# STA 6934: Topics in Basic Analysis

# Contents

# Chapter 1

# An Introduction to Set Theory and the Real Line

## 1.1 Basics of Set Theory

**Definition 1.1.** A **set** is a collection of objects, which are called elements of the set. A set is typically denoted by capital letters A, B, C,...

**Definition 1.2.** Let A and B be sets. A is said to be a **subset** of B iff every element of A is an element of B, i.e. $x \in A \implies x \in B$. This is denoted by $A \subseteq B$.

Operations on Sets: Let A and B be sets

- Union: $A \cup B = \{x | x \in A \text{ or } x \in B\}$
- Intersection: $A \cap B = \{x | x \in A \text{ and } x \in B\}$
- Complementation: $A^c = \{x \in S | x \notin A\}$ ($\bar{A}$ is often used too), where $A \subseteq S$ and $S$ is the largest set in consideration.
- Set difference: $A \setminus B = \{x \in A | x \notin B\} = A \cap B^c$. This is also sometimes written as $A - B$.

**Definition 1.3.** Two sets $A$ and $B$ are equal if and only if $A \subseteq B$ and $B \subseteq A$.

Commutativity: Let A and B be sets. Then

- $A \cup B = B \cup A$
- $A \cap B = B \cap A$

Associativity: Let A and B be sets. Then

- $A \cup (B \cup C) = (A \cup B) \cup C$
- $A \cap (B \cap C) = (A \cap B) \cap C$

Distributive Laws: Let A,B, and C be sets. Then

- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

We will prove the second of the distributive laws below. I leave the rest for you to prove as an exercise.

*Proof of* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

(i) $(A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C))$. Let $x \in A \cap (B \cup C)$. Then we have $x \in A$ and either $x \in B$ or $x \in C$. Given $x \in A$, if $x \in B$, then $x \in A \cap B$. Likewise, given $x \in A$, if $x \in C$, then $x \in A \cap C$. Thus, given $x \in A$, if either $x \in B$ of $x \in C$, then either $x \in A \cap B$ or $x \in A \cap C$. That is, $x \in (A \cap B) \cup (A \cap C)$.

(ii) $(A \cap (B \cup C) \supseteq (A \cap B) \cup (A \cap C))$. Let $x \in (A \cap B) \cup (A \cap C)$. Then we have $x \in A \cap B$ or $x \in A \cap C$. If $x \in A \cap B$, then $x \in A$ and $x \in B$. If $x \in B$ then $x \in B \cup C$. Hence, $x \in A \cap (B \cup C)$. Similarly, if $x \in A \cap C$, then $x \in A \cap (B \cup C)$. Consequently, $x \in A \cap (B \cup C)$.

□

DeMorgan's Law: Let A and B be sets. Then

- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$

**Definition 1.4.** Two sets A and B are said to be **disjoint** (or mutally exclusive) if $A \cap B = \emptyset$. Sets $A_1, A_2, \ldots$ are pairwise disjoint if $A_i \cap A_j = \emptyset$ for all $i \neq j$.

In general, let $\{A_i, i \in I\}$ be a collection of sets. The union and intersection of an arbitrary collection of sets is define as

$$\cup_{i \in I} A_i = \{x | x \in A_i \text{ for some } i \in I\}$$

and

$$\cap_{i \in I} A_i = \{x | x \in A_i \text{ for all } i \in I\}.$$

**Example 1.1.** Find following sets:

(a) $(-\infty, a)^c$

(b) $\cap_{n=1}^{\infty} [a, b + 1/n)$

(c) $\cup_{n=1}^{\infty} [a + 1/n, b - 1/n]$

(d) $\cap_{n=1}^{\infty} [a - 1/n, a]$

**Definition 1.5.** Let $A$ and $B$ be two sets. The Cartesian product of $A$ and $B$, denoted $A \times B$, is the set of all ordered pairs $(a, b)$ such that $a \in A$ and $b \in B$. That is,

$$A \times B = \{(a, b) | a \in A \text{ and } b \in B\}.$$

The word "ordered" means that if $a, c \in A$ and $b, d \in B$ then $(a, b) = (c, d)$ if and only inf $a = c$ and $b = d$.

If $A_i$ is a collection of sets for $i \in \{1, 2, \ldots, n\}$, then the Cartesian product $\bigtimes_{i=1}^{n} A_i$ is

$$\bigtimes_{i=1}^{n} A_i = \{(a_1, \ldots, a_n) | a_i \in A_i, i = 1, \ldots, n\}.$$

If the sets $A_i$ are all equal, then we write $A^n$ for the Cartesian product. The following results can be easily verified:

- $A \times B = \emptyset$ if and only if $A = \emptyset$ or $B\emptyset$.
- $(A \cup B) \times C = (A \times C) \cup (B \times C)$
- $(A \cap B) \times C = (A \times C) \cap (B \times C)$
- $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$

**Definition 1.6.** Let $A$ and $B$ be sets and let $f : A \mapsto B$ be a mapping from $A$ to $B$. $A$ is called the **domain** of $f$ and the **range** of $f$ is $f(A) = \{y \in B | y = f(x) \text{ for some } x \in A\}$. Given $D \subseteq A$, the image of $D$ under $f$ is $f(D) = \{y \in B | y = f(x) \text{ for some } x \in D\}$. Given a $E \subseteq B$, the **pre-image** of $E$ under $f$ is $f^{-1}(E) = \{x \in A | f(x) \in E\}$.

**Definition 1.7.** A mapping $f : A \mapsto B$ is said to be **one-to-one** or **injective** if whenever $f(x) = f(y)$, then $x = y$. Equivalently, $f$ is one-to-one if whenever $x \neq y$, then $f(x) \neq f(y)$.

**Definition 1.8.** A mapping $f : A \mapsto B$ is said to be **onto** or **surjective** if $f(A) = B$. That is, $\forall y \in B$, $\exists x \in A$ such that $f(x) = y$.

## 1.2 The Natural Numbers and the Principle of Mathematical Induction

**Definition 1.9.** The natural numbers are the set of non-negative integers, denoted $\mathbb{N}$. That is,

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

The natural numbers are a well ordered set. This means, that for all $n, m \in \mathbb{N}$ exactly one of the following holds

(1) $n < m$

(2) $n > m$

(3) $n = m$

The natural numbers also has the following properties

(a) 1 is the smallest element of $\mathbb{N}$.

(b) Every non-empty subset of $\mathbb{N}$ has a smallest element. That is, if $A \subseteq \mathbb{N}$ and $A \neq \emptyset$, then $\exists a \in A$ such that $a \leq s$ for all $s \in A$.

(c) If $n \in \mathbb{N}$, then $n + 1 \in \mathbb{N}$. Note that this implies that $\mathbb{N}$ has no greatest element.

The last property is the basis of mathematical induction:

(i) Principle of mathematical induction: Let $P_k$ be a property indexed by $k$. We say that $P_n$ is true if property $P_k$ holds for $k = n \in \mathbb{N}$. The principle of mathematical induction is used to prove that the property holds for all $n \in \mathbb{N}$ as follows: if $P_1$ is true and for all $k \in \mathbb{N}$, $P_k \implies P_{k+1}$, then $P_n$ is true for all $n \in \mathbb{N}$.

(ii) Strong induction: If $P_1$ is true and for all $k \in \mathbb{N}$ , $\{P_1, P_2, \ldots, P_k\} \implies P_{k+1}$, then $P_n$ is true for all $n \in \mathbb{N}$.

**Example 1.2.** Use induction to prove that $\sum_{k=1}^{n} k = \dfrac{n(n+1)}{2}$ for all $n \in \mathbb{N}$.

*Proof.* Let $P_n : \sum_{k=1}^{n} k = \dfrac{n(n+1)}{2}$, and note that $P_1$ is true since

$$\sum_{k=1}^{1} k = 1 = \frac{1(1+1)}{2}.$$

Now, let $n \geq 1$ and suppose that $P_n$ is true. Then

$$\begin{aligned}\sum_{k=1}^{n+1} k &= \sum_{k=1}^{n} k + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) \qquad \text{(by the induction hypothesis)} \\ &= \frac{n(n+1)}{2} + \frac{2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

Thus, if $P_n$ is true, then $P_{n+1}$ is true, so by the principle of mathematical induction, the formula holds for all $n \in \mathbb{N}$. □

## 1.3 Properties of Real Numbers

The absolute value function is defined as

$$|x| = \begin{cases} x, & \text{if } x \geq 0 \\ -x, & \text{if } x < 0 \end{cases}.$$

**Proposition 1.1.** *Let $a, x \in \mathbb{R}$ with $a \geq 0$. Then $|x| \leq a$ if and only if $-a \leq x \leq a$.*

*Proof.* Let $a, x \in \mathbb{R}$ and $a \geq 0$.
($\implies$) Suppose $|x| \leq a$. If $x \geq 0$, then $-a \leq 0 \leq x = |x| \leq a$. If $x < 0$, then $-x = |x| \leq a \implies -a \leq x < 0 \leq a$.
($\impliedby$) Suppose $-a \leq x \leq a$. If $x \geq 0$, then $-a \leq 0 \leq x = |x| \leq a \implies |x| \leq a$. If $x < 0$, then $-a \leq x \leq a \implies -a \leq x \implies a \geq -x = |x|$. □

The absolute value function also satisfies the triangle inequality:

**Proposition 1.2.** *For all* $x, y \in \mathbb{R}$

$$|x + y| \leq |x| + |y|.$$

*Proof.* Let $x, y \in \mathbb{R}$. Note that

$$-|x| \leq x \leq |x| \text{ and } -|y| \leq y \leq |y|.$$

Together, these imply that

$$-(|x| + |y|) = -|x| - |y| \leq x + y \leq |x| + |y|.$$

Then by the previous proposition, $|x + y| \leq |x| + |y|$. □

**Corollary 1.1** (Reverse Triangle Inequality)**.** *For all* $x, y \in \mathbb{R}$

$$||x| - |y|| \leq |x - y|.$$

*Proof.* Exercise. □

**Definition 1.10.** Let $S \subset \mathbb{R}$.

(1) $M \in \mathbb{R}$ is said to be an **upper bound** of the set $S$ if $x \leq M$ for all $x \in S$.

(2) $L \in \mathbb{R}$ is said to be an **lower bound** of the set $S$ if $x \geq L$ for all $x \in S$.

(3) $S$ is said to be **bounded** if it is both bounded above and below. That is, if there exists and $M \in \mathbb{R}$ such that $|x| \leq M$ for all $x \in S$.

**Example 1.3.** Determine if the following sets are bounded above, bounded below or neither.

(a) $S = [0, 1)$

- $S$ is a bounded set.
- $S$ is bounded above by 1
- $S$ is bounded below by 0.

(b) $S = \mathbb{N}$

- $S$ is bounded below but not above.
- $S$ is bounded below by 1
- $S$ has no upper bound. If $M$ is an upper bound, then $x \leq \lfloor M \rfloor$ for all $x \in \mathbb{N}$, but $M < \lfloor M + 1 \rfloor \in \mathbb{N}$, a contradiction.

**Definition 1.11.** Let $S \subseteq \mathbb{R}$ be bounded above. The **supremum** of $S$ is the least upper bound. That is $\beta = \sup S \in \mathbb{R}$ if

a) $x \leq \beta$ for all $x \in S$.

b) $\forall \varepsilon > 0$, $\exists x \in S$ such that $\beta - \varepsilon < x \leq \beta$.

If the set is not bounded above, then we define $\sup S = \infty$.

**Note 1.1.** Part a) of the definition says that $\beta$ is an uppper bound. While part b) says that nothing smaller than $\beta$ can be an upper bound for $S$.

**Definition 1.12.** Let $S \subseteq \mathbb{R}$ be bounded below. The **infimum** of $S$ is the greatest lower bound. That is $\alpha = \inf S \in \mathbb{R}$ if

a) $x \geq \alpha$ for all $x \in S$.

b) $\forall \varepsilon > 0$, $\exists x \in S$ such that $\alpha \leq x < \alpha + \varepsilon$.

If $S$ is not bounded below, then we define $\inf S = -\infty$.

**Note 1.2.** Part a) of the definition says that $\alpha$ is a lower bound. While part b) says that nothing bigger than $\alpha$ can be a lower bound for $S$.

The question now becomes whether a supremum or infimum even exist given a bounded set? The completeness axiom guarantees their existence for bounded sets in the real line.

**Proposition 1.3** (Completeness Axiom of $\mathbb{R}$)**.** *Let $S$ be a nonempty subset of $\mathbb{R}$.*

*a) If $S$ is bounded above, then $S$ has a least upper bound.*

*b) If $S$ is bounded below, then $S$ has a greatest lower bound.*

**Example 1.4.** What are the supremum and infimum of the following sets:

a) $[0, 1)$

- $\inf[0, 1) = 0$
- $\sup[0, 1) = 1$ (Note that the supremum is not part of $S$.)

b) $\mathbb{N}$

- $\inf \mathbb{N} = 1$
- $\sup \mathbb{N} = \infty$.

For a set $A \subset \mathbb{R}$. Define $-A = \{-x | x \in A\}$.

**Proposition 1.4** (Properties of sup and inf)**.** *Let $A \subseteq B \subseteq \mathbb{R}$ be nonempty sets. Then*

*a)* $\sup(-A) = -\inf A$ *and* $\inf(-A) = -\sup A$

*b)* $\inf B \leq \inf A \leq \sup A \leq \sup B$.

*Proof.*

a) Let $\beta = \sup A$. We will show that $-\sup A = \inf(-A)$. To show this, recall that we need to prove two items:

   (i) $-\beta$ is a lower bound for $-A$.

   (ii) $\forall \varepsilon > 0$, $\exists y \in -A$ such that $-\beta \leq y < -\beta + \varepsilon$.

$\beta = \sup A \implies x \leq \beta\ \forall x \in A$, so $-\beta \leq -x, \forall x \in A$. This implies that $-\beta \leq y, \forall y \in -A$. Let $\varepsilon > 0$. Then $\exists x \in A$ such that $\beta - \varepsilon < x \leq \beta$, so $-\beta \leq -x < -\beta + \varepsilon$. Thus, $-\beta = \inf(-A)$.

The proof of $\sup(-A) = -\inf A$ is similar and is left as an exercise.

b) Homework.

□

A common fact that we will use frequently is that for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$0 < \frac{1}{N} < \varepsilon.$$

This follows from the Archimedean property of the real line.

**Proposition 1.5** (The Archimedean Property)**.** *If* $a, b \in \mathbb{R}$ *and* $a > 0$ *and* $b > 0$*, then* $\exists n \in \mathbb{N}$ *such that* $na > b$*.*

*Proof.* We will argue by contradiction. Let $a, b \in \mathbb{R}$, $a > 0$ and $b > 0$, and suppose that $\forall n \in \mathbb{N}, na \leq b$ or equivalently, $n \leq \frac{b}{a}, \forall n \in \mathbb{N}$. This implies that $\mathbb{N}$ is bounded above by $b/a$, so by the completeness axiom, $\beta = \sup \mathbb{N}$ exists and is finite. Then, there exists an $n \in \mathbb{N}$ such that

$$\beta - 1 < n \leq \beta \implies \beta < n + 1 \leq \beta + 1.$$

But $\beta$ is an upper bound for $\mathbb{N}$ and $n + 1 \in \mathbb{N}$, a contradiction. □

**Corollary 1.2.** $\forall \varepsilon > 0$*,* $\exists n \in \mathbb{N}$ *such that* $1/n < \varepsilon$*.*

*Proof.* Let $\varepsilon > 0$. By the Archimedean property, $\exists n \in \mathbb{N}$ such that

$$n\varepsilon > 1 \implies \frac{1}{n} < \varepsilon.$$

□

**Corollary 1.3.** *Let* $S = \{\frac{1}{n} | n \in \mathbb{N}\}$*. Then* $\max S = \sup S = 1$ *and* $\inf S = 0$*.*

*Proof.* Exercise. □

The rational numbers are $\mathbb{Q} = \left\{\frac{m}{n} | m \in \mathbb{Z}, n \in \mathbb{N}\right\}$, and it is a very important subset of $\mathbb{R}$. The following are two important properties of $\mathbb{Q}$ that will be useful.

**Proposition 1.6.**

*1)* $\mathbb{Q}$ *is a countable set. This means that there exists and one-to-one function* $f : \mathbb{N} \mapsto \mathbb{Q}$*.*

*2)* $\mathbb{Q}$ *is dense in* $\mathbb{R}$*. This means that* $\forall a, b \in \mathbb{R}$ *with* $a < b$*,* $\exists r \in \mathbb{Q}$ *such that* $a < r < b$*. This also implies that* $\forall x \in \mathbb{R}$ *and* $\forall \varepsilon > 0$*,* $\exists r \in \mathbb{Q}$ *such that*

$$|x - r| < \varepsilon.$$

*This says that we can approximate any real number by a rational number.*

*Proof.* We will only prove 2). Let $a, b \in \mathbb{R}$ with $a < b$. Then $b - a > 0$, and by Corollary 1.2, $\exists n \in \mathbb{N}$ such that $b - a > 1/n$. Without loss of generality (WLOG) assume $a > 0$. Then by the Archimedean property, $\exists m \in \mathbb{N}$ such that $m(1/n) > a$. Choose $m_0 \in \mathbb{N}$ to be the least such $m$ (this exists due to the well ordering property of $\mathbb{N}$). Then $\frac{m_0-1}{n} \leq a < \frac{m_0}{n}$. Together, this implies that

$$a < \frac{m_0}{n} = \frac{m_0 - 1}{n} + \frac{1}{n} \leq a + (b - a) = b.$$

□

It may seem like $\mathbb{Q}$ is a sufficient number system for analysis. You are used to seeing solutions approximated, e.g. $\sqrt{2} \approx 1.414$. We will see that $\sqrt{2} \notin \mathbb{Q}$. We can approximate $\sqrt{2}$ as well as we want by a rational number, but without the real line, an equation of the form $x^2 - 2 = 0$ has no solution. Before we can show that this equation has no rational solution, we prove the following proposition that such an equation does indeed have some solution:

**Proposition 1.7.** *Let $n \in \mathbb{N}$, $n \geq 2$. Then $\forall x > 0$, $\exists! y > 0$ such that $y^n = x$.*

*Proof.* Let $0 \leq s < t$. Then $(t^n - s^n) = (t - s)\sum_{k=0}^{n-1} t^k s^{n-1-k}$. This implies that

$$(t - s)ns^{n-1} \leq t^n - s^n \leq (t - s)nt^{n-1}. \qquad (\star\star)$$

(Uniqueness) Let $x > 0$ and suppose that $y_1, y_2 > 0$ satisfy, $y_1^n = y_2^n = x$. WLOG assume $0 < y_1 < y_2$. Then $(\star\star)$ implies

$$0 \leq (y_2 - y_1)ny_1^{n-1} \leq y_2^n - y_1^n = x - x = 0.$$

Since $y_2 > y_1 > 0$, this implies that $y_2 = y_1$. (Existence) Let $x > 0$. Define the following two sets

$$L = \{s \geq 0 | s^n < x\} \ \text{ and } \ U = \{t \geq 0 | t^n > x\}.$$

Note that $L \neq \emptyset$ since $0 \in L$. To see that $U \neq \emptyset$, we consider three cases: $0 < x < 1$, $x = 1$, and $x > 1$. First suppose $x > 1$. Then

$$x^n - x = x(x^{n-1} - 1^{n-1}) = x(x - 1)\sum_{k=0}^{n-2} x^k > x(x - 1) > 0 \implies x^n > x.$$

Thus $x > 1$ implies $x \in U$. Similarly, if $0 < x < 1$, then $1 \in U$. If $x = 1$, then $2^n > 1$ and $2 \in U$. Thus for any $x > 0$, $U \neq \emptyset$. Now, note that $\forall s \in L$ and $\forall t \in U$, $s^n < x < t^n \implies s < t$, since

$$0 < t^n - s^n = (t - s)\underbrace{\sum_{k=0}^{n-1} t^k s^{n-1-k}}_{\geq 0} \implies t - s > 0.$$

Thus, every $s \in L$ is a lower bound for $U$, and every $t \in U$ is an upper bound for $L$. Then, by the completeness axiom, $y = \sup L$ and $z = \inf U$ exist and are finite, and $y \leq z$ ($y \leq t$, $\forall t \in U$, since every $t \in U$ is an upper bound for $L$ and $y$ is the least upper bound of $L$. This means that $y$ is a lower bound for $U$ and hence $y \leq z$, since $z$ is the greatest lower bound of $U$.)

We want to show that $y = z$ and $y^n = x$. We will prove this by establishing the following claims:

1. $y^n \leq x$

   *Proof.* Suppose that $y^n > x$. Then $\frac{y^n - x}{ny^{n-1}} > 0$, and by definition of $\sup L$, $\exists s \in L$ such that
   $$y - \frac{y^n - x}{ny^{n-1}} < s \leq y \implies y - s < \frac{y^n - x}{ny^{n-1}}.$$
   Combining this with $(\star\star)$, we have
   $$y^n - s^n \leq (y - s)ny^{n-1} < \left(\frac{y^n - x}{ny^{n-1}}\right) ny^{n-1} < y^n - x.$$
   This implies that $-s^n < -x$, so $s^n > x$. But this can't happen, since $s \in L \implies s^n < x$. □

2. $s^n \leq x \iff s \leq y$

   *Proof.* ($\impliedby$) Let $s \leq y$. Then $s^n \leq y^n \leq x$ by claim 1.

   ($\implies$) Let $s^n \leq x$. If $s^n < x$, then $s \in L \implies s \leq y$. If $s^n = x$, then for all $0 \leq w < s$, $w^n < s^n = x$. Thus $[0, s) \subset L$, so $s = \sup[0, s) \leq \sup L = y$. □

3. $z^n \geq x$

   *Proof.* Suppose $z^n < x$. Let $\varepsilon = \min\left\{1, \frac{x - z^n}{n(1+z)^{n-1}}\right\} > 0$. Then by definition of inf, there exists a $t \in U$ such that
   $$z \leq t < z + \varepsilon \leq z + 1.$$
   Then by $(\star\star)$
   $$t^n - z^n \leq (t - z)nt^{n-1} < (t - z)n(1 + z)^{n-1}.$$
   Since $t - z < \varepsilon$, we have
   $$t^n - z^n < \left(\frac{x - z^n}{n(1 + z)^{n-1}}\right) n(1 + z)^{n-1} = x - z^n.$$
   This implies that $t^n < x$, but $t \in U \implies t^n > x$, a contradiction. □

4. $t^n \geq x \iff t \geq z$

   *Proof.* ($\impliedby$) Let $t \geq z$. Then $t^n \geq z^n \geq x$ by claim 3.

   ($\implies$) Let $t^n \geq x$. If $t^n > x$, then $t \in U \implies t \geq z = \inf U$. If $t^n = x$, then $w^n > x$ for all $w > t$. Thus $(t, \infty) \subseteq U \implies t = \inf(t, \infty) \geq \inf U = z$. □

We now claim that $y = z$. Suppose that $y \neq z$, then $y < z$. Let $u \in \mathbb{R}$ be such that $y < u < z$ (we know we can always find such a $u \in \mathbb{Q}$). Then $y < u \implies u^n > x$ by claim 2. Then $u \in U$, so $u \geq \inf U = z$, a contradiction. Thus, $y = z$. Now, by claims 1 and 3, $y^n \leq x \leq z^n \implies y^n = z^n = x$. □

Now that we know that such an equation has a solution, we will show that not all such equations can be solved by rational solutions.

**Definition 1.13.** Let $a, b \in \mathbb{Z}$. We say that $a$ divides $b$ if $\exists m \in \mathbb{Z}$ such that $b = ma$.

**Theorem 1.1** (Rational Roots Theorem)**.** *Let $a_0, a_1, \ldots, a_n \in \mathbb{Z}$ with $a_n \neq 0$. If*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

*has a rational solution, $x = p/q$, where $p \in \mathbb{Z}, q \in \mathbb{N}$, and $gcd(p, q) = 1$, then*

1. *$q$ divides $a_n$.*
2. *$p$ divides $a_0$.*

*Proof.* Suppose $x = p/q$, where $p \in \mathbb{Z}, q \in \mathbb{N}$, and $\gcd(p, q) = 1$ is a solution of

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

where $a_n \neq 0$. Then

$$a_n \left(\frac{p}{q}\right)^n + \cdots + a_1 \left(\frac{p}{q}\right) + a_0 = 0.$$

Multiplying both sides by $q^n$, we get

$$a_n p^n + \cdots + a_1 p q^{n-1} + a_0 q^n = 0,$$

so

$$-a_n p^n = q(a_{n-1} p^{n-1} + \cdots + a_1 p q^{n-2} + a_0 q^{n-1}).$$

Then $q$ divides either $a_n$ or $p$. Since $p$ and $q$ share no common factors, $q$ must divide $a_n$. Similarly,

$$-a_0 q^n = p(a_n p^{n-1} + \cdots + a_1 q^{n-1}),$$

so $p$ divides either $a_0$ or $q$. Since $\gcd(p, q) = 1$, $p$ must divide $a_0$. □

**Example 1.5.** Consider the equation $x^2 - 2 = 0$. $a_2 = 1$ and $a_0 = 2$, so if $x = p/q$ with $gcd(p, q) = 1$ is a solution of this equation, then $p$ must divide 2, which means $p$ must be one of $\pm 1$ or $\pm 2$, and $q$ must divide 1, which means $q$ must be $\pm 1$. This means the only possible rational solutions to this equation are $x = \pm 2, \pm 1$. We can easily check that none of these are solutions to this equation. Thus, the solution $x = \sqrt{2}$ must be a member of the irrational numbers, $\mathbb{R} \setminus \mathbb{Q}$.

**Proposition 1.8.** *The irrationals are dense in $\mathbb{R}$. $\forall a, b \in \mathbb{R}$ with $a < b$, $\exists x \in \mathbb{R} \setminus \mathbb{Q}$ such that $a < x < b$.*

*Proof.* Let $a, b \in \mathbb{R}$ with $a < b$. Then $a + \sqrt{2} < b + \sqrt{2}$. By the density of $\mathbb{Q}$, $\exists q \in \mathbb{Q}$ such that $a + \sqrt{2} < q < b + \sqrt{2} \implies a < q - \sqrt{2} < b$. We now claim that $q - \sqrt{2}$ is an irrational number. Suppose that $r = q - \sqrt{2} \in \mathbb{Q}$. Then $\sqrt{2} = q - r \in \mathbb{Q}$, but we know $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$, a contradiction. □

## 1.4 Applications in Probability and Statistics

### 1.4.1 Sigma Algebras

### 1.4.2 Measurable Mappings

# Chapter 2

# Sequences, Series, and Basic Topology

## 2.1 Sequences

**Definition 2.1.** A **sequence** in $\mathbb{R}$ is a function $f : \mathbb{N} \mapsto \mathbb{R}$. We write $x_n = f(n)$ for the $n$th term of the sequence and denote the whole sequence by $\{x_n, n \geq 1\}$, $\{x_n\}_{n=1}^{\infty}$, or $(x_n)_n$. The **range** of the sequence is the set of values attained by the sequence: $\{x \in \mathbb{R} | x = x_n \text{ for some } n \in \mathbb{N}\}$.

**Example 2.1.** The following are examples of real sequences:

a) $x_n = \dfrac{3n-5}{2n+3}$, $n \geq 1$. $\{x_n, n \geq 0\} = \{-\frac{2}{5}, \frac{1}{2}, \frac{4}{9}, \ldots\}$

b) $x_n = (-1)^n,\ n \geq 1$. $\{x_n, n \geq 1\} = \{-1, 1, -1, 1, -1, \ldots\}$

c) $x_n = \cos(n\pi/3),\ n \geq 0$. $\{x_n, n \geq 0\} = \{1, 1/2, -1/2, -1. -1/2, 1/2.1, \ldots\}$.

**Definition 2.2.** A real sequence $\{x_n, n \geq 1\}$ is said to converge to $x \in \mathbb{R}$ if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that

$$n \geq N \implies |x_n - x| < \varepsilon.$$

In such a case, we write $\lim\limits_{n \to \infty} x_n = x$, $x_n \to x$ as $n \to \infty$, or just $x_n \to x$ when it is clear that we are letting $n \to \infty$..

**Example 2.2.** Prove that $\lim\limits_{n \to \infty} \frac{1}{n} = 0$.

*Solution.* Let $\varepsilon > 0$. By the Archimedean property, $\exists N \in \mathbb{N}$ such that $N > \frac{1}{\varepsilon} \iff \frac{1}{N} < \varepsilon$. Then for $n \geq N$

$$0 < \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Thus, $n \geq N \implies |1/n - 0| < \varepsilon$. □

**Example 2.3.** Prove that $\lim\limits_{n \to \infty} \frac{3n-5}{2n+3} = \frac{3}{2}$.
Before we write the formal argument, let's work backwards to think about how it should go. We need to show that for all $\varepsilon < 0$

$$\left| \frac{3n-5}{2n+3} - \frac{3}{2} \right| < \varepsilon$$

for all large $n$. To see this, note that

$$\begin{aligned}\left|\frac{3n-5}{2n+3}-\frac{3}{2}\right| &= \left|\frac{6n-10-6n-9}{2(2n+3)}\right| \\ &= \left|\frac{-19}{2(2n+3)}\right| \\ &\leq \frac{20}{2(2n+3)} = \frac{10}{2n+3} \\ &\leq \frac{5}{n}\end{aligned}$$

Now we are ready to write the formal proof.

*Solution.* Let $\varepsilon > 0$. By the Archimedean property, choose an $N \in \mathbb{N}$ such that $\frac{1}{N} < \frac{\varepsilon}{5}$. Then by work above, $n \geq N$ implies

$$\left|\frac{3n-5}{2n+3}-\frac{3}{2}\right| \leq \frac{5}{n} \leq \frac{5}{N} < 5 \cdot \frac{\varepsilon}{5} = \varepsilon.$$

□

We next prove that limits are unique.

**Proposition 2.1.** *Let $\{x_n, n \geq 1\}$ be a real sequence and let $x, y \in \mathbb{R}$. If the $x_n \to x$ and $x_n \to y$ as $n \to \infty$, then $x = y$.*

*Proof.* First, note that $x = y \iff |x-y| = 0 \iff |x-y| < \varepsilon$, $\forall \varepsilon > 0$. Let $\varepsilon > 0$. Since $x_n \to x$, $\exists N_1 \in \mathbb{N}$ such that $n \geq N_1 \implies |x_n - x| < \varepsilon/2$, and since $x_n \to y$, $\exists N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |x_n - y| < \varepsilon/2$. Let $N = \max\{N_1, N_2\}$, then $n \geq N$ implies

$$|x-y| = |x - x_n + x_n - y| \leq |x_n - x| + |x_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

**Example 2.4.** Show that the sequence $\{(-1)^n, n \geq 1\}$ does not converge.

*Solution.* Suppose that $(-1)^n \to a$ for some $a \in \mathbb{R}$. Let $\varepsilon = 1$. Then $\exists N \in \mathbb{N}$ such that $n \geq N \implies |-(1)^n - a| < 1$. This implies that

$$|1-a| < 1 \text{ and } |-1-a| < 1.$$

Then

$$2 = |1+1| = |1-a+a+1| \leq |1-a| + |-1-a| < 1+1 = 2,$$

a contradiction. Thus $(-1)^n \not\to a$ for any $a \in \mathbb{R}$. □

**Proposition 2.2.** *Let $\{x_n, n \geq 1\}$ be a real sequence and let $x \in \mathbb{R}$:*

*a) If $\exists N \in \mathbb{N}$ such that $x_n \geq 0$ for all $n \geq N$ and $x_n \to x$, then $x \geq 0$.*

*b) $x_n \to x \iff |x_n - x| \to 0$.*

*Proof.* a) Suppose $x_n \geq 0$ for all $n \geq N$ and that $x_n \to x$. Suppose that $x < 0$. Let $\varepsilon = |x|$. Then $n \geq N$ implies

$$|x_n - x| = x_n - x = x_n + |x| \geq |x| = \varepsilon.$$

Thus, $x_n \not\to x$, a contradiction.

b) ( $\Longrightarrow$ ) Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon$. Then $||x_n - x| - 0| = |x_n - x| < \varepsilon$ for all $n \geq N$. Since $\varepsilon > 0$ was arbitrary, $|x_n - x| \to 0$.

( $\Longleftarrow$ ) Same as the above argument.

□

**Proposition 2.3** (Squeeze Theorem)**.** *Let $\{a_n, n \geq 1\}$ and $\{b_n, n \geq 1\}$ be real sequences such that $a_n \to 0$ and $|b_n| \leq |a_n|$ for all $n \geq N$ for some $N \in \mathbb{N}$. Then $b_n \to 0$.*

*Proof.* Let $\varepsilon > 0$. Choose $N_1 \geq N$ such that $|a_n| < \varepsilon$. Then $n \geq N_1$ implies

$$|b_n| \leq |a_n| < \varepsilon.$$

□

**Example 2.5.** Show that $\lim_{n\to\infty} \frac{2n+3}{5n-7} = \frac{2}{5}$.

*Solution.* First note that,

$$\begin{aligned} \left|\frac{2n+3}{5n-7} - \frac{2}{5}\right| &= \left|\frac{10+15-10+14}{5(5n-7)}\right| \\ &= \left|\frac{29}{5(5n-7)}\right|. \end{aligned}$$

Now $5n - 7 > n \iff 4n - 7 > 0 \iff n > 7/4$. Thus for $n \geq 2$, $5n - 7 > n$ and

$$\left|\frac{2n+3}{5n-7} - \frac{2}{5}\right| = \left|\frac{29}{5(5n-7)}\right| < \frac{30}{5n} = \frac{6}{n}.$$

Since $6/n \to 0$,

$$\left|\frac{2n+3}{5n-7} - \frac{2}{5}\right| \to 0$$

by the squeeze theorem. □

**Definition 2.3.** A sequence $\{x_n, n \geq 1\}$ is said to be **bounded** if $\exists M > 0$ such that $|x_n| \leq M$ for all $n \in \mathbb{N}$. That is $-M \leq x_n \leq M$ for all $n \geq 1$. Otherwise, the sequence is unbounded.

**Example 2.6.** Determine if the following sequences are bounded:

a) $x_n = \frac{3n-5}{2n+3}$, $n \geq 1$.

*Solution.* Note that for all $n \geq 1$

$$\begin{aligned} |x_n| &= \left|\frac{3n-5}{2n+3}\right| = \left|\frac{3-5/n}{2+3/n}\right| \\ &\leq \frac{|3|+|5/n|}{2+3/n} \quad \text{(by the triangle inequality)} \\ &\leq \frac{3+5}{2} = 4. \end{aligned}$$

Thus, $|x_n| \leq 4$ for all $n \geq 1$, so $\{x_n, n \geq 1\}$ is a bounded sequence. □

b) $x_n = n,\ n \geq 1$.

*Solution.* Let $M > 0$. By the Archimedean property, there exists and $n \in \mathbb{N}$ such that $|x_n| = n > M$. Since $M$ was arbitrary, $\{x_n\}_n$ is not a bounded sequence. □

**Proposition 2.4.** *Every real convergent sequence is bounded.*

*Proof.* Let $\{x_n, n \geq 1\}$ be a real sequence such that $x_n \to x \in \mathbb{R}$. Let $N \in \mathbb{N}$ be such that $|x_n - x| < 1$. Then for $n \geq N$

$$|x_n| = |x_n - x + x| \leq |x_n - x| + |x| < 1 + |x|.$$

Let $M = \max\{|x| + 1, |x_1|, |x_2|, \dots, |x_{N-1}|\}$. Then $|x_n| \leq M$ for all $n \geq 1$. That is $\{x_n, n \geq 1\}$ is bounded. □

**Proposition 2.5.** *If $x_n \to 0$ and $(y_n)_n$ is a bounded sequence, then $x_n y_n \to 0$.*

*Proof.* Let $\varepsilon > 0$ and let $M > 0$ be such that $|y_n| \leq M$ for all $n \geq 1$. Choose $N \in \mathbb{N}$ such that $|x_n| < \varepsilon/M$. Then for $n \geq N$

$$|x_n y_n| \leq M|x_n| < M \cdot \frac{\varepsilon}{M} = \varepsilon.$$

□

## 2.2 Limit Theorems for Sequences

**Proposition 2.6.** *Let $(x_n)_n$ and $(y_n)_n$ be real sequences and let $x, y \in \mathbb{R}$. Suppose $x_n \to x$ and $y_n \to y$. Then*

*a)* $x_n + y_n \to x + y$

*b)* $x_n y_n \to xy$

*c)* $cx_n \to cx$ *and* $x_n + c \to x + c$ *for all* $c \in \mathbb{R}$

*d)* $1/x_n \to 1/x$ *provided* $x_n \neq 0$ *for all* $n \geq 1$ *and* $x \neq 0$.

*Proof.* a) Let $\varepsilon > 0$. Choose $N_1 \in \mathbb{N}$ such that $n \geq N_1 \implies |x_n - x| < \varepsilon/2$, and choose $N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |y_n - y| < \varepsilon/2$. Let $N = \max\{N_1, N_2\}$. Then $n \geq N$ implies

$$|(x_n + y_n) - (x + y)| = |x_n - x + y_n - y| \leq |x_n - x| + y_n - y| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

b) Let $\varepsilon > 0$. Since $y_n \to y$, $(y_n)_n$ is a bounded sequence. Let $M > 0$ be such that $|y_n| \leq M$. Let $N_1 \in \mathbb{N}$ be such that $n \geq N_1 \implies |x_n - x| < \frac{\varepsilon}{2M}$, and let $N_2 \in \mathbb{N}$ be such that $n \geq N_2 \implies |y_n - y| < \frac{\varepsilon}{2(|x|+1)}$. Then for $n \geq \max\{N_1, N_2\}$

$$\begin{aligned} |x_n y_n - xy| &= |x_n y_n - x y_n + x y_n - xy| \\ &\leq |y_n||x_n - x| + |x||y_n - y| \\ &< M \cdot \frac{\varepsilon}{2M} + |x|\frac{\varepsilon}{2(|x|+1)} \\ &< \varepsilon. \end{aligned}$$

c) Let $z_n = c$ for all $n \geq 1$. Then $z_n \to c$ and the results follow from parts a) and b).

d) Let $\varepsilon > 0$. Let $N_1 \in \mathbb{N}$ be such that $|x_n - x| < |x|/2$. Then $|x_n| > |x|/2$ for all $n \geq N_1$. To see this, suppose by contradiction that $|x_n| \leq |x|/2$ for $n \geq N_1$. Then

$$|x| = |x - x_n + x_n| \leq |x - x_n| + |x_n| < \frac{|x|}{2} + \frac{|x|}{2} = |x|$$

which cannot happen. Now, choose $N_2 \in \mathbb{N}$ such that $n \geq N_2 \implies |x_n - x| < |x|^2\varepsilon/2$. Then, for $n \geq \max\{N_1, N_2\}$, we have

$$\left|\frac{1}{x_n} - \frac{1}{x}\right| = \left|\frac{x_n - x}{x_n \cdot x}\right| = \frac{|x_n - x|}{|x_n||x|} < \frac{2}{|x|^2} \cdot \frac{|x|^2\varepsilon}{2} = \varepsilon.$$

□

**Corollary 2.1.** *Suppose $x_n \leq y_n$ for all $n \geq N$ for some $N \in \mathbb{N}$ and $x_n \to x$ and $y_n \to y$. Then $x \leq y$.*

*Proof.* Let $z_n = y_n - x_n$. Then $z_n \geq 0$ for all $n \geq N$ and $z_n \to y - x$. Then by a previous proposition, $y - x \geq 0$, so $y \geq x$. □

**Note 2.1.** It is not true that $x_n < y_n$ for all $n \geq 1$ and $x_n \to x$ and $y_n \to y$ implies that $x < y$. Consider $x_n = 0$ for all $n$ and $y_n = 1/n$. Then $x_n = 0 < 1/n = y_n$ for all $n$, but $\lim_{n\to\infty} x_n = 0 = \lim_{n\to\infty} y_n$.

The following proposition contains some special limits.

**Proposition 2.7.**

*a)* $\lim_{n\to\infty} \frac{1}{n^p} = 0$ *for all* $p > 0$.

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

*c)* $\lim_{n\to\infty} n^{1/n} = 1$.

*d)* $\lim_{n\to\infty} a^{1/n} = 1$ *for* $a > 0$.

*Proof.* a) Let $\varepsilon > 0$. Let $N \in \mathbb{N}$ be such that $N > 1/\varepsilon^{1/p}$ (such an $N$ exists by the Archimedean property). Then $n \geq N$ implies

$$\left|\frac{1}{n^p} - 0\right| = \frac{1}{n^p} \leq \frac{1}{N^p} < \varepsilon.$$

b) If $a = 0$, then $a^n = 0^n = 0$ for all $n$ and clearly $\lim_{n\to\infty} 0 = 0$. Suppose $0 < |a| < 1$. Then $|a| = \frac{1}{1+b}$ for some $b > 0$. (In particulare, $b = \frac{1}{|a|} - 1 > 0$.) Then $|a|^n = 1/(1+b)^n$. By the Binomial theorem

$$(1+b)^n = \sum_{k=0}^{n} \binom{n}{k} b^k = 1 + nb + \cdots + b^n \geq nb.$$

Then

$$|a^n| = \left|\frac{1}{(1+b)^n}\right| \leq \frac{1}{nb}.$$

Since $\frac{1}{nb} \to 0$, the result follows by the squeeze theorem.

c) Let $s_n = n^{1/n} - 1$. Note that $s_n \geq 0$ for all $n \geq 1$, since $n^{1/n} \geq 1 \iff n \geq 1^n = 1$. Then

$$n = (n^{1/n})^n = (1+s_n)^n = \sum_{k=0}^{n} \binom{n}{k} s_n^k \geq \frac{n(n-1)}{2} s_n^2,$$

so

$$0 \leq s_n \leq \sqrt{\frac{2}{n-1}}.$$

Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $N > 1 + (2/\varepsilon^2)$. Then $n \geq N$ implies

$$|n^{1/n} - 1| = s_n < \varepsilon.$$

d) Let $a > 1$. Then for all $n \geq a$, $1 < a \leq n \implies 1 < a^{1/n} \leq n^{1/n}$, so that

$$0 \leq a^{1/n} - 1 \leq n^{1/n} - 1 \to 0.$$

Thus by the squeeze theorem, $a^{1/n} \to 1$. If $0 < a < 1$, then

$$\lim_{n\to\infty} a^{1/n} = \lim_{n\to\infty} \frac{1}{(1/a)^{1/n}} = \frac{1}{\lim_{n\to\infty}(1/a)^{1/n}} = 1,$$

since $1/a > 1$. The result is clear if $a = 1$.

□

## 2.3 Infinite Limits

**Definition 2.4.** Let $(x_n)_n$ be a real sequence. We say that $\lim_{n\to\infty} x_n = \infty$ if $\forall M > 0$, $\exists N \in \mathbb{N}$ such that $x_n \geq M$ for all $n \geq N$. We say that $\lim_{n\to\infty} x_n = -\infty$ if $\forall M < 0$, $\exists N \in \mathbb{N}$ such that $x_n \leq M$ for all $n \geq N$.

**Note 2.2.** We will now say that a sequence **converges** if it has a finite limit. We will say that the limit **exists** if it either converges or diverges to $+\infty$ or $-\infty$. Otherwise, the limit does not exist.

**Example 2.7.** Determine if the following sequences are bounded or unbounded and whether or not the limit exists.

a) $x_n = n^2$. This sequence is unbounded and $\lim_{n\to\infty} n^2 = \infty$.

b) $x_n = -n$. This sequence is unbounded and $\lim_{n\to\infty}(-n) = -\infty$.

c) $x_n = (-1)^n$. This sequence is bounded, but the limit does not exist.

d) $x_n = n\cos^2(n\pi/2)$. This sequence is unbounded and the limit does not exist.

**Example 2.8.** Prove that $\lim_{n\to\infty}(\sqrt{n}+7) = \infty$.

*Solution.* Let $0 < M \leq 8$. Note that $\{x_n = \sqrt{n}+7, n \geq 1\}$ is bounded below by 8, so $\sqrt{n}+7 \geq M$ for all $n \geq 1$. For $M > 8$, note that

$$\sqrt{n}+7 \geq M \iff n \geq (M-7)^2.$$

Let $M > 8$. Choose $N \in \mathbb{N}$ such that $N \geq (M-7)^2$. Then $n \geq N$ implies

$$\sqrt{n}+7 \geq \sqrt{N}+7 \geq \sqrt{(M-7)^2}+7 = M.$$

□

**Proposition 2.8.** *Let $\{x_n, n \geq 1\}$ and $\{y_n, n \geq 1\}$ be a real sequences.*

*a) If $x_n > 0$ and $x_n \to \infty$, then $1/x_n \to 0$.*

*b) If $x_n > 0$ and $x_n \to 0$, then $1/x_n \to \infty$.*

*c) If $x_n \to \infty$ and $(y_n)_n$ is bounded, then $x_n + y_n \to \infty$.*

*d) If $x_n \to \infty$ and $y_n \to y \neq 0$, then*

$$x_n y_n \to \begin{cases} \infty, & \text{if } y > 0 \\ -\infty, & \text{if } y < 0. \end{cases}$$

*Proof.*

a) Let $\varepsilon > 0$. Since $x_n \to \infty$, there exists an $N \in \mathbb{N}$ such that $n \geq N \implies x_n > 1/\varepsilon$. Then $n \geq N \implies 0 \leq 1/x_n < \varepsilon$.

b) Let $M > 0$. Since $x_n \to 0$, there exists an $N \in \mathbb{N}$ such that $n \geq N \implies 0 \leq x_n = |x_n| < 1/M$. Thus $n \geq N \implies M < 1/x_n$.

c) Let $K > 0$ be such that $|y_n| \leq K$ for all $n \in \mathbb{N}$. Then $y_n \geq -K$ for all $n \geq 1$. Let $M > 0$. Since $x_n \to \infty$, we can choose an $N \in \mathbb{N}$ such that $n \geq N \implies x_n \geq M + K$. Then $n \geq N$ implies

$$x_n + y_n \geq (M + K) - K = M.$$

d) First, suppose $y > 0$. Let $M > 0$ and choose $N_1 \in \mathbb{N}$ such that $n \geq N$ implies

$$|y_n - y| < y/2 \iff y/2 < y_n < 3y/2.$$

Let $N_2 \in \mathbb{N}$ be such that $n \geq N_2 \implies x_n \geq 2M/y$. Then $n \geq \max\{N_1, N_2\}$ implies

$$x_n y_n \geq \frac{2M}{y} \cdot \frac{y}{2} = M.$$

The case where $y < 0$ is similar and left as an exercise.

□

## 2.4 Monotonic Sequences

**Definition 2.5.** A real sequence $\{x_n, n \geq 1\}$ is said to be **nondecreasing** (or increasing) provided $x_n \leq x_{n+1}$ for all $n \geq 1$, and $(x_n)_n$ is said to be **nonincreasing** (or decreasing) if $x_n \geq x_{n+1}$ for all $n \geq 1$. In either case, we call the sequence **monotonic**.

**Example 2.9.**

a) $(1/n)_n$ is a decreasing sequence.

b) $\left(\frac{n}{n+1}\right)_n$ is an increasing sequence.

c) $\left(n + \dfrac{1}{n}\right)_n$ is an increasing sequence. To see this, let $f(x) = x + 1/x$, then $f'(x) = 1 - 1/x^2 > 0$ for all $x > 1$.

d) $(-1)^n$ is not monotonic.

**Proposition 2.9.** *Let $(x_n)_n$ be a real sequence.*

*a) If $(x_n)_n$ is increasing, then $x_n \to \sup_{n\geq 1}\{x_n\}(\leq \infty)$.*

*b) If $(x_n)_n$ is decreasing, then $x_n \to \inf_{n\geq 1}\{x_n\}(\geq -\infty)$.*

*c) If $(x_n)_n$ is monotonic, then $(x_n)_n$ is convergent if and only if $(x_n)_n$ is bounded.*

*Proof.* Let $(x_n)_n$ be a real sequence.

a) First suppose that $(x_n)_n$ is bounded above so that $c = \sup_{n\geq 1}\{x_n\} < \infty$. Let $\varepsilon > 0$. Then by definition of supremum, $\exists N \in \mathbb{N}$ such that

$$c - \varepsilon < x_N \leq c.$$

Since $(x_n)_n$ is increasing, $n \geq N \implies c - \varepsilon < x_N \leq x_n \leq c$, so

$$|x_n - c| < \varepsilon \text{ for all } n \geq N.$$

Now suppose that $(x_n)_n$ is not bounded above, so that $\sup_{n\geq 1}\{x_n\} = \infty$. Let $M > 0$ and let $N \in \mathbb{N}$ be such that $x_N \geq M$ (such and $x_N$ exists since $(x_n)_n$ is not bounded above). Then for $n \geq N$, $x_n \geq x_N \geq M$, since $(x_n)_n$ is increasing. Thus $x_n \to \infty = \sup_{n\geq 1}\{x_n\}$.

b) The proof is similar to a) and left as an exercise.

c) Suppose $(x_n)_n$ is monotonic. We proved earlier that if $(x_n)_n$ is convergent, then it is bounded. Suppose that $(x_n)_n$ is bounded. Since $(x_n)_n$ is monotonic, it is either increasing or decreasing. If $(x_n)_n$ is increasing, then by part *a)* $x_n \to \sup_{n\geq 1}\{x_n\} < \infty$. If $(x_n)_n$ is decreasing, then by part *b)* $x_n \to \inf_{n\geq 1}\{x_n\} > -\infty$.

□

**Example 2.10.** Let $x_1 = 1$ and define $x_n = \sqrt{3x_{n-1}}$ for $n \geq 2$. Show that $(x_n)_n$ converges and find its limit.

*Solution.* Note that $x_2 = \sqrt{3 * 1} \geq 1 = x_1$. Now suppose that $x_k \leq x_{k+1}$ for some $k \in \mathbb{N}$. Then

$$x_{k+1} = \sqrt{3x_k} \leq \sqrt{3x_{k+1}} = x_{k+2}.$$

Then by the prinicple of mathematical induction, we have that $x_n \leq x_{n+1}$ for all $n \geq 1$. That is $(x_n)_n$ is an increasing sequence.

Next, we need to show that $(x_n)_n$ is bounded above. Note that $x_1 = 1 < 10$. Suppose that $x_k < 10$ for some $k \in \mathbb{N}$. Then

$$x_{k+1} = \sqrt{3x_k} < \sqrt{3 * 10} = \sqrt{30} < 10.$$

So by induction $x_n < 10$ for all $n \in \mathbb{N}$.

Since $(x_n)_n$ is increasing and bounded above, $(x_n)_n$ is convergent and $x_n \to \sup_{n\geq 1}\{x_n\}$. Let $c = \sup_{n\geq 1}\{x_n\}$. Then $x_n \to c$ implies $\sqrt{3x_{n-1}} \to \sqrt{3c}$. Since $x_n = \sqrt{3x_{n-1}}$ for all $n \geq 2$, we must have

$$c = \sqrt{3c} \implies c^2 - 3c = c(c-3) = 0,$$

so $c = 0$ or $c = 3$. Since $x_1 = 1$ and $(x_n)_n$ is increasing, $c = 3$. □

**Example 2.11.** Let $A > 0$ and let $x_1 = 1$. Define $x_n = \frac{1}{2}\left(x_{n-1} + \dfrac{A}{x_{n-1}}\right)$ for all $n \geq 2$. Show that $x_n \to \sqrt{A}$.

*Solution.* We will need to use the Cauchy inequality: If $a, b > 0$, then $a + b \geq 2\sqrt{ab}$. We will first show that $x_n$ is bounded below by $\sqrt{A}$ for $n \geq 2$. Note that for $n \geq 2$

$$x_n = \frac{1}{2}\left(x_{n-1} + \frac{A}{x_{n-1}}\right) \geq \frac{1}{2}2\sqrt{x_{n-1}\frac{A}{x_{n-1}}} = \sqrt{A}.$$

Next, note that for $n \geq 2$

$$x_n - x_{n+1} = x_n - \frac{1}{2}\left(x_n + \frac{A}{x_n}\right) = \frac{x_n^2 - A}{2x_n} \geq 0.$$

Thus, $(x_n)_n$ is a decreasing sequence that is bounded below, so $\lim_{n\to\infty} x_n = c$ exists and is finite. Therefore,

$$\lim_{n\to\infty} x_n = \frac{1}{2}\left(\lim_{n\to\infty} x_{n-1} + \frac{A}{\lim_{n\to\infty} x_{n-1}}\right) \implies c = \frac{1}{2}\left(c + \frac{A}{c}\right),$$

so we get the equation

$$c^2 = A \implies c = \sqrt{A},$$

since we know $c \geq \sqrt{A}$. □

## 2.5 Limit Superior and Limit Inferior

**Definition 2.6.** Let $(x_n)_n$ be a real sequence. The **limit superior** is defined by

$$\overline{\lim_{n\to\infty}}\, x_n = \limsup_{n\to\infty} x_n = \lim_{n\to\infty} \sup_{k\geq n} x_k.$$

The **limit inferior** is defined by

$$\underline{\lim_{n\to\infty}}\, x_n = \liminf_{n\to\infty} x_n = \lim_{n\to\infty} \inf_{k\geq n} x_k.$$

**Note 2.3.** The limit superior and inferior always exists. To see this, note that for all $n \geq 1$

$$\sup_{k\geq n} x_k \geq \sup_{k\geq n+1} x_k \quad \text{and} \quad \inf_{k\geq n} x_k \leq \inf_{k\geq n+1} x_k.$$

That is $\{\sup_{k\geq n} x_k\}_{n=1}^{\infty}$ is a decreasing sequence and $\{\inf_{k\geq n} x_k\}_{n=1}^{\infty}$ is an increasing sequence. Thus the limit superior and inferior always exist (possible infinite) by Proposition 2.9. Furthermore, since $\inf_{k\geq n} x_k \leq \sup_{k\geq n} x_k$ for all $n \geq 1$. It is clear that

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$$

is always true.

**Example 2.12.** Let $x_n = (-1)^n + \frac{1}{n}$ for $n \geq 1$. Then $\{x_n, n \geq 1\} = \{0, 3/2, -2/3, 5/4, \ldots\}$. Note that the terms with an even index are

$$x_{2k} = 1 + \frac{1}{2k} = \frac{2k+1}{2k}$$

which are decreasing to 1, and the terms with an odd index are

$$x_{2k+1} = -1 + \frac{1}{2k+1} = \frac{-2k}{2k+1},$$

which are decreasing to $-1$. Then

$$\sup_{k\geq n} x_k = \begin{cases} 1 + \frac{1}{2n}, & \text{if } n \text{ is even} \\ 1 + \frac{1}{2(n+1)}, & \text{if } n \text{ is odd.} \end{cases} \quad \text{and} \quad \inf_{k\geq n} x_k = -1.$$

Thus $\overline{\lim}_{n\to\infty} x_n = 1$ and $\underline{\lim}_{n\to\infty} = -1$.

**Proposition 2.10.** *Let $(x_n)_n$ be a bounded sequence in reals. Then*

*a) $\beta = \overline{\lim\limits_{n\to\infty}}\, x_n$ if and only if*

- *(i) $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $x_n < \beta + \varepsilon$, $\forall n \geq N$.*
- *(ii) $\forall \varepsilon > 0$ and $\forall n \in \mathbb{N}$, $\exists k \geq n$ such that $\beta - \varepsilon < x_k$.*

*b) $\alpha = \underline{\lim\limits_{n\to\infty}}\, x_n$ if and only if*

- *(i) $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\alpha - \varepsilon < x_k$, $\forall n \geq N$.*
- *(ii) $\forall \varepsilon > 0$ and $\forall n \in \mathbb{N}$, $\exists k \geq n$ such that $x_k < \alpha + \varepsilon$.*

*Proof.* Let $t_n = \sup_{k\geq n} x_k,\ n \geq 1$. Note that $(t_n)_n$ is a decreasing sequence, so $t_n \to \inf_{n\geq 1} t_n = \beta$. Then

$$\beta = \inf_{n\geq 1} t_n \iff \begin{cases} \forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } t_N < \beta + \varepsilon \\ \forall \varepsilon > 0,\ \beta - \varepsilon < t_n,\ \forall n \in \mathbb{N} \end{cases}$$

$$\iff \begin{cases} \forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } x_k < \beta + \varepsilon\ \forall k \geq N \\ \forall \varepsilon > 0,\ \exists k \geq n \text{ such that } \beta - \varepsilon < x_k,\ \forall n \in \mathbb{N} \end{cases}$$

To see why, we can argue as follows. ( $\Longrightarrow$ ) To prove (i), let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $|t_n - \beta| < \varepsilon$ for a all $n \geq N$. This implies that for $n \geq N$

$$x_n \leq \sup_{k\geq N} x_k = t_N < \beta + \varepsilon.$$

To prove (ii), let $\varepsilon > 0$ and let $n \in \mathbb{N}$. Then

$$\beta - \varepsilon < \beta = \inf_{k\geq 1} t_k \leq t_n = \sup_{k\geq n} x_k.$$

By definition of supremum, $\exists k \geq n$ such that

$$\beta - \varepsilon < x_k \leq \sup_{j\geq n} x_j.$$

( $\Longleftarrow$ ) To see the reverse direction. Let $\varepsilon > 0$. Then by (i), we can choose $N \in \mathbb{N}$ such that $k \geq N$ implies

$$x_k < \beta + \frac{\varepsilon}{2} \implies t_n \leq t_N = \sup_{k\geq N} x_k \leq \beta + \frac{\varepsilon}{2} < \beta + \varepsilon,\ \forall n \geq N.$$

Note that by definition of supremum, $x_k \leq \sup_{j\geq n} x_j$ whenever $k \geq n$, so (ii) implies that $\beta - \varepsilon < t_n$ for all $n \in \mathbb{N}$. Thus, we have for $n \geq N$, $|t_n - \beta| < \varepsilon$.
The proof of part b) is similar and is left as an exercise. □

**Theorem 2.1.** *Let $(x_n)_n$ be a sequence in $\mathbb{R}$. The limit of $(x_n)_n$ exists if and only if $\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n$. In such a case,*

$$\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = \lim_{n\to\infty} x_n$$

*Proof.* ( $\Longrightarrow$ ) Suppose that $\lim_{n\to\infty} x_n$ exists. Then we need to consider three cases: (i) $x_n \to x \in \mathbb{R}$, (ii) $x_n \to \infty$ (iii) $x_n \to -\infty$.
Let $t_n = \sup_{k\geq n} x_k$ and $s_n = \inf_{k\geq n} x_k$ for $n \geq 1$.

Case i: Suppose $x_n \to x \in \mathbb{R}$. Let $\varepsilon > 0$. Then by the definition of a limit, we can choose an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|x_n - x| < \frac{\varepsilon}{2} \iff x - \frac{\varepsilon}{2} < x_n < x + \frac{\varepsilon}{2}.$$

Then

$$x - \frac{\varepsilon}{2} \leq s_N \leq t_N \leq x + \frac{\varepsilon}{2}.$$

Since $(s_n)_n$ is an increasing sequence and $(t_n)_n$ is a decreasing sequence, we have

$$x - \frac{\varepsilon}{2} \leq s_N \leq \underline{\lim}_{n\to\infty} x_n \quad \text{and} \quad \overline{\lim}_{n\to\infty} x_n \leq t_N \leq x + \frac{\varepsilon}{2}.$$

Combining the previous two inequalities, we get

$$0 \leq \overline{\lim}_{n\to\infty} x_n - \underline{\lim}_{n\to\infty} x_n \leq (x + \frac{\varepsilon}{2}) - (x - \frac{\varepsilon}{2}) = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary,

$$\overline{\lim}_{n\to\infty} x_n = \underline{\lim}_{n\to\infty} x_n.$$

Case ii: Suppose that $x_n \to \infty$. Let $M > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $M < x_n$. Then for all $n \geq N$

$$M \leq s_N \leq s_n.$$

Thus $\lim_{n\to\infty} s_n = \underline{\lim}_{n\to\infty} x_n = \infty$. Since

$$\underline{\lim}_{n\to\infty} x_n \leq \overline{\lim}_{n\to\infty} x_n,$$

the result follows.

Case iii: Similar to case ii. I leave the proof as an exercise.

( $\Longleftarrow$ ). Again, we need to consider three cases.

Case i: Suppose $\underline{\lim}_{n\to\infty} = x_n \overline{\lim}_{n\to\infty} x_n = x \in \mathbb{R}$. Let $\varepsilon > 0$. Then $x = \overline{\lim}_{n\to\infty} x_n$ implies $\exists N_1 \in \mathbb{N}$ such that $x_n < x + \varepsilon$ for all $n \geq N_1$, and $x = \underline{\lim}_{n\to\infty} x_n$ implies $\exists N_2 \in \mathbb{N}$ such that $x_n > x - \varepsilon$ for all $n \geq N_2$. Thus for $n \geq \max\{N_1, N_2\}$

$$x - \varepsilon < x_n < x + \varepsilon \iff |x_n - x| < \varepsilon,$$

so $x_n \to x$.

Case ii: Suppose $\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = \infty$. Let $M > 0$. Then $\exists N \in \mathbb{N}$ such that $n \geq N$ implies

$$M < s_n \implies M < s_N = \inf_{k\geq N} x_k \leq x_n, \ \forall n \geq N.$$

Thus, $x_n \to \infty$.

Case iii: $\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = -\infty$. Similar to case ii. I leave the proof as an exercise

□

**Proposition 2.11.** *Let $(x_n)_n$ and $(y_n)_n$ be real sequences.*

*a) If $x_n \leq y_n$, then $\underline{\lim}_{n\to\infty} x_n \leq \underline{\lim}_{n\to\infty} y_n$ and $\overline{\lim}_{n\to\infty} x_n \leq \overline{\lim}_{n\to\infty} y_n$.*

*b) $\overline{\lim}_{n\to\infty} -x_n = -\underline{\lim}_{n\to\infty} x_n$ and $\underline{\lim}_{n\to\infty} -x_n = -\overline{\lim}_{n\to\infty} x_n$.*

*c) $\overline{\lim}_{n\to\infty} (x_n + y_n) \leq \overline{\lim}_{n\to\infty} x_n + \overline{\lim}_{n\to\infty} y_n$ with equality if either sequence converges.*

*d) $\underline{\lim}_{n\to\infty} x_n + \underline{\lim}_{n\to\infty} y_n \leq \underline{\lim}_{n\to\infty} (x_n + y_n)$ with equality if either sequence converges.*

*e) If $x_n, y_n \geq 0$ for all $n \geq 1$, then $\overline{\lim}_{n\to\infty} (x_n y_n) \leq \overline{\lim}_{n\to\infty} x_n \cdot \overline{\lim}_{n\to\infty} y_n$ with equality if either sequence converges.*

*f) If $x_n, y_n \geq 0$ for all $n \geq 1$, then $\underline{\lim}_{n\to\infty} x_n \cdot \underline{\lim}_{n\to\infty} y_n \leq \underline{\lim}_{n\to\infty} (x_n y_n)$ with equality if either sequence converges.*

*Proof.* We leave the proof as an exercise. □

**Example 2.13.** Consider the sequences $x_n = (-1)^n$ and $y_n = (-1)^{n+1} \cdot 2$ for $n \geq 1$. Then $x_n + y_n = (-1)^{n+1}$ for all $n \geq 1$, and

$$\underline{\lim}_{n\to\infty} x_n = -1, \ \underline{\lim}_{n\to\infty} y_n = -2, \ \overline{\lim}_{n\to\infty} x_n = 1, \text{ and } \overline{\lim}_{n\to\infty} y_n = 2.$$

Thus

$$-3 = \underline{\lim}_{n\to\infty} x_n + \underline{\lim}_{n\to\infty} y_n < -1 = \underline{\lim}_{n\to\infty} (x_n + y_n) < 1 = \overline{\lim}_{n\to\infty} (x_n + y_n) < \overline{\lim}_{n\to\infty} x_n + \overline{\lim}_{n\to\infty} y_n = 3.$$

The following theorem will be useful later for tests of convergence for series.

**Theorem 2.2.** *Let $\{s_n, n \geq 1\}$ be a real sequence with $s_n > 0$ for all $n \geq 1$. Then*

$$\underline{\lim}_{n\to\infty} \frac{s_{n+1}}{s_n} \leq \underline{\lim}_{n\to\infty} s_n^{1/n} \leq \overline{\lim}_{n\to\infty} s_n^{1/n} \leq \overline{\lim}_{n\to\infty} \frac{s_{n+1}}{s_n}.$$

*Proof.* Let $\alpha = \underline{\lim}_{n\to\infty} s_{n+1}/s_n$ and $\beta = \overline{\lim}_{n\to\infty} s_{n+1}/s_n$. The first inequality clearly holds if $\alpha = 0$, and the third inequality clearly holds if $\beta = \infty$. WLOG, assume $\alpha > 0$ and $\beta < \infty$. Choose $\alpha_1, \beta_1 \in \mathbb{R}$ such that

$$0 < \alpha_1 < \alpha \text{ and } \beta < \beta_1 < \infty.$$

Then by Proposition 2.10, we can choose and $n \in \mathbb{N}$ such that $n \geq N$ implies

$$\alpha_1 < \frac{s_{n+1}}{s_n} < \beta_1.$$

Thus,

$$\alpha_1 s_N < s_{N+1} < \beta_1 s_N \text{ and } \alpha_1 s_{N+1} < s_{N+2} < \beta_1 s_{N+1}.$$

Together, these imply that

$$\alpha_1^2 s_N < s_{N+2} < \beta_1^2 s_N.$$

We now proceed by induction. Suppose that

$$\alpha_1^k s_N < s_{N+k} < \beta_1^k s_N$$

for some $k \geq 1$. Then $\alpha_1 s_{N+k} < s_{N+k+1} < \beta_1 s_{N+k}$ with the induction hypothesis implies

$$\alpha_1^{k+1} s_N = \alpha_1(\alpha_1^k s_N) < \alpha_1 s_{N+k} < s_{N+k+1} < \beta_1 s_{N+K} < \beta_1^k s_{N+k} = \beta_1^{k+1} s_N.$$

Thus by induction, $\alpha_1^k s_N < s_{N+k} < \beta_1^k s_N$ for all $k \geq 1$. So for $n > N$,

$$\alpha_1^{n-N} s_N < s_n < \beta_1^{n-N} s_N,$$

and

$$\alpha_1 \left(\frac{s_N}{\alpha_1^N}\right)^{1/n} < s_n^{1/n} < \beta_1 \left(\frac{s_N}{\alpha_1^N}\right)^{1/n}.$$

Since $s_N/\alpha_1^N > 0$ and $s_N/\beta_1^N > 0$, we have

$$\lim_{n\to\infty} \left(\frac{s_N}{\alpha_1^N}\right)^{1/n} = \lim_{n\to\infty} \left(\frac{s_N}{\beta_1^N}\right)^{1/n} = 1,$$

so

$$\alpha_1 = \lim_{n\to\infty} \alpha_1 \left(\frac{s_N}{\alpha_1^N}\right)^{1/n} \leq \varliminf_{n\to\infty} s_n^{1/n} \leq \varlimsup_{n\to\infty} s_n^{1/n} \leq \varlimsup_{n\to\infty} \beta_1 \left(\frac{s_N}{\beta_1^N}\right)^{1/n} = \beta_1.$$

Since $0 < \alpha_1 < \alpha$ and $\beta < \beta_1 < \infty$ were arbitrary, it follows that

$$\alpha \leq \varliminf_{n\to\infty} s_n^{1/n} \leq \varlimsup_{n\to\infty} s_n^{1/n} \leq \beta.$$

□

**Corollary 2.2.** *Suppose $\{s_n, n \geq 1\}$ is a reals sequence with $s_n > 0$ for all $n \geq 1$. If $\lim\limits_{n\to\infty} s_{n+1}/s_n$ exists, then $\lim\limits_{n\to\infty} s_n^{1/n}$ exists and both limits are the same.*

**Example 2.14.** Find the limit of $x_n = (1/n!)^{1/n}$. Let $s_n = 1/n!$ and note that $x_n = s_n^{1/n}$, and

$$\frac{s_{n+1}}{s_n} = \frac{1/(n+1)!}{1/n!} = \frac{1}{n+1} \to 0,$$

so by the previous corrolary, $\lim_{n\to\infty} x_n = 0$.

## 2.6 Subsequences

**Definition 2.7.** If $(x_n)_n$ is a sequence, then a subsequence of $(x_n)_n$ is a sequence of the form $\{x_{n_k}, k \geq 1\}$ where $n_k < n_{k+1}$ for all $k \in \mathbb{N}$. For this subsequence, we write $(x_{n_k})_k$.

**Note 2.4.** Because $(n_k)_k \subseteq \mathbb{N}$ and $n_k < n_{k+1}$, we have $n_k \geq k$ for all $k \in \mathbb{N}$. Given any sequence, there are infinitely many subsequences.

**Example 2.15.** The sequence of all evenly indexed elements of $(x_n)_n$, $(x_{2k})_k$, and the sequence of all oddly indexed elements of $(x_n)_n$, $(x_{2k+1})_k$ are subsequences.

**Proposition 2.12.** *Let $(x_n)_n$ be a real sequence. If the limit of $(x_n)_n$ exists and $(x_{n_k})_k$ is a subsequence, then the limit of $(x_{n_k})_k$ also exists and $\lim_{n\to\infty} x_n = \lim_{k\to\infty} x_{n_k}$.*

*Proof.* We need to consider three cases: (i) $x_n \to x \in \mathbb{R}$, (ii) $x_n \to \infty$, and (iii) $x_n \to -\infty$.

Case i: Suppose $x_n \to x \in \mathbb{R}$. Let $\varepsilon > 0$. Since $x_n \to x$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon$. Then $k \geq N \implies n_k \geq k \geq N$, so that $|x_{n_k} - x| < \varepsilon$ for all $k \geq N$.

Case ii: Suppose $x_n \to \infty$. Let $M > 0$. Since $x_n \to \infty$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies x_n \geq M$. Let $k_0 \in \mathbb{N}$ be such that $k \geq k_0 \implies n_k \geq N$. Then $k \geq k_0 \implies x_{n_k} \geq M$.

Case iii: Similar to case ii. The proof is left as an exercise. □

**Proposition 2.13.** *If $x_n \not\to x \in \mathbb{R}$, then $\exists \varepsilon > 0$ and a subsequence $(x_{n_k})_k \subseteq (x_n)_n$ such that $|x_{n_k} - x| \geq \varepsilon$ for all $k \geq 1$.*

*Proof.* Note that $x_n \not\to x \in \mathbb{R}$ if and only if $\exists \varepsilon > 0$ such that $\forall N \in \mathbb{N}\ \exists n \geq N$ such that $|x_n - x| \geq \varepsilon$. We will construct a subsequence inductively. Choose $\varepsilon > 0$ such that this holds. Let $N = 1$. Then $\exists n_1 \geq 1$ such that $|x_{n_1} - x| \geq \varepsilon$. Suppose that $k \geq 1$ and that $n_1 < n_2 < \cdots < n_k$ have been chosen so that $|x_{n_j} - x| \geq \varepsilon$ for $1 \leq j \leq k$. By hypothesis, $\exists n_{k+1} > n_k$ such that $|x_{n_{k+1}} - x| \geq \varepsilon$. Then by induction, $\exists$ a subsequence $(x_{n_k})_k$ such that $|x_{n_k} - x| \geq \varepsilon$ for all $k \geq 1$. □

**Definition 2.8.** Let $(x_n)_n$ be a sequence in $\mathbb{R}$. Then $x \in \mathbb{R} \cup \{-\infty, \infty\}$ is said to be a **subsequential limit point** of $(x_n)_n$ if $\exists$ a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $x_{n_k} \to x$.

**Note 2.5.** By Proposition 2.12, if $x_n \to x \in \mathbb{R}$, then $x$ is the only subsequential limit point.

**Example 2.16.** Let $x_n = (-1)^n$. Then the only subsequential limit points of $(x_n)_n$ are $\{-1, 1\}$. Note that $x_{2k} = (-1)^{2k} = 1 \to 1$ and $x_{2k+1} = (-1)^{2k+1} = -1 \to -1$.

**Theorem 2.3.** *A point $x \in \mathbb{R}$ is a subsequential limit point of the real sequence $(x_n)_n$ if and only if $\forall \varepsilon > 0$ and $\forall N \in \mathbb{N}$, $\exists n \geq N$ such that $|x_n - x| < \varepsilon$.*

*Proof.* ( $\implies$ ) Suppose that $x \in \mathbb{R}$ is a subsequential limit point of $(x_n)_n$. Then $\exists$ a subsequence $(x_{n_k})_k \subseteq (x_n)_n$ such that $x_{n_k} \to x$ as $k \to \infty$. Let $\varepsilon > 0$, and let $N \in \mathbb{N}$. Since $x_{n_k} \to x$, $\exists K \in \mathbb{N}$ such that $k \geq K \implies |x_{n_k} - x| < \varepsilon$. Let $k \geq \max\{K, N\}$. Then $n_k \geq k \geq N$ and $|x_{n_k} - x| < \varepsilon$.

( $\impliedby$ ) Suppose that $\forall \varepsilon > 0$ and $\forall N \in \mathbb{N}$, $\exists n \geq N$ such that $|x_n - x| < \varepsilon$. We will construct a subsequence $(x_{n_k})_k$ such that $x_{n_k} \to x$. Let $\varepsilon = 1$ and $N = 1$. Then by hypothesis, $\exists n_1 \geq N = 1$ such that $|x_{n_1} - x| < 1$. Suppose that $k \geq 1$ and that $n_1 < n_2 < \cdots < n_k$ have been chosen so that $|x_{n_j} - x| < 1/j$ for all $1 \leq j \leq k$. Then by hypothesis, $\exists n_{k+1} \geq n_k + 1 > n_k$ such that $|x_{n_{k+1}} - x| < 1/(k+1)$. Thus by the principle of mathematical induction, $\exists$ a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $|x_{n_k} - x| < 1/k$ for all $k \geq 1$, and by the squeeze theorem $x_{n_k} \to x$. □

**Corollary 2.3.** *$x \in \mathbb{R}$ is a subsequential limit point of $(x_n)_n$ if and only if $\forall \varepsilon > 0$, $(x - \varepsilon, x + \varepsilon)$ contains infinitely many terms of $(x_n)_n$.*

**Theorem 2.4** (Bolzano-Weierstrass)**.** *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

*Proof.* Let $(x_n)_n$ be a bounded sequence in $\mathbb{R}$. Then $\underline{\lim}_{n\to\infty} x_n = \alpha \in \mathbb{R}$ and $\overline{\lim}_{n\to\infty} x_n = \beta \in$. We now claim that $\alpha$ and $\beta$ are subsequential limit points of $(x_n)_n$. Let $\varepsilon > 0$. Then by Proposition 2.10., $\alpha = \underline{\lim}_{n\to\infty} x_n$ if and only if

(i) $\exists N \in \mathbb{N}$ such that $\alpha - \varepsilon < x_k$, $\forall n \geq N$

(ii) $\forall n \in \mathbb{N}$, $\exists k \geq n$ such that $x_k < \alpha + \varepsilon$.

This implies that $\forall n \geq N$, $\exists k \geq N$ such that $\alpha - \varepsilon < x_k < \alpha + \varepsilon \implies |x_k - \alpha| < \varepsilon$. Thus by Theorem 2.16, $\alpha$ is a subsequential limit point of $(x_n)_n$. A similar argument also shows that $\beta$ is a subsequential limit point of $(x_n)_n$. The proof is left as an exercise. □

We proved the following corollary for the case where $(x_n)_n$ is a bounded real sequence in the proof of the Bolzano-Weierstrass theorem, but it remains true if either the limit superior or limit inferior are $+\infty$ or $-\infty$. We omit the proof of the infinit limit cases here but it can be found in Ross.

**Corollary 2.4.** *Let $(x_n)_n$ be a sequence in $\mathbb{R}$. Then $\underline{\lim}_{n\to\infty} x_n$ and $\overline{\lim}_{n\to\infty} x_n$ are subsequential limit points of $(x_n)_n$.*

**Theorem 2.5.** *Let $(x_n)_n$ be a sequence in $\mathbb{R}$, and let*

$$S = \{x \in \mathbb{R} \cup \{-\infty, \infty\} | x \text{ is a subsequential limit point of } (x_n)_n\}.$$

*Then*

*a)* $\inf S = \underline{\lim}_{n\to\infty} x_n$ *and* $\sup S = \overline{\lim}_{n\to\infty} x_n$.

*b)* $\lim_{n\to\infty} x_n = x$ *if and only if* $S = \{x\}$.

*Proof.*

a) By Corollary 2.4, $\underline{\lim}_{n\to\infty} x_n = \alpha \in S$ and $\overline{\lim}_{n\to\infty} x_n = \beta \in S$. Let $t \in S$. Then $\exists$ a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $x_{n_k} \to t$ as $k \to \infty$. Note that $\inf_{n\geq N} x_n \leq \inf_{k\geq N} x_{n_k}$ since $n_k \geq k \implies \{x_{n_k}, k \geq N\} \subseteq \{x_n, n \geq N\}$. Thus $\alpha = \underline{\lim}_{n\to\infty} x_n \leq \underline{\lim}_{k\to\infty} x_{n_k} = t$. A similar argument shows that $t = \overline{\lim}_{k\to\infty} x_{n_k} \leq \overline{\lim}_{n\to\infty} x_n = \beta$. Thus $\forall t \in S$, $\alpha \leq t \leq \beta$. Since $\alpha, \beta \in S$, the result follows.

b) By part a), $S = \{x\} \iff \underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n = x \iff x_n \to x$.

□

## 2.7 Basics of Topology in Metric Spaces

**Definition 2.9.** Let $S$ be a nonempty set. A **metric** is a function $d : S \times S \mapsto \mathbb{R}$ such that

a) $d(x, y) \geq 0$ for all $x, y \in S$ and $d(x, y) = 0$ if and only if $x = y$.

b) $d(x, y) = d(y, x)$ for all $x, y \in S$.

c) $d(x, y) \leq d(x, z) + d(z, y)$ for all $x, y, z \in S$.

A nonempty set $S$ together with a metric, $(S, d)$, is called a **metric space**.

**Example 2.17.** The distance function $d(x, y) = |x - y|$ is a metric on the real line. Thus $(\mathbb{R}, d)$ is a metric space.

**Note 2.6.** Though we will continue to prove theorems in the case of $\mathbb{R}$, most theorems in this section apply to general metric spaces. In fact, the only theorem in this section that is specific to $\mathbb{R}$ or $\mathbb{R}^k$ is the Heine-Borel theorem. All other results will hold more generally in a metric space. Indeed, many of the results for sequences that we have seen so far hold for a general metric space, and the proofs are similar with the only modification of the proof needed to generalize it to a metric space is to replace $|x - y|$ by $d(x, y)$.

**Definition 2.10.** A sequence $(x_n)_n$ in $\mathbb{R}$ is said to be **Cauchy** if $\forall \varepsilon > 0$, $\exists N \geq \mathbb{N}$ such that $m > n \geq N$ implies

$$|x_m - x_n| < \varepsilon.$$

**Proposition 2.14.** *Let $(x_n)_n$ be a seqeuence in $\mathbb{R}$. Then*

*a) If $(x_n)_n$ is converges, then it is Cauchy.*

*b) If $(x_n)_n$ is Cauchy, then it is bounded.*

*c) If $(x_n)_n$ is Cauchy and $x_{n_k} \to x \in \mathbb{R}$ for some subsequence of $(x_{n_k})_k$ of $(x_n)_n$, then $x_n \to x$.*

*Proof.*

a) Let $\varepsilon > 0$. Since $x_n \to x \in \mathbb{R}$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon/2$. Then for $m > n \geq N$

$$|x_m - x_n| = |x_m - x + x - x_n| \leq |x_m - x| + |x_n - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

b) Let $\varepsilon = 1$. We can choose $N \in \mathbb{N}$ such that $|x_m - x_n| < 1$ whenever $m > n \geq N$. Then for $n \geq N$

$$|x_n| = |x_n - x_N + x_N| \leq |x_n - x_N| + |x_N| < 1 + |x_N|.$$

Let $M = \max\{1 + |x_N|, |x_1|, |x_2|, \ldots, |x_{N-1}|\}$. Then $|x_n| \leq M$ for all $n \geq 1$.

c) Let $(x_n)_n$ be a Cauchy sequence in $\mathbb{R}$ and let $(x_{n_k})_k$ be a subsequence of $(x_n)_n$ such that that $x_{n_k} \to x \in \mathbb{R}$. Let $\varepsilon > 0$. Choose $K \in \mathbb{N}$ such that $k \geq K \implies |x_{n_k} - x| < \varepsilon/2$. Let $N \in \mathbb{N}$ be such that $m > n \geq N \implies |x_m - x_n| < \varepsilon/2$. Fix a $k_0 \geq \max\{K, N\}$. Then $n_{k_0} \geq k_0 \geq \max\{K, N\}$ and $n \geq N$ implies

$$|x_n - x| = |x_n - x_{n_{k_0}} + x_{n_{k_0}} - x| \leq |x_n - x_{n_{k_0}}| + |x_{n_{k_0}} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

**Definition 2.11.** A metric space $(S, d)$ is said to be **complete** if every Cauchy sequence in $S$ converges to an element in $S$.

**Theorem 2.6.** $\mathbb{R}$ *is complete.*

**Note 2.7.** Completeness is a nice property when exploring convergence. If a metric space is complete, then we can show that a sequence converges, without knowing what it converges to, by showing that it is Cauchy.

*Proof.* To show that $\mathbb{R}$ is complete, we need to show that if a sequence is convergent that it is Cauchy, and if a sequence is Cauchy, then it converges to some element of $\mathbb{R}$. The first half was proven in Proposition 2.14. It remains to show that if $(x_n)_n$ is a Cauchy sequence in $\mathbb{R}$, then $\exists x \in \mathbb{R}$ such that $x_n \to x$.

Let $(x_n)_n$ be a Cauchy sequence in $\mathbb{R}$. Then by Proposition 2.14(a), $(x_n)_n$ is bounded, so by the Bolzano-Weierstrass theorem, $(x_n)_n$ has a convergent subsequence $(x_{n_k})_k$. Let $x \in \mathbb{R}$ be such that $x_{n_k} \to x$. Then by Proposition 2.14(c), $x_n \to x$. □

**Definition 2.12.** An $\varepsilon$-neighborhood of a point $x \in \mathbb{R}$ is $\{y \in \mathbb{R} : |x - y| < \varepsilon\} = (x - \varepsilon, x + \varepsilon)$. This is also called an **open ball** of radius $\varepsilon$ and is denoted $B(x, \varepsilon)$.

**Definition 2.13.** A set $A \subseteq \mathbb{R}$ is said to be open if $\forall x \in A$, $\exists \varepsilon > 0$ such that

$$B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon) \subset A.$$

**Example 2.18.** Determine if the following sets are open in $\mathbb{R}$.

a) $A = (0, 1)$. Yes

b) $A = [0, 1)$. No, because of 0.

c) $A = [1, 2]$. No, because of $\{1, 2\}$.

d) $A = \mathbb{Q}$. No. Recall the irrationals are dense in $\mathbb{R}$, so $\forall r \in \mathbb{Q}$ and $\forall \varepsilon > 0$, $\exists q \in \mathbb{R} \setminus \mathbb{Q}$ such that $r - \varepsilon < q < r < r + \varepsilon$. Thus $B(r, \varepsilon) \not\subset \mathbb{Q}$.

**Note 2.8.** The sets $\emptyset$ and $\mathbb{R}$ are both open in $\mathbb{R}$.

**Proposition 2.15.** *An open ball is an open set.*

*Proof.* Let $B(x, \varepsilon) = (x - \varepsilon, x + \varepsilon)$ be an open ball of radius $\varepsilon > 0$, and let $y \in B(x, \varepsilon)$. Then $|x - y| < \varepsilon$. Let $\rho = \varepsilon - |x - y| > 0$, and let $z \in B(y, \rho)$. Then

$$|x - z| = |x - y + y - z| \leq |x - y| + |y - z| < |x - y| + \rho = \varepsilon.$$

Thus $B(y, \rho) \subseteq B(x, \varepsilon)$. □

**Definition 2.14.** The **interior** of a set $A$ is

$$A^\circ = \{x \in A | \exists \varepsilon > 0 \text{ such that } B(x, \varepsilon) \subset A\}$$

**Proposition 2.16.** *$A$ is open if and only if $A = A^\circ$.*

*Proof.* ($\implies$). Suppose that $A$ is an open set. Then $\forall x \in A$, $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset A$. Thus $A \subseteq A^\circ$. It is clear that $A^\circ \subset A$. Thus $A = A^\circ$.

($\impliedby$) Suppose that $A = A^\circ$. Let $x \in A$. Then $x \in A^\circ$, so $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset A$. Since $x \in A$ was arbitrary, $A$ is open. □

**Example 2.19.** Find $A^\circ$

a) $A = [0, 2]$. Then $A^\circ = (0, 2)$

b) $A = (0, 3] \cup \{6\}$. Then $A^\circ = (0, 3)$.

c) $\mathbb{Q}^\circ = \emptyset$

d) $A = \{1/n : n \in \mathbb{N}\}$. Then $A^\circ = \emptyset$.

**Proposition 2.17.** *$A^\circ$ is the largest open set contained in $A$.*

*Proof.* Let $x \in A^\circ$. Then $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset A$. Since $B(x, \varepsilon)$ is open, $\forall y \in B(x, \varepsilon)$, $\exists r > 0$ such that $B(y, r) \subset B(x, \varepsilon)$. Thus $B(x, \varepsilon) \subset A^\circ$, so $A^\circ$ is open. Now, let $B$ be an open set such that $B \subset A$. Then $\forall x \in B$, $\exists \varepsilon > 0$ such that $B(x, r) \subset B \subset A$, so $x \in A^\circ$. Thus $B \subset A^\circ$. □

**Proposition 2.18.**

*a) The union of an arbitrary collection of open sets is open.*

*b) The intersection of a finite number of open sets is open.*

**Note 2.9.** The intersection of an arbitrary collection of open sets may not be open as the following example shows.

**Example 2.20.** Let $A_n = (-1/n, 1/n)$ for $n \geq 1$. Then each $A_n$ is open buy

$$\cap_{n=1}^{\infty} A_n = \cap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\},$$

which is not open.

*Proof.*

a) Let $(A_i)_{i \in I}$ be a collection of open sets, and let $A = \cup_{i \in I} A_i$. Let $x \in A$. Then $x \in A_i$ for some $i \in I$. $A_i$ open implies $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset A_i \subset A$. Thus $A$ is open.

b) Let $A_i, i = 1, \ldots, n$ be open and let $A = \cap_{i=1}^{n} A_i$. Let $x \in A$. Then $x \in A_i$ for all $i = 1, 2, \ldots, n$, so $\exists r_i, i = 1, 2, \ldots, n$ such that $B(x, r_i) \subset A_i, i = 1, 2, \ldots, n$. Let $\varepsilon = \min\{r_1, r_2, \ldots, r_n\}$. Then $B(x, \varepsilon) \subset A_i$ for all $i = 1, 2, \ldots, n$, so $B(x, r) \subset A$.

□

**Definition 2.15.** A set $A \subseteq \mathbb{R}$ is said to be **closed** if $A^c = \mathbb{R} \setminus A$ is open.

**Example 2.21.** Determine if the following sets are closed.

a) $A = [1, 5]$. Yes since $\mathbb{R} \setminus A = (-\infty, 1) \cup (5, \infty)$ is open.

b) $A = [1, 6)$. No, since $\mathbb{R} \setminus A = (-\infty, 1) \cup [6, \infty)$ is not open due to the point $\{6\}$.

c) $A = \{1/n : n \in \mathbb{N}\}$. No. $\mathbb{R} \setminus A$ is not open since $0 \in \mathbb{R} \setminus A$, but every interval around 0 contains a point in $A$, i.e. $\forall \varepsilon > 0$, $\exists x \in A$ such that $x \in B(0, \varepsilon)$. So $\mathbb{R} \setminus A$ is not open.

**Proposition 2.19.** *Let $A \subseteq \mathbb{R}$. Then $A$ is closed if and only if $\forall (x_n)_n \subset A$ and $x_n \to x \in \mathbb{R}$ implies $x \in A$.*

Before we prove this result, let's look at a few application.

**Example 2.22.** Determine if the following sets are closed.

a) $A = \{1/n : n \in \mathbb{N}\}$ is not closed because $(1/n)_n \subset A$ and $1/n \to 0$ but $0 \notin A$.

b) $A = [0, 2)$ is not closed because $(2 - 1/n)_n \subset A$ and $2 - 1/n \to 2$ but $2 \notin A$. Note that $A$ is neither closed nor open.

c) $A = \{x \in \mathbb{R} | x^5 + 3x^4 \leq 2\}$. This set is closed. Let $(x_n)_n \subset A$ and suppose that $x_n \to x \in \mathbb{R}$. Then $x_n^5 + 3x_n^4 \leq 2$ for all $n \geq 1$. This implies that $x_n^5 + 3x_n^4 \to x^5 + 3x^4 \leq 2$, so $x \in A$.

*Proof.* ( $\Longrightarrow$ ) Suppose $A$ is closed and let $(x_n)_n \subset A$ be a convergent sequence with limit $x \in \mathbb{R}$. Suppose that $x \notin A$. Then $x \in \mathbb{R} \setminus A$ which is open. Then $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset \mathbb{R} \setminus A$. Since $x_n \to x$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon \implies x_n \in B(x, \varepsilon)$, a contradiction with $x_n \in A$.

( $\Longleftarrow$ ) Suppose that whenever $(x_n)_n \subset A$ and $x_n \to x$ then $x \in A$. Suppose that $A$ is not closed. Then $\mathbb{R} \setminus A$ is not open, so $\exists x \in \mathbb{R} \setminus A$ such that $\forall \varepsilon > 0$, $B(x, \varepsilon) \not\subset \mathbb{R} \setminus A$. This implies that $B(x, \varepsilon) \cap A \neq \emptyset$ for all $\varepsilon > 0$. Let $\varepsilon = 1/n$ and choose $x_n \in B(x, 1/n) \cap A$ for each $n \geq 1$. Then $(x_n)_n \subset A$ and $|x_n - x| < 1/n \to 0 \implies x_n \to x$. Thus by assumption $x \in A$ but this contradicts $x \in \mathbb{R} \setminus A$, so $A$ must be closed. □

**Proposition 2.20.**

*a) The intersection of an arbitrary collection of closed sets is closed.*

*b) The union of a finite collection of closed sets is closed.*

*Proof.*

a) Let $(A_i)_{i \in I}$ be a collection of closed sets and let $A = \cap_{i \in I} A_i$. Then

$$A^c = (\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c$$

which is open since $A_i^c$ is open for all $i \in I$.

b) Let $A_1, \ldots, A_n$ be closed sets and let $A = \cup_{i=1}^n A_i$. Then

$$A^c = \cap_{i=1}^n A_i^c$$

is open since $A_i^c$ is open for all $i = 1, 2, \ldots, n$.

□

**Definition 2.16.** The **closure** of $A$, denoted $\bar{A}$, is the intersection of all closed sets containing $A$, i.e.

$$\bar{A} = \cap\{B | A \subseteq B \text{ and } B \text{ is closed}\}.$$

**Note 2.10.**

- $\bar{A}$ is closed and is the smallest closed set containing $A$.
- $A$ is closed if and only if $A = \bar{A}$.

**Proposition 2.21.** *The following are equivalent:*

*a)* $x \in \bar{A}$

*b)* $\forall \varepsilon > 0$, $B(x, \varepsilon) \cap A \neq \emptyset$.

*c)* $\exists (x_n)_n \subset A$ *such that* $x_n \to x$.

*Proof.* $(a) \implies b))$. Let $x \in \bar{A}$. Suppose $b)$ is not true. Then $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \cap A = \emptyset \implies A \subset \mathbb{R} \setminus B(x, \varepsilon) = B$. Since $B$ is closed and $A \subset B$, $\bar{A} \subset B$, but $x \notin B$ and $x \in \bar{A}$, a contradiction.

$(b) \implies c))$. Let $\varepsilon = 1/n$. Choose $x_n \in B(x, 1/n) \cap A$. Then $(x_n)_n \subset A$ and $|x_n - x| < 1/n \to 0 \implies x_n \to x$.

$(c) \implies a))$ Suppose $(x_n)_n \subset A$ such that $x_n \to x$, but $x \notin \bar{A}$. Then $x \in \mathbb{R} \setminus \bar{A}$ which is open, so $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \subset \mathbb{R} \setminus \bar{A}$. Since $x_n \to x$, $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon$. This implies that $x_n \in B(x, \varepsilon) \subset \mathbb{R} \setminus \bar{A}$ for $n \geq N$, but $x_n \in A \subset \bar{A}$ for all $n \geq 1$, a contradiction. Thus, $c) \implies a)$. □

**Example 2.23.** Find $\bar{A}$.

a) $A = (0, 1)$. Then $\bar{A} = [0, 1]$.

b) $A = (-1, 2) \cup \{3\}$. Then $\bar{A} = [-1, -2] \cup \{3\}$.

c) $A = \mathbb{N}$. Then $\bar{A} = \mathbb{N}$

d) $A = (-2, 2) \cup (3, 5]$. Then $\bar{A} = [-2, 2] \cup [-3, 5]$

e) $A = \{1/n : n \in \mathbb{N}\}$. Then $\bar{A} = A \cup \{0\}$.

f) $A = \{n/(n+1) | n \in \mathbb{N}\}$. Then $\bar{A} = A \cup \{1\}$.

**Definition 2.17.** A point $x \in \mathbb{R}$ is said to be an **accumulation point** of $A \subset \mathbb{R}$ if $\forall \varepsilon > 0$, $B(x, \varepsilon)$ contains a point in $A$ different from $x$. The set of all accumulation points is denoted $A'$.

**Proposition 2.22.** $x \in A'$ *if and only if* $\forall \varepsilon > 0$, $B(x, \varepsilon) \cap A$ *contains infinitely many points.*

*Proof.* $(\impliedby)$ Clear.

$(\implies)$ Let $x \in A'$. Suppose that $B(x, \varepsilon) \cap A$ is finite for some $\varepsilon > 0$. Let $\{x_1, x_2, \ldots, x_n\}$ be the distinct points in $B(x, \varepsilon) \cap A$ different from $x$ (since $x \in A'$, there is at least one such $x_i$). Let $\rho = \min\{|x - x_1|, \ldots, |x - x_n|\} > 0$. Then $B(x, \rho) \cap A$ is either empty or contains only $x$, a contradiction with $x \in A'$. □

**Proposition 2.23.** *$x \in A'$ if and only if $\exists (x_n)_n \subset A \setminus \{x\}$ of distinct elements such that $x_n \neq x$ and $x_n \to x$.*

*Proof.* ( $\Longleftarrow$ ) Suppose that $(x_n)_n$ is a sequence of distinct elements in $A \setminus \{x\}$ such that $x_n \to x$. Let $\varepsilon > 0$. Then $\exists N \in \mathbb{N}$ such that $n \geq N \implies |x_n - x| < \varepsilon$ for $n \geq N$. Thus

$$B(x, \varepsilon) \cap A \supset \{x_n, n \geq N\},$$

which contains infinitely many points, so $x \in A'$.

( $\Longrightarrow$ ) Let $x \in A'$. Choose and $x_1 \in B(x,1) \cap A$ such that $x_1 \neq x$. Now suppose that for some $k \geq 1$, $x_1, x_2, \dots, x_k$ have been chosen such that $x_i \neq x_j$ whenever $i \neq j$, $x_i \neq x$ and $x_i \in B(x, 1/i) \cap A$, , $i = 1, \dots, k$. Then we can choose an $x_{k+1} \in B(x, 1/(k+1)) \cap A$ such that $x_{k+1} \neq x_i$ for all $i = 1, \dots, k$, and $x_{k+1} \neq x$, since $B(x, 1/(k+1)) \cap A$ contains infinitely many points. Thus, by induction we have a sequence of distinct elments $(x_n)_n \subset A$ such that $x_n \neq x$ and $|x_n - x| < 1/n \to 0 \implies x_n \to x$. □

**Proposition 2.24.** $\bar{A} = A \cup A'$.

*Proof.* Clearly $A \subset \bar{A}$ and $A' \subset \bar{A}$, so $A \cup A' \subset \bar{A}$. Now, let $x \in \bar{A}$. If $x \in A$, then $x \in A \cup A'$. Suppose $x \notin A$. Then by Proposition 2.21(b), $B(x, \varepsilon) \cap A \neq \emptyset$ for all $\varepsilon > 0 \implies B(x, \varepsilon) \cap A$ contains a point distinct from $x$ for all $\varepsilon > 0$, since $x \notin A$. □

**Definition 2.18.** The **boundary** of a set $A$, denoted $\partial A$, is

$$\partial A = \bar{A} \cap \overline{\mathbb{R} \setminus A}.$$

That is $x \in \partial A$ if $\forall \varepsilon > 0$, $B(x, \varepsilon)$ contains a point in $A$ and a point not in $A$.

**Example 2.24.** Find $\partial A$.

a) $A = (0,1)$. Then $\partial A = \{0, 1\}$.

b) $A = \{1/n : n \in \mathbb{N}\}$. Then $\partial A = A \cup \{0\}$.

c) $A = \mathbb{Q}$. Then $\bar{A} = \mathbb{R}$ and $\overline{\mathbb{R} \setminus \mathbb{Q}} = \mathbb{R}$, so $\partial \mathbb{Q} = \mathbb{R}$.

d) $A = \mathbb{N}$. Then $\partial A = \mathbb{N}$.

**Definition 2.19.** A point $x \in A$ is said to be an **isolated point** if $\exists \varepsilon > 0$ such that $B(x, \varepsilon) \cap A = \{x\}$.

**Example 2.25.** Find the isolated points of $A$.

a) $A = (0,1) \cup \{8\}$. The only isolated point is 8.

b) $A = \{1/n : n \in \mathbb{N}\}$. Every point of $A$ is an isolated point.

c) $A = \mathbb{N}$. Every point of $\mathbb{N}$ is an isolated point.

**Definition 2.20.** An **open cover** of a set $A$ is a family of open sets, $(G_i)_{i \in I}$ such that

$$A \subset \cup_{i \in I} G_i.$$

**Definition 2.21.** $A$ is said to be compact if every open cover of $A$ has a finite subcover, i.e. if $(G_i)_{i \in I}$ is an open cover of $A$, then $\exists i_1, i_2, \dots, i_n \in I$ such that $A \subset \cup_{k=1}^{n} G_{i_k}$.

**Proposition 2.25.** *A compact set is bounded.*

*Proof.* Let $A$ be compact and let $x_0 \in S$. Then $A \subset \cup_{n=1}^{\infty} B(x_0, n)$, so $(B(x_0, n))_{n=1}^{\infty}$ is an open cover of $A$. By compactness of $A$, there exists an $n_0 \in \mathbb{N}$ such that $A \subset \cup_{n=1}^{n_0} B(x_0, n_0) = B(x_0, n_0)$. Thus for all $x \in A$

$$|x| = |x - x_0 + x_0| \leq |x - x_0| + |x_0| < n_0 + |x_0|,$$

so $A$ is bounded. □

The following theorem gives a nice characterization of compact sets of metric spaces in terms of sequential compactness.

**Theorem 2.7.** *The following are equivalent:*

*a) $A$ is compact.*

*b) Every sequence in $A$ has a subsequence that converges to a point in $A$.*

*Proof.* We only prove the first half
($a) \implies b)$). Suppose that $A$ is compact. Let $(x_n)_n$ be a sequence in $A$, and suppose that for all $x \in A$, $(x_n)_n$ has no subsequence which converges to an $x \in A$. Then for each $x \in A$, $\exists \varepsilon_x > 0$ such that $x_n \in B(x, \varepsilon_x)$ finitely often (otherwise a $\exists (x_{n_k}))_k$ such that $x_{n_k} \to x$ for some $x \in A$). Clearly $A \subset \cup_{x \in A} B(x, \varepsilon_x)$. Since $A$ is compact, $\exists y_1, y_2, \ldots, y_N \in A$ such that $(x_n)_n \subset A \subset \cup_{i=1}^{N} B(y_i, \varepsilon_{y_i})$. But $(x_n)_n$ is infinite, so at least one of $B(y_i, \varepsilon_{y_i})$ must contain infinitely many of $(x_n)_n$ a contradiction.

($b) \implies a)$) See Rudin Theorem 2.41, p.40, for the case of $\mathbb{R}^k$ or DePree and Schwartz (1988) Theorem 6, p.299 for the general metric space case. □

**Proposition 2.26.** *If $A$ is a compact set, then $A$ is closed.*

*Proof.* Let $(x_n)_n$ be a sequence in $A$ such that $x_n \to x$. Since $A$ is compact, $\exists$ a subseqeuence $(x_{n_k})_k$ of $(x_n)_n$ such that $x_{n_k} \to x_0 \in A$. Since $x_n \to x$, $x_{n_k} \to x$, and $x = x_0 \in A$. Thus, $A$ is closed. □

The following theorem not true in a general metric space, but it does hold in $\mathbb{R}^k$ for all $k \geq 1$. The Heine-Borel theorem gives a nice characterization of compact sets in $\mathbb{R}^k$.

**Theorem 2.8** (Heine-Borel)**.** *Let $A \subseteq \mathbb{R}^k$. Then $A$ is compact if and only if $A$ is closed and bounded.*

*Proof.* We will prove this in the case of $\mathbb{R}$, but the proof is exactly the same for general $\mathbb{R}^k$.

($\implies$) This implication is just Proposition 2.25 and 2.26.

($\impliedby$) Suppose that $A$ is closed and bounded. Let $(x_n)_n$ be a sequence in $A$. Then $(x_n)_n$ is bounded, so by the Bolzano-Weierstrass theorem $(x_n)_n$ has a convergent subsequence $(x_{n_k})_k$ such that $x_{n_k} \to x \in \mathbb{R}$. Since $(x_{n_k})_k \subset A$ and $A$ is closed, $x \in A$. So every sequence in $A$ has a convergent subsequence that converges to a point in $A$, which implies $A$ is compact. □

**Proposition 2.27.** *Let $A$ be a compact set and let $B \subset A$ be closed. Then $B$ is compact.*

*Proof.* Let $(x_n)_n$ be a sequence in $B$. Then $(x_n)_n$ is also a sequence in $A$. Since $A$ is compact, $\exists$ a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $x_{n_k} \to x \in A$. Since $B$ is closed, $x \in B$. Therefore, $B$ is compact. □

**Example 2.26.** Determine if the following sets are compact.

a) $A = \mathbb{Q}$. No, because $\mathbb{Q}$ is not closed.

b) $A = \{1/n : n \in \mathbb{N}\}$. No, because $A$ is not closed.

c) $A = \mathbb{N}$. No, because $A$ is not bounded.

d) $A = [0, 7] \cup \{9\}$. Yes, because $A$ is closed and bounded.

e) $A = (0, 9]$. No, because $A$ is not closed.

f) $A = \{x : x^2 > 1\}$. No, because it is neither closed nor bounded.

## 2.8 Series

**Definition 2.22.** Given a sequence $(a_n)_n$ in $\mathbb{R}$, the $n$th partial sum is

$$s_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n.$$

We say that the infinite series $\sum_{k=1}^{\infty} a_k$ **converges** with sum $s \in \mathbb{R}$ if $s_n \to s$. Otherwise, we say the infinite series **diverges**.

**Example 2.27** (Geomertic Series)**.** Let $a, r \in \mathbb{R}$. Show that the series $\sum_{k=1}^{\infty} ar^{k-1}$ converges for $|r| < 1$ and diverges otherwise.

*Solution.* Let $s_n = \sum_{k=1}^{n} ar^{k-1}$ for $n \geq 1$. For $r \neq 1$,

$$(r-1)s_n = rs_n - s_n = a\left(\sum_{k=1}^{n} r^k - \sum_{k=1}^{n} r^{k-1}\right) = a\left(\sum_{k=1}^{n} r^k - \sum_{k=0}^{n-1} r^k\right) = a(r^n - 1).$$

Then for $r \neq 1$,

$$s_n = a\frac{r^n - 1}{r - 1}.$$

By Proposition 2.7, $r^n \to 0$ for $|r| < 1$. Thus,

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

For $|r| > 1$, then

$$|s_n| = |a| \cdot \frac{|r|^n - 1}{|r| - 1} \to \infty.$$

If $r = 1$, then $s_n = na \to \infty$. If $r = -1$, then

$$s_n = \sum_{k=1}^{n} a(-1)^{k-1} = \begin{cases} a, & \text{n odd} \\ 0, & \text{n even} \end{cases},$$

so $(s_n)_n$ does not converge. □

**Proposition 2.28** (Algebra of Series)**.**

a) *If* $\sum_{k=1}^{\infty} a_k$ *and* $\sum_{k=1}^{\infty} b_k$ *converge, then* $\sum_{k=1}^{\infty}(a_k + b_k)$ *converges and the sum is equal to* $\sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$*.*

b) *If* $\sum_{k=1}^{\infty} a_k$ *converges, then for any* $c \in \mathbb{R}$*,* $\sum_{k=1}^{\infty}(ca_k)$ *converges and* $\sum_{k=1}^{\infty}(ca_k) = c\sum_{k=1}^{\infty} a_k$*.*

*Proof.* Exercise. □

Since $\mathbb{R}$ is complete and a series is just a limit of a sequence of partial sums, we can apply convergence theorems about convergent sequences to series. In particular, we have the following Cauchy criterion for convergence of a series.

**Theorem 2.9** (Cauchy Criterion)**.** *Let* $(a_n)_n$ *be a sequence in* $\mathbb{R}$*. Then* $\sum_{k=1}^{\infty} a_k$ *converges if and only if the seqeuence of partials sums* $(s_n)_n$ *is Cauchy, that is,* $\forall \varepsilon > 0$*,* $\exists N \in \mathbb{N}$ *such that* $m > n \geq N$ *implies*

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m} x_k\right| < \varepsilon.$$

*Proof.* Since $\mathbb{R}$ is complete, the sequence of partial sums $(s_n)_n$ converge if and only if $(s_n)_n$ is Cauchy. □

**Corollary 2.5.** *Let* $(a_n)_n$ *be a real sequence. If* $\sum_{k=1}^{\infty} a_n$ *converges, then* $a_n \to 0$*.*

*Proof.* Let $\varepsilon > 0$. Since $\sum_{k=1}^{\infty} a_k$ converges, the sequence of partial sums $(s_n)_n$ are Cauchy, so $\exists N \in \mathbb{N}$ such that $m > n \geq N \implies |s_m - s_n| < \varepsilon$. In particular for $n \geq N$, we have

$$|a_n| = |s_{n+1} - s_n| < \varepsilon.$$

□

Also recall that monotonic sequences converge if and only if the sequence is bounded.

**Theorem 2.10.** *Let* $(a_n)_n$ *be a sequence in* $\mathbb{R}$ *such that* $a_n \geq 0$ *for all* $n \geq 1$*. Then* $\sum_{k=1}^{\infty} a_k$ *converges if and only if the sequence of partial sums is bounded.*

*Proof.* Since $a_n \geq 0$ for all $n \geq 1$, $s_n \leq s_{n+1}$ for all $n \geq 1$, that is $(s_n)_n$ is an increasing sequence. Thus $(s_n)_n$ converges if and only if it is bounded. □

For the next corollary, we are going to need to rely on some of your prior knowledge of Riemann integration from calculus. We will come back formally later to see why the following argument works, but for now we will rely on a picture.

**Corollary 2.6.** *The series* $\sum_{n=1}^{\infty} \frac{1}{n^p}$ *converges if and only if* $p > 1$*.*

*Proof.* Note that for $p > 0$ and for $n \geq 1$, $1/n^p \geq 0$ and

$$\int_1^{n+1} \frac{1}{x^p}\, dx \leq \sum_{k=1}^{n} \frac{1}{k^p} \leq 1 + \int_1^{n} \frac{1}{x^p}\, dx.$$

Then $\sum_{n=1}^{\infty} 1/n^p$ converges $\iff$ $(s_n)_n$ are bounded $\iff$ $\int_1^{\infty} 1/x^p\, dx$ converges $\iff$ $p > 1$. □

**Example 2.28.** Determine if the following series converge.

a) $\sum_{n=1}^{\infty} 1/n^2$

*Solution.* Note that for each $n \geq 1$

$$\sum_{k=1}^{n} \frac{1}{k^2} \leq 1 + \int_1^n \frac{1}{x^2}\, dx = 1 + (1 - 1/n) \to 2,$$

so $\sum_{n=1}^{\infty} 1/n^2$ converges, since it is monotonic and bounded. □

b) $\sum_{n=1}^{\infty} 1/n$ (Harmonic series)

*Solution.* Note that for all $n \geq 1$

$$\int_1^{n+1} \frac{1}{x}\, dx \leq \sum_{k=1}^{n} \frac{1}{k}.$$

Since $\int_1^{\infty} (1/x)\, dx = \lim_{b\to\infty} \int_1^b (1/x)\, dx = \infty$, $\sum_{n=1}^{\infty} 1/n$ diverges to $\infty$. □

**Proposition 2.29.** *If $\sum_{k=1}^{\infty} |a_k|$ converges, then $\sum_{k=1}^{\infty} a_k$ converges and $|\sum_{k=1}^{\infty} a_k| \leq \sum_{k=1}^{\infty} |a_k|$.*

*Proof.* Suppose that $\sum_{k=1}^{\infty} |a_k|$ converges. We will apply the Cauchy criterion to show that $\sum_{k=1}^{\infty} a_k$ converges. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m > n \geq N$ implies

$$\sum_{k=n+1}^{m} |a_k| < \varepsilon.$$

Then for $m > n \geq N$,

$$\left| \sum_{k=n+1}^{m} a_k \right| \leq \sum_{k=n+1}^{m} |a_k| < \varepsilon.$$

Thus $\sum_{k=1}^{\infty} a_k$ converges by the Cauchy criterion. Furthermore, since for all $n \geq 1$

$$\left| \sum_{k=1}^{n} a_k \right| \leq \sum_{k=1}^{n} |a_k|,$$

we have

$$\left| \sum_{k=1}^{\infty} a_k \right| \leq \sum_{k=1}^{\infty} |a_k|$$

□

**Definition 2.23.** A series $\sum_{k=1}^{\infty} a_k$ is said to be **absolutely convergent** if $\sum_{k=1}^{\infty} |a_k|$ converges. If $\sum_{k=1}^{\infty} a_k$ converges but $\sum_{k=1}^{\infty} |a_k| = \infty$, then the series is called **conditionally convergent**.

**Proposition 2.30.** *Given a sequence $(a_n)_n$ in $\mathbb{R}$, let*

$$p_n = \frac{|a_n| + a_n}{2} = \max\{a_n, 0\} \quad \textit{and} \quad q_n = \frac{|a_n| - a_n}{2} = \max\{-a_n, 0\}.$$

*Then $p_n, q_n \geq 0$, $a_n = p_n - q_n$, and $|a_n| = p_n + q_n$.*

*a) $\sum_{k=1}^{\infty} a_k$ is absolutely convergent if and only if $\sum_{k=1}^{\infty} p_k < \infty$ and $\sum_{k=1}^{\infty} q_k < \infty$.*

*b) If $\sum_{k=1}^{\infty} a_k$ is conditionally convergent, then $\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} q_k = \infty$.*

*Proof.*

a) ( $\implies$ ) Suppose $\sum_{k=1}^{\infty} |a_k| < \infty$. Then for all $n \geq 1$

$$\sum_{k=1}^{n} p_k \leq \sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{\infty} |a_k| < \infty \quad \text{and} \quad \sum_{k=1}^{n} q_k \leq \sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{\infty} |a_k| < \infty$$

Thus, $\sum_{k=1}^{\infty} p_k < \infty$ and $\sum_{k=1}^{\infty} q_k < \infty$.

( $\impliedby$ ) Suppose $\sum_{k=1}^{\infty} p_k < \infty$ and $\sum_{k=1}^{\infty} q_k < \infty$. Then

$$\sum_{k=1}^{\infty} |a_k| = \sum_{k=1}^{\infty} (p_k + q_k) = \sum_{k=1}^{\infty} p_k + \sum_{k=1}^{\infty} q_k < \infty.$$

b) Suppose $\sum_{k=1}^{\infty} a_k$ is convergent but $\sum_{k=1}^{\infty} |a_k| = \infty$. If $\sum_{k=1}^{\infty} p_k < \infty$, then $\sum_{k=1}^{\infty}(p_k - a_k) < \infty$. Since $p_k - a_k = p_k - (p_k - q_k) = q_k$, this implies that $\sum_{k=1}^{\infty} q_k < \infty$, so that $\sum_{k=1}^{\infty} |a_k| < \infty$ by part a), a contradiction. Similary, $\sum_{k=1}^{\infty} q_k < \infty \implies \sum_{k=1}^{\infty} p_k < \infty \implies \sum_{k=1}^{\infty} |a_k| < \infty$, a contradiction. Thus it must be the case that $\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} q_k = \infty$.

□

A key difference between absolutely convergent and conditionally convergent series is how rearrangements behave. A rearrangement of a series $\sum_{k=1}^{\infty} a_k$ is a series of the form $\sum_{k=1}^{\infty} b_k$ where $b_k = a_{p(k)}$ and $p : \mathbb{N} \mapsto \mathbb{N}$ is bijective function. This means that $\sum_{k=1}^{\infty} b_k$ is a series of the same terms as $\sum_{k=1}^{\infty} a_k$, but the sum of these elements occurs in a different order.

**Theorem 2.11.** *If $\sum_{k=1}^{\infty} a_k$ is absolutely convergent, then every rearrangement $\sum_{k=1}^{\infty} b_n$ of $\sum_{k=1}^{\infty} a_k$ is absolutely convergent and $\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} a_k$.*

**Note 2.11.** This theorem says the the sum of an absolutely convergent series independent of the order of the terms in the sum. Consider the expected value of a discrete random variable

$$EX = \sum_{x \in \mathcal{X}} x P(X = x).$$

It would be troublesome if the order in which these terms were summed changed the expected value. This is why we say the expected value exists if $E|X| = \sum_{x \in \mathcal{X}} |x| P(X = x) < \infty$. In this case, the sum is independent of the order.

*Proof.* Suppose $\sum_{k=1}^{\infty} |a_k| < \infty$. Let $p : \mathbb{N} \mapsto \mathbb{N}$ be a bijective map and let $b_k = a_{p(k)}$ for $k \geq 1$. Let $s_n = \sum_{k=1}^{n} a_k$, $t_n = \sum_{k=1}^{n} b_k$, and $s = \sum_{k=1}^{\infty} a_k$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m > n \geq N \implies \sum_{k=n+1}^{m} |a_k| < \varepsilon \implies \sum_{k=N+1}^{\infty} |a_k| \leq \varepsilon$. Since $p$ is bijective, $\exists k_1, k_2, \ldots k_N$ such that $\{1, 2, \ldots N\} = \{p(k_1), p(k_2), \ldots, p(k_N)\}$. Let $M = \max\{k_1, k_2, \ldots, k_N\}$. Then $M \geq N$ and $\{1, 2, \ldots, N\} \subseteq \{p(1), p(2), \ldots, p(M)\}$. Now, for $m \geq M \geq N$

$$|s_m - t_m| = \left| \sum_{k=1}^{m} a_k - \sum_{k=1}^{m} b_k \right| \leq \sum_{k=N+1}^{\infty} |a_k| \leq \varepsilon.$$

The inequality is true since the terms $a_1, \ldots, a_N$ cancel as well as any other $a_k$ for $k > N$ that $s_m$ and $t_m$ share. Thus $|s_n - t_n| \to 0$, and

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

Therefore, $\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} |b_k| < \infty$, since

$$\sum_{k=M+1}^{\infty} |b_k| = \sum_{k=M+1}^{\infty} |a_{p(k)}| \leq \sum_{k=N+1}^{\infty} |a_k| \leq \varepsilon$$

and $\{p(k) | k \geq M + 1\} \subseteq \{n | n \geq N + 1\}$ by choice of $M$. □

**Theorem 2.12** (Riemann Theorem on Conditionally Convergent Series)**.** *Suppose $\sum_{k=1}^{\infty} a_k$ is a conditionally convergent real series. Let $-\infty \leq x \leq y \leq \infty$. Then there exists a rearrangement $\sum_{k=1}^{\infty} b_k$ of $\sum_{k=1}^{\infty} a_k$ such that*

$$\underline{\lim}_{n\to\infty} \sum_{k=1}^{n} b_k = x \quad \textit{and} \quad \overline{\lim}_{n\to\infty} \sum_{k=1}^{n} b_k = y.$$

**Note 2.12.** Theorem 2.12 says that if a series is conditionally convergement, then for any $s \in \mathbb{R}$, it can be rearranged so that it converges to $s$.

**Corollary 2.7.** *$\sum_{k=1}^{\infty} a_k$ is absolutely convergent if and only if every rearrangement has the same sum.*

We now turn to some tests that we can use to determine if a series converges or not.

**Proposition 2.31** (The Comparison Test)**.** *Let $\sum_{k=1}^{\infty} a_k$ be a series where $a_k \geq 0$ for all $k \geq 1$.*

*a) If $\sum_{k=1}^{\infty} a_k$ converges and $|b_k| \leq a_k$ for all $k \geq 1$, then $\sum_{k=1}^{\infty} |b_k|$ converges.*

*b) If $\sum_{k=1}^{\infty} a_k = \infty$ and $a_k \leq b_k$ for all $k \geq 1$, then $\sum_{k=1}^{\infty} b_k = \infty$.*

*Proof.*

a) Let $\varepsilon > 0$. Since $\sum_{k=1}^{\infty} a_k$ converges, $\exists N \in \mathbb{N}$ such that $m > n \geq N$ implies

$$\left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Since $|b_k| \leq a_k$ for all $k \geq 1$, we have $m > n \geq N$ implies

$$\sum_{k=n+1}^{m} |b_k| \leq \sum_{k=n+1}^{m} a_k < \varepsilon.$$

Thus by the Cauchy criterion, $\sum_{k=1}^{\infty} |b_k|$ converges.

b) Suppose that $a_k \leq b_k$ for all $k \geq 1$ and that $\sum_{k=1}^{\infty} a_k = \infty$. Let $M > 0$, and choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$M < \sum_{k=1}^{n} a_k.$$

Then $a_k \leq b_k$ for all $k \geq 1$ implies

$$M < \sum_{k=1}^{n} a_k \leq \sum_{k=1}^{n} b_k$$

for all $n \geq N$. Thus $\sum_{k=1}^{\infty} b_k = \infty$.

□

**Theorem 2.13** (Root Test)**.** *Let $\sum_{k=1}^{\infty} a_k$ be a series.*

*a) $\sum_{k=1}^{\infty} a_k$ converges absolutely if $\overline{\lim}_{k\to\infty} |a_k|^{1/k} < 1$.*

*b) $\sum_{k=1}^{\infty} a_k$ does not converge if $\overline{\lim}_{k\to\infty} |a_k|^{1/k} > 1$.*

*c) If $\overline{\lim}_{k\to\infty} |a_k|^{1/k} = 1$, then the test gives no information about the convergence of $\sum_{k=1}^{\infty} a_k$.*

*Proof.* Let $\alpha = \overline{\lim}_{k\to\infty} |a_k|^{1/k}$.

a) Suppose that $\alpha < 1$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that $\alpha + \delta < 1$. Then $\exists N_1 \in \mathbb{N}$ such that

$$\alpha - \delta < \sup_{k \geq N_1} |a_k|^{1/k} < \alpha + \delta.$$

Then for all $k \geq N_1$, $|a_k|^{1/k} < \alpha + \delta$, so

$$|a_k| < (\alpha + \delta)^k, \ \forall k \geq N_1.$$

Since $0 < \alpha + \delta < 1$ and $\sum_{k=1}^{\infty} (\alpha + \delta)^k$ is a geometric series with $a = 1$ and $r = \alpha + \delta$, it converges. Then $\exists N_2 \in \mathbb{N}$ such that $m > n \geq N_2$ implies

$$\left| \sum_{k=n+1}^{m} (\alpha + \delta)^k \right| < \varepsilon.$$

Then $m > n \geq \max\{N_1, N_2\} \implies$

$$\sum_{k=n+1}^{m} |a_k| \leq \sum_{k=n+1}^{m} (\alpha + \delta)^k < \varepsilon.$$

Thus $\sum_{k=1}^{\infty} a_k$ converges absolutely by the Cauchy criterion.

b) If $\alpha > 1$, then by Corollary 2.4, $\exists$ a subsequence $(|a_{n_k}|^{1/n_k})_k$ of $(|a_n|^{1/n})_n$ such that $|a_{n_k}|^{1/n_k} \to \alpha$ as $k \to \infty$. Let $\delta > 0$ such that $\alpha - \delta > 1$, and choose $K \in \mathbb{N}$ such that $k \geq K$ implies

$$\alpha - \delta < |a_{n_k}|^{1/n_k} < \alpha + \delta.$$

Then for $k \geq K$

$$1 < (\alpha + \delta)^{1/n_k} < |a_{n_k}|.$$

Therefore, $N \geq n_K \geq K \implies \sup_{j \geq N} |a_j| > 1 \implies \overline{\lim}_{n\to\infty} |a_n| \geq 1$. Thus $a_n \not\to 0$, so $\sum_{k=1}^{\infty} a_k$ does not converge by Corollary 2.5.

c) Recall that $n^{1/n} \to 1$. Then

$$\lim_{n\to\infty} \left(\frac{1}{n}\right)^{1/n} = \lim_{n\to\infty} \left(\frac{1}{n^2}\right)^{1/n} = 1,$$

but $\sum_{n=1}^{\infty} 1/n$ diverges and $\sum_{n=1}^{\infty} 1/n^2$ converges.

□

**Theorem 2.14** (Ratio Test)**.** *Let $\sum_{k=1}^{\infty} a_k$ be a series of non-zero terms.*

*a) $\sum_{k=1}^{\infty} a_k$ converges absolutely if $\overline{\lim}_{k\to\infty} |a_{k+1}/a_k| < 1$.*

*b) $\sum_{k=1}^{\infty} a_k$ does not converge if $\underline{\lim}_{k\to\infty} |a_{k+1}/a_k| > 1$.*

*c) If $\underline{\lim}_{k\to\infty} |a_{k+1}/a_k| \leq 1 \leq \overline{\lim}_{k\to\infty} |a_{k+1}/a_k|$, then the test gives no information.*

*Proof.* Recall from Theorem 2.2

$$\underline{\lim}_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| \leq \underline{\lim}_{k\to\infty} |a_k|^{1/k} \leq \overline{\lim}_{k\to\infty} |a_k|^{1/k} \leq \overline{\lim}_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|.$$

a) If $\overline{\lim}_{k\to\infty} |a_{k+1}/a_k| < 1$, then $\overline{\lim}_{k\to\infty} |a_k|^{1/k} < 1$ and the series converges absolutely by the root test.

b) If $\underline{\lim}_{k\to\infty} |a_{k+1}/a_k| > 1$, then $\overline{\lim}_{k\to\infty} |a_k|^{1/k} > 1$ and the series does not converge the root test.

c) Again, consider the series $\sum_{n=1}^{\infty} 1/n$ and $\sum_{n=1}^{\infty} 1/n^2$. The first series diverges and the second series converges but

$$\lim_{n\to\infty} \frac{1/(n+1)}{1/n} = \lim_{n\to\infty} \frac{1/(n+1)^2}{1/n^2} = 1.$$

□

**Note 2.13.** The ratio test is frequently easier to apply than the root test, since it is usually easier to compute ratios than $n$th roots. However, the root test has wider scope. That is, whenever the ratio test shows convergence, then the root test does too, and whenever the root test is inconclusive, the ratio test is too. This can easily by seen using Theorem 2.2.

**Theorem 2.15** (Integral Test)**.** *Suppose $f : [1,\infty) \mapsto [0,\infty)$ is decreasing and $a_n = f(n)$ for $n \in \mathbb{N}$. Then $\sum_{k=1}^{\infty} a_k < \infty$ if and only if $\int_1^{\infty} f(x)\,dx < \infty$. If $\sum_{k=1}^{\infty} a_k = s < \infty$, then $|s - \sum_{k=1}^{n} a_k| < \int_n^{\infty} f(x)\,dx$.*

*Proof.* Suppose $f : [1,\infty) \mapsto [0,\infty)$ is decreasing and $a_n = f(n)$ for $n \geq 1$. Then for $k \geq 1$

$$k \leq x \leq k+1 \implies a_{k+1} \leq f(x) \leq a_k$$

so that

$$a_{k+1} = a_{k+1}\int_k^{k+1} 1\,dx \leq \int_k^{k+1} f(x)\,dx \leq a_k \int_k^{k+1} 1\,dx = a_k.$$

Thus

$$\sum_{k=2}^{n+1} a_k = \sum_{k=1}^{n} a_{k+1} \leq \sum_{k=1}^{n}\int_k^{k+1} f(x)\,dx = \int_1^{n+1} f(x)\,dx \leq \sum_{k=1}^{n} a_k.$$

Therefore for all $n \geq 1$

$$\sum_{k=1}^{n} a_k \leq a_1 + \int_1^{n} f(x)\,dx \quad\text{and}\quad \int_1^{n+1} f(x)\,dx \leq \sum_{k=1}^{n} a_k.$$

Since $f(x) \geq 0$ for all $x \geq 1$, we have for all $n \geq 1$

$$\sum_{k=1}^{n} a_k \leq a_1 + \int_1^{\infty} f(x)\,dx \quad\text{and}\quad \int_1^{n+1} f(x)\,dx \leq \sum_{k=1}^{\infty} a_k,$$

so $\sum_{k=1}^{\infty} a_k < \infty \iff \int_1^{\infty} f(x)\,dx < \infty$. Also,

$$\left|\sum_{k=1}^{\infty} a_k - \sum_{k=1}^{n} a_k\right| = \sum_{k=n+1}^{\infty} a_k = \sum_{k=n}^{\infty} a_{k+1} \leq \sum_{k=n}^{\infty}\int_k^{k+1} f(x)\,dx = \int_n^{\infty} f(x)\,dx.$$

□

**Example 2.29.** Determine if the following series converge.

a) $\sum_{n=1}^{\infty} \left(-\frac{1}{3}\right)^n$.

b) $\sum_{n=1}^{\infty} \frac{n}{n^2+3}$

c) $\sum_{n=1}^{\infty} \frac{1}{n^2+1}$

d) $\sum_{n=1}^{\infty} \frac{n}{3^n}$

e) $\sum_{n=1}^{\infty} \left(\frac{2}{(-1)^n-3}\right)^n$.

The following test provides a criterion for convergence of an alternating series. Note that unlike the other theorems, this test only guarantees convergence but not absolute convergence.

**Theorem 2.16** (Alternating Series Test)**.** *Let $(a_n)_n$ be a decreasing sequence such that $a_n \to 0$, then the series $\sum_{k=1}^{\infty}(-1)^k a_k$ converges.*

*Proof.* Define $B_0 = 0$ and let $B_n = \sum_{k=1}^{n}(-1)^n$. Then

$$B_n = \begin{cases} -1, & \text{n odd} \\ 0, & \text{n even} \end{cases}.$$

In particular, $|B_n| \leq 1$ for all $n \geq 1$. Let $\varepsilon > 0$. Since $a_n \to 0$, choose $N \in \mathbb{N}$ such that $n \geq N \implies |a_n| < \varepsilon/2$. Then for $m > n \geq N$

$$\begin{aligned}
\left|\sum_{k=n+1}^{m}(-1)^k a_k\right| &= \left|\sum_{k=n+1}^{m}(B_k - B_{k-1})a_k\right| \\
&= \left|\sum_{k=n+1}^{m} B_k a_k - \sum_{k=n+1}^{m} B_{k-1}a_k\right| \\
&= \left|\sum_{k=n+1}^{m} B_k a_k - \sum_{k=n}^{m-1} B_k a_{k+1}\right| \\
&= \left|\sum_{k=n+1}^{m-1} B_k(a_k - a_{k+1}) + B_m a_m - B_n a_{n+1}\right| \\
&\leq \sum_{k=n+1}^{m-1}(a_k - a_{k+1}) + a_m + a_{n+1} \\
&= [(a_{n+1} - a_{n+2}) + (a_{n+2} - a_{n+3}) + \cdots + (a_{m-1} - a_m)] + a_m + a_{n+1} \\
&= a_{n+1} - a_m + a_m + a_{n+1} \\
&= 2a_{n+1} < 2 \cdot \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Thus, $\sum_{k=1}^{\infty}(-1)^k a_k$ converges by the Cauchy criterion. □

## 2.9 Applications in Probability and Statistics

- Continuity of Probability Measure
- Borel-Cantelli
- Equivalent conditions for almost sure convergence
- Subsequences of random variables that converge in probability

# Chapter 3

# Continuity

## 3.1 Limits of Functions

**Definition 3.1.** Let $f : S \subset \mathbb{R} \mapsto \mathbb{R}$ and let $a$ be an accumulation point of $S$. Then $\lim_{x\to a} f(x) = L \in \mathbb{R}$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for $x \in S$,

$$0 < |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon.$$

**Note 3.1.** Note that $f$ may not even be defined at $a$, but $\lim_{x\to a} f(x)$ can still exist. Indeed, consider the function $f : (-\infty, 2) \cup (2, \infty)$ defined by

$$f(x) = \frac{x^2 - 4}{x - 2} = \frac{(x-2)(x+2)}{x-2}.$$

Then $f(x) = x + 2$ for all $x \neq 2$ and is not defined at $x = 2$, but $\lim_{x\to 2} f(x) = 2 + 2 = 4$.

**Proposition 3.1.** *Let $f : S \subseteq \mathbb{R} \to \mathbb{R}$ and let $a$ be an accumulation point of $S$. Then $\lim_{x\to a} f(x) = L$ if and only if $f(x_n) \to L$ whenever $(x_n)_n \subset S \setminus \{a\}$ and $x_n \to a$.*

*Proof.* ( $\implies$ ) Suppose that $\lim_{x\to a} f(x) = L$, and let $(x_n)_n$ be a sequence in $S \setminus \{a\}$ such that $x_n \to a$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that

$$0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon.$$

Since $x_n \to a$, $\exists N \in \mathbb{N}$ such that $n \geq N$ implies $|x_n - a| < \delta$. Then $n \geq N \implies 0 < |x_n - a| < \delta \implies |f(x_n) - L| < \varepsilon$.

( $\impliedby$ ) (Proof by contrapositive) Suppose $\lim_{x\to a} f(x) \neq L$. Then $\exists \varepsilon > 0$ such that $\forall \delta > 0$, $\exists x \in S \setminus \{a\}$ such that $0 < |x - a| < \delta$ and $|f(x) - L| \geq \varepsilon$. Let $\delta = 1$. Choose $x_1 \in S \setminus \{a\}$ such that $0 < |x_1 - a| < 1$ and $|f(x_1) - L| \geq \varepsilon$. Suppose for some $k \geq 1$, we have for $1 \leq j \leq k$, $x_j \in S \setminus \{a\}$ such that

$$0 < |x_j - a| < 1/j \text{ and } |f(x_j) - L| \geq \varepsilon.$$

Let $\delta = 1/(k+1)$. Since $\lim_{x\to a} f(x) \neq L$, we can choose an $x_{k+1} \in S \setminus \{a\}$ such that

$$0 < |x_{k+1} - a| < \frac{1}{k+1} \text{ and } |f(x_{k+1}) - L| \geq \varepsilon.$$

Then by induction $\exists (x_n)_n \subset S \setminus \{a\}$ such that for all $n \geq 1$

$$0 < |x_n - a| < \frac{1}{n} \text{ and } |f(x_n) - L| \geq \varepsilon,$$

so $x_n \to a$ but $f(x_n) \not\to L$. □

**Proposition 3.2** (Algebra of Limits)**.** *Let $f, g : S \subset \mathbb{R} \mapsto \mathbb{R}$, let $a$ be an accumulation point of $S$, and suppose that*

$$\lim_{x\to a} f(x) = L \quad \text{and} \quad \lim_{x\to a} g(x) = M.$$

*Then*

*a)* $\lim_{x\to a}(f+g)(x) = L + M$

*b)* $\lim_{x\to a}(fg)(x) = LM$

*c)* $\lim_{x\to a}(f/g)(x) = L/M$ *if* $M \neq 0$.

*Proof.* These properties follow immediately from the analogous properties of sequences. We prove a) as an example. Let $(x_n)_n \subset S \setminus \{a\}$ be such that $x_n \to a$. Then

$$f(x_n) \to L \text{ and } g(x_n) \to M \implies f(x_n) + g(x_n) \to L + M.$$

Thus by Proposition 3.1, $\lim_{x\to a}[f(x) + g(x)] = L + M$. □

**Example 3.1.** Find $\displaystyle\lim_{t\to 1} \frac{\sqrt{t}-1}{t-1}$.

*Solution.* Let $f(t) = \dfrac{\sqrt{t}-1}{t-1}$. Then $f : (0,\infty) \setminus \{1\} \mapsto \mathbb{R}$. Note that $f(1)$ is undefined, but 1 is an accumulation point of $(0,\infty) \setminus \{1\}$ and for $t \neq 1$

$$\frac{\sqrt{t}-1}{t-1} = \frac{\sqrt{t}-1}{t-1} \cdot \frac{\sqrt{t}+1}{\sqrt{t}+1} = \frac{t-1}{(t-1)\sqrt{t}+1} = \frac{1}{\sqrt{t}+1} \to \frac{1}{2} \text{ as } t \to 1.$$

□

**Definition 3.2.** Let $f : S \to \mathbb{R}$. If $a$ is an accumulation point of $S \cap (-\infty, a)$, then the onesided limit as $x$ approaches $a$ from below (or from the left) is $L$, written $f(a-) = \displaystyle\lim_{x\to a^-} f(x) = L$, if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for $x \in S$

$$a - \delta < x < a \implies |f(x) - L| < \varepsilon.$$

Similarly, if $a$ is an accumulation point of $S \cap (a, \infty)$, then the onesided limit as $x$ approaches $a$ from above (or from the right) is $L$, written $f(a+) = \displaystyle\lim_{x\to a^+} f(x) = L$, if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for $x \in S$

$$a < x < a + \delta \implies |f(x) - L| < \varepsilon.$$

**Example 3.2.** Consider the signum function

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

Then

$$\lim_{x \to 0^-} \operatorname{sgn}(x) = -1 \qquad \lim_{x \to 0^+} \operatorname{sgn}(x) = 1 \quad \lim_{x \to 0} \operatorname{sgn}(x) = \text{DNE}.$$

**Proposition 3.3** (Sequential Characterization of one sided limits)**.**

1. *Let $a$ be an accumulation point of $S \cap (-\infty, a)$. Then, $\lim\limits_{x \to a^-} f(x) = L$ if and only if $f(x_n) \to L$ whenever $(x_n)_n \subset S$ such that $x_n < a$ and $x_n \to a$.*

2. *Let $a$ be an accumulation point of $S \cap (a, \infty)$. Then, $\lim\limits_{x \to a^+} f(x) = L$ if and only if $f(x_n) \to L$ whenever $(x_n)_n \subset S$ such that $x_n > a$ and $x_n \to a$.*

*Proof.* The proof is nearly identical to the proof of Proposition 3.1. □

**Note 3.2.** The algebra of one sided limits is the same as in the limit case.

**Proposition 3.4.** *Let $f : S \mapsto \mathbb{R}$ and let $a$ be an accumulation point of both $S \cap (-\infty, a)$ and $S \cap (a, \infty)$. Then, $\lim_{x \to a} f(x)$ exists if and only if $f(a-)$ and $f(a+)$ exist and are equal. In this case, $\lim_{x \to a} f(x) = \lim_{x \to a^-} f(x) = \lim_{x \to a^+} f(x)$.*

*Proof.* Exercise. □

**Example 3.3.** Consider the step function $f(x) = \lfloor x \rfloor$. That is for any $n \in \mathbb{Z}$ and $n \geq x < n+1$, $f(x) = n$. Then for $n \in \mathbb{Z}$

$$\lim_{x \to n^-} f(x) = n - 1 \text{ and } \lim_{x \to n^+} f(x) = n \implies \lim_{x \to n} f(x) = \text{DNE}.$$

## 3.2 Continuous Functions

**Definition 3.3.** A function $f : S \subseteq \mathbb{R} \mapsto \mathbb{R}$ is said to be **continuous at** $x_0 \in S$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for $x \in S$

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

If $f$ is continuous at every $x \in S$, then we say $f$ is **continuous**.

**Note 3.3.** Note that for all $x_0 \in S$, $x$ is either an accumulation point of $S$ or an isolated point of $S$. If $x_0$ is an accumulation point of $S$, then $f$ is continuous at $x_0$ if and only if $\lim_{x \to x_0} f(x) = f(x_0)$. That is, the limit of $f$ as $x \to x_0$ exists and is equal to the value of the function at $f(x_0)$. If $x_0$ is an isolated point, then $f$ is always continuous at $x_0$, since $\exists \delta > 0$ such that $B(x_0, \delta) \cap S = \{x_0\}$. Thus, $x \in S$ and $|x - x_0| < \delta \implies x = x_0$, so $|f(x) - f(x_0)| = 0$, so this $\delta$ works for any $\varepsilon > 0$.

**Theorem 3.1.** *Let $f : S \subseteq \mathbb{R} \mapsto \mathbb{R}$ be a function. Then $f$ is continuous at $x_0 \in S$ if and only if $f(x_n) \to f(x_0)$ whenever $(x_n)_n \subset S$ and $x_n \to x_0$.*

*Proof.* The proof is nearly identical to Proposition 3.1. ☐

**Corollary 3.1.** *Let $f : S \mapsto \mathbb{R}$. If $(x_n)_n \subset S$ such that $x_n \to x_0 \in S$, but $(f(x_n))_n$ is divergent (no real limit), then $f$ is discontinuous at $x_0$.*

**Example 3.4.** Let $a \in \mathbb{R}$. Show that

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

is not continuous at $x = 0$.

*Solution.* Let $x_n = 1/(n\pi/2)$ for $n \geq 1$. Then $x_n \to 0$, but

$$f(x_n) = \{1, 0, -1, 0, 1, 0, -1, \ldots\}$$

which is divergent. Thus $f$ is not continuous at $x = 0$. ☐

**Example 3.5.** Prove that $f : \mathbb{R} \mapsto \mathbb{R}$ where $f(x) = x^2$ is continuous.

*Solution.* We will demonstrate this two ways. 1) Using the $\varepsilon - \delta$ definition of continuity and 2) By using the sequential characterization of continuity.

1) Let $x_0 \in \mathbb{R}$ and let $\varepsilon > 0$. Note that

$$\begin{aligned} |f(x) - f(x_0)| = |x^2 - x_0^2| &= |(x - x_0)(x + x_0)| \\ &\leq |x - x_0|(|x| + |x_0|) \\ &= |x - x_0|(|x - x_0 + x_0| + |x_0|) \\ &\leq |x - x_0|(|x - x_0| + 2|x_0|). \end{aligned}$$

Let $\delta = \min\{1, \varepsilon/(1 + 2|x_0|)\}$. Then $|x - x_0| < \delta$ implies

$$|f(x) - f(x_0)| \leq |x - x_0|(|x - x_0| + 2|x_0|) < \delta(1 + 2|x_0|) \leq \frac{\varepsilon}{1 + 2|x_0|}(1 + 2|x_0|) = \varepsilon.$$

2) Let $x_0 \in \mathbb{R}$ and let $(x_n)_n \subset \mathbb{R}$ be such that $x_n \to x_0$. Then by the algebra of limits

$$f(x_n) = x_n^2 \to x_0^2 = f(x_0).$$

Since $x_0$ and $(x_n)_n$ were arbitrary, $f$ is continuous at all $x_0$, and hence continuous.

☐

**Note 3.4.** Note that in the $\varepsilon - \delta$ definition, $\delta$ depends on both $x_0$ and $\varepsilon$. It is very helpful in some cases to be able to choose a delta that does not depend on $x_0$. This can be done for uniformly continuous functions that we will cover later, but not in general.

**Proposition 3.5** (Algebra of Continuity)**.** *Suppose $f, g : S \to \mathbb{R}$ are each continuous at $x_0 \in S$. Then*

*a) $f \pm g$ is continuous at $x_0$.*

*b)* $fg$ *is continuous at* $x_0$.

*c)* $f/g$ *is continuous at* $x_0$ *provided* $g(x_0) \neq 0$.

*Proof.* This follows immediatly from the sequential characterization of continuity and properties of limits. □

Recall the composition of two functions is written $g \circ f$ where

$$(g \circ f)(x) = g(f(x)).$$

**Theorem 3.2** (Composition of Continuous Funtions are Continuous)**.** *Let* $f : S \to S' \subset \mathbb{R}$ *and* $g : S' \mapsto \mathbb{R}$. *If* $f$ *is continuous at* $x_0 \in S$ *and* $g$ *is continuous at* $f(x_0) \in S'$, *then* $g \circ f$ *is continuous at* $x_0$.

*Proof.* Again, we provide two proofs: 1) Using the $\varepsilon - \delta$ definition of continuity at $x_0$ and 2) using the sequential characterization of continuity at $x_0$.

1) Let $\varepsilon > 0$. (We need to find a $\delta$ such that for $x \in S$ $|x - x_0| < \delta \implies |g(f(x)) - g(f(x_0))| < \varepsilon$.) Since $g$ is continuous at $f(x_0)$, we can choose a $\delta_1 > 0$ such that for $y \in S'$

$$|y - f(x_0)| < \delta_1 \implies |g(y) - g(f(x_0))| < \varepsilon.$$

$f$ continuous at $x_0$ implies $\exists \delta > 0$ such that for $x \in S$,

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \delta_1.$$

Thus for $x \in S$

$$|x - x_0| < \delta \implies |\underbrace{f(x)}_{\in S'} - f(x_0)| < \delta_1 \implies |g(f(x)) - g(f(x_0))| < \varepsilon.$$

2) Let $(x_n)_n \subset S$ be a sequence such that $x_n \to x_0$. Since $f$ is continuous at $x_0$, $f(x_n) \to f(x_0)$. Since $(f(x_n))_n \subset S'$, $f(x_n) \to f(x_0)$ and $g$ is continuous at $f(x_0)$, we have $g(f(x_n)) \to g(f(x_0))$. Thus $(x_n)_n \subset S$ and $x_n \to x_0$ implies

$$(g \circ f)(x_n) \to (g \circ f)(x_0).$$

□

Recall that the pre-image of a mapping $f : S \mapsto \mathbb{R}$ is defined by

$$f^{-1}(V) = \{x \in S | f(x) \in V\}.$$

**Theorem 3.3.** *Let* $f : S \mapsto \mathbb{R}$. *The following are equivalent:*

*a)* $f$ *is continuous.*

*b)* $f^{-1}(C) \subseteq S$ *is closed whenever* $C \subseteq \mathbb{R}$ *is closed.*

*c)* $f^{-1}(V) \subseteq S$ *is open whenever* $V \subseteq \mathbb{R}$ *is open.*

*Proof.* ($a$) $\implies$ $b$)). Let $C \subseteq \mathbb{R}$ be closed, and let $(x_n)_n \subset f^{-1}(C)$ be a sequence such that $x_n \to x \in S$. Then $(f(x_n))_n \subset C$, and since $f$ is continuous and $x_n \to x$, $f(x_n) \to f(x)$. Because $C$ is closed, $f(x) \in C$, so that $x \in f^{-1}(C)$. Thus $f^{-1}(C)$ is closed.

($b$) $\implies$ $c$)) Let $V \subset \mathbb{R}$ be open. Then $V^c$ is closed, and by $b$) $f^{-1}(V^c)$ is closed in $S$. Note that

$$f^{-1}(V^c) = \{x \in S | f(x) \notin V\} = \{x \in S | f(x) \in V\}^c = [f^{-1}(V)]^c.$$

Since $[f^{-1}(V)]^c$ is closed, $f^{-1}(V)$ is open.

($c$) $\implies$ $a$)) Let $\varepsilon > 0$ and let $x_0 \in S$. Since $B(f(x_0), \varepsilon)$ is open in $\mathbb{R}$, $f^{-1}(B(f(x_0), \varepsilon))$ is open in $S$. Note that $x_0 \in f^{-1}(B(f(x_0), \varepsilon)))$, so $\exists \delta > 0$ such that $B(x_0, \delta) \subset f^{-1}(B(f(x_0), \varepsilon)))$. Thus, for $x \in S$

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon,$$

so $f$ is continuous at $x_0$. Since $x_0$ was arbitrary, $f$ is continuous on $S$. □

## 3.3 Properties of Continuous Functions

We will need to recall a couple of facts from topology in metric spaces and $\mathbb{R}$:

a) (Heine-Borel) A set $K \subset \mathbb{R}$ is compact if and only if $K$ is closed and bounded.

b) A set $K$ is compact if and only if every sequence $(x_n)_n$ in $K$ has a convergent subsequence that converges to a point in $K$.

c) A set $C$ is closed if and only if $C$ contains all its accumulation points.

**Proposition 3.6.** *Suppose $f : S \mapsto \mathbb{R}$ is continuous. If $K \subseteq S$ is compact, then $f(K)$ is a compact subset of $\mathbb{R}$.*

*Proof.* Suppose $K \subseteq S$ is compact. Let $(y_n)_n$ be a sequence in $f(K) = \{f(x) : x \in K\}$. Then for each $n \geq 1$, $y_n = f(x_n)$ for some $x_n \in K$. Since $K$ is compact, $\exists$ a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $x_{n_k} \to x_0$ for some $x_0 \in K$. Since $f$ is continuous and $(x_{n_k})_k \subseteq K \subseteq S$,

$$x_{n_k} \to x_0 \implies y_{n_k} = f(x_{n_k}) \to f(x_0) = y_0 \in f(K).$$

Thus $(y_n)_n$ has a convergent subsequence $(y_{n_k})_k$ that converges to a point in $K$, so $f(K)$ is compact. □

**Corollary 3.2.** *If $K$ is compact and $f : K \mapsto \mathbb{R}$ is continuous, then $f$ has an absolute maximum and minimum on $K$.*

*Proof.* Since $K$ is compact and $f$ is continuous, $f(K)$ is compact, so $f(K)$ is closed and bounded. Let $\alpha = \inf_{x \in K} f(x)$ and $\beta = \sup_{x \in K} f(x)$. Since $f(K)$ is bounded, $\alpha$ and $\beta$ exists and are finite. Since $\alpha$ and $\beta$ are limit points of $f(K)$ and $f(K)$ is closed, $\alpha, \beta \in f(K)$. Thus, $\exists a, b \in K$ such that $f(a) = \alpha$ and $f(b) = \beta$. □

**Theorem 3.4** (Intermediate Value Theorem)**.** *Suppose $f : [a, b] \mapsto \mathbb{R}$ is continuous. If $y$ is between $f(a)$ and $f(b)$, then $\exists x_0 \in (a, b)$ such that $f(x_0) = y$.*

*Proof.* Suppose $f : [a, b] \mapsto \mathbb{R}$ is continuous and let $y$ be between $f(a)$ and $f(b)$. WLOG, suppose $f(a) < y < f(b)$. Let $S = \{t \in [a, b] | f(t) \leq y\}$. Then $a \in S$, so $S \neq \emptyset$ and is bounded. Thus $\sup S$ exists and is finite. Let $x_0 = \sup S$. By definition of supremum, $\forall n \in \mathbb{N}$, $\exists x_n \in S$ such that $x_0 - 1/n < x_n \leq x_0$. Thus, $\exists$ a sequence $(x_n)_n \subset S$ such that $|x_n - x_0| < 1/n$, so $x_n \to x_0$. Since $x_n \in S$, $\forall n \geq 1$, $f(x_n) \leq y$ for all $n \geq 1$. By continuity of $f$ at $x_0$,

$$x_n \to x_0 \implies f(x_n) \to f(x_0).$$

Since $f(x_n) \leq y$ for all $n \geq 1$, $f(x_0) \leq y$. If we can show that $f(x_0) \geq y$, then we would have $y \leq f(x_0) \leq y \implies f(x_0) = y$. Suppose, by contradiction, that $f(x_0) < y$. Let $\varepsilon = y - f(x_0) > 0$. Since $f$ is continuous at $x_0$, $\exists \delta > 0$ such that for $x \in [a, b]$

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

Thus for $x \in [a, b]$

$$x_0 - \delta < x < x_0 + \delta \implies 2f(x_0) - y < f(x) < y,$$

which implies $(x_0 - \delta, x_0 + \delta) \subset S$. But $x_0 = \sup S$, so $(x_0, x_0 + \delta) \not\subset S$. Thus $f(x_0) \geq y$, and we are done. □

## 3.4 Monotonic Functions

**Definition 3.4.** $f : S \mapsto \mathbb{R}$ is said to be left continuous at $x_0 \in S$ if $f(x_0-) = f(x_0)$. Similarly, $f$ is said to be right continuous at $x_0$ if $f(x_0+) = f(x_0)$.

**Note 3.5.** If $f(x-)$ and $f(x+)$ both exist, then the jump of $f$ at $x$ is

$$j(x) = |f(x+) - f(x-)|.$$

The jump of $f$ at $x$ is 0 if and only if $f(x+) = f(x-)$ if and only if $\lim_{t \to x} f(t)$ exists, and $j(x) > 0$ if and only if $\lim_{t \to x} f(t)$ DNE.

**Definition 3.5.** A function $f : S \mapsto \mathbb{R}$ is said to be **increasing** if $x \leq y$ implies $f(x) \leq f(y)$. $f$ is said to be **decreasing** if $x \leq y$ implies $f(y) \leq f(x)$. In either case, $f$ is said to be **monotonic**. $f$ is said to be **strictly monontonic** if it is either strictly increasing ($x < y \implies f(x) < f(y)$) or strictly decreasing ($x < y \implies f(y) < f(x)$).

**Proposition 3.7.**

*1) Let* $f : [a, b] \mapsto \mathbb{R}$ *be an increasing function. Then*

*a)* $f(x-) = \sup_{t<x} f(t)$, $\forall a < x \leq b$.

*b)* $f(x+) = \inf_{t>x} f(t)$, $\forall a \leq x < b$.

*That is for all* $a < x < b$, $f(x-)$ *and* $f(x+)$ *exist and*

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

*2) Let* $f : [a, b] \mapsto \mathbb{R}$ *be a decreasing function. Then* *a)* $f(x-) = \inf_{t<x} f(t)$, $\forall a < x \le b$.

*b)* $f(x+) = \sup_{t>x} f(t)$, $\forall a \le x < b$.

*That is for all* $a < x < b$, $f(x-)$ *and* $f(x+)$ *exist and*

$$\sup_{t>x} f(t) = f(x+) \le f(x) \le f(x-) = \inf_{t<x} f(t).$$

*3) If* $f$ *is monotonic, then* $f$ *is continuous at* $x$ *if and only if* $j(x) = 0$.

**Note 3.6.** For $f : [a, b] \mapsto \mathbb{R}$, we define

$$j(a) = |f(a+) - f(a)| \text{ and } j(b) = |f(b) - f(b-)|.$$

**Note 3.7.** If $f$ is monotonic, then the only possible discontinuity is a jump discontinuity.

*Proof.* 1) Suppose $f : [a, b] \mapsto \mathbb{R}$ is increasing and let $a < x \le b$. Then for all $a \le t < x$, $f(t) \le f(x)$, so $\sup_{t<x} f(t) \le f(x)$. Let $\varepsilon > 0$. Then by definition of supremum, $\exists a \le t_0 < x$ such that

$$\sup_{s<x} f(s) - \varepsilon < f(t_0) \le \sup_{s<x} f(s).$$

Let $\delta = x - t_0$. Then for $t_0 = x - \delta < t < x$

$$\sup_{s<x} f(s) - \varepsilon < f(t_0) \le f(t) \le \sup_{s<x} f(x).$$

Thus $x - \delta < t < x$ implies $|f(t) - \sup_{s<x} f(s)| < \varepsilon$. That is

$$\lim_{t\to x^-} f(t) = \sup_{s<x} f(s).$$

Similarly, for $x < t \le b$, $f(x) \le f(t) \implies f(x) \le \inf_{t>x} f(t)$. Let $\varepsilon > 0$. By definition of infimum, $\exists x < t_0 \le b$ such that

$$\inf_{s>x} f(s) \le f(t_0) < \inf_{s>x} f(s) + \varepsilon.$$

Take $\delta = t_0 - x$. Then $x < t < x + \delta = t_0$ implies

$$\inf_{s>x} f(s) \le f(t) \le f(t_0) < \inf_{s>x} f(s) + \varepsilon.$$

Thus $\lim_{t\to x^+} f(t) = \inf_{t>x} f(t)$.

2) Similar to 1).

3) Exercise.

□

**Corollary 3.3.** *If* $f : [a, b] \mapsto \mathbb{R}$ *is monotonic, then* $f$ *has at most countably many discontinuities.*

*Proof.* Suppose that $f$ is increasing, then $f(a) \leq f(x) \leq f(b)$ for $a \leq x \leq b$. By Proposition 3.7 (3),

$$D = \{x \in [a,b] | f \text{ is discontinuous at } x\} = \{x \in [a,b] | j(x) > 0\}.$$

Note that $j(x) > 0$ if and only if $j(x) > 1/n$ for some $n \geq 1$. Thus

$$D = \cup_{n=1}^{\infty} \{x \in [a,b] | j(x) > 1/n\}.$$

Let $D_n = \{x \in [a,b] | j(x) > 1/n\}$. Then

$$|D_n| \cdot \frac{1}{n} < \sum_{x \in D_n} j(x) \leq f(b) - f(a),$$

where $|D_n|$ = cardinality of $D_n$. Then

$$|D_n| \leq n(f(b) - f(a)),$$

so $D_n$ is finite. Thus $D$ is countable, since it is a countable union of countable sets. □

## 3.5 Uniform Continuity

**Definition 3.6.** A function $f : S \mapsto \mathbb{R}$ is said to be **uniformly continuous** on $S$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that for $x, y \in S$

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

**Note 3.8.** Note that $\delta$ depends on $\varepsilon$ but not on $x$ or $y$, whereas in the definition of continuity, $\delta$ depended on both $\varepsilon$ and $x_0$.

**Theorem 3.5.** *$f : S \mapsto \mathbb{R}$ is uniformly continuous if and only if $\forall (x_n)_n, (y_n)_n \subset S$ such that $|x_n - y_n| \to 0$, then $|f(x_n) - f(y_n)| \to 0$.*

*Proof.* ($\implies$) Suppose that $f$ is uniformly continuous. Let $(x_n)_n, (y_n)_n \subset S$ be sequences such that $|x_n - y_n| \to 0$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that for $x, y \in S$

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Choose $N \in \mathbb{N}$ such that $n \geq N$ implies $|x_n - y_n| < \delta$. Then for $n \geq N$

$$|f(x_n) - f(y_n)| < \varepsilon.$$

($\impliedby$) (Contrapositive) Suppose $f$ is not uniformly continuous. Then

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists x, y \in S \text{ such that } |x - y| < \delta \text{ but } |f(x) - f(y)| \geq \varepsilon.$$

Take $\varepsilon > 0$ such that the previous statement holds. Then for each $\delta_n = 1/n$, we can choose an $x_n, y_n \in S$ such that $|x_n - y_n| < 1/n$ but $|f(x_n) - f(y_n)| \geq \varepsilon$. Thus, we can construct sequences $(x_n)_n, (y_n)_n \subset S$ such that $|x_n - y_n| \to 0$ but $|f(x_n) - f(y_n)| \geq \varepsilon$ for all $n \geq 1$, i.e. $|f(x_n) - f(y_n)| \not\to 0$. □

**Example 3.6.** Show that $f(x) = x^2$ is not uniformly continuous.

*Solution.* Let $x_n = n + 1/n$ and $y_n = n$. Then

$$|x_n - y_n| = |(n + \frac{1}{n}) - n| = \frac{1}{n} \to 0,$$

but

$$|f(x_n) - f(y_n)| = |n^2 + 2 + \frac{1}{n^2} - n^2| = 2 + \frac{1}{n^2} \to 2 \neq 0.$$

Thus $f(x) = x^2$ is not uniformly continuous. □

**Theorem 3.6.** *If $f : S \mapsto \mathbb{R}$ is uniformly continuous, then $(f(x_n))_n$ is Cauchy whenever $(x_n)_n \subset S$ is Cauchy.*

*Proof.* Suppose $f : S \mapsto \mathbb{R}$ is uniformly continuous, and let $(x_n)_n \subset S$ be a Cauchy sequence. Let $\varepsilon > 0$ and choose $\delta > 0$ such that for $x, y \in S$

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Since $(x_n)_n$ is Cauchy, $\exists N \in \mathbb{N}$ such that $m > n \geq N$ implies

$$|x_m - x_n| < \delta.$$

Thus, for $m > n \geq N$ implies

$$|f(x_m) - f(x_n)| < \varepsilon,$$

so $(f(x_n))_n$ is Cauchy. □

**Example 3.7.** Show that $f : (0, 1] \mapsto [1, \infty)$ defined by $f(x) = 1/x$ is not uniformly continuous.

*Solution.* Recall that a sequence in $\mathbb{R}$ is Cauchy if and only if it converges to a point in $\mathbb{R}$. Let $x_n = 1/n$ for all $n \geq 1$. Then $x_n \to 0$, so $(x_n)_n$ is Cauchy, but

$$f(x_n) = n \to \infty.$$

Therefore, $(f(x_n))_n$ is not Cauchy, since it does not converge. □

The following propositions provide some conditions on $f$ that imply uniform continuity.

**Definition 3.7.** A function $f : S \mapsto \mathbb{R}$ is said to be **Lipschitz** if $\exists M > 0$ such that for all $x, y \in S$

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

**Proposition 3.8.** *If $f : S \mapsto \mathbb{R}$ is Lipschitz, then $f$ is uniformly continuous.*

*Proof.* Let $\varepsilon > 0$. Choose $M > 0$ such that for all $x, y \in S$

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

Take $\delta = \varepsilon/M$. Then for $x, y \in S$, $|x - y| < \delta$ implies

$$|f(x) - f(y)| \leq M|x - y| < M \cdot \delta = M \cdot \frac{\varepsilon}{M} = \varepsilon.$$

Thus, $f$ is uniformly continuous. □

**Proposition 3.9.** *If $f : [a,b] \mapsto \mathbb{R}$ is continuous and differentiable on $(a,b)$ with $f'$ bounded on $(a,b)$, then $f$ is Lipschitz and thus uniformly continuous.*

*Proof.* Let $M > 0$ be such that $|f'(t)| \leq M$ for all $t \in (a,b)$. Let $x, y \in S$. Then by the Mean Value Theorem, $\exists t = t(x,y)$ between $x$ and $y$ such that

$$f'(t) = \frac{f(y) - f(x)}{y - x} \implies |f(y) - f(x)| = |f'(t)||x - y| \leq M|x - y|.$$

Thus $f$ is Lipschitz, and so it is uniformly continous by the previous Proposition. □

**Proposition 3.10.** *If $S$ is compact and $f : S \mapsto \mathbb{R}$ is continuous, then $f$ is uniformly continuous on $S$.*

*Proof.* Let $f : S \mapsto \mathbb{R}$ be continuous and let $S$ be compact. Suppose that $f$ is not uniformly continuous. Then $\exists \varepsilon > 0$ and sequences $(x_n)_n, (y_n)_n \subset S$ such that $|x_n - y_n| \to 0$ but $|f(x_n) - f(y_n)| \geq \varepsilon$ for all $n \geq 1$. Since $S$ is compact, $\exists$ a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $x_n \to x \in S$. Note that

$$|y_{n_k} - x| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - x| \to 0,$$

so $y_{n_k} \to x$ also. Since $f$ is continuous, $|f(x_{n_k}) - f(x)| \to 0$ and $|f(y_{n_k}) - f(x)| \to 0$ as $k \to \infty$. Therefore,

$$|f(x_{n_k}) - f(y_{n_k})| \leq |f(x_{n_k}) - f(x)| + |f(x) - f(y_{n_k})| \to 0,$$

but for all $k \geq 1$

$$|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon,$$

a contradiction. Hence, $f$ is uniformly continuous. □

## 3.6 Applications in Probability and Statistics

- Continuous mapping theorems and common mistakes with convergence in probability.

# Chapter 4

# Sequences and Series of Functions

## 4.1 Power Series

**Definition 4.1.** Give a sequence $(a_n)_n \subset \mathbb{R}$, then

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

is called a **power series** centered at $x_0$. The domain is given by $\{x \in \mathbb{R} | f(x) \text{ converges}\}$.

**Theorem 4.1.** *Given the power series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$, let*

$$\alpha = \overline{\lim_{n\to\infty}} |a_n|^{1/n}$$

*and set $R = 1/\alpha$, where $R = \infty$ if $\alpha = 0$ and $R = 0$ if $\alpha = \infty$. Then $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ converges for $|x - x_0| < R$ and diverges if $|x - x_0| > R$.*

*Proof.* Fix $x \in \mathbb{R}$. Then by the Root Test $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ converges absolutely if

$$\overline{\lim_{n\to\infty}} |a_n (x - x_0)^n|^{1/n} = |x - x_0| \overline{\lim_{n\to\infty}} |a_n|^{1/n} = |x - x_0|\alpha < 1,$$

which happens

a) if $\alpha = 0$, so the series converges for all $x \in \mathbb{R}$.

b) if $0 < \alpha < \infty$ whenever $|x - x_0| < 1/\alpha = R$, so that the series converges for $x \in (x_0 - R, x_0 + R)$.

c) if $\alpha = \infty$ only for $|x - x_0| = 0$.

Similarly, by the root test, the series diverges if $|x - x_0|\alpha > 1 \iff |x - x_0| > R$. □

**Definition 4.2.** For the power series $\sum_{n=0}^{\infty} a_n (x - x_0)^n$, $R = 1/\overline{\lim}_{n\to\infty} |a_n|^{1/n}$ is called the **radius of convergence**.

**Example 4.1.** Find the radius of convergence of the following power series.

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

*Solution.* Since

$$\overline{\lim_{n\to\infty}} |1|^{1/n} = 1,$$

the radius of convergence os $R = 1$, so the series converges for $|x| < 1$ and diverges for $|x| > 1$. For the case $|x| < 1$, note that this is the geometric series, and we proved earlier that $f(x) = \frac{1}{1-x}$ for $|x| < 1$. The series diverges for both $x = 1$ and $x = -1$, so the series only converges for $|x| < 1$. □

b) $f(x) = \sum_{n=2}^{\infty} \frac{1}{n(n-1)} x^n$

*Solution.* Since

$$\overline{\lim_{n\to\infty}} \left| \frac{1}{n^{1/n}(n-1)^{1/n}} \right| = 1,$$

the radius of converges is $R = 1$, so the series converges for $|x| < 1$ and diverges for $|x| > 1$. If $|x| = 1$, then the series converges since

$$\sum_{n=2}^{\infty} \frac{1}{n(n-1)}$$

converges by the integral test. Thus, the series converges for $|x| \leq 1$. □

c) $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n}$

*Solution.* Since

$$\overline{\lim_{n\to\infty}} \left| \frac{1}{n^{1/n}} \right| = 1,$$

the series converges for $|x| < 1$ and diverges for $|x| > 1$. By the integral test, the series diverges for $x = 1$, but it converges for $x = -1$ by the alternating series test. Thus, the series converges for $x \in [-1, 1)$. □

d) $f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

*Solution.* Since

$$\overline{\lim_{n\to\infty}} \left| \frac{1}{n!} \right|^{1/n} \leq \overline{\lim_{n\to\infty}} \left| \frac{1/(n+1)!}{1/n!} \right| = \overline{\lim_{n\to\infty}} \frac{1}{n+1} = 0,$$

the radius of converges is $R = \infty$, so the series converges for all $x \in \mathbb{R}$. □

e) $f(x) = \sum_{n=0}^{\infty} n! x^n$

*Solution.* Since

$$\overline{\lim_{n\to\infty}} |n!|^{1/n} = \infty,$$

the radius of convergence is $R = 0$, and the series converges only for $x = 0$. □

## 4.2 Uniform Convergence

**Definition 4.3.** Let $f_n : S \subseteq \mathbb{R} \mapsto \mathbb{R}$. We say that the sequence of functions $(f_n)_n$ converges **pointwise** to a function $f : S \mapsto \mathbb{R}$ if

$$f_n(x) \to f(x)$$

for each $x \in S$.

**Example 4.2.** Consider the power series $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$. If $R$ is its radius of convergence, then we showed that the sequence of partial sum functions

$$s_n(x) = \sum_{k=0}^{n} a_k(x - x_0)^k$$

converges pointwise to $f(x)$ for each $x \in \mathbb{R}$ such that $|x - x_0| < R$.

The questions that we might now ask is do the properties of our sequence of functions carry over to the limit function. For examples,

a) If $\{f_n, n \geq 1\}$ are continuous and $f_n \to f$, is $f$ also continuous?

b) If $\{f_n, n \geq 1\}$ are differentiable and $f_n \to f$, is $f$ also differentiable?

c) If $\{f_n, n \geq 1\}$ are integrable and $f_n \to f$, is $f$ also integrable?

Pointwise convergence turns out to be insufficient for these properties to carry over to the limit function in general.

**Example 4.3.** Consider the sequence of functions $f_n : \mathbb{R} \mapsto \mathbb{R}$ defined by

$$f_n(x) = \sum_{k=0}^{n} \frac{x^2}{(1 + x^2)^k}$$

and its limit function

$$f(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1 + x^2)^k}.$$

Note that $f_n(0) = 0 \to 0 = f(0)$. For each $x \neq 0$, the series is a geometric series, which converges to

$$\sum_{n=0}^{\infty} \frac{x^2}{(1 + x^2)^n} = x^2 \cdot \frac{1}{1 - \frac{1}{1+x^2}} = 1 + x^2.$$

Thus $f_n \to f$ pointwise where

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

Note that $f_n(x)$ is continuous but $f$ is not continuous at 0. Thus, we have a sequence of continuous function that converge pointwise, but the limit function is not continuous.

The problem here becomes one of interchanging limts. For example, consider a sequence of continuous functions $f_n$ that converge pointwise to $f$. For $f$ to be continuous at $x \in \mathbb{R}$, we would need

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

Since $f_n$ are continuous at $x$ and converge pointwise to $f$, it is also true that

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

Thus, it becomes a question of when can we interchange the two limit operations and say that

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

The following example shows that we cannot always interchange limits freely.

**Example 4.4.** Consider the double array $\{x_{m,n}, m \geq 1, n \geq 1\}$ defined by

$$x_{m,n} = \frac{m}{m+n}.$$

Then for each $n \geq 1$,

$$\lim_{m\to\infty} x_{m,n} = 1,$$

so that

$$\lim_{n\to\infty} \lim_{m\to\infty} x_{m,n} = 1.$$

However, for each fixed $m \geq 1$,

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

so that

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

Thus

$$\lim_{m\to\infty} \lim_{n\to\infty} x_{m,n} \neq \lim_{n\to\infty} \lim_{m\to\infty} x_{m,n}.$$

We illustrate with one more example that pointwise convergence of functions is not strong enough to guarantee convergence of integrals.

**Example 4.5** (Witch's Hat)**.** Consider the functions $f_n : [0,2] \mapsto \mathbb{R}$ defined by

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

Then $f_n(x) \to 0$ (pointwise) for each each $x \in [0,2]$, but for all $n \geq 1$

$$\int_0^2 f_n(x)\, dx = 1.$$

Thus

$$1 = \int_0^2 f_n(x)\, dx \not\to \int_0^2 f(x)\, dx = 0,$$

so that pointwise convergence of integrable functions does not imply the convergence of their integrals.

We will now introduce a stronger mode of convergence, that will allow us to establish conditions where these properties can carry over from the sequence of functions to the limit.

**Definition 4.4.** Let $f_n : S \mapsto \mathbb{R}$ and $f : S \mapsto \mathbb{R}$ be functions. We say that $f_n$ converges uniformly to $f$ if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $n \geq N$ implies

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

for all $x \in S$.

**Note 4.1.** In the definition of uniform convergence, $N$ depends on $\varepsilon$ but not $x$, whereas in pointwise convergence $N$ would depend on both $\varepsilon$ and $x$. It should also be clear that if a series of functions converges uniformly then it also converges pointwise.

**Proposition 4.1** (Cauchy Criterion for Uniform Convergence)**.** *The sequence of functions $f_n : S \mapsto \mathbb{R}$ converges uniformly to $f : S \mapsto \mathbb{R}$ if and only if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $m > n \geq N$ implies*

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

*for all $x \in S$.*

*Proof.* ($\implies$) Suppse $f_n \to f$ uniformly in $S$. Let $\varepsilon > 0$, and choose $N \in \mathbb{N}$ such that for all $x \in S$,

$$n \geq N \implies |f_n(x) - f(x)| < \frac{\varepsilon}{2}.$$

Then for all $x \in S$, $m > n \geq N$ implies

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

($\impliedby$) Suppse $(f_n)_n$ is uniformly Cauchy. Then for each $x \in \mathbb{R}$, $(f_n(x))_n$ is a Cauchy sequence in $\mathbb{R}$, so it converges to some point $a_x \in \mathbb{R}$. Define $f : S \mapsto \mathbb{R}$ by $f(x) = \lim_{n\to\infty} f_n(x) = a_x$ (pointwise). Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that for all $x \in S$, $m > n \geq N$ implies

$$|f_m(x) - f_n(x)| < \varepsilon.$$

Then for $n \geq N$, we have for all $x \in S$

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

so $f_n \to f$ uniformly. □

**Proposition 4.2.** *Suppose that $f_n, f : S \mapsto \mathbb{R}$ are functions such that $f_n \to f$ pointwise. Let*

$$M_n = \sup_{x\in S} |f_n(x) - f(x)|.$$

*Then $f_n \to f$ uniformly on $S$ if and only if $M_n \to 0$.*

*Proof.* ($\implies$) Suppose $f_n \to f$ uniformly on $S$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that for all $x \in S$ and for all $n \geq N$

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

then for $n \geq N$

$$M_n = \sup_{x \in S} |f_n(x) - f(x)| \leq \varepsilon.$$

( $\Longleftarrow$ ) Suppose $M_n \to 0$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $n \geq N$ implies

$$M_n = \sup_{x \in S} |f_n(x) - f(x)| < \varepsilon.$$

Then for all $x \in S$ and for all $n \geq N$

$$|f_n(x) - f(x)| \leq M_n < \varepsilon.$$

□

Note that all the results for sequences of functions carry over to series of functions

$$f(x) = \sum_{n=1}^{\infty} f_n(x)$$

since the series is defined by the limit of the seqeuence of partial sum functions

$$s_n(x) = \sum_{k=1}^{n} f_k(x).$$

The following theorem provides conditions for which a series of functions converges uniformly on a set $S$.

**Theorem 4.2** (Weierstrass M-test)**.** *Let $(M_n)_n$ be a a sequence of positive real numbers such that $\sum_{n=1}^{\infty} M_n < \infty$. If $f_n : S \mapsto \mathbb{R}$ is a sequence of functions such that for each $n \geq 1$, $|f_n(x)| \leq M_n$ for all $x \in S$, then the series $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly on $S$.*

*Proof.* We will prove uniform convergence by showing that the Cauchy criterion for uniform convergence holds. Let $\varepsilon > 0$ and let $s_n(x) = \sum_{k=1}^{n} f_k(x)$. Since $\sum_{n=1}^{\infty} M_k < \infty$, we can choose an $N \in \mathbb{N}$ such that $m > n \geq N$ implies

$$\sum_{k=n+1}^{m} M_k < \varepsilon.$$

Then for $m > n \geq N$ and for all $x \in S$

$$\left| \sum_{k=n+1}^{m} f_k(x) \right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k < \varepsilon.$$

□

**Example 4.6.** Show that $\sum_{n=1}^{\infty} \frac{x^n}{n^3 + nx^n}$ converges uniformly on $[0, 1]$.

*Solution.* Let
$$f_n(x) = \frac{x^n}{n^3 + nx^n},\ 0 \leq x \leq 1.$$
Then
$$f_n'(x) = \frac{nx^{n-1}}{n^3 + nx^n} - \frac{x^n \cdot n^2 x^{n-1}}{(n^3 + nx^n)^2} = \frac{n^4 x^{n-1}}{(n^3 + nx^n)^2} \geq 0,\ 0 \leq x \leq 1,$$
so each $f_n$ is an increasing function and
$$|f_n(x)| \leq \frac{1}{n^3 + n} := M_n, \quad \text{for all } 0 \leq x \leq 1.$$
Since
$$\sum_{n=1}^{\infty} M_n \leq \sum_{n=1}^{\infty} \frac{1}{n^3},$$
and $\sum_{n=1}^{\infty} 1/n^3$ converges by the integral test, $\sum_{n=1}^{\infty} M_n$ converges by the comparison test. Therefore,
$$\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^3 + nx^n}$$
converges uniformly on $[0, 1]$ by the M-test. □

## 4.3 Uniform Convergence and Continuity

We will come back to conditions under which a limit function is differentiable or integrable if the sequence of functions is differentiable or integrable, respectively. For now, we provide some conditions under which the limit of a sequence of continuous functions is continuous. First, we state a more general result about a uniformly converging sequence of functions and interchanging limits.

**Theorem 4.3.** *Let $f_n, f : S \mapsto \mathbb{R}$ be functions and suppose that $f_n \to f$ uniformly on $S$. Let $x$ be an accumulation point of $S$, and suppose that $\lim_{t \to x} f_n(t) = A_n$. Then $(A_n)_n$ converges and*
$$\lim_{t \to x} f(t) = \lim_{n \to \infty} A_n.$$

**Note 4.2.** The previous theorem states that for a sequence of uniformly convergent functions we have for an accumulation point $x$ such that $\lim_{t \to x} f_n(x)$ exists for each $n \geq 1$
$$\lim_{t \to x} \lim_{n \to \infty} f_n(t) = \lim_{n \to \infty} \lim_{t \to x} f_n(t).$$

*Proof.* Let $\varepsilon > 0$. Since $f_n \to f$ uniformly on $S$, there exists an $N \in \mathbb{N}$ such that $m > n \geq N$ implies for all $t \in S$
$$|f_m(t) - f_n(t)| < \varepsilon.$$
Thus for $m > n \geq N$,
$$|A_m - A_n| = \left| \lim_{t \to x} f_m(t) - \lim_{t \to x} f_n(t) \right| = \lim_{t \to x} |f_m(t) - f_n(t)| \leq \varepsilon.$$
Hence, $(A_n)_n$ is a Cauchy sequence in $\mathbb{R}$, so it is convergent. Let $A = \lim_{n \to \infty} A_n$. Next, note that for any $n \geq 1$
$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|.$$

Let $N_1 \in \mathbb{N}$ be such that $n \geq N_1$ implies

$$|f_n(t) - f(t)| < \frac{\varepsilon}{3}, \text{ for all } t \in S,$$

and let $N_2 \in \mathbb{N}$ be such that $n \geq N_2$ implies

$$|A_n - A| < \frac{\varepsilon}{3}.$$

Now, let $N = \max\{N_1, N_2\}$ and fix $n_0 \geq N$. Since $\lim_{t\to x} f_{n_0}(t) = A_{n_0}$, there exists a $\delta > 0$ such that $0 < |t - x| < \delta$ implies

$$|f_{n_0}(t) - A_{n_0}| < \frac{\varepsilon}{3}.$$

Hence for $0 < |t - x| < \delta$, we have

$$|f(t) - A| \leq |f(t) - f_{n_0}(t)| + |f_{n_0}(t) - A_{n_0}| + |A_{n_0} - A| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

□

**Corollary 4.1.** *If $f_n : S \mapsto \mathbb{R}$ are continuous functions on $S$ and $f_n \to f$ uniformly on $S$, then $f$ is continuous.*

*Proof.* Let $x$ be an accumulation point of $S$. Then by continuity of $f_n$

$$\lim_{t\to x} f_n(t) = f_n(x), \ \forall n \geq 1.$$

Since $f_n \to f$ uniformly on $S$, we have $\lim_{n\to\infty} f_n(x) = f(x)$, and by Theorem 4.3 that

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

Together these imply

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

□

**Example 4.7.** Show that $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^3+nx^n}$ is continuous on $[0, 1]$.

**Example 4.8.** We showed previously that the series converges uniformly on $[0, 1]$ by the M-test. Since for each $n \geq 1$

$$s_n(x) = \sum_{k=1}^{n} \frac{x^n}{n^3 + nx^n}$$

is continuous, we have that $f$ is also continuous on $[0, 1]$ by our Corollary 4.1.

We now return to power series. Recall that power series are series of the form

$$\sum_{n=0}^{\infty} a_n x^n$$

which is a limit of polynomial functions. Since polynomial functions are continuous on their domain, we might expect power series to be continuous as well. This is the topic of the next few results.

**Proposition 4.3.** *Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with radius of convergence $0 < R \leq \infty$. If $0 < \rho < R$, then the power series converges uniformly on $[-\rho, \rho]$.*

*Proof.* Let $R > 0$ be the radius of convergence of $\sum_{n=0}^{\infty} a_n x^n$, and let $0 < \rho < R$. Then the series converges absolutely for all $x \in (-R, R)$, so it also converges absolutely for $x \in [-\rho, \rho]$. Note that for each $n \geq 0$,

$$|a_n x^n| \leq |a_n|\rho^n (:= M_n), \text{ for all } x \in [-\rho, \rho],$$

and $\sum_{n=1}^{\infty} |a_n|\rho^n$ converges. Then by the M-test, $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-\rho, \rho]$. □

**Corollary 4.2.** *The power series $\sum_{n=0}^{\infty} a_n x^n$ with radius of convergence $R > 0$ is a continuous function on $(-R, R)$.*

*Proof.* Let $x_0 \in (-R, R)$. Then $\exists \rho > 0$ such that $x_0 \in [-\rho, \rho] \subset (-R, R)$. Since $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-\rho, \rho]$ and $s_n(x) = \sum_{k=0}^{n} a_k x^k$ are continuous on $[-\rho, \rho]$ for each $n \geq 0$, $\sum_{n=0}^{\infty} a_n x^n$ is continuous on $[-\rho, \rho]$. In particular, since $x_0 \in [-\rho, \rho]$, $\sum_{n=0}^{\infty} a_n x^n$ is continuous at $x_0$. Therefore the series is continuous on $(-R, R)$ since $x_0 \in (-R, R)$ was arbitrary. □

**Note 4.3.** Though it might, the power series need not converge uniformly over $(-R, R)$ itself, even though it converges uniformly for every $[-\rho, \rho] \subset (-R, R)$. Regardless, the conclusion that the series converges to a continuous function over $(-R, R)$ holds.

**Example 4.9.** Show that $\sum_{n=0}^{\infty} x^n/2^n$ is continuous on $(-2, 2)$ but does not converge uniformly on $(-2, 2)$.

*Solution.* Note that this is a power series with $a_n = 2^{-n}$. The radius of convergence is given by

$$R = 1/\overline{\lim_{n\to\infty}} |1/2^n|^{1/n} = 2.$$

Then we immediately have that $\sum_{n=0}^{\infty} 2^{-n} x^n$ is continuous on $(-2, 2)$, but for $m > n$,

$$\sup_{x\in(-2,2)} \left| \sum_{k=n+1}^{m} 2^{-k} x^k \right| = \sum_{k=n+1}^{m} 2^{-k} 2^k = m - n.$$

This shows that the sequence of partial sums is not uniformly Cauchy and hence does not converge uniformly on $(-2, 2)$. □

The next result addresses when we can further conclude continuity at $x = R$ and $x = -R$.

**Proposition 4.4** (Abel's Theorem)**.** *Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$ be a power series with finite radius of convergence $0 < R < \infty$. If the series converges at $x = R$, then $f$ is continuous at $x = R$. If the series converges at $x = -R$, then $f$ is continuous at $x = -R$.*

*Proof.* Case 1: First, suppose $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $R = 1$, and that it converges at $x = 1$. Let $s_n(x) = \sum_{k=0}^{n} a_k x^k$ and let $d_n = \sum_{k=0}^{n} a_k = s_n(1)$ for $n = 0, 1, 2, \ldots$. Let $d = \sum_{k=0}^{\infty} a_k = f(1)$. For $0 < x < 1$, we have

$$s_n(x) = \sum_{k=0}^{n} a_k x^k = d_0 + \sum_{k=1}^{n} (d_k - d_{k-1}) x^k$$

$$= d_0 + \sum_{k=1}^{n} d_k x^k - x \sum_{k=1}^{n} d_{k-1} x^{k-1}$$
$$= d_0 + \sum_{k=1}^{n} d_k x^k - x \sum_{k=0}^{n-1} d_k x^k$$
$$= d_0 + d_n x^n + \sum_{k=1}^{n-1} d_k (1-x) x^k - x d_0$$
$$= \sum_{k=0}^{n-1} d_k (1-x) x^k + d_n x^n$$

Letting $n \to \infty$, we have for $0 < x < 1$

$$f(x) = \lim_{n\to\infty} s_n(x) = \sum_{k=0}^{\infty} d_k (1-x) x^k + d \cdot 0 = \sum_{n=0}^{\infty} d_n (1-x) x^n.$$

Since for any $0 < x < 1$, $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$, we have $1 = \sum_{n=0}^{\infty} (1-x) x^n$ for any $0 < x < 1$, and

$$f(1) = d = d \sum_{n=0}^{\infty} (1-x) x^n = \sum_{n=0}^{\infty} d(1-x) x^n, \ 0 < x < 1.$$

Hence we have for $0 < x < 1$

$$f(1) - f(x) = \sum_{n=0}^{\infty} (d - d_n)(1-x) x^n.$$

Let $\varepsilon > 0$. Since $d_n \to d$, there exists an $N \in \mathbb{N}$ such that $n \geq N$ implies $|d_n - d| < \varepsilon/2$. Let $g_N(x) = \sum_{n=0}^{N} |d - d_n|(1-x) x^n$. Then for $0 < x < 1$, we obtain

$$|f(1) - f(x)| \leq g_N(x) + \sum_{k=N+1}^{\infty} |d - d_k|(1-x) x^k$$
$$\leq g_N(x) + \sum_{k=N+1}^{\infty} \frac{\varepsilon}{2} (1-x) x^k < g_N(x) + \frac{\varepsilon}{2}.$$

Note that $g_N(x)$ is continuous and $g_N(1) = 0$. Thus, there exists a $\delta > 0$ such that $1 - \delta < x < 1$ implies $g_N(x) < \varepsilon/2$. Then, for $1 - \delta < x < 1$

$$|f(1) - f(x)| < g_N(x) + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

so $f$ is continuous at $x = 1$.

Case 2: Now, suppose that $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $0 < R < \infty$, and that the series converges at $x = R$. Let $g(x) = f(Rx)$, and note that

$$g(x) = \sum_{n=0}^{\infty} a_n R^n x^n,$$

which has radius of convergence 1, and it converges at $x = 1$. Then $g(x)$ is continuous at $x = 1$ by case 1. Since $f(x) = g(x/R)$, it follows that $f$ is cotinuous at $x = R$.

Case 3: Suppose $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $0 < R < \infty$, and that the series converges at $x = -R$. Let $h(x) = f(-x)$ and note that

$$h(x) = \sum_{n=0}^{\infty} (-1)^n a_n x^n$$

which has radius of convergence $R$ and converges at $R$, so $h$ is continuous at $x = R$ by case 2. If follows that $f(x) = h(-x)$ is continuous at $x = -R$. □

## 4.4 Applications in Probability and Statistics

- Convergence in distribution and Polya's theorem.

# Chapter 5

# Differentiation

## 5.1 Basic Properties of Derivatives

**Definition 5.1.** A function $f : (a, b) \mapsto \mathbb{R}$ is said to be **differentiable** at a point $c \in (a, b)$ if

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

exists and is finite. In such a case, we write $f'(c)$ for the derivative of $f$ at $c$. If $f$ is differentiable at all $x \in D \subseteq (a, b)$, then we say $f$ is differentiable on $D$ and write $f' : D \subseteq (a, b) \mapsto \mathbb{R}$ for the derivative of $f$ and denote this function $f'$ or $\dfrac{d}{dx} f(x)$.

**Note 5.1.** Note that

$$\frac{f(x) - f(x)}{x - c}$$

is the slope of the secant line through the points $(x, f(x))$ and $(c, f(c))$. As $x \to c$, if $f$ is differentiable at $c$, then this approaches the slope of the tangent line to $f$ at $(c, f(c))$ given by

$$L(x) = f(c) + m(c)(x - c).$$

Thus, $f$ is differentiable at $c$ if and only if there exists a linear function $L(x) = f(c) + m(x - c)$ such that

$$\left| \frac{f(x) - L(c)}{x - c} \right| \to 0 \text{ as } x \to c.$$

In this case, $f'(c) = m$ and

$$\left| \frac{f(x) - L(x)}{x - c} \right| = \left| \frac{f(x) - f(c)}{x - c} - m \right| \to 0 \text{ as } x \to c,$$

and $f'(c) = m$. $L(x)$ is a linear approximation of $f$ at $c$ when $|x - c|$ is small.

**Note 5.2.** An equivalent definition of the derivative of $f$ at $c$ is

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

provided the limit exists and is finite.

**Example 5.1.** Show that $f : \mathbb{R} \mapsto \mathbb{R}$ defined by $f(x) = x^n$, $n \in \mathbb{N}$ is differentiable everywhere and that $f'(c) = nc^{n-1}$.

*Solution.* Note that for $n \in \mathbb{N}$

$$x^n - c^n = (x-c)\sum_{k=0}^{n-1} x^k c^{n-k-1}.$$

Then

$$\lim_{x\to c} \frac{f(x)-f(c)}{x-c} = \lim_{x\to c} \sum_{k=0}^{n-1} x^k c^{n-k-1} = \sum_{k=0}^{n-1} c^{n-1} = nc^{n-1}.$$

□

**Theorem 5.1.** *If $f : (a,b) \mapsto \mathbb{R}$ is differentiable at $c \in (a,b)$, then $f$ is continuous at $c$.*

*Proof.* Let $\varepsilon > 0$. Choose $\delta_1 > 0$ such that $x \in (a,b)$ and $0 < |x-c| < \delta_1$ implies

$$\left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < \frac{\sqrt{\varepsilon}}{2}.$$

Let $\delta = \min\{\delta_1, \sqrt{\varepsilon}, \varepsilon/(2|f'(c)|)\}$, where we define $\varepsilon/0 = \infty$. Then for $x \in (a,b)$, $0 < |x-c| < \delta$ implies

$$\begin{aligned} |f(x)-f(c)| &= \left| \frac{f(x)-f(c)}{x-c} \right| |x-c| \\ &\le \left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| |x-c| + |f'(c)||x-c| \\ &< \frac{\sqrt{\varepsilon}}{2} \cdot \sqrt{\varepsilon} + |f'(c)|\delta \le \varepsilon. \end{aligned}$$

□

**Note 5.3.** Alternatively, we can prove that differentiability implies continuity using the properties of limits of functions. Since $c$ is an accumulation point of $(a,b)$, recall that $f$ is continuous at $c$ if

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

or equivalently,

$$|f(x)-f(c)| \to 0 \quad \text{as} \quad x \to c.$$

Note that $f$ differentiable at $c$ implies

$$\lim_{x\to c} \left| \frac{f(x)-f(c)}{x-c} \right| = |f'(c)|,$$

so for $x \neq c$

$$|f(x)-f(c)| = \left| \frac{f(x)-f(c)}{x-c} \right| |x-c| \to |f'(c)| \cdot 0 = 0 \quad \text{as} \quad x \to c.$$

**Definition 5.2.** For a function $f : [a, b] \mapsto \mathbb{R}$, the **right hand derivative of** $f$ **at** $c$ for $a \leq c < b$ is

$$f'(c+) = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c}$$

exists and is finite. The **left hand derivative of** $f$ **at** $c$, $a < c \leq b$ is

$$f'(c-) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c}$$

exist and is finite. We say that $f$ is differentiable in $[a, b]$ if $f$ is differentiable on $(a, b)$ and $f$ has a left hand derivative at $b$ and right hand derivative at $a$.

**Note 5.4.** For $c \in (a, b)$, $f'(c)$ exists if and only if $f'(c+)$ and $f'(c-)$ both exists, are finite, and equal.

**Definition 5.3.** For $f : (a, b) \mapsto \mathbb{R}$, if for $c \in (a, b)$

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

then we write $f'(c) = \infty$. We define similarly $f'(c) = -\infty$, $f'(c+) = \pm\infty$ and $f'(c-) = \pm\infty$. When we say that $f$ is differentiable at $c$, we allow for the derivative to be infinite. In cases where we require the derivative to be finite, we will say differentiable and finite.

**Example 5.2.** Define $f : \mathbb{R} \mapsto \mathbb{R}$ by $f(x) = x^{1/3}$. Then for $x \neq 0$

$$\frac{f(x) - f(0)}{x - 0} = \frac{x^{1/3} - 0}{x - 0} = x^{-2/3},$$

so

$$f'(0) = \lim_{x \to 0} x^{-2/3} = \infty.$$

If we let $g(x) = x^{-2/3}$, then

$$g'(0-) = -\infty \text{ and } g'(0+) = \infty.$$

**Proposition 5.1.** *Suppose that* $f, g : (a, b) \mapsto \mathbb{R}$ *are differentiable at* $x \in (a, b)$. *Then*

*a)* $(f + g)'(x) = f'(x) + g'(x)$.

*b)* $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.

*c)* $(f/g)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$ *provided* $g(x) \neq 0$

*Proof.* a) Suppose that $f$ and $g$ are differentiable at $x \in (a, b)$. Then, for $h \neq 0$

$$\frac{(f + g)(x + h) - (f + g)(x)}{h} = \frac{f(x + h) - f(x)}{h} + \frac{g(x + h) - g(x)}{h} \to f'(x) + g'(x)$$

as $h \to 0$.

b) Suppose that $f$ and $g$ are differentiable at $x \in (a, b)$. Then $f$ and $g$ are continuous at $x$, so for $h \neq 0$

$$\begin{aligned}\frac{(fg)(x+h)-(fg)(x)}{h} &= \frac{(fg)(x+h)-f(x)g(x+h)+f(x)g(x+h)-(fg)(x)}{h} \\ &= \frac{f(x+h)-f(x)}{h}\cdot g(x+h)+f(x)\cdot\frac{g(x+h)-g(x)}{h} \\ &\to f'(x)g(x)+f(x)g'(x) \quad \text{as } h\to 0.\end{aligned}$$

c) Suppose that $f$ and $g$ are differentiable at $x \in (a, b)$ and that $g(x) \neq 0$. Since $g$ is continuous at $x$ and $g(x) \neq 0$, $\exists \delta > 0$ such that for $|x - t| < \delta$, $g(t) \neq 0$. Thus for $-\delta < h < \delta$,

$$\begin{aligned}\frac{\frac{1}{g(x+h)}-\frac{1}{g(x)}}{h} &= \frac{g(x)-g(x+h)}{hg(x)g(x+h)} \\ &= -\frac{g(x+h)-g(x)}{h}\cdot\frac{1}{g(x)g(x+h)} \\ &\to -\frac{g'(x)}{[g(x)]^2} \quad \text{as } h\to 0.\end{aligned}$$

The result follows by applying part b).

□

**Note 5.5.** By part b), if $g(x) = c$ for some $c \in \mathbb{R}$ and for all $x \in (a, b)$ (i.e. $g$ is a constant function), then $(cf)'(x) = cf'(x)$. Furthermore, combining parts a) and b), differentiation is linear, i.e.

$$(af + bg)'(x) = af'(x) + bg'(x).$$

**Theorem 5.2** (Chain Rule)**.** *Suppose $f : (a, b) \mapsto \mathbb{R}$ is differentiable at $c \in (a, b)$, and $g : (m, M) \mapsto \mathbb{R}$ is differentiable at $f(c) \in (m, M)$, then $g \circ f$ is differentiable at $c$ and*

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

*Proof.* Define $h : dom(g) \mapsto \mathbb{R}$ by

$$h(x) = \begin{cases} \dfrac{g(x)-g(f(c))}{x-f(c)}, & x \neq f(c) \\ g'(f(c)), & x = f(c) \end{cases}.$$

Since $g$ is differentiable at $f(c)$,

$$\lim_{x\to f(c)} h(x) = \lim_{x\to f(c)} \frac{g(x)-g(f(c))}{x-f(c)} = g'(f(c)),$$

so $h$ is continuous at $x = f(c)$. Note that for all $x \in dom(g)$

$$g(x) - g(f(c)) = h(x)(x - f(c)).$$

Since $f$ is differentiable at $c$, $f$ is continuous at $c$, so that $h \circ f$ is continuous at $c$ and

$$\lim_{x\to c} h(f(x)) = h(f(c)) = \lim_{x\to f(c)} h(x) = g'(f(c)).$$

Then for $t \neq c$ such that $f(t) \in dom(g)$

$$\begin{aligned}\frac{(g\circ f)(t)-(g\circ f)(c)}{t-c} &= h(f(t))\cdot\frac{f(t)-f(c)}{t-c}\\ &\to h(f(c))f'(c) = g'(f(c))f'(c) \text{ as } t\to c.\end{aligned}$$

□

**Note 5.6.** It is tempting in the proof of the chain rule to write for $t \neq c$

$$\frac{(g\circ f)(t)-(g\circ f)(c)}{t-c} = \frac{(g\circ f)(t)-(g\circ f)(c)}{f(t)-f(c)}\cdot\frac{f(t)-f(c)}{t-c},$$

but this is only true if $t \neq c$ implies $f(t) \neq f(c)$, which is not true in general. The other subtle point is that it is essential that $h$ is continuous at $f(c)$ and $f$ is continuous at $c$ in order to conclude

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

as the following example shows.

Let

$$h(x) = \begin{cases} 4, & x \neq 1\\ -4, & x = 1\end{cases} \text{ and } f(x) = \begin{cases} 1 + x\sin(\pi/x), & x\neq 0\\ 1, & x = 0\end{cases}.$$

Note that $\lim_{x\to 0} f(x) = 1$, and $\lim_{x\to 1} h(x) = 4$. Recall that $\lim_{x\to 0} h(f(x)) = h(f(0))$ if and only if $\forall (x_n)_n \subset dom(h\circ f)$ such that $x_n \neq 0$ and $x_n \to 0$, $h(f(x_n)) \to h(f(0))$. Let $x_n = 2/n$ for $n \in \mathbb{N}$. Then clearly $x_n \to 0$, but

$$f(x_n) = 1 + \frac{2}{n}\sin\left(\frac{n\pi}{2}\right) = \begin{cases} 1, & n \text{ even}\\ 1 + (-1)^{(n-1)/2}\frac{2}{n}(\neq 1), & n \text{ odd}\end{cases}$$

Then

$$h(f(x_n)) = \begin{cases} -4, & n \text{ even}\\ 4, & n \text{ odd}\end{cases},$$

so $\lim_{x\to 0}(h\circ f)(x)$ does not exists, even though

$$\lim_{x\to 0} f(x) = 1 \text{ and } \lim_{x\to 1} h(x) = 4.$$

## 5.2 The Mean Value Theorem

**Definition 5.4.** A function $f : S \mapsto \mathbb{R}$ has a **local max** at $c \in S$ if $\exists \delta > 0$ such that for $x \in S$,

$$|x - c| < \delta \implies f(x) \leq f(c).$$

Similarly, $f$ has a local min at $c \in S$ if $\exists \delta > 0$ such that for $x \in S$,

$$|x - c| < \delta \implies f(x) \geq f(c).$$

**Proposition 5.2.** *If $f : (a, b) \mapsto \mathbb{R}$ has a local extremum at $c \in (a, b)$ and $f'(c)$ exists, then $f'(c) = 0$.*

*Proof.* We consider two cases: $0 < f'(c) \leq \infty$ and $-\infty \leq f'(c) < 0$.

Case 1: $0 < f'(c) \leq \infty$. Then by the definition of the derivative, $\exists \delta > 0$ such that

$$0 < |x - c| < \delta \implies \frac{f(x) - f(c)}{x - c} > 0.$$

Then for $c - \delta < x < c$, $x - c < 0$, so $f(x) - f(c) < 0$. Similarly, for $c < x < c + \delta$, $x - c > 0$, so $f(x) - f(c) > 0$. Thus, $c$ cannot be a local min or max.

Case 2: $-\infty \leq f'(c) < 0$. Similar to case 1. □

**Theorem 5.3** (Rolle's Theorem)**.** *Suppose $f : [a, b] \mapsto \mathbb{R}$ is continuous on $[a, b]$ and differentiable (finite or infinite) on $(a, b)$. If $f(a) = f(b)$, then $\exists c \in (a, b)$ such that $f'(c) = 0$.*

*Proof.* Since $[a, b]$ is a closed and bounded set in $\mathbb{R}$, it is compact. Since $f$ is continuous on $[a, b]$, by Corollary 3.2 $\exists x_1, x_2 \in [a, b]$ such that

$$f(x_1) = \min\{f(x) : x \in [a, b]\} \ \text{ and } \ f(x_2) = \max\{f(x) : x \in [a, b]\},$$

so $f(x_1) \leq f(x) \leq f(x_2)$ for all $x \in [a, b]$.

Case 1: Suppose $\{x_1, x_2\} \subseteq \{a, b\}$. Then $f(a) = f(b)$ implies

$$f(a) = f(x_1) = f(x_2) = f(b).$$

Since $f(x_1) \leq f(x) \leq f(x_2)$ for all $x \in [a, b]$, $f$ is constant on $[a, b]$, so $f'(c) = 0$ for all $c \in (a, b)$.

Case 2: Suppose $x_1 \in (a, b)$. Since $f(x) \geq f(x_1)$ for all $x \in [a, b]$, $x_1$ is a local minimum. Because $f$ is differentiable at $x_1$, $f'(x_1) = 0$ by Proposition 5.2.

Case 3: Suppose $x_2 \in (a, b)$. Similar to case 2. □

**Theorem 5.4** (Generalized Mean Value Theorem (Cauchy's MVT))**.** *If $f, g : [a, b] \mapsto \mathbb{R}$ are continuous on $[a, b]$ and differentiable on $(a, b)$ with at least one of $f'$ or $g'$ finite for all $x \in (a, b)$, then $\exists c \in (a, b)$ such that*

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

*Proof.* Define $h : [a, b] \mapsto \mathbb{R}$ by

$$h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)].$$

Since $f$ and $g$ are continuous on $[a, b]$, $h$ is also continuous on $[a, b]$. Furthermore, $h$ is differentiable on $(a, b)$ (possibly infinite). Note that

$$h(a) = f(a)[g(b) - g(a)] - g(a)[f(b) - f(a)] = f(a)g(b) - g(a)f(b),$$

and

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

Thus $h(a) = h(b)$, so by Rolle's theorem, $\exists c \in (a, b)$ such that

$$0 = h'(c) = f'(c)[g(b) - g(a)] - g'(c)[f(b) - f(a)].$$

□

**Corollary 5.1** (Mean Value Theorem)**.** *If $f : [a,b] \mapsto \mathbb{R}$ is continuous and differentiable (possible infinite) on $(a,b)$, then $\exists c \in (a,b)$ such that*

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

*Proof.* Take $g(x) = x$ and apply the Generalized MVT. □

**Proposition 5.3.** *Suppose $f : [a,b] \mapsto \mathbb{R}$ is continuous and differentiable (possibly infinite) on $(a,b)$.*

*a) If $f'(x) > 0 \ \forall x \in (a,b)$, then $f$ is strictly increasing on $[a,b]$.*

*b) If $f'(x) < 0 \ \forall x \in (a,b)$, then $f$ is strictly decreasing on $[a,b]$.*

*c) If $f'(x) = 0 \ \forall x \in (a,b)$, then $f$ is constant.*

*Proof.* Let $s, t$ be such that $a \le s < t \le b$. Then by the MVT applied to $f : [s,t] \mapsto \mathbb{R}$,

$$f'(x) = \frac{f(t) - f(s)}{t - s}$$

for some $x \in (s,t)$, so $f(t) - f(s) = f'(x)(t - s)$.

a) If $f'(x) > 0$ for all $x \in (a,b)$, then $f(t) - f(s) > 0$ for all $a \le s < t \le b$. Thus, $a \le s < t \le b \implies f(s) < f(t)$.

b) If $f'(x) < 0$ for all $x \in (a,b)$, then $f(t) - f(s) < 0$ for all $a \le s < t \le b$. Thus, $a \le s < t \le b \implies f(s) > f(t)$.

c) If $f'(x) = 0$ for all $x \in (a,b)$, then $f(t) - f(s) = 0$ for all $a \le s < t \le b$. Thus, $f(a) = f(t)$ for all $t \in [a,b]$.

□

**Theorem 5.5** (Intermediate Value Theorem for Derivatives)**.** *Suppose $f : [a,b] \mapsto \mathbb{R}$ is differentiable on $[a,b]$ (possibly infinite). If $m$ lies between $f'(a+)$ and $f'(b-)$, then $\exists c \in (a,b)$ such that $f'(c) = m$.*

*Proof.* WLOG assume $-\infty \le f'(a+) < m < f'(b-) \le \infty$ (otherwise use $-f$). Define

$$g(x) = \frac{f(x) - f(a)}{x - a} \quad \text{for } a < x \le b$$

and

$$h(x) = \frac{f(x) - f(b)}{x - b} \quad \text{for } a \le x < b.$$

Note that

$$g(b) = \frac{f(b) - f(a)}{b - a} = \frac{f(a) - f(b)}{a - b} = h(a)$$

and $\lim_{x \to a^+} g(x) = f'(a+)$ and $\lim_{x \to b^-} h(x) = f'(b-)$.

Case 1: Suppose that $m = \frac{f(b)-f(a)}{b-a}$. By the MVT, $\exists c \in (a,b)$ such that

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

Case 2: Suppose $m < \frac{f(b)-f(a)}{b-a} = g(b)$. By assumption

$$\lim_{x\to a^+} g(x) = f'(a^+) < m,$$

so $\exists x_1$ such that $a < x_1 < b$ and $g(x_1) < m$. Since $g$ is continuous on $[x_1, b]$ and $g(x_1) < m < g(b)$, by the IVT, $\exists x_2 \in (x_1, b)$ such that

$$m = g(x_2) = \frac{f(x_2)-f(a)}{x_2 - a}.$$

Now, by applying the MVT to $f$ on $[a, x_2]$, $\exists c \in (a, x_2)$ such that

$$f'(c) = \frac{f(x_2)-f(a)}{x_2-a} = m.$$

Case 3: Suppose $m > \frac{f(b)-f(a)}{b-a} = h(a)$. By assumption

$$m < \lim_{x\to b^-} h(x) = f'(b^-),$$

so $\exists x_1$ such that $a < x_1 < b$ and $h(x_1) > m$. Since $h$ is continuous on $[a, x_1]$ and $h(a) < m < h(x_1)$, by the IVT, $\exists x_2 \in (a, x_1)$ such that

$$m = h(x_2) = \frac{f(x_2)-f(b)}{x_2-b}.$$

Now, by applying the MVT to $f$ on $[x_2, b]$, $\exists c \in (x_2, b)$ such that

$$f'(c) = \frac{f(x_2)-f(b)}{x_2-b} = m.$$

□

**Corollary 5.2.** *If $f : (a,b) \mapsto \mathbb{R}$ is differentiable and $f'$ is monotonic on $(a,b)$, then $f'$ is continuous on $(a,b)$.*

*Proof.* Since $f'$ is monotonic, the only type of discontinuities possible are jump discontinuities. Suppose that $f'$ has a jump discontinuity at $x \in (c,d) \subset (a,b)$. WLOG assume $f'$ is increasing. Then $f(c) \le f(x-) < f(x+) \le f(d)$, so by the IVT for derivatives, $\exists t \in (c,d)$ such that

$$f'(x-) < f'(t) < f'(x+),$$

but this can't happen, since $f'$ is increasing. □

**Corollary 5.3.** *If $f : [a,b] \mapsto \mathbb{R}$ is continuous and differentiable (possible infinite) on $(a,b)$ with $f'(x) \neq 0$ $\forall x \in (a,b)$, then $f$ is either strictly increasing on $[a,b]$ or $f$ is strictly decreasing on $[a,b]$.*

*Proof.* If $f'(x) \neq 0$ $\forall x \in (a,b)$, then $f'(x) > 0$ for all $x \in (a,b)$ or $f'(x) < 0$ for all $x \in (a,b)$. If not, then by the IVT for derivatives, $\exists c \in (a,b)$ such that $f'(c) = 0$, contradicting our assumption. The result then follows from Proposition 5.3. □

## 5.3 Uniform Convergence and Differentiation

A question we might now ask is "If $f_n \to f$ and and each $f_n$ is differentiable is $f$ differentiable and does $f'(x) = \lim_{n\to\infty} f_n'(x)$?" This is not true in general as the following example illustrates.

**Example 5.3.** Let $f_n : \mathbb{R} \mapsto \mathbb{R}$ be defined by $f_n(x) = \frac{1}{\sqrt{n}} \sin(nx)$. Then

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

However, $f_n'(x) = \sqrt{n}\cos(nx)$ and $f'(x) = 0$, but for $x \neq \frac{(2k+1)\pi}{2}$, $k \in \mathbb{Z}$,

$$f_n(x) \not\to f(x).$$

For example, $f_n(0) = \sqrt{n} \to \infty \neq 0 = f'(x)$.

**Proposition 5.4.** *Let $\{f_n : [a,b] \mapsto \mathbb{R}, n \geq 1\}$ be differentiable on $[a,b]$ and suppose $\{f_n(x_0)\}_n$ converges for some $x_0 \in [a,b]$. If $(f_n')_n$ converges uniformly on $[a,b]$, then $f_n \to f$ uniformly on $[a,b]$ for some function $f$ and*

$$f'(x) = \lim_{n\to\infty} f_n'(x),\ x \in [a,b].$$

*Proof.* Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $m > n \geq N$ implies

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2}$$

and

$$|f_n'(t) - f_m'(t)| < \frac{\varepsilon}{2(b-a)}$$

for all $t \in [a,b]$. Let $x, t \in [a,b]$. Then by the MVT applied to $f_n - f_m$, we have for some $c$ between $x$ and $t$ and for $m > n \geq N$,

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leq |f_n'(c) - f_m'(c)||x-t| \leq \frac{\varepsilon|x-t|}{2(b-a)} \leq \frac{\varepsilon}{2}.$$

Thus, for all $x \in [a,b]$, $m > n \geq N$ implies

$$\begin{aligned}|f_n(x) - f_m(x)| &\leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\end{aligned}$$

so $(f_n)_n$ is uniformly Cauchy on $[a,b]$. By Cauchy convergence criterion for uniform convergence, $f_n \to f$ uniformly on $[a,b]$ to a function $f$, where

$$f(x) = \lim_{n\to\infty} f_n(x),\ x \in [a,b].$$

Now, fix an $x \in [a,b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t-x} \quad\text{and}\quad \phi(t) = \frac{f(t) - f(x)}{t-x}$$

for all $t \in [a,b]$, $t \neq x$. Then for each $n = 1, 2, 3, \ldots$,

$$\lim_{t\to x} \phi_n(t) = f_n'(x)$$

where the limit it understood to be a right hand limit at $x = a$ and a left hand limit at $x = b$. For $m > n \geq N$

$$|\phi_n(t) - \phi_m(t)| = \left| \frac{f_n(t) - f_m(t) - f_n(x) + f_m(x)}{t - x} \right| \leq \frac{\varepsilon|t-x|}{2(b-a)|t-x|} = \frac{\varepsilon}{2(b-a)},$$

so $(\phi_n)_n$ is uniformly Cauchy and hence uniformly convergent for $t \in [a,b]$, $t \neq x$. Thus, we have $\phi_n \to \phi$ uniformly on $[a,b] \setminus \{x\}$ and $\lim_{t\to x} \phi_n(t) = f_n'(x)$, so by Theorem 4.3

$$f'(x) = \lim_{t\to x} \phi(t) = \lim_{t\to x} \lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f_n'(x).$$

□

We now return to a power series. If a function $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $R > 0$, then what can we say about $f'(x)$? It seems reasonable that since the derivative is linear that

$$f'(x) = \frac{d}{dx} \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} \frac{d}{dx} a_n x^n = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

This turns out to be true a we will now show.

**Proposition 5.5.** *If $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $R > 0$, then $f$ is differentiable on $(-R, R)$ and for $|x| < R$,*

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

*Proof.* Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$ be a power series with radius of convergence $0 < R \leq \infty$. Let $0 < \rho < R$ and define $f_n(x) = \sum_{k=0}^{n} a_k x^k$, $n = 1, 2, 3, \ldots$, on $[-\rho, \rho]$. Then $f_n$ is differentiable on $[-\rho, \rho]$ and $f_n \to f$ uniformly on $[-\rho, \rho]$.

Consider the power series $\sum_{k=1}^{\infty} k a_k x^{k-1}$, and note that this series converges at $x$ if and only if

$$x \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=1}^{\infty} k a_k x^k$$

converges, i.e. these two power series have the same radius of convergence. The radius of convergence is

$$\frac{1}{\overline{\lim}_{n\to\infty} |n a_n|^{1/n}} = \frac{1}{\overline{\lim}_{n\to\infty} |a_n|^{1/n}} = R,$$

since $\lim_{n\to\infty} n^{1/n} = 1$. Thus,

$$f_n'(x) = \sum_{k=1}^{n} k a_k x^{k-1} \to \sum_{k=1}^{\infty} k a_k x^{k-1}$$

uniformly on $[-\rho, \rho]$. Hence by Proposition 5.4

$$f'(x) = \lim_{n\to\infty} f_n'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

□

**Definition 5.5** (Higher Order Derivatives)**.** Let $f : (a, b) \mapsto \mathbb{R}$ and set $f^{(0)} = f$ and $f^{(1)} = f'$. The **nth derivative** of $f$ at $x$ is defined recursively by

$$f^{(n)}(x) = \frac{d}{dx} f^{(n-1)}(x) = \lim_{t\to x} \frac{f^{(n-1)}(t) - f^{(n-1)}(x)}{t - x}$$

provided $f^{(n-1)}(x)$ exists and is finite for $t \in (x - \delta, x + \delta)$ for some $\delta > 0$ and the limit exists.

**Note 5.7.** For $f^{(n)}(x)$ to exists, note that $f^{(n-1)}$ must exists in a neighborhood of $x$.

**Corollary 5.4.** *If $f(x) = \sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $R > 0$, then $f$ is infinitely differentiable on $(-R, R)$ and for $|x| < R$,*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) a_n x^{n-k}, k = 0, 1, 2, \ldots$$

*Proof.* The proof follows by induction using that fact that for each $k$, $f^{(k-1))}$ is a power series with radius of convergence $(-R, R)$, and so is differentiable with derivative

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

□

## 5.4 Taylor's Theorem

**Lemma 5.1.** *Let $p(x) = a_0 + a_1(x - c) + \cdots + a_n(x - c)^n = \sum_{k=0}^{n} a_k(x - c)^k$. Then the mth derivative of $p$ for $m \le n$ is*

$$p^{(m)}(x) = \sum_{k=m}^{n} \frac{a_k k!}{(k-m)!}(x - c)^{k-m}.$$

*For $m > n$, $p^{(m)}(x) = 0$.*

*Proof.* For $m = 0$, $p^{(0)} = \sum_{k=0}^{n} \frac{a_k k!}{k!}(x - c)^{k-0} = \sum_{k=0}^{n} a_k(x - c)^k$, so the formula holds for $m = 0$. Now, suppose that for $0 \le m < n$

$$p^{(m)}(x) = \sum_{k=m}^{n} \frac{a_k k!}{(k-m)!}(x - c)^{k-m}.$$

Then

$$\begin{aligned} p^{(m+1)}(x) &= \frac{d}{dx}\left(a_m m! + \sum_{k=m+1}^{n} \frac{a_k k!}{(k-m)!}(x-c)^{k-m}\right) \\ &= 0 + \sum_{k=m+1}^{n} a_k \frac{k!}{(k-m)!}(k-m)(x-c)^{k-m-1} \\ &= \sum_{k=m+1} a_k \frac{k!}{(k-(m+1))!}(x-c)^{k-(m+1)}, \end{aligned}$$

so the formula holds for $m+1$, and the result follows by induction for $0 \leq m \leq n$. Note that $p^{(n)}(x) = a_n n!$, so $p^{(n+k)}(x) = 0$ for all $k \geq 1$. □

**Corollary 5.5.** *Suppose that* $p(x) = \sum_{k=0}^{n} a_k(x-c)^k$ *is a polynomial, then for all* $m \in \{0, 1, 2, \ldots, n\}$, $a_m = \dfrac{p^{(m)}(c)}{m!}$.

*Proof.* For $0 \leq m \leq n$,

$$p^{(m)}(x) = \sum_{k=m}^{n} a_k \frac{k!}{(k-m)!}(x-c)^{k-m},$$

so $p^{(m)}(c) = a_m m!$. □

**Definition 5.6.** If $f : (a,b) \mapsto \mathbb{R}$ is $n$-times differentiable at $c \in (a,b)$, then the **n-th order Taylor polynomial** centered at $c$ for $f$ is

$$p(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k.$$

**Note 5.8.** The n-th order Taylor polynomial is the unique polynomial of degree at most $n$ such that $p^{(k)}(c) = f^{(k)}(c)\ \forall 0 \leq k \leq n$. The remainder in this approximation is

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k.$$

If $f$ has derivatives of all orders and $R_n(x) \to 0$, then we can write

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k,$$

which is called the **Taylor series** of $f$. Not every function has a Taylor series. In the case, where $c = 0$, the n-th order polynomial (or series) is also referred to as a **McLauren polynomial (or series)**. If a function has derivative of all orders, then we need to know if the remainder goes to zero to be able to write a function as its Taylor series.

**Example 5.4.** Find the McLauren series of $f(x) = \sin(x)$.

*Solution.* Note that

$$f^{(0)}(x) = \sin(x) \quad f^{(1)}(x) = \cos(x) \quad f^{(2)}(x) = -\sin(x) \quad f^{(3)}(x) = -\cos(x) \quad f^{(4)} = \sin(x),$$

so we have

$$f^{(n)}(0) = \begin{cases} 0, & n \text{ is even} \\ (-1)^{(n-1)/2}, & n \text{ is odd} \end{cases}.$$

Then the McLauren series is

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

In order to write

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

we need to show that $R_n(x) \to 0$, but we need a useful form of the remainder. This is the subject of Taylor's theorem. Note that this series converges for all $x \in \mathbb{R}$, but we cannot say at this point whether or not it converges to $\sin(x)$. □

**Theorem 5.6** (Taylor's Theorem)**.** *Suppose $f : [a,b] \mapsto \mathbb{R}$ is $n$ times differentiable on $(a,b)$. Let $c \in (a,b)$ and let $p_{n-1}(x)$ be the Taylor polynomial of degree $n-1$ of $f$ given by*

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

*Then for each $x \in [a,b] \setminus \{c\}$, $\exists \xi$ between $x$ and $c$ such that*

$$f(x) = p_{n-1}(x) + \frac{f^{(n)}(\xi)}{n!}(x-c)^n.$$

**Example 5.5.** Show that that for all $x \in \mathbb{R}$,

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

*Solution.* We now apply Taylor's theorem to show that the remainder term goes to zero. Note that $f(x) = \sin(x)$ has derivatives at of all orders and $\forall x \in \mathbb{R}$. Recall that

$$|f^{(n)}(x)| = \begin{cases} |\sin(x)|, & n \text{ is even} \\ |\cos(x)|, & n \text{ is odd} \end{cases} \leq 1, \ \forall x \in \mathbb{R}.$$

Then by Taylor's theorem for each $x \in \mathbb{R}$, $\exists t$ between $x$ and 0 such that

$$|R_n(x)| = \left| \sin(x) - \sum_{k=0}^{n-1} \frac{f^{(n)}(0)}{n!} \right| = \left| \frac{f^{(n+1)}(t)}{(n+1)!} x^{n+1} \right| \leq \frac{|x|^{n+1}}{(n+1)!} \to 0,$$

so $\forall x \in \mathbb{R}$

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

□

*Proof.* Let $c \in [a,b]$ and fix $x \in [a,b] \setminus \{c\}$. Define

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

and

$$G(t) = (x-c)^n - (x-t)^n$$

for $t$ between $x$ and $c$. Then for $t$ between $x$ and $c$

$$G'(t) = n(x-t)^{n-1},$$

and

$$\begin{aligned}
F'(t) &= \frac{d}{dt}\sum_{k=0}^{n-1} \frac{f^{(k)}(t)}{k!}(x-t)^k \\
&= f^{(1)}(t) + \sum_{k=1}^{n-1} \frac{1}{k!}\left(f^{(k+1)}(t)(x-t)^k - kf^{(k)}(t)(x-t)^{k-1}\right) \\
&= \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} + \sum_{k=0}^{n-2} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=0}^{n-2} \frac{f^{(k+1)}(t)}{k!}(x-t)^k \\
&= \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1}.
\end{aligned}$$

Since $F$ and $G$ are continuous on the closed interval between $c$ and $x$ and differentiable on the open interval between $c$ and $x$, by the generalized MVT $\exists$ a $\xi$ between $c$ and $d$ such that

$$F'(\xi)[G(x) - G(c)] = G'(\xi)[F(x) - F(c)].$$

Note that

$$G(x) - G(c) = (x-c)^n$$

and

$$F(x) - F(c) = f(x) - p_{n-1}(x),$$

so that

$$\frac{f^{(n)}(\xi)}{(n-1)!}(x-\xi)^{n-1}[(x-c)^n] = n(x-\xi)^{n-1}[f(x) - p_{n-1}(x)].$$

Rearranging terms we get

$$\frac{f^{(n)}(\xi)}{n!}(x-\xi)^n = f(x) - p_{n-1}(x).$$

□

## 5.5 Applications in Probability and Statistics

- Probability Generating Functions
- Delta Method

# Chapter 6

# Riemann-Stieltjes Integration

In this section, we seek to define the Riemann-Stieltjes integral. This will allow us to define the expectation of a random variable as an integral in both the discrete and continuous cases. In particular, given a random variable with CDF $F$, we will see that

$$EX = \int_{-\infty}^{\infty} x\, dF = \begin{cases} \int_{-\infty}^{\infty} xf(x)\, dx, & X \text{ is continuous with pdf } f \\ \sum_x xp(x), & X \text{ is discrete with pmf } p. \end{cases}$$

Recall that in the continuous case $f(x) = F'(x)$ and in the discrete case $p(x) = F(x+) - F(x-)$.

## 6.1 Definition and Existence of the Integral

We begin with some notation. Suppose $f : S \mapsto \mathbb{R}$ is a bounded function. Define

$$M(f, S) = \sup\{f(x) : x \in S\} \quad \text{and} \quad m(f, S) = \inf\{f(x) : x \in S\}.$$

Note that if $S_1 \subseteq S_1$ and $f : S_2 \mapsto \mathbb{R}$ is bounded, then

$$M(f, S_1) \leq M(f, S_2) \quad \text{and} \quad m(f, S_1) \geq m(f, S_2).$$

**Lemma 6.1.** *If $f : S \mapsto \mathbb{R}$ is bounded, then*

$$M(f, S) - m(f, S) = \sup\{f(s) - f(t) : s, t \in S\}.$$

*Proof.* Recall that for $A, B \subset \mathbb{R}$, $\inf A = -\sup(-A)$ and $\sup(A + B) = \sup(A) + \sup(B)$. Then

$$\begin{aligned} M(f, S) - m(f, S) &= \sup\{f(s) : s \in S\} - \inf\{f(t) : t \in S\} \\ &= \sup\{f(s) : s \in S\} + \sup\{-f(t) : t \in S\} \\ &= \sup\{f(s) - f(t) : s, t \in S\}. \end{aligned}$$

□

**Definition 6.1.** Given an interval $[a, b] \subset \mathbb{R}$, a **partition** of $[a, b]$ is an ordered set $P = \{t_0, t_1, \ldots, t_n\}$ where $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$. Let $\Delta t_k = t_k - t_{k-1}$ be the length of $[t_{k-1}, t_k]$. We define the **mesh** of $P$ by $\text{mesh}(P) = \max\{\Delta t_k : k = 1, 2, \ldots, n\}$.

**Definition 6.2.** Let $f : [a, b] \mapsto \mathbb{R}$ be bounded. Given a partition $P = \{t_k\}_{k=0}^n$ of $[a, b]$, we define the **upper and lower Darboux sums**

$$U(f, P) = \sum_{k=1}^{n} M(f, [t_{k-1}, t_k])\Delta t_k$$

and

$$L(f, P) = \sum_{k=1}^{n} m(f, [t_{k-1}, t_k])\Delta t_k.$$

The **upper and lower Darboux integral** of $f$ are

$$\overline{\int_a^b} f\ dx = U(f) = \inf_P U(f, P)$$

and

$$\underline{\int_a^b} f\ dx = L(f) = \sup_P L(f, P)$$

where the sup and inf are taken over all partitions $P$ of $[a, b]$. $f$ is said to be **Darboux integrable** on $[a, b]$ if $U(f) = L(f)$. In this case, we write

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

We will see later that the Darboux integral is equivalent to the Riemann integral, but we want a more general integral for probability.

**Definition 6.3.** Let $h : [a, b] \mapsto \mathbb{R}$ be a bounded function and let $F : [a, b] \mapsto \mathbb{R}$ be a monotonically increasing function with $F(a)$ and $F(b)$ finite. Given a partition $P = \{t_k\}_{k=0}^n$ of $[a, b]$, we define the **upper and lower Darboux-Stieltjes sums** with respect of $F$ over $[a, b]$ by

$$U(h, P, F) = \sum_{k=1}^{n} M(h, [t_{k-1}, t_k])\Delta F_k$$

and

$$L(h, P, F) = \sum_{k=1}^{n} m(h, [t_{k-1}, t_k])\Delta F_k.$$

where $\Delta F_k = F(t_k) - F(t_{k-1})$. The **upper and lower Darboux-Stieltjes integrals** of $f$ are

$$\overline{\int_a^b} h\ dF = U(h, F) = \inf_P U(h, P, F)$$

and

$$\underline{\int_a^b} h\ dF = L(h, F) = \sup_P L(h, P, F)$$

where the sup and inf are taken over all partitions $P$ of $[a, b]$. $h$ is said to be **Darboux-Stieltjes integrable** with respect to $F$ on $[a, b]$ if $U(h, F) = L(h, F)$. In this case, we write

$$\int_a^b h\ dF \quad \text{or} \quad \int_a^b h(x)\ dF(x).$$

For ease of notation, we may write

$$M_k = M_k(h) = M(h, [t_{k-1}, t_k]) \text{ and } m_k = m_k(h) = m(h, [t_{k-1}, t_k])$$

**Note 6.1.** First, note that when $F(x) = x$, the Darboux-Stieltjes integral reduces to the Darboux integral over $[a, b]$.

**Note 6.2.** If $M = M(h, [a, b])$ and $m = m(h, [a, b])$, then given a partition $P = \{t_k\}_{k=0}^n$ of $[a, b]$, we have for all $k = 1, 2, \ldots, n$

$$m \leq m_k(h) \leq M_k(h) \leq M.$$

So

$$\begin{aligned} m(F(b) - F(a)) &= m \sum_{k=1}^n \Delta F_k \leq \sum_{k=1}^n m_k \Delta F_k \\ &\leq \sum_{k=1}^n M_k \Delta F_k \leq M \sum_{k=1}^n \Delta F_k \\ &= M(F(b) - F(a)). \end{aligned}$$

This implies that for any partition $P$ of $[a, b]$

$$m(F(b) - F(a)) \leq L(h, P, F) \leq U(h, P, F) \leq M(F(b) - F(a)),$$

so that the upper and lower Darboux-Stieltjes sums are bounded, which implies the upper and lower Darboux-Stieltjes integrals exist. Furthermore, if $t_k^\star \in [t_{k-1}, t_k]$, then

$$L(h, P, F) \leq \underbrace{\sum_{k=1}^n h(t_k^\star) \Delta F_k}_{\text{Riemann-Stieltjes sum}} \leq U(h, P, F).$$

**Note 6.3.** Throughout we will assume $F(a) < F(b)$. If $F(a) = F(b)$, then $F(x) = F(a)$ for all $x \in [a, b]$ and for any partition $P$ of $[a, b]$, $\Delta F_k = 0$ for all $k = 1, 2, \ldots, n$ and

$$\sum_{k=1}^n M_k(h) \Delta F_k = \sum_{k=1}^n m_k(h) \Delta F_k = 0.$$

Thus for any bounded $h : [a, b] \mapsto \mathbb{R}$, if $F(a) = F(b)$, then

$$\int_a^b h \, dF = 0.$$

**Definition 6.4.** If $P_1$ and $P_2$ are partitions of $[a, b]$, then $P_2$ is a **refinement** of $P_1$ if $P_1 \subset P_2$.

**Proposition 6.1.** *Let $h : [a, b] \mapsto \mathbb{R}$ be bounded and let $F : [a, b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. If $P_1, P_2$ are partitions of $[a, b]$ such that $P_2$ is a refinement of $P_1$, then*

$$U(h, P_2, F) \leq U(h, P_1, F)$$

*and*

$$L(h, P_1, F) \leq L(h, P_2, F).$$

**Note 6.4.** If $P_1 \subset P_2 \subset \cdots$ is a sequence of partitions of $[a,b]$ such that $P_{i+1}$ is a refinment of $P_i$, then this proposition implies that $\{U(h,P_n,F)\}_n$ is a decreasing sequence and $\{L(h,P_n,F)\}_n$ is an increasing sequence.

*Proof.* Let $P_1 = \{t_k\}_{k=0}^n$ and let $P_2 = P_1 \cup \{s_j\}_{j=1}^m$. First, assume $m = 1$, so that $P_2 = P_1 \cup \{s_1\}$. Let $1 \le j \le n$ be such that $t_{j-1} < s_1 < t_j$. Then

$$\begin{aligned} M(h,[t_{j-1},t_j])\Delta F_j &= M(h,[t_{j-1},t_j])(F(s_1)-F(t_{j-1})) + M(h,[t_{j-1},t_j])(F(t_j)-F(s_1)) \\ &\ge M(h,[t_{j-1},s_1])(F(s_1)-F(t_{j-1})) + M(h,[s_1,t_j])(F(t_j)-F(s_1)) \end{aligned}$$

Thus,

$$\begin{aligned} U(h,P_1,F) &= \sum_{k\neq j} M_k(f)\Delta F_k + M_j(h)\Delta F_j \\ &\ge \sum_{k\neq j} M_k(f)\Delta F_k + M(h,[t_{j-1},s_1])(F(s_1)-F(t_{j-1})) + M(h,[s_1,t_j])(F(t_j)-F(s_1)) \\ &= U(h,P_2,F). \end{aligned}$$

Repeating this argument $m-1$ times, we obtain the result. The argument is the same to show that $L(h,P_1,F) \le L(h,P_2,F)$. □

**Corollary 6.1.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded, and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. If $P_1$ and $P_2$ are any partitions of $[a,b]$, then*

$$L(h,P_1,F) \le U(h,P_2,F).$$

*That is every upper Darboux-Stieltjes sum is an upper bound for the set of all lower Darboux-Stieltjes sums, and every lower Darboux-Stieltjes sum is a lower bound for the set of all upper Darboux-Stieltjes sums*

*Proof.* Let $Q = P_1 \cup P_2$, then $Q$ is a refinement of both $P_1$ and $P_2$, so

$$L(h,P_1,F) \le L(h,Q,F) \le U(h,Q,F) \le U(h,P_2,F).$$

□

**Proposition 6.2.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded, and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. If $P$ is any partition of $[a,b]$ then*

$$L(h,P,F) \le \underline{\int_a^b} h\, dF \le \overline{\int_a^b} h\, dF \le U(h,P,F).$$

*Proof.* Let $P,Q$ be partitions of $[a,b]$. Then

$$L(h,P,F) \le U(h,Q,F).$$

Since $P$ was arbitrary, this implies that

$$\underline{\int_a^b} h\, dF = \sup_P L(h,P,F) \le U(h,Q,F).$$

Since $Q$ was also arbitrary, this also implies that

$$\underline{\int_a^b} h\ dF \leq \inf_Q U(h,Q,F) = \overline{\int_a^b} h\ dF$$

□

**Theorem 6.1** (Cauchy Critertion for Darbou-Stieltjes Integrability)**.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. Then $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$ if and only if $\forall \varepsilon > 0$, $\exists$ a partition $P$ of $[a,b]$ with*

$$U(h,P,F) - L(h,P,F) < \varepsilon.$$

*Proof.* ( $\Longrightarrow$ ) Suppose that $h$ is integrable. Then

$$\int_a^b h\ dF = \sup_P L(h,P,F) = \inf_Q U(h,Q,F).$$

Let $\varepsilon > 0$. By definition of supremum and infimum, $\exists$ partitions $P_0$ and $Q_0$ of $[a,b]$ such that

$$\sup_P L(h,P,F) - \frac{\varepsilon}{2} < L(h,P_0,F) \leq \sup_P L(h,P,F)$$

and

$$\inf_Q U(h,Q,F) \leq U(j,Q_0,F) < \inf_Q U(h,Q,F) + \frac{\varepsilon}{2}.$$

Let $R = P_0 \cup Q_0$ be the common refinement of $P$ and $Q$. Then

$$L(h,P_0,F) \leq L(h,R,F) \leq U(h,R,F) \leq U(h,P_0,F),$$

so

$$U(h,R,F) - L(h,R,F) < \inf_Q U(h,Q,F) + \frac{\varepsilon}{2} - \sup_P L(h,P,F) + \frac{\varepsilon}{2} = \varepsilon.$$

( $\Longleftarrow$ ) Recall that $\forall$ partitions $P$ of $[a,b]$

$$0 \leq \overline{\int_a^b} h\ dF - \underline{\int_a^b} h\ dF \leq U(h,P,F) - L(h,P,F).$$

Let $\varepsilon > 0$. Then $\exists$ a partition $Q$ of $[a,b]$ such that

$$0 \leq \overline{\int_a^b} h\ dF - \underline{\int_a^b} h\ dF \leq U(h,Q,F) - L(h,Q,F) < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this implies that

$$\overline{\int_a^b} h\ dF = \underline{\int_a^b} h\ dF,$$

so $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$. □

We now turn to some sufficient conditions for $h$ to be integrable over $[a,b]$ with respect to $F$.

**Theorem 6.2.** *Let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. If $h : [a,b] \mapsto \mathbb{R}$ is continuous, then $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$.*

*Proof.* Since $h$ is continuous on $[a,b]$ and $[a,b]$ is compact, $h$ is uniformly continuous on $[a,b]$, and $h$ is bounded on $[a,b]$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that $s,t \in [a,b]$ and $|s-t| < \delta$ implies

$$|h(s) - h(t)| < \frac{\varepsilon}{2(F(b) - F(a))}.$$

Now choose a partition $P = \{t_k\}_{k=0}^n$ of $[a,b]$ such that $\text{mesh}(P) < \delta$. (Note that such a partition exists. We can choose the partition $P = \{a, a+\delta/2, a+\delta, \ldots, b\}$.) Then

$$\begin{aligned} U(h,P,F) - L(h,P,F) &= \sum_{k=1}^n (M_k(h) - m_k(h))\Delta F_k \\ &= \sum_{k=1}^n \sup_{s,t \in [t_{k-1},t_k]} |h(s) - h(t)|\Delta F_k \\ &\leq \frac{\varepsilon}{2(F(b) - F(a))} \sum_{k=1}^n \Delta F_k \\ &= \frac{\varepsilon}{2(F(b) - F(a))}(F(b) - F(a)) < \varepsilon. \end{aligned}$$

Therefore, $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$ by the Cauchy criterion for integrability. □

**Theorem 6.3.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. If $h$ is monotonic and $F$ is continuous on $[a,b]$, then $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$.*

*Proof.* WLOG assume that $h$ is monotonically increasing. If $h(a) = h(b)$, then for any partition $P$ of $[a,b]$, $M_k(h) = m_k(h) = h(a)$, so that

$$U(h,P,F) - L(h,P,F) = 0.$$

Thus $h$ is integrable. Now assume $h(a) < h(b)$. Let $\varepsilon > 0$. Choose a partition $P = \{t_k\}_{k=0}^n$ of $[a,b]$ such that $\max_{k \in \{1,2,\ldots,n\}} \Delta F_k < \varepsilon/[h(b) - h(a)]$. (Such a partition can be chosen by the intermediate value theorem.) Note that, since $h$ is increasing $M_k(h) = h(t_k)$ and $m_k(h) = h(t_{k-1})$, so

$$\begin{aligned} U(h,P,F) - L(h,P,F) &= \sum_{k=1}^n (M_k(h) - m_k(h))\Delta F_k \\ &= \sum_{k=1}^n (h(t_k) - h(t_{k-1}))\Delta F_k \\ &< \frac{\varepsilon}{h(b) - h(a)} \sum_{k=1}^n (h(t_k) - h(t_{k-1})) \\ &= \frac{\varepsilon}{h(b) - h(a)}[h(b) - h(a)] \end{aligned}$$

$$= \varepsilon.$$

Therefore, $h$ is Darboux-Stieltjes integrable over $[a, b]$ by the Cauchy criterion for integrability. □

**Theorem 6.4.** *Let $F : [a, b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite, and let $h : [a, b] \mapsto \mathbb{R}$ be Darboux-Stieltjes integrable with respect to $F$ over $[a, b]$. If $m \leq h \leq M$ and $\phi : [m, M] \mapsto \mathbb{R}$ is continuous, then $\phi \circ h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a, b]$.*

*Proof.* Let $\varepsilon > 0$. Since $\phi$ is continuous on $[m, M]$, which is a compact set, $\phi$ is bounded. Let $K$ be such that $\sup_{t\in[m,M]} |\phi(t)| \leq K$. Furthermore, $\phi$ is uniformly continuous on $[m, M]$, so $\exists\delta > 0$ such that $\delta < \varepsilon/[F(b) - F(a) + 2K]$ and for $s, t \in [m, M]$, $|s - t| < \delta \implies |\phi(s) - \phi(t)| < \varepsilon/[F(b) - F(a) + 2K]$.

Since $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a, b]$, $\exists$ a partition $P = \{t_k\}_{k=0}^n$ of $[a, b]$ such that

$$U(h, P, F) - L(h, P, F) < \delta^2.$$

Define the set $A = \{k : M_k(h) - m_k(h) < \delta\}$ and $B = \{k : M_k(h) - m_k(h) \geq \delta\}$. For $k \in A$, we have for all $x, y \in [t_{k-1}, t_k]$, $|h(x) - h(y)| \leq M_k(h) - m_k(h) < \delta$. Thus for $k \in A$,

$$\begin{aligned} M_k(\phi \circ h) - m_k(\phi \circ h) &= \sup_{x,y\in[t_{k-1},t_k]} \phi(h(x)) - \inf_{y\in[t_{k-1},t_k]} \phi(h(y)) \\ &= \sup_{x,y\in[t_{k-1},t_k]} \phi(h(x)) - \phi(h(y)) \\ &\leq \frac{\varepsilon}{F(b) - F(a) + 2K}, \end{aligned}$$

and for $k \in B$, $M_k(\phi \circ h) - m_k(\phi \circ h) \leq 2K$. Then

$$0 \leq \delta \sum_{k\in B} \Delta F_k \leq \sum_{k\in B} (M_k(h) - m_k(h))\Delta F_k \leq U(h, P, F) - L(h, P, F) < \delta^2,$$

so $\sum_{k\in B} \Delta F_k < \delta$. It follows that

$$\begin{aligned} U(\phi \circ h, P, F) - L(\phi \circ h, P, F) &= \sum_{k=1}^{n} (M_k(\phi \circ h) - m_k(\phi \circ h))\Delta F_k \\ &\leq \frac{\varepsilon}{F(b) - F(a) + 2K} \sum_{k\in A} \Delta F_k + 2K \sum_{k\in B} \Delta F_k \\ &\leq \frac{\varepsilon}{F(b) - F(a) + 2K} (F(b) - F(a)) + 2K\delta \\ &< \frac{\varepsilon}{F(b) - F(a) + 2K} [F(b) - F(a) + 2K] = \varepsilon \end{aligned}$$

□

So far we have discussed Darboux-Stieltjes integrals, but we now turn to show that this integral is the same as the Riemann-Stieltjes integral. The following definition describes the Riemann-Stieltjes integral in a way more familiar from calculus.

**Definition 6.5.** Given a partition $P = \{t_k\}_{k=0}^n$ of $[a,b]$, an **augmentation** of $P$ has the form $P^\star = P \cup \{t_k^\star\}_{k=1}^n$ where $t_k^\star \in [t_{k-1}, t_k], k = 1, 2, \ldots, n$. Given a monotonically increasing function $F : [a,b] \mapsto \mathbb{R}$ such that $F(a)$ and $F(b)$ are finite, define the **F-mesh of P** as

$$\text{mesh}_F(P) = \max\{\Delta F_k = F(t_k) - F(t_{k-1}) : t_{k-1}, t_k \in P\}.$$

Let $h : [a,b] \mapsto \mathbb{R}$ be a bounded function, then **Riemann-Stieltjes sum of** $h$ **with respect to** $F$ **over** $[a,b]$ is

$$R(h, P^\star, F) = \sum_{k=1}^{n} h(t_k^\star)\Delta F_k.$$

$h$ is said to be **Riemann-Stieltjes integrable with respect to** $F$ **over** $[a,b]$ if $\exists I \in \mathbb{R}$ such that $\forall \varepsilon > 0$, $\exists \delta > 0$ and a partition $P$ of $[a,b]$ with $\text{mesh}_F(P) < \delta$ implies

$$|I - R(h, P^\star, F)| < \varepsilon$$

$\forall$ augmentation $P^\star$ of $P$. In such a case, we write

$$I = \int_a^b h\, dF.$$

**Theorem 6.5.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. Then $h$ is Darboux-Stieltjes integrable if and only if $h$ is Riemann-Stieltjes integrable. In this case, both integrals are equal.*

Before we can prove this result, we need a few other preliminary results.

**Lemma 6.2.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. Let $M = \sup_{x\in[a,b]} |h(x)|$. Suppose that $P$ and $Q$ are partitions of $[a,b]$ such that $P \subset Q$, i.e. $Q$ is a refinment of $P$. If $Q$ has $n$ more elements than $P$, then*

$$U(h,Q,F) - L(h,Q,F) \le U(h,P,F) - L(h,P,F) \le U(h,Q,F) - L(h,Q,F) + 4Mn mesh_F(P).$$

*Proof.* Suppose $P = \{t_k\}_{k=0}^m$ and $Q = P \cup \{s\}$ with $t_{j-1} < s < t_j$ for some $0 < j \le m$. Then

$$\begin{aligned}
U(h,P,F) - U(h,Q,F) &= \sum_{k=1}^{m} M_k(h)\Delta F_k - [\sum_{k\neq j} M_k \Delta F_k \\
&+ M(h,[t_{j-1},s])(F(s) - F(t_{j-1})) + M(h,[s,t_j])(F(t_j) - F(s))] \\
&= (\sup_{x\in[t_{j-1},t_j]} h(x) - \sup_{x\in[t_{j-1},s]} h(x))(F(s) - F(t_{j-1})) \\
&+ (\sup_{x\in[t_{j-1},t_j]} h(x) - \sup_{x\in[s,t_j]} h(x))(F(t_j) - F(s)) \\
&\le 2M(F(s) - F(t_{j-1})) + 2M(F(t_j) - F(s)) \\
&= 2M(F(t_j) - F(t_{j-1})) \\
&\le 2M\text{mesh}_F(P).
\end{aligned}$$

Similarly,

$$L(h,Q,F) - L(h,P,F) \leq 2M\text{mesh}_F(P).$$

Combining these two inequalities, we get

$$(U(h,P,F) - U(h,Q,F)) + (L(h,Q,F) - L(h,P,F)) \leq 4M\text{mesh}_F(P)$$

so that

$$U(h,P,F) - L(h,P,F) \leq U(h,Q,F) - L(h,Q,F) + 4M\text{mesh}_F(P)$$

Repeating this argument $n-1$ more times, we obtain the result. □

We omit the proof of the following lemma, but the proof is based on the fact that a monotonic function can only have jump discontinuities. This means it is piecewise continuous. Furthermore, for any $\varepsilon > 0$, there can be at most a finite number of jumps with size greater than $\varepsilon$. Thus we can always constuct such a partition on the parts where $F$ is continuous and place points in $P$ at the jumps that are too big, of which there are only finitely many.

**Lemma 6.3.** *If $F : [a,b] \mapsto \mathbb{R}$ is monotonically increasing with $F(a)$ and $F(b)$ finite, then $\forall\varepsilon > 0$, $\exists$ a partition $P$ of $[a,b]$ with mesh$_F(P) < \varepsilon$.*

**Proposition 6.3** (Second Cauchy Criterion for Darboux-Stieltjes Integrability)**.** *Let $h : [a,b] \mapsto \mathbb{R}$ be bounded and let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite. The $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$ if and only if $\forall\varepsilon > 0$, $\exists\delta > 0$ such that if $P$ is a partition of $[a,b]$ with mesh$_F(P) < \delta$, then*

$$U(h,P,F) - L(h,P,F) < \varepsilon.$$

*Proof.* ( $\Longleftarrow$ ) This clearly implies the Cauchy critertion for integrabilitiy.

( $\Longrightarrow$ ) Let $\varepsilon > 0$, and let $M = \sup_{x\in[a,b]} |h(x)|$. If $M = 0$, then the result is trivial. Assume $M > 0$. By the first Cauchy criterion, $\exists$ an partition $P_0 = \{t_k\}_{k=0}^n$ of $[a,b]$ such that

$$U(h,P_0,F) - L(h,P_0,F) < \frac{\varepsilon}{2}.$$

Let $\delta = \frac{\varepsilon}{8Mn}$ and let $P$ be a partition of $[a,b]$ with $\text{mesh}_F(P) < \delta$. Then for $Q = P \cup P_0$

$$U(h,Q,F) - L(h,Q,F) \leq U(h,P_0,F) - L(h,P_0,F) < \frac{\varepsilon}{2}$$

and by the lemma

$$\begin{aligned} U(h,P,F) - L(h,P,F) &\leq U(h,Q,F) - L(h,Q,F) + 4Mn\text{mesh}_F(P) \\ &\leq \frac{\varepsilon}{2} + 4Mn \cdot \frac{\varepsilon}{8Mn} \\ &= \varepsilon. \end{aligned}$$

□

We are now ready to prove the equivalence of the Riemann-Stieltjes and Darboux-Stieltjes integrals.

*Proof.* ( $\Longrightarrow$ ) Suppose that $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a,b]$. Let $\varepsilon > 0$. By the second Cauchy criterion, $\exists \delta > 0$ such that if $P$ is a partition of $[a,b]$ with $\text{mesh}_F(P) < \delta$, then

$$U(h,P,F) - L(h,P,F) < \varepsilon.$$

Choose such a $\delta > 0$. Let $P$ be a partition of $[a,b]$ with $\text{mesh}_F(P) < \delta$, and let $P^\star$ be any augmentation of $P$, that is $P^\star = \{t_k\}_{k=0}^n \cup \{t_k^\star\}_{k=1}^n$, where $t_k^\star \in [t_{k-1}, t_k]$. Then for $k = 1, 2 \ldots, n$,

$$m_k(h) \leq h(t_k^\star) \leq M_k(h),$$

which implies that

$$L(h,P,F) \leq R(h,P^\star,F) \leq U(h,P,F).$$

Thus

$$\left| R(h,P^\star,F) - (\text{Darboux}) \int_a^b h \; dF \right| < U(h,P,F) - L(h,P,F) < \varepsilon.$$

Thus, $h$ is Riemann-Stieltjes integrable with $I = (\text{Darboux}) \int_a^b h \; dF$.

( $\Longleftarrow$ ) Suppose that $h$ is Riemann-Stieltjes integrable with respect to $F$ over $[a,b]$ with integral $I \in \mathbb{R}$. Let $\varepsilon > 0$. Let $\delta > 0$ be such that if $P$ is a partition of $[a,b]$, then $|R(h,P^\star,F) - I| < \varepsilon/4$ for all augmentations $P^\star$ of $P$. Let $P = \{t_k\}_{k=0}^n$ be a partition of $[a,b]$ with $\text{mesh}_F(P) < \delta$, and let $P_1^\star = P \cup \{t_k^\star\}_{k=1}^n$ be an augmentation of $P$ with $t_k^\star$ chosen such that $h(t_k^\star) < m_k(h) + \frac{\varepsilon}{4(F(b)-F(a))}$. Let $P_2^\star = P \cup \{s_k\}_{k=1}^n$ be an augmentation of $P$ with $s_k^\star$ chosen such that $h(s_k^\star) > M_k(h) - \frac{\varepsilon}{4(F(b)-F(a))}$. Then

$$\begin{aligned}
I - \frac{\varepsilon}{4} &< R(h,P_1^\star,F) \\
&= \sum_{k=1}^n h(t_k^\star) \Delta F_k \\
&< \sum_{k=1}^n \left[ m_k(h) + \frac{\varepsilon}{4(F(b)-F(a))} \right] \Delta F_k \\
&= L(h,P,F) + \frac{\varepsilon}{4},
\end{aligned}$$

which implies $I - \frac{\varepsilon}{2} < L(h,P,F) \leq \underline{\int_a^b} h \; dF$. Similarly,

$$\begin{aligned}
I + \frac{\varepsilon}{4} &> R(h,P_2^\star,F) \\
&= \sum_{k=1}^n h(s_k^\star) \Delta F_k \\
&> \sum_{k=1}^n \left[ M_k(h) - \frac{\varepsilon}{4(F(b)-F(a))} \right] \Delta F_k \\
&= U(h,P,F) - \frac{\varepsilon}{4},
\end{aligned}$$

which implies $I - \frac{\varepsilon}{2} > U(h,P,F) \geq \overline{\int_a^b} h \; dF$. Thus,

$$I - \frac{\varepsilon}{2} < L(h,P,F) \leq \underline{\int_a^b} h \; dF \leq \overline{\int_a^b} h \; dF \leq U(h,P,F) < I + \frac{\varepsilon}{2}.$$

In particular,

$$0 \leq \overline{\int_a^b} h\, dF - \underline{\int_a^b} h\, dF \leq (I + \frac{\varepsilon}{2}) - (I - \frac{\varepsilon}{2}) = \varepsilon,$$

which implies that $h$ is Darboux-Stieltjes integrable with respect to $F$ over $[a, b]$. Furthermore,

$$\left| I - (\text{Darboux}) \int_a^b h\, dF \right| < \varepsilon.$$

□

## 6.2 Properties of the Integral

We will now write $h \in \mathscr{R}(F, [a, b])$ to mean that $h$ is Riemann-Stieltjes integrable with respect to $F$ over $[a, b]$.

**Proposition 6.4.** *Let $h, h_1, h_2 : [a, b] \mapsto \mathbb{R}$ be bounded, and let $F, G : [a, b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$, $F(b)$, $G(a)$, and $G(b)$ finite.*

*a) (**Linearity**) If $h_1, h_2 \in \mathscr{R}(F, [a, b])$, then $c_1 h_1 + c_2 h_2 \in \mathscr{R}(F, [a, b])$ for any $c_1, c_2 \in \mathbb{R}$ and*

$$\int_a^b (c_1 h_1 + c_2 h_2)\, dF = c_1 \int_a^b h_1\, dF + c_2 \int_a^b h_2\, dF.$$

*b) (**Order property**) If $h_1, h_2 \in \mathscr{R}(F, [a, b])$ and $h_1 \leq h_2$, then*

$$\int_a^b h_1\, dF \leq \int_a^b h_2\, dF.$$

*c) (**Additivity**) If $h \in \mathscr{R}(F, [a, b])$ and $a < c < b$, then $h \in \mathscr{R}(F, [a, c])$ and $h \in \mathscr{R}(F, [c, b])$. Moreover,*

$$\int_a^b h\, dF = \int_a^c h\, dF + \int_c^b h\, dF.$$

*d) (**Positive combination**) If $h \in \mathscr{R}(F, [a, b])$, $h \in \mathscr{R}(G, [a, b])$ and $k_1, k_2$ are nonnegative constants, then $h \in \mathscr{R}(k_1 F + k_2 G, [a, b])$ and*

$$\int_a^b h\, d(k_1 F + k_2 G) = k_1 \int_a^b h\, dF + k_2 \int_a^b h\, dG.$$

*e) (**Absolute integrability**) If $h \in \mathscr{R}(F, [a, b])$, then $|h| \in \mathscr{R}(F, [a, b])$, and*

$$\left| \int_a^b h\, dF \right| \leq \int_a^b |h|\, dF.$$

*Proof.* a) First, note that for any partition $P$ of $[a, b]$ that

$$m_k(h_1) + m_k(h_2) \leq m_k(h_1 + h_2) \leq M_k(h_1 + h_2) \leq M_k(h_1) + M_k(h_2),$$

which implies

$$L(h_1, P, F) + L(h_2, P, F) \le L(h_1 + h_2, P, F) \le U(h_1 + h_2, P, F) \le U(h_1, P, F) + U(h_2, P, F). \quad (\star)$$

Let $\varepsilon > 0$. Since $h_1, h_2 \in \mathscr{R}(F, [a, b])$, $\exists$ partitions $P_1$ and $P_2$ of $[a, b]$ such that

$$U(h_1, P_1, F) - L(h_1, P_1, F) < \frac{\varepsilon}{2} \text{ and } U(h_2, P_2, F) - L(h_2, P_2, F) < \frac{\varepsilon}{2}.$$

Let $P = P_1 \cup P_2$ be their common refinement. Then by $(\star)$

$$\begin{aligned} U(h_1 + h_2, P, F) - L(h_1 + h_2, P, F) &\le (U(h_1, P, F) - L(h_1, P, F)) + (U(h_2, P, F) - L(h_2, P, F)) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

so $h_1 + h_2 \in \mathscr{R}(F, [a, b])$ by the Cauchy criterion. Since $(\star)$ holds for any partition, we have for all partitions $P$ of $[a, b]$

$$\int_a^b (h_1 + h_2)\, dF \le U(h_1, P, F) + U(h_2, P, F)$$

and

$$L(h_1, P, F) + L(h_2, P, F) \le \int_a^b (h_1 + h_2)\, dF.$$

Since $h_1, h_2 \in \mathscr{R}(F, [a, b])$, we can choose partitions $P_1, P_2, P_3, P_4$ such that

$$\int_a^b h_1\, dF - \frac{\varepsilon}{2} < L(h_1, P_1, F) \quad \text{and} \quad \int_a^b h_2\, dF - \frac{\varepsilon}{2} < L(h_2, P_2, F)$$

and

$$U(h_1, P_3, F) < \int_a^b h_2\, dF + \frac{\varepsilon}{2} \quad \text{and} \quad U(h_2, P_4, F) < \int_a^b h_2\, dF + \frac{\varepsilon}{2}.$$

Let $P = P_1 \cup P_2 \cup P_3 \cup P_4$. Then

$$\int_a^b h_1\, dF + \int_a^b h_2\, dF - \frac{\varepsilon}{2} < \int_a^b (h_1 + h_2)\, dF < \int_a^b h_1\, dF + \int_a^b h_2\, dF + \frac{\varepsilon}{2}.$$

Since $\varepsilon > 0$ was abritrary, $\int_a^b (h_1 + h_2)\, dF = \int_a^b h_1\, dF + \int_a^b h_2\, dF$.

The proof that $ch \in \mathscr{R}(F, [a, b])$ and $\int_a^b ch\, dF = c \int_a^b h\, dF$ is similar and is left as an exercise.

b) First, suppose that $h \in \mathscr{R}(F, [a, b])$ and $h \ge 0$. Let $P$ be a partition of $[a, b]$. Then $m_k(h) \ge 0$ for all $k = 1, 2, \ldots, n$ and

$$\int_a^b h\, dF \ge L(h, P, F) = \sum_{k=1}^{n} m_k(h) \Delta F_k \ge 0.$$

Now, suppose $h_1, h_2 \in \mathscr{R}(F, [a, b])$ with $h_1 \le h_2$. Then $h_2 - h_1 \in \mathscr{R}(F, [a, b])$ by part a) and

$$0 \le \int_a^b (h_2 - h_1)\, dF = \int_a^b h_2\, dF - \int_a^b h_1\, dF.$$

c) Let $h \in \mathscr{R}(F,[a,b])$ and let $a < c < b$. Let $\varepsilon > 0$. Choose a partition $P_0$ of $[a,b]$ such that
$$U(h,P_0,F) - L(h,P_0,F) < \varepsilon.$$
Let $P = P_0 \cup \{c\}$. Then, $P \cap [a,c]$ is a partition of $[a,c]$, and
$$U(h,P\cap[a,c],F) - L(h,P\cap[a,c],F) \le U(h,P,F) - L(h,P,F) < \varepsilon.$$
Similarly, $P \cap [c,b]$ is a partition of $[c,b]$, and
$$U(h,P\cap[c,b],F) - L(h,P\cap[c,b],F) \le U(h,P,F) - L(h,P,F) < \varepsilon.$$
Thus, $h \in \mathscr{R}(F,[a,c])$ and $h \in \mathscr{R}(F,[c,b])$. Now, let $P$ be any partition of $[a,b]$ and define $P_0 = P \cup \{c\}$, $P_1 = P_0 \cap [a,c]$ and $P_2 = P_0 \cap [c,b]$. Then $P_0 = P_1 \cup P_2$ and
$$U(h,P,F) \ge U(h,P_0,F) = U(h,P_1,F) + U(h,P_2,F) \ge \int_a^c h\;dF + \int_c^b h\;dF$$
and
$$L(h,P,F) \le L(h,P_0,F) = L(h,P_1,F) + L(h,P_2,F) \le \int_a^c h\;dF + \int_c^b h\;dF.$$
Since $P$ was arbitrary, we have
$$\int_a^c h\;dF + \int_c^b h\;dF \le \int_a^b h\;dF \le \int_a^c h\;dF + \int_c^b h\;dF.$$

d) Left as an exercise.

e) Note that for all $x,y \in S \subseteq [a,b]$
$$|h(x)| - |h(y)| \le ||h(x)| - |h(y)|| \le |h(x) - h(y)| \le \sup_{x\in S} h(x) - \inf_{y\in S} h(y).$$
This implies that for any partition $P$ of $[a,b]$
$$|h(x)| - |h(y)| \le M_k(h) - m_k(h),\; \forall x,y \in [t_{k-1},t_k], k = 1,2,\ldots,n.$$
so that
$$M_k(|h|) - m_k(|h|) \le M_k(h) - m_k(h),\; k = 1,2,\ldots,n,$$
which implies
$$U(|h|,P,F) - L(|h|,P,F) \le U(h,P,F) - L(h,P,F).$$
Thus, $h \in \mathscr{R}(F,[a,b])$ implies $|h| \in \mathscr{R}(F,[a,b])$ by the Cauchy criterion. Note that $|h| - h \ge 0$ and $|h| + h \ge 0$, so by the order property in part b) and linearity of part a)
$$0 \le \int_a^b (|h| - h)\;dF = \int_a^b |h|\;dF - \int_a^b h\;dF$$
and
$$0 \le \int_a^b (|h| + h)\;dF = \int_a^b |h|\;dF + \int_a^b h\;dF.$$

Therefore,

$$-\int_a^b |h|\ dF \leq \int_a^b h\ dF \leq \int_a^b |h|\ dF \implies \left|\int_a^b h\ dF\right| \leq \int_a^b |h|\ dF.$$

□

**Example 6.1.** Let $h : [a,b] \mapsto \mathbb{R}$ be defined by

$$h(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ -1, & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}.$$

Then for any partition $P = \{t_k\}_{k=0}^n$, $M_k(h) = 1$ and $m_k(h) = -1$ for all $k = 1, 2, \ldots n$, so

$$L(h, P, F) = \sum_{k=1}^n m_k(h)\Delta F_k = -1[F(b) - F(a)] \implies \underline{\int_a^b} h\ dF = F(a) - F(b)$$

and

$$U(h, P, F) = \sum_{k=1}^n M_k(h)\Delta F_k = F(b) - F(a) \implies \overline{\int_a^b} h\ dF = F(b) - F(a).$$

Thus $h$ is not integrable, but $|h| = 1$ for all $x \in [a,b]$ so

$$\int_a^b h(x)\ dF = F(b) - F(a).$$

This example shows that $|h| \in \mathscr{R}(F, [a,b]) \not\Longrightarrow h \in \mathscr{R}(F, [a,b])$.

**Proposition 6.5.** *Let $F : [a,b] \mapsto \mathbb{R}$ be monotonically increasing with $F(a)$ and $F(b)$ finite, and let $h, g : [a,b] \mapsto \mathbb{R}$ be bounded. If $h, g \in \mathscr{R}(F, [a,b])$, then $hg \in \mathscr{R}(F, [a,b])$.*

*Proof.* Note that

$$hg = \frac{(h+g)^2 - (h-g)^2}{4}.$$

$h, g \in \mathscr{R}(F, [a,b]) \implies h+g, h-g \in \mathscr{R}(F, [a,b])$ and with $\phi(t) = t^2$, $(h+g)^2, (h-g)^2 \in \mathscr{R}(F, [a,b])$ by Theorem 6.4. Thus $hg \in \mathscr{R}(F, [a,b])$. □

We now turn to two special cases of the Riemann-Stieltjes integral relevant to expectations of discrete and continuous random variables. Let $s$ be fixed. Recall the indicator function

$$I(x \geq s) = \begin{cases} 1, & x \geq s \\ 0, & x < s. \end{cases}$$

Before proving the expectation formula in the discrete case, we will prove a special case.

**Proposition 6.6.** *If $a < s < b$, $h : [a,b] \mapsto \mathbb{R}$ is bounded, $h$ is continuous at $s$, and $F(x) = I(x \geq s)$, then*

$$\int_a^b h\ dF = h(s).$$

*Proof.* Note that $F(b) = 1$, $F(a) = 0$, and $F$ is increasing. Let $P = \{t_0, t_1, t_2, t_3\}$ be a partition of $[a, b]$ with $a = t_0 < t_1 < t_2 = s < t_3 = b$. Then $F(t_2) = F(t_3) = 1$ and $F(t_0) = F(t_1) = 0$, so

$$U(h, P, F) = \sum_{k=1}^{3} M_k \Delta F_k = M_2 \text{ and } L(h, P, F) = \sum_{k=1}^{3} m_k \Delta F_k = m_2.$$

Let $\varepsilon > 0$. Since $h$ is continuous at $s$, $\exists \delta > 0$ such that $|s - t| < \delta \implies |h(t) - h(s)| < \varepsilon/2$. Choose $t_1 = s - \delta/2$, then $|t - s| < \delta$ for all $t \in [t_1, s]$, which implies

$$|h(t) - h(s)| < \frac{\varepsilon}{2}, \ \forall t \in [t_1, s].$$

Therfore,

$$\begin{aligned}
M_2 - m_2 &= \sup_{t \in [t_1, s]} h(t) - \inf_{t \in [t_1, s]} h(t) \\
&= \sup_{t \in [t_1, s]} h(t) + \sup_{t \in [t_1, s]} [-h(t)] \\
&= \sup_{u, v \in [t_1, s]} [h(u) - h(v)] \\
&= \sup_{u, v \in [t_1, s]} [h(u) - h(s) + h(s) - h(v)] \\
&\leq \sup_{u \in [t_1, s]} [h(u) - h(s)] + \sup_{v \in [t_1, s]} [h(s) - h(v)] \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.
\end{aligned}$$

Thus $h$ is integrable with respect to $F$ over $[a, b]$. Moreover, for every choice of $a < t_1 < s$

$$m_2 \leq h(s) \leq M_2,$$

so

$$\left| \int_a^b h \, dF - h(s) \right| \leq M_2 - m_2 \leq \varepsilon.$$

Thus,

$$\int_a^b h \, dF = h(s).$$

□

**Proposition 6.7.** *Suppose* $c_n \geq 0$ *for all* $n \geq 1$, $\sum_{n=1}^{\infty} c_n$ *converges,* $(t_n)_n$ *is sequence of distinct points in* $(a, b)$*, and*

$$F(x) = \sum_{n=1}^{\infty} c_n I(x \geq t_n).$$

*If* $h$ *is continuous on* $[a, b]$*, the*

$$\int_a^b h \, dF = \sum_{n=1}^{\infty} c_n h(t_n).$$

*Proof.* First, note that

$$\left|\sum_{n=1}^{\infty} c_n I(x \geq t_n)\right| \leq \sum_{n=1}^{\infty} c_n < \infty,$$

so $F(x)$ converges for all $x \in [a,b]$ by the comparison test. Furthermore, $F$ is clearly increasing with $F(a) = 0$ and $F(b) = \sum_{n=1}^{\infty} c_n$. Also, $h$ continuous on $[a,b]$ implies that $h$ is bounded on $[a,b]$. Let $M = \sup_{x\in[a,b]} |h(x)|$. Now, let $\varepsilon > 0$. Choose an $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} c_n < \frac{\varepsilon}{M}.$$

Put

$$F_1(x) = \sum_{n=1}^{N} c_n I(x \geq t_n) \text{ and } F_2(x) = \sum_{n=N+1}^{\infty} c_n I(x \geq t_n).$$

Then $F = F_1 + F_2$, and by Propositions 6.4c) and 6.6,

$$\int_a^b h\, dF_1 = \sum_{n=1}^{N} c_n h(t_n).$$

Since $F_2(b) - F_2(a) = F_2(b) = \sum_{n=N+1}^{\infty} c_n < \frac{\varepsilon}{M}$, we have

$$\left|\int_a^b h\, dF_2\right| \leq M[F_2(b) - F_2(a)] < M\frac{\varepsilon}{M} = \varepsilon.$$

Thus, by Proposition 6.4 d),

$$\left|\int_a^b h\, dF - \sum_{n=1}^{N} c_n h(t_n)\right| = \left|\int_a^b h\, dF_1 + \int_a^b h\, dF_2 - \sum_{n=1}^{N} c_n h(t_n)\right| = \left|\int_a^b h\, dF_2\right| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we obtain that

$$\sum_{n=1}^{\infty} c_n h(t_n) = \lim_{N\to\infty} \sum_{n=1}^{N} c_n h(t_n) = \int_a^b h\, dF.$$

□

**Proposition 6.8.** *Suppose that $F : [a,b] \mapsto \mathbb{R}$ is monotonically increasing and differentiable on $[a,b]$ with $F(a)$ and $F(b)$ finite. Let $h : [a,b] \mapsto \mathbb{R}$ be bounded. If $F' \in \mathscr{R}([a,b])$, then $h \in \mathscr{R}(F,[a,b])$ if and only if $hF' \in \mathscr{R}([a,b])$. In this case,*

$$\int_a^b h\, dF = \int_a^b hF'\, dt.$$

*Proof.* Let $\varepsilon > 0$. Put $M = \sup_{x\in[a,b]} |h(x)|$. Since $F' \in \mathscr{R}([a,b])$, there exists a partition $P = \{t_0, t_1, \ldots, t_n\}$ such that

$$U(F', P) - L(F', P) < \frac{\varepsilon}{M}.$$

By the MVT, for each $j = 1, 2, \dots, n$, there exists an $s_j \in [t_{j-1}, t_j]$ such that $\Delta F_j = F'(s_j)\Delta t_j$. Let $x_j \in [t_{j-1}, t_j]$. Then

$$\sum_{k=1}^{n} |F'(s_j) - F'(x_j)|\, \Delta t_j \leq U(F', P) - L(F', P) < \frac{\varepsilon}{M}.$$

Since $\Delta F_j = F'(s_j)\Delta t_j$, it follows that

$$\left| \sum_{j=1}^{n} h(x_j)\Delta F_j - \sum_{j=1}^{n} h(x_j)F'(x_j)\Delta t_j \right| \leq M \sum_{k=1}^{n} |F'(s_j) - F'(x_j)|\, \Delta t_j < \varepsilon.$$

In particular,

$$\sum_{j=1}^{n} h(x_j)\Delta F_j \leq U(hF', P) + \varepsilon \quad \text{and} \quad \sum_{j=1}^{n} h(x_j)F'(x_j)\Delta t_j \leq U(h, P, F) + \varepsilon.$$

Since $x_j \in [t_{j-1}, t_j]$ was arbitrary, this implies that

$$U(h, P, F \leq U(hF', P) + \varepsilon \quad \text{and} \quad U(hF', P) \leq U(h, P, F) + \varepsilon,$$

and hence that

$$\overline{\int_a^b} h\; dF \leq U(hF', P) + \varepsilon \quad \text{and} \quad \overline{\int_a^b} hF'\; dt \leq U(h, P, F) + \varepsilon.$$

Note that if $Q$ is another partition of $[a, b]$, then our entire argument so far also holds using the refinement $P \cup Q$. In particular, we obtain

$$\overline{\int_a^b} h\; dF \leq U(hF', P{\cup}Q) + \varepsilon \leq U(hF', Q) + \varepsilon \quad \text{and} \quad \overline{\int_a^b} hF'\; dt \leq U(h, P{\cup}Q, F) + \varepsilon \leq U(h, Q, F) + \varepsilon.$$

Since $Q$ was an arbitrary partition, the previous inequalities imply

$$\overline{\int_a^b} h\; dF \leq \overline{\int_a^b} hF'\; dt + \varepsilon \quad \text{and} \quad \overline{\int_a^b} hF'\; dt \leq \overline{\int_a^b} h\; dF + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this implies that

$$\overline{\int_a^b} h\; dF = \overline{\int_a^b} hF'\; dt.$$

The equality $\underline{\int_a^b} h\; dF = \underline{\int_a^b} h'\; dF$ follows in exactly the same manner. Hence, the theorem follows. □

**Theorem 6.6** (Change of Variables)**.** *Suppose $F : [a, b] \mapsto \mathbb{R}$ is monotonically increasing with $F(b)$ and $F(a)$ finite. Let $\phi : [A, B] \mapsto [a, b]$ be strictly increasing and continuous, and let $h \in \mathscr{R}(F, [a, b])$. Define $\beta, g : [A, B] \mapsto \mathbb{R}$ by*

$$\beta(y) = F(\phi(y)) \quad \text{and} \quad g(y) = h(\phi(y)).$$

*Then $g \in \mathscr{R}(\beta, [A, B])$ and*

$$\int_A^B g\; d\beta = \int_a^b h\; dF.$$

*Proof.* To each partition $P = \{t_0, t_1, \ldots, t_n\}$ of $[a, b]$, there exists a unique partition $Q = \{y_0, y_1, \ldots, y_n\}$ such that $t_j = \phi(y_j)$, and vice versa. Then

$$\Delta F_j = F(t_j) - F(t_{j-1}) = F(\phi(y_j)) - F(\phi(y_{j-1})) = \beta(y_j) - \beta(y_{j-1}) = \Delta\beta_j,$$

for each $j = 1, 2, \ldots, n$. Since the values taken by $h$ on $[t_{j-1}, t_j]$ are the same as those taken by $g$ on $[y_{j-1}, y_j]$, we see that

$$U(g, Q, \beta) = U(h, P, F) \text{ and } L(g, Q, \beta) = L(h, P, F).$$

Let $\varepsilon > 0$. Since $h \in \mathscr{R}(F, [a, b])$, $\exists$ a partition $P$ of $[a, b]$ such that

$$U(h, P, F) - L(h, P, F) < \varepsilon.$$

Then

$$U(g, Q, \beta) - L(g, Q, \beta) < \varepsilon$$

where $Q$ is the unique partition of $[A, B]$ corresponding to $P$ described above. Thus, $g \in \mathscr{R}(\beta, [A, B])$ by the Cauchy criterion for integrability and

$$\int_A^B g\, d\beta = \inf_Q U(g, Q, \beta) = \inf_P U(h, P, F) = \int_a^b h\, dF.$$

□

The change of variables formula for Riemann integrals follows by combining the previous two results.

**Corollary 6.2** (Change of Variable for Riemann Integral)**.** *If $h \in \mathscr{R}([a, b])$ and if $\phi : [A, B] \mapsto [a, b]$ is strictly increasing and differentiable with $\phi' \in \mathscr{R}([A, B])$, then*

$$\int_a^b f(x)\, dx = \int_A^B f(\phi(y))\phi'(y)\, dy,$$

*where $a = \phi(A)$ and $b = \phi(B)$.*

## 6.3 Integration and Differentiation

**Theorem 6.7.** *Let $f \in \mathscr{R}([a, b])$. Define*

$$F(x) = \int_a^x f(t)\, dt,\ x \in [a, b].$$

*Then $F$ is continuous on $[a, b]$. Furthermore, if $f$ is continuous at $c \in [a, b]$, then $F$ is differentiable at $c$ with $F'(c) = f(c)$.*

*Proof.* Let $M = \sup_{x\in[a,b]} |f(x)| < \infty$. If $a \le x < y \le b$, then

$$|F(y) - F(x)| = \left|\int_x^y f(t)\, dt\right| \le M|y - x|.$$

Hence, $F$ is uniformly continuous on $[a,b]$ (and so also continuous). Now, suppose $f$ is continuous $c \in [a,b]$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that

$$|f(t) - f(c)| < \varepsilon, \ \forall t \in [a,b], \ |t-c| < \delta.$$

Then, for any $x \in [a,b]$ such that $c - \delta < x < c$

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| = \frac{1}{x-c} \left| \int_x^c (f(t) - f(c)) \, dt \right| < \varepsilon.$$

Similarly, for $x \in [a,b]$ such that $c < x < c + \delta$,

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| = \frac{1}{c-x} \left| \int_c^x (f(t) - f(c)) \, dt \right| < \varepsilon.$$

Hence, for all $x \in [a,b]$ such that $0 < |x - c| < \delta$

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $F'(c) = f(c)$. □

**Theorem 6.8** (Fundamental Theorem of Calculus)**.** *If $f \in \mathscr{R}([a,b])$, and if there exists a differentiable functions $F$ on $[a,b]$ such that $F' = f$, then*

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

*Proof.* Let $\varepsilon > 0$. Choose a partition $P = \{t_0, t_1, \dots, t_n\}$ of $[a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$. By the MVT, $\exists x_j \in [t_{j-1}, t_j]$ such that

$$F(t_j) - F(t_{j-1}) = f(x_j)\Delta t_j, \ j = 1, 2 \dots, n.$$

Then

$$\sum_{j=1}^n f(x_j)\Delta t_j = \sum_{j=1}^n [F(t_j) - F(t_{j-1})] = F(b) - F(a).$$

But,

$$L(f,P) \le \sum_{j=1}^n f(x_j)\Delta t_j \le U(f,P) \ \text{ and } \ L(f,P) \le \int_a^b f(x) \, dx \le U(f,P).$$

Hence,

$$\left| F(b) - F(a) - \int_a^b f(x) \, dx \right| < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this completes the proof. □

**Theorem 6.9** (Integration by Parts)**.** *Suppose that $F, G : [a,b] \mapsto \mathbb{R}$ are differentiable on $[a,b]$ with $F' = f \in \mathscr{R}([a,b])$ and $G' = g \in \mathscr{R}([a,b])$. Then*

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

*Proof.* Since $F$ and $G$ are differentiable on $[a, b]$, they are continuous on $[a, b]$, and hence, $F, G \in \mathscr{R}([a, b])$. Thus, $Fg \in \mathscr{R}([a, b])$ and $fG \in \mathscr{R}([a, b])$. Let $H(x) = F(x)G(x)$ on $[a, b]$. Then $H'(x) = f(x)G(x) + F(x)g(x)$, so $H' \in \mathscr{R}([a, b])$. Therefore, by the Fundamental Theorem of Calculus

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

completing the proof. □

## 6.4 Improper Riemann-Stieltjes Integrals

Throughout let $F : \mathbb{R} \mapsto \mathbb{R}$ be a monotonically increasing function with $F(b-) = \lim_{t\to b^-} F(t)$ and $F(a+) = \lim_{t\to a^+} F(t)$ finite.

**Definition 6.6.** Let $(a, b) \subseteq \mathbb{R}$, where $-\infty \leq a < b \leq \infty$, and $h : (a, b) \mapsto \mathbb{R}$. We say that $h$ is **locally Riemann-Stieltjes integrable** on $(a, b)$ if $h \in \mathscr{R}(F, [c, d])$ for all $[c, d] \subset (a, b)$. We say that $h$ is **improperly Riemann-Stieltjes integrable** on $(a, b)$ if $h$ is locally integrable on $(a, b)$ and the limit

$$\int_a^b h\ dF = \lim_{c\to a^+} \lim_{d\to b^-} \int_c^d h\ dF \qquad (\star)$$

exists and is finite. In this case, the limit is called the **improper Riemann-Stieltjes integral** of $h$ on $(a, b)$.

**Lemma 6.4.** *The order of the limits in* $(\star)$ *does not matter. In particular, if the limit in* $(\star)$ *exists and is finite, then the limit*

$$\lim_{d\to b^-} \lim_{c\to a^+} \int_c^d h\ dF$$

*exists and is finite and is equal to the limit in* $(star)$.

*Proof.* Let $x_0 \in (a, b)$. Then

$$\lim_{c\to a^+} \lim_{d\to b^-} \int_c^d h\ dF = \lim_{c\to a^+} \int_c^{x_0} h\ dF + \lim_{d\to b^-} \int_{x_0}^d h\ dF.$$

Since for each $c \in (a, b)$, $\lim_{d\to b^-} \int_c^d h\ dF$ exists, we have

$$\begin{aligned}
\lim_{x_0\to b^-} \lim_{d\to b^-} \int_{x_0}^d h\ dF &= \lim_{x_0\to b^-} \left[ \lim_{d\to b^-} \left( \int_c^d h\ dF - \int_c^{x_0} h\ dF \right) \right] \\
&= \lim_{x_0\to b^-} \left[ \lim_{d\to b^-} \int_c^d h\ dF - \int_c^{x_0} h\ dF \right] \\
&= \lim_{d\to b^-} \int_c^d h\ dF - \lim_{x_0\to b^-} \int_c^{x_0} h\ dF \\
&= 0.
\end{aligned}$$

Thus,

$$\lim_{x_0\to b^-}\lim_{c\to a^+}\int_c^{x_0} h\,dF = \lim_{x_0\to b^-}\left[\lim_{c\to a^+}\lim_{d\to b^-}\int_c^d h\,dF - \lim_{d\to b^-}\int_{x_0}^d h\,dF\right] = \lim_{c\to a^+}\lim_{d\to b^-}\int_c^d h\,dF.$$

□

**Note 6.5.** If $h$ is integrable with respect to $F$ on $[c,b]$ for all $c\in(a,b)$, then the improper Riemann-Stieltjes integral of $h$ on $(a,b]$ is also given by

$$\int_a^b h\,dF = \lim_{c\to a^+}\int_c^b h\,dF.$$

If this limit exists and is finite, then we also say that $h$ is **improperly Riemann-Stieltjes integrable with respect to $F$ on** $(a,b]$. A similar situation applies at the endpoint $b$, in which we can say that $h$ is **improperly Riemann-Stieltjes integrable with respect to $F$ on** $[a,b)$.

It is easily seen that $h$ is improperly integrable with respect to $F$ on $(a,b)$ if and only if $h$ is imporperly integrable with respect to $F$ on $(a,c]$ and on $[c,b)$ for all $c\in(a,b)$. In this case, we have

$$\int_a^b h\,dF = \int_a^c h\,dF + \int_c^b h\,dF.$$

**Proposition 6.9.** *The function $h(x)=1/x^p$ is improperly Riemann integrable on $(0,1]$ if and only if $p<1$, and is improperly Riemann integrable on $[1,\infty)$ if and only if $p>1$.*

*Proof.* Exercise. □

**Proposition 6.10** (Linearity)**.** *Let $k,l\in\mathbb{R}$. If $g,h$ are improperly Riemann-Stieltjes integrable with respect to $F$ on $(a,b)$, then $kh+lg$ is improperly Riemann-Stieltjes integrable with respect to $F$ on $(a,b)$, and*

$$\int_a^b [kh+lg]\,dF = k\int_a^b h\,dF + l\int_a^b g\,dF.$$

*Proof.* This follows immediately from the linearity property on each subinterval $[c,d]$ of $(a,b)$. □

**Proposition 6.11** (Comparison Theorem for Improper Integrals)**.** *Suppose that $h,g$ are locally integrable with respect to $F$ on $(a,b)$, and $0\le h(x)\le g(x)$ for all $x\in(a,b)$. If $g$ is improperly Riemann-Stieltjes integrable with respect to $F$ on $(a,b)$, then so is $h$ and*

$$\int_a^b h\,dF \le \int_a^b g\,dF.$$

*Proof.* Fix $c\in(a,b)$. Define $H(d)=\int_c^d h\,dF$ and $G(d)=\int_c^d g\,dF$ for $d\in[c,b)$. Then by the order property for proper Riemann-Stieltjes integrals, $H(d)\le G(d)$ for $d\in[c,b)$. Note that both $H$ and $G$ are increasing on $[c,b)$, and $G(b-)$ exists and is finite. Hence, $H$ is increasing and bounded above, so $H(b-)$ exists and is finite. This shows that $h$ is improperly Riemann-Stieltjes integrable with repsect to $F$ on $[c,b)$ for all $c\in(a,b)$. A similar argument shows that $h$ is improperly Riemann-Stieltjes integrable with repsect to $F$ on $(a, c]$ for all $c \in (a, b)$, so $h$ is improperly Riemann-Stieltjes integrable with repsect to $F$ on $(a, b)$. The order property follows easily from the order property of proper integrals on all subintervals $[c, d]$ of $(a, b)$. □

**Example 6.2.** Show that $h(x) = (\sin x)/x^{3/2}$ is improperly Riemann integrable on $(0, 1]$.

*Solution.* Since $0 \leq \sin x \leq x$ for all $x \in [0, 1]$ (you can use elementary calculus to prove it!), it follows that

$$0 \leq h(x) \leq x \cdot x^{-3/2} = x^{-1/2},\ \forall x \in (0, 1].$$

Since $x^{-1/2}$ is improperly Riemann integrable on $(0, 1]$, $h$ is also improperly Riemann integrable on $(0, 1]$ by the comparison theorem. □

**Example 6.3.** Show that $h(x) = (\ln x)/x^{5/2}$ is improperly Riemann integrable on $[1, \infty)$.

*Solution.* Since $0 \leq \ln x \leq x$ for $x \in [1, \infty)$, it follows that

$$0 \leq h(x) \leq x \cdot x^{-5/2} = x^{-3/2},\ \forall x \in [1, \infty).$$

Since $x^{-3/2}$ is improperly Riemann integrable on $[1, \infty)$, $h$ is also by the comparison theorem. □

**Corollary 6.3.** *If $h$ is bounded and locally integrable with respect to $F$ on $(a, b)$, and $|g|$ is improperly Riemann-Stieltjes integrable with respect to $F$ on $(a, b)$, then $|hg|$ is improperly Riemann-Stieltjes integrable with respect to $F$ on $(a, b)$.*

*Proof.* Note that $0 \leq |hg| \leq M|g|$. The result follows by the comparison theorem. □

**Definition 6.7.** Let $h : (a, b) \mapsto \mathbb{R}$. We say that $h$ is **absolutely integrable with respect to $F$ on** $(a, b)$ if $h$ is locally integrable with respect to $F$ on $(a, b)$ and $|h|$ is improperly integrable with respect to $F$ on $(a, b)$. We say that $h$ is **conditionally improperly integrable with respect to $F$ on** $(a, b)$ if $h$ is improperly integrable on $(a, b)$ but $|h|$ is not improperly integrable on $(a, b)$.

**Proposition 6.12.** *If $h$ is absolutely integrable with respect to $F$ on $(a, b)$, then $h$ is improperly integrable with respect to $F$ on $(a, b)$, and*

$$\left| \int_a^b h\, dF \right| \leq \int_a^b |h|\, dF.$$

*Proof.* Since $0 \leq |h| + h \leq 2|h|$, $|h| + h$ is improperly integrable with respect to $F$ on $(a, b)$ by the comparison theorem. Hence, by linearity of the integral, $h = (|h| + h) - |h|$ is improperly integrable with respect to $F$ on $(a, b)$. Furthermore, for every subinterval $[c, d]$ of $(a, b)$, we have

$$\left| \int_c^d h\, dF \right| \leq \int_c^d |h|\, dF.$$

The result then follows by letting $c \to a^+$ and $d \to b^-$. □

The converse of Proposition 6.12 is false as the following example shows.

**Example 6.4.** Integrating by parts, we have for all $d > 1$

$$\int_1^d \frac{\sin x}{x}\, dx = -\left.\frac{\cos x}{x}\right|_1^d - \int_1^d \frac{\cos x}{x^2}\, dx.$$

Since $1/x^2$ is improperly integrable on $[1, \infty)$, we have that $\cos x/x^2$ is absolutely integrable on $[1, \infty)$, and hence, it is improperly integrable on $[1, \infty)$. Taking the limit as $d \to \infty$, we have

$$\int_1^\infty \frac{\sin x}{x}\, dx = -\cos(1) - \int_1^\infty \frac{\cos x}{x^2}\, dx$$

exists and is finite. This proves that $\sin x/x$ is improperly integrable on $[1, \infty)$.

We now show that $|\sin x|/x$ is not improperly integrable on $[1, \infty)$, which proves that $\sin x/x$ is conditionally integrable on $[1, \infty)$. Note that if $n \in \mathbb{N}$ and $n \geq 2$, then

$$\int_1^{n\pi} \frac{|\sin x|}{x}\, dx \geq \sum_{k=2}^n \int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x}\, dx \geq \sum_{k=2}^n \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin x|\, dx = \frac{2}{\pi} \sum_{k=2}^n \frac{1}{k}.$$

Hence,

$$\lim_{n\to\infty} \int_1^{n\pi} \frac{|\sin x|}{x}\, dx = \infty.$$

We now turn to expectations of discrete and continuous random variables.

**Example 6.5.** Suppose that $X$ is a discrete random variable with CDF $F$. Let $\{x_n, n \geq 1\}$ be the support of $X$ and let $p_n = F(x_n) - F(x_n-) = P(X = x_n)$, and suppose that $E|X| = \sum_{n=1}^\infty |x_n| p_n < \infty$. Since $F$ is a CDF, it is monotonically increasing with $F(\infty) = 1$ and $F(-\infty) = 0$. Furthermore, note that

$$F(x) = \sum_{n=1}^\infty p_n I(x \geq x_n),\ p_n \geq 0,\ \forall n \geq 1 \ \text{ and } \ \sum_{n=1}^\infty p_n = 1 < \infty.$$

Let $h : \mathbb{R} \to \mathbb{R}$ be defined by $h(x) = x$. Then $|h|$ is continuous on $[a, b]$ for all $-\infty < a < b < \infty$, so for all $-\infty < a < b < \infty$

$$\int_a^b |h|\, dF = \sum_{n=1}^\infty |h(x_n)| I(a \leq x_n \leq b) p_n$$

by Proposition 6.7. Since

$$\sum_{n=1}^\infty |h(x_n)| I(a \leq x_n \leq b) p_n \leq \sum_{n=1}^\infty |x_n| p_n < \infty,$$

the series converges uniformly in $a$ and $b$ over $\mathbb{R}$ by the M-test. By applying Theorem 4.3 twice, we have

$$\begin{aligned}\lim_{a\to-\infty} \lim_{b\to\infty} \int_a^b |h|\, dF &= \lim_{a\to-\infty} \lim_{b\to\infty} \sum_{n=1}^\infty |h(x_n)| I(a \leq x_n \leq b) p_n \\ &= \sum_{n=1}^\infty \lim_{a\to-\infty} \lim_{b\to\infty} |h(x_n)| I(a \leq x_n \leq b) p_n\end{aligned}$$

$$= \sum_{n=1}^{\infty} |x_n| p_n$$
$$= E|X| < \infty.$$

Thus, $h$ is improperly Riemann-Stieltjes integrable on $\mathbb{R}$ and by a similar argument

$$\int_{-\infty}^{\infty} h \, dF = \sum_{n=1}^{\infty} x_n p_n = EX.$$

**Example 6.6.** Suppose that $X$ is a continuous random variable with CDF $F$ and pdf $f = F'$. Suppose that

$$E|X| = \int_{-\infty}^{\infty} |x| f(x) \, dx < \infty.$$

Let $h(x) = x$. Since $F$ is a CDF it is monotonically increasing with $F(\infty) = 1$ and $F(-\infty) = 0$. Furthermore, $f = F'$ is a continuous function which is improperly Riemann integrable on $\mathbb{R}$ with

$$\int_{-\infty}^{\infty} f(x) \, dx = 1 \text{ and } F(x) = \int_{-\infty}^{x} f(t) \, dt.$$

In particular, $|h|f$ is absolutely integrable on $\mathbb{R}$ by assumption, so by Proposition 6.8

$$\int_{-\infty}^{\infty} |x| \, dF = \lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b |x| \, dF$$
$$= \lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b |x| f(x) \, dx$$
$$= \int_{-\infty}^{\infty} |x| f(x) \, dx < \infty.$$

Thus, $h(x) = x$ is absolutely integrable with respect to $F$ on $\mathbb{R}$, and hence improperly integrable with respect to $F$ over $\mathbb{R}$. Again, by Proposition 6.8,

$$\int_{-\infty}^{\infty} x \, dF = \lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b x \, dF$$
$$= \lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b x f(x) \, dx$$
$$= \int_{-\infty}^{\infty} x f(x) \, dx$$
$$= EX.$$

## 6.5 Uniform Convergence and Integration

**Proposition 6.13.** *Let $F : [a, b] \mapsto \mathbb{R}$ is monotonically increasing with $F(b)$ and $F(a)$ finite. Suppose that $h_n \in \mathscr{R}(F, [a, b])$, $n \geq 1$, and suppose that $h_n \to h$ uniformly on $[a, b]$. Then $h \in \mathscr{R}(F, [a, b])$, and*

$$\int_a^b h \, dF = \lim_{n \to \infty} \int_a^b h_n \, dF.$$

*Proof.* Let

$$\varepsilon_n = \sup_{x \in [a,b]} |h_n(x) - h(x)|.$$

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

$$h_n - \varepsilon_n \le h \le h_n + \varepsilon_n,$$

so for any partition $P$ of $[a, b]$

$$L(h_n + \varepsilon_n, P, F) \le L(h, P, F) \le U(h, P, F) \le U(h_n + \varepsilon_n, P, F).$$

Thus,

$$\int_a^b (h_n + \varepsilon_n)\, dF \le \underline{\int_a^b} h\, dF \le \overline{\int_a^b} h\, dF \le \int_a^b (h_n + \varepsilon_n)\, dF. \quad (\star)$$

This implies that

$$0 \le \overline{\int_a^b} h\, dF - \underline{\int_a^b} h\, dF \le 2\varepsilon_n [F(b) - F(a)].$$

Since $\varepsilon_n \to 0$, $\underline{\int_a^b} h\ dF = \overline{\int_a^b} h\ dF$, so $h \in \mathscr{R}(F, [a, b])$. By two more applications of $(\star)$, have

$$\left| \int_a^b h\, dF - \int_a^b h_n\, dF \right| \le \varepsilon_n [F(b) - F(a)].$$

Again, $\varepsilon_n \to 0$, so $\int_a^b h\, dF = \lim_{n\to\infty} \int_a^b h_n\, dF$. □

## 6.6 Applications in Probability and Statistics

- Expectation in terms of integral for both discrete and continuous cases.
- Integral inequalities: e.g. Holder, Jensen, Cauchy-Schwarz
- Convergence of moments and Helly-Bray Theorem

# Chapter 7

# Measure Theory and the Lebesgue Integral

## 7.1 Measurable Spaces

## 7.2 Lebesgue-Stieltjes Measures

## 7.3 Measurable Mappings

## 7.4 Modes of Convergence

## 7.5 Lebesgue Integration