Mathematical Analysis

Author: Isaac FEI

Contents

Part I

Elementary Concepts

Chapter 1 Topology

1.1 Metric Spaces

Definition 1.1.1

Let X be a metric space. X is bounded if there exists $p \in X$ such that $d(x, p) < M \ \forall x \in X$. ♣

1.2 Compact Sets

Theorem 1.2.1

Closed subsets of a compact set are compact. ♡

Theorem 1.2.2

Let $\{K_\alpha\}_{\alpha \in I}$ be a family of compact subsets in topological space X where I is an arbitrary index set. If for any ***finite*** *subset $J \subset I$, we have $\bigcap_{\alpha \in J} K_\alpha \neq \emptyset$, then $\bigcap_{\alpha \in I} K_\alpha \neq \emptyset$.* ♡

Chapter 2 Numerical Sequences and Series

2.1 Convergent Sequences

Theorem 2.1.1 (Convergence of Monotonic Sequences)

Suppose a sequence of real numbers $\{a_n\}$ is monotonic. Then $\{a_n\}$ converges if and only if it is bounded.

♡

2.2 Euler's Number e

Definition 2.2.1

The Euler's number e is defined by

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

♣

Remark We can easily verify that e is well-defined, i.e., the series on the right-hand side converges by using the ratio test.

Note that e can be regard as the value of the exponential function e^z with $z = 0$. The definition of the exponential function will be introduced in a later section.

The major goal of this section is to prove a very important limit

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e$$

Lemma 2.2.1

The sequence $\left\{\left(1 + \frac{a}{n}\right)^n\right\}$ is increasing where $a > 0$.

♡

Proof Sorry. ■

Theorem 2.2.1

The sequence $\left\{\left(1 + \frac{1}{n}\right)^n\right\}$ is increasing.

♡

Lemma 2.2.2

The sequence $\left\{\left(1 + \frac{1}{n}\right)^n\right\}$ is bounded above by the sequence of partial sums of $\sum \frac{1}{n!}$, i.e.,

$$\left(1 + \frac{1}{n}\right)^n \leq \sum_{k=0}^{n} \frac{1}{k!}$$

where $n \geq 1$. The inequality is strict when $n \geq 2$.

♡

Proof If $n = 1$, then the equality holds. Now, suppose that $n \geq 2$. Apply the binomial expansion to

$\left(1+\frac{1}{n}\right)^n$, and we will obtain

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k}\\ &= 2+\sum_{k=2}^{n}\frac{1}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\\ &= 2+\sum_{k=2}^{n}\frac{1}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\\ &< 2+\sum_{k=2}^{n}\frac{1}{k!}=\sum_{k=0}^{n}\frac{1}{k!}\end{aligned}$$

■

Theorem 2.2.2

The sequence $\left\{\left(1+\frac{1}{n}\right)^n\right\}$ *converges to* e, *i.e.,*

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$$

♡

Proof Let $e_n = \left(1+\frac{1}{n}\right)^n$. We first apply the upper limits on both sides of the inequality proved in Lemma 2.2.2. We have

$$\limsup_{n\to\infty} e_n \le \limsup_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!}=\sum_{k=0}^{\infty}\frac{1}{k!}=e \tag{2.1}$$

On the other hand, we apply the binomial expansion to e_n ($n\ge 2$) in the same manner as in the proof of Lemma 2.2.2, we obtain

$$e_n = 2+\sum_{k=2}^{n}\frac{1}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\ge 2+\sum_{k=2}^{m}\frac{1}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right) \tag{2.2}$$

where m ($m\ge 2$) is some integer less than or equal to n. Let m be fixed for now, and then take the lower limit concerning n on both sides of (2.2). It follows that

$$\liminf_{n\to\infty} e_n \ge 2+\sum_{k=2}^{m}\frac{1}{k!}=\sum_{k=0}^{m}\frac{1}{k!} \tag{2.3}$$

Then by letting $m\to\infty$ on both sides of (2.3), we have

$$\liminf_{n\to\infty} e_n \ge \sum_{k=0}^{\infty}\frac{1}{k!}=e \tag{2.4}$$

Inequalities (2.1) and (2.4) yield

$$e\le\liminf_{n\to\infty} e_n\le\limsup_{n\to\infty} e_n\le e$$

Therefore, $\{e_n\}$ indeed converges to e. ■

The next theorem shows how rapidly that the series $\sum_{n=0}^{\infty}\frac{1}{n!}$ converges to e. As we can imagine, the tailing terms of this series decrease dramatically. Therefore, we can obtain a fair approximation of e by summing up only the first few terms of the series.

Theorem 2.2.3

The difference between the number e and the sum of the first n terms of $\sum \frac{1}{k!}$ is bounded above by $\frac{1}{n!n}$, i.e.,

$$0 < e - \sum_{k=0}^{n} \frac{1}{k!} < \frac{1}{n!n}$$

Remark When we approximate e with the 11th ($n = 10$) partial sum of this series, the error is less than 2.76×10^{-8}, which makes a rather accurate approximation.

Proof The error $e - \sum_{k=0}^{n} \frac{1}{k!}$ can be estimated as follows.

$$\begin{aligned} e - \sum_{k=0}^{n} \frac{1}{k!} &= \sum_{k=n+1}^{\infty} \frac{1}{k!} \\ &= \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+2)(n+1)} + \cdots\right) \\ &\le \frac{1}{n!}\left(\frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots\right) \\ &= \frac{1}{n!}\sum_{k=1}^{\infty} \frac{1}{(n+1)^k} \\ &= \frac{1}{n!n} \end{aligned}$$

■

It is widely known that e is an irrational number. Theorem 2.2.3 provides a very neat proof of this result.

Theorem 2.2.4

e is an irrational number.

Proof We shall prove by contradiction. Assume e is rational, then e can be written as $e = \frac{p}{q}$ where $p, q \in \mathbb{N}^*$. Since $2 < e < 3$, e is clearly not an integer. Hence, $q \ge 2$. It follows from Theorem 2.2.3 that

$$0 < \frac{p}{q} - \sum_{k=0}^{n} \frac{1}{k!} = e - \sum_{k=0}^{n} \frac{1}{k!} < \frac{1}{n!n}$$

Put $n = q$. We have

$$0 < \frac{p}{q} - \sum_{k=0}^{q} \frac{1}{k!} < \frac{1}{q!q} \tag{2.5}$$

Multiply both sides of (2.5) by $q!$,

$$0 < p(q-1)! - \sum_{k=0}^{q} \frac{q!}{k!} < \frac{1}{q} \le \frac{1}{2} \tag{2.6}$$

Note that $p(q-1)! - \sum_{k=0}^{q} \frac{q!}{k!}$ is an *integer* since $k!$ divides $q!$ for all $0 \le k \le q$. But (2.6) implies that $p(q-1)! - \sum_{k=0}^{q} \frac{q!}{k!}$ should be some number between 0 and $\frac{1}{2}$, which leads to a contradiction. ■

2.3 The Root and Ratio Tests

Theorem 2.3.1 (Root Test)

Given series $\sum a_n$ ($a_n \in \mathbb{C}$), *put* $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.

1. *If* $\alpha < 1$ *then* $\sum a_n$ *converges*
2. *If* $\alpha > 1$ *then* $\sum a_n$ *diverges*
3. *If* $\alpha = 1$ *then the test is inconclusive*

Theorem 2.3.2 (Ratio Test)

The series $\sum a_n$ ($a_n \in \mathbb{C}$)

1. *converges if* $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$
2. *diverges if* $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ *for all* $n \geq N$ *where* $N \in \mathbb{N}^*$ *is fixed*

2.4 Power Series

Lemma 2.4.1

Suppose that the power series $\sum c_n z^n$ *converges at some point* $z = z_0$ ($z_0 \neq 0$). *Then* $\sum c_n z^n$ *converges absolutely for all* z *satisfying* $|z| < |z_0|$.

Proof Because $\sum c_n z_0^n$ converges, the sequence $\{c_n z_0^n\}$ also converges (to 0). Therefore, $\{c_n z_0^n\}$ is bounded, that is,

$$|c_n z_0^n| \leq M \quad \forall n \in \mathbb{N}^*$$

for some $M > 0$. Then each term of the series $\sum c_n z^n$ is bounded by

$$|c_n z^n| = |c_n z_0^n| \left(\frac{|z|}{|z_0|}\right)^n \leq M \left(\frac{|z|}{|z_0|}\right)^n$$

Note that $\sum M \left(\frac{|z|}{|z_0|}\right)^n$ converges if $\frac{|z|}{|z_0|} < 1$. Therefore, the series $\sum |c_n z^n|$ converges by the Comparison Test, i.e., $\sum c_n z^n$ converges absolutely. ■

Every power series is associated with a radius of convergence R. We allow R to take the values of 0 and ∞. By writing $R = 0$ we mean that the series only converges at $z = 0$, and by $R = \infty$ we mean that the series converges on the entire complex plane $\mathbb{C}$.

Theorem 2.4.1

Given power series $\sum c_n z^n$, *put*

$$\alpha = \limsup_{n\to\infty} \sqrt[n]{|c_n|} \qquad\qquad R = \frac{1}{\alpha}$$

($R = \infty$ *if* $\alpha = 0$ *and* $R = 0$ *if* $\alpha = \infty$). *Then the power series converges* ***absolutely*** *for* $|z| < R$, *and diverges for* $|z| > R$.

Remark As soon as we know the power series $\sum c_n z^n$ converges at some non-zero point, we are immediately informed that it has a positive radius of convergence, and it converges *absolutely* at points *interior* to the circle of convergence.

Proof We first prove that $\sum c_n z^n$ converges if $|z| < R$. We intend to apply the Root Test. Taking the n-th root of each term of the series and then taking the upper limit, we obtain

$$\limsup_{n\to\infty} \sqrt[n]{|c_n z^n|} = |z| \limsup_{n\to\infty} \sqrt[n]{|c_n|} = |z|\,\alpha = \frac{|z|}{R}$$

Then the convergence of this series follows from the Root Test.

We now prove the absolute convergence. If $R = \infty$, then $\sum c_n r^n$ converges for any $r > 0$ ($r \in \mathbb{R}$). Then Lemma 2.4.1 implies $\sum c_n z^n$ converges absolutely for $|z| < r$. Since $r > 0$ is arbitrary chosen, $\sum c_n z^n$ converges absolutely for all $z \in \mathbb{C}$ (equivalently, $|z| < R = \infty$). For the case $0 < R < \infty$, we have $\sum c_n (R - \varepsilon)^n$ converges for any given $0 < \varepsilon < R$. Then Lemma 2.4.1 implies $\sum c_n z^n$ converges absolutely for $|z| < R - \varepsilon$. It then follows that $\sum c_n z^n$ converges absolutely for any $|z| < R$. ■

2.5 Addition and Multiplication of Series

2.5.1 Double Series

In general, we may interchange the order of a finite summation and an infinite one provided that all involved limits exist.

Proposition 2.5.1

Let $\{a_{ij}\}$ be a double sequence of complex numbers. If $\sum_{j=1}^{\infty} a_{ij}$ converges, then the series $\sum_{j=1}^{\infty} \sum_{i=1}^{m} a_{ij}$ converges, and

$$\sum_{i=1}^{m} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{m} a_{ij}$$

♠

Theorem 2.5.1 (Interchanging the Order of Summations)

Let $\{a_{ij}\}$ ($a_{ij} \in \mathbb{C}$) be a double series where $i, j \in \mathbb{N}^*$. If

1. $\sum_{j=1}^{\infty} |a_{ij}| = b_i < \infty$
2. $\sum b_i$ *converges*

(The two conditions above are equivalent to that $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{ij}| < \infty$.) Then we have

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}$$

♡

Remark We will provide another more interesting proof later using the continuity of the sequence of functions.

Proof We first verify the following:

$$\sum_{j=1}^{\infty} a_{ij} \text{ converges absolutely } \forall i \in \mathbb{N}^*$$

$$\sum_{i=1}^{\infty} a_{ij} \text{ converges absolutely } \forall j \in \mathbb{N}^*$$

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \text{ converges (as a series consisting of terms } \left\{ \sum_{j=1}^{\infty} a_{ij} \right\}_{i \in \mathbb{N}^*} \text{)}$$

By the given condition $\sum_{j=1}^{\infty} |a_{ij}| = b_i < \infty$, $\sum_{j=1}^{\infty} a_{ij}$ converges absolutely for all $i \in \mathbb{N}^*$. For each $j \in \mathbb{N}^*$, we have

$$a_{ij} \le |a_{ij}| \le \sum_{j=1}^{\infty} |a_{ij}| = b_i$$

Because $\sum b_i$ converges, $\sum_{i=1}^{\infty} |a_{ij}|$ also converges by the Comparison Test. It follows that $\sum_{i=1}^{\infty} a_{ij}$ converges absolutely for each $j \in \mathbb{N}^*$. Finally, we note that

$$\left| \sum_{j=1}^{\infty} a_{ij} \right| \le \sum_{j=1}^{\infty} |a_{ij}| = b_i$$

And $\sum b_i$ converges. Applying the Comparison Test, we conclude $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij}$ converges as a series consisting of terms $\left\{ \sum_{j=1}^{\infty} a_{ij} \right\}_{i \in \mathbb{N}^*}$.

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

$$\sum_{j=N_i}^{\infty} |a_{ij}| < \frac{\varepsilon}{2^{i+2}} \quad \forall i \in \mathbb{N}^* \tag{2.7}$$

since $\sum_{j=1}^{\infty} a_{ij}$ converges absolutely. There also exists $M \in \mathbb{N}^*$ such that

$$\sum_{i=M}^{\infty} b_i < \frac{\varepsilon}{4} \tag{2.8}$$

since $\sum b_i$ converges. Let $N \in \mathbb{N}^*$ be given by

$$N = \max\{M, N_1, \ldots, N_M\} \tag{2.9}$$

We now estimate the difference between $\sum_{i=1}^{m} \sum_{j=1}^{\infty} a_{ij}$ and $\sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}$ where $m > N$. We have

$$\begin{aligned}\left|\sum_{i=1}^{m}\sum_{j=1}^{\infty}a_{ij}-\sum_{i=1}^{N}\sum_{j=1}^{N}a_{ij}\right| &= \left|\sum_{i=1}^{N}\sum_{j=N+1}^{\infty}a_{ij}+\sum_{i=N+1}^{m}\sum_{j=1}^{\infty}a_{ij}\right| \\ &\le \sum_{i=1}^{N}\sum_{j=N+1}^{\infty}|a_{ij}|+\sum_{i=N+1}^{m}\sum_{j=1}^{\infty}|a_{ij}| \\ &< \sum_{i=1}^{N}\frac{\varepsilon}{2^{i+2}}+\sum_{i=N+1}^{m}b_i \quad \text{by (2.7) and (2.9)} \\ &\le \sum_{i=1}^{\infty}\frac{\varepsilon}{2^{i+2}}+\sum_{i=N+1}^{m}b_i \\ &= \frac{\varepsilon}{4}+\sum_{i=N+1}^{m}b_i \\ &< \frac{\varepsilon}{4}+\frac{\varepsilon}{4} \quad \text{by (2.8) and (2.9)} \\ &= \frac{\varepsilon}{2}\end{aligned}$$

In summary, we have

$$\left|\sum_{i=1}^{m}\sum_{j=1}^{\infty}a_{ij}-\sum_{i=1}^{N}\sum_{j=1}^{N}a_{ij}\right| < \frac{\varepsilon}{2} \tag{2.10}$$

We then estimate the difference between $\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}$ and $\sum_{j=1}^{N}\sum_{i=1}^{N}a_{ij}$ where $n > N$ in the similar manner. Likewise, we have

$$\begin{aligned}
\left|\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}-\sum_{j=1}^{N}\sum_{i=1}^{N}a_{ij}\right| &= \left|\sum_{j=1}^{N}\sum_{i=N+1}^{\infty}a_{ij}+\sum_{j=N+1}^{m}\sum_{i=1}^{\infty}a_{ij}\right| \\
&\leq \sum_{j=1}^{N}\sum_{i=N+1}^{\infty}|a_{ij}|+\sum_{j=N+1}^{m}\sum_{i=1}^{\infty}|a_{ij}| \\
&= \sum_{i=N+1}^{\infty}\sum_{j=1}^{N}|a_{ij}|+\sum_{i=1}^{\infty}\sum_{j=N+1}^{m}|a_{ij}| \quad \text{by Proposition 2.5.1} \\
&\leq \sum_{i=N+1}^{\infty}\sum_{j=1}^{\infty}|a_{ij}|+\sum_{i=1}^{\infty}\sum_{j=N+1}^{m}|a_{ij}| \\
&= \sum_{i=N+1}^{\infty}b_i+\sum_{i=1}^{\infty}\sum_{j=N+1}^{m}|a_{ij}| \\
&< \frac{\varepsilon}{4}+\sum_{i=1}^{\infty}\sum_{j=N+1}^{m}|a_{ij}| \quad \text{by (2.8) and (2.9)} \\
&< \frac{\varepsilon}{4}+\sum_{i=1}^{\infty}\frac{\varepsilon}{2^{i+2}} \quad \text{by (2.7) and (2.9)} \\
&= \frac{\varepsilon}{4}+\frac{\varepsilon}{4} \\
&= \frac{\varepsilon}{2}
\end{aligned}$$

Therefore,

$$\left|\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}-\sum_{j=1}^{N}\sum_{i=1}^{N}a_{ij}\right| = \left|\sum_{j=1}^{N}\sum_{i=N+1}^{\infty}a_{ij}+\sum_{j=N+1}^{m}\sum_{i=1}^{\infty}a_{ij}\right| < \frac{\varepsilon}{2} \tag{2.11}$$

We are now ready to estimate the difference between $\sum_{i=1}^{m}\sum_{j=1}^{\infty}a_{ij}$ and $\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}$ $(m,n>N)$, which is the central goal of this proof. Indeed, we have

$$\begin{aligned}
\left|\sum_{i=1}^{m}\sum_{j=1}^{\infty}a_{ij}-\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}\right| &\leq \left|\sum_{i=1}^{m}\sum_{j=1}^{\infty}a_{ij}-\sum_{i=1}^{N}\sum_{j=1}^{N}a_{ij}\right|+\left|\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}-\sum_{j=1}^{N}\sum_{i=1}^{N}a_{ij}\right| \\
&< \frac{\varepsilon}{2}+\frac{\varepsilon}{2} \quad \text{by (2.10) and (2.11)} \\
&= \varepsilon
\end{aligned}$$

In summary,

$$\left|\sum_{i=1}^{m}\sum_{j=1}^{\infty}a_{ij}-\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}\right| < \varepsilon \quad \forall m,n>N \tag{2.12}$$

Letting $m\to\infty$ in (2.12),

$$\left|\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{ij}-\sum_{j=1}^{n}\sum_{i=1}^{\infty}a_{ij}\right| < \varepsilon \quad \forall n>N \tag{2.13}$$

Note that we are allowed to do so (letting $m\to\infty$) because $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}a_{ij}$ converges. Then (2.13) implies that the limit of $\sum_{j=1}^{n}\sum_{i=1}^{\infty} a_{ij}$ exists as $n \to \infty$, and

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij}$$

■

The following example shows that we may not change the order of infinite summations at will in general.

Example 2.1 Let $\{a_{ij}\}_{i,j\in\mathbb{N}^*}$ be given by

$$a_{ij} = \begin{cases} 0 & i < j \\ -1 & i = j \\ 2^{j-i} & i > j \end{cases}$$

We list a few terms to get a better intuition:

$$\begin{array}{ccccccc} -1 & 0 & 0 & 0 & 0 & \cdots \\ \frac{1}{2} & -1 & 0 & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{2} & -1 & 0 & 0 & \cdots \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{2} & -1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{array}$$

Summing up each row, we have

$$\sum_{j=1}^{\infty} a_{ij} = \begin{cases} -1 & i = 1 \\ -1 + \sum_{j=1}^{i-1} 2^{j-i} & i > 1 \end{cases} = -2^{1-i}$$

It then follows that

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} -2^{1-i} = -2 \tag{2.14}$$

On the other hand, if we first sum up each column, then we have

$$\sum_{i=1}^{\infty} a_{ij} = -1 + \sum_{i=j+1}^{\infty} 2^{j-i} = 0$$

Therefore,

$$\sum_{j=1}^{\infty}\sum_{i=1}^{\infty} a_{ij} = \sum_{i=1}^{\infty} 0 = 0 \tag{2.15}$$

Clearly, the right-hand sides of (2.14) and (2.15) are not equal to each other.

We can further check which condition of Theorem 2.5.1 that $\{a_{ij}\}$ fails to satisfy. We have

$$\sum_{j=1}^{\infty} |a_{ij}| = \begin{cases} 1 & i = 1 \\ 1 + \sum_{j=1}^{i-1} 2^{j-i} & i > 1 \end{cases} = 2 - 2^{1-i} =: b_i$$

Therefore, the first condition of Theorem 2.5.1 is satisfied. But

$$\sum_{i=1}^{\infty} b_i = \sum_{i=1}^{\infty}(2 - 2^{1-i}) = \infty$$

which means $\sum b_i$ does not converge.

We can always interchange the order of summations for non-negative terms.

Corollary 2.5.1

Let $\{a_{ij}\}$ *be a sequence of non-negative terms, i.e.,* $a_{i,j} \geq 0$. *Then*

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}$$

The case $\infty = \infty$ *may occur.*

Remark Note that we do not rule out ∞ in this corollary. And actually, we only need to prove the case where ∞ occurs.

Proof We shall consider three cases.

(Case 1) Suppose that $\sum_j a_{ij} = \infty$ for some i. Then, on the one hand, it is clear that

$$\sum_i \sum_j a_{ij} = \infty$$

On the other hand, since $\sum_i a_{ij} \geq a_{ij}$, we have

$$\sum_j \sum_i a_{ij} \geq \sum_j a_{ij} = \infty$$

Therefore, $\sum_j \sum_i a_{ij} = \infty$, and hence

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij} = \infty$$

(Case 2) Suppose that $\sum_j a_{ij} = b_i < \infty$ for all i, and $\sum b_i = \infty$ (i.e., $\sum_i \sum_j a_{ij} = \infty$). We need to show that $\sum_j \sum_i a_{ij} = \infty$. Choose some $m \in \mathbb{N}^*$, we have

$$\begin{aligned} \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} &\geq \sum_{j=1}^{\infty} \sum_{i=1}^{m} a_{ij} \\ &= \sum_{i=1}^{m} \sum_{j=1}^{\infty} a_{ij} \quad \text{by Proposition 2.5.1} \\ &= \sum_{i=1}^{m} b_i \end{aligned}$$

Therefore,

$$\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \geq \sum_{i=1}^{m} b_i \quad \forall m \in \mathbb{N}^* \tag{2.16}$$

Letting $m \to \infty$ on both sides of (2.16), we obtain

$$\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \geq \sum_{i=1}^{\infty} b_i = \infty$$

which further implies $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} = \infty$. Hence,

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij} = \infty$$

(Case 3) The only case left is that $\sum_j a_{ij} = b_i < \infty$ for all i, and $\sum b_i$ converges. Then the conclusion follows directly from Theorem 2.5.1. ■

2.5.2 Cauchy Product

Theorem 2.5.2 (Mertens's Theorem on Cauchy Product)

Suppose that $\sum a_n$ and $\sum b_n$ are both convergent series, and ***at least one*** *of them converges* ***absolutely****. Then their Cauchy product $\sum c_n$ converges and*

$$\sum c_n = \sum a_n \sum b_n$$

♡

Remark Note that the requirement of at least one series being absolutely convergent is only a sufficient condition. We will see in Example 2.2 that it is possible the Cauchy product of two conditionally convergent series also converges.

Proof ■

Example 2.2 Let $\sum a_n$ and $\sum b_n$ both be the series $\sum \frac{(-1)^n}{n}$ (the term index starts from 1), which is a classical conditionally convergent series. But their Cauchy product $\sum c_n$ is actually convergent. By the definition, the formula of each term c_n is

$$\begin{aligned} c_n &= \sum_{k=1}^{n} \frac{(-1)^k}{k} \cdot \frac{(-1)^{n-k+1}}{n-k+1} \\ &= \sum_{k=1}^{n} \frac{(-1)^{n+1}}{k(n-k+1)} \\ &= \frac{(-1)^{n+1}}{n+1} \sum_{k=1}^{n} \left(\frac{1}{k} + \frac{1}{n-k+1} \right) \\ &= \frac{(-1)^{n+1} \cdot 2}{n+1} \sum_{k=1}^{n} \frac{1}{k} \end{aligned}$$

After simplification, the expression for c_n turns to

$$c_n = \frac{(-1)^{n+1} \cdot 2}{n+1} \sum_{k=1}^{n} \frac{1}{k}$$

Hence, $\sum c_n$ is clearly an alternating series. We shall then prove $\sum c_n$ converges by the Alternating Series Test. We need to show two things:

1. $\{|c_n|\}$ decrease monotonically, i.e., $|c_n| \geq |c_{n+1}|$
2. $\lim_{n\to\infty} |c_n| = 0$

Note that

$$\begin{aligned} \frac{|c_{n+1}|}{|c_n|} &= \frac{n+1}{n+2} \cdot \frac{1 + \cdots + \frac{1}{n} + \frac{1}{n+1}}{1 + \cdots + \frac{1}{n}} \\ &= \frac{n+1}{n+2} \cdot \left(1 + \frac{\frac{1}{n+1}}{1 + \cdots + \frac{1}{n}} \right) \\ &= \frac{n+1}{n+2} + \frac{1}{(n+2)(1 + \cdots + \frac{1}{n})} \\ &\leq \frac{n+1}{n+2} + \frac{1}{n+2} \\ &= 1 \end{aligned}$$

Therefore,

$$|c_n| \geq |c_{n+1}| \quad \forall n \in \mathbb{N}^*$$

Moreover, the inequality becomes strict when $n \geq 2$, i.e., $|c_n| > |c_{n+1}|$ $\forall n \geq 2$.

Furthermore, we also need to show that $\{|c_n|\}$ converges to 0. Since $\{c_n\}$ decreases monotonically, and it is bounded below by 0, we know that there exists a limit. Now, we show that this limit is precisely 0. Consider the subsequence $\{|c_{2^{n-1}}|\}$:

$$\begin{aligned}
|c_{2^{n-1}}| &= \frac{2}{2^{n-1}+1}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\cdots+\frac{1}{2^{n-1}}\right) \\
&< \frac{2}{2^{n-1}+1}\left(1+\frac{1}{2}+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\cdots+\frac{1}{2^{n-1}}+\cdots+\frac{1}{2^{n-1}}\right) \\
&= \frac{2}{2^{n-1}+1}\left(1+2\cdot\frac{1}{2}+4\cdot\frac{1}{4}+\cdots+2^{n-1}\cdot\frac{1}{2^{n-1}}\right) \\
&= \frac{2n}{2^{n-1}+1}
\end{aligned}$$

Since $\frac{2n}{2^{n-1}+1} \to 0$ as $n \to \infty$, it is clear that $\lim_{n\to\infty} |c_{2^{n-1}}| = 0$. Therefore, the limit of $\{|c_n|\}$ is also 0 because it converges and the limit of its subsequence is 0.

In summary, we have shown $\{|c_n|\}$ decreases monotonically and converges to 0. Then by applying the Alternating Series Test, we conclude that $\sum c_n$ indeed converges.

Chapter 3 Functional Limits

3.1 Limits at Infinity

Proposition 3.1.1

Let $f : E \subset \mathbb{R} \to \mathbb{R}$ be a function where E is unbounded above. Then the limit

$$\lim_{x\to\infty} f(x)$$

exists if and only if for any positive number ε, there exists some number $A > 0$ such that $A_1, A_2 > A$ ($A_1, A_2 \in E$) implies

$$|f(A_1) - f(A_2)| < \varepsilon$$

Proposition 3.1.2

*Let $f : E \subset \mathbb{R} \to \mathbb{R}$ be a **monotone** function. Suppose that $\{a_n\}$ is a sequence in E with infinite limit, i.e., $\lim_{n\to\infty} a_n = \infty$. Then*

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

Note that $\lim_{n\to\infty} f(a_n)$ may assume the value of ∞ or $-\infty$.

Proposition 3.1.3

Let $f : E \to \mathbb{R}$ be a monotonically increasing (resp. decreasing) function where E is unbounded above (resp. below). If f is bounded above (resp. below) by M, then $f(\infty)$ (resp. $f(-\infty)$) exists.

Proof We only prove the case of monotonically increasing functions. Since E is unbounded above, there exists a **subsequence** $\{k_n\}$ of $\mathbb{N}^*$ such that $k_n \in E$. Note that the numerical sequence $\{f(k_n)\}$ increases monotonically. By Theorem 2.1.1, $\{f(k_n)\}$ converges, say, to l.

We are going to show that the limit of f at infinity is exactly l. Given $\varepsilon > 0$, there exists $N \in \mathbb{N}^*$ such that $|f(k_n) - l| < \varepsilon/2 \ \forall n \geq N$. Then for all $x > k_N$, there exist $n, m \geq N$ ($n < m$) such that $k_n \leq x < k_m$. It then follows that

$$\begin{aligned} |f(x) - l| &\leq |f(x) - f(k_n)| + |f(k_n) - l| \\ &< |f(x) - f(k_n)| + \varepsilon/2 \\ &\leq |f(k_m) - f(k_n)| + \varepsilon/2 \end{aligned}$$

The last inequality holds because f is increasing. Sending $m \to \infty$, we have

$$|f(x) - l| \leq |l - f(k_n)| + \varepsilon/2 < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

In summary, we have shown

$$|f(x) - l| < \varepsilon \quad \forall x > k_N$$

Therefore, $f(\infty) = l$. ■

Chapter 4 Functions of Bounded Variation

4.1 Definition of Functions of Bounded Variation

Definition 4.1.1

Let $f : [a, b] \to \mathbb{R}$ be a function. For a partition $P = \{x_0, x_1, \ldots, x_n\}$ 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 4.1.1

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 4.1).

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 which 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 4.1 Let $f : [0, 1] \to \mathbb{R}$ 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

$$
\begin{aligned}
V(P, f) &> \sum_{k=1}^{n} \left| f\left(\frac{1}{2k}\right) - f\left(\frac{1}{2k+1}\right) \right| \\
&= \sum_{k=1}^{n} \left| \frac{1}{2k}\cos(2k\pi) - \frac{1}{2k+1}\cos((2k+1)\pi) \right| \\
&= \sum_{k=1}^{n} (\frac{1}{2k} + \frac{1}{2k+1}) \\
&= \sum_{k=2}^{2n+1} \frac{1}{k}
\end{aligned}
$$

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 4.1.2

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 $P = \{x_0, \ldots, x_n\}$ 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. ■

4.2 Total Variation

Definition 4.2.1

Let f be of bounded variation on $[a,b]$. The total variation of f on $[a,b]$ is defined by

$$V_a^b(f) := \sup_{P\in\mathcal{P}} V(P,f)$$

where $\mathcal{P}$ is the set of partitions on $[a,b]$. ♣

Remark We will sometimes write $V(f)$ instead of $V_a^b(f)$ to ease the notation if it does not confuse.

Theorem 4.2.1

If f and g are each of bounded variation on $[a,b]$, then $f \pm g$ and fg are also of bounded variation. Moreover,

$$V(f\pm g) \le V(f) + V(g) \tag{4.1}$$

and

$$V(fg) \le \sup_{x\in[a,b]} |g(x)|\, V(f) + \sup_{x\in[a,b]} |f(x)|\, V(g) \tag{4.2}$$

♡

Proof (Proof of $f \pm g$ being of bounded variation) We have

$$\begin{aligned}
V(P, f\pm g) &= \sum_{k=1}^{n} |(f\pm g)(x_k) - (f\pm g)(x_{k-1})| \\
&= \sum_{k=1}^{n} |(f(x_k) - f(x_{k-1})) \pm (g(x_k) - g(x_{k-1}))| \\
&\le \sum_{k=1}^{n} |\Delta f_k| + |\Delta g_k| \\
&= V(P,f) + V(P,g) \\
&\le V(f) + V(g)
\end{aligned}$$

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

$$V(P, f\pm g) \le V(f) + V(g)$$

we obtain (4.1).

(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 4.1.1, 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 (4.2) holds. ■

4.3 Additive Property of Total Variation

Theorem 4.3.1

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

$$V_a^b(f) = V_a^c(f) + V_c^b(f) \tag{4.3}$$

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

$$\infty > V_a^b(f) \geq V_a^b(P, f) = V_a^c(P', f) + V_c^b(P'', f) \tag{4.4}$$

Note that (4.4) 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 (4.4), we obtain

$$V_a^b(f) \geq \sup_{P' \in \mathcal{P}[a,c]} V_a^c(P', f) + \sup_{P'' \in \mathcal{P}[c,b]} V_c^b(P'', f) = V_a^c(f) + V_c^b(f) \tag{4.5}$$

Now, let P be a partition on $[a,b]$. Let

$$P' = (P \cap [a,c]) \cup \{c\} \qquad\qquad P'' = P \cap [c,b] \cup \{c\}$$

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

$$V_a^b(P, f) \leq V_a^c(P', f) + V_c^b(P'', f) \leq V_a^c(f) + V_c^b(f) \tag{4.6}$$

Taking the supremum over all partitions on $[a, b]$ on (4.6), we have

$$V_a^b(f) = \sup_{P \in \mathcal{P}[a,b]} V_a^b(P, f) \leq V_a^c(f) + V_c^b(f) \tag{4.7}$$

(4.3) follows from (4.5) and (4.7). ■

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 (4.3) holds for $c = a$ and $c = b$.

Definition 4.3.1

Suppose that f is of bounded variation on $[a, b]$. Define $V_x^x(f) := 0$ where $a \leq x \leq b$. ♣

4.4 Characterization of Functions of Bounded Variation

In fact, all functions of bounded variation can be written as a difference of two increasing functions.

Lemma 4.4.1

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 4.3.1 and Definition 4.3.1, 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

$$V_x^{x+h}(P, f) = |f(x + h) - f(x)| \leq V_x^{x+h}(f) \tag{4.8}$$

It then follows from (4.8) 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 4.4.1 (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 $f = g - h$.* ♡

Proof (Proof of 1 $\implies$ 2) Let function V be as in Lemma 4.4.1. Then by Lemma 4.4.1, V and $V - f$ are both increasing functions on $[a, b]$. Thus, statement 2 holds since $f = V - (V - f)$.

(Proof of 2 $\implies$ 3) Since $f = g - h$, we have

$$f(x) = (g(x) + x) - (h(x) + x)$$

Then statement 3 follows since $g(x) + x$ and $h(x) + x$ are strictly increasing functions.

(Proof of 3 $\implies$ 1) By Theorem 4.1.2, we know that g and h are both of bounded variation on $[a, b]$. Then, by Theorem 4.2.1, the difference of g and h, the function $f = g - h$, is also of bounded variation. ■

Chapter 5 The Riemann-Stieltjes Integral

5.1 Definition of the Riemann-Stieltjes Integral

Definition 5.1.1

Suppose that f and α 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*

$$S(P, f, \alpha) = \sum_{k=0}^{n} f(t_k)\Delta\alpha_k$$

is called a Riemann-Stieltjes sum of f with respect to α. We say f is Riemann-integrable with respect to α, 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_ε such that

$$|S(P, f, \alpha) - A| < \varepsilon$$

for any refinement P of P_ε 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 the following Proposition 5.1.1) and is denoted by

$$\int_a^b f \, \mathrm{d}\alpha$$

♣

Recall Definition 5.1.1 only requires the existence of A. We now show that such number A is also unique.

Proposition 5.1.1

Let $S(P, f, \alpha)$ be as in Definition 5.1.1. If A and A' both satisfy the property stated in Definition 5.1.1, then $A = A'$.

♠

Proof Given $\varepsilon > 0$, by the property in Definition 5.1.1, there exists a partitions P_1 and P_2 such that

$$|S(P_1, f, \alpha) - A| < \varepsilon/2 \qquad |S(P_2, f, \alpha) - A'| < \varepsilon/2$$

Let $P = P_1 \cup P_2$. We have

$$|S(P, f, \alpha) - A| < \varepsilon/2 \qquad |S(P, f, \alpha) - A'| < \varepsilon/2$$

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 $A = A'$. ■

5.2 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 5.2.1

Suppose that $f, g \in \mathfrak{R}(\alpha)$ on $[a, b]$. Then $c_1 f + c_2 g \in \mathfrak{R}(\alpha)$ on $[a, b]$ where c_1 and c_2 are constants. In that case,

$$\int_a^b c_1 f + c_2 g \, \mathrm{d}\alpha = c_1 \int_a^b f \, \mathrm{d}\alpha + c_2 \int_a^b g \, \mathrm{d}\alpha$$

♡

Proof Let P be a partition on $[a, b]$. The Riemann-Stieltjes sum of $c_1 f + c_2 g$ can be written as

$$\begin{aligned} S(P, c_1 f + c_2 g, \alpha) &= \sum_{k=0} c_1 f(t_k) + c_2 g(t_k) \Delta\alpha_k \\ &= c_1 \sum_{k=0} f(t_k)\Delta\alpha_k + c_2 \sum_{k=0} g(t_k)\Delta\alpha_k \\ &= c_1 S(P, f, \alpha) + c_2 S(P, g, \alpha) \end{aligned}$$

Given $\varepsilon > 0$. Since $f \in \mathfrak{R}(\alpha)$ on $[a, b]$ then there exists a partition P'_ε such that

$$\left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \frac{\varepsilon/2}{1 + |c_1|} \qquad \forall P \supset P'_\varepsilon$$

Similarly, since $g \in \mathfrak{R}(\alpha)$, there exists a partition P''_ε such that

$$\left| S(P, g, \alpha) - \int_a^b g \, \mathrm{d}\alpha \right| < \frac{\varepsilon/2}{1 + |c_2|} \qquad \forall P \supset P''_\varepsilon$$

Let P_ε be the refinement of P'_ε and P''_ε, i.e., $P = P'_\varepsilon \cup P''_\varepsilon$. Then for any $P \supset P_\varepsilon$, we have

$$\begin{aligned} &\left| S(P, c_1 f + c_2 g, \alpha) - c_1 \int_a^b f \, \mathrm{d}\alpha - c_2 \int_a^b g \, \mathrm{d}\alpha \right| \\ &\leq |c_1| \left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| + |c_2| \left| S(P, g, \alpha) - \int_a^b g \, \mathrm{d}\alpha \right| \\ &< |c_1| \frac{\varepsilon/2}{1 + |c_1|} + |c_2| \frac{\varepsilon/2}{1 + |c_2|} \\ &< \varepsilon \end{aligned}$$

This completes the proof. ■

The integral is also linear in the integrator.

Theorem 5.2.2

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,

$$\int_a^b f \, \mathrm{d}(c_1\alpha + c_2\beta) = c_1 \int_a^b f \, \mathrm{d}\alpha + c_2 \int_a^b f \, \mathrm{d}\beta$$

♡

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'_ε such that

$$\left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \frac{\varepsilon/2}{1 + |c_1|} \quad \forall P \supset P'_\varepsilon$$

Similarly, since $f \in \mathfrak{R}(\beta)$, there exists a partition P''_ε such that

$$\left| S(P, f, \beta) - \int_a^b f \, \mathrm{d}\beta \right| < \frac{\varepsilon/2}{1 + |c_2|} \quad \forall P \supset P''_\varepsilon$$

Let $P = P'_\varepsilon \cup P''_\varepsilon$. Then for any $P \supset P_\varepsilon$, we have

$$\begin{aligned}
&\left| S(P, f, c_1\alpha + c_2\beta) - c_1 \int_a^b f \, \mathrm{d}\alpha - c_2 \int_a^b f \, \mathrm{d}\beta \right| \\
&\le |c_1| \left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| + |c_2| \left| S(P, f, \beta) - \int_a^b f \, \mathrm{d}\beta \right| \\
&< |c_1| \frac{\varepsilon/2}{1 + |c_1|} + |c_2| \frac{\varepsilon/2}{1 + |c_2|} \\
&< \varepsilon
\end{aligned}$$

■

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 5.2.1

Suppose $c \in (a, b)$. We have

$$\int_a^b f \, \mathrm{d}\alpha = \int_a^c f \, \mathrm{d}\alpha + \int_c^b f \, \mathrm{d}\alpha \tag{5.1}$$

The existence of two integrals in (5.1) will imply the existence of the third one. ♡

Proof We have

$$S(P, f, \alpha) = S(P', f, \alpha) + S(P'', f, \alpha) \quad \forall P = P' \cup P'' \tag{5.2}$$

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 $\int_a^b f \, \mathrm{d}\alpha$) Assume $\int_a^c f \, \mathrm{d}\alpha$ and $\int_c^b f \, \mathrm{d}\alpha$ exist. Then

$$\left| S(P', f, \alpha) - \int_a^c f \, \mathrm{d}\alpha \right| < \varepsilon/2 \quad \forall P' \supset P'_\varepsilon$$

for some P'_ε on $[a, c]$. And

$$\left| S(P'', f, \alpha) - \int_a^c f \, \mathrm{d}\alpha \right| < \varepsilon/2 \quad \forall P'' \supset P''_\varepsilon$$

for some P''_ε on $[c, b]$. Let

$$P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$$

(Note that $c \in P_\varepsilon$.) Let

$$P \supset P_\varepsilon \qquad P' = P \cap [a, c] \qquad P'' = P \cap [c, b]$$

Observe that

$$P' \supset P'_\varepsilon \qquad\qquad P'' \supset P''_\varepsilon$$

It then follows from (5.2) that

$$\begin{aligned}
&\left|S(P,f,\alpha) - \int_a^c f \,\mathrm{d}\alpha - \int_c^b f \,\mathrm{d}\alpha\right| \\
&\le \left|S(P',f,\alpha) - \int_a^c f \,\mathrm{d}\alpha\right| + \left|S(P'',f,\alpha) - \int_c^b f \,\mathrm{d}\alpha\right| \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon
\end{aligned}$$

Therefore, $f \in \mathfrak{R}(\alpha)$ on $[a,b]$ and (5.1) holds.

(Proof of existence of $\int_c^b f \,\mathrm{d}\alpha$) Assume $\int_a^b f \,\mathrm{d}\alpha$ and $\int_a^c f \,\mathrm{d}\alpha$ exist. Then

$$\left|S(P,f,\alpha) - \int_a^b f \,\mathrm{d}\alpha\right| < \varepsilon/2 \quad \forall P \supset P_\varepsilon$$

for some P_ε on $[a,b]$. And

$$\left|S(P',f,\alpha) - \int_a^c f \,\mathrm{d}\alpha\right| < \varepsilon/2 \quad \forall P' \supset P'_\varepsilon$$

for some P'_ε on $[a,c]$. Let

$$P''_\varepsilon = (P_\varepsilon \cup P'_\varepsilon) \cap [c,b]$$

Let

$$P'' \supset P''_\varepsilon \qquad P' \supset (P_\varepsilon \cup P'_\varepsilon) \cap [a,c] \qquad P = P' \cup P''$$

Observe that

$$P' \supset P'_\varepsilon \qquad P = P' \cup P'' \supset (P_\varepsilon \cap [a,c]) \cup (P_\varepsilon \cap [c,b]) = P_\varepsilon$$

It then follows from (5.2) that

$$\begin{aligned}
&\left|S(P'',f,\alpha) - \int_a^b f \,\mathrm{d}\alpha + \int_a^c f \,\mathrm{d}\alpha\right| \\
&\le \left|S(P,f,\alpha) - \int_a^b f \,\mathrm{d}\alpha\right| + \left|S(P',f,\alpha) - \int_a^c f \,\mathrm{d}\alpha\right| \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon
\end{aligned}$$

Therefore, $f \in \mathfrak{R}(\alpha)$ on $[c,b]$ and (5.1) holds.

(Proof of existence of $\int_a^c f \,\mathrm{d}\alpha$) Assume $\int_a^b f \,\mathrm{d}\alpha$ and $\int_c^b f \,\mathrm{d}\alpha$ exist. Then

$$\left|S(P,f,\alpha) - \int_a^b f \,\mathrm{d}\alpha\right| < \varepsilon/2 \quad \forall P \supset P_\varepsilon$$

for some P_ε on $[a,b]$. And

$$\left|S(P'',f,\alpha) - \int_c^b f \,\mathrm{d}\alpha\right| < \varepsilon/2 \quad \forall P'' \supset P''_\varepsilon$$

for some P''_ε on $[c, b]$. Let

$$P'_\varepsilon = (P_\varepsilon \cup P''_\varepsilon) \cap [a, c]$$

Let

$$P' \supset P'_\varepsilon \qquad P'' \supset (P_\varepsilon \cup P''_\varepsilon) \cap [c, b] \qquad P = P' \cup P''$$

Observe that

$$P'' \supset P''_\varepsilon \qquad P = P' \cup P'' \supset (P_\varepsilon \cap [a, c]) \cup (P_\varepsilon \cap [c, b]) = P_\varepsilon$$

It then follows from (5.2) that

$$\begin{aligned}
&\left| S(P', f, \alpha) - \int_a^b f \, \mathrm{d}\alpha + \int_c^b f \, \mathrm{d}\alpha \right| \\
&\leq \left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| + \left| S(P'', f, \alpha) - \int_c^b f \, \mathrm{d}\alpha \right| \\
&< \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon
\end{aligned}$$

Therefore, $f \in \mathfrak{R}(\alpha)$ on $[a, c]$ and (5.1) holds.

■

5.3 Integration by Parts

Theorem 5.3.1 (Integration by Parts)

If $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, then $\alpha \in \mathfrak{R}(f)$ on $[a, b]$, and

$$\int_a^b f \, \mathrm{d}\alpha + \int_a^b \alpha \, \mathrm{d}f = f(b)\alpha(b) - f(a)\alpha(a)$$

♡

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_ε such that

$$\left| S(P, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon \quad \forall P \supset P_\varepsilon \tag{5.3}$$

Let $P = \{x_0, x_1, \ldots, x_n\} \supset P_\varepsilon$ be any refinement of P_ε. The Riemann-Stieltjes sum of α with respect to f is

$$S(P, \alpha, f) = \sum_{k=1}^n \alpha(t_k)(f(x_k) - f(x_{k-1})) = \sum_{k=1}^n \alpha(t_k) f(x_k) - \sum_{k=1}^n \alpha(t_k) f(x_{k-1}) \tag{5.4}$$

Let

$$P^* = P \cup \{t_k \mid 1 \leq k \leq n\} \tag{5.5}$$

Denote by A the value

$$A = f(b)\alpha(b) - f(a)\alpha(a)$$

Note that A can be written as

$$A = \sum_{k=1}^{n} f(x_k)\alpha(x_k) - \sum_{k=1}^{n} f(x_{k-1})\alpha(x_{k-1}) \tag{5.6}$$

Subtracting (5.4) from (5.4), 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 (5.5), 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 $P^* \supset P \supset P_\varepsilon$, it follows from (5.3) that

$$\left| A - S(P, \alpha, f) - \int_a^b f \, \mathrm{d}\alpha \right| = \left| S(P^*, f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon$$

Recall $A = f(b)\alpha(b) - f(a)\alpha(a)$, we have

$$\left| S(P, \alpha, f) + \int_a^b f \, \mathrm{d}\alpha - f(b)\alpha(b) + f(a)\alpha(a) \right| < \varepsilon \quad \forall P \supset P_\varepsilon$$

This implies that $\alpha \in \mathfrak{R}(f)$ on $[a, b]$, and

$$\int_a^b f \, \mathrm{d}\alpha = -\int_a^b f \, \mathrm{d}\alpha + f(b)\alpha(b) - f(a)\alpha(a)$$

■

5.4 Change of Variables

Theorem 5.4.1

Suppose $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, and g 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*

$$a = g(c) \qquad\qquad b = g(d)$$

Define two composite functions:

$$h = f(g(x)) \qquad\qquad \beta = \alpha(g(x))$$

Then $h \in \mathfrak{R}(\beta)$ on I, and $\int_a^b f \, \mathrm{d}\alpha = \int_c^d h \, \mathrm{d}\beta$, i.e.,

$$\int_a^b f(x) \, \mathrm{d}\alpha(x) = \int_c^d f(g(x)) \, \mathrm{d}\alpha(g(x))$$

♡

Remark The reason why we assume that g is strictly monotonic and continuous is to ensure that it is a bijective function. It is equivalent to assume g is strictly monotonic and injective.

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

$$g^{-1} : [a, b] \to [c, d]$$

For any partition $P' = \{x_0, \ldots, x_n\}$ on $[a, b]$, we can associate it with a partition P on $[c, d]$, which is given by

$$P := g^{-1}(P') := \{g^{-1}(x_0), \ldots, g^{-1}(x_n)\}$$

On the other hand, for any partition $P = \{y_0, \ldots, y_n\}$ on $[c, d]$, we can define a partition P' on $[a, b]$ by

$$P' := g(P) := \{g(y_0), \ldots, g(y_n)\}$$

(Existence of the Integral) Given $\varepsilon > 0$, since $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, there exists a partition P'_ε on $[a, b]$ such that

$$\left| S(P', f, \alpha) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon \quad \forall P' \supset P'_\varepsilon \tag{5.7}$$

Let partition P_ε on $[c, d]$ be given by $P_\varepsilon = g^{-1}(P'_\varepsilon)$. For any refinement $P \supset P_\varepsilon$, 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 $t_i = g(s_i)$. It then follows that

$$\begin{aligned} S(P, h, \beta) &= \sum h(g^{-1}(t_i))(\beta(g^{-1}(t_i)) - \beta(g^{-1}(t_{i-1}))) \\ &= \sum f(t_i)(\alpha(t_i) - \alpha(t_{i-1})) \\ &= S(P', f, \alpha) \end{aligned}$$

where $P' = g(P)$. In summary,

$$S(P, h, \beta) = S(P', g, \alpha) \tag{5.8}$$

What is left to shown is that $P' \supset P'_\varepsilon$. For any point $x \in P'_\varepsilon$, we have $g^{-1}(x) \in P_\varepsilon$ since $P_\varepsilon = g^{-1}(P_\varepsilon)$. Recall that $P \supset P_\varepsilon$. Thus, $g^{-1}(x) \in P$. And because $P' = g(P)$, we have $x = g(g^{-1}(x)) \in P'$. Therefore, indeed $P' \supset P'_\varepsilon$. It then follows from (5.7) and (5.8) that

$$\left| S(P, h, \beta) - \int_a^b f \, \mathrm{d}\alpha \right| < \varepsilon \quad \forall P \supset P_\varepsilon$$

This completes the proof. ■

5.5 Comparison Theorems

Theorem 5.5.1

Suppose that $\alpha \uparrow$ on $[a, b]$. If $f, g \in \mathfrak{R}(\alpha)$ on $[a, b]$, and $f \le g$, then

$$\int_a^b f \, \mathrm{d}\alpha \le \int_a^b g \, \mathrm{d}\alpha$$

♡

5.6 Existence of Riemann-Stieltjes Integrals

5.7 Mean Value Theorem for Riemann-Stieltjes Integrals

Theorem 5.7.1 (First Mean Value Theorem for Riemann-Stieltjes Integrals)

Suppose that $\alpha \uparrow$ on $[a, b]$. If $f \in \mathfrak{R}(\alpha)$ on $[a, b]$, then there exists a number c satisfying

$$\inf_{x\in[a,b]} f(x) \leq c \leq \sup_{x\in[a,b]} f(x)$$

such that

$$\int_a^b f \, \mathrm{d}\alpha = c \int_a^b \mathrm{d}\alpha = \alpha(b) - \alpha(a) \tag{5.9}$$

Furthermore, if f is continuous on $[a, b]$, then there exists $x_0 \in [a, b]$ such that $f(x_0) = c$. ♡

Proof If α is a constant function, then the conclusion is trivial since all values in (5.9) are zeros. In the rest of the proof, we assume α is non-constant. In that case, $\alpha(b) > \alpha(a)$. To simplify the notation, we let $m = \inf_{x\in[a,b]} f(x)$ and $M = \sup_{x\in[a,b]} f(x)$. Since constant functions are clearly integrable, by Theorem 5.5.1, we have

$$\int_a^b m \, \mathrm{d}\alpha \leq \int_a^b f \, \mathrm{d}\alpha \leq \int_a^b M \, \mathrm{d}\alpha$$

It then follows that

$$m \leq \frac{1}{\alpha(b) - \alpha(a)} \int_a^b f \, \mathrm{d}\alpha \leq M$$

Let c be given by

$$c = \frac{1}{\alpha(b) - \alpha(a)} \int_a^b f \, \mathrm{d}\alpha$$

Hence, (5.9) is proved.

If we assume further that f is continuous, then by the Intermediate Value Theorem, $\exists x_0 \in [a, b]$ such that

$$f(x_0) = c$$

■

Theorem 5.7.2 (Second Mean Value Theorem for Riemann-Stieltjes Integrals)

♡

5.8 Integration and Differentiation

Theorem 5.8.1

Let $f \in \mathfrak{R}$ on $[a, b]$. For $x \in [a, b]$, put

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

Then, F is continuous on $[a, b]$ (more precisely, F is uniformly continuous). Furthermore, if f is continuous at some point $x_0 \in [a, b]$, then F is differentiable at x_0, and

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

♡

Proof ■

Example 5.1 Consider the function

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

Let $F(x) = \int_{-1/\pi}^{x} f(t)\,\mathrm{d}t$, then a few steps of computation shows

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x} & -\frac{1}{\pi} \leq x \leq \frac{1}{\pi},\ x \neq 0 \\ 0 & x = 0 \end{cases}$$

Note that even though f is discontinuous at $x = 0$, F remains continuous on the entire interval. But, because of this discontinuity of f, F is not differentiable at $x = 0$.

Theorem 5.8.2 (Fundamental Theorem of Calculus)

If $f \in \mathfrak{R}$ on $[a, b]$ and F is a differentiable function on $[a, b]$ such that $F' = f$, then we have

$$\int_a^b f\,\mathrm{d}x = F(b) - F(a)$$

♡

Proof ■

Theorem 5.8.3 (Integration by Parts)

Let F, G be differentiable functions on $[a, b]$ such that $F' = f \in \mathfrak{R}$ and $G' = g \in \mathfrak{R}$, then

$$\int_a^b F(x)g(x)\,\mathrm{d}x = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\,\mathrm{d}x$$

♡

Proof ■

Chapter 6 Improper Integrals

6.1 Improper Integral of the First Kind

Definition 6.1.1

Suppose $f \in \mathfrak{R}$ on $[a, b]$ for all $b \geq a$ where a is fixed. Define

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{b\to\infty} \int_a^b f(x)\,\mathrm{d}x \tag{6.1}$$

*provided that the limit exists and is **finite**. In that case, we say the integral on the left-hand side of (6.1) **converges**. And this kind of improper integral with infinite integration limits is called the **improper integral of the first kind**. If*

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

*the integral $\int_a^\infty f(x)\,\mathrm{d}x$ is said to converge **absolutely**.* ♣

Theorem 6.1.1

Suppose that $f \in \mathfrak{R}$ on $[a, b]$ for all $b \geq a$, and $f \geq 0$ on $[a, \infty)$. Let $F(b) = \int_a^b f(x)\,\mathrm{d}x$. Then $\int_a^\infty f(x)\,\mathrm{d}x$ converges if and only if $F(b)$ is bounded on $[a, \infty)$. ♡

Proof Assume first $\int_a^\infty f(x)\,\mathrm{d}x$ converges, then $F(\infty) < \infty$. F increases monotonically since $f \geq 0$. It follows that $F(b) \leq F(\infty) < \infty$, hence bounded. Conversely, if $F(b)$ is bounded on $[a, \infty)$, then by Proposition 3.1.3, $F(\infty)$ exists, i.e., $\int_a^\infty f(x)\,\mathrm{d}x$ converges since F is increasing. ■

Theorem 6.1.2 (Comparison Test)

Suppose that $f, g \in \mathfrak{R}$ on $[a, b]$ for all $b \geq a$, and f and g satisfy the following inequality for sufficiently large x:

$$0 \leq f(x) \leq g(x) \quad \forall x > c$$

where c is a constant. Then we have

1. *$\int_a^\infty f(x)\,\mathrm{d}x$ converges if $\int_a^\infty g(x)\,\mathrm{d}x$ converges.*
2. *$\int_a^\infty g(x)\,\mathrm{d}x$ diverges if $\int_a^\infty f(x)\,\mathrm{d}x$ diverges.*

♡

Proof ■

Theorem 6.1.3 (Ratio Test)

Suppose that $f, g \in \mathfrak{R}$ on $[a, b]$ for all $b \geq a$, and $g(x) > 0$ for sufficiently large x. Let l be the limit of ratio of $f(x)$ to $g(x)$, i.e.,

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = l$$

We have the following:

1. *If $0 < l < \infty$, then $\int_a^\infty f(x)\,\mathrm{d}x$ and $\int_a^\infty g(x)\,\mathrm{d}x$ both converge or both diverge. In other words, $\int_a^\infty f(x)\,\mathrm{d}x$ converges if and only if $\int_a^\infty g(x)\,\mathrm{d}x$ converges.*
2. *If $l = 0$, then $\int_a^\infty f(x)\,\mathrm{d}x$ converges if $\int_a^\infty g(x)\,\mathrm{d}x$ converges.*
3. *If $l = \infty$, then $\int_a^\infty f(x)\,\mathrm{d}x$ diverges if $\int_a^\infty g(x)\,\mathrm{d}x$ diverges.*

♡

Remark Note that all the conclusions are about $\int_a^\infty f(x)\,\mathrm{d}x$. Indeed, the test is used to determine the convergence of $\int_a^\infty f(x)\,\mathrm{d}x$ based on the knowledge of convergence of $\int_a^\infty g(x)\,\mathrm{d}x$

Proof (Proof of 1) Choose $0 < \varepsilon < l$. Then there exists $c > a$ such that

$$\left|\frac{f(x)}{g(x)} - l\right| < \varepsilon \quad \forall x > c$$

and $g(x) > 0\ \forall x > c$. It follows that

$$(l-\varepsilon)g(x) < f(x) < (l+\varepsilon)g(x) \quad \forall x > c$$

By the Comparison Test (Theorem 6.1.2),

$$0 < (l-\varepsilon)g(x) < f(x) \quad \forall x > c$$

implies that $\int_a^\infty g(x)\,\mathrm{d}x$ converges if $\int_a^\infty f(x)\,\mathrm{d}x$ converges. Similarly,

$$0 \leq f(x) < (l+\varepsilon)g(x) \quad \forall x > c$$

implies that $\int_a^\infty f(x)\,\mathrm{d}x$ converges if $\int_a^\infty g(x)\,\mathrm{d}x$ converges.

(Proof of 2) Choose a positive number $\varepsilon > 0$. There exists $c > a$ such that

$$\left|\frac{f(x)}{g(x)}\right| < \varepsilon \quad \forall x > c$$

and $g(x) > 0\ \forall x > c$. Rearranging the above inequality, we obtain

$$0 \leq f(x) < \varepsilon g(x)$$

Thus, it follows from the Comparison Test that $\int_a^\infty f(x)\,\mathrm{d}x$ converges if $\int_a^\infty g(x)\,\mathrm{d}x$ converges.

(Proof of 3) Choose a positive number $M > 0$. By the definition of infinite limits, there exists $c > a$ such that

$$\frac{f(x)}{g(x)} > M \quad \forall x > c$$

and $g(x) > 0\ \forall x > c$. It follows that

$$f(x) > Mg(x) > 0 \quad \forall x > c$$

Then $\int_a^\infty f(x)\,\mathrm{d}x$ diverges if $\int_a^\infty g(x)\,\mathrm{d}x$ diverges. ■

Theorem 6.1.4 (Integral Test)

Suppose that $f \geq 0$ and f decreases monotonically on $[1, \infty)$. Then

$$\int_1^\infty f(x)\,\mathrm{d}x$$

converges if and only if

$$\sum_{n=1}^\infty f(n)$$

converges. ♡

Proof ■

6.2 Dirichlet's Test and Abel's Test for Improper Integrals

Chapter 7 Sequences and Series of Functions

In the introduction of pointwise and uniform convergence of functions, we confine ourselves to complex-valued functions whose domains lie in *metric spaces*.

7.1 Discussion of Main Problem

Definition 7.1.1

Let $\{f_n\}$ be a sequence of functions on E. If for any $x \in E$, the limit of the numerical sequence $\{f_n(x)\}$ exists, then the function defined by

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

*is called the **limit** of $\{f_n\}$.* ♣

We say f_n converges *pointwise* to f if f_n only satisfies the above definition. There is a stronger version of convergence called uniform convergence, which will be introduced later.

If we regard $\{\sum_{m=0}^{n} f_m\}_n$ as a sequence of partial sums, we can define the *sum* of series of functions $\sum f_n$ as the *limit* of the sequence of partial sums.

Definition 7.1.2

If $\sum f_n(x)$ converges (i.e., the limit of the partial sum exists) for each $x \in E$, and if we define

$$f(x) = \sum_{n=0}^{\infty} f_n(x)$$

*then f is called the **sum** of series $\sum f_n$.* ♣

7.2 Uniform Convergence

Definition 7.2.1

*We say a sequence of functions $\{f_n\}$ converges **uniformly** on E to a function f if for an arbitrary $\varepsilon > 0$, there exists $N \in \mathbb{N}^*$ such that $n \geq N$ implies*

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

for all $x \in E$. ♣

An analogous definition above exists for the series of functions.

Definition 7.2.2

*We say a series of functions $\sum f_n$ converges **uniformly** on E to a function f if for an arbitrary*

$\varepsilon > 0$, there exists $N \in \mathbb{N}^*$ such that $n \geq N$ implies

$$\left|\sum_{m=0}^{n} f_m(x) - f(x)\right| < \varepsilon$$

for all $x \in E$. ♣

The Cauchy criterion for uniform convergence is as follows.

Theorem 7.2.1 (Cauchy Criterion)

The sequence of functions $\{f_n\}$, defined on E, converges uniformly if and only if for any $\varepsilon > 0$, there exists $N \in \mathbb{N}^*$ such that $m, n \geq N$, $x \in E$ implies

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

♡

Proof ∎

The following theorem is another criterion for uniform convergence, which provides us with insight into measuring the distance between two functions.

Theorem 7.2.2

Suppose the sequence of functions $\{f_n\}$ converges pointwise to f on E. Put

$$M_n = \sup_{x \in E} |f_n(x) - f(x)|$$

Then $f_n \to f$ uniformly if and only if $\lim_{n\to\infty} M_n = 0$. ♡

Remark $\sup_{x\in E} |f_n(x) - f(x)|$ actually defines a distance function $d(f_n, f)$, which we will discuss in more details later.

Proof ∎

For series, there is a very convenient test called Weierstrass's M-Test for uniform convergence.

Theorem 7.2.3 (Weierstrass's M-Test)

Let $\{f_n\}$ be a sequence of functions defined on E, and suppose

$$|f_n(x)| \leq M_n \quad \forall x \in E$$

Then $\sum f_n$ converges uniformly on E if $\sum M_n$ converges. ♡

Proof ∎

7.3 Uniform Convergence and Continuity

The uniform convergence of functions allows us to interchange change limits.

Theorem 7.3.1

Suppose $f_n \to f$ uniformly on $E \subset X$ where X is a metric space. Let x be a limit point of E. If the limit $A_n := \lim_{t\to x} f_n(t)$ exists, then $\{A_n\}$ converges and

$$\lim_{t\to x} f(t) = \lim_{n\to\infty} A_n$$

In other words,

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t) \tag{7.1}$$

♡

Remark One way to remember this theorem is that if the *inner* limits on both sides of (7.1) exist, then the *outer* limits also exist and are equal to each other.

Proof We first prove that $\{A_n\}$ converges. Given $\varepsilon > 0$, since $\{f_n\}$ converges uniformly, there exists $N \in \mathbb{N}^*$ such that $n, m \geq N$ implies that

$$|f_n(t) - f_m(t)| < \varepsilon \quad \forall t \in E \tag{7.2}$$

due to Theorem 7.2.1. Because the limits $\lim_{t\to x} f_n(t)$ and $\lim_{t\to x} f_m(t)$ both exist, the limit of the left-hand side of (7.2) also exists as $t \to x$. Letting $t \to x$, we have

$$|A_n - A_m| = \left|\lim_{t\to x} f_n(t) - \lim_{t\to x} f_m(t)\right| = \lim_{t\to x} |f_n(t) - f_m(t)| \leq \varepsilon$$

Therefore, $\{A_n\}$ indeed converges due to the Cauchy criterion for convergence of sequences.

We now show that the limit of $f(t)$ exists as $t \to x$, and it equals $A := \lim_{n\to\infty} A_n$. Let $\varepsilon > 0$ be arbitrary. Since $f_n \to f$ uniformly on E, there exists $N_1 \in \mathbb{N}^*$ such that

$$|f(t) - f_n(t)| < \frac{\varepsilon}{3} \quad \forall n \geq N_1,\ \forall t \in E \tag{7.3}$$

And since $f_n(t) \to A_n$ as $t \to x$, there exists a neighborhood V of x such that

$$|f_n(t) - A_n| < \frac{\varepsilon}{3} \quad \forall t \in V \cap E \tag{7.4}$$

Moreover, because we have proved $\lim_{n\to\infty} A_n = A$, there exists $N_2 \in Ns$ such that

$$|A_n - A| < \frac{\varepsilon}{3} \quad \forall n \geq N_2 \tag{7.5}$$

Let $N = \max\{N_1, N_2\}$. Suppose $t \in V \cap E$ and $n \geq N$, by (7.3), (7.4) and (7.5), we have

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

Therefore, $\lim_{t\to x} f(t) = A$. ■

An immediate corollary to Theorem 7.3.1 is that if a sequence of continuous functions converges uniformly to some function, then that limit function is also continuous.

Theorem 7.3.2

Let $\{f_n\}$ be a sequence of continuous functions on E. If $f_n \to f$ uniformly on E, then f is also continuous.

♡

Remark However, the converse of this theorem is not true in general. That is, if f is continuous it is not necessary that $f_n \to f$ uniformly. Or in other words, it is possible in some cases that $f_n \to f$ only pointwise, and f is still continuous. The following example is one such case.

Example 7.1 Let

$$f_n(x) = n^2 x(1 - x^2)^n$$

where $0 \leq x \leq 1$ and $n \in \mathbb{N}^*$. Note that

$$f(x) = \lim_{n\to\infty} f_n(x) = 0$$

All f_n's are continuous and f is of course continuous since it is a constant function. But $\{f_n\}$ does not converge to f uniformly on $[0,1]$ due to Theorem 7.2.2. To see this, one can obtain

$$M_n = \sup_{x\in[0,1]} f_n(x) = \frac{n^2}{\sqrt{1+2n}}\left(1 - \frac{1}{1+2n}\right)^n$$

by computing the critical points of $f_n(x)$. Note that

$$\lim_{n\to\infty}\left(1 - \frac{1}{1+2n}\right)^n = \lim_{n\to\infty}\sqrt{\frac{1}{(1+\frac{1}{2n})^{2n}}} = \frac{1}{\sqrt{e}}$$

But $\frac{n^2}{\sqrt{1+2n}} \to \infty$ as $n \to \infty$. Therefore, $\lim_{n\to\infty} M_n = \infty$. Since M_n does not converge to 0, Theorem 7.2.2 implies that $\{f_n\}$ does not converge to f uniformly.

Proof Let a sequence $\{x_n\}_{n\in\mathbb{N}}$ be given by

$$x_0 = 0$$
$$x_n = \frac{1}{n} \quad n \geq 1$$

Let E be the set consisting of terms of this sequence, i.e,

$$E = \{x \in \mathbb{R} \mid \exists n \in \mathbb{N},\ x_n = x\} = \{0\} \cup \left\{\frac{1}{n}\ \middle|\ n \in \mathbb{N}^*\right\}$$

It is clear that $\lim_{n\infty} x_n = x_0$, and $x_n \to x_0$ if and only if $n \to \infty$. We now define a sequence of functions $\{f_i\}_{i\in\mathbb{N}^*}$ on E by specifying the function value at each point:

$$f_i(x_0) := \sum_{j=1}^{\infty} a_{ij}$$
$$f_i(x_n) := \sum_{j=1}^{n} a_{ij} \quad n \geq 1$$

And then let function g be defined by

$$g(x) := \sum_{i=1}^{\infty} f_i(x)$$

We need to verify that functions f_i's and g are all well-defined. $f_i(x_n)$ is well-defined since it is just a finite sum of complex numbers. In the conditions of this theorem, we require that $\sum_{j=1}^{\infty} a_{ij}$ converges absolutely, hence the well-definedness of $f(x_0)$. For function g, we note that

$$f_i(x) \leq |f_i(x)| \leq \sum_{j=1}^{\infty} |a_{ij}| = b_i < \infty \tag{7.6}$$

Since $\sum b_i$ converges, we know $\sum |f_i(x)|$ also converges by the Comparison Test. Therefore, $\sum f_i(x)$ converges (absolutely), and hence $g(x)$ is well-defined. Moreover, (7.6) implies that $\sum f_i(x)$ converges *uniformly* to $g(x)$ by Weierstrass's M-Test (Theorem 7.2.3).

In the following, we show that f_i is *continuous* at x_0. Given $\varepsilon > 0$, there exists $N \in \mathbb{N}^*$ such that

$n \geq N$ implies

$$|f_i(x_n) - f_i(x_0)| = \left| \sum_{j=1}^{n} a_{ij} - \sum_{j=1}^{\infty} a_{ij} \right| < \varepsilon$$

Choose a positive number $\delta < \frac{1}{N}$, and then let $|x - x_0| < \delta$ $(x \in E)$. Since $x \in E$, $x = x_n$ for some $n \in \mathbb{N}$. Then, by the definition of x_n, it is clear $n > N$. It follows that if $|x - x_0| < \delta$, we have

$$|f_i(x_n) - f_i(x_0)| < \varepsilon$$

since $n > N$. Therefore, f_i is indeed continuous at x_0.

Recall $\sum f_i(x)$ converges uniformly to $g(x)$. It follows that g is also continuous at x_0 due to Theorem 7.3.1. Then, we have

$$g(x_0) = \lim_{x \to x_0} g(x) = \lim_{n \to \infty} g(x_n)$$

The last equation (converting limit of function limit to limit of sequence) holds because $\{x_n\}$ is indeed a sequence that converges to x_0. It then follows that

$$\begin{aligned} g(x_0) = g(x) = \lim_{n \to \infty} g(x_n) &= \lim_{n \to \infty} \sum_{i=1}^{\infty} f_i(x_n) \\ &= \lim_{n \to \infty} \sum_{i=1}^{\infty} \sum_{j=1}^{n} a_{ij} \\ &= \lim_{n \to \infty} \sum_{j=1}^{n} \sum_{i=1}^{\infty} a_{ij} \\ &= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \end{aligned}$$

Note that the reason why we have

$$\sum_{i=1}^{\infty} \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} \sum_{i=1}^{\infty} a_{ij}$$

is that $\sum_{i=1}^{\infty} a_{ij}$ converges. In summary, on the one hand, we have shown

$$g(x_0) = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij} \tag{7.7}$$

On the other hand, by definition,

$$g(x_0) = \sum_{i=1}^{\infty} f_i(x_0) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \tag{7.8}$$

We complete the proof by equating the right-hand sides of (7.7) and (7.8). ■

Theorem 7.3.3

Suppose K is a compact set, and

1. $\{f_n\}$ *is a sequence of continuous functions on* K
2. $f_n \to f$ *pointwise where* f *is also continuous*
3. $f_n(x) \geq f_{n+1}(x)$ *for all* $x \in K$, $n \in \mathbb{N}^*$

Then $f_n \to f$ *uniformly on* K.

Proof Let $g_n(x) := f_n(x) - f(x)$. It is clear that $g_n(x) \geq g_{n+1}(x) \geq 0$ and $g_n \to g \equiv 0$ pointwise. Given $\varepsilon > 0$. Define set

$$K_n := g_n^{-1}[\varepsilon, \infty)$$

Since, g_n is continuous and $[\varepsilon, \infty)$ is closed in $\mathbb{R}$, it follows that K_n is a compact subset. Consider the intersection $\bigcap_{n\in\mathbb{N}^*} K_n$. Because $g_n \geq g_{n+1}$, we have

$$K_n \supset K_{n+1}$$

We now claim that $\bigcap_{n\in\mathbb{N}^*} K_n = \emptyset$. To see this, we assume $x \in \bigcap_{n\in\mathbb{N}^*} K_n$, which then implies $g_n(x) \geq \varepsilon\ \forall n \in \mathbb{N}^*$. This leads to a contradiction since $g_n \to g \equiv 0$. Having that $\bigcap_{n\in\mathbb{N}^*} K_n = \emptyset$, by Theorem 1.2.2, it follows that there exists some finite subset $J \subset \mathbb{N}^*$ such that $\bigcap_{n\in J} K_n = \emptyset$. Suppose the largest number in J is N, then we have

$$K_N = \bigcap_{n\in J} K_n = \emptyset$$

since $K_n \supset K_{n+1}$. This further implies that

$$K_N = \emptyset \quad \forall n \geq N \tag{7.9}$$

It then follows from (7.9) and the definition of K_n that

$$0 \leq g_n(x) < \varepsilon \quad \forall n \geq N,\ \forall x \in K$$

Therefore, $g_n \to 0$ uniformly, i.e., $f_n \to f$ uniformly. ■

The compactness is necessary. Consider the following example.

Example 7.2 Let $f_n(x) = \frac{1}{nx}$ where $x \in (0,1)$, which is not a compact set in $\mathbb{R}$. Note that $f_n \to 0$ pointwise, all functions are continuous and $f_n(x) > f_{n+1}(x)$. But clearly $\{f_n\}$ does not converge to 0 uniformly on $(0,1)$.

The condition that $f_n(x) \geq f_{n+1}(x)$ is also crucial to this theorem. To illustrate this, we reconsider example 7.1.

Example 7.3 Let f_n be the same function in example 7.1. The domain $[0,1]$ is compact and f_n is continuous. We have already seen that $\{f_n\}$ does not converge to 0 uniformly on $[0,1]$. What goes wrong is that $\{f_n\}$ fails to satisfy $f_n(x) \geq f_{n+1}(x)$ (for *infinitely* many n). To see this, we evaluate f_n and f_{n+1} at the point $x = \frac{1}{\sqrt{2n+3}}$ (the maximum point of f_{n+1}). We have

$$\frac{f_n(1/\sqrt{2n+3})}{f_{n+1}(1/\sqrt{2n+3})} = \frac{n^2(2n+3)}{2(n+1)^3} = \frac{2n^3+3n^2}{2n^3+6n^2+6n+2} < 1$$

Therefore,

$$f_n(1/\sqrt{2n+3}) < f_{n+1}(1/\sqrt{2n+3}) \quad \forall n \in \mathbb{N}^*$$

If we collect a certain kind of functions in a set, we can then interpret uniform convergence of functions simply as sequential convergence. In this case, each term of the sequence is not some number but a *function*.

Definition 7.3.1

Let X be a metric space. $\mathscr{C}(X)$ will denote the set consisting of all complex-valued, bounded and continuous functions with domain X. ♣

Next, to make $\mathscr{C}(X)$ a metric space, we need to define a distance function. We do so by first defining norms.

Definition 7.3.2

The norm on $\mathscr{C}(X)$ is given by the supremum norm of each function $f \in \mathscr{C}(X)$, i.e.,

$$\|f\| := \sup_{x \in X} |f(x)|$$

♣

Remark We need to verify that Definition 7.3.2 indeed defines a norm. It is clearly positive definite, and

$$\|\alpha f\| = \sup_{x \in X} |\alpha f(x)| = |\alpha| \sup_{x \in X} |f(x)| = |\alpha| \, \|f\|$$

where $\alpha \in \mathbb{C}$. Moreover,

$$\|f + g\| = \sup_{x \in X} |f(x) + g(x)| \le \sup_{x \in X}(|f(x)| + |g(x)|) \le \sup_{x \in X} |f(x)| + \sup_{x \in X} |g(x)| = \|f\| + \|g\|$$

The distance function on $\mathscr{C}(X)$ is defined as the norm of the difference between two functions.

Definition 7.3.3

The distance function on $\mathscr{C}(X)$ is given by

$$d(f, g) := \|f - g\|$$

where $f, g \in \mathscr{C}(X)$ and $\|\cdot\|$ is the norm on $\mathscr{C}(X)$.

♣

Remark With the properties of norms, it is easy to verify that the distance function is well-defined.

Associate with a distance function, we are now able to call $\mathscr{C}(X)$ a metric space. Even better, $\mathscr{C}(X)$ is a *complete* metric space, which we shall prove below.

Theorem 7.3.4

$\mathscr{C}(X)$ is a ***complete*** *metric space.*

♡

Remark Let $\{f_n\}$ be a sequence in $\mathscr{C}(X)$. Then $f_n \to f$ uniformly on X is equivalent to that $\{f_n\}$ converges to f. Moreover, $f \in \mathscr{C}(X)$.

Proof Let $\{f_n\}$ be a Cauchy sequence in $\mathscr{C}(X)$. Then for given $\varepsilon > 0$, there exists $N \in \mathbb{N}^*$ such that $n, m \ge N$ implies

$$\sup_{x \in X} |f_n(x) - f_m(x)| = d(f_n, f_m) < \varepsilon$$

By Theorem 7.2.1, there exists some function f such that $f_n \to f$ uniformly on f. (Note that f is not necessary in $\mathscr{C}(X)$ for now. That is exactly what we need to prove.) We intend to show that $f \in \mathscr{C}(X)$. By Theorem 7.3.2, f is continuous since all f_n's are continuous. We also need to show that f is bounded. By Theorem 7.2.2, we have

$$\sup_{x \in X} |f(x) - f_N(x)| < 1$$

for some $N \in \mathbb{N}^*$. And $\sup_{x \in X} |f_N(x)| < M$ for some $M > 0$ since f_N is bounded. It then follows that

$$\sup_{x\in X}|f(x)| \le \sup_{x\in X}|f(x)-f_N(x)| + \sup_{x\in X}|f_N(x)| < 1+M$$

Therefore, f is indeed bounded. Hence, $f \in \mathscr{C}(X)$ since we have proved f is bounded and continuous (of course, f is also complex-valued). As Theorem 7.2.2 states,

$$\lim_{n\to\infty} d(f_n, f) = 0$$

with $f \in \mathscr{C}(X)$, we may conclude that $\mathscr{C}(X)$ is a complete metric space. ■

7.4 Uniform Convergence and Integration

7.5 Uniform Convergence and Differentiation

Theorem 7.5.1

Let $\{f_n\}$ be a sequence of differentiable functions defined on $[a,b]$. Suppose that the numerical sequence $\{f_n(x_0)\}$ converges where $x_0 \in [a,b]$, and the sequence $\{f_n'\}$ converges uniformly on $[a,b]$. Then $\{f_n\}$ converges uniformly to some function f on $[a,b]$. Moreover, f is differentiable on $[a,b]$, the limit of $\{f_n'(x)\}$ exists for each x, and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \tag{7.10}$$

In other words, we can interchange differentiation and limit, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}x}\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty}\frac{\mathrm{d}}{\mathrm{d}x} f_n(x)$$

♡

Remark Pay attention to the conditions of this theorem. The reason that we do not assume the uniform convergence of $\{f_n\}$ directly is that it is not sufficient to guarantee the interchange of differentiation and limit. Therefore, stronger conditions are needed. We assume the uniform convergence of the sequence of derivatives $\{f_n'\}$, and we also require that f_n converges at one point. And in fact, the uniform convergence of $\{f_n\}$ can be derived based on that.

Proof The first thing we need to show is that $\{f_n\}$ converges uniformly. Given $\varepsilon > 0$, there exists some $N_1 \in \mathbb{N}^*$ such that $n, m \ge N_1$ implies

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2} \tag{7.11}$$

since $\{f_n(x_0)\}$ converges. Furthermore, because $\{f_n'\}$ converges uniformly, there exists some $N_2 \in \mathbb{N}^*$ such that $n, m \ge N_2$ implies

$$|f_n'(x) - f_m'(x)| < \frac{\varepsilon}{2(b-a)} \quad \forall x \in [a,b] \tag{7.12}$$

Let $N = \max\{N_1, N_2\}$ and $n, m \ge N$. If we regard function $f_n(x) - f_m(x)$ as a whole, then by the Mean Value Theorem, we have

$$(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0)) = (x - x_0)\left(f_n'(\xi_x) - f_m'(\xi_x)\right) \tag{7.13}$$

for some ξ_x (depending on x) in between x and x_0. Taking into consideration (7.11), (7.12) and (7.13), it then follows that

$$\begin{aligned}|f_n(x) - f_m(x)| &\le |f_n(x_0) - f_m(x_0)| + |x - x_0|\,|f_n'(\xi_x) - f_m'(\xi_x)| \\ &< \frac{\varepsilon}{2} + (b-a)\frac{\varepsilon}{2(b-a)} \\ &= \varepsilon \quad \forall x \in [a,b]\end{aligned}$$

Therefore, $\{f_n\}$ converges uniformly on $[a,b]$ by Theorem 7.2.1. Let f be the limit of this sequence of functions, i.e., $f(x) = \lim_{n\to\infty} f_n(x)$.

We then show that f is differentiable, the limit of $\{f_n'(x)\}$ exists and (7.10) holds. Fix an x in $[a,b]$. Define

$$\begin{aligned}\phi_n(t) &:= \frac{f_n(x) - f_n(t)}{x - t} \\ \phi(t) &:= \frac{f(x) - f(t)}{x - t}\end{aligned}$$

where $t \in [a,b] \setminus \{x\}$. Let $\varepsilon > 0$ be arbitrary. By applying the same argument, we will again obtain (7.12) and (7.13) (with x replaced by t and x_0 replaced by x). It then follows from (7.13) that

$$|\phi_n(t) - \phi_m(t)| = |f_n'(\xi_t) - f_m'(\xi_t)|$$

Then by letting $n, m \ge N_2$, we have

$$|\phi_n(t) - \phi_m(t)| = |f_n'(\xi_t) - f_m'(\xi_t)| < \frac{\varepsilon}{2(b-a)} \quad \forall t \in [a,b] \setminus \{x\}$$

due to (7.12). Therefore, ϕ_n converges uniformly on $[a,b] \setminus \{x\}$. And since we have proved $f_n \to f$ as $n \to \infty$, it is clear that the limit of $\{\phi_n\}$ is ϕ, i.e., $\phi_n \to \phi$ uniformly on $[a,b] \setminus \{x\}$. Finally, by applying Theorem 7.3.1, we have

$$\lim_{t\to x} \phi(t) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f_n'(x)$$

(The existence of involved limits is implied by Theorem 7.3.1.) Note that $\lim_{t\to x} \phi(t)$ is precisely the definition of $f'(x)$. This completes the proof. ■

Example 7.4

Chapter 8 Some Special Functions

8.1 Power Series

In this section, we study the properties of power series, i.e., the function of the form

$$f(x) = \sum c_n x^n$$

where $c_n, x \in \mathbb{R}$. The reason why we confine ourselves to real values is that we have only defined differentiation and integration in the real field.

Theorem 8.1.1

Suppose that the power series

$$\sum_{n=0}^{\infty} c_n x^n$$

converges for $|x| < R$ where $c_n, x \in \mathbb{R}$ and $R > 0$. Define

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

Then f converges uniformly on $[-R+\varepsilon, R-\varepsilon]$ for any $\varepsilon \in (0, R)$. Moreover, f is differentiable in the interval $(-R, R)$, and

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$$

♡

Proof ■

Corollary 8.1.1

The power series $f(x) = \sum_{n=0}^{\infty} c_n x^n$ (which converges for $x \in (-R, R)$, $R > 0$) has derivatives of all orders in $(-R, R)$, which are given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} \left(\prod_{i=0}^{k-1} (n-i) \right) c_n x^{n-k} = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1) c_n x^{n-k} \tag{8.1}$$

where $k \in \mathbb{N}$ ($f^{(0)}$ means f). Furthermore, putting $x = 0$ in (8.1), we obtain the equality

$$c_n = \frac{f^{(n)}(0)}{n!} \tag{8.2}$$

♡

8.1.1 Negative Binomial Theorem

It is well known that

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \qquad -1 < x < 1$$

The right-hand side is a geometric series, and we can easily compute its explicit value, which equals the left-hand side. The Negative Binomial Theorem is a generalization for it states the power series representation of the function $\frac{1}{(1-x)^n}$.

Theorem 8.1.2 (Negative Binomial Theorem)

We have the following identity:

$$\frac{1}{(1-x)^n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k \tag{8.3}$$

where $x \in \mathbb{R}$, $|x| < 1$ and $n \in \mathbb{N}^*$. Moreover, the power series on right-hand side of (8.3) converges only when $|x| < 1$.

♡

Proof Let

$$f(x) = \frac{1}{1-x} \qquad -1 < x < 1$$

The $(n-1)$-th order of derivative of f is

$$f^{(n-1)}(x) = \frac{(n-1)!}{(1-x)^n} \tag{8.4}$$

On the other hand, applying (8.1) in Corollary 8.1.1, we obtain

$$f^{(n-1)}(x) = \sum_{k=n-1}^{\infty} k(k-1)\cdots(k-n+1)x^{k-n} = \sum_{k=0}^{\infty}(k+n-1)(k+n-2)\cdots(k+1)x^k \tag{8.5}$$

Comparing (8.4) and (8.5), we have

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

Therefore,

$$\frac{1}{(1-x)^n} = \sum_{k=0}^{\infty} \frac{(k+n-1)(k+n-2)\cdots(k+1)}{(n-1)!}x^k = \sum_{k=0}^{\infty}\binom{n+k-1}{k}x^k$$

So far, we have shown (8.3) indeed converges in $|x| < 1$. We also need to show that it cannot converge for $|x| \geq 1$. In fact, we only need to show it does not converge for $|x| = 1$ due to Theorem 2.4.1. Suppose $|x| = 1$. We consider the absolute value of each term of the series.

$$\left|\binom{n+k-1}{k}x^k\right| = \binom{n+k-1}{k} = \binom{n+k-1}{n-1} \tag{8.6}$$

Note that the combinatorial number (8.6) will tend to ∞ as $k \to \infty$. Thus, series (8.3) will of course diverge for $|x| = 1$. ■

We note that $\frac{1}{(1-x)^n}$ is obtained by multiplying $\frac{1}{1-x}$ by itself $n-1$ times. It is tempting to multiply the power series $\sum x^k$. Hence, another strategy for proving this theorem is to compute the *Cauchy Product* of these n series $\sum x^k$.

In the alternative proof, we will need an identity from combinatorial mathematics, which is known as the Hockey-Stick Identity.

Theorem 8.1.3 (Hockey-Stick Identity)

We have the identity:

$$\sum_{i=r}^{n}\binom{i}{r} = \binom{n+1}{r+1}$$

♡

Observe the bold numbers in Pascal's Triangle below.

$$\begin{array}{ccccccccccccccc}
&&&&&&&1&&&&&&& \\
&&&&&&1&&1&&&&&& \\
&&&&&1&&2&&\mathbf{1}&&&&& \\
&&&&1&&3&&\mathbf{3}&&1&&&& \\
&&&1&&4&&\mathbf{6}&&4&&1&&& \\
&&1&&5&&\mathbf{10}&&10&&5&&1&& \\
&1&&6&&15&&\mathbf{20}&&15&&6&&1& \\
1&&7&&21&&35&&35&&21&&7&&1
\end{array}$$

We note that the sum of the numbers on the *slope* $(1+3+6+10)$ is equal to the number at the bottom (20).

The shape of these numbers is like a hockey stick, hence the name of this identity. A quick way to memorize this identity is by associating it with the terminology of the hockey stick. The identity can be then rephrased as the numbers on the *shaft* sum up to the number on the *blade*.

Proof ■

We now provide an alternative proof of the Negative Binomial Theorem as follows.

Proof (**Proof of Theorem 8.1.2 using Cauchy Product**) We shall prove by induction on n.

(Base Case) If $n = 1$, then (8.3) is

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

which is already known.

(Inductive Step) Assume (8.3) holds for $n = m$, we shall prove it also holds for $n = m+1$. We have

$$\begin{aligned}
\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k \\
\frac{1}{(1-x)^m} &= \sum_{k=0}^{\infty} \binom{m+k-1}{k} x^k
\end{aligned}$$

Note that $\sum x^k$ converges absolutely for $|x| < 1$. Then by Theorem 2.5.2, the Cauchy Product of two series on the right-hand sides converges, and

$$\sum_{k=0}^{\infty} \sum_{j=0}^{k} x^{k-j} \binom{m+j-1}{j} x^j = \sum_{k=0}^{\infty} x^k \sum_{k=0}^{\infty} \binom{m+k-1}{k} x^k = \frac{1}{1-x} \frac{1}{(1-x)^m} = \frac{1}{(1-x)^{m+1}} \tag{8.7}$$

We now compute the Cauchy product on the left-hand side of (8.7).

$$\begin{aligned}\sum_{k=0}^{\infty}\sum_{j=0}^{k} x^{k-j}\binom{m+j-1}{j}x^j &= \sum_{k=0}^{\infty}\sum_{j=0}^{k}\binom{m+j-1}{j}x^k \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{k}\binom{m+j-1}{m-1}x^k \\ &= \sum_{k=0}^{\infty}\sum_{i=m-1}^{m+k-1}\binom{i}{m-1}x^k \\ &= \sum_{k=0}^{\infty}\binom{m+k}{m}x^k \quad \text{by Theorem 8.1.2} \\ &= \sum_{k=0}^{\infty}\binom{m+k}{k}x^k\end{aligned}$$

Hence, the Cauchy product equals

$$\sum_{k=0}^{\infty}\sum_{j=0}^{k} x^{k-j}\binom{m+j-1}{j}x^j = \sum_{k=0}^{\infty}\binom{m+k}{k}x^k \tag{8.8}$$

It then follows from (8.7) and (8.8) that

$$\frac{1}{(1-x)^{m+1}} = \sum_{k=0}^{\infty}\binom{m+k}{k}x^k$$

which is exactly (8.3) with $n = m+1$. ■

8.1.2 Taylor's Theorem

The Taylor's Theorem states that we can expand a power series as a Taylor series about some point in the interval of convergence.

Theorem 8.1.4 (Taylor's Theorem)

Suppose power series

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

converges in $|x| < R$ $(R > 0)$*. Let* $a \in (-R, R)$*, then* $f(x)$ *can be expanded in a power series about point* a*:*

$$f(x) = \sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n \tag{8.9}$$

which converges in $|x-a| < R - |a|$*.* ♡

Remark Note that the radius of convergence of the power series (8.9) is *at least* $R - |a|$ (provided that R is the radius of convergence of $\sum c_n x^n$). It may converge in a larger interval about point a. Having said that, there also exists a case that $R - |a|$ is precisely the radius of convergence of (8.9). We will illustrate this in Example 8.1.

Proof To obtain the term $(x-a)$, we apply the Binomial Theorem:

$$f(x)=\sum_{n=0}^{\infty} c_n x^n=\sum_{n=0}^{\infty} c_n\left(a+(x-a)\right)^n=\sum_{n=0}^{\infty}\sum_{k=0}^{n} c_n\binom{n}{k}a^{n-k}(x-a)^k \tag{8.10}$$

Define

$$p_{nk}:=\begin{cases} c_n\binom{n}{k}a^{n-k}(x-a)^k & k\leq n \\ 0 & k>n \end{cases}$$

Then, we can rewrite (8.10) as

$$f(x)=\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} p_{nk} \tag{8.11}$$

We intend to apply Theorem 2.5.1 to interchange the order of summations in (8.11). To do so, we need to verify

1. $\sum_{k=0}^{\infty}|p_{nk}| =: b_n<\infty$
2. $\sum_{n=0}^{\infty} b_n<\infty$

(Checking $\sum_{k=0}^{\infty}|p_{nk}| =: b_n<\infty$) We have

$$\begin{aligned} b_n=\sum_{k=0}^{\infty}|p_{nk}| &= \sum_{k=0}^{n}|p_{nk}| \\ &= \sum_{k=0}^{n}|c_n|\binom{n}{k}|a|^{n-k}|x-a|^k \\ &= |c_n|\left(|a|+|x-a|\right)^n \end{aligned}$$

Therefore,

$$b_n=|c_n|\left(|a|+|x-a|\right)^n \tag{8.12}$$

is indeed a finite number.

(Checking $\sum_{n=0}^{\infty} b_n<\infty$) Suppose that $|x-a|<R-|a|$. Then there exists some $\varepsilon>0$ such that

$$|a|+|x-a|=R-\varepsilon$$

Applying (8.12), b_n is bounded above by

$$b_n=|c_n|\left(|a|+|x-a|\right)^n<|c_n|\left(R-\varepsilon\right)^n$$

Note that Theorem 8.1.1 implies that $\sum c_n(R-\varepsilon)^n$ converges absolutely, i.e., $\sum|c_n|\left(R-\varepsilon\right)^n$ converges. It follows that $\sum b_n$ also converges by the Comparison Test.

Indeed, we are allowed to change the order of summations in (8.11). It then follows that

$$
\begin{aligned}
f(x) &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} p_{nk} \\
&= \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_{nk} \\
&= \sum_{k=0}^{\infty}\sum_{n=k}^{\infty} c_n \binom{n}{k} a^{n-k}(x-a)^k \quad \text{recall definition of } p_{nk} \\
&= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \quad \text{putting } x = a \text{ in (8.1)}
\end{aligned}
$$

■

Example 8.1 Let

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \quad -1 < x < 1$$

The derivatives of f are

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

We now expand $f(x)$ about some point $a \in (-1, 1)$ using Theorem 8.1.4

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} \frac{(x-a)^n}{(1-a)^{n+1}} \tag{8.13}$$

The radius of convergence of power series (8.13) can be computed in two ways. The first way is by the Root Test.

$$\limsup_{n\to\infty} \frac{1}{\sqrt[n]{|1-a|^{n+1}}} = \frac{1}{|1-a|}$$

Hence, the radius of convergence is

$$R = \frac{1}{\limsup_{n\to\infty} \frac{1}{\sqrt[n]{|1-a|^{n+1}}}} = |1-a|$$

On the other hand, observe that (8.13) is a geometric series. Therefore, it converges if and only if

$$\left|\frac{x-a}{1-a}\right| < 1$$

That is, $|x-a| < |1-a|$, which implies that the radius of convergence is $|1-a|$.

If $a \geq 0$, then

$$R = |1-a| = 1 - |a|$$

which means (8.13) only converges in the range that is stated in Theorem 8.1.4.

However, if $a < 0$, then

$$R = |1-a| = 1 + |a| > 1 - |a|$$

Therefore, (8.13) may converge in a larger range.

8.2 Exponential Function

In this section, we study the exponential function e^z and provide a rigorous formulation.

Proposition 8.2.1

*The series $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ converges **absolutely** for all $z \in \mathbb{C}$.*

Proof If $z = 0$, then the conclusion is trivial since the series is equal to 1. For $z \neq 0$, we apply the Ratio Test (Theorem 2.3.2) to the series $\sum \frac{|z|^n}{n!}$, we have

$$\limsup_{n\to\infty} \frac{|z|^{n+1}/(n+1)!}{|z|^n/n!} = \limsup_{n\to\infty} \frac{|z|}{n+1} = 0 < 1$$

■

Definition 8.2.1

Define

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{8.14}$$

where $z \in \mathbb{C}$.

Remark $E(z)$ is well-defined by Proposition 8.2.1.

Since the series in (8.14) converges absolutely, we may compute the product of two such series.

Proposition 8.2.2

Let $E(z)$ be as in Definition 8.2.1. We have

$$E(z)E(w) = E(z+w)$$

Proof By Theorem 2.5.2, the Cauchy product $E(z)E(w)$ converges, and it equals

$$\begin{aligned} E(z)E(w) &= \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{z^{n-k}}{(n-k)!}\frac{w^k}{k!} \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k} z^{n-k}w^k \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}(z+w)^n \\ &= E(z+w) \end{aligned}$$

The second last equality follows from the Binomial Theorem, and the last equality follows from Definition 8.2.1. ■

Proposition 8.2.3

Let $E(z)$ be as in Definition 8.2.1. We have

$$E(z)E(-z) = 1$$

If we confine ourselves to real variables, we have

1. $\lim_{x\to\infty} E(x) = \infty$, *i.e*, $E(\infty) = \infty$
2. $\lim_{x\to-\infty} E(x) = 0$, *i.e.*, $E(-\infty) = 0$

where $x \in \mathbb{R}$.

8.2.1 Complex Exponents

We are going to define

$$e^z := E(z)$$

where $z \in \mathbb{C}$. But keep in mind that we have already defined *rational* exponents. (*Irrational* exponents are not defined.) Therefore, we have to verify the consistency of the definition of complex exponents by checking

$$e^r = E(r)$$

where $r \in Q$.

Proposition 8.2.4

Let $E(z)$ be as in Definition 8.2.1, and $r \in \mathbb{Q}$ a rational number. Then,

$$(E(z))^r = E(rz)$$

In particular,

$$e^r = (E(1))^r = E(r)$$

Proof We first prove this proposition for **integer** exponents, that is,

$$(E(z))^p = E(pz) \tag{8.15}$$

where $p \in \mathbb{Z}$. If $p > 0$, then (8.15) follows from Proposition 8.2.2 and mathematical induction. If $p = 0$, then (8.15) also holds since $E(0) = 1$. Finally, if $p < 0$, then by Proposition 8.2.3 we have

$$E(pz)E(-pz) = 1$$

It follows that

$$E(pz)(E(z))^{-p} = 1$$

since $-p > 0$. Then, multiplying both sides by $E(z)^p$, we have

$$E(pz) = (E(z))^p$$

Therefore, we have proved (8.15) for $p \in \mathbb{Z}$.

Let $r \in \mathbb{Q}$. r can be written as

$$r = \frac{p}{q}$$

where $p \in \mathbb{Z}$ and $q \in \mathbb{N}^*$. By applying (8.15) two times, we obtain

$$(E(rz))^q = (E(\frac{p}{q}z))^q = E(pz) = (E(z))^p$$

Therefore,

$$E(rz) = (E(z))^{p/q} = (E(z))^r$$

■

We are now safe to define e^z as follows.

Definition 8.2.2

Define

$$e^z = E(z)$$

where $E(z)$ is as in Definition 8.2.1. ♣

8.3 Logarithm

8.4 Power Function

Part II

Multivariable Mathematical Analysis

Chapter 9 Functions of Several Variables

9.1 Linear Transformations

9.2 Differentiation

Theorem 9.2.1 (Chain Rule)

Suppose that $\mathbf{f} : E \subset \mathbb{R}^n \to \mathbb{R}^m$ *be differentiable at point* $\mathbf{x}_0 \in E$ *where* E *is open, and* $\mathbf{g} : V \supset \mathbf{f}(E) \to \mathbb{R}^k$ *be differentiable at* $\mathbf{f}(\mathbf{x}_0)$ *where* V *is also open. Then the mapping* $\mathbf{F} : E \to \mathbb{R}^k$ *defined by*

$$\mathbf{F}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$$

is differentiable at $\mathbf{x}_0$*, and*

$$\mathbf{F}'(\mathbf{x}_0) = \mathbf{g}'(\mathbf{f}(\mathbf{x}_0))\mathbf{f}'(\mathbf{x}_0)$$

♡

Proof ■

There is a very important concept associated with real-valued functions called gradient. If f is a real-valued differentiable function, we know from the previous definition its derivative f' is a row vector. The gradient of f, denoted by ∇f, is nothing but the transpose of f' so that it becomes a column vector.

Definition 9.2.1 (Gradient)

Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ *be a real-valued function where* E *is open. If* f *is differentiable at point* $\mathbf{x}_0 \in E$*, then the gradient of* f *at* $\mathbf{x}_0$ *is defined by*

$$\nabla f(\mathbf{x}_0) = (f'(\mathbf{x}_0))^\top$$

♣

Definition 9.2.2 (Directional Derivative)

Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ *be a real-valued function where* E *is open, and* $\mathbf{u} \in \mathbb{R}^n$ *a unit vector, i.e.,* $|\mathbf{u}| = 1$*. The directional derivative of* f *at point* $\mathbf{x}$ *is defined by*

$$D_{\mathbf{u}} f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x})}{t}$$

provided that the limit exists.

♣

A function is not necessarily differentiable at some point even if all the directional derivatives at this point exist and are all equal to each other. (In this case, all directional derivatives must be zeros.)

Example 9.1 Let

$$f(x, y) = \begin{cases} 1 & y = x^2,\ x \neq 0 \\ 0 & \text{elsewhere} \end{cases}$$

We consider the partial derivatives as well as the directional derivatives at the origin $(0, 0)$. The partial derivatives are

$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \frac{\partial f}{\partial y}(0,0) = 0$$

We then compute the directional derivatives. Let $\theta \in [0, 2\pi)$ and consider the direction $\mathbf{u} = (\cos\theta, \sin\theta)$ ($\theta \notin \{0, \pi/2, \pi, 3\pi/2\}$, i.e., $0 < \tan\theta < \infty$). If we choose a number t satisfying $0 < |t| < \tan\theta / \cos\theta$, then

$$f(t\cos\theta, t\sin\theta) = 0$$

since under the constraints of θ and t, $t\sin\theta = (t\cos\theta)^2$ if and only if $t = \tan\theta/\cos\theta$. Note that

$$D_{\mathbf{u}} f(0,0) = \lim_{t\to 0} \frac{f(t\cos\theta, t\sin\theta) - f(0,0)}{t} = 0$$

Thus, all the directional derivatives exist at $(0,0)$ and are equal to 0. However, f is clearly not differentiable at $(0,0)$ since it is not even continuous there.

The next proposition is a simple application of the Chain Rule which establishes a connection between gradients and directional derivatives for real-valued functions.

Proposition 9.2.1

Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ be a real-valued function where E is open. If f is differentiable at point $\mathbf{x}_0 \in E$, then all the directional derivatives at $\mathbf{x}_0$ exist. In this case, suppose $\mathbf{u} \in \mathbb{R}^n$ is a unit vector, then

$$D_{\mathbf{u}} f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{u} \tag{9.1}$$

Proof Let $\mathbf{u} \in \mathbb{R}^n$ be a unit vector. Define

$$\boldsymbol{\gamma}(t) = \mathbf{x}_0 + t\mathbf{u}$$

Since E is open and $\mathbf{x}_0 \in E$, it is possible to chose $\delta > 0$ such that $\boldsymbol{\gamma}(t) \in E \ \forall t \in (-\delta, \delta)$. Hence, $\boldsymbol{\gamma} : (-\delta, \delta) \to E$ is a well-defined function. Moreover, $\boldsymbol{\gamma}$ is differentiable in $(-\delta, \delta)$, and

$$\boldsymbol{\gamma}'(t) = \mathbf{u} \quad \forall t \in (-\delta, \delta) \tag{9.2}$$

Define a function $g : (-\delta, \delta) \to \mathbb{R}$ by

$$g(t) = f(\boldsymbol{\gamma}(t))$$

Because $\boldsymbol{\gamma}$ is differentiable at $t = 0$ and f is differentiable at $x = \boldsymbol{\gamma}(0) = \mathbf{x}_0$, it follows from the Chain Rule (Theorem 9.2.1) that g is differentiable at $t = 0$ with

$$g'(0) = f'(\boldsymbol{\gamma}(0))\boldsymbol{\gamma}'(0) = f'(\mathbf{x}_0)\mathbf{u} = \nabla f(\mathbf{x}_0) \cdot \mathbf{u} \tag{9.3}$$

The second last equality above follows from (9.2).

On the other hand, by the definition of derivatives, we have

$$g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \lim_{t\to 0} \frac{f(\boldsymbol{\gamma}(t)) - f(\boldsymbol{\gamma}(0))}{t} = \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{u}) - f(\mathbf{x}_0)}{t} = D_{\mathbf{u}} f(\mathbf{x}_0) \tag{9.4}$$

Then (9.1) follows from (9.3) and (9.4). ■

If f is not differentiable at $\mathbf{x}_0$, then (9.1) will not hold in general. Consider the following example.

Example 9.2 Let

$$f(x,y) = \begin{cases} \frac{y^3}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

We first show that f is not differentiable at the origin. The partial derivatives are

$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \frac{\partial f}{\partial y}(0,0) = 1$$

Let $\mathbf{h} = (h,k)$. Consider the expression

$$\frac{\left| f(h,k) - f(0,0) - \begin{bmatrix} \partial f/\partial x(0,0) & \partial f/\partial y(0,0) \end{bmatrix} \mathbf{h} \right|}{|\mathbf{h}|} = \frac{|h|^2\,|k|}{(h^2+k^2)^{3/2}} \tag{9.5}$$

Put $k = h$ and let $h \to 0$ in (9.5), we obtain

$$\lim_{h\to 0} \frac{|h|^2\,|k|}{(h^2+k^2)^{3/2}} = \lim_{h\to 0} \frac{|h|^3}{2^{3/2}\,|h|^3} = \frac{1}{2^{3/2}} \neq 0$$

Therefore, the left-hand side of (9.5) will not tend to 0 as $\mathbf{h} \to \mathbf{0}$, which implies f is not differentiable at $(0,0)$.

The directional derivative along the direction $\mathbf{u} = (\cos\theta, \sin\theta)$ ($\theta \in [0, 2\pi)$) at the origin is

$$D_{\mathbf{u}} f(0,0) = \sin^3\theta$$

(In particular, put $\theta = 0$ and $\theta = \pi/2$, we can also obtain the partial derivatives, which are the same as what we have calculated.) Note that this is also an example that shows f is not necessarily differentiable even if all its directional derivatives exist. The inner product of the gradient ∇f at $(0,0)$ and $\mathbf{u}$ is

$$\nabla f(0,0) \cdot \mathbf{u} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \sin\theta$$

We see that $D_{\mathbf{u}} f(0,0)$ and $\nabla f(0,0) \cdot \mathbf{u}$ are not equal to each other in general. Specifically, they are not equal for $\theta \notin \{0, \pi/2, \pi, 3\pi/2\}$.

(9.1) fails because f is not differentiable at $(0,0)$. From another point of view, the failure of (9.1) can also be applied to prove that f is not differentiable.

9.3 Mean Value Theorem

Theorem 9.3.1 (Mean Value Theorem in Several Variables)

Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable function where E is open. If $\mathbf{a}, \mathbf{b} \in E$ and the line segment between $\mathbf{a}$ and $\mathbf{b}$ lies in E, i.e.,

$$\mathbf{a} + t\mathbf{b} \in E \quad \forall t \in [0,1]$$

Then there exists $\xi \in (0,1)$ such that

$$f(\mathbf{b}) - f(\mathbf{a}) = f'(\mathbf{a} + \xi(\mathbf{b} - \mathbf{a}))(\mathbf{b} - \mathbf{a}) \tag{9.6}$$

Remark We do not require that $\mathbf{a}$ and $\mathbf{b}$ be distinct since (9.6) holds trivially if $\mathbf{a} = \mathbf{b}$ (both sides are zeros). If we assume E is *convex*, then (9.6) holds for every pair of points in E.

Proof Let curve $\boldsymbol{\gamma} : [0,1] \to E$ be defined by

$$\boldsymbol{\gamma}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a})$$

And let function $g : [0,1] \to \mathbb{R}$ be defined by

$$g(t) = f(\boldsymbol{\gamma}(t))$$

By the Chain Rule (Theorem 9.2.1), function g is differentiable on $(0,1)$ since $\boldsymbol{\gamma}$ is differentiable on $(0,1)$ and f is differentiable on E, and

$$g'(t) = f'(\boldsymbol{\gamma}(t))\boldsymbol{\gamma}'(t) = f'(\mathbf{a} + t(\mathbf{b} - \mathbf{a}))(\mathbf{b} - \mathbf{a})$$

Apply the Mean Value Theorem for single-variable functions,

$$f(\mathbf{b}) - f(\mathbf{a}) = g(1) - g(0) = g'(\xi)(1-0) = f'(\mathbf{a} + \xi(\mathbf{b} - \mathbf{a}))(\mathbf{b} - \mathbf{a})$$

■

9.4 Continuously Differentiable Functions

9.5 Contraction Principle

9.6 Inverse Function Theorem

9.7 Implicit Function Theorem

We shall prove the implicit function theorem.