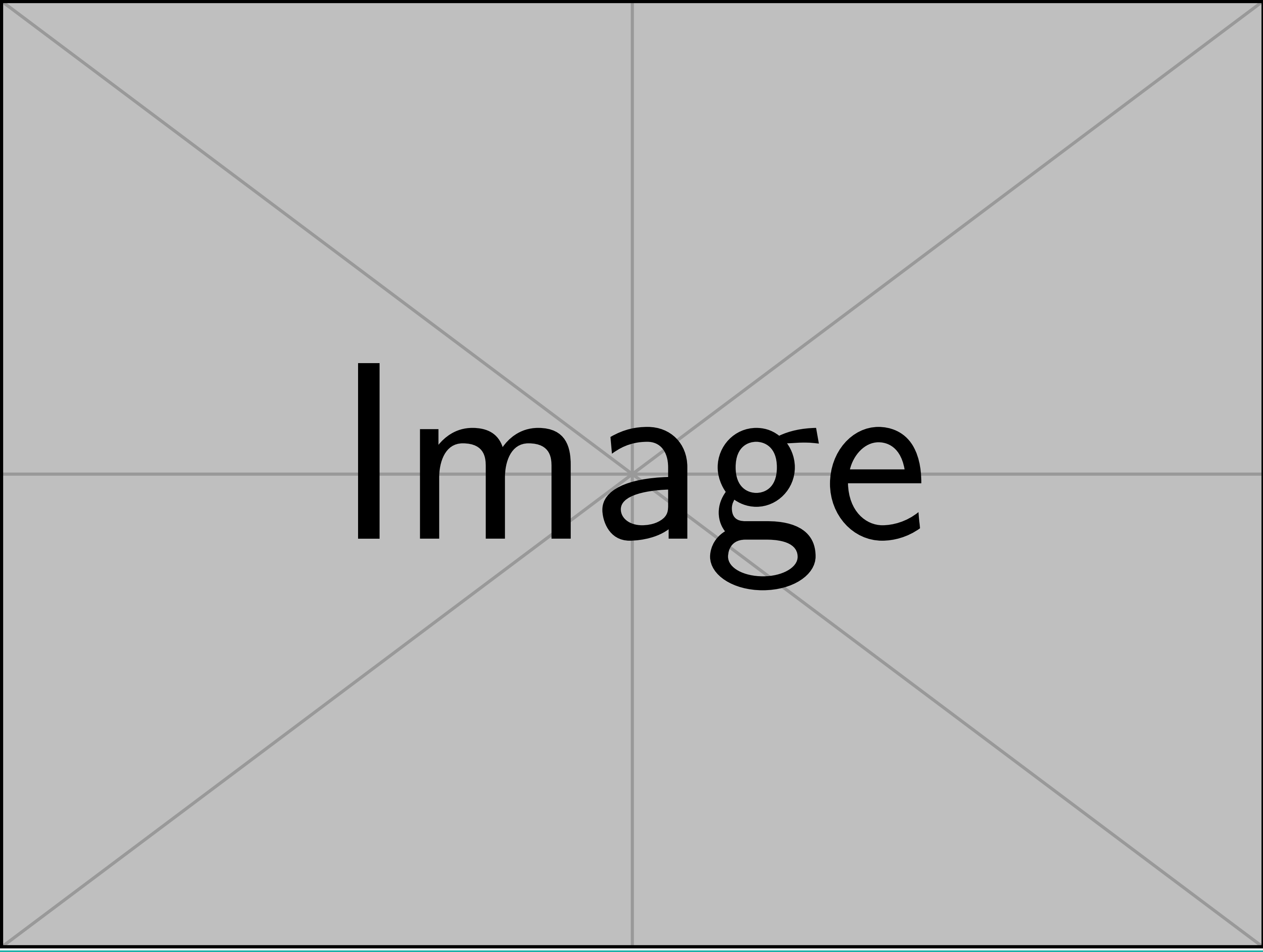


Analyse Mathématique

Author: CatMono

Date: July, 2025

Version: 0.1

Contents

Preface

This is the preface of the book…

Chapter 1 Preliminaries

1.1 Trigonometric Formulas

Product-to-Sum Formulas:

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

Sum and Difference Formulas:

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

Sum-to-Product Formulas:

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

Double Angle Formulas:

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

Half Angle Formulas:

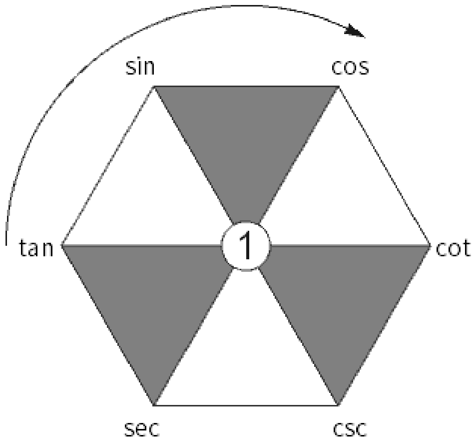
$$\sin\frac{\alpha}{2} = \pm\sqrt{\frac{1-\cos\alpha}{2}}$$
$$\cos\frac{\alpha}{2} = \pm\sqrt{\frac{1+\cos\alpha}{2}}$$
$$\tan\frac{\alpha}{2} = \frac{1-\cos\alpha}{\sin\alpha} = \frac{\sin\alpha}{1+\cos\alpha}$$

Power-Reducing Formulas:

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

Angle Decomposition Formulas:

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



Remark

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

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:

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

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

Chapter 3 Limits and Continuity of Functions

3.1 Limits of Functions

¶ **Definition of Limit**

¶ **Limits of Functions and Sequences**

Theorem 3.1 (Heine Theorem)

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

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

♡

3.2 Continuous Functions

3.3 Infinitesimal and Infinite Quantities

3.4 Continuous Functions on Closed Intervals

¶ **Concerning Theorems**

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

♡

Theorem 3.3 (Zero Point Existence Theorem)

♡

¶ **Uniform Continuity and Lipschitz Continuity**

Definition 3.1 (Uniform Continuity)

♣

Theorem 3.4 (Uniform Continuity Theorem)

♡

Theorem 3.5 (Cantor's Theorem)

♡

Definition 3.2 (Lipschitz Continuity)

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

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

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

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

♣

Remark

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

3.5 Period Three Implies Chaos

3.6 Functional Equations

Chapter 4 Differential

4.1 Differential and Derivative

¶ Basic Differential Rules and Formulas

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

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

4.2 Higher-Order Derivatives

4.3 Differential Mean Value Theorems

Definition 4.1 (Extremum)

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

Lemma 4.1 (Fermat's Lemma)

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

Theorem 4.1 (Rolle's Theorem)

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

Theorem 4.2 (Lagrange's Mean Value Theorem)

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

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

Theorem 4.3 (Cauchy's Mean Value Theorem)

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

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

4.4 Theorems and Applications concerning Derivatives

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

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

Theorem 4.5 (Theorem on the Limit of Derivatives)

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

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

4.5 Taylor Theorem

4.6 Applications of Taylor Theorem

Chapter 5 Indefinite Integral

5.1 Two Common Integration Methods

¶ **Integration Methods**

Definition 5.1 (Integration by Parts)

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

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

♣

¶ **Basic Integration Formulas**

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

Chapter 6 Definite Integral

6.1 Riemann Integral

¶ Riemann Integral

Definition 6.1 (Riemann Integral)

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

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

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

The summation

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

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

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

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

Alternatively, it can also be expressed as:

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

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

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

¶ Riemann-Stieltjes Integral

Definition 6.3 (Riemann-Stieltjes Integral)

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

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

The properties of Darboux sums also apply to Stieltjes sums.

6.2 Integrability Criteria

¶ Common Integrability Criteria

Theorem 6.1 (Integrability Criterion)

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

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

Corollary 1 Continuous functions on closed intervals are necessarily integrable.

Corollary 2 Monotonic functions on closed intervals are necessarily integrable.

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

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

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

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

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

Proof

∎

¶ `Lesbesgue's Theorem`

Theorem 6.2 (Lesbesgue's Theorem)

6.3 Properties of Definite Integrals

¶ `Properties of Riemann Integrals`

¶ `Integrability of Composite Functions`

6.4 Fundamental Theorem of Calculus

¶ `Newton-Leibniz Formula`

¶ `Riemann Lemma`

¶ `Common Questions concerning Integrals`

6.5 Calculation of Definite Integrals

6.6 Integral Inequalities

6.7 Applications of Definite Integrals

Chapter 7 Improper Integral

Chapter 8 Numerical Series

8.1 Convergence of Numerical Series

8.2 Positive Term Series

8.3 General Term Series

¶ `Cauchy Convergence Criterion for Series`

¶ `Alternative Series`

¶ `Abel-Dirichlet Test`

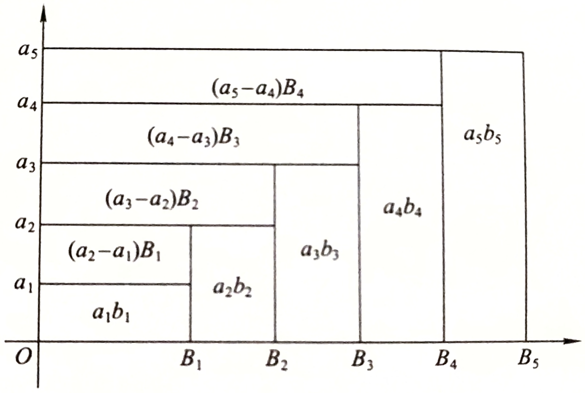
Theorem 8.1 (Abel Transform (Discrete Integration by Parts/Summation by Parts))

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

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



Chapter 9 Series of Functions

Chapter 10 Power Series

Chapter 11 Limits and Continuity in Euclidean Spaces

11.1 Continuous Mappings

¶ Continuous Mappings on Compact Sets

¶ Continuous Mappings on Connected Sets

Definition 11.1 (Connected Set)

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

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

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

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

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

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

Chapter 12 Multi-variable Differential Calculus

12.1 Directional Derivatives and Total Differential

¶ Directional Derivative

Definition 12.1 (Directional Derivative)

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

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

If the derivative of u at $t = 0$

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

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

♣

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

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

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

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

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

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

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

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

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

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

Definition 12.2 (Jacobian Matrix (Gradient))

Let

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

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

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

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

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

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

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

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

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

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

¶ Total Differential

Definition 12.3 (Total Differential)

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

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

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

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

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

Theorem 12.1 (Conditions of Differentiability)

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

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

[a] However, the converse is not true.

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

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

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

Note

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

Proof Taking a function of three variables as an example.

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

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

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

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

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

$$I_1 = f_x(x_0 + \theta_1 \Delta x, y_0 + \Delta y, z_0 + \Delta z)\Delta x,$$
$$I_2 = f_y(x_0, y_0 + \theta_2 \Delta y, z_0 + \Delta z)\Delta y.$$

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($\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$):

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

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

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

Accordingly,

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

Apparently,

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

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

Note *(At some point)*

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

¶ Higher-Order Partial Derivatives and Differential

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

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

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

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

Theorem 12.2 (Conditions for Equality of Mixed Partial Derivatives)

1. Let $U\subset\mathbb{R}^2$ be an open set, and $f:U\to\mathbb{R}$ be a function of two variables. If f_{xy} and f_{yx} are continuous at $(x_0,y_0)\in U$, then
$$f_{xy}(x_0,y_0)=f_{yx}(x_0,y_0).$$
2. Let $U\subset\mathbb{R}^n$ be an open set, and $f:U\to\mathbb{R}$ be a function of n variables. If f has partial derivatives up to order k in D, and all of them are continuous at $\mathbf{x}^0=(x_1^0,x_2^0,\ldots,x_n^0)\in U$, then
$$\frac{\partial^l f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_l}}(\mathbf{x}^0)=\frac{\partial^l f}{\partial x_{i_2}\partial x_{i_1}\cdots\partial x_{i_l}}(\mathbf{x}^0)=\cdots=\frac{\partial^l f}{\partial x_{i_l}\partial x_{i_{l-1}}\cdots\partial x_{i_1}}(\mathbf{x}^0),$$
that is, the order of taking partial derivatives $l(\le k)$ does not affect the result.[a]

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

Proof When $k\neq 0, h\neq 0$, define
$$\varphi(y)=f(x_0+h,y)-f(x_0,y),$$
and
$$\psi(x)=f(x,y_0+k)-f(x,y_0).$$
Applying the Lagrange Mean Value Theorem, we have
$$\begin{aligned}&[f(x_0+h,y_0+k)-f(x_0,y_0+k)]-[f(x_0+h,y_0)-f(x_0,y_0)]\\=&\varphi(y_0+k)-\varphi(y_0)\\=&\varphi'(y_0+\theta_1 k)k\\=&[f_y(x_0+h,y_0+\theta_1 k)-f_y(x_0,y_0+\theta_1 k)]k\\=&f_{yx}(x_0+\theta_2 h,y_0+\theta_1 k)hk,\quad 0<\theta_1,\theta_2<1.\end{aligned}$$
On the other hand,
$$\begin{aligned}&[f(x_0+h,y_0+k)-f(x_0,y_0+k)]-[f(x_0+h,y_0)-f(x_0,y_0)]\\=&[f(x_0+h,y_0+k)-f(x_0+h,y_0)]-[f(x_0,y_0+k)-f(x_0,y_0)]\\=&\psi(x_0+h)-\psi(x_0)\\=&\psi'(x_0+\theta_3 h)h\\=&[f_x(x_0+\theta_3 h,y_0+k)-f_x(x_0+\theta_3 h,y_0)]h\\=&f_{xy}(x_0+\theta_3 h,y_0+\theta_4 k)hk,\quad 0<\theta_3,\theta_4<1.\end{aligned}$$
Therefore,
$$f_{xy}(x_0+\theta_3 h,y_0+\theta_4 k)=f_{yx}(x_0+\theta_2 h,y_0+\theta_1 k).$$

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

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

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

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

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

If z also has continuous second-order partial derivatives, then $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are also differentiable, and thus $\mathrm{d}z$ is differentiable. We call the differential of $\mathrm{d}z$ the second-order differential of z, denoted as

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

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

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

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

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

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

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

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

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

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

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

Similarly, we define

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

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 $u = f(x_1, x_2, \ldots, x_n)$, higher-order differentials can be similarly defined, and the following holds:

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

12.2 Differential of Vector-Valued Functions

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

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

Expressed in coordinate vector form:

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

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

12.3 Derivatives of Composite Mappings (Chain Rule)

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

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

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

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

Note

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

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

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

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

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

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

Proof

■

12.4 Mean Value Theorem and Taylor's Formula

Definition 12.4 (Convex Region)

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

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

then D is called a convex region. ♣

Theorem 12.3 (Lagrange's Mean Value Theorem)

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

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

Proof

■

Theorem 12.4

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

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

then f is constant on D.

Proof

■

Theorem 12.5 (Taylor's Formula)

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

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

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

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

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

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

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

where the matrix

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

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

12.5 Implicit Function Theorem

Theorem 12.6 (Implicit Function Theorem)

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

1. $F \in C^k(U, \mathbb{R})$, where $1 \le k \le +\infty$;
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 $y = f(x)$.

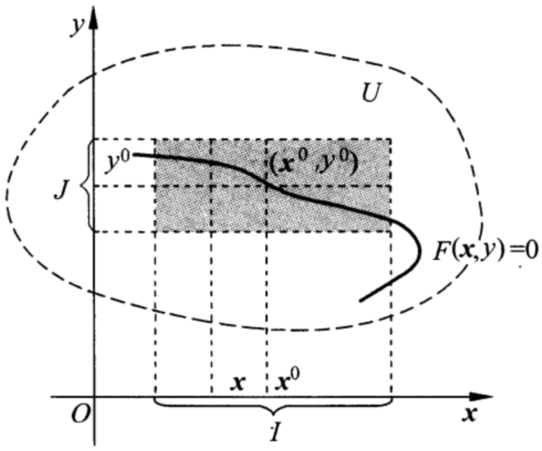


Figure 12.1: Implicit Function

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:

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

where the following holds:

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

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

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

it follows that:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Thus:

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

The proof is complete. ■

Theorem 12.7 (Implicit Mapping Theorem)

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

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

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

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

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

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

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

12.6 Applications of Multi-Variable Differential Calculus

¶ Surface and Tangent Space

Definition 12.5 (Parameterization of Surface)

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

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

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

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

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

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

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

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

Definition 12.6 (Tangent Vector)

♣

¶ **Unconditional Extremum**

¶ **Conditional Extremum**

Chapter 13 Multiple Integrals

Bibliography

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

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

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

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

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

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

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

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

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