



# Mathematical Analysis

Author: Isaac FEI

## Preface

This book mainly follows the structure of [1].

# Contents

<b>Chapter 1 Set Theory</b>	<b>1</b>
1.1 Properties of Suprema and Infima . . . . .	1
<b>Chapter 2 Topology</b>	<b>2</b>
2.1 Compactness . . . . .	2
<b>Chapter 3 Numerical Sequences and Series</b>	<b>3</b>
3.1 Monotonic Sequences of Real Numbers . . . . .	3
3.2 Limit Superior and Limit Inferior . . . . .	3
3.3 Infinite Series . . . . .	10
3.4 Inserting and Removing Parentheses . . . . .	12
3.5 Alternating Series . . . . .	12
3.6 Absolute and Conditional Convergence . . . . .	14
3.7 The Comparison Test . . . . .	15
3.8 The Ratio Test and the Root Test . . . . .	15
3.9 Dirichlet's Test and Abel's Test . . . . .	19
3.10 Subseries . . . . .	20
3.11 Double Series . . . . .	24
3.12 Rearrangement Theorem for Double Series . . . . .	24
3.13 Iterated Series . . . . .	27
3.14 Multiplication of Series . . . . .	28
<b>Chapter 4 Limits and Continuity</b>	<b>31</b>
4.1 Limit of a Function . . . . .	31
4.2 Computing Limits . . . . .	31
4.3 Continuous Functions . . . . .	31
4.4 Continuity of Composite Functions . . . . .	32
4.5 Connectedness . . . . .	32
4.6 Monotonic Functions . . . . .	32
<b>Chapter 5 Differentiation</b>	<b>34</b>
5.1 Definition of Derivative . . . . .	34
5.2 Derivatives and Continuity . . . . .	38
5.3 Computing Derivatives . . . . .	40
5.4 The Mean Value Theorem . . . . .	43
5.5 Intermediate Value Theorem . . . . .	47
5.6 Taylor's Theorem . . . . .	52

---

<b>Chapter 6 Functions of Bounded Variation</b>	<b>56</b>
6.1 Definition of Functions of Bounded Variation . . . . .	56
6.2 Total Variation . . . . .	58
6.3 Additive Property of Total Variation . . . . .	59
6.4 Characterization of Functions of Bounded Variation . . . . .	60
<b>Chapter 7 Riemann-Stieltjes Integral</b>	<b>62</b>
7.1 Definition of the Riemann-Stieltjes Integral . . . . .	62
7.2 Linear Properties . . . . .	63
7.3 Reduction to Riemann-Stieltjes Integrals . . . . .	66
7.4 Integration by Parts . . . . .	68
7.5 Change of Variables . . . . .	69
7.6 Step Functions as Integrators . . . . .	70
7.7 Reduction of Riemann-Stieltjes Integrals to Finite Sums . . . . .	74
7.8 Darboux Integrals with Monotonically Increasing Integrators . . . . .	74
7.9 Riemann's Criterion . . . . .	79
7.10 Properties of Riemann-Stieltjes Integrals With Respect to Increasing Integrators . . . . .	82
7.11 Integrators of Bounded Variation . . . . .	89
7.12 Existence of Riemann-Stieltjes Integrals . . . . .	92
7.13 Mean Value Theorems for Riemann-Stieltjes Integrals . . . . .	93
7.14 Fundamental Theorems of Calculus . . . . .	96
7.15 Change of Variables in Riemann Integrals . . . . .	100
7.16 Lebesgue Criterion for Existence of Riemann Integrals . . . . .	101
7.17 Improper Riemann Integrals . . . . .	108
<b>Chapter 8 Sequences and Series of Functions</b>	<b>109</b>
8.1 Uniform Convergence . . . . .	109
8.2 Uniform Convergence and Continuity . . . . .	111
8.3 Uniform Convergence and Differentiation . . . . .	112
8.4 Power Series . . . . .	115
<b>Chapter 9 Transcendental Functions</b>	<b>121</b>
9.1 Exponential Function . . . . .	121
9.2 Definition of Pi . . . . .	122
9.3 Arcsine and Arccosine . . . . .	124
<b>References</b>	<b>125</b>
<b>Index</b>	<b>126</b>

# Chapter 1 Set Theory

## 1.1 Properties of Suprema and Infima

### Theorem 1.1.1

Let  $f : X \rightarrow \mathbb{R}$  be a real-valued function. We have the following equation.

$$\sup_{x \in X} f(x) - \inf_{x \in X} f(x) = \sup_{x, y \in X} [f(x) - f(y)] \quad (1.1)$$



**Proof** Let  $M$  and  $m$  denote

$$M = \sup_{x \in X} f(x) \quad \text{and} \quad m = \inf_{x \in X} f(x)$$

We have

$$f(x) - f(y) \leq M - m \quad \forall x, y \in X$$

It then follows that

$$\sup_{x, y \in X} [f(x) - f(y)] \leq M - m \quad (1.2)$$

On the other hand, given  $\varepsilon > 0$ , there exists  $x, y \in X$  such that

$$f(x) > M - \varepsilon/2 \quad \text{and} \quad f(y) < m + \varepsilon/2 \quad (1.3)$$

It follows that

$$\sup_{x, y \in X} [f(x) - f(y)] \geq f(x) - f(y) > M - m - \varepsilon$$

Because

$$\sup_{x, y \in X} [f(x) - f(y)] > M - m - \varepsilon$$

holds for any  $\varepsilon > 0$ , we have

$$\sup_{x, y \in X} [f(x) - f(y)] \geq M - m \quad (1.4)$$

Then, equation (1.1) follows from (1.2) and (1.4). ■

# Chapter 2 Topology

## 2.1 Compactness

### Theorem 2.1.1

*Every closed subspace of a compact space is also compact.*



**Proof**



# Chapter 3 Numerical Sequences and Series

We are interested in real or complex-valued sequences and series.

## 3.1 Monotonic Sequences of Real Numbers

### Theorem 3.1.1

A monotonic sequence  $\{a_n\}$  converges if and only if it is bounded. Furthermore, assuming  $\{a_n\}$  does converge,

1.  $\lim_{n \rightarrow \infty} a_n = \sup_{n \in \mathbb{N}^*} a_n$  if  $\{a_n\}$  is increasing, and
2.  $\lim_{n \rightarrow \infty} a_n = \inf_{n \in \mathbb{N}^*} a_n$  if  $\{a_n\}$  is decreasing.



## 3.2 Limit Superior and Limit Inferior

Given a real-valued sequence  $\{a_n\}$ , sometimes the information about the upper and lower bounds of the *entire* sequence, if they exist, may not be useful for we are often interested in the **tail** of the sequence, which is simply the portion of  $\{a_n\}$  after a few indices. To be specific, the  $n$ -th tail of  $\{a_n\}$  is defined by  $\{a_{m+n-1}\}_{m \in \mathbb{N}^*}$ , and is written as  $\{a_m\}_{m \geq n}$  in shorthand.

If  $\{a_n\}$  is bounded, we can define two sequences of upper and lower bounds,  $\{u_n\}$  and  $\{l_n\}$ , of the tail  $\{a_m\}_{m \geq n}$  by

$$u_n = \sup_{m \geq n} a_m \quad \text{and} \quad l_n = \inf_{m \geq n} a_m, \quad n \in \mathbb{N}^*$$

To study the limiting behavior of  $\{a_n\}$ , it is helpful to know how the  $n$ -th tail  $\{a_m\}_{m \geq n}$  is bounded when  $n$  is relatively large.

At a first glance, we observe that  $\{u_n\}$  is decreasing and  $\{l_n\}$  is increasing, which we shall prove in the next lemma.

### Lemma 3.2.1

Let  $\{a_n\}$  be a sequence of real numbers.

1. If  $\{a_n\}$  is bounded above, for each integer  $n \in \mathbb{N}^*$  we can define

$$u_n = \sup_{m \geq n} a_m, \quad n \in \mathbb{N}^*$$

The sequence  $\{u_n\}$  is decreasing.

2. If  $\{a_n\}$  is bounded below, for each integer  $n \in \mathbb{N}^*$  we can define

$$l_n = \inf_{m \geq n} a_m, \quad n \in \mathbb{N}^*$$

The sequence  $\{l_n\}$  is increasing.



**Proof** We only prove 1. For every  $m \in \mathbb{N}^*$  with  $m \geq n + 1$ , it is clear

$$a_m \in \{a_m \mid m \geq n, n \in \mathbb{N}^*\}$$

Note that  $u_n$  is an upper bound of the set on the right-hand side. Hence,

$$a_m \leq u_n \quad \forall m \geq n + 1$$

This implies  $u_n$  is also an upper bound of

$$\{a_m \mid m \geq n + 1, n \in \mathbb{N}^*\}$$

It then follows that  $u_{n+1} \leq u_n$  since  $u_{n+1}$  is the least upper bound of the set above. Therefore, we have shown

$$u_n \geq u_{n+1} \quad \forall n \in \mathbb{N}^*$$

Hence,  $\{u_n\}$  is decreasing. ■

### 3.2.1 Definition

By Theorem 3.1.1, we know that the limits of  $\{u_n\}$  and  $\{l_n\}$  always exist (as finite or infinite numbers). This gives rise to the definitions of limit superior as well as limit inferior. We also take into consideration the cases where  $\{a_n\}$  is not bounded.

#### Definition 3.2.1

Let  $\{a_n\}$  be a sequence of real numbers. The **limit superior** of  $\{a_n\}$ , denoted by  $\limsup_{n \rightarrow \infty} a_n$ , is defined as follows:

1.  $\limsup_{n \rightarrow \infty} a_n := \infty$  if  $\{a_n\}$  is not bounded above, and
2.  $\limsup_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} u_n$  if  $\{a_n\}$  is bounded above where  $u_n = \sup_{m \geq n} a_m$ .

The **limit inferior** of  $\{a_n\}$  is denoted by  $\liminf_{n \rightarrow \infty} a_n$  and is defined similarly. ♣



**Note** The notation  $\limsup$  consists of two parts, the limit symbol and the supremum symbol. Indeed, as we can see from the definition, this notation arises naturally for we can write

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sup_{m \geq n} a_m$$

if  $\{a_n\}$  is bounded above.

**Example 3.1** Let

$$a_n = \left(1 + \frac{10}{n}\right) \sin \frac{n\pi}{32}$$

We claim that the limit superior and limit inferior of  $\{a_n\}$  are 1 and  $-1$ , respectively. See Figure 3.1.

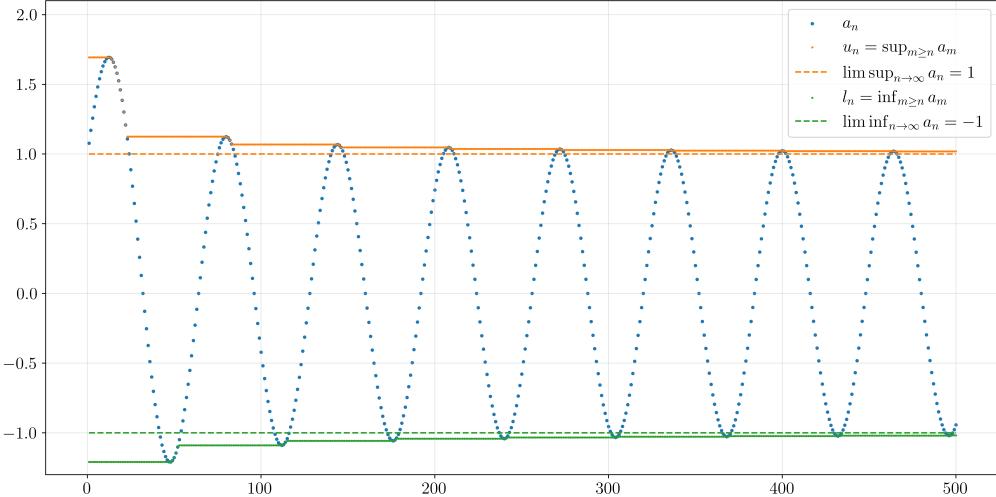
Now, we show that the values of limit superior and limit inferior are indeed as claimed. Note that for any given integer  $n \in \mathbb{N}^*$ , we can choose an integer  $k$  such that  $16 + 64k \geq n$ . Let  $m = 16 + 64k$ . We have

$$a_m = \left(1 + \frac{10}{m}\right) \sin \frac{(16 + 64k)\pi}{32} = \left(1 + \frac{10}{m}\right) \sin \left(\frac{\pi}{2} + 2k\pi\right) = 1 + \frac{10}{m} > 1$$

This implies that

$$u_n = \sup_{m \geq n} a_m > 1 \quad \forall n \in \mathbb{N}^* \tag{3.1}$$

which means the sequence  $\{u_n\}$  is bounded below by 1. We know by Lemma 3.2.1 that  $\{u_n\}$  is decreasing. Thus, by Theorem 3.1.1,  $\{u_n\}$  converges to the infimum of its range,  $\inf_{n \in \mathbb{N}^*} u_n$ . We need



**Figure 3.1:** Limit superior and limit inferior of  $\{a_n\}$  where  $a_n = \left(1 + \frac{10}{n}\right) \sin \frac{n\pi}{32}$ .

to show that  $\inf_{n \in \mathbb{N}^*} u_n = 1$ . Given  $\varepsilon > 0$ , there exists  $n \in \mathbb{N}^*$  such that  $10/n < \varepsilon/2$ . Then, for every  $m \geq n$ , we have

$$1 + \frac{10}{m} < 1 + \varepsilon/2$$

Because the absolute value of the sine function is bounded by 1, it follows that

$$a_m = \left(1 + \frac{10}{m}\right) \sin \frac{m\pi}{32} < 1 + \varepsilon/2 \quad \forall m \geq n$$

Taking the supremum, we obtain

$$u_n = \sup_{m \geq n} a_m \leq 1 + \varepsilon/2 < 1 + \varepsilon$$

Therefore,

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} u_n = \inf_{n \in \mathbb{N}^*} u_n = 1$$

The proof of  $\liminf a_n = -1$  is similar and is left as an exercise.

When the limit superior/inferior of  $\{a_n\}$  is finite, then we have the following equivalent conditions, which may be helpful in proofs.

### Theorem 3.2.1

Let  $\{a_n\}$  be a sequence of real numbers. Suppose there exists a (finite) number  $u$  satisfying the following two conditions:

1. For every  $\varepsilon > 0$ , there exists an integer  $N$  such that

$$a_n < u + \varepsilon \quad \forall n \geq N$$

2. Given  $\varepsilon > 0$  and given an integer  $m$ , there exists an integer  $N \geq m$  such that

$$a_N > u - \varepsilon$$

Then  $u$  is the limit superior of  $\{a_n\}$ , i.e.,  $u = \limsup a_n$ . Conversely, if  $u = \limsup a_n$  is a finite number, then it satisfies the above conditions. An analogous result exists concerning the limit inferior of  $\{a_n\}$ .



Given  $\varepsilon > 0$ , the first condition says  $u + \varepsilon$  is an upper bound for some  $N$ -th tail of  $\{a_n\}$ . In other

words, all the terms of  $\{a_n\}$  will eventually fall below the number  $u + \varepsilon$ . While the second condition says there are *infinitely* many terms of  $\{a_n\}$  that exceed  $u - \varepsilon$ .

 **Note** These two conditions are used to define the limit superior in Apostol's book[1]. The problem is that we cannot assume the existence of such a number  $u$  if we adopt this definition, although the existence can be proved later. But this would increase the difficulty in understanding the concept of limit superior/inferior.

**Proof** (Sufficiency) We first assume  $u$  satisfies the above two conditions and show that it is indeed the limit superior of  $\{a_n\}$ . Note that the first condition implies that  $\{a_n\}$  is bounded above. Hence, it is feasible to define

$$u_n = \sup_{m \geq n} a_m, \quad n \in \mathbb{N}^*$$

By Lemma 3.2.1, we know  $\{u_n\}$  is decreasing. Let  $m \in \mathbb{N}^*$  be fixed for now, by exploiting the second condition, for any  $\varepsilon > 0$ , there exists an integer  $N \geq m$  such that

$$a_N > u - \varepsilon$$

It then follows that

$$u_m = \sup_{n \geq m} a_n \geq a_N > u - \varepsilon$$

Since

$$u_m > u - \varepsilon$$

holds for any  $\varepsilon > 0$ . We have

$$u_m \geq u \tag{3.2}$$

But  $m \in \mathbb{N}^*$  is also chosen arbitrarily. Therefore, the above inequality implies that the sequence  $\{u_n\}$  is bounded below by  $u$ .

Next, we show that  $u$  is the infimum of the range of  $\{u_n\}$ . Given  $\varepsilon > 0$ , by condition 1, there exists  $N \in \mathbb{N}^*$  such that

$$a_n < u + \varepsilon/2 \quad \forall n \geq N$$

Taking the supremum over  $n$ , we obtain

$$u_N = \sup_{n \geq N} a_n \leq u + \varepsilon/2 < u + \varepsilon$$

That is,

$$u_N < u + \varepsilon \tag{3.3}$$

It then follows from (3.2) and (3.3) that

$$u = \inf_{n \in \mathbb{N}^*} u_n$$

Recall  $\{u_n\}$  is decreasing. Then by Theorem 3.1.1,  $\lim u_n$  is precisely the infimum of its range since it is bounded below. Finally, since the limit superior of  $\{a_n\}$  is defined by  $\lim u_n$ , we conclude that  $u = \limsup a_n$ .

(Necessity) Assume that the limit superior of  $\{a_n\}$ ,  $u$ , is a finite number. By definition,  $\{a_n\}$  is

bounded above. Hence, we can again define

$$u_n = \sup_{m \geq n} a_m, \quad n \in \mathbb{N}^*$$

And by definition along with the fact that  $\{u_n\}$  is decreasing, we have

$$u = \lim_{n \rightarrow \infty} u_n = \inf_{n \in \mathbb{N}^*} u_n$$

We need to show  $u$  indeed satisfies the two conditions stated in this theorem.

Given  $\varepsilon > 0$ , since  $u = \inf u_n$ , there exists  $N \in \mathbb{N}^*$  such that

$$u_N < u + \varepsilon$$

But  $u_N = \sup_{n \geq N} a_n$ . Therefore,

$$\sup_{n \geq N} a_n < u + \varepsilon$$

And because  $a_n \leq \sup_{n \geq N} a_n \forall n \geq N$ , we have

$$a_n < u + \varepsilon \quad \forall n \geq N$$

This shows that  $u$  satisfies condition 1.

Given  $\varepsilon > 0$  and  $m \in \mathbb{N}^*$ . Because  $u$  is a lower bound of  $\{u_n\}$ , we have

$$\sup_{n \geq m} a_n = u_m \geq u \tag{3.4}$$

By the property of supremum, there exists an integer  $N \geq m$  such that

$$a_N > \sup_{n \geq m} a_n - \varepsilon \tag{3.5}$$

Combining (3.4) with (3.5), we find

$$a_N > u - \varepsilon$$

which implies that  $u$  also satisfies condition 2. This completes the proof. ■

### 3.2.2 Properties

#### Proposition 3.2.1

Let  $\{a_n\}$  be a sequence of real numbers. Then

1.  $\limsup a_n = -\infty \implies \liminf a_n = -\infty$ , and
2.  $\{a_n\}$  diverges to  $-\infty$  if and only if  $\limsup a_n = \liminf a_n = -\infty$ .



**Proof** (Proof of 1) Note that the fact that the limit superior is not positive infinity implicitly implies that  $\{a_n\}$  is bounded above. Define

$$u_n = \sup_{m \geq n} a_m, \quad n \in \mathbb{N}^* \tag{3.6}$$

We have  $\lim u_n = -\infty$ . Hence, for every  $M > 0$ , there exists  $N \in \mathbb{N}^*$  such that

$$u_n \leq -M \quad \forall n \geq N$$

But  $a_n \leq \sup_{m \geq n} a_m = u_n$ . It follows that

$$a_n \leq -M \quad \forall n \geq N$$

This implies  $\{a_n\}$  diverges to  $-\infty$ , that is,  $\lim a_n = -\infty$ . As a consequence,  $\{a_n\}$  is not bounded below. Therefore,  $\liminf a_n = -\infty$  by definition.

(Proof of 2) Note that in proving the first statement, we have already shown that  $\limsup a_n = \liminf a_n = -\infty$  implies  $\lim a_n = -\infty$ . We only need to prove the other direction. Suppose  $\lim a_n = -\infty$ . For every  $M > 0$ , there exists  $N \in \mathbb{N}^*$  such that

$$a_n \leq -M \quad \forall n \geq N$$

 **Note** Note that the above inequality implies that  $\{a_n\}$  is bounded above. Hence, we can again define  $\{u_n\}$  as in (3.6).

Taking the supremum over  $n$ , we obtain

$$u_N = \sup_{n \geq N} a_n \leq -M$$

Note that  $\{u_n\}$  is decreasing by Lemma 3.2.1. Hence,

$$u_n \leq u_N \leq -M \quad \forall n \geq N$$

This implies that  $\lim u_n = -\infty$ . Therefore, by the definition of limit superior,  $\limsup a_n = -\infty$ . Then by statement 1, which we have proved earlier,  $\liminf a_n = -\infty$ . ■

As an analogous result, we also have the following proposition, the proof of which is similar and is left as an exercise.

### Proposition 3.2.2

Let  $\{a_n\}$  be a sequence of real numbers. Then

1.  $\liminf a_n = \infty \implies \limsup a_n = \infty$ , and
2.  $\{a_n\}$  diverges to  $\infty$  if and only if  $\limsup a_n = \liminf a_n = \infty$ .



### Proposition 3.2.3

Let  $\{a_n\}$  be a sequence of real numbers. Then  $\{a_n\}$  converges if and only if both limit superior and limit inferior are equal to a finite number. That is,

$$\lim a_n = a \iff \limsup a_n = \liminf a_n = a$$

where  $a$  is finite.



**Proof** (Sufficiency) Suppose that the limit superior and limit inferior are equal to a finite number, i.e.,  $\limsup a_n = \liminf a_n = a$ . Let  $\varepsilon > 0$  be chosen arbitrarily. Because  $a$  is the limit superior, by the first condition in Theorem 3.2.1, there exists  $N_1 \in \mathbb{N}^*$  such that

$$a_n < a + \varepsilon \quad \forall n \geq N_1$$

Similarly, if we treat  $a$  as the limit superior, then there exists  $N_2 \in \mathbb{N}^*$  such that

$$a_n > a - \varepsilon \quad \forall n \geq N_2 \tag{3.7}$$

Let  $N = \max\{N_1, N_2\}$ . We have

$$|a_n - a| < \varepsilon \quad \forall n \geq N$$

Therefore,  $\{a_n\}$  converges to  $a$ .

(Necessity) Suppose  $\{a_n\}$  converges to  $a$ . Given  $\varepsilon > 0$ , there exists  $N \in \mathbb{N}^*$  such that

$$|a_n - a| < \varepsilon \quad \forall n \geq N$$

We have

1.  $a_n < a + \varepsilon$  for all  $n \geq N$ , and
2. for any given  $m \in \mathbb{N}^*$ , there exists  $M \geq \max\{m, N\}$  such that  $a_M > a - \varepsilon$ .

By Theorem 3.2.1,  $a = \limsup a_n$ . Similarly, we can also show  $a = \liminf a_n$ . ■

Like the limit and supremum/infimum, limit superior/inferior also preserves the order.

### Theorem 3.2.2

If  $a_n \leq b_n \forall n \in \mathbb{N}^*$ , then

$$\limsup a_n \leq \limsup b_n \quad \text{and} \quad \liminf a_n \leq \liminf b_n$$



**Proof** We only prove  $\limsup a_n \leq \limsup b_n$ .

(Extreme Cases) If  $\limsup a_n = -\infty$  or  $\limsup b_n = \infty$ , then there is nothing to prove. If  $\limsup a_n = \infty$ , then  $\{a_n\}$  is not bounded above and nor is  $\{b_n\}$ , and hence  $\limsup b_n = \infty$ . Suppose  $\limsup b_n = -\infty$ , then  $\lim b_n = -\infty$  by Proposition 3.2.1. It follows that  $\lim a_n = -\infty$ , and hence  $\limsup a_n = -\infty$ .

(Finite Limit Superiors) Now, suppose that  $\limsup a_n$  and  $\limsup b_n$  are both finite. We have

$$\sup_{m \geq n} a_m \leq \sup_{m \geq n} b_m \quad \forall n \in \mathbb{N}^*$$

Taking the limit on both sides yields

$$\limsup a_n = \lim_{n \rightarrow \infty} \sup_{m \geq n} a_m \leq \lim_{n \rightarrow \infty} \sup_{m \geq n} b_m = \limsup b_n$$



The next theorem can be treated as the reciprocal rule of limit superior, which establishes the relation between  $\limsup 1/a_n$  and  $\liminf a_n$ .

### Theorem 3.2.3

Let  $\{a_n\}$  be a sequence of positive numbers, then

1.  $\limsup \frac{1}{a_n} = 1/\liminf a_n$ , and
2.  $\liminf \frac{1}{a_n} = 1/\limsup a_n$ .

Note that, in the first equation, if  $\liminf a_n = 0$  then  $\limsup 1/a_n = \infty$ , and  $\limsup 1/a_n = 0$  if  $\liminf a_n = \infty$ . A similar treatment also applies to the second equation.



**Proof** We only prove 1 since the second equation can be obtained by replacing  $a_n$  with  $1/a_n$  in the first one.

(Case 1:  $\liminf a_n = 0$ .) For every positive number  $M > 0$ . By Theorem 3.2.1 (the second condition for limit inferior), there exists  $N \in \mathbb{N}^*$  such that

$$a_N < 0 + \frac{1}{M} = \frac{1}{M}$$

Taking the reciprocal, we find

$$\frac{1}{a_N} > M$$

This implies that the sequence  $\{1/a_n\}$  is not bounded above. Hence,  $\limsup 1/a_n = \infty$  by definition.

(Case 2:  $\liminf a_n = \infty$ .) We know  $\lim a_n = \infty$  by Proposition 3.2.2. This implies  $\{1/a_n\}$  will converge to zero, i.e.,  $\lim 1/a_n = 0$ . By Proposition 3.2.3,  $\limsup a_n = 0$ .

(Case 3:  $\liminf a_n$  is a finite positive number.) Write  $l = \liminf a_n$ . Let  $\varepsilon > 0$  be chosen arbitrarily. Put

$$\varepsilon_1 = \frac{\varepsilon l^2}{1 + \varepsilon l}$$

By Theorem 3.2.1, there exists  $N \in \mathbb{N}^*$  such that

$$a_n > l - \varepsilon_1 \quad \forall n \geq N$$

We note that  $l - \varepsilon_1 > 0$ . Taking the reciprocal of both sides and then plugging in the value of  $\varepsilon_1$ , we find

$$\frac{1}{a_n} < \frac{1}{l - \varepsilon_1} = \frac{1}{l} + \varepsilon \quad \forall n \geq N \quad (3.8)$$

Now, choose another  $\varepsilon > 0$  so small that  $1 - \varepsilon l > 0$ . Let

$$\varepsilon_2 = \frac{\varepsilon l^2}{1 - \varepsilon l}$$

 **Note** The reason why we choose  $\varepsilon$  such that  $1 - \varepsilon l > 0$  is that we want  $\varepsilon_2 > 0$ . Notice also that the choice of sufficient small  $\varepsilon$  in this way will not affect the generality of the proof.

Given any integer  $m \in \mathbb{N}^*$ , by Theorem 3.2.1, there exists an integer  $N \geq m$  such that

$$a_N < l + \varepsilon_2$$

Taking the reciprocal and then plugging in the value of  $\varepsilon_2$ , we obtain

$$\frac{1}{a_N} > \frac{1}{l + \varepsilon_2} = \frac{1}{l} - \varepsilon \quad (3.9)$$

Therefore, by Theorem 3.2.1, we may conclude  $\limsup a_n = 1/l = 1/\liminf a_n$  from (3.8) and (3.9). ■

### 3.3 Infinite Series

The next theorem tells us that we can add two convergent series term by term. And we can also multiply a convergent series by a constant.

#### Theorem 3.3.1

If  $\sum a_n$  and  $\sum b_n$  are two convergent series, then the series  $\sum(\alpha a_n + \beta b_n)$  also converges where  $\alpha$  and  $\beta$  are constants. Furthermore, we have

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n$$



#### Proof

It is also common that only one of the two series is convergent while the other one diverges to positive or negative infinity. In this case, we have the following results.

**Theorem 3.3.2**

If  $\sum a_n$  diverges to infinity and  $\sum b_n$  converges, then

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \begin{cases} \infty, & \text{if } \alpha > 0 \\ -\infty, & \text{if } \alpha < 0 \end{cases}$$

An analogous result exists when  $\sum a_n = -\infty$ .



Recall the convergence of a series  $\sum a_n$  is defined by the convergence of the sequence of its partial sums  $\{s_n\}$ . Hence, by interpreting Cauchy's criterion for the sequence  $\{s_n\}$  in terms of sums of terms  $a_n$ s, we will obtain a version of **Cauchy's criterion** for series.

**Theorem 3.3.3 (Cauchy's Criterion for Series)**

The series  $\sum a_n$  converges if and only if for every  $\varepsilon > 0$ , there exists an integer  $N \in \mathbb{N}^*$  such that

$$\left| \sum_{k=1}^p a_{n+k} \right| = |a_{n+1} + \cdots + a_{n+p}| < \varepsilon \quad \forall n \geq N \quad \forall p \in \mathbb{N}^*$$

**Corollary 3.3.1**

If  $\sum a_n$  converges, then the absolute value of its term tends to zero. That is,

$$\lim_{n \rightarrow \infty} |a_n| = 0$$

**Theorem 3.3.4**

If series  $\sum a_n$  converges, then

1. the set of its partial sums is bounded, that is,

$$\left| \sum_{k=1}^n a_k \right| \leq M \quad \forall n \in \mathbb{N}^*$$

for some positive constant  $M$ , and

2. even more generally, all finite sums of consecutive terms are bounded by a common constant, that is, there exists  $M > 0$  such that

$$\left| \sum_{k=m}^n a_k \right| \leq M \quad \forall m \in \mathbb{N}^* \quad \forall n \geq m$$

**Proof****Theorem 3.3.5 (Weierstrass M-Test)**

Let  $\{M_n\}$  be a sequence of non-negative terms such that

$$|f_n(x)| \leq M_n \quad \forall n \in \mathbb{N}^* \quad \forall x \in S$$

Then  $\sum f_n(x)$  converges uniformly on  $S$  if  $\sum M_n$  converges.

**Proof**

## 3.4 Inserting and Removing Parentheses

### Definition 3.4.1

Given a series  $\sum a_n$ , we say the series  $\sum A_n$  is obtained by inserting parentheses into  $\sum a_n$  if there is a function  $p : \mathbb{N}^* \rightarrow \mathbb{N}^*$  satisfying

$$n > m \implies p(n) > p(m)$$

and each term  $A_n$  is given by

1.  $A_1 = a_1 + \dots + a_{p(1)}$ , and
2.  $A_n = a_{p(n-1)+1} + \dots + a_{p(n)}$  if  $n \geq 2$ .

The function  $p$  is also referred to as the locations of parentheses.



### Theorem 3.4.1

Let  $\sum A_n$  be a series obtained from  $\sum a_n$  by inserting parentheses. If  $\sum a_n$  converges to sum  $s$ , then  $\sum A_n$  also converges to sum  $s$ .



**Proof** Given  $\varepsilon > 0$ , since  $\sum a_n$  converges to  $s$ , there exists an integer  $N \in \mathbb{N}^*$  such that

$$|s_n - s| < \varepsilon \quad \forall n \geq N$$

where  $s_n$  is the  $n$ -th partial sum of  $\sum a_n$ . Suppose the locations of parentheses are given by function  $p$ . We have

$$\left| \sum_{k=1}^n A_k - s \right| = |s_{p(n)} - s| < \varepsilon \quad \forall n \geq N$$

since  $p(n) \geq n \quad \forall n \in \mathbb{N}^*$  (Exercise 3.1). Therefore,  $\sum A_n$  also converges to  $s$ . ■

☞ **Exercise 3.1** Complete the above proof by showing

$$p(n) \geq n \quad \forall n \in \mathbb{N}^*$$

**Solution** We shall prove by induction.

**Base Case:** It is clear that  $p(1) \geq 1$ .

**Inductive Step:** Assume  $p(k) \geq k$ . We have  $p(k+1) > p(k) \geq k$  by the definition. Therefore, we must have  $p(k+1) \geq k+1$ .

## 3.5 Alternating Series

### Definition 3.5.1

If  $a_n \geq 0$ , then we call the series of the form

$$\sum_{n=1}^{\infty} (-1)^{n+1} a_n$$

the alternating series.



The next theorem is known as the **alternating series test**.

**Theorem 3.5.1 (Alternating Series Test)**

If  $\{a_n\}$  is a sequence such that

1.  $a_n \geq 0 \forall n \geq \mathbb{N}^*$ ,
2. it is decreasing, and
3. it converges to zero, i.e.,  $\lim_{n \rightarrow \infty} a_n = 0$ ,

then the alternating series  $\sum (-1)^{n+1} a_n$  converges. Moreover, suppose the series converges to sum  $s$ , and  $\{s_n\}$  is the sequence of its partial sums, then we have

$$0 \leq (-1)^n (s - s_n) \leq a_{n+1} \quad (3.10)$$



**Note** What (3.10) tells us is that the absolute value of the sum of the tail of the series starting from the  $(n + 1)$ -term is bounded above by the value of the  $(n + 1)$ -th term. This shows how quickly this alternating series will converge. In other words, when we approximate the sum of this series by the sum of first  $n$  terms, the error is bounded by  $a_{n+1}$ .

**Proof** We have

$$s_{n+p} - s_n = (-1)^{n+2} a_{n+1} + (-1)^{n+3} a_{n+2} + \cdots + (-1)^{n+p+1} a_{n+p}$$

Multiplying  $(-1)^n$  on both sides, we find

$$\begin{aligned} & (-1)^n (s_{n+p} - s_n) \\ &= \begin{cases} a_{n+1} - a_{n+2} + a_{n+3} - a_{n+4} + a_{n+5} - \cdots - a_{n+p-1} + a_{n+p} & \text{if } p \text{ is odd} \\ a_{n+1} - a_{n+2} + a_{n+3} - a_{n+4} + a_{n+5} - \cdots - a_{n+p-2} + a_{n+p-1} - a_{n+p} & \text{if } p \text{ is even} \end{cases} \end{aligned} \quad (3.11)$$

By grouping every two consecutive terms on the right-hand side of (3.11), possibly except for the last term when  $p$  is odd, we find

$$\begin{aligned} & (-1)^n (s_{n+p} - s_n) \\ &= \begin{cases} (a_{n+1} - a_{n+2}) + (a_{n+3} - a_{n+4}) + \cdots + (a_{n+p-2} - a_{n+p-1}) + a_{n+p} & \text{if } p \text{ is odd} \\ (a_{n+1} - a_{n+2}) + (a_{n+3} - a_{n+4}) + \cdots + (a_{n+p-1} - a_{n+p}) & \text{if } p \text{ is even} \end{cases} \\ &\geq 0 \end{aligned}$$

since  $\{a_n\}$  is decreasing (and  $a_{n+p} \geq 0$  for the case where  $p$  is odd).

On the other hand, by grouping every two consecutive terms on the right-hand side of (3.11) starting from  $a_{n+2}$ , possibly except for  $a_{n+p}$  when  $p$  is even, we obtain

$$\begin{aligned} & (-1)^n (s_{n+p} - s_n) \\ &= \begin{cases} a_{n+1} - (a_{n+2} - a_{n+3}) - (a_{n+4} - a_{n+5}) - \cdots - (a_{n+p-1} - a_{n+p}) & \text{if } p \text{ is odd} \\ a_{n+1} - (a_{n+2} - (a_{n+3})) - (a_{n+4} - a_{n+5}) - \cdots - (a_{n+p-2} - a_{n+p-1}) - a_{n+p} & \text{if } p \text{ is even} \end{cases} \\ &\leq a_{n+1} \end{aligned}$$

In summary, we have shown that

$$0 \leq (-1)^n (s_{n+p} - s_n) \leq a_{n+1} \quad \forall n \in \mathbb{N}^* \quad \forall p \in \mathbb{N}^* \quad (3.12)$$

Given  $\varepsilon > 0$ , because  $\{a_n\}$  converges to zero, there exists  $N \in \mathbb{N}^*$  such that  $|a_n| < \varepsilon$  for all  $n \geq N$ . It then follows from (3.12) that

$$|s_{n+p} - s_n| < \varepsilon \quad \forall n \geq N \quad \forall p \in \mathbb{N}^*$$

Then Cauchy's criterion implies the series  $\sum (-1)^{n+1} a_n$  indeed converges. Say it converges to sum  $s$ . Letting  $p \rightarrow \infty$  in (3.12) yields (3.10).  $\blacksquare$

## 3.6 Absolute and Conditional Convergence

The next theorem reveals an underlying property about a conditionally convergent series.

### Theorem 3.6.1

Let  $\sum a_n$  be a series of real terms. Define

$$p_n = \frac{|a_n| + a_n}{2} \quad \text{and} \quad q_n = \frac{|a_n| - a_n}{2}, \quad n \in \mathbb{N}^*$$

Then we have the following:

1. If  $\sum a_n$  converges conditionally, then both  $\sum p_n$  and  $\sum q_n$  diverges to infinity, i.e.,

$$\sum_{n=1}^{\infty} p_n = \sum_{n=1}^{\infty} q_n = \infty \tag{3.13}$$

2. If  $\sum a_n$  converges absolutely, then both  $\sum p_n$  and  $\sum q_n$  converge, and we have

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} p_n - \sum_{n=1}^{\infty} q_n \tag{3.14}$$



**Note** Note that

$$p_n = \begin{cases} a_n, & \text{if } a_n \geq 0 \\ 0, & \text{if } a_n < 0 \end{cases} \quad \text{while} \quad q_n = \begin{cases} 0, & \text{if } a_n \geq 0 \\ -a_n, & \text{if } a_n < 0 \end{cases}$$

In other words,  $\sum p_n$  is the sum of all non-negative terms of  $\sum a_n$ , while  $\sum q_n$  is the absolute value of the sum of all negative terms. What this theorem mainly tells us is that for a conditionally convergent series, the sum of all its non-negative terms is infinity, and the sum of all its negative terms is negative infinity.

**Proof** (Proof of 1) We have  $\sum |a_n| = \infty$  since  $\sum a_n$  only converges conditionally. Then (3.13) follows immediately from Theorem 3.3.2.

(Proof of 2) Since  $\sum |a_n|$  also converges, (3.14) then follows from Theorem 3.3.1.  $\blacksquare$

## 3.7 The Comparison Test

### Theorem 3.7.1 (Comparison Test)

Let  $\{a_n\}$  and  $\{b_n\}$  be two sequences with non-negative terms. If there exists an integer  $N \in \mathbb{N}^*$  such that

$$a_n \leq b_n \quad \forall n \geq N$$

then

1.  $\sum a_n$  converges if  $\sum b_n$  converges, and
2.  $\sum b_n$  diverges if  $\sum a_n$  diverges.



### Proof



## 3.8 The Ratio Test and the Root Test

### Theorem 3.8.1 (Ratio Test)

Given a series  $\sum a_n$  with nonzero complex terms, let

$$r = \liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \quad \text{and} \quad R = \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

Then

1.  $\sum a_n$  converges absolutely if  $R < 1$ ,
2.  $\sum a_n$  diverges if  $r > 1$ , and
3. the test is inclusive if  $r \leq 1 \leq R$ .



**Proof** (Proof of 1) Choose a positive number  $\varepsilon$  such that  $R + \varepsilon < 1$ . By the definition of limit superior, there exists an integer  $N \in \mathbb{N}^*$  such that

$$\left| \frac{a_{n+1}}{a_n} \right| < R + \varepsilon \quad \forall n \geq N$$

Multiplying by  $|a_n|$  on both sides yields

$$|a_{n+1}| < (R + \varepsilon) |a_n| \quad \forall n \geq N$$

By applying the above inequality repeatedly, we obtain

$$|a_n| < (R + \varepsilon)^{n-N} |a_N| = [(R + \varepsilon)^{-N} |a_N|] (R + \varepsilon)^n \quad \forall n \geq N + 1$$

Note that  $\sum (R + \varepsilon)^n$  converges since it is a geometric series the absolute values of whose terms are less than 1. Therefore,  $\sum |a_n|$  converges by the comparison test (Theorem 3.7.1).

(Proof of 2) The proof of statement 2 is similar. This time, we can choose a positive number  $\varepsilon$  such that  $r - \varepsilon > 1$ . Then

$$\left| \frac{a_{n+1}}{a_n} \right| > r - \varepsilon \quad \forall n \geq N$$

for some fixed integer  $N \in \mathbb{N}^*$ . By applying a similar argument, we have

$$|a_n| > [(r - \varepsilon)^{-N} |a_N|] (r - \varepsilon)^n \quad \forall n \geq N + 1 \tag{3.15}$$

 **Note** If we apply the comparison test here, then we can only conclude that the series  $\sum |a_n|$  diverges since the geometric series  $\sum(r - \varepsilon)^n$  diverges. But this result is not sufficient enough to conclude that  $\sum a_n$  also diverges.

By observing the right-hand side of (3.15), we note that the sequence  $\{|a_n|\}$  does not converge to zero. In fact,

$$\lim_{n \rightarrow \infty} |a_n| = \infty$$

Therefore,  $\sum a_n$  diverges by Theorem 3.3.1.

(Proof of 3) See Example 3.3 and Example 3.4. ■

Whenever the ratio test is inclusive, one can always try the **root test** as we shall soon introduce. It also has a close connection with the power series and the concept of the radius of convergence, which we shall cover in Section 8.4.

### Theorem 3.8.2 (Root Test)

Let  $\sum a_n$  be a series of complex numbers. Let

$$\rho = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$$

Then

1.  $\sum a_n$  converges absolutely if  $\rho < 1$ ,
2.  $\sum a_n$  diverges if  $\rho > 1$ , and
3. the test is inclusive if  $\rho = 1$ .



**Proof** (Proof of 1) Choose  $\varepsilon > 0$  such that  $\rho + \varepsilon < 1$ . By the property of limit superior, there exists  $N \in \mathbb{N}^*$  such that

$$\sqrt[n]{|a_n|} < \rho + \varepsilon \quad \forall n \geq N$$

Hence,

$$|a_n| < (\rho + \varepsilon)^n \quad \forall n \geq N$$

Note that the geometric series  $\sum(\rho + \varepsilon)^n$  converges. Hence,  $\sum |a_n|$  converges by the comparison test. In other words,  $\sum a_n$  converges absolutely.

(Proof of 2) Choose  $\varepsilon > 0$  such that  $\rho - \varepsilon > 1$ . Again by the property of limit superior, for any integer  $N \in \mathbb{N}^*$ , we can always find an integer  $n \geq N$  such that

$$\sqrt[n]{|a_n|} > \rho - \varepsilon$$

That is,

$$|a_n| > (\rho - \varepsilon)^n \geq (\rho - \varepsilon)$$

This implies that the sequence  $\{|a_n|\}$  cannot converge to zero. Therefore,  $\sum a_n$  diverges by Corollary 3.3.1.

(Proof of 3) See Example 3.3 and Example 3.4. ■

**Example 3.2** Let

$$a_n = \begin{cases} \left(\frac{i}{2}\right)^{n-1}, & \text{if } n \text{ is odd} \\ \left(\frac{i}{2}\right)^n, & \text{if } n \text{ is even} \end{cases}, \quad n \in \mathbb{N}^*$$

Let us investigate whether  $\sum a_n$  converges or not.

We first apply the ratio test. Note that

$$\left| \frac{a_{n+1}}{a_n} \right| = \begin{cases} \frac{1}{4}, & \text{if } n \text{ is odd} \\ 1, & \text{if } n \text{ is even} \end{cases}$$

Hence,

$$r = \liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{4} \quad \text{and} \quad R = \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$$

The ratio test is inconclusive.

Now, let us try the root test. We have

$$\sqrt[n]{|a_n|} = \begin{cases} \left(\frac{1}{2}\right)^{1-1/n}, & \text{if } n \text{ is odd} \\ \frac{1}{2}, & \text{if } n \text{ is even} \end{cases}$$

Taking the limit superior, we obtain

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \frac{1}{2} < 1$$

Therefore,  $\sum a_n$  converges absolutely.

From the above example, the reader may conjecture that the root test is somehow more powerful than the ratio test. This is indeed true as we will see in the next theorem.

### Theorem 3.8.3

Let  $\{a_n\}$  be a sequence of positive numbers. We have

$$\liminf \frac{a_{n+1}}{a_n} \leq \liminf \sqrt[n]{a_n} \leq \limsup \sqrt[n]{a_n} \leq \limsup \frac{a_{n+1}}{a_n}$$



**Proof** Note that the second inequality follows directly from Theorem 3.2.2. What remains to show are the first and the last inequalities. We only prove

$$\limsup \sqrt[n]{a_n} \leq \limsup \frac{a_{n+1}}{a_n}$$

If  $\limsup a_{n+1}/a_n$ , then the above inequality holds trivially. Assume  $r = \limsup a_{n+1}/a_n$  is finite.

Given  $\varepsilon > 0$ , by Theorem 3.2.1, there exists  $N \in \mathbb{N}^*$  such that

$$\frac{a_{n+1}}{a_n} < r + \varepsilon \quad \forall n \geq N$$

It then follows that

$$\frac{a_{N+m}}{a_N} = \frac{a_{N+1}}{a_N} \frac{a_{N+2}}{a_{N+1}} \dots \frac{a_{N+m}}{a_{N+m-1}} < (r + \varepsilon)^m \quad \forall m \in \mathbb{N}^*$$

Multiplying by  $a_N$  on both sides and then taking the  $(N+m)$ -th root, we obtain

$$\sqrt[N+m]{a_{N+m}} < \sqrt[N+m]{a_N} (r + \varepsilon)^{1 - \frac{N}{N+m}} \quad \forall m \in \mathbb{N}^*$$

By Theorem 3.2.2, we can take the limit superior on both sides while preserving the (partial) order.

We have

$$\limsup_{m \rightarrow \infty} \sqrt[N+m]{a_{N+m}} \leq \limsup_{m \rightarrow \infty} \sqrt[N+m]{a_N} (r + \varepsilon)^{1-\frac{N}{N+m}} = r + \varepsilon$$

Hence,

$$\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \leq r + \varepsilon$$

Since the above inequality holds for any  $\varepsilon > 0$ , we finally have

$$\limsup_{n \rightarrow \infty} \sqrt[n]{a_n} \leq r = \limsup_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$$

■

We will see in the following two examples that neither the ratio test nor the root test will be helpful.

**Example 3.3** Let

$$\sum a_n = \sum_{n=1}^{\infty} \frac{1}{n}$$

We know that this series diverges.

We first try the ratio test. Since  $|a_{n+1}/a_n| = n/(n+1)$ , we have

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$$

Because the limit inferior and limit superior of the ratio are both 1, the ratio test is inconclusive.

As for the root test, we find

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n}} = 1$$

Hence, the root test also fails.

**Example 3.4** Let

$$\sum a_n = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

This series converges.

For the ratio test, since  $|a_{n+1}/a_n| = n^2/(n+1)^2$ , we have

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$$

The ratio test fails.

Applying the root test, we find

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n^2}} = \lim_{n \rightarrow \infty} \left( \frac{1}{\sqrt[n]{n}} \times \frac{1}{\sqrt[n]{n}} \right) = 1$$

Hence, the root test also fails.

## 3.9 Dirichlet's Test and Abel's Test

The tests in this section are helpful to determine the convergence of series that may not converge absolutely. And the series we will study throughout this section are of the form

$$\sum_{n=1}^{\infty} a_n b_n$$

The frequently used tests are Dirichlet's test as well as the Abel's test. Both of them are based on (3.16) given in the following theorem, which is known as the formula of **summation by parts**.

### Theorem 3.9.1 (Summation by Parts)

Let  $\{a_n\}$  and  $\{b_n\}$  be two complex sequences. Define

$$A_n = a_1 + \cdots + a_n, \quad n \in \mathbb{N}^*$$

Then we have the identity:

$$\sum_{k=1}^n a_k b_k = A_n b_{n+1} - \sum_{k=1}^n A_k (b_{k+1} - b_k) \quad (3.16)$$



**Note** We can use (3.16) to test convergence of series in the form of  $\sum a_n b_n$ . That is,  $\sum a_n b_n$  converges if both the sequence  $\{A_n b_{n+1}\}$  and the series  $\sum A_n (b_{n+1} - b_n)$  converge.



**Proof** Define  $A_0 = 0$  so that we can write

$$a_k = A_k - A_{k-1} \quad \forall k \in \mathbb{N}^*$$

It then follows that

$$\begin{aligned} \sum_{k=1}^n a_k b_k &= \sum_{k=1}^n (A_k - A_{k-1}) b_k \\ &= \sum_{k=1}^n A_k b_k - \sum_{k=0}^{n-1} A_k b_{k+1} \\ &= \sum_{k=1}^n A_k b_k - A_0 b_1 - \sum_{k=1}^n A_k b_{k+1} + A_n b_{n+1} \\ &= A_n b_{n+1} - \sum_{k=1}^n A_k (b_{k+1} - b_k) \end{aligned}$$

The last equality is exactly (3.16). ■

The **Dirichlet's test** is applicable when all partial sums of  $\sum a_n$  is bounded, and the sequence  $\{b_n\}$  decreases to zero.

### Theorem 3.9.2 (Dirichlet's Test)

Let  $\sum a_n$  be a series of complex terms and  $\{b_n\}$  a real-valued sequence. If

1. the sequence of partial sums of  $\sum a_n$ ,  $\{A_n\}$  is bounded,
2.  $\{b_n\}$  is decreasing, and
3.  $\{b_n\}$  converges to zero,

then  $\sum a_n b_n$  converges.



**Proof** Since  $\{A_n\}$  is bounded and  $\{b_n\}$  converges to zero, the sequence  $\{A_n b_{n+1}\}$  converges to zero.

Next, we are going to show the convergence of  $\sum A_n(b_{n+1} - b_n)$  using the comparison test. There exists  $M > 0$  such that  $|A_n| \leq M \forall n \in \mathbb{N}^*$ . We have

$$|A_n(b_{n+1} - b_n)| \leq M |b_{n+1} - b_n| = M(b_n - b_{n+1})$$

Because

$$\sum_{k=1}^n (b_k - b_{k+1}) = b_1 - b_{n+1}$$

and  $\lim_{n \rightarrow \infty} b_n = 0$ , the series  $\sum(b_n - b_{n+1})$  converges to  $b_1$ . Then by the comparison test (Theorem 3.7.1),  $\sum A_n(b_{n+1} - b_n)$  converges absolutely.

Therefore,  $\sum a_n b_n$  converges by Theorem 3.9.1. ■

## 3.10 Subseries

Think about summing up countably many (real or complex) numbers:

$$a_1, a_2, \dots, a_n, \dots$$

If you do not have to sum them up in order one by one, how many ways can you think of? For example, one way is to first sum up all the terms with odd indices:

$$s_1 = a_1 + a_3 + \dots + a_{2k-1} + \dots$$

and then all the terms with even indices:

$$s_2 = a_2 + a_4 + \dots + a_{2k} + \dots$$

Finally, the total sum is  $s_1 + s_2$ . As a matter of fact, there are more ways of summing these numbers than one can imagine at first thought, some of which may be difficult to describe. The need for a precise description of summing countably many numbers in the way we want gives rise to the concept of **subseries**.

Suppose  $s$  is the sum we will obtain by summing up the numbers in the way we like. A natural question to ask is whether  $\sum a_n$  converges to  $s$ . The answer is yes by Theorem 3.10.2, which we will soon show provided that  $\sum a_n$  converges absolutely.

### Definition 3.10.1

Let  $f : \mathbb{N}^* \rightarrow \mathbb{N}^*$  be an injective/one-to-one function whose domain is  $\mathbb{N}^*$  and whose range  $f(\mathbb{N}^*)$  is a subset of  $\mathbb{N}^*$ . Let  $\sum a_n$  and  $\sum b_n$  be two series such that

$$b_n = a_{f(n)}, \quad n \in \mathbb{N}^*$$

Then we say  $\sum b_n$  is a **subseries** of  $\sum a_n$ .



To understand the definition, we note that a subseries of  $\sum a_n$  can be constructed in the following way. Given a sequence

$$a_1, a_2, \dots, a_n, \dots$$

We first extract a subsequence  $\{a_{k_n}\}$  out of it:

$$a_{k_1}, a_{k_2}, \dots, a_{k_n}, \dots$$

where  $k_n < k_{n+1} \forall n \in \mathbb{N}^*$ . And then we rearrange this subsequence into

$$a_{g(k_1)}, a_{g(k_2)}, \dots, a_{g(k_n)}, \dots$$

where  $g$  is a bijection from  $\{k_n \mid n \in \mathbb{N}^*\}$  to itself. If we define  $f(n) = g(k_n)$ , then  $\sum_{a_{f(n)}}$  is precisely a subseries of  $\sum a_n$ .

**Example 3.5** Consider the series  $\sum a_n = \sum_{n=1}^{\infty} i^n$ . We write down its first several terms:

$$\sum_{n=1}^{\infty} i^n = i + (-1) + (-i) + 1 + i + (-1) + (-i) + 1 + \dots$$

Let

$$f(n) = 2n - 1 + 2 \times (-1)^{n-1}, \quad n \in \mathbb{N}^*$$

Note that  $f$  is injective. The following expression of  $f$  is more intuitive:

$$f(n) = \begin{cases} 2n + 1, & n \text{ is odd} \\ 2n - 3, & n \text{ is even} \end{cases}$$

The first several values of  $f$  are

$$f(1) = 3, \quad f(2) = 1, \quad f(3) = 7, \quad f(4) = 5, \dots$$

Hence, the subseries  $\sum a_{f(n)}$  is

$$\sum a_{f(n)} = (-i) + i + (-i) + i + \dots$$

### Theorem 3.10.1

If  $\sum a_n$  converges absolutely, then every subseries  $\sum b_n$  also converges absolutely. And we have

$$\left| \sum_{n=1}^{\infty} b_n \right| \leq \sum_{n=1}^{\infty} |b_n| \leq \sum_{n=1}^{\infty} |a_n| \tag{3.17}$$

**Proof** Suppose  $b_k = a_{f(k)}$   $k \in \mathbb{N}^*$ . Let integer  $n$  be fixed, and let  $N = \max\{f(1), \dots, f(n)\}$ . Then we have

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

Since

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

holds for every  $n \in \mathbb{N}^*$ , we conclude that  $\sum b_n$  converges absolutely. Then by sending  $n \rightarrow \infty$ , we see that (3.17) indeed holds. ■

To sum up

$$a_1, a_2, \dots, a_n, \dots$$

we first *partition* them into several (possibly countably infinitely many) subseries

$$\sum_{n=1}^{\infty} a_{f_1(n)}, \sum_{n=1}^{\infty} a_{f_2(n)}, \dots, \sum_{n=1}^{\infty} a_{f_k(n)}, \dots$$

After that, we sum up all the subseries. Suppose the final sum is  $s$ , the next theorem says if  $\sum a_n$  converges absolutely, then it must also converge to  $s$ . In other words, we can sum up an absolutely convergent series in whatever way we want without changing its sum!

### Theorem 3.10.2

Let  $\{f_1, f_2, \dots\}$  be a countable collection of functions defined on  $\mathbb{N}^*$ , each having the following properties:

1.  $f_k$  is injective,
2. the range  $f_k(\mathbb{N}^*)$  is a subset of  $\mathbb{N}^*$ , and
3. the collection of all function ranges form a partition of  $\mathbb{N}^*$ , that is,  $\biguplus_{k=1}^{\infty} f_k(\mathbb{N}^*) = \mathbb{N}^*$ .

Suppose that  $\sum a_n$  is an absolutely convergent series. Let  $b_k(n)$  be defined by

$$b_k(n) = a_{f_k(n)}, \quad n \in \mathbb{N}^*, \quad k \in \mathbb{N}^*$$

Then

1. each  $\sum_n b_k(n)$  is an absolutely convergent subseries of  $\sum a_n$ , and
2. if we let  $s_k = \sum_{n=1}^{\infty} b_k(n)$ , then  $\sum s_k$  converges absolutely to the same sum as  $\sum a_n$ , that is,  $\sum_{k=1}^{\infty} s_k = \sum_{n=1}^{\infty} a_n$ .



**Note** Pay attention to how the properties of  $f_k$  being injective and  $\biguplus_{k=1}^{\infty} f_k(\mathbb{N}^*) = \mathbb{N}^*$  are used in the proof below.

**Proof** The first statement that  $\sum_n b_k(n)$  is an absolutely convergent subseries of  $\sum a_n$  immediately follows from Theorem 3.10.1. We now prove 2.

We first show that  $\sum s_k$  converges absolutely. The  $p$ -th partial sum of  $\sum |s_k|$  satisfies

$$\begin{aligned} \sum_{k=1}^p |s_k| &= |s_1| + \dots + |s_p| \\ &= \left| \sum_{n=1}^{\infty} b_1(n) \right| + \dots + \left| \sum_{n=1}^{\infty} b_p(n) \right| \\ &\leq \sum_{n=1}^{\infty} |b_1(n)| + \dots + \sum_{n=1}^{\infty} |b_p(n)| \\ &= \sum_{n=1}^{\infty} (|b_1(n)| + \dots + |b_p(n)|) \end{aligned}$$

where the last equality follows from Theorem 3.3.1 since every series  $\sum_k |b_k(n)|$  converges. For any given  $q \in \mathbb{N}^*$ , we always have

$$\sum_{n=1}^q (|b_1(n)| + \dots + |b_p(n)|) \leq \sum_{n=1}^{\infty} |a_n|$$

because  $\sum_{n=1}^q (|b_1(n)| + \dots + |b_p(n)|)$  is just a sum of finitely many terms of the series  $\sum |a_n|$ . Letting

$q \rightarrow \infty$ , we obtain

$$\sum_{k=1}^p |s_k| \leq \sum_{n=1}^{\infty} (|b_1(n)| + \cdots + |b_p(n)|) \leq \sum_{n=1}^{\infty} |a_n|$$

Hence, all partial sums of  $\sum |s_k|$  are bounded above by  $\sum_{n=1}^{\infty} |a_n|$ . Therefore, the series  $\sum |s_k|$  converges, i.e.,  $\sum s_k$  converges absolutely.

Suppose  $\sum a_n$  converges to sum  $a$ . Given  $\varepsilon > 0$ , there exists  $N \in \mathbb{N}^*$  such that

$$\left| \sum_{n=1}^q a_n - s \right| < \varepsilon/2 \quad \forall q \geq N \quad (3.18)$$

For any integer  $j \in \mathbb{N}^*$ , there exists a *unique* integer  $k \in \mathbb{N}^*$  such that

$$j \in f_k(\mathbb{N}^*)$$

since  $\bigcup_{k=1}^{\infty} f_k(\mathbb{N}^*) = \mathbb{N}^*$ . (Existence of  $k$  is due to  $\bigcup_{k=1}^{\infty} f_k(\mathbb{N}^*) = \mathbb{N}^*$ , and the uniqueness follows from the fact that these function ranges are mutually disjoint.) Furthermore, there exists a *unique* integer  $n \in \mathbb{N}^*$  such that

$$f_k(n) = j$$

since  $f_k$  is injective. Therefore, any integer  $j$  can uniquely determine a pair  $(k, n)$  and hence we can define a function  $g : \mathbb{N}^* \rightarrow \mathbb{N}^* \times \mathbb{N}^*$  by

$$g(j) = (k, n)$$

satisfying

$$f_k(n) = j$$

Recall how we chose the integer  $N$ . Write

$$g(1) = (k_1, n_1), \dots, g(N) = (k_N, n_N)$$

Let

$$K = \max\{k_1, \dots, k_N\} \quad \text{and} \quad M = \max\{n_1, \dots, n_N\}$$

We claim that

$$\left| \sum_{k=1}^p \sum_{n=1}^q b_k(n) - s \right| < \varepsilon/2 \quad \forall p \geq K \quad \forall q \geq M \quad (3.19)$$

The proof of (3.19) is left as an exercise. See Exercise 3.2. Letting  $q \rightarrow \infty$  in (3.19) yields

$$\left| \sum_{k=1}^p \sum_{n=1}^{\infty} b_k(n) - s \right| \leq \varepsilon/2 < \varepsilon \quad \forall p \geq K$$

Recall each subseries  $\sum_n b_k(n)$  converges absolutely, and  $s_k = \sum_{n=1}^{\infty} b_k(n)$ . Hence, we have

$$\left| \sum_{k=1}^p s_k - s \right| < \varepsilon \quad \forall p \geq K$$

Note that the choice of  $K$  only depends on  $\varepsilon$  (through  $N$ ). Therefore, the series  $\sum s_k$  indeed converges to sum  $a$ . ■

☞ **Exercise 3.2** Complete the above proof by proving the inequality (3.19).

**Solution** For each  $j \in \{1, \dots, N\}$ , we have  $g(j) = (k, n)$  where  $k$  and  $n$  satisfy

$$f_k(n) = j, \quad k \leq K \quad \text{and} \quad n \leq M$$

This implies

$$\{f_k(n) \mid 1 \leq k \leq p, 1 \leq n \leq q\} \supseteq \{1, \dots, N\}$$

provided that  $p \geq K$  and  $q \geq M$ . Inequality (3.19) then immediately follows from (3.18) since

$$\sum_{k=1}^p \sum_{n=1}^q b_k(n) = \sum_{k=1}^p \sum_{n=1}^q a_{f_k(n)}$$

is a sum of terms including the first  $N$  terms of  $\sum a_n$ .

## 3.11 Double Series

### Theorem 3.11.1

Suppose double series  $\sum a_{m,n}$  consists of non-negative terms. Then it converges if and only if the set of its partial sums is bounded above.



#### Proof

The absolute convergence of a double series will also imply convergence of itself.

### Theorem 3.11.2

A double series  $\sum a_{m,n}$  converges if it converges absolutely.



#### Proof

## 3.12 Rearrangement Theorem for Double Series

### Definition 3.12.1

Suppose  $\{a_{m,n}\}$  is a double sequence, and  $g : \mathbb{N}^* \rightarrow \mathbb{N}^* \times \mathbb{N}^*$  is a **bijective** function. Let  $b_k$  be defined by

$$b_k = a_{g(k)} = a_{m,n}, \quad k \in \mathbb{N}^*$$

Then function  $g$  is said to be an **arrangement** of  $\{a_{m,n}\}$  into the (single) sequence  $\{b_k\}$ .



### Lemma 3.12.1

Let  $\sum a_{m,n}$  be a double series, and  $g$  an arrangement of double sequence  $\{a_{m,n}\}$  into  $\{b_k\}$ . Then the series  $\sum b_k$  converges absolutely if and only if  $\sum a_{m,n}$  converges absolutely.



**Proof** We first suppose that  $\sum a_{m,n}$  converges absolutely. By Theorem 3.11.1, the set of its partial sums of the series  $\sum |a_{m,n}|$  must be bounded above by some positive number, say  $A$ , that is,

$$\sum_{m=1}^p \sum_{n=1}^q |a_{m,n}| \leq A \quad \forall p, q \in \mathbb{N}^*$$

The  $K$ -th partial sum of  $\sum |b_k|$  satisfies

$$\sum_{k=1}^K |b_k| = \sum_{k=1}^K |a_{g(k)}|$$

where the right-hand side is just a partial sum of  $\sum |a_{m,n}|$ . Hence,

$$\sum_{k=1}^K |b_k| \leq A \quad \forall K \in \mathbb{N}^*$$

Then  $\sum |b_k|$  indeed converges by Theorem 3.11.1, i.e.,  $\sum b_k$  converges absolutely.

Conversely, suppose now the series  $\sum b_k$  converges absolutely. Hence, the set of partial sums of  $\sum |b_k|$  is bounded above by some number  $B > 0$ . It follows that

$$\sum_{m=1}^p \sum_{n=1}^q |a_{m,n}| \leq \sum_{k=1}^K |b_k|$$

where

$$K = \max \{g^{-1}(m, n) \mid 1 \leq m \leq p, 1 \leq n \leq q\}$$

We then have

$$\sum_{m=1}^p \sum_{n=1}^q |a_{m,n}| \leq B \quad \forall p, q \in \mathbb{N}^*$$

Therefore,  $\sum |a_{m,n}|$  converges absolutely by Theorem 3.11.1. ■

### Theorem 3.12.1

Suppose  $\sum a_{m,n}$  is an absolutely convergent double series with sum  $s$ . Let  $g$  be an arrangement of double sequence  $\{a_{m,n}\}$  into  $\{b_k\}$ . Then

1.  $\sum b_k$  converges absolutely to  $s$ .
2. Series  $\sum_n a_{m,n}$  converges absolutely for each  $m \in \mathbb{N}^*$ , and series  $\sum_m a_{m,n}$  converges absolutely for each  $n \in \mathbb{N}^*$ .
3. Let  $r_m = \sum_{n=1}^{\infty} a_{m,n}$  and  $c_n = \sum_{m=1}^{\infty} a_{m,n}$ . Then the series  $\sum r_m$  and  $\sum c_n$  both converge absolutely to the same sum  $s$ , i.e.,

$$\sum_{m=1}^{\infty} r_m = \sum_{n=1}^{\infty} c_n = s$$

If we expand  $r_m$  and  $c_n$ , then we obtain the same equation as above in terms of iterated sums:

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n} = s$$



**Note** The symbols  $r_m$  and  $c_n$  have the meanings of the sum of  $m$ -th row and  $n$ -th column, respectively.

**Proof** By the above lemma (Lemma 3.12.1), we know that  $\sum b_k$  converges absolutely. Now, we show that it converges to  $s$ . Choose an arbitrary  $\varepsilon > 0$ . Let  $s_{p,q} = \sum_{m=1}^p \sum_{n=1}^q a_{m,n}$ . Since  $\sum a_{m,n}$  converges to  $s$ , there exists an integer  $N_1 \in \mathbb{N}^*$  such that

$$|s_{p,q} - s| < \varepsilon/2 \quad \forall p, q \geq N_1$$

On the other hand, since  $\sum |a_{m,n}|$  converges, there exists  $N_2 \in \mathbb{N}^*$  such that

$$S_{p+l,p+l} - S_{p,p} < \varepsilon/2 \quad \forall p \geq N_2 \quad \forall l \in \mathbb{N}^*$$

where  $S_{p,q} = \sum_{m=1}^p \sum_{n=1}^q |a_{m,n}|$ . Let  $N = \max\{N_1, N_2\}$ . In particular, we have

$$|s_{N,N} - s| < \varepsilon/2 \quad \text{and} \quad S_{N+l,N+l} - S_{N,N} < \varepsilon/2 \quad \forall l \in \mathbb{N}^* \quad (3.20)$$

Let

$$K = \max \{g^{-1}(m, n) \mid 1 \leq m \leq N, 1 \leq n \leq N\}$$

Combined with (3.20), we find

$$\left| \sum_{k=1}^K b_k - s \right| \leq \left| \sum_{k=1}^K b_k - s_{N,N} \right| + |s_{N,N} - s| < \varepsilon$$

 **Note** Terms  $b_1, \dots, b_K$  includes all the terms  $a_{m,n}$  with  $1 \leq m \leq N$  and  $1 \leq n \leq N$ . Hence, the difference  $\sum_{k=1}^K b_k - s_{N,N}$  is a sum of terms  $a_{m,n}$ , at least one of whose two indices is greater than  $N$ . Therefore,  $\sum b_k$  indeed converges to sum  $s$ .

Let  $m \in \mathbb{N}^*$  be fixed. Define a function  $h_m : \mathbb{N}^* \rightarrow \mathbb{N}^* \times \mathbb{N}^*$  by

$$h_m(n) = (m, n)$$

And let function  $f_m : \mathbb{N}^* \rightarrow \mathbb{N}^*$  be defined as a composition of  $h_m$  and  $g^{-1}$ , i.e.,

$$f_m = g^{-1} \circ h_m \quad (3.21)$$

Note that  $f_m$  is injective since both  $h_m$  and  $g$  are injective. For a term  $a_{m,n}$ , through arrangement  $g$ , we have

$$a_{m,n} = b_k \quad \text{where } k = g^{-1}(m, n)$$

Applying  $f_m$  to  $n$ , we find that

$$f_m(n) = g^{-1}(h_m(n)) = g^{-1}(m, n) = k$$

This implies  $\sum_n a_{m,n}$  is a subseries of  $\sum b_k$ . Because  $\sum b_k$  converges absolutely, by Theorem 3.10.2,  $\sum_n a_{m,n}$  also converges absolutely. Similarly, each  $\sum_m a_{m,n}$  also converges absolutely.

Write  $r_m = \sum_{n=1}^{\infty} a_{m,n}$ . Define  $f_m$  for each  $m \in \mathbb{N}^*$  as in (3.21). It can be shown that

1.  $f_m$  is injective,
2.  $f_m(\mathbb{N}^*) \subseteq \mathbb{N}^*$ , and
3.  $\bigcup_{m=1}^{\infty} f_m(\mathbb{N}^*) = \mathbb{N}^*$ .

(See Exercise 3.3.) Hence, all the requirements of Theorem 3.10.2 are met. Therefore,  $\sum r_m$  converges absolutely, and

$$\sum_{m=1}^{\infty} r_m = \sum_{k=1}^{\infty} b_k = s$$

Similarly, we can also show that  $\sum c_n$  converges absolutely to  $s$ . ■

 **Exercise 3.3** Complete the above proof by showing

1.  $f_m$  is injective,
2.  $f_m(\mathbb{N}^*) \subseteq \mathbb{N}^*$ , and
3.  $\bigcup_{m=1}^{\infty} f_m(\mathbb{N}^*) = \mathbb{N}^*$ .

Each function  $f_m$  is defined by (3.21).

**Solution** The first two statements are evident by (3.21). We now show

$$\bigcup_{m=1}^{\infty} f_m(\mathbb{N}^*) = \mathbb{N}^*$$

Given  $k \in \mathbb{N}^*$ , we have

$$g(k) = (m, n)$$

It then follows from (3.21) that

$$f_m(n) = g^{-1}(h_m(n)) = g^{-1}(m, n) = k$$

Hence,  $n \in f_m(\mathbb{N}^*)$ . This, along with the fact  $f_m(\mathbb{N}^*) \subseteq \mathbb{N}^*$ , implies the union of all function ranges is the set of natural numbers, i.e.,

$$\bigcup_{m=1}^{\infty} f_m(\mathbb{N}^*) = \mathbb{N}^*$$

We also need to show that every two function ranges are mutually disjoint. Assume that, on the contrary, there exist  $m_1, m_2 \in \mathbb{N}^*$  such that  $m_1 \neq m_2$  and

$$k \in f_{m_1}(\mathbb{N}^*) \cap f_{m_2}(\mathbb{N}^*)$$

for some integer  $k \in \mathbb{N}^*$ . It then follows that there exist  $n_1, n_2 \in \mathbb{N}^*$  satisfying

$$f_{m_1}(n_1) = k \quad \text{and} \quad f_{m_2}(n_2) = k$$

But then

$$f_{m_1}(n_1) = g^{-1}(m_1, n_1) = k = g^{-1}(m_2, n_2) = f_{m_2}(n_2)$$

Since  $g$  is bijective, we must have  $m_1 = m_2$ , which leads to a contradiction. Therefore, every two function ranges are mutually disjoint. This completes the proof.

## 3.13 Iterated Series

In practice, it is very common to encounter situations where interchanging the order of summations of a double series will make the computation much easier. That is, sometimes we favor the right-hand side over the left-hand side in (3.22).

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n} \tag{3.22}$$

We state a sufficient condition in the following theorem so that (3.22) is valid.

### Theorem 3.13.1

Let  $\{a_{m,n}\}$  be a double sequence with complex terms. If

- a.  $\sum_{n=1}^{\infty} a_{m,n}$  converges absolutely for each  $m$ , and
- b.  $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |a_{m,n}|$  converges,

then

1. the double series  $\sum_{m,n} a_{m,n}$  converges absolutely,
2.  $\sum_{m=1}^{\infty} a_{m,n}$  converges absolutely for each  $n$ , and

3.  $\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n}$  (regarded as a series with terms  $\sum_{m=1}^{\infty} a_{m,n}$ ) also converges absolutely with the same sum as  $\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n}$  and  $\sum_{m,n} a_{m,n}$ , that is,

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{m,n} = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{m,n} = \sum_{m,n} a_{m,n}$$



**Proof** (Proof of 1) Let  $g$  be an arrangement of  $\{a_{m,n}\}$  into  $\{b_k\}$ . For any given  $K \in \mathbb{N}^*$ , we have

$$\sum_{k=1}^K |b_k| \leq \sum_{m=1}^M \sum_{n=1}^N |a_{m,n}| \leq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |a_{m,n}|$$

where  $M = \max\{m_1, \dots, m_K\}$  and  $N = \max\{n_1, \dots, n_K\}$  where  $g(k) = (m_k, n_k) \forall k = 1, \dots, K$ . This implies that the sequence of partial sums of  $\sum |b_k|$  is bounded above, which further implies  $\sum b_k$  converges absolutely. And then by Lemma 3.12.1, the double series  $\sum_{m,n} a_{m,n}$  also converges absolutely.

(Proof of 2 and 3) The last two statements follow directly from Theorem 3.12.1. ■

## 3.14 Multiplication of Series

### Definition 3.14.1

Given two complex-valued series  $\sum_{n=0}^{\infty} a_n$  and  $\sum_{n=0}^{\infty} b_n$ . Let

$$c_n = \sum_{k=0}^n a_{n-k} b_k, \quad n \in \mathbb{N}$$

Then the series  $\sum_{n=0}^{\infty} c_n$  is called the **Cauchy product** of  $\sum a_n$  and  $\sum b_n$ .



### Theorem 3.14.1 (Mertens)

Let  $\sum a_n$  and  $\sum b_n$  be two series that converge to sums  $A$  and  $B$ , respectively. Suppose further that  $\sum a_n$  converges absolutely. Then their Cauchy product  $\sum c_n$  also converges with sum  $AB$ . In this case, we can write

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \sum_{k=0}^n a_{n-k} b_k = \sum_{n=0}^{\infty} a_n \sum_{k=0}^{\infty} b_k$$



**Proof** Let  $\varepsilon > 0$  be chosen arbitrarily. Firstly, because  $\sum b_n$  converges, by Theorem 3.3.4, there exists  $M > 0$  such that

$$\left| \sum_{k=m}^n b_k \right| \leq M \quad \forall m \in \mathbb{N}^* \quad \forall n \geq M \quad (3.23)$$

Suppose  $\sum |a_n|$  converges to  $S$ . Moreover, by Cauchy's criterion for series, there exist  $N_1, N_2 \in \mathbb{N}^*$  such that

$$\sum_{k=m}^{m+j} |a_k| < \frac{\varepsilon/2}{|M|} \quad \forall m \geq N_1 \quad \forall j \in \mathbb{N} \quad (3.24)$$

and

$$\left| \sum_{k=m}^{m+j} b_k \right| < \frac{\varepsilon/2}{S+1} \quad \forall m \geq N_2 \quad \forall j \in \mathbb{N} \quad (3.25)$$

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

Let  $A_n, B_n$  and  $C_n$  denote

$$A_n = \sum_{k=0}^n a_k, \quad B_n = \sum_{k=0}^n b_k, \quad \text{and} \quad C_n = \sum_{k=0}^n c_k$$

Let  $n \geq 2N$ . After a few steps of algebraic computation, we obtain

$$\begin{aligned} A_n B_n - C_n = & a_1 b_n \\ & + a_2(b_{n-1} + b_n) \\ & + a_3(b_{n-2} + b_{n-1} + b_n) \\ & + \dots \\ & + a_n(b_1 + b_2 + \dots + b_n) \end{aligned}$$

It then follows that

$$\begin{aligned} |A_n B_n - C_n| \leq & |a_1| |b_n| \\ & + \dots \\ & + |a_{n-N}| |b_{N+1} + \dots + b_n| \\ & + |a_{n-N+1}| |b_N + \dots + b_n| \\ & + \dots \\ & + |a_n| |b_1 + \dots + b_n| \end{aligned} \quad (3.26)$$

Applying (3.25) to the first  $n - N$  terms on right-hand side of (3.26), we find

$$\begin{aligned} |a_1| |b_n| + \dots + |a_{n-N}| |b_{N+1} + \dots + b_n| & < (|a_1| + \dots + |a_{n-N}|) \frac{\varepsilon/2}{S+1} \\ & \leq S \times \frac{\varepsilon/2}{S+1} \\ & < \varepsilon/2 \end{aligned} \quad (3.27)$$

Meanwhile, applying (3.23) to the last  $N$  terms on the right-hand side of (3.26) yields

$$\begin{aligned} |a_{n-N+1}| |b_N + \dots + b_n| + \dots + |a_n| |b_1 + \dots + b_n| \\ & < |a_{n-N+1}| M + \dots + |a_n| M = (|a_{n-N+1}| + \dots + |a_n|) M \end{aligned} \quad (3.28)$$

By using (3.24), the right-hand side of (3.28) can be further enlarged to

$$(|a_{n-N+1}| + \dots + |a_n|) M < \frac{\varepsilon/2}{M} \cdot M = \varepsilon/2$$

Hence,

$$|a_{n-N+1}| |b_N + \dots + b_n| + \dots + |a_n| |b_1 + \dots + b_n| < \varepsilon/2 \quad (3.29)$$

Combined with (3.26), (3.27) and (3.29), we find

$$|A_n B_n - C_n| < \varepsilon \quad \forall n \geq 2N$$

Therefore, the limit of  $\{A_n B_n - C_n\}$  exists, and it equals zero, that is,

$$\lim_{n \rightarrow \infty} A_n B_n - C_n = 0$$

But since  $\lim_{n \rightarrow \infty} A_n = A$  and  $\lim_{n \rightarrow \infty} B_n = B$ , by the algebraic properties of limits, we conclude that the limit of  $\{C_n\}$  also exists, and it equals  $AB$ . In other words,  $\sum c_n$  converges to  $AB$ .  $\blacksquare$

# Chapter 4 Limits and Continuity

## 4.1 Limit of a Function

### Proposition 4.1.1

Let  $f$  and  $g$  be complex-valued functions defined on a subset  $A$  of a metric space  $(X, d)$ , and  $p$  a limit point of  $A$ . If  $\lim_{x \rightarrow p} f(x) = 0$  and  $g$  is bounded, then

$$\lim_{x \rightarrow p} f(x)g(x) = 0$$



**Proof**



## 4.2 Computing Limits

The basic rules of computing limits are listed in the next theorem.

### Theorem 4.2.1

Let  $f$  and  $g$  be complex-valued functions defined on a subset  $S$  of a metric space  $(X, d)$ . Let  $p$  be an accumulation point of  $S$ , and suppose that the following limits exist:

$$\lim_{x \rightarrow p} f(x) = a \quad \text{and} \quad \lim_{x \rightarrow p} g(x) = b$$

We have

1.  $\lim_{x \rightarrow p} [f(x) \pm g(x)] = a \pm b$ ,
2.  $\lim_{x \rightarrow p} f(x)g(x) = ab$ ,
3.  $\lim_{x \rightarrow p} f(x)/g(x) = a/b$  provided that  $b \neq 0$ .



## 4.3 Continuous Functions

The next theorem states that the composite function of continuous functions is also continuous.

### Theorem 4.3.1

Let  $(X, d_X)$ ,  $(Y, d_Y)$  and  $(Z, d_Z)$  be metric spaces. Let  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  be functions, and let  $h = g \circ f$  be the composite function defined on  $X$ . If  $f$  is continuous at  $p \in X$ , and  $g$  is continuous at  $f(p)$ , then  $h$  is continuous at  $p$ .



**Proof** Given  $\varepsilon > 0$ , since  $g$  is continuous at  $f(p)$ , there exists some  $\delta' > 0$  such that

$$d_Y(y, f(p)) < \delta' \implies d_Z(g(y), g(f(p))) < \varepsilon \quad (4.1)$$

And since  $f$  is continuous at  $p$ , there exists some  $\delta > 0$  such that

$$d_X(x, p) < \delta \implies d_Y(f(x), f(p)) < \delta' \quad (4.2)$$

Now, choosing  $x \in X$  satisfying  $d_X(x, p) < \delta$ , and then applying (4.2) followed by (4.1), we have

$$d_Z(g(f(x)), g(f(p))) < \varepsilon$$

This is precisely  $d_Z(h(x), h(p)) < \varepsilon$ . Hence,  $h$  is continuous at  $p$ . ■

### Theorem 4.3.2

*Let  $f : X \rightarrow \mathbb{R}$  be a real-valued continuous function from a metric space  $X$  to  $\mathbb{R}$ . Let  $K \subseteq X$  be a compact subset. Then there exist points  $p, q \in K$  such that*

$$f(p) = \sup f(K) \quad \text{and} \quad f(q) = \inf f(K)$$



## 4.4 Continuity of Composite Functions

### Theorem 4.4.1

*Let  $(X, d_X)$ ,  $(Y, d_Y)$  and  $(Z, d_Z)$  be metric spaces. Let  $f : X \rightarrow Y$  and  $g : f(X) \rightarrow Z$  be functions. If  $f$  is continuous at point  $p \in X$  and  $g$  is continuous at  $f(p)$ , then the composite function  $f \circ g$  is continuous at  $p$ .* ♥

## 4.5 Connectedness

### Theorem 4.5.1 (Intermediate Value Theorem)

*Let  $f : [a, b] \rightarrow \mathbb{R}$  be a real-valued continuous function. Suppose that  $f(a) \neq f(b)$ , and  $y$  is some number in between  $f(a)$  and  $f(b)$ . Then there exists an interior point  $c \in (a, b)$  such that  $f(c) = y$ .* ♥

### Proof

## 4.6 Monotonic Functions

### Theorem 4.6.1

*If  $f$  is increasing on  $[a, b]$ , then  $f(c+)$  and  $f(c-)$  both exists for every  $c \in (a, b)$ , and we have*

$$f(c-) \leq f(c) \leq f(c+)$$

*As the endpoints, we have  $f(a+)$  and  $f(b-)$  both exists and*

$$f(a) \leq f(a+) \quad \text{and} \quad f(b-) \leq f(b) \tag{4.3}$$

*Analogous results exist for  $f$  being a decreasing function.* ♥

**Proof** We only prove this theorem for increasing functions. Let  $c \in (a, b)$  be an interior point.

(Existence of  $f(c-)$ ) Define a set

$$A = \{f(x) \mid a \leq x < c\}$$

Note that  $A$  is bounded above by  $f(c)$  since  $f$  is increasing. Then  $A$  has the least upper bound, say  $\alpha = \sup A$ . We want to show that  $f(c-)$  exists and equals  $\alpha$ . Given  $\varepsilon > 0$ , by the property of  $\alpha$ , there exists  $x_\varepsilon = (c - \delta) \in [a, c]$  such that ( $\delta$  depends on  $\varepsilon$ )

$$f(c - \delta) > \alpha - \varepsilon$$

Because  $f$  is increasing, we have

$$\alpha \geq f(x) \geq f(c - \delta) > \alpha - \varepsilon \quad \forall x \in (c - \delta, c)$$

This implies that

$$|f(x) - \alpha| < \varepsilon \quad \forall x \in (c - \delta, c)$$

Hence,  $\alpha$  is the left-hand limit of  $f$  at  $x = c$ , i.e.,  $f(c-) = \alpha$ . Recall  $A$  is bounded above by  $f(c)$ . We have  $\alpha \leq f(c)$ , i.e.,

$$f(c-) \leq f(c)$$

(Existence of  $f(c+)$ ) Define

$$B = \{f(x) \mid c < x \leq b\}$$

Let  $\beta$  be the greatest lower bound of  $B$ . A similar argument shows that  $f(c+)$ , and  $f(c+) = \beta$  with

$$f(c) \leq f(c+)$$

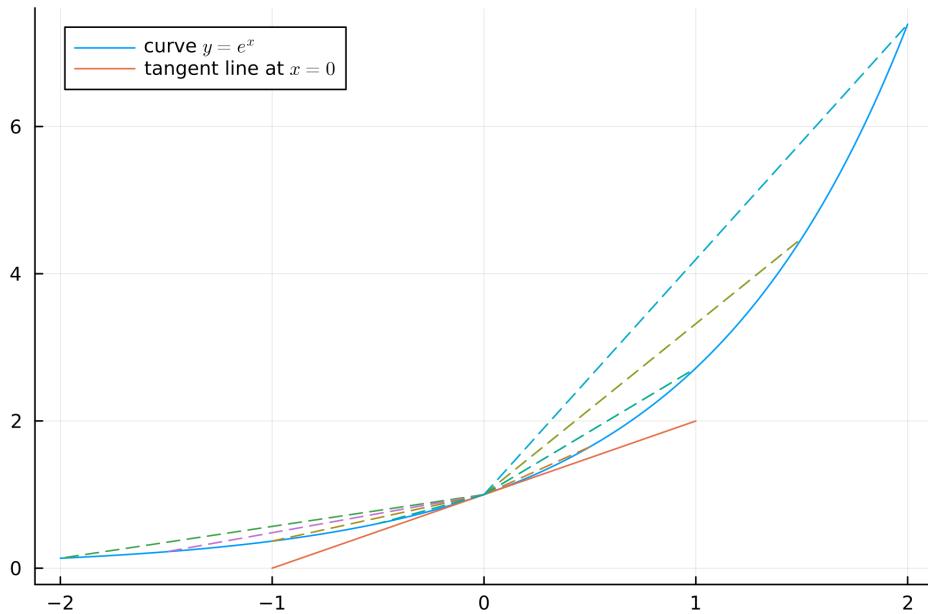
(Existence of  $f(a+)$  and  $f(b-)$ ) The existence of  $f(a+)$  and  $f(b-)$  along with (4.3) can be proved by the exact argument as before. ■

# Chapter 5 Differentiation

Let our exploration begins with real-valued functions defined on real numbers.

## 5.1 Definition of Derivative

Observe the slopes of secant lines (displayed as dashed lines) of the curve shown in Figure 5.1



**Figure 5.1:** Secant lines and a tangent line of curve  $y = e^x$ .

All secant lines have a common point of intersection,  $(0, 1)$ . And as the other intersection point gets closer and closer to  $(0, 1)$ , the slope of the secant line tends to converge to some number. This number, the limiting slope, is precisely the **derivative** of this curve/function at  $x = 0$ , which we shall soon define. We can draw a straight line passing through the point  $(0, 1)$  with the limiting slope. This line is then called the tangent of the curve at point  $(0, 1)$ , as illustrated in Figure 5.1.

### Definition 5.1.1

Let  $f$  be defined on an open interval  $(a, b)$ , and let  $c \in (a, b)$  an interior point. Then  $f$  is said to be **differentiable** at  $c$  if the limit

$$\lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}$$

exists. This limit, denoted by  $f'(c)$ , is called the **derivative** of  $f$  at point  $c$ .



If  $f'(c)$  exists  $\forall c \in I$  for some interval  $I$ , then we can define a function  $f'$  on  $I$ , which is called the derivative of  $f$ . Here the word *derivative* means a function instead of just a number.

The definition of derivatives we present here involves the limit obtained by letting one point approach the other one,  $x \rightarrow c$ . For computational convenience, we can treat the derivative of  $f$  at  $c$  as

the limit of the fraction  $[f(c + h) - f(c)]/h$  as  $h \rightarrow 0$ . That is, the limit is reached when the distance between points  $x$  and  $c$ ,  $h = |x - c|$ , is small. The following proposition states this idea formally.

### Proposition 5.1.1

*Let  $f$  be defined on  $(a, b)$ . Then  $f$  is differentiable at  $c \in (a, b)$  if and only if*

$$\lim_{h \rightarrow 0} \frac{f(c + h) - f(c)}{h} \quad (5.1)$$

*exists. In that case,  $f'(c) = \lim_{h \rightarrow 0} [f(c + h) - f(c)]/h$ .*



**Proof** Suppose  $f$  is differentiable at  $c$ . For an arbitrary  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$|x - c| < \delta \implies \left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < \varepsilon$$

Let number  $h$  be such that  $|h| < \delta$ . Then since  $|(c + h) - c| = |h| < \delta$ , we have

$$\left| \frac{f(c + h) - f(h)}{h} - f'(c) \right| = \left| \frac{f(c + h) - f(h)}{(c + h) - c} - f'(c) \right| < \varepsilon$$

This implies the limit (5.1) exists, and it equals  $f'(c)$ .

Conversely, suppose the limit (5.1) exists, say it equals  $L$ . Then for  $\varepsilon > 0$  there exists  $\delta > 0$  such that

$$|h| < \delta \implies \left| \frac{f(c + h) - f(c)}{h} - L \right| < \varepsilon$$

Choose  $x$  such that  $|x - c| < \delta$ , we have

$$\left| \frac{f(x) - f(c)}{x - c} - L \right| = \left| \frac{f(c + (x - c)) - f(c)}{x - c} - L \right| < \varepsilon$$

Therefore, the limit  $[f(x) - f(c)]/(x - c)$  exists, which is equal to  $L$ . By the definition of derivatives,  $f$  is differentiable at  $c$ , and  $f'(c) = L$ . ■

☞ **Exercise 5.1** Calculate the derivative of  $e^x$  at  $x = 0$ .

[Hint: You may use the fact  $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots$ ]

☞ **Exercise 5.2** Calculate the derivative of

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

at  $x = 0$ .

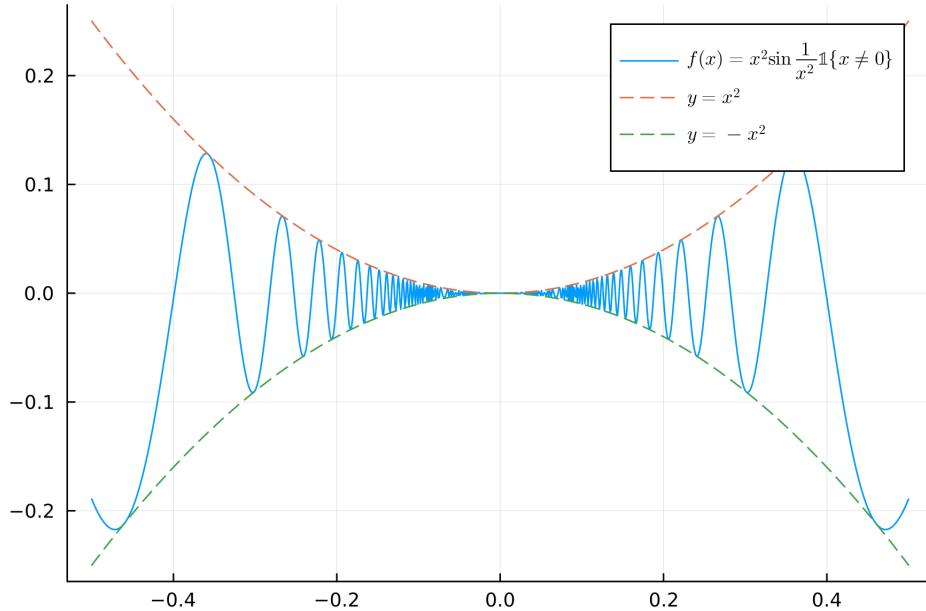
**Solution** We have

$$\frac{f(0 + h) - f(0)}{h} = h \sin \frac{1}{h^2}$$

Note the limit of the above expression is zero as  $h \rightarrow 0$  by Proposition 4.1.1 since  $\lim_{h \rightarrow 0} h = 0$  and  $\sin(1/h^2)$  is bounded by 1. Hence,  $f'(0) = 0$  by Proposition 5.1.1.

The graph of  $f(x)$  is shown in Figure 5.2. As we can see,  $f(x)$  is clamped between curves  $y = x^2$  and  $y = -x^2$  with the same tangents  $y = 0$  at  $x = 0$ . And we have shown  $f'(0) = 0$ , which means that the tangent of  $f(x)$  is also  $y = 0$  at  $x = 0$ . What conjecture can you make?

Moreover, we observe that  $f(x)$  crosses its tangent line at  $x = 0$  infinitely many times. This example shows that the tangent line does not have to touch the curve only at one point.



**Figure 5.2:** Graph of  $f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}_{\{x \neq 0\}}$ .

### 5.1.1 One-Sided Derivatives and Infinite Derivatives

So far, the point at which the derivative is defined has to be an *interior* point. Sometimes, we are required to consider the derivatives at the endpoints of the interval. For example, as we shall see in more detail in Chapter 7, we need to take the derivative of function  $F(x) = \int_a^x f(t)dt$  at  $x = a$ . Hence, we are motivated to define the **one-sided derivatives**.

If we consider the derivative of  $f$ ,  $f'$  as a function, then  $f'(x)$  may tend to infinity as  $x$  approaches an endpoint. In addition, sometimes we need to interpret the meaning of vertical tangent lines. This leads to the definition of **infinite derivatives**.

#### Definition 5.1.2

Let  $f$  be defined on a closed interval  $I$ . Suppose  $f$  is continuous at point  $c \in I$ . Then  $f$  is said to have a **right-hand derivative** at  $c$  if the right-hand limit

$$\lim_{x \rightarrow c^+} \frac{f(x) - f(c)}{x - c}$$

exists as a finite number, or the limit is  $\infty$  or  $-\infty$ . This right-hand derivative shall be denoted by  $f'_+(c)$ . The **left-hand derivative**, denoted by  $f'_-(c)$ , is defined similarly. In addition, if  $c$  is an interior point, and  $f'_+(c) = f'_-(c) = \infty$ , then we write  $f'(c) = \infty$ .  $f'(c) = -\infty$  is similarly defined.



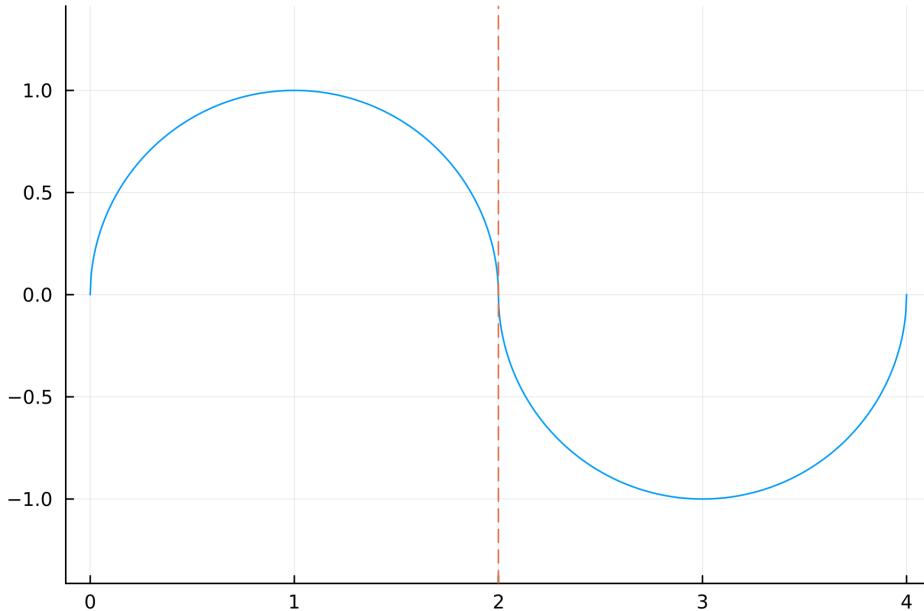
**Remark** Note that though we write  $f'(c) = \infty$  or  $f'(c) = -\infty$ , we do not say  $f$  is differentiable there.

It is not very intuitive to imagine, at first thought, a continuous function having infinite derivatives. But there are quite a lot of such examples.

**Example 5.1** The first example is constructed by cutting a circle in the middle. If we cut a circle in half and stick the lower half to the right of the upper one, then we have a function with a vertical

tangent line in the middle. The explicit expression of this function is

$$f(x) = \begin{cases} \sqrt{1 - (x - 1)^2} & 0 \leq x \leq 2 \\ -\sqrt{1 - (x - 3)^2} & 2 < x \leq 4 \end{cases}$$



**Figure 5.3:** Graph of  $f(x) = \sqrt{1 - (x - 1)^2} \mathbb{1}\{0 \leq x \leq 2\} - \sqrt{1 - (x - 3)^2} \mathbb{1}\{2 < x \leq 4\}$ .

The function is continuous and one can show that  $f'(2) = -\infty$ .

**Example 5.2** The following next example includes both positive and negative intuitive derivatives, and a point at which the right and left-hand side derivatives are  $\infty$  and  $-\infty$ , respectively. Consider the function

$$f(x) = \sqrt[3]{x^2(x - 1)(x - 2)}$$

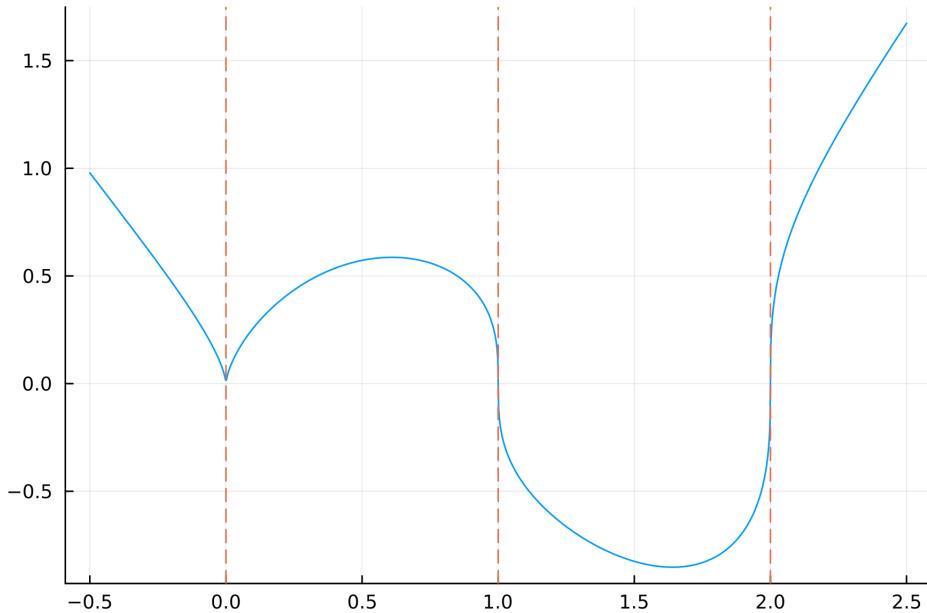
Its graph is shown in Figure 5.4.

It is an exercise to show  $f'(1) = -\infty$  and  $f'(2) = \infty$ . We now consider the one-sided derivatives at  $x = 0$ . We have

$$\frac{f(x) - f(0)}{x - 0} = \sqrt[3]{\frac{(x - 1)(x - 2)}{x}}$$

Letting  $x \rightarrow 0^+$  leads to  $f'_+(0) = \infty$ , while  $x \rightarrow 0^-$  yields  $f'_-(0) = -\infty$ . Hence, we say the derivative of  $f$  does not exist at  $x = 0$ .

Sometimes, when in some proofs, we need to consider the *corner cases* where function  $f$  (defined on  $[a, b]$ ) is required to be *differentiable* at the endpoints. Well, we cannot really say  $f$  is differentiable at  $a$  or  $b$  since the point where a function is differentiable must be an interior point from our definition. However, we can make  $a$  and  $b$  be interior points by extending the definition of  $f$  to a larger interval, say  $[a - \delta, b + \delta]$  ( $\delta > 0$ ).



**Figure 5.4:** Graph of  $f(x) = \sqrt[3]{x^2(x-1)(x-2)}$ .

### Lemma 5.1.1

Let function  $f$  be defined on  $[a, b]$ . If the one-sided derivatives  $f'_+(a)$  and  $f'_-(b)$  both exist as finite numbers at endpoints. Then we can extend  $f$ , to  $\bar{f}$  on a larger closed interval  $[a - \delta, b + \delta]$  ( $\delta > 0$ ) such that  $\bar{f}$  is differentiable at  $a$  and  $b$  with

$$\bar{f}'(a) = f'_+(a) \quad \text{and} \quad \bar{f}'(b) = f'_-(b)$$

To be specific,  $\bar{f}$  can be defined by

$$\bar{f}(x) = \begin{cases} x, & a \leq x \leq b \\ f'_+(a)(x-a) + f(a), & a - \delta \leq x < a \\ f'_-(b)(x-b) + f(b), & b < x \leq b + \delta \end{cases}$$



The geometric interpretation of extension  $\bar{f}$  is simply that we draw two straight lines at the two ends of  $f$  such that their slopes are exactly the one-sided derivatives of  $f$  at  $a$  and  $b$ , respectively. See Figure 5.5.

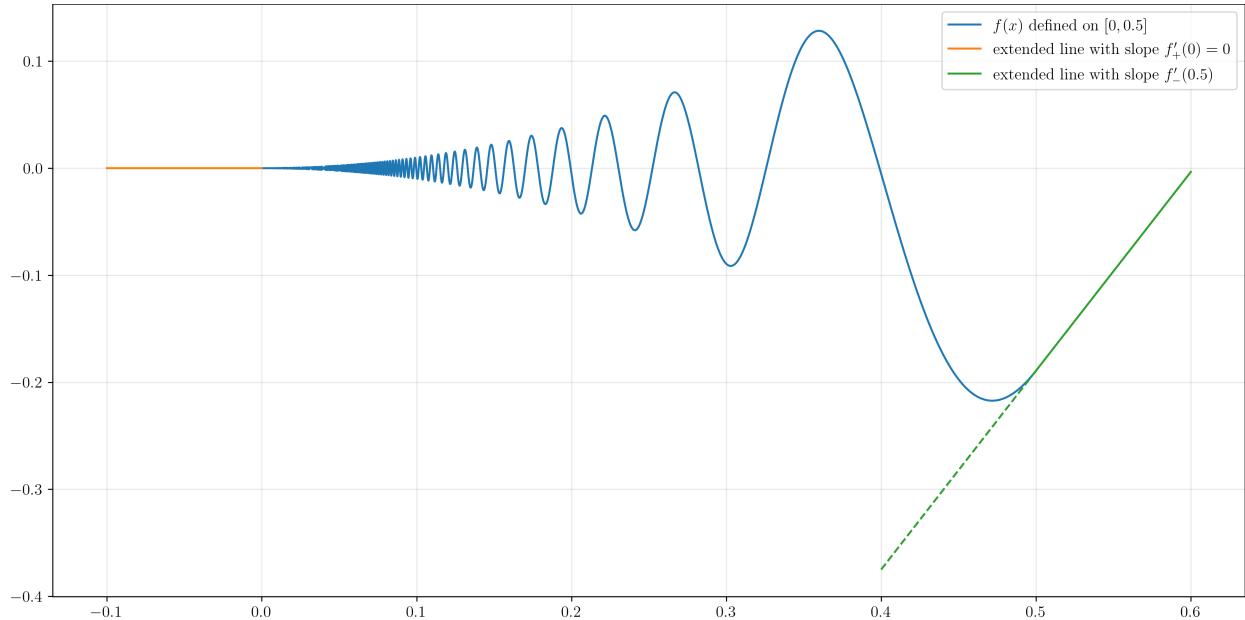
## 5.2 Derivatives and Continuity

The next theorem helps to reduce some theorems on derivatives to theorems on continuity.

### Theorem 5.2.1

Let  $f$  be a function defined on  $(a, b)$ , and  $c \in (a, b)$  a fixed point in that interval. We have the following statements:

1. If  $f$  is differentiable at  $c$ , then there exists a function  $\phi$  on  $(a, b)$  which is continuous at  $c$ ,



**Figure 5.5:**  $f(x) = x^2 \sin \frac{1}{x^2}$  defined on  $[0, 0.5]$  whose one-sided derivatives both exist and are finite.

and satisfies

$$f(x) - f(c) = (x - c)\phi(x) \quad \forall x \in (a, b) \quad (5.2)$$

And  $f'(c) = \phi(c)$ .

2. Conversely, if there exists a function  $\phi$  on  $(a, b)$  which is continuous at  $c$ , and satisfies (5.2), then  $f$  is differentiable at  $c$  with  $f'(c) = \phi(c)$ .



The function  $\phi$  is precisely the slope of function  $f$ 's secant line.

**Proof** (Proof of 1) Suppose  $f$  is differentiable at  $c$ . Define

$$\phi(x) = \frac{f(x) - f(c)}{x - c}, \quad x \neq c \quad \text{and} \quad \phi(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}$$

Note that  $\phi(c)$  is well-defined since the limit (which equals  $f'(c)$ ) indeed exists by the definition of the derivative of  $f$  at  $c$ . Notice also that by defining function  $\phi$  in this way, it is automatically continuous at  $c$ . And of course, the equation (5.2) holds.

(Proof of 2) It follows from (5.2) that

$$\phi(x) = \frac{f(x) - f(c)}{x - c} \quad \forall x \neq c$$

Since  $\phi$  is continuous at  $c$ , by sending  $x \rightarrow c$ , we have

$$\phi(c) = \lim_{x \rightarrow c} \phi(x) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}$$

This implies  $f$  is differentiable at  $c$ , and  $f'(c) = \phi(c)$ . ■

With the help of Theorem 5.2.1, we can easily show that  $f$  must be continuous at  $c$  if it is differentiable there.

### Theorem 5.2.2

Let  $f$  be defined on  $(a, b)$ . If  $f$  is differentiable at  $c \in (a, b)$ , then it is continuous at  $c$ .



**Proof** By Theorem 5.2.1, there exists a function  $\phi$  on  $(a, b)$  such that  $\phi$  is continuous at  $c$ , and

$$f(x) - f(c) = (x - c)\phi(x) \quad \forall x \in (a, b) \quad (5.3)$$

Since the limit of the right-hand side in (5.3) exists as  $x \rightarrow c$ , the limit of the left-hand side also exists, and

$$\lim_{x \rightarrow c} f(x) - f(c) = \lim_{x \rightarrow c} (x - c)\phi(x) = 0$$

This implies

$$\lim_{x \rightarrow c} f(x) = f(c)$$

Hence,  $f$  is continuous at  $c$ . ■

## 5.3 Computing Derivatives

It is difficult in practice to compute the derivative of any function only using the definition. In this section, we shall introduce some theorems and standard rules that are helpful with computation. If one knows the derivatives of elementary functions (e.g.,  $x^a$ ,  $e^x$ ,  $\ln x$ ,  $\sin x$ ,  $\cos x$ , etc.), then by applying these theorems, one should be able to compute the derivatives of any functions that are built up out of elementary functions using summations, differences, products, quotients, and compositions.

 **Note** *In fact, these so-called elementary functions are not elementary at all. We shall present their definitions as well as their derivatives rigorously in later chapters.*

### 5.3.1 Algebra of Derivatives

#### Theorem 5.3.1

Suppose  $f$  and  $g$  are defined on  $(a, b)$  and are both differentiable at  $c \in (a, b)$ . Then the functions  $f + g$ ,  $f - g$ ,  $fg$  are also differentiable at  $c$ . If  $g(c) \neq 0$ , then the quotient  $f/g$  is also differentiable at  $c$ . And the formulas of their derivatives are given by

1.  $(f \pm g)'(c) = f'(c) \pm g'(c)$ ,
2.  $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$ ,
3.  $(f/g)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g^2(c)}$ , provided  $g(c) \neq 0$ .



**Remark** One might be worried about the possibility that the quotient  $f/g$  may only be defined at point  $c$  if we only require  $g(c) \neq 0$ . Then how can we say  $f/g$  is differentiable at some point? However, the condition that  $g$  is differentiable at  $c$  implicitly implies that  $g(x) \neq 0$  for  $x$  in some interval about  $c$  if we assume  $g(c) \neq 0$ . Take a few seconds and think about why this is true. Note that  $g$  is continuous at  $c$  by Theorem 5.2.2. Then, if  $g(c) \neq 0$ , by the continuity,  $g$  is also nonzero in some neighborhood of  $c$ . Hence, in this case,  $f/g$  is actually well-defined in an interval instead of at just one point.

The proof we present here is done by exploiting Theorem 5.2.1, and it may differ from a traditional proof that is solely based on the definition.

**Proof** Since  $f$  and  $g$  are differentiable at  $c$ . It follows from Theorem 5.2.1 that there exists functions

$\phi_1$  and  $\phi_2$  on  $(a, b)$  that are continuous at  $c$ , and satisfy

$$f(x) - f(c) = (x - c)\phi'_1(x) \quad (5.4)$$

$$g(x) - g(c) = (x - c)\phi'_2(x) \quad (5.5)$$

with

$$\phi_1(c) = f'(c) \quad \text{and} \quad \phi_2(c) = g'(c) \quad (5.6)$$

(Proof of 1) Applying (5.4) and (5.5) to  $(f \pm g)(x) - (f \pm g)(c)$ , we have

$$(f \pm g)(x) - (f \pm g)(c) = (x - c)(\phi_1(x) \pm \phi_2(x))$$

Since  $\phi_1(x) \pm \phi_2(x)$  is continuous at  $c$ ,  $(f \pm g)$  is differentiable at  $c$  by Theorem 5.2.1. Then by applying (5.6), we obtain

$$(f \pm g)'(c) = \phi_1(c) \pm \phi_2(c) = f'(c) \pm g'(c)$$

(Proof of 2) Again by applying (5.4) and (5.5), we have

$$(fg)(x) - (fg)(c) = (x - c)[\phi_1(x)g(c) + f(c)\phi_2(x) + (x - c)\phi_1(x)\phi_2(x)]$$

after a few steps of algebraic calculation. It then follows from Theorem 5.2.1 that  $fg$  is differentiable at  $c$  since the function

$$\phi_1(x)g(c) + f(c)\phi_2(x) + (x - c)\phi_1(x)\phi_2(x)$$

is continuous at  $c$ . And

$$(fg)'(c) = \phi_1(c)g(c) + f(c)\phi_2(c) + (c - c)\phi_1(c)\phi_2(c) = f'(c)g(c) + f(c)g'(c)$$

(Proof of 3) Note that the quotient  $(f/g)$  can be regarded as a product of  $f$  and  $1/g$ . Hence, we are going to show  $1/g$  is differentiable at  $c$ , and then apply statement 2, which we have just proved, to conclude the proof. As explained in the remark of this theorem,  $1/g$  is well-defined in a neighborhood of  $c$ . It follows from (5.5) that

$$\frac{1}{g(x)} - \frac{1}{g(c)} = (x - c) \frac{-\phi_2(x)}{g(x)g(c)}$$

Since  $-\phi_2(x)/[g(x)g(c)]$  is continuous at  $c$ , we conclude from Theorem 5.2.1 that  $1/g$  is differentiable at  $c$ , and

$$\left(\frac{1}{g}\right)'(c) = \frac{-\phi_2(c)}{g(c)g(c)} = \frac{-g'(c)}{g^2(c)}$$

Then by applying statement 2, we conclude that  $f/g$  is also differentiable at  $c$  with

$$(f/g)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g^2(c)}$$

■

### 5.3.2 The Chain Rule

Taking compositions is a more fundamental and natural way of combining two functions apart from summations, products, etc. The next result, the **chain rule**, provides a method of computing the derivative of a composite function.

**Theorem 5.3.2**

Suppose  $f$  is a function defined on an open interval  $I$ . Let  $c \in I$  be an interior point, and  $g$  a function defined on some open interval  $J$  containing  $f(c)$ . We consider the composite function  $g \circ f$ . If  $f$  is differentiable at  $c$ , and  $g$  is differentiable at  $f(c)$ , then  $g \circ f$  is differentiable at  $c$  with

$$(g \circ f)'(c) = g'(f(c))f'(c) \quad (5.7)$$



With the help of Theorem 5.2.1, we can reduce the proof of this theorem to that of Theorem 4.3.1.

**Proof** Since  $f$  is differentiable at  $c$ , using Theorem 5.2.1, we have

$$f(x) - f(c) = (x - c)\phi_1(x) \quad \forall x \in I \quad (5.8)$$

where  $\phi_1$  is some function that is continuous at  $c$  with  $\phi_1(c) = f'(c)$ . Using the same argument, since  $g$  is differentiable at  $f(c)$ , we have

$$g(y) - g(f(c)) = (y - f(c))\phi_2(y) \quad \forall y \in J \quad (5.9)$$

where  $\phi_2$  is some function that is continuous at  $f(c)$  with  $\phi_2(f(c)) = g'(f(c))$ . Replacing  $y$  with  $f(x)$  in (5.9) yields

$$g(f(x)) - g(f(c)) = (f(x) - f(c))\phi_2(f(x)) \quad \forall x \in I$$

Then by replacing  $f(x) - f(c)$  with the right-hand side of (5.8), we have

$$g(f(x)) - g(f(c)) = (x - c)\phi_1(x)\phi_2(f(x)) \quad \forall x \in I$$

Note that  $\phi_2(f(x))$  is continuous at  $c$  by Theorem 4.3.1. And the product  $\phi_1(x)\phi_2(f(x))$  is also continuous. This implies  $g \circ f$  is differentiable at  $c$  by Theorem 5.2.1. And its derivative at  $x = c$  equals

$$(g \circ f)'(c) = \phi_1(c)\phi_2(f(c)) = f'(c)g'(f(c))$$

which is exactly (5.7). ■

☞ **Exercise 5.3** Calculate the derivative of

$$g(x) = x \sin \frac{1}{x} \mathbb{1}\{x \neq 0\}$$

for  $x \neq 0$ . Does the derivative exist at  $x = 0$ ?

☞ **Exercise 5.4** Calculate the derivative of

$$f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}\{x \neq 0\}$$

for all  $x \in \mathbb{R}$ . Recall we have already calculated its derivative at  $x = 0$  in Exercise 5.2, which is  $f'(0) = 0$ .

[Hint: Make use of the result in Exercise 5.3 and the chain rule (Theorem 5.3.2).]

## 5.4 The Mean Value Theorem

### 5.4.1 Local Extrema

One may be familiar with the fact that if the derivative of function  $f$  is positive at some point, i.e.,  $f'(c) > 0$ , then  $f$  is increasing near  $c$ , and if  $f'(c) < 0$ , it is decreasing. However, we can extend this result a little further for we have defined one-sided and infinite derivatives.

#### Theorem 5.4.1

*Let  $f$  be defined on a subset  $A \subseteq \mathbb{R}$  and  $c \in A$ . If  $f'_+(c) > 0$ , possibly  $f'_+(c) = \infty$ , then there exists a half-open ball  $(c, c + \delta) \subseteq A$  in which*

$$x > c \implies f(x) > f(c) \quad (5.10)$$

*Similarly, if  $f'_-(c) > 0$ , then  $\exists (c - \delta, c) \subseteq A$  in which*

$$x < c \implies f(x) < f(c)$$

*Analogous results hold if we instead assume  $f'_+(c) < 0$  or  $f'_-(c) < 0$ .*



**Proof** We only prove (5.10) since the proofs of the rest of the results are similar. First suppose  $f'_+(c) < \infty$ . Because the right-hand limit exists, and it equals  $f'_+(c)$ , there exists  $\delta > 0$  such that

$$\left| \frac{f(x) - f(c)}{x - c} - f'_+(c) \right| < \frac{f'_+(c)}{2} \quad \forall x \in (c, c + \delta)$$

This implies

$$\frac{f(x) - f(c)}{x - c} > \frac{f'_+(c)}{2} > 0 \quad \forall x \in (c, c + \delta)$$

Then  $f(x) > f(c)$  provided  $x > c$  since they have the same sign.

We also need to consider  $f'_+(c) = \infty$ . Because the limit is infinity, there exists  $\delta > 0$  such that the following fraction exceeds any predetermined number, in particular,

$$\frac{f(x) - f(c)}{x - c} > 0 \quad \forall x \in (c, c + \delta)$$

One can instead choose any large number (non-negative of course, to prove this theorem). This completes the proof since both the numerator and denominator share the same sign. ■

The following corollary is the version that is more familiar to the reader and is more common in practice.

#### Corollary 5.4.1

*Let  $f$  be defined on  $(a, b)$ , and  $c \in (a, b)$  an interior point. If  $f'(c) > 0$ , possibly  $f'(c) = \infty$ , then there exists an open ball  $B_\delta(c) \subseteq (a, b)$  in which*

$$x > c \implies f(x) > f(c) \quad \text{and} \quad x < c \implies f(x) < f(c) \quad (5.11)$$

*An analogous result holds if  $f'(c) < 0$ .*



☞ **Exercise 5.5** Prove Corollary 5.4.1 assuming  $f'(c) < \infty$  using Theorem 5.2.1.

**Solution** It follows from Theorem 5.2.1 that

$$f(x) - f(c) = (x - c)\phi(x) \quad \forall x \in (a, b) \quad (5.12)$$

where  $\phi(x)$  is continuous at  $c$  with  $\phi(c) = f'(c) > 0$ . Since  $\phi(c) > 0$  and it is continuous there, there exists some neighborhood  $B_\delta(c) \subseteq (a, b)$  such that  $\phi(x) > 0 \quad \forall x \in B_\delta(c)$ . And then (5.11) follows from (5.12).

The **local extremum** of a function is the largest or smallest value within some *open ball*.

### Definition 5.4.1

Let  $f$  be a real-valued function defined on a subset  $A$  of a metric space  $(X, d)$ , and suppose  $p \in A$ . Then  $f$  is said to have a **local maximum** at  $p$  if

$$f(x) \leq f(p) \quad \forall x \in B_\delta(p) \cap A$$

for some open ball  $B_\delta(p)$ . An analogous definition exists for **local minimum**.



The reader should be very familiar with the exercise of finding a local extremum by solving the equation in which the derivative vanishes. The following Theorem is the theory behind that.

### Theorem 5.4.2

Let  $f$  be defined on  $(a, b)$ , and suppose  $f$  has a local extremum at an interior point  $c \in (a, b)$ . If  $f'(c)$  exists as a finite or infinite number, then  $f'(c) = 0$ .



**Remark** Since the conclusion is  $f'(c) = 0$ , we may as well just assume  $f$  is differentiable at  $c$ . The reason why we suppose  $f'(c)$  exists is that it makes the assumption somewhat looser.

**Proof** We shall prove by contradiction. Assume  $f'(c) \neq 0$ , then either  $f'(c) > 0$  or  $f'(c) < 0$ . (We do not exclude the possibilities that  $f'(c) = \infty$  or  $f'(c) = -\infty$ .) If  $f'(c) > 0$ , then follows from Corollary 5.4.1 that there acts some open ball  $B_\delta(c)$  in which  $f(x) > f(c)$  for  $x > c$  and  $f(x) < f(c)$  for  $x < c$ . This contradicts the fact that  $f(c)$  is a local extremum. Similarly,  $f'(c) < 0$  will also lead to a contradiction. ■

The converse of this Theorem is not true.

**Example 5.3** Consider  $f(x) = x^3$ . Its derivative is  $f'(x) = 3x^2$ . We note that  $f'(0) = 0$ . But it is clear that  $f(0) = 0$  is neither a local maximum nor a local minimum. In fact,  $f(x)$  does not have any local extremum.

☞ **Exercise 5.6** Recall the function

$$f(x) = x^2 \sin \frac{1}{x^2} \mathbb{1}_{\{x \neq 0\}}$$

defined in Exercise 5.2. Show that  $f(0)$  is not a local extremum. Recall we have shown that  $f'(0) = 0$ . This is another example that the converse of Theorem 5.4.2 does not hold.

Moreover, it should be emphasized that we conclude in Theorem 5.4.2 that  $f'(c) = 0$  under the condition that  $f'(c)$  exists.  $f(c)$  may be a local extremum without having a derivative there.

**Example 5.4** The simplest example is  $f(x) = |x|$ . Note that  $f(0) = 0$  is a local minimum. But the derivative does not exist at  $x = 0$  since  $f'_+(0) = 1 \neq -1 = f'_-(0)$ .

## 5.4.2 Rolle's Theorem and Mean Value Theorem

If a function starts from point  $a$  and ends at point  $b$  with the same level of height as its initial position, then there must be a turning point somewhere in between.

### Theorem 5.4.3 (Rolle)

*Let  $f$  be defined on  $[a, b]$ . Suppose  $f'(x)$  exists (as finite or infinite number) for each  $x \in (a, b)$ , and  $f$  is continuous at the endpoints  $a$  and  $b$ . If  $f(a) = f(b)$ , then there exists  $c \in (a, b)$  such that*

$$f'(c) = 0$$



**Proof** If  $f$  is a constant function, i.e.,  $f(x) = f(a) = f(b) \forall x \in [a, b]$ , then the conclusion is trivial since  $f'(x) = 0 \forall x \in (a, b)$ .

We now suppose  $f$  is non-constant. We know that  $f$  is continuous on  $[a, b]$  since  $f'$  exists throughout  $(a, b)$  and  $f$  is continuous at endpoints. It then follows from Theorem 4.3.2 that  $f$  attains its maximum and minimum on  $[a, b]$ . Because we assume  $f$  is non-constant and  $f(a) = f(b)$ , at least one of the maximum and the minimum of  $f$  is reached at an interior point  $c \in (a, b)$ . Note that  $f(c)$  is a global extremum, and of course also a local extremum. Then by applying Theorem 5.4.2, we conclude  $f'(c) = 0$ . ■

**Example 5.5** Review Example 5.1. Note that  $f(0) = f(4) = 0$ . And the derivative vanishes at  $x = 1$  and  $x = 3$ .

The **mean value theorem** is a generalization of Rolle's theorem. Roughly speaking it says  $f$  has a derivative value that equals the slope of the secant line joining two endpoints of the graph of  $f$ . The mean value theorem itself can be treated as a special case of an even more generalized version, the **generalized mean value theorem** (Theorem 5.4.5).

### Theorem 5.4.4 (Mean Value Theorem)

*Suppose  $f$  has a derivative (finite or infinite) at each point of  $(a, b)$ , and suppose  $f$  is continuous at endpoints  $a$  and  $b$ . Then there exists  $c \in (a, b)$  such that*

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



### Theorem 5.4.5 (Generalized Mean Value Theorem)

*Let  $f$  and  $g$  be two functions, each having a derivative (finite or infinite) at each point in  $(a, b)$ . Suppose  $f$  and  $g$  are both continuous at endpoints  $a$  and  $b$ , and  $f'(x)$  and  $g'(x)$  do not assume infinite values simultaneously. Then there exists  $c \in (a, b)$  such that*

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)] \quad (5.13)$$



**Remark** By taking  $g(x) = x$ , this theorem reduces to the standard mean value theorem, Theorem 5.4.4.

Though it is a generalized version of the mean value theorem, and hence Rolle's theorem, it can be proved easily using Rolle's theorem by constructing a function out of  $f$  and  $g$ .

**Proof** Define

$$\psi(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$$

It is clear that  $\psi$  is continuous on  $[a, b]$ . Since  $f'(x)$  and  $g'(x)$  can never both be infinity, the following equation holds:

$$\psi'(x) = f'(x)[g(b) - g(a)] - g'(x)[f(b) - f(a)] \quad \forall x \in (a, b)$$

Furthermore, we note that

$$\psi(a) = f(a)g(b) - f(b)g(a) = \psi(b)$$

Applying Rolle's theorem (Theorem 5.4.3) to  $\psi$  leads to  $\psi'(c) = 0$  for some  $c \in (a, b)$ , which is exactly (5.13). ■

Sometimes, functions  $f$  and  $g$  in Theorem 5.4.5 may not be defined at endpoints. But all the one-sided limits exist as finite values. In this case, we still wish to apply the theorem. The following is a slight extension of Theorem 5.4.5.

### Theorem 5.4.6

Let  $f$  and  $g$  be two functions, each having a derivative (finite or infinite) at each point in  $(a, b)$ .

Suppose that  $f'(x)$  and  $g'(x)$  do not assume infinite values simultaneously, and one-sided limits  $f(a+)$ ,  $f(b-)$ ,  $g(a+)$  and  $g(b-)$  all exist as finite values. Then there exists  $c \in (a, b)$  such that

$$f'(c)[g(b-) - g(a+)] = g'(c)[f(b-) - f(a+)] \quad (5.14)$$



The proof is done by simply defining the function values at the endpoints.

**Proof** Define

$$\tilde{f}(x) = f(x) \quad \forall x \in (a, b), \quad \tilde{f}(a) = f(a+), \quad \text{and} \quad \tilde{f}(b) = f(b-)$$

$$\tilde{g}(x) = g(x) \quad \forall x \in (a, b), \quad \tilde{g}(a) = g(a+), \quad \text{and} \quad \tilde{g}(b) = g(b-)$$

It is evident that  $\tilde{f}$  and  $\tilde{g}$  are both continuous on  $[a, b]$ . Furthermore, the derivatives  $\tilde{f}'(x) = f'(x)$  and  $\tilde{g}'(x) = g'(x)$  for each  $x \in (a, b)$ . Applying Theorem 5.4.5 to functions  $\tilde{f}$  and  $\tilde{g}$  leads to (5.14). ■

Of course, by letting  $g(x) = x$  we will also obtain a version concerning only one function.

### Theorem 5.4.7

Let  $f$  be a function having a derivative (finite or infinite) at each point in  $(a, b)$ . Suppose that one-sided limits  $f(a+)$  and  $f(b-)$  at the endpoints both exist as finite values. Then there exists  $c \in (a, b)$  such that

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



The following theorem is an immediate result of the mean value theorem, which provides a sufficient condition for determining strictly increasing and decreasing functions. It also says a function is constant if its derivative vanishes everywhere.

**Theorem 5.4.8**

Suppose  $f$  has a derivative (finite or infinite) at each point in  $(a, b)$ , and itself is continuous at the endpoints  $a$  and  $b$ .

1. If  $f'(x) > 0 \forall x \in (a, b)$ , then  $f$  is strictly increasing on  $[a, b]$ .
2. If  $f'(x) < 0 \forall x \in (a, b)$ , then  $f$  is strictly decreasing on  $[a, b]$ .
3. If  $f'(x) = 0 \forall x \in (a, b)$ , then  $f$  is constant on  $[a, b]$ .



**Proof** (Proof of 1) Let  $x_1, x_2 \in [a, b]$  with  $x_1 < x_2$ . Applying Theorem 5.4.4 to  $f$  on the interval  $[x_1, x_2]$  leads to

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$$

for some  $c \in (x_1, x_2)$ . Since  $f'(c) > 0$  and  $x_2 - x_1 > 0$ , we have  $f(x_2) > f(x_1)$ . Because  $x_1$  and  $x_2$  are arbitrarily chosen, this implies  $f$  is strictly increasing on  $[a, b]$ .

(Proof of 2) Similarly, for any  $x_1, x_2 \in [a, b]$  with  $x_1 < x_2$ , we have

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) < 0$$

(Proof of 3) For any  $x_1, x_2 \in [a, b]$  with  $x_1 < x_2$ , by applying Theorem 5.4.4, we have

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) = 0$$

This implies  $f$  is a constant function. ■

If  $f$  and  $g$  have the same derivatives everywhere in  $(a, b)$ , then they only differ by a constant.

**Corollary 5.4.2**

If  $f$  and  $g$  are continuous on  $[a, b]$ , and  $f'(x) = g'(x) \forall x \in (a, b)$ , then  $f - g$  is constant on  $[a, b]$ .



**Proof** Note that the difference  $f - g$  is continuous on  $[a, b]$  and  $(f - g)'(x) = 0 \forall x \in (a, b)$ . It then follows from Theorem 5.4.8 that  $f - g$  is a constant. ■

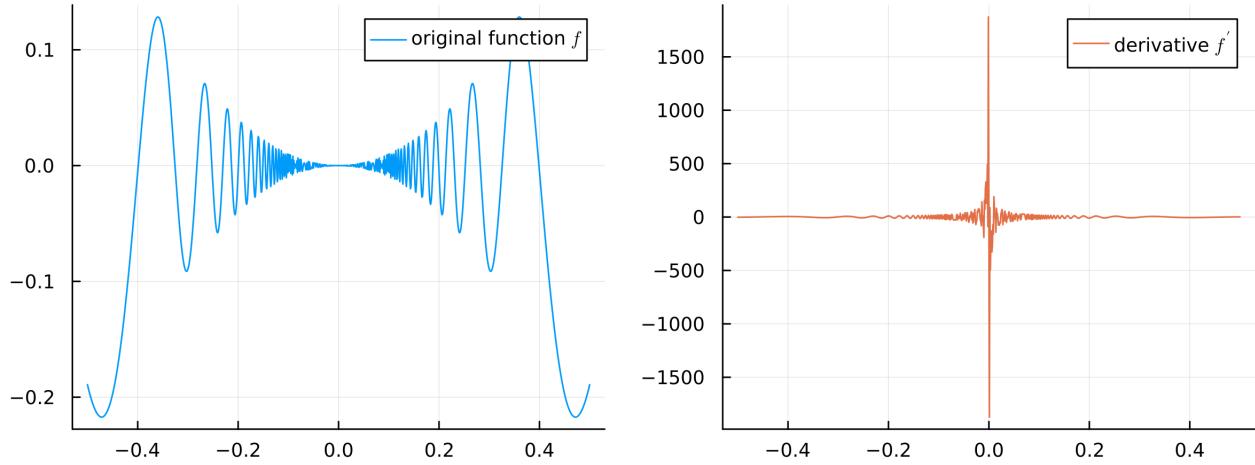
## 5.5 Intermediate Value Theorem

In Exercise 5.4, we have shown that the derivative of function  $f(x) = x^2 \sin(1/x^2) \mathbb{1}\{x \neq 0\}$  is given by

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

We plot the original function  $f$  along with its derivative  $f'$  side by side in Figure 5.6.

Note that the derivative  $f'$  is not continuous at  $x = 0$  since the limit of  $f'(x)$  does not exist as  $x \rightarrow 0$ . This example shows that the derivative of a function may not be continuous, and hence theorems for continuous functions do not apply to derivatives in general. However, the **intermediate value theorem** is an exception. The intermediate value theorem for derivatives is also known as the **Darboux's theorem**.



**Figure 5.6:** Left: original function. Right: derivative.

### Theorem 5.5.1 (Darboux)

Suppose  $f$  is defined on  $[a, b]$ , and it has a derivative (finite or infinite) at each point in  $(a, b)$ . Assume also that the one-sided derivatives  $f'_+(a)$  and  $f'_-(b)$  both exist as finite numbers, and  $f'_+(a) \neq f'_-(b)$ . If  $k$  is a number in between  $f'_+(a)$  and  $f'_-(b)$ , then there exists an interior point  $c \in (a, b)$  such that

$$f'(c) = k$$



**Proof** Note that  $f$  is continuous on the entire closed interval  $[a, b]$ .

We first consider the case where  $k = [f(b) - f(a)]/(b - a)$ . Then by applying the mean value theorem (Theorem 5.4.4), there exists  $c \in (a, b)$  such that  $f'(c) = [f(b) - f(a)]/(b - a) = k$ .

Now, we suppose  $k \neq [f(b) - f(a)]/(b - a)$ . We note that either  $f'_+(a) \neq [f(b) - f(a)]/(b - a)$  or  $f'_-(b) \neq [f(b) - f(a)]/(b - a)$  since  $f'_+(a) \neq f'_-(b)$ . In other words,  $k$  is either in between  $f'_+(a)$  and  $[f(b) - f(a)]/(b - a)$  or it is in between  $f'_-(b)$  and  $[f(b) - f(a)]/(b - a)$ .

Without loss of generality, we may assume  $k$  is in between  $f'_+(a)$  and  $[f(b) - f(a)]/(b - a)$ . Define a function  $\phi$  on  $[a, b]$  by

$$\phi(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} & x \neq a \\ f'_+(a) & x = a \end{cases}$$

Note that  $\phi$  is continuous on  $[a, b]$ , and  $k$  is in between  $\phi(a)$  and  $\phi(b)$ . Hence, we may apply the intermediate values theorem for continuous functions (Theorem 4.5.1) to  $\phi$ , there is a number  $d \in (a, b)$  such that

$$k = \phi(d) = \frac{f(d) - f(a)}{d - a}$$

Then applying the mean value theorem (Theorem 5.4.4) to the fraction on the right-hand side of the above equation, we have

$$k = \frac{f(d) - f(a)}{d - a} = f'(c)$$

for some  $c$  in between  $a$  and  $d$ .

For the case where  $k$  lying in between  $f'_-(b)$  and  $[f(b) - f(a)]/(b - a)$ , it can be proved with a

similar argument, started by defining

$$\phi(x) = \begin{cases} \frac{f(x)-f(b)}{x-b} & x \neq b \\ f'_-(b) & x = b \end{cases}$$

■

Essentially, Darboux's theorem tells us that though the derivative may not be continuous, it cannot have any *jump discontinuities*. The following is another classical example in addition to the one in Exercise 5.2.

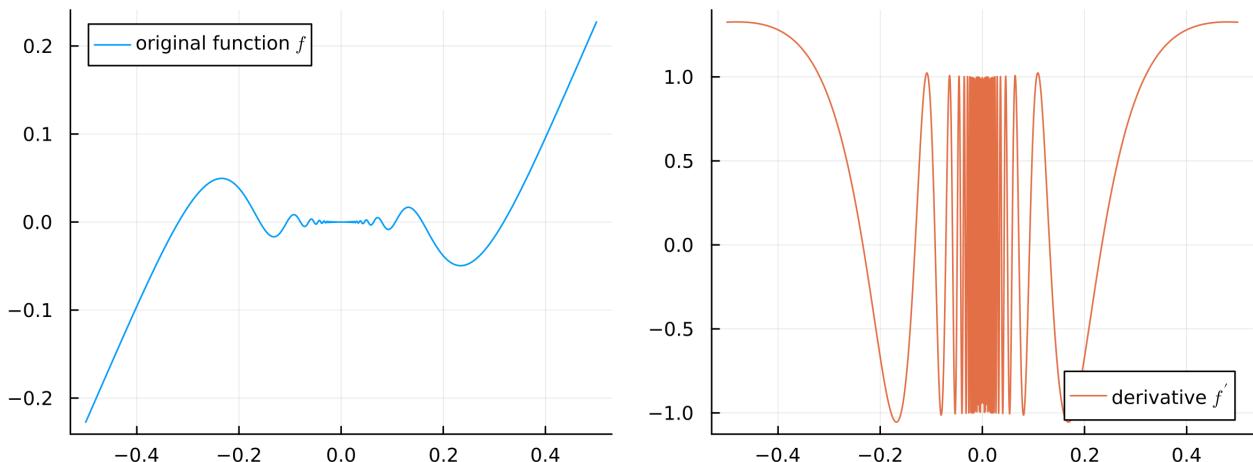
**Example 5.6** Consider the function

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

Its derivative is given by

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Note that the derivative  $f'$  is not continuous at  $x = 0$ . The graphs of  $f$  and  $f'$  are shown in Figure 5.7.



**Figure 5.7:** Left: original function. Right: derivative.

Actually, there is no need to require that  $f'_+(a)$  and  $f'_-(b)$  to be finite numbers in Theorem 5.5.1. The following theorem is an extension, and its proof is somewhat neater and more interesting than that of Theorem 5.5.1.

### Theorem 5.5.2

Suppose  $f$  is defined on  $[a, b]$ , and it has a derivative (finite or infinite) at each point in  $(a, b)$ . Assume also that the one-sided derivatives  $f'_+(a)$  and  $f'_-(b)$  both exist (each of the two can either be finite or infinite), and  $f'_+(a) \neq f'_-(b)$ . If  $k$  is a number in between  $f'_+(a)$  and  $f'_-(b)$ , then there exists an interior point  $c \in (a, b)$  such that

$$f'(c) = k$$



**Proof** Without loss of generality, we may assume  $f'_+(a) < k < f'_-(b)$ . Define a function  $g$  on  $[a, b]$  by the equation

$$g(x) = f(x) - kx$$

Note that  $g$  is continuous on  $[a, b]$ , and the derivative of  $g$  exists at each point in  $(a, b)$ . Furthermore, the one-sided derivatives of  $g$ ,  $g'_+(a)$  and  $g'_-(b)$ , both exist, which are given by

$$g'_+(a) = f'_+(a) - k \quad \text{and} \quad g'_-(b) = f'_-(b) - k \quad (5.15)$$

It should be emphasized that (5.15) also holds for infinite values.

Because we assume  $f'_+(a) < k < f'_-(b)$ , (5.15) yields

$$g'_+(a) < 0 \quad \text{and} \quad g'_-(b) > 0$$

It then follows from Theorem 5.4.1 that there exist points  $t_1$  and  $t_2$  at which

$$g(t_1) < g(a) \quad \text{and} \quad g(t_2) < g(b)$$

This implies that  $g$  must attain its minimum at some interior point  $c \in (a, b)$  rather than at the endpoints. Then by applying Theorem 5.4.2, we have

$$g'(c) = 0$$

Because  $g'(x) = f'(x) - k \forall x \in (a, b)$ , it follows that

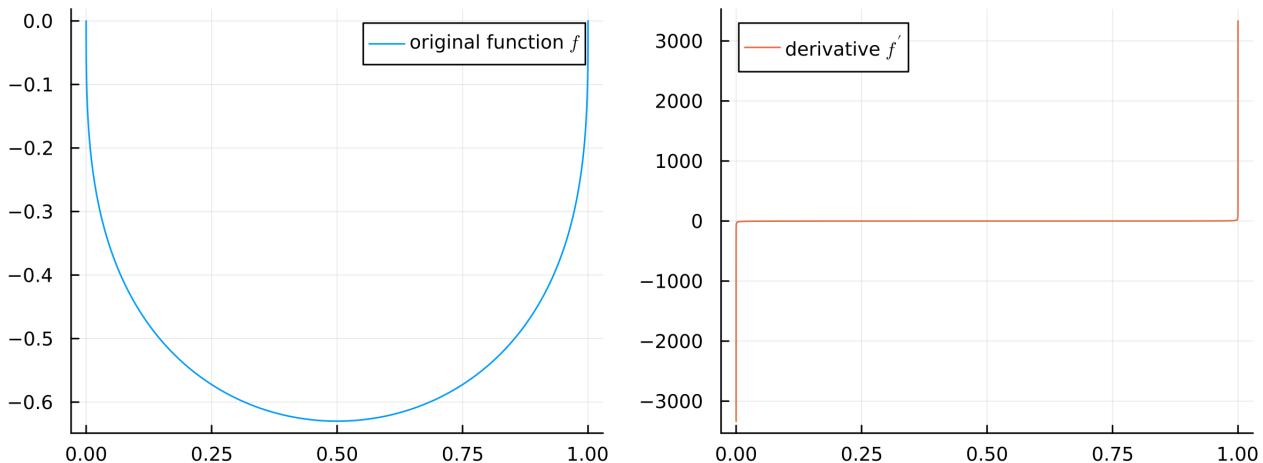
$$f'(c) = k$$

This completes the proof. ■

**Example 5.7** Consider a function defined on  $[0, 1]$  given by

$$f(x) = \sqrt[3]{x(x-1)}, \quad x \in [0, 1]$$

Note that the derivative exists as a finite number in  $(0, 1)$ . And it can be shown that  $f'_+(0) = -\infty$  and  $f'_-(1) = \infty$ . Then Theorem 5.5.2 tells us  $f'$  takes all the real numbers in  $(0, 1)$ . See Figure 5.8.



**Figure 5.8:** Left: original function. Right: derivative.



**Note** Theoretically, the graph of  $f'$  in Figure 5.8 will tend to  $-\infty$  and  $\infty$  as  $x$  approaches 0 and 1, respectively. The reason why the absolute values of the derivatives near the endpoints shown in the graph are only about 3000 is because of the limited computational precision of the computer.

By Theorem 5.4.8, we know a function is strictly increasing if its derivative is always positive, and it is strictly decreasing if its derivative is always negative. Suppose now we only know that  $f'$  is nonzero in some open interval, what can we conclude? Thanks to the intermediate value theorem, or Darboux's theorem,  $f'$  is either always positive or always negative. Hence, we can conclude that  $f$  must be strictly monotonic.

### Theorem 5.5.3

*Suppose  $f$  has a derivative (finite or infinite) in  $(a, b)$ , and is continuous at endpoints  $a$  and  $b$ .*

*If  $f'(x) \neq 0 \forall x \in (a, b)$ , then  $f$  is strictly monotonic on  $[a, b]$ .*



**Proof** Pick a point  $x_1 \in (a, b)$ . Suppose  $f'(x_0) > 0$ . Let  $x_0$  be fixed, and choose an arbitrary point  $x \in (a, b)$  other than  $x_0$ . We claim that  $f'(x) > 0$ . Otherwise, if  $f'(x) < 0$ , then by Theorem 5.5.2, there exists some point  $c$  in between  $x_0$  and  $x$  such that  $f'(c) = 0$ , which contradicts the given condition that  $f'$  is nonzero in  $(a, b)$ . In this case, we have  $f'(x) > 0 \forall x \in (a, b)$ . It then follows from Theorem 5.4.8 that  $f$  is strictly increasing on  $[a, b]$ .

If  $f'(x_0) < 0$ , then a similar argument will show that  $f$  is strictly decreasing on  $[a, b]$ . ■

It is because the derivatives cannot have jump discontinuities that monotonic derivatives are necessarily continuous.

### Theorem 5.5.4

*Suppose  $f'$  exists, and is monotonic in  $(a, b)$ . Then  $f'$  is continuous in  $(a, b)$ .*



**Proof** Without loss of generality, suppose that  $f'$  is increasing. We shall prove by contradiction. Assume that  $f'$  is discontinuous at  $x = c \in (a, b)$ . By Theorem 4.6.1, we have  $f'(c-)$  and  $f'(c+)$  both exist, and

$$f'(c-) \leq f(c) \leq f'(c+)$$

But because we assume  $f'$  is not continuous at  $x = c$ , it holds that

$$f'(c-) < f'(c+)$$

Let  $[c - \delta, c + \delta]$  ( $\delta > 0$ ) be a closed interval contained in  $(a, b)$ . We have

$$f'(c - \delta) \leq f'(c-) < f'(c+) \leq f'(c + \delta)$$

Let  $k$  be a number in between  $f'(c-)$  and  $f'(c+)$  other than  $f'(c)$ . That is,

$$f'(c-) < k < f'(c+) \quad \text{and} \quad k \neq f'(c)$$

Applying Theorem 5.5.2 to  $f'$  on  $[c - \delta, c + \delta]$ , we conclude that there exists  $x_0 \in (c - \delta, c + \delta)$  such that  $f'(x_0) = k$ . But since  $f'$  is increasing,  $f'(c-) < k = f'(x_0)$  and  $k \neq f'(c)$ , we have  $x_0 > c$ . On the other hand, because  $f'(x_0) = k < f'(c+)$  and  $k \neq f'(c)$ , we have  $x_0 < c$ . This leads to a contradiction. ■

## 5.6 Taylor's Theorem

If  $f$  is differentiable at  $x = c$ , then it can be approximated by

$$f(x) \approx f(c) + f'(c)(x - c) \quad (5.16)$$

when  $|x - c|$  is small. Geometrically,  $f$  is close to its tangent line at  $x = c$  near the point of tangency. Another way to look at (5.16) is that  $f$  is approximated by a polynomial of degree one whose first-order derivative at  $x = c$  is exactly  $f'(c)$ . And the value of this polynomial at  $x = c$  is  $f(c)$ , that is, its zeroth-order derivative is  $f(c)$ . We can extend this idea by approximating  $f$  with higher degree polynomials whose derivatives of all orders match with that of  $f$  at  $x = c$ .

**Example 5.8** Suppose we know any order derivative of the function  $f(x) = e^x$  is itself. That is,

$$f^{(k)}(x) = f(x) = e^x \quad \forall k = 0, 1, 2, \dots$$

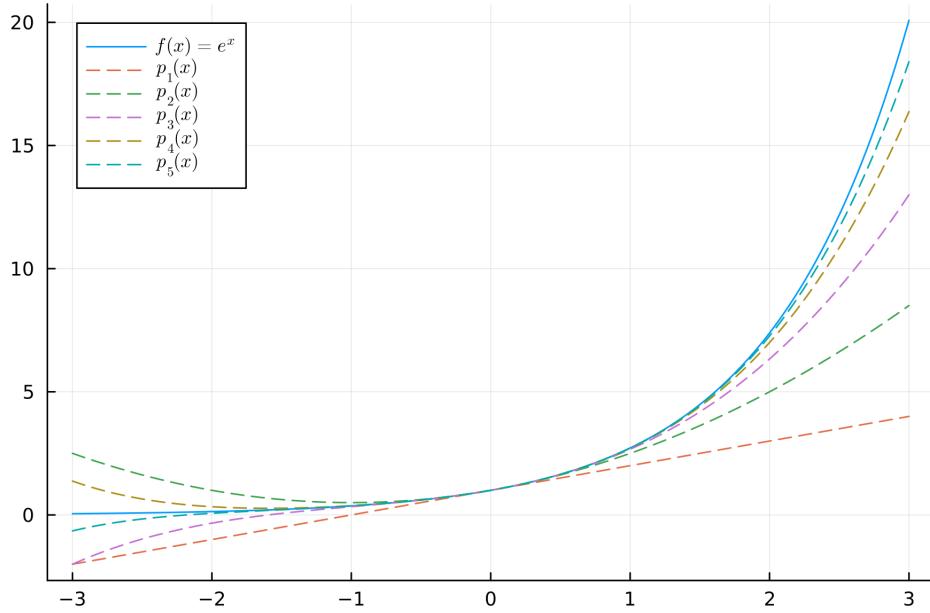
Consider the polynomial

$$p_n(x) = \sum_{k=0}^n \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \dots + \frac{x^n}{n!}$$

One can check that

$$p_n^{(k)}(0) = f^{(k)}(0) \quad \forall k = 0, 1, \dots, n$$

Hence, it is reasonable to approximate  $f(x)$  with  $p_n(x)$  near point  $x = 0$ . Figure 5.9 compares  $f$  with  $p_n$  with several choices of  $n$ .



**Figure 5.9:** Approximations of  $f$  at  $x = 0$ .

In general, a function  $f$  can be approximated by the polynomial

$$\sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k} (x - c)^k = f(c) + f'(c)(x - c) + \dots + \frac{f^{(n)}(c)}{n!} (x - c)^n$$

at point  $c$  provided that  $f$  has up to  $(n + 1)$ -th order derivative. The reason why we approximate  $f$

with a polynomial of  $n$  degree is that we have a remainder term, which makes use of the  $(n + 1)$ -th order derivative. The remainder tells us how accurate the approximation is. The formal statement is described in the following theorem, the **Taylor's theorem**.

### Theorem 5.6.1 (Taylor)

*Let  $f$  be a function having  $(n + 1)$ -th order derivative everywhere in  $(a, b)$ , and suppose  $f^{(n)}$  is continuous on  $[a, b]$ . Let  $c \in (a, b)$  be an interior point. Then for every  $x \in [a, b]$  other than  $c$ , we have*

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!} (x - c)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - c)^{n+1} \quad (5.17)$$

where  $\xi$  is some number in between  $x$  and  $c$ .



This theorem can be obtained as an immediate corollary of the following extension of the mean value theorem.

### Theorem 5.6.2

*Let  $f$  and  $g$  be two functions each having  $(n + 1)$ -th order derivative everywhere in  $(a, b)$ . Suppose that  $f^{(n)}$  and  $g^{(n)}$  are both continuous on  $[a, b]$ . Let  $c \in (a, b)$  be an interior point. Then for every  $x \in [a, b]$  other than  $c$ , we have*

$$\left[ f(x) - \sum_{k=0}^n \frac{f^{(k)}(c)}{k!} (x - c)^k \right] g^{(n+1)}(\xi) = f^{(n+1)}(\xi) \left[ g(x) - \sum_{k=0}^n \frac{g^{(k)}(c)}{k!} (x - c)^k \right] \quad (5.18)$$

where  $\xi$  is some number in between  $x$  and  $c$ .



**Remark** Taking  $g(x) = (x - c)^{n+1}$ , we have

$$g^{(n+1)}(\xi) = (n+1)! \quad \text{and} \quad \sum_{k=0}^n \frac{g^{(k)}(c)}{k!} (x - c)^k = 0$$

Then (5.18) reduces to (5.17), and hence this theorem reduces to Theorem 5.6.1.

**Proof** Define functions  $F(t)$  and  $G(t)$  on  $[a, b]$  by

$$F(t) = \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x - t)^k \quad \text{and} \quad G(t) = \sum_{k=0}^n \frac{g^{(k)}(t)}{k!} (x - t)^k \quad \forall t \in [a, b]$$

Note that  $F$  and  $G$  are continuous on  $[a, b]$ , and they have derivatives everywhere in  $(a, b)$ . Applying the generalized mean value theorem (Theorem 5.4.5) to  $F(t)$  and  $G(t)$  on the closed interval joining points  $x$  and  $c$ , we have

$$[F(x) - F(c)]G'(\xi) = F'(\xi)[G(x) - G(c)] \quad (5.19)$$

for some number  $\xi$  in between  $x$  and  $c$ .

After a few steps of computation, we obtain derivatives  $F'(t)$  and  $G'(t)$  as follows.

$$F'(t) = \frac{f^{(n+1)}(t)}{n!} (x - t)^n \quad \text{and} \quad G'(t) = \frac{g^{(n+1)}(t)}{n!} (x - t)^n \quad \forall t \in (a, b) \quad (5.20)$$

We also note that

$$F(x) = f(x), \quad F(c) = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!}(x-c)^k, \quad G(x) = g(x), \quad \text{and} \quad G(c) = \sum_{k=0}^n \frac{g^{(k)}(c)}{k!}(x-c)^k \quad (5.21)$$

Plugging (5.20) and (5.21) into (5.19) leads to (5.18). ■

We can rewrite (5.17) as

$$f(x) = p_n(x) + r_n(x)$$

where

$$p_n(x) = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!}(x-c)^k$$

is the polynomial approximation, and

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

is the remainder. The smaller the absolute value of  $r_n(x)$  is, the more accurate the approximation will be.

**Exercise 5.7** Use Taylor's theorem (Theorem 5.6.1) to approximate the value of  $\ln(0.6)$  so that the absolute error is less than 0.1.

**Solution** Putting  $f(x) = \ln(x)$  in (5.17), we have

$$\begin{aligned} p_n(x; c) &= \sum_{k=0}^n \frac{f^{(k)}(c)}{k!}(x-c)^k = \ln(c) + \sum_{k=1}^n \frac{(-1)^{k+1}}{kc^n}(x-c)^k \\ r_n(x; c) &= \frac{(-1)^{n+2}}{(n+1)\xi^{n+1}}(x-c)^{n+1} \end{aligned}$$

Since the value of  $\ln(x)$  at  $x = 1$  is known, i.e.,  $\ln(1) = 0$ , we can expand it about that point by putting  $c = 1$ . Then we have

$$p_n(x) = \sum_{k=1}^n \frac{(-1)^{k+1}}{k}(x-1)^k \quad \text{and} \quad r_n(x) = \frac{(-1)^{n+2}}{(n+1)\xi^{n+1}}(x-1)^{n+1}$$

To approximate  $\ln(0.6)$  by  $p_n(0.6)$  within the error bounds, we need to estimate the remainder  $r_n(0.6)$ . Note that  $\xi$  is now some number in between 0.6 and 1 since  $x = 0.6$  and  $c = 1$ . The absolute value of  $r_n(0.6)$  is bounded above by

$$|r_n(0.6)| = \frac{1}{n+1} \left| \frac{0.6-1}{\xi} \right|^{n+1} < \frac{1}{n+1} \left( \frac{0.4}{0.6} \right)^{n+1} = \frac{1}{n+1} \left( \frac{2}{3} \right)^{n+1} \quad (5.22)$$

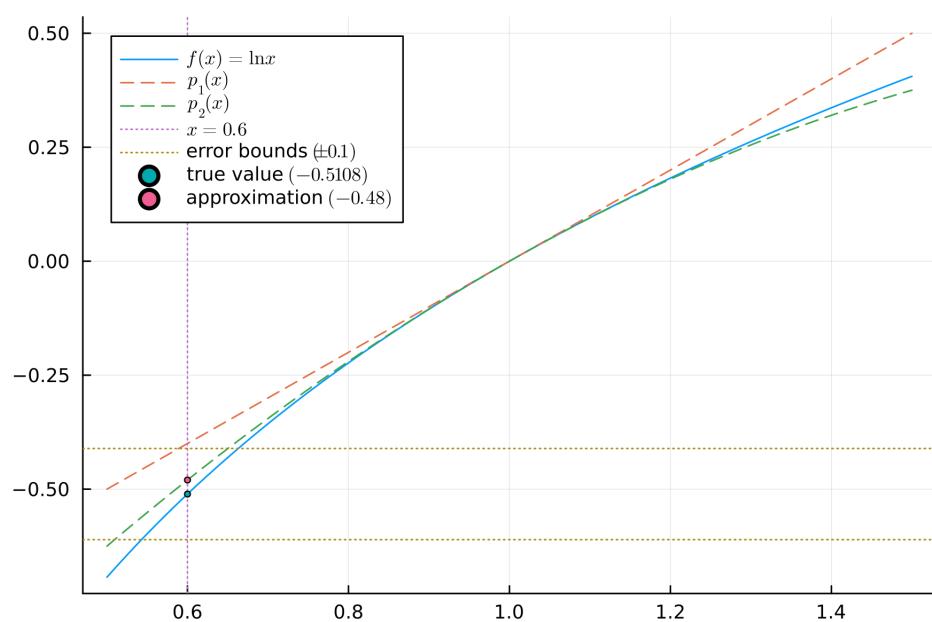
We then estimate  $|r_n(0.6)|$  using (5.22) by trying the first few values of  $n$ . We have

$$r_1(0.6) = \frac{2}{9} > 0.1 \quad \text{and} \quad r_2(0.6) = \frac{8}{81} < 0.1$$

Hence, it is adequate to approximate  $\ln(0.6)$  by  $p_2(0.6)$ . The approximated value is

$$\ln(0.6) \approx p_2(0.6) = (0.6-1) - \frac{1}{2} \times (0.6-1)^2 = -0.48$$

We can also visualize the result in Figure 5.10. As we can see, the approximated value is quite close to the true value. And the absolute error is actually much less than the required bound, 0.1.



**Figure 5.10:** Approximation of  $\ln(0.6)$ .

# Chapter 6 Functions of Bounded Variation

## 6.1 Definition of Functions of Bounded Variation

### Definition 6.1.1

Let  $f : [a, b] \rightarrow \mathbb{R}$  be a function. For a partition  $P = \{x_0, x_1, \dots, x_n\}$  on  $[a, b]$ , write

$$V(P, f) := \sum_{k=1}^n |\Delta f_k| = \sum_{k=1}^n |f(x_k) - f(x_{k-1})|$$

If there exists  $M > 0$  such that

$$V(P, f) < M$$

holds for all partition  $P$ , then we say  $f$  is of bounded variation on  $[a, b]$ . 

### Theorem 6.1.1

If  $f$  is of bounded variation on  $[a, b]$ , then  $f$  is bounded on  $[a, b]$ . 

**Remark** The converse is not true. That is, if  $f$  is a bounded function, then it may not be of bounded variation (see Example 6.1).

**Proof** Pick any  $x \in [a, b]$ . Consider the partition  $P = \{a, b\} \cup \{x\}$ . Since  $f$  is of bounded variation, we have

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)| \leq V(P, f) < M$$

where  $M$  is a constant that is independent of  $x$ . It follows that

$$|f(x)| \leq |f(a)| + M \quad \forall x \in [a, b]$$

Therefore,  $f$  is bounded on  $[a, b]$ . 

**Example 6.1** Let  $f : [0, 1] \rightarrow \mathbb{R}$  be given by

$$f(x) = \begin{cases} x \cos \frac{\pi}{x} & 0 < x \leq 1 \\ 0 & x = 0 \end{cases}$$

Note that  $f$  is a bounded continuous function on  $[0, 1]$ . We now show that  $f$  is however not of bounded

variation. Consider a sequence of partitions

$$\begin{aligned} P_1 &= \left\{ 0, \frac{1}{3}, \frac{1}{2}, 1 \right\} \\ P_2 &= \left\{ 0, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1 \right\} \\ P_3 &= \left\{ 0, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1 \right\} \\ &\vdots \\ P_n &= \{0, 1\} \cup \bigcup_{k=1}^n \left\{ \frac{1}{2k+1}, \frac{1}{2k} \right\} \\ &\vdots \end{aligned}$$

For partition  $P_n$ , the sum

$$\begin{aligned} V(P, f) &> \sum_{k=1}^n \left| f\left(\frac{1}{2k}\right) - f\left(\frac{1}{2k+1}\right) \right| \\ &= \sum_{k=1}^n \left| \frac{1}{2k} \cos(2k\pi) - \frac{1}{2k+1} \cos((2k+1)\pi) \right| \\ &= \sum_{k=1}^n \left( \frac{1}{2k} + \frac{1}{2k+1} \right) \\ &= \sum_{k=2}^{2n+1} \frac{1}{k} \end{aligned}$$

Since the series  $\sum \frac{1}{k} = \infty$ , we can always find a partition such that the sum  $\sum |\Delta f_k|$  exceeds any given positive number. Therefore,  $f$  is not of bounded variation on  $[0, 1]$  though it is bounded.

Recall the definition of uniform continuity. The condition of uniform continuity, to some extent, eliminates the functions that tend to infinity at some points while the condition of bounded variation eliminates the functions that *oscillate*.

### Theorem 6.1.2

If  $f$  is a monotonic function on  $[a, b]$ , then  $f$  is of bounded variation on  $[a, b]$ .



**Proof** Without loss of generality, we assume  $f$  is increasing. For any partition  $P = \{x_0, \dots, x_n\}$  on  $[a, b]$ , we have

$$\sum_{k=1}^n |\Delta f_k| = \sum_{k=1}^n |f(x_k) - f(x_{k-1})| = \sum_{k=1}^n f(x_k) - f(x_{k-1}) = f(b) - f(a)$$

This completes the proof since  $f(b) - f(a)$  is independent of  $P$ . ■

## 6.2 Total Variation

### Definition 6.2.1

Let  $f$  be of bounded variation on  $[a, b]$ . The total variation of  $f$  on  $[a, b]$  is defined by

$$V_a^b(f) := \sup_{P \in \mathcal{P}[a,b]} V(P, f)$$

where  $\mathcal{P}[a, b]$  is the collection of all partitions on  $[a, b]$ .



**Note** We will sometimes write  $V(f)$  instead of  $V_a^b(f)$  to ease the notation if it does not confuse.

### Theorem 6.2.1

If  $f$  and  $g$  are each of bounded variation on  $[a, b]$ , then  $f \pm g$  and  $fg$  are also of bounded variation. Moreover,

$$V(f \pm g) \leq V(f) + V(g) \quad (6.1)$$

and

$$V(fg) \leq \sup_{x \in [a,b]} |g(x)| V(f) + \sup_{x \in [a,b]} |f(x)| V(g) \quad (6.2)$$



**Proof** (Proof of  $f \pm g$  being of bounded variation) We have

$$\begin{aligned} V(P, f \pm g) &= \sum_{k=1}^n |(f \pm g)(x_k) - (f \pm g)(x_{k-1})| \\ &= \sum_{k=1}^n |(f(x_k) - f(x_{k-1})) \pm (g(x_k) - g(x_{k-1}))| \\ &\leq \sum_{k=1}^n |\Delta f_k| + |\Delta g_k| \\ &= V(P, f) + V(P, g) \\ &\leq V(f) + V(g) \end{aligned}$$

Since  $f$  and  $g$  are both of bounded variation,  $V(f), V(g) < \infty$ . Therefore,  $V(P, f \pm g)$  is bounded above by a finite number for any partition  $P$ , which implies the function  $f \pm g$  is of bounded variation. Taking supremum over all  $P$ 's on

$$V(P, f \pm g) \leq V(f) + V(g)$$

we obtain (6.1).

(Proof of  $fg$  being of bounded variation) For any partition  $P$  on  $[a, b]$ , we have

$$\begin{aligned} V(P, fg) &= \sum_{k=1}^n |f(x_k)g(x_k) - f(x_{k-1})g(x_{k-1})| \\ &= \sum_{k=1}^n |f(x_k)g(x_k) - f(x_{k-1})g(x_k) + f(x_{k-1})g(x_k) - f(x_{k-1})g(x_{k-1})| \\ &= \sum_{k=1}^n |g(x_k)\Delta f_k + f(x_{k-1})\Delta g_k| \\ &\leq \sup |g| V(P, f) + \sup |f| V(P, g) \\ &\leq \sup |g| V(f) + \sup |f| V(g) \end{aligned}$$

Similarly, because  $f$  and  $g$  are of bounded variation,  $V(f)$  and  $V(g)$  are finite. Moreover, by Theorem 6.1.1,  $f$  and  $g$  are bounded on  $[a, b]$ , which implies  $\sup |f|, \sup |g| < \infty$ . Hence,  $V(P, fg)$  is bounded above by a finite number for any partition, which implies  $fg$  is also of bounded variation, and (6.2) holds.  $\blacksquare$

## 6.3 Additive Property of Total Variation

### Theorem 6.3.1

*Let  $f$  be of bounded variation on  $[a, b]$ . For a point  $c \in (a, b)$ , we have  $f$  is of bounded variation on  $[a, c]$  and  $[c, b]$ , and*

$$V_a^b(f) = V_a^c(f) + V_c^b(f) \quad (6.3)$$



**Proof** Let  $P'$  be a partition on  $[a, c]$  and  $P''$  a partition on  $[c, b]$ . Note that  $P = P' \cup P''$  forms a partition on  $[a, b]$ . We have

$$V_a^b(P, f) = V_a^c(P', f) + V_c^b(P'', f)$$

Since  $f$  is of bounded variation on  $[a, b]$ , it follows that

$$\infty > V_a^b(f) \geq V_a^b(P, f) = V_a^c(P', f) + V_c^b(P'', f) \quad (6.4)$$

Note that (6.4) holds for any partition  $P'$  on  $[a, c]$  and any partition  $P''$  on  $[c, b]$ . Therefore,  $f$  is of bounded variation on  $[a, c]$  and  $[c, b]$ , i.e.,  $V_a^c(f), V_c^b(f) < \infty$ .

Taking the supremum over all partitions on  $[a, c]$  followed by taking the supremum over all partitions on  $[a, c]$  on (6.4), we obtain

$$V_a^b(f) \geq \sup_{P' \in \mathcal{P}[a, c]} V_a^c(P', f) + \sup_{P'' \in \mathcal{P}[c, b]} V_c^b(P'', f) = V_a^c(f) + V_c^b(f) \quad (6.5)$$

Now, let  $P$  be a partition on  $[a, b]$ . Let

$$P' = (P \cap [a, c]) \cup \{c\} \quad P'' = P \cap [c, b] \cup \{c\}$$

There are two cases. If  $c \in P$ , then

$$V_a^b(P, f) = V_a^c(P', f) + V_c^b(P'', f)$$

If  $c \notin P$ , then

$$V_a^b(P, f) \leq V_a^c(P', f) + V_c^b(P'', f)$$

Either way, it holds that

$$V_a^b(P, f) \leq V_a^c(P', f) + V_c^b(P'', f) \leq V_a^c(f) + V_c^b(f) \quad (6.6)$$

Taking the supremum over all partitions on  $[a, b]$  on (6.6), we have

$$V_a^b(f) = \sup_{P \in \mathcal{P}[a,b]} V_a^b(P, f) \leq V_a^c(f) + V_c^b(f) \quad (6.7)$$

Equation (6.3) then follows from (6.4) and (6.7). ■

The total variation of a function with the same lower and upper limits is defined as zero, i.e.,  $V_x^x(f) = 0$  so that (6.3) holds for  $c = a$  and  $c = b$ .

### Definition 6.3.1

Suppose that  $f$  is of bounded variation on  $[a, b]$ . The total variation

$$V_a^x(f) \quad a \leq x \leq b$$

can be regarded as a function of  $x$  on  $[a, b]$  with an addition definition  $V_a^a(f) := 0$ . ♣

## 6.4 Characterization of Functions of Bounded Variation

In fact, all functions of bounded variation can be written as a difference between two increasing functions.

### Lemma 6.4.1

Let  $f$  be of bounded variation on  $[a, b]$ . Define  $V(x) := V_a^x(f)$  ( $a \leq x \leq b$ ). Then

- 1.  $V$  is increasing on  $[a, b]$
  - 2.  $V - f$  is increasing on  $[a, b]$
- ♡

**Proof** (Proof of 1) By Theorem 6.3.1 and Definition 6.3.1, we know  $V(x)$  is well-defined, and

$$V(x+h) - V(x) = V_x^{x+h}(f) \geq 0$$

where  $h > 0$ . Therefore,  $V$  is indeed increasing on  $[a, b]$ .

(Proof of 2) Fix  $x \in [a, b]$ . Let  $h$  be such that  $0 \leq x < x+h \leq b$ . Let a partition  $P$  on  $[x, x+h]$  be given by

$$P = \{x, x+h\}$$

We have

$$V_x^{x+h}(P, f) = |f(x+h) - f(x)| \leq V_x^{x+h}(f) \quad (6.8)$$

It then follows from (6.8) that

$$\begin{aligned} (V - f)(x+h) - (V - f)(x) &= V_x^{x+h}(f) - (f(x+h) - f(x)) \\ &\geq V_x^{x+h}(f) - |f(x+h) - f(x)| \\ &\geq 0 \end{aligned}$$

Hence,  $V - f$  is also an increasing function on  $[a, b]$ . ■

**Theorem 6.4.1 (Characterization of Functions of Bounded Variation)**

Let  $f$  be a real-valued function on  $[a, b]$ . Then, the following statements are equivalent.

1.  $f$  is of bounded variation on  $[a, b]$ .
2. There exist two increasing functions  $g$  and  $h$  on  $[a, b]$  such that  $f = g - h$ .
3. There exist two **strictly** increasing functions  $g$  and  $h$  on  $[a, b]$  such that  $f = g - h$ .



# Chapter 7 Riemann-Stieltjes Integral

## 7.1 Definition of the Riemann-Stieltjes Integral

### Definition 7.1.1

Suppose that  $f$  and  $\alpha$  are **real-valued bounded functions** on  $[a, b]$ . Let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition on  $[a, b]$  and  $t_k$  a point in the sub-interval  $[x_{k-1}, x_k]$ . A sum of the form

$$S(P, f, \alpha) = \sum_{k=1}^n f(t_k) \Delta \alpha_k$$

is called a **Riemann-Stieltjes sum** of  $f$  with respect to  $\alpha$ . We say  $f$  is **Riemann-integrable** with respect to  $\alpha$ , and write  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  if there exists a number  $A$  having the following property: for any given  $\varepsilon > 0$ , there exists a partition  $P_\varepsilon$  such that

$$|S(P, f, \alpha) - A| < \varepsilon$$

for any refinement  $P$  of  $P_\varepsilon$  and for any choice of points  $t_k$ . (Note that  $S(P, f, \alpha)$  depends on  $t_k$ .) Moreover, the number  $A$  is **uniquely determined** if it exists (this is proved in Proposition 7.1.1) and is denoted by

$$\int_a^b f \, d\alpha$$



**Note** To ease the nation, we will sometimes write

$$\sum_k f(t_k) \Delta \alpha_k$$

without specifying the lower and upper limits of  $k$  if no confusion will be caused.

Recall Definition 7.1.1 only requires the existence of  $A$ . We now show that such number  $A$  is also unique.

### Proposition 7.1.1

Let  $S(P, f, \alpha)$  be as in Definition 7.1.1. If  $A$  and  $A'$  both satisfy the property stated in Definition 7.1.1, then  $A = A'$ .



**Proof** Given  $\varepsilon > 0$ , by the property in Definition 7.1.1, there exist partitions  $P_1$  and  $P_2$  such that

$$|S(P_1, f, \alpha) - A| < \varepsilon/2 \quad |S(P_2, f, \alpha) - A'| < \varepsilon/2$$

Let  $P = P_1 \cup P_2$ . We have

$$|S(P, f, \alpha) - A| < \varepsilon/2 \quad |S(P, f, \alpha) - A'| < \varepsilon/2$$

It then follows that

$$|A - A'| \leq |S(P, f, \alpha) - A| + |S(P, f, \alpha) - A'| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

Since  $\varepsilon > 0$  is arbitrary, we must have  $A = A'$ . ■

## 7.2 Linear Properties

The integral is linear in the integrand. In other words, the integral of a linear combination of functions is equal to the linear combination of integrals of each function.

### Theorem 7.2.1

Suppose that  $f, g \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . Then  $c_1f + c_2g \in \mathfrak{R}(\alpha)$  on  $[a, b]$  where  $c_1$  and  $c_2$  are constants. In that case,

$$\int_a^b c_1f + c_2g \, d\alpha = c_1 \int_a^b f \, d\alpha + c_2 \int_a^b g \, d\alpha$$



**Proof** Let  $P$  be a partition on  $[a, b]$ . The Riemann-Stieltjes sum of  $c_1f + c_2g$  can be written as

$$\begin{aligned} S(P, c_1f + c_2g, \alpha) &= \sum_k c_1f(t_k) + c_2g(t_k)\Delta\alpha_k \\ &= c_1 \sum_k f(t_k)\Delta\alpha_k + c_2 \sum_k g(t_k)\Delta\alpha_k \\ &= c_1S(P, f, \alpha) + c_2S(P, g, \alpha) \end{aligned}$$

Given  $\varepsilon > 0$ . Since  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  then there exists a partition  $P'_\varepsilon$  such that

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \frac{\varepsilon/2}{1 + |c_1|} \quad \forall P \supseteq P'_\varepsilon$$

Similarly, since  $g \in \mathfrak{R}(\alpha)$ , there exists a partition  $P''_\varepsilon$  such that

$$\left| S(P, g, \alpha) - \int_a^b g \, d\alpha \right| < \frac{\varepsilon/2}{1 + |c_2|} \quad \forall P \supseteq P''_\varepsilon$$

Let  $P_\varepsilon$  be the refinement of  $P'_\varepsilon$  and  $P''_\varepsilon$ , i.e.,  $P = P'_\varepsilon \cup P''_\varepsilon$ . Then for any  $P \supseteq P_\varepsilon$ , we have

$$\begin{aligned} &\left| S(P, c_1f + c_2g, \alpha) - c_1 \int_a^b f \, d\alpha - c_2 \int_a^b g \, d\alpha \right| \\ &\leq |c_1| \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| + |c_2| \left| S(P, g, \alpha) - \int_a^b g \, d\alpha \right| \\ &< |c_1| \frac{\varepsilon/2}{1 + |c_1|} + |c_2| \frac{\varepsilon/2}{1 + |c_2|} \\ &< \varepsilon \end{aligned}$$

This completes the proof. ■

The integral is also linear in the integrator.

### Theorem 7.2.2

If  $f \in \mathfrak{R}(\alpha)$  and  $f \in \mathfrak{R}(\beta)$  on  $[a, b]$ , then  $f \in \mathfrak{R}(c_1\alpha + c_2\beta)$  where  $c_1$  and  $c_2$  are constants. In that case,

$$\int_a^b f \, d(c_1\alpha + c_2\beta) = c_1 \int_a^b f \, d\alpha + c_2 \int_a^b f \, d\beta$$



**Proof** Let  $P$  be a partition on  $[a, b]$ . We have

$$\begin{aligned} S(P, f, c_1\alpha + c_2\beta) &= \sum_{k=0} f(t_k) + \Delta(c_1\alpha + c_2\beta)_k \\ &= c_1 \sum_{k=0} f(t_k)\Delta\alpha_k + c_2 \sum_{k=0} f(t_k)\Delta\beta_k \\ &= c_1 S(P, f, \alpha) + c_2 S(P, f, \beta) \end{aligned}$$

Given  $\varepsilon > 0$ . Since  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  then there exists a partition  $P'_\varepsilon$  such that

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \frac{\varepsilon/2}{1 + |c_1|} \quad \forall P \supseteq P'_\varepsilon$$

Similarly, since  $f \in \mathfrak{R}(\beta)$ , there exists a partition  $P''_\varepsilon$  such that

$$\left| S(P, f, \beta) - \int_a^b f \, d\beta \right| < \frac{\varepsilon/2}{1 + |c_2|} \quad \forall P \supseteq P''_\varepsilon$$

Let  $P = P'_\varepsilon \cup P''_\varepsilon$ . Then for any  $P \supseteq P_\varepsilon$ , we have

$$\begin{aligned} &\left| S(P, f, c_1\alpha + c_2\beta) - c_1 \int_a^b f \, d\alpha - c_2 \int_a^b f \, d\beta \right| \\ &\leq |c_1| \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| + |c_2| \left| S(P, f, \beta) - \int_a^b f \, d\beta \right| \\ &< |c_1| \frac{\varepsilon/2}{1 + |c_1|} + |c_2| \frac{\varepsilon/2}{1 + |c_2|} \\ &< \varepsilon \end{aligned}$$

■

If we divide the interval  $[a, b]$  into two parts with some point  $c \in (a, b)$  in the middle, then the integral over the entire interval is the sum of the integrals on these two sub-intervals. This is also a kind of linearity of integrals considering the interval of integration.

### Lemma 7.2.1

Suppose  $c \in (a, b)$ . We have

$$\int_a^b f \, d\alpha = \int_a^c f \, d\alpha + \int_c^b f \, d\alpha \quad (7.1)$$

The existence of two integrals in (7.1) will imply the existence of the third one.



**Proof** We have

$$S(P, f, \alpha) = S(P', f, \alpha) + S(P'', f, \alpha) \quad \forall P = P' \cup P'' \quad (7.2)$$

where  $P$ ,  $P'$  and  $P''$  are partitions on  $[a, b]$ ,  $[a, c]$  and  $[c, b]$ , respectively. Let  $\varepsilon > 0$  be chosen arbitrarily.

(Proof of existence of  $\int_a^b f \, d\alpha$ ) Assume  $\int_a^c f \, d\alpha$  and  $\int_c^b f \, d\alpha$  exist. Then

$$\left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| < \varepsilon/2 \quad \forall P' \supseteq P'_\varepsilon$$

for some  $P'_\varepsilon$  on  $[a, c]$ . And

$$\left| S(P'', f, \alpha) - \int_c^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P'' \supseteq P''_\varepsilon$$

for some  $P''_\varepsilon$  on  $[c, b]$ . Let

$$P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$$

(Note that  $c \in P_\varepsilon$ .) Let

$$P \supseteq P_\varepsilon \quad P' = P \cap [a, c] \quad P'' = P \cap [c, b]$$

Observe that

$$P' \supseteq P'_\varepsilon \quad P'' \supseteq P''_\varepsilon$$

It then follows from (7.2) that

$$\begin{aligned} & \left| S(P, f, \alpha) - \int_a^c f \, d\alpha - \int_c^b f \, d\alpha \right| \\ & \leq \left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| + \left| S(P'', f, \alpha) - \int_c^b f \, d\alpha \right| \\ & < \varepsilon/2 + \varepsilon/2 \\ & = \varepsilon \end{aligned}$$

Therefore,  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  and (7.1) holds.

(Proof of existence of  $\int_c^b f \, d\alpha$ ) Assume  $\int_a^b f \, d\alpha$  and  $\int_a^c f \, d\alpha$  exist. Then

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P \supseteq P_\varepsilon$$

for some  $P_\varepsilon$  on  $[a, b]$ . And

$$\left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| < \varepsilon/2 \quad \forall P' \supseteq P'_\varepsilon$$

for some  $P'_\varepsilon$  on  $[a, c]$ . Let

$$P''_\varepsilon = (P_\varepsilon \cup P'_\varepsilon) \cap [c, b]$$

Let

$$P'' \supseteq P''_\varepsilon \quad P' \supseteq (P_\varepsilon \cup P'_\varepsilon) \cap [a, c] \quad P = P' \cup P''$$

Observe that

$$P' \supseteq P'_\varepsilon \quad P = P' \cup P'' \supseteq (P_\varepsilon \cap [a, c]) \cup (P'_\varepsilon \cap [c, b]) = P_\varepsilon$$

It then follows from (7.2) that

$$\begin{aligned} & \left| S(P'', f, \alpha) - \int_a^b f \, d\alpha + \int_a^c f \, d\alpha \right| \\ & \leq \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| + \left| S(P', f, \alpha) - \int_a^c f \, d\alpha \right| \\ & < \varepsilon/2 + \varepsilon/2 \\ & = \varepsilon \end{aligned}$$

Therefore,  $f \in \mathfrak{R}(\alpha)$  on  $[c, b]$  and (7.1) holds.

(Proof of existence of  $\int_a^c f \, d\alpha$ ) Assume  $\int_a^b f \, d\alpha$  and  $\int_c^b f \, d\alpha$  exist. Then

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P \supseteq P_\varepsilon$$

for some  $P_\varepsilon$  on  $[a, b]$ . And

$$\left| S(P'', f, \alpha) - \int_c^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P'' \supseteq P_\varepsilon''$$

for some  $P_\varepsilon''$  on  $[c, b]$ . Let

$$P'_\varepsilon = (P_\varepsilon \cup P_\varepsilon'') \cap [a, c]$$

Let

$$P' \supseteq P'_\varepsilon \quad P'' \supseteq (P_\varepsilon \cup P_\varepsilon'') \cap [c, b] \quad P = P' \cup P''$$

Observe that

$$P'' \supseteq P_\varepsilon'' \quad P = P' \cup P'' \supseteq (P_\varepsilon \cap [a, c]) \cup (P_\varepsilon \cap [c, b]) = P_\varepsilon$$

It then follows from (7.2) that

$$\begin{aligned} & \left| S(P', f, \alpha) - \int_a^b f \, d\alpha + \int_c^b f \, d\alpha \right| \\ & \leq \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| + \left| S(P'', f, \alpha) - \int_c^b f \, d\alpha \right| \\ & < \varepsilon/2 + \varepsilon/2 \\ & = \varepsilon \end{aligned}$$

Therefore,  $f \in \mathfrak{R}(\alpha)$  on  $[a, c]$  and (7.1) holds. ■

### Theorem 7.2.3

*The existence of two integrals in (7.3) will imply the existence of the third one.*

$$\int_a^b f \, d\alpha = \int_a^c f \, d\alpha + \int_c^b f \, d\alpha \quad (7.3) \quad \heartsuit$$

## 7.3 Reduction to Riemann-Stieltjes Integrals

In the definition of Riemann-Stieltjes integral,  $d\alpha$  is merely a symbol which emphasizes that the integrator is  $\alpha$ . However, it is tempting to associate this symbol with the concept of differentials, and then replace it with  $\alpha'(x) dx$ . As stated in the following theorem, we can indeed do this under some conditions.

### Theorem 7.3.1

*Suppose  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , and  $\alpha$  has a continuous derivative  $\alpha'$  on  $[a, b]$ . Then the Riemann integral  $f\alpha'$  exists, i.e.,  $f\alpha' \in \mathfrak{R}$  on  $[a, b]$ , and we have*

$$\int_a^b f(x) \, d\alpha(x) = \int_a^b f(x)\alpha'(x) \, dx \quad (7.4) \quad \heartsuit$$

**Proof** We first do some preliminary work. Suppose  $f$  is bounded by a positive number  $M > 0$ , i.e.,  $|f(x)| \leq M \forall x \in [a, b]$ , and let  $\varepsilon > 0$  be chosen arbitrarily.

Then, we make use of the continuity of  $\alpha'$ . Note that  $\alpha'$  is uniformly continuity on  $[a, b]$  since the

closed interval is compact. Then there exists  $\delta > 0$  such that

$$|s - t| < \delta \implies |\alpha'(s) - \alpha'(t)| < \frac{\varepsilon/2}{M(b-a)} \quad (7.5)$$

Let  $P'_\varepsilon$  be a partition on  $[a, b]$  with mesh less than  $\delta$ , i.e.,  $\|P'\| < \delta$ .

Because  $f \in \alpha$ , there exists a partition, say  $P''_\varepsilon$ , such that

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/2 \quad \forall P \supseteq P''_\varepsilon \quad (7.6)$$

holds for any refinement  $P$ .

Let partition  $P_\varepsilon$  be the refinement of both partitions  $P'_\varepsilon$  and  $P''_\varepsilon$ , i.e.,  $P = P'_\varepsilon \cup P''_\varepsilon$ . Then for any refinement  $P$  of  $P_\varepsilon$ , we have

$$\begin{aligned} \left| S(P, f\alpha') - \int_a^b f \, d\alpha \right| &\leq |S(P, f\alpha') - S(P, f, \alpha)| + \left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| \\ &< |S(P, f\alpha') - S(P, f, \alpha)| + \varepsilon/2 \end{aligned} \quad (7.7)$$

where the last inequality follows from (7.6).

We now work on  $|S(P, f\alpha') - S(P, f, \alpha)|$ . We have

$$|S(P, f\alpha') - S(P, f, \alpha)| = \left| \sum_k f(t_k) \alpha'(t_k)(x_k - x_{k-1}) - \sum_k f(t_k) [\alpha(x_k) - \alpha(x_{k-1})] \right|$$

Applying the mean value theorem (Theorem 5.4.4) to  $\alpha$  on each sub-interval  $[x_{k-1}, x_k]$  leads to

$$\sum_k f(t_k) \alpha'(t_k)(x_k - x_{k-1}) - \sum_k f(t_k) [\alpha(x_k) - \alpha(x_{k-1})] = \sum_k f(t_k) [\alpha'(t_k) - \alpha'(\xi_k)](x_k - x_{k-1})$$

where  $\xi_k \in (x_{k-1}, x_k)$ . It then follows that

$$\begin{aligned} |S(P, f\alpha') - S(P, f, \alpha)| &= \left| \sum_k f(t_k) [\alpha'(t_k) - \alpha'(\xi_k)](x_k - x_{k-1}) \right| \\ &\leq \sum_k |f(t_k)| |\alpha'(t_k) - \alpha'(\xi_k)| \Delta x_k \\ &\quad \text{recall } |f| \leq M \\ &\leq M \sum_k |\alpha'(t_k) - \alpha'(\xi_k)| \Delta x_k \end{aligned}$$

Note that  $|t_k - \xi_k| \leq \|P\| \leq \|P'_\varepsilon\| < \delta$ . Hence, by (7.5), we have

$$|\alpha'(t_k) - \alpha'(\xi_k)| < \frac{\varepsilon/2}{M(b-a)}$$

Then,  $|S(P, f\alpha') - S(P, f, \alpha)|$  is finally bounded by

$$|S(P, f\alpha') - S(P, f, \alpha)| < M \frac{\varepsilon/2}{M(b-a)} \sum_k \Delta x_k = M \frac{\varepsilon/2}{M(b-a)} (b-a) = \varepsilon/2 \quad (7.8)$$

Combining (7.7) and (7.8), we obtain

$$\left| S(P, f\alpha') - \int_a^b f \, d\alpha \right| < \varepsilon$$

where  $P$  is an arbitrary refinement of the chosen  $P_\varepsilon$ . This implies  $f\alpha'$  is Riemann integrable, and (7.4) indeed holds. ■

## 7.4 Integration by Parts

### Theorem 7.4.1 (Integration by Parts)

If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then  $\alpha \in \mathfrak{R}(f)$  on  $[a, b]$ , and

$$\int_a^b f \, d\alpha + \int_a^b \alpha \, df = f(b)\alpha(b) - f(a)\alpha(a)$$



**Remark** This can be treated as the *reciprocity law* for integrals.

**Proof** Given  $\varepsilon > 0$ , since  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , there exists a partition  $P_\varepsilon$  such that

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon \quad \forall P \supseteq P_\varepsilon \quad (7.9)$$

Let  $P = \{x_0, x_1, \dots, x_n\} \supseteq P_\varepsilon$  be any refinement of  $P_\varepsilon$ . The Riemann-Stieltjes sum of  $\alpha$  with respect to  $f$  is

$$S(P, \alpha, f) = \sum_{k=1}^n \alpha(t_k)(f(x_k) - f(x_{k-1})) = \sum_{k=1}^n \alpha(t_k)f(x_k) - \sum_{k=1}^n \alpha(t_k)f(x_{k-1}) \quad (7.10)$$

Let

$$P^* = P \cup \{t_k \mid 1 \leq k \leq n\} \quad (7.11)$$

Denote by  $A$  the value

$$A = f(b)\alpha(b) - f(a)\alpha(a)$$

Note that  $A$  can be written as

$$A = \sum_{k=1}^n f(x_k)\alpha(x_k) - \sum_{k=1}^n f(x_{k-1})\alpha(x_{k-1}) \quad (7.12)$$

Subtracting (7.10) from (7.12), we obtain

$$A - S(P, \alpha, f) = \sum_{k=1}^n f(x_k)(\alpha(x_k) - \alpha(t_k)) + \sum_{k=1}^n f(x_{k-1})(\alpha(t_k) - \alpha(x_{k-1}))$$

By recalling the construction of  $P^*$  in (7.11), we observe that the right-hand side of the above equation is precisely the Riemann-Stieltjes sum  $S(P^*, f, \alpha)$ . That is,

$$A - S(P, \alpha, f) = S(P^*, f, \alpha)$$

Since  $P^* \supseteq P \supseteq P_\varepsilon$ , it follows from (7.9) that

$$\left| A - S(P, \alpha, f) - \int_a^b f \, d\alpha \right| = \left| S(P^*, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon$$

Recall  $A = f(b)\alpha(b) - f(a)\alpha(a)$ , we have

$$\left| S(P, \alpha, f) + \int_a^b f \, d\alpha - f(b)\alpha(b) + f(a)\alpha(a) \right| < \varepsilon \quad \forall P \supseteq P_\varepsilon$$

This implies that  $\alpha \in \mathfrak{R}(f)$  on  $[a, b]$ , and

$$\int_a^b \alpha \, df = - \int_a^b f \, d\alpha + f(b)\alpha(b) - f(a)\alpha(a)$$



☞ **Exercise 7.1** Evaluate the integral

$$\int_a^b x \cos(x) \, dx$$

## 7.5 Change of Variables

### Theorem 7.5.1

Suppose  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , and  $\phi$  is a **strictly monotonic continuous function** on a closed interval  $I$  with endpoints  $c$  and  $d$ . ( $I$  is either  $[c, d]$  or  $[d, c]$ .) Assume

$$a = \phi(c) \quad b = \phi(d)$$

Define two composite functions:

$$h = f(\phi(x)) \quad \beta = \alpha(\phi(x))$$

Then  $h \in \mathfrak{R}(\beta)$  on  $I$ , and  $\int_a^b f \, d\alpha = \int_c^d h \, d\beta$ , i.e.,

$$\int_a^b f(x) \, d\alpha(x) = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(x)) \, d\alpha(\phi(x))$$



**Remark** The reason why we assume that  $\phi$  is strictly monotonic and continuous is to ensure that it is a bijective function. It is equivalent to assuming  $\phi$  is strictly monotonic and injective.

The main idea of this proof is exploiting the one-to-one relation between the partitions on  $[a, b]$  and  $[c, d]$ .

**Proof** (One-To-One Relation of Partitions) Without loss of generality, we may assume that  $g$  is strictly *increasing* and continuous. Then  $I = [c, d]$ . From the conditions of  $g$ , we can immediately conclude that it has a bijective inverse function

$$\phi^{-1} : [a, b] \rightarrow [c, d]$$

For any partition  $P' = \{x_0, \dots, x_n\}$  on  $[a, b]$ , we can associate it with a partition  $P$  on  $[c, d]$ , which is given by

$$P := \phi^{-1}(P') := \{\phi^{-1}(x_0), \dots, \phi^{-1}(x_n)\}$$

On the other hand, for any partition  $P = \{y_0, \dots, y_n\}$  on  $[c, d]$ , we can define a partition  $P'$  on  $[a, b]$  by

$$P' := \phi(P) := \{\phi(y_0), \dots, \phi(y_n)\}$$

(Existence of the Integral) Given  $\varepsilon > 0$ , since  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , there exists a partition  $P'_\varepsilon$  on  $[a, b]$  such that

$$\left| S(P', f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon \quad \forall P' \supseteq P'_\varepsilon \tag{7.13}$$

Let partition  $P_\varepsilon$  on  $[c, d]$  be given by  $P_\varepsilon = \phi^{-1}(P'_\varepsilon)$ . For any refinement  $P \supseteq P_\varepsilon$ , we have

$$S(P, h, \beta) = \sum h(s_i) \Delta \beta_i = \sum h(s_i) (\beta(s_i) - \beta(s_{i-1}))$$

For each point  $s_i$ , we can map it to  $[a, b]$  by  $t_i = \phi(s_i)$ . It then follows that

$$\begin{aligned} S(P, h, \beta) &= \sum h(\phi^{-1}(t_i))(\beta(\phi^{-1}(t_i)) - \beta(\phi^{-1}(t_{i-1}))) \\ &= \sum f(t_i)(\alpha(t_i) - \alpha(t_{i-1})) \\ &= S(P', f, \alpha) \end{aligned}$$

where  $P' = \phi(P)$ . In summary,

$$S(P, h, \beta) = S(P', \phi, \alpha) \quad (7.14)$$

What is left to show is that  $P' \supseteq P'_\varepsilon$ . For any point  $x \in P'_\varepsilon$ , we have  $\phi^{-1}(x) \in P_\varepsilon$  since  $P_\varepsilon = \phi^{-1}(P'_\varepsilon)$ .

Recall that  $P \supseteq P_\varepsilon$ . Thus,  $\phi^{-1}(x) \in P$ . And because  $P' = \phi(P)$ , we have  $x = \phi(\phi^{-1}(x)) \in P'$ . Therefore, indeed  $P' \supseteq P'_\varepsilon$ . It then follows from (7.13) and (7.14) that

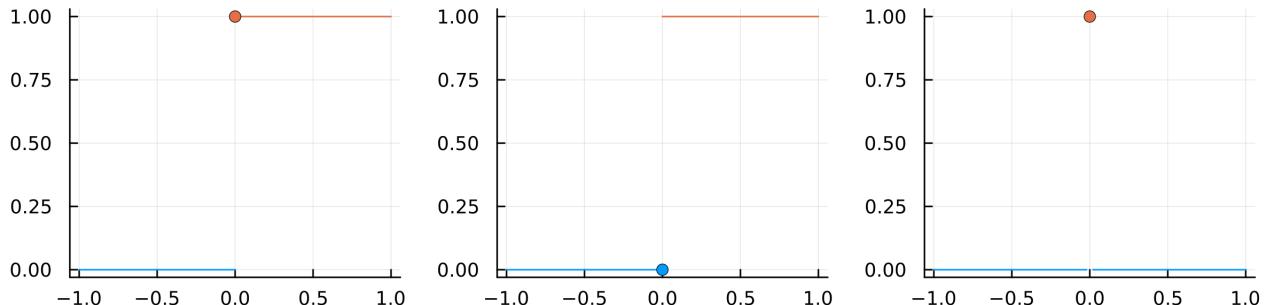
$$\left| S(P, h, \beta) - \int_a^b f \, d\alpha \right| < \varepsilon \quad \forall P \supseteq P_\varepsilon$$

This completes the proof. ■

## 7.6 Step Functions as Integrators

If the integrator  $\alpha$  is a constant function on  $[a, b]$ , then the  $f$  is always integrable with  $\int_a^b f \, d\alpha = 0$  since the Stieltjes sum  $S(P, f, \alpha) = 0$  for any partition.

What if  $\alpha$  is constant except for one jump or removable discontinuity at  $c \in (a, b)$ ? For example,  $\alpha$  may look like functions shown in Figure 7.1.



**Figure 7.1:** Left:  $\alpha(x)$  is continuous from the right at  $x = 0$ . Middle:  $\alpha(x)$  is continuous from the left at  $x = 0$ . Right:  $\alpha(x)$  has a removable discontinuity at  $x = 0$ .

In this case,  $f$  may not be even integrable, and if it is,  $\int_a^b f \, d\alpha$  may not be zero.

### Theorem 7.6.1

Let  $c \in (a, b)$ . Define a function  $\alpha$  on  $[a, b]$  as follows. The values of  $\alpha(a)$ ,  $\alpha(b)$  and  $\alpha(c)$  are chosen arbitrarily, and elsewhere

$$\alpha(x) = \begin{cases} \alpha(a) & a \leq x < c \\ \alpha(b) & c < x \leq b \end{cases}$$

Suppose  $f$  is a bounded function defined on  $[a, b]$  in such a way that

1. at least one of  $f$  and  $\alpha$  is continuous from the left at  $c$ , and

2. at least one of  $f$  and  $\alpha$  is continuous from the right at  $c$ .

Then,  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  with

$$\int_a^b f \, d\alpha = f(c)[\alpha(b) - \alpha(a)] \quad (7.15)$$



**Proof** (Possible expressions of  $S(P, f, \alpha)$ ) We first investigate the expression of the Stieltjes sum  $S(P, f, \alpha)$ . Given any partition  $P$  on  $[a, b]$ . If  $c \notin P$ , then we have

$$S(P, f, \alpha) = f(c)[\alpha(b) - \alpha(a)]$$

And if  $c \in P$ , we have

$$S(P, f, \alpha) = f(c - \delta_1)[\alpha(c) - \alpha(a)] + f(c + \delta_2)[\alpha(b) - \alpha(c)]$$

where  $c - \delta_1$  and  $c + \delta_2$  ( $\delta_1, \delta_2 \geq 0$ ) are two points in the adjacent sub-intervals sharing the same endpoint  $c$ . Note that  $\delta_1$  and  $\delta_2$  can not be too large. In fact, they satisfy that

$$0 \leq \delta_1, \delta_2 \leq \|P\|$$

Taking into consideration both cases, we have the following inequality for any partition  $P$ :

$$\begin{aligned} |S(P, f, \alpha) - f(c)[\alpha(b) - \alpha(a)]| &\leq |f(c - \delta_1) - f(c)| |\alpha(c) - \alpha(a)| + |f(c + \delta_2) - f(c)| |\alpha(b) - \alpha(c)| \end{aligned} \quad (7.16)$$

where of course  $0 \leq \delta_1, \delta_2 \leq \|P\|$ .



**Note** In fact, the left-hand side of (7.16) reduces to zero when  $c \notin P$  and the equality holds if  $c \in P$ .

There are altogether four scenarios concerning the continuity of  $f$  and  $\alpha$  at point  $c$ , which are

1.  $\alpha(x)$  is continuous at  $x = c$ .
2.  $\alpha(x)$  is continuous from the right at  $x = c$ , and  $f(x)$  is continuous from the left at  $x = c$ .
3.  $\alpha(x)$  is continuous from the left at  $x = c$ , and  $f(x)$  is continuous from the right at  $x = c$ .
4.  $f(x)$  is continuous at  $x = c$ .

(Proof of 1) If  $\alpha(x)$  is continuous at  $x = c$ , then  $\alpha$  is a constant function. In this case,  $f$  is certainly integrable with respect to  $\alpha$  with  $\int_a^b f \, d\alpha = 0$ . And (7.15) holds since its right-hand side is also zero for  $\alpha(a) = \alpha(b)$ .

(Proof of 2) Because  $\alpha(x)$  is continuous from the right at  $x = c$ , we have  $\alpha(c) = \alpha(b)$ , and the inequality (7.16) reduces to

$$|S(P, f, \alpha) - f(c)[\alpha(b) - \alpha(a)]| \leq |f(c - \delta_1) - f(c)| |\alpha(c) - \alpha(a)| \quad (7.17)$$

Given  $\varepsilon > 0$ , since  $f(x)$  is continuous from the left at  $x = c$ , there exists  $\delta > 0$  such that

$$-\delta < x - c < 0 \implies |f(x) - f(c)| < \frac{\varepsilon}{|\alpha(c) - \alpha(a)| + 1} \quad (7.18)$$

Choose a partition  $P_\varepsilon$  on  $[a, b]$  with  $\|P_\varepsilon\| < \delta$ . Let  $P \supseteq P_\varepsilon$  be an arbitrary refinement. Applying (7.17) to  $P$ , we have  $\delta_1 < \|P\| \leq \|P_\varepsilon\| < \delta$ . We may then apply (7.18), and obtain

$$\begin{aligned} |S(P, f, \alpha) - f(c)[\alpha(b) - \alpha(a)]| &\leq |f(c - \delta_1) - f(c)| |\alpha(c) - \alpha(a)| \\ &< \frac{\varepsilon}{|\alpha(c) - \alpha(a)| + 1} |\alpha(c) - \alpha(a)| < \varepsilon \end{aligned}$$

This implies  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , and (7.15) holds.

(Proof of 3) The proof, in this case, is similar to that of 2.

(Proof of 4) Given  $\varepsilon > 0$ . Because  $f(x)$  is continuous at  $x = c$ , there exists  $\delta > 0$  such that

$$|x - c| < \delta \implies |f(x) - f(c)| < \frac{\varepsilon}{|\alpha(c) - \alpha(a)| + |\alpha(b) - \alpha(c)| + 1} \quad (7.19)$$

Choose a partition  $P_\varepsilon$  with  $\|P_\varepsilon\| < \delta$ . Let  $P \supseteq P_\varepsilon$ . Applying (7.16) to  $P$ , we have  $0 \leq \delta_1, \delta_2 \leq \|P\| \leq \|P_\varepsilon\| < \delta$ . Then (7.16) and (7.19) yield

$$\begin{aligned} |S(P, f, \alpha) - f(c)[\alpha(b) - \alpha(a)]| \\ &< \frac{\varepsilon}{|\alpha(c) - \alpha(a)| + |\alpha(b) - \alpha(c)| + 1} (|\alpha(c) - \alpha(a)| + |\alpha(b) - \alpha(c)|) < \varepsilon \end{aligned}$$

Hence,  $f \in \mathfrak{R}(\alpha)$ , and (7.15) holds. This completes the proof.  $\blacksquare$

It should be emphasized that there are cases where  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  while the conditions in the above theorem are not satisfied.

**Example 7.1** Let

$$f(x) = \alpha(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0, \quad x \in [-1, 1] \\ 1 & x > 0 \end{cases}$$

Note that neither  $f$  nor  $\alpha$  is continuous from the left at 0. (And also neither  $f$  nor  $\alpha$  is continuous from the right at 0.) Hence, the conditions in Theorem 7.6.1 are not satisfied.

However, we now show that  $f$  is integrable. Let  $P$  be any partition on  $[-1, 1]$ . If  $0 \notin P$ , then

$$S(P, f, \alpha) = f(0)[\alpha(1) - \alpha(-1)] = 0$$

And if  $0 \in P$ , we then have

$$S(P, f, \alpha) = f(-\delta_1)[\alpha(0) - \alpha(-1)] + f(\delta_2)[\alpha(1) - \alpha(0)] = (-1)[0 - (-1)] + 1 \cdot [1 - 0] = 0$$

where  $\delta_1, \delta_2 \geq 0$ . Hence,  $S(P, f, \alpha) = 0$  no matter what partition and representative points we choose. It follows that  $f \in \mathfrak{R}(\alpha)$  on  $[-1, 1]$  with  $\int_{-1}^1 f \, d\alpha = 0$ .

**Example 7.2** Let

$$f(x) = \begin{cases} \frac{|\sin(x)|}{x} & x \neq 0 \\ -1 & x = 0 \end{cases} \quad \text{and} \quad \alpha(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}, \quad x \in [-3\pi, 3\pi]$$

Let us take for granted that the following limit exists:

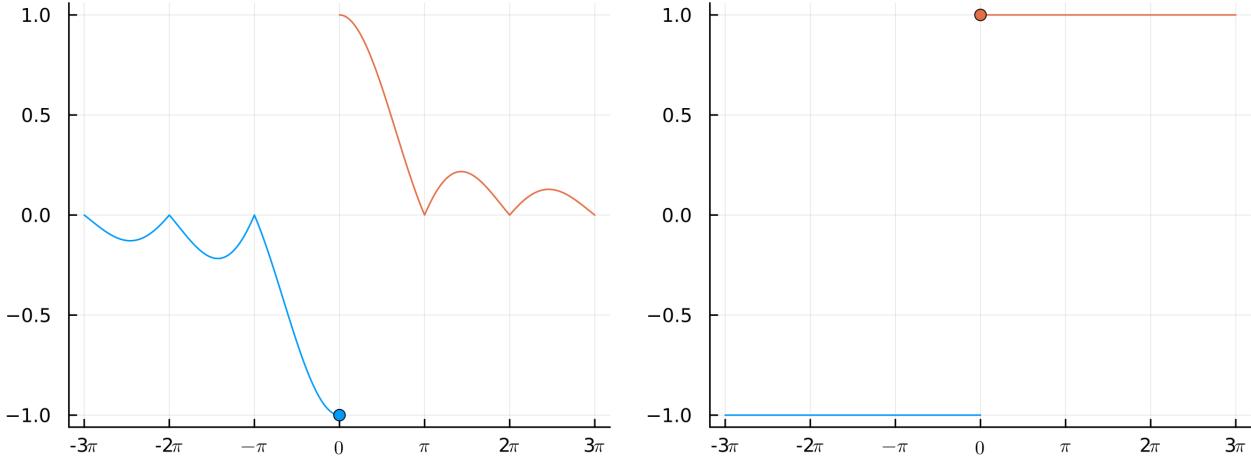
$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1$$

Then we note that the left and right limits of  $f$  both exist but are not equal. Figure 7.2 exhibits the graphs of these two functions.

The integrator  $\alpha$  is constant except for a jump discontinuity at 0. We note that  $\alpha$  is continuous from the *right* at 0, and  $f$  is continuous from the *left* at 0. Hence, the conditions in Theorem 7.6.1 are satisfied, and we conclude  $f \in \mathfrak{R}(\alpha)$  on  $[-3\pi, 3\pi]$  with

$$\int_{-3\pi}^{3\pi} f \, d\alpha = f(0)[\alpha(3\pi) - \alpha(-3\pi)] = f(0) \cdot 2 = -2$$

**Example 7.3** If we alter the function  $f$  in the example above (Example 7.2) by re-defining  $f(0) = 1$



**Figure 7.2:** Left: integrand  $f$  is continuous from the left at 0. Right: integrator  $\alpha$  is continuous from the right at 0.

while keeping the integrator  $\alpha$  unchanged, then  $f$  is no longer integrable w.r.t.  $\alpha$ . Now,

$$f(x) = \begin{cases} \frac{|\sin(x)|}{x} & x \neq 0 \\ 1 & x = 0 \end{cases} \quad \text{and} \quad \alpha(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}, \quad x \in [-3\pi, 3\pi]$$

Note that we cannot apply Theorem 7.6.1 here since neither the integrand nor the integrator is continuous from the left of 0.

To show that  $f$  is indeed not integrable, we consider  $f$  and  $\alpha$  on  $[-3\pi, 0]$  and  $[0, 3\pi]$  separately. It is evident that  $f \in \mathfrak{R}(\alpha)$  on  $[0, 3\pi]$  with  $\int_0^{3\pi} f \, d\alpha = 0$ . We now consider the other half interval  $[-3\pi, 0]$ . Given any partition  $P = \{x_0, x_1, \dots, x_n\}$  on  $[-3\pi, 0]$ , we have

$$S(P, f, \alpha) = f(t_n)[\alpha(x_n) - \alpha(x_{n-1})] = 2f(t_n) = \begin{cases} 2 |\sin(t_n)| / t_n & t_n < 0 \\ 2 & t_n = 0 \end{cases}$$



**Note** Intuitively, we know  $f$  is not integrable w.r.t.  $\alpha$  on  $[-3\pi, 0]$  since for any partition  $P$ , we can always choose the last point  $t_n$  such that the Stieltjes sum  $S(P, f, \alpha)$  does not fall in any neighborhood of some number. But we still have to prove this rigorously by the definition.

We shall prove by contradiction. Assume  $f \in \mathfrak{R}(\alpha)$  on  $[-3\pi, 0]$  with  $\int_{-3\pi}^0 f \, d\alpha = A$ . Since  $2 |\sin x| / x \rightarrow -2$  as  $x \rightarrow 0^-$ , there exists  $\delta > 0$  such that

$$-3 < \frac{2 |\sin x|}{x} < -1 \quad \forall x \in (-\delta, 0)$$

Choose  $\varepsilon = 1$ . For all partitions  $P_\varepsilon$ , we can always find a refinement  $P = \{x_0, x_1, \dots, x_n\}$  such that  $x_{n-1} > -\delta$ . We choose the last representative  $t_n \in [x_{n-1}, x_n]$  in two ways. Letting  $t_n$  be some number in that sub-interval other than 0 yields a Riemann-Stieltjes sum in between  $-3$  and  $-1$ , i.e.,

$$S(P, f, \alpha) = 2 |\sin(t_n)| / t_n \in (-3, -1)$$

since  $-\delta < t_n < 0$ . On the other hand, letting  $t'_n = 0$  leads to another Riemann-Stieltjes sum

$$S'(P, f, \alpha) = 2$$

We have the following inequality:

$$|S(P, f, \alpha) - A| + |S'(P, f, \alpha) - A| \geq |S(P, f, \alpha) - S'(P, f, \alpha)| > 3 > 2\epsilon$$

This implies that at least one of  $|S(P, f, \alpha) - A|$  and  $|S'(P, f, \alpha) - A|$  is greater than  $\epsilon$ , which leads to a contradiction. Hence,  $f$  is not integrable on  $[-3\pi, 0]$ .

Finally, by Theorem 7.2.3,  $f$  cannot be integrable with respect to  $\alpha$  on the whole interval  $[-3\pi, 3\pi]$ .

As a matter of fact,  $f \notin \mathfrak{R}(\alpha)$  on  $[-3\pi, 3\pi]$  if we re-define  $f(0)$  other than 1 while keep  $\alpha$  unchanged.

 **Exercise 7.2** Let  $f$  be a constant function on  $[a, b]$  except for *finitely* many points. Show that  $f$  is Riemann integrable on  $[a, b]$ .

[Hint: You may find the theorem of integration by parts (Theorem 7.4.1) helpful.]

## 7.7 Reduction of Riemann-Stieltjes Integrals to Finite Sums

## 7.8 Darboux Integrals with Monotonically Increasing Integrators

In this section, we shall introduce another formulation of integrals, which is known as the **Darboux integrals**. As we noted before, the Riemann-Stieltjes sum  $S(P, f, \alpha)$  also depends on the choice of representatives,  $t_k$ s. This uncertainty increases the difficulty of proving assertions related to integrals. For example, as we have seen in Example 7.3, it takes quite a lot of work to prove that integral does not exist. Hence, we need Darboux's formulation, which has the advantage of being easier to apply in computations or proofs.

### Definition 7.8.1

Let  $f$  be a bounded function on  $[a, b]$ , and  $P = \{a = x_0, x_1, \dots, x_n = b\}$  a partition. In each sub-interval  $[x_{k-1}, x_k]$ , we use the symbols  $M_k$  and  $m_k$  to denote

$$M_k = \sup \{f(x) \mid x \in [x_{k-1}, x_k]\} \quad \text{and} \quad m_k = \inf \{f(x) \mid x \in [x_{k-1}, x_k]\}$$

(The supremum and infimum are well-defined since we assume  $f$  is bounded.) The **upper Riemann-Stieltjes sum**  $U(P, f, \alpha)$  and **lower Riemann-Stieltjes sum**  $L(P, f, \alpha)$  are defined by

$$U(P, f, \alpha) = \sum_{k=1}^n M_k \Delta \alpha_k \quad \text{and} \quad L(P, f, \alpha) = \sum_{k=1}^n m_k \Delta \alpha_k$$



If  $\alpha(x) = x$ , we write  $U(P, f)$  and  $L(P, f)$  with the specification of  $\alpha$  dropped. And in this case,  $U(P, f)$  and  $L(P, f)$  are called the **upper and lower Darboux sum**.

 **Note** From the definition, we note that, unlike  $S(P, f, \alpha)$ ,  $U(P, f, \alpha)$  and  $L(P, f, \alpha)$  only depend on  $P$ ,  $f$  and  $\alpha$ .

On each sub-interval  $[x_{k-1}, x_k]$ , by definition, we have

$$m_k \leq f(x) \leq M_k \quad \forall x \in [x_{k-1}, x_k] \quad (7.20)$$

It follows that

$$L(P, f, \alpha) \leq U(P, f, \alpha) \quad (7.21)$$

If we pick a representative  $t_k \in [x_{k-1}, x_k]$  in each interval and assume  $\alpha$  is monotonically increasing, then (7.20) yields

$$L(P, f, \alpha) \leq S(P, f, \alpha) \leq U(P, f, \alpha)$$

The following theorem states that not only  $L(P, f, \alpha) \leq U(P, f, \alpha)$  for the same partition  $P$ , but also  $L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$  for any two partitions  $P_1$  and  $P_2$ . The theorem also says the upper sums will decrease as the partition gets finer and finer while the lower sums will increase.

### Theorem 7.8.1

*Suppose  $\alpha$  is increasing on  $[a, b]$ . We have the following statements.*

1. If  $P' \supseteq P$ , then

$$U(P', f, \alpha) \leq U(P, f, \alpha) \quad \text{and} \quad L(P', f, \alpha) \geq L(P, f, \alpha)$$

2. For any two partitions  $P_1$  and  $P_2$  on  $[a, b]$ , we have

$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$$



**Proof** (Proof of 1) We only prove  $U(P', f, \alpha) \leq U(P, f, \alpha)$  since the proof for lower sums is similar. Note that it suffices to prove the case where  $P'$  has exactly one point, say  $c$ , more than  $P$ .

**Note** If  $P'$  has  $m$  points more than  $P$ , then we can construct a chain of partitions:

$$P' = P_1 \supseteq P_2 \cdots \supseteq P_m = P$$

in such a way that  $P_i$  has one point more than  $P_{i+1}$ . Then by applying the inequality  $U(P_i, f, \alpha) \leq U(P_{i+1}, f, \alpha)$ , which is assumed proved, we obtain

$$U(P', f, \alpha) \leq U(P_2, f, \alpha) \leq \cdots \leq U(P, f, \alpha)$$

Suppose  $P = \{x_0, x_1, \dots, x_n\}$ , and point  $c$  is in between  $x_{i-1}$  and  $x_i$ . The upper sum  $U(P', f, \alpha)$  is then

$$U(P', f, \alpha) = \sum_{k=1, k \neq i}^n M_k \Delta \alpha_k + M'[\alpha(c) - \alpha(x_{i-1})] + M''[\alpha(x_i) - \alpha(c)]$$

where

$$M' = \sup \{f(x) \mid x \in [x_{i-1}, c]\} \quad \text{and} \quad M'' = \sup \{f(x) \mid x \in [c, x_i]\}$$

Since  $M_i \geq M', M''$ , we have

$$M'[\alpha(c) - \alpha(x_{i-1})] + M''[\alpha(x_i) - \alpha(c)] \leq M_i \Delta \alpha_i$$

Hence,  $U(P', f, \alpha) \leq U(P, f, \alpha)$ .

(Proof of 2) Let  $P$  be a common refinement of  $P_1$  and  $P_2$ , say  $P = P_1 \cup P_2$ . Applying what we have just proved, it follows that

$$L(P, f, \alpha) \geq L(P_1, f, \alpha) \quad \text{and} \quad U(P, f, \alpha) \leq U(P_2, f, \alpha)$$

As noted before in (7.21),  $L(P, f, \alpha) \leq U(P, f, \alpha)$ . Therefore,

$$L(P_1, f, \alpha) \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq U(P_2, f, \alpha)$$

■

In particular, since every partition is a refinement of the trivial partition  $P_0 = \{a, b\}$ , Theorem 7.8.1 leads to

$$\inf_{x \in [a, b]} f(x)[\alpha(b) - \alpha(a)] \leq L(P_1, f, \alpha) \leq U(P_2, f, \alpha) \leq \sup_{x \in [a, b]} f(x)[\alpha(b) - \alpha(a)] \quad (7.22)$$

What (7.22) implies is that all the upper and lower Riemann-Stieltjes sums are bounded. Therefore, the infimum of upper sums and supremum of lower sums both exist as *finite* numbers, which gives rise to the definition of **upper and lower Riemann-Stieltjes integrals**.

### Definition 7.8.2

Suppose  $f$  is bounded and  $\alpha$  is monotonically increasing on  $[a, b]$ . The upper and lower Riemann-Stieltjes integrals, written as  $\bar{\int}_a^b f d\alpha$  and  $\underline{\int}_a^b f d\alpha$ , are defined by

$$\underline{\int}_a^b f d\alpha = \inf \{U(P, f, \alpha) \mid P \in \mathcal{P}[a, b]\} \quad \text{and} \quad \bar{\int}_a^b f d\alpha = \sup \{L(P, f, \alpha) \mid P \in \mathcal{P}[a, b]\}$$



Sometimes, we will just say upper and lower integrals for short.

As indicated by the names, one should expect that the lower integral should be less than or equal to the upper integral. This is indeed the case.

### Theorem 7.8.2

Suppose  $\alpha$  is increasing on  $[a, b]$ . We have

$$\underline{\int}_a^b f d\alpha \leq \bar{\int}_a^b f d\alpha \quad (7.23)$$



**Proof** Given  $\varepsilon > 0$ , there exists a partition  $P_\varepsilon$  such that

$$U(P_\varepsilon, f, \alpha) < \bar{\int}_a^b f d\alpha + \varepsilon$$

By Theorem 7.8.1, for any partition  $P$ , we have

$$L(P, f, \alpha) \leq U(P_\varepsilon, f, \alpha)$$

It then follows that

$$L(P, f, \alpha) < \bar{\int}_a^b f d\alpha + \varepsilon \quad \forall P \in \mathcal{P}[a, b]$$

This implies that  $\bar{\int}_a^b f d\alpha + \varepsilon$  is an upper bound of all the lower Riemann-Stieltjes sums. Thus, it must be no less than their supremum, that is,

$$\underline{\int}_a^b f d\alpha \leq \bar{\int}_a^b f d\alpha + \varepsilon$$

Since the above inequality holds for any positive number  $\varepsilon > 0$ , we have

$$\underline{\int}_a^b f d\alpha \leq \bar{\int}_a^b f d\alpha$$



There exist cases where the inequality (7.23) is strict.

**Example 7.4** Consider the Dirichlet function

$$\mathbb{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

restricted on  $[0, 1]$ . For any partition  $P = \{0 = x_0, x_1, \dots, x_n = 1\}$ , the infimum of  $\mathbb{1}_{\mathbb{Q}}$  on each sub-interval  $[x_{k-1}, x_k]$  is 0. Hence,  $L(P, \mathbb{1}_{\mathbb{Q}}) = \sum_{k=1}^n 0 \cdot (x_k - x_{k-1}) = 0$ , which implies the lower Darboux sum is always 0. Therefore, the lower integral is 0, i.e.,  $\underline{\int}_0^1 \mathbb{1}_{\mathbb{Q}}(x) dx = 0$ . Similarly, because the supremum of the Dirichlet function is 1 in each sub-interval, the upper Darboux sum is always 1. Hence,  $\bar{\int}_0^1 \mathbb{1}_{\mathbb{Q}}(x) dx = 1$ . In this case, the lower integral is strictly less than the upper integral.

In the following theorems, we introduce some properties of upper and lower integrals.

### Theorem 7.8.3

Suppose that  $f$  is bounded and  $\alpha$  is increasing on  $[a, b]$ . Let  $c \in (a, b)$ . We have

1.  $\bar{\int}_a^b f d\alpha = \bar{\int}_a^c f d\alpha + \bar{\int}_c^b f d\alpha$
2.  $\underline{\int}_a^b f d\alpha = \underline{\int}_a^c f d\alpha + \underline{\int}_c^b f d\alpha$



**Proof** We only prove 1, the equality concerning the upper integrals, since 2 can be proved similarly.

Given  $\varepsilon > 0$ , there exist a partition  $P_1$  on  $[a, c]$  and a partition  $P_2$  on  $[c, b]$  such that

$$U(P_1, f, \alpha) < \bar{\int}_a^c f d\alpha + \varepsilon/2 \quad \text{and} \quad U(P_2, f, \alpha) < \bar{\int}_c^b f d\alpha + \varepsilon/2 \quad (7.24)$$

Let  $P = P_1 \cup P_2$ . Note that  $P$  is a partition on  $[a, b]$ . Furthermore, the upper Riemann-Stieltjes sum  $U(P, f, \alpha)$  is given by

$$U(P, f, \alpha) = U(P_1, f, \alpha) + U(P_2, f, \alpha) \quad (7.25)$$

But

$$U(P, f, \alpha) \geq \bar{\int}_a^b f d\alpha \quad (7.26)$$

Combining (7.24), (7.25) and (7.26) together, we have

$$\bar{\int}_a^b f d\alpha < \bar{\int}_a^c f d\alpha + \bar{\int}_c^b f d\alpha + \varepsilon$$

Since the above inequality holds for any  $\varepsilon > 0$ . It follows that

$$\bar{\int}_a^b f d\alpha \leq \bar{\int}_a^c f d\alpha + \bar{\int}_c^b f d\alpha \quad (7.27)$$

On the other hand, letting  $\varepsilon > 0$  be chosen arbitrarily, there exists a partition  $P$  on  $[a, b]$  such that

$$U(P, f, \alpha) < \bar{\int}_a^b f d\alpha + \varepsilon \quad (7.28)$$

We want to split  $P$  into two partitions, one on  $[a, c]$  and the other on  $[c, b]$ . But we need to ensure that  $c$  is in  $P$ . To do so, we consider a refinement  $P' = P \cup \{c\}$ . By Theorem 7.8.1, we know that the

upper Riemann-Stieltjes sum concerning  $P'$  is reduced, i.e.,

$$U(P', f, \alpha) \leq U(P, f, \alpha) < \int_a^b f \, d\alpha + \varepsilon \quad (7.29)$$

where the second inequality follows from (7.28). Let  $P_1$  and  $P_2$  be given by

$$P_1 = \{x \in P' \mid x \leq c\} \quad \text{and} \quad P_2 = \{x \in P' \mid x \geq c\}$$

Note that indeed  $P_1$  and  $P_2$  are partitions on  $[a, c]$  and  $[c, b]$ , respectively. We have

$$U(P', f, \alpha) = U(P_1, f, \alpha) + U(P_2, f, \alpha) \quad (7.30)$$

Moreover, The two upper Riemann-Stieltjes sums are bounded below by the corresponding upper integrals, that is,

$$U(P_1, f, \alpha) \geq \int_a^c f \, d\alpha \quad \text{and} \quad U(P_2, f, \alpha) \geq \int_c^b f \, d\alpha \quad (7.31)$$

Then, by combining (7.29), (7.30) and (7.31), we obtain

$$\int_a^b f \, d\alpha + \varepsilon > \int_a^c f \, d\alpha + \int_c^b f \, d\alpha$$

Similarly, since the above inequality holds for any  $\varepsilon > 0$ , we have

$$\int_a^b f \, d\alpha \geq \int_a^c f \, d\alpha + \int_c^b f \, d\alpha \quad (7.32)$$

Finally, the equation

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

follows from (7.27) and (7.32). ■

However, some equations of integrals will not be valid for upper and lower integrals. For example, the following equation is a special case of Theorem 7.2.1.

$$\int_a^b (f + g) \, d\alpha = \int_a^b f \, d\alpha + \int_a^b g \, d\alpha$$

It will not hold if we replace the integrals with upper and lower integrals. To fix this, we also need to replace equality with inequality, as stated in the next theorem.

#### Theorem 7.8.4

Suppose that  $f$  and  $g$  are bounded and  $\alpha$  is increasing on  $[a, b]$ . We have

1.  $\bar{\int}_a^b (f + g) \, d\alpha \leq \bar{\int}_a^b f \, d\alpha + \bar{\int}_a^b g \, d\alpha$
2.  $\underline{\int}_a^b (f + g) \, d\alpha \geq \underline{\int}_a^b f \, d\alpha + \underline{\int}_a^b g \, d\alpha$



**Proof** We only prove 1. Given  $\varepsilon > 0$ , there exist partitions  $P_1$  and  $P_2$  on  $[a, b]$  such that

$$U(P_1, f, \alpha) < \int_a^b f \, d\alpha + \varepsilon/2 \quad \text{and} \quad U(P_2, g, \alpha) < \int_a^b g \, d\alpha + \varepsilon/2$$

Let  $P = P_1 \cup P_2$ . It follows from Theorem 7.8.1 that

$$U(P, f, \alpha) \leq U(P_1, f, \alpha) < \int_a^b f \, d\alpha + \varepsilon/2 \quad \text{and} \quad U(P, g, \alpha) \leq U(P_2, g, \alpha) < \int_a^b g \, d\alpha + \varepsilon/2$$

Adding the two inequalities above yields

$$U(P, f, \alpha) + U(P, g, \alpha) < \int_a^b f \, d\alpha + \int_a^b g \, d\alpha + \varepsilon \quad (7.33)$$

Write  $P = \{x_0, x_1, \dots, x_n\}$ . Let  $M_k$ ,  $M'_k$  and  $M''_k$  denote

$$\begin{aligned} M_k &= \sup \{f(x) + g(x) \mid x \in [x_{k-1}, x_k]\} \\ M'_k &= \sup \{f(x) \mid x \in [x_{k-1}, x_k]\} \\ M''_k &= \sup \{g(x) \mid x \in [x_{k-1}, x_k]\} \end{aligned}$$

It is clear that

$$M_k \leq M'_k + M''_k \quad \forall k = 1, \dots, n$$

since  $f(x) + g(x) \leq M'_k + M''_k \forall x \in [x_{k-1}, x_k]$ . It then follows that

$$U(P, f + g, \alpha) \leq U(P, f, \alpha) + U(P, g, \alpha)$$

Because  $U(P, f + g, \alpha)$  is bounded below by its upper integral, we further have

$$\int_a^b (f + g) \, d\alpha \leq U(P, f + g, \alpha) \leq U(P, f, \alpha) + U(P, g, \alpha) \quad (7.34)$$

Combining inequalities (7.33) and (7.34), we obtain

$$\int_a^b (f + g) \, d\alpha < \int_a^b f \, d\alpha + \int_a^b g \, d\alpha + \varepsilon$$

Because the above inequality holds for any  $\varepsilon > 0$ , we finally have the inequality

$$\int_a^b (f + g) \, d\alpha \leq \int_a^b f \, d\alpha + \int_a^b g \, d\alpha$$

as desired. ■

The following is an example where both inequalities are strict.

**Example 7.5** Let functions

$$f = \mathbb{1}_{\mathbb{Q}} \quad \text{and} \quad g = -\mathbb{1}_{\mathbb{Q}}$$

be restricted on  $[0, 1]$ . We have calculated the upper and lower integrals of  $f$  in Example 7.4. We can do the same for  $g$  similarly. The results are as follows.

$$\int_0^1 f(x) \, dx = 1, \quad \underline{\int}_0^1 f(x) \, dx = 0, \quad \overline{\int}_0^1 f(x) \, dx = 0, \quad \text{and} \quad \int_0^1 g(x) \, dx = -1,$$

Note that the sum of the two functions is zero, i.e.,  $f + g = 0$ . Hence, we have

$$\begin{aligned} \int_0^1 (f + g)(x) \, dx &= 0 < 1 = \int_0^1 f(x) \, dx + \int_0^1 g(x) \, dx \\ \int_0^1 (f + g)(x) \, dx &= 0 > -1 = \underline{\int}_0^1 f(x) \, dx + \overline{\int}_0^1 g(x) \, dx \end{aligned}$$

## 7.9 Riemann's Criterion

In Darboux's formulation, the integral exists if and only if the upper and lower integrals are equal to each other. We now show that this is equivalent to the criterion of the existence of integrals from our definition (Definition 7.1.1). Furthermore, we will introduce another equivalent condition,

**Riemann's criterion**, which is convenient to use in proofs.

**Definition 7.9.1 (Riemann's Criterion)**

Suppose  $f$  is bounded and  $\alpha$  is increasing on  $[a, b]$ . Function  $f$  is said to satisfy **Riemann's criterion** w.r.t.  $\alpha$  on  $[a, b]$  if for any  $\varepsilon > 0$ , there exists a partition  $P_\varepsilon$  on  $[a, b]$  such that

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

for any refinement  $P \supseteq P_\varepsilon$ .



Essentially, Riemann's criterion requires that the upper and lower sums should be arbitrarily close.

The following states that the existence of Riemann-Stieltjes integrals, the existence of Darboux integrals and Riemann's criterion are equivalent.

**Theorem 7.9.1**

Suppose  $f$  is bounded and  $\alpha$  is increasing on  $[a, b]$ . Then the following statements are equivalent.

1.  $f$  is integrable w.r.t.  $\alpha$  on  $[a, b]$ , i.e.,  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ .
2.  $f$  satisfies Riemann's criterion w.r.t.  $\alpha$  on  $[a, b]$ .
3. The upper and lower integrals are equal to each other, i.e.,  $\bar{\int}_a^b f \, d\alpha = \int_a^b f \, d\alpha$ .



**Proof** If  $\alpha$  is constant, then this theorem holds trivially. We assume  $\alpha$  is non-constant in the rest of the proof.

(Proof of 1  $\implies$  2) Given  $\varepsilon > 0$ , since  $f \in \mathfrak{R}(\alpha)$ , there exists a partition  $P_\varepsilon$  on  $[a, b]$  such that

$$\left| S(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/6 \quad (7.35)$$

for any refinement  $P \supseteq P_\varepsilon$  and for any choice of representative  $t_k$  in each sub-interval  $[x_{k-1}, x_k]$ . Let this partition  $P$  be fixed for now.

Let  $M_k$  and  $m_k$  be as in Definition 7.8.1. We can choose each  $t'_k \in [x_{k-1}, x_k]$  such that

$$M_k - \frac{\varepsilon/3}{\alpha(b) - \alpha(a)} < f(t'_k) \leq M_k$$

Multiplying by  $\Delta\alpha_k$  and then summing up over  $k$  yields

$$U(P, f, \alpha) - \varepsilon/3 < S'(P, f, \alpha) \leq U(P, f, \alpha) \quad (7.36)$$

where  $S'(P, f, \alpha) = \sum_k f(t'_k) \Delta\alpha_k$ . We can instead choose each  $t''_k \in [x_{k-1}, x_k]$  such that

$$m_k + \frac{\varepsilon/3}{\alpha(b) - \alpha(a)} > f(t''_k) \geq m_k$$

Similarly, we then have

$$L(P, f, \alpha) + \varepsilon/3 > S''(P, f, \alpha) \geq L(P, f, \alpha) \quad (7.37)$$

where  $S''(P, f, \alpha) = \sum_k f(t''_k) \Delta\alpha_k$ . By combining (7.36) and (7.37), we will obtain the following two inequalities.

$$0 \leq U(P, f, \alpha) - S'(P, f, \alpha) < \varepsilon/3 \quad (7.38)$$

$$0 \leq S''(P, f, \alpha) - L(P, f, \alpha) < \varepsilon/3 \quad (7.39)$$

Recall the Riemann-Stieltjes sums  $S'(P, f, \alpha)$  and  $S''(P, f, \alpha)$  are formed with the same partition  $P$ . Hence, they both satisfy (7.36). Their difference is then bounded by

$$|S'(P, f, \alpha) - S''(P, f, \alpha)| \leq \left| S'(P, f, \alpha) - \int_a^b f \, d\alpha \right| + \left| S''(P, f, \alpha) - \int_a^b f \, d\alpha \right| < \varepsilon/3 \quad (7.40)$$

Then, the difference between the upper and lower sums can be bounded as follows.

$$\begin{aligned} & U(P, f, \alpha) - L(P, f, \alpha) \\ & \leq |U(P, f, \alpha) - S'(P, f, \alpha)| + |S'(P, f, \alpha) - S''(P, f, \alpha)| + |S''(P, f, \alpha) - L(P, f, \alpha)| \\ & < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ & = \varepsilon \end{aligned}$$

where the last inequality follows from (7.38), (7.39) and (7.40). Therefore, Riemann's criterion is satisfied since the above inequality holds for any refinement  $P$  of  $P_\varepsilon$ .

(Proof of 2  $\implies$  3) Let  $\varepsilon > 0$  be arbitrary. By Riemann's criterion, there exists a partition  $P_\varepsilon$  such that

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

for any refinement  $P \supseteq P_\varepsilon$ . Because  $\bar{\int}_a^b f \, d\alpha \leq U(P, f, \alpha)$  and  $\underline{\int}_a^b f \, d\alpha \geq L(P, f, \alpha)$ , it follows that

$$0 \leq \bar{\int}_a^b f \, d\alpha - \underline{\int}_a^b f \, d\alpha < \varepsilon$$

Therefore, the upper and lower integrals are equal to each other since the above inequality holds for any  $\varepsilon > 0$ .

(Proof of 3  $\implies$  1) Given  $\varepsilon > 0$ . There exist partitions  $P_1$  and  $P_2$  on  $[a, b]$  such that

$$U(P_1, f, \alpha) < \bar{\int}_a^b f \, d\alpha + \varepsilon \quad \text{and} \quad L(P_2, f, \alpha) > \underline{\int}_a^b f \, d\alpha - \varepsilon$$

Let  $P_\varepsilon = P_1 \cup P_2$  and  $P$  be any of its refinements. By Theorem 7.8.1 and (7.20), it is clear

$$\bar{\int}_a^b f \, d\alpha - \varepsilon < L(P, f, \alpha) \leq S(P, f, \alpha) \leq U(p, f, \alpha) < \bar{\int}_a^b f \, d\alpha + \varepsilon$$

holds for any choice of  $t_k$  in each sub-interval  $[x_{k-1}, x_k]$ . Since the upper and lower integrals have the same value, say  $A$ , it then follows that

$$|S(P, f, \alpha) - A| < \varepsilon$$

Therefore, indeed  $f \in \mathfrak{R}(\alpha)$ . This completes the proof. ■

As one should expect, the integral  $\int_a^b f \, d\alpha$  has to be equal to the upper and lower integrals whenever  $f$  is integrable.

### Corollary 7.9.1

Suppose  $\alpha$  be increasing on  $[a, b]$ . If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then

$$\int_a^b f \, d\alpha = \bar{\int}_a^b f \, d\alpha = \underline{\int}_a^b f \, d\alpha$$



**Proof** Given  $\varepsilon > 0$ , there exist partitions  $P_1$  and  $P_2$  on  $[a, b]$  such that

$$U(P_1, f, \alpha) < \int_a^b f \, d\alpha + \varepsilon \quad \text{and} \quad L(P_2, f, \alpha) > \int_a^b f \, d\alpha - \varepsilon$$

Let  $P_\varepsilon = P_1 \cup P_2$ . For any refinement  $P \supseteq P_\varepsilon$ , we have

$$\int_a^b f \, d\alpha - \varepsilon < L(P, f, \alpha) \leq S(P, f, \alpha) \leq U(P, f, \alpha) < \int_a^b f \, d\alpha + \varepsilon$$

The upper and lower integrals are equal, say to  $A$ , by Theorem 5.4.4. It follows that

$$|S(P, f, \alpha) - A| < \varepsilon$$

By Definition 7.1.1,  $A = \int_a^b f \, d\alpha$ . ■

As an exercise, we reconsider Example 7.3 and prove that  $f$  is not integrable w.r.t.  $\alpha$ .

**Exercise 7.3** Let functions  $f$  and  $\alpha$  be given by

$$f(x) = \begin{cases} \frac{|\sin(x)|}{x} & x \neq 0 \\ 1 & x = 0 \end{cases} \quad \text{and} \quad \alpha(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}, \quad x \in [-3\pi, 3\pi]$$

Show that  $f \notin \mathfrak{R}(\alpha)$  on  $[-3\pi, 3\pi]$ .

[Hint:  $\sin x < x \forall x > 0$ .]

**Solution** It suffices to prove  $f$  is not integrable w.r.t.  $\alpha$  on  $[-3\pi, 0]$  since  $f \in \mathfrak{R}(\alpha)$  on  $[0, 3\pi]$ .

Note that For any partition  $P = \{x_0, x_1, \dots, x_n\}$  on  $[-3\pi, 0]$ , only the last sub-integral  $[x_{n-1}, x_n = 0]$  contributes to the upper and lower sums. Furthermore, using the hint, we know

$$M_n = 1 \quad \text{and} \quad m_n = -1$$

Then the upper and lower sums are given by

$$U(P, f, \alpha) = M_n[\alpha(0) - \alpha(x_{n-1})] = 2 \quad \text{and} \quad L(P, f, \alpha) = m_n[\alpha(0) - \alpha(x_{n-1})] = -2$$

Since the above equations hold for any partition  $P$  on  $[-3\pi, 0]$ , we have

$$\int_{-3\pi}^0 = 2 \quad \text{and} \quad \int_{-3\pi}^0 = -2$$

Therefore, by Theorem 7.9.1,  $f$  is not integrable w.r.t.  $\alpha$  on  $[-3\pi, 0]$ , and hence on  $[-3\pi, 3\pi]$ .

This proof is way cleaner than the one given in Example 7.3 thanks to Darboux's formulation of integrals.

## 7.10 Properties of Riemann-Stieltjes Integrals With Respect to Increasing Integrators

It is much easier to study and prove some properties of the Riemann-Stieltjes integrals if we require that  $\alpha$  is monotonically increasing.

Thanks to Theorem 7.11.1, which we will prove in Section 7.11, almost all the properties introduced in this section can be extended with ease to integrators of bounded variation.

### 7.10.1 Integrability on Sub-Intervals

Let  $c \in (a, b)$  be a point that splits the interval into two parts. If  $f$  is integrable w.r.t.  $\alpha$  on the entire interval  $[a, b]$ , then it is also integrable on both of the sub-intervals.

#### Lemma 7.10.1

*Suppose  $\alpha$  is increasing on  $[a, b]$ . Let  $c \in (a, b)$  be an interior point. If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then  $f \in \mathfrak{R}(\alpha)$  on  $[a, c]$  and  $f \in \mathfrak{R}(\alpha)$  on  $[c, b]$*



**Proof** We only prove  $f \in \mathfrak{R}(\alpha)$  on  $[a, c]$ . Given  $\varepsilon > 0$ . By Theorem 7.9.1, there exists a partition  $P_\varepsilon$  on  $[a, b]$  such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon \quad \forall P \supseteq P_\varepsilon$$

Let  $P'_\varepsilon = (P_\varepsilon \cup \{c\}) \cap [a, c]$ . Note that  $P'_\varepsilon$  is a partition on  $[a, c]$ . Now, let  $P'$  be a refinement of  $P'_\varepsilon$  on  $[a, c]$ , and  $P$  a refinement of  $P_\varepsilon$  on  $[a, b]$ . It is evident that

$$P'' = P' \cup (P \cap [c, b])$$

is a refinement of  $P_\varepsilon$ . Therefore, we have

$$U(P'', f, \alpha) - L(P'', f, \alpha) < \varepsilon$$

**Note**  $P''$  can be regarded as a partition formed by adding points to  $P'$  that are greater than  $c$ .

Then it is clear

$$\begin{aligned} U(P', f, \alpha) - L(P', f, \alpha) &= \sum_{k \text{ such that } x_k \leq c} (M_k - m_k) \Delta \alpha_k \\ &\leq \sum_k (M_k - m_k) \Delta \alpha_k \leq U(P'', f, \alpha) - L(P'', f, \alpha) \end{aligned}$$

Hence,

$$U(P', f, \alpha) - L(P', f, \alpha) < \varepsilon$$

which implies  $f \in \mathfrak{R}(\alpha)$  on  $[a, c]$  by Theorem 7.9.1. ■

In fact,  $f$  is integrable on all sub-intervals of  $[a, b]$ .

#### Theorem 7.10.1

*Suppose  $\alpha$  is increasing on  $[a, b]$ . Let  $[c, d] \subseteq [a, b]$  be a sub-interval. If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then  $f \in \mathfrak{R}(\alpha)$  on  $[c, d]$ .*



**Proof** Without loss of generality, assume  $a < c < d < b$ . Lemma 7.10.1 implies that  $f$  is integrable w.r.t.  $\alpha$  on  $[a, c]$  and  $[a, d]$ . We then conclude the proof using Theorem 7.2.1. ■

### 7.10.2 Comparison Theorem

It is very common in practice to compare the values of two integrals. Intuitively,

$$\int_a^b f \, d\alpha \leq \int_a^b g \, d\alpha \tag{7.41}$$

provided that  $f \leq g$  and  $\alpha$  is increasing. The proof of this inequality follows from the following lemma, which says the integral of  $f$  w.r.t.  $\alpha$  is non-negative whenever  $f$  stays non-negative and  $\alpha$  is increasing on  $[a, b]$ . Note that this is a special case of inequality (7.41).

### Lemma 7.10.2

*Suppose that  $\alpha$  is increasing and  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . If  $f(x) \geq 0 \forall x \in [a, b]$ , then*

$$\int_a^b f \, d\alpha \geq 0$$



**Proof** Consider the trivial partition  $P = \{a, b\}$ . We have

$$L(P, f, \alpha) = \inf_{x \in [a, b]} f(x)[\alpha(b) - \alpha(a)] \geq 0$$

It then follows that

$$\int_a^b f \, d\alpha \geq L(P, f, \alpha) \geq 0$$

Note that  $\int_a^b f \, d\alpha = \int_a^b f \, d\alpha$  since  $f \in \mathfrak{R}(\alpha)$ . The proof is then complete. ■

Inequality (7.41) can be proved easily by applying the above lemma and Theorem 7.2.1.

### Theorem 7.10.2

*Suppose that  $\alpha$  is increasing and  $f, g \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . If  $f(x) \leq g(x) \forall x \in [a, b]$ , then we have the inequality*

$$\int_a^b f \, d\alpha \leq \int_a^b g \, d\alpha$$



**Proof** Note that the difference  $(g - f)(x) \geq 0 \forall x \in [a, b]$ . Furthermore,  $(g - f) \in \mathfrak{R}(\alpha)$  on  $[a, b]$  by Theorem 7.2.1. Then Lemma 7.10.2 gives

$$\int_a^b (g - f) \, d\alpha \geq 0$$

Using Theorem 7.2.1 again to split the integrand, we have

$$\int_a^b g \, d\alpha - \int_a^b f \, d\alpha \geq 0$$



## 7.10.3 Applying Continuous Functions to Integrable Functions

Suppose we know  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . The following theorem tells us that we can produce more integrable functions by applying continuous functions to  $f$ .

### Theorem 7.10.3

*Suppose that  $\alpha$  is increasing and  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . Let  $M$  and  $m$  denote the supremum and infimum of  $f$  on  $[a, b]$ . If function  $\phi : [m, M] \rightarrow \mathbb{R}$  is continuous, then the composite function  $\phi \circ f$  is also integrable w.r.t.  $\alpha$  on  $[a, b]$ , i.e.,  $\phi \circ f \in \mathfrak{R}(\alpha)$ .*



**Proof** In the following proof, we assume that  $\alpha$  is non-constant.

Given  $\varepsilon > 0$ . Since  $\phi$  is uniformly continuous on  $[m, M]$ , there exists  $\delta > 0$  such that

$$|s - t| < \delta \implies |\phi(s) - \phi(t)| < \frac{\varepsilon/2}{\alpha(b) - \alpha(a)} \quad (7.42)$$

Suppose that  $\phi$  is bounded by a positive number  $A > 0$ , i.e.,

$$|\phi| < A$$

Because  $f \in \mathfrak{R}(\alpha)$ , there exists a partition  $P_\varepsilon$  on  $[a, b]$  such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \frac{\varepsilon\delta}{4A} \quad (7.43)$$

for any refinement  $P \supseteq P_\varepsilon$ .

Write  $P = \{x_0, x_1, \dots, x_n\}$ . We now separate the indices of sub-intervals into two sets in the following way.

$$\begin{aligned} I &= \{k = 1, \dots, n \mid M_k(f) - m_k(f) < \delta\} \\ J &= \{k = 1, \dots, n \mid M_k(f) - m_k(f) \geq \delta\} \end{aligned}$$

We first consider the sub-intervals whose indices belong to  $I$ . Suppose  $k \in I$ . Since  $M_k(f) - m_k(f) < \delta$ , it follows from (7.42) that

$$|\phi[f(s)] - \phi[f(t)]| < \frac{\varepsilon/2}{\alpha(b) - \alpha(a)} \quad \forall s, t \in [x_{k-1}, x_k]$$

Then by Theorem 1.1.1, we have

$$M_k(\phi \circ f) - m_k(\phi \circ f) = \sup_{s, t \in [x_{k-1}, x_k]} |\phi[f(s)] - \phi[f(t)]| \leq \frac{\varepsilon/2}{\alpha(b) - \alpha(a)}$$

Summing up over  $k \in I$  yields

$$\sum_{k \in I} [M_k(\phi \circ f) - m_k(\phi \circ f)] \Delta \alpha_k \leq \varepsilon/2 \quad (7.44)$$

Now, suppose  $k \in J$ . Note that

$$\sum_{k \in I} [M_k(\phi \circ f) - m_k(\phi \circ f)] \Delta \alpha_k \leq U(P, f, \alpha) < \frac{\varepsilon\delta}{4A}$$

where the last inequality follows from (7.43). Because  $M_k(f) - m_k(f) \geq \delta$ , we have

$$\delta \sum_{k \in I} \Delta \alpha_k \leq \sum_{k \in I} [M_k(\phi \circ f) - m_k(\phi \circ f)] \Delta \alpha_k < \frac{\varepsilon\delta}{4A}$$

which further implies

$$\sum_{k \in I} \Delta \alpha_k < \frac{\varepsilon}{4A} \quad (7.45)$$

Recall that  $|\phi| < A$ . Hence, the difference  $M_k(\phi \circ f) - m_k(\phi \circ f)$  is bounded above by

$$M_k(\phi \circ f) - m_k(\phi \circ f) < 2A \quad (7.46)$$

Combining inequalities (7.45) and (7.46), we obtain

$$\sum_{k \in I} [M_k(\phi \circ f) - m_k(\phi \circ f)] \Delta \alpha_k < \varepsilon/2 \quad (7.47)$$

Note that the difference  $U(P, \phi \circ f, \alpha) - L(P, \phi \circ f, \alpha)$  can be summed up as follows.

$$\begin{aligned} U(P, \phi \circ f, \alpha) - L(P, \phi \circ f, \alpha) &= \sum_{k=1}^n [M_k(\phi \circ f) - m_k(\phi \circ f)] \Delta \alpha_k \\ &= \sum_{k \in I} [M_k(\phi \circ f) - M_k(\phi \circ f)] \Delta \alpha_k + \sum_{k \in J} [M_k(\phi \circ f) - m_k(\phi \circ f)] \Delta \alpha_k \end{aligned}$$

Finally, by applying (7.44) and (7.47), the difference of upper and lower sums are bounded by

$$U(P, \phi \circ f, \alpha) - L(P, \phi \circ f, \alpha) < \varepsilon$$

Therefore, the composite function  $\phi \circ f$  satisfies Riemann's criterion, and hence  $\phi \circ f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  by Theorem 7.9.1. ■

Note that the absolute value function  $\phi(x) = |x|$  is continuous. Applying it to an integrable function  $f \in \mathfrak{R}(\alpha)$ , we immediately conclude that  $|f|$  is also integrable.

#### Theorem 7.10.4

Suppose  $\alpha$  is increasing. If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then  $|f| \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , and we have

$$\left| \int_a^b f \, d\alpha \right| \leq \int_a^b |f| \, d\alpha \quad (7.48)$$



**Proof** First, we note that  $|f| \in \mathfrak{R}(\alpha)$  on  $[a, b]$  by Theorem 7.10.3 since  $\phi(x) = |x|$  is continuous.

It then follows from the inequality

$$-|f| \leq f \leq |f|$$

and Theorem 7.10.2 that

$$-\int_a^b |f| \, d\alpha = \int_a^b -|f| \, d\alpha \leq \int_a^b f \, d\alpha \leq \int_a^b |f| \, d\alpha$$

This completes the proof. ■

The converse of this theorem is not true. That is,  $f$  is not necessarily integrable if  $|f|$  is integrable.

**Example 7.6** Let

$$f(x) = \mathbb{1}_{\mathbb{Q}}(x) - \mathbb{1}_{\mathbb{R} \setminus \mathbb{Q}}(x), \quad x \in [a, b]$$

and  $\alpha$  be any *non-constant* increasing function on  $[a, b]$ .

Note that the absolute value of  $f$ ,  $|f| = 1$  is constant, and hence  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . And

$$\int_a^b f \, d\alpha = 0$$

But  $f$  is not integrable w.r.t.  $\alpha$  on  $[a, b]$  since

$$\int_a^b f \, d\alpha = \alpha(b) - \alpha(a) > 0 \quad \text{and} \quad \int_a^b f \, d\alpha = \alpha(a) - \alpha(b) < 0$$

The upper and lower integrals are not equal to each other for we have assumed  $\alpha$  is non-constant.

☞ **Exercise 7.4** Prove the existence of  $\int_a^b |f| \, d\alpha$  in Theorem 7.10.4 without using Theorem 7.10.3.

**Solution**

Put  $\phi(x) = x^2$  in Theorem 7.10.3, and we will have another remarkable conclusion which says the square of  $f$ ,  $f^2$ , is integrable w.r.t.  $\alpha$  whenever  $f \in \mathfrak{R}(\alpha)$ .

**Theorem 7.10.5**

Suppose  $\alpha$  is increasing. If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then its square  $f^2$  is also integrable w.r.t.  $\alpha$ , i.e.,  $f^2 \in \mathfrak{R}(\alpha)$  on  $[a, b]$ .



**Proof** Because  $\phi(x) = x^2$  is continuous on  $\mathbb{R}$ , it is, of course, continuous on the closed interval joining the infimum and supremum of  $f$ ,  $[m, M]$ . Then by Theorem 7.10.3,  $f^2 \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . ■

An alternative proof without using Theorem 7.10.3 is left as an exercise.

☞ **Exercise 7.5** Prove Theorem 7.10.5 without using Theorem 7.10.3. But feel free to use Theorem 7.10.4.

**Solution**

An immediate consequence of the above theorem is as follows, which states the product of two integrable functions is also integrable.

**Theorem 7.10.6**

Suppose  $\alpha$  is increasing. If  $f, g \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then  $fg \in \mathfrak{R}(\alpha)$  on  $[a, b]$ .



**Proof** Note that the product  $fg$  can be written as

$$fg = \frac{1}{4}[(f+g)^2 - (f-g)^2]$$

Then this theorem follows directly from Theorem 7.2.1 and Theorem 7.10.5. ■

## 7.10.4 Riemann-Stieltjes Integral with a Variable Upper Limit

We know by Theorem 7.10.6 that the product  $fg$  is integrable w.r.t.  $\alpha$  whenever  $f$  and  $g$  are integrable. The next theorem provides an alternative view of the integral  $\int_a^b fg d\alpha$ . As we will see, we can define a function  $G$  in such a way that the symbol  $g d\alpha$  in the original integral can be replaced with  $dG$ . Or we can replace  $f d\alpha$  with  $dF$  by defining a function  $F$ .

**Theorem 7.10.7**

Suppose that  $\alpha$  is increasing, and  $f, g \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . Define

$$F(x) = \int_a^x f d\alpha \quad \text{and} \quad G(x) = \int_a^x g d\alpha, \quad x \in [a, b]$$

Then  $f \in \mathfrak{R}(G)$  and  $g \in \mathfrak{R}(F)$  on  $[a, b]$  (of course  $fg \in \mathfrak{R}(\alpha)$  by Theorem 7.10.6), and we have

$$\int_a^b fg d\alpha = \int_a^b f dG = \int_a^b g dF$$



**Proof** We only prove  $f \in \mathfrak{R}(G)$  and  $\int_a^b fg d\alpha = \int_a^b f dG$ . Suppose  $g$  is bounded by some positive number, say  $A$ , i.e.,  $|g(x)| < A \forall x \in [a, b]$ . Given  $\varepsilon > 0$ , since  $fg \in \mathfrak{R}(\alpha)$ , there exists a partition  $P_\varepsilon$  on  $[a, b]$  such that

$$|U(P, f, \alpha) - L(P, f, \alpha)| < \varepsilon/A \tag{7.49}$$

## 7.10 Properties of Riemann-Stieltjes Integrals With Respect to Increasing Integrators

holds for any refinement  $P$  of  $P_\varepsilon$ . Let  $P = \{x_0, x_1, \dots, x_n\} \supseteq P_\varepsilon$  and  $t_k \in [x_{k-1}, x_k]$ . We have

$$\begin{aligned} S(P, f, G) &= \sum_k f(t_k) \Delta G \\ &= \sum_k f(t_k) \int_{x_{k-1}}^{x_k} g \, d\alpha \\ &= \sum_k \int_{x_{k-1}}^{x_k} f(t_k) g(t) \, d\alpha(t) \end{aligned} \tag{7.50}$$

On the other hand, we can write

$$\int_a^b f g \, d\alpha = \sum_k \int_{x_{k-1}}^{x_k} f(t) g(t) \, d\alpha(t) \tag{7.51}$$

Subtracting (7.50) and (7.51) yields

$$S(P, f, G) - \int_a^b f g \, d\alpha = \sum_k \int_{x_{k-1}}^{x_k} [f(t_k) - f(t)] g(t) \, d\alpha(t)$$

It then follows from Theorem 7.10.4 that

$$\begin{aligned} \left| S(P, f, G) - \int_a^b f g \, d\alpha \right| &\leq \sum_k \left| \int_{x_{k-1}}^{x_k} [f(t_k) - f(t)] g(t) \, d\alpha(t) \right| \\ &\leq \sum_k \int_{x_{k-1}}^{x_k} |f(t_k) - f(t)| |g(t)| \, d\alpha(t) \\ &\leq \sum_k \int_{x_{k-1}}^{x_k} (M_k - m_k) |g(t)| \, d\alpha(t) \\ &= \sum_k (M_k - m_k) \int_{x_{k-1}}^{x_k} |g(t)| \, d\alpha(t) \end{aligned}$$

Recall  $|g| \leq A$ . Using Theorem 7.10.2, we obtain

$$\begin{aligned} \left| S(P, f, G) - \int_a^b f g \, d\alpha \right| &\leq \sum_k (M_k - m_k) \int_{x_{k-1}}^{x_k} A \, d\alpha(t) \\ &= A \sum_k (M_k - m_k) \Delta \alpha_k \\ &= A(U(P, f, \alpha) - L(P, f, \alpha)) \end{aligned}$$

Hence,

$$\left| S(P, f, G) - \int_a^b f g \, d\alpha \right| < \varepsilon$$

by (7.49). Therefore,  $f$  is integrable w.r.t.  $G$  on  $[a, b]$  with  $\int_a^b f \, dG = \int_a^b f g \, d\alpha$ . ■

**Example 7.7** Let

$$\alpha(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}, \quad -1 \leq x \leq 1$$

Suppose  $f$  and  $g$  are any two functions defined on  $[-1, 1]$  being continuous from the left at  $x = 0$ .

By Theorem 7.6.1 and Theorem 7.10.1, we know both functions are integrable w.r.t.  $\alpha$  on  $[-1, x]$  if

$-1 < x \leq 1$  with

$$F(x) = \int_{-1}^x f \, d\alpha = \begin{cases} f(0) & x \geq 0 \\ 0 & x < 0 \end{cases} \quad \text{and} \quad G(x) = \int_{-1}^x g \, d\alpha = \begin{cases} g(0) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Note that  $F$  and  $G$  are also step functions that are continuous from the right at  $x = 0$ . Hence, by applying Theorem 7.6.1 again, we have

$$\int_{-1}^1 f \, dG = f(0)g(0) \quad \text{and} \quad \int_{-1}^1 g \, dF = f(0)g(0)$$

On the other hand, since the product  $fg$  is continuous from the left at  $x = 0$ , it follows from Theorem 7.6.1 that

$$\int_{-1}^1 fg \, d\alpha = f(0)g(0)$$

The results agree with Theorem 7.10.7 as expected.

## 7.11 Integrators of Bounded Variation

If  $f \in \mathfrak{R}(\alpha_1)$  and  $f \in \mathfrak{R}(\alpha_2)$  on  $[a, b]$  where  $\alpha_1$  and  $\alpha_2$  are both increasing functions, then we know from Theorem 7.2.2 that  $f \in \mathfrak{R}(\alpha)$  where  $\alpha = \alpha_1 - \alpha_2$ . Moreover, by Theorem 6.4.1, we are immediately informed that  $\alpha$  is of bounded variation.

Now, let us think about the converse of this statement. Suppose  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  where  $\alpha$  is of bounded variation. Can we assert that  $f \in \mathfrak{R}(\alpha_1)$  and  $f \in \mathfrak{R}(\alpha_2)$  on  $[a, b]$  where  $\alpha_1$  and  $\alpha_2$  are any increasing functions satisfying  $\alpha = \alpha_1 - \alpha_2$ ? Unfortunately, this is not true in general. For example, the Dirichlet function  $\mathbb{1}_{\mathbb{Q}}$  is integrable w.r.t.  $\alpha(x) = 0$  on  $[0, 1]$ . We can write  $\alpha(x) = 0 = x - x$ . But clearly  $\mathbb{1}_{\mathbb{Q}} \notin \mathfrak{R}$  on  $[0, 1]$ .

However, as the next theorem will state,  $f$  is always integrable w.r.t. the total variation  $V_a^x(\alpha)$  of  $\alpha$  provided that  $f \in \mathfrak{R}(\alpha)$ .

### Theorem 7.11.1

If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then  $f \in \mathfrak{R}(V_a^x(\alpha))$  on  $[a, b]$  where  $V_a^x(\alpha)$  is the total variation of  $\alpha$ , which is regarded as a function of  $x$  as in Definition 6.3.1.



Before presenting the proof, we first introduce some notations we shall use as well as the general strategy of this proof.

For a sub-interval  $[x_{k-1}, x_k]$  in the partition  $P = \{x_0, \dots, x_n\}$  on  $[a, b]$ , we adopt the following notation:

$$M_k = \sup_{x \in [x_{k-1}, x_k]} f(x), \quad m_k = \inf_{x \in [x_{k-1}, x_k]} f(x) \quad \text{and} \quad \Delta V_k = V_{x_{k-1}}^{x_k}(\alpha)$$

And we write  $V(x)$  instead of  $V_a^x(\alpha)$  to emphasize that the total variation is a function of  $x$ .

Observe that  $V(x)$  is monotonically increasing, hence we shall prove  $f$  is integrable w.r.t.  $V(x)$  mainly relies on Riemann's criterion (Theorem 7.9.1). We intend to bound the difference between the

upper and the lower sum, which is shown in the following, with an arbitrarily small positive number.

$$\begin{aligned} U(P, f, V) - L(P, f, V) &= \sum_k (M_k - m_k) \Delta V_k \\ &= \sum_k (M_k - m_k) (\Delta V_k - |\Delta \alpha_k|) + \sum_k (M_k - m_k) |\Delta \alpha_k| \end{aligned}$$

Therefore, we shall focus on bounding each of the two terms on the right-hand side of the above equation.

We now present the proof as follows.

**Proof** If  $f$  is a constant function, then this theorem holds trivially. We assume that  $f$  is non-constant in the rest of the proof. Let  $A$  be given by

$$A = \sup_{x \in [a, b]} f(x) - \inf_{x \in [a, b]} f(x)$$

Since  $f$  is non-constant, it is clear that  $A > 0$ .

Let  $\varepsilon > 0$  be chosen arbitrarily. Since  $\alpha$  is of bounded variation on  $[a, b]$ , by the definition of total variations, there exists a partition  $P'_\varepsilon$  on  $[a, b]$  such that

$$V_a^b(\alpha) < V(P, \alpha) + \frac{\varepsilon}{3A} \quad \forall P \supseteq P'_\varepsilon \quad (7.52)$$

We shall use this inequality later.

Now, we prepare another inequality that will be in use. Because  $f \in \mathfrak{R}(\alpha)$ , there exists a partition  $P''_\varepsilon$  such that

$$\left| \sum (f(s_k) - f(t_k)) \Delta \alpha_k \right| < \frac{\varepsilon}{3} \quad \forall P \supseteq P''_\varepsilon \quad (7.53)$$

where  $s_k$  and  $t_k$  are any two points in the  $k$ -th subinterval of  $P$ .

Let  $P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$ , and let  $P = \{x_0, \dots, x_n\} \supseteq P_\varepsilon$  be a refinement. We have

$$\begin{aligned} U(P, f, V) - L(P, f, V) &= \sum_{k=1}^n (M_k - m_k) \Delta V_k \\ &= \sum_{k=1}^n (M_k - m_k) (\Delta V_k - |\Delta \alpha_k|) + \sum_{k=1}^n (M_k - m_k) |\Delta \alpha_k| \end{aligned}$$

Recall  $A = \sup f - \inf f$ . Hence, the right-hand side of the equality above can be further enlarged to

$$U(P, f, V) - L(P, f, V) \leq A \sum_{k=1}^n (\Delta V_k - |\Delta \alpha_k|) + \sum_{k=1}^n (M_k - m_k) |\Delta \alpha_k| \quad (7.54)$$

Note that  $P \supseteq P_\varepsilon \supseteq P'_\varepsilon$ . It then follows from (7.52) that

$$\sum_{k=1}^n \Delta V_k = V_a^b(\alpha) < V(P, \alpha) + \frac{\varepsilon}{3A} = \sum_{k=1}^n |\Delta \alpha_k| + \frac{\varepsilon}{3A}$$

Rearranging the terms, we obtain

$$\sum_{k=1}^n (\Delta V_k - |\Delta \alpha_k|) < \frac{\varepsilon}{3A} \quad (7.55)$$

By combining (7.54) and (7.55), we have

$$U(P, f, V) - L(P, f, V) < \frac{\varepsilon}{3} + \sum_{k=1}^n (M_k - m_k) |\Delta \alpha_k| \quad (7.56)$$

It requires much more work to bound the term  $\sum_{k=1}^n (M_k - m_k) |\Delta \alpha_k|$  in (7.56). Let  $h$  be a

positive number given by

$$h = \frac{\varepsilon}{3(V_a^b(\alpha) + 1)}$$

Let  $I(P), J(P) \subseteq \{0, 1, \dots, n\}$  be defined as

$$I(P) = \{k \mid \Delta\alpha_k \geq 0\} \quad J(P) = \{k \mid \Delta\alpha_k < 0\}$$

For  $k \in I(P)$ , we can choose a pair of points,  $s_k$  and  $t_k$ , in the  $k$ -th subinterval of  $P$  such that

$$M_k - m_k < f(s_k) - f(t_k) + h$$

In this case,

$$(M_k - m_k) |\Delta\alpha_k| < (f(s_k) - f(t_k) + h) |\Delta\alpha_k| = (f(s_k) - f(t_k) + h) \Delta\alpha_k \quad (7.57)$$

On the other hand, for  $k \in J(P)$ , we can choose a pair of points,  $s_k$  and  $t_k$ , in the  $k$ -th subinterval of  $P$  such that

$$M_k - m_k < -f(s_k) + f(t_k) + h$$

It then follows that

$$(M_k - m_k) |\Delta\alpha_k| < (-f(s_k) + f(t_k) + h) |\Delta\alpha_k| = (f(s_k) - f(t_k) - h) \Delta\alpha_k \quad (7.58)$$

Compare (7.57) and (7.58), we conclude

$$(M_k - m_k) |\Delta\alpha_k| < (f(s_k) - f(t_k)) \Delta\alpha_k + h |\Delta\alpha_k| \quad \forall k \in \{0, 1, \dots, n\}$$

Therefore,

$$\begin{aligned} \sum_{k=1}^n (M_k - m_k) |\Delta\alpha_k| &< \sum_{k=1}^n (f(s_k) - f(t_k)) \Delta\alpha_k + h \sum_{k=1}^n |\Delta\alpha_k| \\ &\leq \sum_{k=1}^n (f(s_k) - f(t_k)) \Delta\alpha_k + h V_a^b(\alpha) \\ &< \frac{\varepsilon}{3} + \sum_{k=1}^n (f(s_k) - f(t_k)) \Delta\alpha_k \\ &\leq \frac{\varepsilon}{3} + \left| \sum_{k=1}^n (f(s_k) - f(t_k)) \Delta\alpha_k \right| \end{aligned}$$

In summary,

$$\sum_{k=1}^n (M_k - m_k) |\Delta\alpha_k| < \frac{\varepsilon}{3} + \left| \sum_{k=1}^n (f(s_k) - f(t_k)) \Delta\alpha_k \right| \quad (7.59)$$

Since  $P \supseteq P_\varepsilon \supseteq P_\varepsilon''$ , we may apply (7.53) to (7.59) and obtain that

$$\sum_{k=1}^n (M_k - m_k) |\Delta\alpha_k| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3} \quad (7.60)$$

Finally, by combining (7.56) and (7.60), we have

$$U(P, f, V) - L(P, f, V) < \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon$$

Therefore,  $f \in \mathfrak{R}(V(x))$  by Theorem 7.9.1 since  $V(x)$  is increasing. ■

Recall that all the theorems proved in previous section are under the common assumption that the integrator  $\alpha$  is increasing. Now, thanks to Theorem 7.11.1, we can easily show that some of

them are also valid if we only assume  $\alpha$  is of bounded variation. For example, we can generalize Theorem 7.10.1 as follows.

**Theorem 7.11.2**

*Suppose  $\alpha$  is of bounded variation on  $[a, b]$ . Let  $[c, d] \subseteq [a, b]$  be a subinterval. If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then we also have  $f \in \mathfrak{R}(\alpha)$  on  $[c, d]$ .*



**Proof** Let  $V(x)$  denote  $V_a^x(\alpha)$ . By Theorem 7.11.1, we know  $f \in \mathfrak{R}(V)$  and  $f \in \mathfrak{R}(V - \alpha)$  on  $[a, b]$ . Since both  $V$  and  $V - \alpha$  are increasing functions, by applying Theorem 7.10.1 twice, we have  $f \in \mathfrak{R}(V)$  and  $f \in \mathfrak{R}(V - \alpha)$  on  $[c, d]$ . Finally, by Theorem 7.2.2, it follows that  $f \in \mathfrak{R}(\alpha)$  on  $[c, d]$ .



Theorem 7.10.3, 7.10.5, 7.10.6 and 7.10.7 can also be generalized similarly.

## 7.12 Existence of Riemann-Stieltjes Integrals

### 7.12.1 Sufficient Conditions

In a Calculus course, the first theorem about integrability one might be taught is that  $f$  is Riemann integrable if it is continuous. When considering the Riemann-Stieltjes integrals, we impose one more condition that the integrator  $\alpha$  is of bounded variation.

**Theorem 7.12.1**

*If  $f$  is continuous, and  $\alpha$  is of bounded variation on  $[a, b]$ , then  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ .*



**Proof** If  $\alpha$  is constant, then the conclusion is trivial. Assume now  $\alpha$  is non-constant. By Theorem 6.4.1 and Theorem 7.2.1, it suffices to prove this theorem when  $\alpha$  is increasing.

Given  $\varepsilon > 0$ , since  $f$  is continuous on  $[a, b]$ , there exists  $\delta > 0$  such that

$$|s - t| < \delta \implies |f(s) - f(t)| < \frac{\varepsilon/2}{\alpha(b) - \alpha(a)}$$

Let  $P_\varepsilon$  be a partition on  $[a, b]$  with  $\|P_\varepsilon\| < \delta$ . Let  $P$  be any refinement of  $P_\varepsilon$ . On each subinterval  $[x_{k-1}, x_k]$ , we have

$$|f(s) - f(t)| < \frac{\varepsilon/2}{\alpha(b) - \alpha(a)} \quad \forall s, t \in [x_{k-1}, x_k]$$

Taking the supremum on both sides yields

$$\sup_{s,t \in [x_{k-1}, x_k]} |f(s) - f(t)| \leq \frac{\varepsilon/2}{\alpha(b) - \alpha(a)} < \frac{\varepsilon}{\alpha(b) - \alpha(a)}$$

Applying Theorem 1.1.1 to the left-hand side, we find

$$M_k - m_k < \frac{\varepsilon}{\alpha(b) - \alpha(a)}$$

It then follows that

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_k (M_k - m_k) \Delta \alpha_k < \frac{\varepsilon}{\alpha(b) - \alpha(a)} \sum_k \Delta \alpha_k = \varepsilon$$

Therefore,  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  since it satisfies Riemann's criterion (Theorem 7.9.1). ■

Thanks to integration by parts, we can get for free an additional theorem of existence of Riemann integrals from Theorem 7.12.1. That is,  $\int f \, dx$  exists if  $f$  is of bounded variation.

### Theorem 7.12.2

*The Riemann integral  $\int_a^b f \, dx$  exists if one of the following conditions is satisfied.*

1.  *$f$  is continuous on  $[a, b]$ .*
2.  *$f$  is of bounded variation on  $[a, b]$ .*



**Proof** If  $f$  is continuous on  $[a, b]$ , then  $f \in \mathfrak{R}$  on  $[a, b]$  by Theorem 7.12.1 since the identity function is of bounded variation.

If  $f$  is of bounded variation on  $[a, b]$ , then by Theorem 7.12.1,  $x \in \mathfrak{R}(f)$  on  $[a, b]$  since the identity function is continuous. Then by applying Theorem 7.4.1, we also have  $f \in \mathfrak{R}$  on  $[a, b]$ . ■

## 7.12.2 Necessary Conditions

## 7.13 Mean Value Theorems for Riemann-Stieltjes Integrals

In the case of Riemann integrals, the **average** of function  $f$  over an interval  $[a, b]$  is defined as the *average value* of the integral  $\int_a^b f(x) \, dx$ , i.e.,

$$\frac{1}{b-a} \int_a^b f(x) \, dx$$

We can extend this concept for Riemann-Stieltjes integrals.

### Theorem 7.13.1 (First Mean Value Theorem)

*Suppose  $\alpha$  is increasing and  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . Denote the infimum and the supremum of  $f$  on  $[a, b]$  by  $m$  and  $M$ , respectively. Then there exists a number  $A$  satisfying  $m \leq A \leq M$  such that*

$$\int_a^b f \, d\alpha = A \int_a^b \alpha = A[\alpha(b) - \alpha(a)] \quad (7.61)$$

*In particular, if  $f$  is continuous on  $[a, b]$ , then  $f$  can attain this average value, that is, there exists  $c \in [a, b]$  such that  $f(c) = A$ .*



**Proof** If  $\alpha$  is constant, then  $\int_a^b f \, d\alpha = 0$ . Then the theorem can be proved by taking  $A = f(a)$ .

Now, we assume  $\alpha$  is non-constant. Applying Theorem 7.10.2, we have

$$m \int_a^b \alpha \leq \int_a^b f \, d\alpha \leq M \int_a^b \alpha$$

Let

$$A = \frac{1}{\alpha(b) - \alpha(a)} \int_a^b f \, d\alpha$$

Clearly,  $m \leq A \leq M$  and (7.61) holds.

If  $f$  is continuous, then by the intermediate value theorem, there exists  $c \in [a, b]$  such that  $f(c) = A$ . ■

☞ **Exercise 7.6** Suppose  $f$  is continuous, and the integrator  $\alpha$  is non-constant and increasing on  $[a, b]$ .

If  $\int_a^b f \, d\alpha = 0$ , show that  $f$  has at least one zero on  $[a, b]$ .

**Solution** First, note that  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$  by Theorem 7.12.1 since  $f$  is continuous and  $\alpha$  is of bounded variation. Then by applying Theorem 7.13.1, there exists  $c \in [a, b]$  such that  $f(c) = 0$ .

A discontinuous function need not assume its average value. The following is an example.

**Example 7.8** Let

$$f(x) = \begin{cases} 1, & \text{if } x \leq 0 \\ 0, & \text{if } x > 0 \end{cases}, \quad x \in [-1, 1]$$

Applying Theorem 7.6.1 and Theorem 7.4.1, we can evaluate the integral:

$$\int_{-1}^1 f \, dx = 1$$

Hence, the average of  $f$  over  $[-1, 1]$  is  $1/2$ . But clearly  $f$  does not assume  $1/2$ .

Now, by replacing the integrator in the previous example with a step function, we will have a different result. See the following example.

**Example 7.9** Let

$$f(x) = \begin{cases} 1, & \text{if } x \leq 0 \\ 0, & \text{if } x > 0 \end{cases} \quad \text{and} \quad \alpha(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases}, \quad x \in [-1, 1]$$

We have

$$\int_{-1}^1 f \, d\alpha = f(0)[\alpha(1) - \alpha(-1)] = 1$$

The average of  $f$ , w.r.t.  $\alpha$  is now

$$\frac{1}{\alpha(1) - \alpha(-1)} \int_{-1}^1 f \, d\alpha = 1$$

In this example, though  $f$  is not continuous, we still have  $f(0) = 1$ .

Review Theorem 7.6.1. Actually, in general, if the integrator  $\alpha$  is constant except at point  $c$ , and one of  $f$  and  $\alpha$  is continuous from the left at  $c$  and one of  $f$  and  $\alpha$  is continuous from the right at  $c$ , then we always have

$$\int_a^b f \, d\alpha = f(c)[\alpha(b) - \alpha(a)]$$

That is,  $f(c)$  is exactly the average of  $f$  w.r.t.  $\alpha$  on  $[a, b]$ .

By applying the first mean value theorem as well as integration by parts, we immediately obtain the following theorem, which is known as the second mean value theorem for Riemann-Stieltjes integrals

### Theorem 7.13.2 (Second Mean Value Theorem)

Suppose  $f$  is monotonic and  $\alpha$  is continuous on  $[a, b]$ . If  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , then there exists  $c \in [a, b]$  such that

$$\int_a^b f \, d\alpha = f(a) \int_a^c d\alpha + f(b) \int_c^b d\alpha \tag{7.62}$$

Recall the geometric interpretation of the Riemann-Stieltjes integral. Then this theorem mainly

says that we can find two rectangles with maximum and minimum heights of the  $(x, \alpha, f)$ -fence, respectively, such that the sum of their areas equals the area of the shadow of the fence, i.e.,  $\int_a^b f d\alpha$ . See Figure 7.3.

**Proof** It suffices to prove this theorem when  $f$  is increasing. By Theorem 7.4.1, we know  $\alpha \in \mathfrak{R}(f)$  on  $[a, b]$ , and

$$\int_a^b f d\alpha = - \int_a^b \alpha df + f(b)\alpha(b) - f(a)\alpha(a) \quad (7.63)$$

Since  $\alpha$  is continuous and  $f$  is increasing, it then follows from Theorem 7.13.1 that

$$\int_a^b \alpha df = \alpha(c)[f(b) - f(a)] \quad (7.64)$$

Equation (7.62) then follows immediately from (7.63) and (7.64).  $\blacksquare$

**Example 7.10** Let

$$f(x) = \frac{1}{1 + e^x} \quad \text{and} \quad \alpha(x) = e^x, \quad x \in [-1, 1]$$

It is known that

$$\int_{-1}^1 f d\alpha = 1$$

Since  $f$  is decreasing and  $\alpha$  is continuous, the second mean value theorem (Theorem 7.13.2) is then applicable. And in this example, we can find the value of  $c$  in (7.62) explicitly.

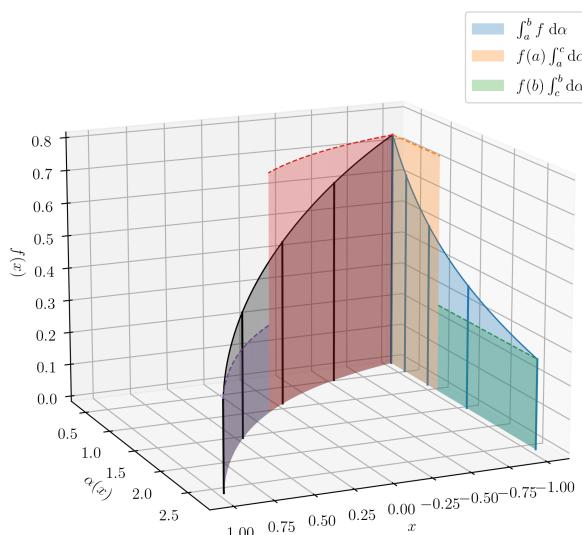
Note that the values of  $f$  and  $\alpha$  at the endpoints are

$$f(-1) = \frac{e}{e+1}, \quad f(1) = \frac{1}{e+1}, \quad \alpha(-1) = \frac{1}{e}, \quad \text{and} \quad \alpha(1) = e$$

By solving (7.62), we find

$$c = \ln \frac{2}{e-1}$$

Figure 7.3 depicts the  $(x, \alpha, f)$ -fence of this example.



**Figure 7.3:** Geometric interpretation of the second mean value theorem for Riemann-Stieltjes integral. The blue area equals the sum of the areas of the orange and green rectangles.

## 7.14 Fundamental Theorems of Calculus

The **first fundamental theorem of calculus** establishes the relation between the antiderivative of function  $f$  and its definite integral. The following theorem is a slightly generalized version where the integrator is any function of bounded variation.

### Theorem 7.14.1

Let  $\alpha$  be of bounded variation on  $[a, b]$ . Suppose that  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ . Define

$$F(x) = \int_a^x f \, d\alpha, \quad x \in [a, b]$$

(Note that  $F$  is well-defined by Theorem 7.11.2.) We have the following:

1.  $F$  is of bounded variation on  $[a, b]$ .
2.  $F$  is continuous at the point where  $\alpha$  is continuous.
3. If  $\alpha$  is increasing, then  $F'(x)$  exists at each  $x \in (a, b)$  where  $\alpha'(x)$  exists and  $f$  is continuous. For such  $x$ , we have

$$F'(x) = f(x)\alpha'(x)$$



**Note** Let us consider the case in statement 3 where  $x$  is one of the endpoints, i.e.,  $x = a$  or  $x = b$ . Suppose  $x = a$ . If one-sided derivative  $\alpha'_+(a)$  and the one-sided limit  $f(a+)$  both exist and are finite, then  $F'(a+)$  also exists, and

$$F'_+(a) = f(a+)\alpha'_+(a)$$

This can be proved by changing the limit processes to the one-sided limits in the following proof we are about to present. An analogous result also holds for  $x = b$ .

Notice also that if  $F$  is defined as

$$F(x) = \int_c^x f \, d\alpha, \quad x \in [a, b]$$

where  $c$  is any point in  $[a, b]$ , then all the conclusions also hold, as an immediate consequence of this theorem. To see this, we write

$$F(x) = \int_a^x f \, d\alpha + \int_c^a f \, d\alpha$$

**Proof** It suffices to prove this theorem when  $\alpha$  is increasing.

(Proof of 1) Let  $P$  be any partition on  $[a, b]$ . We have

$$V(P, F) = \sum_k \left| \int_{x_{k-1}}^{x_k} f \, d\alpha \right| \leq \sum_k \int_{x_{k-1}}^{x_k} |f| \, d\alpha = \int_a^b |f| \, d\alpha$$

Hence,  $F$  is of bounded of variation on  $[a, b]$  since the right-hand side does not depend on  $P$ .

(Proof of 2) There exists  $M > 0$  such that  $|f| \leq M$ . Suppose  $\alpha$  is continuous at  $c$ . We have

$$|F(x) - F(c)| = \left| \int_c^x f \, d\alpha \right| \leq \int_c^x |f| \, d\alpha \leq M[\alpha(x) - \alpha(c)]$$

By taking  $x \rightarrow c$  on both sides, we see that the limit  $\lim_{x \rightarrow c} F(x)$  exists, and it equals  $F(c)$ . Hence,  $F$  is also continuous at  $c$ .

(Proof of 3) Suppose  $\alpha'(c)$  exists and  $f$  is continuous at  $c$ . Applying the first mean value theorem

(Theorem 7.13.1), there exists a number  $\xi$  in the closed interval joining  $x$  and  $c$  ( $x \neq c$ ) such that

$$F(x) - F(c) = \int_c^x f \, d\alpha = f(\xi)[\alpha(x) - \alpha(c)]$$

Dividing by  $(x - c)$  on both sides yields

$$\frac{F(x) - F(c)}{x - c} = f(\xi) \frac{\alpha(x) - \alpha(c)}{x - c}$$

Since  $f$  is continuous at  $c$  and  $\alpha'(c)$  exists. The limit of the left-hand side exists as  $x \rightarrow c$ . Hence,  $F'(c)$  exists. Letting  $x \rightarrow c$ , we have

$$F'(c) = \lim_{x \rightarrow c} \frac{F(x) - F(c)}{x - c} = f(c)\alpha'(c)$$

■

In particular, let  $\alpha(x) = x$ , and Theorem 7.14.1 will reduce to the traditional form the first fundamental theorem of calculus.

### Theorem 7.14.2 (First Fundamental Theorem of Calculus)

Suppose  $f \in \mathfrak{R}$  on  $[a, b]$ . Define

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

Then we have the following:

1.  $F$  is of bounded variation on  $[a, b]$ .
2.  $F$  is continuous on  $[a, b]$ .
3. If  $f$  is continuous at  $x \in (a, b)$ , then  $F'(x)$  exists and

$$F'(x) = f(x)$$



**Note** We have similar comments for this theorem to that of the preceding one.

In particular, if  $f$  is continuous on  $[a, b]$ , which is very common in practice, then the derivative of  $F$  exists everywhere (including the one-sided derivatives at the endpoints).

**Example 7.11** Let

$$f(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases}, \quad x \in [-1, 1]$$

Then

$$F(x) = \int_{-1}^x f(t) \, dt = 0$$

is constantly zero. Note that  $f$  is not continuous at 0. Though the derivative of  $F$  exists at 0, it does not equal  $f(0)$ . We have  $F'(0) = 0 \neq 1 = f(0)$ .

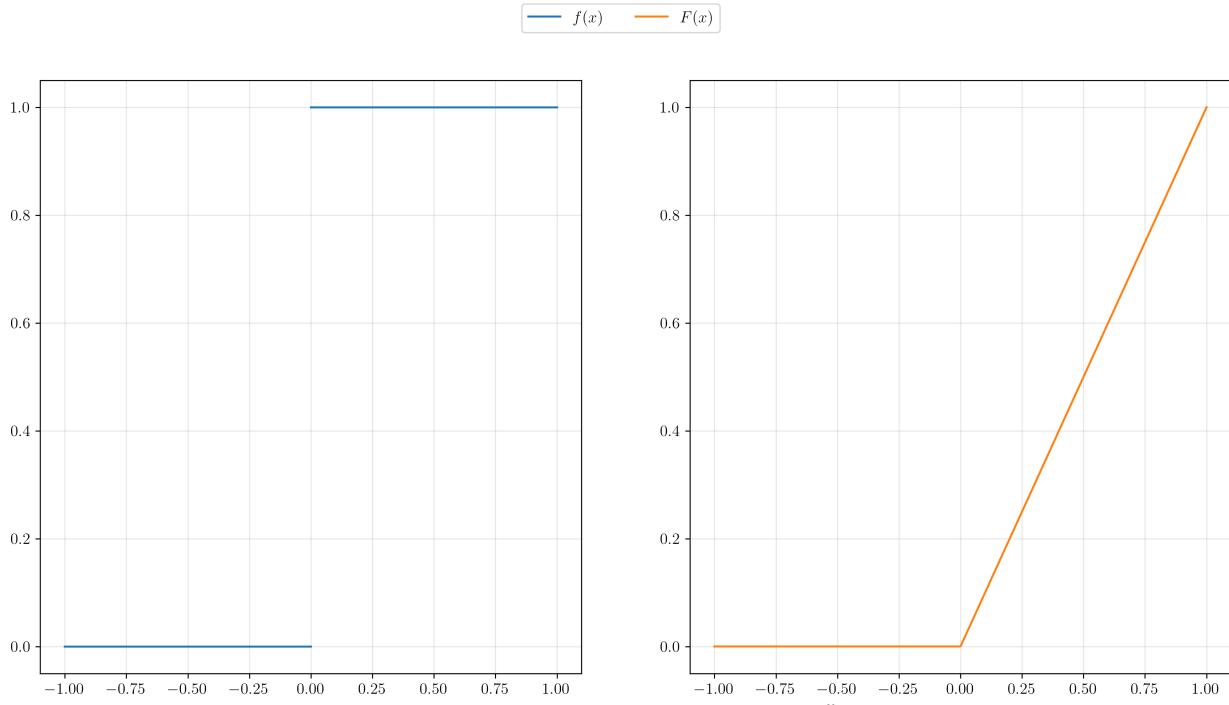
**Example 7.12** Let

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}, \quad x \in [-1, 1]$$

Then

$$F(x) = \int_{-1}^x f(t) dt = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Since  $f$  is not continuous at 0, in this example, we see that  $F'(0)$  does not even exist. See Figure 7.4



**Figure 7.4:** Left:  $f(x)$ . Right:  $F(x) = \int_{-1}^x f(t) dt$ .

The **second fundamental theorem of calculus** mainly tells us how to integrate a derivative.

### Theorem 7.14.3 (Second Fundamental Theorem of Calculus)

Suppose that  $f \in \mathfrak{R}$  on  $[a, b]$ . Let  $F$  be a function defined on  $[a, b]$  that is differentiable in  $(a, b)$  with

$$F'(x) = f(x) \quad \forall x \in (a, b)$$

Suppose also at the endpoints, the one-sided limits  $F(a+)$  and  $F(b-)$  both exist. Then we have

$$\int_a^b f(x) dx = F(b-) - F(a+) \tag{7.65}$$



**Proof** Given  $\varepsilon > 0$ , since  $f$  is integrable, it satisfies Riemann's criterion. That is, there exists a partition  $P = \{x_0, \dots, x_n\}$  on  $[a, b]$  such that

$$U(P, f) - L(P, f) < \varepsilon$$

On each subinterval  $[x_{k-1}, x_k]$ , applying the mean value theorem for derivatives (Theorem 5.4.7), we have

$$F(x_k-) - F(x_{k-1}+) = F'(\xi_k)\Delta x_k = f(\xi_k)\Delta x_k$$

Note that  $F(x_k+) = F(x_k-) = F(x_k)$  except possibly for two endpoints (when  $x_k$  equals  $a$  or  $b$ ) since  $F$  is continuous in  $(a, b)$  for it is differentiable there.

Summing up over  $k$  yields

$$F(b-) - F(a+) = \sum_{k=1}^n F(x_k-) - F(x_{k-1}+) = \sum_{k=1}^n f(\xi_k) \Delta x_k$$

It then follows that

$$\begin{aligned} U(P, f) - [F(b-) - F(a+)] &= \sum_{k=1}^n [M_k - f(\xi_k)] \Delta x_k \\ &\leq \sum_{k=1}^n (M_k - m_k) \Delta x_k = U(P, f) - L(P, f) < \varepsilon \end{aligned}$$

where  $M_k = \sup_{x \in [x_{k-1}, x_k]} f(x)$ . Note that the above inequality holds for any refinement  $P$  of  $P_\varepsilon$ .

Taking the infimum over partition  $P$ , we obtain

$$\int_a^b f(x) dx - [F(b-) - F(a+)] \leq \varepsilon$$

This proves (7.65) since  $\varepsilon > 0$  is arbitrary and that the upper integral equals  $\int_a^b f(x) dx$ . ■

Recall Theorem 7.3.1 allows us to replace  $d\alpha$  with  $\alpha' dx$ :

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

under the assumption that  $\alpha$  has a continuous derivative on  $[a, b]$ . Thanks to the second fundamental theorem of calculus, we now present a stronger version of this theorem without assuming  $\alpha'$  is continuous.

#### Theorem 7.14.4

Suppose  $f \in \mathfrak{R}$  on  $[a, b]$ . Let  $\alpha$  be continuous function on  $[a, b]$  whose derivative  $\alpha' \in \mathfrak{R}$  on  $[a, b]$ . Then  $f\alpha' \in \mathfrak{R}$  and  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , and

$$\int_a^b f d\alpha = \int_a^b f(x) \alpha'(x) dx \tag{7.66}$$



**Proof** First, we note that  $f\alpha' \in \mathfrak{R}$  on  $[a, b]$  by Theorem 7.10.6 since both  $f$  and  $\alpha'$  are integrable.

Now, we want to apply Theorem 7.10.7 to merge the symbol  $\alpha'(x)dx$ .



**Note** To apply Theorem 7.10.7, we need to define a function  $\int_a^x \alpha'(t) dt$ . Note that we are basically integrating a derivative. Hence, the second fundamental theorem may help.

By the second fundamental theorem of calculus (Theorem 7.14.3), we have

$$\int_a^x \alpha(t) dt = \alpha(x) - \alpha(a) \quad \forall x \in [a, b]$$

Then Theorem 7.10.7 implies that  $f \in \mathfrak{R}(\alpha(x) - \alpha(a))$  on  $[a, b]$ , and

$$\int_a^b f(x) \alpha'(x) dx = \int_a^b f(x) d[\alpha(x) - \alpha(a)]$$

Of course, from the right-hand side of the above equation, we immediately find that  $f \in \mathfrak{R}(\alpha)$  on  $[a, b]$ , and the right-hand side may reduce to  $\int_a^b f d\alpha$ . Therefore, (7.66) is proved. ■

## 7.15 Change of Variables in Riemann Integrals

### Theorem 7.15.1 (Change of Variables)

Suppose that function  $\phi$  with a continuous derivative on  $[c, d]$ , and that  $f$  is continuous on  $\phi([c, d])$ . Define

$$F(x) = \int_{\phi(c)}^x f(t) dt, \quad x \in \phi([c, d])$$

Then for each  $x$  in  $[c, d]$ , the integral  $\int_c^x f[\phi(t)]\phi'(t) dt$  exists and has the value  $F[\phi(x)]$ . In particular, putting  $x = d$  yields

$$\int_{\phi(c)}^{\phi(d)} f(t) dt = \int_c^x f[\phi(t)]\phi'(t) dt \quad (7.67)$$



Note that the integrand on right-hand side of (7.67) looks like the expression that the chain rule will produce. Indeed, the proof of this theorem differs from other theorems in this chapter in that it mainly relies on the chain rule, which is a technique from differentiation. And as one will expect, we shall use fundamental theorems of calculus to establish this connection.

**Proof** Consider the composite function  $F[\phi(x)]$ . Note that  $\phi'(x)$  exists (and is finite) for each  $x$  in  $[c, d]$  and that the derivative of  $F$  exists (and is finite) at  $\phi(x)$  by the first fundamental theorem of calculus (Theorem 7.14.2) since  $f$  is continuous there. Then the chain rule (Theorem 5.3.2) is applicable. We have

$$\frac{d}{dx} F[\phi(x)] = f[\phi(x)]\phi'(x) =: g(x)$$

Furthermore, note that  $g(x)$  is continuous on  $[c, d]$  since both  $f[\phi(x)]$  and  $\phi'(x)$  are. Hence,  $g \in \mathfrak{R}$  on  $[c, d]$ , which further implies that  $g$  is integrable on each subinterval  $[c, x]$ . Then by the second fundamental theorem of calculus (Theorem 7.14.3), we have

$$\int_c^x g(t) dt = F[\phi(x)] - F[\phi(c)] = F[\phi(x)]$$

Expanding expression of  $g(t)$  proves the theorem. ■

**Example 7.13** Consider the integral

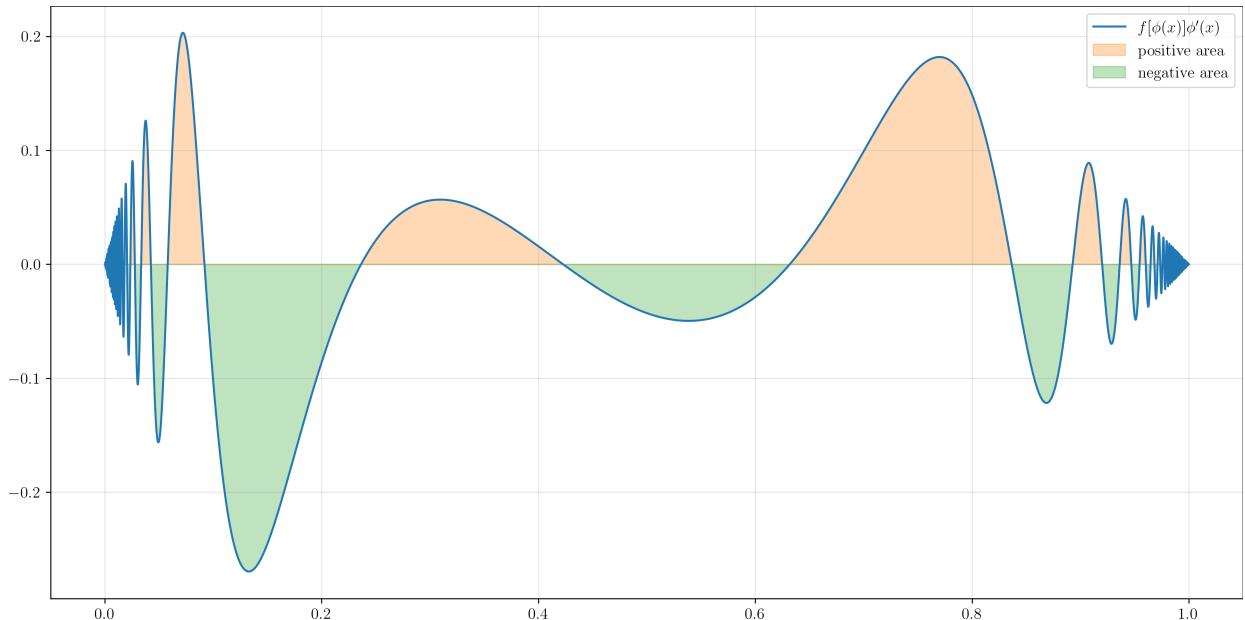
$$\int_0^1 f[x(x-1)(x-2)](3x^2 - 6x + 2) dx$$

where  $f$  is any continuous function on  $[0, 1]$ . Regarding the entire expression passed to  $f$  as a new variable, that is, letting  $\phi(x) = x(x-1)(x-2)$ , we find

$$\int_0^1 f[x(x-1)(x-2)](3x^2 - 6x + 2) dx = \int_0^1 f[\phi(x)]\phi'(x) dx = \int_{\phi(0)=0}^{\phi(1)=0} f(x) dx = 0$$

Figure 7.5 depicts the graph of  $f[\phi(x)]$  where  $f(x) = x \sin \frac{1}{x}$ .

Note that the function  $\phi(x) = x(x-1)(x-2)$  can be decomposed into finitely many piece-wise monotonic continuous functions. In fact, it is increasing on  $[0, (3 + \sqrt{3})/2]$  and is decreasing on  $[(3 + \sqrt{3})/2, 1]$ . Therefore, we can compute the integral by applying Theorem 7.5.1 two times on each subintervals without even using Theorem 7.15.1. Hence, it seems that we did not exploit the full power of Theorem 7.15.1



**Figure 7.5:** The integral is zero where  $f(x) = x \sin \frac{1}{x} \mathbb{1}\{x \neq 0\}$  and  $\phi(x) = x(x-1)(x-2)$ .

In the following, the function  $\phi$  cannot be decomposed into finitely many monotonic functions. In this case, Theorem 7.5.1 is no longer applicable.

**Example 7.14** Consider the integral

$$\int_0^{1/\pi} f(x^3 \sin \frac{1}{x} \mathbb{1}\{x \neq 0\})(3x^2 \sin \frac{1}{x} - x \cos \frac{1}{x}) \mathbb{1}\{x \neq 0\} dx$$

where  $f$  can be any continuous functions on  $[0, 1/\pi]$ . Let

$$\phi(x) = x^3 \sin \frac{1}{x} \mathbb{1}\{x \neq 0\}$$

Note that

$$\phi'(x) = \begin{cases} 3x^2 \sin \frac{1}{x} - x \cos \frac{1}{x}, & x > 0 \\ 0, & x = 0 \end{cases}, \quad x \in [0, 1/\pi]$$

Then by applying Theorem 7.15.1, the integral reduces to

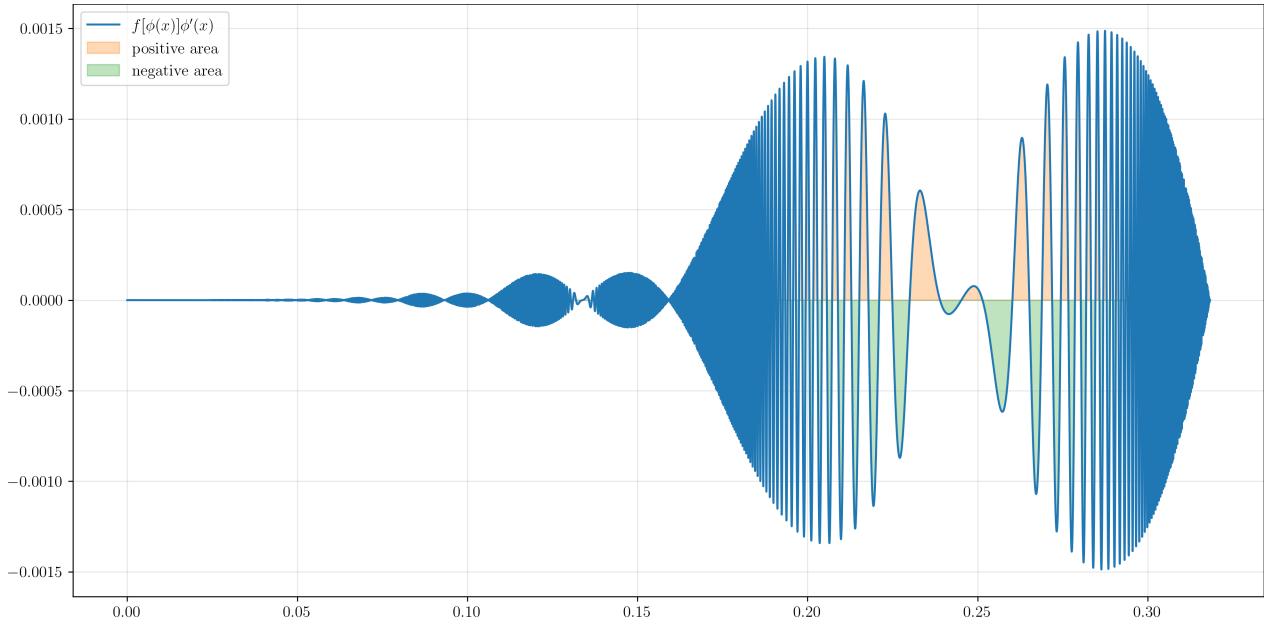
$$\int_{\phi(0)=0}^{\phi(1/\pi)=0} f(x) dx = 0$$

See Figure 7.6.

## 7.16 Lebesgue Criterion for Existence of Riemann Integrals

### Definition 7.16.1

A set  $S$  in  $\mathbb{R}$  is said to have measure zero, or a **zero-measure** set if for any given  $\varepsilon > 0$ , there exists a countable covering of  $S$  by open intervals the sum of whose lengths is bounded by  $\varepsilon$ . Mathematically, that is to say, there exists a family of intervals  $\{(a_k, b_k)\}_{k \in \Lambda}$  where the index



**Figure 7.6:** The integral is zero where  $f(x) = x \sin \frac{1}{x} \mathbb{1}\{x \neq 0\}$  and  $\phi(x) = x^3 \sin \frac{1}{x} \mathbb{1}\{x \neq 0\}$ .

set  $\Lambda$  is countable such that

$$S \subseteq \bigcup_{k \in \Lambda} (a_k, b_k) \quad \text{and} \quad \sum_{k \in \Lambda} (b_k - a_k) < \varepsilon \quad (7.68)$$



If the index set  $\Lambda$  is countably infinite, then we often take  $\Lambda = \mathbb{N}^*$ . In this way, (7.68) is written as

$$S \subseteq \bigcup_{k=1}^{\infty} (a_k, b_k) \quad \text{and} \quad \sum_{k=1}^{\infty} (b_k - a_k) < \varepsilon \quad (7.69)$$

On the other hand, if  $\Lambda$  is finite, then in order to write (7.68) as (7.69) for consistency, we append countably infinitely many degenerate open intervals, i.e., the empty sets  $\emptyset$  to the covering. Since taking the union with empty sets will not change the origin set, and the sum of lengths of empty sets is zero, (7.69) remains valid.

It is evident that the union of two zero-measure sets also has measure zero. By Mathematical induction, any finite unions of zero-measure sets also have measure zero. In fact, this is also the case for the countable union of zero-measure sets, as stated in the following theorem.

### Theorem 7.16.1

Let  $\{S_k\}$  be a countable family of zero-measure sets. Then their union also has measure zero.

**Proof** Without loss of generality, we consider countably infinite family  $\{S_k\}_{k=1}^{\infty}$  of zero-measure sets. Given  $\varepsilon > 0$ , for each  $S_k$ , there exists a family of open intervals  $\{I_{k,j} = (a_{k,j}, b_{k,j})\}_{j=1}^{\infty}$  (including degenerate ones), such that

$$S_k \subseteq \bigcup_{j=1}^{\infty} I_{k,j} \quad \text{and} \quad \sum_{j=1}^{\infty} (b_{k,j} - a_{k,j}) < \frac{\varepsilon}{2^{k+1}}$$

We have

$$\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} (b_{k,j} - a_{k,j}) \leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2} < \varepsilon$$

By Theorem 3.13.1, the double series  $\sum_{k,j} (b_{k,j} - a_{k,j})$  also converges (absolutely) and its sum is less than  $\varepsilon$ .

We can rearrange the doubly-indexed sets  $I_{k,j}$  into  $I_n$  by defining a bijection between  $\mathbb{N}^* \times \mathbb{N}^*$  and  $\mathbb{N}^*$ ,  $n = g(k, j)$ . It is clear that

$$\bigcup_{k=1}^{\infty} S_k \subseteq \bigcup_n I_n$$

Moreover,

$$\sum_{n=1}^{\infty} (b_n - a_n) = \sum_{k,j} (b_{k,j} - a_{k,j}) < \varepsilon$$

since the double series converges absolutely (Theorem 3.13.1). Therefore,  $\bigcup_{k=1}^{\infty} S_k$  indeed has measure zero.  $\blacksquare$

Next, we introduce the concept of oscillation, which has actually been already applied in several proofs from previous sections.

### Definition 7.16.2

*Let  $f$  be defined and bounded on an interval  $I$ . ( $I$  can be open, closed or half-open.) Let  $S \subseteq I$  be a subset. The number*

$$\Omega_f(S) = \sup_{x,y \in S} [f(x) - f(y)]$$

*is called the **oscillation of  $f$  on  $S$** .*



The oscillation of  $f$  on  $S$  measures the greatest difference between two function values. Recall Theorem 1.1.1. Equivalently, it is the difference between the supremum and the infimum. We have

$$\Omega_f(S) = \sup_{x \in S} f(x) - \inf_{x \in S} f(x)$$

Another simple observation is that if  $S \subseteq T$ , then

$$\Omega_f(S) \leq \Omega_f(T)$$

We can also define the **oscillation of  $f$  at a point  $x$** . Consider a neighborhood  $B_r(x) \cap I$  of  $x$  in  $I$ . Note that the oscillation of  $f$  on  $B_r(x) \cap I$ ,  $\Omega_f(B_r(x) \cap I)$ , will get smaller and smaller as the radius  $r$  decreases. Hence, we can regard

$$\Omega_f(B_r(x) \cap I) =: g(r)$$

as an increasing function of  $r$ . Then by Theorem 4.6.1, we know  $g(0+)$  exists.



**Note** Well, to apply Theorem 4.6.1, function  $g$  needs to be defined on a closed interval. Since the oscillation is always greater than or equal to zero, we may define  $g(0) = 0$  without violating the hypothesis that  $g$  is increasing. Then, we may apply Theorem 4.6.1 to  $g$  on  $[0, R]$  where  $R$  is any fixed positive number.

The oscillation of  $f$  at  $x$  is defined as precisely this limit.

**Definition 7.16.3**

Let  $f$  be defined and bounded on an interval  $I$ . The oscillation of  $f$  at point  $x \in I$  is defined by

$$\omega_f(x) = \lim_{r \rightarrow 0^+} \Omega_f(B_r(x) \cap I)$$



**Note** Beware of the difference between the symbols  $\Omega_f(\{x\})$  and  $\omega_f(x)$ . The former always equals zero while the latter does not. Hence, we do not describe the oscillation of  $f$  at a point  $x$  with the oscillation on the singleton  $\{x\}$ .

The following are some simple observations, which we shall use in the proofs without quoting:

1.  $\Omega_f(B_r(x) \cap I)$  decreases as  $r$  decreases (which we have already noted), and

$$\Omega_f(B_r(x) \cap I) > \omega_f(x) \quad \forall r > 0$$

2. The oscillation of  $f$  on a set is clearly no less than the oscillation at a point in it. Mathematically, if  $S \subseteq I$  and  $x \in S$ , then

$$\Omega_f(S) \geq \omega_f(x)$$

**Example 7.15** It is frequently stated that

$$f(x) = \sin \frac{1}{x} \mathbb{1}_{\{x \neq 0\}}$$

oscillates at 0. We can now measure the degree of this oscillation. For any  $r > 0$ , we can always find a large enough integer  $n \in \mathbb{N}^*$  such that

$$x_n = \frac{1}{\pi/2 + 2n\pi} \quad \text{and} \quad y_n = \frac{1}{-\pi/2 + 2n\pi}$$

are both in  $B_r(0)$ . Note that  $f(x_n) = 1$  and  $f(y_n) = -1$  are the supremum and infimum of  $f$ , respectively. Therefore,

$$\Omega_f(B_r(0)) = 1 - (-1) = 2 \quad \forall r > 0$$

And the oscillation of  $f$  at 0 is exactly

$$\omega_f(0) = 2$$

$\omega_f(x) = 0$  means  $f$  has no oscillation at  $x$ . Intuitively, this implies that  $f$  is continuous there.

**Theorem 7.16.2**

Let  $f$  be defined and bounded on interval  $I$ , and  $x \in I$  a point in it. Then  $\omega_f(x) = 0$  if and only if  $f$  is continuous at  $x$ .


**Proof**
**Theorem 7.16.3 (Lebesgue's Number Lemma)**


The following theorem states that if the oscillation of  $f$  at each point  $x$  in some closed interval (compact) is bounded by a small number  $\varepsilon$ , then there exists  $\delta > 0$  such that the oscillation of  $f$  on any subinterval whose length is less than  $\delta$  is also bounded by  $\varepsilon$ .

**Lemma 7.16.1**

Let  $f$  be defined and bounded on an closed interval  $I$ . Given  $\varepsilon > 0$ . Suppose that  $\omega_f(x) < \varepsilon$  for every  $x$  in  $I$ . Then there exists  $\delta > 0$  (only depending on  $\varepsilon$ ) such that

$$\Omega_f(J) < \varepsilon$$

where  $J$  is any subinterval in  $I$  whose length is less than  $\delta$ .


**Proof**

**Exercise 7.7** Provide a simpler proof of Theorem 7.16.1 without using Lebesgue's number.

**Solution** For any  $x \in I$ , by the definition of oscillation of  $f$  at a point, there exists  $\delta_x > 0$  (depending on  $x$ ) such that

$$\Omega_f(B_r(x) \cap I) < \varepsilon \quad \forall r \leq \delta_x$$

Note that the family of sets  $\{B_{\delta_x/2}(x) \mid x \in I\}$  forms an open cover of  $I$ . Since  $I$  is a closed interval, hence compact, there exists a finite subcover  $\{B_{\delta_i/2}(x_i) \mid i = 1, \dots, n\}$ .

**Note** If we adopt the previous notation, the radius of each open ball should be  $\delta_{x_i}/2$ . Here, we simply write  $\delta_i/2$  to ease the notation.

Let

$$\delta = \min \{\delta_i/2 \mid i = 1, \dots, n\}$$

Suppose that  $J$  is a subinterval of  $I$  of length less than  $\delta$ . Denote the left and right endpoints of  $J$  by  $c$  and  $d$ , respectively. ( $J$  may not include these endpoints.) Since  $\{B_{\delta_i/2}(x_i) \mid i = 1, \dots, n\}$  is an open cover of  $I$ , point  $c$  must fall in some member of the open over, say  $B_{\delta_i/2}(x_i)$ . Because

$$|d - x_i| \leq |d - c| + |c - x_i| < \delta + \delta_i/2 \leq \delta_i/2 + \delta_i/2 = \delta_i$$

point  $d$  is in  $B_{\delta_i}(x_i)$ . (Of course,  $c$  is also in  $B_{\delta_i}(x_i)$ .) It then follows that  $J \subseteq B_{\delta_i}(x_i)$ . Therefore,

$$\Omega_f(J) \leq \Omega_f(B_{\delta_i}(x_i) \cap I) < \varepsilon$$

This completes the proof.

**Lemma 7.16.2**

Let  $f$  be defined and bounded on a closed interval  $I$ . For each  $\varepsilon > 0$ , the set

$$J_\varepsilon = \{x \in I \mid \omega_f(x) \geq \varepsilon\}$$

is closed in  $I$ .



**Proof** Consider the complement of  $J_\varepsilon$  in  $I$ ,

$$J_\varepsilon^C = \{x \in I \mid \omega_f(x) < \varepsilon\}$$

Let  $x$  be an arbitrary point in  $J_\varepsilon^C$ . There exists  $\delta > 0$  such that

$$\Omega_f(B_\delta(x) \cap I) < \varepsilon$$

For any  $y \in B_\delta(x) \cap I$ , there exists  $r_y > 0$  such that

$$B_{r_y}(y) \subset B_\delta(x)$$

It then follows that

$$\omega_f(y) \leq \Omega_f(B_{r_y}(y) \cap I) \leq \Omega_f(B_\delta(x) \cap I) < \varepsilon$$

Note that this implies

$$B_\delta(x) \cap I \subseteq J_\varepsilon^C$$

by the definition of  $J_\varepsilon^C$ . Since  $B_\delta(x) \cap I$  is open (in  $I$ ) and contained in  $J_\varepsilon^C$  for each  $x$ ,  $J_\varepsilon^C$  is open in  $I$ . Therefore, equivalently,  $J_\varepsilon$  is closed in  $I$ .  $\blacksquare$

#### Theorem 7.16.4 (Lebesgue's Criterion for Reimann-Integrability)

*Let  $f$  be defined and bounded on  $[a, b]$ . Let  $D$  denote the set of discontinuities of  $f$  on  $[a, b]$ .*

*Then  $f \in \mathfrak{R}$  on  $[a, b]$  if and only if set  $D$  has measure zero.*



**Proof** (Necessity) We shall prove the necessity by proving the contrapositive. That is, we assume  $D$  does not have measure zero, and then show that  $f$  cannot be integrable.

Note that  $D$  can be written as

$$D = \bigcup_{r=1}^{\infty} D_r$$

where

$$D_r = \left\{ x \in [a, b] \mid \omega_f(x) \geq \frac{1}{r} \right\}$$

This is left as an exercise. (See Exercise 7.8.) Since  $D$  does not have measure zero, by Theorem 7.16.1, one member of the above union, say  $D_r$  does not have measure zero. Then there exists a positive number  $\varepsilon_0 > 0$  such that the sum of lengths of any countable open cover of  $D_r$  is greater than or equal to  $\varepsilon_0$ . That is,

$$D_r \subseteq \bigcup_{k=1}^{\infty} I_k \implies \sum_{k=1}^{\infty} I_k \geq \varepsilon_0 \quad (7.70)$$

Let  $\varepsilon = \varepsilon_0/r$ . Let  $P = \{x_0, \dots, x_n\}$  be any partition on  $[a, b]$ . Define

$$\Lambda = \{k = 1, \dots, n \mid (x_{k-1}, x_k) \cap D_r \neq \emptyset\}$$

Note that  $\{(x_{k-1}, x_k)\}_{k \in \Lambda}$  covers  $D_r$  except for possibly finitely many points. More specifically, we have

$$D_r \setminus P \subseteq \bigcup_{k \in \Lambda} (x_{k-1}, x_k)$$

Note that  $D_r \setminus P$  still does not have measure zero since  $P$  is only a finite set. It then follows from (7.70) that

$$\sum_{k \in \Lambda} \Delta x_k = \sum_{k \in \Lambda} (x_k - x_{k-1}) \geq \varepsilon_0$$

Moreover, for  $k \in \Lambda$ , we have

$$M_k(f) - m_k(f) = \Omega_f([x_{k-1}, x_k]) \geq \frac{1}{r}$$

since  $(x_{k-1}, x_k)$  contains some point in  $D_r$  at which the oscillation of  $f$  is greater than or equal to  $1/r$ .

Consider the difference  $U(P, f) - L(P, f)$ . We have

$$U(P, f) - L(P, f) \geq \sum_{k \in \Lambda} [M_k(f) - m_k(f)]\Delta x_k \geq \frac{1}{r} \sum_{k \in \Lambda} \Delta x_k \geq \frac{\varepsilon_0}{r} = \varepsilon$$

In summary, we have chosen an  $\varepsilon > 0$  such that

$$U(P, f) - L(P, f) \geq \varepsilon$$

for any partition  $P$  on  $[a, b]$ . Therefore,  $f$  fails to satisfy Riemann's criterion, and hence not integrable on  $[a, b]$ .

(Sufficiency) Now, we assume  $D$  has measure zero and show that  $f$  satisfies Riemann's criterion. Recall  $D$  can be written as  $D = \bigcup_{r=1}^{\infty} D_r$ . Given  $\varepsilon > 0$ , there exists  $r \in \mathbb{N}^*$  such that

$$r > (\Omega_f([a, b]) + b - a)/\varepsilon \quad (7.71)$$

Consider the set  $D_r$ . Note that  $D_r$  also has measure zero since  $D$  has measure zero. Thus, there exists a countable covering of  $D_r$  by open intervals the sum of whose lengths is less than  $1/r$ . Moreover, since  $D_r$  is compact by Lemma 7.16.2 and Theorem 7.16.2, there exists a *finite* subcover of  $D_r$ , say  $U_1, \dots, U_N$ . Note that the union  $\bigcup_{k=1}^n U_k$  can be written as a finite union of *disjoint* open intervals, say  $I_1, \dots, I_n$ . And the sum their lengths is of course also less than  $1/r$  since it is actually less than the sum of lengths of all  $U_k$ 's. In summary, we have

$$D_r \subseteq \bigcup_{k=1}^n I_k, \quad \text{and} \quad \sum_{k=1}^n \ell(I_k) < \frac{1}{r}$$

Let  $S$  denote the complement of  $\bigcup_{k=1}^n I_k$  in  $[a, b]$ . We claim that there exists  $\delta > 0$  such that  $S$  can be written as a finite union of closed intervals,  $J_1, \dots, J_m$ , satisfying the following:

1. Each pair of different closed intervals intersect at most one point, i.e.,  $J_i \cap J_j$  is either  $\emptyset$  or a singleton whenever  $i \neq j$ ,
2. the length of each closed interval is less than  $\delta$ , i.e.,  $\ell(J_k) < \delta$ , and
3.  $\Omega_f(J_k) < 1/r$  for each  $k$ .

The proof of this claim is left as an exercise. (See Exercise 7.9.)

Note that the union of endpoints of intervals  $I_k$ 's and  $J_k$ 's for a partition  $P = \{x_0, \dots, x_{m+n}\}$  on  $[a, b]$ .

 **Note** Since there are  $n$  open intervals  $I_1, \dots, I_n$  and  $m$  closed intervals  $J_1, \dots, J_m$ , one can verify that the number of set of endpoints is indeed  $m + n$ .

We are going to show  $U(P, f) - L(P, f) < \varepsilon$ . Let

$$\Lambda_1 = \{k \mid (x_{k-1}, x_k) = I_j \text{ for some } j\} \quad \text{and} \quad \Lambda_2 = \{k \mid [x_{k-1}, x_k] = J_j \text{ for some } j\}$$

Then the  $U(P, f) - L(P, f)$  can be decomposed into the following two sums:

$$U(P, f) - L(P, f) = \sum_{k \in \Lambda_1} \Omega_f([x_{k-1}, x_k])\Delta x_k + \sum_{k \in \Lambda_2} \Omega_f([x_{k-1}, x_k])\Delta x_k \quad (7.72)$$

We consider each of these two sums separately.

For  $k \in \Lambda_1$ , we have

$$\sum_{k \in \Lambda_1} \Omega_f([x_{k-1}, x_k])\Delta x_k \leq \Omega_f([a, b]) \sum_{k \in \Lambda_1} \Delta x_k < \Omega_f([a, b]) \cdot \frac{1}{r} \quad (7.73)$$

where the last inequality holds because  $\sum_{k \in \Lambda_1} \Delta x_k$  is the exactly the sum of lengths of all  $I_k$ 's, which

is less than  $1/r$ .

On the other hand, for  $k \in \Lambda_2$ , we have

$$\sum_{k \in \Lambda_2} \Omega_f([x_{k-1}, x_k]) \Delta x_k < \frac{1}{r} \sum_{k \in \Lambda_2} \Delta x_k \leq \frac{1}{r} \cdot (b - a) \quad (7.74)$$

where the second last inequality holds because  $\Omega_f(J_k) < 1/r$  for each  $k = 1, \dots, m$ .

Finally, by combining (7.71), (7.72), (7.73) and (7.74) together, we find that indeed  $U(P, f) - L(P, f) < \varepsilon$ . Therefore,  $f$  satisfies Riemann's criterion on  $[a, b]$  and hence is integrable. This completes the proof. ■

☞ **Exercise 7.8** Complete the above proof by showing that  $D = \bigcup_{r=1}^{\infty} D_r$

☞ **Exercise 7.9** Complete the above proof by showing the set  $S$  can be written as a finite union of closed intervals  $J_k$ 's satisfying the properties as claimed.

## 7.17 Improper Riemann Integrals

### Definition 7.17.1

If  $f \in \mathfrak{R}$  on  $[a, b]$  for every  $b \geq a$ , and the limit

$$\lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

exists, then  $f$  is said to be **improper Riemann integrable** on  $[a, \infty)$ , and the **improper Riemann integral**, denoted by  $\int_a^{\infty} f dx$ , is defined by the equation

$$\int_a^{\infty} f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

The improper Riemann integrals

$$\int_{-\infty}^b f(x) dx \quad \text{and} \quad \int_{-\infty}^{\infty} f(x) dx$$

are defined in a similar fashion. In fact, the above improper Riemann integrals are categorized as the **first kind** for the upper or lower limit of the integral is infinity.

The **second kind** improper integrals are those whose upper and lower limits are finite whereas the integrand is not defined at either of the endpoints. Formally, if  $f \in \mathfrak{R}$  on  $[a, c]$  for every  $c \leq b$ , the function value at  $b$  is not defined, and the limit

$$\lim_{c \rightarrow b^-} \int_a^c f(x) dx$$

exists, then the improper Riemann integral, denoted by  $\int_a^b f(x) dx$ , is defined by

$$\int_a^b f(x) dx = \lim_{c \rightarrow b^-} \int_a^c f(x) dx$$

The improper integral whose integrand is not defined at  $a$  or at both points is defined similarly. ♣

# Chapter 8 Sequences and Series of Functions

## 8.1 Uniform Convergence

As mentioned, only assuming  $f_n(x) \rightarrow f(x)$  for each  $x \in S$  is not enough to preserve certain properties of  $f_n$ . Hence, we need a stronger version of convergence than pointwise convergence. The condition  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  says at a given point  $x$ ,  $f_n(x)$  will get closer and closer to  $f(x)$ . Chances are that  $f_n(x_1)$  and  $f_n(x_2)$  may not converge to  $f(x_1)$  and  $f(x_2)$  at the same speed. Now, we want  $f_n(x)$  to get close to  $f(x)$  *uniformly* at the same pace for all  $x \in S$ . This gives rise to the definition of **uniform convergence**.

### Definition 8.1.1

Let  $S$  be a subset of a metric space. A sequence of functions  $\{f_n\}$  is said to converge to  $f$  uniformly on  $S$  if for every  $\varepsilon > 0$ , there exists an integer  $N \in \mathbb{N}^*$  such that

$$|f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N \quad \forall x \in S$$

Symbolically, we write  $f_n \rightarrow f$  uniformly on  $S$ .



The next theorem is useful to check whether  $\{f_n\}$  converges uniformly.

### Theorem 8.1.1

Suppose  $f_n \rightarrow f$  on  $S$ . Then  $f_n \rightarrow f$  uniformly on  $S$  if and only if

$$M_n = \sup_{x \in S} |f_n(x) - f(x)| < \infty \quad \forall n \geq N_0$$

for some  $N_0 \in \mathbb{N}^*$ , and

$$\lim_{n \rightarrow \infty} M_n = 0$$



**Note** To be clear,  $M_n$  is only defined when  $n \geq N_0$  for it has to be a real number.

**Proof** (Sufficiency) If  $\lim_{n \rightarrow \infty} M_n = 0$ , then for any  $\varepsilon > 0$  there exists  $N \in \mathbb{N}^*$  (of course,  $N \geq N_0$ ) such that

$$\sup_{x \in S} |f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N$$

Therefore,  $f_n \rightarrow f$  uniformly on  $S$  by definition since  $|f_n(x) - f(x)| \leq \sup_{x \in S} |f_n(x) - f(x)|$  for all  $x \in S$ .

(Necessity) Suppose  $f_n \rightarrow f$  uniformly on  $S$ . Given  $\varepsilon > 0$ , there exists  $N \in \mathbb{N}^*$  such that

$$|f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N \quad \forall x \in S$$

Taking the supremum over  $x$ , we have

$$M_n = \sup_{x \in S} |f_n(x) - f(x)| \leq \varepsilon \quad \forall n \geq N$$

This implies that the sequence  $\{M_n\}$  converges to 0. ■

**Example 8.1** Let

$$f_n(x) = \frac{1}{nx}, \quad x \in (0, 1)$$

The limit of  $\{f_n(x)\}$  at each  $x$  is zero as  $n \rightarrow \infty$ . Hence,  $f_n \rightarrow f = 0$ . Note that

$$\sup_{x \in (0,1)} |f_n(x) - f(x)| = \sup_{x \in (0,1)} \frac{1}{nx} = \infty$$

for every  $n \in \mathbb{N}^*$ . Therefore, by Theorem 8.1.1, the convergence is not uniform.

**Example 8.2** Let

$$f_n(x) = \frac{1}{x} + \frac{1}{n}, \quad x \in (0, 1)$$

Note that  $f(x) = \lim_{n \rightarrow \infty} f_n(x) = 1/x$ . We have

$$M_n = \sup_{x \in (0,1)} |f_n(x) - f(x)| = \frac{1}{n}$$

Since  $\lim_{n \rightarrow \infty} M_n = 0$ ,  $f_n \rightarrow f$  uniformly on  $(0, 1)$ .

A sequence of functions  $\{f_n\}$  is said to be **uniformly bounded** on  $S$  if there exists a positive constant  $M > 0$  such that

$$|f_n(x)| \leq M \quad \forall n \in \mathbb{N}^* \quad \forall x \in S$$

### Theorem 8.1.2

Suppose  $\{f_n\}$  is uniformly bounded on  $S$ . Then  $f_n \rightarrow f$  uniformly on  $S$  if and only if

$$\lim_{n \rightarrow \infty} \|f_n - f\| = 0$$

where  $\|\cdot\|$  is the uniform norm on the set of bounded functions on  $S$ ,  $\mathcal{B}(S)$ . ♡

Similar to the situation we encountered when dealing with numerical sequences and series, sometimes we want to prove that  $\{f_n\}$  converges uniformly without knowing its limit.

### Theorem 8.1.3 (Cauchy's Criterion)

A sequence of functions  $\{f_n\}$  converges uniformly on  $S$  if and only if for any  $\varepsilon > 0$ , there exists  $N \in \mathbb{N}^*$  such that

$$|f_m(x) - f_n(x)| < \varepsilon \quad \forall m, n \geq N \quad \forall x \in S$$

We say  $\{f_n\}$  satisfies **Cauchy's criterion** if the above condition is met. ♡

**Proof** (Sufficiency) Suppose  $\{f_n\}$  satisfies Cauchy's criterion. Fix  $x \in S$ , we know that the numerical sequence  $\{f_n(x)\}$  is a Cauchy sequence. Hence, it has a limit, and we denote this limit by  $f(x)$ . In this way, we can define a function  $f$  on  $S$  such that

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad \forall x \in S$$

Given  $\varepsilon > 0$ . By Cauchy's criterion, there exists  $N \in \mathbb{N}^*$  such that

$$|f_m(x) - f_n(x)| < \varepsilon/2 \quad \forall m, n \geq N \quad \forall x \in S$$



**Note** The reason why we choose  $\varepsilon/2$  is that when  $m \rightarrow \infty$ , the strict inequality will become a non-strict one.

Letting  $m \rightarrow \infty$ , we have

$$|f(x) - f_n(x)| = \lim_{m \rightarrow \infty} |f_m(x) - f_n(x)| \leq \varepsilon/2 < \varepsilon \quad \forall n \geq N \quad \forall x \in S$$

Therefore,  $f_n \rightarrow f$  uniformly on  $S$ .

(Necessity) Conversely, suppose  $f_n \rightarrow f$  uniformly on  $S$  for some function  $f$ . Given  $\varepsilon > 0$ , by definition, there exists  $N \in \mathbb{N}^*$  such that

$$|f_n(x) - f(x)| < \varepsilon/2 \quad \forall n \geq N \quad \forall x \in S$$

Let  $m, n \geq N$ , we have

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f(x)| + |f_n(x) - f(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon \quad \forall x \in S$$

Hence, Cauchy's criterion is satisfied. ■

## 8.2 Uniform Convergence and Continuity

If we assume  $f_n$  converges to  $f$  uniformly, then we can interchange the two limit processes, that is,

$$\lim_{x \rightarrow p} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow p} f_n(x)$$

### Theorem 8.2.1

Suppose  $f_n \rightarrow f$  uniformly on  $S$ , and let  $p$  be an accumulation point of  $S$ . If the limit

$$A_n = \lim_{x \rightarrow p} f(x)$$

exist for all  $n \in \mathbb{N}^*$ , then the sequence  $\{A_n\}$  converges, say to  $A$ . Furthermore, the limit  $\lim_{x \rightarrow p} f(x)$  also exists, and we have

$$\lim_{x \rightarrow p} f(x) = A$$

In other words, we can write

$$\lim_{x \rightarrow p} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow p} f_n(x) \tag{8.1}$$



**Note** One way to remember this theorem is that if the inner limits on both sides of (8.1) exist, then the outer limits also exist and are equal to each other.

**Proof** We first show that the sequence of limits  $\{A_n\}$  converges. Given  $\varepsilon > 0$ , by Cauchy's criterion (Theorem 8.1.3), there exists an integer  $N \in \mathbb{N}^*$  such that

$$|f_m(x) - f_n(x)| < \varepsilon \quad \forall m, n \geq N \quad \forall x \in S$$

Note that the absolute value function is continuous and both  $f_m(x)$  and  $f_n(x)$  have limits at point  $p$ . Using the rules of computing limits, we obtain

$$\varepsilon \geq \lim_{x \rightarrow p} |f_m(x) - f_n(x)| = \left| \lim_{x \rightarrow p} f_m(x) - \lim_{x \rightarrow p} f_n(x) \right| = |A_m - A_n|$$

Therefore  $\{A_n\}$  indeed converges by Cauchy's criterion for sequences. Let  $A$  denote the limit of  $\{A_n\}$ .

Next, we shall prove  $f(x)$  has a limit as  $x \rightarrow p$ . We choose (another) arbitrary positive number

$\varepsilon > 0$ . Since  $f_n \rightarrow f$  uniformly on  $S$ , there exist  $N_1 \in \mathbb{N}^*$  such that

$$|f(x) - f_n(x)| < \varepsilon \quad \forall n \geq N_1 \quad (8.2)$$

Meanwhile, since we have proved  $\{A_n\}$  converges to  $A$ , there exist  $N_2 \in \mathbb{N}^*$  such that

$$|A_n - A| < \varepsilon/3 \quad \forall n \geq N_2 \quad (8.3)$$

Let  $N = \max\{N_1, N_2\}$ . Because the limit of  $\lim_{x \rightarrow p} f_N(x) = A_N$ , there exist a neighborhood  $V$  of  $p$  in which

$$|f_N(x) - A_n| < \varepsilon/3 \quad \forall x \in V \cap S \quad (8.4)$$

We have

$$|f(x) - A| \leq |f(x) - f_N(x)| + |f_N(x) - A_N| + |A_N - A| < \varepsilon \quad \forall x \in V \cap S$$

where the last inequality follows from (8.2), (8.3) and (8.4). This completes the proof. ■

The version of Theorem 8.2.1 concerning the series of functions is as follows.

### Theorem 8.2.2

Suppose  $\sum f_n \rightarrow f$  uniformly on  $S$ . If  $f_n$  is continuous on  $S$  for every  $n \in \mathbb{N}^*$ , then  $f$  is also continuous on  $S$ . In particular, if  $p$  is an accumulation point of  $S$ , then

$$\lim_{x \rightarrow p} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \lim_{x \rightarrow p} f_n(x)$$



## 8.3 Uniform Convergence and Differentiation

In this section, we study the conditions that allow us to swap the limit and derivative symbols:

$$\frac{d}{dx} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \frac{d}{dx} f_n(x) \quad (8.5)$$

However, assuming  $f_n \rightarrow f$  uniformly on  $S$  is not sufficient to guarantee that (8.5) holds. The following is an example.

**Example 8.3** Let

$$f_n(x) = \frac{\sin nx}{\sqrt{n}}, \quad x \in \mathbb{R}$$

Because  $|\sin nx|$  is bounded by 1 and  $1/\sqrt{n} \rightarrow 0$  as  $n \rightarrow \infty$ , the limit of  $f_n$  is  $f(x) = 0$ . Note that

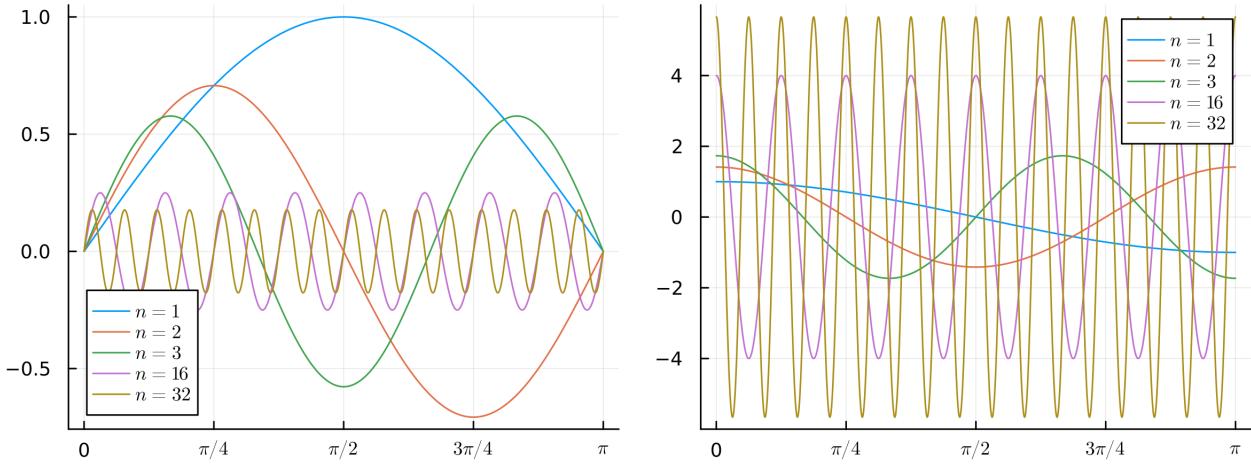
$$M_n = \sup_{x \in \mathbb{R}} |f_n(x) - f(x)| = \sup_{x \in \mathbb{R}} \frac{|\sin nx|}{\sqrt{n}} = \frac{1}{\sqrt{n}}$$

Since  $\lim_{n \rightarrow \infty} M_n = 0$ ,  $f_n \rightarrow f$  uniformly on  $\mathbb{R}$ .

It is clear  $f'(x) = 0$ . We now show that the limiting derivative of  $f_n$  is not  $f'$  in general. We have

$$f'_n(x) = \sqrt{n} \cos nx$$

We note that the limit of  $f'_n(x)$  will not exist as a finite number for most of  $x$ s as  $n \rightarrow \infty$ . In particular,  $f'_n(0) = \sqrt{n}$  and hence  $\lim_{n \rightarrow \infty} f'_n(0) = \infty$ . A plot of  $f_n$  and  $f'_n$  is shown in Figure 8.1.



**Figure 8.1:** Left:  $f_n(x) = (\sin nx)/\sqrt{n}$ . Right:  $f'_n(x)$ .

### Theorem 8.3.1

Let  $\{f_n\}$  be a sequence of real-valued functions each of which has a finite derivative at every point in  $(a, b)$ . Suppose there exists a point  $x_0 \in (a, b)$  such that the numerical sequence  $\{f_n(x_0)\}$  converges. Suppose further that there exists a function  $g$  such that  $f'_n \rightarrow g$  uniformly on  $(a, b)$ . Then

1. there exists a function  $f$  such that  $f_n \rightarrow f$  uniformly on  $(a, b)$ , and
2.  $f'(x)$  exists everywhere in  $(a, b)$  with  $f'(x) = g(x)$ .

In this case, we can write

$$\frac{d}{dx} \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \frac{d}{dx} f_n(x)$$



**Proof** We first show that  $\{f_n\}$  converges uniformly on  $(a, b)$ . Given  $\varepsilon > 0$ , since  $\{f_n(x_0)\}$  converges, by Cauchy's criterion for sequences,

$$|f_m(x_0) - f_n(x_0)| < \varepsilon/2 \quad \forall m, n \geq N_1 \quad (8.6)$$

for some integer  $N_1 \in \mathbb{N}^*$ . Meanwhile, since  $f'_n \rightarrow g$  uniformly on  $(a, b)$ , it follows from Cauchy's criterion for uniform convergence (Theorem 8.1.3) that there exists  $N_2 \in \mathbb{N}^*$  such that

$$|f'_m(x) - f'_n(x)| < \frac{\varepsilon/2}{|x - x_0| + 1} \quad \forall m, n \geq N_2 \quad \forall x \in (a, b) \quad (8.7)$$

Let  $N = \max\{N_1, N_2\}$  and let  $m, n \geq N$ . Define function  $h$  by

$$h(x) = f_m(x) - f_n(x)$$

Applying the mean value theorem (Theorem 5.4.4) to  $h$ , we obtain

$$[f_m(x) - f_n(x)] - [f_m(x_0) - f_n(x_0)] = [f'_m(\xi) - f'_n(\xi)](x - x_0)$$

where  $x \neq x_0$  and  $\xi$  is some number in between  $x$  and  $x_0$ . We can then use (8.6) and (8.7) to bound

the value of  $|f_m(x) - f_n(x)|$  ( $x \neq x_0$ ) as follows:

$$\begin{aligned} |f_m(x) - f_n(x)| &\leq |f_m(x_0) - f_n(x_0)| + |f'_m(\xi) - f'_n(\xi)| |x - x_0| \\ &\leq \varepsilon/2 + \frac{\varepsilon/2}{|x - x_0| + 1} \cdot |x - x_0| \\ &< \varepsilon \quad \forall m, n \geq N \quad \forall x \in (a, b) \setminus \{x_0\} \end{aligned}$$

But since  $m, n \geq N \geq N_1$ , the inequality

$$|f_m(x_0) - f_n(x_0)| < \varepsilon/2 < \varepsilon$$

also holds by (8.6). Therefore, we have

$$|f_m(x) - f_n(x)| < \varepsilon \quad \forall m, n \geq N \quad \forall x \in (a, b)$$

which implies that  $\{f_n\}$  converges uniformly on  $(a, b)$ . Suppose that the limit function is  $f$ .

Let  $c \in (a, b)$ . By Theorem 5.2.1, there exists a continuous function  $\phi$  defined on  $(a, b)$  such that

$$f_n(x) - f_n(c) = (x - c)\phi_n(x)$$

with  $\phi_n(c) = f'_n(c)$ , i.e.,

$$\lim_{x \rightarrow c} \phi_n(x) = f'_n(c)$$

 **Note** We want to show

$$\lim_{x \rightarrow c} \lim_{n \rightarrow \infty} \phi_n(x) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow c} \phi_n(x)$$

using Theorem 8.2.1.

Note that  $\lim_{n \rightarrow \infty} \phi_n(x)$  exists, and

$$\lim_{n \rightarrow \infty} \phi_n(x) = \frac{f(x) - f(c)}{x - c} =: \phi(x)$$

We claim that  $\phi_n \rightarrow \phi$  uniformly on  $(a, b)$ . The proof of this assertion is left as an exercise. (See Exercise 8.1.) Since both the limits  $\lim_{x \rightarrow c} \phi_n(x)$  and  $\lim_{n \rightarrow \infty} \phi_n(x)$  exist, and  $\phi_n(x)$  converges uniformly on  $(a, b)$ , by Theorem 8.2.1, the limits  $\lim_{n \rightarrow \infty} f'_n(c)$  and  $\lim_{x \rightarrow c} [f(x) - f(c)]/(x - c)$  both exist and are equal to each other, that is,

$$\lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c} = \lim_{n \rightarrow \infty} f'_n(c)$$

Therefore,  $f$  has a derivative at  $c$ , and

$$f'(c) = \lim_{n \rightarrow \infty} f'_n(c) = g(c)$$

This completes the proof. ■

 **Exercise 8.1** Complete the above proof by showing  $\phi_n \rightarrow \phi$  uniformly on  $(a, b)$ .

**Solution** Given  $\varepsilon > 0$ , since  $f_n \rightarrow f$  uniformly on  $(a, b)$ , there exist an integer  $N \in \mathbb{N}^*$  such that

$$|f_n(x) - f(x)| < (b - a)\varepsilon/2 \quad \forall n \geq N \quad \forall x \in (a, b)$$

Let  $n \geq N$ . We then have

$$\begin{aligned} |\phi_n(x) - \phi(x)| &\leq \frac{|[f_n(x) - f(x)] - [f_n(c) - f(c)]|}{|x - c|} \\ &\leq \frac{|f_n(x) - f(x)| + |f_n(c) - f(c)|}{b - a} \\ &< \frac{(b - a)\varepsilon/2 + (b - a)\varepsilon/2}{b - a} \\ &= \varepsilon \end{aligned}$$

Because

$$|\phi_n(x) - \phi(x)| < \varepsilon$$

holds for all  $n \geq N$  and for all  $x \in (a, b)$ , we conclude that  $\phi_n \rightarrow \phi$  uniformly on  $(a, b)$ .

The following theorem is another version of Theorem 8.3.1 in terms of a series of functions  $\sum f_n$ .

### Theorem 8.3.2

Let  $\sum f_n$  be a series of real-valued functions each of which has a finite derivative at every point in  $(a, b)$ . Suppose there exists a point  $x_0 \in (a, b)$  such that the numerical series  $\sum f_n(x_0)$  converges. Suppose further that there exists a function  $g$  such that  $\sum f'_n \rightarrow g$  uniformly on  $(a, b)$ . Then

1. there exists a function  $f$  such that  $\sum f_n \rightarrow f$  uniformly on  $(a, b)$ , and
2.  $f'(x)$  exists everywhere in  $(a, b)$  with  $f'(x) = g(x)$ .

In this case, we can write

$$\frac{d}{dx} \sum f_n(x) = \sum \frac{d}{dx} f_n(x) \quad (8.8)$$



**Note** Equation (8.8) is also known as **term-by-term differentiation**.

## 8.4 Power Series

A **power series** (in variable  $z \in \mathbb{C}$ ) has the form

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n$$

where  $a_n \in \mathbb{C}$  is called the coefficient of the  $n$ -th term, and  $z_0 \in \mathbb{C}$  is often referred to as the center of the series.

With every power series, there is associated a **disk of convergence** such that the series converges if  $|z - z_0| < r$  whereas diverges if  $|z - z_0| > r$ . The number  $r$  is called **radius of convergence**. If  $r$  is a *finite positive* number, then the disk of convergence is exactly the open ball in the complex plane with radius  $r$ ,  $B_r(z_0)$ . However, we allow  $r$  to take the value of 0 or  $\infty$  in extreme cases. When the power series is nowhere convergent except for  $z = z_0$ , then we say the radius of convergence is zero, i.e.,  $r = 0$ . And if the power series converges everywhere in the complex plane, then we write  $r = \infty$ .

**Theorem 8.4.1**

Given a power series  $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ , let

$$\lambda = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \quad \text{and} \quad r = \frac{1}{\lambda}$$

(where  $r = 0$  if  $\lambda = \infty$  and  $r = \infty$  if  $\lambda = 0$ .) Then

1. the power series converges absolutely if  $|z - z_0| < r$ , and
2. diverges if  $|z - z_0| > r$ .

Furthermore, the power series, regarded as a series of functions in variable  $z$ , converges uniformly on any compact subset  $K$  interior to the disk of convergence.



**Proof** (Proof of 1 and 2) Applying the root test (Theorem 3.8.2), we find

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n(z - z_0)^n|} = |z - z_0| \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = |z - z_0| \lambda$$

We investigate the value of  $|z - z_0| \lambda$  case by case.

(Case Where  $\lambda = \infty$  and  $r = 0$ ) If  $\lambda = \infty$ , then for any  $z$  such that  $|z - z_0| > r = 0$ , we have  $|z - z_0| \lambda = \infty > 1$ . Hence, the power series diverges.



**Note** Note that, in this case, statement 1 is vacuously true since the hypothesis  $|z - z_0| < r = 0$  is false. In fact, this theorem does not conclude that  $\sum a_n(z - z_0)^n$  converges (absolutely) if  $z = z_0$  when  $\lambda = \infty$ . But we know this is true.

(Case Where  $\lambda = 0$  and  $r = \infty$ ) In this case,  $|z - z_0| \lambda = 0$ . Hence, the power series converges absolutely for any  $z \in \mathbb{C}$ . And we note that statement 2 is vacuously true in this case.

(Case Where  $0 < r < \infty$ ) If  $|z - z_0| < r$ , then

$$|z - z_0| \lambda < r\lambda = \frac{1}{\lambda} \cdot \lambda = 1$$

Hence, the power series converges absolutely. And if  $|z - z_0| > r$ , we then have

$$|z - z_0| \lambda > r\lambda = \frac{1}{\lambda} \cdot \lambda = 1$$

Therefore, it diverges when  $|z - z_0| > r$ .

(Proof of Uniform Convergence) We intend to prove that the power series converges uniformly using the Weierstrass M-test (Theorem 3.3.5). Define

$$g(z) = |z - z_0|, \quad z \in K$$

Note that  $g$  is a continuous real-valued function defined on a compact set. Hence,  $g$  attains its maximum on  $K$  by Theorem 4.3.2, that is, there exist  $p \in K$  such that

$$g(p) = \sup_{z \in K} g(z) = \sup_{z \in K} |z - z_0|$$

It then follows that

$$|a_n(z - z_0)^n| = |a_n| |z - z_0|^n \leq |a_n| [g(p)]^n = |a_n| |p - z_0|^n$$

Let

$$M_n = |a_n| |p - z_0|^n$$

Then the power series converges uniformly on  $K$  if we can show that  $\sum M_n$  converges. Indeed, by

applying the root test, we have

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|M_n|} = |p - z_0| \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = |p - z_0| \lambda < r\lambda = 1$$

This completes the proof. ■

### 8.4.1 Derivatives of Power Series

In fact, derivatives of the power series  $f(z) = \sum a_n(z - z_0)^n$  exist throughout its disk of convergence. To show  $f'(z_1)$  exists, we need to consider the fraction

$$\frac{f(z) - f(z_1)}{z - z_1}$$

The difficulty is that  $f$  consists of terms in powers of  $z - z_0$  while the denominator is  $z - z_1$ . What if we can rewrite  $f$  by expanding it about  $z_1$ ? Indeed, we are allowed to do so, as stated in the following theorem.

#### Theorem 8.4.2

Suppose the series  $\sum_{n=0}^{\infty} a_n(z - z_0)^n$  converges (absolutely) for  $z \in B_R(z_0)$ . Define

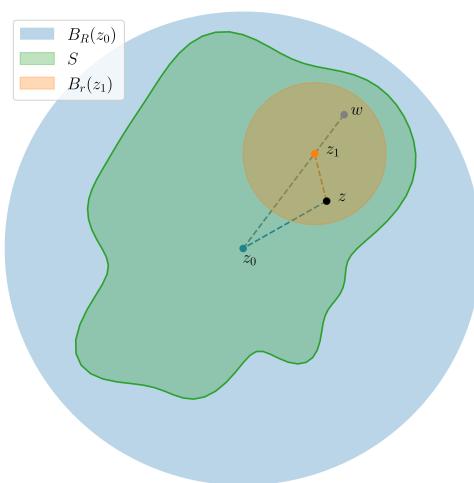
$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n, \quad z \in B_R(z_0)$$

Let  $S$  be an open subset in  $B_R(z_0)$  and  $z_1 \in S$ . Then there exists an open disk  $B_r(z_1) \subset S$  in which  $f$  has a power series expansion about  $z_1$  of the form

$$f(z) = \sum_{k=0}^{\infty} b_k(z - z_1)^k$$

where

$$b_k = \sum_{n=k}^{\infty} \binom{n}{k} a_n(z_1 - z_0)^{n-k}$$



**Figure 8.2:** Domains of power series expansions about different points.

**Proof** Since  $S$  is open, we can find a neighborhood  $B_r(z_1)$  contained in  $S$ . Let  $z \in B_r(z_1)$ . Expanding the expression  $(z - z_0)^n = [(z_1 - z_0) + (z - z_1)]^n$  in  $f(z)$  using the binomial theorem, we

have

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n = \sum_{n=0}^{\infty} a_n \sum_{k=0}^n \binom{n}{k} (z_1 - z_0)^{n-k} (z - z_1)^k \quad (8.9)$$

Define

$$c_{n,k} = \begin{cases} a_n \binom{n}{k} (z_1 - z_0)^{n-k} (z - z_1)^k, & 0 \leq k \leq n \\ 0, & \text{otherwise} \end{cases}$$

In this way, we can rewrite the right-hand side of (8.9) as an iterated series:

$$f(z) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} c_{n,k} \quad (8.10)$$

We desire to interchange the order of summations in (8.10) using Theorem 3.13.1. To do so, we need to show

1.  $\sum_{k=0}^{\infty} |c_{n,k}|$  converges, and
2.  $\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |c_{n,k}|$

For the first one, we note that  $\sum_{k=0}^{\infty} |c_{n,k}| = \sum_{k=0}^n |a_n| \binom{n}{k} |z_1 - z_0|^{n-k} |z - z_1|^k$  is just a finite sum. Hence, it of course converges. As for the second one, we have

$$\begin{aligned} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |c_{n,k}| &= \sum_{n=0}^{\infty} |a_n| \sum_{k=0}^n \binom{n}{k} |z_1 - z_0|^{n-k} |z - z_1|^k \\ &= \sum_{n=0}^{\infty} |a_n| (|z_1 - z_0| + |z - z_1|)^n = \sum_{n=0}^{\infty} |a_n| (z_2 - z_0)^n \end{aligned}$$

where

$$z_2 = |z_1 - z_0| + |z - z_1| + z_0 \quad (8.11)$$

Since the series  $\sum_{n=0}^{\infty} a_n (z - z_0)^n$  converges absolutely for all  $z \in B_R(z_0)$ ,  $\sum_{n=0}^{\infty} |a_n| |z_2 - z_0|^n$  converges since  $z_2 \in B_R(z_0)$  (Exercise 8.2). Note that  $(z_2 - z_0)^n = |z_2 - z_0|^n$  since  $z_2 - z_0$  is in fact a non-negative number. Hence, indeed  $\sum_{n=0}^{\infty} \sum_{k=0}^{\infty} |c_{n,k}|$  converges and Theorem 3.12.1 is applicable. We conclude that  $\sum_{n=0}^{\infty} c_{n,k}$  converges absolutely, and  $\sum_{k=0}^{\infty} \sum_{n=0}^{\infty} c_{n,k}$  also converges absolutely. By interchanging the summations in (8.10), we obtain

$$f(z) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} c_{n,k} = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} c_{n,k} = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} a_n \binom{n}{k} (z_1 - z_0)^{n-k} (z - z_1)^k = \sum_{k=0}^{\infty} b_k (z - z_1)^k$$

This completes the proof. ■

**Exercice 8.2** Complete the above proof by showing  $z_2 \in B_R(z_0)$ . By (8.11), it is equivalent to show  $|z_1 - z_0| + |z - z_1| < R$ .

Though it is intuitive to see  $|z_1 - z_0| + |z - z_1| < R$ , the proof is not trivial.

**Solution** Without loss of generality, we may assume  $z_1 \neq z_0$ . Define

$$w = z_1 + \frac{|z - z_1|}{|z_1 - z_0|} (z_1 - z_0)$$

Refer to Figure 8.2 to see the geometric meaning of  $w$ . We have

$$|w - z_1| = |z - z_1| < r$$

since  $z \in B_r(z_1)$ . This implies that  $w \in B_r(z_1)$ . But

$$w \in B_r(z_1) \subset S \subset B_R(z_0)$$

We have  $|w - z_0| < R$ . Plugging the expression of  $w$ , we find

$$\begin{aligned} |w - z_0| &= \left| (z_1 - z_0) + \frac{|z - z_1|}{|z_1 - z_0|}(z_1 - z_0) \right| \\ &= |z_1 - z_0| \left| 1 + \frac{|z - z_1|}{|z_1 - z_0|} \right| \\ &= |z_1 - z_0| \left( 1 + \frac{|z - z_1|}{|z_1 - z_0|} \right) \\ &= |z_1 - z_0| + |z - z_1| \end{aligned}$$

Therefore, indeed

$$|z_1 - z_0| + |z - z_1| < R$$

### Theorem 8.4.3

Suppose the series  $\sum_{n=0}^{\infty} a_n(z - z_0)^n$  converges (absolutely) for  $z \in B_r(z_0)$ . Define

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n, \quad z \in B_r(z_0)$$

Then  $f$  is differentiable in  $B_r(z_0)$  with the derivative given by

$$f'(z) = \sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1}, \quad z \in B_r(z_0) \tag{8.12}$$



**Note** Note that the power series expansions of  $f$  and  $f'$  have the same radius of convergence since

$$\limsup \sqrt[n]{|a_n|} = \limsup \sqrt[n]{|na_n|}$$

Notice also that (8.12) is referred to as the term-by-term differentiation formula for power series.

**Proof** Pick a point  $z_1 \in B_r(z_0)$ . We can expand  $f$  in some neighborhood  $B_\varepsilon(z_1) \subseteq B_r(z_0)$  about  $z_1$  by Theorem 8.4.2. We have

$$f(z) = \sum_{k=0}^{\infty} b_k (z - z_1)^k$$

where  $b_k = \sum_{n=k}^{\infty} \binom{n}{k} a_n (z_1 - z_0)^{n-k}$ . Note that  $f(z_1) = b_0$ . Subtracting  $f(z_1)$  from  $f(z)$ , we find

$$f(z) - f(z_1) = \sum_{k=1}^{\infty} b_k (z - z_1)^k = b_1 (z - z_1) + \sum_{k=2}^{\infty} b_k (z - z_1)^k$$

Dividing both sides by  $z - z_1$  ( $z \in B_\varepsilon(z_1) \setminus \{z_1\}$ ) yields

$$\frac{f(z) - f(z_1)}{z - z_1} = \sum_{k=1}^{\infty} b_k (z - z_1)^{k-1} = b_1 + \sum_{k=2}^{\infty} b_k (z - z_1)^{k-1} \tag{8.13}$$

Because, by Theorem 8.4.1, we can find a closed disk  $D \subseteq B_\varepsilon(z_1)$  containing  $z_1$  on which the series  $\sum_{k=2}^{\infty} b_k (z - z_1)^{k-1}$  in  $z$  converges uniformly, the limit exists as  $z \rightarrow z_1$  (Theorem 8.2.2). Letting

$z \rightarrow z_1$  on both sides of (8.13), we find

$$f'(z_1) = b_1 = \sum_{n=1}^{\infty} \binom{n}{1} a_n (z_1 - z_0)^{n-1} = \sum_{n=1}^{\infty} n a_n (z_1 - z_0)^{n-1}$$

This completes the proof. ■

If we treat  $f'$  as the original function, then we may differentiate it using Theorem 8.4.3 to obtain the second derivative of  $f$  in the same disk of convergence. Hence, by applying Theorem 8.4.3 repeatedly, we can obtain the derivative of  $f$  of any order in the form of

$$f^{(k)}(z) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n (z - z_0)^{n-k}, \quad z \in B_r(z_0) \quad (8.14)$$

In particular, putting  $z = z_0$  in (8.14), and we will obtain a very important formula

$$f^{(k)}(z_0) = n! a_k \quad \text{or equivalently} \quad a_k = \frac{f^{(k)}(z_0)}{n!}, \quad \forall k \in \mathbb{N}$$

This tells us that if a function  $f$  has a power series expansion, then each coefficient  $a_k$  is determined. In other words, the power series expansion of a function is *unique*.

## 8.4.2 Multiplication of Power Series

### Theorem 8.4.4

Suppose functions  $f$  and  $g$  both have power series expansions about the origin, which are given by

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \quad z \in B_r(0)$$

and

$$g(z) = \sum_{n=0}^{\infty} b_n z^n, \quad z \in B_R(0)$$

Then their product  $f(z)g(z)$  also has a power series expansion with

$$f(z)g(z) = \sum_{n=0}^{\infty} c_n z^n, \quad z \in B_r(0) \cap B_R(0)$$

where

$$c_n = \sum_{k=0}^n a_{n-k} b_k$$



**Proof** Both series  $\sum_{n=0}^{\infty} a_n z^n$  and  $\sum_{n=0}^{\infty} b_n z^n$  converge absolutely for  $z \in B_r(0) \cap B_R(0)$  by Theorem 8.4.1. (In fact, we only need one of them to converge absolutely for Theorem 3.14.1 to hold.) Hence, their Cauchy product converges by Theorem 3.14.1, and it is given by

$$f(z)g(z) = \sum_{n=0}^{\infty} \sum_{k=0}^n a_{n-k} b_k z^{n-k} = \sum_{n=0}^{\infty} \sum_{k=0}^n a_{n-k} b_k z^n$$

This completes the proof. ■

# Chapter 9 Transcendental Functions

## 9.1 Exponential Function

Consider the power series  $\sum a_n(z - z_0)^n$  given by

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} \quad (9.1)$$

It can be regarded as a expansion of a function in  $z$  about the origin. Recall we can find its radius of convergence using methods based on either the root test or the ratio test. By applying the root test, we find

$$\limsup \sqrt[n]{|a_n|} = \limsup \frac{1}{\sqrt[n]{n!}} = 0$$

Therefore, the radius of convergence is  $\infty$ . On the other hand, the ratio test shows that

$$\liminf \left| \frac{a_n}{a_{n+1}} \right| = \liminf \frac{1/n!}{1/(n+1)!} = \liminf (n+1) = \infty$$

This implies that the sequence  $\{|a_n/a_{n+1}|\}$  diverges to infinity, which means the radius of convergence is infinity.

Either way, by referring to Theorem 8.4.1, we see that the power series (9.1) converges absolutely for each  $z \in \mathbb{C}$ . And it converges uniformly on the entire complex plane if we regard it as a function of  $z$ . It is by this power series that we define the well-known exponential function, which is personally I think the most fundamental and the most important function in mathematics.

### Definition 9.1.1

The **exponential function** is defined by

$$\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad z \in \mathbb{C}$$



The following addition formula captures the central behavior of the exponential function.

### Theorem 9.1.1

Given any two complex numbers  $z, w \in \mathbb{C}$ , we have

$$\exp(z + w) = \exp(z) \exp(w)$$



**Proof** Since the power series expansion of the exponential function converges absolutely for every  $z \in \mathbb{C}$ , we may multiply these two power series together by applying Mertens Theorem (Theorem 8.4.4).

We have

$$\begin{aligned}
 \exp(z) \exp(w) &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{(n-k)! k!} z^{n-k} w^k \\
 &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{n!} \binom{n}{k} z^{n-k} w^k \\
 &= \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n \\
 &= \exp(z+w)
 \end{aligned}$$

■

## 9.2 Definition of Pi

$\pi$  is a mathematical constant that represents the ratio of the circumference,  $C$ , of a circle to its diameter,  $d$ .

$$\pi = \frac{C}{d} \quad (9.2)$$

It is approximately equal to 3.14159 and is a non-repeating, non-terminating decimal [5]. The value of pi has been known for thousands of years and has been studied extensively by mathematicians throughout history.

The earliest known written approximations of  $\pi$  date back to ancient Egypt and Babylon, around 1900-1600 BCE [4]. The Greek mathematician Archimedes is credited with determining an accurate approximation of  $\pi$  using inscribed and circumscribed polygons around a circle in the 3rd century BCE. Over time, many mathematicians have contributed to our understanding of  $\pi$ , including John Wallis, Leonhard Euler, and Carl Friedrich Gauss.

$\pi$  has many interesting properties and is used in many areas of mathematics, science, and engineering. It is an irrational number, meaning it cannot be expressed as a ratio of two integers [2].  $\pi$  is also a transcendental number, meaning it is not the solution to any algebraic equation with rational coefficients. These properties of  $\pi$  have been proven mathematically and have important implications in many areas of mathematics.

The symbol for  $\pi$  was first used by Welsh mathematician William Jones in 1706 and was popularized by the Swiss mathematician Leonhard Euler in the 18th century. Today,  $\pi$  is widely used in mathematics, physics, engineering, and many other fields, and its value has been calculated to trillions of digits using computer algorithms [5].

The way of defining  $\pi$  by (9.2) is perhaps the most widely known and accepted method. But this leads to a few problems:

1. What is a circle?
2. How is the circumference of a circle calculated or even defined?
3. Why is the ratio  $C/d$  a constant?

Of course, we can answer these questions rigorously by building up the theory of geometry. But this

requires a lot of work. Hence, we shall define  $\pi$  from another approach with a Riemann integral.

Note that the function  $f(x) = \sqrt{1 - x^2}$  is bounded on  $[0, 1]$  and it is Riemann-integrable for it is continuous. Therefore, its integral on  $[0, 1]$  is well-defined.

### Definition 9.2.1

*The number  $\pi$  is defined as the value of the integral*

$$\pi = 4 \int_0^1 \sqrt{1 - x^2} dx$$



The geometric interpretation of the function value  $y = f(x) = \sqrt{1 - x^2}$  is the  $y$ -coordinate of the unit circle in the first quadrant given its  $x$ -coordinate. Hence, the meaning of the integral  $\int_0^1 \sqrt{1 - x^2} dx$  is the area of a quarter of the unit circle. See Figure 9.1.

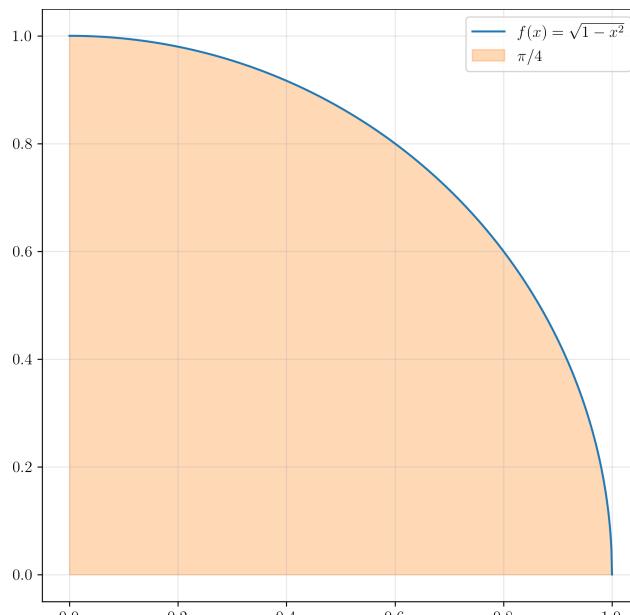


Figure 9.1: Geometric interpretation of  $\pi$ .

Let us prove an alternative definition of  $\pi$  as an Improper Riemann integral, which will be useful in defining arcsine and arccosine functions.

### Theorem 9.2.1

We have

$$\pi = 2 \int_0^1 \frac{1}{\sqrt{1 - x^2}} dx \quad (9.3)$$



**Note** Equation (9.3) serves as the definition of  $\pi$  in Bloch's book [3], which paves the way to defining the arcsine function right after. But this approach lacks geometry interpretation.

**Proof** Consider the closed interval  $[0, a]$  where  $a$  is any positive number in  $(0, 1)$ . By applying integration by parts (Theorem 7.4.1), we have

$$\int_0^a \sqrt{1 - x^2} dx = - \int_0^a x d\sqrt{1 - x^2} \quad (9.4)$$

The derivative of the integrator  $\alpha(x) = \sqrt{1-x^2}$  on the right-hand side of (9.4) is

$$\alpha'(x) = -\frac{x}{\sqrt{1-x^2}}, \quad x \in [0, a]$$

Note that  $\alpha'$  is clearly integrable on  $[0, a]$  for it is continuous. Hence, Theorem 7.14.4 is applicable. We have

$$\int_0^a x \, d\sqrt{1-x^2} = \int_0^a -\frac{x^2}{\sqrt{1-x^2}} \, dx = \int_0^a \sqrt{1-x^2} \, dx - \int_0^a \frac{1}{\sqrt{1-x^2}} \, dx \quad (9.5)$$

By combining (9.4) and (9.5), we find

$$\int_0^a \frac{1}{\sqrt{1-x^2}} \, dx = 2 \int_0^a \sqrt{1-x^2} \, dx \quad (9.6)$$

Since (9.6) holds for every  $a \in (0, 1)$  and the limit of the right-hand side of (9.6) exists as  $a \rightarrow 1^-$ . It follows that

$$\int_0^1 \frac{1}{\sqrt{1-x^2}} \, dx = 2 \int_0^1 \sqrt{1-x^2} \, dx = 2 \cdot \frac{\pi}{4} = \frac{\pi}{2}$$

■

## 9.3 Arcsine and Arccosine

## References

- [1] Tom M. Apostol. *Mathematical Analysis: A Modern Approach to Advanced Calculus*. 2. ed. Addison-Wesley Series in Mathematics. Reading, Mass.: Addison-Wesley, 1974. ISBN: 978-0-201-00288-1.
- [2] Jörg Arndt and Christoph Haenel. *Pi-Unleashed*. Berlin ; New York: Springer, 2001. ISBN: 978-3-540-66572-4.
- [3] Ethan D. Bloch. *The Real Numbers and Real Analysis*. New York: Springer, 2011. ISBN: 978-0-387-72176-7.
- [4] Henry W. Gould and Petr Beckmann. “A History of  $\pi$ (Pi)”. In: *Mathematics of Computation* 28.125 (Jan. 1974), p. 325. ISSN: 00255718. DOI: [10.2307/2005843](https://doi.org/10.2307/2005843). JSTOR: [2005843](https://www.jstor.org/stable/2005843). (Visited on 03/21/2023).
- [5] Eric W. Weisstein. *Pi*. <https://mathworld.wolfram.com/>. Text. (Visited on 03/21/2023).

# Index

<b>Symbols</b>			
$\pi$	123	limit inferior	4
		limit superior	4
		local extremum	44
		local maximum	44
		local minimum	44
		lower Darboux sum	74
		lower Riemann-Stieltjes integral	76
		lower Riemann-Stieltjes sum	74
<b>A</b>			
alternating series	12		
alternating series test	13		
arrangement of double sequences	24		
average of a function over an interval	93		
<b>C</b>			
Cauchy's criterion for series	11	mean value theorem	45
Cauchy's criterion for uniform convergence	110		
chain rule	41		
<b>D</b>			
Darboux integrals	74	one-sided derivatives	36
Darboux's theorem	47	oscillation of a function at a point	103
derivative	34	oscillation of a function on a set	103
differentiable functions	34		
Dirichlet's test	19		
disk of convergence	115		
<b>E</b>			
exponential function	121	power series	115
<b>F</b>			
first fundamental theorem of calculus	96	radius of convergence	115
		Riemann's criterion	80
		right-hand derivative	36
		root test	16
<b>G</b>			
generalized mean value theorem	45		
<b>I</b>			
improper Riemann integral of the first kind	108		
improper Riemann integral of the second kind	108		
<b>L</b>			
left-hand derivative	36	tail of a sequence	3
intermediate value theorem for derivatives	47	Taylor's theorem	53
		term-by-term differentiation	115
<b>U</b>			
		uniform convergence	109

<b>uniformly bounded functions</b>	110	<b>upper Riemann-Stieltjes sum</b>	74
<b>upper Darboux sum</b>	74	<b>Z</b>	
<b>upper Riemann-Stieltjes integral</b>	76	<b>zero-measure sets</b>	101