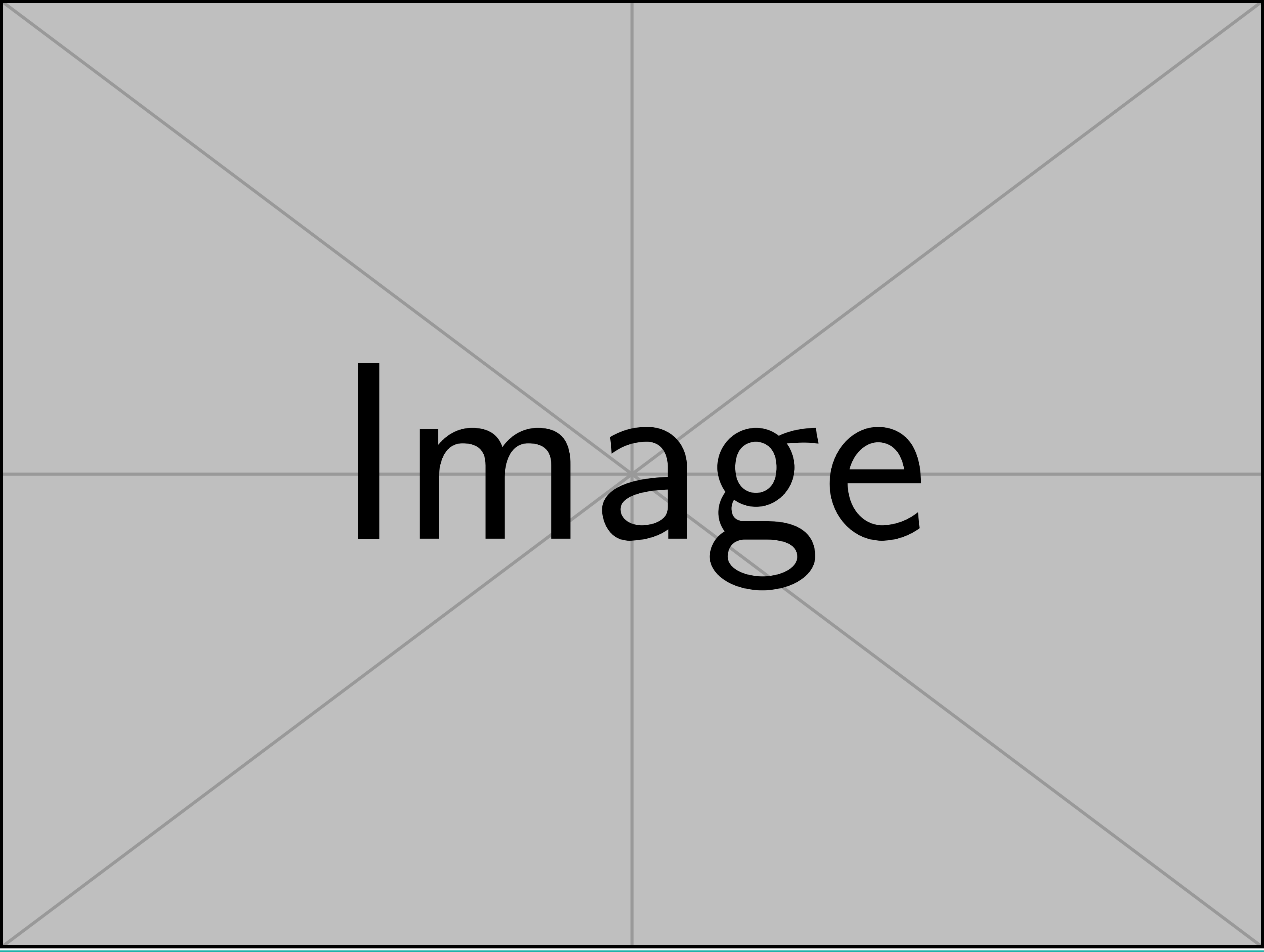


Analyse Mathématique

Author: CatMono

Date: July, 2025

Version: 0.1

Contents

Preface

This is the preface of the book...

Chapter 1 Preliminaries

1.1 Trigonometric Formulas

Product-to-Sum Formulas:

$$\sin\alpha\cos\beta = \frac{1}{2}\left[\sin(\alpha+\beta)+\sin(\alpha-\beta)\right]$$
$$\cos\alpha\sin\beta = \frac{1}{2}\left[\sin(\alpha+\beta)-\sin(\alpha-\beta)\right]$$
$$\cos\alpha\cos\beta = \frac{1}{2}\left[\cos(\alpha+\beta)+\cos(\alpha-\beta)\right]$$
$$\sin\alpha\sin\beta = -\frac{1}{2}\left[\cos(\alpha+\beta)-\cos(\alpha-\beta)\right]$$

Sum and Difference Formulas:

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta+\cos\alpha\sin\beta$$
$$\sin(\alpha-\beta) = \sin\alpha\cos\beta-\cos\alpha\sin\beta$$
$$\cos(\alpha+\beta) = \cos\alpha\cos\beta-\sin\alpha\sin\beta$$
$$\cos(\alpha-\beta) = \cos\alpha\cos\beta+\sin\alpha\sin\beta$$

Sum-to-Product Formulas:

$$\sin\alpha+\sin\beta = 2\sin\left(\frac{\alpha+\beta}{2}\right)\cos\left(\frac{\alpha-\beta}{2}\right)$$
$$\sin\alpha-\sin\beta = 2\sin\left(\frac{\alpha-\beta}{2}\right)\cos\left(\frac{\alpha+\beta}{2}\right)$$
$$\cos\alpha+\cos\beta = 2\cos\left(\frac{\alpha+\beta}{2}\right)\cos\left(\frac{\alpha-\beta}{2}\right)$$
$$\cos\alpha-\cos\beta = -2\sin\left(\frac{\alpha+\beta}{2}\right)\sin\left(\frac{\alpha-\beta}{2}\right)$$

Double Angle Formulas:

$$\sin 2\alpha = 2\sin\alpha\cos\alpha$$
$$\cos 2\alpha = \cos^2\alpha-\sin^2\alpha = 2\cos^2\alpha-1 = 1-2\sin^2\alpha$$
$$\tan 2\alpha = \frac{2\tan\alpha}{1-\tan^2\alpha}$$

Half Angle Formulas:

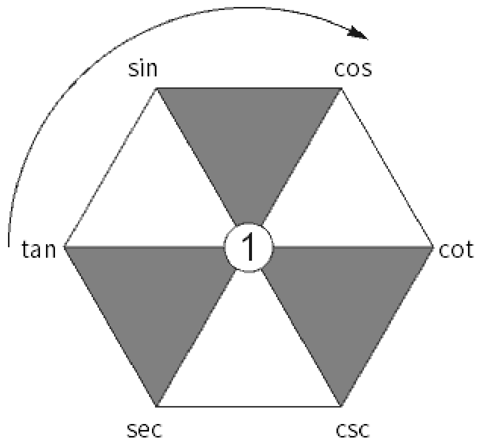
$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}$$
$$\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}}$$
$$\tan\frac{\alpha}{2} = \frac{1-\cos\alpha}{\sin\alpha} = \frac{\sin\alpha}{1+\cos\alpha}$$

Power-Reducing Formulas:

$$\sin^2\alpha = \frac{1-\cos 2\alpha}{2}$$
$$\cos^2\alpha = \frac{1+\cos 2\alpha}{2}$$

Angle Decomposition Formulas:

$$\sin^2\alpha - \sin^2\beta = \sin(\alpha+\beta)\sin(\alpha-\beta)$$
$$\cos^2\alpha - \sin^2\beta = \cos(\alpha+\beta)\cos(\alpha-\beta)$$



Remark

- On the gray triangle, the sum of the squares of the two numbers above is equal to the square of the number below, for instance, $\tan^2 x + 1 = \sec^2 x$
- The three trigonometric functions in the clockwise direction have the following properties: $\tan x = \frac{\sin x}{\cos x}$, etc.

1.2 Factorial Power

Definition 1.1

Rising factorials and falling factorials can be expressed in multiple notations.

The Pochhammer symbol, introduced by Leo August Pochhammer, is one of the commonly used notations, represented as $x^{(n)}$ or $(x)_n$.

Ronald Graham, Donald Ervin Knuth, and Oren Patashnik introduced the symbols $x^{\bar{n}}$ and $x^{\underline{n}}$ in their book *Concrete Mathematics*.

Definitions:

- **Rising factorial:**

$$x^{\bar{n}} = x(x+1)(x+2)\dots(x+n-1) = \frac{(x+n-1)!}{(x-1)!}.$$

- **Falling factorial:**

$$x^{\underline{n}} = x(x-1)(x-2)\dots(x-n+1) = \frac{x!}{(x-n)!}.$$

Relationships:

- Relationship between rising and falling factorials:

$$x^{\bar{n}} = (x+n-1)^{\underline{n}}.$$

- Relationship with factorial:

$$1^{\bar{n}} = n^{\underline{n}} = n!.$$

Chapter 2 Limits of Sequences and Continuity of Real Number System

2.1 Convergent Sequences

¶ Convergent Sequences

¶ Properties of Convergent Sequences

¶ Cauchy Proposition and Fitting Method

Proposition 2.1 (Cauchy Proposition)

Let $\lim_{n\to\infty} x_n = l$, then:

$$\lim_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n} = l.$$

Note

1. *In the proposition, l can be $+\infty$ or $-\infty$.*
2. *Let $\lim_{n\to\infty} x_n = l$, then:*

$$\lim_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n} = \lim_{n\to\infty} \sqrt[n]{x_1 x_2 \cdots x_n} = \lim_{n\to\infty} \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} = l.$$

It can be proved directly by Stolz theorem 2.1. On top of that, it can also be proved by the **fitting method**.

Proof

■

Remark To prove $\lim_{n\to\infty} x_n = A$, the key is to show that $|x_n - A|$ can be arbitrarily small. For this purpose, it is generally recommended to simplify the expression of x_n as much as possible. However, in some cases, A can also be transformed into a form similar to x_n. This method is called the fitting method. The core idea behind the method of fitting is to appropriately divide into units of 1 for analysis.

2.2 Indeterminate Form

¶ Infinitely Large Quantities and Infinitesimal Quantities

¶ Indeterminate Forms

Theorem 2.1 (Stolz-Cesàro theorem)

Type $\frac{0}{0}$ Let $\{a_n\}$, $\{b_n\}$ be two infinitesimal sequences, where $\{a_n\}$ is also a strictly monotonic decreasing sequence. If

$$\lim_{n\to\infty} \frac{b_{n+1} - b_n}{a_{n+1} - a_n} = l \text{ (finite or } \pm\infty),$$

then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = l.$$

Type $\frac{*}{\infty}$ Let $\{a_n\}$ be a strictly monotonic increasing sequence of divergent large quantities. If

$$\lim_{n\to\infty} \frac{b_{n+1} - b_n}{a_{n+1} - a_n} = l \text{ (finite or } \pm\infty),$$

then

$$\lim_{n\to\infty} \frac{a_n}{b_n} = l.$$

Note

1. *The inverse proposition of Stolz's Theorem does not hold.*
2. *If a_1 is an undefined infinite quantity ∞, Stolz Theorem does not hold.*

Theorem 2.2 (Silverman-Toeplitz Theorem)

Let

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \\ \vdots \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \vdots \end{bmatrix},$$

where the infinite triangular matrix satisfies:

1. $\forall j, \lim_{n\to\infty} a_{nj} = 0$.(Every column sequence converges to 0.)
2. $\sup_{i\in\mathbb{N}} \sum_{j=1}^{i} |a_{ij}| < \infty$.(The absolute row sums are bounded.)

And $\lim_{n\to\infty} x_n = l$. We denote y_n as the weighted sum sequence: $y_n = \sum_{j=1}^{n} a_{nj}x_j$. Then the following results hold:

1. If $l = 0$, then $\lim_{n\to\infty} y_n = 0$.
2. If $l \neq 0$ and $\lim_{n\to\infty} \sum_{j=1}^{n} a_{ij} = 1$, then $\lim_{n\to\infty} y_n = l$.

2.3 Subsequences

¶ `Subsequences`

¶ `Upper Limits and Lower Limits`

2.4 Completeness of The Real Numbers

¶ `Dedkind Completeness`

¶ `Least Upper Bound Property`

¶ `Monotone Convergence Theorem`

¶ `Bolzano-Weierstrass Theorem`

¶ `Nested Interval Theorem`

¶ `Cauchy Completeness`

Definition 2.1 (Cauchy Sequence)

A sequence $\{x_n\}$ is called a **Cauchy sequence** if for any $\varepsilon > 0$, there exists a positive integer N such that when $m, n > N$,

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

Theorem 2.3 (Cauchy Convergence Criterion for Sequences)

A sequence $\{x_n\}$ converges if and only if it is a Cauchy sequence. ♡

¶ Heine-Borel Theorem

2.5 Iterative Sequences

Formally, x_0 is a **fixed point** of the function f if $f(x_0) = x_0$.

Theorem 2.4 (Banach Fixed-Point Theorem (Contraction Mapping Theorem))

There exists a contraction mapping (in 3.2) f on an interval I, which admits a unique fixed point $x^* \in I$. Furthermore, x^* can be found as follows: start with an arbitrary point $x_0 \in I$ and define the iterative sequence $x_{n+1} = f(x_n)$ for $n = 0, 1, 2, \cdots$. Then $\lim_{n\to\infty} x_n = x^*$. ♡

Remark The following inequalities are equivalent and describe the speed of convergence:

$$|x_n - x^*| \leqslant \frac{L^n}{1-L} |x_1 - x_0|,$$
$$|x_{n+1} - x^*| \leqslant \frac{L}{1-L} |x_{n+1} - x_n|,$$
$$|x_{n+1} - x^*| \leqslant L |x_n - x^*|.$$

Any such value of $L < 1$ is the Lipschitz constant for f, and the smallest one is sometimes called **the best Lipschitz constant** of L.

Chapter 3 Limits and Continuity of Functions

3.1 Limits of Functions

¶ **Definition of Limit**

¶ **Limits of Functions and Sequences**

Theorem 3.1 (Heine Theorem)

Let f be a function defined on a deleted neighborhood $\mathring{U}(x_0)$ of x_0. The following two statements are equivalent:

1. $\lim_{x\to x_0} f(x) = A$.
2. For any sequence $\{x_n\} \subset \mathring{U}(x_0)$ with $\lim_{n\to\infty} x_n = x_0$, we have $\lim_{n\to\infty} f(x_n) = A$ for the sequence $\{f(x_n)\}$.

♡

3.2 Continuous Functions

3.3 Infinitesimal and Infinite Quantities

3.4 Continuous Functions on Closed Intervals

¶ **Concerning Theorems**

Theorem 3.2 (The Bolzano-Cauchy Intermediate-Value Theorem)

♡

Theorem 3.3 (Zero Point Existence Theorem)

♡

¶ **Uniform Continuity and Lipschitz Continuity**

Definition 3.1 (Uniform Continuity)

♣

Theorem 3.4 (Uniform Continuity Theorem)

♡

Theorem 3.5 (Cantor's Theorem)

♡

Definition 3.2 (Lipschitz Continuity)

If there exists a constant $L > 0$ such that for any $x_1, x_2 \in I$,

$$|f(x_1) - f(x_2)| \leq L\,|x_1 - x_2|\,,$$

then f is called **Lipschitz continuous** on I.

Specially, if $L < 1$, then f is called a **contraction mapping** on I.

♣

Remark

- If f is Lipschitz continuous on I, then f is uniformly continuous on I. ($\forall \varepsilon > 0$, just let $\delta = \frac{\varepsilon}{L}$)
- If f is uniformly continuous on I, then f is continuous on I.
- The converse of the above two statements does not hold.

3.5 Period Three Implies Chaos

3.6 Functional Equations

Chapter 4 Differential

4.1 Differential and Derivative

¶ Basic Differential Rules and Formulas

	Derivative Rules	Differential Rules
Linear Combination	$(c_1 f + c_2 g)' = c_1 f' + c_2 g'$	$\mathrm{d}(c_1 f + c_2 g) = c_1 \mathrm{d}f + c_2 \mathrm{d}g$
Product Rule	$(fg)' = f'g + fg'$	$\mathrm{d}(fg) = g\mathrm{d}f + f\mathrm{d}g$
Quotient Rule	$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$	$\mathrm{d}\left(\frac{f}{g}\right) = \frac{g\mathrm{d}f - f\mathrm{d}g}{g^2}$
Inverse Function	$[f^{-1}(y)]' = \frac{1}{f'(x)}$	$\mathrm{d}x = \frac{\mathrm{d}y}{f'(x)} = [f^{-1}(y)]'\mathrm{d}y$
Chain Rule	$[f(g(x))]' = f'(u)g'(x)$	$\mathrm{d}[f(g(x))] = f'(u)g'(x)\mathrm{d}x$

Derivative	Differential
$(C)' = 0$	$\mathrm{d}(C) = 0 \cdot \mathrm{d}x = 0$
$(x^\alpha)' = \alpha x^{\alpha-1}$	$\mathrm{d}(x^\alpha) = \alpha x^{\alpha-1}\mathrm{d}x$
$(\sin x)' = \cos x$	$\mathrm{d}(\sin x) = \cos x\mathrm{d}x$
$(\cos x)' = -\sin x$	$\mathrm{d}(\cos x) = -\sin x\mathrm{d}x$
$(\tan x)' = \sec^2 x$	$\mathrm{d}(\tan x) = \sec^2 x\mathrm{d}x$
$(\cot x)' = -\csc^2 x$	$\mathrm{d}(\cot x) = -\csc^2 x\mathrm{d}x$
$(\sec x)' = \tan x \sec x$	$\mathrm{d}(\sec x) = \tan x \sec x\mathrm{d}x$
$(\csc x)' = -\cot x \csc x$	$\mathrm{d}(\csc x) = -\cot x \csc x\mathrm{d}x$
$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$	$\mathrm{d}(\arcsin x) = \frac{1}{\sqrt{1-x^2}}\mathrm{d}x$
$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$	$\mathrm{d}(\arccos x) = -\frac{1}{\sqrt{1-x^2}}\mathrm{d}x$
$(\arctan x)' = \frac{1}{1+x^2}$	$\mathrm{d}(\arctan x) = \frac{1}{1+x^2}\mathrm{d}x$
$(\mathrm{arccot} x)' = -\frac{1}{1+x^2}$	$\mathrm{d}(\mathrm{arccot} x) = -\frac{1}{1+x^2}\mathrm{d}x$
$(a^x)' = \ln a \cdot a^x,\ (e^x)' = e^x$	$\mathrm{d}(a^x) = \ln a \cdot a^x\mathrm{d}x,\ \mathrm{d}(e^x) = e^x\mathrm{d}x$
$(\log_a x)' = \frac{1}{x\ln a},\ (\ln x)' = \frac{1}{x}$	$\mathrm{d}(\log_a x) = \frac{1}{x\ln a}\mathrm{d}x,\ \mathrm{d}(\ln x) = \frac{1}{x}\mathrm{d}x$
$(\mathrm{sh}\, x)' = \mathrm{ch}\, x$	$\mathrm{d}(\mathrm{sh}\, x) = \mathrm{ch}\, x\mathrm{d}x$
$(\mathrm{ch}\, x)' = \mathrm{sh}\, x$	$\mathrm{d}(\mathrm{ch}\, x) = \mathrm{sh}\, x\mathrm{d}x$
$(\mathrm{th}\, x)' = \mathrm{sech}^2\, x$	$\mathrm{d}(\mathrm{th}\, x) = \mathrm{sech}^2\, x\mathrm{d}x$
$(\mathrm{cth}\, x)' = -\mathrm{csch}^2\, x$	$\mathrm{d}(\mathrm{cth}\, x) = -\mathrm{csch}^2\, x\mathrm{d}x$
$(\mathrm{arcsh}\, x)' = \frac{1}{\sqrt{1+x^2}}$	$\mathrm{d}(\mathrm{arcsh}\, x) = \frac{1}{\sqrt{1+x^2}}\mathrm{d}x$
$(\mathrm{arcch}\, x)' = \frac{1}{\sqrt{x^2-1}}$	$\mathrm{d}(\mathrm{arcch}\, x) = \frac{1}{\sqrt{x^2-1}}\mathrm{d}x$
$(\mathrm{arcth}\, x)' = (\mathrm{arccth}\, x)' = \frac{1}{1-x^2}$	$\mathrm{d}(\mathrm{arcth}\, x) = \mathrm{d}(\mathrm{arccth}\, x) = \frac{1}{1-x^2}\mathrm{d}x$
$\ln(x + \sqrt{x^2+a^2})' = \frac{1}{\sqrt{x^2+a^2}}$	$\mathrm{d}[\ln(x + \sqrt{x^2+a^2})] = \frac{\mathrm{d}x}{\sqrt{x^2+a^2}}$

4.2 Higher-Order Derivatives

4.3 Differential Mean Value Theorems

Definition 4.1 (Extremum)

Let $f(x)$ is defined on (a,b), $x_0 \in (a,b)$. If there exists $U(x_0,\delta) \subset (a,b)$ such that $f(x) \leqslant f(x_0)$ on it, then x_0 is called a local maximum point of f, and $f(x_0)$ is referred to as the corresponding local maximum value. The definition of the minimum value is analogous.

Lemma 4.1 (Fermat's Lemma)

If f is differentiable at x_0 which is a local extremum, then $f'(x_0) = 0$.

Theorem 4.1 (Rolle's Theorem)

If $f \in C[a,b]$, $f \in D(a,b)$ and $f(a) = f(b)$, then there exists $\xi \in (a,b)$ such that $f'(\xi) = 0$.
Enhanced Version: If $f \in D(a,b)$(finite or infinite interval), and $\lim_{x\to a^+} f(x) = \lim_{x\to b^-} f(x)$, then there exists $\xi \in (a,b)$ such that $f'(\xi) = 0$.

Theorem 4.2 (Lagrange's Mean Value Theorem)

If $f \in C[a,b]$, $f \in D(a,b)$, then there exists $\xi \in (a,b)$ such that

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

Theorem 4.3 (Cauchy's Mean Value Theorem)

If $f,g \in C[a,b]$, $f,g \in D(a,b)$ and $g'(x) \neq 0$ for all $x \in (a,b)$, then there exists $\xi \in (a,b)$ such that

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

4.4 Theorems and Applications concerning Derivatives

Theorem 4.4 (Darboux's Intermediate Value Theorem for Derivatives)

If $f(x) \in D[a,b]$, and $f'_+(a) \cdot f'_-(b) < 0$, then there at least exists $\xi \in (a,b)$ such that $f'(\xi) = 0$.

Theorem 4.5 (Theorem on the Limit of Derivatives)

If $f(x) \in C(U(x_0))$, $\mathring{D}(U(x_0))$, and $\lim_{x\to x_0} f'(x) = A$, then f is differentiable at x_0 and $f'(x_0) = A$.

Remark In fact, $\lim_{x\to x_0} f'(x) = A$ has already been shown to imply that $f \in \mathring{D}(U(x_0))$.

4.5 Taylor Theorem

4.6 Applications of Taylor Theorem

Chapter 5 Indefinite Integral

5.1 Two Common Integration Methods

¶ **Integration Methods**

Definition 5.1 (Integration by Parts)

Let $u(x)$ and $v(x)$ be two differentiable functions, and at least one of them has an antiderivative. Then the integration by parts formula states that:

$$\int u(x)\,\mathrm{d}v(x) = u(x)v(x) - \int v(x)\,\mathrm{d}u(x).$$

♣

¶ **Basic Integration Formulas**

Integral	Result
$\int a\,\mathrm{d}x$	$ax + C \quad (a \text{ is constant})$
$\int x^n\,\mathrm{d}x$	$\frac{x^{n+1}}{n+1} + C \quad (n \neq -1)$
$\int \frac{1}{x}\,\mathrm{d}x$	$\ln\lvert x\rvert + C$
$\int e^x\,\mathrm{d}x$	$e^x + C$
$\int a^x\,\mathrm{d}x$	$\frac{a^x}{\ln a} + C \quad (a > 0, a \neq 1)$
$\int \sin x\,\mathrm{d}x$	$-\cos x + C$
$\int \cos x\,\mathrm{d}x$	$\sin x + C$
$\int \tan x\,\mathrm{d}x$	$-\ln\lvert\cos x\rvert + C$
$\int \cot x\,\mathrm{d}x$	$\ln\lvert\sin x\rvert + C$
$\int \sec x\,\mathrm{d}x$	$\ln\lvert\sec x + \tan x\rvert + C$
$\int \csc x\,\mathrm{d}x$	$\ln\lvert\csc x - \cot x\rvert + C$
$\int \sec x \tan x\,\mathrm{d}x$	$\sec x + C$
$\int \csc x \cot x\,\mathrm{d}x$	$-\csc x + C$
$\int \sec^2 x\,\mathrm{d}x$	$\tan x + C$
$\int \csc^2 x\,\mathrm{d}x$	$-\cot x + C$
$\int \frac{1}{\sqrt{a^2-x^2}}\,\mathrm{d}x$	$\arcsin\left(\frac{x}{a}\right) + C$
$\int \frac{-1}{\sqrt{a^2-x^2}}\,\mathrm{d}x$	$\arccos\left(\frac{x}{a}\right) + C$
$\int \frac{1}{a^2+x^2}\,\mathrm{d}x$	$\frac{1}{a}\arctan\left(\frac{x}{a}\right) + C$
$\int \frac{-1}{a^2+x^2}\,\mathrm{d}x$	$\frac{1}{a}\operatorname{arccot}\left(\frac{x}{a}\right) + C$
$\int \frac{1}{\sqrt{x^2+a^2}}\,\mathrm{d}x$	$\ln\lvert x + \sqrt{x^2+a^2}\rvert + C$
$\int \frac{1}{\sqrt{x^2-a^2}}\,\mathrm{d}x$	$\ln\lvert x + \sqrt{x^2-a^2}\rvert + C \quad (x > a \text{ or } x < -a)$
$\int \sinh x\,\mathrm{d}x$	$\cosh x + C$
$\int \cosh x\,\mathrm{d}x$	$\sinh x + C$

Chapter 6 Definite Integral

6.1 Riemann Integral

¶ **Riemann Integral**

Definition 6.1 (Riemann Integral)

Let $f(x)$ be a bounded function defined on $[a,b]$. Take any set of division points $\{x_i\}_{i=0}^n$ on $[a,b]$ to form a partition $P: a = x_0 < x_1 < \cdots < x_n = b$, and choose arbitrary points $\xi_i \in [x_{i-1}, x_i]$. Denote the length of the sub-interval $[x_{i-1}, x_i]$ as $\Delta x_i = x_i - x_{i-1}$, and let $\lambda = \max\limits_{1 \leqslant i \leqslant n}(\Delta x_i)$. If the limit

$$\lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i) \Delta x_i$$

exists as $\lambda \to 0$, and the limit is independent of the partition P and the choice of ξ_i, then $f(x)$ is said to be **Riemann integrable** on $[a,b]$.

The summation

$$S_n = \sum_{i=1}^{n} f(\xi_i) \Delta x_i$$

is called the Riemann sum, and its limit I is called the definite integral of $f(x)$ on $[a,b]$, denoted as:

$$I = \int_a^b f(x)\,\mathrm{d}x,$$

where a and b are called the lower and upper limits of the definite integral, respectively.

Alternatively, it can also be expressed as:

$$\exists I, \forall \varepsilon > 0, \exists \delta > 0, \text{s.t.} \forall P(\lambda = \max_{1 \leqslant i \leqslant n}(\Delta x_i) < \delta), \forall \{\xi_i\} : \left| \sum_{i=1}^{n} f(\xi_i) \Delta x_i - I \right| < \varepsilon.$$

Then $f(x)$ is said to be Riemann integrable on $[a,b]$, and I is the definite integral of $f(x)$ on $[a,b]$. ♣

Remark Partition → Intermediate points → Summation → Take the limit.

¶ **Darboux Sum**

Definition 6.2 (Darboux Sum)

Let the supremum and infimum of $f(x)$ on $[a,b]$ be M and m, respectively. Clearly, $m \leqslant f(x) \leqslant M$. Let the supremum and infimum of $f(x)$ on $[x_{i-1}, x_i]$ be M_i and m_i $(i = 1, 2, \ldots, n)$, respectively, i.e.,

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

After fixing the partition P, define the sums:

$$\bar{S}(P) = \sum_{i=1}^{n} M_i \Delta x_i, \quad \underline{S}(P) = \sum_{i=1}^{n} m_i \Delta x_i,$$

which are called the Darboux upper sum and Darboux lower sum corresponding to the partition P, respectively. ♣

Property

1. $\underline{S}(P) \leqslant \sum_{i=1}^{n} f(\xi_i) \Delta x_i \leqslant \bar{S}(P)$.
2. *If a new partition is formed by adding division points to the original partition, the upper sum does not increase, and the lower sum does not decrease.*

3. *Let $\bar{S}$ denote the set of Darboux upper sums and $\underline{S}$ denote the set of Darboux lower sums. For any $\bar{S}(P_1) \in \bar{\boldsymbol{S}}, \underline{S}(P_2) \in \underline{\boldsymbol{S}}$, it always holds that:*
$$m(b-a) \leqslant \underline{S}(P_2) \leqslant \bar{S}(P_1) \leqslant M(b-a).$$
4. *Let $L = \inf\{\bar{S}(P) \mid \bar{S}(P) \in \bar{\boldsymbol{S}}\}, l = \sup\{\underline{S}(P) \mid \underline{S}(P) \in \underline{\boldsymbol{S}}\}$, which are called the upper integral and lower integral, respectively. It always holds that: $l \leqslant L$.*
5. ***Darboux's Theorem**: For any $f(x) \in B[a,b]$, it always holds that:*
$$\lim_{\lambda \to 0} \bar{S}(P) = L, \quad \lim_{\lambda \to 0} \underline{S}(P) = l.$$

¶ **Riemann-Stieltjes Integral**

Definition 6.3 (Riemann-Stieltjes Integral)

Let α be a bounded, monotonically increasing function on $[a,b]$. For every partition P of $[a,b]$, let $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1})$ (clearly $\Delta\alpha_i \geqslant 0$). For a bounded real function $f(x)$ on $[a,b]$, define the Stieltjes upper sum and lower sum as:
$$\bar{S}(P,\alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i, \quad \underline{S}(P,\alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i,$$
and define the upper and lower integrals as:
$$L = \inf\{\bar{S}(P,\alpha) \mid \bar{S}(P,\alpha) \in \bar{\boldsymbol{S}}\}, \quad l = \sup\{\underline{S}(P,\alpha) \mid \underline{S}(P,\alpha) \in \underline{\boldsymbol{S}}\},$$
where $\bar{\boldsymbol{S}}, \underline{\boldsymbol{S}}$ are the sets of Stieltjes upper and lower sums respectively.
If $L = l$, then:
$$\int_a^b f(x)\, \mathrm{d}\alpha(x) = L = l,$$
and $f(x)$ is said to be **Riemann-Stieltjes integrable** on $[a,b]$ with respect to α, or simply Stieltjes integrable. ♣

When $\alpha(x) = x$, this reduces to the Riemann integral. However, in general, $\alpha(x)$ does not even need to be continuous.

The properties of Darboux sums also apply to Stieltjes sums.

6.2 Integrability Criteria

¶ **Common Integrability Criteria**

Theorem 6.1 (Integrability Criterion)

A bounded function $f(x)$ is Riemann integrable on $[a,b]$ if and only if:

- The upper and lower integrals are equal, i.e.,
$$\forall P(\lambda = \max_{1 \leqslant i \leqslant n} (\Delta x_i) < \delta) : \lim_{\lambda \to 0} \bar{S}(P) = L = l = \lim_{\lambda \to 0} \underline{S}(P).$$
- Let $\omega_i = M_i - m_i$ be the oscillation of $f(x)$ on $[x_{i-1}, x_i]$. Then: The limit of the sum of oscillations is zero, i.e.,
$$\forall P(\lambda = \max_{1 \leqslant i \leqslant n} (\Delta x_i) < \delta) : \lim_{\lambda \to 0} \sum_{i=1}^{n} \omega_i \Delta x_i = 0.$$

Corollary 1 Continuous functions on closed intervals are necessarily integrable.

Corollary 2 Monotonic functions on closed intervals are necessarily integrable.

- For all $\varepsilon > 0$, there exists a partition P such that:

$$\sum_{i=1}^{n} \omega_i \Delta x_i < \varepsilon.$$

Corollary 1 The total length of intervals where oscillation ω cannot be arbitrarily small can be made arbitrarily small, i.e.,

$$\forall \varepsilon, \eta > 0, \exists P, \text{s.t.} \sum_{\omega \geqslant \eta} \Delta x_i < \varepsilon.$$

Corollary 2 Bounded functions with only finitely many discontinuities on closed intervals are necessarily integrable.

Proof

■

¶ Lesbesgue's Theorem

Theorem 6.2 (Lesbesgue's Theorem)

6.3 Properties of Definite Integrals

¶ Properties of Riemann Integrals

Property

Linearity *Let $f(x), g(x) \in R[a,b]$, and k_1, k_2 are constants. Then the function $k_1 f(x) + k_2 g(x) \in R[a,b]$, and:*

$$\int_a^b [k_1 f(x) + k_2 g(x)] \mathrm{d}x = k_1 \int_a^b f(x) \mathrm{d}x + k_2 \int_a^b g(x) \mathrm{d}x.$$

Multiplicative Integrability *Let $f(x), g(x) \in R[a,b]$, and k_1, k_2. Then $f(x) \cdot g(x) \in R[a,b]$. In general,*

$$\int_a^b f(x) g(x) \mathrm{d}x \neq \left(\int_a^b f(x) \mathrm{d}x \right) \cdot \left(\int_a^b g(x) \mathrm{d}x \right).$$

Monotonicity *Let $f(x), g(x) \in R[a,b]$, and $f(x) \geqslant g(x)$ (or $f(x) > g(x)$) on $[a,b]$. Then:*

$$\int_a^b f(x) \mathrm{d}x \geqslant \int_a^b g(x) \mathrm{d}x \quad \left(\int_a^b f(x) \mathrm{d}x > \int_a^b g(x) \mathrm{d}x \right).$$

Corollary 1 *If $f(x) \in C[a,b]$, $f(x) \geqslant 0$, $f(x) \not\equiv 0$, then:*

$$\int_a^b f(x)\, \mathrm{d}x > 0.$$

Corollary 2 *If $f(x) \in R[a,b]$, $f(x) > 0$, then:*

$$\int_a^b f(x)\, \mathrm{d}x > 0.$$

Absolute Value Integrability *Let $f(x) \in R[a,b]$. Then $|f(x)| \in R[a,b]$, and:*

$$\left| \int_a^b f(x) \mathrm{d}x \right| \leqslant \int_a^b |f(x)| \mathrm{d}x.$$

The inverse statement of this property is not true.

Additivity Over Intervals *Let $f(x) \in R[a,b]$. For any point $c \in [a,b]$, $f(x)$ is integrable on $[a,b]$ and $[c,d]$. Conversely, if $f \in R[a,c] \cup [c,b]$, then $f(x)$ is integrable on $[a,b]$, and:*

$$\int_a^b f(x) \mathrm{d}x = \int_a^c f(x) \mathrm{d}x + \int_c^b f(x) \mathrm{d}x.$$

Theorem 6.3 (Integral Mean Value Theorem)

First Integral Mean Value Theorem Let $f(x), g(x) \in R[a,b]$, and $g(x)$ does not change sign on $[a,b]$. Then there exists $\eta \in [m, M]$ such that:

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

where m, M represent the infimum and supremum of $f(x)$ on $[a,b]$, respectively.
In particular, if $f(x) \in C[a,b]$, then there exists $\xi \in [a,b]$ such that:

$$\int_a^b f(x)g(x)\mathrm{d}x = f(\xi) \int_a^b g(x)\mathrm{d}x.$$

A special case arises when $f(x) \in C[a,b]$ and $g(x) \equiv 1$, then:

$$\int_a^b f(x)g(x)\mathrm{d}x = f(\xi) \int_a^b g(x)\mathrm{d}x.$$

Corollary If $f(x) \in C[a,b]$, then there exists $\xi \in (a,b)$ such that:

$$\int_a^b f(x)g(x)\mathrm{d}x = f(\xi) \int_a^b g(x)\mathrm{d}x.$$

Second Integral Mean Value Theorem (Bonnet Formula) Let $f(x) \in R[a,b]$,

- If $g(x)$ is decreasing on $[a,b]$ and $g(x) \geqslant 0$ $(x \in [a,b])$:

$$\exists \xi \in [a,b]: \quad \int_a^b f(x)g(x)\mathrm{d}x = g(a) \int_a^\xi f(x)\mathrm{d}x.$$

- If $g(x)$ is increasing on $[a,b]$ and $g(x) \geqslant 0$ $(x \in [a,b])$:

$$\exists \eta \in [a,b]: \quad \int_a^b f(x)g(x)\mathrm{d}x = g(b) \int_\eta^b f(x)\mathrm{d}x.$$

The general form is: Let $f(x) \in R[a,b]$, and $g(x)$ be a monotonic function. Then:

$$\exists \xi \in [a,b], \quad \int_a^b f(x)g(x)\mathrm{d}x = g(a) \int_a^\xi f(x)\mathrm{d}x + g(b) \int_\xi^b f(x)\mathrm{d}x.$$

Note *For the first integral mean value theorem,*

- *If $f(x) \in C[a,b]$ is replaced with $f(x) \in R[a,b]$, the conclusion does not hold.*
- *If $f(x) \in R[a,b]$ and $\int f(x)\mathrm{d}x$ exists, the conclusion holds.*

¶ `Integrability of Composite Functions`

Outer Continuity, Inner Integrability Let $f(x) \in R[a,b]$, $A \leqslant f(x) \leqslant B$, and $g(u) \in C[A,B]$. Then the composite function $g(f(x)) \in R[a,b]$.

Outer Integrability, Inner Continuity In this case, the composite function may not be integrable.

Both Inner and Outer Integrability In this case, the composite function may not be integrable. In fact, even if both the inner and outer functions are not integrable, the composite function may still be integrable.

6.4 Fundamental Theorem of Calculus

¶ `Newton-Leibniz Formula`

Definition 6.4 (Variable Limit Integrals)

Let $f(x) \in R[a,b]$. Define:

$$F(x) = \int_a^x f(t)\,\mathrm{d}t \quad \text{and} \quad F(x) = \int_x^b f(t)\,\mathrm{d}t,$$

which are referred to as the variable upper limit integral and variable lower limit integral, respectively.

Property

Continuity of Antiderivative $F(x) \in C[a,b]$ *(The variable upper limit integral satisfies the Lipschitz condition and is uniformly continuous on the closed interval).*

Fundamental Theorem of Calculus *Let $x_0 \in [a,b]$ be a point where $f(x)$ is continuous. Then:*

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

Existence of Antiderivatives *If $f(x) \in C[a,b]$, then $F(x) \in D[a,b]$ and $F'(x) = f(x)$.*

Rule of Derivation *If $F(x) = \int_{u(x)}^{v(x)} f(t)\,\mathrm{d}t$, then:*

$$F'(x) = f(v(x))v'(x) - f(u(x))u'(x).$$

*In fact, the formula is the simplified version of the **Leibniz's law**.*

Remark Differentiation can reduce the smoothness of functions (the original function may be differentiable, while the derivative may have second-type discontinuities), whereas integration can improve smoothness.

Theorem 6.4 (Newton-Leibniz Formula)

Let $f(x) \in C[a,b]$, and $F(x)$ be an antiderivative of $f(x)$ on $[a,b]$. Then:

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

Generalized Newton-Leibniz Formula Let $f(x) \in R[a,b]$, $F(x) \in C[a,b]$, and $F'(x) = f(x)$ holds except for finitely many points. Then:

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

¶ Riemann Lemma

¶ Common Questions concerning Integrals

6.5 Calculation of Definite Integrals

6.6 Integral Inequalities

Theorem 6.5 (Integral Inequalities)

Hadamard Inequality Let $f(x)$ be a convex function on (a,b). Then for any pair $x_1, x_2 \in (a,b)$ with $x_1 < x_2$, we have:

$$f\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} f(t)\,\mathrm{d}t \leqslant \frac{f(x_1) + f(x_2)}{2}.$$

Schwarz Inequality Let $f(x), g(x) \in R[a,b]$. Then:

$$\left(\int_a^b f(x)g(x)\,\mathrm{d}x\right)^2 \leqslant \int_a^b f^2(x)\,\mathrm{d}x \int_a^b g^2(x)\,\mathrm{d}x.$$

Hölder Inequality Let $f(x), g(x) \in R[a, b]$, and p, q are conjugate numbers (i.e., $p > 0, q > 0, \frac{1}{p} + \frac{1}{q} = 1$). Then:

$$\int_a^b |f(x)g(x)| \, dx \leqslant \left(\int_a^b |f(x)|^p \, dx\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q \, dx\right)^{\frac{1}{q}}.$$

Young Inequality Let $y = f(x) \in C[0, +\infty)$, strictly increasing, and $f(0) = 0$. Denote its inverse function as $x = f^{-1}(y)$. Then:

$$\int_0^a f(x) \, dx + \int_0^b f^{-1}(y) \, dy \geqslant ab \quad (a > 0, b > 0).$$

Minkowski Inequality Let $f(x), g(x) \in R[a, b]$. Then:

$$\left\{\int_a^b [f(x) + g(x)]^2 \, dx\right\}^{\frac{1}{2}} \leqslant \left[\int_a^b f^2(x) \, dx\right]^{\frac{1}{2}} + \left[\int_a^b g^2(x) \, dx\right]^{\frac{1}{2}}.$$

Chebyshev Inequality Let $f(x), g(x)$ be similarly ordered functions, i.e., $\forall x_1, x_2 : (f(x_1) - f(x_2))(g(x_1) - g(x_2)) \geqslant 0$. Then:

$$\int_a^b f(x) \, dx \int_a^b g(x) \, dx \leqslant (b - a) \int_a^b f(x)g(x) \, dx.$$

Discrete Form Let sequences $\{a_n\}, \{b_n\}$ be similarly ordered, i.e., $\forall i, j : (a_i - a_j)(b_i - b_j) \geqslant 0$. Then:

$$\left(\sum_{i=1}^n a_i\right)\left(\sum_{i=1}^n b_i\right) \leqslant n \sum_{i=1}^n a_i b_i.$$

If the sequences are oppositely ordered, the inequality reverses.

6.7 Applications of Definite Integrals

¶ `Arc Length`

¶ `Curvature`

¶ `Polar Coordinate System`

Category	Explicit Cartesian Equation	Parametric Cartesian Equation	Polar Equation
Equation	$y = f(x), x \in [a, b]$	$\begin{cases} x = x(t), t \in [T_1, T_2], \\ y = y(t), \end{cases}$	$r = r(\theta), \theta \in [\alpha, \beta]$
Area of Plane Shape	$\int_a^b f(x) \, dx$	$\int_{T_1}^{T_2} \lvert y(t)x'(t) \rvert \, dt$	$\frac{1}{2} \int_\alpha^\beta r^2(\theta) \, d\theta$
Infinitesimal Arc Length	$dl = \sqrt{1 + [f'(x)]^2} \, dx$	$dl = \sqrt{[x'(t)]^2 + [y'(t)]^2} \, dt$	$dl = \sqrt{r^2(\theta) + r'^2(\theta)} \, d\theta$
Curve Length	$\int_a^b \sqrt{1 + [f'(x)]^2} \, dx$	$\int_{T_1}^{T_2} \sqrt{[x'(t)]^2 + [y'(t)]^2} \, dt$	$\int_\alpha^\beta \sqrt{r^2(\theta) + r'^2(\theta)} \, d\theta$
Volume of Solid of Revolution	$\pi \int_a^b [f(x)]^2 \, dx$	$\pi \int_{T_1}^{T_2} y^2(t)x'(t) \, dt$	$\frac{2}{3}\pi \int_\alpha^\beta r^3(\theta) \sin\theta \, d\theta$
Surface Area of Solid of Revolution	$2\pi \int_a^b f(x)\sqrt{1 + [f'(x)]^2} \, dx$	$2\pi \int_{T_1}^{T_2} y(t)\sqrt{[x'(t)]^2 + [y'(t)]^2} \, dt$	$2\pi \int_\alpha^\beta r(\theta) \sin\theta \sqrt{r^2(\theta) + r'^2(\theta)} \, d\theta$

Chapter 7 Improper Integral

7.1 Infinite and Defective Integrals

7.2 Convergence Tests for Improper Integrals

Definition 7.1 (Absolute and Conditional Convergence)

Let $f(x) \in R[a, A] \subset [a, +\infty)$, and suppose $\int_a^{+\infty} |f(x)| \, \mathrm{d}x$ converges. Then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ is said to be **absolutely convergent** (or $f(x)$ is **absolutely integrable** on $[a, +\infty)$).
If $\int_a^{+\infty} f(x) \, \mathrm{d}x$ converges but is not absolutely convergent, then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ is said to be **conditionally convergent.**

♣

¶ **Infinite Integrals**

Theorem 7.1 (Cauchy Convergence Criterion for Infinite Integrals)

The necessary and sufficient condition for the convergence of the infinite integral $\int_a^{+\infty} f(x) \, \mathrm{d}x$ is:

$$\forall \varepsilon > 0, \exists A_0 > \max\{a, 0\}, \forall A', A'' > A_0 : \left| \int_a^{A'} f(x) \, \mathrm{d}x - \int_a^{A''} f(x) \, \mathrm{d}x \right| = \left| \int_{A'}^{A''} f(x) \, \mathrm{d}x \right| < \varepsilon.$$

♡

From this, we can conclude that if $\int_a^{+\infty} f(x) \, \mathrm{d}x$ is absolutely convergent, then it must be convergent.

Theorem 7.2 (Comparison Test for Infinite Integrals)

Comparison Test Let $f(x), g(x)$ be functions defined on $[a, +\infty)$, and suppose $f(x) \leq K g(x)$ (where K is a positive constant). Then:

i) If $\int_a^{+\infty} g(x) \, \mathrm{d}x$ converges, then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ also converges.
ii) If $\int_a^{+\infty} f(x) \, \mathrm{d}x$ diverges, then $\int_a^{+\infty} g(x) \, \mathrm{d}x$ also diverges.

Limit Form Let $f(x), g(x) > 0$ be functions defined on $[a, +\infty)$, and suppose:

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

Then:

i) If $0 < l < +\infty$, and $\int_a^{+\infty} g(x) \, \mathrm{d}x$ converges, then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ also converges.
ii) If $0 < l < +\infty$, and $\int_a^{+\infty} g(x) \, \mathrm{d}x$ diverges, then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ also diverges.
iii) If $l = +\infty$, $\int_a^{+\infty} g(x) \, \mathrm{d}x$ and $\int_a^{+\infty} f(x) \, \mathrm{d}x$ both converge or both diverge.

Comparison with p-Integrals Let $f(x) > 0$ be defined on $[a, +\infty)$, and suppose:

$$\frac{f(x)}{x^p} \leq \frac{K}{x^p}, \quad \text{where } p > 0.$$

Then:

i) If $p > 1$, then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ converges.
ii) If $p \leq 1$, then $\int_a^{+\infty} f(x) \, \mathrm{d}x$ diverges.

Limit Form Let $f(x) > 0$ be defined on $[a, +\infty)$, and suppose:

$$\lim_{x \to +\infty} x^p f(x) = l, \quad \text{where } l > 0.$$

Then:

i) If $p > 1$, then $\int_a^{+\infty} f(x)\,\mathrm{d}x$ converges.

ii) If $p \leq 1$, then $\int_a^{+\infty} f(x)\,\mathrm{d}x$ diverges.

Theorem 7.3 (Abel-Dirichlet Test)

The infinite integral $\int_a^{+\infty} f(x)g(x)\,\mathrm{d}x$ converges if either of the following two conditions is satisfied:

Abel $\int_a^{+\infty} f(x)\,\mathrm{d}x$ converges, and $g(x)$ is monotonic and bounded on $[a, +\infty)$.

Dirichlet $F(A) = \int_a^A f(x)\,\mathrm{d}x$ is bounded on $[a, +\infty)$, $g(x)$ is monotonic on $[a, +\infty)$, in the meanwhile $\lim_{x \to +\infty} g(x) = 0$.

¶ `Defective Integrals`

7.3 Special Integrals

¶ `Definite Integrals`

Dirichlet Integral

$$\int_0^{\pi} \frac{\sin\left(n + \frac{1}{2}\right)x}{\sin\frac{x}{2}}\,\mathrm{d}x = \pi, \quad n \in \mathbb{N},$$

where integrand $D_n(x)$ is called the Dirichlet kernel.

Fejèr Integral

$$\int_0^{\pi} \left(\frac{\sin\frac{nx}{2}}{\sin\frac{x}{2}}\right)^2 \mathrm{d}x = n\pi, \quad n \in \mathbb{N},$$

¶ `Improper Integrals`

Euler Integral

$$\int_0^{\frac{\pi}{2}} \ln\sin x\,\mathrm{d}x = -\frac{\pi}{2}\ln 2.$$

Froullani Integral

$$\int_0^{+\infty} \frac{f(ax) - f(bx)}{x}\,\mathrm{d}x = [f(0) - f(+\infty)]\ln\frac{b}{a}, \quad a, b > 0,$$

where $f(x)$ is continuous on $(0, +\infty)$, and both limits $f(0)$ and $f(+\infty)$ exist.

Dirichlet Integral

$$\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}.$$

Euler-Poisson Integral

$$\int_0^{+\infty} e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}.$$

Poisson Integral

$$\int_{-\pi}^{\pi} \frac{1 - r^2}{1 - 2r\cos x + r^2}\,\mathrm{d}x, \quad (0 < r < 1)$$

Special Integral

$$\int_0^{+\infty} \frac{1}{1 + x^a \sin^b x} \, \mathrm{d}x \quad (a > b, b > 0 \text{and even})$$

7.4 Common Questions

Chapter 8 Numerical Series

8.1 Convergence of Numerical Series

8.2 Positive Term Series and Its Convergence Tests

Definition 8.1 (Positive Term Series)

If all terms of the series $\sum_{n=1}^{\infty} x_n$ are non-negative real numbers, i.e., $x_n \geqslant 0\,(x_n > 0), n = 1, 2, \ldots$, then this series is called a **positive term series** (or strictly positive term series).

Note *The positive term series converges if and only if the partial sums of the sequence are bounded. If the partial sums are unbounded, the series must diverge to $+\infty$.*

¶ `Comparison Test`

Theorem 8.1 (Comparison Test)

Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be positive term series. If $\exists N \in \mathbb{N}$, s.t. $\forall n > N : a_n \leqslant b_n$, then:

1. If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ also converges.
2. If $\sum_{n=1}^{\infty} a_n$ diverges, then $\sum_{n=1}^{\infty} b_n$ also diverges.

Limit Form Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be positive term series, and suppose $\lim_{n\to\infty} \frac{a_n}{b_n}$ exists. Then:

1. If $0 < l < +\infty$, $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ have the same convergence or divergence behavior.
2. If $l = 0$, $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ also converges.
3. If $l = +\infty$, $\sum_{n=1}^{\infty} b_n$ diverges, then $\sum_{n=1}^{\infty} a_n$ also diverges.

Theorem 8.2

Cauchy Test Let $\sum_{n=1}^{\infty} a_n$ be a positive term series.

1. If $\exists q \in [0, 1)$, s.t. $\sqrt[n]{a_n} \leqslant q < 1 \quad (n \geqslant N, N \in \mathbb{N})$, then the series converges.
2. If $\sqrt[n]{a_n} \geqslant 1$ for infinitely many n, then the series diverges.

Limit Form Let $\sum_{n=1}^{\infty} a_n$ be a positive term series, and suppose $\overline{\lim}_{n\to+\infty} \sqrt[n]{a_n} = r$. Then:

1. If $0 \leqslant r < 1$, the series $\sum_{n=1}^{\infty} a_n$ converges.
2. If $r > 1$, the series $\sum_{n=1}^{\infty} a_n$ diverges.
3. If $r = 1$, the test fails.

D'Alembert Test Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series.

1. If $\exists q \in [0, 1)$, s.t. $\frac{a_{n+1}}{a_n} \leqslant q < 1 \quad (n \geqslant N, N \in \mathbb{N})$, then the series converges.
2. If $\frac{a_{n+1}}{a_n} \geqslant 1 \quad (n \geqslant N, N \in \mathbb{N})$, then the series diverges.

Limit Form Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series. Then:

1. If $\overline{\lim}_{n\to+\infty} \frac{a_{n+1}}{a_n} = r \in (0, 1)$, the series converges.
2. If $\underline{\lim}_{n\to+\infty} \frac{a_{n+1}}{a_n} = r' > 1$, the series diverges.
3. If $r = 1$ or $r' = 1$, the test fails.

Raabe Test Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series.

1. If $\exists r > 1, \exists N_0 \in \mathbb{N}$ s.t. $\forall n > N_0 : n\left(\frac{a_n}{a_{n+1}} - 1\right) \geqslant r$, then the series converges.
2. If $\exists N_0 \in \mathbb{N}$, s.t. $\forall n > N_0 : n\left(\frac{a_n}{a_{n+1}} - 1\right) \leqslant 1$, then the series diverges.

Limit Form Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series. Then:

1. If $\underline{\lim}_{n\to+\infty} n\left(\frac{a_n}{a_{n+1}}-1\right)=l>1$, the series converges.
2. If $\overline{\lim}_{n\to+\infty} n\left(\frac{a_n}{a_{n+1}}-1\right)=l'<1$, the series diverges.
3. If $l=1$ or $l'=1$, the test fails.

Bertrand Test Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series.

1. If $\underline{\lim}_{n\to+\infty} \ln n\left[n\left(\frac{a_n}{a_{n+1}}-1\right)\right]=l>1$, the series converges.
2. If $\overline{\lim}_{n\to+\infty} \ln n\left[n\left(\frac{a_n}{a_{n+1}}-1\right)\right]=l'<1$, the series diverges.
3. If $l=1$ or $l'=1$, the test fails.

Gauss Test Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series, and suppose:

$$\frac{a_n}{a_{n+1}}=1+\frac{1}{n}+\frac{\delta}{n\ln n}+o\left(\frac{1}{n\ln n}\right),\quad (n\to+\infty).$$

Then:

1. If $\delta>1$, the series converges.
2. If $\delta<1$, the series diverges.
3. If $\delta=1$, the criterion fails.

Generalized Form Let $\sum_{n=1}^{\infty} a_n$ be a strictly positive term series, and suppose:

$$\frac{a_n}{a_{n+1}}=1+\frac{1}{n}+\frac{\delta_n}{n\ln n}+o\left(\frac{1}{n\ln n}\right),\quad (n\to+\infty).$$

If $\lim_{n\to\infty}\delta_n=\delta\in\mathbb{R}$, then:

1. If $\delta>1$, the series converges.
2. If $\delta<1$, the series diverges.
3. If $\delta=1$, the criterion fails.

Note *The Bertrand test can be refined by considering series such as:*

$$\sum_{n=3}^{\infty}\frac{1}{n\ln n(\ln\ln n)^p},\quad \sum_{n=9}^{\infty}\frac{1}{n\ln n\ln\ln n(\ln\ln n)^p},\cdots$$

These refinements are collectively known as the Bertrand test.

Remark All the aforementioned criteria are derived from the Comparison Criterion.

- By comparing positive term series with the geometric series (or equal ratio series), the Cauchy Criterion and d'Alembert Criterion are derived.
- By comparing positive term series with the slower-converging series $\sum_{n=1}^{\infty}\frac{1}{n^\alpha}$ $(\alpha>1)$, the Raabe Criterion is derived.
- By comparing positive term series with the even slower-converging series $\sum_{n=1}^{\infty}\frac{1}{n\ln^\alpha n}$ $(\alpha>1)$, the Gauss Criterion is derived.

General Observation The slower the convergence of the series used for comparison, the more precise the derived criterion.

¶ `Integral Test`

Theorem 8.3 (Cauchy Integral Test)

Let $f(x)$ be defined on $[a, +\infty)$, where $f(x) \geqslant 0$, and $f(x)$ is Riemann integrable on any finite interval $[a, A]$. Consider a monotonic increasing sequence $\{a_n\}$ such that $a = a_1 < a_2 < \cdots < a_n < \ldots$, and let:

$$u_n = \int_{a_n}^{a_{n+1}} f(x)\,\mathrm{d}x.$$

Then the improper integral $\int_a^{+\infty} f(x)\,\mathrm{d}x$ and the positive term series $\sum_{n=1}^{\infty} u_n$ converge or diverge to $+\infty$ simultaneously. Moreover:

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} \int_{a_n}^{a_{n+1}} f(x)\,\mathrm{d}x.$$

¶ `Other Tests`

Theorem 8.4 (Cauchy Condensation Test)

Let $\{a_n\}$ be a monotonically decreasing sequence of positive numbers. Then the positive term series $\sum_{n=1}^{\infty} a_n$ converges if and only if the condensed series:

$$\sum_{n=0}^{\infty} 2^n a_{2^n} = a_1 + 2a_2 + 4a_4 + \cdots + 2^n a_{2^n} + \ldots$$

converges.

8.3 General Term Series and Its Convergence Tests

¶ `Cauchy Convergence Criterion for Series`

Theorem 8.5 (Cauchy Convergence Criterion for Series)

The necessary and sufficient condition for the convergence of the series $\sum_{n=1}^{\infty} x_n$ is:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n > N : |x_{n+1} + x_{n+2} + \cdots + x_m| = \left| \sum_{k=n+1}^{m} x_k \right| < \varepsilon.$$

¶ `Alternative Series`

Definition 8.2 (Alternative Series)

A series of the form:

$$\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} (-1)^{n-1} u_n \quad (u_n > 0),$$

is called an **alternative series.**

Moreover, if u_n is a monotonically decreasing sequence and $\lim_{n\to\infty} u_n = 0$, then the series is called a **Leibniz series.**

Theorem 8.6 (Leibniz Test)

Leibniz series converges.

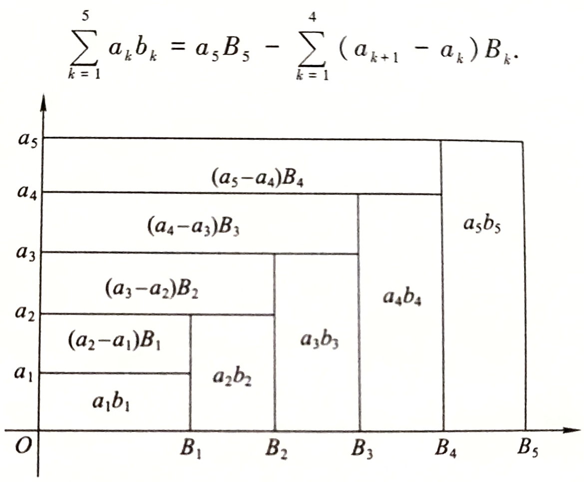
¶ Abel-Dirichlet Test

Theorem 8.7 (Abel Transform (Discrete Integration by Parts/Summation by Parts))

Let $\{a_n\}$, $\{b_n\}$ be two sequences, then for any $n \in \mathbb{N}^+$,

$$\sum_{k=1}^{n} a_k b_k = a_n B_n + \sum_{k=1}^{n-1} (a_{k+1} - a_k) B_k,$$

where $B_n = \sum_{k=1}^{n} b_k$.



Lemma 8.1 (Abel Lemma (Discrete Second Integral Mean Value Theorem))

Let $\{a_n\}$, $\{b_n\}$ be two sequences, if $\{a_n\}$ is a monotonic sequence and $\{B_k\} = \sum_{k=1}^{n} b_k$ is a bounded sequence with bound M, then for any $p \in \mathbb{N}^+$,

$$\left| \sum_{k=1}^{p} a_k b_k \right| \leq M \left(|a_1| + 2|a_p| \right).$$

Theorem 8.8 (Abel-Dirichlet Test)

The series $\sum_{n=1}^{\infty} a_n b_n$ converges if one of the following two conditions is satisfied:

Abel $\{a_n\}$ is a bounded monotonic sequence and $\sum_{n=1}^{\infty} b_n$ converges.

Dirichlet $\{a_n\}$ is a monotonic sequence, $\lim_{n\to\infty} a_n = 0$, and the partial sums $B_n = \sum_{k=1}^{n} b_k$ are bounded.

8.4 Absolute and Conditional Convergence of Series

Definition 8.3 (Absolute and Conditional Convergence of Series)

If the series $\sum_{n=1}^{\infty} |x_n|$ converges, then the series $\sum_{n=1}^{\infty} x_n$ is said to be **absolutely convergent**.
If the series $\sum_{n=1}^{\infty} x_n$ converges but is not absolutely convergent, then the series $\sum_{n=1}^{\infty} x_n$ is said to be **conditionally convergent**.

♣

8.5 Comparison of Convergence Speed of Series

The series $\sum_{n=1}^{\infty} a_n$ is said to converge faster than the series $\sum_{n=1}^{\infty} b_n$ if:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

Theorem 8.9 (Du Bois-Reymond Theorem)

For a given convergent positive term series $\sum_{n=1}^{\infty} a_n$, there always exists a convergent strictly positive term series $\sum_{n=1}^{\infty} b_n$ such that:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

♡

Theorem 8.10 (Abel Theorem)

For a given divergent positive term series $\sum_{n=1}^{\infty} a_n$, there always exists a divergent positive term series $\sum_{n=1}^{\infty} b_n$ such that:

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 0.$$

♡

Remark The above two theorems imply that the slowest converging positive term series does not exist.

8.6 Infinite Products

8.7 Special Series

Chapter 9 Series of Functions

9.1 Pointwise and Uniform Convergence

¶ Pointwise Convergence

Definition 9.1 (Function Term Series)

Let $u_n(x)$ $(n = 1, 2, 3, \ldots)$ be a sequence of functions with a common domain E. The sum of these infinitely many functions $u_1(x) + u_2(x) + \cdots + u_n(x) + \ldots$ is called a **function term series**, denoted as:

$$\sum_{n=1}^{\infty} u_n(x).$$

For any fixed point $x_0 \in E$, if the numerical series $\sum_{n=1}^{\infty} u_n(x_0)$ converges, then the function term series $\sum_{n=1}^{\infty} u_n(x)$ is said to converge at x_0, or equivalently, x_0 is called a **convergence point** of $\sum_{n=1}^{\infty} u_n(x)$. The set of all convergence points is called the **domain of convergence** of $\sum_{n=1}^{\infty} u_n(x)$. ♣

Definition 9.2 (Pointwise Convergence)

Let the domain of convergence of the function term series $\sum_{n=1}^{\infty} u_n(x)$ be $D \subset E$. Then $\sum_{n=1}^{\infty} u_n(x)$ defines a function $S(x)$ on the set D, where:

$$S(x) = \sum_{n=1}^{\infty} u_n(x), \quad x \in D.$$

The function $S(x)$ is called the **sum function** of the series, and $\sum_{n=1}^{\infty} u_n(x)$ is said to **converge pointwise** to $S(x)$ on D. ♣

Define the **partial sum function** of the series as:

$$S_n(x) = \sum_{k=1}^{n} u_k(x).$$

It is evident that the set of all x for which $\{S_n(x)\}$ converges is precisely D. Therefore, on D, we have:

$$S(x) = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \sum_{k=1}^{n} u_k(x).$$

Conversely, if a sequence of functions $\{S_n(x)\}$ $(x \in E)$ is given, we can define:

$$\begin{cases} u_1(x) = S_1(x), \\ u_{n+1}(x) = S_{n+1}(x) - S_n(x), \quad n = 1, 2, \ldots \end{cases}$$

to obtain the corresponding function term series.

Thus, the convergence behavior of a function term series and the corresponding sequence of partial sum functions is essentially the same.

However, it is important to note that the pointwise convergence has certain limitations.

Continuity The sum of finitely many continuous functions satisfies additive continuity:

$$\lim_{x\to x_0}[u_1(x)+u_2(x)+\cdots+u_n(x)]=\lim_{x\to x_0}u_1(x)+\lim_{x\to x_0}u_2(x)+\cdots+\lim_{x\to x_0}u_n(x).$$

If this property can be extended to infinitely many functions, that is: If $u_n(x)$ is continuous on D, the sum function $S(x)=\sum_{n=1}^{\infty}u_n(x)$ is also continuous on D. Moreover:

$$\lim_{x\to x_0}\sum_{n=1}^{\infty}u_n(x)=\sum_{n=1}^{\infty}\lim_{x\to x_0}u_n(x),$$

meaning that the limit operation and infinite summation can be interchanged (also known as the fact that function term series can be evaluated termwise).

For the sequence of partial sums $\{S_n(x)\}$, the corresponding conclusion is that the limit function $S(x)=\lim_{n\to\infty}S_n(x)$ is also continuous on D, and:

$$\lim_{x\to x_0}\lim_{n\to\infty}S_n(x)=\lim_{n\to\infty}\lim_{x\to x_0}S_n(x),$$

meaning that the two limit operations can be interchanged.

Unfortunately, in the case of pointwise convergence, this property does not hold.

Derivability The sum of finitely many differentiable functions satisfies additive differentiability:

$$\frac{\mathrm{d}}{\mathrm{d}x}[u_1(x)+u_2(x)+\cdots+u_n(x)]=\frac{\mathrm{d}}{\mathrm{d}x}u_1(x)+\frac{\mathrm{d}}{\mathrm{d}x}u_2(x)+\cdots+\frac{\mathrm{d}}{\mathrm{d}x}u_n(x).$$

If this property can be extended to infinitely many functions, that is: If $u_n(x)$ is differentiable on D, the sum function $S(x)=\sum_{n=1}^{\infty}u_n(x)$ is also differentiable on D. Moreover:

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=1}^{\infty}u_n(x)=\sum_{n=1}^{\infty}\frac{\mathrm{d}}{\mathrm{d}x}u_n(x),$$

meaning that the differentiation operation and infinite summation can be interchanged (also known as the fact that function term series can be differentiated termwise).

For the sequence of partial sums $\{S_n(x)\}$, the corresponding conclusion is that the limit function $S(x)=\lim_{n\to\infty}S_n(x)$ is also differentiable on D, and:

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

meaning that the two operations can be interchanged.

Unfortunately, in the case of pointwise convergence, this property does not hold.

Integrability The sum of finitely many integrable functions satisfies additive integrability:

$$\int_a^b[u_1(x)+u_2(x)+\cdots+u_n(x)]\,\mathrm{d}x=\int_a^b u_1(x)\,\mathrm{d}x+\int_a^b u_2(x)\,\mathrm{d}x+\cdots+\int_a^b u_n(x)\,\mathrm{d}x.$$

If this property can be extended to infinitely many functions, that is: If $u_n(x)$ is integrable on $[a,b]\subset D$, the sum function $S(x) = \sum_{n=1}^{\infty} u_n(x)$ is also integrable on $[a, b] \subset D$. Moreover:

$$\int_a^b \sum_{n=1}^{\infty} u_n(x) \, \mathrm{d}x = \sum_{n=1}^{\infty} \int_a^b u_n(x) \, \mathrm{d}x,$$

meaning that the integration operation and infinite summation can be interchanged (also known as the fact that function term series can be integrated termwise).

For the sequence of partial sums $\{S_n(x)\}$, the corresponding conclusion is that the limit function $S(x) = \lim_{n\to\infty} S_n(x)$ is also integrable on $[a, b] \subset D$, and:

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

meaning that the two operations can be interchanged.

Unfortunately, in the case of pointwise convergence, this property does not hold.

¶ Uniform Convergence

Definition 9.3 (Uniform Convergence)

Let $\{S_n(x)\}(x \in D)$ be a sequence of functions. If:

$$\forall \varepsilon > 0, \exists N(\varepsilon) \in \mathbb{N}^+, \forall n > N(\varepsilon) : |S_n(x) - S(x)| < \varepsilon \quad (\forall x \in D),$$

then $\{S_n\}$ is said to **converge uniformly** to $S(x)$ on D, denoted as:

$$S_n(x) \overset{D}{\rightrightarrows} S(x).$$

If the partial sum sequence $\{S_n(x)\}$ of the function term series $\sum_{n=1}^{\infty} u_n(x)(x \in D)$ converges uniformly to $S(x)$ on D, then $\sum_{n=1}^{\infty} u_n(x)$ is said to converge uniformly to $S(x)$ on D. ♣

Obviously, if the partial sum sequence $\{S_n(x)\}$ of $\sum_{n=1}^{\infty} u_n(x)$ satisfies:

$$S_n(x) \overset{D}{\rightrightarrows} S(x),$$

then:

$$u_n(x) \overset{D}{\rightrightarrows} 0.$$

Theorem 9.1 (Cauchy Criterion for Uniform Convergence)

The necessary and sufficient condition for the sequence of functions $\{S_n(x)\}$ to converge uniformly on D is:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}^*, \forall m > n > N : |S_m(x) - S_n(x)| < \varepsilon \quad (\forall x \in D).$$

Correspondingly, the necessary and sufficient condition for the function term series $\sum_{n=1}^{\infty} u_n(x)$ to converge uniformly on D is:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}^*, \forall m > n > N : \left| \sum_{i=n+1}^{m} u_i(x) \right| < \varepsilon \quad (\forall x \in D).$$

♡

Theorem 9.2 (Necessary and Sufficient Condition for Uniform Convergence)

Let $\{S_n(x)\}$ converge pointwise to $S(x)$ on D. The necessary and sufficient conditions for $S_n(x) \overset{D}{\rightrightarrows} S(x)$ are:

1.
$$\lim_{n\to\infty} d(S_n, S) = \lim_{n\to\infty} \sup_{x\in D} |S_n(x) - S(x)| = 0.$$

2. For any sequence $\{x_n\}$ where $x_n \in D$, the following holds:
$$\lim_{n\to\infty} \big(S_n(x_n) - S(x_n)\big) = 0.$$

With the concept of uniform convergence, the flaws of pointwise convergence can be remedied, and the following properties can be established:

Property

Continuity *Let $f_n(x) \overset{I\subset\mathbb{R}}{\rightrightarrows} f(x)$. If $f_n(x)$ is continuous at $x_0 \in I$ for $n = 1, 2, 3, \ldots$, then $f(x)$ is also continuous at x_0.*

In particular, if $f_n(x) \in C(I)$, then $f(x) \in C(I)$.

Termwise Limit *If $\sum_{n=1}^{\infty} u_n(x) \overset{I\subset\mathbb{R}}{\rightrightarrows} S(x)$ and $u_n(x) \in C(I)$, then the sum function $S(x) \in C(I)$.*

Integrability *Let $f_n(x) \overset{[a,b]}{\rightrightarrows} f(x)$. If $f_n(x) \in R[a, b]$, then $f(x) \in R[a, b]$, and:*

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

Termwise Integration: *If $\sum_{n=1}^{\infty} u_n(x) \overset{[a,b]}{\rightrightarrows} S(x)$ and $u_n(x) \in R[a, b]$, then $S(x) \in R[a, b]$.*

Differentiability *Let $f_n'(x) \overset{[a,b]}{\rightrightarrows} \sigma(x)$. If there exists $x_0 \in [a, b]$ such that:*

$$\lim_{n\to\infty} f_n(x_0) = a,$$

then there exists a function $f(x)$ such that $f_n(x) \overset{[a,b]}{\rightrightarrows} f(x)$ and $f'(x) = \sigma(x)$.

Termwise Differentiation *If $\sum_{n=1}^{\infty} u_n'(x) \overset{[a,b]}{\rightrightarrows} \sigma(x)$ and there exists $x_0 \in [a, b]$ such that:*

$$\sum_{n=1}^{\infty} u_n(x_0) \to a,$$

then there exists a function $S(x)$ such that $\sum_{n=1}^{\infty} u_n(x) \overset{[a,b]}{\rightrightarrows} S(x)$ and $S'(x) = \sigma(x)$.

Corollary *Obviously, if we add the condition $f_n'(x) \in C[a, b]$, the conclusion still holds, and the proof becomes simpler.*

Note *Since continuity and differentiability are both local properties, it suffices to have **internally closed uniform convergence** of (a, b) to ensure that $f(x)$ is continuous/differentiable.*

¶ `Quasi-Uniform Convergence`

Definition 9.4 (Quasi-Uniform Convergence)

The sequence of functions $\{S_n(x)\}$ is said to **converge quasi-uniformly** on the interval $[a, b]$ if it converges pointwise to $S(x)$ on $[a, b]$, and the following condition is satisfied:

$$\forall \varepsilon > 0, \forall N \in \mathbb{N}^*, \exists N_0 > N, \text{ s.t. } \forall x \in [a, b], \exists n_x \in [N, N_0] \ (n_x \in \mathbb{N}^*) : |S_{n_x}(x) - S(x)| < \varepsilon.$$

♣

9.2 Uniform Convergence Tests

¶ `Weierstrass Test (M-Test)`

¶ `Abel-Dirichlet Test`

¶ `Dini Theorem`

9.3 Special Cases

Chapter 10 Power Series

Chapter 11 Limits and Continuity in Euclidean Spaces

11.1 Continuous Mappings

¶ **Continuous Mappings on Compact Sets**

¶ **Continuous Mappings on Connected Sets**

Definition 11.1 (Connected Set)

Let S be a set of points in $\mathbb{R}^n$. If a continuous mapping

$$\gamma : [0, 1] \to \mathbb{R}^n$$

satisfies that the range of $\gamma([0, 1])$ lies entirely within S, we call γ a **path** in S, where $\gamma(0)$ and $\gamma(1)$ are referred to as the starting point and ending point of the path, respectively.

If for any two points $\mathbf{x}, \mathbf{y} \in S$, there exists a path in S with $\mathbf{x}$ as the starting point and $\mathbf{y}$ as the ending point, S is called path-connected, or equivalently, S is called a **connected set**.

A connected open set is called an **(open) region**. The closure of an (open) region is referred to as a closed region. ♣

Remark Intuitively, this means that any two points in S can be connected by a curve lying entirely within S. Clearly, a connected subset of $\mathbb{R}$ is an interval, and a connected subset of $\mathbb{R}$ is compact if and only if it is a closed interval.

Chapter 12 Multi-variable Differential Calculus

12.1 Directional Derivatives and Total Differential

¶ Directional Derivative

Definition 12.1 (Directional Derivative)

Let $U \subset \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}^1$, $\mathbf{e}$ is a unit vector in $\mathbb{R}^n$, $\mathbf{x}^0 \in U$. Define

$$u(t) = f(\mathbf{x}^0 + t\mathbf{e}).$$

If the derivative of u at $t = 0$

$$u'(0) = \lim_{t \to 0} \frac{u(t) - u(0)}{t} = \lim_{t \to 0} \frac{f(\mathbf{x}^0 + t\mathbf{e}) - f(\mathbf{x}^0)}{t}$$

exists and is finite, it is called the **directional derivative** of f at $\mathbf{x}^0$ in the direction $\mathbf{e}$, denoted by $\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}^0)$. It is the rate of change of f at $\mathbf{x}^0$ in the direction $\mathbf{e}$. ♣

Consider the following set of unit coordinate vectors: $\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n$. Let $\mathbf{e}_i = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$ denote the standard orthonormal basis in $\mathbb{R}^n$, where the 1 appears in the i-th position. That is,

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

For a function f, the directional derivative of f at the point $\mathbf{x}_0$ in the direction of $\mathbf{e}_i$ is called the ith first-order **partial derivative** of f at $\mathbf{x}^0$, denoted by

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^0) \quad \text{or} \quad \mathrm{D}_i f(\mathbf{x}^0) \quad \text{or} \quad f_{x_i}(\mathbf{x}^0) \quad (i = 1, 2, \cdots, n).$$

$\mathrm{D}_i = \frac{\partial}{\partial x_i}$ is called the ith partial differential operator ($i = 1, 2, \cdots, n$).

Let $\mathbf{e}_i = \sum_{i=0}^{n} \mathbf{e}_i \cos\alpha$ be a unit vector, where $\sum_{i=0}^{n} \cos^2\alpha = 1$. If $\frac{\partial f}{\partial x_i}$ is continuous at $\mathbf{x}^0$, then the directional derivative of f at $\mathbf{x}^0$ along the direction $\mathbf{e}$ is given by:

$$\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}^0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}^0) \cos\alpha_i.$$

This is the formula for **expressing a directional derivative using partial derivatives**.

Note *Let* $\mathbf{e}$ *be a direction, then* $\|-\mathbf{e}\| = \|\mathbf{e}\| = 1$, *which implies that* $-\mathbf{e}$ *is also a direction. At this point, we have:*

$$\frac{\partial f}{\partial (-\mathbf{e})}(\mathbf{x}^0) = -\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}^0).$$

Definition 12.2 (Jacobian Matrix (Gradient))

Let

$$Jf(\mathbf{x}) = (\mathrm{D}_1 f(\mathbf{x}), \mathrm{D}_2 f(\mathbf{x}), \ldots, \mathrm{D}_n f(\mathbf{x})),$$

which is called the **Jacobian matrix** of the function f at the point $\mathbf{x}$, (a $1 \times n$ matrix) whose counterpart is the first-order derivative of a single-variable function.

Henceforth, we represent the point $\mathbf{x}$ in $\mathbb{R}^n$ and its increments $\mathbf{h}$ as column vectors. In this way, the differential of the function can be expressed using matrix multiplication as follows:

$$\mathrm{d}f(\mathbf{x}^0)(\mathbf{\Delta x}) = Jf(\mathbf{x}^0)\mathbf{\Delta x}.$$

The Jacobian matrix of the function f is also frequently denoted as grad f (or ∇f), that is,

$$\operatorname{grad} f(\mathbf{x}) = Jf(\mathbf{x}),$$

which is called the **gradient** of the scalar function f. ♣

Note *Let $U \subset \mathbb{R}^n$ be an open set, and $\mathbf{f} : U \to \mathbb{R}^m$ be a C^k mapping:*

- $k = 0$, *$\mathbf{f}$ is a continuous mapping;*
- $0 < k < +\infty$, *f_i has continuous partial derivatives up to order k, $i = 1, 2, \ldots, m$;*
- $k = +\infty$, *f_i has continuous partial derivatives of all orders, $i = 1, 2, \ldots, m$;*
- $k = \omega$, *f_i is really analytic, i.e., in the neighborhood of any point $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in U$, f_i can be expanded into a convergent (n-dimensional) power series, $i = 1, 2, \ldots, m$.*

Let $C^k(U, \mathbb{R}^m)$ denote the totality of C^k mappings from U to $\mathbb{R}^m$.

¶ Total Differential

Definition 12.3 (Total Differential)

Let $U \subset \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}^1, \mathbf{x}^0 \in U, \mathbf{\Delta x} = (\Delta x_1, \Delta x_2, \cdots, \Delta x_n) \in \mathbb{R}^n$. If

$$f(\mathbf{x}^0 + \mathbf{\Delta x}) - f(\mathbf{x}^0) = \sum_{i=1}^{n} A_i \Delta x_i + o(\|\mathbf{\Delta x}\|) \qquad (\|\mathbf{\Delta x}\| \to 0),$$

where $A_1, A_2, \ldots, A_n$ are constants independent of $\mathbf{\Delta x}$, then the function f is said to be **differentiable** at the point $\mathbf{x}^0$, and the linear main part $\sum_{i=1}^{n} A_i \Delta x_i$ is called the **total differential** of f at $\mathbf{x}^0$, denoted as

$$df(\mathbf{x}^0)(\mathbf{\Delta x}) = \sum_{i=1}^{n} A_i \Delta x_i.$$

If f is differentiable at every point in the open set U, then f is called a differentiable function on U. ♣

Theorem 12.1 (Conditions of Differentiability)

Necessary Condition If an n-variable function f is differentiable at the point $\mathbf{x}_0$, then f is continuous at $\mathbf{x}^0$ and possesses first-order partial derivatives $\frac{\partial f}{\partial x_i}(\mathbf{x}^0)$ at $\mathbf{x}^0$ for $i = 1, 2, \ldots, n$, and

$$\mathbf{A} = (A_1, A_2, \ldots, A_n) = Jf(\mathbf{x}^0) = \left(\mathrm{D}_1 f(\mathbf{x}^0), \mathrm{D}_2 f(\mathbf{x}^0), \ldots, \mathrm{D}_n f(\mathbf{x}^0)\right).$$

[a] However, the converse is not true.

Sufficient Condition Let $U \subset \mathbb{R}^n$ be an open set, and let $f : U \to \mathbb{R}^1$ be an n-variable function. If $Jf = (\mathrm{D}_1 f, \mathrm{D}_2 f, \ldots, \mathrm{D}_n f)$ is continuous at $\mathbf{x}^0$ (i.e., $\frac{\partial f}{\partial x_i}$ is continuous at $\mathbf{x}^0$ for $i = 1, 2, \ldots, n$), then f is differentiable at $\mathbf{x}^0$. However, the converse is not necessarily true.

[a] It is referred to as the total differential formula, and the more common form is

$$\mathrm{d}f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\,\mathrm{d}x + \frac{\partial f}{\partial y}(x_0, y_0)\,\mathrm{d}y.$$

Note

- *The continuity of the derivative function at $\mathbf{x}^0$ implies that the original function f is differentiable in some neighborhood of $\mathbf{x}^0$.*
- *In fact, this condition can be relaxed to require that one partial derivative exists at the point, while the remaining $n-1$ partial derivative functions are continuous at that point.*

Proof Taking a function of three variables as an example.

Assume the 3-ary function $f : \mathbb{R}^3 \to \mathbb{R}$ meets:

1. There exists $f_z(x_0, y_0, z_0)$.
2. The partial derivative functions $f_x(x, y, z)$ and $f_y(x, y, z)$ are continuous at (x_0, y_0, z_0), i.e. there are partial derivatives in some neighborhood of (x_0, y_0, z_0).

Consider the total increment of f at the point (x_0, y_0, z_0):

$$\begin{aligned}\Delta f = &\underbrace{[f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) - f(x_0, y_0 + \Delta y, z_0 + \Delta z)]}_{I_1} \\ &+ \underbrace{[f(x_0, y_0 + \Delta y, z_0 + \Delta z) - f(x_0, y_0, z_0 + \Delta z)]}_{I_2} \\ &+ \underbrace{[f(x_0, y_0, z_0 + \Delta z) - f(x_0, y_0, z_0)]}_{I_3}.\end{aligned}$$

For I_1, I_2, by the Lagrange's Mean Value Theorem of unary functions, there exist $\theta_1, \theta_2 \in (0, 1)$ such that

$$\begin{aligned}I_1 &= f_x(x_0 + \theta_1 \Delta x, y_0 + \Delta y, z_0 + \Delta z)\Delta x, \\ I_2 &= f_y(x_0, y_0 + \theta_2 \Delta y, z_0 + \Delta z)\Delta y.\end{aligned}$$

Then by the continuity of the their partial derivatives at (x_0, y_0, z_0), we have

$$\lim_{\Delta x, \Delta y, \Delta z \to 0} I_1 = f_x(x_0, y_0, z_0)\Delta x, \qquad \lim_{\Delta x, \Delta y, \Delta z \to 0} I_2 = f_y(x_0, y_0, z_0)\Delta y.$$

They can be expressed in terms of infinitesimals($\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$):

$$\begin{aligned}I_1 &= f_x(x_0, y_0, z_0)\Delta x + \alpha_1 \Delta x, \quad \alpha_1 \to 0 (\rho \to 0), \\ I_2 &= f_y(x_0, y_0, z_0)\Delta y + \alpha_2 \Delta y, \quad \alpha_2 \to 0 (\rho \to 0).\end{aligned}$$

For I_3, by the definition of the partial derivative $f_z(x, y, z)$ at (x_0, y_0, z_0), we have

$$I_3 = f_z(x_0, y_0, z_0)\Delta z + \alpha_3 \Delta z, \quad \alpha_3 \to 0(\rho \to 0).$$

Accordingly,

$$\begin{aligned}\Delta f &= I_1 + I_2 + I_3 \\ &= [f_x(x_0, y_0, z_0)\Delta x + \alpha_1 \Delta x] + [f_y(x_0, y_0, z_0)\Delta y + \alpha_2 \Delta y] + [f_z(x_0, y_0, z_0)\Delta z + \alpha_3 \Delta z] \\ &= f_x(x_0, y_0, z_0)\Delta x + f_y(x_0, y_0, z_0)\Delta y + f_z(x_0, y_0, z_0)\Delta z + [\alpha_1 \Delta x + \alpha_2 \Delta y + \alpha_3 \Delta z]\,.\end{aligned}$$

Apparently,

$$\lim_{\rho \to 0} \frac{\alpha_1 \Delta x + \alpha_2 \Delta y + \alpha_3 \Delta z}{\rho} = 0,$$

i.e. $\alpha_1 \Delta x + \alpha_2 \Delta y + \alpha_3 \Delta z = o(\rho)$. Therefore, $f(x, y, z)$ is differentiable at (x_0, y_0, z_0), which completes the proof. ■

Note *(At some point)*

1. *Differentiable*
   - $\Longrightarrow$ *Continuous*
   - $\Longrightarrow$ *Partial derivatives exist:* $\mathrm{D}_{\vec{u}} = \nabla f \cdot \vec{u}$
2. *Directional Derivative*
   - *All directional derivatives exist* $\not\Longrightarrow$ *differentiable or continuous.*
   - *All directional derivatives exist and are equal* $\not\Longrightarrow$ *differentiable.*
3. *Partial Derivative*
   - *The continuity and existence of directional/partial derivatives are mutually exclusive.*

¶ Higher-Order Partial Derivatives and Differential

If the first-order partial derivative of f, $\frac{\partial f}{\partial x_i}$, itself possesses partial derivatives, then the second-order partial derivative of f is defined, and is denoted as follows(the first is also called the mixed partial derivative):

$$f_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right), \quad f_{x_i x_i} = \frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right), \quad i, j = 1, 2, \ldots, n.$$

Similarly, higher-order partial derivatives of order $3, 4, \cdots m, \cdots$ can be defined.

The following theorem provides the conditions under which mixed partial derivatives are equal.

Theorem 12.2 (Conditions for Equality of Mixed Partial Derivatives)

1. Let $U \subset \mathbb{R}^2$ be an open set, and $f : U \to \mathbb{R}$ be a function of two variables. If f_{xy} and f_{yx} are continuous at $(x_0, y_0) \in U$, then
$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$
2. Let $U \subset \mathbb{R}^n$ be an open set, and $f : U \to \mathbb{R}$ be a function of n variables. If f has partial derivatives up to order k in D, and all of them are continuous at $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in U$, then
$$\frac{\partial^l f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_l}}(\mathbf{x}^0) = \frac{\partial^l f}{\partial x_{i_2} \partial x_{i_1} \cdots \partial x_{i_l}}(\mathbf{x}^0) = \cdots = \frac{\partial^l f}{\partial x_{i_l} \partial x_{i_{l-1}} \cdots \partial x_{i_1}}(\mathbf{x}^0),$$

that is, the order of taking partial derivatives $l(\leq k)$ does not affect the result.[a]

[a] If the condition "f_{xy} and f_{yx} are continuous at (x_0, y_0)", is not satisfied, then the conclusion "$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$" does not necessarily hold.

Proof When $k \neq 0, h \neq 0$, define

$$\varphi(y) = f(x_0 + h, y) - f(x_0, y),$$

and

$$\psi(x) = f(x, y_0 + k) - f(x, y_0).$$

Applying the Lagrange Mean Value Theorem, we have

$$\begin{aligned}
&[f(x_0+h, y_0+k) - f(x_0, y_0+k)] - [f(x_0+h, y_0) - f(x_0, y_0)] \\
=&\varphi(y_0+k) - \varphi(y_0) \\
=&\varphi'(y_0+\theta_1 k)k \\
=&[f_y(x_0+h, y_0+\theta_1 k) - f_y(x_0, y_0+\theta_1 k)]k \\
=&f_{yx}(x_0+\theta_2 h, y_0+\theta_1 k)hk, \quad 0<\theta_1,\theta_2<1.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&[f(x_0+h, y_0+k) - f(x_0, y_0+k)] - [f(x_0+h, y_0) - f(x_0, y_0)] \\
=&[f(x_0+h, y_0+k) - f(x_0+h, y_0)] - [f(x_0, y_0+k) - f(x_0, y_0)] \\
=&\psi(x_0+h) - \psi(x_0) \\
=&\psi'(x_0+\theta_3 h)h \\
=&[f_x(x_0+\theta_3 h, y_0+k) - f_x(x_0+\theta_3 h, y_0)]h \\
=&f_{xy}(x_0+\theta_3 h, y_0+\theta_4 k)hk, \quad 0<\theta_3,\theta_4<1.
\end{aligned}$$

Therefore,

$$f_{xy}(x_0+\theta_3 h, y_0+\theta_4 k) = f_{yx}(x_0+\theta_2 h, y_0+\theta_1 k).$$

Since f_{xy} and f_{yx} are continuous at (x_0, y_0), letting $h \to 0, k \to 0$, we obtain

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

By applying 12.1 and the principle of mathematical induction, one can immediately derive the following result.
■

Suppose $z = f(x, y)$ has continuous partial derivatives in the domain $U \subset \mathbb{R}^2$. Then z is differentiable, and

$$\mathrm{d}z = \frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y.$$

If z also has continuous second-order partial derivatives, then $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are also differentiable, and thus $\mathrm{d}z$

is differentiable. We call the differential of $\mathrm{d}z$ the second-order differential of z, denoted as

$$\mathrm{d}^2 z = \mathrm{d}(\mathrm{d}z).$$

In general, based on the k-th order differential $\mathrm{d}^k z$ of z, its $(k+1)$-th order differential (if it exists) is defined as

$$\mathrm{d}^{k+1} z = \mathrm{d}(\mathrm{d}^k z), \quad k = 1, 2, \cdots.$$

Due to the fact that for the independent variables x and y, we always have

$$\mathrm{d}^2 x = \mathrm{d}(\mathrm{d}x) = 0, \qquad \mathrm{d}^2 y = \mathrm{d}(\mathrm{d}y) = 0,$$

the second-order differential of $z = f(x, y)$ is given by

$$\begin{aligned} \mathrm{d}^2 z = \mathrm{d}(\mathrm{d}z) &= \mathrm{d}\left(\frac{\partial z}{\partial x}\mathrm{d}x + \frac{\partial z}{\partial y}\mathrm{d}y\right) \\ &= \mathrm{d}\left(\frac{\partial z}{\partial x}\right)\mathrm{d}x + \frac{\partial z}{\partial x}\mathrm{d}^2 x + \mathrm{d}\left(\frac{\partial z}{\partial y}\right)\mathrm{d}y + \frac{\partial z}{\partial y}\mathrm{d}^2 y \\ &= \left(\frac{\partial^2 z}{\partial x^2}\mathrm{d}x + \frac{\partial^2 z}{\partial x \partial y}\mathrm{d}y\right)\mathrm{d}x + \left(\frac{\partial^2 z}{\partial y \partial x}\mathrm{d}x + \frac{\partial^2 z}{\partial y^2}\mathrm{d}y\right)\mathrm{d}y \\ &= \frac{\partial^2 z}{\partial x^2}(\mathrm{d}x)^2 + 2\frac{\partial^2 z}{\partial x \partial y}\mathrm{d}x\mathrm{d}y + \frac{\partial^2 z}{\partial y^2}(\mathrm{d}y)^2, \end{aligned}$$

where $(\mathrm{d}x)^2$ and $(\mathrm{d}y)^2$ denote $\mathrm{d}^2 x$ and $\mathrm{d}^2 y$ respectively. If we treat $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ as operators for partial differentiation and define

$$\left(\frac{\partial}{\partial x}\right)^2 = \frac{\partial^2}{\partial x^2}, \quad \left(\frac{\partial}{\partial y}\right)^2 = \frac{\partial^2}{\partial y^2}, \quad \left(\frac{\partial}{\partial x}\frac{\partial}{\partial y}\right) = \frac{\partial^2}{\partial x \partial y},$$

then the formulas for the first and second differentials can be written as

$$\mathrm{d}z = \left(\mathrm{d}x\frac{\partial}{\partial x} + \mathrm{d}y\frac{\partial}{\partial y}\right) z,$$

$$\mathrm{d}^2 z = \left(\mathrm{d}x\frac{\partial}{\partial x} + \mathrm{d}y\frac{\partial}{\partial y}\right)^2 z.$$

Similarly, we define

$$\left(\frac{\partial}{\partial x}\right)^p \left(\frac{\partial}{\partial y}\right)^q = \frac{\partial^{p+q}}{\partial x^p \partial y^q} = \frac{\partial^q}{\partial y^q}\left(\frac{\partial}{\partial x}\right)^p, \quad (p, q = 1, 2, \dots)$$

It is easy to use mathematical induction to prove the formula for higher-order differentials:

$$\mathrm{d}^k z = \left(\mathrm{d}x\frac{\partial}{\partial x} + \mathrm{d}y\frac{\partial}{\partial y}\right)^k z, \quad k = 1, 2, \cdots.$$

For an n-variable function $u = f(x_1, x_2, \dots, x_n)$, higher-order differentials can be similarly defined, and the following holds:

$$\mathrm{d}^k u = \left(\mathrm{d}x_1\frac{\partial}{\partial x_1} + \mathrm{d}x_2\frac{\partial}{\partial x_2} + \cdots + \mathrm{d}x_n\frac{\partial}{\partial x_n}\right)^k u, \quad k = 1, 2, \dots.$$

12.2 Differential of Vector-Valued Functions

Consider an n-dimensional vector-valued function defined on a domain $U \subset \mathbb{R}^n$:

$$f : U \to \mathbb{R}^m,$$
$$\mathbf{x} \mapsto \mathbf{y} = f(\mathbf{x})$$

Expressed in coordinate vector form:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in U$$

1. If each component function $f_i(x_1, x_2, \ldots, x_n)$ $(i = 1, 2, \ldots, m)$ is partially differentiable at $\mathbf{x}^0$, then the vector-valued function **f** is differentiable at $\mathbf{x}^0$, and we define the matrix

$$\left(\frac{\partial f}{\partial x_j}(\mathbf{x}^0)\right)_{m\times n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}^0) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}^0) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}^0) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}^0) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}^0) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}^0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}^0) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}^0) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}^0) \end{pmatrix}$$

   This matrix is called the Jacobian matrix of **f** at $\mathbf{x}^0$, denoted by $f'(\mathbf{x}^0)$ (or $\mathrm{D}f(\mathbf{x}^0)$, $J_f(\mathbf{x}^0)$).
   For the special case $m = 1$, i.e., n-variable scalar function $z = f(x_1, x_2, \ldots, x_n)$, the derivative at $\mathbf{x}^0$ is

$$f'(\mathbf{x}^0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}^0), \frac{\partial f}{\partial x_2}(\mathbf{x}^0), \cdots, \frac{\partial f}{\partial x_n}(\mathbf{x}^0)\right)$$

   If the vector-valued function **f** is differentiable at every point in U, then **f** is said to be differentiable on U, and the corresponding relationship is

$$\mathbf{x} \in U \mapsto f'(\mathbf{x}) = J_f(\mathbf{x})$$

   where $f'(\mathbf{x})$ (or $\mathrm{D}f(\mathbf{x})$, $J_f(\mathbf{x})$) denotes the derivative of **f** at $\mathbf{x}$ in U.
2. If every component function $f_i(x_1, x_2, \ldots, x_n)$ $(i = 1, 2, \ldots, m)$ of **f** has continuous partial derivatives at $\mathbf{x}^0$, then every element of the Jacobian matrix of **f** is continuous at $\mathbf{x}^0$. In this case, **f** is said to have a continuous derivative at $\mathbf{x}^0$ as a vector-valued function.
   If the derivative of a vector-valued function **f** is continuous at every point in U, then **f** is said to have a continuous derivative on U.
3. If there exists an $m \times n$ matrix A that depends only on $\mathbf{x}^0$ (and not on $\Delta\mathbf{x}$), such that in the neighborhood of $\mathbf{x}^0$,

$$\Delta\mathbf{y} = f(\mathbf{x}^0 + \Delta\mathbf{x}) - f(\mathbf{x}^0) = A\Delta\mathbf{x} + o(\|\Delta\mathbf{x}\|)$$

   (where $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)^T$ is a column vector and $\|\Delta\mathbf{x}\|$ denotes its norm), then f is said to be differentiable at $\mathbf{x}^0$ as a vector-valued function, and $A\Delta\mathbf{x}$ is called the differential of f at $\mathbf{x}^0$, denoted as $\mathrm{d}\mathbf{y}$. If we denote $\Delta\mathbf{x}$ by $\mathrm{d}\mathbf{x}$ ($\mathrm{d}\mathbf{x} = (\mathrm{d}x_1, \mathrm{d}x_2, \ldots, \mathrm{d}x_n)^T$), then

$$\mathrm{d}\mathbf{y} = A\,\mathrm{d}\mathbf{x}.$$

If the vector-valued function $\mathbf{f}$ is differentiable at every point in U, then $\mathbf{f}$ is said to be differentiable on U.

Combining the above three points, we obtain the following unified statement:

A vector-valued function $\mathbf{f}$ is continuous, differentiable, and has derivatives if and only if each of its coordinate component functions $f_i(x_1, x_2, \ldots, x_n)$ $(i = 1, 2, \ldots, m)$ is continuous, differentiable, and has derivatives.

12.3 Derivatives of Composite Mappings (Chain Rule)

Let $U \subset \mathbb{R}^l$ and $V \subset \mathbb{R}^n$ be open sets, and let

$$\mathbf{g} : U \to V \quad \text{and} \quad \mathbf{f} : V \to \mathbb{R}^m$$

be mappings. If $\mathbf{g}$ is derivative at $\mathbf{u}^0 \in U$ and $\mathbf{f}$ is differentiable at $\mathbf{x}^0 = \mathbf{g}(\mathbf{u}^0)$, then the composite mapping $\mathbf{f} \circ \mathbf{g}$ is differentiable at $\mathbf{u}^0$, and:

$$J(\mathbf{f} \circ \mathbf{g})(\mathbf{u}^0) = J\mathbf{f}(\mathbf{x}^0) J\mathbf{g}(\mathbf{u}^0).$$

Note

1. *outer differentiable + inner derivative = total derivative*
2. *outer differentiable + inner differentiable = total differentiable*
3.

Specially, define $z = f(x, y), (x, y) \subset D_f \subset \mathbb{R}^2$, $\mathbf{g} : D_g \to \mathbb{R}^2, (u, v) \mapsto (x(u, v), y(u, v))$, and $g(D_g) \subset D_f$, then we have composite function

$$z = f \circ \mathbf{g} = f\left[x(u, v), y(u, v)\right], \quad (u, v) \in D_g.$$

$$\mathbb{R}^2 \xrightarrow{\mathbf{g}\text{:derivative}} \mathbb{R}^2 \xrightarrow{f\text{:differentiable}} \mathbb{R}$$

If $\mathbf{g}$ is derivative at $(u_0, v_0) \in D_g$, and f is differentiable at $(x_0, y_0) = \mathbf{g}(u_0, v_0)$, then $z = f \circ \mathbf{g}$ is differentiable at (u_0, v_0), and at the point,

$$\begin{bmatrix} \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$

Proof

■

12.4 Mean Value Theorem and Taylor's Formula

Definition 12.4 (Convex Region)

Let $D \subseteq \mathbb{R}^n$ be a region. If every line segment connecting any two points $\mathbf{x}_0, \mathbf{x}_1 \in D$ (denoted by $\overline{\mathbf{x}_0\mathbf{x}_1}$) is entirely contained in D, i.e., for any $\lambda \in [0, 1]$, we have

$$\mathbf{x}_0 + \lambda(\mathbf{x}_1 - \mathbf{x}_0) \in D,$$

then D is called a convex region. ♣

Theorem 12.3 (Lagrange's Mean Value Theorem)

Let f be differentiable on a convex region $D \subseteq \mathbb{R}^n$. For any two points $\mathbf{a}, \mathbf{b} \in D$, there exists a point $\xi \in \overline{\mathbf{ab}}$ such that:

$$f(\mathbf{b}) - f(\mathbf{a}) = Jf(\xi)(\mathbf{b} - \mathbf{a}).$$

♡

Proof

■

Theorem 12.4

Let D be a region in $\mathbb{R}^n$. If for any $\mathbf{x} \in D$, we have

$$Jf(\mathbf{x}) = 0,$$

then f is constant on D. ♡

Proof

■

Theorem 12.5 (Taylor's Formula)

Lagrange's Remainder Let $D \subseteq \mathbb{R}^n$ be a convex region, and let $f : D \to \mathbb{R}$ have $m + 1$ continuous partial derivatives. For $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in D$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in D$, there exists $\xi \in \overline{\mathbf{x}^0\mathbf{x}}$ such that:

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \sum_{k=1}^{m} \frac{1}{k!} \left(\sum_{i=1}^{n} (x_i - x_i^0) \frac{\partial}{\partial x_i} \right)^k f(\mathbf{x}^0) + \frac{1}{(m+1)!} \left(\sum_{i=1}^{n} (x_i - x_i^0) \frac{\partial}{\partial x_i} \right)^{m+1} f(\xi).$$

Peano's Remainder Let $D \subseteq \mathbb{R}^n$ be a convex region, and let $f : D \to \mathbb{R}$ have m continuous partial derivatives. Then:

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \sum_{k=1}^{m} \frac{1}{k!} \sum_{i_1, i_2, \ldots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}(\mathbf{x}^0) \prod_{j=1}^{k} (x_{i_j} - x_{i_j}^0) + R_m(\mathbf{x} - \mathbf{x}^0),$$

where $R_m(\mathbf{x} - \mathbf{x}^0) = O(\|\mathbf{x} - \mathbf{x}^0\|^{m+1})$ or $o(\|\mathbf{x} - \mathbf{x}^0\|^m)$, as $\|\mathbf{x} - \mathbf{x}^0\| \to 0$. ♡

In applications, particularly important is the expression of the first three terms in Taylor's formula, which is given as (let $x_1 - x_1^0$ be denoted by Δx_1, and similarly for other variables; $\Delta\mathbf{x} = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)$):

$$f(\mathbf{x}) = f(\mathbf{x}^0) + Jf(\mathbf{x}^0)(\Delta\mathbf{x}) + \frac{1}{2!}(\Delta\mathbf{x})Hf(\mathbf{x}^0)(\Delta\mathbf{x})^{\mathrm{T}} + \cdots,$$

where the matrix

$$Hf(\mathbf{x}^0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}_{\mathbf{x}^0}$$

is called the **Hessian matrix** of the function f. It is an $n \times n$ symmetric matrix.

12.5 Implicit Function Theorem

Theorem 12.6 (Implicit Function Theorem)

Let $U \subset \mathbb{R}^{n+1}$ be an open set, and $F : U \to \mathbb{R}$ be an $n+1$-variable function. If:

1. $F \in C^k(U, \mathbb{R})$, where $1 \leq k \leq +\infty$;
2. $F(\mathbf{x}^0, y^0) = 0$, where $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in \mathbb{R}^n$, $y^0 \in \mathbb{R}$, and $(\mathbf{x}^0, y^0) \in U$ (i.e., the equation $F(\mathbf{x}, y) = 0$ has a solution $(\mathbf{x}^0, y^0)$);
3. $F'_y(\mathbf{x}^0, y^0) \neq 0$.

Then there exists an open interval $I \times J$ containing $(\mathbf{x}^0, y^0)$ (I being an open interval in $\mathbb{R}^n$ containing $\mathbf{x}^0$, and J being an open interval in $\mathbb{R}$ containing y^0), as shown in Fig. 12.1, such that:

1. $\forall x \in I$, the equation $F(\mathbf{x}, y) = 0$ has a unique solution $y = f(\mathbf{x})$, where $f : I \to J$ is an n-variable function (called the **implicit function** f, hidden within the equation $F(\mathbf{x}, f(\mathbf{x})) = 0$, though not necessarily explicitly expressed);
2. $y^0 = f(\mathbf{x}^0)$;
3. $f \in C^k(I, \mathbb{R})$;
4. When $x \in I$, $\frac{\partial f}{\partial x_i} = \frac{\partial y}{\partial x_i} = -\frac{F_x(\mathbf{x},y)}{F_y(\mathbf{x},y)}$, $i = 1, 2, \ldots, n$, where $y = f(x)$.

♡

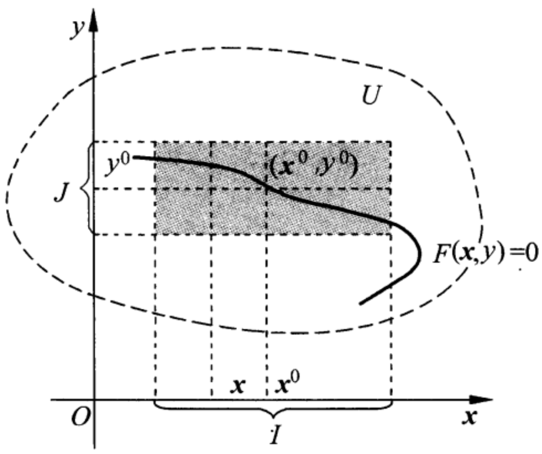


Figure 12.1: Implicit Function

Proof Only the single-variable implicit function theorem is proved; the multi-variable case can be derived using mathematical induction.

Without loss of generality, assume $F_y(x^0, y^0) > 0$.

First, prove the existence of the implicit function. From the continuity of $F_y(x^0, y^0) > 0$ and $F_y(x, y)$, it is known that there exist closed rectangle:

$$D^* = \{(x, y) \mid |x - x_0| \leq \alpha, |y - y_0| \leq \beta\} \subset U,$$

where the following holds:

$$F_y(x, y) > 0.$$

Thus, for fixed x_0, the function $F(x^0, y)$ is strictly monotonically increasing within $[y^0 - \beta, y^0 + \beta]$. Furthermore, since:

$$F(x^0, y^0) = 0,$$

it follows that:

$$F(x^0, y^0 - \beta) < 0, \quad F(x^0, y^0 + \beta) > 0.$$

Due to the continuity of $F(x, y)$ within D^*, there exists $\rho > 0$ such that along the line segment:

$$x = x^0 + \rho,\ y = y^0 + \beta,$$

we have $F(x, y) > 0$, and along the line segment:

$$x = x^0 + \rho,\ y = y^0 - \beta,$$

we have $F(x, y) < 0$. Therefore, for any point $\bar{x} \in (x^0 - \rho, x^0 + \rho)$, treat $F(x, y)$ as a single-variable function of y. Within $[y^0 - \beta, y^0 + \beta]$, this function is continuous. From the previous discussion, we know:

$$F(\bar{x}, y^0 - \beta) < 0, \quad F(\bar{x}, y^0 + \beta) > 0.$$

According to the zero point existence theorem 3.3, there must exist a unique $\bar{y} \in [y^0 - \beta, y^0 + \beta]$ such that $F(\bar{x}, \bar{y}) = 0$. Furthermore, because $F_y(x, y) > 0$ within D^*, this $\bar{y}$ is unique. Denote the corresponding relationship as $\bar{y} = f(\bar{x})$, then the function $y = f(x)$ is defined within $(x^0 - \rho, x^0 + \rho)$, satisfying $F(x, f(x)) = 0$, and clearly:

$$y^0 = f(x^0).$$

Further proving the continuity of the implicit function $y = f(x)$ on $(x^0 - \rho, x^0 + \rho)$: Let $\bar{x} \in (x^0 - \rho, x^0 + \rho)$ be any point. For any given $\varepsilon > 0$ (ε being sufficiently small), since $F(\bar{x}, \bar{y}) = 0$ ($\bar{y} = f(\bar{x})$), from the previous discussion we know:

$$F(\bar{x}, \bar{y} - \varepsilon) < 0, \quad F(\bar{x}, \bar{y} + \varepsilon) > 0.$$

Furthermore, due to the continuity of $F(x, y)$ on D^*, there exists $\delta > 0$ such that:

$$F(x, \bar{y} - \varepsilon) < 0, \quad F(x, \bar{y} + \varepsilon) > 0, \quad \text{when} \quad x \in O(x^0, \delta).$$

By reasoning similar to the previous discussion, it can be obtained that when $x \in O(x^0, \delta)$, the corresponding implicit function value must satisfy $f(x) \in (\bar{y} - \varepsilon, \bar{y} + \varepsilon)$, i.e.,

$$\left|f(x) - f(x^0)\right| < \varepsilon.$$

This implies that $y = f(x)$ is continuous on $(x^0 - \rho, x^0 + \rho)$.

Finally, prove the differentiability of $y = f(x)$ on $(x^0 - \rho, x^0 + \rho)$: Let $\bar{x} \in (x^0 - \rho, x^0 + \rho)$ be any point. Take Δx sufficiently small such that $\bar{x} = x + \Delta x \in (x^0 - \rho, x^0 + \rho)$. Denote $\bar{y} = f(\bar{x})$ and $\bar{y} + \Delta y = f(\bar{x})$. Clearly,

$$F(\bar{x}, \bar{y}) = 0 \quad \text{and} \quad F(\bar{x}, \bar{y} + \Delta y) = 0.$$

Using the multi-variable function's mean value theorem 12.3, we obtain:

$$\begin{aligned} 0 &= F(\bar{x}, \bar{y} + \Delta y) - F(\bar{x}, \bar{y}) \\ &= F_x(\bar{x} + \theta\Delta x, \bar{y} + \theta\Delta y)\Delta x + F_y(\bar{x} + \theta\Delta x, \bar{y} + \theta\Delta y)\Delta y, \end{aligned}$$

where $0 < \theta < 1$. Note that $F_y \neq 0$ on D^*, hence:

$$\frac{\Delta y}{\Delta x} = -\frac{F_x(\bar{x} + \theta\Delta x, \bar{y} + \theta\Delta y)}{F_y(\bar{x} + \theta\Delta x, \bar{y} + \theta\Delta y)}.$$

Let $\Delta x \to 0$. Considering the continuity of F_x and F_y, we obtain:

$$\left.\frac{dy}{dx}\right|_{x=\bar{x}} = -\frac{F_x(\bar{x}, \bar{y})}{F_y(\bar{x}, \bar{y})}.$$

Thus:

$$f'(\bar{x}) = -\frac{F_x(\bar{x}, \bar{y})}{F_y(\bar{x}, \bar{y})}.$$

The proof is complete. ∎

Theorem 12.7 (Implicit Mapping Theorem)

Let $U \subset \mathbb{R}^{n+m}$ be an open set, and $\mathbf{F} : U \to \mathbb{R}^m$ be a mapping. If:

1. $\mathbf{F} \in C^k(U, \mathbb{R}^m), 1 \le k \le \infty$;
2. $\mathbf{F}(\mathbf{x}^0, \mathbf{y}^0) = 0$, where $\mathbf{x}^0 = (x_1, x_2, \dots, x_n)$, $\mathbf{y}^0 = (y_1, y_2, \dots, y_m)$, $(\mathbf{x}^0, \mathbf{y}^0) \in U$ (implying $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ has a solution at $(\mathbf{x}^0, \mathbf{y}^0)$);
3. The determinant

$$\det \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}_{(\mathbf{x}^0, \mathbf{y}^0)} = \det J_{\mathbf{y}}\mathbf{F}(\mathbf{x}^0, \mathbf{y}^0) \neq 0,$$

then there exists an open neighborhood $I \times J \subset U \subset \mathbb{R}^{n+m}$ containing $(\mathbf{x}^0, \mathbf{y}^0)$, such that:

1. For all $\mathbf{x} \in I$, the system $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ has a unique solution $\mathbf{y} = \mathbf{f}(\mathbf{x})$, where $\mathbf{f} : I \to J$ is a mapping (called $\mathbf{f}$ the implicit function hidden in $\mathbf{F}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = \mathbf{0}$);
2. $\mathbf{y}^0 = \mathbf{f}(\mathbf{x}^0)$;

3. $\mathbf{f} \in C^k(I, \mathbb{R}^m)$;
4. For $x \in I$,

$$J_{\mathbf{f}}(x) = -(J_{\mathbf{y}}\mathbf{F}(\mathbf{x}, \mathbf{f}(x)))^{-1} J_{\mathbf{x}}\mathbf{F}(\mathbf{x}, \mathbf{f}(x)) = -\begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix},$$

where $\mathbf{y} = \mathbf{f}(\mathbf{x})$.

12.6 Applications of Multi-Variable Differential Calculus

¶ Surface and Tangent Space

Definition 12.5 (Parameterization of Surface)

Let Δ be an open subset in $\mathbb{R}^s$, and $\mathbf{x} : \Delta \to \mathbb{R}^n$ be a mapping, where $\mathbf{u} = (u_1, u_2, \ldots, u_s) \to \mathbf{x}(\mathbf{u}) = (x_1(u_1, u_2, \ldots, u_s), x_2(u_1, u_2, \ldots, u_s), \ldots, x_n(u_1, u_2, \ldots, u_s))$. Then $M = \mathbf{x}(\Delta) = \{\mathbf{x}(\mathbf{u}) \mid \mathbf{u} \in \Delta\}$ is called an s-dimensional surface, and $\mathbf{x}(\mathbf{u})$ is referred to as the parameterization of M. When $\mathbf{x}(\mathbf{u}) \in C^k$ $(k \geq 0)$, $\mathbf{x}$ or M is called an s-dimensional C^k surface.

If $\mathbf{x} \in C^k$ $(k \geq 1)$, $\mathbf{x}$ or M is called an s-dimensional C^k smooth surface. When

$$\operatorname{rank}(x_1'(\mathbf{u}^0), x_2'(\mathbf{u}^0), \ldots, x_s'(\mathbf{u}^0)) = \operatorname{rank} \begin{pmatrix} \frac{\partial x_1}{\partial u_1} & \cdots & \frac{\partial x_1}{\partial u_s} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial u_1} & \cdots & \frac{\partial x_n}{\partial u_s} \end{pmatrix}_{\mathbf{u}^0} = s,$$

we call $\mathbf{u}^0$ or $\mathbf{x}(\mathbf{u}^0)$ a regular point of the surface M. Otherwise, it is called a singular point. Every point that is a regular point of the surface is referred to as an s-dimensional C^k regular surface. At such points, $\{x_1', \ldots, x_s'\}$ are linearly independent.

When $s = 1$, t represents the parameter, a one-dimensional surface is commonly referred to as a curve. Considering a C^k $(k \geq 1)$ curve $\mathbf{x}(t)$, we have:

$$\mathbf{x}'(t) = \left(x_1'(t), x_2'(t), \cdots, x_n'(t)\right).$$

If t is a regular point, then $\operatorname{rank}(\mathbf{x}'(t)) = \operatorname{rank}(x_1'(t), x_2'(t), \ldots, x_n'(t)) = 1$; this is equivalent to $\mathbf{x}'(t) \neq 0$, which means $x_1'(t), x_2'(t), \ldots, x_n'(t)$ are not all zero.

We refer to $\mathbf{x}'(t)$ as the tangent vector of the curve $\mathbf{x}(t)$ at point t. When t varies, a tangent vector field along the curve $\mathbf{x}(t)$ is obtained. If $\mathbf{x}(t)$ is a regular curve, $\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$ is the unit tangent vector field along the curve $\mathbf{x}(t)$. It should be emphasized that $\mathbf{x}'(t)$ or $\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$ always points outward from point t.

Definition 12.6 (Tangent Vector)

¶ Unconditional Extremum

¶ Conditional Extremum

Chapter 13 Multiple Integrals

Bibliography

[1] 徐森林, 薛春华. *数学分析 (1st edition)* . 清华大学出版社, 2005.

[2] 陈纪修, 於崇华. *数学分析 (3rd edition)* . 高等教育出版社, 2019.

[3] 常庚哲, 史济怀. *数学分析教程 (3rd edition)* . 中国科学技术大学出版社, 2012.

[4] 裴礼文. *数学分析中的典型问题与方法 (3rd edition)* . 高等教育出版社, 2021.

[5] 汪林. *数学分析中的问题与反例 (1st edition)* . 高等教育出版社, 2015.

[6] 谢惠民, 恽自求, 易法槐, 钱定边. *数学分析习题课讲义 (2nd edition)* . 高等教育出版社, 2019.

[7] Walter Rudin. *Principles of Mathematical Analysis (3rd edition)* . McGraw-Hill, 1976.

[8] 菲赫金哥尔茨. *微积分学教程 (8th edition)* . 高等教育出版社, 2006.

[9] Wikipedia. https://en.wikipedia.org/wiki/.