid x \in [x^*, x_i]\}. \end{aligned}$$

Clearly $w_1 \geqslant m_i$ and $w_2 \geqslant m_i$, where as before

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

Hence,

$$\begin{aligned} L(Q, f, \alpha) - L(P, f, \alpha) &= w_1[\alpha(x^*) - \alpha(x_{i-1}] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_{i-1}] \\ &= (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \\ &\geqslant 0. \end{aligned}$$

Similar argument applies to the second inequality. □

Theorem 7.1.9.

$$L(f, \alpha) \leqslant U(f, \alpha).$$

Proof. For any partitions P_1, P_2 with common refinement $Q = P_1 \cup P_2$, we have

$$L(P_1, f, \alpha) \leqslant L(Q, f, \alpha) \leqslant U(Q, f, \alpha) \leqslant U(P_2, f, \alpha).$$

Then taking the sup over P_1 and the inf over P_2 gives

$$L(f, \alpha) \leqslant U(f, \alpha).$$

□

Theorem 7.1.10 (Cauchy Criterion)**.** $f \in \mathcal{R}(\alpha)$ on $[a, b]$ if and only if for every $\epsilon > 0$ there exists a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

Proof. By definition of sup and inf, for every partition P, we have

$$L(P, f, \alpha) \leqslant L(f, \alpha) \leqslant U(f, \alpha) \leqslant U(P, f, \alpha),$$

which implies

$$0 \leqslant U(f, \alpha) - L(f, \alpha) \leqslant U(P, f, \alpha) - L(P, f, \alpha).$$

Since for every ϵ there exists a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

Hence for every $\epsilon > 0$, we have

$$0 \leqslant U(f, \alpha) - L(f, \alpha) < \epsilon,$$

which implies that $U(f, \alpha) = L(f, \alpha)$, that is, $f \in \mathcal{R}(\alpha)$.
Conversely, suppose $f \in \mathcal{R}(\alpha)$, and let $\epsilon > 0$ be given. Since

$$\int f d\alpha = \sup{}_P L(P, f, \alpha) = \inf{}_P U(P, f, \alpha),$$

there exists P_1, P_2 such that

$$\int f d\alpha - L(P_1, f, \alpha) < \frac{\epsilon}{2}$$
$$U(P_2, f, \alpha) - \int f d\alpha < \frac{\epsilon}{2}.$$

Now let $P = P_1 \cup P_2$ be the common refinement. Then we have

$$\int f d\alpha - L(P, f, \alpha) < \frac{\epsilon}{2}$$
$$U(P, f, \alpha) - \int f d\alpha < \frac{\epsilon}{2},$$

which implies

$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

□

Theorem 7.1.11. Let $U_P = U(P, f, \alpha)$ and $L_P = L(P, f, \alpha)$.

(i) If $U_P - L_P < \epsilon$, then for any Q, refinement of P, we have

$$U_Q - L_Q < \epsilon.$$

(ii) If $U_P - L_P < \epsilon$, and let $s_i, t_i \in [x_{i-1}, x_i]$, then

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

(iii) If $f \in \mathcal{R}(\alpha)$, and $U_P - L_P < \epsilon$, $s_i \in [x_{i-1}, x_i]$, then

$$\left| \sum_{i=1}^{n} f(t_i) \Delta\alpha_i - \int_a^b f d\alpha \right| < \epsilon.$$

Proof of (ii). Since $|f(s_i) - f(t_i)| \leqslant M_i - m_i$, we have

$$\begin{aligned}\sum_{i=1}^{n} |f(s_i) - f(t_i)|\Delta\alpha_i &\leqslant \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &= U_P - L_P \\ &< \epsilon.\end{aligned}$$

□

Theorem 7.1.12. If f is continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$ on $[a, b]$.

Proof. Let $\epsilon > 0$ be given. Since f is continuous on a compact set, f is uniformly continuous. Hence, for every $\eta > 0$, there eixsts $\delta(\eta) > 0$ such that $|x - y| < \delta(\eta)$ implies $|f(x) - f(y)| < \eta$.

Take a partition P where $\Delta x_i < \delta(\eta)$ so that

$$M_i - m_i \leqslant \eta.$$

Hence,

$$\begin{aligned}U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &\leqslant \sum_{i=1}^{n} \eta\Delta\alpha_i \\ &= \eta(\alpha(b) - \alpha(a)).\end{aligned}$$

Choose η such that $\eta(\alpha(b) - \alpha(a)) < \epsilon$. □

Theorem 7.1.13. If f is monotonic on $[a, b]$ and α is also monotonic and continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$.

Proof. Let $\epsilon > 0$. For any $n \in \mathbb{N}$, choose a partition P such that

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

This is possible by the continuity of α and intermediate value theorem. Then

$$\begin{aligned}U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &= \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \cdot \frac{\alpha(b) - \alpha(a)}{n} \\ &= \frac{\alpha(b) - \alpha(a)}{n} \cdot (f(b) - f(a)).\end{aligned}$$

Then n large enough so that $U_P - L_P < \epsilon$. □

Theorem 7.1.14. Suppose f is bounded on $[a, b]$, f has only finitely many points of discontinuity on $[a, b]$ and α is continuous at every points at which f is dicontinuous. Then $f \in \mathcal{R}(\alpha)$.

Proof. Fix $\epsilon > 0$. Let $E = \{c_1 < c_2 < \cdots < c_m\}$ be the set of discontinuities for f. WLOG, assume $E \subset (a, b)$. Since α is continuous at c_i, we have

$$\alpha(c_i) = \lim_{t \to c_i^-} \alpha(t) = \lim_{t \to c_i^+} \alpha(t).$$

Hence we can take (u_i, v_i) around c_i such that

$$\begin{aligned} \alpha(v_i) - \alpha(c_i) &\leqslant \frac{\epsilon}{2m}, \\ \alpha(c_i) - \alpha(u_i) &\leqslant \frac{\epsilon}{2m}. \end{aligned}$$

Then we have

$$\alpha(u_i) - \alpha(v_i) \leqslant \frac{\epsilon}{m},$$

which implies that

$$\sum_{i=1}^{m} \alpha(u_i) - \alpha(v_i) \leqslant \epsilon.$$

Let $K = [a, b] \backslash \bigcup_{j=1}^{m} (u_i, v_i)$, a finite disjoint union of closed interval. Since f is continuous on K and K is compact, f is uniformly continuous on K. Hence there exists $\delta > 0$ such that for any $x, y \in K$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$.

Now let P be a partition of $[a, b]$ satisfying

(i) $[u_i, v_i]$ are intervals in P (jump interval or *bad* interval),

(ii) If $I_i = [x_{i-1}, x_i]$ is not a jump interval (*good* interval), i.e., $I_i \subset K$, then $|x_i - x_{i-1}| < \delta$.

Then

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i) \Delta\alpha_i \\ &= \sum_{I_i:\text{good}} (M_i - m_i)\Delta\alpha_i + \sum_{I_i:\text{bad}} (M_i - m_i)\Delta\alpha_i \\ &\leqslant \sum_{I_i:\text{good}} \epsilon\Delta\alpha_i + \sum_{I_i:\text{bad}} (M - m)\Delta\alpha_i \\ &\leqslant \epsilon[\alpha(b) - \alpha(a)] + (M - m)\epsilon \\ &= \epsilon[\alpha(b) - \alpha(a) + M - m]. \end{aligned}$$

Since ϵ is arbitrary, by the Cauchy criterion, we have $f \in \mathcal{R}(\alpha)$. □

Theorem 7.1.15. Suppose $f \in \mathcal{R}(\alpha)$ on $[a, b]$, where $m \leqslant f \leqslant M$, and ϕ is continuous on $[m, M]$, and $h = \phi \circ f$ on $[a, b]$. Then $h \in \mathcal{R}(\alpha)$ on $[a, b]$.

Proof. Fix $\epsilon > 0$. Since ϕ is uniformly continuous, there exists $\delta > 0$ such that for any $x, y \in [m, M]$, $|x - y| < \delta$ implies $|\phi(x) - \phi(y)| < \epsilon$. Let $K = \sup |\phi(x)|$ for any $x \in [m, M]$.

Since $f \in \mathcal{R}(\alpha)$, there exists partition P of $[a, b]$ such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \delta^2.$$

Let $M_i = \sup_{I_i} f(x)$, $m_i = \inf_{I_i} f(x)$, where $I_i = [x_{i-1}, x_i]$. Similarly, let $M_i^* = \sup_{I_i} h(x)$, $m_i^* = \inf_{I_i} h(x)$. Divide into two classes:

1. $i \in G$ if $M_i - m_i < \delta$,
2. $i \in B$ if $M_i - m_i \geqslant \delta$.

For $i \in G$, our choice of δ implies $M_i^* - m_i^* \leqslant \epsilon$. For $i \in B$, $M_i^* - m_i^* \leqslant 2K$. Then we have

$$\begin{aligned}
\delta^2 &\geqslant U(P, f\alpha) - L(P, f, \alpha) \\
&\geqslant \sum_{i \in B} (M_i - m_i)\Delta\alpha_i \\
&\geqslant \sum_{i \in B} \delta \Delta\alpha_i.
\end{aligned}$$

Hence,

$$\sum_{i \in B} \Delta\alpha_i \leqslant \delta.$$

Thus,

$$\begin{aligned}
U(P, h, \alpha) - L(P, h, \alpha) &= \sum_{i \in G} (M_i^* - m_i^*)\Delta\alpha_i + \sum_{iB} (M_i^* - m_i^*)\Delta\alpha_i \\
&\leqslant \epsilon[\alpha(b) - \alpha(a)] + 2K\delta \\
&< \epsilon[\alpha(b) - \alpha(a) + 2K].
\end{aligned}$$

Since ϵ is arbitrary, by Cauchy criterion, we have $h \in \mathcal{R}(\alpha)$. $\square$

7.2 Properties of the Integral

Theorem 7.2.1 (Properties of integrals). The integration operation has the following properties

(i) If $f_1, f_2 \in \mathcal{R}(\alpha)$ on $[a,b]$ and for any constant c, then

$$f_1 + f_2 \in \mathcal{R}(\alpha), \quad cf \in \mathcal{R}(\alpha),$$

$$\int_a^b (f_1 + f_2) d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha,$$
$$\int_a^b cf d\alpha = c \int_a^b f d\alpha.$$

(ii) If $f_1(x) \leqslant f_2(x)$ on $[a,b]$, then

$$\int_a^b f_1 d\alpha \leqslant \int_a^b f_2 d\alpha.$$

(iii) If $f \in \mathcal{R}(\alpha)$ on $[a,b]$ and if $a < c < b$, then $f \in \mathcal{R}(\alpha)$ on $[a,c]$ and on $[c,b]$, and

$$\int_a^c f d\alpha + \int_c^b f d\alpha = \int_a^b f d\alpha.$$

(iv) If $f \in \mathcal{R}(\alpha)$ and if $|f(x)| \leqslant M$ on $[a,b]$, then

$$\left| \int_a^b f d\alpha \right| \leqslant M[\alpha(b) - \alpha(a)].$$

(v) If $f \in \mathcal{R}(\alpha_1)$ and $f \in \mathcal{R}(\alpha_2)$, then $f \in \mathcal{R}(\alpha_1 + \alpha_2)$ and

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2$$
$$\int_a^b f d(c\alpha) = c \int_a^b f d\alpha.$$

Theorem 7.2.2. If $f \in \mathcal{R}(\alpha)$ and $g \in \mathcal{R}(\alpha)$ on $[a,b]$, then

(i) $fg \in \mathcal{R}(\alpha)$;

(ii) $|f| \in \mathcal{R}(\alpha)$ and $\left|\int_a^b f d\alpha\right| \leqslant \int_a^b |f| d\alpha$.

Proof. For (i), let $\phi(t) = t^2$, then $f^2 = \phi \circ f \in \mathcal{R}(\alpha)$ by previous theorem. Since $fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$, where the RHS is integrable with respect to α, $fg \in \mathcal{R}(\alpha)$ as well.

For (ii), let $\phi(t) = |t|$, then $|f| = \phi \circ f \in \mathcal{R}(\alpha)$. Choose $c = \pm 1$, so that

$$c \int f d\alpha \geqslant 0.$$

Then

$$\left|\int f d\alpha\right| = c\int f d\alpha = \int cf d\alpha \leqslant \int |f| d\alpha,$$

since $cf \leqslant |f|$. □

Definition 7.2.3 (Unit Step Function)**.** The **unit step function** I is defined by

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

Theorem 7.2.4. If $f : [a, b] \to \mathbb{R}$ is bounded and is continuous at $s \in (a, b)$, and $\alpha(x) = I(x-s)$, then

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

Proof. Consider partitions $P = \{a = x_0, s = x_1, x_2, x_3 = b\}$. Then

$$\begin{aligned} U(P, f, \alpha) &= \sup\{f(x) \mid x \in [s, x_2]\} \cdot 1 = \\ L(P, f, \alpha) &= \inf\{f(x) \mid x \in [s, x_2]\} \cdot 1. \end{aligned}$$

Since f is continuous at s, we see that $U_p, L_p \to f(s)$ as $x_2 \to s$. □

Theorem 7.2.5. Suppose $c_n \geqslant 0$ for $n = 1, 2, 3, \ldots, \sum c_n$ converges, $\{s_n\}$ is a sequence of distinct points in (a, b). and

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

Let f be continuous on $[a, b]$. Then

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} c_n f(s_n).$$

Theorem 7.2.6. Suppose α increases monotonically and $\alpha' \in \mathcal{R}$ on $[a, b]$. Let f be a bounded real function on $[a, b]$. Then $f \in \mathcal{R}(\alpha)$ if and only if $f\alpha' \in \mathcal{R}$. In that case

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

Theorem 7.2.7 (Change of Variable)**.** Suppose φ is a strictly increasing continuous function that maps an interval $[A, B]$ onto $[a, b]$. Suppose α is monotonically increasing on $[a, b]$ and $f \in \mathcal{R}(\alpha)$ on $[a, b]$. Define β and g on $[A, B]$ by

$$\beta(y) = \alpha(\varphi(y)), \quad g(y) = f(\varphi(y))$$

Then $g \in \mathcal{R}(\beta)$ and

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

7.3 Integration and Differentiation

Theorem 7.3.1 (Fundamental Theorem of Calculus I)**.** Let $f \in \mathscr{R}$ on $[a, b]$. For $a \leqslant x \leqslant b, put$

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

Then F is continuous on $[a, b]$; furthermore, if f is continuous at a point x_0 of $[a, b]$, then F is differentiable at x_0, and

$$F'(x_0) = f(x_0).$$

Theorem 7.3.2 (Fundamental Theorem of Calculus II)**.** If $f \in \mathcal{R}$ on $[a, b]$ and if there is a differentiable function F on $[a, b]$ such that $F' = f$, then

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

Theorem 7.3.3 (Integration by Parts)**.** Suppose F and G are differentiable functions on $[a, b]$, $F' = f \in \mathcal{R}$, and $G' = g \in \mathcal{R}$. Then

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

7.4 Uniform Convergence and Integration

Theorem 7.4.1. Let α be monotonically increasing on $[a, b]$. Suppose $f_n \in \mathcal{R}(\alpha)$ on $[a, b]$, for $n = 1, 2, 3, \ldots$, and suppose $f_n \to f$ uniformly on $[a, b]$. Then $f \in \mathcal{R}(\alpha)$ on $[a, b]$, and (23)

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

Corollary 7.4.2. If $f_n \in \mathcal{R}(\alpha)$ on $[a, b]$ and if

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad (a \leqslant x \leqslant b)$$

the series converging uniformly on $[a, b]$, then

$$\int_a^b f d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n d\alpha.$$

In other words, the series may be integrated term by term.

Chapter 8

Special Functions

8.1 The Gamma Function

Definition 8.1.1 (Gamma function)**.** For $0 < x < \infty$,

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Theorem 8.1.2. Properties of the gamma function:

(i) If $0 < x < \infty$,
$$\Gamma(x+1) = x\Gamma(x).$$

(ii) For $n \in \mathbb{N}$,
$$\Gamma(n+1) = n!.$$

(iii) $\log \Gamma$ is convex on $(0, \infty)$.

Theorem 8.1.3. If f is a positive function on $(0, \infty)$ such that

(i) $f(x+1) = xf(x)$,

(ii) $f(1) = 1$,

(iii) $\log f$ is convex,

then $f(x) = \Gamma(x)$.

8.1.1 Beta function

Theorem 8.1.4. If $x > 0$, and $y > 0$, then

$$\int_0^1 t^{x-1}(1-t)^{y-1}dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

where the integral is the **beta function** $B(x, y)$.

Chapter 9

The Lebesgue Theory

9.1