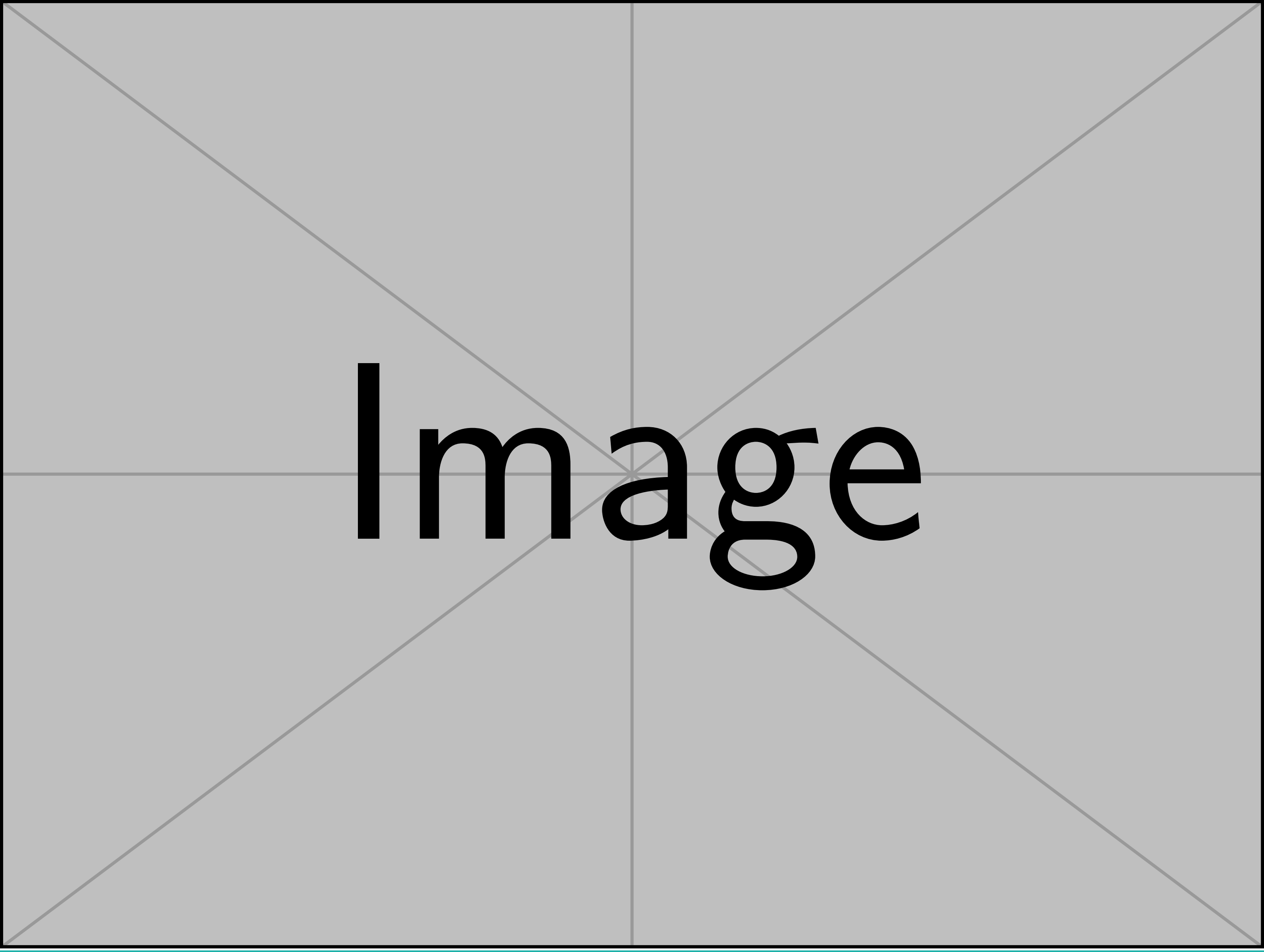


Analyse Mathématique

Author: CatMono

Date: July, 2025

Version: 0.1

Contents

Preface

For an interval I, a open interval (a, b) and a closed interval $[a, b]$, we denote $C(I)$, $C(a, b)$ and $C[a, b]$ as the set of continuous univariate functions on I, (a, b) and $[a, b]$ respectively. Similarly, the following notations are used[1]:

Notation	Meaning
$D(I)$	Set of derivative (differential) functions on I
$D(a, b)$	Set of derivative (differential) functions on (a, b)
$D[a, b]$	Set of derivative (differential) functions on $[a, b]$
$D^k(I)$	Set of k-th order derivative (differential) functions on I

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 set of C^k mappings from U to $\mathbb{R}^m$.

Sometimes, we use subscripts i to denote the partial derivative with respect to the i-th variable, for example, for function $f(x^2 + y^2 + z^2, xyz)$, $f_2 := \frac{\partial f}{\partial (xyz)}$, and similarly for higher-order partial derivatives, e.g., $f_{12} := \frac{\partial^2 f(u,v)}{\partial v \partial u}$.

[1] Other notations include: $R[a, b]$ (denoting Riemann integrable functions on $[a, b]$), $B[a, b]$ (denoting bounded functions on $[a, b]$), etc.

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:

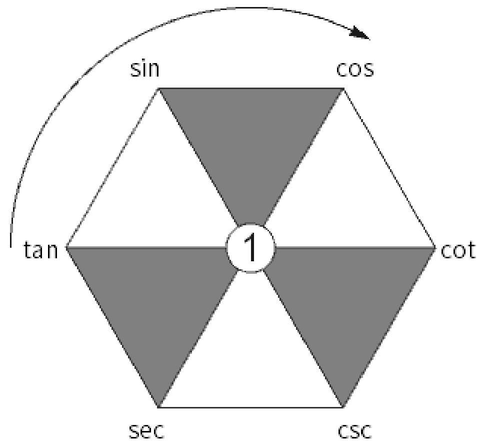
$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}$$
$$\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}}$$
$$\tan\frac{\alpha}{2} = \frac{1-\cos\alpha}{\sin\alpha} = \frac{\sin\alpha}{1+\cos\alpha}$$

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.

Theorem 1.1 (Weierstrass Substitution (All-Powerful Formula))

Let $t = \tan\frac{x}{2}$, then:

$$\sin x = \frac{2t}{1+t^2},$$
$$\cos x = \frac{1-t^2}{1+t^2},$$
$$\mathrm{d}x = \frac{2}{1+t^2}\mathrm{d}t.$$

1.2 Common Inequalities

Some common inequalities:

$$\frac{x}{1+x} < \ln(1+x) < x, \quad x > 0;$$

1.3 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:

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

♣

1.4 Combination

Definition 1.2 (Combination)

The number of ways to choose k elements from a set of n distinct elements, denoted as C_n^k or $\binom{n}{k}$, is given by:

$$\mathrm{C}_n^k = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

♣

Property

$$\mathrm{C}_n^k = \frac{\mathrm{A}_n^k}{k!} = \frac{n!}{(n-k)!k!}$$
$$\mathrm{C}_n^k = \mathrm{C}_n^{n-k}$$
$$\mathrm{C}_n^k = \mathrm{C}_{n-1}^{k-1} + \mathrm{C}_{n-1}^k$$

Remark The third property can be understood that to choose k elements from $n+1$, you can first take one element A:

1. The number of ways that include A is C_n^{k-1};
2. The number of ways that do not include A is C_n^k.

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:

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

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 $x_1, x_2 \in I$,

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

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

Some useful formulas of higher-order derivatives:

$$\begin{aligned}(a^x)^{(n)} &= (\ln a)^n a^x, \\ (\sin \alpha x)^{(n)} &= \alpha^n \sin\left(\alpha x + \frac{n\pi}{2}\right), \\ (\cos \alpha x)^{(n)} &= \alpha^n \cos\left(\alpha x + \frac{n\pi}{2}\right), \\ (\ln x)^{(n)} &= \frac{(-1)^{n-1}(n-1)!}{x^n}, \\ (x^\alpha)^{(n)} &= \alpha(\alpha-1)\cdots(\alpha-n+1)x^{\alpha-n}.\end{aligned}$$

In order to obtain the higher-order derivative of two or more functions' linear combination and product, we need to use the following theorems.

Theorem 4.1 (Linear Operation of Higher-Order Derivatives)

If $f, g \in D^{(n)}(I)$, then for any constants $c_1, c_2 \in \mathbb{R}$,

$$(c_1 f + c_2 g)^{(n)} = c_1 f^{(n)} + c_2 g^{(n)}.$$

♡

Theorem 4.2 (Leibniz's Formula)

If $f, g \in D^{(n)}(I)$, then

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

♡

Caution *Note the distinction:*

- $\mathrm{d}x^2$ *represents the square of the differential of the independent variable, i.e.,* $(\mathrm{d}x)^2$*;*
- d^2x *represents the second differential of the independent variable,* $\mathrm{d}(\mathrm{d}x)$*;*
- $\mathrm{d}(x^2)$ *represents the differential of* x^2*, which is* $2x\mathrm{d}x$*.*

4.3 Differential Mean Value Theorems

Definition 4.1 (Argmax and Argmin)

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 arguments of the maxima point of f, and $f(x_0)$ is referred to as the corresponding arguments of the maxima (abbreviated arg max or argmax).

The definition of the argmin 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 4.3 (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 4.4 (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 4.5 (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

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

Note *The following types of problems commonly appear in proofs related to intermediate values in differential calculus:*

1. *Prove the existence of a point ξ such that $F(\xi, f(\xi), f'(\xi)) = 0$. Problems of this type generally involve constructing auxiliary functions and applying Rolle's theorem. The commonly used auxiliary functions include:*

$$\begin{aligned}
\xi f'(\xi) + f(\xi) = 0, &\quad x f(x), \\
\xi f'(\xi) + n f(\xi) = 0, &\quad x^n f(x), \\
\xi f'(\xi) - f(\xi) = 0, &\quad e^x f(x), \\
f'(\xi) + \lambda f(\xi) = 0, &\quad e^{-x} f(x), \\
f'(\xi) + f(\xi) = 0, &\quad x^n f(x), \\
f'(\xi) - f(\xi) = 0, &\quad x f(x).
\end{aligned}$$

2. *Prove the existence of two points ξ, η (i.e., two intermediate values) such that $F(\xi, f(\xi), f'(\xi), \eta, f(\eta), f'(\eta)) = 0$. These problems can be divided into the following categories:*

   $\xi \neq \eta$ *Problems of this type usually occur in the same interval $[a, b]$ and employ theorems of double differentiation intermediate values such as the Lagrange mean value theorem or Cauchy's mean value theorem. The specific choice of auxiliary functions often includes terms like ξ and other variables determined after decomposition.*

   $\xi = \eta$ *Such problems cannot occur within the same interval $[a, b]$. They use double differentiation mean value theorems by splitting $[a, b]$ into two intervals $[a, c]$ and $[c, b]$, applying the Lagrange mean value theorem separately to each interval. Here, the selection of ξ and η is key.*

3. *As a rule, when conditions in a theorem involve additional constraints about higher-order derivatives, it is necessary to use Taylor's intermediate value theorem.*

4.4 Theorems about Derivatives

Theorem 4.6 (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 4.7 (Theorem on the Limit of Derivatives)

If $f(x) \in C(U(x_0)), D(\mathring{U}(x_0))$, 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 $f \in D(\mathring{U}(x_0))$.

The mnemonic for this theorem is: Continuous function + limit of derivative ⇒ derivative at the point.

4.5 Taylor Theorem

¶ L'Hôpital's Rule

¶ Taylor Formula

¶ Maclaurin Formula

Lemma 4.2

If $f(x)$ has $n+2$ derivatives in some neighborhood of x_0, then the derivative of its $n+1$th degree Taylor polynomial is exactly the nth degree Taylor polynomial of $f'(x)$.

Taylor formula at $x_0 = 0$ is called the **Maclaurin formula**. Some common Maclaurin formulas are as follows:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n),$$
$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + o(x^n),$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!} + o(x^{2n}),$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n\frac{x^{2n}}{(2n)!} + o(x^{2n+1}),$$

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^{n-1}\frac{x^{2n-1}}{2n-1} + o(x^{2n}),$$
$$\arcsin x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} + \cdots + \frac{(2n-1)!!}{(2n)!!}\frac{x^{2n+1}}{2n+1} + o(x^{2n+2}).$$

Specially,

$$(1+x)^\alpha = \sum_{k=0}^{\alpha}\binom{\alpha}{k}x^k + o(x^n),$$

- if $\alpha = n \in \mathbb{N}^+$, that is Newton's binomial formula $(1+x)^n = 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \cdots + \binom{n}{n}x^n$;
- if $\alpha = \frac{1}{2}$, then $(1+x)^{\frac{1}{2}} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \cdots$;
- if $\alpha = -1$, then $(1+x)^{-1} = 1 - x + x^2 - x^3 + \cdots$;
- if $\alpha = -\frac{1}{2}$, then $(1+x)^{-\frac{1}{2}} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \cdots$.

¶ Euler and Bernoulli Numbers

Definition 4.2 (Euler Numbers)

The Euler numbers E_n are defined by the Taylor series expansion of the secant function:

$$\operatorname{sech} x = \frac{2}{e^x + e^{-x}} = \sum_{n=0}^{\infty} E_n \frac{x^n}{n!}.$$

The odd-indexed Euler numbers are all zero. The even-indexed ones have alternating signs. Some values are:

$$E_0 = 1, \quad E_2 = -1, \quad E_4 = 5, \quad E_6 = -61, \quad E_8 = 1385.$$

♣

Definition 4.3 (Bernoulli Numbers)

The Bernoulli numbers B_n are defined by the Taylor series expansion of the function $\frac{x}{e^x-1}$:

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}.$$

Some values are:

$$B_0 = 1, \quad B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad B_8 = -\frac{1}{30}.$$

Notably, all odd-indexed Bernoulli numbers (except $B_1 = -\frac{1}{2}$) are zero.

♣

Remark Euler and Bernoulli numbers are widely used in number theory, combinatorics, and numerical analysis. For example, in the infinite series:

$$\sum_{n=1}^{\infty} \frac{1}{n^{2k}} = (-1)^{k-1} \frac{(2\pi)^{2k}}{2(2k)!} B_{2k}, \quad k \in \mathbb{N}^+,$$

when $k = 1$, it gives the famous Basel problem result:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

With the help of Bernoulli numbers, we have

$$\tan x = \sum_{n=0}^{\infty} \frac{B_{2n}}{2n} \frac{x^{2n}}{(2n)!} = x + \frac{x^3}{3} + \frac{2}{15} x^5 + \cdots.$$

4.6 Properties of Functions

¶ Monotonicity and Convexity

Definition 4.4 (Convex Function)

A function f is called **convex** on an interval I if for any $x_1, x_2 \in I$ and $t \in [0, 1]$, the following inequality holds:

$$f(tx_1 + (1-t)x_2) \leqslant tf(x_1) + (1-t)f(x_2).$$

If the inequality is strict for $x_1 \neq x_2$ and $t \in (0, 1)$, then f is called **strictly convex** on I. Conversely, if the inequality is reversed, then f is called **concave** or **concave down** on I.

♣

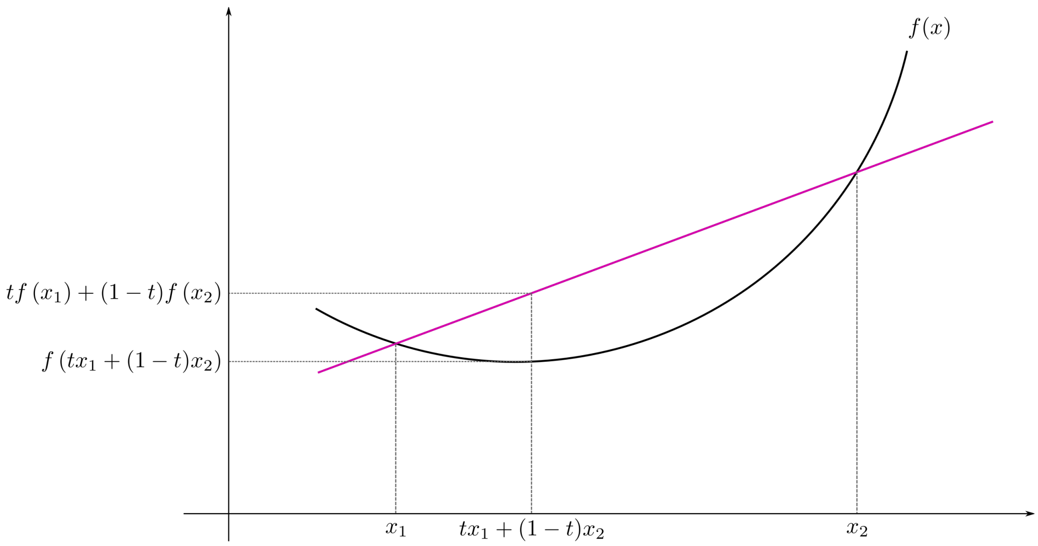
A related concept is that of **inflection points**: a point on the graph of a function at which the concavity changes.

Theorem 4.8

Mark above definition as Definition 1, give the following statements:

2. (**Jensen Definition**) A function f is called convex on an interval I if for any $x_1, x_2 \in I$:

$$f\left(\frac{x_1 + x_2}{2}\right) \leqslant \frac{f(x_1) + f(x_2)}{2}.$$



3. A function f is called convex on an interval I if for any $x_1, x_2, \cdots, x_n \in I$:
$$f\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \leqslant \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n}.$$
4. A function f is called convex on an interval I if the tangent line at any point lies below the graph of the function.

Then,

- Definitions 2 and 3 are equivalent.
- When f is continuous, Definition 1, 2, 3 is equivalent.
- When f is differentiable, all four definitions are equivalent.

Theorem 4.9 (Jensen Inequality)

If f is convex on an interval I, then for any $x_1, x_2, \cdots, x_n \in I$ and any $t_1, t_2, \cdots, t_n > 0$ such that $t_1 + t_2 + \cdots + t_n = 1$, the following inequality holds:
$$f(t_1x_1 + t_2x_2 + \cdots + t_nx_n) \leqslant t_1f(x_1) + t_2f(x_2) + \cdots + t_nf(x_n).$$
Specially, when $t_1 = t_2 = \cdots = t_n = \frac{1}{n}$, it reduces to Definition 3.

Next, we present derivative-based criteria for monotonicity and convexity:

Theorem 4.10

1. If $f \in D(I)$, then f is increasing (decreasing) on I if and only if $f'(x) \geq 0$ ($f'(x) \leqslant 0$) for all $x \in I$.
2. If $f \in D^{(2)}(I)$, then f is convex (concave) on I if and only if $f''(x) \geq 0$ ($f''(x) \leqslant 0$) for all $x \in I$.

Note *If $f'(x) > 0$ ($f''(x) > 0$) for all $x \in I$, then f is strictly increasing (convex) on I. Even though the condition weakens to holding except at finitely many points, the conclusion of strict monotonicity (convexity) still holds. For example, $f(x) = x^3$ is strictly increasing on $\mathbb{R}$ despite $f'(0) = 0$.*

¶ `Argmax and Argmin`

Definition 4.5 (Stationary Point)

Stationary points are points where the first derivative of a function is zero or non-existent. ♣

Stationary points can be classified into three types:

Argmax and argmin points Points where the function attains its maximum or minimum values.

Inflection points Points where the function changes concavity.

Trivial points Points that are neither local maxima nor local minima.

¶ `Asymptote`

4.7 Applications

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:

$$\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u.$$

♣

Definition 5.2 (Substitution Method)

♣

Some common substitutions are as follows:

Trigonometric Substitution When restoring variables, auxiliary right triangles is often utilized.

Sine $\sqrt{a^2 - x^2}$: $x = a\sin t$ or $x = a\cos t$

Tangent $\sqrt{a^2 + x^2}$: $x = a\tan t$ or $x = a\sinh t$

Secant $\sqrt{x^2 - a^2}$: $x = a\sec t$ or $x = a\cosh t$

Irreational Substitution

- If the integrand contains $\sqrt[n]{x}$, one can use the substitution $t = \sqrt[n]{x}$ to simplify the expression.
- If the integrand contains $\sqrt[n]{\frac{\alpha x+\beta}{\gamma x+\delta}}$, one can use the substitution $t = \sqrt[n]{\frac{\alpha x+\beta}{\gamma x+\delta}}$ to simplify the expression.

Reciprocal Substitution If the degree of the numerator is lower than that of the denominator according to x one can use the substitution $x = \frac{1}{t}$ to reduce the degree.

¶ **Basic Integration Formulas**

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

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 $\to$ Intermediate points $\to$ Summation $\to$ 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

Definition 6.4 (Null Set)

A set $E \subset \mathbb{R}$ is called a **null set** (or measure zero set) if for any $\varepsilon > 0$, there exists a countable collection of open intervals $\{I_n | n \in \mathbb{N}^*\}$ such that:

$$E \subset \bigcup_{i=1}^{\infty} I_n \quad \text{and} \quad \sum_{i=1}^{\infty} |I_n| < \varepsilon.$$

If some property holds for all $x \in A$ except for a null set $E \subset A$, we say that the property holds **almost everywhere** on A.

Lemma 6.1

1. Let ω be the oscillation of bounded function $f(x)$ on $[a, b]$, then:

$$\omega = \sup\{f(y_1) - f(y_0) \mid y_0, y_1 \in [a, b]\}.$$

2. $f(x)$ is continuous at point x_0 if and only if the oscillation of $f(x)$ at x_0 is zero, i.e., $\omega_f(x_0) = 0$.
3. Let $D(f)$ be the set of discontinuities of bounded function $f(x)$ on $[a, b]$. For $\delta > 0$, denote $D_\delta = \{x \in [a, b] \mid \omega_f(x) \geqslant \delta\}$. Then

$$D(f) = \bigcup_{n=1}^{\infty} D_{\frac{1}{n}}.$$

4. If there exists a series of open intervals (α_j, β_j) $(j = 1, 2, \cdots)$ such that $D(f) \subset \bigcup_{j=1}^{\infty} (\alpha_j, \beta_j)$, and let $K = [a, b] \setminus \bigcup_{j=1}^{\infty} (\alpha_j, \beta_j)$. Then:

$$\forall \varepsilon > 0, \exists \delta > 0, \text{s.t.} \forall x \in K, y \in [a, b] (|x - y| < \delta) : |f(x) - f(y)| < \varepsilon.$$

Theorem 6.2 (Lesbesgue's Theorem)

Let $f(x) \in B[a, b]$, then $f(x)$ is Riemann integrable on $[a, b]$ if and only if $f(x)$ is continuous almost everywhere on $[a, b]$.

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 6.5 (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:

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

¶ Common Questions concerning Integrals

6.5 Calculation of Definite Integrals

Example 6.1 Prove the ignition formula (Wallis formula) with recursion method:

$$\int_0^{\frac{\pi}{2}} \sin^n x\,\mathrm{d}x = \int_0^{\frac{\pi}{2}} \cos^n x\,\mathrm{d}x = \begin{cases} \frac{(n-1)!!}{n!!} \cdot \frac{\pi}{2}, & n \text{ is even;} \\ \frac{(n-1)!!}{n!!}, & n \text{ is odd.} \end{cases}$$

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:

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

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:

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

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

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

Чебышёв 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:

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

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.

Example 6.2 Let $f(t)$ be convex on $[0,1]$, prove that:

$$\int_0^1 t(1-t)f(t)\,\mathrm{d}t \leqslant \frac{1}{3}\int_0^1 \left(t^3 + (1-t)^3\right) f(t)\,\mathrm{d}t.$$

Proof Since $f(t)$ is convex on $[0,1]$, for any $t \in (0,1)$, we have:

$$t = (1-t)(tx) + t(1-x+tx),$$

then

$$f(t) \leqslant (1-t)f(tx) + tf(1-x+tx).$$

Integrating both sides from 0 to 1 with respect to x, we get:

$$f(t) \leqslant (1-t)\int_0^1 f(tx)\,\mathrm{d}x + t\int_0^1 f(1-x+tx)\,\mathrm{d}x = \frac{1-t}{t}\int_0^t f(x)\mathrm{d}x + \frac{t}{1-t}\int_t^1 f(x)\mathrm{d}x.$$

Multiplying both sides by $t(1-t)$ and integrating from 0 to 1 with respect to t, we have:

$$\int_0^1 t(1-t)f(t)\,\mathrm{d}t \leqslant \int_0^1 \left[(1-t)^2\int_0^t f(x)\,\mathrm{d}x\right]\mathrm{d}t + \int_0^1 t^2\left[\int_t^1 f(x)\,\mathrm{d}x\right]\mathrm{d}t.$$

Change the order of integration in the right side:

$$\int_0^1 \left[(1-t)^2\int_0^t f(x)\,\mathrm{d}x\right]\mathrm{d}t + \int_0^1 t^2\left[\int_t^1 f(x)\,\mathrm{d}x\right]\mathrm{d}t = \frac{1}{3}\int_0^1 \left(t^3 + (1-t)^3\right) f(t)\,\mathrm{d}t.$$

Thus, the desired inequality is proven. ■

6.7 Applications of Definite Integrals

¶ 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	$\int_a^b f(x)\,\mathrm{d}x$	$\int_{T_1}^{T_2} \lvert y(t)x'(t)\rvert\,\mathrm{d}t$	$\frac{1}{2}\int_\alpha^\beta r^2(\theta)\,\mathrm{d}\theta$
Infinitesimal Arc Length	$\mathrm{d}l = \sqrt{1+[f'(x)]^2}\,\mathrm{d}x$	$\mathrm{d}l = \sqrt{[x'(t)]^2+[y'(t)]^2}\,\mathrm{d}t$	$\mathrm{d}l = \sqrt{r^2(\theta)+r'^2(\theta)}\,\mathrm{d}\theta$
Curve Length	$\int_a^b \sqrt{1+[f'(x)]^2}\,\mathrm{d}x$	$\int_{T_1}^{T_2} \sqrt{[x'(t)]^2+[y'(t)]^2}\,\mathrm{d}t$	$\int_\alpha^\beta \sqrt{r^2(\theta)+r'^2(\theta)}\,\mathrm{d}\theta$
Volume of Solid of Revolution	$\pi\int_a^b [f(x)]^2\,\mathrm{d}x$	$\pi\int_{T_1}^{T_2} y^2(t)x'(t)\,\mathrm{d}t$	$\frac{2}{3}\pi\int_\alpha^\beta r^3(\theta)\sin\theta\,\mathrm{d}\theta$
Surface Area of Solid of Revolution	$2\pi\int_a^b f(x)\sqrt{1+[f'(x)]^2}\,\mathrm{d}x$	$2\pi\int_{T_1}^{T_2} y(t)\sqrt{[x'(t)]^2+[y'(t)]^2}\,\mathrm{d}t$	$2\pi\int_\alpha^\beta r(\theta)\sin\theta\sqrt{r^2(\theta)+r'^2(\theta)}\,\mathrm{d}\theta$

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 $f(x) \leqslant K g(x)$ (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:

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

Then:

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

¶ `Examples`

Example 7.1 Discuss the convergence of the following improper integrals:

1. $$\int_0^{+\infty} \frac{\sin x}{x^p}\,\mathrm{d}x$$

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

3. $$\int_0^{1} \frac{1}{x^p \ln x}\,\mathrm{d}x$$

4. $$\int_0^{+\infty} \frac{1}{x^p} \sin \frac{1}{x}\,\mathrm{d}x$$

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

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

When $a=6, b=2$, the figure is shown as Fig 7.1.

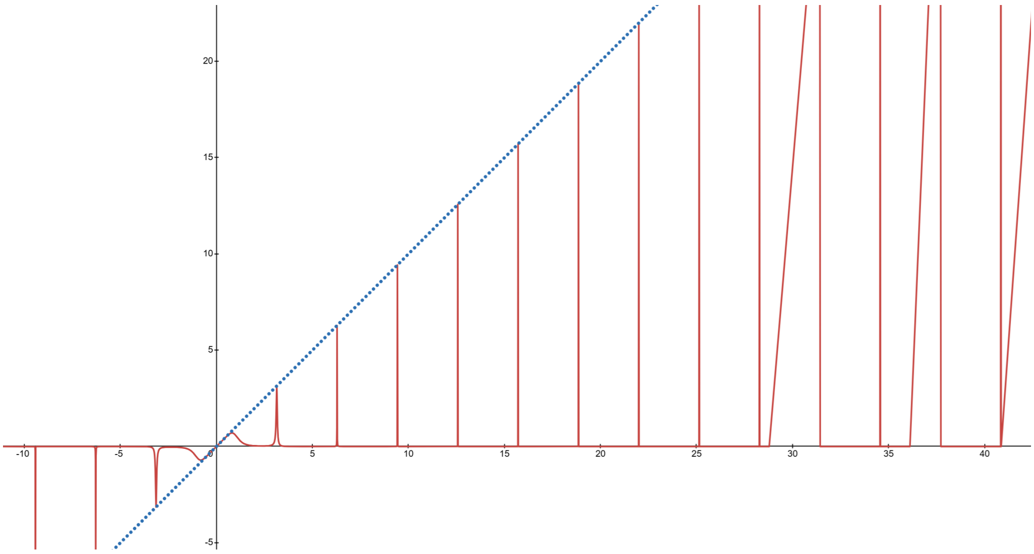


Figure 7.1: Graph of $y=\frac{1}{1+x^6\sin^2 x}$

7.4 Common Questions

¶ **Square Integrable**

Definition 7.2 (Square Integrable Function)

If $f(x) \in R[a,+\infty)$ and $\int_a^{+\infty}[f(x)]^2\,\mathrm{d}x$ converges, then $f(x)$ is called a **square integrable function** on $[a,+\infty)$. For defective integrals, the definition is similar. ♣

Property

¶ **Properties of the Integrand of the Convergent Infinite Integral at Infinity**

For the infinite integral

$$\int_0^{+\infty} \frac{1}{1 + x^6 \sin^2 x} \, \mathrm{d}x,$$

whose integrand is shown in Fig 7.1, we can deduce that even if the integral converges, $f(+\infty)$ is not necessarily equal to 0. Moreover, it is possible that $\overline{\lim}_{x\to+\infty} f(x) = +\infty$.

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 $\varliminf_{n\to+\infty} n\left(\frac{a_n}{a_{n+1}}-1\right)=l>1$, the series converges.
2. If $\varlimsup_{n\to+\infty} n\left(\frac{a_n}{a_{n+1}}-1\right)=l'<1$, 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 $\varliminf_{n\to+\infty} \ln n\left[n\left(\frac{a_n}{a_{n+1}}-1\right)\right]=l>1$, the series converges.
2. If $\varlimsup_{n\to+\infty} \ln n\left[n\left(\frac{a_n}{a_{n+1}}-1\right)\right]=l'<1$, 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

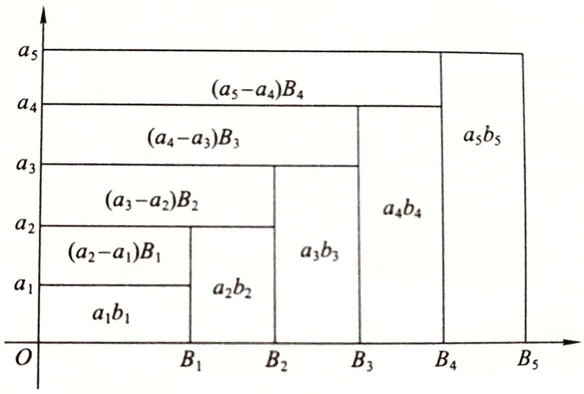
Theorem 8.7 (Abel Transform (Discrete Integration by Parts/Summation by Parts))

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

$$\sum_{k=1}^{5} a_k b_k = a_5 B_5 - \sum_{k=1}^{4} (a_{k+1} - a_k) B_k.$$



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 $p \in \mathbb{N}^+$,

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

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

¶ `Infinite Products`

¶ `Two Formulas`

Theorem 8.11 (Wallis Formula)

$$\lim_{n\to\infty} \frac{1}{2n+1}\left[\frac{(2n)!!}{(2n-1)!!}\right]^2 = \frac{\pi}{2}.$$

Equivalently $(n \to +\infty)$,

$$\frac{(2n)!!}{(2n-1)!!} \sim \sqrt{\pi n},$$
$$\frac{(n!)^2 2^{2n}}{(2n)!} \sim \sqrt{\pi n}.$$

♡

Theorem 8.12 (Stirling Formula)

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} - \frac{1}{288n^2} + \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \cdots + \frac{B_{2n}}{2k(2k-1)n^k} + \cdots\right),$$

where B_{2k} are Bernoulli numbers of order $2k$. Simplified form:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \quad (n \to +\infty),$$

or

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\theta_n}, \quad \frac{1}{12n+1} < \theta_n < \frac{1}{12n}.$$

♡

8.7 Special Series

Geometric Series

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q},$$

it converges when $|q| < 1$, diverges otherwise.

Telescoping Series

$$\sum_{n=1}^{\infty}(a_n - a_{n+1}) = a_1 - \lim_{n\to\infty} a_{n+1},$$

it converges when $\lim_{n\to\infty} a_n$ exists, diverges otherwise.

p-Series/Hyperharmonic Series

$$\sum_{n=1}^{\infty} \frac{1}{n^p},$$

it converges when $p > 1$, diverges otherwise.

q-Series

$$\sum_{n=1}^{\infty} \frac{1}{n(\ln n)^q},$$

it converges when $q > 1$, diverges otherwise.

Generalized q-Series

$$\sum_{n=3}^{\infty} \frac{1}{n \ln n (\ln\ln n) \cdots (\ln^{(k-1)} n)(\ln^{(k)} n)^q},$$

where $\ln^{(k)} n$ denotes the k-th iterated logarithm, it converges when $q > 1$, diverges otherwise.

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 Conditions 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*

$$f_n(x) \overset{[a,b]}{\rightrightarrows} f(x) \text{ and } f'(x) = \sigma(x).$$

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*

$$\sum_{n=1}^{\infty} u_n(x) \overset{[a,b]}{\rightrightarrows} S(x) \text{ and } S'(x) = \sigma(x).$$

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)

Theorem 9.3 (Weierstrass Test (M-Test))

If there exists a convergent positive term series $\sum_{n=1}^{\infty} a_n$ such that:

$$|u_n(x)| \leqslant a_n, \quad \forall x \in E, n = 1, 2, 3, \ldots$$

then the function term series $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly on E.

The positive term series $\sum_{n=1}^{\infty} a_n$ is called a **majorant series** of $\sum_{n=1}^{\infty} u_n(x)$.

If replace the convergent positive term series $\sum_{n=1}^{\infty} a_n$ with a uniform convergent series of functions $\sum_{n=1}^{\infty} a_n(x)$, the conclusion still holds.

¶ Abel-Dirichlet Test

Theorem 9.4 (Abel-Dirichlet Test)

If the series of functions $\sum_{n=1}^{\infty} a_n(x) b_n(x)$ $(x \in E)$ satisfies at least one of the following two conditions, then it converges uniformly on E.

Abel $\{a_n(x_0)\}$ $(\forall x_0 \in E)$ is monotonic and the series of functions $\{a_n(x)\}$ is bounded uniformly on E. Simultaneously, the series $\sum_{n=1}^{\infty} b_n(x)$ converges uniformly on E.

Dirichlet $\{a_n(x_0)\}$ $(\forall x_0 \in E)$ is a monotonic and and $a_n(x) \to 0$ uniformly convergent on E with limit 0. Simultaneously, the partial sums $B_n(x) = \sum_{k=1}^{n} b_k(x)$ are uniformly bounded on E.

¶ Dini Theorem

Theorem 9.5 (Dini Theorem)

Let the sequence of functions $\{S_n(x)\}$ converges pointwise to $S(x)$ on the closed interval $[a,b]$, if

1. $S_n(x) \in C[a,b]$ $(n = 1, 2, 3, \ldots)$;
2. $S(x) \in C[a,b]$;
3. $\{S_n(x_0)\}$ $(\forall x_0 \in [a,b])$ is monotonic;

then $S_n(x) \overset{[a,b]}{\rightrightarrows} S(x)$.

Remark Removing the condition of monotonicity, the Arzelà-Borel theorem (**??**) becomes the result of quasi-uniform convergence.

9.3 Special Cases

Chapter 10 Power Series

10.1 Power Series and Its Convergence Radius

10.2 Expanding Functions into Power Series

10.3 Smooth Appropriation of Functions

First, we use continuous functions to approximate Riemann integrable functions and smooth functions to approximate continuous functions, respectively.

Theorem 10.1

Let $f(x) \in R[a,b]$. For any $\varepsilon > 0$, there exists a function $g(x) \in C[a,b]$ such that:

$$\int_a^b |f(x) - g(x)| < \varepsilon.$$

Theorem 10.2

Let $f(x) \in C[a,b]$. For any $\varepsilon > 0$, there exists a function $g(x) \in C^\infty[a,b]$ such that:

$$|f(x) - g(x)| < \varepsilon, \quad \forall x \in [a,b].$$

Then, Weierstrass approximation theorem is stated as follows:

Theorem 10.3 (Weierstrass First Approximation Theorem)

Let $f(x) \in C[a,b]$. For any $\varepsilon > 0$, there exists a polynomial $P(x)$ such that:

$$|f(x) - P(x)| < \varepsilon, \quad \forall x \in [a,b].$$

Theorem 10.4 (Weierstrass Second Approximation Theorem)

Let $f(x)$ be a continuous periodic function with period 2π. For any $\varepsilon > 0$, there exists a trigonometric polynomial sequence

$$\{T_n(x) = \frac{A_0}{2} + \sum_{k=1}^{n} A_k \cos(kx) + B_k \sin(kx)\}$$

such that:

$$T_n(x) \rightrightarrows f(x).$$

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 $\mathbf{e}_i = \sum_{i=0}^{n} \mathbf{e}_i \cos\alpha_i$ be a unit vector, where $\sum_{i=0}^{n} \cos^2\alpha_i = 1$. 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

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

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:

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

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

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

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

¶ Total Differential

Definition 12.3 (Total Differential)

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

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

where $A_1, A_2, \dots, A_n$ are constants independent of $\boldsymbol{\Delta}\mathbf{x}$, 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

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

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 $i = 1, 2, \dots, n$, and[a]

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

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 $Jf = (\mathrm{D}_1 f, \mathrm{D}_2 f, \dots, \mathrm{D}_n f)$ is continuous at $\mathbf{x}^0$ (i.e., $\frac{\partial f}{\partial x_i}$ is continuous at $\mathbf{x}^0$ for $i = 1, 2, \dots, n$), then f is differentiable at $\mathbf{x}^0$[b].

However, the converse is not necessarily true.

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

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

[b] 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.

Note *The continuity of the derivative function at* $\mathbf{x}^0$ *implies that the original function* f *is differentiable in some neighborhood of* $\mathbf{x}^0$.

Proof Taking the 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. *The existence of partial derivatives at a point does not necessarily imply their continuity at that point. A classic counterexample is:*

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

Here, $f_x(0, 0) = 0$ and $f_y(0, 0) = 0$, but $f_x(x, y)$ and $f_y(x, y)$ are not continuous at $(0, 0)$.

2. *(partial derivatives bounded $\Rightarrow$ continuous) If the partial derivatives exist and are bounded in a neighborhood of a point, then they are continuous at that point.*
3. *Even if all directional derivatives exist at a point and the function is continuous at that point, it does not necessarily imply that the function is differentiable at that point. A classic counterexample is:*

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

Here, all directional derivatives of f exist at $(0, 0)$, and f is continuous at $(0, 0)$, but f is not differentiable at $(0, 0)$. Another counterexample is:

$$f(x, y) = \sqrt{|xy|},$$

which is continuous at $(0, 0)$ and has all directional derivatives equal to 0 at $(0, 0)$, but is not differentiable at $(0, 0)$.

Proof Take the function of two variables as an example. Assume the bivariate function $f : \mathbb{R}^2 \to \mathbb{R}$ meets: $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist and are bounded in some neighborhood of (x_0, y_0).

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

$$\begin{aligned}\Delta f &= [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)] \\ &\quad + [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] .\end{aligned}$$

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

$$\Delta f = f_x(x_0 + \theta_1 \Delta x, y_0 + \Delta y)\Delta x + f_y(x_0, y_0 + \theta_2 \Delta y)\Delta y.$$

Since $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are bounded in some neighborhood of (x_0, y_0),

$$\lim_{(\Delta x, \Delta y)\to(0,0)} \Delta f = 0,$$

i.e. $f(x, y)$ is continuous at (x_0, y_0), which completes the proof. ■

12.2 Higher-Order Partial Derivatives and Differentiability

¶ Higher-Order Partial Derivatives

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 the partial derivatives f_x, f_y and f_{xy} exist in some neighborhood of $(x_0, y_0) \in U$, and f_{xy} is continuous at (x_0, y_0), then f_{yx} also exists at (x_0, y_0), and
$$f_{yx}(x_0, y_0) = f_{xy}(x_0, y_0).$$
2. Let $U \subset \mathbb{R}^n$ be an open set, and $f : U \to \mathbb{R}$ be a function of n variables. If the partial derivatives f_{x_i}, f_{x_j} and $f_{x_i x_j}$ exist in some neighborhood of $\mathbf{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0) \in U$, and $f_{x_i x_j}$ is continuous at $\mathbf{x}^0$, then $f_{x_j x_i}$ exist at $\mathbf{x}^0$, and
$$f_{x_j x_i}(\mathbf{x}^0) = f_{x_i x_j}(\mathbf{x}^0).$$

♡

Proof

■

¶ Higher-Order Differentiability

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

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

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:

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

12.3 Differential of Vector-Valued Functions

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

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

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 $\mathbf{f}$ 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 $\mathbf{f}$ 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 $\mathbf{f}$ is differentiable at every point in U, then $\mathbf{f}$ 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 $\mathbf{f}$ 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 $\mathbf{f}$ has continuous partial derivatives at $\mathbf{x}^0$, then every element of the Jacobian matrix of $\mathbf{f}$ is continuous at $\mathbf{x}^0$. In this case, $\mathbf{f}$ is said to have a continuous derivative at $\mathbf{x}^0$ as a vector-valued function.
   If the derivative of a vector-valued function $\mathbf{f}$ is continuous at every point in U, then $\mathbf{f}$ 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 coor-

dinate component functions $f_i(x_1, x_2, \ldots, x_n)$ $(i = 1, 2, \ldots, m)$ is continuous, differentiable, and has derivatives.

12.4 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*

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

■

Applications

As an important application of the chain rule, we have the following theorem on the differentiation of determinants.

Theorem 12.3

For

$$\Delta(t) = \begin{vmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{vmatrix},$$

where each element $a_{ij}(t)$ is differentiable with respect to t, then $\Delta(t)$ is differentiable with respect to t, and

$$\frac{\mathrm{d}\Delta(t)}{\mathrm{d}t} = \sum_{j=1}^{n} \begin{vmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{d}}{\mathrm{d}t}a_{1j}(t) & \frac{\mathrm{d}}{\mathrm{d}t}a_{2j}(t) & \cdots & \frac{\mathrm{d}}{\mathrm{d}t}a_{nj}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{vmatrix}$$

where in each determinant on the right-hand side, the j-th column is replaced by the derivative of the j-th column of $\Delta(t)$.

Another important application is homogeneous functions.

Proposition 12.1

The following statements can be generalized for n variables:

1. Let $f(x,y) \in C^1$, then f is a homogeneous function of degree m if and only if

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = mf(x,y).$$

2. Let $f(x,y) \in C^2$ be a homogeneous function of degree m, then

$$\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f(x,y) = m(m-1)f(x,y),$$

where

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

which is just a formal notation, not an operator multiplication.

3. Let $f(x,y) \in C^2$ be a homogeneous function of degree m, then $f_x(x,y), f_y(x,y)$ are homogeneous functions of degree $m-1$.
4. Let $f(x,y) \in C(\mathbb{R}^2 \setminus \{(0,0)\})$ be a homogeneous function of degree m, then

$$|f(x,y)| \leqslant C\rho^m, \quad \rho = \sqrt{x^2+y^2},$$

where $C = \max_{\rho=1} |f(x,y)|$.

Example 12.1 Let $f(x,y)$ be a differential function on $\mathbb{R}^2$, and satisfy the equation

$$x\frac{\mathrm{d}f}{\mathrm{d}x} + y\frac{\mathrm{d}f}{\mathrm{d}y} = 0,$$

prove that $f(x,y)$ is always constant.

12.5 Mean Value Theorem and Taylor's Formula

¶ **Mean Value Theorem**

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 12.4 (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

■

For mappings, Lagrange's mean value theorem can not be generalized directly, we need introduce inner product:

Theorem 12.5 (Lagrange's Mean Value Theorem for Mappings)

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

$$\mathbf{a} \cdot [\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})] = \mathbf{a} \cdot [J\mathbf{f}(\xi)(\mathbf{b} - \mathbf{a})], \quad \forall \mathbf{a} \in \mathbb{R}^m.$$

♡

Note *If it does not contain the inner product, then it is not necessarily true. For example, let*

$$\mathbf{f}(t) = (\cos t, \sin t), \quad t \in [0, 2\pi],$$

then

$$J\mathbf{f}(t) = (-\sin t, \cos t),$$

note that $\mathbf{f}(2\pi) = \mathbf{f}(0)$, then there does not exist $\theta \in (0,1)$ such that

$$\mathbf{f}(2\pi) - \mathbf{f}(0) = J\mathbf{f}(\theta \cdot 2\pi)(2\pi - 0).$$

In fact,

$$J\mathbf{f}(t) \not\equiv 0, \quad \forall t \in [0, 2\pi].$$

And we have global estimation for the difference of mappings:

Theorem 12.6 (Quasi-Differential Mean Value Theorem for Mappings)

Let $\mathbf{f} : D \to \mathbb{R}^m$ 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:

$$\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\| \leq \|J\mathbf{f}(\xi)\| \cdot \|\mathbf{b} - \mathbf{a}\|.$$

Corollary 12.1

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

$$J\mathbf{f}(\mathbf{x}) = 0,$$

then $\mathbf{f}$ is constant mapping on D.

Proof

■

¶ Taylor's Formula

Theorem 12.7 (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.

12.6 Implicit Function Theorem

¶ Implicit Mapping

Theorem 12.8 (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 $1 \leqslant k \leqslant +\infty$;
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 $y = f(x)$. ♡

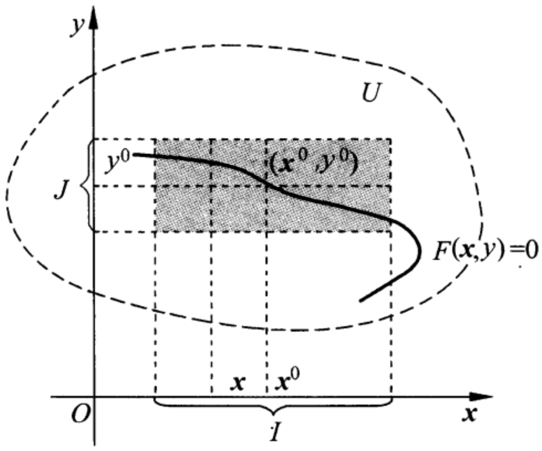


Figure 12.1: Implicit Function

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:

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

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 12.4, 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. ■

Note *From the proof process of the implicit function theorem, it can be observed that if only require the continuity of the implicit function $y = f(x)$, then the theorem can be restated as follows:*

If

1. $F \in C(U, \mathbb{R})$;
2. $F(\mathbf{x}^0, y^0) = 0$;
3. *For fixed $\mathbf{x} = \mathbf{x}^0$, $F(\mathbf{x}^0, y)$ is strictly monotonic with respect to y.*

Then the function derived from the implicit function $F(\mathbf{x}, y) = 0$, i.e., $y = f(\mathbf{x})$, is continuous at I.

Theorem 12.9 (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. $\mathbf{F} \in C^k(U, \mathbb{R}^m), 1 \leqslant k \leqslant \infty$;
2. $\mathbf{F}(\mathbf{x}^0, \mathbf{y}^0) = 0$, where $\mathbf{x}^0 = (x_1, x_2, \ldots, x_n)$, $\mathbf{y}^0 = (y_1, y_2, \ldots, y_m)$, $(\mathbf{x}^0, \mathbf{y}^0) \in U$ (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 $x \in I$,

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

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

Example 12.2

$$\begin{cases} x = x(z), \\ y = y(z), \end{cases}$$

is an mapping solved from the implicit function defined by the equations:

$$\begin{cases} F(y - z, x + z) = 0, \\ G(\frac{y}{z}, xz) = 0, \end{cases}$$

where $F, G \in C^1$. Find $\frac{\mathrm{d}x}{\mathrm{d}z}$ and $\frac{\mathrm{d}y}{\mathrm{d}z}$.

Remark Here, we use F_1 to represent the partial derivative of F with respect to its first variable, which is equivalent to F_u in $F(u, v)$. Other notations follow similarly.

Solution

Method 1: Direct Derivative Derivative both sides of the equations with respect to z:

$$F_1(y' - 1) + F_2(x' + 1) = 0,$$
$$G_1(\frac{y'z - y}{z^2}) + G_2(x'z + x) = 0.$$

Solve the above equations to get:

$$\frac{\mathrm{d}x}{\mathrm{d}z} = \frac{zG_1(F_1 - F_2) - F_1(yG_1 - xz^2G_2)}{z(F_2G_1F_1G_2z^2)},$$
$$\frac{\mathrm{d}y}{\mathrm{d}z} = \frac{F_2(yG_1 - xz^2G_2) - G_2z^3(F_1 - F_2)}{z(F_2G_1 - F_1G_2z^2)}.$$

Method 2: Implicit Function Theorem By the implicit function theorem, we have:

$$\begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}z} \\ \frac{\mathrm{d}y}{\mathrm{d}z} \end{pmatrix} = -\begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F}{\partial z} \\ \frac{\partial G}{\partial z} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{zG_1(F_1 - F_2) - F_1(yG_1 - xz^2G_2)}{z(F_2G_1 - F_1G_2z^2)} \\ \frac{F_2(yG_1 - xz^2G_2) - G_2z^3(F_1 - F_2)}{z(F_2G_1 - F_1G_2z^2)} \end{pmatrix}.$$

□

Example 12.3 Let $u(x,y)$ is a function solved from the implicit function defined by the equation:

$$\begin{cases} u = f(x,y,z,t), \\ g(y,z,t) = 0, \\ h(z,t) = 0, \end{cases}$$

where $f, g, h \in C^1$ and $\frac{\partial(g,h)}{\partial(z,t)} \neq 0$. Find $\frac{\partial u}{\partial y}$.

Solution

Method 1. Since $\frac{\partial(g,h)}{\partial(z,t)} \neq 0$, and $g, h \in C^1$, $g(y,z,t) = 0, h(z,t) = 0$, by the implicit mapping theorem 12.9, we can express z and t as functions of y:

$$\begin{cases} z = z(y), \\ t = t(y). \end{cases}$$

Derivative both sides with respect to y:

$$g_y + g_z \frac{\mathrm{d}z}{\mathrm{d}y} + g_t \frac{\mathrm{d}t}{\mathrm{d}y} = 0,$$
$$h_z \frac{\mathrm{d}z}{\mathrm{d}y} + h_t \frac{\mathrm{d}t}{\mathrm{d}y} = 0.$$

And u is a function of x and y: $u = u(x,y) = f(x,y,z(y),t(y))$. Thus:

$$\frac{\partial u}{\partial y} = f_2 + f_3 \frac{\mathrm{d}z}{\mathrm{d}y} + f_4 \frac{\mathrm{d}t}{\mathrm{d}y}.$$

Solve the above equations to get:

$$\frac{\partial u}{\partial y} = f_y - g_y(f_z h_t - f_t h_z) \left(\frac{\partial(g,h)}{\partial(z,t)} \right)^{-1}.$$

Method 2. Considering

$$\begin{cases} F(x,y,z,t,u) = u - f(x,y,z,t) = 0, \\ g(y,z,t) = 0, \\ h(z,t) = 0. \end{cases}$$

Since $\frac{\partial(F,g,h)}{\partial(u,z,t)} = \frac{\partial(g,h)}{\partial(z,t)} \neq 0$, by the implicit mapping theorem 12.9, we have

$$\begin{cases} u = u(x,y), \\ z = z(x,y), \\ t = t(x,y). \end{cases}$$

Derivative both sides with respect to y:

$$u_y - f_y - f_z z_y - f_t t_y = 0,$$
$$g_y + g_z z_y + g_t t_y = 0,$$
$$h_z z_y + h_t t_y = 0.$$

Solve the above equations to get the same result. □

¶ **Inverse Mapping**

Theorem 12.10 (Local Inverse Mapping Theorem)

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

1. $\mathbf{f} \in C^k(U, \mathbb{R}^n), 1 \leqslant k \leqslant +\infty$;
2. At point $\mathbf{x}^0 \in U$, the Jacobian determinant

$$\det J\mathbf{f}(\mathbf{x}^0) \neq 0.$$

Then there exist open neighborhoods $V \subset U$ of $\mathbf{x}^0$ and $W \subset \mathbb{R}^n$ of $\mathbf{f}(\mathbf{x}^0 = \mathbf{y}^0)$, such that:

1. The restriction of $\mathbf{f}$ to V, denoted as $\mathbf{f}|_V : V \to W$, is a bijection;
2. The inverse mapping $\mathbf{f}^{-1} : W \to V$ exists and belongs to $C^k(W, \mathbb{R}^n)$;
3. For any $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in W$,

$$J\mathbf{f}^{-1}(\mathbf{y}) = [J\mathbf{f}(\mathbf{x})]^{-1},$$

   where $\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y})$.

At this time, $\mathbf{f}$ is called a C^k diffeomorphism. ♡

If the conditions are strengthened, then a global inverse mapping theorem can be established.

Theorem 12.11 (Inverse Mapping Theorem)

Let $U \subset \mathbb{R}^n$ be a convex region, and $\mathbf{f} : U \to \mathbb{R}^n$ be a mapping. If:

1. $\mathbf{f} \in C^k(U, \mathbb{R}^n), 1 \leqslant k \leqslant +\infty$;
2. For any $\mathbf{x} \in U$, the Jacobian determinant

$$\det J\mathbf{f}(\mathbf{x}) \neq 0.$$

Then $\mathbf{f} : U \to \mathbf{f}(U)$ is a bijection, and the inverse mapping $\mathbf{f}^{-1} : \mathbf{f}(U) \to U$ exists and belongs to $C^k(\mathbf{f}(U), \mathbb{R}^n)$. ♡

Example 12.4 Here are substitutions:

$$x = t, y = \frac{t}{1+tu}, z = \frac{t}{1+tv}.$$

Transform the following equation to the form of dependent variables v and independent variables t, u:

$$x^2 \frac{\partial z}{\partial x} + y^2 \frac{\partial z}{\partial y} = z^2.$$

12.7 Extremum of Multi-variable Functions

¶ Unconditional Extremum

Proposition 12.2 (Fermat's Three Villiges Problem)

There are three villages located at points A, B, and C on a flat plane. A supply station needs to be established at point P on the plane, such that the total distance from P to the three villages A, B, and C is minimized. Such a point P is called the Fermat point of triangle ABC, which can be determined as follows:

1. If any angle of triangle ABC is greater than or equal to $120°$, then the Fermat point is the vertex of that angle.
2. If all angles of triangle ABC are less than $120°$, then the Fermat point P is located inside triangle ABC, and the angles between the segments PA, PB, and PC are all equal to $120°$.

♠

¶ Conditional Extremum

Definition 12.5 (Conditional Extremum)

Let $f : D \to \mathbb{R}$ be a function with $n + m$ variables defined on a open set $D \subseteq \mathbb{R}^{n+m}$, and let $\mathbf{\Phi} : D \to \mathbb{R}^m$ be a mapping, $M = \{\mathbf{x} \in D \mid \mathbf{\Phi}(\mathbf{x}) = 0\}$. If there exists $\mathbf{x}^0 \in M$ satisfying the constraints such that:

$$f(\mathbf{x}^0) \leq f(\mathbf{x}) \quad (\text{or } f(\mathbf{x}^0) \geq f(\mathbf{x})),$$

for all $\mathbf{x} \in M$ that also satisfy the constraints, then f is said to have a conditional minimum (or maximum) at point $\mathbf{x}^0$ under the given constraints.

♣

Theorem 12.12 (Lagrange Multiplier Method)

Let $f : D \to \mathbb{R}$ be a function with $n + m$ variables defined on an open set $D \subseteq \mathbb{R}^{n+m}$, and let $\mathbf{\Phi} : D \to \mathbb{R}^m$ be a mapping, $M = \{\mathbf{x} \in D \mid \mathbf{\Phi}(\mathbf{x}) = 0\}$. If:

1. $f \in C^1(D, \mathbb{R})$, $\mathbf{\Phi} \in C^1(D, \mathbb{R}^m)$;
2. $\operatorname{rank}(J\mathbf{\Phi}(\mathbf{x}^0)) = m$;
3. $\mathbf{x}^0$ is a conditional extremum point of f on M;

then there exist $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R}$, such that:

$$\nabla f(\mathbf{x}^0) + \sum_{i=1}^{m} \lambda_i \nabla \Phi_i(\mathbf{x}^0) = 0.$$

♡

Chapter 13 Multiple Integrals

13.1 Multiple Integrals on Bounded Closed Regions

How to define a region with measurable area? Generally speaking, there are two approaches to define regions with measurable area:

1. Consider the integral over a closed rectangle, and then extend it to a bounded closed region within the rectangle with the help of characteristic functions;
2. Define that a bounded closed region D is measurable if $\forall \varepsilon > 0$, there exist two polygonal regions Σ_1 and Σ_2 consisting of finite rectangles, such that $\Sigma_1 \subset D \subset \Sigma_2$ and the area of $\Sigma_2 \setminus \Sigma_1$ is less than ε.

¶ **Definition of Multiple Integral**

Here, we introduce the definition of double integrals using the first approach.

Initially, we define the double integral on a closed interval (rectangle).

Definition 13.1 (Double Integral on a Closed Interval)

Let $I = [a, b] \times [c, d]$ be a closed interval in $\mathbb{R}^2$, (i.e., each boundary is parallel to the coordinate axes). Partition $[a, b]$:

$$T_x : a = x_0 < x_1 < \cdots < x_n = b.$$

Partition $[c, d]$:

$$T_y : c = y_0 < y_1 < \cdots < y_m = d.$$

Two sets of parallel lines $x = x_i \, (i = 0, 1, \ldots, n)$ and $y = y_j \, (j = 0, 1, \ldots, m)$ divide I into $n \times m$ subrectangles:

$$[x_{i-1}, x_i] \times [y_{j-1}, y_j], \quad i = 1, \ldots, n, \; j = 1, \ldots, m.$$

The union of these k subrectangles forms a partition $T = T_x \times T_y = \{I_1, I_2, \ldots, I_k\}$. For each $\xi^i \in I_i \, (i = 1, 2, \ldots, k)$, define the **Riemann sum** (also called a sum of integrals) as:

$$\sum_{i=1}^{k} f(\boldsymbol{\xi}^i) v(I_i),$$

where $v(I_i)$ is the area of the rectangle I_i, i.e., the product of its length and width. Denote:

$$\lambda = \max(\operatorname{diam}(I_1), \operatorname{diam}(I_2), \ldots, \operatorname{diam}(I_k)),$$

where $\operatorname{diam}(I)$ is the diagonal length of the rectangle I, and λ is called the modulus or width of the partition T. The points $\boldsymbol{\xi} = (\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \ldots, \boldsymbol{\xi}^k) \in I_1 \times I_2 \times \cdots \times I_k$ are called sampling points for the Riemann sum.

If there exists $J \in \mathbb{R}$, such that $\forall \varepsilon > 0$, there exists $\delta > 0$, such that when $\lambda < \delta$, for all $\boldsymbol{\xi} \in I_1 \times I_2 \times \cdots \times I_k$, we have:

$$\left| \sum_{i=1}^{k} f(\boldsymbol{\xi}^i) v(I_i) - J \right| < \varepsilon,$$

then f is said to be Riemann integrable on I, and:

$$J = \lim_{\lambda \to 0} \sum_{i=1}^{k} f(\boldsymbol{\xi}^i) v(I_i) =: \iint_I f(x,y)\,\mathrm{d}x\mathrm{d}y \quad \text{or} \quad \int_I f\,\mathrm{d}v \quad \text{or} \quad \int_I f.$$

The function f is said to have a double integral on I, or simply f is integrable on I. Here f is called the integrand, I is called the integration region, and $\mathrm{d}v = \mathrm{d}x\mathrm{d}y$ is called the integration element. ♣

The defined double integral possesses properties similar to those of single-variable integrals.

On the basis of the above definition, we can extend it to the case of a bounded set.

Definition 13.2 (Double Integral on a Bounded Set)

Let $\Omega \subset \mathbb{R}^2$ be a bounded set, and $f : \Omega \to \mathbb{R}$ a two-dimensional function. Define:

$$f_\Omega(\mathbf{x}) = f_\Omega(x,y) = \begin{cases} f(x,y), & \text{if } \mathbf{x} = (x,y) \in \Omega, \\ 0, & \text{if } \mathbf{x} = (x,y) \notin \Omega, \end{cases}$$

and call this the **zero extension** (or **characteristic function**) of f. For any closed interval $I \supset \Omega$, if f_Ω is Riemann integrable on I, then f is said to be **Riemann integrable** on Ω (abbreviated as integrable). The integral of f on Ω, denoted as:

$$\iint_\Omega f(x,y)\,\mathrm{d}x\mathrm{d}y = \int_\Omega f\,\mathrm{dV} = \int_\Omega f = \int_\Omega f_\Omega = \iint_I f_\Omega(x,y)\,\mathrm{d}x\mathrm{d}y,$$

represents the Riemann integral of f on Ω. ♣

In above definition, the integral $\int_\Omega f$ is independent of the choice of the closed interval I containing Ω (this confirms the consistency of the definition).

It is worth noting that all the definitions and properties of double integrals can be extended to triple integrals and higher-dimensional integrals without excessive inconvenience.

¶ `About the Second Approach`

Definition 13.3 (Set with Zero Area and Set with Zero Measure (Null Set))

Let $A \subset \mathbb{R}^2$. If for any $\varepsilon > 0$, there exist finitely many closed intervals $I_1, I_2, \ldots, I_k$ such that:

$$\bigcup_{i=1}^{k} I_i \supset A, \quad \text{and} \quad \sum_{i=1}^{k} v(I_i) < \varepsilon,$$

then A is called a **set with zero area.**

Let $A \subset \mathbb{R}^2$. If for any $\varepsilon > 0$, there exist at most countably many closed intervals $I_1, I_2, \ldots, I_k, \ldots$ such that:

$$\bigcup_{i=1}^{\infty} I_i \supset A, \quad \text{and} \quad \sum_{i=1}^{\infty} v(I_i) < \varepsilon,$$

then A is called a **set with zero measure** (**null set**). ♣

Definition 13.4 (Set with Finite Area)

Let $\Omega \subset \mathbb{R}^2$ be a bounded set. If the constant function 1 is integrable on Ω, then Ω is called a **set with finite area**, and the area of Ω is defined as:

$$v(\Omega) = \int_{\Omega} 1 = \iint_{\Omega} \mathrm{d}x\mathrm{d}y = \int_I 1_{\Omega}.$$

♣

Obviously, Ω is a set with zero area if and only if Ω has finite area, and $v(\Omega) = \int_{\Omega} 1 = 0$.

Proposition 13.1

A bounded closed region $\Omega \subset \mathbb{R}^2$ is measurable if and only if its boundary $\partial\Omega$ is a set with zero area.

♠

In the definition of multiple integrals derived from the second approach, the key point is the division T of the bounded closed region Ω into two polygonal regions Σ_1 and Σ_2. With above statements, we can see that the division T is implemented by infinitely many curves net with zero area.

¶ **Necessary and Sufficient Conditions for Integrability**

Proposition 13.2

Let non-negative function $f \in R(D)$, then $\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = 0$ if and only if for any continuous points $(x,y) \in D$, $f(x,y) = 0$.

♠

13.2 Properties of Multiple Integrals

¶ **Reduction of Double Integral to Iterated Integral**

Theorem 13.1 (Reduction of Double Integral to Iterated Integral on a Closed Interval)

Let f be integrable on the closed interval $I = [a,b] \times [c,d]$.
If $\forall x \in [a,b]$, the integral $\phi(x) = \int_c^d f(x,y)\,\mathrm{d}y$ exists, then ϕ is integrable on $[a,b]$, and:

$$\iint_I f = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x =: \int_a^b \mathrm{d}x \int_c^d f(x,y)\,\mathrm{d}y.$$

Similarly, if $\forall y \in [c,d]$, the integral $\psi(y) = \int_a^b f(x,y)\,\mathrm{d}x$ exists, then ψ is integrable on $[c,d]$, and:

$$\iint_I f = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y =: \int_c^d \mathrm{d}y \int_a^b f(x,y)\,\mathrm{d}x.$$

♡

Note *That is, if $f \in C(I)$, then two iterated integrals above all exist, and they are equal to the double integral of f on I (they can exchange the order of integration).*

On the basis of the above theorem, we can extend it to the case of a bounded region.

Theorem 13.2 (Reduction of Double Integral to Iterated Integral on a Bounded Set)

Let $\Omega \subset \mathbb{R}^2$ be a set with infinite area, and $f : \Omega \to \mathbb{R}$ be bounded and continuous (13.1). Denote the vertical projection of Ω onto the x-axis as:

$$I = \{x \in \mathbb{R} \mid \exists y, \text{ s.t. } (x,y) \in \Omega\}.$$

If $\forall x \in I$, let $\Omega_x = \{y \in \mathbb{R} \mid (x, y) \in \Omega\}$ be an interval (possibly reducing to a single point), then:

$$\int_\Omega f = \int_I \mathrm{d}y \int_{\Omega_x} f(x, y)\, \mathrm{d}x.$$

Similarly, denote the vertical projection of Ω onto the y-axis as:

$$J = \{y \in \mathbb{R} \mid \exists x, \text{ s.t. } (x, y) \in \Omega\}.$$

If $\forall y \in J$, let $\Omega_y = \{x \in \mathbb{R} \mid (x, y) \in \Omega\}$ be an interval (possibly reducing to a single point), then:

$$\int_\Omega f = \int_J \mathrm{d}y \int_{\Omega_y} f(x, y)\, \mathrm{d}x.$$

♡

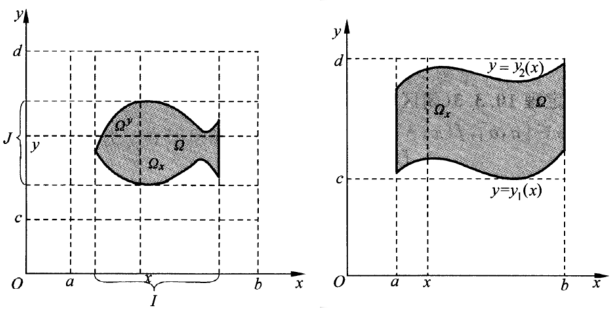


Figure 13.1: Double Integral on a Bounded Set

Specially, Let:

$$\Omega = \{(x, y) \in \mathbb{R}^2 \mid y_1(x) \leqslant y \leqslant y_2(x),\ a \leqslant x \leqslant b\},$$

where the functions y_1 and y_2 are continuous on $[a, b]$ (13.1) and the function f is integrable on Ω. If $\forall x \in [a, b]$, the single-variable integral:

$$\int_{y_1(x)}^{y_2(x)} f(x, y)\, \mathrm{d}y$$

exists, then:

$$\int_\Omega f = \int_a^b \mathrm{d}x \int_{y_1(x)}^{y_2(x)} f(x, y)\, \mathrm{d}y.$$

This area called the **type X region**, similarly, we can define the **type Y region**.

According to 13.1, we can derive the formula of multiplicative property for double integral.

Theorem 13.3 (Formula of Multiplicative Property for Double Integral)

Let $f \in C([a, b])$, $g \in C([c, d])$. Then the function $h(x, y) = f(x)g(y)$ is integrable on the closed interval $I = [a, b] \times [c, d]$, and:

$$\iint_I h(x, y)\, \mathrm{d}x\mathrm{d}y = \left(\int_a^b f(x)\, \mathrm{d}x\right)\left(\int_c^d g(y)\, \mathrm{d}y\right).$$

♡

Example 13.1 Let $p(x) \in R[a,b], p(x) > 0, x \in [a,b]$, the monotonicity of $f(x), g(x)$ is same, prove that

$$\int_a^b p(x)f(x)\mathrm{d}x \int_a^b p(x)g(x)\mathrm{d}x \leqslant \int_a^b p(x)\mathrm{d}x \int_a^b p(x)f(x)g(x)\mathrm{d}x$$

Proof Let

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

then

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

similarly,

$$I = \int_a^b \int_a^b p(x)p(y)g(x)(f(x) - f(y))\mathrm{d}x\mathrm{d}y.$$

Then

$$2I = \int_a^b \int_a^b p(x)p(y)(g(y) - g(x))(f(x) - f(y))\mathrm{d}x\mathrm{d}y \geqslant 0,$$

which implies

$$I \geqslant 0.$$

The proof is complete. ■

13.3 Calculation of Multiple Integrals

¶ Variable Substitution in Multiple Integrals

Theorem 13.4 (Variable Substitution in Double Integral)

Let $\Omega \subset \mathbb{R}^2$ be an open set, and let the mapping:

$$\mathbf{F} : \Omega \to \mathbb{R}^2, \quad (u,v) \mapsto \mathbf{F}(u,v) = (x(u,v), y(u,v))$$

satisfy the following conditions:

1. $\mathbf{F} \in C^1(\Omega, \mathbb{R}^2)$;
2. $\frac{\partial(x,y)}{\partial(u,v)} = \det J\mathbf{F}(u,v) = \det J\mathbf{F}(\mathbf{p}) \neq 0, \quad \mathbf{p} = (u,v) \in \Omega$;
3. $\mathbf{F}$ is injective.

If the set Δ is a set with finite area and $\overline{\Delta} \subset \Omega$, and f is continuous on $\mathbf{F}(\Omega)$, then $\mathbf{F}(\Delta)$ is also a set with finite area, and:

$$\iint_{\mathbf{F}(\Delta)} f = \iint_{\Delta} f \circ \mathbf{F} \left|\det J\mathbf{F}\right|,$$

i.e.,

$$\iint_{F(\Delta)} f(x,y)\,\mathrm{d}x\mathrm{d}y = \iint_{\Delta} f(x(u,v), y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \mathrm{d}u\mathrm{d}v.$$

♡

For triple and higher-dimensional integrals, the variable substitution theorem is similar to the above theorem.

Some common variable substitutions in multiple integrals are as follows:

Polar Coordinates

$$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta, \end{cases} \qquad \begin{cases} r = \sqrt{x^2+y^2}, & r \geqslant 0 \\ \theta = \arctan\left(\frac{y}{x}\right) & x \neq 0, \theta \in [0, 2\pi]. \end{cases}$$

and

$$\frac{\partial(x,y)}{\partial(r,\theta)} = r.$$

Cylindrical Coordinate System

$$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta, \\ z = z, \end{cases} \qquad \begin{cases} r = \sqrt{x^2+y^2}, & r \geqslant 0 \\ \theta = \arctan\left(\frac{y}{x}\right) & x \neq 0, \theta \in [0, 2\pi], \\ z = z. \end{cases}$$

and

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = r.$$

Spherical Coordinate System

$$\begin{cases} x = r\sin\varphi\cos\theta, \\ y = r\sin\varphi\sin\theta, \\ z = r\cos\varphi, \end{cases} \qquad \begin{cases} r = \sqrt{x^2+y^2+z^2}, & r \geqslant 0 \\ \varphi = \arccos\left(\frac{z}{r}\right) & r \neq 0, \varphi \in [0, \pi], \\ \theta = \arctan\left(\frac{y}{x}\right) & x \neq 0, \theta \in [0, 2\pi]. \end{cases}$$

and

$$\frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = r^2\sin\varphi.$$

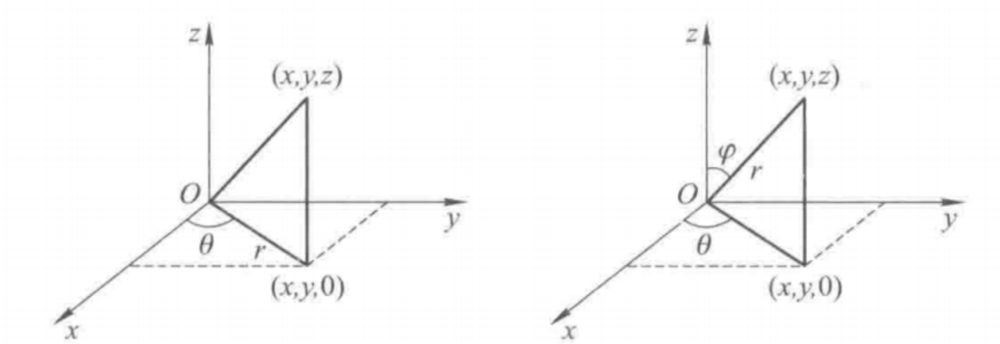


Figure 13.2: Cylindrical and Spherical Coordinate Systems

¶ Calculation of Triple Integrals

Example 13.2 Calculating $I = \iiint_\Omega z^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z$, where Ω is the cone defined by $z^2 = \frac{h^2}{R^2}(x^2+y^2)$ and $z = h$ (13.3).

Example 13.3 Calculating $I = \iiint_\Omega xy \mathrm{d}x\mathrm{d}y\mathrm{d}z$, where Ω is the region defined by $0 \leqslant z \leqslant xy, 0 \leqslant y \leqslant 1-x, 0 \leqslant x \leqslant 1$ (13.4).

With the help of examples above, we can derive **two methods for calculating triple integrals**.

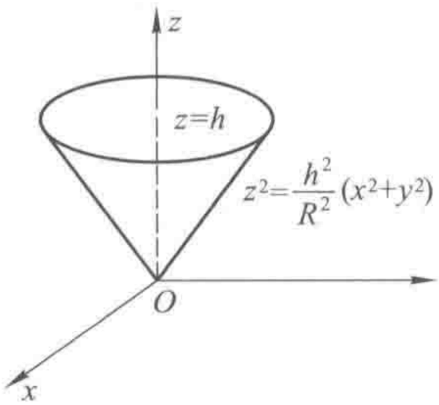


Figure 13.3: Cone Example.

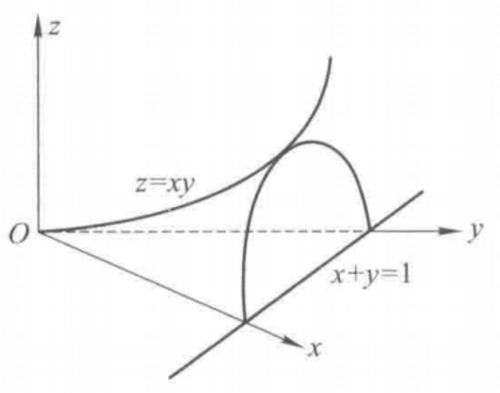


Figure 13.4: Project Method Example.

First 2 then 1 (Section Method) Fix one variable (e.g., z), first perform a double integral over the other two variables (e.g., x, y) on the "section region" corresponding to the fixed variable, and then perform a definite integral over the fixed variable (z) within its range of values.

This method is convenient when the area of the section region is easy to calculate, or when the integrand is only related to the "later-integrated variable" (e.g., only related to z).

In the example 13.2, the following steps are taken:

1. Determine the range of z: $z \in [0, h]$.
2. Determine the section region D_z: For a fixed z, D_z is the region on the xy-plane satisfying $\frac{h^2}{R^2}(x^2+y^2) \leqslant z^2$, which is a circle with radius $\frac{R}{h}z$.
3. Split the integral:
$$I = \int_0^h \left(\iint_{D_z} z^2 \, \mathrm{d}x\mathrm{d}y \right) \mathrm{d}z.$$
Since z^2 is independent of x and y, it can be factored out: $I = \int_0^h z^2 \left(\iint_{D_z} \mathrm{d}x\mathrm{d}y \right) dz$.
4. Calculate the double integral (area of the section):
$$\iint_{D_z} \mathrm{d}x\mathrm{d}y = \pi \left(\frac{R}{h} z \right)^2 = \pi \frac{R^2}{h^2} z^2.$$

5. Calculate the definite integral:

$$I = \int_0^h z^2 \cdot \pi \frac{R^2}{h^2} z^2 \, dz = \frac{\pi R^2 h^3}{5}.$$

First 1 then 2 (Project Method) Fix two variables (e.g., x, y), first perform a definite integral over the third variable (e.g., z) on the "vertical line segment" corresponding to the fixed variables, and then perform a double integral over the fixed two variables (x, y) on their "projection region.

This method is convenient when the projection region of the integral region on a certain coordinate plane (e.g., xy-plane) is easy to determine, and the upper and lower limits of a single variable (e.g., z) can be easily expressed by the other two variables.

In the example 13.3, the following steps are taken:

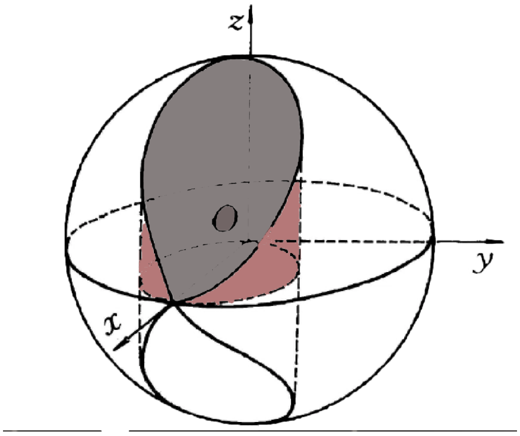
1. Determine the projection region D_{xy}:D_{xy} is the region on the xy-plane bounded by $x + y \leqslant 1$, $x \geq 0$, and $y \geq 0$, which can be expressed as $0 \leqslant x \leqslant 1$ and $0 \leqslant y \leqslant 1 - x$.
2. Determine the range of z: $z \in [0, xy]$ (since z is bounded below by $z = 0$ and above by $z = xy$).
3. Split the integral:
$$I = \iint_{D_{xy}} \left(\int_0^{xy} xy \mathrm{d}z \right) \mathrm{d}x\mathrm{d}y,$$
split the double integral on D_{xy} as: $I = \int_0^1 \, \mathrm{d}x \int_0^{1-x} \, \mathrm{d}y \int_0^{xy} xy \, \mathrm{d}z$. (Since xy is independent of z, it can be factored out without affecting the integral: $I = \int_0^1 \, \mathrm{d}x \int_0^{1-x} xy \, \mathrm{d}y \int_0^{xy} \, \mathrm{d}z$.)
4. Calculate the inner integral (with respect to z): $\int_0^{xy} xy \, dz = xy \cdot \int_0^{xy} dz = xy \cdot z|_0^{xy} = xy \cdot xy = x^2 y^2$.
5. Calculate the middle integral (with respect to y): Substitute the result of the inner integral,
$$\int_0^{1-x} x^2 y^2 \, dy = x^2 \cdot \left. \frac{y^3}{3} \right|_0^{1-x} = \frac{x^2 (1-x)^3}{3}.$$
6. Calculate the outer integral (with respect to x):Substitute the result of the middle integral:
$$\begin{aligned} \int_0^1 \frac{x^2 (1-x)^3}{3} \, dx &= \frac{1}{3} \int_0^1 (x^2 - 3x^3 + 3x^4 - x^5) \, dx \\ &= \frac{1}{3} \left(\frac{x^3}{3} - \frac{3x^4}{4} + \frac{3x^5}{5} - \left. \frac{x^6}{6} \right|_0^1 \right) \\ &= \frac{1}{3} \left(\frac{1}{3} - \frac{3}{4} + \frac{3}{5} - \frac{1}{6} \right) \\ &= \frac{1}{180}. \end{aligned}$$

Some tips for choosing between the two methods (take the above two examples as reference):

First 2 then 1 (Section Method)	**First 1 then 2 (Project Method)**
Section area D_z is easy to calculate	Projection region D_{xy} is easy to determine
Integrand is only related to z	Upper and lower limits z can be easily expressed by the other two variables x, y

Example 13.4 Find the volume of region bounded by the half Viviani's curve: sphere $x^2 + y^2 + z^2 \leqslant a^2$ and

cylinder $x^2 + y^2 \leqslant ax$ $(a > 0)$.



13.4 Improper Multiple Integrals

Improper multiple integrals can be also classified into two types, infinite integrals and defective integrals.

Definition 13.5 (Infinite Multiple Integral)

Let $D \subset \mathbb{R}^2$ be an unbounded region, whose boundary consists of finite or countably many smooth curves, and $f : D \to \mathbb{R}$ be a function, which is integrable on any measurable bounded closed set $D' \subset D$. If there exists an increasing sequence of bounded closed regions $\{D_k\}$ such that:

$$D_1 \subset D_2 \subset \cdots \subset D_k \subset \cdots, \quad \bigcup_{k=1}^{\infty} D_k = D,$$

which is called an **exhaustion** of D, and for each k, the integral $I(D_k) = \iint_{D_k} f$ exists, and the limit:

$$I = \lim_{k \to \infty} I(D_k)$$

exists, then I is called the **improper multiple integral** of f on D, denoted as:

$$I = \iint_D f = \lim_{k \to \infty} \iint_{D_k} f.$$

♣

Remark There are also other ways to define improper multiple integrals, such as using limit definitions based on distance to infinity. They are equivalent to the above definition.

Theorem 13.5

Improper multiple integral is integrable if and only if it is absolutely integrable.

♡

Example 13.5 Calculate

$$\iint_{\mathbb{R}^2} e^{-(x^2+y^2)} \, dx dy,$$

and find the value of Poisson integral

$$\int_{-\infty}^{+\infty} e^{-x^2} \, dx.$$

Chapter 14 Introduction to Surface Theory

14.1 Parameterization of Surface

Definition 14.1 (Parameterization of Surface)

Let Δ be an open subset in $\mathbb{R}^s$, and $\mathbf{r} : \Delta \to \mathbb{R}^n$ 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 $M = \mathbf{r}(\Delta) = \{\mathbf{r}(\mathbf{u}) \mid \mathbf{u} \in \Delta\}$ is called an s-dimensional **surface (patch)**, and $\mathbf{r}(\mathbf{u})$ is referred to as the parameterization of M.

When $\mathbf{r}(\mathbf{u}) \in C^k$ $(k \geq 0)$, $\mathbf{r}$ or M is called an s-dimensional C^k surface.

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

When

$$\operatorname{rank}(r_1'(\mathbf{u}^0), r_2'(\mathbf{u}^0), \ldots, r_s'(\mathbf{u}^0)) = \operatorname{rank}\begin{pmatrix} \frac{\partial r_1}{\partial u_1} & \cdots & \frac{\partial r_1}{\partial u_s} \\ \vdots & \ddots & \vdots \\ \frac{\partial r_n}{\partial u_1} & \cdots & \frac{\partial r_n}{\partial u_s} \end{pmatrix}_{\mathbf{u}^0} = s,$$

we call $\mathbf{u}^0$ or $\mathbf{r}(\mathbf{u}^0)$ 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 regular points, $\{r_1', \ldots, r_s'\}$ 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 $\mathbf{r}(t)$, we have:

$$\mathbf{r}'(t) = \left(r_1'(t), r_2'(t), \cdots, r_n'(t)\right).$$

If t is a regular point, then $\operatorname{rank}(\mathbf{r}'(t)) = \operatorname{rank}(r_1'(t), r_2'(t), \ldots, r_n'(t)) = 1$; this is equivalent to $\mathbf{r}'(t) \neq 0$, which means $r_1'(t), r_2'(t), \ldots, r_n'(t)$ are not all zero.

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

14.2 Tangent Space and Normal Space

Definition 14.2 (Tangent Space and Normal Space)

M is an s-dimensional smooth surface in $\mathbb{R}^n$ defined above, and $\mathbf{u}^0$ is a regular point of M. The **tangent space** of M at point $\mathbf{r}(\mathbf{u}^0)$ is the linear space spanned by s tangent vectors:

$$T_{\mathbf{u}^0}M = \operatorname{span}\{r_1'(\mathbf{u}^0), r_2'(\mathbf{u}^0), \ldots, r_s'(\mathbf{u}^0)\}.$$

Accordingly, the **normal space** of M at point $\mathbf{r}(\mathbf{u}^0)$ is the orthogonal complement of the tangent space:

$$N_{\mathbf{u}^0}M = (T_{\mathbf{u}^0}M)^{\perp}.$$

♣

Some special cases of tangent space and normal space expressions are given below:

¶ **Curve**

When $n=3, s=1$, M is a curve in three-dimensional space.

1. If the curve is parameterized as

$$\mathbf{r}(t)=(x(t),y(t),z(t)),\quad t\in I\subseteq\mathbb{R}.$$

At the regular point $\mathbf{r}(t^0)=(x(t^0),y(t^0),z(t^0))$, the tangent line and normal plane are:

$$T_{t^0}M=\operatorname{span}\{\mathbf{r}'(t^0)\}:\ \frac{x-x(t^0)}{x'(t^0)}=\frac{y-y(t^0)}{y'(t^0)}=\frac{z-z(t^0)}{z'(t^0)},$$

$$\begin{aligned}N_{t^0}M:\quad & x'(t^0)(x-x(t^0))+y'(t^0)(y-y(t^0))+z'(t^0)(z-z(t^0))=0\\ \Leftrightarrow & \mathbf{r}'(t^0)\cdot(\mathbf{r}-\mathbf{r}(t^0))=0.\end{aligned}$$

2. If the curve is described by:

$$\begin{cases}F(x,y,z)=0,\\ G(x,y,z)=0,\end{cases}$$

and the regular point is $\mathbf{x}^0=(x^0,y^0,z^0)$.
For the Jacobian matrix:

$$J=\begin{pmatrix}F_x(\mathbf{x}^0) & F_y(\mathbf{x}^0) & F_z(\mathbf{x}^0)\\ G_x(\mathbf{x}^0) & G_y(\mathbf{x}^0) & G_z(\mathbf{x}^0)\end{pmatrix},$$

since $\operatorname{rank} J=2$, without loss of generality, assume:

$$\frac{\partial(F,G)}{\partial(y,z)}=\begin{vmatrix}F_y(\mathbf{x}^0) & F_z(\mathbf{x}^0)\\ G_y(\mathbf{x}^0) & G_z(\mathbf{x}^0)\end{vmatrix}\neq 0.$$

By the implicit mapping theorem (12.9), we can express:

$$y=f(x),\quad z=g(x),\quad x\in U(x^0)\subseteq\mathbb{R}.$$

Then

$$f'(x^0)=\frac{\frac{\partial(F,G)}{\partial(z,x)}(\mathbf{x}^0)}{\frac{\partial(F,G)}{\partial(y,z)}(\mathbf{x}^0)},\quad g'(x^0)=\frac{\frac{\partial(F,G)}{\partial(x,y)}(\mathbf{x}^0)}{\frac{\partial(F,G)}{\partial(y,z)}(\mathbf{x}^0)}.$$

Therefore, the tangent line and normal plane at point $\mathbf{x}^0$ are:

$$\begin{aligned}T_{x^0}M:\quad & \frac{x-x^0}{1}=\frac{y-y^0}{f'(x^0)}=\frac{z-z^0}{g'(x^0)}\Leftrightarrow\frac{x-x^0}{\frac{\partial(F,G)}{\partial(y,z)}(\mathbf{x}^0)}=\frac{y-y^0}{\frac{\partial(F,G)}{\partial(z,x)}(\mathbf{x}^0)}=\frac{z-z^0}{\frac{\partial(F,G)}{\partial(x,y)}(\mathbf{x}^0)},\\ N_{x^0}M:\quad & \frac{\partial(F,G)}{\partial(y,z)}(\mathbf{x}^0)(x-x^0)+\frac{\partial(F,G)}{\partial(z,x)}(\mathbf{x}^0)(y-y^0)+\frac{\partial(F,G)}{\partial(x,y)}(\mathbf{x}^0)(z-z^0)=0.\end{aligned}$$

¶ **Surface**

When $n=3, s=2$, M is a surface in three-dimensional space.

1. If the surface can be described explicitly as:

$$z = f(x, y), \quad (x, y) \in D \subseteq \mathbb{R}^2,$$

at the regular point $\overline{\mathbf{x}}^0 = (x^0, y^0, z^0)$ ($\mathbf{x}^0 = (x^0, y^0)$), the tangent plane and normal line are:

$$\begin{aligned} T_{\mathbf{x}^0}M: &\quad z - z^0 = f_x(\mathbf{x}^0)(x - x^0) + f_y(\mathbf{x}^0)(y - y^0), \\ N_{\mathbf{x}^0}M: &\quad \frac{x - x^0}{f_x(\mathbf{x}^0)} = \frac{y - y^0}{f_y(\mathbf{x}^0)} = \frac{z - z^0}{-1}, \end{aligned}$$

where the expression of $T_{\mathbf{x}^0}M$ is derived from the total differential of $z = f(x, y)$ at point $\mathbf{x}^0$:

$$\mathrm{d}z = f_x(\mathbf{x}^0)\mathrm{d}x + f_y(\mathbf{x}^0)\mathrm{d}y.$$

2. If the surface is parameterized as

$$\mathbf{r}(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in D \subseteq \mathbb{R}^2,$$

at the regular point $\mathbf{x}^0 = (x^0, y^0, z^0)$.
For the Jacobian matrix:

$$J = \begin{pmatrix} x_u(\mathbf{x}^0) & x_v(\mathbf{x}^0) \\ y_u(\mathbf{x}^0) & y_v(\mathbf{x}^0) \\ z_u(\mathbf{x}^0) & z_v(\mathbf{x}^0) \end{pmatrix},$$

since $\operatorname{rank} J = 2$, without loss of generality, assume:

$$\frac{\partial(x, y)}{\partial(u, v)}(\mathbf{x}^0) = \begin{vmatrix} x_u(\mathbf{x}^0) & x_v(\mathbf{x}^0) \\ y_u(\mathbf{x}^0) & y_v(\mathbf{x}^0) \end{vmatrix} \neq 0.$$

By the inverse mapping theorem (12.11), we can determine the inverse mapping of

$$\begin{cases} x = x(u, v), \\ y = y(u, v), \end{cases}$$

in a neighborhood of point $\mathbf{x}^0$:

$$\begin{cases} u = u(x, y), \\ v = v(x, y), \end{cases}$$

where $u^0 = u(x^0, y^0)$, $v^0 = v(x^0, y^0)$. Then we obtain the explicit representation of the surface:

$$z = z(u(x, y), v(x, y)) = f(x, y), \quad (x, y) \in U(x^0, y^0) \subseteq \mathbb{R}^2.$$

Therefore, the tangent plane and normal line at point $\mathbf{x}^0$ are:

$$T_{\mathbf{x}^0}M: \quad \left.\frac{\partial(y,z)}{\partial(u,v)}\right|_{(u^0,v^0)}(x-x^0)+\left.\frac{\partial(z,x)}{\partial(u,v)}\right|_{(u^0,v^0)}(y-y^0)+\left.\frac{\partial(x,y)}{\partial(u,v)}\right|_{(u^0,v^0)}(z-z^0)=0,$$
$$N_{\mathbf{x}^0}M: \quad \frac{x-x^0}{\left.\frac{\partial(y,z)}{\partial(u,v)}\right|_{(u^0,v^0)}}=\frac{y-y^0}{\left.\frac{\partial(z,x)}{\partial(u,v)}\right|_{(u^0,v^0)}}=\frac{z-z^0}{\left.\frac{\partial(x,y)}{\partial(u,v)}\right|_{(u^0,v^0)}}.$$

14.3 Intrinsic Geometry

This two sections will introduce the first and second fundamental forms of surfaces, which can be all generalized to higher-dimensional manifolds; here, we only discuss the case of two-dimensional surfaces in three-dimensional space.

Let $\Delta \in \mathbb{R}^2$ be an open set, and $\mathbf{r}: \Delta \to \mathbb{R}^3$ be a C^k ($k \geqslant 2$) smooth regular surface parameterization, $M = \mathbf{r}(\Delta)$, where $\mathbf{u} = (u,v) \to \mathbf{r}(u,v) = (x(u,v), y(u,v), z(u,v))$. We can obtain that:

1. $\mathbf{r} \in C^k(\Delta, \mathbb{R}^3)$;
2. For any $p = (u,v) \in \Delta$, $\operatorname{rank}(\mathbf{r}'_u(u,v), \mathbf{r}'_v(u,v)) = 2$, that is, $\mathbf{r}'_u(u,v)$ and $\mathbf{r}'_v(u,v)$ are linearly independent, where
$$\mathbf{r}'_u(u,v) = \left(\frac{\partial x}{\partial u}, \frac{\partial y}{\partial u}, \frac{\partial z}{\partial u}\right), \quad \mathbf{r}'_v(u,v) = \left(\frac{\partial x}{\partial v}, \frac{\partial y}{\partial v}, \frac{\partial z}{\partial v}\right).$$

At this time, the tangent space $T_pM = \operatorname{span}(\mathbf{r}'_u(u,v), \mathbf{r}'_v(u,v))$, which is a subspace of $\mathbb{R}^3$. Hence, it inherits the inner product from $\mathbb{R}^3$.

The first fundamental form is the metric that a surface inherits from its ambient Euclidean space $\mathbb{R}^3$. It is essentially a symmetric positive-definite bilinear form defined on the tangent space, which allows us to measure lengths, angles, and areas on the surface.

Definition 14.3 (The First Fundamental Form)

In the above conditions, for any point $p = (u,v) \in \Delta$, the **first fundamental form** of the surface M at point p is defined as: for any tangent vector $\mathbf{w}_1, \mathbf{w}_2 \in T_pM$,

$$\mathrm{I}_p(\mathbf{w}_1, \mathbf{w}_2) := \mathbf{w}_1 \cdot \mathbf{w}_2,$$

which is a symmetric positive-definite bilinear form on the tangent space T_pM. This form is also called the **Riemann metric** of **metric tensor**, denoted as I_p or g_p. ♣

For convenience, we express I_p in the basis $\{\mathbf{r}'_u, \mathbf{r}'_v\}$ of the tangent space T_pM. Define:

$$E(u,v) := \mathrm{I}_p(\mathbf{r}_u, \mathbf{r}_u) = \mathbf{r}_u \cdot \mathbf{r}_u = \|\mathbf{r}_u\|^2,$$
$$F(u,v) := \mathrm{I}_p(\mathbf{r}_u, \mathbf{r}_v) = \mathbf{r}_u \cdot \mathbf{r}_v,$$
$$G(u,v) := \mathrm{I}_p(\mathbf{r}_v, \mathbf{r}_v) = \mathbf{r}_v \cdot \mathbf{r}_v = \|\mathbf{r}_v\|^2,$$

which are called the **Gauß coefficients**.

Then the matrix representation of the first fundamental form I_p under the basis $\{\mathbf{r}'_u, \mathbf{r}'_v\}$ is:

$$\mathrm{I}_p = \begin{pmatrix} E & F \\ F & G \end{pmatrix},$$

which is symmetric and positive-definite.

The quadratic form corresponding to this bilinear form is also commonly called the first fundamental form, denoted as $\mathrm{d}s^2$. For a tangent vector $\mathbf{w} \in T_pS$, it represents the square of the length of that vector:

$$\mathrm{d}s^2 := \mathrm{I}_p(\mathbf{w}, \mathbf{w}) = \|\mathbf{w}\|^2.$$

If $\mathbf{w}$ is the tangent vector to the curve $\gamma(t) = \mathbf{r}(u(t), v(t))$, given by $\gamma'(t) = \mathbf{r}_u u'(t) + \mathbf{r}_v v'(t)$, then $\mathrm{d}s^2$ is conventionally written as:

$$\mathrm{d}s^2 = E\,\mathrm{d}u^2 + 2F\,\mathrm{d}u\,\mathrm{d}v + G\,\mathrm{d}v^2.$$

Here,$\mathrm{d}u$ and $\mathrm{d}v$ are the coordinates under the basis $\{\mathrm{d}u, \mathrm{d}v\}$, representing the components of the tangent vector (u', v'). This is a long-standing notation, and strictly speaking, it represents the value of the quadratic form on the vector (u', v').

¶ Arc Length

Definition 14.4 (Arc Length)

Let $C = \overset{\frown}{AB}$ be a curve on the $\mathbb{R}^2$ plane[a], take any partition $A = P_0, P_1, \ldots, P_n = B$, which divides the curve C into n segments, denoted as T. Then connect every two adjacent points P_{i-1} and P_i with a straight line segment, obtaining n chords $\overline{P_{i-1}P_i}(i = 1, 2, \ldots, n)$, which in turn form an inscribed polygonal line C. Let

$$\|T\| = \max_{1 \leqslant i \leqslant n} \|P_{i-1}P_i\|, \quad s_T = \sum_{i=1}^{n} \|P_{i-1}P_i\|.$$

If the limit

$$\lim_{\|T\| \to 0} s_T = s,$$

namely,

$$\forall \varepsilon > 0, \exists \delta > 0, \text{s.t.} \forall T(\|T\| < \delta) : |s_T - s| < \varepsilon,$$

and the limit is independent of the choice of partition T, then C is said to be rectifiable, and the limit s is called the arc length of the curve C.

[a] Or in $\mathbb{R}^3$ space, even in a higher-dimensional Euclidean space.

♣

Theorem 14.1 (Sufficient Condition for Rectifiability of Curves)

Let the curve C in $\mathbb{R}^2$ be given by the parametric equations

$$(x, y) = (x(t), y(t)), \quad t \in [\alpha, \beta],$$

and let it be a C^1 smooth regular curve[a] Then C is rectifiable, and its arc length is

$$s = \int_{\alpha}^{\beta} \sqrt{x'^2(t) + y'^2(t)}\,\mathrm{d}t.$$

[a] I.e., $x(t)$ and $y(t)$ are continuously differentiable, and $x'^2(t) + y'^2(t) \neq 0$; a curve C satisfying this condition is called a regular point. Also see Definition 14.1

¶ **Area**

14.4 Extrinsic Geometry

The second fundamental form is a symmetric bilinear form defined on the tangent space that measures the change in the normal vector of a surface, thereby describing the extrinsic curvature of the surface relative to its ambient space $\mathbb{R}^3$.

On the regular surface patch M defined in the beginning of last section, we can define a continuous unit normal vector field $\mathbf{n} : M \to \mathbb{S}^2$, where $\mathbb{S}^2$ is the unit sphere in $\mathbb{R}^3$:

$$\mathbf{n}(p) = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}(p).$$

This mapping $\mathbf{n}$ from the surface to the unit sphere is called the **Gauß map**. The second fundamental form is defined by studying the differential of the Gauß map.

Definition 14.5 (The Second Fundamental Form)

Under the above conditions, for any point $p = (u, v) \in \Delta$, the **second fundamental form** of the surface M at point p is a symmetric bilinear form on the tangent space T_pM, which is defined as: for any tangent vector $\mathbf{w}_1, \mathbf{w}_2 \in T_pM$,

$$\mathrm{II}_p(\mathbf{w}_1, \mathbf{w}_2) := -\mathrm{d}_p\mathbf{n}(\mathbf{w}_1) \cdot \mathbf{w}_2,$$

[a] where $\mathrm{d}_p\mathbf{n} : T_pM \to T_{\mathbf{n}(p)}\mathbb{S}^2$ is the differential (or Jacobian) of the Gauß map at point p.
The linear operator associated with $\mathrm{d}_p\mathbf{n}$, defined as $W_p(\mathbf{w}) = -\mathrm{d}_p\mathbf{n}(\mathbf{w})$, is called the Weingarten map or shape operator, and it is a linear operator from T_pM to itself. Therefore, the second fundamental form can also be written as:

$$\mathrm{II}_p(\mathbf{w}_1, \mathbf{w}_2) = W_p(\mathbf{w}_1) \cdot \mathbf{w}_2.$$

[a] About the formula,

- since $\mathbf{n}(p)$ is a unit vector, $T_{\mathbf{n}(p)}\mathbb{S}^2$ is the plane orthogonal to $\mathbf{n}(p)$, and T_pM itself is also orthogonal to $\mathbf{n}(p)$, it follows that $\mathrm{d}_p\mathbf{n}(\mathbf{w}_1)$ and $\mathbf{w}_2$ lie in the same plane, and their dot product is well-defined.
- the negative sign in this definition is a convention, which makes the principal curvatures of a convex surface (like a sphere) positive.

For convenience, we express II_p in the basis $\{\mathbf{r}'_u, \mathbf{r}'_v\}$ of the tangent space T_pM. Define:

$$\begin{aligned}
L(u, v) &:= \mathrm{II}_p(\mathbf{r}_u, \mathbf{r}_u) = W_p(\mathbf{r}_u) \cdot \mathbf{r}_u = \mathbf{r}_{uu} \cdot \mathbf{n};\\
M(u, v) &:= \mathrm{II}_p(\mathbf{r}_u, \mathbf{r}_v) = W_p(\mathbf{r}_u) \cdot \mathbf{r}_v = \mathbf{r}_{uv} \cdot \mathbf{n};\\
N(u, v) &:= \mathrm{II}_p(\mathbf{r}_v, \mathbf{r}_v) = W_p(\mathbf{r}_v) \cdot \mathbf{r}_v = \mathbf{r}_{vv} \cdot \mathbf{n}.
\end{aligned}$$

Then the matrix representation of the second fundamental form II_p under the basis $\{\mathbf{r}'_u, \mathbf{r}'_v\}$ is:

$$\mathrm{II}_p = \begin{pmatrix} L & M \\ M & N \end{pmatrix},$$

which is symmetric, but not necessarily positive-definite. And its sign reflects the way the surface is curved.

The associated second fundamental form, also denoted by II, is an expression for the normal curvature:

$$\mathrm{II} = L\,\mathrm{d}u^2 + 2M\,\mathrm{d}u\,\mathrm{d}v + N\,\mathrm{d}v^2.$$

For a unit tangent vector $\mathbf{w} \in T_pM$, the value of $\mathrm{II}_p(\mathbf{w}, \mathbf{w})$ is the normal curvature of the surface in the direction of $\mathbf{w}$, denoted $k_n(\mathbf{w})$.

¶ **Curvature**

Curvature is a mathematical quantity describing the "bending" degree of a geometric object, such as a curve or a surface.

The meaning of curvature varies for geometric objects of different dimensions:

- Curvature on a curve: Describes the degree to which the curve deviates from a straight line.
- Description of curvature by a surface: Is more complex, involving directionality—the bending of a surface can be completely different in different directions.

The curvature of a surface is usually classified into the following typical types: normal curvature, principal curvatures, mean curvature, Gaussian curvature, etc.

Definition 14.6 (Curvature of Curve)

Let C be a C^2 smooth regular curve in $\mathbb{R}^3$, parameterized by arc length t:

$$\mathbf{r}(t) = (x(t), y(t), z(t)), \quad t \in [a, b].$$

The unit tangent vector of the curve at point t is:

$$\mathbf{T}(t) = \mathbf{r}'(t) = (x'(t), y'(t), z'(t)).$$

The **curvature** of the curve at point t is defined as the magnitude of the derivative of the unit tangent vector with respect to arc length:

$$\kappa(t) = \left\| \frac{\mathrm{d}\mathbf{T}(t)}{\mathrm{d}t} \right\| = \frac{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|}{\|\mathbf{r}'(t)\|^3}.$$

Geometrically, curvature measures how quickly the curve changes direction at point t.
If the best-fit circle is found based on the tangent and normal at a certain point, the radius of this circle is called the radius of curvature R, and the curvature is its reciprocal:

$$\kappa = \frac{1}{R}.$$

This fitted circle is called the **osculating circle** of the curve at that point. ♣

Some special cases of curvature are given below:

1. For a plane curve given by $y = f(x)$, the curvature at point x is:

$$\kappa(x) = \frac{|f''(x)|}{(1 + (f'(x))^2)^{3/2}}.$$

2. For a circle with radius R, the curvature is constant:

$$\kappa = \frac{1}{R}.$$

14.5 Oriented Surface

Chapter 15 Line Integrals and Surface Integrals

15.1 Line Integrals and Surface Integrals of scalar fields

¶ **Line Integral of Scalar Field**

Definition 15.1 (Line Integral of Scalar Field)

Let L is a rectifiable continuous curve in $\mathbb{R}^3$, whose endpoints are A and B, and $f(x, y, z)$ is bounded on L. Partition L into n segments by points $A = P_0, P_1, \ldots, P_n = B$, and select a point $\boldsymbol{\xi}_i$ on each segment $P_{i-1}P_i$ $(i = 1, 2, \ldots, n)$. Remark that the length of segment $P_{i-1}P_i$ is Δs_i $(i = 1, 2, \cdots n)$, and make the sum:

$$\sum_{i=1}^{n} f(\boldsymbol{\xi}_i)\Delta s_i.$$

If when λ (the length of the longest segment) tends to 0, the above sum tends to a limit I independent of the partition and the choice of points $\boldsymbol{\xi}_i$, then I is called the **line integral of the scalar field f along the curve L**, denoted as:

$$\int_L f \, \mathrm{d}s.$$

That is,

$$I = \int_L f(\boldsymbol{\xi}) \, \mathrm{d}s = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\boldsymbol{\xi}_i)\Delta s_i.$$

♣

Theorem 15.1

Let L be a C^1 smooth regular curve parameterized by $\mathbf{x}(t) = (x(t), y(t), z(t)), t \in [\alpha, \beta]$, and f be continuous on L. Then:

$$\int_L f \, \mathrm{d}s = \int_\alpha^\beta f(\mathbf{x}(t))\|\mathbf{x}'(t)\| \, \mathrm{d}t. = \int_\alpha^\beta f(x(t), y(t), z(t))\sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2} \, \mathrm{d}t.$$

♡

Specially, if the plane curve L is given by $y = y(x), x \in [a, b]$, then:

$$\int_L f \, \mathrm{d}s = \int_a^b f(x, y(x))\sqrt{1 + (y'(x))^2} \, \mathrm{d}x.$$

¶ **Surface Integrals of Scalar Fields**

Definition 15.2 (Surface Integral of Scalar Field)

Let Σ be a piecewise smooth surface in $\mathbb{R}^3$, and $f(x, y, z)$ be bounded on Σ. Partition Σ into n small pieces $\Delta\Sigma_1, \Delta\Sigma_2, \ldots, \Delta\Sigma_n$ with smooth curve webs, and select a point $\boldsymbol{\xi}_i$ on each piece $\Delta\Sigma_i$ $(i = 1, 2, \ldots, n)$. Remark that the area of piece $\Delta\Sigma_i$ is ΔS_i $(i = 1, 2, \cdots n)$, and make the sum:

$$\sum_{i=1}^{n} f(\boldsymbol{\xi}_i)\Delta S_i.$$

If when λ (the area of the largest piece) tends to 0, the above sum tends to a limit I independent of the partition and the choice of points $\boldsymbol{\xi}_i$, then I is called the **surface integral of the scalar field f over the surface Σ**, denoted as:

$$\iint_\Sigma f \,\mathrm{d}S.$$

That is,

$$I = \iint_\Sigma f(\boldsymbol{\xi}) \,\mathrm{d}S = \lim_{\lambda \to 0} \sum_{i=1}^{n} f(\boldsymbol{\xi}_i) \Delta S_i.$$

Theorem 15.2

Let Σ be a piecewise smooth closed surface parameterized by $\mathbf{r}(u,v) = (x(u,v), y(u,v), z(u,v)), (u,v) \in D$, and f be continuous on Σ. x, y, z have continuous first-order partial derivatives with respect to u and v on D, and according Jacobian matrix

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

is of full rank. Then:

$$\iint_\Sigma f \,\mathrm{d}S = \iint_D f(\mathbf{r}(u,v)) \left\| \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right\| \mathrm{d}u\mathrm{d}v = \iint_D f(x(u,v), y(u,v), z(u,v)) \sqrt{EG - F^2} \,\mathrm{d}u\mathrm{d}v,$$

where E, G, F are the Gauß coefficients of the surface Σ.

Specially, if the surface Σ is given by $z = z(x,y), (x,y) \in D$, then:

$$\iint_\Sigma f \,\mathrm{d}S = \iint_D f(x, y, z(x,y)) \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \,\mathrm{d}x\mathrm{d}y.$$

15.2 Differential Form and Exterior Differentiation

Let $\mathrm{d}x_i, \mathrm{d}x_j$ be differentials of independent variables x_i, x_j.
In $\mathbb{R}^1$:

0-form: $f(x)$,

1-form: $\omega = f(x)\mathrm{d}x$,

k-form ($k \geqslant 2$): $\omega = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n} f_{i_1 i_2 \cdots i_k}(x_1, x_2, \cdots, x_n) \mathrm{d}x_{i_1} \wedge \mathrm{d}x_{i_2} \wedge \cdots \wedge \mathrm{d}x_{i_k} = 0.$

In $\mathbb{R}^2$:

0-form: $f(x,y)$,

1-form: $\omega = P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y$,

2-form: $\omega = f(x,y)\mathrm{d}x \wedge \mathrm{d}y$,

k-form ($k \geqslant 3$): $\omega = \sum\limits_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} f_{i_1 i_2\cdots i_k}(x_1,x_2,\cdots,x_n)\mathrm{d}x_{i_1}\wedge \mathrm{d}x_{i_2}\wedge\cdots\wedge \mathrm{d}x_{i_k} = 0.$

In $\mathbb{R}^3$:

0-form: $f(x,y,z)$,

1-form: $\omega = P(x,y,z)\mathrm{d}x + Q(x,y,z)\mathrm{d}y + R(x,y,z)\mathrm{d}z$,

2-form: $\omega = P(x,y,z)\mathrm{d}y\wedge \mathrm{d}z + Q(x,y,z)\mathrm{d}z\wedge \mathrm{d}x + R(x,y,z)\mathrm{d}x\wedge \mathrm{d}y$,

3-form: $\omega = f(x,y,z)\mathrm{d}x\wedge \mathrm{d}y\wedge \mathrm{d}z$,

k-form ($k \geqslant 4$): $\omega = \sum\limits_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} f_{i_1 i_2\cdots i_k}(x_1,x_2,\cdots,x_n)\mathrm{d}x_{i_1}\wedge \mathrm{d}x_{i_2}\wedge\cdots\wedge \mathrm{d}x_{i_k} = 0.$

Here, $\wedge$ is called the **wedge product**, which satisfies:

1. Skew symmetric: $\mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i$,
2. Associative: $(\mathrm{d}x_i \wedge \mathrm{d}x_j)\wedge \mathrm{d}x_k = \mathrm{d}x_i \wedge (\mathrm{d}x_j \wedge \mathrm{d}x_k)$,
3. In a fixed dimension, the wedge product of k-forms becomes zero (as higher forms are not defined), for example, in 3-dimensional space, a 4-form is equal to 0.

Differential form is a skew symmetric tensor on vector field.

Definition 15.3 (Exterior Differentiation)

Let ω be a k-form on $\mathbb{R}^n$,

$$\omega = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} f_{i_1 i_2\cdots i_k}(x_1,x_2,\cdots,x_n)\mathrm{d}x_{i_1}\wedge \mathrm{d}x_{i_2}\wedge\cdots\wedge \mathrm{d}x_{i_k},$$

where $f_{i_1 i_2\cdots i_k}$ are functions with continuous first-order partial derivatives. The exterior differentiation of ω is defined as:

$$\mathrm{d}\omega = \sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n} \mathrm{d}f_{i_1 i_2\cdots i_k}(x_1,x_2,\cdots,x_n)\wedge\mathrm{d}x_{i_1}\wedge \mathrm{d}x_{i_2}\wedge\cdots\wedge \mathrm{d}x_{i_k},$$

where

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\mathrm{d}x_2 + \cdots + \frac{\partial f}{\partial x_n}\mathrm{d}x_n.$$

Note that the exterior differentiation of a k-form is a $k+1$-form. ♣

Property

Linearity $\mathrm{d}(\alpha\omega + \beta\eta) = \alpha\mathrm{d}\omega + \beta\mathrm{d}\eta$, *where α, β are constants.*

Leibniz Rule $\mathrm{d}(\omega\wedge\eta) = \mathrm{d}\omega\wedge\eta + (-1)^k\omega\wedge\mathrm{d}\eta$, *where ω is a k-form.*

Nilpotency $\mathrm{d}(\mathrm{d}\omega) = 0.$

15.3 Line Integrals and Surface Integrals of Vector Fields

¶ Line Integral of Vector Field

Definition 15.4 (Line Integral of Vector Field)

Let $\overrightarrow{L}$ be a orientated smooth curve in $\mathbb{R}^3$, whose endpoints are A and B. Take unit tangent vector $\boldsymbol{\tau} = (\cos\alpha, \cos\beta, \cos\gamma)$ at each point of $\overrightarrow{L}$, making it consistent with the orientation of $\overrightarrow{L}$. Let $\mathbf{f}(x, y, z) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ be a vector-valued function on $\overrightarrow{L}$, then

$$\int_{\overrightarrow{L}} \mathbf{f} \cdot \boldsymbol{\tau} \mathrm{d}s = \int_{\overrightarrow{L}} [P\cos\alpha + Q\cos\beta + R\cos\gamma]\, \mathrm{d}s$$

is called the **line integral of the vector field f along the oriented curve** $\overrightarrow{L}$ (if the right-hand side exists). ♣

Consider a differential arc length element $\mathrm{d}s$ at a point (x, y, z) on the curve L. We form the vector $\mathrm{d}\mathbf{s} = \boldsymbol{\tau}\mathrm{d}s$, where $\boldsymbol{\tau} = (\cos\alpha, \cos\beta, \cos\gamma)$ represents the unit tangent vector of curve L at (x, y, z), pointing along the direction of L. The projection of $\mathrm{d}s$ onto the x-axis is given by $\cos\alpha\, \mathrm{d}s$. Therefore, we denote:

$$\mathrm{d}x = \cos\alpha\, \mathrm{d}s, \quad \mathrm{d}y = \cos\beta\, \mathrm{d}s, \quad \mathrm{d}z = \cos\gamma\, \mathrm{d}s.$$

Thus, the second type of line integral can be expressed as:

$$\int_{\overrightarrow{L}} \mathbf{f} \cdot \boldsymbol{\tau} \mathrm{d}s = \int_{\overrightarrow{L}} \mathbf{f}\, \mathrm{d}\mathbf{s} = \int_{\overrightarrow{L}} P(x, y, z)\mathrm{d}x + Q(x, y, z)\mathrm{d}y + R(x, y, z)\mathrm{d}z.$$

This line integral is also referred to as the integral of the 1-form:

$$\omega = P(x, y, z)\mathrm{d}x + Q(x, y, z)\mathrm{d}y + R(x, y, z)\mathrm{d}z.$$

The second type of line integral of ω along the curve L is denoted as:

$$\int_{\overrightarrow{L}} \omega.$$

Theorem 15.3

Let $\overrightarrow{L}$ be a C^1 smooth regular oriented curve parameterized by $\mathbf{x}(t) = (x(t), y(t), z(t)), t \in [\alpha, \beta]$, and $\mathbf{f} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ be continuous on $\overrightarrow{L}$. Then:

$$\int_{\overrightarrow{L}} \mathbf{f} \cdot \boldsymbol{\tau} \mathrm{d}s = \int_{\alpha}^{\beta} \mathbf{f}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)\, \mathrm{d}t$$
$$= \int_{\alpha}^{\beta} [P(x(t), y(t), z(t))x'(t) + Q(x(t), y(t), z(t))y'(t) + R(x(t), y(t), z(t))z'(t)]\, \mathrm{d}t.$$

♡

Specially, if the plane curve $\overrightarrow{L}$ is given by $y = y(x), x : a \to b$, then:

$$\int_{\overrightarrow{L}} \mathbf{f} \cdot \boldsymbol{\tau} \mathrm{d}s = \int_{a}^{b} \mathbf{f}(x, y(x)) \cdot (1, y'(x))\sqrt{1 + (y'(x))^2}\, \mathrm{d}x.$$

¶ Surface Integral of Vector Field

Definition 15.5 (Surface Integral of Vector Field)

Let $\vec{\Sigma}$ be an orientated smooth surface in $\mathbb{R}^3$, and $\mathbf{f}(x,y,z) = P(x,y,z)\mathbf{i} + Q(x,y,z)\mathbf{j} + R(x,y,z)\mathbf{k}$ be a vector-valued function on $\vec{\Sigma}$. Each point appoints a unit normal vector $\mathbf{n} = (\cos\alpha, \cos\beta, \cos\gamma)$. Then

$$\iint_{\vec{\Sigma}} \mathbf{f}\cdot\mathbf{n}\mathrm{d}S = \iint_{\vec{\Sigma}} [P\cos\alpha + Q\cos\beta + R\cos\gamma]\ \mathrm{d}S$$

is called the **surface integral of the vector field f over the oriented surface** $\vec{\Sigma}$ (if the right-hand side exists). ♣

Consider a differential area element $\mathrm{d}S$ at a point (x,y,z) on the surface Σ. We form the vector $\mathrm{d}\mathbf{S} = \mathbf{n}\mathrm{d}S$, where $\mathbf{n} = (\cos\alpha, \cos\beta, \cos\gamma)$ represents the unit normal vector of surface Σ at (x,y,z), pointing along the orientation of Σ. The projection of $\mathrm{d}S$ onto the x-axis is given by $\cos\alpha\,\mathrm{d}S$. Therefore, we denote:

$$\mathrm{d}y\wedge\mathrm{d}z = \cos\alpha\,\mathrm{d}S, \quad \mathrm{d}z\wedge\mathrm{d}x = \cos\beta\,\mathrm{d}S, \quad \mathrm{d}x\wedge\mathrm{d}y = \cos\gamma\,\mathrm{d}S.$$

Thus, the surface integral can be expressed as:

$$\iint_{\vec{\Sigma}} \mathbf{f}\cdot\mathbf{n}\mathrm{d}S = \iint_{\vec{\Sigma}} P\mathrm{d}y\wedge\mathrm{d}z + Q\mathrm{d}z\wedge\mathrm{d}x + R\mathrm{d}x\wedge\mathrm{d}y = \iint_{\Sigma} P\mathrm{d}y\mathrm{d}z + Q\mathrm{d}z\mathrm{d}x + R\mathrm{d}x\mathrm{d}y,$$

where $\mathrm{d}y\mathrm{d}z$ is the simplified notation for $\mathrm{d}y\wedge\mathrm{d}z$, etc. This surface integral is also referred to as the integral of the 2-form:

$$\omega = P(x,y,z)\mathrm{d}y\wedge\mathrm{d}z + Q(x,y,z)\mathrm{d}z\wedge\mathrm{d}x + R(x,y,z)\mathrm{d}x\wedge\mathrm{d}y.$$

The second type of surface integral of ω over the surface Σ is denoted as:

$$\iint_{\vec{\Sigma}} \omega.$$

Theorem 15.4

Let $\vec{\Sigma}$ be a smooth oriented surface parameterized by $\mathbf{r}(u,v) = (x(u,v), y(u,v), z(u,v))$, $(u,v)\in D$, where D is a closed region with piecewise smooth boundary in uv-plane, and $\mathbf{f} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ be continuous on $\vec{\Sigma}$. x, y, z have continuous first-order partial derivatives with respect to u and v on D, and according Jacobian matrix is of full rank. Then:

$$\begin{aligned}
&\iint_{\vec{\Sigma}} \mathbf{f}\cdot\mathbf{n}\mathrm{d}S \\
&= \iint_{\vec{\Sigma}} [P\cos\alpha + Q\cos\beta + R\cos\gamma]\,\mathrm{d}S \\
&= \iint_{D} \mathbf{f}(\mathbf{r}(u,v))\cdot\left(\frac{\partial\mathbf{r}}{\partial u}\times\frac{\partial\mathbf{r}}{\partial v}\right)\mathrm{d}u\mathrm{d}v \\
&= \pm\iint_{D}\left[P(x(u,v),y(u,v),z(u,v))\cdot\frac{\partial(y,z)}{\partial(u,v)} + Q(x(u,v),y(u,v),z(u,v))\cdot\frac{\partial(z,x)}{\partial(u,v)}\right. \\
&\quad \left. + R(x(u,v),y(u,v),z(u,v))\cdot\frac{\partial(x,y)}{\partial(u,v)}\right]\mathrm{d}u\mathrm{d}v,
\end{aligned}$$

where the sign $\pm$ depends on whether the orientation of $\vec{\Sigma}$ is consistent with the direction of $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$.

Specially, if the surface $\vec{\Sigma}$ is given by $z = z(x, y), (x, y) \in D_{xy}$, where D_{xy} is a closed region with piecewise smooth boundary in xy-plane, and $R(x, y, z)$ is continuous on D_{xy}, then:

$$\iint_{\vec{\Sigma}} R(x, y, z)\mathrm{d}x\mathrm{d}y = \pm \iint_{D_{xy}} R(x, y, z(x, y))\, \mathrm{d}x\mathrm{d}y,$$

where the sign $\pm$ depends on whether the orientation of $\vec{\Sigma}$ is upward or downward.

15.4 Stokes' Formula

¶ **Newton-Leibniz Formula**

¶ **Green's Formula**

Consider two kinds of special orientated closed regions in xy-plane as shown in Figure 15.1. As for the first region $\vec{M}$, it consists of four orientated curves:
$\vec{C_1}$ $y = \varphi_1(x), x \in [a, b]$,
$\vec{C_2}$ $x = b, y \in [\varphi_1(b), \varphi_2(b)]$, can be reduced to a point,
$\vec{C_3}$ $y = \varphi_2(x), x \in [a, b]$,
$\vec{C_4}$ $x = a, y \in [\varphi_1(a), \varphi_2(a)]$, can be reduced to a point.
The second region is similar.

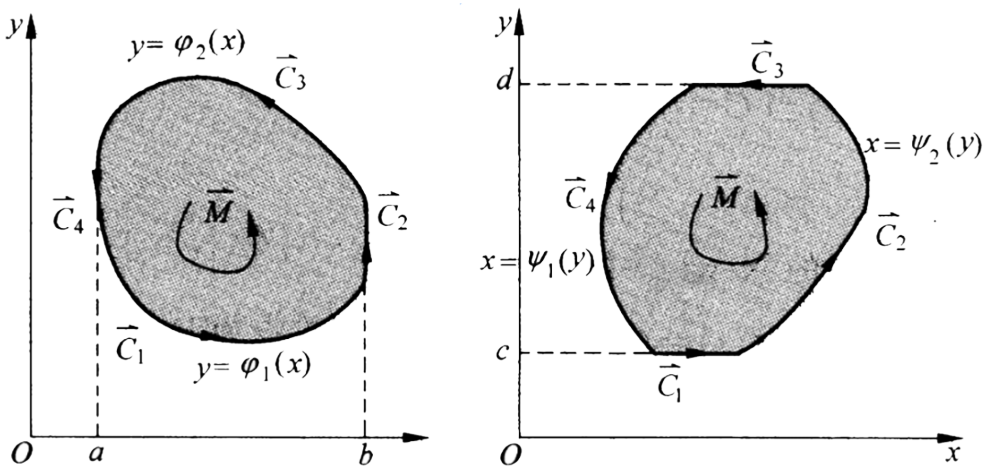


Figure 15.1: Two special orientated closed regions.

Denote $\oint_{\partial \vec{M}}$ as the line integral along the boundary of region $\vec{M}$, then we have the following lemma.

Lemma 15.1

1. Let $\partial \vec{M}$ be the first region in Fig 15.1, $P(x, y) \in C^1(M)$, then:

$$\oint_{\partial \vec{M}} P\, \mathrm{d}x = -\iint_{\vec{M}} \frac{\partial P}{\partial y}\, \mathrm{d}x \wedge \mathrm{d}y,$$

2. Let $\overrightarrow{\partial M}$ be the second region in Fig 15.1, $Q(x, y) \in C^1(M)$, then:

$$\oint_{\overrightarrow{\partial M}} Q \,\mathrm{d}y = \iint_{\overrightarrow{M}} \frac{\partial Q}{\partial x} \,\mathrm{d}x \wedge \mathrm{d}y.$$

Theorem 15.5 (Green's Theorem)

Let $\overrightarrow{M}$ be an orientated closed region in $\mathbb{R}^2$, and $\omega = P\mathrm{d}x + Q\mathrm{d}y \in C^1(M)$. If $\overrightarrow{\partial M}$ can be split into finitely many first and second regions in Fig 15.1 simultaneously (non-overlapping, no shared interior points), then:

$$\oint_{\overrightarrow{\partial M}} P \,\mathrm{d}x + Q \,\mathrm{d}y = \iint_{\overrightarrow{M}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x \wedge \mathrm{d}y = \iint_{M} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x\mathrm{d}y,$$

[a] or equivalently,

$$\oint_{\overrightarrow{\partial M}} \omega = \iint_{\overrightarrow{M}} \mathrm{d}\omega,$$

where $\overrightarrow{\partial M}$ is the induced orientation of $\overrightarrow{M}$.

[a] Note that $\mathrm{d}x \wedge \mathrm{d}y$ is directed area element, while $\mathrm{d}x\mathrm{d}y$ is unsigned area element.

¶ Gauß's Formula

Consider three kinds of special orientated closed surfaces in $\mathbb{R}^3$ as shown in Figure 15.2. As for the first surface $\overrightarrow{M}$ ($\overrightarrow{M}$ adopts a positive orientation (right-hand system), and $\overrightarrow{\partial M}$ adopts the outward normal orientation), it consists of three orientated surfaces:

$\overrightarrow{\Sigma_1}$ $z = \varphi_1(x, y), (x, y) \in \Delta_1,$

$\overrightarrow{\Sigma_2}$ $z = \varphi_2(x, y), (x, y) \in \Delta_1,$

$\overrightarrow{\Sigma_3}$ A cylindrical taking $\partial\Delta_1$ as the directrix, with the generatrix paralleling to the Oz-axis. Of course, it can also be reduced as a closed curve.

The second and third surfaces are similar.

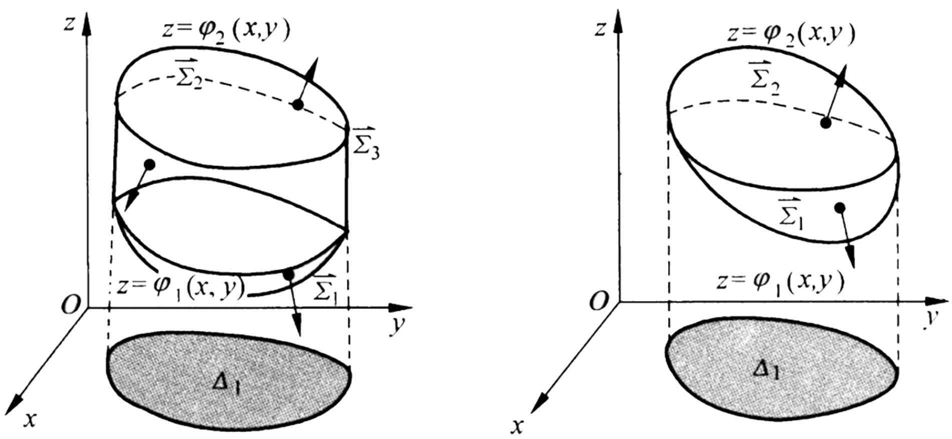


Figure 15.2: Three special orientated closed surfaces (only the first two are shown).

Denote $\oiint_{\overrightarrow{\partial M}}$ as the surface integral over the boundary of region $\overrightarrow{M}$, then we have the following lemma.

Lemma 15.2

1. Let $\overrightarrow{\partial M}$ be the first surface in Fig 15.2, $R(x,y,z) \in C^1(M)$, then:

$$\oiint_{\overrightarrow{\partial M}} R\,\mathrm{d}x \wedge \mathrm{d}y = \iiint_{\overrightarrow{M}} \frac{\partial R}{\partial z}\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z,$$

2. Let $\overrightarrow{\partial M}$ be the second surface in Fig 15.2, $P(x,y,z) \in C^1(M)$, then:

$$\oiint_{\overrightarrow{\partial M}} P\,\mathrm{d}y \wedge \mathrm{d}z = \iiint_{\overrightarrow{M}} \frac{\partial P}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z,$$

3. Let $\overrightarrow{\partial M}$ be the third surface in Fig 15.2, $Q(x,y,z) \in C^1(M)$, then:

$$\oiint_{\overrightarrow{\partial M}} Q\,\mathrm{d}z \wedge \mathrm{d}x = \iiint_{\overrightarrow{M}} \frac{\partial Q}{\partial y}\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z.$$

♡

Theorem 15.6 (Gauß's Theorem)

Let $\overrightarrow{M}$ be an orientated closed region in $\mathbb{R}^3$, and $\omega = P\mathrm{d}y \wedge \mathrm{d}z + Q\mathrm{d}z \wedge \mathrm{d}x + R\mathrm{d}x \wedge \mathrm{d}y \in C^1(M)$. If $\overrightarrow{\partial M}$ can be split into finitely many first, second and third regions in Fig 15.1 simultaneously (non-overlapping, no shared interior points), then: then:

$$\oiint_{\overrightarrow{\partial M}} P\,\mathrm{d}y \wedge \mathrm{d}z + Q\,\mathrm{d}z \wedge \mathrm{d}x + R\,\mathrm{d}x \wedge \mathrm{d}y = \iiint_{\overrightarrow{M}} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z,$$

or equivalently,

$$\oiint_{\overrightarrow{\partial M}} \omega = \iiint_{\overrightarrow{M}} \mathrm{d}\omega,$$

where $\overrightarrow{\partial M}$ is the induced orientation of $\overrightarrow{M}$.

♡

¶ Stokes' Formula

Theorem 15.7 (Stokes' Theorem)

Let $\overrightarrow{M}$ be an orientated smooth surface in $\mathbb{R}^3$ with boundary $\overrightarrow{\partial M}$, and $\omega = P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z \in C^1(\Sigma)$. Then:

$$\begin{aligned}
&\oint_{\overrightarrow{\partial M}} P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z \\
&= \iint_{\overrightarrow{M}} \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathrm{d}y \wedge \mathrm{d}z + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathrm{d}z \wedge \mathrm{d}x + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x \wedge \mathrm{d}y \\
&= \iint_{\overrightarrow{M}} \begin{vmatrix} \mathrm{d}y \wedge \mathrm{d}z & \mathrm{d}z \wedge \mathrm{d}x & \mathrm{d}x \wedge \mathrm{d}y \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \\
&= \iint_{\overrightarrow{M}} \begin{vmatrix} \cos\alpha & \cos\beta & \cos\gamma \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \mathrm{d}S,
\end{aligned}$$

or equivalently,

$$\oint_{\overrightarrow{\partial M}} \omega = \iint_{\overrightarrow{M}} \mathrm{d}\omega,$$

where $\overrightarrow{\partial M}$ is the induced orientation of $\overrightarrow{M}$. ♡

15.5 Closed and Exact Differential Forms

Definition 15.6 (Closed and Exact Differential Forms)

Let $U \subset \mathbb{R}^n$ be an open set and ω be a $C^r\,(r \geqslant 1)$ k-form on U.

1. If $\mathrm{d}\omega = 0$, then ω is called a **closed form.**
2. If there exists a C^{r+1} $(k-1)$-form η such that $\omega = \mathrm{d}\eta$, then ω is called an **exact differential form.** ♣

Theorem 15.8 (Necessary Condition for Exactness)

Let $U \subset \mathbb{R}^n$ be an open set and ω be a C^1 k-form on U. If ω is exact, then ω is closed. The converse is not necessarily true. ♡

We only discuss the case of 1-forms in $\mathbb{R}^2$ below.

Let $\omega = P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y$ be a C^1 1-form on an open set $U \subset \mathbb{R}^2$. For any points $A, B \in U$, a piecewise smooth simple closed curve on U is called a **path** from A to B if it starts at A and ends at B.

For any path $\overrightarrow{L}$ from A to B, if

$$\int_{\overrightarrow{L}} \omega = \int_A^B \omega,$$

where the right-hand side is independent of the choice of path $\overrightarrow{L}$, then the line integral of ω is said to be **path-independent** on U.

Theorem 15.9

Let $U \in \mathbb{R}^2$ is a simply connected open region, and $\omega = P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y$ be a C^1 1-form on U. Then the following statements are equivalent:

(i) ω is exact on U, i.e., there exists a C^2 function $F(x,y)$ on U, such that

$$\mathrm{d}F = \omega = P\,\mathrm{d}x + Q\,\mathrm{d}y.$$

At this time, $F(x,y)$ is called a **potential function** of ω on U and

$$F(x,y) = \int_{(x_0,y_0)}^{(x,y)} \omega + C = \int_{x_0}^{x} P(t,y_0)\,\mathrm{d}t + \int_{y_0}^{y} Q(x,s)\,\mathrm{d}s + C,$$

where (x_0,y_0) is a fixed point in U and C is an arbitrary constant.

(ii) ω is closed on U, i.e.,

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

(iii) The line integral of ω is path-independent on U.

(iv) For any piecewise smooth simple closed curve $\vec{L}$ on U,

$$\oint_{\vec{L}} \omega = 0.$$

Example 15.1 Calculate

$$I = \oint_{\vec{C}} \frac{\cos(\mathbf{r}, \mathbf{n})}{r} \, \mathrm{d}s,$$

where $\vec{C}$ is piecewise smooth simple closed curve, $\mathbf{r} = (x, y)$, $r = \|\mathbf{r}\| = \sqrt{x^2 + y^2}$, and $\mathbf{n}$ is the unit outward normal vector of $\vec{C}$.

Chapter 16 Integrals with Variable Parameters

16.1 Definite Integrals with Variable Parameters

Definition 16.1 (Definite Integral with Variable Parameters)

Let $f(x, y)$ be defined on $[a, b] \times [c, d]$. For each fixed $y \in [c, d]$, if the definite integral

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

exists, then $I(y)$ is called a **definite integral with variable parameter** y. ♣

16.2 Elliptic Integrals

16.3 Improper Integrals with Variable Parameters

There are two types of improper integrals with variable parameters: on infinite interval and with unbounded integrand. Here we only give the definition of improper integrals on infinite interval with variable parameters.

Definition 16.2 (Improper Integral with Variable Parameters)

Let $f(x, y)$ be defined on $[a, +\infty) \times [c, d]$. For some fixed $y_0 \in [c, d]$, if the improper integral $I(y_0) = \int_a^{+\infty} f(x, y_0)\,\mathrm{d}x$ converges, then $\int_a^{+\infty} f(x, y)\,\mathrm{d}x$ is called convergent at y_0, and y_0 is called its convergence point.

Let the set of all convergence points be E, then E is the domain of definition of the improper integral with variable parameters

$$I(y) = \int_a^{+\infty} f(x, y)\,\mathrm{d}x,$$

also called the convergence domain of the improper integral $\int_a^{+\infty} f(x, y)\,\mathrm{d}x$. ♣

¶ **Uniform Convergence and Its Tests**

Definition 16.3 (Uniform Convergence of Improper Integrals with Variable Parameters)

Let $f(x, y)$ be defined on $[a, +\infty) \times [c, d]$, where $[c, d]$ is the convergence domain of the improper integral $\int_a^{+\infty} f(x, y)\,\mathrm{d}x$. If for every $\varepsilon > 0$, there exists a number $A_0 > a$ independent of y, such that for all $A > A_0$ and for all $y \in [c, d]$,

$$\left|\int_a^A f(x, y)\,\mathrm{d}x - I(y)\right| = \left|\int_A^{+\infty} f(x, y)\,\mathrm{d}x\right| < \varepsilon,$$

then the improper integral $\int_a^{+\infty} f(x, y)\,\mathrm{d}x$ is said to be **uniformly convergent** on $[c, d]$. ♣

Theorem 16.1 (Cauchy Criterion for Uniform Convergence of Improper Integrals with Variable Parameters)

Let $f(x,y)$ be defined on $[a,+\infty)\times[c,d]$, where $[c,d]$ is the convergence domain of the improper integral $\int_a^{+\infty} f(x,y)\,\mathrm{d}x$. The improper integral $\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ is uniformly convergent on $[c,d]$ if and only if for every $\varepsilon>0$, there exists a number $A_0>a$ independent of y, such that for all $A_1, A_2 > A_0$ and for all $y\in[c,d]$,

$$\left|\int_{A_1}^{A_2} f(x,y)\,\mathrm{d}x\right| < \varepsilon.$$

16.4 Analysis Properties of Uniform Convergence

Lemma 16.1

Theorem 16.2 (Uniform Convergence and Continuity)

Let $f(x,y)$ be continuous on $[a,+\infty)\times[c,d]$, and $\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ is uniformly convergent on $[c,d]$ with respect to y, then:

(i)

$$I(y)=\int_a^{+\infty} f(x,y)\,\mathrm{d}x$$

is continuous on $[c,d]$, i.e.,

$$\lim_{y\to y_0}\int_a^{+\infty} f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\lim_{y\to y_0} f(x,y)\,\mathrm{d}x,\quad y_0\in[c,d],$$

that is, the limit and the integral can be interchanged.

(ii)

$$\int_c^d \mathrm{d}y\int_a^{+\infty} f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_c^d f(x,y)\,\mathrm{d}y,$$

that is, the order of integration can be interchanged.

When $[c,d]$ is replaced by $[c,+\infty)$, the above theorem fails, but we have the following theorem.

Theorem 16.3

On the region $D=[a,+\infty)\times[c,+\infty)$,

1. if $f(x,y)$ satisfies:
   (a). $f(x,y)\in C(D)$;
   (b). $\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ internally closed uniformly converges with respect to y; $\int_c^{+\infty} f(x,y)\,\mathrm{d}y$ internally closed uniformly converges with respect to x;
   (c). One of the two integrals $\int_a^{+\infty}\mathrm{d}x\int_c^{+\infty}|f(x,y)|\,\mathrm{d}y$ or $\int_c^{+\infty}\mathrm{d}y\int_a^{+\infty}|f(x,y)|\,\mathrm{d}x$ converges;

   then
   $$\int_c^{+\infty}\mathrm{d}y\int_a^{+\infty} f(x,y)\,\mathrm{d}x=\int_a^{+\infty}\mathrm{d}x\int_c^{+\infty} f(x,y)\,\mathrm{d}y$$
2. if $f(x,y)$ satisfies:
   (a). $f(x,y)\in C(D)$ and $f(x)\geqslant 0$ on D;

(b). $\int_a^{+\infty} f(x,y)\,\mathrm{d}x \in C[c,+\infty)$; $\int_c^{+\infty} f(x,y)\,\mathrm{d}y \in C[a,+\infty)$;
(c). One of the two integrals $\int_a^{+\infty}\mathrm{d}x\int_c^{+\infty} f(x,y)\,\mathrm{d}y$ or $\int_c^{+\infty}\mathrm{d}y\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ converges;

then
$$\int_c^{+\infty}\mathrm{d}y\int_a^{+\infty} f(x,y)\,\mathrm{d}x = \int_a^{+\infty}\mathrm{d}x\int_c^{+\infty} f(x,y)\,\mathrm{d}y$$

Remark One of the two integrals exists implies the other exists as well as the equality holds.

Theorem 16.4 (Uniform Convergence and Differentiation)

On the region $D=[a,+\infty]\times[c,d]$, if the following conditions are satisfied:
(i) $\frac{\partial}{\partial y} f(x,y)\in C(D)$;
(ii) $\int_a^{+\infty}\frac{\partial}{\partial y} f(x,y)\,\mathrm{d}x$ converges uniformly with respect to y on $[c,d]$;
(iii) There exists a point $y_0\in[c,d]$, such that $\int_a^{+\infty} f(x,y_0)\,\mathrm{d}x$ converges;
(iv) For any $[\alpha,\beta]\subset[a,+\infty)$, $\int_\alpha^\beta f(x,y)\,\mathrm{d}x$ exists.

Then $I(y)=\int_a^{+\infty} f(x,y)\,\mathrm{d}x$ is differentiable on $[c,d]$, and
$$\frac{\mathrm{d}}{\mathrm{d}y}\int_a^{+\infty} f(x,y)\,\mathrm{d}x = \int_a^{+\infty}\frac{\partial}{\partial y} f(x,y)\,\mathrm{d}x.$$

Example 16.1 Let
$$F(a)=\int_0^{+\infty}\frac{1}{t}(1-e^{-at})\cos bt\,\mathrm{d}t,\quad b\neq 0.$$

1. Prove that $F(a)\in C[0,+\infty)\cap D(0,+\infty)$
2. Find the expression of $F(a)$.

¶ `Imbedding Method`

If $I=\int_a^b f(x)\mathrm{d}x$ is difficult to calculate directly, we can introduce a parameter y and consider the integral
$$I(y)=\int_a^b f(x,y)\mathrm{d}x,$$
and let $I=I(y_0)$ for some specific y_0. If we can calculate $I(y)$ and then take $y=y_0$, then we can get the value of I. This method is called **imbedding method**.

Example 16.2 Compute the integral:
$$I=\int_0^1\frac{\ln(1+x)}{1+x^2}\,\mathrm{d}x.$$

Example 16.3 Compute Dirichlet's integral:
$$I=\int_0^{+\infty}\frac{\sin x}{x}\,\mathrm{d}x.$$

Solution

□

16.5 Euler Integrals

Beta Function

Beta function can be defined in the following equivalent forms:

1. For $p > 0, q > 0$:
$$B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,\mathrm{d}t.$$
2. Via substitution $t = \frac{u}{1+u}$:
$$B(p,q) = \int_0^{+\infty} \frac{u^{p-1}}{(1+u)^{p+q}}\,\mathrm{d}u = \int_0^{+\infty} \frac{u^{q-1}}{(1+u)^{p+q}}\,\mathrm{d}u.$$
3. Via substitution $t = \sin^2\theta$:
$$B(p,q) = 2\int_0^{\frac{\pi}{2}} \sin^{2p-1}\theta\cos^{2q-1}\theta\,\mathrm{d}\theta.$$

Then we have:
$$B(\frac{1}{2},\frac{1}{2}) = \pi, \quad B(\frac{3}{2},\frac{1}{2}) = \frac{\pi}{2}, \quad B(1,1) = 1.$$

Property

Continuity $B(p,q) \in C(U)$, *where* $U = \{(p,q)|p > 0, q > 0\}$.

Symmetry $B(p,q) = B(q,p)$.

Recurrence Relation $B(p,q) = \frac{q-1}{p+q-1}B(p,q-1)$ *for* $p > 0, q > 1$.

Gamma Function

Gamma function can be defined in the following equivalent forms:

1. For $s > 0$:
$$\Gamma(s) = \int_0^{+\infty} x^{s-1}e^{-x}\,\mathrm{d}x.$$
2. Via the limit:
$$\Gamma(s) = \lim_{n\to\infty} \frac{n!}{s(s+1)(s+2)\cdots(s+n)}.$$
3. Via substitution $x = t^2$:
$$\Gamma(s) = \frac{1}{2}\int_0^{+\infty} t^{2s-1}e^{-t^2}\,\mathrm{d}t.$$

Then we have:
$$\Gamma(\frac{1}{2}) = \sqrt{\pi}, \quad \Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}, \quad \Gamma(1) = 1.$$

Property

Continuity $\Gamma(s) \in C(0,+\infty)$.

Recurrence Relation $\Gamma(s+1) = s\Gamma(s)$ *for* $s > 0$.

Gamma function can be extended to the whole complex plane except for non-positive integers, where it has simple poles.

Relation between Beta and Gamma Functions

Theorem 16.5

There holds the following relation between Beta and Gamma functions:

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p>0, q>0.$$

Next, we give three important formulas about Gamma function, which can be extended to the complex domain as well.

Theorem 16.6 (Bohr-Mollerup Theorem)

The Gamma function is the unique function defined on $(0,+\infty)$ satisfying the following three conditions:

(i) $f(x)>0$ and $f(1)=1$;

(ii) $f(x+1)=xf(x)$ for all $x>0$;

(iii) $\ln f(x)$ is convex on $(0,+\infty)$.

Theorem 16.7 (Legendre's Duplication Formula)

For $s>0$, there holds:

$$\Gamma(s)\Gamma(s+\frac{1}{2}) = \frac{\sqrt{\pi}}{2^{2s-1}}\Gamma(2s).$$

Theorem 16.8 (Reflection Formula)

For $0<s<1$, there holds:

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}.$$

Theorem 16.9 (Stirling's Formula)

$$\Gamma(s+1) = \sqrt{2\pi s}\left(\frac{s}{e}\right)^s \exp\left(\frac{\theta}{12s}\right),$$

where $0<\theta<1$.

Specially, when $s=n\in\mathbb{N}$,

$$\Gamma(n+1) = n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(\frac{\theta}{12n}\right),$$

where $0<\theta<1$.

Example 16.4 Prove the integral form of Riemann ζ function:

$$\zeta(s) = \sum_{n=1}^{\infty}\frac{1}{n^s} = \frac{1}{\Gamma(s)}\int_0^{+\infty}\frac{x^{s-1}}{e^x-1}\,\mathrm{d}x, \quad s>1.$$

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/.