Calculus 微積分

Ching-hsiao Arthur Cheng 鄭經斅

Chapter 0

Preliminary

0.1 Functions and Their Graphs

Definition 0.1: Real-Valued Functions of a Real Variable

Let $X, Y \subseteq \mathbb{R}$ be subsets of real numbers. A real-valued function f of a real variable x from X to Y is a correspondence that assigns to each element x in X exactly one number y in Y. Here X is called the domain of f and is usually denoted by $\mathrm{Dom}(f)$, Y is called "the" co-domain of f, the number y is called the image of x under f and is usually denoted by $f(x)$, which is called the value of f at x. The range of f, denoted by $\mathrm{Ran}(f)$, is a subset of Y consisting of all images of numbers in X. In other words,

$$\mathrm{Ran}(f) \equiv \text{the range of } f \equiv \big\{ f(x) \,\big|\, x \in X \big\}\,.$$

Remark 0.2. Given a way of assignment $x \mapsto f(x)$ without specifying where x is chosen from, we still treat f as a function and $\mathrm{Dom}(f)$ is considered as the collection of all $x \in \mathbb{R}$ such that $f(x)$ is well-defined. For example, $f(x) = x + 1$ and $g(x) = \dfrac{x^2 - 1}{x - 1}$ are both considered as functions with

$$\mathrm{Dom}(f) = \mathbb{R} \qquad \text{and} \qquad \mathrm{Dom}(g) = \mathbb{R} \backslash \{1\}\,.$$

Since $\mathrm{Dom}(f) \neq \mathrm{Dom}(g)$, f and g are considered as different functions even though $f(x) = g(x)$ for all $x \neq 1$.

Terminologies:

1. Explicit form of a function: $y = f(x)$;

2. Implicit form of a function: $F(x, y) = 0$. (参考影片)

Definition 0.3

A function f is a polynomial function if f takes the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \,,$$

where $a_0, a_1, a_2, \cdots, a_n$ are real numbers, called coefficients of the polynomial, and n is a non-negative integer. If $a_n \neq 0$, then a_n is called the leading coefficient, and n is called the degree of the polynomial. A rational function is the quotient of two polynomials.

Definition 0.4

The graph of the function $y = f(x)$ consists of all points $\big(x, f(x)\big)$, where x is in the domain of f. In other words,

$$\mathrm{G}(f) \equiv \text{the graph of } f \equiv \Big\{\big(x, f(x)\big) \,\Big|\, x \in \mathrm{Dom}(f)\Big\} \,.$$

Definition 0.5: Composite Functions

Let f and g be functions. The function $f \circ g$, read f circle g, is the function defined by $(f \circ g)(x) = f\big(g(x)\big)$. The domain of $f \circ g$ is the set of all x in the domain of g such that $g(x)$ is in the domain of f. In other words,

$$\mathrm{Dom}(f \circ g) = \big\{x \in \mathrm{Dom}(g) \,\big|\, g(x) \in \mathrm{Dom}(f)\big\} \,.$$

0.2 Trigonometric Functions

Definition 0.6

An angle consists of an initial ray, a terminal ray and a vertex where two rays intersects. An angle is in standard position when its initial ray coincides with the positive x-axis and its vertex is at the origin. Positive angles are measured counterclockwise, and negative angles are measured clockwise.

Let θ be a central angle of a circle of radius 1. The radian measure of θ is defined to be the length of the arc of the sector.

Remark 0.7. Using radian measure of θ, the length s of a circular arc of radius r is given by $s = r\theta$.

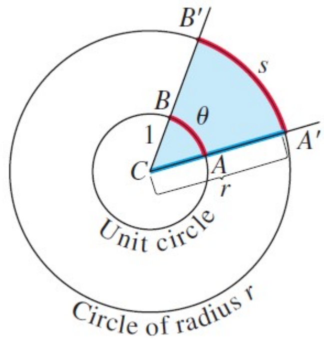


Figure 1: The radian measure of the central angle $A'CB'$ is the number $u = s/r$. For a unit circle of radius $r = 1$, u is the length of arc AB that central angle ACB cuts from the unit circle.

Remark 0.8. For a point P on the plane with Cartesian coordinate (x, y), let $r = \sqrt{x^2 + y^2}$ and θ be the angle in standard position with $\overrightarrow{OP}$ as the terminal ray. The ordered pair (r, θ) is called the ***polar coordinate*** of the point P.

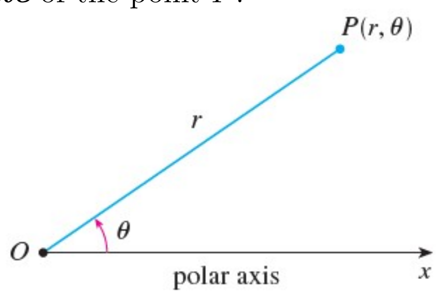


Figure 2: Polar coordinate

Definition 0.9

Let θ be an angle in standard position, and the terminal ray intersects the circle centered at the origin of radius r at point (x, y). The trigonometric functions sine, cosine, tangent, cotangent, secant and cosecant, abbreviated as sin, cos, tan, cot, sec and csc, respectively, of angle θ are defined by

$$\sin\theta = \frac{y}{r}, \quad \cos\theta = \frac{x}{r}, \quad \tan\theta = \frac{y}{x}, \quad \cot\theta = \frac{x}{y}, \quad \sec\theta = \frac{r}{x} \quad \text{and} \quad \csc\theta = \frac{r}{y},$$

provided that the quotients make sense.

Remark 0.10. Suppose that a point P has polar coordinate (r, θ). Then the Cartesian coordinate of P is $(r\cos\theta, r\sin\theta)$.

Proposition 0.11: Properties of Trigonometric Functions

1. For all real numbers θ,

$$\sin^2\theta + \cos^2\theta = 1\,, \quad 1 + \tan^2\theta = \sec^2\theta\,, \quad 1 + \cot^2\theta = \csc^2\theta\,.$$

2. For all real numbers θ,

$$\begin{aligned} &\sin(-\theta) = -\sin\theta\,, \quad \cos(-\theta) = \cos\theta\,, \quad \tan(-\theta) = -\tan\theta\,,\\ &\cot(-\theta) = -\cot\theta\,, \quad \sec(-\theta) = \sec\theta\,, \quad \csc(-\theta) = -\csc\theta\,. \end{aligned}$$

3. For all real numbers θ,

$$\begin{aligned} &\sin\left(\theta + \frac{\pi}{2}\right) = \cos\theta\,, \quad \cos\left(\theta + \frac{\pi}{2}\right) = -\sin\theta\,, \quad \tan\left(\theta + \frac{\pi}{2}\right) = -\cot\theta\,,\\ &\sin(\theta + \pi) = -\sin\theta\,, \quad \cos(\theta + \pi) = -\cos\theta\,, \quad \tan(\theta + \pi) = \tan\theta\,. \end{aligned}$$

4. (**Law of Cosines**): Let a, b, c be the length of sides of a triangle, and θ be the angle opposite to the side with length c. Then $c^2 = a^2 + b^2 - 2ab\cos\theta$.

5. (**Sum and Difference Formulas**): Let θ, ϕ be real numbers. Then

$$\sin(\theta \pm \phi) = \sin\theta\cos\phi \pm \sin\phi\cos\theta\,, \quad \cos(\theta \pm \phi) = \cos\theta\cos\phi \mp \sin\theta\sin\phi\,.$$

6. (**Double-Angle Formulas**): For all real numbers θ,

$$\sin(2\theta) = 2\sin\theta\cos\theta\,, \quad \cos(2\theta) = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta\,.$$

7. (**Half-Angle Formulas**): For all real numbers θ,

$$\cos^2\frac{\theta}{2} = \frac{1+\cos\theta}{2}\,, \quad \sin^2\frac{\theta}{2} = \frac{1-\cos\theta}{2}\,, \quad \tan\frac{\theta}{2} = \frac{\sin\theta}{1+\cos\theta}\,.$$

8. (**Triple-Angle Formulas**): For all real numbers θ,

$$\cos(3\theta) = 4\cos^3\theta - 3\cos\theta\,, \quad \sin(3\theta) = 3\sin\theta - 4\sin^3\theta\,.$$

9. (**Sum-to-Product Formulas**): For all real numbers θ and ϕ,

$$\begin{aligned} &\sin\theta + \sin\phi = 2\sin\frac{\theta+\phi}{2}\cos\frac{\theta-\phi}{2}\,, \quad \sin\theta - \sin\phi = 2\sin\frac{\theta-\phi}{2}\cos\frac{\theta+\phi}{2}\,,\\ &\cos\theta + \cos\phi = 2\cos\frac{\theta+\phi}{2}\cos\frac{\theta-\phi}{2}\,, \quad \cos\theta - \cos\phi = 2\sin\frac{\theta+\phi}{2}\sin\frac{\phi-\theta}{2}\,. \end{aligned}$$

Theorem 0.12: de Moivre (棣美弗)

For each real number θ and natural number n,

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)\,. \tag{0.2.1}$$

Proof. Clearly (0.2.1) holds for $n = 1$. Suppose that (0.2.1) holds for $n = k$ for some natural number k. Then by the sum and difference formulas,

$$\begin{aligned}(\cos\theta + i\sin\theta)^{k+1} &= (\cos\theta + i\sin\theta)^k \cdot (\cos\theta + i\sin\theta)\\ &= \big[\cos(k\theta) + i\sin(k\theta)\big] \cdot (\cos\theta + i\sin\theta)\\ &= \cos(k\theta)\cos\theta - \sin(k\theta)\sin\theta + i\big[\sin(k\theta)\cos\theta + \cos(k\theta)\sin\theta\big]\\ &= \cos[(k+1)\theta] + i\sin[(k+1)\theta]\end{aligned}$$

which shows that (0.2.1) holds for $n = k+1$. By induction, we find that (0.2.1) holds for all natural number n. □

Theorem 0.13

Let θ be a real number and $0 \leqslant \theta < \dfrac{\pi}{2}$. Then

$$\sin\theta \leqslant \theta \leqslant \tan\theta\,. \tag{0.2.2}$$

Proof. Inequality (0.2.2) follows from the following figure

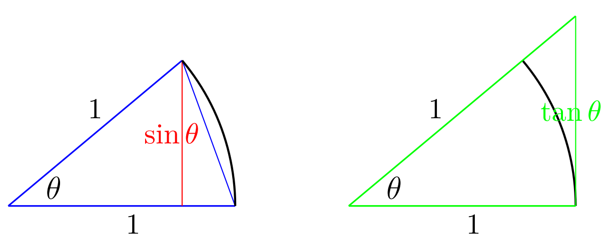


Figure 3: The area of the sector is larger than the area of the blue triangle but is smaller than the green triangle

which shows $\dfrac{1}{2}\sin\theta \leqslant \dfrac{1}{2}\theta \leqslant \dfrac{1}{2}\tan\theta$. □

0.3 Exercise

Problem 0.1. Let θ be a real number such that $t = \tan\dfrac{\theta}{2}$ also be a real number. Show that

$$\sin\theta = \frac{2t}{1+t^2} \qquad \text{and} \qquad \cos\theta = \frac{1-t^2}{1+t^2}\,.$$

Chapter 1

Limits and Continuity

1.1 Limits of Functions

Goal: Given a function f defined "near c", find the value of f at x when x is "arbitrarily close" to c. (給定一函數 f，我們想知道「當除 c 之外的點到 c 的距離愈來愈近時，其函數值是否向某數集中」)

Notation: When there exists such a value, the value is denoted by $\lim\limits_{x \to c} f(x)$.

Example 1.1. Consider the function $g(x) = \dfrac{x^2-1}{x-1}$ given in Remark 0.2, and

$$h(x) = \begin{cases} \dfrac{x^2-1}{x-1} & \text{if } x \neq 1\,, \\ 0 & \text{if } x = 1\,. \end{cases}$$

Then the limit of g at 1 should be the same as the limit of h at 1. Therefore, to consider the limit of a function at a point c, the value of the function at c is not important at all.

Example 1.2. Let $g(x) = \dfrac{x^2-1}{x-1}$. Then $\text{Dom}(g) = \mathbb{R}\backslash\{1\}$ and $g(x) = x+1$ if $x \neq 1$. Therefore, the graph of g is given by

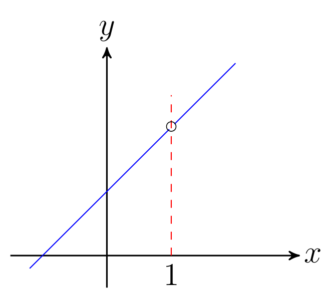


Figure 1.1: The graph of function $g(x) = \dfrac{x^2-1}{x-1}$

Then (by looking at the graph of g we find that) $\lim_{x \to 1} g(x) = 2$.

Example 1.3. Let $f(x) = \begin{cases} 1 & \text{if } x \neq 2\,, \\ 0 & \text{if } x = 2\,. \end{cases}$ The graph of f is given by

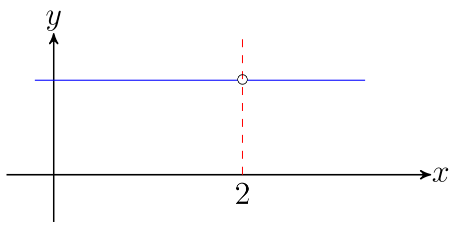


Figure 1.2: The graph of function $f(x)$

Then (by looking at the graph of f we find that) $\lim_{x \to 2} f(x) = 1$.

Next we give some examples in which the limit of functions (at certain points) do not exist.

Example 1.4. （詳見影片）Let $f(x) = \sin \frac{1}{x}$. Then $\mathrm{Dom}(f) = \mathbb{R} \backslash \{0\}$. For the graph of f, we note that if $x \in I_n \equiv \left[\frac{1}{2n\pi + 2\pi}, \frac{1}{2n\pi}\right]$ for some $n \in \mathbb{N}$, the graph of f on I_n must touch $x = 1$ and $x = -1$ once. Therefore, the graph of f looks like

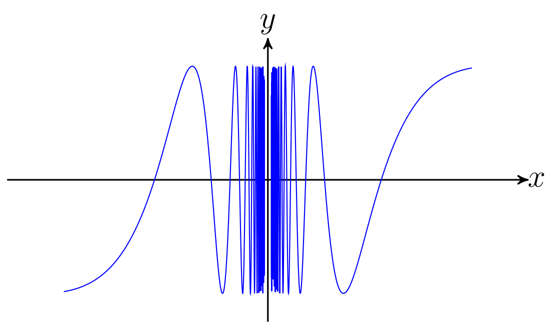


Figure 1.3: The graph of function $f(x) = \sin \frac{1}{x}$

In any interval containing 0, there are infinitely many points whose image under f is 1, and there are always infinitely many points whose image under f is -1. In fact, in any interval containing 0 and $L \in [-1, 1]$ there are infinitely many points whose image under f is L. Therefore, $\lim_{x \to 0} f(x)$ D.N.E. (does not exist).

Example 1.5. Let $f(x) = \frac{|x|}{x}$. Then $f(x) = 1$ if $x > 0$, $f(x) = -1$ if $x < 0$, and the graph of f is given by

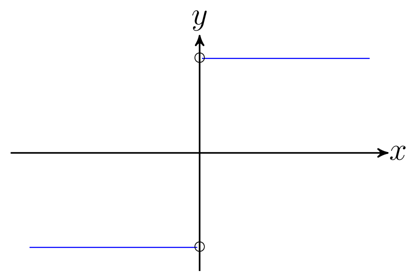


Figure 1.4: The graph of function $f(x) = \dfrac{|x|}{x}$

By observation (that is, looking at the graph of f), $\lim_{x\to 0} f(x)$ D.N.E.

Example 1.6. (詳見影片) Consider the Dirichlet function

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q}\,, \\ 1 & \text{if } x \notin \mathbb{Q}\,, \end{cases}$$

where $\mathbb{Q}$ denotes the collection of rational numbers (有理數). Then $\lim_{x\to c} f(x)$ D.N.E. for all c.

Example 1.7. (詳見影片) Let $f : (0,\infty) \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} \dfrac{1}{p} & \text{if } x = \dfrac{q}{p}, \text{ where } p, q \in \mathbb{N} \text{ and } (p,q) = 1\,, \\ 0 & \text{if } x \text{ is irrational (無理數)}\,. \end{cases}$$

Then $\lim_{x\to c} f(x) = 0$ for all $c \in (0,\infty)$.

Definition 1.8

Let f be a function defined on an open interval containing c (except possibly at c), and L be a real number. The statement

$$\lim_{x\to c} f(x) = L\,, \quad \text{read "the limit of } f \text{ at } c \text{ is } L\text{"},$$

means that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\big|f(x) - L\big| < \varepsilon \quad \text{whenever} 0 < |x - c| < \delta\,.$$

Explanation: (詳見影片) 因為 $|f(x) - L| < \varepsilon$ 等價於 $f(x) \in (L-\varepsilon, L+\varepsilon)$,所以定義敘述中的 ε 可視為用來度量 $f(x)$ 向 L 這個數集中的程度。定義所述是指對於任意給定的集中程度 $\varepsilon > 0$,一定可以找到在 c 附近的一個範圍(以到 c 的距離小於 δ 來表示),滿足

此範圍中的點之函數值落入想要其落入的集中區域 $(L-\varepsilon, L+\varepsilon)$ 之內。此即「當除 c 之外的點到 c 的距離愈來愈近時，其函數值向 L 集中」的意思。

Example 1.9. In this example we show that $\lim\limits_{x\to 1}(x+1) = 2$ using Definition 1.8.

Let $\varepsilon > 0$ be given. Define $\delta = \varepsilon$. Then $\delta > 0$ and if $0 < |x-1| < \delta$, we have

$$|(x+1)-2| = |x-1| < \delta = \varepsilon\,.$$

One could also pick $\delta = \dfrac{\varepsilon}{2}$ so that if $0 < |x-1| < \delta$,

$$|(x+1)-2| = |x-1| < \delta = \frac{\varepsilon}{2} < \varepsilon\,.$$

Example 1.10. Show that $\lim\limits_{x\to 2} x^2 = 4$. If $\varepsilon = 1$, we can choose $\delta = \min\big\{\sqrt{5}-2, 2-\sqrt{3}\big\}$ so that $\delta > 0$ and if $0 < |x-2| < \delta$ we must have $|x^2-4| < 1$.

For general ε, we can choose $\delta = \min\big\{\sqrt{4+\varepsilon}-2, 2-\sqrt{4-\varepsilon}\big\}$ so that $\delta > 0$ and if $0 < |x-2| < \delta$ we must have $|x^2-4| < \varepsilon$.

Example 1.11 (Proof of Example 1.7)**.** Let $\varepsilon > 0$ be given. Then there exists a prime number p such that $\dfrac{1}{p} < \varepsilon$. Let $q_1, q_2, \cdots, q_n$ be rational numbers in $\big(\dfrac{c}{2}, \dfrac{3c}{2}\big)$ satisfying

$$q_j = \frac{s}{r}\,, (r,s) = 1, 1 \leqslant r \leqslant p\,,$$

and define $\delta = \dfrac{1}{2}\min\Big(\big\{|c-q_1|, |c-q_2|, \cdots, |c-q_n|\big\} - \{0\}\Big)$. Then $\delta > 0$. Suppose that x satisfies that $0 < |x-c| < \delta$.

1. If $x \in \mathbb{Q}^\complement$, then $f(x) = 0$ which shows that $|f(x)| < \varepsilon$.

2. If $x \in \mathbb{Q}$, then $x = \dfrac{s}{r}$ for some natural numbers r, s satisfying $(r,s) = 1$. By the choice of δ, we find that $r > p$; thus

$$\big|f(x)\big| = \frac{1}{r} < \frac{1}{p} < \varepsilon\,.$$

In either case, $\big|f(x)\big| < \varepsilon$; thus we establish that

$$\big|f(x) - 0\big| < \varepsilon \quad \text{whenever} \quad 0 < |x-c| < \delta\,.$$

Therefore, $\lim\limits_{x\to c} f(x) = 0$.

Proposition 1.12

Let f, g be functions defined on an open interval containing c (except possibly at c), and $f(x) = g(x)$ if $x \neq c$. If $\lim_{x\to c} g(x) = L$, then $\lim_{x\to c} f(x) = L$.

Proof. Let $\varepsilon > 0$ be given. Since $\lim_{x\to c} g(x) = L$, there exists $\delta > 0$ such that

$$|g(x) - L| < \varepsilon \quad \text{if} \quad 0 < |x - c| < \delta.$$

Since $f(x) = g(x)$ if $x \neq c$, we must have if $0 < |x - c| < \delta$,

$$|f(x) - L| = |g(x) - L| < \varepsilon\,. \qquad \square$$

Example 1.13. Let $f(x) = x + 1$ and $g(x) = \dfrac{x^2 - 1}{x - 1}$. Since $f(x) = g(x)$ if $x \neq 1$, the proposition above implies that

$$\lim_{x\to 1} g(x) = \lim_{x\to 1} f(x) = 2\,.$$

1.2 Properties of Limits

Theorem 1.14

Let b, c be real numbers, f, g be functions defined on an open interval containing c (except possibly at c) with $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} g(x) = K$.

1. $\lim_{x\to c} b = b$, $\lim_{x\to c} x = c$, $\lim_{x\to c} |x| = |c|$;
2. $\lim_{x\to c} \big[f(x) \pm g(x)\big] = L + K$;（和或差的極限等於極限的和或差）
3. $\lim_{x\to c} \big[f(x)g(x)\big] = LK$;（乘積的極限等於極限的乘積）
4. $\lim_{x\to c} \dfrac{f(x)}{g(x)} = \dfrac{L}{K}$ if $K \neq 0$.（若分母極限不為零，則商的極限等於極限的商）

Proof. 1. Let $\varepsilon > 0$ be given.

(a) Define $\delta = 1$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we have $|b - b| = 0 < \varepsilon\,.$

(b) Define $\delta = \varepsilon$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we have $|x - c| < \delta = \varepsilon\,.$

(c) Define $\delta = \varepsilon$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, by the triangle inequality we have

$$\big||x| - |c|\big| \leqslant |x - c| < \delta = \varepsilon\,.$$

2. Let $\varepsilon > 0$ be given. Since $\lim\limits_{x\to c} f(x) = L$ and $\lim\limits_{x\to c} g(x) = K$, there exist $\delta_1, \delta_2 > 0$ such that

$$\big|f(x) - L\big| < \frac{\varepsilon}{2} \quad \text{whenever} \quad 0 < |x - c| < \delta_1$$

and

$$\big|g(x) - K\big| < \frac{\varepsilon}{2} \quad \text{whenever} \quad 0 < |x - c| < \delta_2\,.$$

Define $\delta = \min\{\delta_1, \delta_2\}$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we have

$$\big|f(x) + g(x) - (L + K)\big| \leqslant \big|f(x) - L\big| + \big|g(x) - K\big| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

3. Let $\varepsilon > 0$ be given. Since $\lim\limits_{x\to c} f(x) = L$, there exist δ_1, $\delta_2 > 0$ such that

$$\big|f(x) - L\big| < 1 \quad \text{whenever} \quad 0 < |x - c| < \delta_1$$

and

$$\big|f(x) - L\big| < \frac{\varepsilon}{2(|K| + 1)} \quad \text{whenever} \quad 0 < |x - c| < \delta_2\,.$$

Moreover, since $\lim\limits_{x\to c} g(x) = K$, there exists $\delta_3 > 0$ such that

$$\big|g(x) - K\big| < \frac{\varepsilon}{2(|L| + 1)} \quad \text{whenever} \quad 0 < |x - c| < \delta_3\,.$$

Define $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we have

$$\begin{aligned}
\big|f(x)g(x) - LK\big| &= \big|f(x)g(x) - f(x)K + f(x)K - LK\big| \\
&\leqslant \big|f(x)\big|\big|g(x) - K\big| + |K|\big|f(x) - L\big| \\
&< (|L| + 1)\frac{\varepsilon}{2(|L| + 1)} + |K|\frac{\varepsilon}{2(|K| + 1)} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.
\end{aligned}$$

4. W.L.O.G. (Without loss of generality), we can assume that $K > 0$ for otherwise we have $\lim\limits_{x\to c}(-g)(x) = -K > 0$ and

$$\lim_{x\to c}\Big(\frac{f}{g}\Big)(x) = \lim_{x\to c}\Big(\frac{-f}{-g}\Big)(x) = \frac{\lim\limits_{x\to c}(-f)(x)}{-K} = \frac{-L}{-K} = \frac{L}{K}\,.$$

Let $\varepsilon > 0$ be given. Since $\lim_{x\to c} g(x) = K$, there exist $\delta_1, \delta_2 > 0$ such that

$$|g(x) - K| < \frac{K}{2} \quad \text{whenever} \quad 0 < |x - c| < \delta_1$$

and

$$|g(x) - K| < \frac{K^2\varepsilon}{4(|L| + 1)} \quad \text{whenever} \quad 0 < |x - c| < \delta_2\,.$$

Moreover, since $\lim_{x\to c} f(x) = L$, there exists $\delta_3 > 0$ such that

$$|f(x) - L| < \frac{K\varepsilon}{4} \quad \text{whenever} \quad 0 < |x - c| < \delta_3\,.$$

Define $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we have

$$\begin{aligned}
\Big|\frac{f(x)}{g(x)} - \frac{L}{K}\Big| &= \frac{|Kf(x) - Lg(x)|}{K|g(x)|} \leqslant \frac{1}{|g(x)|}\frac{|Kf(x) - KL| + |KL - Lg(x)|}{K} \\
&\leqslant \frac{2}{K}\Big(|f(x) - L| + \frac{|L|}{K}|g(x) - K|\Big) \\
&< \frac{2}{K}\Big(\frac{K\varepsilon}{4} + \frac{|L|}{K}\frac{K^2\varepsilon}{4(|L| + 1)}\Big) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,,
\end{aligned}$$

where we have used $\dfrac{2}{K} \leqslant \dfrac{1}{|g(x)|}$ if $0 < |x - c| < \delta$ to conclude the inequality. Therefore, we conclude that $\lim_{x\to c} \dfrac{f(x)}{g(x)} = \dfrac{L}{K}$ if $K > 0$. □

Example 1.15. Find $\lim_{x\to 3} x^2$. By 1 of Theorem 1.14 $\lim_{x\to 3} x = 3$; thus 3 of Theorem 1.14 implies that

$$\lim_{x\to 3} x^2 = \Big(\lim_{x\to 3} x\Big)\Big(\lim_{x\to 3} x\Big) = 9\,.$$

The above equality further shows that

$$\lim_{x\to 3} x^3 = \Big(\lim_{x\to 3} x^2\Big)\Big(\lim_{x\to 3} x\Big) = 27\,.$$

In particular, if n is a positive integer, then (by induction) $\lim_{x\to c} x^n = c^n$.

Corollary 1.16

Assume the assumptions in Theorem 1.14, and let n be a positive integer.

1. $\lim_{x \to c} \left[f(x)^n \right] = L^n$.
2. If p is a polynomial function, then $\lim_{x \to c} p(x) = p(c)$.
3. If r is a rational function given by $r(x) = \dfrac{p(x)}{q(x)}$ for some polynomials p and q, and $q(c) \neq 0$, then $\lim_{x \to c} r(x) = r(c)$.

An illustration of why 2 in Corollary 1.16 is correct: Suppose that $p(x) = 3x^2 + 5x - 10$. Then applying 1-3 in Theorem 1.14, we obtain that

$$\begin{aligned}\lim_{x \to c} p(x) &= \lim_{x \to c}(3x^2 + 5x) - \lim_{x \to c}(10) = \lim_{x \to c}(3x^2 + 5x) - 10 \\ &= \Big(\lim_{x \to c}(3)\Big)\Big(\lim_{x \to c} x^2\Big) + \Big(\lim_{x \to c}(5)\Big)\Big(\lim_{x \to c} x\Big) - 10 \\ &= 3c^2 + 5c - 10 = p(c)\,.\end{aligned}$$

Theorem 1.17

If $c > 0$ and n is a positive integer, then $\lim_{x \to c} x^{\frac{1}{n}} = c^{\frac{1}{n}}$.

Proof. Let $\varepsilon > 0$ be given. Define $\delta = \min\left\{\dfrac{c}{2}, \dfrac{nc^{\frac{n-1}{n}}\varepsilon}{2}\right\}$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we must have

$$x^{\frac{n-1}{n}} + x^{\frac{n-2}{n}}c^{\frac{1}{n}} + x^{\frac{n-3}{n}}c^{\frac{2}{n}} + \cdots + x^{\frac{1}{n}}c^{\frac{n-2}{n}} + c^{\frac{n-1}{n}} \geqslant \frac{n}{2}c^{\frac{n-1}{n}}\,.$$

Therefore, if $0 < |x - c| < \delta$,

$$\begin{aligned}\left|x^{\frac{1}{n}} - c^{\frac{1}{n}}\right| &= \left|\frac{x - c}{x^{\frac{n-1}{n}} + x^{\frac{n-2}{n}}c^{\frac{1}{n}} + x^{\frac{n-3}{n}}c^{\frac{2}{n}} + \cdots + x^{\frac{1}{n}}c^{\frac{n-2}{n}} + c^{\frac{n-1}{n}}}\right| \\ &\leqslant \frac{2}{n}c^{-\frac{n-1}{n}}|x - c| < \frac{2}{n}c^{-\frac{n-1}{n}}\delta \leqslant \frac{2}{n}c^{-\frac{n-1}{n}}\frac{nc^{\frac{n-1}{n}}\varepsilon}{2} = \varepsilon\end{aligned}$$

which implies that $\lim_{x \to c} x^{\frac{1}{n}} = c^{\frac{1}{n}}$. $\square$

Theorem 1.18

If f and g are functions (defined on open intervals) such that $\lim_{x\to c} g(x) = K$, $\lim_{x\to K} f(x) = L$ and $L = f(K)$, then

$$\lim_{x\to c}(f\circ g)(x) = L\,.$$

Proof. Let $\varepsilon > 0$ be given. Since $\lim_{x\to L} f(x) = L$, there exists $\delta_1 > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - K| < \delta_1\,.$$

Since $L = f(K)$, the statement above implies that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad |x - K| < \delta_1\,.$$

Fix such δ_1. Since $\lim_{x\to c} g(x) = K$, there exists $\delta > 0$ such that

$$|g(x) - K| < \delta_1 \quad \text{whenever} \quad 0 < |x - c| < \delta\,.$$

Therefore, if $0 < |x - c| < \delta$, $|(f\circ g)(x) - L| = |f(g(x)) - L| < \varepsilon$ which concludes the theorem. $\square$

Example 1.19. Find $\lim_{x\to 0}\dfrac{\sqrt{x+1}-1}{x}$.

Let $f(x) = \dfrac{\sqrt{x+1}-1}{x}$. If $x \neq 0$,

$$f(x) = \frac{(\sqrt{x+1}-1)(\sqrt{x+1}+1)}{x(\sqrt{x+1}+1)} = \frac{1}{\sqrt{x+1}+1} \equiv g(x)\,.$$

To see the limit of g, note that

$$\lim_{x\to 0}\sqrt{x+1} = 1 \qquad \text{(by Theorem 1.18);}$$

thus by Theorem 1.14 $\lim_{x\to 0} g(x) = \dfrac{1}{2}$.

Remark 1.20. In Theorem 1.18, the condition $L = f(K)$ is important, even though intuitively if $g(x) \to K$ as $x \to c$ and $f(x) \to L$ as $x \to K$ then $(f\circ g)(x)$ should approach L as x approaches c. A counter-example is given by the following two functions: f is the function given in Example 1.3 and g is a constant function with value 2. This example/theorem demonstrates an important fact: intuition could be wrong! That is the reason why mathematicians develop the ε-δ language in order to explain ideas of limits rigorously.

Theorem 1.21: Squeeze Theorem（夾擠定理）

Let f, g, h be functions defined on an open interval containing c (except possibly at c), and $h(x) \leqslant f(x) \leqslant g(x)$ if $x \neq c$. If $\lim\limits_{x\to c} h(x) = \lim\limits_{x\to c} g(x) = L$, then $\lim\limits_{x\to c} f(x)$ exists and is equal to L.

Proof. Let $\varepsilon > 0$. Since $\lim\limits_{x\to c} h(x) = \lim\limits_{x\to c} g(x) = L$, there exist $\delta_1, \delta_2 > 0$ such that

$$|h(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta_1$$

and

$$|g(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta_2\,.$$

Define $\delta = \min\{\delta_1, \delta_2\}$. Then $\delta > 0$ and if $0 < |x - c| < \delta$,

$$L - \varepsilon < h(x) \leqslant f(x) \leqslant g(x) < L + \varepsilon$$

which implies that $|f(x) - L| < \varepsilon$ whenever $0 < |x - c| < \delta$. $\square$

Example 1.22. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x & \text{if } x \text{ is rational}\,, \\ -x & \text{if } x \text{ is irrational}\,. \end{cases}$$

Then $\lim\limits_{x\to c} f(x)$ D.N.E. if $c \neq 0$ and $\lim\limits_{x\to 0} f(x) = 0$.

1. If $c \neq 0$, then as $x \neq c$ approaches c and $x \in \mathbb{Q}$, $f(x)$ approaches c, while as $x \neq c$ approaches c and $x \notin \mathbb{Q}$, $f(x)$ approaches $-c$. This implies that as x approaches c, $f(x)$ does not approaches a fixed number; thus $\lim\limits_{x\to c} f(x)$ D.N.E.

2. Note that $\big|f(x)\big| = |x|$; thus $-|x| \leqslant f(x) \leqslant |x|$ for all $x \in \mathbb{R}$. Since $\lim\limits_{x\to 0} |x| = 0$, the Squeeze theorem implies that $\lim\limits_{x\to 0} f(x) = 0$.

Example 1.23. In this example we consider the limit of the sine function at a real number c. Before proceeding, let us first establish a fundamental inequality

$$|\sin x| \leqslant |x| \qquad \text{for all real numbers } x \text{ (in radian unit).} \tag{1.2.1}$$

Recall (0.2.2) that

$$\sin x \leqslant x \leqslant \tan x \qquad \forall\, 0 \leqslant x \leqslant \frac{\pi}{2}\,. \tag{0.2.2}$$

To see (1.2.1), it suffices to consider the case when $x \notin \big[0, \frac{\pi}{2}\big]$. Nevertheless,

1. it trivially holds that $|\sin x| \leqslant x$ if $x \geqslant \frac{\pi}{2}$;

2. if $x < 0$, then $|\sin x| = |\sin(-x)| \leqslant |-x| = |x|$.

Having establish (1.2.1), now note the sum-to-product formula implies that

$$|\sin x - \sin c| = 2\Big|\sin\frac{x-c}{2}\cos\frac{x+c}{2}\Big| \leqslant 2\Big|\sin\frac{x-c}{2}\Big| \leqslant |x-c| \qquad \text{for all real number } x.$$

Therefore, $\sin c - |x-c| \leqslant \sin x \leqslant \sin c + |x-c|$ for all real number x, and the Squeeze Theorem then implies that $\lim\limits_{x\to c}\sin x = \sin c$ since $\lim\limits_{x\to c}|x-c| = 0$.

Similarly, using the sum-to-product formula

$$\cos x - \cos c = -2\sin\frac{x+c}{2}\sin\frac{x-c}{2},$$

we can also conclude that $\lim\limits_{x\to c}\cos x = \cos c$. The detail is left as an exercise.

By Theorem 1.14, Example 1.23 shows the following

Theorem 1.24

Let c be a real number in the domain of the given trigonometric functions.

1. $\lim\limits_{x\to c}\sin x = \sin c$; 2. $\lim\limits_{x\to c}\cos x = \sin c$; 3. $\lim\limits_{x\to c}\tan x = \tan c$;

4. $\lim\limits_{x\to c}\cot x = \cot c$; 5. $\lim\limits_{x\to c}\sec x = \sec c$; 6. $\lim\limits_{x\to c}\csc x = \csc c$.

Example 1.25. In this example we compute $\lim\limits_{x\to 0} x\sin\frac{1}{x}$ if it exists. Note that if the limit exists, we cannot apply 3 of Theorem 1.14 to find the limit since $\lim\limits_{x\to 0}\sin\frac{1}{x}$ does not exist. On the other hand, since $\big|x\sin\frac{1}{x}\big| \leqslant |x|$ if $x \neq 0$, $-|x| \leqslant x\sin\frac{1}{x} \leqslant |x|$ if $x \neq 0$. By the fact that $\lim\limits_{x\to 0}|x| = \lim\limits_{x\to 0}(-|x|) = 0$, the Squeeze Theorem implies that $\lim\limits_{x\to 0} x\sin\frac{1}{x} = 0$.

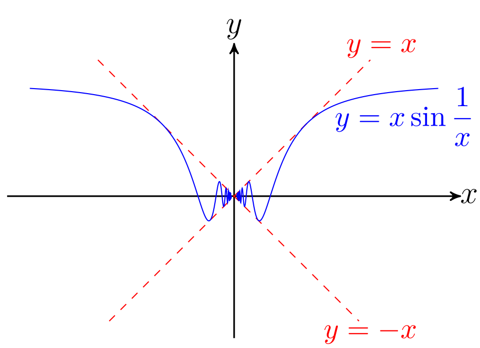


Figure 1.5: The graph of function $y = x\sin\frac{1}{x}$

1.2.1 One-sided limits and limits as $x \to \pm\infty$

Suppose that f is a function defined (only) on one side of a point c, it is also possible to consider the one-sided limit $\lim\limits_{x\to c^+} f(x)$ or $\lim\limits_{x\to c^-} f(x)$, where the notation $x \to c^+$ and $x \to c^-$ means that x is taken from the right-hand side and left-hand side of c, respectively, and becomes arbitrarily close to c. In other words, $\lim\limits_{x\to c^+} f(x)$ means the value to which $f(x)$ approaches as x approaches to c from the right, while $\lim\limits_{x\to c^-} f(x)$ means the value to which $f(x)$ approaches as x approaches to c from the left.

Definition 1.26: One-sided limits

Let f be a function defined on an interval with c as the left/right end-point (except possibly at c), and L be a real number. The statement

$$\lim_{x\to c^+} f(x) = L \Big/ \lim_{x\to c^-} f(x) = L\,,$$

read "the right/left(-hand) limit of f at c is L" or "the limit of f at c from the right/left is L", means that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\big|f(x) - L\big| < \varepsilon \quad \text{whenever} \quad 0 < x - c < \delta \big/ -\delta < x - c < 0\,.$$

Example 1.27. In this example we show that $\lim\limits_{x\to 0^+} x^{\frac{1}{n}} = 0$. Let $\varepsilon > 0$ be given. Define $\delta = \varepsilon^n$. Then $\delta > 0$ and if $0 < x < \delta$, we have

$$|x^{\frac{1}{n}} - 0| = x^{\frac{1}{n}} < \delta^{\frac{1}{n}} = \epsilon\,.$$

We note that Theorem 1.14, Theorem 1.17 and 1.21 are also valid when the limits are replaced by one-sided limits, and the precise statements are provided below.

Theorem 1.28

Let b, c be real numbers, f, g be functions with $\lim\limits_{x\to c^+} f(x) = L$ and $\lim\limits_{x\to c^+} g(x) = K$.

1. $\lim\limits_{x\to c^+} b = b$, $\lim\limits_{x\to c^+} x = c$, $\lim\limits_{x\to c^+} |x| = |c|$;
2. $\lim\limits_{x\to c^+} \big[f(x) \pm g(x)\big] = L + K$;
3. $\lim\limits_{x\to c^+} \big[f(x)g(x)\big] = LK$;
4. $\lim\limits_{x\to c^+} \dfrac{f(x)}{g(x)} = \dfrac{L}{K}$ if $K \neq 0$.

The conclusions above also hold for the case of left limits (that is, with $x \to c^+$ replaced by $x \to c^-$).

Theorem 1.29

If $c > 0$ and n is a positive integer, then $\lim\limits_{x\to c^+} x^{\frac{1}{n}} = c^{\frac{1}{n}}$ and $\lim\limits_{x\to c^-} x^{\frac{1}{n}} = c^{\frac{1}{n}}$

Theorem 1.30

If f and g are functions such that $\lim\limits_{x\to c^+} g(x) = K$, $\lim\limits_{x\to K} f(x) = L$ and $L = f(K)$, then

$$\lim_{x\to c^+} (f\circ g)(x) = L\,.$$

The conclusions above also hold for the case of left limits (that is, with $x \to c^+$ replaced by $x \to c^-$).

Remark 1.31. Theorem 1.30 is not true if one only has the one-sided limit $\lim\limits_{x\to K^+} f(x) = L$ instead of the full limit $\lim\limits_{x\to K} f(x) = L$. For example, consider $g(x) = -x$ and $f(x)$ be the function

$$f(x) = \begin{cases} 1 & \text{if } x \geqslant 0\,, \\ 0 & \text{if } x < 0\,. \end{cases}$$

Then $\lim\limits_{x\to 0^+} g(x) = 0$ and $\lim\limits_{x\to 0^+} f(x) = f(0)$; however,

$$(f\circ g)(x) = \begin{cases} 0 & \text{if } x > 0\,, \\ 1 & \text{if } x \leqslant 0\,, \end{cases}$$

which implies that $\lim\limits_{x\to 0^+} (f\circ g)(x) = 0 \neq f(0)$.

Theorem 1.32: Squeeze Theorem（夾擠定理）

1. Let f, g, h be functions defined on an interval with c as the left end-point (except possible at c), and $h(x) \leqslant f(x) \leqslant g(x)$ if $x > c$. If $\lim\limits_{x\to c^+} h(x) = \lim\limits_{x\to c^+} g(x) = L$, then $\lim\limits_{x\to c^+} f(x)$ exists and is equal to L.
2. Let f, g, h be functions defined on an interval with c as the right end-point (except possible at c), and $h(x) \leqslant f(x) \leqslant g(x)$ if $x < c$. If $\lim\limits_{x\to c^-} h(x) = \lim\limits_{x\to c^-} g(x) = L$, then $\lim\limits_{x\to c^-} f(x)$ exists and is equal to L.

The following theorem shows the relation between the limit and one-sided limits of functions.

Theorem 1.33

Let f be a function defined on an open interval containing c (except possibly at c). The limit $\lim\limits_{x \to c} f(x)$ exists if and only if $\lim\limits_{x \to c^+} f(x)$ and $\lim\limits_{x \to c^-} f(x)$ both exist and are identical. In either case,

$$\lim_{x \to c} f(x) = \lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x)\,.$$

Explanation on "A if and only if B" in Theorem 1.33: It should be clear that "A if B" means "A happens when B happens" (which is the same as "B implies A"). The statement "A only if B" means that "A happens only when B happens"; thus "A only if B" means that "A implies B".

Proof of Theorem 1.33. ($\Rightarrow$) - the "only if" part: Suppose that $\lim\limits_{x \to c} f(x) = L$, and let $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta\,.$$

Therefore, there exists $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < x - c < \delta\,;$$

thus $\lim\limits_{x \to c^+} f(x) = L$. Similarly, $\lim\limits_{x \to c^-} f(x) = L$.

($\Leftarrow$) - the "if" part: Suppose that $\lim\limits_{x \to c^+} f(x) = \lim\limits_{x \to c^-} f(x) = L$. Let $\varepsilon > 0$. Then there exist $\delta_1, \delta_2 > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < x - c < \delta_1$$

and

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad -\delta_2 < x - c < 0\,.$$

Define $\delta = \min\{\delta_1, \delta_2\}$. Then $\delta > 0$ and if $0 < |x - c| < \delta$, we must have $0 < x - c < \delta_1$ and $-\delta_2 < x - c < 0$; thus if $0 < |x - c| < \delta$, we must have $|f(x) - L| < \varepsilon$. $\square$

Example 1.34. In this example we compute a very important limit

$$\lim_{x \to 0} \frac{\sin x}{x} = 1\,. \tag{1.2.2}$$

To see this, we recall (0.2.2) that

$$\sin x \leqslant x \leqslant \tan x \quad \text{for all } 0 \leqslant x \leqslant \frac{\pi}{2}\,. \tag{0.2.2}$$

Now using (0.2.2), we find that

$$\cos x \leqslant \frac{\sin x}{x} \leqslant 1 \quad \text{for all } 0 < x < \frac{\pi}{2}\,.$$

The Squeeze Theorem (Theorem 1.32) then implies that $\lim\limits_{x\to 0^+} \dfrac{\sin x}{x} = 1$. On the other hand,

$$\lim_{x\to 0^-} \frac{\sin x}{x} = \lim_{x\to 0^-} \frac{\sin(-x)}{-x} = \lim_{x\to 0^+} \frac{\sin x}{x} = 1\,;$$

thus Theorem 1.33 implies that $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$.

Remark 1.35. The function $\dfrac{\sin x}{x}$ is the famous (unnormalized) sinc function; that is, $\operatorname{sinc}(x) = \dfrac{\sin x}{x}$ and $\operatorname{sinc}(0) = 1$. The example above shows that $\lim\limits_{x\to 0} \operatorname{sinc}(x) = \operatorname{sinc}(0)$.

Example 1.36. In this example we compute the limit $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2}$. By the half-angle formula, $1-\cos x = 2\sin^2 \dfrac{x}{2}$; thus

$$\frac{1-\cos x}{x^2} = \frac{2\sin^2\frac{x}{2}}{x^2} = \frac{1}{2}\frac{\sin^2\frac{x}{2}}{\left(\frac{x}{2}\right)^2} = \frac{1}{2}\operatorname{sinc}^2\left(\frac{x}{2}\right).$$

Therefore, Theorem 1.18 implies that $\lim\limits_{x\to 0} \dfrac{1-\cos x}{x^2} = \dfrac{1}{2}$.

An open interval in the real number system can be unbounded. When the open interval on which f is defined is not bounded from above (which means there is no real number which is larger than all the numbers in this interval), we can also consider the behavior of $f(x)$ as x becomes increasingly large and eventually outgrow all finite bounds.

Definition 1.37: Limits as $x \to \pm\infty$

Let f be a function defined on an infinite interval bounded from below/above, and L be a real number. The statement

$$\lim_{x\to\infty} f(x) = L \Big/ \lim_{x\to-\infty} f(x) = L\,,$$

read "the right/left(-hand) limit of f at c is L" or "the limit of f at c from the right/ left is L", means that for each $\varepsilon > 0$ there exists a real number $M > 0$ such that

$$\big|f(x) - L\big| < \varepsilon \quad \text{whenever} \quad x > M \big/ x < -M\,.$$

Similar to the case of one-sided limit, Theorem 1.28, Theorem 1.30 and 1.32 are also valid when the notation $x \to c^\pm$ are replaced by $x \to \pm\infty$.

Example 1.38. In this example we show that $\lim\limits_{x\to\infty}\dfrac{1}{|x|}=0$ and $\lim\limits_{x\to-\infty}\dfrac{1}{|x|}=0$.

Let $\varepsilon>0$ be given. Define $M=\dfrac{1}{\varepsilon}$. Then if $x>M$ or $x<-M$, we must have $|x|>M$; thus if $x>M$ or $x<-M$,

$$\Big|\frac{1}{x}-0\Big|=\frac{1}{|x|}<\frac{1}{M}<\varepsilon\,.$$

Example 1.39. Recall that the sinc function is defined by

$$\operatorname{sinc}(x)=\begin{cases}\dfrac{\sin x}{x} & \text{if } x\neq 0\,,\\ 1 & \text{if } x=0\,.\end{cases}$$

Then $\Big|\dfrac{\sin x}{x}\Big|\leqslant\dfrac{1}{|x|}$ for all $x\neq 0$ and this provides the inequality $-\dfrac{1}{|x|}\leqslant\dfrac{\sin x}{x}\leqslant\dfrac{1}{|x|}$ for all $x\neq 0$. By the Squeeze Theorem and the previous example, we find that

$$\lim_{x\to\infty}\operatorname{sinc}(x)=\lim_{x\to-\infty}\operatorname{sinc}(x)=0\,.$$

Theorem 1.40

Let f be a function defined on an open interval, and $g(x)=f\big(\frac{1}{x}\big)$ if $x\neq 0$.

1. Suppose that the open interval is not bounded from above. Then $\lim\limits_{x\to\infty}f(x)$ exists if and only if $\lim\limits_{x\to 0^+}g(x)$ exists. In either case,
$$\lim_{x\to\infty}f(x)=\lim_{x\to 0^+}g(x)\,.$$
2. Suppose that the open interval is not bounded from below. Then $\lim\limits_{x\to-\infty}f(x)$ exists if and only if $\lim\limits_{y\to 0^-}g(x)$ exists. In either case,
$$\lim_{x\to-\infty}f(x)=\lim_{x\to 0^-}g(x)\,.$$

The theorem above should be very intuitive, and the proof is left as an exercise.

Example 1.41. Find the limit $\lim\limits_{x\to\infty}\dfrac{x+\sin x}{x+1}$.

By Theorem 1.40, we have

$$\begin{aligned}\lim_{x\to\infty}\frac{x+\sin x}{x+1}&=\lim_{x\to 0^+}\frac{\frac{1}{x}+\sin\frac{1}{x}}{\frac{1}{x}+1}=\lim_{x\to 0^+}\frac{1+x\sin\frac{1}{x}}{1+x}\\&=\lim_{x\to 0^+}\frac{1}{1+x}+\Big(\lim_{x\to 0^+}\frac{1}{x+1}\Big)\Big(\lim_{x\to 0^+}x\sin\frac{1}{x}\Big)=1+1\cdot 0=1\,.\end{aligned}$$

Here we note that in the process of computing the limit we have used results analogous to Theorem 1.28. We can also apply the Squeeze theorem to the inequality $\frac{x-1}{x+1} \leqslant \frac{x+\sin x}{x+1} \leqslant 1$ for all $x > 0$ and obtain the same limit.

Corollary 1.42

Let p and q be polynomial functions.

1. If the degree of p is smaller than the degree of q, then
$$\lim_{x\to\infty} \frac{p(x)}{q(x)} = \lim_{x\to-\infty} \frac{p(x)}{q(x)} = 0\,.$$

2. If the degree of p is the same as the degree of q, then
$$\lim_{x\to\infty} \frac{p(x)}{q(x)} = \lim_{x\to-\infty} \frac{p(x)}{q(x)} = \frac{\text{the leading coefficient of } p}{\text{the leading coefficient of } q}\,.$$

1.3 Continuity of Functions

Definition 1.43

Let f be a function defined on an interval I, and $c \in I$.

1. f is said to be right-continuous at c (or continuous from the right at c) if
$$\lim_{x\to c^+} f(x) = f(c)\,.$$

2. f is said to be left-continuous at c (or continuous from the left at c) if
$$\lim_{x\to c^-} f(x) = f(c)\,.$$

3. If c is the left end-point of I, f is said to be continuous at c if f is right-continuous at c.

4. If c is the right end-point of I, f is said to be continuous at c if f is left-continuous at c.

5. If c is an interior point of I; that is, c is neither the left end-point nor the right end-point of I, then f is said to be continuous at c if $\lim\limits_{x\to c} f(x) = f(c)$.

f is said to be discontinuous at c if f is not continuous at c, and in this case c is called a point of discontinuity (or simply a discontinuity) of f. f is said to be continuous (or a continuous function) on I if f is continuous at each point of I.

Example 1.44. Consider the the greatest integer function (also known as the Gauss function or the floor function) $[\![\cdot]\!] : \mathbb{R} \to \mathbb{R}$ defined by

$$[\![x]\!] = \text{the greatest integer which is not greater than } x.$$

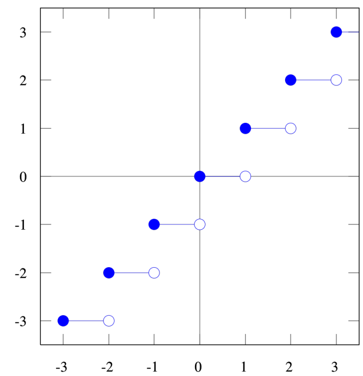


Figure 1.6: The greatest integer function $y = [\![x]\!]$

For example, $[\![2.5]\!] = 2$ and $[\![-2.5]\!] = -3$. If c is not an integer, $\lim\limits_{x\to c} [\![x]\!] = c$, while if c is an integer, we have

$$\lim_{x\to c^+} [\![x]\!] = c \qquad \text{and} \qquad \lim_{x\to c^-} [\![x]\!] = c - 1\,.$$

Let $f : [0, 2] \to \mathbb{R}$ be given by $f(x) = [\![x]\!]$. Then the conclusion above shows that f is continuous at every non-integer number, while f is not continuous at 1 (since $\lim\limits_{x\to 1} f(x)$ does not exist) and 2 (since $\lim\limits_{x\to 2^-} f(x) \neq f(2)$). On the other hand, $\lim\limits_{x\to 0^+} f(x) = f(0)$, so f is continuous at 0.

Therefore, f is continuous at c if c is not an integer, and f is right-continuous at c if c is an integer.

Example 1.45. Let $f(x) = x^n$, where n is a positive integer. We have shown that

$$\lim_{x\to c} x^n = c^n$$

for all real numbers c; thus f is continuous on $\mathbb{R}$. In general, polynomial functions are continuous on $\mathbb{R}$ (because of Corollary 1.16).

Example 1.46. Let n be a positive integer, and $f : [0, \infty) \to \mathbb{R}$ be defined by $f(x) = x^{\frac{1}{n}}$. By Theorem 1.17 and Example 1.27,

$$\lim_{x\to c} x^{\frac{1}{n}} = c^{\frac{1}{n}} \quad \text{if } c > 0 \qquad \text{and} \qquad \lim_{x\to 0^+} x^{\frac{1}{n}} = 0\,;$$

thus f is continuous on $[0, \infty)$.

Example 1.47. Recall the Dirichlet function $f : \mathbb{R} \to \mathbb{R}$ in Example 1.6 given by

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q}\,, \\ 1 & \text{if } x \notin \mathbb{Q}\,, \end{cases}$$

We have explained (but not proven) that the limit $\lim\limits_{x \to c} f(x)$ does not exist for all $c \in (0, \infty)$; thus f is discontinuous at all real numbers.

Example 1.48. Recall the function $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q}\,, \\ -x & \text{if } x \notin \mathbb{Q}\,. \end{cases}$$

in Example 1.22. We have shown that $\lim\limits_{x \to 0} f(x) = 0$; thus f is continuous at 0.

Example 1.49. Recall the function $f : (0, \infty) \to \mathbb{R}$ in Example 1.7 given by

$$f(x) = \begin{cases} \dfrac{1}{p} & \text{if } x = \dfrac{q}{p}, \text{ where } p, q \in \mathbb{N} \text{ and } (p, q) = 1\,, \\ 0 & \text{if } x \text{ is irrational}\,. \end{cases}$$

We have shown that $\lim\limits_{x \to c} f(x) = 0$ for all $c \in (0, \infty)$. Therefore, f is continuous at all irrational numbers but is discontinuous at all rational numbers.

Example 1.50. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous, and $f(x) = 2$ if $x \in \mathbb{Q}$. Then intuitively $f(x) = 2$ for all $x \in \mathbb{R}$. We now prove this using the definition of continuity.

Suppose the contrary that there exists $c \in \mathbb{R}$ such that $f(c) \neq 2$. Define $\varepsilon = \big|f(c) - 2\big|$. Then $\varepsilon > 0$. Since f is continuous at c, there exists $\delta > 0$ such that

$$\big|f(x) - f(c)\big| < \varepsilon \quad \text{whenever} \quad |x - c| < \delta\,.$$

Choose $x \in \mathbb{Q}$ such that $|x - c| < \delta$. Then the triangle inequality implies that

$$\varepsilon = \big|f(c) - 2\big| \leqslant \big|f(c) - f(x)\big| + \big|f(x) - 2\big| < \varepsilon$$

which is a contradiction.

Remark 1.51. Let I be an interval, $c \in I$, and $f : I \to \mathbb{R}$ be a function. The continuity of f at c is equivalent to that for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon \quad \text{whenever} \quad |x - c| < \delta \;\text{ and }\; x \in I\,.$$

To see this, we first consider the case that c is an interior point of I. Then by the definition, f is continuous at c if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon \quad \text{whenever} \quad 0 < |x - c| < \delta\,.$$

Since $|f(x) - f(c)| < \varepsilon$ automatically holds if $|x - c| = 0$, the statement above is equivalent to that

$$|f(x) - f(c)| < \varepsilon \quad \text{whenever} \quad |x - c| < \delta\,.$$

Now let us look at the case when c is the left end-point of I (so in this case $c \in I$). Then by definition, f is continuous at c if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon \quad \text{whenever} \quad 0 < x - c < \delta\,.$$

Again $|f(x) - f(c)| < \varepsilon$ automatically holds if $x - c = 0$, the statement above is equivalent to that

$$|f(x) - f(c)| < \varepsilon \quad \text{whenever} \quad c \leqslant x < c + \delta\,.$$

Note that since c is the left end-point, the set $\{x \mid c \leqslant x < c+\delta\}$ is the same as $\{x \mid |x-c| < \delta, x \in I\}$; thus the statement above is equivalent to that

$$|f(x) - f(c)| < \varepsilon \quad \text{whenever} \quad |x - c| < \delta \;\text{ and }\; x \in I\,.$$

Similar argument can be applied to the case when c is the right end-point of I.

Remark 1.52. Discontinuities of functions can be classified into different categories: removable discontinuities and non-removable discontinuities. Let c be a discontinuity of a function f. Then either (1) $\lim\limits_{x\to c} f(x)$ exists but $\lim\limits_{x\to c} f(x) \neq f(c)$ or (2) $\lim\limits_{x\to c} f(x)$ does not exist. If it is the first case, then c is called a ***removable discontinuity*** and that means we can adjust/re-define the value of f at c to make it continuous at c. For the second case, no matter what $f(c)$ is, f cannot be continuous at c.

If $\lim\limits_{x\to c^+} f(x)$ and $\lim\limits_{x\to c^-} f(x)$ both exist but are not identical, c is also called a ***jump discontinuity***.

Proposition 1.53

Let f, g be defined on an interval I, $c \in I$, and f, g be continuous at c. Then

1. $f \pm g$ is continuous at c.
2. fg is continuous at c.
3. $\dfrac{f}{g}$ is continuous at c if $g(c) \neq 0$.

Corollary 1.54

Let f, g be continuous functions on an interval I. Then

1. $f \pm g$ is continuous on I.
2. fg is continuous on I
3. $\dfrac{f}{g}$ is continuous (on its domain).

Theorem 1.55

Let I, J be ~~open~~ intervals,, $g : I \to \mathbb{R}$, $f : J \to \mathbb{R}$ be functions, and J contains the range of g. If g is continuous at c, then $f \circ g$ is continuous at c.

Proof. Let $\varepsilon > 0$ be given. Since f is continuous at $g(c)$, there exists $\delta_1 > 0$ such that

$$\big|f(y) - f\big(g(c)\big)\big| < \varepsilon \qquad \text{whenever} \quad \big|y - g(c)\big| < \delta_1 \;\text{ and }\; y \in J\,.$$

For such a δ_1, by the continuity of g at c there exists $\delta > 0$ such that

$$\big|g(x) - g(c)\big| < \delta_1 \qquad \text{whenever} \quad |x - c| < \delta \;\text{ and }\; x \in I\,.$$

Suppose that $|x - c| < \delta$ and $x \in I$. Let $y = g(x)$. By the condition that J contains the range of g,

$$\big|y - g(c)\big| < \delta_1 \;\text{ and }\; y \in J\,.$$

Therefore, if $|x - c| < \delta$ and $x \in I$,

$$\big|f(g(x)) - f(g(c))\big| < \varepsilon$$

which shows the continuity of $f \circ g$ at c. $\square$

Corollary 1.56

Let I, J be ~~open~~ intervals, and $g : I \to \mathbb{R}$, $f : J \to \mathbb{R}$ be continuous functions. If J contains the range of g, then $f \circ g$ is continuous on I.

Example 1.57. Let g be continuous on an interval I, and n be a positive integer. We show that g^n and $|g|^{\frac{1}{n}}$ are also continuous on I. Note that g^n is the function given by $g^n(x) = g(x)^n$ and $|g|^{\frac{1}{n}}$ is the function given by $|g|^{\frac{1}{n}} = |g(x)|^{\frac{1}{n}}$.

1. Let $f(x) = x^n$. Then Theorem 1.14 (or Corollary 1.16) implies that f is continuous on $\mathbb{R}$. Since $\mathbb{R}$ contains the range of g, by the corollary above we find tat $f \circ g \,(\equiv g^n)$ is continuous on I.

2. Let $h(x) = |x|$. Then Theorem 1.14 implies that h is continuous on $\mathbb{R}$. Since $\mathbb{R}$ contains the range of g, by the corollary above we find that $h \circ g \,(\equiv |g|)$ is continuous on I.

   Let $f(x) = x^{\frac{1}{n}}$. Then Theorem 1.17 and Example 1.27 imply that f is continuous on the non-negative real axis $[0, \infty)$. Since $[0, \infty)$ contains the range of $|g|$, the corollary above shows that $f \circ |g| (\equiv |g|^{\frac{1}{n}})$ is continuous on I.

Theorem 1.58: Intermediate Value Theorem - 中間值定理

If f is continuous on the closed interval $[a, b]$, $f(a) \neq f(b)$, and k is any number between $f(a)$ and $f(b)$, then there is at least one number c in $[a, b]$ such that $f(c) = k$.

Example 1.59 (Bisection method of finding zeros of continuous functions)**.** Let f be a function and $f(a)f(b) < 0$. Then the intermediate value theorem implies that there exists a zero c of f between a and b. How do we "find" (one of) this c? Consider the middle point $\dfrac{a+b}{2}$ of a and b. If $f\big(\dfrac{a+b}{2}\big) = 0$, then we find this zero, or otherwise we either have

$$f(a)f\big(\frac{a+b}{2}\big) < 0 \qquad \text{or} \qquad f(b)f\big(\frac{a+b}{2}\big) < 0$$

and only one of them can happen. In either case we can consider the middle point of the two points at which the value of f have different sign. Continuing this process, we can locate one zero as accurate as possible.

Example 1.60. Let $f : [0, 1] \to [0, 1]$ be a continuous function. In the following we prove that there exists $c \in [0, 1]$ such that $f(c) = c$. To see this, W.L.O.G. we assume that $f(0) \neq 0$ and $f(1) \neq 1$ for otherwise we find c (which is 0 or 1) such that $f(c) = c$.

Define $g(x) = f(x) - x$. Then g is continuous (by Proposition 1.53). Since $f : [0, 1] \to [0, 1]$, $f(0) \neq 0$ and $f(1) \neq 1$, we must have $g(0) > 0$ and $g(1) < 0$. By the intermediate value theorem, there exists $c \in (0, 1)$ such that $g(c) = 0$, and this implies that there exists $c \in (0, 1)$ such that $f(c) = c$. So either (1) $f(0) = 0$, (2) $f(1) = 1$, or (3) there ia $c \in (0, 1)$ such that $f(c) = c$.

1.4 Infinite Limits and Asymptotes

Definition 1.61

Let f be defined on an open interval containing c (except possible at c). The statement

$$\lim_{x\to c} f(x) = \infty\,,$$

read "$f(x)$ approaches infinity as x approaches c", means that for every $N > 0$ there exists $\delta > 0$ such that

$$f(x) > N \quad \text{whenever} \quad 0 < |x - c| < \delta\,.$$

The statement

$$\lim_{x\to c} f(x) = \infty\,,$$

read "$f(x)$ approaches minus infinity as x approaches c", means that for every $N > 0$ there exists $\delta > 0$ such that

$$f(x) < -N \quad \text{whenever} \quad 0 < |x - c| < \delta\,.$$

To define the infinite limit from the left/right, replace $0 < |x - c| < \delta$ by $c < x < c + \delta / c - \delta < x < c$. To define the infinite limit as $x \to \infty / x \to -\infty$, replace $0 < |x - c| < \delta$ by $x > \delta / x < -\delta$.

Note that the statement $\lim\limits_{x\to c} f(x) = \infty$ does **not** mean that the limit exists. It is a simple notation for saying that the value of f becomes unbounded as x approaches c and the limit fail to exist.

Example 1.62. $\lim\limits_{x\to 1} \dfrac{1}{(x-1)^2} = \infty$, $\lim\limits_{x\to 1^+} \dfrac{1}{x-1} = \infty$, and $\lim\limits_{x\to 1^-} \dfrac{1}{x-1} = -\infty$.

Example 1.63. Later we will talk about the exponential function in detail. In the mean time, assume that you know the graph of $y = 2^x$. Then $\lim\limits_{x\to\infty} 2^x = \infty$ and $\lim\limits_{x\to-\infty} 2^x = 0$.

• **Asymptotes**（漸近線）: If the distance between the graph of a function and some fixed straight line approaches zero as a point on the graph moves increasingly far from the origin, we say that the graph approaches the line asymptotically and that the line is an asymptote of the graph.

Definition 1.64: Vertical Asymptotes - 垂直漸近線

If f approaches infinity (or minus infinity) as x approaches c from the left or from the right, then the line $x = c$ is called a vertical asymptote of the graph of f.

Definition 1.65: Horizontal and Slant (Oblique) Asymptotes - 水平與斜漸近線

The straight line $y = mx + k$ is an asymptote of the graph of the function $y = f(x)$ if

$$\lim_{x\to\infty} \big[f(x) - mx - k\big] = 0 \qquad \text{or} \qquad \lim_{x\to-\infty} \big[f(x) - mx - k\big] = 0\,.$$

The straight line $y = mx + k$ is called a horizontal asymptote of the graph of f if $m = 0$, and is called a slant (oblique) asymptote of the graph of f if $m \neq 0$.

By the definition of horizontal asymptotes, it is clear that if $\lim\limits_{x\to\infty} f(x) = k$ or $\lim\limits_{x\to-\infty} f(x) = k$, then $y = k$ is a horizontal asymptote of the graph of f.

Example 1.66. Let $f(x) = \dfrac{x^2+3}{3x^2-4x+5}$. Then $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to-\infty} f(x) = \dfrac{1}{3}$; thus $y = \dfrac{1}{3}$ is a horizontal asymptote of the graph of f.

Example 1.67. Let $f(x) = \dfrac{x^3+3}{3x^2-4x+5}$. Then $\lim\limits_{x\to\infty} f(x) = \infty$ and $\lim\limits_{x\to-\infty} f(x) = -\infty$; thus the graph of f has no horizontal asymptote. However,

$$\lim_{x\to\infty} \Big[f(x) - \frac{x}{3}\Big] = \lim_{x\to\infty} \Big[\frac{3x^3+9}{3(3x^2-4x+5)} - \frac{x(3x^2-4x+5)}{3(3x^2-4x+5)}\Big] = \lim_{x\to\infty} \frac{4x^2-5x+9}{3(3x^2-4x+5)} = \frac{4}{9}\,;$$

thus $\lim\limits_{x\to\infty} \Big[f(x) - \dfrac{x}{3} - \dfrac{4}{9}\Big] = 0$. Therefore, $y = \dfrac{x}{3} + \dfrac{4}{9}$ is a slant asymptote of the graph of f.

Theorem 1.68

Let f and g be continuous on an open interval containing c. If $f(c) \neq 0$, $g(c) = 0$, and there exists an open interval containing c such that $g(x) \neq 0$ for all $x \neq c$ in the interval, then the graph of the function $h(x) = \dfrac{f(x)}{g(x)}$ has a vertical asymptote at $x = c$.

Example 1.69. Let $f(x) = \tan x$. Note that $\tan x = \dfrac{\sin x}{\cos x}$. For $n \in \mathbb{Z}$, $\sin\big(n\pi + \dfrac{\pi}{2}\big) \neq 0$ and $\cos\big(n\pi + \dfrac{\pi}{2}\big) = 0$. Moreover, $\cos x \neq 0$ for every x in the open interval $\big(n\pi + \dfrac{\pi}{4}, n\pi + \dfrac{3\pi}{4}\big)$ except $n\pi + \dfrac{\pi}{2}$. Therefore, by the theorem above we find that $x = n\pi + \dfrac{\pi}{2}$ is a vertical asymptote of the graph of the tangent function for all $n \in \mathbb{Z}$.

Theorem 1.70

If $y = mx + k$ is a slant asymptote of the graph of the function $y = f(x)$, then

$$m = \lim_{x\to\infty} \frac{f(x)}{x} \qquad \text{or} \qquad m = \lim_{x\to-\infty} \frac{f(x)}{x}$$

and

$$k = \lim_{x\to\infty} \big[f(x) - mx\big] \qquad \text{of} \qquad k = \lim_{x\to-\infty} \big[f(x) - mx\big]\,.$$

Proof. It suffices to shows that $m = \lim\limits_{x\to\infty} \dfrac{f(x)}{x}$ or $m = \lim\limits_{x\to-\infty} \dfrac{f(x)}{x}$. W.L.O.G., we assume that $\lim\limits_{x\to\infty} \big[f(x) - mx - k\big] = 0$. Then

$$\lim_{x\to\infty} \frac{f(x) - mx - k}{x} = 0\,.$$

On the other hand, $\lim\limits_{x\to\infty} \dfrac{mx+k}{x} = m$. By the fact that $\dfrac{f(x)}{x} = \dfrac{f(x) - mx - k}{x} + \dfrac{mx+k}{x}$, we find that $\lim\limits_{x\to\infty} \dfrac{f(x)}{x}$ exists and

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \Big[\frac{f(x) - mx - k}{x}\Big] + \lim_{x\to\infty} \frac{mx+k}{x} = m\,. \qquad \square$$

Example 1.71. In this example, we find all asymptotes of the graph of the function

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

Since the denominator vanishes at $x = \pm 1$, there are two possible vertical asymptotes $x = 1$ or $x = -1$. Since the denominator also vanishes at $x = 1$, we need to check further the behavior of $f(x)$ as x approaches 1. Note that for $x \neq \pm 1$,

$$\frac{x - \sqrt[3]{x^3 - x^2 + x}}{x^2 - 1} = \frac{x}{(x+1)\big[x^2 + x\sqrt[3]{x^3 - x^2 + x} + (x^3 - x^2 + x)^{\frac{2}{3}}\big]}\,;$$

thus for $x \neq \pm 1$,

$$f(x) = \frac{3x^4}{(x+1)\big[x^2 + x\sqrt[3]{x^3 - x^2 + x} + (x^3 - x^2 + x)^{\frac{2}{3}}\big]}\,.$$

Therefore, $\lim\limits_{x\to 1} f(x) = 0$ exists which shows that $x = 1$ is not a vertical asymptote of the graph of f. On the other hand,

$$\lim_{x\to -1^+} f(x) = \infty \qquad \text{and} \qquad \lim_{x\to -1^-} f(x) = -\infty\,,$$

we find that $x=-1$ is the only vertical asymptote of the graph of f.

For slant or horizontal asymptotes, we note that for $x \neq \pm 1, 0$,

$$\frac{f(x)}{x} = \frac{3}{\left(1+\frac{1}{x}\right)\left[1+\left(1-\frac{1}{x}+\frac{1}{x^2}\right)^{\frac{1}{3}}+\left(1-\frac{1}{x}+\frac{1}{x^2}\right)^{\frac{2}{3}}\right]}\,. \tag{1.4.1}$$

Since $\lim\limits_{x\to\pm\infty}\frac{1}{x}=0$, we find that $\lim\limits_{x\to\infty}\frac{f(x)}{x}=1$ and $\lim\limits_{x\to-\infty}\frac{f(x)}{x}=1$. It remains to find the limit $\lim\limits_{x\to\infty}\big[f(x)-x\big]$ and $\lim\limits_{x\to-\infty}\big[f(x)-x\big]$. Using (1.4.1),

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

Noting that the denominator approaches 3 as x approaches $\pm\infty$, we only focus on the limit of the numerator. Since

$$\begin{aligned}
&3x-(x+1)\Big[1+\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}+\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{2}{3}}\Big]\\
&\quad = 3x-(x+1)\Big[3+\Big(\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}-1\Big)+\Big(\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{2}{3}}-1\Big)\Big]\\
&\quad = -3-\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}-1\Big]\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}+2\Big]\\
&\qquad -x\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}-1\Big]\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}+2\Big],
\end{aligned}$$

to find the limit of the numerator as $x\to\pm\infty$ it suffices to find the limit

$$\lim_{x\to\infty} x\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}-1\Big] \qquad \text{and} \qquad \lim_{x\to-\infty} x\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}-1\Big].$$

Now, by Theorem 1.40,

$$\begin{aligned}
\lim_{x\to\infty} x\Big[\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}-1\Big] &= \lim_{x\to 0^+}\frac{\big(1-x+x^2\big)^{\frac{1}{3}}-1}{x}\\
&= \lim_{x\to 0^+}\frac{x-1}{\big(1-x+x^2\big)^{\frac{2}{3}}+\big(1-x+x^2\big)^{\frac{1}{3}}+1} = -\frac{1}{3}
\end{aligned}$$

and similarly, $\lim\limits_{x\to-\infty} x\Big[\big(1-\frac{1}{x}+\frac{1}{x^2}\big)^{\frac{1}{3}}-1\Big]=-\frac{1}{3}$. Therefore,

$$\lim_{x\to\pm\infty}\Big[3x-(x+1)\Big[1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{1}{3}}+\Big(1-\frac{1}{x}+\frac{1}{x^2}\Big)^{\frac{2}{3}}\Big]\Big] = -3+\frac{1}{3}\cdot 3 = -2\,;$$

thus $\lim\limits_{x\to\pm\infty}\big[f(x)-x\big]=-\frac{2}{3}$ which implies that $y=x-\frac{2}{3}$ is the only slant asymptote of the graph of f.

1.5 Exercise

Problem 1.1. Let f be given in Example 1.7 and $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(x) = \begin{cases} f(x) & \text{if } x > 0\,, \\ f(-x) & \text{if } x < 0\,, \\ 1 & \text{if } x = 0\,. \end{cases}$$

Find $\lim\limits_{x\to 0} g(x)$.

Problem 1.2. Let f be a function defined on an open interval containing c (except possibly at c).

1. Prove that if $\lim\limits_{x\to c} f(x) = L$, then $\lim\limits_{x\to c} \big|f(x)\big| = |L|$.
2. Prove that $\lim\limits_{x\to c} f(x) = L$ if and only if $\lim\limits_{x\to c} \big|f(x) - L\big| = 0$.

Problem 1.3. Let f be a function defined on an open interval containing c and $\lim\limits_{x\to c} f(x)$ exists. Show that there exist $\delta > 0$ and $M > 0$ such that

$$\big|f(x)\big| \leqslant M \qquad \text{whenever} \quad |x - c| < \delta\,.$$

Problem 1.4. Let f, g be a function defined on an open interval containing c (except possibly at c), and $f(x) \leqslant g(x)$ for all $x \neq c$. Prove that if $\lim\limits_{x\to c} f(x) = L$ and $\lim\limits_{x\to c} g(x) = K$ both exist, then $L \leqslant K$.

Problem 1.5. 1. Suppose that $\lim\limits_{x\to 2} \dfrac{f(x) - 5}{x - 2} = 3$. Find $\lim\limits_{x\to 2} f(x)$.

2. Suppose that $\lim\limits_{x\to 2} \dfrac{f(x) - 5}{x - 2} = 4$. Find $\lim\limits_{x\to 2} f(x)$.
3. Suppose that $\lim\limits_{x\to c} \dfrac{f(x) - p(x)}{x - c} = L$ exists, where p is a polynomial function. Find $\lim\limits_{x\to c} f(x)$.

Problem 1.6. Suppose that you are given $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$. Compute the following limits:

1. $\lim\limits_{x\to 0} \dfrac{\sin(x^2)}{x}$. 2. $\lim\limits_{x\to 0} \dfrac{x \sin x}{1 - \cos x}$. 3. $\lim\limits_{x\to 0} \dfrac{\sin(\sin(\sin x))}{x}$.
4. $\lim\limits_{x\to 0} \dfrac{\sin(x + c) - \sin c}{x}$, where c is a real number.

Problem 1.7. Show that $\lim\limits_{x\to 0^+} x^{\frac{3}{4}}\cos\dfrac{1}{x^2} = 0$ using (1) ε-δ definition and (2) the Squeeze theorem.

Problem 1.8.
1. Find the limits $\lim\limits_{x\to 2^+}\dfrac{x^2+x-6}{|x-2|}$ and $\lim\limits_{x\to 2^-}\dfrac{x^2+x-6}{|x-2|}$. Determine whether the limit $\lim\limits_{x\to 2}\dfrac{x^2+x-6}{|x-2|}$ exists or not.
2. Let $f:\mathbb{R}\to\mathbb{R}$ be defined by
$$f(x)=\begin{cases} a+\sin(x-2) & \text{if } x>2\,,\\ x^2-3x+b & \text{if } x\leqslant 2\,.\end{cases}$$
Find the relation between a and b so that the limit $\lim\limits_{x\to 2} f(x)$ exists.
3. Let $g:\mathbb{R}\to\mathbb{R}$ be defined by
$$g(x)=\begin{cases} 1+\sin(x-2) & \text{if } x>2\,,\\ x^2-3x+3 & \text{if } x\leqslant 2\,.\end{cases}$$
Find the limit $\lim\limits_{x\to 2}\dfrac{g(x)-1}{x-2}$ using the left limit and right limit criteria.

Problem 1.9. Let I be an open interval in $\mathbb{R}$, $c\in I$, and $f:I\to\mathbb{R}$ be a function. Show that f is continuous at c if and only if $\lim\limits_{h\to 0} f(c+h)=f(c)$.

Problem 1.10. Let $f:\mathbb{R}\to\mathbb{R}$ be a function satisfying $f(a+b)=f(a)f(b)$ for all $a,b\in\mathbb{R}$.

1. Show that $f(x)\geqslant 0$ for all $x\in\mathbb{R}$.
2. Show that if f is continuous at 0, then f is continuous on $\mathbb{R}$ (that is, f is continuous at every point of $\mathbb{R}$).

Problem 1.11. Let I be an interval in $\mathbb{R}$ and $f,g:I\to\mathbb{R}$ be continuous functions. Show that if $f(x)=g(x)$ for all $x\in\mathbb{Q}\cap I$, then $f(x)=g(x)$ for all $x\in I$.

Problem 1.12. Let I be an interval, $c\in I$, and $f:I\to\mathbb{R}$ be a continuous function. Show that if $f(c)\neq 0$, there exists $\delta>0$ such that $f(x)f(c)>0$ whenever $|x-c|<\delta$ and $x\in I$.

Problem 1.13. Construct a function $f:\mathbb{R}\to\mathbb{R}$ so that f is continuous at all integers but nowhere else.

Problem 1.14. Find the following limits:

1. $\lim_{x\to-\infty}(2x+\sqrt{4x^2+3x-2})$.

2. $\lim_{x\to\infty}\left(x-\sqrt[3]{x^3+2x-3}\right)$.

3. $\lim_{x\to\infty}\frac{[\![x]\!]}{x}$, where $[\![\cdot]\!]$ is the floor function.

Problem 1.15. Show that the equation $x^3-15x+1=0$ has three solutions in the interval $[-4,4]$.

Problem 1.16. Suppose that a and b are positive constants. Show that the equation

$$\frac{a}{x^3+2x^2-1}+\frac{b}{x^3+x-2}=0$$

has at least one solution in the interval $(-1,1)$.

Problem 1.17. True or False: Determine whether the following statements are true or false. If it is true, prove it. Otherwise, give a counter-example.

1. If f and g are functions such that $\lim_{x\to c^+} g(x)=K$, $\lim_{y\to K^+} f(y)=f(K)$, then

   $$\lim_{x\to c^+}(f\circ g)(x)=f(K)\,.$$

   How about if $x\to c^+$ and $y\to K^+$ are replaced by $x\to c^-$ and $y\to K^-$, respectively?

2. Let f,g be a function defined on an open interval containing c (except possibly at c), and $f(x)<g(x)$ for all $x\neq c$. If $\lim_{x\to c} f(x)=L$ and $\lim_{x\to c} g(x)=K$ both exist, then $L<K$.

3. If $|f|$ is continuous at c, so is f.

4. Let I be an interval and $f:I\to\mathbb{R}$ be a continuous function. If $f(x)\neq 0$ for all $x\in I$, then f never change signs; that is, either $f(x)>0$ for all $x\in I$ or $f(x)<0$ for all $x\in I$.

5. If $\lim_{x\to c} f(x)=\infty$ and $\lim_{x\to c}\big[f(x)-g(x)\big]=0$, then $\lim_{x\to c} g(x)=\infty$.

Chapter 2

Differentiation

2.1 The Derivatives of Functions

Definition 2.1

Let f be a function defined on an open interval containing c. If the limit $\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = m$ exists, then the line passing through $\big(c, f(c)\big)$ with slope m is the tangent line to the graph of f at point $\big((c, f(c)\big)$.

Definition 2.2

Let f be a function defined on an open interval I containing c. f is said to be differentiable at c if the limit

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x}$$

exists. If the limit above exists, the limit is denoted by $f'(c)$ and called the derivative of f at c. When the derivative of f at each point of I exists, f is said to be differentiable on I and the derivative of f is a function denoted by f'.

• **Notation**: The prime notation $'$ is associated with a function (of one variable) and is used to denote the derivative of that function. For a given function f defined on an open interval I and x being the name of the variable, the limit operation

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

is denoted by $\dfrac{d}{dx}f(x)$ $\big($or $\dfrac{df(x)}{dx}$ or even $\dfrac{dy}{dx}$ if $y = f(x)\big)$, and the limit

$$\lim_{\Delta x \to 0} \frac{f(c+\Delta) - f(c)}{\Delta x}$$

is denoted by $\dfrac{d}{dx}\Big|_{x=c} f(x)$ but not $\dfrac{d}{dx}f(c)$ $\big(\frac{d}{dx}f(c)$ is in fact $0\big)$. The operator $\dfrac{d}{dx}$ is a differential operator called the differentiation and is applied to functions of variable x. However, for historical (and convenient) reason, $\dfrac{d}{dx}f(x)$ is sometimes denoted by $(f(x))'$ (so that $'$ is treated as the differential operator $\dfrac{d}{dx}$) and f' is sometimes denoted by $\dfrac{df}{dx}$ (so that f is always treated as a function of variable x).

Remark 2.3. Letting $x = c + \Delta x$ in the definition of the derivatives, then

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

if the limit exists.

Example 2.4. Let f be a constant function. Then f' is the zero function.

Example 2.5. Let $f(x) = x^n$, where n is a positive integer. Then

$$f(x+\Delta x) = x^n + C^n_1 x^{n-1}\Delta x + C^n_2 x^{n-2}(\Delta x)^2 + \cdots + C^n_{n-1} x(\Delta x)^{n-1} + (\Delta x)^n\,;$$

thus if $\Delta x \neq 0$,

$$\frac{f(x+\Delta x) - f(x)}{\Delta x} = nx^{n-1} + C^n_2 x^{n-2}\Delta x + \cdots + C^n_{n-1} x(\Delta x)^{n-2} + (\Delta x)^{n-1}\,.$$

The limit on the right-hand side is clearly nx^{n-1}, so we establish that

$$\frac{d}{dx}x^n = nx^{n-1}\,.$$

Example 2.6. Now suppose that $f(x) = x^{-n}$, where n is a positive integer. Then if $x + \Delta x \neq 0$,

$$f(x+\Delta x) = \frac{1}{x^n + C^n_1 x^{n-1}\Delta x + C^n_2 x^{n-2}(\Delta x)^2 + \cdots + C^n_{n-1} x(\Delta x)^{n-1} + (\Delta x)^n}\,;$$

thus if $x \neq 0$, $\Delta x \neq 0$, and $x + \Delta x \neq 0$ (which can be achieved if $|\Delta x| \ll 1$),

$$\frac{f(x+\Delta x) - f(x)}{\Delta x} = \frac{-\big[C^n_1 x^{n-1} + C^n_2 x^{n-2}\Delta x + \cdots + C^n_{n-1} x(\Delta x)^{n-2} + (\Delta x)^{n-1}\big]}{x^n\big[x^n + C^n_1 x^{n-1}\Delta x + C^n_2 x^{n-2}(\Delta x)^2 + \cdots + C^n_{n-1} x(\Delta x)^{n-1} + (\Delta x)^n\big]}\,.$$

Therefore, if $x \neq 0$,

$$\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \frac{-nx^{n-1}}{x^{2n}} = -nx^{-n-1}$$

which shows $\dfrac{d}{dx}x^{-n} = -nx^{-n-1}$.

Combining the previous three examples, we conclude that

$$\frac{d}{dx}x^n = \begin{cases} nx^{n-1} & \forall\, x \in \mathbb{R} \quad \text{if } n \in \mathbb{N} \cup \{0\}\,, \\ nx^{n-1} & \forall\, x \neq 0 \quad \text{if } n \in \mathbb{Z} \text{ and } n < 0\,. \end{cases} \tag{2.1.1}$$

Combining Example 2.4-2.6, we conclude that

$$\frac{d}{dx}x^n = \begin{cases} nx^{n-1} & \forall\, x \in \mathbb{R} \quad \text{if } n \in \mathbb{N} \cup \{0\}\,, \\ nx^{n-1} & \forall\, x \neq 0 \quad \text{if } n \in \mathbb{Z} \text{ and } n < 0\,. \end{cases} \tag{2.1.2}$$

我們注意到當 n 是負整數時，在計算 $\left.\dfrac{d}{dx}\right|_{x=c} x^n$ 時，已經必須先假設 $c \neq 0$ 才能計算導數，並非最後算出來 $\left.\dfrac{d}{dx}\right|_{x=c} x^n = nc^{n-1}$ 時發現 c 不可為零所以不能代入。這是一個非常重要的觀念！不能搞錯順序！

Example 2.7. Let $f(x) = \sin x$. By the sum and difference formula,

$$\begin{aligned} f(x+\Delta x) - f(x) &= \sin(x+\Delta x) - \sin x = \sin x \cos \Delta x + \sin \Delta x \cos x - \sin x \\ &= \sin x(\cos \Delta x - 1) + \sin \Delta x \cos x\,; \end{aligned}$$

thus by the fact that $\displaystyle\lim_{x \to 0} \frac{\sin x}{x} = 1$ and $\displaystyle\lim_{x \to 0} \frac{\cos x - 1}{x} = 0$, we find that

$$\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \Big[\sin x \frac{\cos \Delta x - 1}{\Delta x} + \frac{\sin \Delta x}{\Delta x} \cos x\Big] = \cos x\,. \tag{2.1.3}$$

In other words, the derivative of the sine function is cosine.

On the other hand, let $g(x) = \cos x$. Then $g(x) = -f\big(x - \frac{\pi}{2}\big)$. Then if $\Delta x \neq 0$,

$$\frac{g(x+\Delta x) - g(x)}{\Delta x} = -\frac{f\big(x - \frac{\pi}{2} + \Delta x\big) - f\big(x - \frac{\pi}{2}\big)}{\Delta x}\,;$$

thus

$$\lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} = -\cos\big(x - \frac{\pi}{2}\big) = -\sin x\,.$$

In other words, the derivative of the cosine function is minus sine. To summarize,

$$\frac{d}{dx}\sin x = \cos x \qquad \text{and} \qquad \frac{d}{dx}\cos x = -\sin x\,. \tag{2.1.4}$$

Example 2.8. Consider the function $g : \mathbb{R} \to \mathbb{R}$ defined by

$$g(x) = \left\{ \begin{array}{cl} x^2 & \text{if } x \text{ is rational}\,, \\ -x^2 & \text{if } x \text{ is irrational}\,. \end{array} \right.$$

Then $g(x) = xf(x)$, where f is given in Example 1.22. By the fact that $\lim\limits_{x \to 0} f(x) = 0$,

$$\lim_{\Delta x \to 0} \frac{g(\Delta x) - g(0)}{\Delta x} = \lim_{\Delta x \to 0} f(\Delta x) = 0\,.$$

In other words, g is differentiable at 0. Moreover, similar argument used to explain that the function f in Example 1.22 is only continuous at 0 can be used to show that the function g is only continuous at 0. Therefore, we obtain a function which is differentiable at one point but discontinuous elsewhere.

Remark 2.9. If f is a function defined on a interval I, and c is one of the end-point. Then it is possible to define the one-sided derivative. For example, if c is the left end-point of I, then we can consider the limit

$$\lim_{\Delta x \to 0^+} \frac{f(c + \Delta x) - f(c)}{\Delta x} = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c}$$

if it exists. The limit above, if exists, is called the derivatives of f at c from the right.

Theorem 2.10: 可微必連續

Let f be a function defined on an open interval I, and $c \in I$. If f is differentiable at c, then f is continuous at c.

Proof. If $x \neq c$, $f(x) - f(c) = \dfrac{f(x) - f(c)}{x - c}(x - c)$. Since the limit $\lim\limits_{x \to c} \dfrac{f(x) - f(c)}{x - c}$ exists and $\lim\limits_{x \to c}(x - c) = 0$, by Theorem 1.14 we conclude that

$$\lim_{x \to c} \big[f(x) - f(c)\big] = \Big(\lim_{x \to c} \frac{f(x) - f(c)}{x - c}\Big)\Big(\lim_{x \to c}(x - c)\Big) = 0\,.$$

Therefore, $\lim\limits_{x \to c} f(x) = f(c)$ which shows that f is continuous at c. $\square$

Remark 2.11. When f is continuous on an open interval I, f is **not** necessary differentiable on I. For example, consider $f(x) = |x|$. Then Theorem 1.14 implies that f is continuous on I, but $\lim\limits_{\Delta x \to 0} \dfrac{f(\Delta x) - f(0)}{\Delta x} = \lim\limits_{\Delta x \to 0} \dfrac{|\Delta x|}{\Delta x}$ D.N.E.

2.2 Rules of Differentiation

Theorem 2.12

We have the following differentiation rules:

1. If k is a constant, then $\dfrac{d}{dx}k = 0$.
2. If n is a non-zero integer, then $\dfrac{d}{dx}x^n = nx^{n-1}$ (whenever x^{n-1} makes sense).
3. $\dfrac{d}{dx}\sin x = \cos x$, $\dfrac{d}{dx}\cos x = -\sin x$.
4. If k is a constant and $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$, then kf is differentiable at c and
$$\frac{d}{dx}\Big|_{x=c}\big[kf(x)\big] = kf'(c)\,.$$
5. If $f, g : (a, b) \to \mathbb{R}$ are differentiable at $c \in (a, b)$, then $f \pm g$ is differentiable at c and
$$\frac{d}{dx}\Big|_{x=c}\big[f(x) \pm g(x)\big] = f'(c) \pm g'(c)\,.$$

Proof of 5. Let $h(x) = f(x) + g(x)$. Then if $\Delta x \neq 0$,

$$\frac{h(c + \Delta x) - h(c)}{\Delta x} = \frac{f(c + \Delta x) - f(c)}{\Delta x} + \frac{g(c + \Delta x) - g(c)}{\Delta x}\,.$$

Since f, g are differentiable at c,

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x} = f'(c) \qquad \text{and} \qquad \lim_{\Delta x \to 0} \frac{g(c + \Delta x) - g(c)}{\Delta x}$$

exist. Therefore, by Theorem 1.14,

$$h'(c) = f'(c) + g'(c)\,.$$

The conclusion for the difference can be proved in the same way. □

Example 2.13. Let $f(x) = 3x^2 - 5x + 7$. Then

$$\begin{aligned}\frac{d}{dx}f(x) &= \frac{d}{dx}(3x^2 - 5x) + \frac{d}{dx}7 = \frac{d}{dx}(3x^2) - \frac{d}{dx}(5x)\\ &= 3\frac{d}{dx}x^2 - 5\frac{d}{dx}x = 3 \cdot (2x) - 5 = 6x - 5\,.\end{aligned}$$

In general, for a polynomial function

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \equiv \sum_{k=0}^{n} a_k x^k\,,$$

where $a_0, a_1, \cdots, a_n \in \mathbb{R}$, by induction we can show that

$$\frac{d}{dx}p(x) = n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1 = \sum_{k=1}^{n} k a_k x^{k-1}\,.$$

Theorem 2.14: Product Rule

Let $f, g : (a,b) \to \mathbb{R}$ be real-valued functions, and $c \in (a,b)$. If f and g are differentiable at c, then fg is differentiable at c and

$$\frac{d}{dx}\Big|_{x=c}(fg)(x) = f'(c)g(c) + f(c)g'(c)\,.$$

Proof. Let $h(x) = f(x)g(x)$. Then

$$\begin{aligned} h(c+\Delta x) - h(c) &= f(c+\Delta x)g(c+\Delta x) - f(c)g(c) \\ &= f(c+\Delta x)g(c+\Delta x) - f(c)g(c+\Delta x) + f(c)g(c+\Delta x) - f(c)g(c) \\ &= \big[f(c+\Delta x) - f(c)\big]g(c+\Delta x) + f(c)\big[g(c+\Delta x) - g(c)\big]\,. \end{aligned}$$

Therefore, if $\Delta x \neq 0$,

$$\frac{h(c+\Delta x) - h(c)}{\Delta x} = \frac{f(c+\Delta x) - f(c)}{\Delta x}g(c+\Delta x) + f(c)\frac{g(c+\Delta x) - g(c)}{\Delta x}\,.$$

Since f, g are differentiable at c,

$$\lim_{\Delta x \to 0}\frac{f(c+\Delta x) - f(c)}{\Delta x} = f'(c)\,,\ \lim_{\Delta x \to 0}\frac{g(c+\Delta x) - g(c)}{\Delta x}\,,\ \text{and}\ \lim_{\Delta x \to 0} g(c+\Delta x) = g(c)$$

exist. By Theorem 1.14,

$$h'(c) = f'(c)g(c) + f(c)g'(c)$$

which concludes the product rule. $\square$

Example 2.15. Let $f(x) = x^3 \sin x$. Then the product rule implies that

$$f'(x) = 3x^2 \sin x + x^3 \cos x\,.$$

Theorem 2.16: Quotient Rule

Let $f, g : (a, b) \to \mathbb{R}$ be real-valued functions, and $c \in (a, b)$. If f and g are differentiable at c and $g(c) \neq 0$, then $\dfrac{f}{g}$ is differentiable at c and

$$\frac{d}{dx}\Big|_{x=c} \frac{f}{g}(x) = \frac{f'(c)g(c) - f(c)g'(c)}{g(c)^2}\,.$$

Proof. Let $h(x) = \dfrac{f(x)}{g(x)}$. Then

$$\begin{aligned} h(c+\Delta x) - h(c) &= \frac{f(c+\Delta x)}{g(c+\Delta x)} - \frac{f(c)}{g(c)} = \frac{f(c+\Delta x)g(c) - f(c)g(c+\Delta x)}{g(c)g(c+\Delta x)} \\ &= \frac{f(c+\Delta x)g(c) - f(c)g(c) + f(c)g(c) - f(c)g(c+\Delta x)}{g(c)g(c+\Delta x)} \\ &= \frac{\big[f(c+\Delta x) - f(c)\big]g(c) - f(c)\big[g(c+\Delta x) - g(c)\big]}{g(c)g(c+\Delta x)}\,. \end{aligned}$$

Therefore, if $\Delta x \neq 0$,

$$\frac{h(c+\Delta x) - h(c)}{\Delta x} = \frac{1}{g(c)g(c+\Delta x)}\Big[\frac{f(c+\Delta x) - f(c)}{\Delta x} g(c) - f(c)\frac{g(c+\Delta x) - g(c)}{\Delta x}\Big]\,.$$

Since f, g are differentiable at c,

$$\lim_{\Delta x \to 0} \frac{f(c+\Delta x) - f(c)}{\Delta x} = f'(c)\,, \lim_{\Delta x \to 0} \frac{g(c+\Delta x) - g(c)}{\Delta x}\,, \text{ and } \lim_{\Delta x \to 0} g(c+\Delta x) = g(c)$$

exist. By Theorem 1.14,

$$h'(c) = \frac{1}{g(c)^2}\Big[f'(c)g(c) - f(c)g'(c)\Big]$$

which concludes the quotient rule. □

Remark 2.17. Suppose that in addition to the assumption in Theorem 2.16 one has already known that $h = f/g$ is differentiable at c, then applying the product rule to $f = gh$ one finds that

$$f'(c) = g'(c)h(c) + g(c)h'(c) = g'(c)\frac{f(c)}{g(c)} + g(c)h'(c)$$

which, after rearranging terms, shows the quotient rule. The proof of Theorem 2.16 indeed is based on the fact that we do not know the differentiability of h at c yet.

Example 2.18. Let n be a positive integer and $f(x) = x^{-n}$. We have shown by definition that $f'(x) = -nx^{-n-1}$ if $x \neq 0$. Now we use Theorem 2.16 to compute the derivative of f: if $x \neq 0$,

$$\frac{d}{dx}x^{-n} = \frac{d}{dx}\frac{1}{x^n} = -\frac{\frac{d}{dx}x^n}{x^{2n}} = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1}\,.$$

Example 2.19. Since $\tan x = \dfrac{\sin x}{\cos x}$, by Theorem 2.16 we have

$$\frac{d}{dx}\tan x = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x\,.$$

Similarly, we also have

$$\begin{aligned}
\frac{d}{dx}\cot x &= \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} = -\csc^2 x\,,\\
\frac{d}{dx}\sec x &= -\frac{-\sin x}{\cos^2 x} = \sec x \tan x\,,\\
\frac{d}{dx}\csc x &= -\frac{\cos x}{\sin^2 x} = -\csc x \cot x\,.
\end{aligned}$$

We note that without using the quotient rule, the derivative of the tangent function can be found using the sum-and-difference formula

$$\tan(x-y) = \frac{\tan x - \tan y}{1 + \tan x \tan y}\,. \tag{2.2.1}$$

Using (2.2.1), we find that

$$\tan(x+\Delta x) - \tan x = \tan \Delta x\big[1 + \tan(x+\Delta x)\tan x\big]\,;$$

thus if $\Delta x \neq 0$,

$$\frac{\tan(x+\Delta x) - \tan x}{\Delta x} = \frac{\sin \Delta x}{\Delta x}\cdot\frac{1 + \tan(x+\Delta x)\tan x}{\cos \Delta x}$$

which, using (1.2.2), shows that

$$\lim_{\Delta x \to 0}\frac{\tan(x+\Delta x) - \tan x}{\Delta x} = \Big(\lim_{\Delta x \to 0}\frac{\sin \Delta x}{\Delta x}\Big)\Big(\lim_{\Delta x \to 0}\frac{1 + \tan(x+\Delta x)\tan x}{\cos \Delta x}\Big) = \sec^2 x\,.$$

- **Higher-order derivatives**:

Let f be defined on an open interval $I = (a,b)$. If f' exists on I and possesses derivatives at every point in I, by definition we use f'' to denote the derivative of f'.In other words,

$$f''(x) = \frac{d}{dx}f'(x) = \frac{d}{dx}\frac{d}{dx}f(x) \equiv \frac{d^2}{dx^2}f(x) = \frac{d^2 f(x)}{dx^2}\Big(= \frac{d^2 y}{dx^2} \text{ if } y = f(x)\Big)\,.$$

The function f'' is called the second derivative of f. Similar as the "first" derivative case, $f''(c) = \frac{d^2}{dx^2}\Big|_{x=c} f(x)$.

The third derivatives and even higher-order derivatives are denoted by the following: if $y = f(x)$,

$$\begin{array}{rccc} \text{Third derivative: } y''' & f'''(x) & \dfrac{d^3}{dx^3}f(x) & \dfrac{d^3 f(x)}{dx^3} \\ \text{Fourth derivative: } y^{(4)} & f^{(4)}(x) & \dfrac{d^4}{dx^4}f(x) & \dfrac{d^4 f(x)}{dx^4} \\ & \vdots & & \\ \text{n-th derivative: } y^{(n)} & f^{(n)}(x) & \dfrac{d^n}{dx^n}f(x) & \dfrac{d^n f(x)}{dx^n}\,. \end{array}$$

2.3 The Chain Rule

The chain rule is used to study the derivative of composite functions.

Theorem 2.20: Chain Rule - 連鎖律

Let I, J be open intervals, $f : J \to \mathbb{R}$, $g : I \to \mathbb{R}$ be real-valued functions, and the range of g is contained in J. If g is differentiable at $c \in I$ and f is differentiable at $g(c)$, then $f \circ g$ is differentiable at c and

$$\frac{d}{dx}\Big|_{x=c}(f \circ g)(x) = f'(g(c))g'(c)\,.$$

Proof. To simplify the notation, we set $d = g(c)$.

Let $\varepsilon > 0$ be given. Since f is differentiable at d and g is differentiable at c, there exist $\delta_1, \delta_2 > 0$ such that

$$\begin{aligned} \Big|\frac{f(d+k) - f(d)}{k} - f'(d)\Big| &< \frac{\varepsilon}{2(1 + |g'(c)|)} \quad \text{whenever} \quad 0 < |k| < \delta_1\,, \\ \Big|\frac{g(c+h) - g(c)}{h} - g'(c)\Big| &< \min\Big\{1, \frac{\varepsilon}{2(1 + |f'(d)|)}\Big\} \quad \text{whenever} \quad 0 < |h| < \delta_2\,. \end{aligned}$$

Therefore,

$$\begin{aligned} \big|f(d+k) - f(d) - f'(d)k\big| &\leqslant \frac{\varepsilon}{2(1 + |g'(c)|)}|k| \quad \text{whenever} \quad |k| < \delta_1\,, \\ \big|g(c+h) - g(c) - g'(c)h\big| &\leqslant \min\Big\{1, \frac{\varepsilon}{2(1 + |f'(d)|)}\Big\}|h| \quad \text{whenever} \quad |h| < \delta_2\,. \end{aligned}$$

By Theorem 2.10, g is continuous at c; thus $\lim_{h\to 0} g(c+h) = g(c)$. This fact provides $\delta_3 > 0$ such that

$$\big|g(c+h) - g(c)\big| < \delta_1 \quad \text{whenever} \quad |h| < \delta_3\,.$$

Define $\delta = \min\{\delta_2, \delta_3\}$. Then $\delta > 0$. Moreover, if $|h| < \delta$, the number $k \equiv g(c+h) - g(c)$ satisfies $|k| < \delta_1$. As a consequence, if $|h| < \delta$,

$$\begin{aligned}
&\big|(f\circ g)(c+h) - (f\circ g)(c) - f'(d)g'(c)h\big| = \big|f(g(c+h)) - f(d) - f'(d)g'(c)h\big| \\
&\quad = \big|f(d+k) - f(d) - f'(d)g'(c)h\big| \\
&\quad = \big|f(d+k) - f(d) - f'(d)k + f'(d)k - f'(d)g'(c)h\big| \\
&\quad \leqslant \big|f(d+k) - f(d) - f'(d)k\big| + \big|f'(d)\big|\big|k - g'(c)h\big| \\
&\quad \leqslant \frac{\varepsilon}{2(1+|g'(c)|)}|k| + \big|f'(d)\big|\big|g(c+h) - g(c) - g'(c)h\big| \\
&\quad \leqslant \frac{\varepsilon}{2(1+|g'(c)|)}\big(|k - g'(c)h| + |g'(c)||h|\big) + \big|f'(d)\big|\frac{\varepsilon}{2(1+|f'(d)|)} \\
&\quad \leqslant \frac{\varepsilon}{2(1+|g'(c)|)}\big(|h| + |g'(c)||h|\big) + \big|f'(d)\big|\frac{\varepsilon|h|}{2(1+|f'(d)|)} \\
&\quad = \frac{\varepsilon}{2}|h| + \frac{\big|f'(d)\big|}{2(1+|f'(d)|)}\varepsilon|h|\,.
\end{aligned}$$

The inequality above implies that if $0 < |h| < \delta$,

$$\Big|\frac{(f\circ g)(c+h) - (f\circ g)(c)}{h} - f'(d)g'(c)\Big| \leqslant \frac{\varepsilon}{2} + \frac{\big|f'(d)\big|}{2(1+|f'(d)|)}\varepsilon < \varepsilon$$

which concludes the chain rule. □

How to memorize the chain rule? Let $y = g(x)$ and $u = f(y)$. Then the derivative $u = (f\circ g)(x)$ is $\dfrac{du}{dx} = \dfrac{du}{dy}\dfrac{dy}{dx}$.

Example 2.21. Let $f(x) = (3x - 2x^2)^3$. Then $f'(x) = 3(3x - 2x^2)^2(3 - 4x)$.

Example 2.22. Let $f(x) = \Big(\dfrac{3x-1}{x^2+3}\Big)^2$. Then

$$\begin{aligned}
f'(x) &= 2\Big(\frac{3x-1}{x^2+3}\Big)^{2-1}\frac{d}{dx}\frac{3x-1}{x^2+3} = \frac{2(3x-1)}{x^2+3}\cdot\frac{3(x^2+3) - 2x(3x-1)}{(x^2+3)^2} \\
&= \frac{2(3x-1)(-3x^2+2x+9)}{(x^2+3)^3}\,.
\end{aligned}$$

Example 2.23. Let $f(x) = \tan^3\left[(x^2-1)^2\right]$. Then

$$\begin{aligned} f'(x) &= \Big\{3\tan^2\left[(x^2-1)^2\right]\sec^2\left[(x^2-1)^2\right]\Big\} \times \left[2(x^2-1)\cdot(2x)\right] \\ &= 12x(x^2-1)\tan^2\left[(x^2-1)^2\right]\sec^2\left[(x^2-1)^2\right]. \end{aligned}$$

Example 2.24. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x^2 \sin\dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0\,. \end{cases}$$

Then if $x \neq 0$, by the chain rule we have

$$\begin{aligned} f'(x) &= \Big(\frac{d}{dx}x^2\Big)\sin\frac{1}{x} + x^2\Big(\frac{d}{dx}\sin\frac{1}{x}\Big) = 2x\sin\frac{1}{x} + x^2\cos\frac{1}{x}\Big(\frac{d}{dx}\frac{1}{x}\Big) \\ &= 2x\sin\frac{1}{x} + x^2\cos\frac{1}{x}\Big(-\frac{1}{x^2}\Big) = 2x\sin\frac{1}{x} - \cos\frac{1}{x}\,. \end{aligned}$$

Next we compute $f'(0)$. If $\Delta x \neq 0$, we have

$$\Big|\frac{f(\Delta x) - f(0)}{\Delta x}\Big| = \Big|\Delta x \sin\frac{1}{\Delta x}\Big| \leqslant |\Delta x|\,;$$

thus $-|\Delta x| \leqslant \dfrac{f(\Delta x) - f(0)}{\Delta x} \leqslant |\Delta x|$ for all $\Delta x \neq 0$ and the Squeeze Theorem implies that

$$f'(0) = \lim_{\Delta x \to 0} \frac{f(\Delta x) - f(0)}{\Delta x} = 0\,.$$

Therefore, we conclude that

$$f'(x) = \begin{cases} 2x\sin\dfrac{1}{x} - \cos\dfrac{1}{x} & \text{if } x \neq 0\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

Definition 2.25

Let f be a function defined on an open interval I. f is said to be continuously differentiable on I if f is differentiable on I and f' is continuous on I.

The function f given in Example 2.24 is differentiable on $\mathbb{R}$ but not continuously differentiable since $\lim\limits_{x \to 0} f'(x)$ D.N.E.

2.4 Implicit Differentiation

An implicit function is a function that is defined implicitly by an equation that x and y satisfy, by associating one of the variables (the value y) with the others (the arguments x). For example, $x^2 + y^2 = 1$ and $x = \cos y$ are implicit functions. Sometimes we know how to express y in terms of x from the equation (such as the first case above $y = \sqrt{1-x^2}$ or $y = -\sqrt{1-x^2}$), while in most cases there is no way to know what the function y of x exactly is.

Given an implicit function (without solving for y in terms of x from the equation), can we find the derivative of y? This is the main topic of this section. We first focus on implicit functions of the form $f(x) = g(y)$. If $f(a) = g(b)$, we are interested in how the set $\{(x,y) \,|\, f(x) = g(y)\}$ looks like "mathematically" near (a,b).

Theorem 2.26: Implicit Function Theorem - 隱函數定理簡單版

Let f, g be continuously differentiable functions defined on some open intervals, and $f(a) = g(b)$. If $g'(b) \neq 0$, then there exists a unique continuously differentiable function $y = h(x)$, defined in an open interval containing a, satisfying that $b = h(a)$ and $f(x) = g(h(x))$.

Example 2.27. Let us compute the derivative of $h(x) = x^r$, where $r = \dfrac{p}{q}$ for some $p, q \in \mathbb{N}$ and $(p,q) = 1$. Write $y = h(x)$. Then $y^q = x^p$. Since $\dfrac{d}{dy} y^q = qy^{q-1} \neq 0$ if $y \neq 0$, by the Implicit Function Theorem we find that h is differentiable at every x satisfying $x \neq 0$. Since $h(x)^q = x^p$, by the chain rule we find that

$$qh(x)^{q-1}h'(x) = px^{p-1} \qquad \forall\, x \neq 0\,;$$

thus

$$h'(x) = \frac{p}{q} h(x)^{1-q} x^{p-1} = \frac{p}{q} x^{\frac{p}{q}(1-q)+p-1} = rx^{r-1} \qquad \forall\, x \neq 0\,.$$

If r is a negative rational number, we can apply the quotient and find that

$$\frac{d}{dx} x^r = \frac{d}{dx} \frac{1}{x^{-r}} = \frac{rx^{-r-1}}{x^{-2r}} = rx^{r-1} \qquad \forall\, x \neq 0\,.$$

Therefore, we conclude that

$$\frac{d}{dx} x^r = rx^{r-1} \qquad \forall\, x \neq 0\,. \tag{2.4.1}$$

Remark 2.28. The derivative of x^r can also be computed by first finding the derivative of $x^{\frac{1}{p}}$ $\Big($that is, find the limit $\displaystyle\lim_{\Delta x \to 0} \frac{(x+\Delta x)^{\frac{1}{p}} - x^{\frac{1}{p}}}{\Delta x}\Big)$ and then apply the chain rule.

Example 2.29. Suppose that y is an implicit function of x given that $y^3+y^2-5y-x^2 = -4$.

1. Find $\dfrac{dy}{dx}$.
2. Find the tangent line passing through the point $(3,-1)$.

Let $f(x) = x^2 - 4$ and $g(y) = y^3 + y^2 - 5y$. Then $g'(y) = 3y^2 + 2y - 5$; thus if $y \neq 1$ or $y \neq -\dfrac{5}{3}$ $\Big($or equivalently, $x \neq \pm 1$ or $x \neq \pm\sqrt{\dfrac{283}{27}}\Big)$,

$$\frac{dy}{dx} = \frac{2x}{3y^2 + 2y - 5}\,.$$

Since $(1,-3)$ satisfies the relation $y^3 + y^2 - 5y - x^2 = -4$, the slope of the tangent line passing through $(3,-1)$ is $\dfrac{2\cdot 3}{3(-1)^2 + 2(-1) - 5} = -\dfrac{3}{2}$; thus the desired tangent line is

$$y = -\frac{3}{2}(x-3) - 1\,.$$

Example 2.30. Find $\dfrac{dy}{dx}$ implicitly for the equation $\sin y = x$.

Let $f(x) = x$ and $g(y) = \sin y$. Then $g'(y) = \cos y$; thus if $y \neq n\pi + \dfrac{\pi}{2}$ (or equivalently, $x \neq \pm 1$),

$$\frac{dy}{dx} = \frac{1}{\cos y}\,. \tag{2.4.2}$$

Similarly, for function y defined implicitly by $\cos y = x$, we find that if $y \neq n\pi$ (or equivalently, $x \neq \pm 1$),

$$\frac{dy}{dx} = -\frac{1}{\sin y}\,. \tag{2.4.3}$$

Remark 2.31. The curve consisting of points (x, y) satisfying the relation $\sin y = x$ cannot be the graph of a function since one x may corresponds to several y; however, the curve consisting of points (x, y) satisfying the relation $\sin y = x$ as well as $-\dfrac{\pi}{2} < y < \dfrac{\pi}{2}$ is the graph of a function called arcsin. In other words, for each $x \in (-1,1)$, there exists a unique $y \in \big(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\big)$ satisfying $\sin y = x$, and such y is denoted by $\arcsin x$. Since for $y \in \big(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\big)$ we must have $\cos y > 0$, by the fact that $\sin^2 y + \cos^2 y = 1$, using (2.4.2) we find that

$$\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}} \qquad \forall\, x \in (-1,1)\,. \tag{2.4.4}$$

Similarly, the curve consisting of points (x, y) satisfying the relation $\cos y = x$ as well as $0 < y < \pi$ is the graph of a function called arccos, and (2.4.3) implies that

$$\frac{d}{dx}\arccos x = -\frac{1}{\sqrt{1-x^2}} \qquad \forall\, x \in (-1, 1)\,. \tag{2.4.5}$$

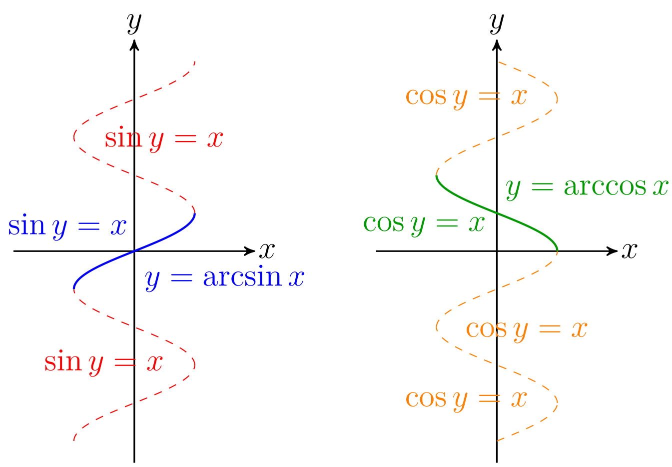


Figure 2.1: The graph of functions $y = \arcsin x$ and $y = \arccos x$

There are, unfortunately, many implicit functions that are not given by the equation of the form $f(x) = g(y)$. Nevertheless, there is a more powerful version of the Implicit Function Theorem that guarantees the continuous differentiability of the implicit functions defined through complicated relations between x and y (written in the form $f(x, y) = 0$). In the following, we always assume that the implicit function given by the equation that x and y satisfy is differentiable.

Example 2.32. Find the second derivative of the implicit function given by the equation $y = \cos(5x - 3y)$.

Differentiate in x once, we find that $\dfrac{dy}{dx} = -\sin(5x-3y)\cdot\big(5 - 3\dfrac{dy}{dx}\big)$; thus

$$\frac{dy}{dx} = \frac{-5\sin(5x-3y)}{1-3\sin(5x-3y)} = \frac{5}{3}\Big[1 - \frac{1}{1-3\sin(5x-3y)}\Big]\,. \tag{2.4.6}$$

Differentiate the equation above in x, we obtain that

$$\frac{d^2y}{dx^2} = -\frac{5}{3}\cdot\frac{3\cos(5x-3y)(5-3y')}{\big[1-3\sin(5x-3y)\big]^2} = -\frac{5\cos(5x-3y)(5-3y')}{\big[1-3\sin(5x-3y)\big]^2}$$

and (2.4.6) further implies that $\dfrac{d^2y}{dx^2} = -\dfrac{25\cos(5x-3y)}{\big[1-3\sin(5x-3y)\big]^3}\,.$

Example 2.33. Show that if it is possible to draw three normals from the point $(a, 0)$ to the parabola $x = y^2$, then $a > \frac{1}{2}$.

Suppose that the line L connecting $(a, 0)$ and (b^2, b), where $b \neq 0$, is normal to the parabola $x = y^2$. The derivative of the function defined implicitly by $x = y^2$ satisfies that

$$1 = 2y\frac{dy}{dx};$$

thus the slope of the tangent line passing through (b^2, b) is $\frac{1}{2b}$. Since line L is perpendicular to the tangent line passing through (b^2, b), we must have

$$\frac{1}{2b} \cdot \frac{b-0}{b^2-a} = -1\,.$$

Therefore, $a = \frac{1}{2} + b^2$. Since $b \neq 0$, $a > \frac{1}{2}$.

2.5 Exercise

Problem 2.1. Let f be a function defined on an open interval containing c. Show that f is differentiable at c if and only if there exists a real number L satisfying that for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\big|f(c+h) - f(c) - Lh\big| \leqslant \varepsilon|h| \qquad \text{whenever} \quad |h| < \delta\,.$$

Hint: See the (first part of the) proof of the chain rule for reference.

Problem 2.2. Let f, g be functions defined on an open interval, and $n \in \mathbb{N}$. Show that if the n-th derivatives of f and g exist on I, then

$$\begin{aligned}\frac{d^n}{dx^n}(fg)(x) &= f^{(n)}(x)g(x) + C^n_1 f^{(n-1)}(x)g'(x) + C^n_2 g^{(n-2)}(x)g''(x) + \cdots \\ &\qquad + C^n_{n-2} f''(x)g^{(n-2)}(x) + C^n_{n-1} f'(x)g^{(n-1)}(x) + f(x)g^{(n)}(x) \\ &= \sum_{k=0}^{n} C^n_k f^{(n-k)}(x)g^{(k)}(x)\,,\end{aligned}$$

where $C^n_k = \frac{n!}{k!(n-k)!}$ is "n choose k".

Hint: Prove by induction.

Problem 2.3. Let I be an open interval and $c \in I$. The left-hand and right-hand derivative of f at c, denoted by $f'(c^+)$ and $f'(c^-)$, respectively, are defined by

$$f'(c^+) = \lim_{h \to 0^+} \frac{f(c+h) - f(c)}{h} \qquad \text{and} \qquad f'(c^-) = \lim_{h \to 0^-} \frac{f(c+h) - f(c)}{h}$$

provides the limits exist.

1. Show that if f is differentiable at c if and only if $f'(c^+) = f'(c^-)$, and in either case we have $f'(c) = f'(c^+) = f'(c^-)$.

2. Let $f(x) = \begin{cases} x^2 & \text{if } x \leqslant 2\,, \\ mx + k & \text{if } x > 2\,. \end{cases}$ Find the value of m and k such that f is differentiable at 2.

3. Is there a value of b that will make

$$g(x) = \begin{cases} x + b & \text{if } x < 0\,, \\ \cos x & \text{if } x \geqslant 0\,. \end{cases}$$

continuous at 0? Differentiable at 0? Give reasons for your answers.

Problem 2.4. 1. Let $n \in \mathbb{N}$. Show that $\sum\limits_{k=1}^{n-1} kx^{k-1} = \dfrac{(n-1)x^n - nx^{n-1} + 1}{(x-1)^2}$ if $x \neq 1$.

2. Show that $\sum\limits_{k=1}^{n} k\cos(kx) = \dfrac{-1 + (2n+1)\sin\frac{x}{2}\sin(n+\frac{1}{2})x + \cos\frac{x}{2}\cos(n+\frac{1}{2})x}{4\sin^2\frac{x}{2}}$ if $x \in (-\pi, \pi)$.

Hint 1. Find the sum $\sum\limits_{k=1}^{n-1} x^k$ first and then observe that $\sum\limits_{k=1}^{n-1} kx^{k-1} = \sum\limits_{k=1}^{n-1} \dfrac{d}{dx} x^k$.

2. Find the sum $\sum\limits_{k=1}^{n} \sin(kx)$ first and then observe that $\sum\limits_{k=1}^{n} k\cos(kx) = \sum\limits_{k=1}^{n} \dfrac{d}{dx}\sin(kx)$.

Problem 2.5. For a fixed constant $a > 1$, consider the function $f(x) = \log_a x$. Suppose that you are given the fact that the limit

$$\lim_{h \to 0} \frac{\log_{10}(1+h)}{h} \approx 0.43429$$

exists.

1. Show that f is differentiable on $(0, \infty)$ for all $a > 1$.

2. Show that there exists $a > 1$ such that $f'(x) = \dfrac{1}{x}$ for all $x \in (0, \infty)$.

Hint: 1. Use the "change of base formula"（换底公式）for logarithm.

2. Define $g(a) = \dfrac{d}{dx}\Big|_{x=1} \log_a x$. Apply the intermediate value theorem to g.

Problem 2.6. Let $f(x) = a_1 \sin x + a_2 \sin(2x) + a_3 \sin(3x) + \cdots + a_n \sin(nx)$, where $a_1, a_2, \cdots, a_n$ are real numbers and $n \in \mathbb{N}$. Show that if $|f(x)| \leqslant |\sin x|$ for all $x \in \mathbb{R}$, then

$$|a_1 + 2a_2 + 3a_3 + \cdots + na_n| \leqslant 1\,.$$

Problem 2.7. Let $k \in \mathbb{N}$. Suppose that $\dfrac{d^n}{dx^n}\dfrac{1}{x^k - 1} = \dfrac{p_n(x)}{(x^k - 1)^{n+1}}$. Find the degree of p_n and $p_n(1)$.

Problem 2.8. Let $f_1, f_2, \cdots, f_n : \mathbb{R} \to \mathbb{R}$ be differentiable functions (that is, f_j is differentiable on $\mathbb{R}$ for all $1 \leqslant j \leqslant n$), and

$$h(x) = (f_n \circ f_{n-1} \circ \cdots \circ f_2 \circ f_1)(x)\,.$$

Show that

$$h'(x) = f_n'\big(g_{n-1}(x)\big) \cdot f_{n-2}'\big(g_{n-2}(x)\big) \cdot \cdots \cdot f_2'\big(g_1(x)\big) \cdot f_1'(x)\,.$$

where $g_k = f_k \circ f_{k-1} \circ \cdots \circ f_2 \circ f_1$.

Hint: Prove by induction.

Problem 2.9. 1. Let $r \in \mathbb{Q}$, and $f : (0, \infty) \to \mathbb{R}$ be defined by $f(x) = x^r$. Find the derivative of f.

2. Find the derivatives of $y = x^{\frac{1}{4}}$ and $y = x^{\frac{3}{4}}$ by the fact that $x^{\frac{1}{4}} = \sqrt{\sqrt{x}}$ and $x^{\frac{3}{4}} = \sqrt{x\sqrt{x}}$.

3. Let $g : (a, b) \to \mathbb{R}$ be differentiable. Find the derivative of $y = |g(x)|$.

Problem 2.10. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and satisfy $f\big(\dfrac{x^2 - 1}{x^2 + 1}\big) = x$ for all $x > 0$. Find $f'(0)$.

Problem 2.11. 1. Let $n \in \mathbb{N}$. Show that $\dfrac{d}{dx}\big[\sin^n x \cos(nx)\big] = n \sin^{n-1} x \cos(n+1)x\,.$

2. Find a similar formula for the derivative of $\cos^n x \cos(nx)$.

Problem 2.12. Find the derivative of the following functions:

1. $y = \cos\sqrt{\sin(\tan(\pi x))}$. 2. $y = \left[x + (x + \sin^2 x)^3\right]^4$.

Problem 2.13. Note that in class we have introduced two new functions "arcsin" and "arccos" whose graphs are (the blue and green) part of the curve consisting of points (x, y) satisfying $\sin y = x$ and $\cos y = x$, respectively, given below

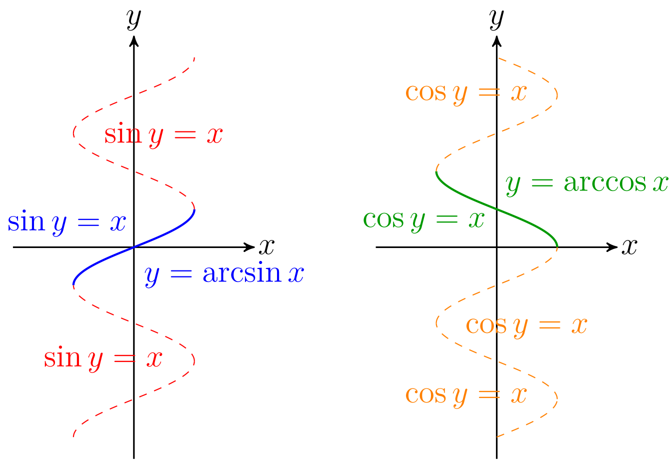


Figure 2.2: The graph of functions $y = \arcsin x$ and $y = \arccos x$

1. Find the domain and the range of the two functions arcsin and arccos.

2. Show that $\sin(\arcsin x) = x$ for all x in the domain of arcsin and $\cos(\arccos x) = x$ whenever x in the domain of arccos.

3. Is it true that $\arcsin(\sin x) = x$ or $\arccos(\cos x) = x$?

4. Find $\sin(\arccos x)$ and $\cos(\arcsin x)$.

5. Show that $\left.\frac{d}{dx}\right|_{x=c} \big(\arcsin x + \arccos x\big) = 0$ for all c in both domains.

6. Find $\frac{d}{dx} \arcsin \frac{1}{x}$ and $\frac{d}{dx}(\arccos x)^2$.

Problem 2.14. The function arctan is defined similarly to functions arcsin and arccos: consider the collection of all points (x, y) satisfying $\tan y = x$ (see the figure below), and the blue part is the graph of a function called "arctan".

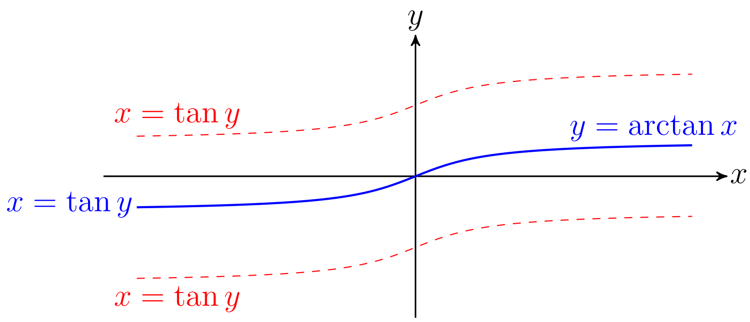


Figure 2.3: The graph of function $y = \arctan x$

1. Find the domain and the range of the function arctan.

2. Show that $\tan(\arctan x) = x$ for all x in the domain of arctan.

3. Is is true that $\arctan(\tan x) = x$ for all x in the domain of tan?

4. Find $\dfrac{d}{dx} \arctan x$.

Problem 2.15. Find $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2}$ if $\sin(x + y) = y^2 \cos x$.

Problem 2.16. The line that is normal to the curve $x^2 + 2xy - 3y^2 = 0$ at $(1, 1)$ intersects the curve at what other point?

Problem 2.17. Show that the sum of the x- and y-intercepts of any tangent line to the curve $\sqrt{x} + \sqrt{y} = \sqrt{c}$ is equal to c.

Problem 2.18. The Bessel function of order 0, denoted by $y = J_0(x)$, satisfies the differential equation

$$xy'' + y' + xy = 0$$

for all values of x and its value at 0 is $J_0(0) = 1$.

1. Find $J_0'(0)$.

2. Use implicit differentiation to find $J_0''(0)$.

Chapter 3

Applications of Differentiation

3.1 Extrema on an Interval

Definition 3.1

Let f be defined on an interval I containing c.

1. $f(c)$ is the minimum of f on I when $f(c) \leqslant f(x)$ for all x in I.
2. $f(c)$ is the maximum of f on I when $f(c) \geqslant f(x)$ for all x in I.

The minimum and maximum of a function on an interval are the extreme values, or extrema (the singular form of extrema is extremum), of the function on the interval. The minimum and maximum of a function on an interval are also called the absolute minimum and absolute maximum, or the global minimum and global maximum, on the interval. Extrema can occur at interior points or end-points of an interval. Extrema that occur at the end-points are called end-point extrema.

Theorem 3.2: Extreme Value Theorem - 極值定理

If f is continuous on a closed interval $[a, b]$, then f has both a minimum and a maximum on the interval.（連續函數在閉區間上必有最大最小值）

When f is continuous on an open interval (a, b) (or a half-open half-closed interval), it is still possibly that f attains its maximum or minimum but there is no guarantee. Moreover, it is also possible that f does not attain its extrema when f is continuous on an interval which is not closed.

Definition 3.3

Let f be defined on an interval I containing c.

1. If there is an open interval containing c on which $f(c)$ is a maximum, then $f(c)$ is called a relative maximum of f, or you can say that f has a relative maximum at $\big(c, f(c)\big)$.
2. If there is an open interval containing c on which $f(c)$ is a minimum, then $f(c)$ is called a relative minimum of f, or you can say that f has a relative minimum at $\big(c, f(c)\big)$.

The plural of relative maximum is relative maxima, and the plural of relative minimum is relative minima. Relative maximum and relative minimum are sometimes called local maximum and local minimum, respectively.

Definition 3.4

Let f be defined on an open interval containing c. The number/point c is called a critical number or critical point of f if $f'(c) = 0$ or if f is not differentiable at c.

Theorem 3.5

If f has a relative minimum or relative maximum at $x = c$, then c is a critical point of f.

Proof. W.L.O.G., we assume that f is differentiable at c. If $f'(c) > 0$, then there exists $\delta_1 > 0$ such that

$$\Big|\frac{f(x) - f(c)}{x - c} - f'(c)\Big| < \frac{f'(c)}{2} \qquad \text{if} \quad 0 < |x - c| < \delta_1\,;$$

thus

$$\frac{f'(c)}{2} < \frac{f(x) - f(c)}{x - c} < \frac{3f'(c)}{2} \qquad \text{if} \quad 0 < |x - c| < \delta_1\,.$$

1. If $0 < x - c < \delta_1$,

$$f(c) + \frac{f'(c)}{2}(x - c) < f(x) < f(c) + \frac{3f'(c)}{2}(x - c)$$

which implies that f cannot attain a relative maximum at $x = c$ since $f(x) > f(c)$ on the right-hand side of c.

2. if $-\delta < x - c < 0$,

$$f(c) + \frac{f'(c)}{2}(x-c) > f(x) > f(c) + \frac{3f'(c)}{2}(x-c)$$

   which implies that f cannot attain a relative minimum at $x = c$ since $f(c) > f(x)$ on the left-hand side of c.

Therefore, we conclude that if $f'(c) > 0$, then f cannot attain either a relative maximum or minimum at $x = c$. Similar conclusion can be drawn for the case $f'(c) < 0$; thus if f attains a relative extremum at $x = c$, then $f'(c) = 0$. □

Remark 3.6. A more strict version of Theorem 3.5 is called ***Fermat's Theorem*** which is stated as follows:

If f has a local maximum or minimum at c, and if $f'(c)$ exists, then $f'(c) = 0$.

The way to find extrema of a continuous function f on a closed interval $[a, b]$:

1. Find the critical points of f in (a, b).

2. Evaluate f at each critical points in (a, b).

3. Evaluate f at the end-points of $[a, b]$.

4. The least of these values is the minimum, and the greatest is the maximum.

Example 3.7. Find the extrema of $f(x) = 2\sin x - \cos 2x$ on the interval $[0, 2\pi]$.

Since f is differentiable on $(0, 2\pi)$, a critical point c satisfies

$$0 = f'(c) = 2\cos c + 2\sin 2c = 2\cos c(1 + 2\sin c)\,.$$

Therefore, $c = \frac{\pi}{2}$, $c = \frac{3\pi}{2}$, $c = \frac{7\pi}{6}$ or $c = \frac{11\pi}{6}$, and the values of f at these critical points are

$$f\big(\frac{\pi}{2}\big) = 2\cdot 1 - (-1) = 3\,, \qquad f\big(\frac{3\pi}{2}\big) = 2\cdot(-1) - (-1) = -1\,,$$
$$f\big(\frac{7\pi}{6}\big) = 2\cdot\big(-\frac{1}{2}\big) - \frac{1}{2} = -\frac{3}{2}\,, \qquad f\big(\frac{11\pi}{6}\big) = 2\cdot\big(-\frac{1}{2}\big) - \frac{1}{2} = -\frac{3}{2}\,.$$

On the other hand, the values of f at the end-points are

$$f(0) = 2\cdot 0 - 1 = -1 \quad \text{and} \quad f(2\pi) = 2\cdot 0 - 1 = -1\,.$$

Therefore, $f\big(\frac{\pi}{2}\big) = 3$ is the maximum of f on $[0, 2\pi]$, while the minimum of f on $[0, 2\pi]$ occurs at $c = \frac{7\pi}{6}$ and $c = \frac{11\pi}{6}$ and the minimum is $-\frac{3}{2}$.

3.2 Rolle's Theorem and the Mean Value Theorem

Theorem 3.8: Rolle's Theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuous function and f is differentiable on (a, b). If $f(a) = f(b)$, then there is at least one point $c \in (a, b)$ such that $f'(c) = 0$.

Proof. If f is a constant function, then $f'(x) = 0$ for all $x \in (a, b)$. Now suppose that f is not a constant function on $[a, b]$, by the Extreme Value Theorem implies that f has a maximum and a minimum on $[a, b]$, and the maximum and the minimum of f on $[a, b]$ are different. Therefore, there must be a point $c \in (a, b)$ at which f attains its extreme value. By Theorem 3.5, $f'(c) = 0$. ☐

Theorem 3.9: Mean Value Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous and f is differentiable on (a, b), then there exists a point $c \in (a, b)$ such that

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

Proof. Define $g : [a, b] \to \mathbb{R}$ by $g(x) = \big[f(x) - f(a)\big](b - a) - \big[f(b) - f(a)\big](x - a)$. Then $g : [a, b] \to \mathbb{R}$ is continuous and g is differentiable on (a, b). Moreover, $g(a) = g(b) = 0$; thus the Rolle Theorem implies that there exists $c \in (a, b)$ such that $g'(c) = 0$. On the other hand,

$$0 = g'(c) = (b - a)f'(c) - \big[f(b) - f(a)\big];$$

thus there exists $c \in (a, b)$ satisfying $f'(c) = \dfrac{f(b) - f(a)}{b - a}$. ☐

Remark 3.10. In fact, by modifying the proof of the mean value theorem a little bit, we can show the following: Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $g'(x) \neq 0$ for all $x \in (a, b)$, then there exists $c \in (a, b)$ such that

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

The statement above is a generalization of the mean value theorem and is called the Cauchy mean value theorem (see Theorem 5.45).

Example 3.11. Note that the sine function is continuous on any closed interval $[a, b]$ and is differentiable on (a, b). Therefore, the mean value theorem implies that there exists $c \in (a, b)$ such that

$$\cos c = \frac{d}{dx}\Big|_{x=c} \sin x = \frac{\sin b - \sin a}{b - a}$$

which implies that $|\sin a - \sin b| = |\cos c||a - b| \leqslant |a - b|$. Therefore,

$$|\sin x - \sin y| \leqslant |x - y| \qquad \forall\, x, y \in \mathbb{R}\,.$$

Similarly,

$$|\cos x - \cos y| \leqslant |x - y| \qquad \forall\, x, y \in \mathbb{R}\,.$$

3.3 Monotone Functions and the First Derivative Test

Definition 3.12

Let f be defined on an interval I.

1. f is said to be increasing on I if
$$f(x_1) \leqslant f(x_2) \qquad \forall\, x_1, x_2 \in I \text{ and } x_1 < x_2\,.$$
2. f is said to be decreasing on I if
$$f(x_1) \geqslant f(x_2) \qquad \forall\, x_1, x_2 \in I \text{ and } x_1 < x_2\,.$$
3. f is said to be strictly increasing on I if
$$f(x_1) < f(x_2) \qquad \forall\, x_1, x_2 \in I \text{ and } x_1 < x_2\,.$$
4. f is said to be strictly decreasing on I if
$$f(x_1) > f(x_2) \qquad \forall\, x_1, x_2 \in I \text{ and } x_1 < x_2\,.$$

When f is either increasing on I or decreasing on I, then f is said to be monotone. When f is either strictly increasing on I or strictly decreasing on I, then f is said to be strictly monotone on I.

Remark 3.13. Note that f is increasing on I if

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} \geqslant 0 \qquad \forall\, x_1, x_2 \in I \text{ and } x_1 \neq x_2\,.$$

Therefore, f is increasing on I if the slope of each secant line of the graph of f is non-negative. Similar conclusions hold for the other cases.

Example 3.14. The function $f(x) = x^3$ is strictly increasing on $\mathbb{R}$, and $f(x) = -x^3$ is strictly decreasing on $\mathbb{R}$.

Example 3.15. The sine function is strictly increasing on $\left[2n\pi - \frac{\pi}{2}, 2n\pi + \frac{\pi}{2}\right]$ for all $n \in \mathbb{Z}$, but decreasing on $\left[2n\pi - \frac{\pi}{2}, 2n\pi + \frac{3\pi}{2}\right]$ for all $n \in \mathbb{Z}$. However, the sine function is **not** strictly increasing on $\bigcup\limits_{n=-\infty}^{\infty} \left[2n\pi - \frac{\pi}{2}, 2n\pi + \frac{\pi}{2}\right]$ and is **not** strictly decreasing on $\bigcup\limits_{n=-\infty}^{\infty} \left[2n\pi - \frac{\pi}{2}, 2n\pi + \frac{3\pi}{2}\right]$.

Theorem 3.16

Let $f : [a, b] \to \mathbb{R}$ be continuous and f is differentiable on (a, b).

1. If $f'(x) \geqslant 0$ for all $x \in (a, b)$, then f is increasing on $[a, b]$.
2. If $f'(x) \leqslant 0$ for all $x \in (a, b)$, then f is decreasing on $[a, b]$.
3. If $f'(x) > 0$ for all $x \in (a, b)$, then f is strictly increasing on $[a, b]$.
4. If $f'(x) < 0$ for all $x \in (a, b)$, then f is strictly decreasing on $[a, b]$.

Proof. We only prove 1 since all the other conclusion can be proved in a similar fashion.

Suppose that $f'(x) \geqslant 0$, and $x_1 < x_2$. By the Mean Value Theorem, there exists $c \in (x_1, x_2)$ such that

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(c) \geqslant 0\,;$$

thus $f(x_1) \leqslant f(x_2)$ if $x_1 < x_2$. □

Remark 3.17. The condition $f'(x) > 0$ is just a sufficient condition for that f is strictly increasing, but not a necessary condition. For example, $f(x) = x^3$ is strictly increasing on $\mathbb{R}$, but $f'(0) = 0$.

Example 3.18. Show that

$$\cos x \geqslant 1 - \frac{x^2}{2} \qquad \forall\, x \geqslant 0\,. \tag{3.3.1}$$

Let $f(x) = \cos x - 1 + \frac{x^2}{2}$. In order to show (3.3.1), we need to show that $f(x) \geqslant 0$ for all $x \geqslant 0$. Since $f'(x) = -\sin x + x$, by Theorem 0.13 we find that f' is non-negative on $[0, \infty)$. Therefore, Theorem 3.16 implies that f is increasing on $[0, \infty)$ which further shows that $f(x) \geqslant f(0) = 0$ for all $x \geqslant 0$.

Example 3.19. Using (3.3.1), we can show that

$$\sin x \geqslant x - \frac{x^3}{6} \qquad \forall\, x \geqslant 0\,.$$

In fact, by defining $g(x) = \sin x - x + \dfrac{x^3}{6}$, using (3.3.1) we find that

$$g'(x) = \cos x - 1 + \frac{x^2}{2} \geqslant 0 \qquad \forall\, x \geqslant 0\,;$$

thus g is increasing on $[0, \infty)$ which shows that $g(x) \geqslant g(0) = 0$ for all $x \geqslant 0$. Similar argument then shows that

$$\cos x \leqslant 1 - \frac{x^2}{2} + \frac{x^4}{24} \qquad \forall\, x \geqslant 0$$

and the inequality above in turn implies that

$$\sin x \leqslant x - \frac{x^3}{6} + \frac{x^5}{120} \qquad \forall\, x \geqslant 0\,.$$

By induction, we can show that for all $k \in \mathbb{N} \cup \{0\}$,

$$x - \frac{x^3}{3!} + \cdots + \frac{x^{4k+1}}{(4k+1)!} - \frac{x^{4k+3}}{(4k+3)!} \leqslant \sin x \leqslant x - \frac{x^3}{3!} + \cdots + \frac{x^{4k+1}}{(4k+1)!} \qquad \forall\, x \geqslant 0\,,$$

$$1 - \frac{x^2}{2!} + \cdots + \frac{x^{4k}}{(4k)!} - \frac{x^{4k+2}}{(4k+2)!} \leqslant \cos x \leqslant 1 - \frac{x^2}{2} + \cdots + \frac{x^{4k}}{(4k)!} \qquad \forall\, x \geqslant 0\,.$$

Theorem 3.20: The First Derivative Test

Let f be a continuous function defined on an open interval I containing c. If f is differentiable on I, except possibly at c, then

1. If f' changes from negative to positive at c, then $f(c)$ is a local minimum of f.
2. If f' changes from positive to negative at c, then $f(c)$ is a local maximum of f.
3. If f' is sign definite on $I\backslash\{c\}$, then $f(c)$ is neither a relative minimum or relative maximum of f.

Proof. We only prove 1. Assume that f' changes from negative to positive at c. Then there exists a and b in I such that

$$f'(x) < 0 \text{ for all } x \in (a, c) \quad \text{and} \quad f'(x) > 0 \text{ for all } x \in (c, b)\,.$$

By Theorem 3.16, f is decreasing on (a, c) and is increasing on (c, b). Therefore, $f(c)$ is a minimum on an open interval (a, b); thus is a relative minimum on I. $\square$

Example 3.21. Find the relative extrema of $f(x) = \frac{1}{2}x - \sin x$ in the interval $(0, 2\pi)$.

By Theorem 3.5 the relative extrema occurs at critical points. Since f is differentiable on $(0, 2\pi)$, a critical point x satisfies

$$0 = f'(x) = \frac{1}{2} - \cos x$$

which implies that $c = \frac{\pi}{3}$ and $c = \frac{5\pi}{3}$ are the only critical points. To determine if $f\left(\frac{\pi}{3}\right)$ or $f\left(\frac{5\pi}{3}\right)$ is a relative minimum, we apply Theorem 3.20 and found that, since f' changes from negative to positive at $\frac{\pi}{3}$ and changes from positive to negative at $\frac{5\pi}{3}$, $f\left(\frac{\pi}{3}\right)$ is a relative minimum of f on $(0, 2\pi)$.

Remark 3.22. When a differentiable function f attains a local minimum at an interior point c, it is not necessary that f' changes from positive to negative. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} x^2\left(1 + \sin\frac{1}{x}\right) & \text{if } x \neq 0\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

Then

$$f'(x) = \begin{cases} 2x\left(1 + \sin\frac{1}{x}\right) - \cos\frac{1}{x} & \text{if } x \neq 0\,, \\ 0 & \text{if } x = 0\,. \end{cases}$$

Therefore,

1. 0 is a critical point of f.
2. f attains a (global) minimum at 0 since obviously $f(x) \geqslant 0 = f(0)$ for all $x \in \mathbb{R}$.
3. It is impossible to determine if f' changes "from negative to positive" or "from positive to negative" at 0.

3.4 Concavity（凹性）and the Second Derivative Test

Definition 3.23

Let f be differentiable on an open interval I. The graph of f is concave upward（凹向上）on I if f' is strictly increasing on the interval and concave downward（凹向下）on I if f' is strictly decreasing on the interval.

Remark 3.24. It does not really matter if f' has to be strictly monotone, instead of just monotone, in order to define the concavity of the graph of f. Here we define the concavity by the strict monotonicity.

- Graphical interpretation of concavity: Let f be differentiable on an open interval I.

  1. If the graph of f is concave upward on I, then the graph of f lies above all of its tangent lines on I.

  2. If the graph of f is concave downward on I, then the graph of f lies below all of its tangent lines on I.

The following theorem is a direct consequence of Theorem 3.16.

Theorem 3.25: Test for Concavity

Let f be a twice differentiable function on an open interval I.

1. If $f''(x) > 0$ for all x in I, then the graph of f is concave upward on I.
2. If $f''(x) < 0$ for all x in I, then the graph of f is concave downward on I.

Example 3.26. Determine the open intervals on which the graph of $f(x) = \dfrac{6}{x^2+3}$ is concave upward or concave downward.

First we compute the second derivative of f:

$$f'(x) = \frac{-12x}{(x^2+3)^2} \Rightarrow f''(x) = -12\frac{(x^2+3)^2 - 2(x^2+3)(2x)x}{(x^2+3)^4} = \frac{36(x^2-1)}{(x^2+3)^3}.$$

Therefore, the graph of f is concave upward if $x > 1$ and is concave downward if $x < 1$.

Definition 3.27: Point of inflection（反曲點）

Let f be a differentiable function on an open interval containing c. The point $\big(c, f(c)\big)$ is called a point of inflection (or simply an inflection point) of the graph of f if the concavity of f changes from upward to downward or downward to upward at this point.

Example 3.28. Recall Example 3.26 $\Big(f(x) = \dfrac{6}{x^2+3}$ with $f''(x) = \dfrac{36(x^2-1)}{(x^2+3)^3}\Big)$. Since f'' changes sign at $x = \pm 1$, $\Big(\pm 1, \dfrac{3}{2}\Big)$ are both points of inflection of the graph of f.

Theorem 3.29

Let f be a differentiable function on an open interval containing c. If $\big(c, f(c)\big)$ is a point of inflection of the graph of f, then either $f''(c) = 0$ or $f''(c)$ does not exist.

Remark 3.30. A point $\big(c, f(c)\big)$ may not be an inflection point of the graph of f even if $f''(c) = 0$. For example, the point $(0,0)$ is not an inflection point of $f(x) = x^4$ since $f''(x) > 0$ for all $x \neq 0$ which implies that the concavity of f does not change at $c = 0$.

Example 3.31. Determine the points of inflection and discuss the concavity of the graph of $f(x) = x^4 - 4x^3$. Note that the zero of f'' is $x = 0$ or $x = 2$ (since $f''(x) = 12x^2 - 24x$). Since $f''(x) > 0$ if $x < 0$ or $x > 2$, and $f''(x) > 0$ if $0 < x < 2$, we find that $(0,0)$ and $(2,-16)$ are points of inflection of the graph of f.

Theorem 3.32

Let f be a twice differentiable function on an open interval I containing c, and c is a critical point of f.

1. If $f''(c) > 0$, then $f(c)$ is a relative minimum of f on I.
2. If $f''(c) < 0$, then $f(c)$ is a relative maximum of f on I.

Remark 3.33. If $f''(c) = 0$ for some critical point c of f, then f may have a relative maximum, a relative minimum, or neither at c. In such cases, you should use the First Derivative Test.

Proof of Theorem 3.32. Since $f''(c) > 0$, there exist $\delta > 0$ such that

$$\Big|\frac{f'(x) - f'(c)}{x - c} - f''(c)\Big| < \frac{f''(c)}{2} \quad \text{if } 0 < |x - c| < \delta.$$

Since c is a critical point of f, $f'(c) = 0$; thus the inequality above implies that

$$\frac{f''(c)}{2} < \frac{f'(x)}{x - c} < \frac{3f''(c)}{2} \quad \text{if } 0 < |x - c| < \delta.$$

In particular,

$$0 < \frac{f'(c)}{2}(x - c) < f'(x) \qquad \text{if } 0 < x - c < \delta\,,$$
$$f'(x) < \frac{f'(c)}{2}(x - c) < 0 \qquad \text{if } -\delta < x - c < 0\,.$$

Therefore, f' changes from negative to positive at c; thus $f(c)$ is a relative minimum of f on I. $\square$

Example 3.34. Recall Example 3.21 $\left(f(x) = \frac{1}{2}x - \sin x\right)$. We have established that $f\left(\frac{\pi}{3}\right)$ is a relative minimum of f on $(0, 2\pi)$ using the First Derivative Test. Note that $f''(x) = \sin x$; thus $f''\left(\frac{\pi}{3}\right) = \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2} > 0$. Therefore, without using the First Derivative Test, we can still conclude that $f\left(\frac{\pi}{3}\right)$ is a relative minimum of f on $(0, 2\pi)$ by the second derivative test.

Example 3.35. Show that for all $1 < p, q < \infty$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q} \qquad \forall\, a, b > 0\,. \tag{3.4.1}$$

The inequality above is called Young's inequality. We remark that if $\frac{1}{p} + \frac{1}{q} = 1$, then $q = \frac{p}{p-1}$.

For the moment we only show (3.4.1) for the case that $p, q \in \mathbb{Q}$ (because we have not talked about what it means by the p-th power if p is irrational). To show (3.4.1), we prove that for each given $b > 0$, the function $f : (0, \infty) \to \mathbb{R}$

$$f(x) \equiv \frac{x^p}{p} - bx + \frac{b^q}{q}$$

is non-negative. In other words, we have to show that the "minimum" of f is non-negative.

To find the minimum of f, we differentiate and find that $f'(x) = x^{p-1} - b$ which implies that $c = b^{\frac{1}{p-1}}$ is the only critical point. Since

$$f''(c) = (p-1)c^{p-2} = (p-1)b^{\frac{p-2}{p-1}} > 0\,,$$

the second derivative test implies that f attains a local minimum at c. Since there is no other critical points, f must attain its global minimum at c; thus

$$f(x) \geqslant f(c) \qquad \forall\, x \in (0, \infty)$$

and (3.4.1) is established since $f(c) = \frac{b^{\frac{p}{p-1}}}{p} - b^{1+\frac{1}{p-1}} + \frac{b^q}{q} = \frac{b^q}{p} - b^q + \frac{b^q}{q} = 0$.

Remark 3.36. Suppose that c is a critical point of a differentiable function f with $f''(c) = 0$. For f to attain a local extremum at c, $f'''(c)$ must be zero if the third derivative of f is continuous. If in addition $f^{(4)}$ is continuous, then

1. f attains a local maximum at c provided that $f^{(4)}(c) < 0$.

2. f attains a local minimum at c provided that $f^{(4)}(c) > 0$.

In general, if f is $2k$-times continuously differentiable (which means $f^{(2k)}$ exists everywhere and is continuous) and $f'(c) = f''(c) = \cdots = f^{(2k-1)}(c) = 0$, then

1. f attains a local maximum at c provided that $f^{(2k)}(c) < 0$.

2. f attains a local minimum at c provided that $f^{(2k)}(c) > 0$.

On the other hand, if f is $(2k+1)$-times continuously differentiable and $f'(c) = f''(c) = \cdots = f^{(2k)}(c) = 0$ but $f^{(2k+1)}(c) \neq 0$, then f cannot attain its local extremum at c.

3.5 A Summary of Curve Sketching

When sketching the graph of functions, you need to have the following on the plot.

1. x-intercepts and y-intercepts;

2. asymptotes;

3. absolution extrema and relative extrema;

4. points of inflection.

Example 3.37. Sketch the graph of the function $f(x) = \dfrac{3x-2}{\sqrt{2x^2+1}}$.

First, we note that the x-intercepts and y-intercepts are $\left(\frac{3}{2}, 0\right)$ and $(0, f(0)) = (0, -2)$. To determine the asymptotes, since $\sqrt{2x^2+1}$ are never zero, there is no vertical asymptote. As for the horizontal and slant asymptotes, by the fact that

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{\dfrac{3x-2}{x}}{\dfrac{\sqrt{2x^2+1}}{x}} = \lim_{x\to\infty} \frac{3-\dfrac{2}{x}}{\sqrt{2+\dfrac{1}{x^2}}} = \lim_{y\to 0^+} \frac{3-2y}{\sqrt{2+y^2}} = \frac{3}{\sqrt{2}}$$

and

$$\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(-x) = \lim_{x\to\infty} \frac{-3x-2}{\sqrt{2x^2+1}} = \lim_{x\to\infty} \frac{-3-\dfrac{2}{x}}{\sqrt{2+\dfrac{1}{x^2}}} = \lim_{y\to 0^+} \frac{3-2y}{-\sqrt{2+y^2}} = -\frac{3}{\sqrt{2}},$$

we find that there are two horizontal asymptotes $y = \pm\dfrac{3}{\sqrt{2}}$.

By the quotient rule,

$$\begin{aligned} f'(x) &= \frac{3\sqrt{2x^2+1} - (3x-2)\dfrac{d}{dx}(2x^2+1)^{\frac{1}{2}}}{2x^2+1} = \frac{3\sqrt{2x^2+1} - (3x-2)\dfrac{1}{2}(2x^2+1)^{-\frac{1}{2}} \cdot (4x)}{2x^2+1} \\ &= \frac{3(2x^2+1) - 2x(3x-2)}{(2x^2+1)^{\frac{3}{2}}} = \frac{4x+3}{(2x^2+1)^{\frac{3}{2}}} \end{aligned}$$

and

$$\begin{aligned} f''(x) &= \frac{4(2x^2+1)^{\frac{3}{2}} - (4x+3)\dfrac{3}{2}(2x^2+1)^{\frac{1}{2}} \cdot (4x)}{(2x^2+1)^3} = \frac{4(2x^2+1) - 6x(4x+3)}{(2x^2+1)^{\frac{5}{2}}} \\ &= \frac{-16x^2 - 18x + 4}{(2x^2+1)^{\frac{5}{2}}} = \frac{-2(8x^2+9x-2)}{(2x^2+1)^{\frac{5}{2}}}. \end{aligned}$$

Therefore, $x = -\dfrac{3}{4}$ is the only critical point and since f' changes from negative to positive at $-\dfrac{3}{4}$, $f\big(-\dfrac{3}{4}\big)$ is a relative minimum of f.

$f''(x) = 0$ occurs at $x_1 = \dfrac{-9-\sqrt{145}}{16}$ and $x_2 = \dfrac{-9+\sqrt{145}}{16}$. Since f'' changes sign at x_1 and x_2, $(x_1, f(x_1))$ and $(x_2, f(x_2))$ are inflection points of the graph of f.

3.6 Optimization Problems

Explanation of examples in Section 3.7 in the textbook:

1. 一製造商想設計一個底部為正方形、表面積 108 平方公分且上方有開口的箱子。要怎麼設計才能讓箱子容積最大？

2. Which points on the graph of $y = 4 - x^2$ are closest to the point $(0, 2)$? 拋物線 $y = 4 - x^2$ 上哪個點到 $(0, 2)$ 最近？

3. 試找出最小面積的方形頁面使之能上下留白三公分、左右留白兩公分且要包含 216 平方公分的長方形區域可用於印刷。

4. 兩根分別為 12 公尺及 28 公尺高的桿子相距 30 公尺。找出地面上一點使之到兩桿之頂端之距離和最小。

5. Four meters of wire is to be used to form a square and a circle. How much of the wire should be used for the square and how much should be used for the circle to enclose the maximum total area? 一總長 4 公尺的線要被分為兩段用來圍出一個正方形和一個圓形。要怎麼分段才能圍出最大的面積。

6. **Application in Physics**: Let v_1 be the velocity of light in air and v_2 the velocity of light in water. According to Fermat's Principle, a ray of light will travel from a point A in the air to a point B in the water by a path ACB that minimizes the time taken. Show that
$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{v_1}{v_2}\,,$$
where θ_1 (the angle of incidence) and θ_2 (the angle of refraction) are as shown. This equation is known as Snell's Law.

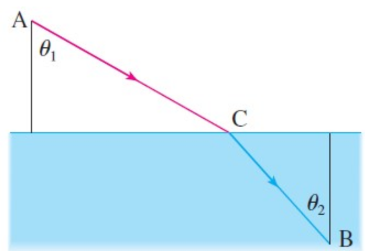


Figure 3.1: Snell's law

Proof. Assume that $A = (0, a)$ and $B = (b, -c)$. The goal is to find $C = (x, 0)$ so that
$$f(x) = \frac{\sqrt{x^2 + a^2}}{v_1} + \frac{\sqrt{(x-b)^2 + c^2}}{v_2}$$
is minimized. Differentiating f, we find that a critical point x of f satisfies
$$\frac{1}{v_1}\frac{x}{\sqrt{x^2+a^2}} = \frac{1}{v_2}\frac{b-x}{\sqrt{(x-b)^2+c^2}}\,.$$
Snell's law then is concluded from the fact that $\sin\theta_1 = \dfrac{x}{\sqrt{x^2+a^2}}$ and $\sin\theta_2 = \dfrac{b-x}{\sqrt{(b-x)^2+c^2}}$. $\square$

7. **Application in Economics**: Suppose that

$$\begin{aligned}
r(x) &= \text{the revenue from selling } x \text{ items,}\\
c(x) &= \text{the cost of producing the } x \text{ items,}\\
p(x) &= r(x) - c(x) = \text{the profit from producing and selling } x \text{ items.}
\end{aligned}$$

Although x is usually an integer in many applications, we can learn about the behavior of these functions by defining them for all nonzero real numbers and by assuming they are differentiable functions. Economists use the terms *marginal revenue* (邊際收益), *marginal cost* (邊際成本), and *marginal profit* (邊際利潤) to name the derivatives $r'(x)$, $c'(x)$, and $p'(x)$ of the revenue, cost, and profit functions. Let us consider the relationship of the profit p to these derivatives. If $r(x)$ and $c(x)$ are differentiable for x in some interval of production possibilities, and if $p(x) = r(x) - c(x)$ has a maximum value there, it occurs at a critical point of $p(x)$ or at an end-point of the interval. If it occurs at a critical point, then $p'(x) = r'(x) - c'(x) = 0$ and we see that $r'(x) = c'(x)$. In economic terms, this last equation means that

> At a production level yielding maximum profit, marginal revenue equals marginal cost.

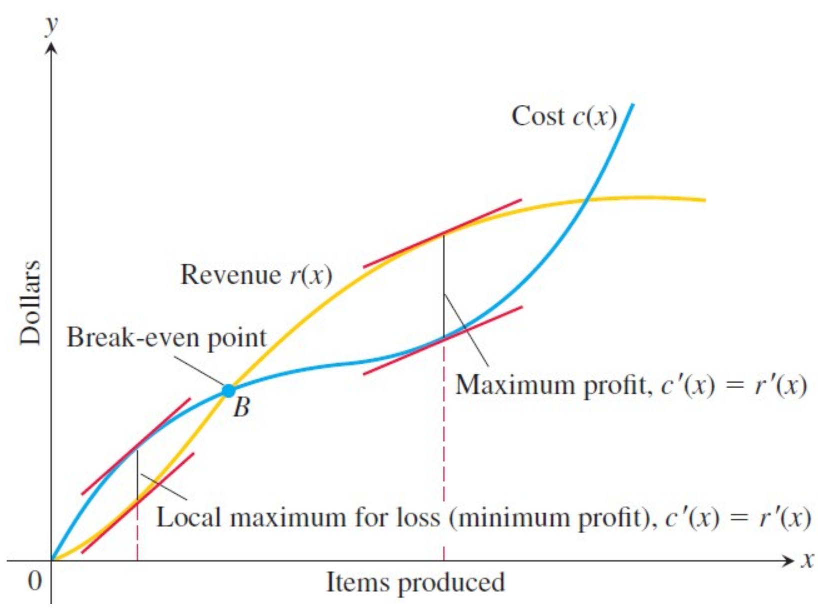


Figure 3.2: The graph of a typical cost function starts concave down and later turns concave up. It crosses the revenue curve at the break-even point B. To the left of B, the company operates at a loss. To the right, the company operates at a profit, with the maximum profit occurring where $c'(x) = r'(x)$. Farther to the right, cost exceeds revenue (perhaps because of a combination of rising labor and material costs and market saturation) and production levels become unprofitable again.

3.7 Newton's Method

The Newton method is a numerical method for finding zeros of differentiable functions. Let $f : (a, b) \to \mathbb{R}$ be a differentiable function, and $c \in (a, b)$ is a zero of f. To find an approximated value of c, the Newton method is the following iterative scheme:

1. Make an initial estimate $x_1 \in (a, b)$ that is close to c.
2. Determine a new approximation using the iterative relation:
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}\,.$$

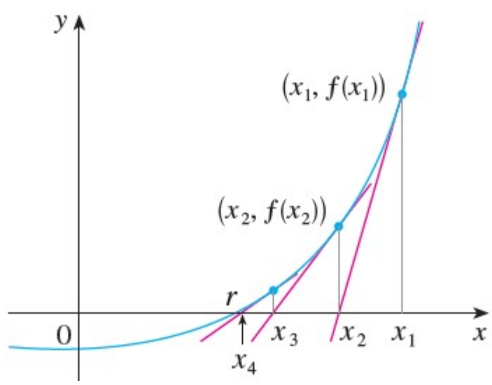


Figure 3.3: Sequence of approximated zeros by Newton's method

3. When $|x_n - x_{n+1}|$ is within the desired accuracy, let x_{n+1} serve as the final approximation.

Example 3.38. To find the square root of a positive number A is equivalent to finding zeros of the function $f(x) = x^2 - A$ in $(0, \infty)$. The Newton method provides the iterative scheme
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - A}{2x_n} = \frac{x_n}{2} + \frac{A}{2x_n}$$
to find approximated value of $\sqrt{A}$.

Example 3.39. To find the precise value of π we can look for a zero of the function $f(x) = \cos\frac{x}{2}$. The Newton method provides the iterative scheme
$$x_{n+1} = x_n - \frac{\cos\frac{x_n}{2}}{-\frac{1}{2}\sin\frac{x_n}{2}} = x_n + 2\cot\frac{x_n}{2}$$

to find the value of zeros of f. Starting the iteration with $x_1 = 3$, then $x_2 \approx 3.141829688605305$, $x_3 \approx 3.141592653588683$ and $x_4 \approx 3.141592653589793$. We note that x_4 has already been very close to π.

It can be shown that when $\left|\dfrac{f(x)f''(x)}{f'(x)^2}\right| < 1$ for all $x \in (a,b)$, then the Newton method produces a convergent sequence which approaches a zero in (a,b).

3.8 Exercise

Problem 3.1. 1. Let $f, g : (a,b) \to \mathbb{R}$ be functions and $f'(x) = g'(x)$. Show that there exists a constant C such that $f(x) = g(x) + C$.

2. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a differentiable function satisfying that $f'(x) = 3x^2 + 4\cos x$ and $f(0) = 0$. Find $f(x)$.

Problem 3.2. Let $f : [a,b] \to \mathbb{R}$ be a continuous function such that f has only one critical point $c \in (a,b)$.

1. Show that if $f(c)$ is a local extremum of f, then $f(c)$ is an absolute extremum of f.

2. Show that if $f(c)$ is the absolute minimum of f, then $f(x) > f(c)$ for all $x \in [a,b]$ and $x \neq c$. Similarly, show that if $f(c)$ is the absolute maximum of f, then $f(x) < f(c)$ for all $x \in [a,b]$ and $x \neq c$.

Problem 3.3. Let I, J be intervals, $g : I \to \mathbb{R}$ and $f : J \to \mathbb{R}$ be increasing functions. Show that if J contains the range of g, then $f \circ g$ is increasing on I.

Problem 3.4. 1. If the function $f(x) = x^3 + ax^2 + bx$ has the local minimum value $-\dfrac{2\sqrt{3}}{9}$ at $x = \dfrac{1}{\sqrt{3}}$, what are the values of a and b?

2. Which of the tangent lines to the curve in part (1) has the smallest slope?

Problem 3.5. A number a is called a fixed point of a function f if $f(a) = a$. Prove that if $f'(x) \neq 1$ for all real numbers x, then f has at most one fixed point.

Problem 3.6. Suppose f is an odd function (that is, $f(-x) = -f(x)$ for all $x \in \mathbb{R}$) and is differentiable everywhere. Prove that for every positive number b, there exists a number c in $(-b, b)$ such that $f'(c) = \dfrac{f(b)}{b}$.

Problem 3.7. Show that $2\sqrt{x} > 3 - \dfrac{1}{x}$ for all $x > 1$.

Problem 3.8. Show that $\sqrt{b} - \sqrt{a} < \dfrac{b-a}{2\sqrt{a}}$ for all $0 < a < b$.

Problem 3.9. Show that for all (rational numbers) $p, q \in (1, \infty)$ satisfying $\dfrac{1}{p} + \dfrac{1}{q} = 1$, we have

$$ac + bd \leqslant (a^p + b^p)^{\frac{1}{p}}(c^q + d^q)^{\frac{1}{q}} \qquad \forall\, a, b, c, d > 0\,.$$

Hint: Let $x = \dfrac{a}{b}$ and $y = \dfrac{d}{c}$.

Problem 3.10. Show that for all $k \in \mathbb{N} \cup \{0\}$,

$$x - \frac{x^3}{3!} + \cdots + \frac{x^{4k+1}}{(4k+1)!} - \frac{x^{4k+3}}{(4k+3)!} \leqslant \sin x \leqslant x - \frac{x^3}{3!} + \cdots + \frac{x^{4k+1}}{(4k+1)!} \qquad \forall\, x \geqslant 0\,,$$
$$1 - \frac{x^2}{2!} + \cdots + \frac{x^{4k}}{(4k)!} - \frac{x^{4k+2}}{(4k+2)!} \leqslant \cos x \leqslant 1 - \frac{x^2}{2} + \cdots + \frac{x^{4k}}{(4k)!} \qquad \forall\, x \geqslant 0\,.$$

Problem 3.11. （不要用交叉相乘）Show that for all $k \in \mathbb{N} \cup \{0\}$,

$$1 - x + x^2 - x^3 + \cdots + x^{2k} - x^{2k+1} \leqslant \frac{1}{1+x} \leqslant 1 - x + x^2 - x^3 + \cdots + x^{2k} \qquad \forall\, x \geqslant 0\,.$$

Problem 3.12. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function satisfying that $f'(x) = f(x)$ for all $x \in \mathbb{R}$, and $f(0) = 1$.

1. （不要試著找出 f 而是直接用 f 的性質）Show that f is increasing on $\mathbb{R}$.
2. Show that if $k \in \mathbb{N} \cup \{0\}$, then $f(x) \geqslant 1 + x + \dfrac{x^2}{2!} + \cdots + \dfrac{x^k}{k!}$ for all $x \geqslant 0$.
3. Show that if $k \in \mathbb{N} \cup \{0\}$, then

$$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{2k}}{(2k)!} + \frac{x^{2k+1}}{(2k+1)!} \leqslant f(x) \leqslant 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{2k}}{(2k)!} \qquad \forall\, x \leqslant 0\,.$$

Hint: 1. Show that f^2 is increasing on $\mathbb{R}$ and argue that f is also increasing on $\mathbb{R}$.

Problem 3.13. 1. The function

$$f(x) = \begin{cases} 0 & \text{if } x = 0\,, \\ 1 - x & \text{if } 0 < x \leqslant 1 \end{cases}$$

is differentiable on $(0, 1)$ and satisfies $f(0) = f(1)$. However, its derivative is never zero on $(0, 1)$. Does this contradict Rolle's Theorem? Explain.

2. Can you find a function f such that $f(-2) = -2$, $f(2) = 6$, and $f'(x) < 1$ for all x? Why or why not?

Problem 3.14. Find the minimum value of

$$\left| \sin x + \cos x + \tan x + \cot x + \sec x + \csc x \right|$$

for real numbers x.
Hint: Let $t = \sin x + \cos x$.

Problem 3.15. Let $f, g : (a, b) \to \mathbb{R}$ be twice differentiable functions such that $f''(x) \neq 0$ and $g''(x) \neq 0$ for all $x \in (a, b)$. Prove that if f and g are positive, increasing, and concave upward on the interval (a, b), then fg is also concave upward on (a, b).

Problem 3.16. For what values of a and b is $(2, 2.5)$ an inflection point of the curve $x^2 + ax + by = 0$? What additional inflection points does the curve have?

Chapter 4

Integration

- **The Σ notation**: The sum of n-terms $a_1, a_2, \cdots, a_n$ is written as $\sum_{i=1}^{n} a_i$. In other words,

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n \,.$$

Here i is called the index of summation, a_i is the i-th terms of the sum. We note that i in the sum $\sum_{i=1}^{n} a_i$ is a dummy index which can be replaced by other indices such as j, k, and etc. Therefore, $\sum_{i=1}^{n} a_i = \sum_{j=1}^{n} a_j = \sum_{k=1}^{n} a_k$, and so on.

- **Basic properties of sums**: $\displaystyle\sum_{i=1}^{n}(ca_i + b_i) = c\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i$.

Theorem 4.1: Summation Formula

1. $\sum_{i=1}^{n} c = cn$ if c is a constant;
2. $\sum_{i=1}^{n} i = \dfrac{n(n+1)}{2}$;
3. $\sum_{i=1}^{n} i^2 = \dfrac{n(n+1)(2n+1)}{6}$;
4. $\sum_{i=1}^{n} i^3 = \dfrac{n^2(n+1)^2}{4}$.

4.1 The Area under the Graph of a Non-negative Continuous Function

Let $f : [a, b] \to \mathbb{R}$ be a non-negative continuous function, and R be the region enclosed by the graph of the function f, the x-axis and straight lines $x = a$ and $x = b$. We consider computing $\mathcal{A}(\mathrm{R})$, the area of R. Generally speaking, since the graph of $y = f(x)$ is in general not a straight line, the computation of $\mathcal{A}(\mathrm{R})$ is not straight-forward. How do we compute the area $\mathcal{A}(\mathrm{R})$?

Partition $[a,b]$ into n sub-intervals with equal length, and let $\Delta x = \dfrac{b-a}{n}$, $x_i = a + i\Delta x$. By the Extreme Value Theorem, for each $1 \leqslant i \leqslant n$ f attains its maximum and minimum on $[x_{i-1}, x_i]$; thus for $1 \leqslant i \leqslant n$, there exist $M_i, m_i \in [x_{i-1}, x_i]$ such that

$$f(M_i) = \text{the maximum of } f \text{ on } [x_{i-1}, x_i]$$

and

$$f(m_i) = \text{the minimum of } f \text{ on } [x_{i-1}, x_i].$$

The sum $S(n) \equiv \sum\limits_{i=1}^{n} f(M_i)\Delta x$ is called the upper sum of f for the partition $\{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$, and $s(n) \equiv \sum\limits_{i=1}^{n} f(m_i)\Delta x$ is called the lower sum of f for the partition $\{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$. By the definition of the upper sum and lower sum, we find that for each $n \in \mathbb{N}$,

$$\sum_{i=1}^{n} f(m_i)\Delta x \leqslant \mathcal{A}(\mathrm{R}) \leqslant \sum_{i=1}^{n} f(M_i)\Delta x\,.$$

If the limits of the both sides exist and are identical as Δx approaches 0 (which is the same as n approaches infinity), by the Squeeze Theorem we can conclude that $\mathcal{A}(\mathrm{R})$ is the same as the limit.

Example 4.2. Let $f(x) = x^2$, and R be the region enclosed by the graph of $y = f(x)$, the X-axis, and the straight lines $x = a$ and $x = b$, where we assume that $0 \leqslant a < b$. Then the lower sum is obtained by the "left end-point rule" approximation of $\mathcal{A}(\mathrm{R})$

$$\sum_{i=1}^{n} \Big(a + \frac{(i-1)(b-a)}{n}\Big)^2 \frac{b-a}{n}$$

and the upper sum is obtained by the "right end-point rule" approximation

$$\sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n}\,.$$

By Theorem 4.1,

$$\begin{aligned}
\sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n} &= \sum_{i=1}^{n} \Big[a^2 + \frac{2a(b-a)i}{n} + \frac{a^2(b-a)^2 i^2}{n^2}\Big] \frac{b-a}{n} \\
&= a^2(b-a) + \frac{a(b-a)^2 n(n+1)}{n^2} + \frac{a^2(b-a)^3}{n^3}\frac{n(n+1)(2n+1)}{6} \\
&= a^2(b-a) + a(b-a)^2\Big(1 + \frac{1}{n}\Big) + \frac{a^2(b-a)^3}{6}\Big(1 + \frac{1}{n}\Big)\Big(2 + \frac{1}{n}\Big)\,.
\end{aligned}$$

Letting $n \to \infty$, we find that

$$\lim_{n\to\infty} \sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n} = a^2(b-a) + a(b-a)^2 + \frac{a^2(b-a)^3}{3} = \frac{b^3-a^3}{3}\,.$$

Similarly,

$$\begin{aligned}\sum_{i=1}^{n} \Big(a + \frac{(i-1)(b-a)}{n}\Big)^2 \frac{b-a}{n} &= \frac{a^2(b-a)}{n} + \sum_{i=1}^{n} \Big(a + \frac{i(b-a)}{n}\Big)^2 \frac{b-a}{n} - \frac{b^2(b-a)}{n} \\ &= a^2(b-a) + \frac{a(b-a)^2 n(n+1)}{n^2} + \frac{a^2(b-a)^3}{n^3}\frac{n(n+1)(2n+1)}{6} + \frac{(a^2-b^2)(b-a)}{n}\,;\end{aligned}$$

thus

$$\lim_{n\to\infty} \sum_{i=1}^{n} \Big(a + \frac{(i-1)(b-a)}{n}\Big)^2 \frac{b-a}{n} = \frac{b^3-a^3}{3}\,.$$

Therefore, $\mathcal{A}(\mathrm{R}) = \dfrac{b^3-a^3}{3}$.

Remark 4.3. Let R_1 be the region enclosed by $f(x) = x^2$, the x-axis and $x = a$, the R_2 be the region enclosed by $f(x) = x^2$, the x-axis and $x = b$, then intuitively $\mathcal{A}(\mathrm{R}) = \mathcal{A}(\mathrm{R}_2) - \mathcal{A}(\mathrm{R}_1)$ and this is true since $\mathcal{A}(\mathrm{R}_1) = \dfrac{a^3}{3}$ and $\mathcal{A}(\mathrm{R}_2) = \dfrac{b^3}{3}$.

If f is not continuous, then f might not attain its extrema on the interval $[x_{i-1}, x_i]$. In this case, it might be impossible to form the upper sum or the lower sum for a given partition. On the other hand, the left end-point rule $\sum\limits_{i=1}^{n} f(x_{i-1})\Delta x$ and the right end-point rule $\sum\limits_{i=1}^{n} f(x_i)\Delta x$ of approximating the area are always possible. We can even consider the "mid-point rule" approximation given by

$$\sum_{i=1}^{n} f\big(\frac{x_{i-1}+x_i}{2}\big)\Delta x$$

and consider the limit of the expression above as n approaches infinity.

4.2 Riemann Sums and Definite Integrals

In general, in order to find an approximation of $\mathcal{A}(\mathrm{R})$, the interval $[a, b]$ does not have to be divided into sub-intervals with equal length. Assume that $[a, b]$ are divided into n sub-intervals and the end-points of those sub-intervals are ordered as $a = x_0 < x_1 < x_2 < \cdots <$

$x_n = b$, here the collection of end-points $\mathcal{P} = \{x_0, x_1, \cdots, x_n\}$ is called a ***partition*** of $[a, b]$. Then the "left end-point rule" approximation for the partition $\mathcal{P}$ is given by

$$\ell(\mathcal{P}) = \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1})$$

and the "right end-point rule" approximation for the partition $\mathcal{P}$ is given by

$$r(\mathcal{P}) = \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}),$$

and the limit process as $n \to \infty$ in the discussion above is replaced by the limit process as the norm of partition $\mathcal{P}$, denoted by $\|\mathcal{P}\|$ and defined by $\|\mathcal{P}\| \equiv \max\left\{x_i - x_{i-1} \,\middle|\, 1 \leqslant i \leqslant n\right\}$, approaches 0. Before discussing what the limits above mean, let us look at the following examples.

Example 4.4. Consider the region bounded by the graph of $f(x) = \sqrt{x}$ and the x-axis for $0 \leqslant x \leqslant 1$. Let $x_i = \dfrac{i^2}{n^2}$ and $\mathcal{P} = \{x_0 = 0 < x_1 < \cdots < x_n = 1\}$. We note that

$$\|\mathcal{P}\| = \max\left\{\frac{i^2 - (i-1)^2}{n^2} \,\middle|\, 1 \leqslant i \leqslant n\right\} = \max\left\{\frac{2i-1}{n^2} \,\middle|\, 1 \leqslant i \leqslant n\right\} = \frac{2n-1}{n^2}$$

thus $\|\mathcal{P}\| \to 0$ is equivalent to that $n \to \infty$.

Using the right end-point rule (which is the same as the upper sum),

$$\begin{aligned} S(\mathcal{P}) &= \sum_{i=1}^{n} f(x_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} \frac{i}{n}\frac{2i-1}{n^2} = \frac{1}{n^3}\sum_{i=1}^{n}(2i^2 - i) \\ &= \frac{1}{n^3}\Big[\frac{n(n+1)(2n+1)}{3} - \frac{n(n+1)}{2}\Big] \\ &= \frac{1}{3}\Big(1+\frac{1}{n}\Big)\Big(2+\frac{1}{n}\Big) - \frac{1}{2n}\Big(1+\frac{1}{n}\Big); \end{aligned}$$

thus

$$\lim_{\|\mathcal{P}\|\to 0} S(\mathcal{P}) = \lim_{n\to\infty}\Big[\frac{1}{3}\Big(1+\frac{1}{n}\Big)\Big(2+\frac{1}{n}\Big) - \frac{1}{2n}\Big(1+\frac{1}{n}\Big)\Big] = \frac{2}{3}.$$

Using the left end-point rule (which is the same as the lower sum),

$$\begin{aligned} s(\mathcal{P}) &= \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1}) = \sum_{i=1}^{n} \frac{i-1}{n}\frac{2i-1}{n^2} = \frac{1}{n^3}\sum_{i=1}^{n}(2i^2 - 3i + 1) \\ &= \frac{1}{n^3}\Big[\frac{n(n+1)(2n+1)}{3} - \frac{3n(n+1)}{2} + n\Big] \\ &= \frac{1}{3}\Big(1+\frac{1}{n}\Big)\Big(2+\frac{1}{n}\Big) - \frac{3}{2n}\Big(1+\frac{1}{n}\Big) + \frac{1}{n^2}; \end{aligned}$$

thus

$$\lim_{\|\mathcal{P}\|\to 0} s(\mathcal{P}) = \lim_{n\to\infty} \Big[\frac{1}{3}\Big(1+\frac{1}{n}\Big)\Big(2+\frac{1}{n}\Big) - \frac{3}{2n}\Big(1+\frac{1}{n}\Big) + \frac{1}{n^2}\Big] = \frac{2}{3}\,.$$

Therefore, the area of the region of interest is $\dfrac{2}{3}$.

Example 4.5. In this example we use a different approach to compute $\mathcal{A}(\mathrm{R})$ in Example 4.2. Assume that $0 < a < b$. Let $r = \big(\frac{b}{a}\big)^{\frac{1}{n}}$, $x_i = ar^i$, and $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$.
Claim: If $c > 1$, then $c^{\frac{1}{n}} = 1$ as n approaches infinity.
Proof of the claim: If $c > 1$, then $c^{\frac{1}{n}} > 1$. Let $y_n = c^{\frac{1}{n}} - 1$. Then $c = (1+y_n)^n \geqslant 1 + ny_n$ which implies that $0 < y_n \leqslant \dfrac{c-1}{n}$ for all $n \in \mathbb{N}$. By the Squeeze Theorem, $c^{\frac{1}{n}} \to 1$ as $n \to \infty$. □

Note that the claim above implies that $r \to 1$ as $n \to \infty$. Moreover, $x_i - x_{i-1} = a(r^i - r^{i-1}) = ar^{i-1}(r-1)$; thus

$$0 < a(r-1) = x_1 - x_0 \leqslant \|\mathcal{P}\| = x_n - x_{n-1} = ar^{n-1}(r-1) < b(r-1)\,.$$

Therefore, $\|\mathcal{P}\| \to 0$ is equivalent to that $n \to \infty$.

Using the "left end-point rule" approximation of the area,

$$\begin{aligned}\mathcal{A}(\mathrm{R}) &= \lim_{n\to\infty}\sum_{i=1}^n x_{i-1}^2(x_i - x_{i-1}) = \lim_{n\to\infty}\sum_{i=1}^n a^2 r^{2(i-1)} ar^{i-1}(r-1) = a^3 \lim_{n\to\infty}(r-1)\sum_{i=1}^n r^{3(i-1)}\\ &= a^3 \lim_{n\to\infty}(r-1)\frac{r^{3n}-1}{r^3-1} = a^3 \lim_{n\to\infty}\frac{\frac{b^3}{a^3}-1}{r^2+r+1} = \frac{b^3-a^3}{3}\,.\end{aligned}$$

Similarly, when applying the "right end-point rule" approximation, we obtain that

$$\lim_{n\to\infty}\sum_{i=1}^n x_i^2(x_i - x_{i-1}) = a^3 \lim_{n\to\infty}(r-1)\sum_{i=1}^n r^{3i} = a^3 \lim_{n\to\infty}(r-1)\frac{r^{3n+3}-r^3}{r^3-1} = \frac{b^3-a^3}{3}\,.$$

This also gives the area of the region R.

To compute an approximated value of $\mathcal{A}(\mathrm{R})$, there is no reason for evaluating the function at the left end-points or the right end-points like what we have discussed above. For example, we can also consider the "mid-point rule"

$$m(\mathcal{P}) = \sum_{i=1}^n f\Big(\frac{x_i + x_{i-1}}{2}\Big)(x_i - x_{i-1})$$

to approximate the value of $\mathcal{A}(\mathrm{R})$, and compute the limit of the sum above as $\|\mathcal{P}\|$ approaches 0 in order to obtain $\mathcal{A}(\mathrm{R})$. In fact, we should be able to consider any point $c_i \in [x_{i-1}, x_i]$ and consider the limit of the sum

$$\lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{n} f(c_i)(x_i - x_{i-1})$$

if the region R does have area.

Now let us forget about the concept of the area. For a general function $f : [a,b] \to \mathbb{R}$, we can also consider the limit above as $\|\mathcal{P}\|$ approaches 0, if the limit exists. The discussion above motivates the following definitions.

Definition 4.6: Partition of Intervals and Riemann Sums

A finite set $\mathcal{P} = \{x_0, x_1, \cdots, x_n\}$ is said to be a partition of the closed interval $[a,b]$ if $a = x_0 < x_1 < \cdots < x_n = b$. Such a partition $\mathcal{P}$ is usually denoted by $\{a = x_0 < x_1 < \cdots < x_n = b\}$. The norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$, is the number $\max\big\{x_i - x_{i-1} \,\big|\, 1 \leqslant i \leqslant n\big\}$; that is,

$$\|\mathcal{P}\| \equiv \max\big\{x_i - x_{i-1} \,\big|\, 1 \leqslant i \leqslant n\big\}.$$

A partition $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ is called regular if $x_i - x_{i-1} = \|\mathcal{P}\|$ for all $1 \leqslant i \leqslant n$.

Let $f : [a,b] \to \mathbb{R}$ be a function. A Riemann sum of f for the partition $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ of $[a,b]$ is a sum which takes the form

$$\sum_{i=1}^{n} f(c_i)(x_i - x_{i-1}),$$

where the set $\Xi = \{c_0, c_1, \cdots, c_{n-1}\}$ satisfies that $x_{i-1} \leqslant c_i \leqslant x_i$ for each $1 \leqslant i \leqslant n$.

Definition 4.7: Riemann Integrals - 黎曼積分

Let $f : [a,b] \to \mathbb{R}$ be a function. f is said to be Riemann integrable on $[a,b]$ if there exists a real number A such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is partition of $[a,b]$ satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums for the partition $\mathcal{P}$ belongs to the interval $(A - \varepsilon, A + \varepsilon)$. Such a number A (is unique and) is called the Riemann integral of f on $[a,b]$ and is denoted by $\displaystyle\int_{[a,b]} f(x)\,dx$.

Remark 4.8. For conventional reason, the Riemann integral of f over the interval with left end-point a and right-end point b is written as $\displaystyle\int_a^b f(x)\,dx$, and is called the definite integral of f from a to b. The function f sometimes is called the integrand of the integral.

We also note that here in the representation of the integral, x is a dummy variable; that is, we can use any symbol to denote the independent variable; thus

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(u)\,du$$

and etc.

The following example shows that no all functions are Riemann integrable.

Example 4.9. Consider the Dirichlet function

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is rational}\,, \\ 1 & \text{if } x \text{ is irrational}\,, \end{cases}$$

on the interval $[1,2]$. By partitioning $[1,2]$ into n sub-intervals with equal length, the Riemann sum given by the right end-point rule is always zero since the right end-point of each sub-interval is rational. On the other hand, by partitioning $[1,2]$ into n sub-intervals using geometric sequence $1, r, r^2, \cdots, r^{n-1}, 2$, where $r = 2^{\frac{1}{n}}$, by the fact that $r^i \notin \mathbb{Q}$ for each $1 \leqslant i \leqslant n-1$ the Riemann sum of f for this partition given by the right end-point rule is

$$\begin{aligned}\sum_{i=1}^{n} f(r^i)(r^i - r^{i-1}) &= \sum_{i=1}^{n-1}(r^i - r^{i-1}) = r^1 - r^0 + r^2 - r^1 + \cdots + r^{n-1} - r^{n-2} \\ &= r^{n-1} - r^0 = \frac{2}{r} - 1\end{aligned}$$

which approaches 1 as r approaches 1. Therefore, f is not integrable on $[1,2]$ since there are two possible limits of Riemann sums which means that the Riemann sums cannot concentrate around any firxed real number.

Theorem 4.10

If $f : [a,b] \to \mathbb{R}$ is continuous, then f is Riemann integrable on $[a,b]$.

Example 4.11. In this example we compute $\displaystyle\int_a^b x^q\,dx$ when $q \neq -1$ is a rational number and $0 < a < b$. Since $f(x) = x^q$ is continuous on $[a,b]$, by Theorem 4.10 to find the integral it suffices to find the limit of the Riemann sum given by the left end-point rule as $\|\mathcal{P}\|$ approaches 0.

We follow the idea in Example 4.5. Let $r = \left(\frac{b}{a}\right)^{\frac{1}{n}}$ and $x_i = ar^i$, as well as the partition $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$. Then the Riemann sum of f for the partition $\mathcal{P}$ given by left end-point rule is

$$\begin{aligned}L(\mathcal{P}) &= \sum_{i=1}^{n}(ar^{i-1})^q(ar^i - ar^{i-1}) = a^{q+1}(r-1)\sum_{i=1}^{n} r^{(i-1)(q+1)} = a^{q+1}(r-1)\frac{r^{n(q+1)}-1}{r^{q+1}-1} \\ &= \frac{r-1}{r^{q+1}-1}\big(b^{q+1} - a^{q+1}\big)\,.\end{aligned}$$

Since $\left.\frac{d}{dr}\right|_{r=1} r^{q+1} = (q+1)$, we have

$$\lim_{r\to 1}\frac{r^{q+1}-1}{r-1} = \left.\frac{d}{dr}\right|_{r=1} r^{q+1} = q+1\,;$$

thus by the fact that $r \to 1$ as $n \to \infty$ (or $\|\mathcal{P}\| \to 0$), we find that

$$\lim_{\|\mathcal{P}\|\to 0} L(\mathcal{P}) = \lim_{\|\mathcal{P}\|\to 0} L(\mathcal{P}) = \frac{b^{q+1}-a^{q+1}}{q+1}\,.$$

Therefore, $\displaystyle\int_a^b x^q\,dx = \frac{b^{q+1}-a^{q+1}}{q+1}$ if $q \neq 1$ is a rational number and $0 < a < b$.

Example 4.12. Since the sine function is continuous on any closed interval $[a, b]$, to find $\displaystyle\int_a^b \sin x\,dx$ we can partition $[a, b]$ into sub-intervals with equal length, use the right end-point rule to find an approximated value of the integral, and finally find the integral by passing the number of sub-intervals to the limit.

Let $\Delta x = \frac{b-a}{n}$ and $x_i = a + i\Delta x$. The right end-point rule gives the approximation

$$\sum_{i=1}^{n}\sin x_i\Delta x = \sum_{i=1}^{n}\sin(a+i\Delta x)\Delta x = \Delta x\sum_{i=1}^{n}\sin(a+i\Delta x)$$

of the integral.

Using the sum and difference formula, we find that

$$\cos\big[a + \big(i - \frac{1}{2}\big)\Delta x\big] - \cos\big[a + \big(i + \frac{1}{2}\big)\Delta x\big] = 2\sin(a+i\Delta x)\sin\frac{\Delta x}{2}\,;$$

thus if $\sin\dfrac{\Delta x}{2} \neq 0$,

$$\begin{aligned}\sum_{i=1}^{n}\sin(a+i\Delta x) = \frac{1}{2\sin\frac{\Delta x}{2}}\Big[\Big(\cos\big(a+\frac{1}{2}\Delta x\big) - \cos\big(a+\frac{3}{2}\Delta x\big)\Big) + \Big(\cos\big(a+\frac{3}{2}\Delta x\big) \\ - \cos\big(a+\frac{5}{2}\Delta x\big)\Big) + \cdots + \cos\big[a+(n-\frac{1}{2})\Delta x\big] \\ - \cos\big[a+\big(n+\frac{1}{2}\big)\Delta x\big]\Big]\end{aligned}$$

which, by the fact that $a+\big(n+\dfrac{1}{2}\Delta x\big) = b+\dfrac{1}{2}\Delta x$, implies that

$$\sum_{i=1}^{n}\sin x_i\Delta x = \frac{\frac{\Delta x}{2}}{\sin\frac{\Delta x}{2}}\Big[\cos\big(a+\frac{1}{2}\Delta x\big) - \cos\big(b+\frac{1}{2}\Delta x\big)\Big]\,.$$

By the fact that $\lim\limits_{x\to 0}\dfrac{\sin x}{x} = 1$ and the continuity of the cosine function, we conclude that

$$\int_a^b \sin x\,dx = \lim_{n\to\infty}\sum_{i=1}^{n}\sin x_i\Delta x = \cos a - \cos b\,.$$

Theorem 4.13

Let $f:[a,b]\to\mathbb{R}$ be a non-negative and continuous function. The area of the region enclosed by the graph of f, the x-axis, and the vertical lines $x=a$ and $x=b$ is $\displaystyle\int_a^b f(x)\,dx$.

Example 4.14. In this example we use the integral notation to denote the areas of some common geometric figures (without really doing computations):

1. $\displaystyle\int_{-2}^{2}\sqrt{4-x^2}\,dx = 2\pi\,;$ 2. $\displaystyle\int_{-1}^{1}\sqrt{4-x^2}\,dx = \frac{2\pi}{3}+\sqrt{3}\,;$ 3. $\displaystyle\int_{-1}^{\sqrt{3}}\sqrt{4-x^2}\,dx = \pi+\sqrt{3}.$

4.2.1 Properties of Definite Integrals

Definition 4.15

1. If f is defined at $x=a$, then $\displaystyle\int_a^a f(x)\,dx = 0$.
2. If f is integrable on $[a,b]$, then $\displaystyle\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx = -\int_{[a,b]} f(x)\,dx$.

Remark 4.16. By the definition above, if f is Riemann integrable on $[a, b]$, $\displaystyle\int_b^a f(x)\,dx$ is the limit of the sum

$$\sum_{i=1}^{n} f(x_i)(x_i - x_{i-}) \quad \textit{and} \quad \sum_{i=1}^{n} f(x_{i-1})(x_i - x_{i-1})$$

as $\max\left\{|x_i - x_{i-1}| \,\middle|\, 1 \leqslant i \leqslant n\right\} \to 0$, where $x_0 = b > x_1 > x_2 > \cdots > x_n = a$.

Theorem 4.17

If f is Riemann integrable on the three closed intervals determined by a, b and c, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,.$$

Theorem 4.18

Let $f, g : [a, b] \to \mathbb{R}$ be Riemann integrable on $[a, b]$ and k be a constant. Then the function $kf \pm g$ are Riemann integrable on $[a, b]$, and

$$\int_a^b (kf \pm g)(x)\,dx = k\int_a^b f(x)\,dx \pm \int_a^b g(x)\,dx\,.$$

Theorem 4.19

If f is non-negative and Riemann integrable on $[a, b]$, then $\displaystyle\int_a^b f(x)\,dx \geqslant 0$.

Corollary 4.20

If f, g are Riemann integrable on $[a, b]$ and $f(x) \leqslant g(x)$ for all $a \leqslant x \leqslant b$, then

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

Theorem 4.21

If f is Riemann integrable on $[a, b]$, then $|f|$ is Riemann integrable on $[a, b]$ and

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b \big|f(x)\big|\,dx\,.$$

Theorem 4.22: 可積必有界

Let $f : [a, b] \to \mathbb{R}$ be a function. If f is Riemann integrable on $[a, b]$, then f is bounded on $[a, b]$; that is, there exists $M > 0$ such that

$$\big|f(x)\big| \leqslant M \qquad \text{whenever} \quad x \in [a, b]\,.$$

Proof. Let f be Riemann integrable on $[a, b]$. Then there exists $A \in \mathbb{R}$ and $\delta > 0$ such that if $\mathcal{P}$ is a partition of $[a, b]$ satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sum of f for $\mathcal{P}$ belongs to $(A-1, A+1)$. Choose $n \in \mathbb{N}$ so that $\dfrac{b-a}{n} < \delta$. Then the regular partition $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$, where $x_i = a + \dfrac{b-a}{n} i$, satisfies $\|\mathcal{P}\| < \delta$.

Suppose the contrary that f is not bounded. Then there exists $x^* \in [a, b]$ such that

$$\big|f(x^*)\big| > \frac{n(|A|+1)}{b-a} + \sum_{i=1}^{n} \big|f(x_i)\big|\,.$$

Suppose that $x^* \in [x_{k-1}, x_k]$. By the fact that $\sum\limits_{\substack{i=1\\ i\neq k}}^{n} f(x_i)(x_i - x_{i-1}) + f(x^*)(x_k - x_{k-1})$ is a Riemann sum of f for $\mathcal{P}$, we have

$$A - 1 < \sum_{\substack{i=1\\ i\neq k}}^{n} f(x_i)(x_i - x_{i-1}) + f(x^*)(x_k - x_{k-1}) < A + 1\,.$$

Since $x_i - x_{i-1} = \dfrac{b-1}{n}$ for all $1 \leqslant i \leqslant n$, the inequality above shows that

$$\frac{n(A-1)}{b-a} - \sum_{\substack{i=1\\ i\neq k}}^{n} f(x_i) < f(x^*) < \frac{n(A+1)}{b-a} - \sum_{\substack{i=1\\ i\neq k}}^{n} f(x_i)$$

and the triangle inequality further implies that

$$-\Big[\frac{n(|A|+1)}{b-a} + \sum_{\substack{i=1\\ i\neq k}}^{n} \big|f(x_i)\big|\Big] < f(x^*) < \frac{n(|A|+1)}{b-a} + \sum_{\substack{i=1\\ i\neq k}}^{n} \big|f(x_i)\big|\,.$$

Therefore, we conclude that

$$\big|f(x^*)\big| < \frac{n(|A|+1)}{b-a} + \sum_{\substack{i=1\\ i\neq k}}^{n} \big|f(x_i)\big| \leqslant \frac{n(|A|+1)}{b-a} + \sum_{i=1}^{n} \big|f(x_i)\big|\,,$$

a contradiction. $\square$

Example 4.23. Let $f : [0, 1] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} \dfrac{1}{x} & \text{if } x \in (0, 1], \\ 0 & \text{if } x = 0. \end{cases}$$

Then f has only one discontinuity in $[0, 1]$ but f is not Riemann integrable on $[0, 1]$ since f is not bounded.

4.3 The Fundamental Theorem of Calculus

In this section, we develop a theory which shows a systematic way of finding integrals if the integrand is a continuous function.

Definition 4.24

A function F is an anti-derivative of f on an interval I if $F'(x) = f(x)$ for all x in I.

Theorem 4.25

If F is an anti-derivative of f on an interval I, then G is an anti-derivative of f on the interval I if and only if G is of the form $G(x) = F(x) + C$ for all x in I, where C is a constant. (導函數相同的函數相差一常數)

Proof. It suffices to show the "$\Rightarrow$" (only if) direction. Suppose that $F' = G' = f$ on I. Then the function $h = F - G$ satisfies $h'(x) = 0$ for all $x \in I$. By the mean value theorem, for any $a, b \in I$ with $a \neq b$, there exists c in between a and b such that

$$h(b) - h(a) = h'(c)(b - a).$$

Since $h'(x) = 0$ for all $x \in I$, $h(a) = h(b)$ for all $a, b \in I$; thus h is a constant function. □

Theorem 4.26: Mean Value Theorem for Integrals - 積分均值定理

Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there exists $c \in [a, b]$ such that

$$\int_a^b f(x)\, dx = f(c)(b - a).$$

Proof. By the Extreme Value Theorem, f has a maximum and a minimum on $[a, b]$. Let $M = f(x_1)$ and $m = f(x_2)$, where $x_1, x_2 \in [a, b]$, denote the maximum and minimum of f

on $[a,b]$, respectively. Then $m \leqslant f(x) \leqslant M$ for all $x \in [a,b]$; thus Corollary 4.20 implies that

$$m(b-a) = \int_a^b m\,dx \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b M\,dx = M(b-a)\,.$$

Therefore, the number $\dfrac{1}{b-a}\displaystyle\int_a^b f(x)\,dx \in [m,M]$. By the Intermidiate Value Theorem, there exists c in between x_1 and x_2 such that $f(c) = \dfrac{1}{b-a}\displaystyle\int_a^b f(x)\,dx$. □

Theorem 4.27: Fundamental Theorem of Calculus - 微積分基本定理

Let $f : [a,b] \to \mathbb{R}$ be a continuous function, and F be an anti-derivative of f on $[a,b]$. Then

$$\int_a^b f(x)\,dx = F(b) - F(a)\,.$$

Moreover, if $G(x) = \displaystyle\int_a^x f(t)\,dt$ for $x \in [a,b]$, then G is an anti-derivative of f.

We note that for $x \in [a,b]$, f is continuous on $[a,x]$; thus f is Riemann integrable on $[a,x]$ which shows that $G(x) = \displaystyle\int_a^x f(t)\,dt$ is well-defined.

Proof of the Fundamental Theorem of Calculus. Note that for $h \neq 0$ such that $x+h \in [a,b]$, we have

$$\frac{G(x+h)-G(x)}{h} = \frac{1}{h}\Big[\int_a^{x+h} f(t)\,dt - \int_a^x f(t)\,dt\Big] = \frac{1}{h}\int_x^{x+h} f(t)\,dt\,.$$

By the Mean Value Theorem for Integrals, there exists $c = c(h)$ in between x and $x+h$ such that $\dfrac{1}{h}\displaystyle\int_x^{x+h} f(t)\,dt = f(c)$. Since f is continuous on $[a,b]$, $\lim\limits_{h\to 0} f(c) = \lim\limits_{c\to x} f(c) = f(x)$; thus

$$\lim_{h\to 0}\frac{G(x+h)-G(x)}{h} = \lim_{h\to 0}\frac{1}{h}\int_x^{x+h} f(t)\,dt = \lim_{h\to 0} f(c) = f(x)$$

which shows that G is an anti-derivative of f on $[a,b]$.

By Theorem 4.25, $G(x) = F(x) + C$ for all $x \in [a,b]$. By the fact that $G(a) = 0$, $C = -F(a)$; thus

$$\int_a^b f(x)\,dx = G(b) = F(b) - F(a)$$

which concludes the theorem. □

Example 4.28. Since an anti-derivative of the function $y = x^q$, where $q \neq -1$ is a rational number, is $y = \dfrac{x^{q+1}}{q+1}$, we find that

$$\int_a^b x^q\,dx = \frac{x^{q+1}}{q+1}\Big|_{x=b} - \frac{x^{q+1}}{q+1}\Big|_{x=a} = \frac{b^{q+1}-a^{q+1}}{q+1}\,.$$

Example 4.29. Since an anti-derivative of the sine function is negative of cosine, we find that

$$\int_a^b \sin x\,dx = (-\cos)(b) - (-\cos)(b) = \cos b - \cos a\,.$$

Example 4.30. Find $\dfrac{d}{dx}\displaystyle\int_0^{\sqrt{x}} \sin^{100} t\,dt$ for $x > 0$.

Let $F(x) = \displaystyle\int_0^x \sin^{100} t\,dt$. Then by the chain rule,

$$\frac{d}{dx}F(\sqrt{x}) = F'(\sqrt{x})\frac{d}{dx}\sqrt{x} = \frac{1}{2\sqrt{x}}F'(\sqrt{x})\,.$$

By the Fundamental Theorem of Calculus, $F'(x) = \sin^{100} x$; thus

$$\frac{d}{dx}\int_0^{\sqrt{x}} \sin^{100} t\,dt = \frac{d}{dx}F(\sqrt{x}) = \frac{\sin^{100}\sqrt{x}}{2\sqrt{x}}\,.$$

Theorem 4.31

Let $f : [a,b] \to \mathbb{R}$ be continuous and f is differentiable on (a,b). If f' is Riemann integrable on $[a,b]$, then

$$\int_a^b f'(x)\,dx = f(b) - f(a)\,.$$

Proof. Let $\varepsilon > 0$ be given, and define $A = \displaystyle\int_a^b f'(x)\,dx$. By the definition of the integrability there exists $\delta > 0$ such that if $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ is a partition of $[a,b]$ satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums of f for $\mathcal{P}$ belongs to the interval $(A-\varepsilon, A+\varepsilon)$.

Let $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[a,b]$ satisfying that $\|\mathcal{P}\| < \delta$. Then by the mean value theorem, for each $1 \leqslant i \leqslant n$ there exists $x_{i-1} < c < x_i$ such that $f(x_i) - f(x_{i-1}) = f'(c_i)(x_i - x_{i-1})$. Since

$$\sum_{i=1}^n f'(c_i)(x_i - x_{i-1})$$

is a Riemann sum of f for $\mathcal{P}$, we must have

$$\Big|\sum_{i=1}^{n} f'(c_i)(x_i - x_{i-1}) - A\Big| < \varepsilon\,.$$

On the other hand, by the fact that

$$\begin{aligned}\sum_{i=1}^{n} f'(c_i)(x_i - x_{i-1}) &= \sum_{i=1}^{n} \big[f(x_i) - f(x_{i-1})\big] \\ &= f(x_1) - f(x_0) + f(x_2) - f(x_1) + \cdots + f(x_n) - f(x_{n-1}) \\ &= f(x_n) - f(x_0) = f(b) - f(a)\,,\end{aligned}$$

we conclude that

$$\Big|f(b) - f(a) - \int_a^b f'(x)\,dx\Big| < \varepsilon\,.$$

Since $\varepsilon > 0$ is chosen arbitrarily, we find that $\displaystyle\int_a^b f'(x)\,dx = f(b) - f(a)$. ☐

Remark 4.32. If f' is continuous on $[a,b]$, then the theorem above is simply a direct consequence of the Fundamental Theorem of Calculus. The theorem above can be viewed as a generalization of the Fundamental Theorem of Calculus.

Theorem 4.27 and 4.31 can be combined as follows:

Theorem 4.33

Let $f : [a,b] \to \mathbb{R}$ be a Riemann integrable function and F be an anti-derivative of f on $[a,b]$. Then

$$\int_a^b f(x)\,dx = F(b) - F(a)\,.$$

Moreover, if in addition f is continuous on $[a,b]$, then $G(x) = \displaystyle\int_a^x f(t)\,dt$ is differentiable on $[a,b]$ and

$$G'(x) = f(x) \qquad \text{for all } x \in [a,b]\,.$$

Definition 4.34

An anti-derivative of f, if exists, is denoted by $\displaystyle\int f(x)\,dx$, and sometimes is also called an indefinite integral of f.

- Basic Rules of Integration:

Differentiation Formula	Anti-derivative Formula
$\frac{d}{dx}C = 0$	$\int 0\,dx = C$
$\frac{d}{dx}x^r = rx^{r-1}$	$\int x^q\,dx = \frac{x^{q+1}}{q+1} + C$ if $q \neq -1$
$\frac{d}{dx}\sin x = \cos x$	$\int \cos x\,dx = \sin x + C$
$\frac{d}{dx}\cos x = -\sin x$	$\int \sin x\,dx = -\cos x + C$
$\frac{d}{dx}\tan x = \sec^2 x$	$\int \sec^2 x\,dx = \tan x + C$
$\frac{d}{dx}\sec x = \sec x \tan x$	$\int \sec x \tan x\,dx = \sec x + C$
$\frac{d}{dx}\big[kf(x) + g(x)\big] = kf'(x) + g'(x)$	$\int \big[kf'(x) + g'(x)\big]\,dx = kf(x) + g(x) + C$

4.4 Integration by Substitution - 變數變換

Suppose that $g : [a,b] \to \mathbb{R}$ is differentiable, and $f : \text{range}(g) \to \mathbb{R}$ is differentiable. Then the chain rule implies that $f \circ g$ is an anti-derivative of $(f' \circ g)g'$; thus provided that

1. $(f \circ g)'$ is Riemann integrable on $[a,b]$,
2. f' is Riemann integrable on the range of g,

then Theorem 4.31 implies that

$$\begin{aligned}\int_a^b f'\big(g(x)\big)g'(x)\,dx &= \int_a^b (f \circ g)'(x)\,dx = (f \circ g)(b) - (f \circ g)(a)\\ &= f\big(g(b)\big) - f\big(g(a)\big) = \int_{g(a)}^{g(b)} f'(u)\,du\,. \end{aligned} \tag{4.4.1}$$

Replacing f' by f in the identity above shows the following

Theorem 4.35

If the function $u = g(x)$ has a continuous derivative on the closed interval $[a,b]$, and f is continuous on the range of g, then

$$\int_a^b f\big(g(x)\big)g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du\,.$$

The anti-derivative version of Theorem 4.36 is stated as follows.

Theorem 4.36

Let g be a function with range I and f be a continuous function on I. If g is differentiable on its domain and F is an anti-derivative of f on I, then

$$\int f\big(g(x)\big)g'(x)\,dx = F\big(g(x)\big) + C$$

Letting $u = g(x)$ gives $du = g'(x)\,dx$ and

$$\int f(u)\,du = F(u) + C\,.$$

Example 4.37. Find $\displaystyle\int (x^2+1)^2(2x)\,dx$.

Let $u = x^2+1$. Then $du = 2xdx$; thus

$$\int (x^2+1)^2(2x)\,dx = \int u^2\,du = \frac{1}{3}u^3 + C = \frac{1}{3}(x^2+1)^3 + C\,.$$

Example 4.38. Find $\displaystyle\int \cos(5x)\,dx$.

Let $u = 5x$. Then $du = 5dx$; thus

$$\int \cos(5x)\,dx = \frac{1}{5}\int \cos u\,du = \frac{1}{5}\sin u + C = \frac{1}{5}\sin(5x) + C\,.$$

Example 4.39. Find $\displaystyle\int \sec^2 x(\tan x + 3)\,dx$.

Let $u = \tan x$. Then $du = \sec^2 x dx$; thus

$$\int \sec^2 x(\tan x+3)\,dx = \int (u+3)\,du = \frac{1}{2}u^2 + 3u + C = \frac{1}{2}\tan^2 x + 3\tan x + C\,.$$

On the other hand, let $v = \tan x + 3$. Then $dv = \sec^2 x\,dx$; thus

$$\begin{aligned}\int \sec^2 x(\tan x+3)\,dx &= \int v\,dv = \frac{1}{2}v^2 + C = \frac{1}{2}(\tan x+3)^2 + C\\ &= \frac{1}{2}\tan^2 x + 3\tan x + \frac{9}{2} + C\,.\end{aligned}$$

We note that even though the right-hand side of the two indefinite integrals look different, they are in fact the same since C could be any constant, and $\dfrac{9}{2} + C$ is also any constant.

Example 4.40. Find $\displaystyle\int \frac{2z\,dz}{\sqrt[3]{z^2+1}}$.

Method 1: Let $x = z^2+1$. Then $dx = 2z\,dz$; thus

$$\int \frac{2z\,dz}{\sqrt[3]{z^2+1}} = \int \frac{dx}{\sqrt[3]{x}} = \int x^{-\frac{1}{3}}\,dx = \frac{3}{2}x^{\frac{2}{3}} + C = \frac{3}{2}(z^2+1)^{\frac{2}{3}} + C\,.$$

Method 2: Let $y = \sqrt[3]{z^2+1}$. Then $y^3 = z^2+1$; thus $3y^2dy = 2z\,dz$. Therefore,

$$\int \frac{2z\,dz}{\sqrt[3]{z^2+1}} = \int \frac{3y^2dy}{y} = \int 3y\,dy = \frac{3}{2}y^2 + C = \frac{3}{2}(z^2+1)^{\frac{2}{3}} + C\,.$$

Example 4.41. Find $\displaystyle\int \frac{18\tan^2 x\sec^2 x}{(2+\tan^3 x)^2}\,dx$.

Let $u = 2+\tan^3 x$. Then $du = 3\tan^2 x\sec^x\,dx$; thus

$$\int \frac{18\tan^2 x\sec^2 x}{(2+\tan^3 x)^2}\,dx = \int \frac{6du}{u^2} = 6\int u^{-2}\,du = -6u^{-1} + C = -\frac{6}{2+\tan^3 x} + C\,.$$

Sometimes an definite integral can be evaluated even though the anti-derivative of the integrand cannot be found. In such kind of cases, we have to look for special structures so that we can simplify the integrals. There is no general rule for this, and we have to do this case by case.

Example 4.42. Find $\displaystyle\int_0^{\pi} \frac{2x\sin x}{3+\cos(2x)}\,dx$.

Let the integral be I. By the substitution $u = \pi - x$, we find that

$$\begin{aligned}
\mathrm{I} &= \int_{\pi}^{0} \frac{2(\pi-u)\sin(\pi-u)}{3+\cos(2(\pi-u))}(-1)\,du = \int_0^{\pi} \frac{2(\pi-u)\sin u}{3+\cos 2u}\,du \\
&= \int_0^{\pi} \frac{2\pi\sin u}{3+\cos 2u}\,du - \int_0^{\pi} \frac{2u\sin u}{3+\cos 2u}\,du = 2\pi\int_0^{\pi} \frac{\sin u}{3+\cos 2u}\,du - \mathrm{I}\,;
\end{aligned}$$

thus

$$\begin{aligned}
\mathrm{I} &= \pi\int_0^{\pi} \frac{\sin u}{3+\cos 2u}\,du = -\pi\int_0^{\pi} \frac{d(\cos u)}{3+2\cos^2 u-1} = -\frac{\pi}{2}\int_1^{-1} \frac{dv}{v^2+1} \\
&= \frac{\pi}{2}\int_{-1}^{1} \frac{dv}{v^2+1} = \frac{\pi}{2}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \frac{\sec^2 y}{\tan^2 y+1}\,dy = \frac{\pi}{2}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} dy = \frac{\pi^2}{4}\,.
\end{aligned}$$

4.5 Exercise

Problem 4.1. Let $f : [a,b] \to \mathbb{R}$ be a function, and f is Riemann integrable on $[a,b]$. Show that f must be bounded on $[a,b]$; that is, there exists a real number $M > 0$ such that $|f(x)| \leqslant M$ for all $a \leqslant x \leqslant b$.

Problem 4.2. Let $a < b$ be real numbers. Compute $\displaystyle\int_a^b \cos x \, dx$ by the following steps.

(a) Partition $[a,b]$ into n sub-intervals with equal length. Write down the Riemann sum using the right end-point rule.

(b) Prove that

$$\sum_{i=1}^{n} \cos(a + id) = \frac{\sin\left[a + \left(n + \frac{1}{2}\right)d\right] - \sin\left(a + \frac{d}{2}\right)}{2 \sin \dfrac{d}{2}}. \tag{4.5.1}$$

Hint: Use the sum and difference formula $\sin(\vartheta + \varphi) - \sin(\vartheta - \varphi) = 2 \sin\vartheta \cos\varphi$.

(c) Use (4.5.1) to simplify the Riemann sum in (a), and find the limit of the Riemann sum as n approaches infinity. Show that

$$\int_a^b \cos x \, dx = \sin b - \sin a\,.$$

Problem 4.3. Let $a < b$ be real numbers. Compute $\displaystyle\int_a^b x^N \, dx$, where N is a non-negative integer, by the following steps.

(a) Let $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a regular partition of $[a,b]$. Show that the Riemann sum using the right end-point rule is given by

$$I_n = \sum_{k=0}^{N} \left[C_k^N a^{N-k} (b-a)^{k+1} \Big(\frac{1}{n^{k+1}} \sum_{i=1}^{n} i^k\Big)\right],$$

where $C_k^N = \dfrac{N!}{k!(N-k)!}$.

(b) Show that

$$\sum_{i=1}^{n} i^k = \frac{1}{k+1}(n+1)^{k+1} - \frac{1}{k+1}\left[C_{k-1}^{k+1} \sum_{i=1}^{n} i^{k-1} + \cdots + C_1^{k+1} \sum_{i=1}^{n} i + (n+1)\right]. \tag{4.5.2}$$

Hint: Expand $(j+1)^k$ for $j = 0, 1, 2, \cdots, n$ by the binomial expansion formula, and sum over j to obtain the equality above.

(c) Use (4.5.2) to show that $\displaystyle\lim_{n\to\infty}\frac{1}{n^{k+1}}\sum_{i=1}^{n} i^k = \frac{1}{k+1}$ for each $k \in \mathbb{N}$.

(d) Use the limit in (c) to find the limit of the Riemann sum in (a) by passing to the limit as n approaches infinity. Simplify the result to show that

$$\int_a^b x^N\,dx = \frac{b^{N+1}-a^{N+1}}{N+1}\,.$$

Hint: (c) By induction!

Problem 4.4. In class we have used the limit of Riemann sums to compute the integral $\displaystyle\int_0^\pi x\cos x\,dx$. Find this integral by completing what we did in class.

Problem 4.5. Determine the following limits by identifying the limits as limits of certain Riemann sums so that the limits are the same as certain integrals.

1. $\displaystyle\lim_{n\to\infty}\frac{\sqrt{1}+\sqrt{2}+\sqrt{3}+\cdots+\sqrt{n}}{n^{\frac{3}{2}}}$.

2. $\displaystyle\lim_{n\to\infty}\frac{1}{\sqrt{n}}\Big(1+\frac{1}{\sqrt{2}}+\cdots+\frac{1}{\sqrt{n}}\Big)$.

3. $\displaystyle\lim_{n\to\infty}\Big[\frac{1}{\sqrt{n^2+2n}}+\frac{1}{\sqrt{n^2+4n}}+\frac{1}{\sqrt{n^2+6n}}+\cdots+\frac{1}{\sqrt{n^2+2n^2}}\Big]$.

Problem 4.6. Let $f:[a,b]\to\mathbb{R}$ be Riemann integrable on $[a,b]$, and $m\leqslant f(x)\leqslant M$ for all $x\in[a,b]$. Show that

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

Problem 4.7. Let $f:[0,1]\to\mathbb{R}$ be a function satisfying that

$$\big|f(x)-f(y)\big|\leqslant M|x-y|\qquad\forall\,x,y\in[0,1]\,.$$

Under the fact that f is Riemann integrable on $[0,1]$, show that

$$\Big|\int_0^1 f(x)\,dx-\frac{1}{n}\sum_{i=1}^{n} f\big(\frac{i}{n}\big)\Big|<\frac{M}{2n}\,.$$

Problem 4.8. Suppose that $f,g:[a,b]\to\mathbb{R}$ are Riemann integrable on $[a,b]$. Under the fact that fg is Riemann integrable on $[a,b]$, show that

$$\int_a^b f(x)g(x)\,dx\leqslant\Big(\int_a^b\big|f(x)\big|^2\,dx\Big)^{\frac{1}{2}}\Big(\int_a^b\big|g(x)\big|^2\,dx\Big)^{\frac{1}{2}}\,.$$

Problem 4.9. Recall that in Problem 2.5 we have "shown" that there exists a number $e > 1$ such that

$$\frac{d}{dx}\log_e x = \frac{1}{x} \qquad \forall\, x > 0\,.$$

In this example you need to compute $\displaystyle\int_1^b \log_e x\, dx$ by the following steps.

(a) Partition $[1, b]$ into n sub-intervals by $x_i = r^i$, where $1 \leqslant i \leqslant n$ and $r = b^{\frac{1}{n}}$. Show that the Riemann sum given by the right end-point rule is

$$(r-1)\log_e r \sum_{i=1}^{n} i r^{i-1}\,. \tag{4.5.3}$$

(b) Use (4.5.3) and the formula in Problem 4 of Exercise 4 to simplify the Riemann sum given above and show that the Riemann sum is

$$\frac{nbr - nb - b + 1}{n(r-1)}\log_e b = \left[b - \frac{b-1}{n(r-1)}\right]\log_e b\,.$$

(c) Pass the Riemann sum above to the limit as $n \to \infty$ to show that

$$\int_1^b \log_e x\, dx = b\log_e b - b + 1\,.$$

(d) Verify that $f(x) = x\log_e x - x$ is an anti-derivative of $y = \log_e x$.

Problem 4.10. Use Problem 2.14 to find the integral $\displaystyle\int_1^{\sqrt{3}} \frac{1}{x^2+1}\, dx$.

Problem 4.11. Find an anti-derivative of the function $y = x\sin x$ (using Riemann sums).
Hint: See Problem 2.4 for reference.

Chapter 5

Logarithmic, Exponential, and other Transcendental Functions

5.1 Inverse Functions

Definition 5.1

A function g is the inverse function of the function f if

$$f(g(x)) = x \quad \text{for all } x \text{ in the domain of } g \tag{5.1.1}$$

and

$$g(f(x)) = x \quad \text{for all } x \text{ in the domain of } f. \tag{5.1.2}$$

The inverse function of f is usually denoted by f^{-1}.

Some important observations about inverse functions:

1. If g is the inverse function of f, then f is the inverse function of g.

2. Note that (5.1.1) implies that

   (a) the domain of g is contained in the range of f,

   (b) the domain of f contains the range of g,

   (c) g is one-to-one since if $g(x_1) = g(x_2)$, then $x_1 = f(g(x_1)) = f(g(x_2)) = x_2$

   and (5.1.2) implies that

   (a) the domain of f is contained in the range of g,

(b) the domain of g contains the range of f,

(c) f is one-to-one since if $f(x_1) = f(x_2)$, then $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$.

According to the statements above, the domain of f^{-1} is the range of f, and the range of f^{-1} is the domain of f.

3. A function need not have an inverse function, but when it does, the inverse function is unique: Suppose that g and h are inverse function of f, then

   (a) the domain of g is identical to the domain of h (since they are both the range of f);

   (b) for each x in the range of f,
   $$f(g(x)) = x = f(h(x))$$
   thus by the fact that f is one-to-one, $g(x) = h(x)$ for all x in the range of f.

   Therefore, g and h are identical functions.

Example 5.2. The functions
$$f(x) = 2x^3 - 1 \qquad \text{and} \qquad g(x) = \sqrt[3]{\frac{x+1}{2}}$$
are inverse functions of each other since
$$f(g(x)) = 2\Big[\sqrt[3]{\frac{x+1}{2}}\Big]^3 - 1 = 2\frac{x+1}{2} - 1 = x$$
and
$$g(f(x)) = \sqrt[3]{\frac{2x^3 - 1 + 1}{2}} = \sqrt[3]{x^3} = x\,.$$

Theorem 5.3

A function f has an inverse function if and only if f is one-to-one.

Proof. It suffices to show the "$\Leftarrow$" direction. Suppose that f is one-to-one. Then for each x in the range of f, there exists only a unique y in the domain of f such that $f(y) = x$. Denote the map $x \mapsto y$ by g; that is,
$$y = g(x) \quad \text{if} \quad f(y) = x \text{ and } x \in \text{Range}(f)\,.$$

Then $f(g(x)) = x$ for all x in the range of f. Since the domain of g is the range of f, we find that

$$f(g(x)) = x \quad \text{for all } x \text{ in the domain of } g.$$

On the other hand, by the definition of g we must also have

$$g(f(x)) = x \qquad \text{for all } x \text{ in the domain of } f;$$

thus f has an inverse function. $\square$

Theorem 5.4

Let f be a function with inverse f^{-1}. The graph of f contains the point (a, b) if and only if the graph of f^{-1} contains the point (b, a).

Proof. Let (a, b) be on the graph of f. Then $b = f(a)$ which implies that $f^{-1}(b) = f^{-1}(f(a)) = a$. Therefore, (b, a) is on the graph of f^{-1}. $\square$

Remark 5.5. Theorem 5.4 implies that the graph of f and the graph of f^{-1} is symmetric above the straight line $y = x$.

Theorem 5.6

Let f be a function defined on an interval I and have an inverse function. Then

1. if f is continuous on I, then f^{-1} is continuous on its domain;
2. if f is strictly increasing on I, then f^{-1} is strictly increasing on the range of f;
3. if f is strictly decreasing on I, then f^{-1} is strictly decreasing on the range of f;
4. if f is differentiable on an interval containing c and $f'(c) \neq 0$, then f^{-1} is differentiable at $f(c)$.

Proof. We only show 2 (and the proof of 3 is similar).

To show that f^{-1} is strictly increasing on the range of f, we need to show that

$$f^{-1}(x_1) < f^{-1}(x_2) \quad \text{if } x_1 < x_2 \text{ are in the range of } f.$$

Nevertheless, if f is increasing on I and $x_1 < x_2$ are in the range of f, there exists $y_1 = f^{-1}(x_1)$ and $y_2 = f^{-1}(x_2)$ in I such that $f(y_1) = x_1$ and $f(y_2) = x_2$. Since $x_1 < x_2$, $y_1 \not\geq y_2$; thus the trichotomy law implies that $y_1 < y_2$. $\square$

Remark 5.7. If I is not an interval, then even if $f : I \to \mathbb{R}$ is one-to-one and continuous, f^{-1} might be discontinuous. For example, let $I = \big[0, \frac{\pi}{2}\big) \cup \big(\frac{\pi}{2}, \pi\big)$ and $f(x) = \tan x$. Then clearly $f : I \to \mathbb{R}$ is one-to-one, onto and continuous; however, the inverse function is not continuous at 0: you can check this by looking at the graph of f^{-1}.

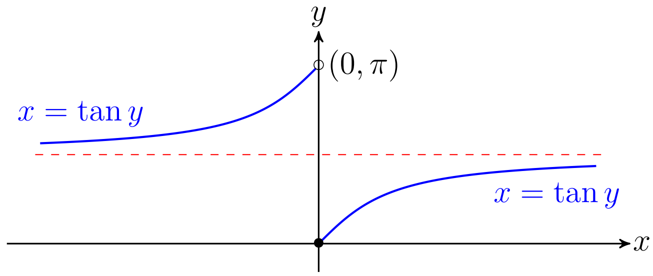


Figure 5.1: The graph of f^{-1}

From the graph of f^{-1}, we find that $\lim\limits_{x\to 0^+} f^{-1}(x) = 0$ while $\lim\limits_{x\to 0^-} f^{-1}(x) = \pi$; thus f is not continuous at 0.

Theorem 5.8: Inverse Function Differentiation

Let f be a function that is differentiable on an interval I. If f has an inverse function g, then g is differentiable at any x for which $f'(g(x)) \neq 0$. Moreover,

$$g'(x) = \frac{1}{f'(g(x))} \qquad \text{for all } x \text{ with } f'(g(x)) \neq 0.$$

Proof. Suppose that f is differentiable at $g(c) \in I$ and $f'(g(c)) \neq 0$. We show that g is differentiable at c. If $k \neq 0$ is small enough, $g(c+k) - g(c) = h$. Then $c + k = f(g(c) + h)$. Moreover, $h \to 0$ as $k \to 0$ since g is continuous (by Theorem 5.6). Therefore,

$$\frac{g(c+k) - g(c)}{k} = \frac{h}{f(g(c)+h) - f(g(c))} = \frac{h}{f(g(c)+h) - f(g(c))}$$

which approaches $\dfrac{1}{f'(g(c))}$ as k approaches zero. Therefore, $g'(c) = \dfrac{1}{f'(g(c))}$. $\square$

5.2 The Function $y = \ln x$

Recall Example 4.11 that $\displaystyle\int_a^b x^q\,dx = \frac{b^{q+1} - a^{q+1}}{q+1}$ if $q \neq -1$ is a rational number and $0 < a < b$. What happened to the case $\displaystyle\int_a^b x^{-1}\,dx$? In the following, we define a new function which can be used to compute this integral.

Definition 5.9

The function $\ln : (0, \infty) \to \mathbb{R}$ is defined by

$$\ln x = \int_1^x \frac{1}{t}\, dt \qquad \forall\, x > 0\,.$$

We emphasize again that we **cannot** write $\ln x = \displaystyle\int_1^x \frac{1}{x}\, dx$ since the upper limit in the integral is some arbitrary but fixed number (denoted by x) and the variable of the integrand should be really arbitrary.

Remark 5.10. For historical reason, when the variable is clear we should ignore the parentheses and write $\ln x$ instead of $\ln(x)$. On the other hand, if the variable is product of several variables such as xy, for the sake of clarity we should still write $\ln(xy)$ instead of $\ln xy$.

5.2.1 Properties of $y = \ln x$

- **Differentiability**

Since the function $y = \dfrac{1}{x}$ is continuous on $(0, \infty)$, the Fundamental Theorem of Calculus implies the following

Theorem 5.11

$\dfrac{d}{dx} \ln x = \dfrac{1}{x}$ for all $x > 0$.

In particular, the function $y = \ln x$ is continuous on $(0, \infty)$.

Corollary 5.12

The function $\ln : (0, \infty) \to \mathbb{R}$ is strictly increasing on $(0, \infty)$, and the graph of $y = \ln x$ is concave downward on $(0, \infty)$.

Example 5.13. In this example we prove that

$$x - \frac{x^2}{2} \leqslant \ln(1 + x) \leqslant x \qquad \forall\, x > 0\,. \tag{5.2.1}$$

Let $f(x) = \ln(1 + x) - x + \dfrac{x^2}{2}$ and $g(x) = \ln(1 + x) - x$. Then for $x > 0$,

$$f'(x) = \frac{1}{1 + x} - 1 + x = \frac{x^2}{1 + x} > 0\,, \qquad g'(x) = \frac{1}{1 + x} - 1 = \frac{-x}{1 + x} < 0\,.$$

The two identities above shows that f is strictly increasing on $[0, \infty)$ and g is strictly decreasing on $[0, \infty)$. Therefore,

$$f(x) > f(0) = 0 \quad \text{and} \quad g(x) < g(0) = 0 \qquad \forall\, x > 0\,.$$

These inequalities lead to (5.2.1).

- **The range**

Next we show that $\lim\limits_{x\to\infty} \ln x = \infty$ and $\lim\limits_{x\to-\infty} \ln x = -\infty$. To see this, we note that

$$\begin{aligned}
\ln(2^n) &= \int_1^{2^n} \frac{1}{t}\,dt = \int_1^2 \frac{1}{t}\,dt + \int_2^4 \frac{1}{t}\,dt + \int_4^8 \frac{1}{t}\,dt + \cdots + \int_{2^{n-1}}^{2^n} \frac{1}{t}\,dt \\
&= \sum_{i=1}^n \int_{2^{i-1}}^{2^i} \frac{1}{t}\,dt \geqslant \sum_{i=1}^n \int_{2^{i-1}}^{2^i} \frac{1}{2^i}\,dt = \sum_{i=1}^n \frac{2^i - 2^{i-1}}{2^i} = \sum_{i=1}^n \frac{1}{2} = \frac{n}{2}
\end{aligned}$$

and

$$\begin{aligned}
\ln(2^{-n}) &= \int_1^{2^{-n}} \frac{1}{t}\,dt = -\int_{2^{-n}}^1 \frac{1}{t}\,dt = -\Big[\int_{2^{-n}}^{2^{-n+1}} \frac{1}{t}\,dt + \int_{2^{-n+1}}^{2^{-n+2}} \frac{1}{t}\,dt + \cdots + \int_{\frac{1}{2}}^1 \frac{1}{t}\,dt\Big] \\
&= -\sum_{i=1}^n \int_{2^{-i}}^{2^{1-i}} \frac{1}{t}\,dt \leqslant -\sum_{i=1}^n \int_{2^{-i}}^{2^{1-i}} \frac{1}{2^{1-i}}\,dt = -\sum_{i=1}^n \frac{2^{1-i} - 2^{-i}}{2^{1-i}} = -\sum_{i=1}^n \frac{1}{2} = -\frac{n}{2}\,;
\end{aligned}$$

thus we have $\lim\limits_{x\to\infty} \ln x = \infty$ and $\lim\limits_{x\to-\infty} \ln x = -\infty$. By the continuity of ln and the Intermediate Value Theorem, for each $b \in \mathbb{R}$ there exists one $a \in (0, \mathbb{R})$ such that $b = \ln a$. By the strict monotonicity $\ln : (0, \infty) \to \mathbb{R}$ is one-to-one and onto.

Remark 5.14. In particular, there exists one unique number e such that $\ln e = 1$. We note that

$$\ln 2 = \int_1^2 \frac{1}{t}\,dt = \int_1^{1.5} \frac{1}{t}\,dt + \int_{1.5}^2 \frac{1}{t}\,dt \leqslant \frac{0.5}{1} + \frac{0.5}{1.5} = \frac{5}{6} < 1$$

and

$$\begin{aligned}
\ln 3 &= \int_1^3 \frac{1}{t}\,dt = \Big(\int_1^{1.25} + \int_{1.25}^{1.5} + \int_{1.5}^{1.75} + \int_{1.75}^2 + \int_2^{2.5} + \int_{2.5}^3\Big)\frac{1}{t}\,dt \\
&\geqslant \frac{0.25}{1.25} + \frac{0.25}{1.5} + \frac{0.25}{1.75} + \frac{0.25}{2} + \frac{0.5}{2.5} + \frac{0.5}{3} \\
&= \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{5} + \frac{1}{6} = \frac{841}{840} > 1\,.
\end{aligned}$$

Therefore, $2 < e < 3$. In fact, $e \approx 2.718281828459$.

Example 5.15. In this example we show that there is no slant/horizontal asymptote of the graph of $y = \ln x$. Recall that if the graph of $y = \ln x$ has a slant/horizontal asymptote $y = mx + k$, then $m = \lim\limits_{x\to\infty} \dfrac{\ln x}{x}$ and $k = \lim\limits_{x\to\infty} (\ln x - mx)$. We first show that $\lim\limits_{x\to\infty} \dfrac{\ln x}{x} = 0$.

Let $\varepsilon > 0$. Choose $M = \max\left\{\frac{\varepsilon}{2}, 1\right\}$. Then if $x > M$, for all $1 < c < x$ we have

$$0 < \frac{\ln x}{x} = \frac{1}{x}\int_1^x \frac{1}{t}\,dt = \frac{1}{x}\Big[\int_1^c \frac{1}{t}\,dt + \int_c^x \frac{1}{t}\,dt\Big] \leqslant \frac{c-1}{x} + \frac{1}{x}\int_c^x \frac{1}{t}\,dt\,.$$

By the mean value theorem for integrals (Theorem 4.26), there exists $c \leqslant d \leqslant x$ such that $\displaystyle\int_c^x \frac{1}{t}\,dt = \frac{x-c}{d}$; thus if $x > M$ and $1 < c < x$,

$$0 < \frac{\ln x}{x} = \frac{1}{x}\int_1^x \frac{1}{t}\,dt \leqslant \frac{c-1}{x} + \frac{x-c}{dx} \leqslant \frac{c-1}{M} + \frac{1}{M} \leqslant \frac{\varepsilon c}{2} < \varepsilon\,,$$

where the last inequality is concluded by choosing $1 < c < x$ and $c < 2$. Therefore, for every $\varepsilon > 0$ there exists $M > 0$ such that

$$\Big|\frac{\ln x}{x} - 0\Big| < \varepsilon \qquad \text{whenever} \quad x > M\,.$$

This is exactly the definition of $\lim\limits_{x\to\infty} \dfrac{\ln x}{x} = 0$. However, since the range of ln is $\mathbb{R}$, $\lim\limits_{x\to\infty} \ln x = \infty$ which implies that

$$\lim_{x\to 0}(\ln x - 0\cdot x) \text{ D.N.E.}$$

Therefore, there is no slant/horizontal asymptote of the graph of $y = \ln x$.

- **Logarithmic Laws**

The most important property of the function $y = \ln x$ is the relation among $\ln a$, $\ln b$ and $\ln(ab)$. By the property of integration,

$$\ln(ab) = \int_1^{ab} \frac{1}{t}\,dt = \int_1^a \frac{1}{t}\,dt + \int_a^{ab} \frac{1}{t}\,dt = \ln a + \int_a^{ab} \frac{1}{t}\,dt\,.$$

By the substitution $t = au$, $dt = adu$; thus

$$\int_a^{ab} \frac{1}{t}\,dt = \int_1^b \frac{1}{au} adu = \int_1^b \frac{1}{u}\,du = \ln b\,.$$

Therefore, we obtain the identity:

$$\ln(ab) = \ln a + \ln b \qquad \forall\, a, b > 0\,. \tag{5.2.2}$$

Having established (5.2.2), we can show that the function ln is a logarithmic function for the following reason. First, we observe that for all $a > 0$ and $n \in \mathbb{N}$,

$$\ln(a^n) = \ln(a^{n-1}a) = \ln(a^{n-1}) + \ln a = \ln(a^{n-2}a) + \ln a = \ln(a^{n-2}) + 2\ln a = \cdots = n\ln a\,.$$

Moreover, by the definition of ln, $0 = \ln(1) = \ln(a^0) = 0\ln a$; thus

$$\ln(a^n) = n\ln a \qquad \forall\, a > 0\,, n \in \mathbb{N} \cup \{0\}\,.$$

Next, by the law of exponents, for $a > 0$ and $n \in \mathbb{N}$ we have

$$0 = \ln(a^0) = \ln(a^n \cdot a^{-n}) = \ln(a^n) + \ln(a^{-n}) = n\ln a + \ln(a^{-n})\,.$$

Therefore, for all $n \in \mathbb{N}$, we also have $\ln(a^{-n}) = -n\ln a$; hence

$$\ln(a^n) = n\ln a \qquad \forall\, a > 0\,, n \in \mathbb{Z}\,.$$

The identity above also implies that if $k, n \in \mathbb{Z}$ and $n \neq 0$,

$$n\ln(a^{\frac{k}{n}}) = \ln((a^{\frac{k}{n}})^n) = \ln(a^k) = k\ln a\,,$$

and this shows that

$$\ln(a^{\frac{k}{n}}) = \frac{k}{n}\ln a \qquad \forall\, a > 0\,, n, k \in \mathbb{Z}\,, n \neq 0\,.$$

As a consequence,

$$\ln(a^r) = r\ln a \qquad \forall\, a > 0\,, r \in \mathbb{Q}\,.$$

Finally, we find that $\ln(e^r) = r\ln e = r$, so $\ln x$ is indeed the logarithm of x to the base e. In other words, we obtain that

$$\log_e x = \ln x = \int_1^x \frac{1}{t}\,dt \qquad \forall\, x > 0\,. \tag{5.2.3}$$

Theorem 5.16: Logarithmic properties of $y = \ln x$

Let a, b be positive numbers and r be a rational number. Then

1. $\ln 1 = 0$;
2. $\ln(ab) = \ln a + \ln b$;
3. $\ln(a^r) = r\ln a$;
4. $\ln\left(\frac{a}{b}\right) = \ln a - \ln b$.

Remark 5.17. Since the function $y = \ln x$ has the logarithmic property, it is called the **natural logarithmic** function.

Example 5.18. Let $f(x) = \dfrac{(x^2+3)^2}{x\sqrt[3]{x^2+1}}$. Since $\ln f(x) = 2\ln(x^2+3) - \ln x - \dfrac{1}{3}\ln(x^2+1)$ for $x > 0$, by the chain rule we find that

$$\frac{f'(x)}{f(x)} = \frac{d}{dx}\ln f(x) = \frac{4x}{x^2+3} - \frac{1}{x} - \frac{2x}{3(x^2+1)};$$

thus

$$f'(x) = \frac{(x^2+3)^2}{x\sqrt[3]{x^2+1}}\Big[\frac{d}{dx}\ln f(x) = \frac{4x}{x^2+3} - \frac{1}{x} - \frac{2x}{3(x^2+1)}\Big].$$

Theorem 5.19

If f is a differentiable function on an interval I, then $\ln|f|$ is differentiable at those point $x \in I$ satisfying $f(x) \neq 0$. Moreover,

$$\frac{d}{dx}\ln\big|f(x)\big| = \frac{f'(x)}{f(x)} \qquad \text{for all } x \in I \text{ with } f(x) \neq 0\,.$$

Proof. Note that the function $y = |x|$ is differentiable at non-zero points, and

$$\frac{d}{dx}|x| = \frac{d}{dx}(x^2)^{\frac{1}{2}} = \frac{1}{2}(x^2)^{-\frac{1}{2}} \cdot 2x = \frac{x}{|x|} \qquad \forall\, x \neq 0\,.$$

If $f(c) \neq 0$, by the fact that the natural logarithmic function ln is differentiable at $|f(c)|$, the absolute function $|\cdot|$ is differentiable at $f(c)$ and f is differentiable at c, the chain rule implies that $y = \ln\big|f(x)\big|$ is differentiable at c and

$$\frac{d}{dx}\Big|_{x=c}\ln\big|f(x)\big| = \frac{1}{\big|f(c)\big|}\frac{f(c)}{\big|f(c)\big|}f'(c) = \frac{f'(c)}{f(c)}\,. \qquad \square$$

Example 5.20. $\dfrac{d}{dx}\ln|\cos x| = \dfrac{-\sin x}{\cos x} = -\tan x$ for all x with $\cos x \neq 0$.

Example 5.21. Compute the derivative of $f(x) = \dfrac{(x^2+3)^2}{x\sqrt[3]{x^2+1}}$ for $x > 0$.

Let $h(x) = \ln f(x)$. Then

$$\begin{aligned}\frac{f'(x)}{f(x)} = h'(x) &= \frac{d}{dx}\Big[2\ln(x^2+3) - \ln x - \frac{1}{3}\ln(x^2+1)\Big]\\ &= 2\frac{d}{dx}\ln(x^2+3) - \frac{d}{dx}\ln x - \frac{1}{3}\frac{d}{dx}\ln(x^2+1)\\ &= \frac{4x}{x^2+3} - \frac{1}{x} - \frac{2x}{3(x^2+1)};\end{aligned}$$

thus

$$f'(x) = \frac{(x^2+3)^2}{x\sqrt[3]{x^2+1}}\left[\frac{4x}{x^2+3} - \frac{1}{x} - \frac{2x}{3(x^2+1)}\right].$$

5.3 Integrations Related to $y = \ln x$

Theorem 5.19 implies the following

Theorem 5.22

1. $\displaystyle\int \frac{1}{x}\,dx = \ln|x| + C$; 2. $\displaystyle\int \frac{f'(x)}{f(x)}\,dx = \ln\big|f(x)\big| + C$.

Example 5.23. Compute $\displaystyle\int \frac{x}{x^2+1}\,dx$. From observation, the numerator is a half of the derivative of the denominator, so

$$\int \frac{x}{x^2+1}\,dx = \frac{1}{2}\int \frac{2x}{x^2+1}\,dx = \frac{1}{2}\ln(x^2+1) + C\,.$$

Example 5.24. Compute $\displaystyle\int \frac{1}{x\ln x}\,dx$. Let $u = \ln x$. Then $du = \dfrac{1}{x}\,dx$; thus

$$\int \frac{1}{x\ln x}\,dx = \int \frac{1}{u}\,du = \ln|u| + C = \ln|\ln x| + C\,.$$

Theorem 5.25

1. $\displaystyle\int \sin x\,dx = -\cos x + C$; 2. $\displaystyle\int \cos x\,dx = \sin x + C$;
3. $\displaystyle\int \tan x\,dx = -\ln|\cos x| + C = \ln|\sec x| + C$;
4. $\displaystyle\int \sec x\,dx = \ln|\sec x + \tan x| + C$.

Proof. We only prove 4. Let $t = \tan\dfrac{x}{2}$. Then $\sin x = \dfrac{2t}{1+t^2}$, $\cos x = \dfrac{1-t^2}{1+t^2}$ and $dx = \dfrac{2dt}{1+t^2}$; thus

$$\begin{aligned}\int \sec x\,dx &= \int \frac{1+t^2}{1-t^2}\frac{2}{1+t^2}\,dt = \int \frac{2}{1-t^2}\,dt = \int \frac{-2}{(t-1)(t+1)}\,dt\\ &= \int \left[\frac{1}{t+1} - \frac{1}{t-1}\right]dt = \ln|t+1| - \ln|t-1| + C = \ln\left|\frac{t+1}{t-1}\right| + C\,.\end{aligned}$$

The conclusion then follows from the identity

$$\begin{aligned}
\frac{t+1}{t-1} &= \frac{\sin\frac{x}{2}+\cos\frac{x}{2}}{\sin\frac{x}{2}-\cos\frac{x}{2}} = \frac{\left(\sin\frac{x}{2}+\cos\frac{x}{2}\right)^2}{\sin^2\frac{x}{2}-\cos^2\frac{x}{2}} = \frac{1+2\sin\frac{x}{2}\cos\frac{x}{2}}{-\cos x} \\
&= -\frac{1+\sin x}{\cos x} = -(\sec x + \tan x)\,. \qquad \square
\end{aligned}$$

Finally we compute $\displaystyle\int_1^a \ln x\,dx$ for $a>0$. Suppose first that $a>1$. Following the idea of Example 4.5, we let $r=a^{\frac{1}{n}}$ and $x_i=r^i$, as well as a partition $\mathcal{P}=\{1=x_0<x_1<\cdots<x_n=a\}$ of $[1,a]$. Then the Riemann sum of f for the partition $\mathcal{P}$ given by the right end-point rule, which happens to be the upper sum of f for the partition $\mathcal{P}$, is

$$S(\mathcal{P})=\sum_{i=1}^n \ln(x_i)(x_i-x_{i-1}) = \sum_{i=1}^n \ln(r^i)(r^i-r^{i-1}) = (r-1)\ln r\sum_{i=1}^n ir^{i-1}\,.$$

Note that $ir^{i-1}=\dfrac{d}{dr}r^i$; thus

$$\begin{aligned}
\sum_{i=1}^n ir^{i-1} &= \sum_{i=1}^n \frac{d}{dr}r^i = \frac{d}{dr}\sum_{i=1}^n r^i = \frac{d}{dr}\frac{r^{n+1}-r}{r-1} = \frac{\big[(n+1)r^n-1\big](r-1)-r^{n+1}+r}{(r-1)^2} \\
&= \frac{nr^{n+1}-(n+1)r^n+1}{(r-1)^2} = \frac{nar-(n+1)a+1}{(r-1)^2}\,.
\end{aligned}$$

By the fact that $n=\dfrac{\ln a}{\ln r}$,

$$S(\mathcal{P})=\frac{ra\ln a - a\ln a - a\ln r + \ln r}{r-1}\,.$$

Since $\|\mathcal{P}\|\to 0$ is equivalent to that $r\to 1$,

$$\begin{aligned}
\lim_{\|\mathcal{P}\|\to 0} S(\mathcal{P}) &= \lim_{r\to 1}\frac{ra\ln a - a\ln a - a\ln r + \ln r}{r-1} = \frac{d}{dr}\Big|_{r=1}\big(ra\ln a - a\ln a - a\ln r + \ln r\big) \\
&= a\ln a - a + 1\,.
\end{aligned}$$

If $0<a<1$, by Remark 4.16 it suffices to show that $a^{\frac{1}{n}}\to 1$ as n approaches infinity. Nevertheless, $a^{\frac{1}{n}}=1/(1/a)^{\frac{1}{n}}$ and the denominator approaches 1 as n approaches infinity; thus $\lim\limits_{n\to\infty} a^{\frac{1}{n}}=1$ even if $0<a<1$.

Theorem 5.26

1. $\displaystyle\int_1^a \ln x\,dx = a\ln a - a + 1$ for all $a>0$;
2. $\displaystyle\int \ln x\,dx = x\ln x - x + C$.

Example 5.27. Find the limit $\lim\limits_{n\to\infty}\left(\dfrac{n!}{n^n}\right)^{\frac{1}{n}}$.

Consider the sum $\sum\limits_{k=1}^{n}\dfrac{1}{n}\ln\dfrac{k}{n}$. This sum looks like a Riemann sum of the "integral" $\displaystyle\int_0^1 \ln x\,dx$; however, since $\ln x$ blows up at $x=0$, $\ln x$ is not Riemann integrable on $[0,1]$. In other words, the sum is not a Riemann sum for a particular integral.

On the other hand, by the monotonicity of the function $y=\ln x$, we find that

$$\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n}=\sum_{k=1}^{n-1}\frac{1}{n}\ln\frac{k}{n}\leqslant\int_{\frac{1}{n}}^{1}\ln x\,dx\leqslant\sum_{k=2}^{n}\frac{1}{n}\ln\frac{k}{n}=-\frac{1}{n}\ln\frac{1}{n}+\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n};$$

thus by Theorem 5.26,

$$\frac{1}{n}-1\leqslant\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n}\leqslant-\frac{1}{n}\ln\frac{1}{n}+\frac{1}{n}-1\,.$$

Therefore, by the fact that $\lim\limits_{n\to\infty}\dfrac{\ln n}{n}=0$, we conclude from the Squeeze Theorem that

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n}=-1\,.$$

Finally, note that

$$\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n}=\frac{1}{n}\sum_{k=1}^{n}\ln\frac{k}{n}=\frac{1}{n}\ln\frac{n!}{n^n}=\ln\left(\frac{n!}{n^n}\right)^{\frac{1}{n}};$$

thus the continuity and strict monotonicity of $y=\ln x$ implies that

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

5.4 Exponential Functions

In the previous section we have shown that the natural logarithmic function $\ln:(0,\infty)\to\mathbb{R}$ is one-to-one and onto. Therefore, for each $a\in\mathbb{R}$ there exists a unique $b\in(0,\infty)$ satisfying $a=\ln b$. The map $a\mapsto b$ is called the natural exponential function. To be more precise, we have the following

Definition 5.28

The natural exponential function $\exp:\mathbb{R}\to(0,\infty)$ is a function defined by

$$\exp(x)=y\qquad\text{if and only if}\qquad x=\ln y\,.$$

By the definition of the natural exponential function, we have

$$\exp(\ln x) = x \quad \forall\, x \in (0, \infty) \qquad \text{and} \qquad \ln(\exp(x)) = x \quad \forall\, x \in \mathbb{R}\,. \tag{5.4.1}$$

Therefore, exp and ln are inverse functions to each other; thus $\exp : \mathbb{R} \to (0, \infty)$ is one-to-one, onto, and strictly increasing. Note that by the definition, $\exp(0) = 1$.

Let $a > 0$ be a real number. If $r \in \mathbb{Q}$, a^r is a well-defined positive number and the logarithmic laws implies that

$$\ln a^r = r \ln a\,.$$

By the definition of the natural exponential function, $a^r = \exp(r \ln a)$ for all $r \in \mathbb{Q}$. Since $\exp : \mathbb{R} \to (0, \infty)$ is continuous, for a real number x, we shall defined a^x as $\exp(x \ln a)$ and this induces the following

Definition 5.29

Let $a > 0$ be a real number. For each $x \in \mathbb{R}$, the exponential function to the base a, denote by $y = a^x$, is defined by $a^x \equiv \exp(x \ln a)$. In other words,

$$a^x = \exp(x \ln a) \qquad \forall\, x \in \mathbb{R}\,.$$

Remark 5.30. For each $x \in \mathbb{R}$, the number 1^x is 1 since $1^x = \exp(x \ln 1) = \exp(0) = 1$.

Remark 5.31. The function $y = e^x$ is identical to the function $y = \exp(x)$ since

$$e^x = \exp(x \ln e) = \exp(x) \qquad \forall\, x \in \mathbb{R}\,.$$

Therefore, we often write $\exp(x)$ as e^x as well (even though e^x, when x is a irrational number, has to be defined through the natural exponential function), and write $a^x = e^{x \ln a}$. Moreover, by the definition of the natural exponential function,

$$\ln(a^x) = \ln(\exp(x \ln a)) = x \ln a \qquad \forall\, a > 0 \text{ and } x \in \mathbb{R}\,. \tag{5.4.2}$$

5.4.1 Properties of Exponential Functions

- **The range and the strict monotonicity of the exponential functions**

Note that Theorem 5.6 implies that $\exp : \mathbb{R} \to (0, \infty)$ is strictly increasing. Suppose that $a > 1$. Then $\ln a > 0$ which further implies that

$$a^{x_1} = \exp(x_1 \ln a) < \exp(x_2 \ln a) = a^{x_2} \qquad \forall\, x_1 < x_2\,.$$

Similarly, if $0 < a < 1$, the exponential function to the base a is a strictly decreasing function.

Moreover, since $\exp : \mathbb{R} \to (0, \infty)$ is onto, we must have that for $0 < a \neq 1$, the range of the exponential function to the base a is also $\mathbb{R}$. Therefore, for $0 < a \neq 1$, the exponential function $a^{\cdot} : \mathbb{R} \to (0, \infty)$ is one-to-one and onto.

- **The law of exponentials**

(a) If $a > 0$, then $a^{x+y} = a^x a^y$ for all $x, y \in \mathbb{R}$: First we show the case when $a = e$. Let $\exp(x) = c$ and $\exp(y) = d$ or equivalently, $x = \ln c$ and $y = \ln d$. Then

$$e^{x+y} = \exp(x+y) = \exp(\ln c + \ln d) = \exp(\ln(cd)) = cd = e^x e^y \,.$$

For general $a > 0$, by the definition of exponential functions, for $x, y \in \mathbb{R}$,

$$a^{x+y} = \exp((x+y)\ln a) = e^{x\ln a + y\ln a} = e^{x\ln a} e^{y\ln a} = \exp(x\ln a)\exp(y\ln a) = a^x a^y \,.$$

(b) If $a > 0$, then $a^{x-y} = \dfrac{a^x}{a^y}$ for all $x, y \in \mathbb{R}$: Using (a), we obtain that

$$a^{x-y} a^y = a^{x-y+y} = a^x \qquad \forall\, x, y \in \mathbb{R}\,;$$

thus $a^{x-y} = \dfrac{a^x}{a^y}$ for all $x, y \in \mathbb{R}$.

(c) If $a, b > 0$, then $(ab)^x = a^x b^x$ for all $x \in \mathbb{R}$: By the definition of the exponential functions,

$$(ab)^x = e^{x\ln(ab)} = e^{x(\ln a + \ln b)} = e^{x\ln a + x\ln b} = e^{x\ln a} e^{x\ln b} = a^x b^x \,.$$

(d) If $a, b > 0$, then $\left(\dfrac{a}{b}\right)^x = \dfrac{a^x}{b^x}$ for all $x \in \mathbb{R}$: Using (b), we obtain that

$$\left(\frac{a}{b}\right)^x = e^{x\ln\frac{a}{b}} = e^{x\ln(ab^{-1})} = e^{x(\ln a - \ln b)} = \frac{e^{x\ln a}}{e^{x\ln b}} = \frac{a^x}{b^x} \,.$$

(e) If $a > 0$, then $(a^x)^y = a^{xy}$ for all $x, y \in \mathbb{R}$: Using (5.4.2),

$$(a^x)^y = e^{y\ln a^x} = e^{xy\ln a} = a^{xy} \,.$$

- **The differentiation of the exponential functions**

Theorem 5.32

$\frac{d}{dx}e^x = e^x$ for all $x \in \mathbb{R}$.

Proof. Define $f : (0, \infty) \to \mathbb{R}$ and $g : \mathbb{R} \to (0, \infty)$ by $f(x) = \ln x$ and $g(x) = \exp(x) = e^x$. Then f and g are inverse functions to each other, and the Inverse Function Differentiation implies that

$$g'(x) = \frac{1}{f'(g(x))} \qquad \forall\, x \in \mathbb{R} \text{ with } f'(g(x)) \neq 0\,.$$

Since $f'(x) = \frac{1}{x}$, $f'(g(x)) = \frac{1}{g(x)} = \exp(-x) \neq 0$ for all $x \in \mathbb{R}$; thus

$$g'(x) = g(x) \qquad \forall\, x \in \mathbb{R}\,.$$

□

Corollary 5.33

1. $\int_0^a e^x\,dx = e^a - 1$ for all $a > 0$;
2. $\int e^x\,dx = e^x + C$.

The following corollary is a direct consequence of Theorem 5.32 and the chain rule.

Corollary 5.34

Let f be a differentiable function defined on an interval I. Then

$$\frac{d}{dx}e^{f(x)} = e^x f'(x) \qquad \forall\, x \in I\,.$$

Corollary 5.35

1. For $a > 0$, $\frac{d}{dx}a^x = a^x \ln a$ for all $x \in \mathbb{R}$ $\Big($so $\int a^x\,dx = \frac{a^x}{\ln a} + C\Big)$.
2. Let r be a real number. Then $\frac{d}{dx}x^r = rx^{r-1}$ for all $x > 0$.
3. Let f, g be differentiable functions defined on an interval I. Then

$$\frac{d}{dx}|f(x)|^{g(x)} = |f(x)|^{g(x)}\Big[g'(x)\ln|f(x)| + \frac{f'(x)}{f(x)}g(x)\Big] \qquad \forall\, x \in I \text{ with } f(x) \neq 0\,.$$

Proof. The corollary holds because $a^x = e^{x\ln a}$, $x^r = e^{r\ln x}$, and $|f(x)|^{g(x)} = e^{g(x)\ln|f(x)|}$. □

Example 5.36. $\dfrac{d}{dx}e^{-\frac{3}{x}} = e^{-\frac{3}{x}}\dfrac{d}{dx}\Big(-\dfrac{3}{x}\Big) = \dfrac{3e^{-3/x}}{x^2}$ for all $x \neq 0$.

Example 5.37. Let $f : (0,\infty) \to \mathbb{R}$ be defined by $f(x) = x^x$. Then

$$f'(x) = \frac{d}{dx}e^{x\ln x} = e^{x\ln x}\frac{d}{dx}(x\ln x) = x^x(\ln x + 1)\,.$$

Example 5.38. Find the indefinite integral $\displaystyle\int 5xe^{-x^2}\,dx$.

Let $u = -x^2$. Then $du = -2xdx$; thus

$$\int 5xe^{-x^2}\,dx = -\frac{5}{2}\int e^{-x^2}(-2x)\,dx = -\frac{5}{2}\int e^u\,du = -\frac{5}{2}e^u + C = -\frac{5}{2}e^{-x^2} + C\,.$$

Example 5.39. Compute the definite integral $\displaystyle\int_{-1}^{0} e^x\cos(e^x)\,dx$.

Let $u = e^x$. Then $du = e^x\,dx$; thus

$$\int_{-1}^{0} e^x\cos(e^x)\,dx = \int_{e^{-1}}^{1}\cos u\,du = \sin u\Big|_{u=e^{-1}}^{u=1} = \sin 1 - \sin(e^{-1})\,.$$

5.4.2 The number e

By the mean value theorem for integrals, for each $x > 0$ there exists $c \in [1, 1+x]$ such that

$$\frac{\ln(1+x)}{x} = \frac{1}{x}\int_1^{1+x}\frac{1}{t}\,dt = \frac{1}{c}$$

which implies that

$$(1+x)^{\frac{1}{x}} = \exp\big(\ln(1+x)^{\frac{1}{x}}\big) = \exp\Big(\frac{\ln(1+x)}{x}\Big) = \exp\big(\frac{1}{c}\big)\,.$$

By the fact that the natural exponential function is continuous, we find that

$$\lim_{x\to 0^+}(1+x)^{\frac{1}{x}} = \lim_{x\to 0^+}\exp\big(\frac{1}{c}\big) = \lim_{c\to 1^+}\exp\big(\frac{1}{c}\big) = e\,.$$

Note that the limit above also shows that

$$e = \lim_{x\to\infty}\big(1+\frac{1}{x}\big)^x\,. \tag{5.4.3}$$

Example 5.40. Let $f(x) = (1+x)^{\frac{1}{x}} = e^{\frac{\ln(1+x)}{x}}$. Then

$$f'(x) = (1+x)^{\frac{1}{x}}\cdot\frac{\dfrac{x}{1+x} - \ln(1+x)}{x^2} = \frac{(1+x)^{\frac{1}{x}}}{x^2}\Big(1 - \frac{1}{1+x} - \ln(1+x)\Big)\,.$$

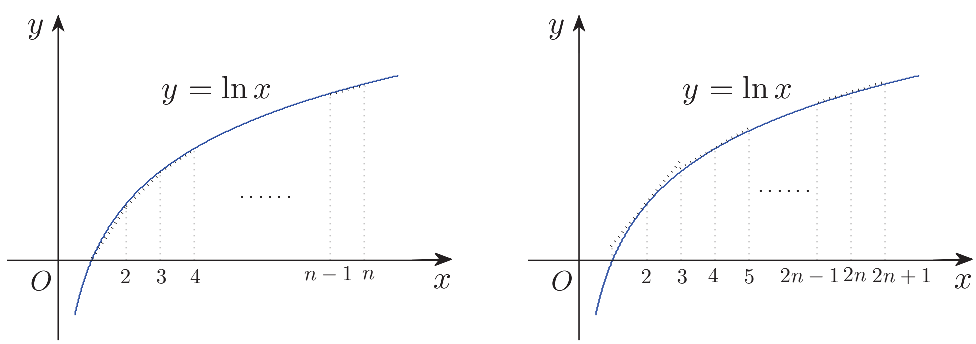
Let $g(x) = 1 - \dfrac{1}{1+x} - \ln(1+x)$. Then

$$g'(x) = \frac{1}{(1+x)^2} - \frac{1}{1+x} = \frac{-x}{(1+x)^2} < 0 \qquad \text{if } x > 0\,.$$

Therefore, $g(x) < g(0) = 0$ if $x > 0$; thus $f'(x) < 0$ for $x > 0$. This implies that f is strictly decreasing on $(0,\infty)$. This fact then implies that the function $h(x) = \big(1 + \dfrac{1}{x}\big)^x$ is strictly increasing on $(0,\infty)$.

Example 5.41. From Example 5.27 we find that for large n we have $\Big(\dfrac{n!}{n^n}\Big)^{\frac{1}{n}} \approx \dfrac{1}{e}$ which seems to imply that $n! \approx n^n e^{-n}$. This is in fact not true since the n-root of any constant, or even n, converges to 1. In this example, we try to determine how $n!$ behaves as $n \to \infty$.

Recall that the graph of $y = \ln x$ is concave downward. Therefore, we have the two figures below



(a) Under-estimate of $\displaystyle\int_1^n \ln x dx$ (b) Over-estimate of $\displaystyle\int_1^{2n+1} \ln x dx$

and find that

$$\int_1^n \ln x\,dx \geqslant \sum_{k=2}^{n} \frac{\ln k + \ln(k-1)}{2} = \frac{1}{2}\sum_{k=2}^{n} \ln k + \frac{1}{2}\sum_{k=1}^{n-1} \ln k = \ln(n!) - \frac{1}{2}\ln n$$

and

$$\int_1^{2n+1} \ln x\,dx \leqslant \sum_{k=1}^{n} 2\ln(2k) = 2n\ln 2 + 2\sum_{k=1}^{n} \ln k = 2n\ln 2 + 2\ln(n!)\,.$$

Theorem 5.26 then shows that

$$\ln(n!) - \frac{1}{2}\ln n \leqslant n\ln n - n + 1 \qquad \text{and} \qquad \big(n + \frac{1}{2}\big)\ln\big(n + \frac{1}{2}\big) + \frac{1}{2}\ln 2 - n \leqslant \ln(n!)\,.$$

As a consequence, we conclude that

$$\sqrt{2}\big(1 + \frac{1}{2n}\big)^{n+0.5} \leqslant \frac{n!}{n^{n+0.5}e^{-n}} \leqslant e \qquad \forall\, n \in \mathbb{N}\,. \tag{5.4.4}$$

Note that the function $f(x) = \left(1 + \frac{1}{2x}\right)^{x+0.5}$ is decreasing on $(0, \infty)$ since (5.2.1) shows that

$$f'(x) = f(x)\frac{d}{dx}\Big[\big(x + \frac{1}{2}\big)\ln\big(1 + \frac{1}{2x}\big)\Big] = f(x)\Big[\ln\big(1 + \frac{1}{2x}\big) - \frac{1}{2x}\Big] \leqslant 0 \quad \text{for all } x > 0\,;$$

thus (5.4.3) and (5.4.4) imply that

$$\sqrt{2e} \leqslant \frac{n!}{n^{n+0.5}e^{-n}} \leqslant e \qquad \forall\, n \in \mathbb{N}\,. \tag{5.4.5}$$

5.5 Logarithmic Functions to Bases Other than e

Definition 5.42

Let $0 < a \neq 1$ be a real number. The logarithmic function to the base a, denoted by $\log_a$, is the inverse function of the exponential function to the base a. In other words,

$$y = \log_a x \qquad \text{if and only if} \qquad a^y = x\,.$$

Theorem 5.43

Let $0 < a \neq 1$. Then $\log_a x = \dfrac{\ln x}{\ln a}$ for all $x > 0$.

Proof. Let $y = \log_a x$. Then $a^y = x$; thus (5.4.2) implies that

$$y \ln a = \ln(a^y) = \ln x$$

which shows $y = \dfrac{\ln x}{\ln a}$. □

5.5.1 Properties of logarithmic functions

- **Logarithmic laws**

The following theorem is a direct consequence of Theorem 5.16 and 5.43.

Theorem 5.44: Logarithmic properties of $y = \log_a x$

Let a, b, c be positive numbers, $a \neq 1$, and r is rational. Then

1. $\log_a 1 = 0$;
2. $\log_a(bc) = \log_a b + \log_b c$;
3. $\log_a(a^x) = x$ for all $x \in \mathbb{R}$;
4. $a^{\log_a x} = x$ for all $x > 0$;
5. $\log_a\left(\dfrac{b}{c}\right) = \log_a b - \log_a c$.

- **The change of base formula**

We have the following identity

$$\log_a c = \frac{\log_b c}{\log_b a} \qquad \forall\, a, b, c > 0, a, b \neq 1\,.$$

In fact, if $d = \log_a c$, then $c = a^d$; thus $\log_b c = d \log_b a$ which implies the identity above.

- **The differentiation of $y = \log_a x$**

By Theorem 5.43, we find that

$$\frac{d}{dx} \log_a x = \frac{1}{x \ln a} \qquad \forall\, x > 0\,.$$

Similar to Theorem 5.19, if f is differentiable on an interval I, we also have

$$\frac{d}{dx} \log_a \big|f(x)\big| = \frac{f'(x)}{f(x)\ln a} \qquad \text{for all } x \in I \text{ with } f(x) \neq 0\,.$$

5.6 Indeterminate Forms and L'Hôspital's Rule

Theorem 5.45: Cauchy Mean Value Theorem

Let $f, g : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $g'(x) \neq 0$ for all $x \in (a, b)$, then there exists $c \in (a, b)$ such that

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

Proof. Let $h : [a, b] \to \mathbb{R}$ be defined by

$$h(x) = \big(f(x) - f(a)\big)\big(g(b) - g(a)\big) - \big(f(b) - f(a)\big)\big(g(x) - g(a)\big)\,.$$

Then $h(a) = h(b) = 0$, and h is differentiable on (a, b). Then Rolle's Theorem implies that there exists $c \in (a, b)$ such that $h'(c) = 0$; thus for some $c \in (a, b)$,

$$f'(c)\big(g(b) - g(a)\big) - \big(f(b) - f(a)\big)g'(c) = 0\,.$$

Since $g'(x) \neq 0$ for all $x \in (a, b)$, the Mean Value Theorem implies that $g(b) \neq g(a)$. Therefore, the equality above implies that

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

for some $c \in (a, b)$. $\square$

Theorem 5.46: L'Hôspital's Rule

Let f, g be differentiable on (a,b), and $\dfrac{f(x)}{g(x)}$ and $\dfrac{f'(x)}{g'(x)}$ be defined on (a,b). If $\lim\limits_{x\to a^+}\dfrac{f'(x)}{g'(x)}$ exists, and one of the following conditions holds:

1. $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = 0$;
2. $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = \infty$,

then $\lim\limits_{x\to a^+}\dfrac{f(x)}{g(x)}$ exists, and

$$\lim_{x\to a^+}\frac{f(x)}{g(x)} = \lim_{x\to a^+}\frac{f'(x)}{g'(x)}\,.$$

Proof. We first prove L'Hôspital's rule for the case that $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = 0$. Define $F, G : (a,b) \to \mathbb{R}$ by

$$F(x) = \begin{cases} f(x) & \text{if } x \in (a,b)\,, \\ 0 & \text{if } x = a\,, \end{cases} \qquad \text{and} \qquad G(x) = \begin{cases} g(x) & \text{if } x \in (a,b)\,, \\ 0 & \text{if } x = a\,. \end{cases}$$

Then for all $x \in (a,b)$, F, G are continuous on the closed $[a,x]$, and differentiable on the open interval with end-points (a,x). Therefore, the Cauchy Mean Value Theorem implies that there exists a point c between a and x such that

$$\frac{f'(c)}{g'(c)} = \frac{F'(c)}{G'(c)} = \frac{F(x)-F(a)}{G(x)-G(a)} = \frac{F(x)}{G(x)} = \frac{f(x)}{g(x)}\,.$$

Since c approaches a as x approaches a, we have

$$\lim_{x\to a^+}\frac{f'(c)}{g'(c)} = \lim_{c\to a^+}\frac{f'(c)}{g'(c)} = \lim_{x\to a^+}\frac{f'(x)}{g'(x)}\,;$$

thus

$$\lim_{x\to a^+}\frac{f(x)}{g(x)} = \lim_{x\to a^+}\frac{f'(c)}{g'(c)} = \lim_{x\to a^+}\frac{f'(x)}{g'(x)}\,.$$

Next we prove L'Hôspital's rule for the case that $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^+} g(x) = \infty$. Let $L = \lim\limits_{x\to a^+}\dfrac{f'(x)}{g'(x)}$ and $\varepsilon > 0$ be given. Then there exists $\delta_1 > 0$ such that

$$\left|\frac{f'(x)}{g'(x)} - L\right| < \frac{\varepsilon}{2} \qquad \text{whenever} \quad a < x < a + \delta_1\,(< b)\,.$$

Let $d = a + \delta_1$. For $a < x < d$, the Cauchy mean value theorem implies that for some c in (x, d) such that

$$\frac{f(x) - f(d)}{g(x) - g(d)} = \frac{f'(c)}{g'(c)}\,.$$

Note that the quotient above belongs to $\left(L - \frac{\varepsilon}{2}, L + \frac{\varepsilon}{2}\right)$ (if $a < x < d$). Moreover,

$$\begin{aligned}\frac{f(x) - f(d)}{g(x) - g(d)} - \frac{f(x)}{g(x)} &= \frac{\big(f(x) - f(d)\big)g(x) - \big(g(x) - g(d)\big)f(x)}{\big(g(x) - g(d)\big)g(x)} \\ &= \frac{\big(f(x) - f(d)\big)g(d) - \big(g(x) - g(d)\big)f(x)}{\big(g(x) - g(d)\big)g(d)} = \frac{f'(c)}{g'(c)}\frac{g(d)}{g(x)} - \frac{f(d)}{g(x)}\,;\end{aligned}$$

thus

$$\Big|\frac{f(x) - f(d)}{g(x) - g(d)} - \frac{f(x)}{g(x)}\Big| \leqslant \Big(|L| + \frac{\varepsilon}{2}\Big)\Big|\frac{g(d)}{g(x)}\Big| + \Big|\frac{f(d)}{g(x)}\Big| \qquad \text{whenever} \quad a < x < d\,.$$

Since $\lim\limits_{x \to a^+} g(x) = \infty$, the right-hand side of the inequality above approaches zero as x approaches a from the right. Therefore, there exists $0 < \delta < \delta_1$, such that

$$\Big|\frac{f(x) - f(d)}{g(x) - g(d)} - \frac{f(x)}{g(x)}\Big| < \frac{\varepsilon}{2} \qquad \text{whenever} \quad a < x < a + \delta\,(< d < b)\,.$$

As a consequence, if $a < x < a + \delta$,

$$\Big|\frac{f(x)}{g(x)} - L\Big| \leqslant \Big|\frac{f(x) - f(d)}{g(x) - g(d)} - \frac{f(x)}{g(x)}\Big| + \Big|\frac{f(x) - f(d)}{g(x) - g(d)} - L\Big| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which concludes the theorem. $\square$

Remark 5.47. 1. L'Hôspital Rule can also be applied to the case when $\lim\limits_{x \to b^-}$ replaces $\lim\limits_{x \to a^+}$ in the theorem. Moreover, the one-sided limit can also be replaced by full limit $\lim\limits_{x \to c}$ if $c \in (a, b)$ (by considering L'Hôspital's Rule on (a, c) and (c, b), respectively). See Example 5.48 for more details on the full limit case.

2. L'Hôspital Rule can also be applied to limits as $x \to \infty$ or $x \to -\infty$ (and here b or a has to be changed to ∞ or $-\infty$ as well). To see this, we note that if $F(x) = f\big(\frac{1}{x}\big)$ and $G(x) = g\big(\frac{1}{x}\big)$, then either $\lim\limits_{x \to 0^+} F(x) = \lim\limits_{x \to 0^+} G(x) = 0$ or $\lim\limits_{x \to 0^+} F(x) = \lim\limits_{x \to 0^+} G(x) = \infty$; thus L'Hôspital Rule implies that

$$\lim_{x \to \infty} \frac{f'(x)}{g'(x)} = \lim_{y \to 0^+} \frac{f'\big(\frac{1}{y}\big)}{g'\big(\frac{1}{y}\big)} = \lim_{y \to 0^+} \frac{f'\big(\frac{1}{y}\big)\frac{-1}{y^2}}{g'\big(\frac{1}{y}\big)\frac{-1}{y^2}} = \lim_{y \to 0^+} \frac{F'(y)}{G'(y)} = \lim_{y \to 0^+} \frac{F(y)}{G(y)} = \lim_{x \to \infty} \frac{f(x)}{g(x)}\,.$$

3. L'Hôspital's rule only states that under suitable assumptions, if the limit of $\frac{f'(x)}{g'(x)}$ exists, so does the limit of $\frac{f(x)}{g(x)}$ and the limits are identical, but not the other way around. In other words, under the same assumptions in the statement of L'Hôspital's rule, the existence of the limit of $\frac{f(x)}{g(x)}$ does **NOT** implies the existence of the limit of $\frac{f'(x)}{g'(x)}$. For example, consider the case $f(x) = xe^{-x^{-2}}\sin(x^{-4})$ and $g(x) = e^{-x^{-2}}$. Then the Squeeze Theorem implies that $\lim\limits_{x\to 0} f(x) = \lim\limits_{x\to 0} g(x) = 0$, and

$$\lim_{x\to 0}\frac{f(x)}{g(x)} = \lim_{x\to 0} x\sin(x^{-4}) = 0\,.$$

However, since $f'(x) = \big[(1+2x^{-2})\sin(x^{-4}) - 4x^{-4}\cos(x^{-4})\big]e^{-x^{-2}}$ and $g'(x) = 2x^{-3}e^{-x^{-2}}$, we have

$$\frac{f'(x)}{g'(x)} = \frac{1}{2}(x^3+2x)\sin(x^{-4}) - \frac{2}{x}\cos(x^{-4})$$

whose limit, as x approaches 0, does not exist.

• **Indeterminate form** $\frac{0}{0}$

Example 5.48. Compute $\lim\limits_{x\to 0}\frac{e^{2x}-1}{x}$.

Let $f(x) = e^{2x}-1$ and $g(x) = x$. Then f, g are differentiable on $(0,1)$ and $g(x) \neq 0, g'(x) \neq 0$ for all $x \in (0,1)$. Moreover,

$$\lim_{x\to 0^+}\frac{f'(x)}{g'(x)} = \lim_{x\to 0^+}\frac{2e^{2x}}{1} = 2$$

and $\lim\limits_{x\to 0^+} f(x) = \lim\limits_{x\to 0^+} g(x) = 0$. Therefore, L'Hôspital's Rule implies that

$$\lim_{x\to 0^+}\frac{f(x)}{g(x)} = \lim_{x\to 0^+}\frac{f'(x)}{g'(x)} = 2\,.$$

Similarly, by the fact that

1. f, g are differentiable on $(-1,0)$ and $g(x) \neq 0, g'(x) \neq 0$ for all $x \in (-1,0)$,

2. $\lim\limits_{x\to 0^-}\frac{f'(x)}{g'(x)} = \lim\limits_{x\to 0^-}\frac{2e^{2x}}{1} = 2$,

3. $\lim\limits_{x\to 0^+} f(x) = \lim\limits_{x\to 0^+} g(x) = 0$,

L'Hôspital's Rule implies that $\lim\limits_{x\to 0^-}\dfrac{f(x)}{g(x)} = \lim\limits_{x\to 0^-}\dfrac{f'(x)}{g'(x)} = 2$. Theorem 1.33 then shows that $\lim\limits_{x\to 0}\dfrac{f(x)}{g(x)} = 2$ exists.

From the discussion in Example 5.48, using L'Hôspital's Rule in Theorem 5.46 we deduce the following L'Hôspital's Rule for the full limit case.

Theorem 5.46*

Let $a < c < b$, and f, g be differentiable functions on $(a,b)\backslash\{c\}$. Assume that $g'(x) \neq 0$ for all $x \in (a,b)\backslash\{c\}$. If the limit of $\dfrac{f(x)}{g(x)}$ as x approaches c produces the indeterminate form $\dfrac{0}{0}$ $\big($or $\dfrac{\infty}{\infty}\big)$; that is, $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} g(x) = 0$ $\big($or $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} g(x) = \infty\big)$, then

$$\lim_{x\to c}\frac{f(x)}{g(x)} = \lim_{x\to c}\frac{f'(x)}{g'(x)}$$

provided the limit on the right exists.

- **Indeterminate form** $\dfrac{\infty}{\infty}$

Example 5.49. In this example we compute $\lim\limits_{x\to\infty}\dfrac{\ln x}{x}$. Note that $\lim\limits_{x\to\infty}\dfrac{\frac{d}{dx}\ln x}{\frac{d}{dx}x} = \lim\limits_{x\to\infty}\dfrac{1}{x} = 0$, so L'Hôspital's Rule implies that

$$\lim_{x\to\infty}\frac{\ln x}{x} = \lim_{x\to\infty}\frac{\frac{d}{dx}\ln x}{\frac{d}{dx}x} = 0\,.$$

In fact, the logarithmic function $y = \ln x$ grows slower than any power function; that is,

$$\lim_{x\to\infty}\frac{\ln x}{x^p} = 0 \qquad \forall\, p > 0\,.$$

To see this, note that $\lim\limits_{x\to\infty}\dfrac{\frac{d}{dx}\ln x}{\frac{d}{dx}x^p} = \lim\limits_{x\to\infty}\dfrac{\frac{1}{x}}{px^{p-1}} = \dfrac{1}{p}\lim\limits_{x\to\infty}\dfrac{1}{x^p} = 0$, so L'Hôspital's Rule implies that

$$\lim_{x\to\infty}\frac{\ln x}{x^p} = \lim_{x\to\infty}\frac{\frac{d}{dx}\ln x}{\frac{d}{dx}x^p} = 0\,.$$

- **Indeterminate form** $0\cdot\infty$

Example 5.50. Compute $\lim\limits_{x\to\infty} e^{-x}\sqrt{x}$. Rewrite $e^{-x}\sqrt{x}$ as $\dfrac{\sqrt{x}}{e^x}$ and note that

$$\lim_{x\to\infty}\frac{\frac{d}{dx}\sqrt{x}}{\frac{d}{dx}e^x} = \lim_{x\to\infty}\frac{\frac{1}{2\sqrt{x}}}{e^x} = \lim_{x\to\infty}\frac{1}{2\sqrt{x}e^x} = 0\,.$$

Therefore, L'Hôspital's Rule implies that

$$\lim_{x\to\infty}\frac{\sqrt{x}}{e^x}=\lim_{x\to\infty}\frac{\frac{d}{dx}\sqrt{x}}{\frac{d}{dx}e^x}=0\,.$$

In fact, the natural exponential function $y=e^x$ grows faster than any power function; that is,

$$\lim_{x\to\infty}\frac{x^p}{e^x}=0\qquad\forall\,p>0\,.$$

The proof is left as an exercise.

- **Indeterminate form** 1^∞

Example 5.51. In this example we compute $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}}$. Rewrite $(1+x)^{\frac{1}{x}}$ as $e^{\frac{\ln(1+x)}{x}}$. If the limit $\lim\limits_{x\to 0}\dfrac{\ln(1+x)}{x}$ exists, then the continuity of the exponential function implies that

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}}=\exp\big(\lim_{x\to 0}\frac{\ln(1+x)}{x}\big)\,.$$

Nevertheless, since $\lim\limits_{x\to 0}\ln(1+x)=0$, $\lim\limits_{x\to 0}x=0$ and

$$\lim_{x\to 0}\frac{\frac{d}{dx}\ln(1+x)}{\frac{d}{dx}x}=\lim_{x\to 0}\frac{1}{1+x}=1$$

L'Hôspital's Rule implies that

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=\lim_{x\to 0}\frac{\frac{d}{dx}\ln(1+x)}{\frac{d}{dx}x}=1\,;$$

thus $\lim\limits_{x\to 0}(1+x)^{\frac{1}{x}}=\exp(1)=e$.

- **Indeterminate form** 0^0

Example 5.52. In this example we compute $\lim\limits_{x\to 0^+}(\sin x)^x$. When $\sin x>0$, we have

$$(\sin x)^x=e^{x\ln\sin x}=e^{\frac{\ln\sin x}{1/x}}\,.$$

Since

$$\lim_{x\to 0^+}\frac{\frac{d}{dx}\ln\sin x}{\frac{d}{dx}\frac{1}{x}}=\lim_{x\to 0^+}\frac{\frac{\cos x}{\sin x}}{-\frac{1}{x^2}}=-\lim_{x\to 0^+}\frac{x}{\sin x}x\cos x=0\,,$$

by L'Hôspital's Rule and the continuity of the natural exponential function we find that

$$\lim_{x\to 0^+}(\sin x)^x=\lim_{x\to 0^+}e^{\frac{\ln\sin x}{1/x}}=e^0=1\,.$$

- **Indeterminate form** $\infty - \infty$

Example 5.53. Compute $\lim_{x\to 1+}\left(\frac{1}{\ln x} - \frac{1}{x-1}\right)$.

Rewrite $\frac{1}{\ln x} - \frac{1}{x-1} = \frac{x-1-\ln x}{(x-1)\ln x}$ and note that the right-hand side produces indeterminate form $\frac{0}{0}$ as x approaches from the right. Also note that

$$\frac{\frac{d}{dx}(x-1-\ln x)}{\frac{d}{dx}(x-1)\ln x} = \frac{1-\frac{1}{x}}{\ln x + \frac{x-1}{x}} = \frac{x-1}{x\ln x + x - 1}$$

which, as x approaches 1 from the right, again produces indeterminate form $\frac{0}{0}$. In order to find the limit of the right-hand side we compute

$$\lim_{x\to 1^+} \frac{\frac{d}{dx}(x-1)}{\frac{d}{dx}(x\ln x + x - 1)} = \lim_{x\to 1^+} \frac{1}{\ln x + 1 + 1} = \frac{1}{2}\,;$$

thus L'Hôspital's Rule implies that

$$\lim_{x\to 1^+} \frac{x-1}{x\ln x + x - 1} = \lim_{x\to 1^+} \frac{\frac{d}{dx}(x-1)}{\frac{d}{dx}(x\ln x + x - 1)} = \frac{1}{2}\,.$$

This in turm shows that

$$\lim_{x\to 1^+} \frac{x-1-\ln x}{(x-1)\ln x} = \lim_{x\to 1^+} \frac{\frac{d}{dx}(x-1-\ln x)}{\frac{d}{dx}(x-1)\ln x} = \lim_{x\to 1^+} \frac{x-1}{x\ln x + x - 1} = \frac{1}{2}\,.$$

5.7 The Inverse Trigonometric Functions: Differentiation

Definition 5.54

The arcsin, arccos, and arctan functions are the inverse functions of the function $f : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to \mathbb{R}$, $g : [0,\pi] \to \mathbb{R}$, and $h : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$, respectively, where $f(x) = \sin x$, $g(x) = \cos x$ and $h(x) = \tan x$. In other words,

1. $y = \arcsin x$ if and only if $\sin y = x$, where $-\frac{\pi}{2} \leqslant y \leqslant \frac{\pi}{2}$, $-1 \leqslant x \leqslant 1$.
2. $y = \arccos x$ if and only if $\cos y = x$, where $0 \leqslant y \leqslant \pi$, $-1 \leqslant x \leqslant 1$.
3. $y = \arctan x$ if and only if $\tan y = x$, where $-\frac{\pi}{2} < y < \frac{\pi}{2}$, $-\infty < x < \infty$.

Remark 5.55. Since arcsin, arccos and arctan look like the inverse function of sin, cos and tan, respectively, often times we also write arcsin as $\sin^{-1}$, arccos as $\cos^{-1}$, and arctan as $\tan^{-1}$.

Example 5.56. $\arcsin\frac{1}{2} = \frac{\pi}{6}$, $\arccos\big(\frac{-\sqrt{2}}{2}\big) = \frac{3\pi}{4}$, and $\arctan 1 = \frac{\pi}{4}$.

Example 5.57. Suppose that $y = \arcsin x$. Then $y \in \big[-\frac{\pi}{2}, \frac{\pi}{2}\big]$ which implies that $\cos y \geqslant 0$. Therefore, by the fact that $\sin^2 y + \cos^2 y = 1$, we have

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2} \qquad \text{if} \quad y = \arcsin x\,.$$

In other words, $\cos(\arcsin x) = \sqrt{1 - x^2}$.

Similarly, if $y = \arccos x$, then $y \in (0, \pi)$ which implies that $\sin y \geqslant 0$. Therefore,

$$\sin y = \sqrt{1 - \cos^2 y} = \sqrt{1 - x^2} \qquad \text{if} \quad y = \arccos x$$

or equivalently, $\sin(\arccos x) = \sqrt{1 - x^2}$.

Example 5.58. Suppose that $y = \arctan x$ for some $x \in \mathbb{R}$. Then $y \in \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$ which implies that $\cos y > 0$. Therefore,

$$\cos y = \frac{1}{\sec y} = \frac{1}{\sqrt{1 + \tan^2 y}} = \frac{1}{\sqrt{1 + x^2}}\,.$$

As for $\sin y$, we note that $y > 0$ if and only if $x > 0$; thus $\sin y = \dfrac{x}{\sqrt{1 + x^2}}$ $\Big($instead of $\dfrac{-x}{\sqrt{1 + x^2}}\Big)$. Therefore,

$$\sin(\arctan x) = \frac{x}{\sqrt{1 + x^2}} \quad \text{and} \quad \cos(\arctan x) = \frac{1}{\sqrt{1 + x^2}}\,.$$

Theorem 5.59: Differentiation of Inverse Trigonometric Functions

1. $\dfrac{d}{dx}\arcsin x = \dfrac{1}{\sqrt{1 - x^2}}$ for all $-1 < x < 1$.
2. $\dfrac{d}{dx}\arccos x = -\dfrac{1}{\sqrt{1 - x^2}}$ for all $-1 < x < 1$.
3. $\dfrac{d}{dx}\arctan x = \dfrac{1}{1 + x^2}$ for all $x \in \mathbb{R}$.

Proof. By Inverse Function Differentiation,

$$\frac{d}{dx}\arcsin x = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}} \qquad \forall\, x \in (-1,1)\,,$$
$$\frac{d}{dx}\arccos x = \frac{1}{-\sin(\arccos x)} = -\frac{1}{\sqrt{1-x^2}} \qquad \forall\, x \in (-1,1)\,,$$

and

$$\frac{d}{dx}\arctan x = \frac{1}{\sec^2(\arctan x)} = \frac{1}{1+\tan^2(\arctan x)} = \frac{1}{1+x^2} \qquad \forall\, x \in \mathbb{R}\,. \qquad \square$$

Remark 5.60. By Theorem 5.59,

$$\frac{d}{dx}\big(\arcsin x + \arccos x\big) = \frac{1}{\sqrt{1-x^2}} - \frac{1}{\sqrt{1-x^2}} = 0 \qquad \forall\, -1 < x < 1\,.$$

Therefore, the function $y = \arcsin x + \arccos x$ is constant on the interval $(-1,1)$. The constant can be obtained by testing with $x = 0$ and we find that

$$\arcsin x + \arccos x = \frac{\pi}{2} \qquad \forall\, x \in [-1,1]\,, \tag{5.7.1}$$

where the value of the left-hand side at $x = \pm 1$ are computed separately.

Example 5.61. Find the derivative of $y = \arcsin x + x\sqrt{1-x^2}$.

By Theorem 5.59 and the chain rule, for $-1 < x < 1$ we have

$$\frac{dy}{dx} = \frac{1}{\sqrt{1-x^2}} + \sqrt{1-x^2} - x \cdot \frac{1}{2}(1-x^2)^{-\frac{1}{2}}(2x) = 2\sqrt{1-x^2}\,.$$

Example 5.62. Find the derivative of $y = \arctan\sqrt{x}$.

By the chain rule,

$$\frac{dy}{dx} = \frac{1}{1+\sqrt{x}^2}\frac{d}{dx}\sqrt{x} = \frac{1}{1+x}\frac{1}{2\sqrt{x}} = \frac{1}{2\sqrt{x}(1+x)}\,.$$

5.8 Inverse Trigonometric Functions: Integration

Theorem 5.63

Let a be a positive real number. Then

1. $\displaystyle\int \frac{dx}{\sqrt{a^2-x^2}} = \arcsin\frac{x}{a} + C.$ 2. $\displaystyle\int \frac{dx}{a^2+x^2} = \frac{1}{a}\arctan\frac{x}{a} + C.$

Proof. 1. Let $x = a\sin u$. Then $dx = a\cos u du$; thus

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \int \frac{a\cos u}{\sqrt{a^2(1-\sin^2 u)}}\, du = \int du = u + C = \arcsin\frac{x}{a} + C\,.$$

2. Let $x = a\tan u$. Then $dx = a\sec^2 u du$; thus

$$\int \frac{dx}{a^2 + x^2} = \int \frac{a\sec^2 u}{a^2(1+\tan^2 u)}\, du = \frac{1}{a}\int du = \frac{u}{a} + C = \frac{1}{a}\arctan\frac{x}{a} + C\,. \qquad \square$$

Example 5.64. Find the indefinite integral $\displaystyle\int \frac{dx}{\sqrt{x^2 - a^2}}$, where $a > 0$ is a constant.

Let $x = a\sec u$. Then $dx = a\sec u\tan u du$; thus

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2 - a^2}} &= \int \frac{a\sec u\tan u}{\sqrt{a^2(\sec^2 u - 1)}}\, du = \int \sec u\, du = \ln|\sec u + \tan u| + C\\ &= \ln\Big|\frac{x}{a} + \frac{\sqrt{x^2-a^2}}{a}\Big| + C = \ln\big|x + \sqrt{x^2-a^2}\big| + C\,.\end{aligned}$$

Example 5.65. Find the indefinite integral $\displaystyle\int \frac{dx}{\sqrt{x^2 + a^2}}$, where $a > 0$ is a constant.

Let $x = a\tan u$. Then $dx = a\sec^2 u du$; thus

$$\begin{aligned}\int \frac{dx}{\sqrt{a^2 + x^2}} &= \int \frac{a\sec^2 u}{\sqrt{a^2(\tan^2 u + 1)}}\, du = \int \sec u\, du = \ln|\sec u + \tan u| + C\\ &= \ln\Big|\frac{\sqrt{x^2+a^2}}{a} + \frac{x}{a}\Big| + C = \ln\big|x + \sqrt{x^2+a^2}\big| + C\,.\end{aligned}$$

Example 5.66. Find the indefinite integral $\displaystyle\int \frac{dx}{x\sqrt{x^2 - a^2}}$, where $a > 0$ is a constant.

Let $x = a\sec u$. Then $dx = a\sec u\tan u$; thus

$$\int \frac{dx}{x\sqrt{x^2-a^2}} = \int \frac{a\sec u\tan u}{a\sec u\sqrt{a^2(\sec^2 u - 1)}} = \frac{1}{a}\int du = \frac{u}{a} + C\,.$$

If $x = a\sec u$, then $\tan u = \dfrac{\sqrt{x^2-a^2}}{a}$; thus $u = \arctan\dfrac{\sqrt{x^2-a^2}}{a}$ which implies that

$$\int \frac{dx}{x\sqrt{x^2-a^2}} = \frac{1}{a}\arctan\frac{\sqrt{x^2-a^2}}{a} + C\,.$$

Example 5.67. Find the indefinite integral $\displaystyle\int \frac{dx}{\sqrt{e^{2x}-1}}$.

Let $u = e^x$. Then $du = e^x dx$; thus $dx = \dfrac{du}{u}$ which implies that

$$\int \frac{dx}{\sqrt{e^{2x}-1}} = \int \frac{du}{u\sqrt{u^2-1}} = \arctan\sqrt{u^2-1} + C = \arctan\sqrt{e^{2x}-1} + C\,.$$

Example 5.68. Find the indefinite integral $\displaystyle\int \frac{x+2}{\sqrt{4-x^2}}\,dx$.

Let $x = 2\sin u$. Then $dx = 2\cos u du$; thus

$$\begin{aligned}\int \frac{x+2}{\sqrt{4-x^2}}\,dx &= \int \frac{2\sin u+2}{\sqrt{4-4\sin^2 u}}\cdot 2\cos u\,du = \int (2\sin u+2)\,du = 2u - 2\cos u + C\\ &= 2\arcsin\frac{x}{2} - 2\sqrt{1-\left(\frac{x}{2}\right)^2} + C = 2\arcsin\frac{x}{2} - \sqrt{4-x^2} + C\,.\end{aligned}$$

Example 5.69. Find the indefinite integral $\displaystyle\int \frac{dx}{x^2-4x+7}$.

First we complete the square and obtain that $x^2-4x+7 = (x-2)^2+3$. Let $x-2 = \sqrt{3}\tan u$. Then $dx = \sqrt{3}\sec^2 u du$; thus

$$\int \frac{dx}{x^2-4x+7} = \int \frac{\sqrt{3}\sec^2 u}{3\tan^2 u+3}\,du = \frac{1}{\sqrt{3}}\int du = \frac{1}{\sqrt{3}}u + C = \frac{1}{\sqrt{3}}\arctan\frac{x-2}{\sqrt{3}} + C\,.$$

Example 5.70. Find the indefinite integral $\displaystyle\int \sqrt{\frac{1-x}{1+x}}\,dx$.

Note that the integrand can be rewritten as $\dfrac{1-x}{\sqrt{1-x^2}}$. Therefore,

$$\begin{aligned}\int \sqrt{\frac{1-x}{1+x}}\,dx &= \int \frac{1-x}{\sqrt{1-x^2}}\,dx = \int \frac{1}{\sqrt{1-x^2}}\,dx - \int \frac{x}{\sqrt{1-x^2}}\,dx\\ &= \arcsin x + \sqrt{1-x^2} + C\,.\end{aligned}$$

Example 5.71. In this example, we compute $\displaystyle\int \arcsin x\,dx$. Note the by the substitution $x = \sin u$,

$$\int \arcsin x\,dx = \int u\cos u\,du\,;$$

thus it suffices to compute the anti-derivative of the function $y = x\cos x$. We first compute the definite integral $\displaystyle\int_0^a x\cos x\,dx$.

By Example 4.12, for $0 < x < \pi$ we have

$$\sum_{i=1}^{n} \sin(ix) = \frac{1}{2\sin\frac{x}{2}}\Big[\cos\frac{x}{2} - \cos\big((n+\tfrac{1}{2})x\big)\Big]\,.$$

Therefore, if $0 < x < \pi$,

$$\begin{aligned}
\sum_{i=1}^{n} i\cos(ix) = \frac{d}{dx}\sum_{i=1}^{n}\sin(ix) &= \frac{d}{dx}\frac{1}{2\sin\frac{x}{2}}\Big[\cos\frac{x}{2} - \cos\big((n+\frac{1}{2})x\big)\Big] \\
&= \frac{-\cos\frac{x}{2}}{4\sin^2\frac{x}{2}}\Big[\cos\frac{x}{2} - \cos\big((n+\frac{1}{2})x\big)\Big] \\
&\quad + \frac{1}{2\sin\frac{x}{2}}\Big[-\frac{1}{2}\sin\frac{x}{2} + (n+\frac{1}{2})\sin\big((n+\frac{1}{2})x\big)\Big].
\end{aligned}$$

By partitioning $[0, a]$ into n sub-intervals with equal length, the Riemann sum of $y = x\cos x$ for this partition given by the right end-point rule is

$$I_n = \sum_{i=1}^{n}\frac{ia}{n}\cos\frac{ia}{n}\frac{a}{n} = \frac{a^2}{n^2}\sum_{i=1}^{n} i\cos\frac{ia}{n}.$$

Letting $r = \dfrac{a}{2n}$, we find that

$$\begin{aligned}
I_n &= 4r^2\sum_{i=1}^{n} i\cos(2ir) \\
&= \frac{-r^2\cos r}{\sin^2 r}\Big[\cos r - \cos(a+r)\Big] + \frac{r}{\sin r}\Big[-r\sin r + (a+r)\sin(a+r)\Big]
\end{aligned}$$

which, by the fact that $\dfrac{\sin x}{x} \to 1$ as $x \to 0$ and $r \to 0$ as $n \to \infty$, implies that

$$\int_0^a x\cos x\,dx = \lim_{n\to\infty} I_n = -(1-\cos a) + a\sin a = a\sin a + \cos a - 1.$$

The identity above further implies that

$$\int x\cos x\,dx = x\sin x + \cos x + C;$$

thus with the substitution $x = \sin u$,

$$\int \arcsin x\,dx = \int u\cos u\,du = u\sin u + \cos u + C = x\arcsin x + \sqrt{1-x^2} + C.$$

Using (5.7.1), we also find that

$$\begin{aligned}
\int \arccos x\,dx = \int \big(\frac{\pi}{2} - \arcsin x\big)\,dx &= \frac{\pi}{2}x - x\arcsin x - \sqrt{1-x^2} + C \\
&= x\big(\frac{\pi}{2} - \arcsin x\big) - \sqrt{1-x^2} + C = x\arccos x - \sqrt{1-x^2} + C.
\end{aligned}$$

5.9 Hyperbolic Functions

Definition 5.72: Hyperbolic Functions

The hyperbolic functions sinh, cosh, tanh, coth, sech and csch are defined by

$$\sinh x = \frac{e^x - e^{-x}}{2}\,, \qquad \cosh x = \frac{e^x + e^{-x}}{2}\,, \qquad \tanh x = \frac{\sinh x}{\cosh x}\,,$$
$$\coth x = \frac{1}{\tanh x}\,, \qquad \operatorname{sech} x = \frac{1}{\cosh x}\,, \qquad \operatorname{csch} x = \frac{1}{\sinh x}\,.$$

Motivation: The Euler identity provides the following relation

$$e^{ix} = \cos x + i \sin x \qquad \forall\, x \in \mathbb{R}\,. \tag{5.9.1}$$

This implies that

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \qquad \text{and} \qquad \cos x = \frac{e^{ix} + e^{-ix}}{2} \qquad \forall\, x \in \mathbb{R}\,.$$

For a complex number $z = x + iy$, where $x, y \in \mathbb{R}$, define $\sin z$ and $\cos z$ by

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} = \frac{e^{i(x+iy)} - e^{-i(x+iy)}}{2i} = \frac{e^{-y}e^{ix} - e^{y}e^{-ix}}{2i}\,, \tag{5.9.2a}$$
$$\cos z = \frac{e^{iz} + e^{-iz}}{2} = \frac{e^{i(x+iy)} + e^{-i(x+iy)}}{2} = \frac{e^{-y}e^{ix} + e^{y}e^{-ix}}{2}\,. \tag{5.9.2b}$$

Then on the imaginary axis, we have

$$\sin(iy) = \frac{e^{-y} - e^{y}}{2i} = i \sinh y \qquad \forall\, y \in \mathbb{R}\,, \tag{5.9.3a}$$
$$\cos(iy) = \frac{e^{-y} + e^{y}}{2} = \cosh y \qquad \forall\, y \in \mathbb{R}\,. \tag{5.9.3b}$$

The hyperbolic functions, roughly speaking, can be viewed as trigonometric functions on the imaginary axis (by ignoring i in the output).

We also note that by definition, for $z = x + iy$ with $x, y \in \mathbb{R}$,

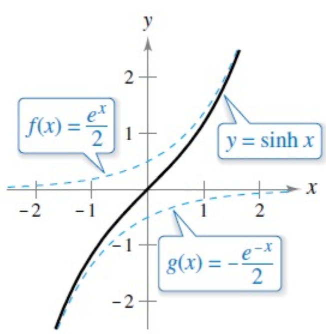
$$\sin^2 z + \cos^2 z = \frac{e^{-2y}e^{2ix} - 2 + e^{2y}e^{-2ix}}{-4} + \frac{e^{-2y}e^{2ix} + 2 + e^{2y}e^{-2ix}}{4} = 1\,.$$

Moreover, if z_1, z_2 are complex numbers,

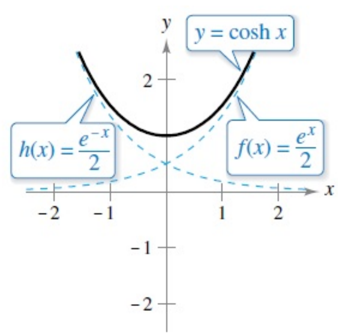
$$
\begin{aligned}
&\cos z_1 \cos z_2 - \sin z_1 \sin z_2 \\
&\quad = \frac{e^{iz_1} + e^{-iz_1}}{2} \cdot \frac{e^{iz_2} + e^{-iz_2}}{2} - \frac{e^{iz_1} - e^{-iz_1}}{2i} \cdot \frac{e^{iz_2} - e^{-iz_2}}{2i} \\
&\quad = \frac{e^{i(z_1+z_2)} + e^{i(z_1-z_2)} + e^{i(z_2-z_1)} + e^{-i(z_1+z_2)}}{4} + \frac{e^{i(z_1+z_2)} - e^{i(z_1-z_2)} - e^{i(z_2-z_1)} + e^{-i(z_1+z_2)}}{4} \\
&\quad = \frac{e^{i(z_1+z_2)} + e^{i(z_1+z_2)}}{2} = \cos(z_1 + z_2)\,.
\end{aligned}
$$

The above computations show that trigonometric identities are still valid even for complex arguments.

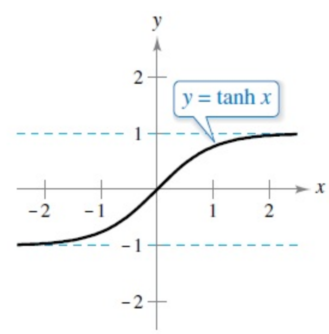
- **The graph of hyperbolic functions**



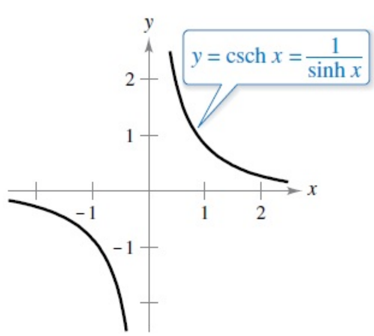
Domain: $(-\infty, \infty)$
Range: $(-\infty, \infty)$



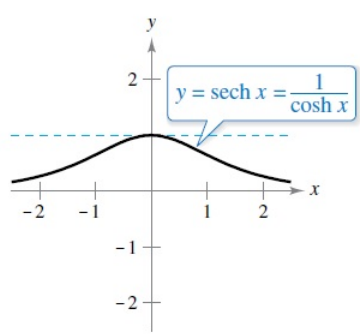
Domain: $(-\infty, \infty)$
Range: $[1, \infty)$



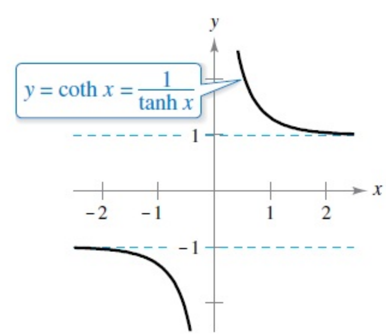
Domain: $(-\infty, \infty)$
Range: $(-1, 1)$



Domain: $(-\infty, 0) \cup (0, \infty)$
Range: $(-\infty, 0) \cup (0, \infty)$



Domain: $(-\infty, \infty)$
Range: $(0, 1]$



Domain: $(-\infty, 0) \cup (0, \infty)$
Range: $(-\infty, -1) \cup (1, \infty)$

Theorem 5.73: Hyperbolic identities

1. $\cosh^2 x - \sinh^2 x = 1$; 2. $\tanh^2 x + \operatorname{sech}^2 x = 1$;
3. $\coth^2 x - \operatorname{csch}^2 x = 1$;
4. $\sinh(x \pm y) = \sinh x \cosh y \pm \sinh y \cosh x$;
5. $\cosh(x \pm y) = \cosh x \cosh y \pm \sinh x \sinh y$;
6. $\sinh^2 = \dfrac{-1 + \cosh(2x)}{2}$; $\cosh^2 x = \dfrac{1 + \cosh(2x)}{2}$;
7. $\sinh(2x) = 2 \sinh x \cosh x$; $\cosh(2x) = \cosh^2 x + \sinh^2 x$.

Remark 5.74. By the definition (5.9.2), one can easily check that $\sin^2 z + \cos^2 z = 1$ for all complex z and this further implies that

$$1 = \sin^2(iy) + \cos^2(iy) = (i \sinh y)^2 + \cosh^2 y = \cosh^2 y - \sinh^2 y \qquad \forall\, y \in \mathbb{R}\,.$$

All the other hyperbolic identities can be memorized/derived in the same way.

Theorem 5.75: Differentiation and integration of hyperbolic functions

1. $\dfrac{d}{dx} \sinh x = \cosh x$; $\displaystyle\int \cosh x \, dx = \sinh x + C$;
2. $\dfrac{d}{dx} \cosh x = \sinh x$; $\displaystyle\int \sinh x \, dx = \cosh x + C$;
3. $\dfrac{d}{dx} \tanh x = \operatorname{sech}^2 x$; $\displaystyle\int \operatorname{sech}^2 x \, dx = \tanh x + C$;
4. $\dfrac{d}{dx} \coth x = -\operatorname{csch}^2 x$; $\displaystyle\int \operatorname{csch}^2 x \, dx = -\coth x + C$;
5. $\dfrac{d}{dx} \operatorname{sech} x = -\operatorname{sech} x \tanh x$; $\displaystyle\int \operatorname{sech} x \tanh x \, dx = -\operatorname{sech} x + C$;
6. $\dfrac{d}{dx} \operatorname{csch} x = -\operatorname{csch} x \coth x$; $\displaystyle\int \operatorname{csch} x \coth x \, dx = \operatorname{csch} x + C$;
7. $\displaystyle\int \tanh x \, dx = \ln \cosh x + C$;
8. $\displaystyle\int \operatorname{sech} x \, dx = 2 \arctan e^x + C$.

Proof. We only prove 7 and 8. By Theorem 5.22, it is easy to see that

$$\int \tanh x \, dx = \int \frac{\sinh x}{\cosh x} \, dx = \int \frac{\frac{d}{dx}\cosh x}{\cosh x} \, dx = \ln \cosh x + C \,,$$

so we focus on 8.

Let $u = e^x$. Then $du = e^x \, dx$ or equivalently, $\frac{du}{u} = dx$; thus

$$\int \operatorname{sech} x \, dx = \int \frac{2}{u + u^{-1}} \cdot \frac{du}{u} = \int \frac{2}{u^2 + 1} \, du = 2 \arctan u + C = 2 \arctan e^x + C \,. \qquad \square$$

Remark 5.76. Assuming that one knows that $\frac{d}{dx} f(ix) = i f'(ix)$ (that is, the rule of differentiation $\frac{d}{dx} f(ax) = a f'(ax)$ can also be applied for complex a), we have

$$\begin{aligned} \frac{d}{dx} \tanh x &= \frac{d}{dx} \frac{\sinh x}{\cosh x} = \frac{1}{i} \frac{d}{dx} \frac{\sin(ix)}{\cos(ix)} = \frac{1}{i} \tan(ix) = \sec^2(ix) \\ &= \frac{1}{\cos^2(ix)} = \frac{1}{\cosh^2 x} = \cosh^2 x \,. \end{aligned}$$

All the other derivatives of hyperbolic functions can be memorized/derived in the same way.

• Inverse hyperbolic functions

Similar to inverse trigonometric functions, we can also talk about the inverse function of hyperbolic functions. Note that

$$\begin{aligned} \sinh &: (-\infty, \infty) \xrightarrow[\textit{onto}]{1-1} (-\infty, \infty) \,, \\ \tanh &: (-\infty, \infty) \xrightarrow[\textit{onto}]{1-1} (-1, 1) \,, \end{aligned}$$

while

$$\begin{aligned} &\cosh : (-\infty, \infty) \to [1, \infty) \text{ is onto but not one-to-one} \,, \\ &\operatorname{sech} : (-\infty, \infty) \to (0, 1] \text{ is onto but not one-to-one} \,. \end{aligned}$$

We first find the inverse function of sinh and tanh.

1. Let $y = \sinh x = \frac{e^x - e^{-x}}{2}$. Then $e^{2x} - 2ye^x - 1 = 0$; thus by the fact that $e^x > 0$,

$$e^x = \frac{2y + \sqrt{4y^2 + 4}}{2} = y + \sqrt{y^2 + 1}$$

   which further implies that $x = \ln(y + \sqrt{y^2 + 1})$. Therefore,

$$\sinh^{-1} x = \ln(x + \sqrt{x^2 + 1}) \qquad \forall \, x \in \mathbb{R} \,.$$

2. Let $y = \tanh x = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$. Then $e^{2x}(1-y) = 1+y$; thus $x = \dfrac{1}{2}\ln\dfrac{1+y}{1-y}$. Therefore,

$$\tanh^{-1} x = \frac{1}{2}\ln\frac{1+x}{1-x} \qquad \forall\, x \in (-1,1)\,.$$

To find the inverse of cosh, we note that $\cosh : [0,\infty) \xrightarrow[onto]{1-1} [1,\infty)$. Let $x \geqslant 0$ and $y = \cosh x = \dfrac{e^x + e^{-x}}{2}$. Then $e^{2x} - 2ye^x + 1 = 0$ which implies that

$$e^x = y + \sqrt{y^2-1}\,.$$

As a consequence,

$$\cosh^{-1} x = \ln(x + \sqrt{x^2-1}) \qquad \forall\, x \in [1,\infty)\,.$$

Since $\operatorname{sech} x = \dfrac{1}{\cosh x}$, we find that

$$\operatorname{sech}^{-1} x = \cosh^{-1}\frac{1}{x} = \ln\Big(\frac{1}{x} + \sqrt{\frac{1}{x^2}-1}\Big) = \ln\frac{1+\sqrt{1-x^2}}{x}\,.$$

We summarize these inverse hyperbolic functions in the following

Theorem 5.77

1. $\sinh^{-1} x = \ln(x + \sqrt{x^2+1}) \qquad \forall\, x \in \mathbb{R}$.
2. $\cosh^{-1} x = \ln(x + \sqrt{x^2-1}) \qquad \forall\, x \in [1,\infty)$.
3. $\tanh^{-1} x = \dfrac{1}{2}\ln\dfrac{1+x}{1-x} \qquad \forall\, x \in (-1,1)$.
4. $\operatorname{sech}^{-1} x = \ln\dfrac{1+\sqrt{1-x^2}}{x} \qquad \forall\, x \in (0,1]$.

• **Differentiation and integration of inverse hyperbolic functions**

Theorem 5.78

1. $\dfrac{d}{dx}\sinh^{-1} x = \dfrac{1}{\sqrt{x^2+1}}; \qquad \displaystyle\int \frac{dx}{\sqrt{x^2+1}} = \sinh^{-1} x + C;$
2. $\dfrac{d}{dx}\cosh^{-1} x = \dfrac{1}{\sqrt{x^2-1}}; \qquad \displaystyle\int \frac{dx}{\sqrt{x^2-1}} = \cosh^{-1} x + C;$
3. $\dfrac{d}{dx}\tanh^{-1} x = \dfrac{1}{1-x^2}; \qquad \displaystyle\int \frac{dx}{1-x^2} = \tanh^{-1} x + C.$

Proof. By the chain rule,

$$\frac{d}{dx}\sinh^{-1}x = \frac{1}{x+\sqrt{x^2+1}}\frac{d}{dx}(x+\sqrt{x^2+1}) = \frac{1}{\sqrt{x^2+1}}\,,$$
$$\frac{d}{dx}\cosh^{-1}x = \frac{1}{x+\sqrt{x^2-1}}\frac{d}{dx}(x+\sqrt{x^2-1}) = \frac{1}{\sqrt{x^2-1}}\,,$$

as well as

$$\frac{d}{dx}\tanh^{-1}x = \frac{1}{2}\Big[\frac{1}{1+x}+\frac{1}{1-x}\Big] = \frac{1}{1-x^2}\,.$$

□

Example 5.79. Find the indefinite integral $\displaystyle\int\frac{dx}{x\sqrt{a^2-x^2}}$, where $a>0$.

First we use trigonometric substitution $x = a\cos u$ to compute the integral. Since $dx = -a\sin u\,du$, we have

$$\begin{aligned}\int\frac{dx}{x\sqrt{a^2-x^2}} &= \int\frac{-a\sin u}{a\cos u\cdot a\sin u}\,du = -\frac{1}{a}\int\sec u\,du = -\frac{1}{a}\ln|\sec u+\tan u| + C\\ &= -\frac{1}{a}\ln\frac{a+\sqrt{a^2-x^2}}{|x|}+C\,.\end{aligned}$$

Now we use hyperbolic functions substitution to compute the integral. Let $x = a\,\mathrm{sech}\,u$ (we note that when using this substitution, we have already restrict ourself to the case $x>0$). Then $dx = -a\,\mathrm{sech}\,u\tanh u\,du$; thus

$$\begin{aligned}\int\frac{dx}{x\sqrt{a^2-x^2}} &= \int\frac{-a\,\mathrm{sech}\,u\tanh u}{a\,\mathrm{sech}\,u\sqrt{a^2-a^2\,\mathrm{sech}^2u}}\,du = -\frac{1}{a}\int\frac{\mathrm{sech}\,u\tanh}{\mathrm{sech}\,u\sqrt{1-\mathrm{sech}^2u}}\,du\\ &= -\frac{1}{a}\int\frac{\mathrm{sech}\,u\tanh u}{\mathrm{sech}\,u\sqrt{a^2\tanh^2u}}\,du = -\frac{1}{a}\int du = -\frac{1}{a}u+C\\ &= \frac{1}{a}\,\mathrm{sech}^{-1}\frac{x}{a}+C = \frac{1}{a}\ln\frac{1+\sqrt{1-\frac{x^2}{a^2}}}{\frac{x}{a}}+C\\ &= -\frac{1}{a}\ln\frac{a+\sqrt{a^2-x^2}}{x}+C\,.\end{aligned}$$

5.10 Exercise

Problem 5.1. Compute $\displaystyle\int_{-\sqrt{\frac{3}{2}}}^{1}\frac{dx}{\sqrt{2-x^2}}$ using the following substitution of variables:

1. $x=\sqrt{2}\sin t$. 2. $x=-\sqrt{2}\sin t$. 3. $x=\sqrt{2}\cos t$. 4. $x=-\sqrt{2}\cos t$.

Problem 5.2. Find the definite integral $\displaystyle\int_0^{\frac{\pi}{2}} \frac{\sin x dx}{1+\cos^2 x}$.

Problem 5.3. Find the following indefinite integrals.

1. $\displaystyle\int \frac{\cos\sqrt{x}}{\sqrt{x}}\,dx$. 2. $\displaystyle\int \frac{1}{x^3}\cos^2\frac{1}{x^2}\,dx$. 3. $\displaystyle\int \frac{1}{x^2}\cos^3\frac{1}{x}\,dx$. 4. $\displaystyle\int \frac{\sin\sqrt{x}}{\sqrt{x\cos^3\sqrt{x}}}\,dx$.

Problem 5.4. Find the following indefinite integrals.

1. $\displaystyle\int \tan^3 x\sec^2 x\,dx\,,\ \int \frac{\sec^2 x}{\tan^2 x}\,dx\,,\ \int \tan^2 x\sec^2 x\,dx$. Use your experience on these three integrals to find $\displaystyle\int \tan^m x\sec^2 x\,dx$ for $m \neq -1$.

2. $\displaystyle\int \sec^3 x\tan x\,dx\,,\ \int \sec^5 x\tan x\,dx\,,\ \int \frac{\tan x}{\sec^3 x}\,dx$. Use your experience on these three integrals to find $\displaystyle\int \sec^m x\tan x\,dx$ for $m \neq 0$.

Problem 5.5. Compute the indefinite integral $\displaystyle\int \sin^6 x\,dx$ by the following steps:

1. Write $\sin^6 x = \sin^3 x\cdot\sin^3 x$, and use the triple and double angle formula, as well as the product to sum formula, to show that
$$\sin^6 x = -\frac{1}{32}\cos 6x + \frac{3}{16}\cos 4x - \frac{15}{32}\cos 2x + \frac{5}{16}\,.$$

2. Find the indefinite integral $\displaystyle\int \sin^6 x\,dx$ using the identity above.

Problem 5.6. Compute the indefinite integral $\displaystyle\int \cos^5 x\,dx$ by the following steps:

1. Write $\cos^5 x = \cos^3 x\cdot\cos^2 x$, and use the triple and double angle formula, as well as the product to sum formula, to show that
$$\cos^5 x = \frac{1}{16}\cos 5x + \frac{5}{16}\cos 3x + \frac{5}{8}\cos x\,.$$

2. Find the indefinite integral $\displaystyle\int \cos^5 x\,dx$ using the identity above.

Problem 5.7. Let y be a twice differentiable function satisfying
$$\frac{d^2y}{dx^2} = 4\sec^{-2} 2x\tan 2x\,,\quad y(0) = -1, y'(0) = 4\,.$$

Find y.

Problem 5.8. 1. Let $f : [0, a] \to \mathbb{R}$ be a continuous function. Find $\displaystyle\int_0^a \frac{f(x)}{f(x) + f(a - x)}\, dx$.

2. Find $\displaystyle\int_0^1 \frac{\sin x}{\sin x + \sin(1 - x)}\, dx$. 3. Find $\displaystyle\int_0^3 \frac{\sqrt{x}}{\sqrt{x} + \sqrt{3 - x}}\, dx$.

Problem 5.9. Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function.

1. Show that
$$\int_0^{\frac{\pi}{2}} f(\sin x)\, dx = \int_0^{\frac{\pi}{2}} f(\cos x)\, dx\,.$$

2. Use the identity above to find the integrals $\displaystyle\int_0^{\frac{\pi}{2}} \cos^2 x\, dx$ and $\displaystyle\int_0^{\frac{\pi}{2}} \sin^2 x\, dx$.

Problem 5.10. Let a and b be positive (rational) numbers. Show that
$$\int_0^1 x^a (1 - x)^b\, dx = \int_0^1 x^b (1 - x)^a\, dx\,.$$

Problem 5.11. Let $a_0, a_1, \cdots, a_n$ be real numbers satisfying
$$\frac{a_0}{1} + \frac{a_1}{2} + \frac{a_2}{3} + \cdots + \frac{a_n}{n + 1} = 0\,.$$
Show that the polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ has at least one real zero.

Problem 5.12. Let I be an interval, and $f : I \to \mathbb{R}$ be one-to-one, onto and continuous. Show that if $g : \mathbb{N} \to \mathbb{R}$ is a function satisfying that $\lim\limits_{n\to\infty} f(g(n)) = b$, then $\lim\limits_{n\to\infty} g(n) = f^{-1}(b)$.

Problem 5.13. Show that the following functions (defined by integrals) are one-to-one and find $(f^{-1})'(0)$.

1. $f(x) = \displaystyle\int_2^x \sqrt{1 + t^2}\, dt$. 2. $f(x) = \displaystyle\int_2^x \frac{dt}{\sqrt{1 + t^4}}$.

Problem 5.14. Let f be an one-to-one, twice differentiable function with an inverse function g.

1. Show that g is twice differentiable function and find g''.

2. Show that if in addition f is strictly increasing and the graph of f is concave upward, then the graph of g is concave downward.

Problem 5.15. Find the limit $\lim_{n\to\infty}\left(\frac{n!}{n^n}\right)^{\frac{1}{n}}$ through the following steps.

(1) Show that $\sum_{k=1}^{n-1}\frac{1}{n}\ln\frac{k}{n} \leqslant \int_{\frac{1}{n}}^{1}\ln x\,dx \leqslant \sum_{k=2}^{n}\frac{1}{n}\ln\frac{k}{n}$.

(2) Find $\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{n}\ln\frac{k}{n}$.

(3) Find $\lim_{n\to\infty}\left(\frac{n!}{n^n}\right)^{\frac{1}{n}}$.

Hint: (1) Use the property of integrals.
(3) Using (1).

Problem 5.16. Show that for all natural number n,

$$\sum_{k=1}^{2n}\frac{(-1)^{k-1}x^k}{k} \leqslant \ln(1+x) \leqslant \sum_{k=1}^{2n-1}\frac{(-1)^{k-1}x^k}{k} \qquad \forall\, x>0\,.$$

Problem 5.17. Find the derivative of the following functions by first taking the logarithm (base e) and then differentiating.

1. $y=\frac{x(x-1)^{\frac{3}{2}}}{\sqrt{x+1}}$, $x>1$. 2. $y=\frac{(x+1)(x-2)}{(x-1)(x+2)}$, $x>2$

Problem 5.18. Use implicit differentiation to find $\frac{dy}{dx}$, where (x,y) satisfies the relation $4xy+\ln x^2y=7$.

Problem 5.19. Locate any relative extrema and points of inflection of the function $y=x^2\ln\frac{x}{4}$.

Problem 5.20. Use the substitution of variable $t=\tan\frac{x}{2}$ to find the integral $\int\csc x\,dx$.

Problem 5.21. Find the following indefinite integrals.

1. $\int\frac{(\ln x)^2}{x}\,dx$. 2. $\int\frac{\ln\sqrt{x}}{x}\,dx$. 3. $\int\frac{dx}{x(\ln x^2)^3}$. 4. $\int\frac{(1+\ln x)^2}{x}\,dx$.

5. $\int\frac{\sin(\ln x)}{x}\,dx$. 6. $\int\frac{\sin 2x}{1+\cos^2 x}\,dx$.

Problem 5.22. Show that $\frac{1}{y}<\frac{\ln x-\ln y}{x-y}<\frac{1}{x}$ for all $0<x<y$.

Problem 5.23. Show that the following functions are decreasing on $(0, \infty)$.

1. $y = \left(1 + \dfrac{1}{2x}\right)^{x+0.5}$. 2. $y = \left(1 + \dfrac{1}{x}\right)^{x+0.5}$.

Problem 5.24. In this example you are asked to compute the integral of $y = xe^x$ by the Riemann sum. Complete the following.

1. Show that if $r \neq 1$, then $\sum\limits_{k=1}^{n} kr^k = \dfrac{r(1-r^n)}{(1-r)^2} - \dfrac{nr^{n+1}}{1-r}$.

2. Compute $\displaystyle\int_0^a xe^x\,dx$ by the limit the Riemann sum of $y = xe^x$ for regular partition using the right end-point rule.

3. Find an anti-derivative of $y = xe^x$.

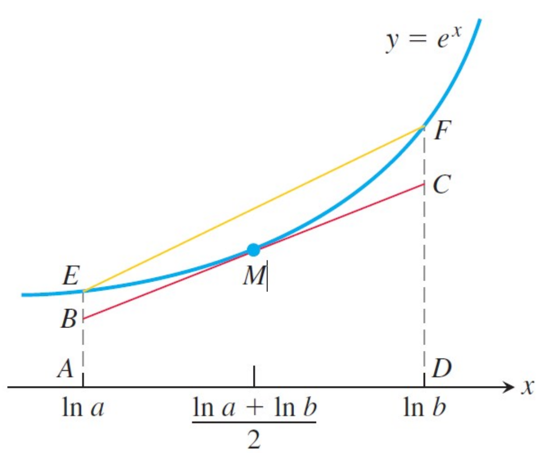
Problem 5.25 (Integrating Factor).

1. Let $f, g : [a, b] \to \mathbb{R}$ be a continuous function, F be an anti-derivative of f, and $y : [a, b] \to \mathbb{R}$ satisfies that
$$y' + f(x)y = g(x)\,. \tag{$\star$}$$
Find an expression of y.

2. Find the function y satisfying $y' + x^2y = 2x^3$ and $y(0) = 1$.

Hint: Multiply both sides of $(\star)$ by $\exp(F(x))$ and observe that the left-hand side is the derivative of a certain function.

Problem 5.26. 1. Show that for $0 < a < b$,
$$e^{\frac{\ln a + \ln b}{2}} \cdot (\ln b - \ln a) < \int_{\ln a}^{\ln b} e^x\,dx < \frac{e^{\ln a} + e^{\ln b}}{2} \cdot (\ln b - \ln a)\,.$$



2. Using the result above to show that for $0 < a < b$,

$$\sqrt{ab} < \frac{b-a}{\ln b - \ln a} < \frac{a+b}{2}.$$

Problem 5.27. Prove the following inequalities.

1. $e^x > 1 + \ln(1+x)$ for all $x > 0$.

2. $e^x > 1 + (1+x)\ln(1+x)$ for all $x > 0$.

3. $e^x \geqslant 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \cdots + \dfrac{x^n}{n!}$ for all $x \geqslant 0$ and $n \in \mathbb{N}$.

Problem 5.28. Let a, b be two positive numbers, p, q any nonzero numbers, and $p < q$. Prove that

$$\left[\theta a^p + (1-\theta)b^p\right]^{\frac{1}{p}} \leqslant \left[\theta a^q + (1-\theta)b^q\right]^{\frac{1}{q}} \qquad \forall\, \theta \in (0,1).$$

Hint: Show that the function $f(p) = \left[\theta a^p + (1-\theta)b^p\right]^{\frac{1}{p}}$ is an increasing function of p.

Problem 5.29. 1. Find an equation for the line through the origin tangent to the graph of $y = \ln x$.

2. Show that $\ln x < \dfrac{x}{e}$ for all $x \neq e$.

3. Show that $x^e < e^x$ for all $x \neq e$.

4. Show that if $e \leqslant A < B$, then $A^B > B^A$.

Problem 5.30 (Implicit Differentiation).

1. Find y' if $e^{\frac{x}{y}} = x - y$.

2. Find an equation of the tangent line to the curve $xe^y + ye^x = 1$ at the point $(0,1)$.

3. Find an equation of the tangent line to the curve $1 + \ln xy = e^{x-y}$ at the point $(1,1)$

Problem 5.31. Evaluate the following limits. Use L'Hôspital's Rule where appropriate. If L'Hôspital's Rule does not apply, explain why.

1. $\displaystyle\lim_{x\to 0^+} \frac{\arctan(2x)}{\ln x}$.
2. $\displaystyle\lim_{x\to 0^+} \frac{x^x - 1}{\ln x + x - 1}$.
3. $\displaystyle\lim_{x\to 0} \frac{\ln(1+x)}{\cos x + e^x - 1}$.
4. $\displaystyle\lim_{x\to 0} \frac{x^a - 1}{x^b - 1}$, where $b \neq 0$.
5. $\displaystyle\lim_{x\to 0} \frac{e^x - e^{-x} - 2x}{x - \sin x}$.
6. $\displaystyle\lim_{x\to a^+} \frac{\cos x \cdot \ln(x-a)}{\ln(e^x - e^a)}$.

7. $\lim\limits_{x\to 0^+}\left(\frac{1}{x}-\frac{1}{\arctan x}\right)$. 8. $\lim\limits_{x\to\infty}(x-\ln x)$. 9. $\lim\limits_{x\to 1^+}\ln(x^7-1)-\ln(x^5-1)$.

10. $\lim\limits_{x\to\infty}x^{\frac{\ln 2}{1+\ln x}}$. 11. $\lim\limits_{x\to\infty}x^{e^{-x}}$. 12. $\lim\limits_{x\to 1}(2-x)^{\tan(\pi x/2)}$. 13. $\lim\limits_{x\to 0^+}(\sin x)(\ln x)$.

Problem 5.32. Evaluate the following limits:

1. $\lim\limits_{x\to\infty}x\Big[\big(1+\frac{1}{x}\big)^x-e\Big]$. 2. $\lim\limits_{x\to\infty}\Big\{\frac{e}{2}x+x^2\Big[\big(1+\frac{1}{x}\big)^x-e\Big]\Big\}$.

3. $\lim\limits_{x\to\infty}x\Big[\big(1+\frac{1}{x}\big)^x-e\ln\big(1+\frac{1}{x}\big)^x\Big]$. 4. $\lim\limits_{x\to 0}\big(\frac{\sin x}{x}\big)^{\frac{1}{x^2}}$. 5. $\lim\limits_{x\to\infty}\big(\frac{\sin x}{x}\big)^{\frac{1}{x^2}}$.

6. $\lim\limits_{x\to\infty}\left(x-x^2\ln\frac{1+x}{x}\right)$. 7. $\lim\limits_{x\to\infty}\left[\frac{1}{x}\cdot\frac{a^x-1}{a-1}\right]^{\frac{1}{x}}$, where $a>0$ and $a\neq 1$.

Problem 5.33. For what values of a and b is the following equations true?

1. $\lim\limits_{x\to 0}\left(\frac{\sin 2x}{x^3}+a+\frac{b}{x^2}\right)=0.$

2. $\lim\limits_{x\to 0}\left(\frac{\tan 2x}{x^3}+\frac{a}{x^2}+\frac{\sin bx}{x}\right)=0.$

Problem 5.34. Show that $\lim\limits_{x\to\infty}x^{x^{-n}}=1$ for every positive integer n.

Problem 5.35. Let $f(x)=\begin{cases} e^{-1/x^2} & \text{if } x\neq 0\,,\\ 0 & \text{if } x=0\,.\end{cases}$

1. Find $f'(0)$. Is f continuously differentiable?

2. Show that f has derivatives of all orders on $\mathbb{R}$; that is, f is infinitely many times differentiable on $\mathbb{R}$.

Hint: First show by induction that there is a polynomial $p_n(x)$ and a non-negative integer k_n such that $f^{(n)}(x)=\frac{p_n(x)f(x)}{x^{k_n}}$ for $x\neq 0$.

Problem 5.36. Find $\frac{d}{dx}\arcsin(\sin x)$, $\frac{d}{dx}\arccos(\sin x)$ and $\frac{d}{dx}\arctan(\tan x)$.

Problem 5.37. Show that $2\arcsin x=\arccos(1-2x^2)$ for all $x\geqslant 0$.

Problem 5.38. Prove the identity $\arcsin\frac{x-1}{x+1}=2\arctan\sqrt{x}-\frac{\pi}{2}$ for all $x\geqslant 0$.

Problem 5.39. Prove that $\frac{x}{1+x^2}<\arctan x<x$ for all $x>0$.

Problem 5.40. Evaluate $\displaystyle\int_0^1 \arcsin x\,dx$ by interpreting it as an area and integrating with respect to y instead of x.

Problem 5.41. Evaluate the following definite integrals.

1. $\displaystyle\int_0^{\frac{1}{\sqrt{2}}} \frac{\arcsin x}{\sqrt{1-x^2}}\,dx.$ 2. $\displaystyle\int_0^{\frac{1}{\sqrt{2}}} \frac{\arccos x}{\sqrt{1-x^2}}\,dx.$ 3. $\displaystyle\int_{\ln 2}^{\ln 4} \frac{e^{-x}}{\sqrt{1-e^{-2x}}}\,dx.$

4. $\displaystyle\int_2^3 \frac{2x-3}{\sqrt{4x-x^2}}\,dx.$ 5. $\displaystyle\int_3^4 \frac{dx}{(x-1)\sqrt{x^2-2x}}.$

Problem 5.42. Find the following indefinite integrals.

1. $\displaystyle\int \sqrt{e^x-3}\,dx.$ 2. $\displaystyle\int \frac{\sqrt{x-2}}{x+1}\,dx.$ 3. $\displaystyle\int \frac{dx}{\sqrt{-2x^2+8x+4}}.$

4. $\displaystyle\int \frac{2x\arctan(x^2+1)}{x^4+2x^2+2}\,dx.$ 5. $\displaystyle\int \frac{\sqrt{x}}{4+x^3}\,dx.$ 6. $\displaystyle\int \sqrt{\frac{x}{4+x^3}}\,dx,\ x>0.$

Problem 5.43. Find the function y satisfying $(1+x^2)y'+xy=1$ and $y(0)=1$.

Problem 5.44. Show (by contradiction) that π is an irrational number through the following steps.

1. Assume (the contrary) that $\pi = \dfrac{a}{b}$ for some $a, b \in \mathbb{N}$. Define $f(x) = \dfrac{x^n(a-bx)^n}{n!}$. Show that $f(x) = f(\pi - x)$, and $0 < f(x)\sin x < \dfrac{a^n\pi^n}{n!}$ for all $0 < x < \pi$.

2. Show that $f^{(k)}(0)$ and $f^{(k)}(\pi)$ are all integers for all $k \in \mathbb{N}$, where $f^{(k)}(0) = \dfrac{d^k}{dx^k}\Big|_{x=0} f(x)$ and $f^{(k)}(\pi) = \dfrac{d^k}{dx^k}\Big|_{x=\pi} f(x)$.

3. Define
$$g(x) = \sum_{k=0}^{n} (-1)^k f^{(2k)}(x) = f(x) - f^{(2)}(x) + f^{(4)}(x) + \cdots + (-1)^n f^{(2n)}(x)\,.$$
Show that $\dfrac{d}{dx}\big(g'(x)\sin x - g(x)\cos x\big) = f(x)\sin x$ and conclude that $\displaystyle\int_0^\pi f(x)\sin x\,dx = g(0)+g(\pi)$.

4. Conclude that $0 < \displaystyle\int_0^\pi f(x)\sin x\,dx < \frac{a^n\pi^{n+1}}{n!}$ and use (5.4.5) to lead to a contradiction.

Chapter 7

Applications of Integration

7.1 Area of a Region between Two Curves

The motivation of integration of functions is finding areas. Let us recall that if $f : [a, b] \to \mathbb{R}$ is non-negative, then the area A of the region bounded by the graph of f, the x-axis and vertical lines $x = a$ and $x = b$ is the integral of f on $[a, b]$ or in notation,

$$A = \int_a^b f(x)\,dx\,.$$

The idea above can be extended to the following statement: Let $f, g : [a, b] \to \mathbb{R}$ be continuous and $g(x) \leqslant f(x)$ for all $x \in [a, b]$, then the area of the region bounded by the graphs of f and g and the vertical lines $x = a$ and $x = b$ is

$$A = \int_a^b \big[f(x) - g(x)\big]\,dx\,.$$

• **How about if the graphs of two continuous functions intersect?**

Suppose that $f, g : [a, b] \to \mathbb{R}$ are continuous functions but neither $g(x) \leqslant f(x)$ for all $x \in [a, b]$ nor $f(x) \leqslant g(x)$ for all $x \in [a, b]$. In other words, the graphs of f and g intersect (and transverse). In this case, the area of the region bounded by the graphs of f and g, as well as the vertical lines $x = a$ and $x = b$, is given by

$$A = \int_a^b \big|f(x) - g(x)\big|\,dx\,.$$

To find this integral, in general we need to find all the zeros of the function $h(x) = f(x) - g(x)$ and write the integral as sum of integrals on sub-intervals. To be more precise, suppose that the distinct zeros of h is given by $\{c_k\}_{k=1}^n$, where $a \leqslant c_1 < c_2 < \cdots < c_n \leqslant b$, then

$$\begin{aligned} A &= \int_a^b \big|f(x) - g(x)\big|\,dx \\ &= \int_a^{c_1} \big|f(x) - g(x)\big|\,dx + \sum_{k=1}^n \int_{c_{k-1}}^{c_k} \big|f(x) - g(x)\big|\,dx + \int_{c_n}^b \big|f(x) - g(x)\big|\,dx \\ &= \Big|\int_a^{c_1} \big[f(x) - g(x)\big]\,dx\Big| + \sum_{k=1}^n \Big|\int_{c_{k-1}}^{c_k} \big[f(x) - g(x)\big]\,dx\Big| + \Big|\int_{c_n}^b \big[f(x) - g(x)\big]\,dx\Big|. \end{aligned}$$

When f, g are continuous function on $\mathbb{R}$ and $h = f - g$ has finitely many distinct zeros $\{c_k\}_{k=1}^n$, we can also talk about the area of the (bounded) region bounded by the graph of f and g. This area is given by

$$A = \sum_{k=1}^n \Big|\int_{c_{k-1}}^{c_k} \big[f(x) - g(x)\big]\,dx\Big|$$

7.2 Volume: The Disk Method（圓盤法）

In the following two sections, the main focus is to develop ways of finding the volume of the so-called solids of revolution（旋轉體）, a solid formed by revolving a certain region about a line called the axis of revolution (and usually a line parallel to the x-axis or y-axis).

Example 7.1. The ball centered at the origin with radius r (usually denoted by $B(0, r)$ or $B_r(0)$), is a solid of revolution. It can be formed by revolving the region

$$\mathrm{R} = \Big\{(x, y)\,\Big|\,0 \leqslant y \leqslant \sqrt{r^2 - x^2}\Big\}$$

about the x-axis.

Example 7.2. A solid torus can be formed by revolving a disk

$$\mathrm{D} = \Big\{(x, y)\,\Big|\,(x - a)^2 + y^2 = r^2\Big\} \qquad \text{(where } 0 < a < r)$$

about the y-axis.

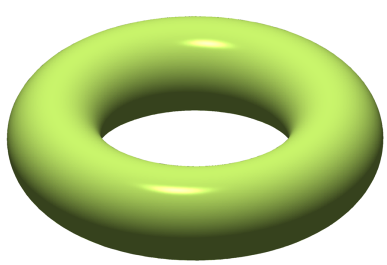


Figure 7.1: A solid torus

Consider the volume of a solid D formed by revolving a region R about the line $y = y_0$, where the region R is given by

$$\mathrm{R} = \left\{(x,y) \,\middle|\, x \in [a,b]\,, y_0 \leqslant y \leqslant f(x)\right\}$$

for some continuous function $f : [a,b] \to \mathbb{R}$ with $\min\limits_{x\in[a,b]} f(x) \geqslant y_0$. Note that the function $y = \pi\big[f(x) - c\big]^2$ is also continuous on $[a,b]$ thus integrable on $[a,b]$.

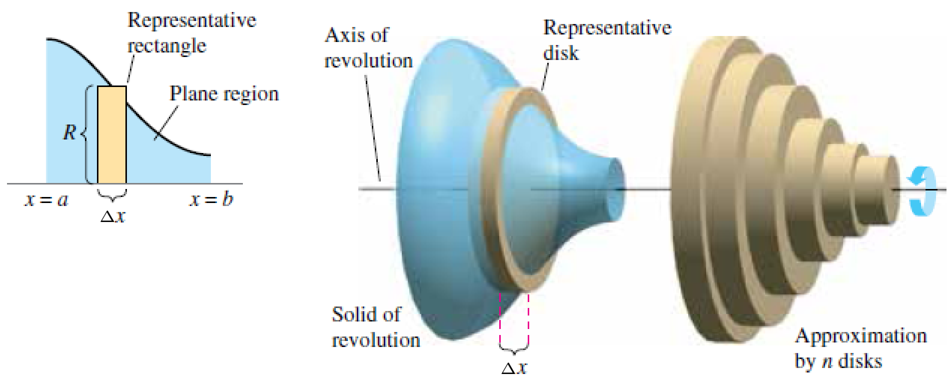


Figure 7.2: Disk method

Let $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[a,b]$, and $\Delta x_i = x_i - x_{i-1}$. Then the volume of D is approximated by

$$\sum_{i=1}^{n} \pi\big[f(\xi_i) - y_0\big]^2 \Delta x_i\,,$$

where $\xi_i \in [x_{i-1}, x_i]$ for each $1 \leqslant i \leqslant n$. Note that the sum above is a Riemann sum of the function $y = \pi\big[f(x) - y_0\big]^2$ for partition $\mathcal{P}$.

When $\|\mathcal{P}\|$ approaches 0, we expect that the sum above approaches the volume of D. Since f is continuous on $[a,b]$, the function $y = \pi\big[f(x) - y_0\big]^2$ is Riemann integrable on $[a,b]$; thus for any given $\varepsilon > 0$, there exists $\delta > 0$ such that if $\|P\| < \delta$, then any Riemman sum of the function $y = \pi\big[f(x) - y_0\big]^2$ for $\mathcal{P}$ lies in the interval

$$\Big(\int_a^b \pi\big[f(x) - y_0\big]^2\,dx - \varepsilon,\, \int_a^b \pi\big[f(x) - y_0\big]^2\,dx + \varepsilon\Big)\,.$$

In particular, if $\max\left\{x_i - x_{i-1} \,\middle|\, 1 \leqslant i \leqslant n\right\} < \delta$,

$$\Big|\sum_{i=1}^{n} \pi\big[f(\xi_i) - y_0\big]^2 \Delta x_i - \int_a^b \pi\big[f(x) - y_0\big]^2\,dx\Big| < \varepsilon\,.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that the volume of D can be computed by

$$\pi \int_a^b \big[f(x) - y_0\big]^2\, dx\,.$$

Example 7.3. The volume of the ball $B(0, r)$ is given by

$$\pi \int_{-r}^{r} \big(\sqrt{r^2 - x^2}\big)^2\, dx = \pi \int_{-r}^{r} (r^2 - x^2)\, dx = \pi \Big[r^2 x - \frac{1}{3}x^3\Big]\Big|_{x=-r}^{x=r} = \frac{4}{3}\pi r^3\,.$$

Example 7.4. The volume of the solid formed by revolving the region bounded by the graphs of $f(x) = 2 - x^2$ and $g(x) = 1$ about the line $y = 1$ is given by

$$\begin{aligned}\pi \int_{-1}^{1} \big[(2 - x^2) - 1\big]^2\, dx &= \pi \int_{-1}^{1} (1 - x^2)^2\, dx = \pi \int_{-1}^{1} (1 - 2x^2 + x^4)\, dx \\ &= \pi \Big[x - \frac{2}{3}x^3 + \frac{1}{5}x^5\Big]\Big|_{x=-1}^{x=1} = \frac{16\pi}{15}\,.\end{aligned}$$

Similarly, if D is a solid formed by revolving a region R about the line $x = x_0$, where R is given by

$$\mathrm{R} = \big\{(x, y)\,\big|\, y \in [c, d]\,, x_0 \leqslant x \leqslant g(y)\big\}$$

for some continuous function $g : [c, d] \to \mathbb{R}$ with $\min\limits_{y\in[c,d]} g(y) \geqslant x_0$, then the volume of D is

$$\pi \int_c^d \big[g(y) - x_0\big]^2\, dy\,.$$

A solid of revolution may be formed by revolving a region away from the axis of revolution. In this case, the solid will have holes and the volume of

Suppose that the region R is given by

$$\mathrm{R} = \big\{(x, y)\,\big|\, a \leqslant x \leqslant b\,, y_0 \leqslant g(x) \leqslant y \leqslant f(x)\big\}\,,$$

where $f, g : [a, b] \to \mathbb{R}$ are continuous functions with $\max\limits_{x\in[a,b]} g(x) \leqslant \min\limits_{x\in[a,b]} f(x)$. Let R_1 and R_2 be given by

$$\mathrm{R}_1 = \big\{(x, y)\,\big|\, a \leqslant x \leqslant b\,, y_0 \leqslant y \leqslant f(x)\big\} \quad \text{and} \quad \mathrm{R}_2 = \big\{(x, y)\,\big|\, a \leqslant x \leqslant b\,, y_0 \leqslant y \leqslant g(x)\big\}\,.$$

Then the volume of the solid formed by revolving R about the line $y = y_0$ is the volume of the solid formed by revolving R_1 about the line $y = y_0$ minus the volume of the solid formed by revolving R_2 about the line $y = y_0$ and is given by

$$\pi \int_a^b \big[(f(x) - y_0)^2 - (g(x) - y_0)^2\big]\, dx\,.$$

Similarly, if R is given by

$$\mathrm{R} = \left\{(x,y) \,\middle|\, c \leqslant y \leqslant d \,, x_0 \leqslant g(y) \leqslant x \leqslant f(y)\right\},$$

where $f, g : [c,d] \to \mathbb{R}$ are continuous functions with $\max\limits_{y\in[c,d]} g(y) \leqslant \min\limits_{y\in[c,d]} f(y)$.. Then the volume of the solid formed by revolving R about the line $x = x_0$ is given by

$$\pi \int_c^d \left[(f(y) - x_0)^2 - (g(y) - x_0)^2\right] dy \,.$$

Example 7.5. Find the volume of the solid formed by revolving the region bounded by the graphs of $y = \sqrt{x}$ and $y = x^2$ about the x-axis.

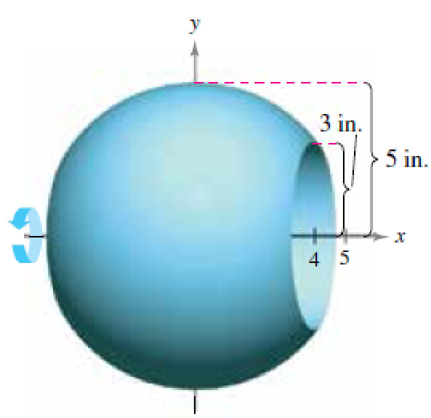
The points of intersection of the graphs of the two functions are $x = 0$ and $x = 1$, and $0 \leqslant x^2 \leqslant x$ on $[0,1]$. Therefore, the volume of the solid described above is given by

$$\pi \int_0^1 \left[\sqrt{x}^2 - (x^2)^2\right] dx = \pi \int_0^1 \left(x - x^4\right) dx = \pi \Big(\frac{1}{2}x^2 - \frac{1}{5}x^4\Big)\Big|_{x=0}^{x=1} = \frac{3\pi}{10} \,.$$

Example 7.6. The volume of the solid torus given in Example 7.2 is given by

$$\begin{aligned} &\pi \int_{-r}^{r} \left[(a + \sqrt{r^2 - y^2} - 0)^2 - (a - \sqrt{r^2 - y^2} - 0)^2\right] dy \\ &\qquad = 4a\pi \int_{-r}^{r} \sqrt{r^2 - y^2}\, dy = 4a\pi \cdot \frac{\pi r^2}{2} = 2\pi^2 a r^2 \,. \end{aligned}$$

Example 7.7. Find the volume of the solid formed by a ball with 5 inch radius having a cylindrical hole as shown in the following figure.



Solid of revolution

The volume of the solid described above is given by

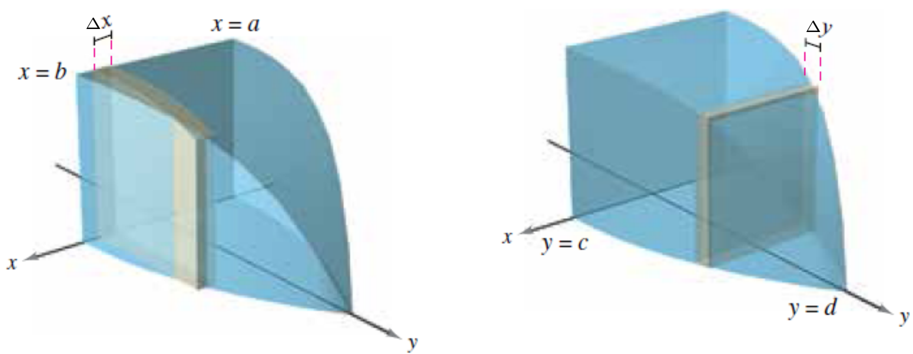
$$\pi \int_{-4}^{4} \left[(\sqrt{25 - x^2} - 0)^2 - (3 - 0)^2\right] dx = \frac{256\pi}{3} \,.$$

In general, the disk method can be used to compute a solid whose area of cross sections along a particular axis is known. Let D be a solid lies between two planes $x = a$ and $x = b$ ($a < b$), and the area of the cross section of D taken perpendicular to the x-axis is $A(x)$, then

$$\text{the volume of D} = \int_a^b A(x)\,dx\,.$$

Similarly, if D lies between $y = c$ and $y = d$ ($c < d$), and the area of the cross section of D taken perpendicular to the y-axis is $A(y)$, then

$$\text{the volume of D} = \int_c^d A(y)\,dy\,.$$

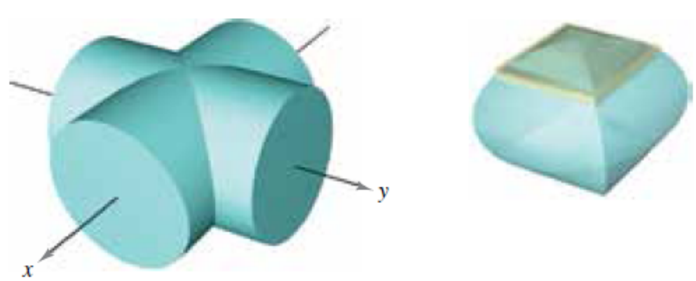


(a) Cross sections perpendicular to x-axis **(b)** Cross sections perpendicular to y-axis

Example 7.8. The volume of a cone with height h and base area A is given by

$$\int_0^h \frac{A(h-y)^2}{h^2}\,dy = -\frac{A}{h^2}\frac{1}{3}(h-y)^3\Big|_{y=0}^{y=h} = \frac{1}{3}Ah\,.$$

Example 7.9. Find the volume of the solid of intersection of the two right circular cylinders of radius r whose axes meet at right angles.



Two intersecting cylinders Solid of intersection

The area of cross sections taken perpendicular to the z-axis is given by

$$A(z) = (2\sqrt{r^2 - z^2})^2 = 4(r^2 - z^2)\,.$$

Therefore, the volume of the solid of intersection is given by

$$\int_{-r}^{r} 4(r^2 - z^2)\,dz = \frac{16}{3}r^3\,.$$

7.3 Volume: The Shell Method（剝殼法）

Consider the volume of a solid D formed by revolving a region R about line $x = L$, where R is given by

$$\mathrm{R} = \big\{(x, y)\,\big|\,x \in [a, b]\,, 0 \leqslant y \leqslant f(x)\big\}\,,$$

for some $a \geqslant L$ and continuous function $f : [a, b] \to [0, \infty)$. When f is one-to-one, let $g : [c, d] \to \mathbb{R}$ be the inverse function of f $\big($note that $c = \min\limits_{x\in[a,b]} f(x)$ and $d = \max\limits_{x\in[a,b]} f(x)\big)$. Then the volume of D computed using the disk method is given by

$$\pi \int_c^d \big[(g(y) - L)^2 - (a - L)^2\big]^2\,dy + \pi \int_0^c \big[(b - L)^2 - (a - L)^2\big]\,dy\,.$$

On the other hand, if f is not one-to-one, then sometimes it will be not so easy to find the volume of D using the disk method. How do we compute the volume of D in this case?

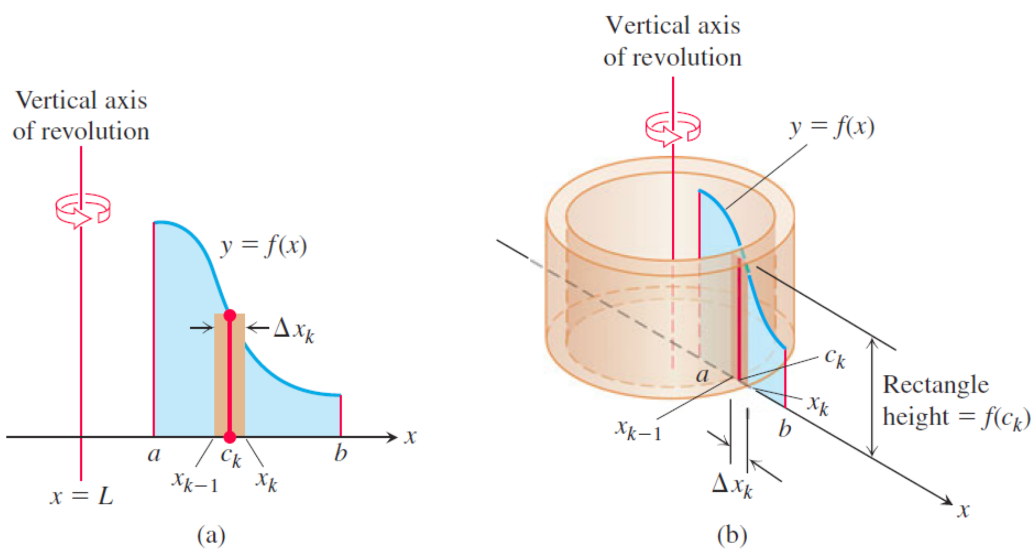


Figure 7.3: The shell method

Let $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[a, b]$, $\Delta x_k = x_k - x_{k-1}$ and $c_k = \dfrac{x_{k-1} + x_k}{2}$; that is, c_k is the middle point of the interval $[x_{k-1}, x_k]$. Then Figure 7.3 implies that the volume of D can be approximated by the sum of these cylindrical shells (one cylindrical shell is shown above in orange color). The volume of the cylindrical shell shown in Figure 7.3 is given by

$$\begin{aligned}\pi(x_k - L)^2 f(c_k) - \pi(x_{k-1} - L)^2 f(c_k) &= \pi f(c_k)\big[(x_k - L)^2 - (x_{k-1} - L)^2\big] \\ &= \pi f(c_k)(x_k - L + x_{k-1} - L)\big[x_k - L - (x_{k-1} - L)\big] = 2\pi(c_k - L)f(c_k)\Delta x_k\,.\end{aligned}$$

Therefore, the volume of D can be approximated by the sum

$$\sum_{k=1}^{n} 2\pi(c_k - L)f(c_k)\Delta x_k\,.$$

We note that the sum above is a Riemann sum of the function $y = 2\pi(x - L)f(x)$ for partition $\mathcal{P}$ using the mid-point rule. Therefore, similar to the argument in Section 7.2, we find that the volume of D is given by

$$2\pi \int_a^b (x - L)f(x)\,dx\,.$$

This way of computing the volume of D is called the shell method.

Similarly, let D be formed by revolving a region R about the line $x = L$, where the region R is given by

$$\mathrm{R} = \big\{(x, y)\,\big|\,x \in [a, b]\,, g(x) \leqslant y \leqslant f(x)\big\}$$

for some $a > L$ and continuous functions $f, g : [a, b] \to \mathbb{R}$ with $\min\limits_{x\in[a,b]} f(x) \geqslant \max\limits_{x\in[a,b]} g(x)$. Then the volume of D is given by

$$2\pi \int_a^b (x - L)\big[f(x) - g(x)\big]\,dx\,.$$

Example 7.10. In this example we compute the volume of the ball $B(0, r)$ by the shell method. Note that $B(0, r)$ can be formed by revolving the region

$$\mathrm{R} = \big\{(x, y)\,\big|\,x \in [0, r]\,, -\sqrt{r^2 - x^2} \leqslant y \leqslant \sqrt{r^2 - x^2}\big\}$$

about the y-axis. Therefore, by the shell method, the volume of $B(0, r)$ is given by

$$\begin{aligned}&\pi \int_0^r (x - 0)\big[\sqrt{r^2 - x^2} - (-\sqrt{r^2 - x^2})\big]\,dx \\ &\quad = 4\pi \int_0^r x(r^2 - x^2)^{\frac{1}{2}}\,dx = 4\pi\Big[-\frac{3}{4}(r^2 - x^2)^{\frac{2}{3}}\Big]\Big|_{x=0}^{x=r} = \frac{4\pi}{3}r^3\,.\end{aligned}$$

Example 7.11. In this example we compute the volume of the solid torus given in Example 7.2 by the shell method. Note that this solid torus can also be formed by revolving the region

$$\mathrm{R} = \left\{(x,y) \,\middle|\, x \in [a-r, a+r]\,, -\sqrt{r^2-(x-a)^2} \leqslant y \leqslant \sqrt{r^2-(x-a)^2}\right\}$$

about the y-axis. Therefore, by the shell method, the volume of the solid torus given in Example 7.2 is given by

$$2\pi \int_{a-r}^{a+r} (x-0)\big[\sqrt{r^2-(x-a)^2} - (-\sqrt{r^2-(x-a)^2})\big]\,dx = 4\pi \int_{a-r}^{a+r} x\sqrt{r^2-(x-a)^2}\,dx\,.$$

By the substitution $x = a + r\sin u$, we find that

$$\begin{aligned}
\int_{a-r}^{a+r} x\sqrt{r^2-(x-a)^2}\,dx &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (a+r\sin u)\sqrt{r^2-r^2\sin^2 u}\cdot r\cos u\,du \\
&= r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (a+r\sin u)\cos^2 u\,du = r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \Big[a\frac{1+\cos(2u)}{2} + r\sin u\cos^2 u\Big]\,du \\
&= r^2\Big[\frac{au}{2} + \frac{a\sin(2u)}{4} - \frac{r}{3}\cos^3 u\Big]\Big|_{u=-\frac{\pi}{2}}^{u=\frac{\pi}{2}} = \frac{\pi}{2}ar^2\,;
\end{aligned}$$

thus the volume of the solid torus is $4\pi\cdot\dfrac{\pi}{2}ar^2 = 2\pi^2 ar^2$ which agrees with what Example 7.6 shows.

Example 7.12. A solid D is formed by rotating the region bounded by the graph of $y = 1 - \dfrac{x^2}{16}$ and $y = 0$ about the x-axis. Then the volume of D computed by the disk method is given by

$$\pi\int_{-4}^{4}\Big(1-\frac{x^2}{16}\Big)^2\,dx = \pi\Big[x - \frac{x^3}{24} + \frac{x^5}{5\cdot 256}\Big]\Big|_{x=-4}^{x=4} = \frac{64\pi}{15}$$

while the volume of D computed by the shell method is given by

$$\begin{aligned}
2\pi\int_0^1 y\big[\sqrt{16(1-y)} - (-\sqrt{16(1-y)})\big]\,dy &= 16\pi\int_0^1 y\sqrt{1-y}\,dy = 16\pi\int_1^0 (1-u)u^{\frac{1}{2}}(-\,du) \\
&= 16\pi\int_0^1 \big(u^{\frac{1}{2}} - u^{\frac{3}{2}}\big)\,du = 16\pi\Big[\frac{2}{3}u^{\frac{3}{2}} - \frac{2}{5}u^{\frac{5}{2}}\Big]_{u=0}^{u=1} = 16\pi\Big(\frac{2}{3}-\frac{2}{5}\Big) = \frac{64\pi}{15}\,.
\end{aligned}$$

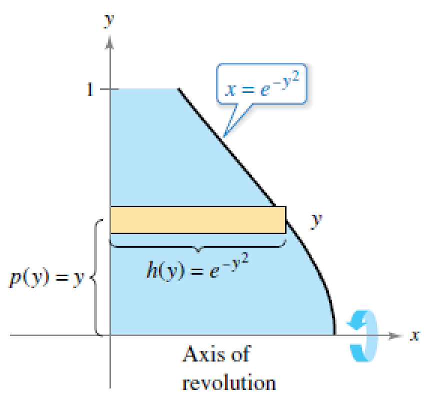
Now consider the volume of a solid D formed by revolving a region R about the line $y = L$, where

$$\mathrm{R} = \left\{(x,y) \,\middle|\, y \in [c,d]\,, g(y) \leqslant x \leqslant f(y)\right\}$$

for some $c \geqslant L$ and continuous functions $f, g : [c,d] \to \mathbb{R}$ with $\min\limits_{y\in[c,d]} f(y) \geqslant \max\limits_{y\in[c,d]} g(y)$. Similar to the argument above, the volume of D is given by

$$2\pi \int_c^d (y-L)\big[f(y)-g(y)\big]\,dy\,.$$

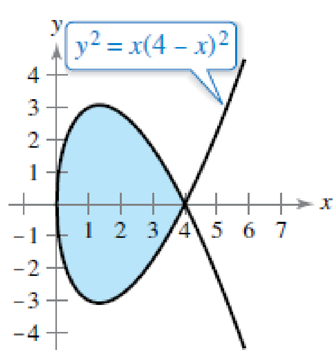
Example 7.13. Find the volume of the solid formed by revolving the region R about the x-axis, where R is the region bounded by the graph of $x = e^{-y^2}$, $y = 0$, $y = 1$ and the y-axis.



Using the shell method, the volume of this solid is given by

$$2\pi \int_0^1 (y-0)e^{-y^2}\,dy = 2\pi\Big(-\frac{e^{-y^2}}{2}\Big)\Big|_{y=0}^{y=1} = \pi(1-e^{-1})\,.$$

Example 7.14. Let R be the region bounded by the graph of $y^2 = x(4-x)^2$.



Find the volume of the solid D formed by revolving R about

(a) the x-axis, (b) the y-axis, and (c) the line $x = 4$.

(a) Using the disk method, the volume of D is given by

$$\pi \int_0^4 x(4-x)^2\,dx = \pi \int_0^4 \big(x^3 - 8x^2 + 16x\big)\,dx = \pi\Big(\frac{1}{4}x^4 - \frac{8}{3}x^3 + 8x^2\Big)\Big|_{x=0}^{x=4} = \frac{64\pi}{3}\,.$$

It will not be easy to compute the volume of D using the shell method since it requires solving for x (in terms of y) from $y^2 = x(4-x)^2$.

(b) Using the shell method, the volume of D is given by

$$\begin{aligned} 2\pi \int_0^4 x\big[\sqrt{x(4-x)^2} - (-\sqrt{x(4-x)^2})\big]\,dx &= 4\pi \int_0^4 x(4-x)x^{\frac{1}{2}}\,dx \\ = 4\pi \int_0^4 \big(4x^{\frac{3}{2}} - x^{\frac{5}{2}}\big)\,dx &= 4\pi\Big(\frac{8}{5}x^{\frac{5}{2}} - \frac{2}{7}x^{\frac{7}{2}}\Big)\Big|_{x=0}^{x=4} = \frac{2048\pi}{35}\,. \end{aligned}$$

(c) Using the shell method, the volume of D is given by

$$\begin{aligned} 2\pi \int_0^4 (4-x)\big[\sqrt{x(4-x)^2} - (-\sqrt{x(4-x)^2})\big]\,dx &= 4\pi \int_0^4 (x-4)^2 x^{\frac{1}{2}}\,dx \\ = 4\pi \int_0^4 \big(x^{\frac{5}{2}} - 8x^{\frac{3}{2}} + 16x^{\frac{1}{2}}\big)\,dx &= 4\pi\Big(\frac{2}{7}x^{\frac{7}{2}} - \frac{16}{5}x^{\frac{5}{2}} + \frac{32}{3}x^{\frac{3}{2}}\Big)\Big|_{x=0}^{x=4} = \frac{8192\pi}{105}\,. \end{aligned}$$

7.4 Arc Length and Surfaces of Revolution

7.4.1 Arc length

In this sub-section we consider the arc length of a curve which is the graph of a function on a closed interval. Let $f : [a,b] \to \mathbb{R}$ be a continuous function, and $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[a,b]$. Then the arc length of the graph of f on $[a,b]$ can be approximated by

$$\sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2}\,,$$

where for each k, the number $\sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2}$ is the length of the line segment joining points $\big(x_k, f(x_k)\big)$ and $\big(x_{k-1}, f(x_{k-1})\big)$.

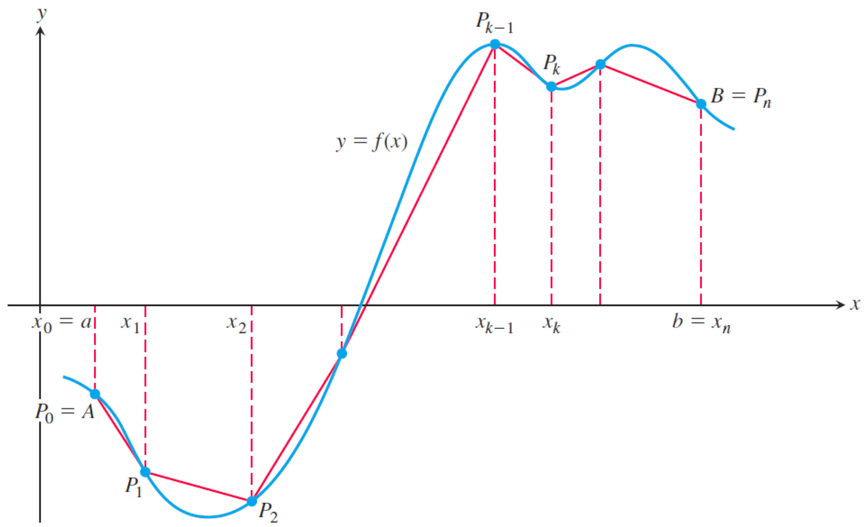


Figure 7.4: The length of the polygonal path $P_0P_1P_2\cdots P_n$ approximates the arc length of the graph of f on $[a, b]$

With Δx_k denoting $x_k - x_{k-1}$, then

$$\sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2} = \sum_{k=1}^{n} \sqrt{1 + \Big(\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}\Big)^2} \Delta x_k \,.$$

If f is differentiable on (a, b), then the Mean Value Theorem implies that for each $1 \leqslant k \leqslant n$ there exists $c_k \in (x_{k-1}, x_k)$ such that

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} = f'(c_k)\,;$$

thus

$$\sum_{k=1}^{n} \sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2} = \sum_{k=1}^{n} \sqrt{1 + f'(c_k)^2} \Delta x_k$$

which is a Riemann sum of the function $y = \sqrt{1 + f'(x)^2}$ for partition $\mathcal{P}$. Therefore, if f is continuously differentiable on $[a, b]$; that is, f' is continuous on $[a, b]$, then $\sqrt{1 + f'(x)^2}$ is Riemann integrable on $[a, b]$. Therefore, using the arguments in Section 7.2, we find that the arc length of the graph of a continuously differentiable function f on $[a, b]$ is

$$\int_a^b \sqrt{1 + f'(x)^2}\, dx \,.$$

Example 7.15. Compute the perimeter of a circle with radius r.

Let $f(x) = \sqrt{r^2 - x^2}$. Then the perimeter of a circle with radius r is the same as twice the arc length of the graph of f on $[-r, r]$. Therefore, the perimeter of a circle with radius r is

$$\begin{aligned} 2\int_{-r}^{r} \sqrt{1 + f'(x)^2}\, dx = 2\int_{-r}^{r} \sqrt{1 + \frac{x^2}{r^2 - x^2}}\, dx = 2r\int_{-r}^{r} \frac{1}{\sqrt{r^2 - x^2}}\, dx \\ = 2r\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{\sqrt{r^2 - r^2\sin^2 u}} r\cos u\, du \\ = 2r\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{r\cos u}{r\cos u}\, du = 2r\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} du = 2\pi r\,. \end{aligned}$$

Example 7.16. The arc length of the graph of $y = \dfrac{x^3}{6} + \dfrac{1}{2x}$ on the interval $\Big[\dfrac{1}{2}, 2\Big]$ is

$$\begin{aligned} \int_{\frac{1}{2}}^{2} \sqrt{1 + \Big[\frac{d}{dx}\Big(\frac{x^3}{6} + \frac{1}{2x}\Big)\Big]^2}\, dx &= \int_{\frac{1}{2}}^{2} \sqrt{1 + \Big[\frac{x^2}{2} - \frac{1}{2x^2}\Big]^2}\, dx \\ &= \int_{\frac{1}{2}}^{2} \sqrt{1 + \frac{x^4}{4} - \frac{1}{2} + \frac{1}{4x^4}}\, dx = \int_{\frac{1}{2}}^{2} \sqrt{\Big(\frac{x^2}{2} + \frac{1}{2x^2}\Big)^2}\, dx \\ &= \int_{\frac{1}{2}}^{2} \Big(\frac{x^2}{2} + \frac{1}{2x^2}\Big)\, dx = \Big(\frac{x^3}{6} - \frac{1}{2x}\Big)\Big|_{x=\frac{1}{2}}^{x=2} = \frac{33}{16}\,. \end{aligned}$$

Example 7.17. The arc length of the graph of $y = \ln(\cos x)$ from $x = 0$ to $x = \dfrac{\pi}{4}$ is

$$\begin{aligned} \int_0^{\frac{\pi}{4}} \sqrt{1 + \Big(\frac{d}{dx}\ln(\cos x)\Big)^2}\, dx = \int_0^{\frac{\pi}{4}} \sqrt{1 + \frac{\sin^2 x}{\cos^2 x}}\, dx = \int_0^{\frac{\pi}{4}} \sqrt{1 + \tan^2 x}\, dx \\ = \int_0^{\frac{\pi}{4}} \sec x\, dx = \ln|\sec x + \tan x|\Big|_{x=0}^{x=\frac{\pi}{4}} \\ = \ln(\sqrt{2} + 1) - \ln 1 = \ln(\sqrt{2} + 1)\,. \end{aligned}$$

Let f be continuously differentiable on $[a, b]$. Then the arc length of the graph of f on $[a, x]$, where $x \in [a, b]$, is given by

$$s(x) = \int_a^x \sqrt{1 + f'(t)^2}\, dx\,.$$

The fundamental theorem of Calculus then shows that

$$s'(x) = \frac{ds}{dx}(x) = \sqrt{1 + f'(x)^2}$$

or equivalently,

$$\frac{ds}{dx} = \sqrt{1 + \Big(\frac{dy}{dx}\Big)^2}\,.$$

Symbolically, $ds = \sqrt{dx^2 + dy^2}$; thus the arc length of the graph of a function is $\displaystyle\int ds$. This variable s is usually called the arc length parameter.

7.4.2 Surface of Revolution

In this section we consider the surface area of a surface formed by revolving a curve about a line (again, this line is called the axis of revolution and usually is a line parallel to the x-axis or y-axis). Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and $\mathcal{S}$ be the surface formed by revolving the graph of f on $[a, b]$ about the x-axis. The general procedures shown in the previous sections is first finding a way to compute an approximated value of the surface area and then see what is the limit of this approximation as $\|\mathcal{P}\|$ approaches 0.

We first try the following idea: let $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_n = b\}$ be a partition of $[a, b]$ and $\Delta x_k = x_k - x_{k-1}$. Consider the sum

$$\sum_{k=1}^{n} 2\pi f(c_k)\Delta x_k\,, \qquad c_k \in [x_{k-1}, x_k]$$

which is the sum of the area of cylinders formed by revolving the graph of the constant function $y = f(c_k)$ on $[x_{k-1}, x_k]$ about the x-axis. Since the sum above is a Riemann sum of the function $y = 2\pi f(x)$ for partition $\mathcal{P}$, we expect that if $f : [a, b] \to \mathbb{R}$ is continuous, then as $\|\mathcal{P}\| \to 0$ the sum approaches

$$2\pi \int_a^b f(x)\,dx\,.$$

If this is true, then the surface of the sphere with radius r is given by

$$2\pi \int_{-r}^{r} \sqrt{r^2 - x^2}\,dx = 2\pi \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cos^2 u\,du = 2\pi r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1 + \cos(2u)}{2}\,du = \pi^2 r^2$$

which is definitely not the correct area of the sphere with radius r. What is wrong with this idea?

The mistake is due to the fact that the area of surface of revolution has to be approximated by the sum of the lateral surface area of frustum of right circular cones rather than sum of lateral surface area of cylinders. The lateral area of the frustum in Figure 7.5 below

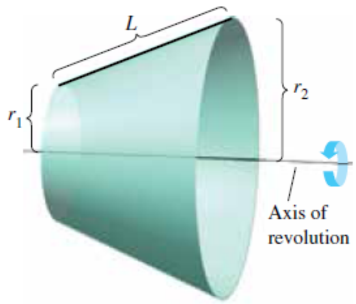


Figure 7.5

is given by $2\pi rL$, where $r = \dfrac{r_1 + r_2}{2}$; thus the surface area of $\mathcal{S}$ can be approximated by

$$\begin{aligned}\sum_{k=1}^{n} 2\pi \frac{f(x_k) + f(x_{k-1})}{2} \sqrt{(x_k - x_{k-1})^2 + \big(f(x_k) - f(x_{k-1})\big)^2} \\ = \sum_{k=1}^{n} 2\pi \frac{f(x_k) + f(x_{k-1})}{2} \sqrt{1 + f'(c_k)^2}\Delta x_k\end{aligned}$$

It can be shown that the sum above approaches $\displaystyle\int_a^b 2\pi f(x)\sqrt{1+f'(x)^2}\,dx$ as $\|\mathcal{P}\|$ approaches 0. Therefore, the area of the surface formed by revolving the graph of f on $[a, b]$ about the x-axis is given by

$$2\pi \int_a^b \big|f(x)\big|\sqrt{1+f'(x)^2}\,dx\,.$$

In general, the area of the surface formed by revolving the graph of f on $[a, b]$ about $y = L$ is given by

$$2\pi \int_a^b \big|f(x) - L\big|\sqrt{1+f'(x)^2}\,dx\,.$$

Example 7.18. The surface area of a sphere with radius r is given by

$$2\pi \int_{-r}^{r} \big(\sqrt{r^2 - x^2} - 0\big)\sqrt{1 + \frac{x^2}{r^2 - x^2}}\,dx = 2\pi \int_{-r}^{r} \sqrt{r^2}\,dx = 4\pi r^2\,,$$

where we treat the sphere as a surface formed by revolving the graph of $y = \sqrt{r^2 - x^2}$ about the x-axis.

Example 7.19. In this example we consider the area of the surface formed by revolving the (upper part) ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ (or the graph of $y = \dfrac{b}{a}\sqrt{a^2 - x^2}$ on $[-a, a]\big)$ about the

x-axis. Using the formula above, we find that the surface area is given by

$$\begin{aligned}
&2\pi\int_{-a}^{a}\frac{b}{a}\sqrt{a^2-x^2}\sqrt{1+\frac{b^2}{a^2}\frac{x^2}{a^2-x^2}}\,dx=\frac{2\pi b}{a}\int_{-a}^{a}\sqrt{a^2-x^2+\frac{b^2}{a^2}x^2}\,dx\\
&\quad=\frac{2\pi b}{a^2}\int_{-a}^{a}\sqrt{a^4-(a^2-b^2)x^2}\,dx=2\pi b\int_{-a}^{a}\sqrt{1-\frac{a^2-b^2}{a^4}x^2}\,dx\,.
\end{aligned}$$

1. Suppose that $a>b$; that is, x-axis is the major axis. Let $c=\dfrac{\sqrt{a^2-b^2}}{a^2}$. Then the substitution $x=\dfrac{1}{c}\sin u$ implies that

$$\begin{aligned}
&2\pi\int_{-a}^{a}\frac{b}{a}\sqrt{a^2-x^2}\sqrt{1+\frac{b^2}{a^2}\frac{x^2}{a^2-x^2}}\,dx=2\pi b\int_{-a}^{a}\sqrt{1-c^2x^2}\,dx\\
&\quad=2\pi b\int_{-\arcsin(ac)}^{\arcsin(ac)}\sqrt{1-\sin^2 u}\cdot\frac{1}{c}\cos u\,du\\
&\quad=\frac{2\pi b}{c}\int_{-\arcsin(ac)}^{\arcsin(ac)}\cos^2 u\,du=\frac{2\pi b}{c}\int_{-\arcsin(ac)}^{\arcsin(ac)}\frac{1+\cos(2u)}{2}\,du\\
&\quad=\frac{2\pi b}{c}\Big(\frac{u}{2}+\frac{\sin(2u)}{4}\Big)\Big|_{u=-\arcsin(ac)}^{u=\arcsin(ac)}\\
&\quad=\frac{2\pi a^2 b}{\sqrt{a^2-b^2}}\Big[\arcsin\frac{\sqrt{a^2-b^2}}{a}+\frac{b\sqrt{a^2-b^2}}{a^2}\Big]\\
&\quad=\frac{2\pi a^2 b}{\sqrt{a^2-b^2}}\arcsin\frac{\sqrt{a^2-b^2}}{a}+2\pi b^2\,.
\end{aligned}$$

2. Suppose that $a<b$; that is, x-axis is the minor axis. Let $c=\dfrac{\sqrt{b^2-a^2}}{a^2}$. Then similar to the previous case, the substitution $x=\dfrac{1}{c}\sinh u$ implies that

$$\begin{aligned}
&2\pi\int_{-a}^{a}\frac{b}{a}\sqrt{a^2-x^2}\sqrt{1+\frac{b^2}{a^2}\frac{x^2}{a^2-x^2}}\,dx=2\pi b\int_{-\sinh^{-1}(ac)}^{\sinh^{-1}(ac)}\sqrt{1+\sinh^2 u}\cdot\frac{1}{c}\cosh u\,du\\
&\quad=\frac{2\pi b}{c}\int_{-\sinh^{-1}(ac)}^{\sinh^{-1}(ac)}\cosh^2 u\,du=\frac{2\pi b}{c}\int_{-\sinh(ac)}^{\sinh^{-1}(ac)}\frac{1+\cosh(2u)}{2}\,du\\
&\quad=\frac{2\pi b}{c}\Big(\frac{u}{2}+\frac{\sinh(2u)}{4}\Big)\Big|_{u=-\sinh^{-1}(ac)}^{u=\sinh^{-1}(ac)}\\
&\quad=\frac{2\pi a^2 b}{\sqrt{b^2-a^2}}\Big[\sinh^{-1}\frac{\sqrt{a^2-b^2}}{a}+\frac{\sqrt{a^2-b^2}}{a}\cosh\Big(\sinh^{-1}\frac{\sqrt{b^2-a^2}}{a}\Big)\Big]\\
&\quad=\frac{2\pi a^2 b}{\sqrt{b^2-a^2}}\sinh^{-1}\frac{\sqrt{a^2-b^2}}{a}+2\pi b^2\,.
\end{aligned}$$

7.5 Moments, Centers of Mass, and Centroids

• Center of mass in a one-dimensional system

Let $m_1, m_2, \cdots, m_n$ be n point masses located at $x_1, x_2, \cdots, x_n$ on a (massless) rigid x-axis supported by a fulcrum at the origin.

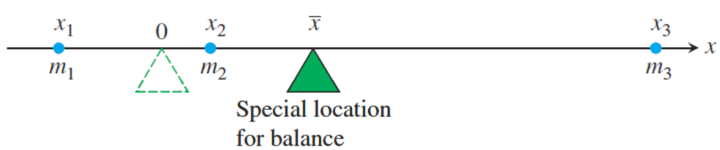
Each mass m_k exerts a downward force $m_k g$ (which is negative), and each of these forces has a tendency to turn the x-axis about the origin. This turning effect, called a torque, is measured by multiplying the force $m_k g$ by the signed distance x_k from the point of application to the origin. By convention, a positive torque induces a counterclockwise turn.

The sum of these torques measures the tendency of the system to rotate about the fultrum/origin. This sum is called the system torque; thus

$$\text{System torque} = m_1 g x_1 + m_2 g x_2 + \cdots + m_n g x_n = g(m_1 x_1 + m_2 x_2 + \cdots + m_n x_n)\,.$$

The system will balance if and only if its torque is zero. The number $M_0 \equiv m_1 x_1 + m_2 x_2 + \cdots + m_n x_n$ is called the moment of the system about the origin, and is the sum of moments $m_1 x_1$, $m_2 x_2$, $\cdots$, $m_n x_n$ of individual masses. If M_0 is 0, then the system is said to be in equilibrium.

For a system that is not in equilibrium, the center of mass (of the system) is defined as the point $\bar{x}$ at which the fulcrum could be relocated to attain equilibrium.



Such an $\bar{x}$ must satisfy

$$0 = m_1(x_1 - \bar{x}) + m_2(x_2 - \bar{x}) + \cdots + m_n(x_n - \bar{x})$$

which implies that

$$\bar{x} = \frac{\sum\limits_{i=1}^{n} m_i x_i}{\sum\limits_{i=1}^{n} m_i} = \frac{\text{moment of system about the origin}}{\text{total mass of system}}\,.$$

Definition 7.20

Let the point masses $m_1, m_2, \cdots, m_n$ be located at x_1, x_2, $\cdots$, x_n (on a coordinate line).

1. The moment about the origin is
$$M_0 = m_1x_1 + m_2x_2 + \cdots + m_nx_n\,.$$
2. The center of mass $\bar{x}$ is $\dfrac{M_0}{m}$, where $m = m_1 + m_2 + \cdots + m_n$ is the total mass of the system.

- **Center of mass in a two-dimensional system**

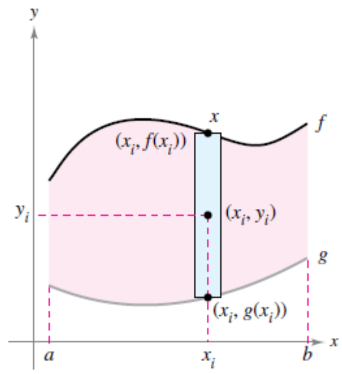
Definition 7.21

Let the point masses $m_1, m_2, \cdots, m_n$ be located at (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n) (on a plane).

1. The moment about the y-axis is
$$M_y = m_1x_1 + m_2x_2 + \cdots + m_nx_n\,.$$
2. The moment about the x-axis is
$$M_x = m_1y_1 + m_2y_2 + \cdots + m_ny_n\,.$$
3. The center of mass $(\bar{x}, \bar{y})$ is
$$\bar{x} = \frac{M_y}{m} \quad \text{and} \quad \bar{y} = \frac{M_x}{m}\,,$$
where $m = m_1 + m_2 + \cdots + m_n$ is the total mass of the system.

- **Center of mass of a planar lamina**

Consider an irregularly shaped thin flat plate of material (called lamina) of uniform density ϱ (a measure of mass per unit of area), bounded by the graphs of $y = f(x)$, $y = g(x)$, and $x = a$, $x = b$, as shown in the following figure.



Then the density of this region is

$$m = \varrho \int_a^b \big[f(x) - g(x)\big]\, dx = \varrho A\,,$$

where A is the area of this region.

Partition $[a, b]$ into n sub-intervals with equal width Δx, and let x_i be the mid-point of the i-th sub-interval. The area of the portion on the i-th sub-interval can be approximated by $\big[f(x_i) - g(x_i)\big]\Delta x$; thus the mass of the portion on the i-th sub-interval can be approximated by $\varrho\big[f(x_i) - g(x_i)\big]\Delta x$. Now, considering this mass to be located at the center $\big(x_i, \dfrac{f(x_i) + g(x_i)}{2}\big)$, the moment of this mass about the x-axis is

$$\varrho\big[f(x_i) - g(x_i)\big]\Delta x \frac{f(x_i) + g(x_i)}{2}\,.$$

Summing all the moments and passing to the limit as $n \to \infty$ suggest the following

Definition 7.22

Let $f, g : [a, b] \to \mathbb{R}$ be continuous such that $f(x) \geqslant g(x)$ for all $x \in [a, b]$, and consider the lamina of uniform density ϱ bounded by the graphs of f, g and the lines $x = a$, $x = b$.

1. The moment about the x-axis and the y-axis are

$$M_x = \frac{\varrho}{2} \int_a^b \big[f(x)^2 - g(x)^2\big]\, dx \quad \text{and} \quad M_y = \varrho \int_a^b x\big[f(x) - g(x)\big]\, dx\,.$$

2. The center of mass $(\bar{x}, \bar{y})$ is given by $\bar{x} = \dfrac{M_y}{m}$ and $\bar{y} = \dfrac{M_x}{m}$, where $m = \varrho \displaystyle\int_a^b \big[f(x) - g(x)\big]\, dx$ is the mass of the lamina.

The center of mass of a lamina of uniform density depends only on the shape of the lamina but not on its density. For this reason, the center of mass of a region in the plain is also called the centroid of the region.

Example 7.22. Compute the centroid of a triangle with vertex $(0,0)$, (a,b_1) and (a,b_2), where $a>0$ and $b_1<b_2$.

Let $f(x)=\dfrac{b_2}{a}x$ and $g(x)=\dfrac{b_1}{a}x$. Then the triangle given above is the region bounded by the graphs of f, g and $x=a$. Assume uniform density $\varrho=1$. Then the moment of the region about the x-axis is

$$M_x=\frac{1}{2}\int_0^a\Big(\frac{b_2^2}{a^2}-\frac{b_1^2}{a^2}\Big)x^2\,dx=\frac{a(b_2^2-b_1^2)}{6}$$

and the moment of the region about the y-axis is

$$M_y=\int_0^a x\Big[\frac{b_2}{a}-\frac{b_1}{a}\Big]x\,dx=\frac{a^2(b_2-b_1)}{3}\,,$$

as well as the total mass

$$m=\int_0^a\Big[\frac{b_2}{a}-\frac{b_1}{a}\Big]x\,dx=\frac{a(b_2-b_1)}{2}\,.$$

Therefore, the centroid of the given triangle is

$$(\bar{x},\bar{y})=\Big(\frac{2a}{3},\frac{b_1+b_2}{3}\Big)\,.$$

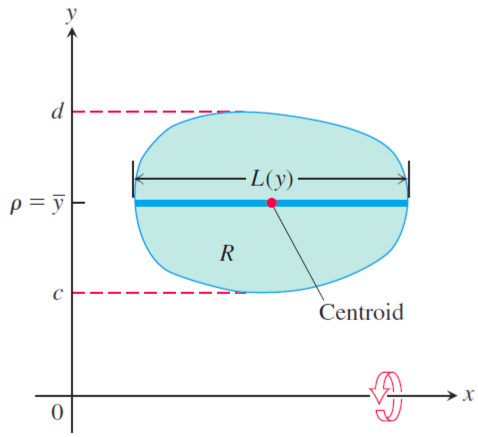
Theorem 7.23: Pappus

Let R be a region in a plane and L be a line in the same plane such that L does not intersect the interior of R. If r is the distance between the centroid of R and the line, then the volume V of the solid of revolution formed by revolving R about the line is

$$V=2\pi rA\,,$$

where A is the area of R.

Proof. We draw the axis of revolution as the x-axis with the region R in the first quadrant (see figure below).



Let $L(y)$ be the length of the cross section of R perpendicular to the y-axis at y, and we assume that L is continuous on $[c, d]$. Then the area of R is given by

$$A = \int_c^d L(y)\,dy\,,$$

and the shell method implies that the volume of the solid formed by revolving R about the x-axis is

$$V = 2\pi \int_c^d yL(y)\,dy\,.$$

On the other hand, if r denotes the distance between the centroid of R and the x-axis, then r is the y-coordinate of the centroid of R and is given by

$$r = \frac{\text{the moment of the region about the } x\text{-axis}}{\text{the total mass of the region}} = \frac{\displaystyle\int_c^d yL(y)\,dy}{\displaystyle\int_c^d L(y)\,dy}$$

which validates the relation $V = 2\pi rA$. $\square$

Example 7.24. Using the Pappus theorem, the volume of the solid torus given in Example 7.2 is

$$2\pi a(\pi r^2) = 2\pi^2 ar^2$$

since the centroid of a disk is the center of the disk.

Chapter 8

Integration Techniques and Improper Integrals

8.1 Basic Integration Rules

We recall the following formula:

1. Let f, g be functions and k be a constant. Then

$$\int kf(x)\,dx = k\int f(x)\,dx\,, \qquad \int (f+g)(x)\,dx = \int f(x)\,dx + \int g(x)\,dx\,.$$

2. Let r be a real number. Then

$$\int x^r\,dx = \begin{cases} \dfrac{1}{r+1}x^{r+1} + C & \text{if } r \neq -1\,, \\ \ln x + C & \text{if } r = -1\,. \end{cases}$$

3. If $a > 0$, then $\displaystyle\int a^x\,dx = \frac{1}{\ln a}a^x + C$. In particular, $\displaystyle\int e^x\,dx = e^x + C$.

4. If $a \neq 0$, $\displaystyle\int \sin(ax)\,dx = -\frac{1}{a}\cos(ax) + C$, $\displaystyle\int \cos(ax)\,dx = \frac{1}{a}\sin(ax) + C$,
$\displaystyle\int \tan(ax)\,dx = \frac{1}{a}\ln|\sec(ax)| + C$, $\displaystyle\int \cot(ax)\,dx = \frac{1}{a}\ln|\sin(ax)| + C$,
$\displaystyle\int \sec(ax)\,dx = \frac{1}{a}\ln|\sec(ax)+\tan(ax)|+C$, $\displaystyle\int \csc x\,dx = -\frac{1}{a}\ln|\csc(ax)+\cot(ax)|+C$.

5. $\displaystyle\int \sec^2 x\,dx = \tan x + C$, $\displaystyle\int \sec x\tan x\,dx = \sec x + C$.

6. If $a > 0$, then

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + C\,, \qquad \int \frac{dx}{a^2 + x^2} = \frac{1}{a} \arctan \frac{x}{a} + C$$

$$\int \frac{dx}{x\sqrt{x^2 - a^2}} = \frac{1}{a} \arctan \frac{\sqrt{x^2 - a^2}}{a} + C\,.$$

Example 8.1. Find the indefinite integrals $\displaystyle\int \frac{4}{x^2+9}\,dx$, $\displaystyle\int \frac{4x}{x^2+9}\,dx$ and $\displaystyle\int \frac{4x^2}{x^2+9}\,dx$.

From the formula above, it is easy to see that

$$\int \frac{4}{x^2+9}\,dx = \frac{4}{3} \arctan \frac{x}{3} + C\,.$$

Noting that $\displaystyle\frac{4x}{x^2+9} = 2\frac{\frac{d}{dx}(x^2+9)}{x^2+9}$, using the formula $\displaystyle\frac{d}{dx} \ln |f(x)| = \frac{f'(x)}{f'(x)}$, we find that

$$\int \frac{4x}{x^2+9}\,dx = 2\ln|x^2+9| + C = 2\ln(x^2+9) + C\,.$$

Finally, noting that $\displaystyle\frac{4x^2}{x^2+9} = \frac{4(x^2+9)-36}{x^2+9} = 4 - \frac{36}{x^2+9}$, by the formula above we find that

$$\int \frac{4x^2}{x^2+9}\,dx = 4x - 12\arctan\frac{x}{3} + C\,.$$

Example 8.2. Find the indefinite integrals $\displaystyle\int \frac{3}{\sqrt{4-x^2}}\,dx$, $\displaystyle\int \frac{3x}{\sqrt{4-x^2}}\,dx$ and $\displaystyle\int \frac{3x^2}{\sqrt{4-x^2}}\,dx$.

From the formula above,

$$\int \frac{3}{\sqrt{4-x^2}}\,dx = 3\arcsin\frac{x}{2} + C\,.$$

For the second integral, we let $4 - x^2 = u$. Then $-2xdx = du$; thus

$$\int \frac{3x}{\sqrt{4-x^2}}\,dx = -\frac{3}{2}\int u^{-\frac{1}{2}}\,du = -\frac{3}{2}\frac{1}{1-\frac{1}{2}}u^{\frac{1}{2}} + C = -3(4-x^2)^{\frac{1}{2}} + C\,.$$

For the third integral, first we observe that

$$\int \frac{3x^2}{\sqrt{4-x^2}}\,dx = \int \frac{3(x^2-4)}{\sqrt{4-x^2}}\,dx + \int \frac{12}{\sqrt{4-x^2}}\,dx = -3\int \sqrt{4-x^2}\,dx + 12\arcsin\frac{x}{2}\,.$$

Let $x = 2\sin u$. Then $dx = 2\cos u\,du$; thus

$$\begin{aligned}\int \sqrt{4-x^2}\,dx &= \int \sqrt{4(1-\sin^2 u)}\cdot 2\cos u\,du = \int 4\cos^2 u\,du = \int \big[2+2\cos(2u)\big]\,du\\ &= 2u + \sin(2u) + C = 2u + 2\sin u\cos u + C\\ &= 2\arcsin\frac{x}{2} + x\sqrt{1-\frac{x^2}{4}} + C = 2\arcsin\frac{x}{2} + \frac{x\sqrt{4-x^2}}{2} + C\,.\end{aligned}$$

Therefore,

$$\int \frac{3x^2}{\sqrt{4-x^2}}\,dx = 6\arcsin\frac{x}{2} - \frac{3}{2}x\sqrt{4-x^2} + C\,.$$

Remark 8.3. One should add

$$\int \frac{x}{\sqrt{a^2-x^2}}\,dx = -\sqrt{a^2-x^2} + C \qquad \text{and} \qquad \int \frac{x}{\sqrt{a^2+x^2}}\,dx = \sqrt{a^2+x^2} + C$$

into the table of integrations.

Example 8.4. Find the indefinite integral $\displaystyle\int \frac{dx}{1+e^x}$.

Let $u = 1 + e^x$. Then $du = e^x dx$ which implies that $dx = \dfrac{du}{u-1}$. Therefore,

$$\begin{aligned}\int \frac{dx}{1+e^x} &= \int \frac{du}{u(u-1)} = \int \Big(\frac{1}{u-1} - \frac{1}{u}\Big)\,du = \ln|u-1| - \ln|u| + C \\ &= x - \ln(1+e^x) + C\,.\end{aligned}$$

Another way of finding the integral is by observing that

$$\frac{1}{1+e^x} = \frac{1+e^x}{1+e^x} - \frac{e^x}{1+e^x} = 1 - \frac{\frac{d}{dx}(1+e^x)}{1+e^x}\,;$$

thus using the formula $\dfrac{d}{dx}\ln|f(x)| = \dfrac{f'(x)}{f(x)}$, we find that

$$\int \frac{dx}{1+e^x} = x - \ln(1+e^x) + C\,.$$

8.2 Integration by Parts - 分部積分

Suppose that u, v are two differentiable functions of x. Then the product rule implies that

$$\frac{d}{dx}(uv) = \frac{du}{dx}v + u\frac{dv}{dx}\,.$$

Therefore, if $\dfrac{du}{dx}v$ and $u\dfrac{dv}{dx}$ are Riemann integrable (on the interval of interests),

$$\int \frac{du}{dx}v\,dx + \int u\frac{dv}{dx}\,dx = (uv)(x) + C\,.$$

Symbolically, we write $\dfrac{du}{dx}v\,dx$ ad $v\,du$ and $u\dfrac{dv}{dx}\,dx$ as $u\,dv$, the formula above implies that

$$\int udv = uv - \int vdu\,.$$

Theorem 8.5: Integration by Parts

If u and v are functions of x and have continuous derivatives, then

$$\int u\,dv = uv - \int v\,du\,.$$

Example 8.6. Find the indefinite integral $\displaystyle\int \ln x\,dx$. Recall that we have shown that

$$\int \ln x\,dx = x\ln x - x + C$$

using the Riemann sum. Let $u = \ln x$ and $v = x$ (so that $dv = dx$). Then integration by parts shows that

$$\int \ln x\,dx = x\ln x - \int x\,d(\ln x) = x\ln x - \int x\cdot\frac{1}{x}\,dx = x\ln x - \int dx = x\ln x - x + C\,.$$

Example 8.7. Find the indefinite integral $\displaystyle\int x\cos x\,dx$. Recall that we have shown that

$$\int x\cos x\,dx = x\sin x + \cos x + C$$

using the Riemann sum. Let $u = x$ and $v = \sin x$ (so that $dv = \cos x\,dx$). Then integration by parts shows that

$$\int x\cos x\,dx = x\sin x - \int \sin x\,dx = x\sin x + \cos x + C\,.$$

Principles of applying integration by parts: Choose u and v such that $v\,du$ has simpler form than $u\,dv$, and this is usually achieved by

1. finding u such that the derivative of u is a function simpler than u, or
2. finding v such that the derivative of v is more complicate than v.

Example 8.8. Find the indefinite integral $\displaystyle\int xe^x\,dx$.

Let $u = x$ and $v = e^x$ (so that $dv = e^x dx$). Then integration by parts shows that

$$\int xe^x\,dx = xe^x - \int e^x\,dx = (x-1)e^x + C\,.$$

Example 8.9. Find the indefinite integral $\displaystyle\int x^r \ln x\,dx$, where r is a real number.

Suppose first that $r \neq -1$. Let $u = \ln x$ and $v = \dfrac{1}{r+1}x^{r+1}$. Then integration by parts shows that

$$\begin{aligned}\int x^r \ln x\,dx &= \frac{1}{r+1}x^{r+1}\ln x - \int \frac{1}{r+1}x^{r+1}\cdot\frac{1}{x}\,dx = \frac{1}{r+1}x^{r+1}\ln x - \frac{1}{r+1}\int x^r\,dx \\ &= \frac{1}{r+1}x^{r+1}\ln x - \frac{1}{(r+1)^2}x^{r+1} + C\,.\end{aligned}$$

Now if $r = -1$. Let $u = v = \ln x$. Then integration by parts implies that

$$\int x^{-1}\ln x\,dx = (\ln x)^2 - \int \ln x\cdot\frac{1}{x}\,dx = (\ln x)^2 - \int x^{-1}\ln x\,dx$$

which implies that

$$\int x^{-1}\ln x\,dx = \frac{1}{2}(\ln x)^2 + C\,.$$

Therefore,

$$\int x^r \ln x\,dx = \begin{cases} \dfrac{1}{r+1}x^{r+1}\ln x - \dfrac{1}{(r+1)^2}x^{r+1} + C & \text{if } r \neq -1\,, \\ \dfrac{1}{2}(\ln x)^2 + C & \text{if } r = -1\,. \end{cases}$$

Example 8.10. Find the indefinite integral $\displaystyle\int x^2\cos x\,dx$.

Let $u = x^2$ and $v = \sin x$ (so that $dv = \cos x\,dx$). Then integration by parts shows that

$$\int x^2\cos x\,dx = x^2\sin x - \int \sin x\cdot 2x\,dx = x^2\sin x - 2\int x\sin x\,dx\,.$$

Integrating by parts again, we find that

$$\int x\sin x\,dx = -x\cos x + \int\cos x\,dx = -x\cos x + \sin x + C\,;$$

thus we obtain the

$$\int x^2\cos x\,dx = x^2\sin x + 2x\cos x - 2\sin x + C\,.$$

Example 8.11. Find the indefinite integrals $\displaystyle\int e^{ax}\sin(bx)\,dx$ and $\displaystyle\int e^{ax}\cos(bx)\,dx$, where a, b are non-zero constants.

Let $u = \sin(bx)$ (or $u = \cos(ax)$) and $v = a^{-1}e^{ax}$ (so that $dv = e^{ax}\,dx$). Then

$$\int e^{ax}\sin(bx)\,dx = \frac{1}{a}e^{ax}\sin(bx) - \frac{b}{a}\int e^{ax}\cos(bx)\,dx\,,$$
$$\int e^{ax}\cos(bx)\,dx = \frac{1}{a}e^{ax}\cos(bx) + \frac{b}{a}\int e^{ax}\cos(bx)\,dx\,.$$

The two identities above further imply that

$$\begin{aligned}\int e^{ax}\sin(bx)\,dx &= \frac{1}{a}e^{ax}\sin(bx) - \frac{b}{a}\int e^{ax}\cos(bx)\,dx\\ &= \frac{1}{a}e^{ax}\sin(bx) - \frac{b}{a}\Big[\frac{1}{a}e^{ax}\cos(bx) + \frac{b}{a}\int e^{ax}\sin(bx)\,dx\Big]\\ &= \frac{1}{a}e^{ax}\sin(bx) - \frac{b}{a^2}e^{ax}\cos(bx) - \frac{b^2}{a^2}\int e^{ax}\sin(bx)\,dx\,;\end{aligned}$$

thus

$$\int e^{ax}\sin(bx)\,dx = \frac{1}{a^2+b^2}\big[ae^{ax}\sin(bx) - be^{ax}\cos(bx)\big] + C\,. \tag{8.2.1}$$

Similarly,

$$\int e^{ax}\cos(bx)\,dx = \frac{1}{a^2+b^2}\big[ae^{ax}\cos(bx) + be^{ax}\sin(bx)\big] + C\,. \tag{8.2.2}$$

Remark 8.12. By the Euler identity (5.9.1), $\int e^{ax}\sin(bx)\,dx$ and $\int e^{ax}\cos(bx)\,dx$ are the real and imaginary part of the integral $\int e^{ax}e^{ibx}\,dx$. By the fact that $e^{ax}e^{ibx} = e^{(a+ib)x}$ and pretending that $\int e^{cx}\,dx = \frac{1}{c}e^{cx} + C$ for complex number c, we find that

$$\begin{aligned}\int e^{ax}e^{ibx}\,dx &= \frac{1}{a+ib}e^{(a+ib)x} + C = \frac{1}{a+ib}e^{ax}\big[\cos(bx) + i\sin(bx)\big] + C\\ &= \frac{a-ib}{a^2+b^2}e^{ax}\big[\cos(bx) + i\sin(bx)\big] + C\\ &= \frac{e^{ax}}{a^2+b^2}\big[a\cos(bx) + b\sin(bx) + i\big(a\sin(bx) - b\cos(bx)\big)\big] + C\,;\end{aligned}$$

thus we conclude (8.2.1) and (8.2.2).

Example 8.13. Find the indefinite $\int x^n e^{ax}\,dx$, $\int x^n\sin(ax)\,dx$ and $\int x^n\cos(ax)\,dx$, where $a > 0$ is a constant.

Let $u = x^n$ and $v = a^{-1}e^{ax}$ (so that $dv = e^{ax}\,dx$), $v = -a^{-1}\cos(ax)$ (so that $dv = \sin(ax)$) and $v = a^{-1}\sin(ax)$ (so that $dv = \cos(ax)$) in these three cases. Then

$$\int x^n e^{ax}\,dx = \frac{1}{a}x^n e^{ax} - \int \frac{1}{a}e^{ax}\cdot nx^{n-1}\,dx = \frac{1}{a}x^n e^{ax} - \frac{n}{a}\int x^{n-1}e^{ax}\,dx\,.$$

Moreover,

$$\int x^n \sin(ax)\,dx = -\frac{1}{a}x^n\cos(ax) + \frac{n}{a}\int x^{n-1}\cos(ax)\,dx\,,$$
$$\int x^n \cos(ax)\,dx = \frac{1}{a}x^n\sin(ax) - \frac{n}{a}\int x^{n-1}\sin(ax)\,dx\,.$$

The two identities above further imply that the following recurrence relations

$$\int x^n \sin(ax)\,dx = -\frac{1}{a}x^n\cos(ax) + \frac{n}{a^2}x^{n-1}\sin(ax) - \frac{n(n-1)}{a^2}\int x^{n-2}\cos(ax)\,dx\,,$$
$$\int x^n \cos(ax)\,dx = \frac{1}{a}x^n\sin(ax) + \frac{n}{a^2}x^{n-1}\cos(ax) - \frac{n(n-1)}{a^2}\int x^{n-2}\sin(ax)\,dx\,.$$

Example 8.14. Using integration by parts, we have

$$\begin{aligned}\int \cos^n x\,dx &= \int \cos^{n-1}x\,d(\sin x) = \sin x\cos^{n-1}x - \int \sin x\,d(\cos^{n-1}x)\\ &= \sin x\cos^{n-1}x + (n-1)\int \sin^2 x\cos^{n-2}x\,dx\\ &= \sin x\cos^{n-1}x + (n-1)\int (1-\cos^2 x)\cos^{n-2}x\,dx\\ &= \sin x\cos^{n-1}x + (n-1)\int \cos^{n-2}x\,dx - (n-1)\int \cos^n x\,dx\,;\end{aligned}$$

thus rearranging terms, we conclude that

$$\int \cos^n x\,dx = \frac{\sin x\cos^{n-1}x}{n} + \frac{n-1}{n}\int \cos^{n-2}x\,dx\,. \tag{8.2.3}$$

Similarly,

$$\int \sin^n x\,dx = -\frac{\cos x\sin^{n-1}x}{n} + \frac{n-1}{n}\int \sin^{n-2}x\,dx\,. \tag{8.2.4}$$

Theorem 8.15: Wallis's Formulas

If n is a non-negative integer, then

$$\int_0^{\frac{\pi}{2}} \sin^{2n+1}x\,dx = \int_0^{\frac{\pi}{2}} \cos^{2n+1}x\,dx = \frac{(2^n n!)^2}{(2n+1)!}$$

and

$$\int_0^{\frac{\pi}{2}} \sin^{2n}x\,dx = \int_0^{\frac{\pi}{2}} \cos^{2n}x\,dx = \frac{(2n)!}{(2^n n!)^2}\cdot\frac{\pi}{2}\,.$$

Proof. Note that (8.2.3) implies that

$$\int_0^{\frac{\pi}{2}} \cos^n x\,dx = \frac{\sin x \cos^{n-1} x}{n}\Big|_{x=0}^{x=\frac{\pi}{2}} + \frac{n-1}{n}\int_0^{\frac{\pi}{2}} \cos^{n-2} x\,dx = \frac{n-1}{n}\int_0^{\frac{\pi}{2}} \cos^{n-2} x\,dx\,.$$

Therefore,

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos^{2n+1} x\,dx &= \frac{2n}{2n+1}\int_0^{\frac{\pi}{2}} \cos^{2n-1} x\,dx = \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\int_0^{\frac{\pi}{2}} \cos^{2n-3} x\,dx = \cdots \\ &= \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdot\frac{2n-4}{2n-3}\cdots\frac{2}{3}\int_0^{\frac{\pi}{2}} \cos x\,dx = \frac{2}{3}\cdot\frac{4}{5}\cdots\frac{2n}{2n+1} \\ &= \frac{2^2 4^2\cdots(2n)^2}{(2n+1)!} = \frac{(2^n n!)^2}{(2n+1)!}\end{aligned}$$

and

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos^{2n} x\,dx &= \frac{2n-1}{2n}\int_0^{\frac{\pi}{2}} \cos^{2n-2} x\,dx = \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\int_0^{\frac{\pi}{2}} \cos^{2n-4} x\,dx = \cdots \\ &= \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdot\frac{2n-5}{2n-4}\cdots\frac{1}{2}\int_0^{\frac{\pi}{2}} \cos^0 x\,dx = \frac{1}{2}\cdot\frac{3}{4}\cdots\frac{2n-1}{2n}\cdot\frac{\pi}{2} \\ &= \frac{(2n)!}{2^2 4^2\cdots(2n)^2}\cdot\frac{\pi}{2} = \frac{(2n)!}{(2^n n!)^2}\cdot\frac{\pi}{2}\,.\end{aligned}$$

The substitution $x = \frac{\pi}{2} - u$ shows that

$$\int_0^{\frac{\pi}{2}} \sin^n x\,dx = \int_0^{\frac{\pi}{2}} \cos^n x\,dx \qquad \text{for all non-negative integers } n\,,$$

so we conclude the theorem. □

Theorem 8.16: Stirling's Formula

$$\lim_{n\to\infty} \frac{n!}{n^{n+0.5}e^{-n}} = \sqrt{2\pi}.$$

Proof. Let $I_n = \int_0^{\frac{\pi}{2}} \sin^n x\,dx$. Then Wallis's formula shows that

$$I_{2n} = \frac{(2n)!}{(2^n n!)^2}\cdot\frac{\pi}{2} \qquad \text{and} \qquad I_{2n+1} = \frac{(2^n n!)^2}{(2n+1)!}\,.$$

Moreover, since $\sin^{2n+2} x \leqslant \sin^{2n+1} x \leqslant \sin^{2n} x$ on $\big[0, \frac{\pi}{2}\big]$, we also have $I_{2n+2} \leqslant I_{2n+1} \leqslant I_{2n}$ for all $n \geqslant 0$. Therefore,

$$\frac{I_{2n+2}}{I_{2n}} \leqslant \frac{I_{2n+1}}{I_{2n}} \leqslant 1 \qquad \forall\, n \geqslant 0\,.$$

Note that

$$\frac{I_{2n+2}}{I_{2n}} = \frac{I_{2(n+1)}}{I_{2n}} = \frac{\dfrac{(2(n+1))!}{2^{2(n+1)}((n+1)!)^2}}{\dfrac{(2n)!}{2^{2n}(n!)^2}} = \frac{2n+1}{2(n+1)}\,;$$

thus $\lim\limits_{n\to\infty} \dfrac{I_{2n+2}}{I_{2n}} = 1$. As a consequence, the Squeeze Theorem implies that $\lim\limits_{n\to\infty} \dfrac{I_{2n+1}}{I_{2n}} = 1\,.$

Let $s_n = \dfrac{n!}{n^{n+0.5}e^{-n}}$. Then the fact that the function $y = \big(1+\dfrac{1}{x}\big)^{x+0.5}$ is decreasing on $(0,\infty)$ (left as an exercise) and (5.4.3) show that $s_n \geqslant s_{n+1} \geqslant 0$ for all $n \in \mathbb{N}$. Therefore, the completeness of the real number (see Theorem 9.20) implies that $\lim\limits_{n\to\infty} s_n = s$ exists. Moreover,

$$\begin{aligned}\frac{I_{2n+1}}{I_{2n}} &= \frac{\dfrac{2^{2n}(n!)^2}{(2n+1)!}}{\dfrac{(2n)!}{2^{2n}(n!)^2}\dfrac{\pi}{2}} = \frac{2^{4n}(n!)^4}{(2n)!(2n+1)!}\cdot\frac{2}{\pi}\\ &= \frac{2^{4n}(s_n n^{n+0.5}e^{-n})^4}{s_{2n}(2n)^{2n+0.5}e^{-2n}s_{2n+1}(2n+1)^{2n+1.5}e^{-2n-1}}\cdot\frac{2}{\pi}\\ &= \frac{s_n^4}{s_{2n}s_{2n+1}}\frac{e}{2\pi}\big(1+\frac{1}{2n}\big)^{-2n-1.5}\,;\end{aligned}$$

thus (5.4.3) implies that

$$1 = \lim_{n\to\infty}\frac{I_{2n+1}}{I_{2n}} = \lim_{n\to\infty}\frac{s_n^4}{s_{2n}s_{2n+1}}\cdot\frac{1}{2\pi} = \frac{s^2}{2\pi}\,.$$

The theorem is then concluded by the fact that $s \geqslant 0$. □

8.3 Trigonometric Integrals

In this section, we are concerned with the integrals

$$\int \sin^m x\cos^n x\,dx \qquad \text{and} \qquad \int \sec^m x\tan^n x\,dx\,,$$

where m,n are non-negative integers.

8.3.1 The integral of $\sin^m x\cos^n x$

- **The case when one of m and n is odd**

Suppose $m = 2k+1$ or $n = 2\ell+1$. Write

$$\int \sin^{2k+1} x\cos^n x\,dx = \int \cos^n x(1-\cos^2 x)^k \sin x\,dx = -\int \cos^n x(1-\cos^x x)^k\,d(\cos x)$$

and

$$\int \sin^m x \cos^{2\ell+1} x\, dx = \int \sin^m x (1-\sin^2 x)^\ell \cos x\, dx = \int \sin^m x (1-\sin^2 x)^\ell\, d(\sin x)$$

so that the integral can be obtained by integrating polynomials.

Example 8.17. Find the indefinite integral $\displaystyle\int \sin^3 x \cos^4 x\, dx$.

Let $u = \cos x$. Then $du = -\sin x\, dx$; thus

$$\begin{aligned}\int \sin^3 x \cos^4 x\, dx &= \int (1-\cos^2 x)\cos^4 x \sin x\, dx = -\int (1-u^2)u^4\, du \\ &= -\frac{1}{5}u^5 + \frac{1}{7}u^7 + C = -\frac{1}{5}\cos^5 x + \frac{1}{7}\cos^7 x + C\,.\end{aligned}$$

We also write

$$\begin{aligned}\int \sin^3 x \cos^4 x\, dx &= \int (1-\cos^2 x)\cos^4 x \sin x\, dx = -\int (1-\cos^2 x)\cos^4 x\, d(\cos x) \\ &= -\frac{1}{5}\cos^5 x + \frac{1}{7}\cos^7 x + C\,.\end{aligned}$$

- **The case when m and n are both even**

First we talk about how to integrate $\cos^n x$. We have shown the recurrence relation (8.2.3) in previous section, and there are other ways of finding the integral of $\cos^n x$ without using integration by parts. The case when $n = 2\ell+1$ can be dealt with the previous case, so we focus on the case $n = 2\ell$. Make use of the half angle formula

$$\cos^2 x = \frac{1+\cos(2x)}{2}\,,$$

we can write

$$\int \cos^{2\ell} x\, dx = \int \Big(\frac{1+\cos(2x)}{2}\Big)^\ell dx = \sum_{i=0}^{\ell} \frac{C^\ell_i}{2^\ell} \int \cos^i(2x)\, dx \overset{(u=2x)}{=} \sum_{i=0}^{\ell} \frac{C^\ell_i}{2^{\ell+1}} \int \cos^i u\, du$$

which is a linear combination of integrals of the form $\displaystyle\int \cos^i u\, du$, while the power i is at most half of n. Keeping on applying the half angle formula for even powers of cosine, eventually integral $\displaystyle\int \cos^i u\, du$ will be reduced to sum of integrals of cosine with odd powers (which can be evaluated by the previous case).

Example 8.18. Find the indefinite integral $\int \cos^6 x\, dx$.

By the half angle formula,

$$\begin{aligned}\int \cos^6 x\, dx &= \int \Big(\frac{1+\cos(2x)}{2}\Big)^3 dx = \frac{1}{8}\int \big[1+3\cos(2x)+3\cos^2(2x)+\cos^3(2x)\big]\, dx\\ &= \frac{1}{8}\int \Big[1+3\cos(2x)+\frac{3}{2}\big(1+\cos(4x)\big)+\big(1-\sin^2(2x)\big)\cos(2x)\Big]\, dx\\ &= \frac{1}{8}\int \Big(\frac{5}{2}+4\cos(2x)+\frac{3}{2}\cos(4x)\Big)\, dx - \frac{1}{16}\int \sin^2(2x)\, d\big(\sin(2x)\big)\\ &= \frac{1}{8}\Big[\frac{5x}{2}+2\sin(2x)+\frac{3}{8}\sin(4x)\Big] - \frac{1}{48}\sin^3(2x) + C\,.\end{aligned}$$

Now suppose that $m = 2k$ and $n = 2\ell$. Make use of the half angle formulas

$$\sin^2 x = \frac{1-\cos(2x)}{2} \qquad \text{and} \qquad \cos^2 x = \frac{1+\cos(2x)}{2}$$

to write

$$\int \sin^{2k} x \cos^{2\ell} x\, dx = \frac{1}{2^{k+\ell}} \int \big(1-\cos(2x)\big)^k \big(1+\cos(2x)\big)^\ell\, dx\,.$$

Expanding parenthesis, the integral above becomes the linear combination of integrals of the form $\int \cos^i(2x)\, dx$.

Example 8.19. Find the indefinite integral $\int \sin^2 x \cos^4 x\, dx$.

By the half angle formula,

$$\begin{aligned}\int \sin^2 x \cos^4 x\, dx &= \int \frac{1-\cos(2x)}{2}\Big(\frac{1+\cos(2x)}{2}\Big)^2 dx\\ &= \frac{1}{8}\int \big[1-\cos(2x)\big]\big[1+2\cos(2x)+\cos^2(2x)\big]\, dx\\ &= \frac{1}{8}\int \big[1+\cos(2x)-\cos^2(2x)-\cos^3(2x)\big]\, dx\\ &= \frac{1}{8}\int \Big(\frac{1-\cos(4x)}{2}+\sin^2(2x)\cos(2x)\Big]\, dx\\ &= \frac{1}{8}\Big[\frac{x}{2}-\frac{\sin(4x)}{8}\Big] + \frac{1}{48}\sin^3(2x) + C\,.\end{aligned}$$

8.3.2 The integral of $\sec^m x \tan^n x$

Rule of thumb: make use of $\frac{d}{dx}\tan x = \sec^2 x$ and $\frac{d}{dx}\sec x = \sec x \tan x$.

- **The case when m is even**

Suppose that $m = 0$ and $n \geqslant 2$. Then we obtain the recurrence relation

$$\begin{aligned}\int \tan^n x\,dx &= \int \tan^{n-2} x \tan^2 x\,dx = \int \tan^{n-2}(\sec^2 x - 1)\,dx \\ &= \int \tan^{n-2}\ d(\tan x) - \int \tan^{n-2} x\,dx = \frac{1}{n-1}\tan^{n-1} x - \int \tan^{n-2} x\,dx\,.\end{aligned}$$

Suppose that $m = 2k$ is even and positive. Using the substitution $u = \tan x$, we have

$$\int \sec^{2k} x \tan^n x\,dx = \int \sec^{2(k-1)} x \tan^n x \sec^2 x\,dx = \int (1 + \tan^2 x)^{k-1} \tan^n x\, d(\tan x)$$

which can be obtained by integrating polynomials.

- **The case when n is odd**

Suppose that $n = 2\ell + 1$ is odd and $m \geqslant 1$. Then

$$\int \sec^m x \tan^{2\ell+1} x\,dx = \int \sec^{m-1} x \tan^{2\ell} \sec x \tan x\,dx = \int \sec^{m-1} x(\sec^2 x - 1)^\ell\, d(\sec x)$$

which can be obtained by integrating polynomials.

- **The case when m is odd and n is even**

Suppose that $m = 2k + 1$ and $n = 2\ell$. Then

$$\int \sec^{2k+1} x \tan^{2\ell} x\,dx = \int \sec^{2k+1} x(\sec^2 x - 1)^\ell\,dx\,;$$

thus it suffices to know how to compute $\displaystyle\int \sec^m x\,dx$.

Using integration by parts,

$$\begin{aligned}\int \sec^m x\,dx = \int \sec^{m-2} x\, d(\tan x) &= \tan x \sec^{m-2} x - \int \tan x\, d(\sec^{m-2} x) \\ &= \tan x \sec^{m-2} x - (m-2)\int \tan^2 x \sec^{m-2} x\,dx \\ &= \tan x \sec^{m-2} x - (m-2)\int (\sec^2 x - 1)\sec^{m-2} x\,dx\end{aligned}$$

thus rearranging terms we obtain the recurrence relation

$$\int \sec^m x\,dx = \frac{m-2}{m-1}\tan x \sec^{m-2} x + \frac{m-2}{m-1}\int \sec^{m-2} x\,dx\,.$$

Example 8.20. Find the indefinite integral $\displaystyle\int \sec^4(3x)\tan^3(3x)\,dx$.

By the discussion above,

$$\begin{aligned}\int \sec^4(3x)\tan^3(3x)\,dx &= \frac{1}{3}\int \sec^2(3x)\tan^3(3x)d(\tan(3x))\\ &= \frac{1}{3}\int \big[\tan^2(3x)+1\big]\tan^3(3x)d(\tan(3x))\\ &= \frac{1}{3}\Big[\frac{1}{6}\tan^6(3x)+\frac{1}{4}\tan^4(3x)\Big]+C\,.\end{aligned}$$

Example 8.21. Find the indefinite integral $\displaystyle\int \sqrt{a^2+x^2}\,dx$.

By the substitution of variable $x = a\tan\theta$ (so that $dx = a\sec^2\theta d\theta$), we find that

$$\begin{aligned}\int \sqrt{a^2+x^2}\,dx = \int a^2\sec^3\theta\,d\theta &= a^2\Big(\frac{1}{2}\tan\theta\sec\theta+\frac{1}{2}\int\sec\theta\,d\theta\Big)\\ &= \frac{a^2}{2}\big(\tan\theta\sec\theta+\ln|\sec\theta+\tan\theta|\big)+C\\ &= \frac{a^2}{2}\Big(\frac{x}{a}\cdot\frac{\sqrt{a^2+x^2}}{a}+\ln\Big|\frac{x+\sqrt{a^2+x^2}}{a}\Big|\Big)+C\\ &= \frac{x\sqrt{a^2+x^2}}{2}+\frac{a^2}{2}\ln\big(x+\sqrt{a^2+x^2}\big)+C\,. \qquad (8.3.1)\end{aligned}$$

8.3.3 Other techniques of integration involving trigonometric functions

- Integration by substitution (for integrand with special structures):

Example 8.22. Find the indefinite integral $\displaystyle\int \frac{\cos^3 x}{\sqrt{\sin x}}\,dx$.

Let $u = \sin x$. Then $du = \cos x\,dx$; thus

$$\begin{aligned}\int \frac{\cos^3 x}{\sqrt{\sin x}}\,dx = \int \frac{(1-u^2)}{\sqrt{u}}\,du &= \int \big(u^{-\frac{1}{2}}-u^{\frac{3}{2}}\big)\,du\\ = \frac{1}{1-\frac{1}{2}}u^{\frac{1}{2}}-\frac{1}{1+\frac{3}{2}}u^{\frac{5}{2}}+C &= 2\sqrt{\sin x}-\frac{5}{2}\sin^{\frac{5}{2}}x+C\,.\end{aligned}$$

Example 8.23. Find the indefinite integral $\displaystyle\int \frac{\sec x}{\tan^2 x}\,dx$.

Rewrite the integrand into a fraction of sine and cosine, we find that

$$\int \frac{\sec x}{\tan^2 x}\,dx = \int \frac{\cos x}{\sin^2 x}\,dx = \int \frac{1}{\sin^2 x}\,d(\sin x) = -\sin^{-1}x+C = -\csc x+C\,.$$

Example 8.24. Find the indefinite integral $\displaystyle\int \frac{\tan^3 x}{\sqrt{\sec x}}\,dx$.

Let $u = \sec x$. Then $du = \sec x \tan x\,dx$; thus

$$\begin{aligned}\int \frac{\tan^3 x}{\sqrt{\sec x}}\,dx &= \int \frac{(\sec^2 x - 1)\sec x \tan x}{\sec^{\frac{3}{2}} x}\,dx = \int \frac{u^2-1}{u^{\frac{3}{2}}}\,du = \int \big(u^{\frac{1}{2}} - u^{-\frac{3}{2}}\big)\,du \\ &= \frac{2}{3}u^{\frac{3}{2}} + 2u^{-\frac{1}{2}} + C = \frac{2}{3}\sec^{\frac{3}{2}} x + 2\cos^{\frac{1}{2}} x + C\,.\end{aligned}$$

• When the angular variable are different, making use of the sum and difference formula:

Example 8.25. Find the indefinite integral $\displaystyle\int \sin^3(5x)\cos(4x)\,dx$.

Using the sum and difference formula

$$\sin\theta\cos\phi = \frac{1}{2}\big[\sin(\theta+\phi) + \sin(\theta-\phi)\big]\,, \quad \sin\theta\sin\phi = \frac{1}{2}\big[\cos(\theta-\phi) - \sin(\theta+\phi)\big]\,,$$

we find that

$$\begin{aligned}\int \sin^3(5x)\cos(4x)\,dx &= \frac{1}{2}\int \sin^2(5x)\big[\sin(9x) + \sin x\big]\,dx \\ &= \frac{1}{4}\int \sin(5x)\big[\cos(4x) - \cos(14x) + \cos(4x) - \cos(6x)\big]\,dx \\ &= \frac{1}{8}\int \big[2\sin(9x) + 2\sin x - \sin(19x) + \sin(9x) - \sin(11x) + \sin x\big]\,dx \\ &= \frac{1}{8}\Big[-\frac{1}{3}\cos(9x) - 3\cos x + \frac{1}{19}\cos(19x) + \frac{1}{11}\cos(11x)\Big] + C\,.\end{aligned}$$

8.4 Partial Fractions - 部份分式

In this section, we are concerned with the integrals $\displaystyle\int \frac{N(x)}{D(x)}\,dx$, where N, D are polynomial functions.

Write $N(x) = D(x)Q(x) + R(x)$, where Q, R are polynomials such that the degree of R is less than the degree of D (such an R is called a remainder). Then $\dfrac{N(x)}{D(x)} = R(x) + \dfrac{R(x)}{D(x)}$. Since it is easy to find the indefinite integral of R, it suffices to consider the case when the degree of the numerator is less than the degree of the denominator.

W.L.O.G., we assume that N and D no common factor, $\deg(N) < \deg(D)$, and the leading coefficient of D is 1. Since D is a polynomial with real coefficients,

$$D(x) = \Big(\prod_{j=1}^{m}(x+q_j)^{r_j}\Big)\Big(\prod_{j=1}^{n}(x^2+b_jx+c_j)^{d_j}\Big)\,,$$

where $r_j, d_j \in \mathbb{N}$, $q_j \neq q_k$ for all $j \neq k$, $b_j \neq b_k$ or $c_j \neq c_k$ for all $j \neq k$, and $b_j^2 - 4c_j < 0$ for all $1 \leqslant j \leqslant m$. In other words, $-q_j$ are zeros of D with multiplicity r_j, and $\dfrac{-b_j \pm i\sqrt{4c_j - b_j^2}}{2}$ are zeros of D with multiplicity d_j, here $i = \sqrt{-1}$. Therefore,

$$\frac{N(x)}{D(x)} = \sum_{j=1}^{m} \Big[\sum_{\ell=1}^{r_j} \frac{A_{j\ell}}{(x+q_j)^\ell} \Big] + \sum_{j=1}^{n} \Big[\sum_{\ell=1}^{r_j} \frac{B_{j\ell}x + C_{j\ell}}{(x^2 + b_j x + c_j)^\ell} \Big] \tag{8.4.1}$$

for some constants $A_{j\ell}$, $B_{j\ell}$ and $C_{j\ell}$. Note that there are $\sum\limits_{j=1}^{m} r_j + 2 \sum\limits_{j=1}^{n} d_j \equiv \deg(D)$ constants to be determined, and this can be done by the comparison of coefficients after the reduction of common denominator.

Remark 8.26. In this remark we "show" that a rational function N/D with $\deg(N) < \deg(D)$ can always be written as the sum of partial fractions (8.4.1). Suppose that α is a zero of D with multiplicity k so that $D(x) = (x-\alpha)^k f(x)$, where $f(x)$ is a polynomial and $f(\alpha) \neq 0$. Since

$$\frac{N(x)}{D(x)} - \frac{N(\alpha)}{(x-\alpha)^k f(\alpha)} = \frac{N(x)f(\alpha) - f(x)N(\alpha)}{(x-\alpha)^k f(x) f(\alpha)} = \frac{g(x)}{(x-\alpha)^k f(x)}\,,$$

where $g(x) = N(x) - f(x)\dfrac{N(\alpha)}{f(\alpha)}$. Since g vanishes at $x = \alpha$, $g(x) = (x-\alpha)^m h(x)$ for some polynomial h satisfying $h(\alpha) \neq 0$ (and we remark that here m is not necessarily less than k). Therefore, with β denoting the constant $\dfrac{N(\alpha)}{f(\alpha)}$, we obtain that

$$\frac{N(x)}{D(x)} - \frac{\beta}{(x-\alpha)^k} = \frac{(x-\alpha)^m h(x)}{(x-\alpha)^k f(x)} = \frac{h_1(x)}{(x-\alpha)^{k_1} f(x)}\,,$$

where $k_1 \geqslant 0$ and $h_1(\alpha) \neq 0$ if $k_1 > 0$. We note that f and h_1 are both polynomials satisfying $\deg h_1 < k_1 + \deg(f)$ and $f(\alpha) \neq 0$. Applying the process continuously, we obtain that

$$\frac{N(x)}{D(x)} = \sum_{i=1}^{k} \frac{C_k}{(x-\alpha)^k} + \frac{N_1(x)}{D_1(x)}$$

for some polynomials N_1, $D_1 (= f)$ with $\deg(N_1) < \deg(D_1) = \deg(D) - k$ and some sequence of constants $C_1, C_2, \cdots, C_k$, where $D_1(\alpha) \neq 0$. This explains the presence of the first sum on the right-hand side of (8.4.1) (and also shows how to find the coefficient A_{jr_j} in the highest order term $\dfrac{1}{(x+q_j)^{r_j}}$ for each j).

Example 8.27. Write $\dfrac{5x^2+20x+6}{x^3+2x^2+x}$ in the form of (8.4.1).

Note that $x^3+2x^2+x = x(x^2+2x+1) = x(x+1)^2$; thus to write the rational function above in the form of (8.4.1), we must have

$$\frac{5x^2+20x+6}{x^3+2x^2+x} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{(x+1)^2}$$

for some constant A, B, C.

Multiplying both sides of the equality above by $x(x+1)^2$, we find that

$$5x^2+20x+6 = A(x+1)^2 + Bx(x+1) + Cx = (A+B)x^2 + (2A+B+C)x + A\,;$$

thus A, B, C satisfy

$$\begin{aligned} A+B &= 5 \\ 2A+B+C &= 20 \\ A &= 6\,. \end{aligned}$$

Therefore, $A=6$, $B=-1$ and $C=9$; thus

$$\frac{5x^2+20x+6}{x^3+2x^2+x} = \frac{6}{x} - \frac{1}{x+1} + \frac{9}{(x+1)^2}\,.$$

Example 8.28. Write $\dfrac{1}{x^4+1}$ in the form of (8.4.1).

Note that $x^4+1 = (x^2+\sqrt{2}x+1)(x^2-\sqrt{2}x+1)$, so

$$\frac{1}{x^4+1} = \frac{Ax+B}{x^2+\sqrt{2}x+1} + \frac{Cx+D}{x^2-\sqrt{2}x+1}\,.$$

Multiplying both sides of the equality above by x^4+1, we have

$$\begin{aligned} 1 &= (Ax+B)(x^2-\sqrt{2}x+1) + (Cx+D)(x^2+\sqrt{2}x+1) \\ &= (A+C)x^3 + (-\sqrt{2}A+B+\sqrt{2}C+D)x^2 + (A-\sqrt{2}B+C+\sqrt{2}D)x + (B+D)\,; \end{aligned}$$

thus comparing the coefficients, we find that A, B, C, D satisfy

$$\begin{aligned} A+C &= 0 \\ -\sqrt{2}A+B+\sqrt{2}C+D &= 0 \\ A-\sqrt{2}B+C+\sqrt{2}D &= 0 \\ B+D &= 1\,. \end{aligned}$$

Therefore, the first and the third equations imply that $A = -C$ and $B = D$; thus the second and the fourth equation shows that $A = -C = \dfrac{1}{2\sqrt{2}}$ and $B = D = \dfrac{1}{2}$. As a consequence,

$$\frac{1}{x^4+1} = \frac{1}{2\sqrt{2}}\Big[\frac{x+\sqrt{2}}{x^2+\sqrt{2}x+1} + \frac{-x+\sqrt{2}}{x^2-\sqrt{2}x+1}\Big].$$

In order to find the integral of $\dfrac{N(x)}{D(x)}$, by writing $\dfrac{N(x)}{D(x)}$ in the form of (8.4.1), it suffices to find the integral of $\dfrac{B_{j\ell}x + C_{j\ell}}{(x^2+b_jx+c_j)^\ell}$ for

$$\int \frac{A_{j\ell}}{(x+q_j)^\ell}\,dx = \begin{cases} \dfrac{A_{j\ell}}{1-\ell}(x+q_j)^{1-\ell} + C & \text{if } \ell \neq 1\,, \\ A_{j\ell}\ln|x+q_j| + C & \text{if } \ell = 1\,. \end{cases}$$

Note that

$$\frac{B_{j\ell}x + C_{j\ell}}{(x^2+b_jx+c)^\ell} = \frac{B_{j\ell}}{2}\frac{2x+b_j}{(x^2+b_jx+c_j)^\ell} + \Big(C_{j\ell} - \frac{b_jB_{j\ell}}{2}\Big)\frac{1}{(x^2+b_jx+c_j)^\ell}$$

and

$$\int \frac{2x+b_j}{(x^2+b_jx+c_j)^\ell}\,dx = \begin{cases} \dfrac{1}{1-\ell}(x^2+b_jx+c_j)^{1-\ell} + C & \text{if } \ell \neq 1\,, \\ \ln(x^2+b_jx+c_j) + C & \text{if } \ell = 1\,; \end{cases}$$

thus to find the integral of $\dfrac{B_{j\ell}x + C_{j\ell}}{(x^2+b_jx+c_j)^\ell}$, it suffices to compute $\displaystyle\int \frac{1}{(x^2+b_jx+c_j)^\ell}\,dx$. Nevertheless, with a denoting the number $\dfrac{4c_j - b_j^2}{4}$,

$$\int \frac{1}{(x^2+b_jx+c_j)^\ell}\,dx = \int \frac{1}{\big[\big(x-\frac{b_j}{2}\big)^2 + \frac{4c_j-b_j^2}{4}\big]^\ell}\,dx = \int \frac{1}{\big[\big(x-\frac{b_j}{2}\big)^2 + a^2\big]^\ell}\,d\big(x-\tfrac{b_j}{2}\big)$$

which can be computed through the substitution $x - \dfrac{b_j}{2} = a\tan u$:

$$\int \frac{1}{\big[\big(x-\frac{b_j}{2}\big)^2 + a^2\big]^\ell}\,d\big(x-\tfrac{b_j}{2}\big) = a^{1-2\ell}\int \cos^{2\ell-2} u\,du\,.$$

Example 8.29. Find the indefinite integral $\displaystyle\int \frac{dx}{x^4+1}$.

Using the conclusion from Example 8.28, we find that

$$\begin{aligned}
\int \frac{dx}{x^4+1} &= \frac{1}{2\sqrt{2}} \int \Big[\frac{x+\sqrt{2}}{x^2+\sqrt{2}x+1} + \frac{-x+\sqrt{2}}{x^2-\sqrt{2}x+1}\Big]\, dx \\
&= \frac{1}{2\sqrt{2}} \int \Big[\frac{1}{2}\cdot\frac{2x+\sqrt{2}}{x^2+\sqrt{2}x+1} - \frac{1}{2}\cdot\frac{2x-\sqrt{2}}{x^2-\sqrt{2}x+1}\Big]\, dx \\
&\quad + \frac{1}{2\sqrt{2}} \int \Big[\frac{1}{2}\cdot\frac{\sqrt{2}}{x^2+\sqrt{2}x+1} + \frac{1}{2}\cdot\frac{\sqrt{2}}{x^2-\sqrt{2}x+1}\Big]\, dx \\
&= \frac{1}{4\sqrt{2}} \int \Big[\frac{2x+\sqrt{2}}{x^2+\sqrt{2}x+1} + \frac{\sqrt{2}}{\big(x+\frac{1}{\sqrt{2}}\big)^2+(\frac{1}{\sqrt{2}})^2} - \frac{2x-\sqrt{2}}{x^2-\sqrt{2}x+1} + \frac{\sqrt{2}}{\big(x-\frac{1}{\sqrt{2}}\big)^2+(\frac{1}{\sqrt{2}})^2}\Big]\, dx \\
&= \frac{1}{4\sqrt{2}} \Big[\ln\frac{x^2+\sqrt{2}x+1}{x^2-\sqrt{2}x+1} + 2\arctan(\sqrt{2}x+1) + 2\arctan(\sqrt{2}x-1)\Big] + C\,.
\end{aligned}$$

Example 8.30. Find the indefinite integral $\displaystyle\int \frac{\sec x}{\tan^3 x}\, dx$.

Let $u = \sec x$. Then $du = \sec x \tan x$; thus

$$\int \frac{\sec x}{\tan^3 x}\, dx = \int \frac{\sec x \tan x}{\tan^4 x}\, dx = \int \frac{du}{(u^2-1)^2} = \int \frac{du}{(u+1)^2(u-1)^2}\,.$$

Write $\displaystyle\frac{1}{(u+1)^2(u-1)^2}$ is the form of (8.4.1):

$$\frac{1}{(u+1)^2(u-1)^2} = \frac{A}{u+1} + \frac{B}{(u+1)^2} + \frac{C}{u-1} + \frac{D}{(u-1)^2}\,,$$

where A, B, C, D satisfy

$$A(u+1)(u-1)^2 + B(u-1)^2 + C(u-1)(u+1)^2 + D(u+1)^2 = 1\,.$$

Therefore, A, B, C, D satisfy

$$\begin{aligned}
A + C &= 0 \\
-A + B + C + D &= 0 \\
-A - 2B - C + 2D &= 0 \\
A + B - C + D &= 1
\end{aligned}$$

which implies that $A = B = -C = D = \dfrac{1}{4}$. As a consequence,

$$\begin{aligned}\int \frac{du}{(u+1)^2(u-1)^2} &= \frac{1}{4}\int \Big[\frac{1}{u+1} + \frac{1}{(u+1)^2} - \frac{1}{u-1} + \frac{1}{u-1)^2}\Big]\,du \\ &= \frac{1}{4}\Big[\ln|u+1| - \frac{1}{u+1} - \ln|u-1| - \frac{1}{u-1}\Big] + C \\ &= \frac{1}{4}\Big[\ln\Big|\frac{u+1}{u-1}\Big| - \frac{2u}{u^2-1}\Big] + C\,;\end{aligned}$$

thus

$$\int \frac{\sec x}{\tan^3 x}\,dx = \frac{1}{4}\Big[\ln\Big|\frac{\sec x+1}{\sec x-1}\Big| - \frac{2\sec x}{\tan^2 x}\Big] + C\,.$$

Example 8.31. Find the indefinite integral $\displaystyle\int \sqrt{\tan x}\,dx$.

Let $u = \sqrt{\tan x}$. Then $u^2 = \tan x$ which implies that $2udu = \sec^2 x\,dx$ or $\dfrac{2udu}{1+u^4} = dx$. Therefore,

$$\begin{aligned}\int \sqrt{\tan x}\,dx = \int \frac{2u^2}{1+u^4}\,du &= \frac{1}{\sqrt{2}}\int \Big[\frac{u}{u^2-\sqrt{2}u+1} - \frac{u}{u^2+\sqrt{2}u+1}\Big]\,du \\ &= \frac{1}{2\sqrt{2}}\ln\Big|\frac{u^2-\sqrt{2}u+1}{u^2+\sqrt{2}u+1}\Big| + \frac{1}{2}\int \Big[\frac{1}{u^2-\sqrt{2}u+1} + \frac{1}{u^2+\sqrt{2}u+1}\Big]\,du \\ &= \frac{1}{2\sqrt{2}}\ln\Big|\frac{u^2-\sqrt{2}u+1}{u^2+\sqrt{2}u+1}\Big| + \frac{\sqrt{2}}{2}\arctan(\sqrt{2}u-1) + \arctan(\sqrt{2}u+1) + C \\ &= \frac{1}{2\sqrt{2}}\ln\Big|\frac{\tan x-\sqrt{2\tan x}+1}{\tan x+\sqrt{2\tan x}+1}\Big| + \frac{\sqrt{2}}{2}\arctan\frac{\sqrt{2\tan x}}{1-\tan x} + C\,,\end{aligned}$$

where we have use the fact that

$$\arctan x + \arctan y = \arctan\frac{x+y}{1-xy} + C$$

to conclude the last equality.

Example 8.32. Find the indefinite integral $\displaystyle\int \frac{dx}{(1+x^n)^{\frac{1}{n}}}$, where n is a positive integer.

Let $1 + x^{-n} = u^n$. Then $x^n = \dfrac{1}{u^n-1}$ and $-x^{-n-1}\,dx = u^{n-1}\,du$; thus

$$\int \frac{dx}{(1+x^n)^{\frac{1}{n}}} = \int \frac{dx}{x(1+x^{-n})^{\frac{1}{n}}} = \int \frac{-x^n}{(1+x^{-n})^{\frac{1}{n}}}(-x^{-n-1})\,dx = -\int \frac{u^{n-2}}{u^n-1}\,du$$

which is the indefinite integral of a rational function of u and we know how to compute it. In particular, when $n = 4$,

$$\frac{u^2}{u^4-1} = \frac{u^2}{(u-1)(u+1)(u^2+1)} = \frac{1}{4}\cdot\frac{1}{u-1} - \frac{1}{4}\cdot\frac{1}{u+1} + \frac{1}{2}\cdot\frac{1}{u^2+1}\,;$$

thus

$$\int \frac{u^2}{u^4-1}\,du = \frac{1}{4}\ln|u-1| - \frac{1}{4}\ln|u+1| + \frac{1}{2}\arctan u + C$$

which further implies that

$$\int \frac{dx}{(1+x^4)^{\frac{1}{4}}} = \frac{1}{4}\ln\left|\frac{(1+x^{-4})^{\frac{1}{4}}-1}{(1+x^{-4})^{\frac{1}{4}}+1}\right| + \frac{1}{2}\arctan\left[(1+x^{-4})^{\frac{1}{4}}\right] + C\,.$$

- **The substitution of** $t = \tan\dfrac{x}{2}$

In Section 5.3 we have introduced the substitution $t = \tan\frac{x}{2}$ to find the anti-derivative of trigonometric functions. We recall that if $t = \tan\frac{x}{2}$, then

$$\sin x = \frac{2t}{1+t^2}\,, \qquad \cos x = \frac{1-t^2}{1+t^2} \qquad \text{and} \qquad dx = \frac{2dt}{1+t^2}\,.$$

Using this substitution, the anti-derivative of rational functions of sine and cosine can be computed via the integration of rational functions.

Example 8.33. Find the indefinite integral $\displaystyle\int \frac{\sec x}{\tan^3 x}\,dx$.

Rewriting the integrand, we have

$$\int \frac{\sec x}{\tan^3 x}\,dx = \int \frac{\cos^2 x}{\sin^3 x}\,dx\,.$$

Let $t = \tan\frac{x}{2}$. Then $\sin x = \dfrac{2t}{1+t^2}$, $\cos x = \dfrac{1-t^2}{1+t^2}$ and $dx = \dfrac{2dt}{1+t^2}$; thus

$$\begin{aligned}\int \frac{\sec x}{\tan^3 x}\,dx &= \int \frac{\frac{(1-t^2)^2}{(1+t^2)^2}}{\frac{(2t)^3}{(1+t^2)^3}}\frac{2dt}{1+t^2} = \frac{1}{4}\int \frac{(1-t^2)^2}{t^3}\,dt = \frac{1}{4}\int \left(t^{-3} - 2t^{-1} + t\right)dt \\ &= \frac{1}{4}\left[-\frac{1}{2}t^{-2} - 2\ln|t| + \frac{1}{2}t^2\right] + C \\ &= \frac{1}{8}\left[\tan^2\frac{x}{2} - \cot^2\frac{x}{2}\right] - \frac{1}{2}\ln\left|\tan\frac{x}{2}\right| + C\,.\end{aligned}$$

Example 8.34. Find the indefinite integral $\displaystyle\int \frac{1}{2+\sin x}\,dx$.

Let $t = \tan\dfrac{x}{2}$. Then $\sin x = \dfrac{2t}{1+t^2}$, $\cos x = \dfrac{1-t^2}{1+t^2}$ and $dx = \dfrac{2dt}{1+t^2}$; thus

$$\begin{aligned}
\int \frac{1}{2+\sin x}\,dx &= \int \frac{1}{2+\frac{2t}{1+t^2}}\,\frac{2dt}{1+t^2} = \int \frac{dt}{t^2+t+1} = \int \frac{dt}{\left(t+\frac{1}{2}\right)^2+\left(\frac{\sqrt{3}}{2}\right)^2} \\
&= \frac{2}{\sqrt{3}}\arctan\frac{t+\frac{1}{2}}{\frac{\sqrt{3}}{2}} + C = \frac{2}{\sqrt{3}}\arctan\frac{2t+1}{\sqrt{3}} + C \\
&= \frac{2}{\sqrt{3}}\arctan\Big(\frac{2}{\sqrt{3}}\tan\frac{x}{2} + \frac{1}{\sqrt{3}}\Big) + C\,.
\end{aligned}$$

8.5 Improper Integrals - 瑕積分

Recall that given a non-negative continuous function $f : [a, b] \to \mathbb{R}$, the area of the region enclosed by the graph of f, the x-axis and lines $x = a$, $x = b$ is given by $\displaystyle\int_a^b f(x)\,dx$.What happened when

1. the function under consideration is non-negative and continuous on the whole real line and we would like to know, for example, the area of the region enclosed by the graph of f and the x-axis and is on the right-hand (or left-hand) side of the line $x = c$?

2. the function under consideration blows up at a point $c \in [a, b]$; that is, $\displaystyle\lim_{x\to c^\pm} f(x)$ diverges to ∞ or $-\infty$ (so that f is not continuous at c but everywhere else) and we would like to know the area of the region enclosed by the graph of f, the x-axis and lines $x = a$ and $x = b$?

Note that the definition of a definite integral $\displaystyle\int_a^b f(x)\,dx$ requires that the interval $[a, b]$ be finite and f be bounded. Therefore, $\displaystyle\int_a^\infty f(x)\,dx$, $\displaystyle\int_{-\infty}^b f(x)\,dx$ and $\displaystyle\int_a^b f(x)\,dx$ when f is unbounded are meaningless in the sense of Riemann integrals. How do we compute the area of those unbounded regions?

Definition 8.34: Improper Integrals with Infinite Integration Limits

1. If f is Riemann integrable on the interval $[a,b]$ for all $a<b$, then
$$\int_a^{\infty} f(x)\,dx \equiv \lim_{b\to\infty}\int_a^b f(x)\,dx\,.$$
2. If f is Riemann integrable on the interval $[a,b]$ for all $a<b$, then
$$\int_{-\infty}^{b} f(x)\,dx \equiv \lim_{a\to-\infty}\int_a^b f(x)\,dx\,.$$
3. If f is Riemann integrable on the interval $[a,b]$ for all $a<b$, then
$$\int_{-\infty}^{\infty} f(x)\,dx \equiv \int_{-\infty}^{c} f(x)\,dx + \int_c^{\infty} f(x)\,dx\,,$$
where c is any real number.

In the first two cases, the improper integral converges when the limit exists. Otherwise, the improper integral diverges. If the limits, as b approaches ∞ (or a approaches $-\infty$), approaches ∞ or $-\infty$, then the improper integral diverges to ∞ or $-\infty$. In the third case, the improper integral on the left converges when both of the improper integrals on the right converges, and diverges when either of the improper integrals on the right diverges. The improper integral on the left diverges to ∞ (or $-\infty$) if it diverges and the improper integrals on the right is $\infty+\infty$, $\infty+C$ or $C+\infty$ (or $(-\infty)+(-\infty)$, $(-\infty)+C$ or $C+(-\infty)$).

Example 8.35. Evaluate $\displaystyle\int_0^{\infty} e^{-x}\,dx$ and $\displaystyle\int_0^{\infty}\frac{1}{x^2+1}\,dx$.

Since an anti-derivative of the function $y=e^{-x}$ and $y=\dfrac{1}{x^2+1}$ is $y=-e^{-x}$ and $y=\arctan x$, the Fundamental Theorem of Calculus implies that

$$\int_0^{\infty} e^{-x}\,dx = \lim_{b\to\infty}\int_0^b e^{-x}\,dx = \lim_{b\to\infty}(-e^{-x})\Big|_{x=0}^{x=b} = \lim_{b\to\infty}(1-e^{-b}) = 1-\lim_{b\to\infty}e^{-b} = 1$$

and
$$\int_0^{\infty}\frac{1}{x^2+1}\,dx = \lim_{b\to\infty}\int_0^b\frac{1}{x^2+1}\,dx = \lim_{b\to\infty}\arctan x\Big|_{x=0}^{x=b} = \lim_{b\to\infty}\arctan b = \frac{\pi}{2}\,.$$

Example 8.36. Evaluate $\displaystyle\int_1^{\infty}(1-x)e^{-x}\,dx$.

Let $u=1-x$ and $v=-e^{-x}$ (so that $dv=e^{-x}\,dx$). For any real number b, integration by parts implies that

$$\int_1^b (1-x)e^{-x}\,dx = \big[(1-x)(-e^{-x})\big]\Big|_{x=1}^{x=b} - \int_1^b (-e^{-x})(-dx) = -(1-b)e^{-b} - \int_1^b e^{-x}\,dx$$
$$= -(1-b)e^{-b} + e^{-x}\Big|_{x=1}^{x=b} = -(1-b)e^{-b} + e^{-b} - e^{-1} = be^{-b} - e^{-1}\,.$$

Therefore,

$$\int_1^\infty (1-x)e^{-x}\,dx = \lim_{b\to\infty}\int_1^b (1-x)e^{-x}\,dx = \lim_{b\to\infty}(be^{-b} - e^{-1}) = -e^{-1}\,.$$

Example 8.37. Evaluate $\displaystyle\int_{-\infty}^\infty \frac{e^x}{1+e^{2x}}\,dx$.

To evaluate the integral above, we evaluate the two integrals

$$\int_0^\infty \frac{e^x}{1+e^{2x}}\,dx \quad\text{and}\quad \int_{-\infty}^0 \frac{e^x}{1+e^{2x}}\,dx\,.$$

By the substitution of variable $u = e^x$, we find that $du = e^x\,dx$; thus

$$\int \frac{e^x}{1+e^{2x}}\,dx = \int \frac{1}{1+u^2}\,du = \arctan u + C = \arctan(e^x) + C\,.$$

Therefore,

$$\int_0^\infty \frac{e^x}{1+e^{2x}}\,dx = \lim_{b\to\infty}\int_0^b \frac{e^x}{1+e^{2x}}\,dx = \lim_{b\to\infty}\arctan(e^x)\Big|_{x=0}^{x=b}$$
$$= \lim_{b\to\infty}\big[\arctan(e^b) - \arctan 1\big] = \frac{\pi}{4}$$

and similarly,

$$\int_{-\infty}^0 \frac{e^x}{1+e^{2x}}\,dx = \lim_{a\to-\infty}\int_a^0 \frac{e^x}{1+e^{2x}}\,dx = \lim_{a\to-\infty}\arctan(e^x)\Big|_{x=a}^{x=0}$$
$$= \lim_{a\to-\infty}\big[\arctan 1 - \arctan(e^a)\big] = \frac{\pi}{4}\,.$$

The two integrals above implies that $\displaystyle\int_{-\infty}^\infty \frac{e^x}{1+e^{2x}}\,dx = \frac{\pi}{4} + \frac{\pi}{4} = \frac{\pi}{2}$.

Example 8.38. The improper integral $\displaystyle\int_0^\infty x\,dx$ diverges to ∞, and the improper integral $\displaystyle\int_{-\infty}^\infty (\sin x - 1)\,dx$ diverges to $-\infty$. The improper integral $\displaystyle\int_0^\infty \sin x\,dx$ diverges, but not diverges to ∞ or $-\infty$, and the improper integrals $\displaystyle\int_{-\infty}^\infty x\,dx$ diverges but not diverges to ∞ or $-\infty$.

Example 8.39. The improper integral $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$ converges although it is not obvious what its value is. In fact,

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}\,.$$

Theorem 8.40

1. If f is Riemann integrable on the interval $[a,b]$ for all $a<b$, then
$$\int_a^\infty f(x)\,dx = \int_a^c f(x)\,dx + \int_c^\infty f(x)\,dx \qquad \forall\, a<c\,,$$
provided that the improper integrals on both sides converge or diverge to ∞ (or $-\infty$).

2. If f is Riemann integrable on the interval $[a,b]$ for all $a<b$, then
$$\int_{-\infty}^b f(x)\,dx = \int_{-\infty}^c f(x)\,dx + \int_c^b f(x)\,dx \qquad \forall\, c<b\,,$$
provided that the improper integrals on both sides converge or diverge to ∞ (or $-\infty$).

3. If f is Riemann integrable on the interval $[a,b]$ for all $a<b$ and $\displaystyle\int_{-\infty}^\infty f(x)\,dx$ converges or diverges to ∞ (or $-\infty$), then
$$\int_{-\infty}^a f(x)\,dx + \int_a^\infty f(x)\,dx = \int_{-\infty}^b f(x)\,dx + \int_b^\infty f(x)\,dx \qquad \forall\, a,b\in\mathbb{R}\,.$$

Proof. We only prove 1 and 3, for the proof of 2 is similar to the proof of 1.

1. By the properties of the definite integrals, for $a<c$ we have
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,;$$
thus
$$\begin{aligned}\int_a^\infty f(x)\,dx &= \lim_{b\to\infty}\int_a^b f(x)\,dx = \lim_{b\to\infty}\Big[\int_a^c f(x)\,dx + \int_c^b f(x)\,dx\Big]\\ &= \int_a^c f(x)\,dx + \lim_{b\to\infty}\int_c^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^\infty f(x)\,dx\,.\end{aligned}$$

3. If $\int_{-\infty}^{\infty} f(x)\,dx$ converges or diverges to ∞ (or $-\infty$), then both improper integrals $\int_{c}^{\infty} f(x)\,dx$ and $\int_{-\infty}^{c} f(x)\,dx$ converge or diverge to ∞ (or $-\infty$). Therefore,

$$\begin{aligned}\int_{-\infty}^{b} f(x)\,dx + \int_{b}^{\infty} f(x)\,dx &= \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{b} f(x)\,dx + \int_{b}^{\infty} f(x)\,dx \\ &= \int_{-\infty}^{a} f(x)\,dx + \int_{a}^{\infty} f(x)\,dx\,.\end{aligned}$$

□

Definition 8.41: Improper integrals with Infinite Discontinuities

1. If f is Riemann integrable on $[a, c]$ for all $a < c < b$, and f has an infinite discontinuity at b; that is, $\lim\limits_{x\to b^-} f(x) = \infty$ or $-\infty$, then
$$\int_a^b f(x)\,dx \equiv \lim_{c\to b^-} \int_a^c f(x)\,dx\,.$$
2. If f is Riemann integrable on $[c, b]$ for all $a < c < b$, and f has an infinite discontinuity at a; that is, $\lim\limits_{x\to a^+} f(x) = \infty$ or $-\infty$, then
$$\int_a^b f(x)\,dx \equiv \lim_{c\to a^+} \int_c^b f(x)\,dx\,.$$
3. Suppose that $a < c < b$. If f is Riemann integrable on $[a, c-\epsilon]$ and $[c+\epsilon, b]$ for all $0 < \epsilon \ll 1$, and f has an infinite discontinuity at c; that is $\lim\limits_{x\to c^+} f(x) = \infty$ or $-\infty$ and $\lim\limits_{x\to c^-} f(x) = \infty$ or $-\infty$, then
$$\int_a^b f(x)\,dx \equiv \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,.$$

The convergence and divergence of the improper integrals with infinite discontinuities are similar to the statements in Definition 8.34.

Example 8.42. Evaluate $\int_0^1 x^{-\frac{1}{3}}\,dx$.

We observe that the integrand has an infinite discontinuity at 0. Therefore,

$$\int_0^1 x^{-\frac{1}{3}}\,dx = \lim_{a\to 0^+} \int_a^1 x^{-\frac{1}{3}}\,dx = \lim_{a\to 0^+} \frac{3}{2} x^{\frac{2}{3}}\Big|_{x=a}^{x=1} = \lim_{a\to 0^+} \frac{3}{2}(1 - a^{\frac{2}{3}}) = \frac{3}{2}\,.$$

Example 8.43. Evaluate $\int_0^2 x^{-3}\,dx$.

We observe that the integrand has an infinite discontinuity at 0. Therefore,

$$\int_0^2 x^{-3}\,dx = \lim_{a\to 0^+}\int_a^2 x^{-3}\,dx = \lim_{a\to 0^+}\Big(\frac{-x^{-2}}{2}\Big)\Big|_{x=a}^{x=2} = \lim_{a\to 0^+}\Big(-\frac{1}{8}+\frac{1}{2a^2}\Big) = \infty\,;$$

thus the improper integral $\displaystyle\int_0^2 x^{-3}\,dx$ diverges to ∞.

Example 8.44. Evaluate $\displaystyle\int_{-1}^2 x^{-3}\,dx$.

Since the integrand has an infinite discontinuity at 0,

$$\int_{-1}^2 x^{-3}\,dx = \int_{-1}^0 x^{-3}\,dx + \int_0^2 x^{-3}\,dx\,.$$

We have shown in previous example that the second integral on the right-hand side diverges to ∞. Similarly, the first integral on the right-hand side diverges to $-\infty$ since

$$\int_{-1}^0 x^{-3}\,dx = \lim_{b\to 0^-}\int_{-1}^b x^{-3}\,dx = \lim_{b\to 0^-}\frac{-x^{-2}}{2}\Big|_{x=-1}^{x=b} = \lim_{b\to 0^-}\Big(-\frac{1}{2b^2}+\frac{1}{2}\Big) = -\infty\,;$$

thus the improper integral $\displaystyle\int_{-1}^2 x^{-3}\,dx$ diverges (but not diverges to ∞ or $-\infty$).

Remark 8.45. Even though $y = -\dfrac{x^{-2}}{2}$ is an anti-derivative of the function $y = x^{-3}$, you cannot apply the "Fundamental Theorem of Calculus" to conclude that

$$\int_{-1}^2 x^{-3}\,dx = \frac{x^{-2}}{-2}\Big|_{x=-1}^{x=2} = -\frac{1}{8}+\frac{1}{2} = \frac{3}{8}$$

since $y = x^{-3}$ is not Riemann integrable on $[-1, 2]$.

Similar to Theorem 8.40, we also have the following

Theorem 8.46

If f is Riemann integrable on $[a, c]$ for all $a < c < b$, and f has an infinite discontinuity at a or b, then

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \qquad \forall\, a < c < b\,,$$

provided that the improper integrals on both sides converge or diverge to ∞ (or $-\infty$).

We can also consider improper integral $\displaystyle\int_a^b f(x)\,dx$ in which $a=-\infty$ or $b=\infty$, and f has an infinite discontinuity at c for $a<c<b$. In this case, we define

$$\int_a^\infty f(x)\,dx = \int_a^d f(x)\,dx + \int_d^\infty f(x)\,dx \qquad \forall\, d>c\,,$$
$$\int_{-\infty}^b f(x)\,dx = \int_{-\infty}^d f(x)\,dx + \int_d^b f(x)\,dx \qquad \forall\, d<c\,,$$

and etc. In other words, when the integrand and the domain of integration are unbounded, we divide the integral into improper integrals of one type and compute those integrals separately, pretending that the summing rule

$$\int_a^b f(x)\,dx = \int_a^{c_1} f(x)\,dx + \int_{c_1}^{c_2} f(x)\,dx + \cdots + \int_{c_{n-1}}^{c_n} f(x)\,dx + \int_{c_n}^b f(x)\,dx$$

also holds for improper integrals.

Example 8.47. Evaluate $\displaystyle\int_0^\infty \frac{dx}{\sqrt{x}(x+1)}$.

We observe that the integrand has an infinite discontinuity at 0, and the domain of integration is unbounded. Therefore,

$$\int_0^\infty \frac{dx}{\sqrt{x}(x+1)} = \int_0^1 \frac{dx}{\sqrt{x}(x+1)} + \int_1^\infty \frac{dx}{\sqrt{x}(x+1)}\,.$$

By the substitution $u=\sqrt{x}$, $du = \dfrac{dx}{2\sqrt{x}}$; thus

$$\int \frac{dx}{\sqrt{x}(x+1)} = \int \frac{2du}{u^2+1} = 2\arctan u + C = 2\arctan\sqrt{x} + C\,.$$

Therefore,

$$\begin{aligned}\int_0^1 \frac{dx}{\sqrt{x}(x+1)} &= \lim_{a\to 0^+}\int_a^1 \frac{dx}{\sqrt{x}(x+1)} = \lim_{a\to 0^+} 2\arctan\sqrt{x}\Big|_{x=a}^{x=1} \\ &= \lim_{a\to 0^+}\Big(2\cdot\frac{\pi}{4} - 2\arctan\sqrt{a}\Big) = \frac{\pi}{2}\end{aligned}$$

and

$$\begin{aligned}\int_1^\infty \frac{dx}{\sqrt{x}(x+1)} &= \lim_{b\to\infty}\int_1^b \frac{dx}{\sqrt{x}(x+1)} = \lim_{b\to\infty} 2\arctan\sqrt{x}\Big|_{x=1}^{x=b} \\ &= \lim_{b\to\infty}\Big(2\arctan\sqrt{b} - 2\cdot\frac{\pi}{4}\Big) = \pi - \frac{\pi}{2} = \frac{\pi}{2}\,.\end{aligned}$$

As a consequence,

$$\int_0^\infty \frac{dx}{\sqrt{x}(x+1)} = \frac{\pi}{2} + \frac{\pi}{2} = \pi\,.$$

Definition 8.48

Let $\displaystyle\int_a^b f(x)\,dx$, where a, b could be infinite, be an improper integral.

1. The improper integral $\displaystyle\int_a^b f(x)\,dx$ is said to be absolutely convergent or converge absolutely if $\displaystyle\int_a^b \big|f(x)\big|\,dx$ converges.
2. The improper integral $\displaystyle\int_a^b f(x)\,dx$ is said to be conditionally convergent or converge conditionally if $\displaystyle\int_a^b f(x)\,dx$ converges but $\displaystyle\int_a^b \big|f(x)\big|\,dx$ diverges (to ∞).

Remark 8.49. Even though it is not required in the definition that an absolutely convergent improper integral has to converge, it is in fact true an absolutely convergent improper integral converges.

Example 8.50. The improper integral $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$ is conditionally convergent but not absolutely convergent. To see that the improper integral is not absolutely convergent, we note that if $n \in \mathbb{N}$,

$$\begin{aligned}
\int_0^{2n\pi} \Big|\frac{\sin x}{x}\Big|\,dx &= \sum_{k=1}^n \int_{2(k-1)\pi}^{2k\pi} \Big|\frac{\sin x}{x}\Big|\,dx = \sum_{k=1}^n \int_0^{2\pi} \Big|\frac{\sin\big[x+2(k-1)\pi\big]}{x+2(k-1)\pi}\Big|\,dx \\
&= \sum_{k=1}^n \int_0^{2\pi} \frac{|\sin x|}{|x+2(k-1)\pi|}\,dx = \sum_{k=1}^n \int_0^{2\pi} \frac{|\sin x|}{2k\pi}\,dx \geqslant \frac{2}{\pi}\sum_{k=1}^n \frac{1}{k}\,;
\end{aligned}$$

thus by the fact that

$$\begin{aligned}
\sum_{k=1}^{2^n} \frac{1}{k} &= 1 + \frac{1}{2} + \Big(\frac{1}{3}+\frac{1}{4}\Big) + \Big(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\Big) + \cdots + \Big(\frac{1}{2^{n-1}+1} + \frac{1}{2^{n-1}+2} + \cdots + \frac{1}{2^n}\Big) \\
&\geqslant 1 + \frac{1}{2} + \Big(\frac{1}{4}+\frac{1}{4}\Big) + \Big(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\Big) + \cdots + \Big(\underbrace{\frac{1}{2^n} + \frac{1}{2^n} + \cdots + \frac{1}{2^n}}_{2^{n-1}\text{ terms}}\Big) \\
&= 1 + \underbrace{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}}_{n\text{ terms}} = \frac{n}{2} + 1 \geqslant \frac{n}{2}\,,
\end{aligned}$$

we find that

$$\int_0^{2^{n+1}\pi} \Big|\frac{\sin x}{x}\Big|\,dx \geqslant \frac{2}{\pi}\sum_{k=1}^{2^n}\frac{1}{k} \geqslant \frac{n}{\pi}$$

which approaches ∞ as $n \to \infty$.

Theorem 8.51: A special type of improper integral

$$\int_1^\infty \frac{dx}{x^p} = \begin{cases} \dfrac{1}{p-1} & \text{if } p > 1\,, \\ \text{diverges to } \infty & \text{if } p \leqslant 1\,. \end{cases}$$

- **Comparison Test for Improper Integrals**

In the last part of this section, we consider some criteria which can be used to judge if an improper integral converges or diverges, without evaluating the exact value of the improper integral.

Theorem 8.52: Direct Comparison Test

Let f and g be continuous functions and $0 \leqslant g(x) \leqslant f(x)$ on the interval $[a, \infty)$.

1. If the improper integral $\displaystyle\int_a^\infty f(x)\,dx$ converges, then the improper integral $\displaystyle\int_a^\infty g(x)\,dx$ converges.
2. If the improper integral $\displaystyle\int_a^\infty g(x)\,dx$ diverges to ∞, then the improper integral $\displaystyle\int_a^\infty f(x)\,dx$ diverges.

Similar result also holds for improper integrals given by other two cases in Definition 8.34 and the case with infinite discontinuities.

Proof. For $b > a$, define $G(b) = \displaystyle\int_a^b g(x)\,dx$ and $F(b) = \displaystyle\int_a^b f(x)\,dx$. By the Fundamental Theorem of Calculus, $F, G : [a, \infty) \to \mathbb{R}$ is differentiable (hence continuous). Since $0 \leqslant g(x) \leqslant f(x)$ on $[a, \infty)$, for all $b > a$ we have $0 \leqslant G(b) \leqslant F(b)$, and F, G are monotone increasing.

1. If the improper integral $\displaystyle\int_a^\infty f(x)\,dx$ converges, the $\lim\limits_{b\to\infty} F(b) = M$ exists. Since F is monotone increasing, $F(b) \leqslant M$ for all $b > a$; thus $G(b) \leqslant M$ for all $b > a$. By the monotonicity of G, $\lim\limits_{b\to\infty} G(b)$ exists.
2. If the improper integral $\displaystyle\int_a^\infty g(x)\,dx$ diverges to ∞, $\lim\limits_{b\to\infty} G(b) = \infty$; thus the fact that $G(b) \leqslant F(b)$ implies that $\lim\limits_{b\to\infty} F(b) = \infty$. $\square$

Example 8.53. Consider the improper integral $\displaystyle\int_1^\infty e^{-x^2}\,dx$. Note that $e^{-x^2} \leqslant e^{-x}$ for all

$x \in [1, \infty)$. Since

$$\int_1^\infty e^{-x}\,dx = \lim_{b\to\infty}\int_1^b e^{-x}\,dx = \lim_{b\to\infty}(-e^{-x})\Big|_{x=1}^{x=b} = \lim_{b\to\infty}(e^{-b} - e^{-1}) = -e^{-1}\,,$$

by Theorem 8.52 we find that the improper integral $\displaystyle\int_1^\infty e^{-x^2}\,dx$ converges.

Example 8.54. Consider the improper integral $\displaystyle\int_1^\infty \frac{\sin^2 x}{x^2}\,dx$. Note that $\dfrac{\sin^2 x}{x^2} \leqslant \dfrac{1}{x^2}$ for all $x \in [1, \infty)$. Since

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{b\to\infty}\int_1^b \frac{1}{x^2}\,dx = \lim_{b\to\infty}\Big(-\frac{1}{x}\Big)\Big|_{x=1}^{x=b} = \lim_{b\to\infty}\Big(\frac{1}{b} - 1\Big) = -1\,,$$

by Theorem 8.52 we find that the improper integral $\displaystyle\int_1^\infty e^{-x^2}\,dx$ converges.

Example 8.55 (The Gamma Function)**.** The Gamma function $\Gamma : (0, \infty) \to \mathbb{R}$ is defined by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\,dt\,.$$

We note that for each $x \in \mathbb{R}$, the integrand $f(t) = t^{x-1}e^{-t}$ is positive on $[0, \infty)$.

1. If $x \geqslant 1$, the function $y = t^{x-1}e^{-\frac{t}{2}}$ is differentiable on $[0, \infty)$ and has a maximum at the point $t = 2(x-1)$. Therefore,

   $$0 \leqslant f(t) \leqslant 2^{x-1}(x-1)^{x-1}e^{-\frac{t}{2}} \qquad \forall\, t \geqslant 0\,.$$

   By the fact that

   $$\int_0^\infty e^{-\frac{t}{2}}\,dt = \lim_{b\to\infty}\int_0^b e^{-\frac{t}{2}}\,dt = \lim_{b\to\infty}\Big(-2e^{-\frac{t}{2}}\Big)\Big|_{t=0}^{t=b} = \lim_{b\to\infty}\Big(2 - 2e^{-\frac{b}{2}}\Big) = 2,$$

   we find that the improper integral $\displaystyle\int_0^\infty t^{x-1}e^{-t}\,dt$ converges.

2. If $0 < x < 1$, the function f has an infinite discontinuity at 0. Therefore,

   $$\int_0^\infty t^{x-1}e^{-t}\,dt = \int_0^1 t^{x-1}e^{-t}\,dt + \int_1^\infty t^{x-1}e^{-t}\,dt\,.$$

   Again, the function $y = t^{x-1}e^{-\frac{t}{2}}$ is bounded from above by $2^{x-1}(x-1)^{x-1}$; thus the same reason as above show that the improper integral $\displaystyle\int_1^\infty t^{x-1}e^{-t}\,dt$ converges.

On the other hand, note that $f(t) \leqslant t^{x-1}$ for all $t \in [0,1]$. By the fact that

$$\int_0^1 t^{x-1}\,dt = \lim_{a\to 0^+}\int_a^1 t^{x-1}\,dx = \lim_{a\to 0^+}\frac{t^x}{x}\Big|_{t=a}^{t=1} = \lim_{a\to 0^+}\frac{1-a^x}{x} = \frac{1}{x}\,,$$

we find that the improper integral $\displaystyle\int_0^1 t^{x-1}e^{-t}\,dt$ converges. Therefore, the improper integral $\displaystyle\int_0^\infty t^{x-1}e^{-t}\,dt$ converges.

3. If $x \leqslant 0$, then $t^{x-1}e^{-t} \geqslant t^{x-1}e^{-1}$ for all $t \in [0,1]$. By the fact that

$$\int_0^1 t^{x-1}e^{-1}\,dt = \lim_{a\to 0^+}\int_a^1 t^{x-1}e^{-1}\,dt = \infty\,,$$

Theorem 8.52 implies that the improper integral $\displaystyle\int_0^1 t^{x-1}e^{-t}\,dt$ diverges to ∞. This implies that the improper integral $\displaystyle\int_0^\infty t^{x-1}e^{-t}\,dt$ diverges to ∞ as well.

Theorem 8.56: Limit Comparison Test

Let f and g be positive continuous functions on the interval $[a,\infty)$. If the limit $\displaystyle\lim_{x\to\infty}\frac{f(x)}{g(x)} = L$ for some $0 < L < \infty$, then

$$\int_a^\infty f(x)\,dx \;\text{ converges if and only if }\; \int_a^\infty g(x)\,dx \;\text{ converges.}$$

Similar result also holds for improper integrals given by other two cases in Definition 8.34 and the case with infinite discontinuities.

Proof. By the fact $\displaystyle\lim_{x\to\infty}\frac{f(x)}{g(x)} = L$, there exists $M > a$ such that

$$\Big|\frac{f(x)}{g(x)} - L\Big| < \frac{L}{2} \qquad \text{whenever} \quad x > M\,.$$

Therefore,

$$0 < \frac{L}{2}g(x) < f(x) < \frac{3L}{2}g(x) \qquad \text{whenever} \quad x > M\,.$$

By the direct comparison test,

$$\int_M^\infty f(x)\,dx \;\text{ converges if and only if }\; \int_M^\infty g(x)\,dx \;\text{ converges.}$$

The theorem is then concluded since $\int_a^M f(x)\,dx$ and $\int_a^M g(x)\,dx$ are both finite. $\square$

Example 8.57. Consider the improper integral $\int_1^\infty \frac{1+e^{-x}}{x}\,dx$. Since $\lim_{x\to\infty}\frac{(1+e^{-x})/x}{1/x}=1$, the limit comparison test implies that

$$\int_1^\infty \frac{1+e^{-x}}{x}\,dx \text{ converges if and only if } \int_1^\infty \frac{dx}{x} \text{ converges.}$$

By Theorem 8.51, we find that the integral $\int_1^\infty \frac{dx}{x}$ diverges; thus the improper integral $\int_1^\infty \frac{1+e^{-x}}{x}\,dx$ diverges.

Example 8.58. Consider the improper integral $\int_0^{\frac{\pi}{4}} \frac{dx}{x+\tan x}$. Note that this is an improper integral with infinite discontinuity at $x=0$. Since

$$\lim_{x\to 0^+}\frac{x+\tan x}{x}=1+\lim_{x\to 0^+}\frac{\tan x}{x}=1+\lim_{x\to 0^+}\frac{\sin x}{x\cos x}=2\,,$$

the limit comparison test implies that

$$\int_0^{\frac{\pi}{4}} \frac{dx}{x+\tan x} \text{ converges if and only if } \int_0^{\frac{\pi}{4}} \frac{dx}{x} \text{ converges.}$$

Since the improper integral $\int_0^{\frac{\pi}{4}} \frac{dx}{x}$ diverges (to ∞), we must have $\int_0^{\frac{\pi}{4}} \frac{dx}{x+\tan x}$ diverges.

Example 8.59. Determine the convergence of the improper integral $\int_0^\infty \frac{dx}{\sqrt[3]{x^4-x^2}}$.

Note that $\frac{1}{\sqrt[3]{x^4-x^2}}=x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}$. In the interval $[0,\infty)$, the integrand has singular points at 0 and 1. Write

$$\int_0^\infty \frac{dx}{\sqrt[3]{x^4-x^2}}=\int_0^{\frac{1}{2}} \frac{dx}{\sqrt[3]{x^4-x^2}}+\int_{\frac{1}{2}}^1 \frac{dx}{\sqrt[3]{x^4-x^2}}+\int_1^2 \frac{dx}{\sqrt[3]{x^4-x^2}}+\int_2^\infty \frac{dx}{\sqrt[3]{x^4-x^2}}\,. \tag{8.5.1}$$

1. Let $f(x)=-x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}$ and $g(x)=x^{-\frac{2}{3}}$. Then f,g are positive continuous on $\big[a,\frac{1}{2}\big]$ for all $a>0$. Moreover,

$$\lim_{x\to 0^+}\frac{f(x)}{g(x)}=\lim_{x\to 0^+}\big[-(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}\big]=1>0\,,$$

and

$$\int_0^{\frac{1}{2}} g(x)\,dx = \lim_{a\to 0^+}\int_a^{\frac{1}{2}} x^{-\frac{2}{3}}\,dx = \lim_{a\to 0^+} 3x^{\frac{1}{3}}\Big|_{x=a}^{x=\frac{1}{2}} = \frac{3}{\sqrt[3]{2}}$$

which shows that the improper integral $\int_0^{\frac{1}{2}} g(x)\,dx$ converges. Therefore, the limit comparison test implies that $\int_0^{\frac{1}{2}} f(x)\,dx = -\int_0^{\frac{1}{2}} \frac{dx}{\sqrt[3]{x^4-x^2}}$ converges.

2. Let $f(x) = -x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}$ and $g(x) = -(x-1)^{\frac{1}{3}}$. Then f, g are positive continuous on $\big[\frac{1}{2}, b\big]$ for all $\frac{1}{2} < b < 1$. Moreover,

$$\lim_{x\to 1^-}\frac{f(x)}{g(x)} = \lim_{x\to 1^-} x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}} = 2^{-\frac{1}{3}} > 0\,,$$

and

$$\int_{\frac{1}{2}}^1 g(x)\,dx = -\lim_{b\to 1^-}\int_{\frac{1}{2}}^b (x-1)^{-\frac{1}{3}}\,dx = -\lim_{b\to 1^-}\frac{3}{2}(x-1)^{\frac{2}{3}}\Big|_{x=\frac{1}{2}}^{x=b} = \frac{3}{2\sqrt[3]{4}}$$

which shows that the improper integral $\int_{\frac{1}{2}}^1 g(x)\,dx$ converges. Therefore, the limit comparison test implies that $\int_{\frac{1}{2}}^1 f(x)\,dx = -\int_{\frac{1}{2}}^1 \frac{dx}{\sqrt[3]{x^4-x^2}}$ converges.

3. Similar to the previous case, we let $f(x) = x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}$ and $g(x) = (x-1)^{\frac{1}{3}}$. Then f, g are positive continuous on $[a, 2]$ for all $1 < a < 2$. Moreover,

$$\lim_{x\to 1^+}\frac{f(x)}{g(x)} = \lim_{x\to 1^+} x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}} = 2^{-\frac{1}{3}} > 0\,,$$

and

$$\int_1^2 g(x)\,dx = \lim_{a\to 1^+}\int_a^2 (x-1)^{-\frac{1}{3}}\,dx = -\lim_{a\to 1^+}\frac{3}{2}(x-1)^{\frac{2}{3}}\Big|_{x=a}^{x=2} = \frac{3}{2}$$

which shows that the improper integral $\int_1^2 g(x)\,dx$ converges. Therefore, the limit comparison test implies that $\int_1^2 f(x)\,dx = \int_1^2 \frac{dx}{\sqrt[3]{x^4-x^2}}$ converges.

4. Let $f(x) = x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}$ and $g(x) = x^{-\frac{4}{3}}$. Then f, g are positive continuous on $[2, b]$ for all $b > 2$. Moreover,

$$\lim_{x\to\infty}\frac{f(x)}{g(x)} = \lim_{x\to\infty}\frac{x^{-\frac{2}{3}}(x+1)^{-\frac{1}{3}}(x-1)^{-\frac{1}{3}}}{x^{-\frac{4}{3}}} = \lim_{x\to\infty}\sqrt[3]{\frac{x^2}{(x-1)(x+1)}} = 1 > 0\,,$$

and

$$\int_2^\infty g(x)\,dx = \lim_{b\to\infty}\int_2^b x^{-\frac{4}{3}}\,dx = -\lim_{b\to\infty} 3x^{-\frac{1}{3}}\Big|_{x=2}^{x=b} = 3$$

which shows that the improper integral $\displaystyle\int_2^\infty g(x)\,dx$ converges. Therefore, the limit comparison test implies that $\displaystyle\int_2^\infty f(x)\,dx = \int_2^\infty \frac{dx}{\sqrt[3]{x^4-x^2}}$ converges.

Since the four improper integrals on the right-hand side of (8.5.1) converges, we find that the improper integral $\displaystyle\int_0^\infty \frac{dx}{\sqrt[3]{x^4-x^2}}$ converges.

8.5.1 The Laplace transform（補充，不考）

Definition 8.60: Laplace Transform

Let $f:[0,\infty)\to\mathbb{R}$ be continuous. The Laplace transform of f, denoted by $\mathscr{L}(f)$, is the function defined by

$$\mathscr{L}(f)(s) = \int_0^\infty e^{-st}f(t)\,dt\ \left(=\lim_{R\to\infty}\int_0^R e^{-st}f(t)dt\right),$$

and the domain of $\mathscr{L}(f)$ is the set consisting of all numbers s for which the integral converges.

Remark 8.61. In general, the Laplace transform of f can be defined, without assuming that f is continuous on $[0,\infty)$, as long as the integral $\displaystyle\int_0^\infty e^{-st}f(t)\,dt$ makes sense. Moreover, if f is continuous and satisfies

$$\big|f(t)\big| \leqslant Me^{\alpha t} \qquad \forall\, t\in[0,\infty)\,, \tag{8.5.2}$$

then $\mathscr{L}(f)(s)$ exists for all $s>\alpha$. A function f is said to be of exponential order α if there exist $M>0$ such that the growth condition (8.5.2) holds.

Example 8.62. Let $f:[0,\infty)\to\mathbb{R}$ be given by $f(t)=t^p$ for some $p>-1$. Recall that the Gamma function $\Gamma:(0,\infty)\to\mathbb{R}$ is defined by

$$\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}\,dt\,.$$

We note that if $-1 < p < 0$, f is not of exponential order a for all $a \in \mathbb{R}$; however, the Laplace transform of f still exists. In fact, for $s > 0$,

$$\mathscr{L}(f)(s) = \lim_{R\to\infty} \int_0^R e^{-st} t^p \, dt = \lim_{R\to\infty} \int_0^{sR} e^{-t} \Big(\frac{t}{s}\Big)^p \frac{dt}{s} = \frac{\Gamma(p+1)}{s^{p+1}} \, .$$

In particular, if $p = n \in \mathbb{N} \cup \{0\}$, then

$$\mathscr{L}(f)(s) = \frac{n!}{s^{n+1}} \qquad \forall\, s > 0 \, .$$

Example 8.63. Let $g : [0, \infty) \to \mathbb{R}$ be given by $g(t) = e^{at} \sin(bt)$ for some $b \neq 0$. Using (8.2.1), we find that

$$\int e^{(a-s)t} \sin(bt) \, dt = \frac{1}{(s-a)^2 + b^2} \Big[(a-s) e^{(a-s)t} \sin(bt) - b e^{(a-s)t} \cos(bt) \Big] + C \, .$$

Therefore, for $s > a$,

$$\begin{aligned} \mathscr{L}(g)(s) &= \int_0^\infty e^{(a-s)t} \sin(bt) \, dt \\ &= \lim_{b\to\infty} \frac{1}{(s-a)^2 + b^2} \Big[(a-s) e^{(a-s)t} \sin(bt) - b e^{(a-s)t} \cos(bt) \Big] \Big|_{t=0}^{t=b} \\ &= \frac{b}{(s-a)^2 + b^2} \, . \end{aligned}$$

Similarly, if $h(t) = e^{at} \cos(bt)$, using (8.2.2) we find that for $s > a$,

$$\begin{aligned} \mathscr{L}(h)(s) &= \int_0^\infty e^{(a-s)t} \cos(bt) \, dt \\ &= \lim_{b\to\infty} \frac{1}{(s-a)^2 + b^2} \Big[(a-s) e^{(a-s)t} \cos(bt) + b e^{(a-s)t} \sin(bt) \Big] \Big|_{t=0}^{t=b} \\ &= \frac{s-a}{(s-a)^2 + b^2} \, . \end{aligned}$$

Theorem 8.65: Linearity of the Laplace transform

Let $f, g : [0, \infty) \to \mathbb{R}$ be functions whose Laplace transform exist for $s > \alpha$ and c be a constant. Then for $s > \alpha$,

1. $\mathscr{L}(f+g)(s) = \mathscr{L}(f)(s) + \mathscr{L}(g)(s)$.
2. $\mathscr{L}(cf)(s) = c\mathscr{L}(f)(s)$.

Theorem 8.66

Suppose that $f : [0,\infty) \to \mathbb{R}$ is a function such that $f, f', f'', \cdots, f^{(n-1)}$ are continuous of exponential order α, and $f^{(n)}$ is piecewise continuous. Then $\mathscr{L}(f^{(n)})(s)$ exists for all $s > \alpha$, and

$$\mathscr{L}(f^{(n)})(s) = s^n \mathscr{L}(f)(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \cdots - s f^{(n-2)}(0) - f^{(n-1)}(0)\,. \quad (8.5.3)$$

Proof. We prove by induction. Suppose that f is continuously differentiable on $[0,\infty)$ and is of exponential order α. Then for $s > \alpha$,

$$\begin{aligned}\int_0^\infty e^{-st} f'(t)\,dt &= \lim_{b\to\infty} \int_0^b e^{-st} f'(t)\,dt = \lim_{b\to\infty} \Big[e^{-st} f(t)\Big|_{t=0}^{t=b} + s\int_0^b e^{-st} f(t)\,dt \Big] \\ &= s\int_0^\infty e^{-st} f(t)\,dt\Big] - f(0) + \lim_{b\to\infty} e^{-sb} f(b) = s\mathscr{L}(f)(s) - f(0)\end{aligned}$$

which shows that (8.5.3) holds for $n = 1$ and all continuously differentiable f.

Now suppose that (8.5.3) holds for all k-times continuously differentiable function f. Then if $s > \alpha$ and f is $(k+1)$-times continuously differentiable function on $[0,\infty)$,

$$\begin{aligned}\mathscr{L}(f^{(k+1)})(s) &= \mathscr{L}\big((f')^{(k)}\big)(s) \\ &= s^k \mathscr{L}(f')(s) - s^{k-1} f'(0) - s^{k-2} (f')'(0) - \cdots - s(f')^{(n-2)}(0) - (f')^{(n-1)}(0) \\ &= s^k \big[s\mathscr{L}(f)(s) - f(0)\big] - s^{k-1} f'(0) - s^{k-2} f''(0) - \cdots - s f^{(n-1)}(0) - f^{(n)}(0) \\ &= s^{k+1} \mathscr{L}(f)(s) - s^k f(0) - s^{k-1} f'(0) - s^{k-2} f''(0) - \cdots - s f^{(n-1)}(0) - f^{(n)}(0)\end{aligned}$$

which implies that (8.5.3) holds for the case $n = k+1$. The theorem is then concluded by induction. □

- **Applications in solving the ordinary differential equations**

Let $a_0, a_1, \cdots, a_{n-1}$, $y_0, y_1, \cdots, y_{n-1}$ be given numbers, and $g : [0,\infty) \to \mathbb{R}$ be a continuous function of exponential order. The idea of solving an ordinary differential equation (here y is the unknown function to be solved) of the form

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = g(s)\,, \quad (8.5.4a)$$

$$y(0) = y_0, y'(0) = y_1, \cdots, y^{(n-1)}(0) = y_{n-1}\,, \quad (8.5.4b)$$

using the method of the Laplace transform is based on the following facts:

1. The Laplace transform is a one-to-one mapping in the sense that if f and g are continuous function such that $\mathscr{L}(f) = \mathscr{L}(g)$ for $s > \alpha$, then $f = g$ on $[0, \infty)$.

2. The solution of (8.5.4) is of exponential order α (so that the Laplace transform of derivatives of y can be computed using Theorem 8.66).

Under these two facts, we then take the Laplace transform of (8.5.4a) and apply Theorem 8.65 and 8.66 to obtain, by letting $Y(s) = \mathscr{L}(y)(s)$, that

$$\begin{aligned}
&a_n\big[s^n Y(s) - s^{n-1}y_0 - s^{n-2}y_1 - \cdots - sy_{n-2} - y_{n-1}\big] \\
&\quad + a_{n-1}\big[s^{n-1}Y(s) - s^{n-2}y_0 - s^{n-3}y_1 - \cdots - sy_{n-3} - y_{n-2}\big] \\
&\quad + a_{n-2}\big[s^{n-2}Y(s) - s^{n-3}y_0 - s^{n-4}y_1 - \cdots - sy_{n-4} - y_{n-3}\big] \\
&\quad + \cdots + a_1\big[sY(s) - y_0\big] + a_0 Y(s) = \mathscr{L}(g)(s)\,;
\end{aligned}$$

thus

$$\begin{aligned}
Y(s) &= \frac{1}{a_n s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-2} + \cdots + a_1 s + a_0} \times \\
&\quad \times \Big[\mathscr{L}(g)(s) + y_0(a_n s^{n-1} + a_{n-1}s^{n-2} + \cdots + a_2 s + a_1) \\
&\qquad + y_1(a_n s^{n-2} + a_{n-1}s^{n-2} + \cdots + a_3 s + a_2) + \cdots\cdots + y_{n-2}(a_n s + a_{n-1}) + y_{n-1}\Big] \\
&= \frac{1}{a_n s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-3} + \cdots + a_1 s + a_0}\Big[\mathscr{L}(g)(s) + \sum_{j=0}^{n-1} y_j \sum_{\ell=0}^{n-j-1} a_{n-\ell}s^{n-j-\ell-1}\Big].
\end{aligned}$$

The final step is to identify which function gives the Laplace transform above.

Example 8.64. Find the function y satisfying

$$y'' + 2y' + 5y = \sin t\,, \qquad y(0) = 1\,,\ y'(0) = 0\,.$$

Using the result in Example 8.63 and Theorem 8.66, with Y denoting $\mathscr{L}(y)$ we find that

$$s^2 Y(s) - s + 2\big[sY(s) - 1\big] + 5Y(s) = \frac{1}{s^2+1} \qquad \forall\, s > a$$

for some a. Therefore,

$$Y(s) = \frac{1}{s^2+2s+5}\Big(\frac{1}{s^2+1} + s + 2\Big) = \frac{s+2}{(s+1)^2+2^2} + \frac{1}{(s^2+2s+5)(s^2+1)}\,.$$

Writing the last term as the sum of partial fractions, we have

$$\frac{1}{(s^2+2s+5)(s^2+1)} = \frac{1}{10}\Big(\frac{s}{s^2+2s+5} - \frac{s-2}{s^2+1}\Big);$$

thus

$$\begin{aligned} Y(s) &= \frac{s+2}{(s+1)^2+2^2} + \frac{1}{10}\frac{s}{(s+1)^2+2^2} - \frac{1}{10}\frac{s-2}{s^2+1} \\ &= \frac{11}{10}\frac{s+1}{(s+1)^2+2^2} + \frac{9}{20}\frac{2}{(s+1)^2+2^2} - \frac{1}{10}\frac{s}{s^2+1} + \frac{1}{5}\frac{1}{s^2+1}. \end{aligned}$$

Therefore, Fact 1 and Example 8.63 imply that

$$y(t) = \frac{11}{10}e^{-t}\cos(2t) + \frac{9}{20}e^{-t}\sin(2t) - \frac{1}{10}\cos t + \frac{1}{5}\sin t.$$

8.6 Exercise

Problem 8.1. Find the following indefinite integrals.

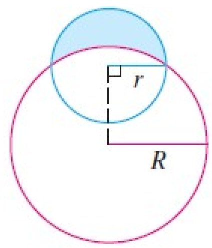
1. $\displaystyle\int x\csc x\cot x\,dx$
2. $\displaystyle\int \frac{\sqrt{1+\ln x}}{x\ln x}\,dx$
3. $\displaystyle\int x\sin^2 x\,dx$
4. $\displaystyle\int \exp(\sqrt[3]{x})\,dx$
5. $\displaystyle\int x\arcsin x\,dx$
6. $\displaystyle\int x\arctan x\,dx$
7. $\displaystyle\int x^2\arctan x\,dx$
8. $\displaystyle\int \ln(x^2-1)\,dx$
9. $\displaystyle\int \sin\sqrt{ax}\,dx$
10. $\displaystyle\int x\tan^2 x\,dx$
11. $\displaystyle\int x^5 e^{-x^3}\,dx$
12. $\displaystyle\int \frac{x\ln x}{\sqrt{x^2-1}}\,dx$
13. $\displaystyle\int \sqrt{x}e^{\sqrt{x}}\,dx$
14. $\displaystyle\int \frac{\arctan\sqrt{x}}{\sqrt{x}}\,dx$
15. $\displaystyle\int \frac{\ln(x+1)}{x^2}\,dx$
16. $\displaystyle\int \frac{dx}{x^4\sqrt{x^2-2}}$
17. $\displaystyle\int \sqrt{\tan x}\,dx$
18. $\displaystyle\int x\sin^2 x\cos x\,dx$

Problem 8.2. The function $y = e^{x^2}$ and $y = x^2e^{x^2}$ don't have elementary anti-derivatives, but $y = (2x^2+1)e^{x^2}$ does. Find the indefinite integral $\displaystyle\int (2x^2+1)e^{x^2}\,dx$.

Problem 8.3. Obtain a recursive formula for $\displaystyle\int x^p(ax^n+b)^q\,dx$ and use this relation to find the indefinite integral $\displaystyle\int x^3(x^7+1)^4\,dx$.

Problem 8.4. Obtain a recursive formula for $\displaystyle\int x^m(\ln x)^n\,dx$ and use this relation to find the indefinite integral $\displaystyle\int x^4(\ln x)^3\,dx$.

Problem 8.5. Find the area of the crescent-shaped region (called a lune) bounded by arcs of circles with radii r and R. (See the figure)



Problem 8.6. Complete the following.

1. Let $f : [a,b] \to [c,d]$ be a continuously differentiable increasing function. Suppose that f has an inverse f^{-1}. Show that
$$\int_a^b f(x)\,dx + \int_c^d f^{-1}(y)\,dy = bf(b) - af(a)\,. \tag{8.6.1}$$

2. How about if f is decreasing?

3. Use (8.6.1) to compute $\displaystyle\int_0^1 \arcsin x\,dx$ and $\displaystyle\int_0^1 \arctan x\,dx$.

4. Let F be an anti-derivative of a continuously differentiable function f with inverse f^{-1}. Find an anti-derivative of f^{-1} in terms of f and F.

Problem 8.7. For $n \in \mathbb{N} \cup \{0\}$, the Legendre polynomial of degree n, denoted by P_n, is defined by
$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n\,.$$

1. Show that $\displaystyle\int_{-1}^1 P_n(x)P_m(x)\,dx = 0$ if $m \neq n$.

2. Show that $\displaystyle\int_{-1}^1 P_n(x)^2\,dx = \frac{2}{2n+1}$ for all $n \in \mathbb{N} \cup \{0\}$.

3. Show that $\displaystyle\int_{-1}^1 x^m P_n(x)\,dx = 0$ if $m < n$.

4. Evaluate $\displaystyle\int_{-1}^1 x^n P_n(x)\,dx$.

Problem 8.8. Let $\alpha_1, \alpha_2, \cdots, \alpha_n$ be distinct real numbers, and

$$g(x) = \prod_{k=1}^{n}(x - \alpha_k) \equiv (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n)\,.$$

Use the partial fraction expansion to prove Newton's formula

$$\frac{\alpha_1^k}{g'(\alpha_1)} + \frac{\alpha_2^k}{g'(\alpha_2)} + \cdots + \frac{\alpha_n^k}{g'(\alpha_n)} = \begin{cases} 0 & \text{for } k = 0, 1, 2, \cdots, n-2\,, \\ 1 & \text{for } k = n-1\,. \end{cases}$$

Hint: By partial fraction, for $k < n-1$

$$\frac{x^k}{(x - \alpha_2)(x - \alpha_3)\cdots(x - \alpha_n)} = \frac{A_2}{x - \alpha_2} + \frac{A_3}{x - \alpha_3} + \cdots + \frac{A_n}{x - \alpha_n}\,.$$

Show that $A_j = \dfrac{\alpha_j^k(\alpha_j - \alpha_1)}{g'(\alpha_j)}$ and conclude from here. Do the same for the case $k = n-1$.

Problem 8.9. Find at least two ways to compute the following integrals.

1. $\displaystyle\int \frac{x-1}{x^2 - 4x - 5}\,dx$
2. $\displaystyle\int \frac{3x^2 - 2}{x^3 - 2x - 1}\,dx$
3. $\displaystyle\int \frac{1 + 4\cot x}{4 - \cot x}\,dx$
4. $\displaystyle\int \frac{1}{x(x^4 + 1)}\,dx$
5. $\displaystyle\int \frac{4}{\tan x - \sec x}\,dx$
6. $\displaystyle\int \frac{2}{x^6 + x}\,dx$

Problem 8.10. Find the following indefinite integrals using the techniques of partial fractions.

1. $\displaystyle\int \frac{x}{x^4 - 1}\,dx$
2. $\displaystyle\int \frac{x}{x^4 + 4x^2 + 3}\,dx$
3. $\displaystyle\int \frac{x-1}{x^2 - 4x + 5}\,dx$
4. $\displaystyle\int \frac{x^3 + 1}{x^3 - x^2}\,dx$
5. $\displaystyle\int \frac{1}{x^6 + 1}\,dx$
6. $\displaystyle\int \frac{1}{(x-2)(x^2 + 4)}\,dx$
7. $\displaystyle\int \frac{1}{x + 4 + 4\sqrt{x+1}}\,dx$
8. $\displaystyle\int \frac{1}{x\sqrt{4x+1}}\,dx$
9. $\displaystyle\int \frac{1}{x^2\sqrt{4x+1}}\,dx$
10. $\displaystyle\int \frac{1}{x + \sqrt[3]{x}}\,dx$
11. $\displaystyle\int \frac{1}{1 + 2e^x - e^{-x}}\,dx$
12. $\displaystyle\int \frac{1}{e^{3x} - e^x}\,dx$
13. $\displaystyle\int \frac{\sin x \cos x}{\sin^4 x + \cos^4 x}\,dx$
14. $\displaystyle\int \frac{1}{3 - 2\sin x}\,dx$
15. $\displaystyle\int \frac{1}{1 + \sin\theta + \cos\theta}\,d\theta$

Problem 8.11. Determine if the following improper integral converges or not.

1. $\displaystyle\int_0^\infty \frac{dx}{\sqrt[3]{x^4 - x^2}}$
2. $\displaystyle\int_1^\infty \frac{dx}{x(\ln x)^\alpha}$
3. $\displaystyle\int_1^\infty \frac{\ln x}{x^\alpha}\,dx$
4. $\displaystyle\int_{10}^\infty \frac{dx}{x(\ln\ln x)^\alpha}$
5. $\displaystyle\int_0^\pi \frac{dx}{\sqrt{x} + \sin x}$
6. $\displaystyle\int_0^\pi \frac{dx}{x - \sin x}$
7. $\displaystyle\int_0^{\ln 2} x^{-2} e^{-\frac{1}{x}}\,dx$
8. $\displaystyle\int_0^1 \frac{e^{-\sqrt{x}}}{\sqrt{x}}\,dx$

9. $\displaystyle\int_1^\infty \frac{dx}{\sqrt{e^x - x}}$ 10. $\displaystyle\int_{-\infty}^\infty \frac{dx}{e^x + e^{-x}}$ 11. $\displaystyle\int_\pi^\infty \frac{1+\sin x}{x^2}\,dx$ 12. $\displaystyle\int_{-1}^1 \ln|x|\,dx.$

Problem 8.12. Compute $\displaystyle\int_0^1 \frac{\ln(x+1)}{x^2+1}\,dx.$

Hint: Let $I(t) = \displaystyle\int_0^1 \frac{\ln(tx+1)}{x^2+1}\,dx$. Use the fact that $\displaystyle\frac{d}{dt}\int_0^1 \frac{\ln(tx+1)}{x^2+1}\,dx = \int_0^1 \frac{\partial}{\partial t}\frac{\ln(tx+1)}{x^2+1}\,dx$, where $\displaystyle\frac{\partial}{\partial t}f(x,t)$ is the derivative of f w.r.t. t variable by treating x as a constant.

Problem 8.13. Compute $\displaystyle\int_0^1 \frac{x-1}{\ln x}\,dx.$

Hint: Let $I(t) = \displaystyle\int_0^1 \frac{x^t-1}{\ln x}\,dx$. Use the fact that $\displaystyle\frac{d}{dt}I(t) = \int_0^1 \frac{\partial}{\partial t}\frac{x^t-1}{\ln x}\,dx.$

Problem 8.14. Compute $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx.$

Hint: Let $I(t) = \displaystyle\int_0^\infty \frac{e^{-tx}\sin x}{x}\,dx$. Use the fact that $I'(t) = \displaystyle\int_0^\infty \frac{\partial}{\partial t}\frac{e^{-tx}\sin x}{x}\,dx$ and use the fact that $\displaystyle\lim_{t\to\infty} I(t) = 0$.

Chapter 9

Infinite Series

9.1 Sequences

Definition 9.1: Sequence

A ***sequence*** of real numbers (or simply a real sequence) is a function $f : \mathbb{N} \to \mathbb{R}$. The collection of numbers $\big\{f(1), f(2), f(3), \cdots\big\}$ are called ***terms*** of the sequence and the value of f at n is called the ***n-th term*** of the sequence. We usually use f_n to denote the n-th term of a sequence $f : \mathbb{N} \to \mathbb{R}$, and this sequence is usually denoted by $\{f_n\}_{n=1}^{\infty}$ or simply $\{f_n\}$.

Example 9.2. Let $f : \mathbb{N} \to \mathbb{R}$ be the sequence defined by $f(n) = 3 + (-1)^n$. Then f is a real sequence. Its terms are $\{2, 4, 2, 4, \cdots\}$.

Example 9.3. A sequence can also be defined recursively. For example, let $\{a_n\}_{n=1}^{\infty}$ be defined by

$$a_{n+1} = \sqrt{2a_n}\,, \qquad a_1 = \sqrt{2}\,.$$

Then $a_2 = \sqrt{2\sqrt{2}}$, $a_3 = \sqrt{2\sqrt{2\sqrt{2}}}$, and etc. The general form of a_n is given by

$$a_n = 2^{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\frac{1}{2^n}} = 2^{\frac{2^n-1}{2^n}}\,.$$

There are also sequences that are defined recursively but it is difficult to obtain the general form of the sequence. For example, let $\{b_n\}_{n=1}^{\infty}$ be defined by

$$b_{n+1} = \sqrt{2 + b_n}\,, \qquad b_1 = \sqrt{2}\,.$$

Then $b_2 = \sqrt{2+\sqrt{2}}$, $b_3 = \sqrt{2+\sqrt{2+\sqrt{2}}}$, and etc.

Remark 9.4. Occasionally, it is convenient to begin a sequence with the 0-th term or even the k-th term. In such cases, we write $\{a_n\}_{n=0}^{\infty}$ and $\{a_n\}_{n=k}^{\infty}$ to denote the sequences.

Similar to the concept of the limit of functions, we would like to consider the limit of sequences; that is, we would like to know to which value the n-th term of a sequence approaches as n become larger and larger.

Definition 9.5

A sequence of real numbers $\{a_n\}_{n=1}^{\infty}$ is said to ***converge to*** L if for every $\varepsilon > 0$, there exists $N > 0$ such that

$$|a_n - L| < \varepsilon \quad \text{whenever} \quad n \geqslant N .$$

Such an L (must be a real number and) is called a ***limit*** of the sequence. If $\{a_n\}_{n=1}^{\infty}$ converges to L, we write $a_n \to x$ as $n \to \infty$.
A sequence of real number $\{a_n\}_{n=1}^{\infty}$ is said to be ***convergent*** if there exists $L \in \mathbb{R}$ such that $\{a_n\}_{n=1}^{\infty}$ converges to L. If no such L exists we say that $\{a_n\}_{n=1}^{\infty}$ ***does not converge*** or simply ***diverges***.

Motivation: Intuitively, we expect that a sequence of real numbers $\{a_n\}_{n=1}^{\infty}$ converges to a number L if "outside any open interval containing L there are only finitely many $a_n's$". The statement inside " " can be translated into the following mathematical statement:

$$\forall\, \varepsilon > 0, \#\big\{n \in \mathbb{N} \,\big|\, a_n \notin (L - \varepsilon, L + \varepsilon)\big\} < \infty\,, \tag{9.1.1}$$

where $\#A$ denotes the number of points in the set A. One can easily show that the convergence of a sequence defined by (9.1.1) is equivalent to Definition 9.5.

In the definition above, we do not exclude the possibility that there are two different limits of a convergent sequence. In fact, this is never the case because of the following

Proposition 9.6

If $\{a_n\}_{n=1}^{\infty}$ is a sequence of real numbers, and $a_n \to a$ and $a_n \to b$ as $n \to \infty$, then $a = b$. (若收斂則極限唯一).

We will not prove this proposition and treat it as a fact.

• **Notation**: Since the limit of a convergent sequence is unique, we use $\lim\limits_{n\to\infty} a_n$ to denote this unique limit of a convergent sequence $\{a_n\}_{n=1}^{\infty}$.

Theorem 9.7

Let L be a real number, and $f : [1, \infty) \to \mathbb{R}$ be a function of a real variable such that $\lim_{x\to\infty} f(x) = L$. If $\{a_n\}_{n=1}^{\infty}$ is a sequence such that $f(n) = a_n$ for every positive integer n, then

$$\lim_{n\to\infty} a_n = L\,.$$

Example 9.8. The limit of the sequence $\{e_n\}_{n=1}^{\infty}$ defined by $e_n = \left(1 + \frac{1}{n}\right)^n$ is e.

When a sequence $\{a_n\}_{n=1}^{\infty}$ is given by evaluating a differentiable function $f : [1, \infty) \to \mathbb{R}$ on $\mathbb{N}$, sometimes we can use L'Hôspital's rule to find the limit of the sequence.

Example 9.9. The limit of the sequence $\{a_n\}_{n=1}^{\infty}$ defined by $a_n = \dfrac{n^2}{2^n - 1}$ is

$$\lim_{x\to\infty} \frac{x^2}{2^x - 1} = \lim_{x\to\infty} \frac{2x}{2^x \ln 2} = \lim_{x\to\infty} \frac{2}{2^x (\ln 2)^2} = 0\,.$$

There are cases that a sequence cannot be obtained by evaluating a function defined on $[1, \infty)$. In such cases, the limit of a sequence cannot be computed using L'Hôspital's rule and it requires more techniques to find the limit.

Example 9.10. The limit of the sequence $\{s_n\}_{n=1}^{\infty}$ defined by $s_n = \dfrac{n!}{n^{n+\frac{1}{2}} e^{-n}}$ is $\sqrt{2\pi}$; that is,

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\, n^n e^{-n}} = 1\,. \tag{9.1.2}$$

Similar to Theorem 1.14, we have the following

Theorem 9.11

Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences of real numbers such that $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} b_n = K$. Then

1. $\lim_{n\to\infty} (a_n \pm b_n) = L \pm K$.

2. $\lim_{n\to\infty} (a_n b_n) = LK$. In particular, $\lim_{n\to\infty} (ca_n) = cL$ if c is a real number.

3. $\lim_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{L}{K}$ if $K \neq 0$.

Theorem 9.12: Squeeze Theorem

Let $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ and $\{c_n\}_{n=1}^{\infty}$ be sequences of real numbers such that $a_n \leqslant c_n \leqslant b_n$ for all $n \geqslant N$. If $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n = L$, then $\lim\limits_{n\to\infty} c_n = L$.

Theorem 9.13: Absolute Value Theorem

Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers. If $\lim\limits_{n\to\infty} |a_n| = 0$, then $\lim\limits_{n\to\infty} a_n = 0$.

Proof. Let $\{b_n\}_{n=1}^{\infty}$ and $\{c_n\}_{n=1}^{\infty}$ be sequence of real numbers defined by $b_n = -|a_n|$ and $c_n = |a_n|$. Then $b_n \leqslant a_n \leqslant c_n$ for all $n \in \mathbb{N}$. Since $\lim\limits_{n\to\infty} |a_n| = 0$, Theorem 9.11 implies that $\lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} c_n = 0$ and the Squeeze Theorem further implies that $\lim\limits_{n\to\infty} a_n = 0$. $\square$

Definition 9.14: Monotonicity of Sequences

A sequence $\{a_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ is said to be

1. (**monotone**) **increasing** if $a_{n+1} \geqslant a_n$ for all $n \in \mathbb{N}$;
2. (**monotone**) **decreasing** if $a_{n+1} \leqslant a_n$ for all $n \in \mathbb{N}$;
3. ***monotone*** if $\{a_n\}_{n=1}^{\infty}$ is an increasing sequence or a decreasing sequence.

Example 9.15. The sequence $\{s_n\}_{n=2}^{\infty}$ defined in Example 9.10 is a monotone decreasing sequence.

Definition 9.16: Boundedness of Sequences

Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers.

1. $\{a_n\}_{n=1}^{\infty}$ is said to be ***bounded***（有界的）if there exists $M \in \mathbb{R}$ such that $|a_n| \leqslant M$ for all $n \in \mathbb{N}$.
2. $\{a_n\}_{n=1}^{\infty}$ is said to be ***bounded from above***（有上界）if there exists $B \in \mathbb{R}$, called an ***upper bound*** of the sequence, such that $a_n \leqslant B$ for all $n \in \mathbb{N}$. Such a number B is called an upper bound of the sequence.
3. $\{a_n\}_{n=1}^{\infty}$ is said to be ***bounded from below***（有下界）if there exists $A \in \mathbb{R}$, called a ***lower bound*** of the sequence, such that $A \leqslant a_n$ for all $n \in \mathbb{N}$. Such a number A is called a lower bound of the sequence.

Example 9.17. The sequence $\{a_n\}_{n=1}^{\infty}$ defined by $a_n = n$ is bounded from below by 0 by not bounded from above.

Proposition 9.18

A convergent sequence of real numbers is bounded（數列收斂必有界）.

Proof. Let $\{a_n\}_{n=1}^{\infty}$ be a convergent sequence with limit L. Then by the definition of limits of sequences, there exists $N > 0$ such that

$$a_n \in (L-1, L+1) \qquad \forall\, n \geqslant N.$$

Let $M = \max\big\{|a_1|, |a_2|, \cdots, |a_{N-1}|, |L|+1\big\}$. Then $|a_n| \leqslant M$ for all $n \in \mathbb{N}$. $\square$

Remark 9.19. A bounded sequence might not be convergent. For example, let $\{a_n\}_{n=1}^{\infty}$ be defined by $a_n = 3 + (-1)^n$. Then

$$a_1 = a_3 = a_5 = \cdots = a_{2k-1} = \cdots = 2 \quad \text{and} \quad a_2 = a_4 = a_6 = \cdots = a_{2k} = \cdots = 4\,.$$

Therefore, the only possible limits are $\{2, 4\}$; however, by the fact that

$$\#\big\{n \in \mathbb{N} \,\big|\, a_n \notin (1,3)\big\} = \#\big\{n \in \mathbb{N} \,\big|\, a_n \notin (3,5)\big\} = \infty\,,$$

we find that 2 and 4 are not the limit of $\{a_n\}_{n=1}^{\infty}$. Therefore, $\{a_n\}_{n=1}^{\infty}$ does not converge.

- **Completeness of Real Numbers**:

One important property of the real numbers is that they are ***complete***. The completeness axiom for real numbers states that "every bounded sequence of real numbers has a ***least upper bound*** and a ***greatest lower bound***"; that is, if $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence of real numbers, then there exists an upper bound M and a lower bound m of $\{a_n\}_{n=1}^{\infty}$ such that there is no smaller upper bound nor greater lower bound of $\{a_n\}_{n=1}^{\infty}$.

Theorem 9.20: Monotone Sequence Property (MSP)

Let $\{a_n\}_{n=1}^{\infty}$ be a monotone sequence of real numbers. Then $\{a_n\}_{n=1}^{\infty}$ converges if and only if $\{a_n\}_{n=1}^{\infty}$ is bounded.

Proof. It suffices to show the "$\Leftarrow$" direction.

Without loss of generality, we can assume that $\{a_n\}_{n=1}^{\infty}$ is increasing and bounded. By the completeness of real numbers, there exists a least upper bound M for the sequence $\{a_n\}_{n=1}^{\infty}$.

Let $\varepsilon > 0$ be given. Since M is the least upper bound for $\{a_n\}_{n=1}^{\infty}$, $M - \varepsilon$ is not an upper bound; thus there exists $N \in \mathbb{N}$ such that $a_N > M - \varepsilon$. Since $\{a_n\}_{n=1}^{\infty}$ is increasing, $a_n \geqslant a_N$ for all $n \geqslant N$. Therefore,

$$M - \varepsilon < a_n \leqslant M \qquad \forall\, n \geqslant N$$

which implies that

$$|a_n - M| < \varepsilon \qquad \forall\, n \geqslant N\,.$$

The statement above shows that $\{a_n\}_{n=1}^{\infty}$ converges to M. □

Remark 9.21. A sequence of real numbers $\{a_n\}_{n=1}^{\infty}$ is called a ***Cauchy sequence*** if for every $\varepsilon > 0$ there exists $N > 0$ such that

$$|a_n - a_m| < \varepsilon \qquad \text{whenever} \quad n, m \geqslant N\,.$$

A convergent sequence must be a Cauchy sequence. Moreover, the completeness of real numbers is equivalent to that "every Cauchy sequence of real number converges".

9.2 Series and Convergence

An infinite series is the "sum" of an infinite sequence. If $\{a_n\}_{n=1}^{\infty}$ is a sequence of real numbers, then

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \cdots + a_n + \cdots$$

is an infinite series (or simply series). The numbers a_1, a_2, a_3, $\cdots$ are called the terms of the series. For convenience, the sum could begin the index at $n = 0$ or some other integer.

Definition 9.22

The series $\sum\limits_{k=1}^{\infty} a_k$ is said to be convergent or converge to S if the sequence of the partial sum, denoted by $\{S_n\}_{n=1}^{\infty}$ and defined by

$$S_n \equiv \sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n,$$

converges to S. S_n is called the n-th partial sum of the series $\sum\limits_{k=1}^{\infty} a_k$.

When the series converges, we write $S = \sum\limits_{k=1}^{\infty} a_k$ and $\sum\limits_{k=1}^{\infty} a_k$ is said to be convergent. If $\{S_n\}_{n=1}^{\infty}$ diverges, the series is said to be divergent or diverge. If $\lim\limits_{n\to\infty} S_n = \infty$ (or $-\infty$), the series is said to diverge to ∞ (or $-\infty$).

Example 9.23. The n-th partial sum of the series $\sum\limits_{k=1}^{\infty} \dfrac{1}{k(k+1)}$ is

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \Big(\frac{1}{k} - \frac{1}{k+1}\Big) = \Big(1 - \frac{1}{2}\Big) + \Big(\frac{1}{2} - \frac{1}{3}\Big) + \cdots + \Big(\frac{1}{n} - \frac{1}{n+1}\Big) \\ &= 1 - \frac{1}{n+1}\,; \end{aligned}$$

thus the series $\sum\limits_{k=1}^{\infty} \dfrac{1}{k(k+1)}$ converges to 1, and we write $\sum\limits_{k=1}^{\infty} \dfrac{1}{k(k+1)} = 1$.

Example 9.24. The n-th partial sum of the series $\sum\limits_{k=1}^{\infty} \dfrac{2}{4k^2-1}$ is

$$\begin{aligned} \sum_{k=1}^{n} \frac{2}{4k^2-1} &= \sum_{k=1}^{n} \frac{2}{(2k-1)(2k+1)} = \sum_{k=1}^{n} \Big(\frac{1}{2k-1} - \frac{1}{2k+1}\Big) \\ &= \Big(1 - \frac{1}{3}\Big) + \Big(\frac{1}{3} - \frac{1}{5}\Big) + \cdots + \Big(\frac{1}{2n-1} - \frac{1}{2n+1}\Big) = 1 - \frac{1}{2n+1}\,; \end{aligned}$$

thus the series $\sum\limits_{k=1}^{\infty} \dfrac{2}{4k^2-1}$ converges to 1, and we write $\sum\limits_{k=1}^{\infty} \dfrac{2}{4k^2-1} = 1$.

The series in the previous two examples are series of the form

$$\sum_{k=1}^{n} (b_k - b_{k+1}) = (b_1 - b_2) + (b_2 - b_3) + \cdots + (b_n - b_{n+1}) + \cdots\,,$$

and are called telescoping series. A telescoping series converges if and only if $\lim\limits_{n\to\infty} b_n$ converges.

Example 9.25. The series $\sum\limits_{k=1}^{\infty} r^k$, where r is a real number, is called a geometric series (with ratio r). Note that the n-th partial sum of the series is

$$S_n = \sum_{k=1}^{n} r^k = 1 + r + r^2 + \cdots + r^n = \begin{cases} \dfrac{1 - r^{n+1}}{1 - r} & \text{if } r \neq 1\,, \\ n + 1 & \text{if } r = 1\,. \end{cases}$$

Therefore, the geometric series converges if and only if the common ratio r satisfies $|r| < 1$.

Theorem 9.26

Let $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty}$ be convergent series, and c is a real number. Then

1. $\sum_{k=1}^{\infty} ca_k = c\sum_{k=1}^{\infty} a_k$.

2. $\sum_{k=1}^{\infty} (a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k$.

3. $\sum_{k=1}^{\infty} (a_k - b_k) = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} b_k$.

Theorem 9.27: Cauchy Criteria

A series $\sum_{k=1}^{\infty} a_k$ converges if and only if for every $\varepsilon > 0$, there exists $N > 0$ such that

$$\Big|\sum_{k=n}^{n+\ell} a_k\Big| < \varepsilon \qquad \text{whenever} \quad n \geqslant N, \ell \geqslant 0\,.$$

Proof. Let S_n be the n-th partial sum of the series $\sum_{k=1}^{\infty} a_k$. Then by Remark 9.21,

$$\begin{aligned}
\sum_{k=1}^{\infty} a_k \text{ converges} &\Leftrightarrow \{S_n\}_{n=1}^{\infty} \text{ is a convergent sequence}\\
&\Leftrightarrow \{S_n\}_{n=1}^{\infty} \text{ is a Cauchy sequence}\\
&\Leftrightarrow \text{for every } \varepsilon > 0, \text{ there exists } N > 0 \text{ such that}\\
&\qquad\qquad |S_n - S_m| < \varepsilon \text{ whenever } n, m \geqslant N\\
&\Leftrightarrow \text{for every } \varepsilon > 0, \text{ there exists } N > 0 \text{ such that}\\
&\qquad\qquad |a_n + a_{n+1} + \cdots + a_{n+\ell}| < \varepsilon \text{ whenever } n \geqslant N \text{ and } \ell \geqslant 0.
\end{aligned}$$

□

Corollary 9.28: n-th Term Test

If the series $\sum_{k=1}^{\infty} a_k$ converges, then $\lim_{k\to\infty} a_k = 0$.

Remark 9.29. It is not true that $\lim_{n\to\infty} a_n = 0$ implies the convergence of $\sum_{k=1}^{\infty} a_k$. For example, we have shown in Example 8.50 that the harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges to ∞ while we know that $\lim_{n\to\infty} \frac{1}{n} = 0$.

Corollary 9.30: n-th term test for divergence

Let $\{a_n\}_{n=1}^{\infty}$ be a sequence. If $\lim\limits_{n\to\infty} a_n \neq 0$ or does not exist, then the series $\sum\limits_{k=1}^{\infty} a_k$ diverges.

9.3 The Integral Test and p-Series

9.3.1 The integral test

Suppose that the sequence $\{a_n\}_{n=1}^{\infty}$ is obtained by evaluating a non-negative continuous decreasing function $f : [1,\infty) \to \mathbb{R}$ on $\mathbb{N}$; that is, $f(n) = a_n$. Then

$$\int_1^{n+1} f(x)\,dx \leqslant S_n \equiv \sum_{k=1}^{n} a_k \leqslant a_1 + \int_1^{n} f(x)\,dx\,. \tag{9.3.1}$$

Since the sequence of partial sums $\{S_n\}_{n=1}^{\infty}$ of the series $\sum\limits_{k=1}^{\infty} a_k$ is increasing, the completeness of real numbers implies that $\{S_n\}_{n=1}^{\infty}$ converges if and only if the improper integral $\displaystyle\int_1^{\infty} f(x)\,dx$ converges.

Theorem 9.31

Let $f : [1,\infty) \to \mathbb{R}$ be a non-negative continuous decreasing function. The series $\sum\limits_{k=1}^{\infty} f(k)$ converges if and only if the improper integral $\displaystyle\int_1^{\infty} f(x)\,dx$ converges.

Example 9.32. The series $\sum\limits_{k=1}^{\infty} \dfrac{1}{k^2+1}$ converges since

$$\int_1^{\infty} \frac{dx}{x^2+1} = \lim_{b\to\infty} \int_1^{b} \frac{dx}{x^2+1} = \lim_{b\to\infty} \arctan x\Big|_{x=1}^{x=b} = \lim_{b\to\infty} (\arctan b - \arctan 1) = \frac{\pi}{4}$$

and the function $f(x) = \dfrac{1}{x^2+1}$ is non-negative continuous and decreasing on $[1,\infty)$.

Example 9.33. The series $\sum\limits_{k=1}^{\infty} \dfrac{k}{k^2+1}$ diverges since

$$\int_1^{\infty} \frac{x}{x^2+1}\,dx = \lim_{b\to\infty} \int_1^{b} \frac{x}{x^2+1}\,dx = \lim_{b\to\infty} \frac{\ln(x^2+1)}{2}\Big|_{x=1}^{x=b} = \frac{1}{2} \lim_{b\to\infty} \big[\ln(b^2+1) - \ln 2\big] = \infty$$

and the function $f(x) = \dfrac{x}{x^2+1}$ is non-negative continuous and decreasing on $[1,\infty)$.

Example 9.34. The series $\sum\limits_{k=2}^{\infty} \frac{1}{k \ln k}$ converges since

$$\begin{aligned}\int_2^\infty \frac{dx}{x\ln x} &= \lim_{b\to\infty} \int_2^b \frac{dx}{x\ln x} \overset{(x=e^u)}{=} \lim_{b\to\infty} \int_{\ln 2}^{\ln b} \frac{e^u du}{e^u \ln e^u} = \lim_{b\to\infty} \int_{\ln 2}^{\ln b} \frac{du}{u} = \lim_{b\to\infty} \ln u\Big|_{u=\ln 2}^{u=\ln b} \\ &= \lim_{b\to\infty} (\ln\ln b - \ln\ln 2) = \infty\end{aligned}$$

and the function $f(x) = \frac{1}{x\ln x}$ is non-negative continuous and decreasing on $[2,\infty)$.

9.3.2 p-series

A series of the form

$$\sum_{k=1}^{\infty} \frac{1}{k^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots$$

is called a p-series. The series is a function of p, and this function is usually called the ***Riemann zeta function***; that is,

$$\zeta(s) \equiv \sum_{n=1}^{\infty} \frac{1}{n^s}\,.$$

A harmonic series is the p-series with $p = 1$, and a general harmonic series is of the form

$$\sum_{k=1}^{\infty} \frac{1}{ak+b}\,.$$

By Theorem 8.51 and 9.31, the p-series converges if and only if $p > 1$.

Remark 9.35. It can be shown that $\sum\limits_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$. In fact, for all integer $k \geqslant 2$, the number $\sum\limits_{k=1}^{\infty} \frac{1}{n^k}$ can be computed by hand (even though it is very time consuming).

Remark 9.36. Using (9.3.1), we find that

$$\ln(n+1) \leqslant \sum_{k=1}^{n} \frac{1}{k} \leqslant 1 + \ln n \qquad \forall\, n \in \mathbb{N}\,.$$

Therefore, the sequence $\{a_n\}_{n=1}^{\infty}$ defined by

$$a_n = \sum_{k=1}^{n} \frac{1}{k} - \ln n$$

is bounded. Moreover,

$$a_n - a_{n+1} = \sum_{k=1}^{n} \frac{1}{k} - \ln n - \sum_{k=1}^{n+1} \frac{1}{k} + \ln(n+1) = \ln\Big(1 + \frac{1}{n}\Big) - \frac{1}{n+1}\,.$$

Since the derivative of the function $f(x) = \ln(1+x) - \dfrac{x}{x+1}$ is positive on $[0,1]$, we find that f is increasing on $[0,1]$; thus

$$\ln\Big(1 + \frac{1}{n}\Big) - \frac{1}{n+1} = f\Big(\frac{1}{n}\Big) \geqslant f(0) = \ln 1 - \frac{0}{1} = 0 \qquad \forall\, n \in \mathbb{N}$$

which shows that $a_n \geqslant a_{n+1}$. Therefore, $\{a_n\}_{n=1}^{\infty}$ is monotone decreasing and bounded from below (by 0). The completeness of real numbers then implies the convergence of the sequence $\{a_n\}_{n=1}^{\infty}$. The limit

$$\lim_{n\to\infty} \Big(\sum_{k=1}^{n} \frac{1}{k} - \ln n \Big)$$

is called Euler's constant. Euler's constant is approximated 0.5772.

9.3.3 Error estimates

Similar to (9.3.1), under the same setting we have

$$S_n + \int_{n+1}^{\infty} f(x)\,dx \leqslant S \leqslant S_n + \int_{n}^{\infty} f(x)\,dx \qquad \forall\, n \in \mathbb{N}\,. \tag{9.3.2}$$

The inequality above shows the following

Theorem 9.37: Bounds for the Remainder in the Integral Test

Let $f : [1,\infty) \to \mathbb{R}$ be a non-negative continuous decreasing function such that the series $S = \sum\limits_{k=1}^{\infty} f(k)$ converges. Then the remainder $R_n = S - S_n$, where $S_n = \sum\limits_{k=1}^{n} f(k)$, satisfies the inequality

$$\int_{n+1}^{\infty} f(x)\,dx \leqslant R_n \leqslant \int_{n}^{\infty} f(x)\,dx\,.$$

Example 9.38. Estimate the sum of the series $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^2}$ using the inequalities in (9.3.2) and $n = 10$.

Since

$$\int_{n}^{\infty} \frac{1}{x^2}\,dx = \lim_{b\to\infty} \frac{-1}{x}\Big|_{x=n}^{x=b} = \frac{1}{n}\,,$$

using (9.3.2) we find that

$$S_{10} + \frac{1}{11} \leqslant \sum_{k=1}^{\infty} \frac{1}{k^2} \leqslant S_{10} + \frac{1}{10}\,.$$

Computing S_{10}, we obtain that

$$S_{10} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots + \frac{1}{81} + \frac{1}{100} \approx 1.54977\,;$$

thus

$$1.64068 \leqslant \sum_{k=1}^{\infty} \frac{1}{k^2} \leqslant 1.64977\,.$$

9.4 Comparisons of Series

When the sequence $\{a_n\}_{n=1}^{\infty}$ is not obtained by $a_n = f(n)$ for some decreasing function $f : [1,\infty) \to \mathbb{R}$, the convergence of the series $\sum\limits_{k=1}^{\infty} a_k$ cannot be judged by the convergence of the improper integral $\displaystyle\int_1^{\infty} f(x)\,dx$. To determine the convergence of this kind of series, usually one uses comparison tests.

9.4.1 Direct Comparison Test

Theorem 9.39

Let $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ be sequences of real numbers, and $0 \leqslant a_n \leqslant b_n$ for all $n \in \mathbb{N}$.

1. If $\sum\limits_{k=1}^{\infty} b_k$ converges, then $\sum\limits_{k=1}^{\infty} a_k$ converges.
2. If $\sum\limits_{k=1}^{\infty} a_k$ diverges, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.

Proof. Let S_n and T_n be the n-th partial sum of the series $\sum\limits_{k=1}^{\infty} a_k$ and $\sum\limits_{k=1}^{\infty} b_k$, respectively; that is,

$$S_n = \sum_{k=1}^{n} a_k \qquad \text{and} \qquad T_n = \sum_{k=1}^{n} b_k\,.$$

Then by the assumption that $0 \leqslant a_n \leqslant b_n$ for all $n \in \mathbb{N}$, we find that $0 \leqslant S_n \leqslant T_n$ for all $n \in \mathbb{N}$, and $\{S_n\}_{n=1}^{\infty}$ and $\{T_n\}_{n=1}^{\infty}$ are monotone increasing sequences.

1. If $\sum\limits_{k=1}^{\infty} b_k$ converges, $\lim\limits_{n\to\infty} T_n = T$ exists; thus $0 \leqslant S_n \leqslant T_n \leqslant T$ for all $n \in \mathbb{N}$. Since $\{S_n\}_{n=1}^{\infty}$ is increasing, the monotone sequence property shows that $\lim\limits_{n\to\infty} S_n$ exists; thus $\sum\limits_{k=1}^{\infty} a_k$ converges.

2. If $\sum\limits_{k=1}^{\infty} a_k$ diverges, $\lim\limits_{n\to\infty} S_n = \infty$; thus by the fact that $S_n \leqslant T_n$ for all $n \in \mathbb{N}$, we find that $\lim\limits_{n\to\infty} T_n = \infty$. Therefore, $\sum\limits_{k=1}^{\infty} b_k$ diverges (to ∞). $\square$

Remark 9.40. It does not require that $0 \leqslant a_n \leqslant b_n$ for all $n \in \mathbb{N}$ for the direct comparison test to hold. The condition can be relaxed by that "$0 \leqslant a_n \leqslant b_n$ for all $n \geqslant N$" for some N since the sum of the first $N-1$ terms does not affect the convergence of the series.

Example 9.41. The series $\sum\limits_{k=1}^{\infty} \dfrac{1+\sin k}{k^2}$ converges since $\dfrac{1+\sin n}{n^2} \leqslant \dfrac{2}{n^2}$ for all $n \in \mathbb{N}$ and the p-series $\sum\limits_{k=1}^{\infty} \dfrac{2}{k^2}$ converges.

Example 9.42. The series $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+3^k}$ converges since $\dfrac{1}{2+3^n} \leqslant \dfrac{1}{3^n}$ for all $n \in \mathbb{N}$ and the geometric series $\sum\limits_{k=1}^{\infty} \dfrac{1}{3^k}$ converges.

Example 9.43. The series $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+\sqrt{k}}$ diverges since $\dfrac{1}{2+\sqrt{n}} \geqslant \dfrac{1}{3\sqrt{n}}$ for all $n \in \mathbb{N}$ and the p-series $\sum\limits_{k=1}^{\infty} \dfrac{1}{3\sqrt{k}} = \dfrac{1}{3}\sum\limits_{k=1}^{\infty} \dfrac{1}{\sqrt{k}}$ diverges.

One can also use the fact that $\dfrac{1}{2+\sqrt{n}} \geqslant \dfrac{1}{n}$ for all $n \geqslant 4$ and $\sum\limits_{k=1}^{\infty} \dfrac{1}{k}$ diverges to conclude that $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+\sqrt{k}}$ diverges.

9.4.2 Limit Comparison Test

Theorem 9.44

Let $\{a_n\}_{n=1}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ be sequences of real numbers, $a_n, b_n > 0$ for all $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L\,,$$

where L is a non-zero real number. Then $\sum\limits_{k=1}^{\infty} a_k$ converges if and only if $\sum\limits_{k=1}^{\infty} b_k$ converges.

Proof. We first note that if $L \neq 0$, then $L > 0$ since $\dfrac{a_n}{b_n} > 0$ for all $n \in \mathbb{N}$. By the fact that $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = L$, there exists $N > 0$ such that $\left|\dfrac{a_n}{b_n} - L\right| < \dfrac{L}{2}$ whenever $n \geqslant N$. In other words, $\dfrac{L}{2} < \dfrac{a_n}{b_n} < \dfrac{3L}{2}$ for all $n \geqslant N$; thus

$$0 < a_n < \frac{3L}{2} b_n \quad \text{and} \quad 0 < b_n < \frac{2}{L} a_n \qquad \text{whenever} \quad n \geqslant N\,.$$

By Theorem 9.39 and Remark 9.40, we find that $\sum\limits_{k=1}^{\infty} a_k$ converges if and only if $\sum\limits_{k=1}^{\infty} b_k$ converges. □

Remark 9.45. 1. If $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = 0$, then the convergence of $\sum\limits_{k=1}^{\infty} b_k$ implies the convergence of $\sum\limits_{k=1}^{\infty} a_k$, but not necessary the reverse direction.

2. The condition "$a_n, b_n > 0$ for all $n \in \mathbb{N}$" can be relaxed by "a_n and b_n are sign-definite for $n \geqslant N$, where a sequence $\{c_n\}_{n=1}^{\infty}$ is called sign-definite for $n \geqslant N$ if $c_n > 0$ for all $n \geqslant N$ or $c_n < 0$ for all $n \geqslant N$.

Example 9.46. Recall that in Example 9.42 and 9.43 we have shown that the series $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+3^k}$ converges and the series $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+\sqrt{k}}$ diverges using the direct comparison test. Note that since

$$\lim_{n\to\infty} \frac{\dfrac{1}{2+3^n}}{\dfrac{1}{3^n}} = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{\dfrac{1}{2+\sqrt{n}}}{\dfrac{1}{\sqrt{n}}} = 1\,,$$

using the convergence of the p-series and the limit comparison test we can also conclude that $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+3^k}$ converges and $\sum\limits_{k=1}^{\infty} \dfrac{1}{2+\sqrt{k}}$ diverges.

Example 9.47. The general harmonic series $\sum\limits_{k=1}^{\infty} \dfrac{1}{ak+b}$ diverges for the following reasons:

1. if $a = 0$, then clearly $\sum\limits_{k=1}^{\infty} \dfrac{1}{b}$ diverges.

2. if $a \neq 0$, then $\sum\limits_{k=1}^{\infty} \dfrac{1}{ak}$ diverges and $\lim\limits_{n\to\infty} \dfrac{\dfrac{1}{ak}}{\dfrac{1}{ak+b}} = 1$.

9.5 The Ratio and Root Tests

9.5.1 The Ratio Test

Theorem 9.48: Ratio Test

Let $\sum\limits_{k=1}^{\infty} a_k$ be a series with positive terms.

1. The series $\sum\limits_{k=1}^{\infty} a_k$ converges if $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} < 1$.

2. The series $\sum\limits_{k=1}^{\infty} a_k$ diverges (to ∞) if $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} > 1$.

Proof. Suppose that $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = L$ exists. Define $r = \frac{L+1}{2}$.

1. Assume that $L < 1$. Then for $\varepsilon = \frac{1-L}{2}$, there exists $N > 0$ such that
$$\Big|\frac{a_{n+1}}{a_n} - L\Big| < \frac{1-L}{2} \qquad \text{whenever } n \geqslant N\,;$$
thus
$$0 < \frac{a_{n+1}}{a_n} < r \qquad \text{whenever } n \geqslant N\,.$$
Note that $0 < r < 1$, and the inequality above implies that if $n \geqslant N$, $a_{n+1} < ra_n$. Therefore,
$$0 < a_n \leqslant a_N r^{n-N} \qquad \text{for all } n \geqslant N\,.$$
Now, since the series $\sum\limits_{k=1}^{\infty} a_N r^k$ converges, the comparison test implies that $\sum\limits_{k=1}^{\infty} a_k$ converges as well.

2. Assume that $L > 1$. Then for $\varepsilon = \frac{L-1}{2}$, there exists $N > 0$ such that
$$\Big|\frac{a_{n+1}}{a_n} - L\Big| < \frac{L-1}{2} \qquad \text{whenever } n \geqslant N\,;$$
thus
$$r < \frac{a_{n+1}}{a_n} \qquad \text{whenever } n \geqslant N\,.$$
Note that $r > 1$, and the inequality above implies that if $n \geqslant N$, $a_{n+1} > ra_n$. Therefore,
$$0 < a_N r^{n-N} \leqslant a_n \qquad \text{for all } n \geqslant N\,.$$

Now, since the series $\sum\limits_{k=1}^{\infty} a_N r^{k-N}$ diverges, the comparison test implies that $\sum\limits_{k=1}^{\infty} a_k$ diverges as well. $\square$

Remark 9.49. When $\lim\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n} = 1$, the convergence or divergence of $\sum\limits_{n=1}^{\infty} a_k$ cannot be concluded. For example, the p-series could converge or diverge depending on how large p is, but no matter what p is,

$$\lim_{n\to\infty} \frac{(n+1)^p}{n^p} = 1\,.$$

Example 9.50. The series $\sum\limits_{k=1}^{\infty} \dfrac{2^k}{k!}$ converges since

$$\lim_{n\to\infty} \frac{2^{n+1}/(n+1)!}{2^n/n!} = \lim_{n\to\infty} \frac{2}{n+1} = 0 < 1\,.$$

Example 9.51. The series $\sum\limits_{k=1}^{\infty} \dfrac{k^2 2^{k+1}}{3^k}$ converges since

$$\lim_{n\to\infty} \frac{(n+1)^2 2^{n+2}/3^{n+1}}{n^2 2^{n+1}/3^n} = \lim_{n\to\infty} \frac{2}{3}\frac{(n+1)^2}{n^2} = \frac{2}{3} < 1\,.$$

Example 9.52. The series $\sum\limits_{k=1}^{\infty} \dfrac{k^k}{k!}$ diverges since

$$\lim_{n\to\infty} \frac{(n+1)^{n+1}/(n+1)!}{n^n/n!} = \lim_{n\to\infty} \Big(1+\frac{1}{n}\Big)^n = e > 1\,.$$

9.5.2 The Root Test

Theorem 9.53: Root Test

Let $\sum\limits_{k=1}^{\infty} a_k$ be a series with positive terms.

1. The series $\sum\limits_{k=1}^{\infty} a_k$ converges if $\lim\limits_{n\to\infty} \sqrt[n]{a_n} < 1$.

2. The series $\sum\limits_{k=1}^{\infty} a_k$ diverges (to ∞) if $\lim\limits_{n\to\infty} \sqrt[n]{a_n} > 1$.

Proof. Suppose that $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = L$ exists. Define $r = \dfrac{L+1}{2}$.

1. Assume that $L < 1$. Then for $\varepsilon = \frac{1-L}{2}$, there exists $N > 0$ such that
$$\left|\sqrt[n]{a_n} - L\right| < \frac{1-L}{2} \qquad \text{whenever } n \geqslant N\,;$$
thus
$$0 < \sqrt[n]{a_n} < r \qquad \text{whenever } n \geqslant N$$
or equivalently,
$$0 < a_n \leqslant r^n \qquad \text{whenever } n \geqslant N\,.$$
By the fact that $0 < r < 1$, the series $\sum\limits_{k=1}^{\infty} r^k$ converges; thus the comparison test implies that $\sum\limits_{k=1}^{\infty} a_k$ converges as well.

2. Left as an exercise. □

Remark 9.54. When $\lim\limits_{n\to\infty} \sqrt[n]{a_n} = 1$, the convergence or divergence of $\sum\limits_{n=1}^{\infty} a_k$ cannot be concluded. For example, the p-series could converge or diverge depending on how large p is, but no matter what p is,
$$\lim_{n\to\infty} \sqrt[n]{n^p} = \Big(\lim_{n\to\infty} \sqrt[n]{n}\Big)^p = 1\,.$$

Example 9.55. The series $\sum\limits_{k=1}^{\infty} \frac{e^{2k}}{k^k}$ converges since
$$\lim_{n\to\infty} \Big(\frac{e^{2n}}{n^n}\Big)^{\frac{1}{n}} = \lim_{n\to\infty} \frac{e^2}{n} = 0 < 1\,.$$
We also note that the convergence of this series can be obtained through the ratio test:
$$\lim_{n\to\infty} \frac{e^{2(n+1)}/(n+1)^{n+1}}{e^{2n}/n^n} = \lim_{n\to\infty} \frac{e^2}{n+1}\Big(1+\frac{1}{n}\Big)^{-n} = 0 < 1\,.$$

Example 9.56. The series $\sum\limits_{k=1}^{\infty} \frac{k^2 2^{k+1}}{3^k}$ converges since
$$\lim_{n\to\infty} \Big(\frac{n^2 2^{n+1}}{3^n}\Big)^{\frac{1}{n}} = \lim_{n\to\infty} \frac{2(2n^2)^{\frac{1}{n}}}{3} = \frac{2}{3} < 1\,.$$

Example 9.57. The series $\sum\limits_{k=1}^{\infty} \frac{k^k}{k!}$ diverges since
$$\lim_{n\to\infty} \Big(\frac{n^n}{n!}\Big)^{\frac{1}{n}} = \lim_{n\to\infty} \Big(\frac{n^n}{\sqrt{2\pi n}\, n^n e^{-n}} \frac{\sqrt{2\pi n}\, n^n e^{-n}}{n!}\Big)^{\frac{1}{n}} = \lim_{n\to\infty} \Big(\frac{e^n}{\sqrt{2\pi n}}\Big)^{\frac{1}{n}} = e > 1\,,$$
here we have used Stirling's formula (9.1.2) to compute the limit.

Remark 9.58. Observe from Example 9.51, 9.52, 9.56 and 9.57, we see that as long as $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ and $\lim_{n\to\infty} \sqrt[n]{a_n}$ exists, then the limits are the same. This is in fact true in general, but we will not prove it since this is not our focus.

9.6 Absolute and Conditional Convergence

In the previous three sections we consider the convergence of series whose terms do not have different signs. How about the convergence of series like

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^p}, \qquad \sum_{k=1}^{\infty} \frac{\sin k}{k^p} \qquad \text{and etc.}$$

In the following two sections, we will focus on how to judge the convergence of a series that has both positive and negative terms.

Definition 9.59

An infinite series $\sum_{k=1}^{\infty} a_k$ is said to be absolutely convergent or converge absolutely if the series $\sum_{k=1}^{\infty} |a_k|$ converges. An infinite series $\sum_{k=1}^{\infty} a_k$ is said to be conditionally convergent or converge conditionally if $\sum_{k=1}^{\infty} a_k$ converges but $\sum_{k=1}^{\infty} |a_k|$ diverges (to ∞).

Example 9.60. The series $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^p}$ converge absolutely for $p > 1$ but does not converge absolutely for $p \leqslant 1$ since the p-series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ converges for $p > 1$ and diverges for $p \leqslant 1$.

Example 9.61. The series $\sum_{k=1}^{\infty} \frac{\sin k}{k^p}$ converges absolutely for $p > 1$ since

$$0 \leqslant \left|\frac{\sin n}{n^p}\right| \leqslant \frac{1}{n^p} \qquad \forall\, n \in \mathbb{N}$$

and the p-series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ converges for $p > 1$.

Theorem 9.62

An absolutely convergent series is convergent.（絕對收斂則收斂）

Proof. Let $\sum\limits_{k=1}^{\infty} a_k$ be an absolutely convergent series, and $\varepsilon > 0$ be given. Since $\sum\limits_{k=1}^{\infty} |a_k|$ converges, the Cauchy criteria implies that there exists $N > 0$ such that

$$\Big|\sum_{k=n}^{n+p} |a_k|\Big| < \varepsilon \qquad \text{whenever } n \geqslant N \text{ and } p \geqslant 0\,.$$

Therefore, if $n \geqslant N$ and $p \geqslant 0$,

$$\Big|\sum_{k=n}^{n+p} a_k\Big| \leqslant \sum_{k=n}^{n+p} |a_k| < \varepsilon$$

thus the Cauchy criteria implies that $\sum\limits_{k=1}^{\infty} a_k$ converges. □

Corollary 9.63: Ratio and Root Tests

The series $\sum\limits_{k=1}^{\infty} a_k$ converges if $\lim\limits_{n\to\infty} \dfrac{|a_{n+1}|}{|a_n|} < 1$ or $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} < 1$.

Example 9.64. The series $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^k 2^k}{k!}$ converges since

$$\lim_{n\to\infty} \frac{\left|\dfrac{(-1)^{n+1}2^{n+1}}{(n+1)!}\right|}{\left|\dfrac{(-1)^n 2^n}{n!}\right|} = \lim_{n\to\infty} \frac{\dfrac{2^{n+1}}{(n+1)!}}{\dfrac{2^n}{n!}} = \lim_{n\to\infty} \frac{2}{n+1} = 0 < 1$$

which shows the absolute convergence of the series the series $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^k 2^k}{k!}$.

Example 9.65. The series $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^{k+1} k!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2k+1)}$ converges since

$$\lim_{n\to\infty} \frac{\left|\dfrac{(-1)^{n+2}(n+1)!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2n+3)}\right|}{\left|\dfrac{(-1)^{n+1}n!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2n+1)}\right|} = \lim_{n\to\infty} \frac{\dfrac{(n+1)!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2n+3)}}{\dfrac{n!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2n+1)}} = \lim_{n\to\infty} \frac{n+1}{2n+3} = \frac{1}{2} < 1$$

which shows the absolute convergence of the series $\sum\limits_{k=1}^{\infty} \dfrac{(-1)^{k+1} k!}{1\cdot 3\cdot 5\cdot\cdots\cdot(2k+1)}$.

Example 9.66. Consider the series $\sum_{k=1}^{\infty} \frac{(k^2 \sin k)^k}{(k!)^k}$. Since

$$\lim_{n\to\infty} \left[\frac{n^{2n}}{(n!)^n}\right]^{\frac{1}{n}} = \lim_{n\to\infty} \frac{n^2}{n!} = \lim_{n\to\infty} \frac{n}{n-1}\frac{1}{(n-2)!} = 0 < 1\,,$$

the series $\sum_{k=1}^{\infty} \frac{k^{2k}}{(k!)^k}$ converges absolutely. By the fact that

$$\left|\frac{(n^2 \sin n)^n}{(n!)^n}\right| \leqslant \frac{(n^2)^n}{(n!)^n} \qquad \forall\, n \in \mathbb{N}\,,$$

the comparison test implies that the series $\sum_{k=1}^{\infty} \frac{(k^2 \sin k)^k}{(k!)^k}$ converges absolutely.

9.6.1 Alternating Series

In the previous two sections we consider the convergence of series whose terms do not have different signs. How about the convergence of series like

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\,, \quad \sum_{k=1}^{\infty} \frac{\sin k}{k} \quad \text{and etc.}$$

In the following two sections, we will focus on how to judge the convergence of a series that has both positive and negative terms.

Theorem 9.67: Dirichlet's Test

Let $\{a_n\}_{n=1}^{\infty}$, $\{p_n\}_{n=1}^{\infty}$ be sequences of real numbers such that

1. the sequence of partial sums of the series $\sum_{k=1}^{\infty} a_k$ is bounded; that is, there exists $M \in \mathbb{R}$ such that $\left|\sum_{k=1}^{n} a_k\right| \leqslant M$ for all $n \in \mathbb{N}$.
2. $\{p_n\}_{n=1}^{\infty}$ is a decreasing sequence, and $\lim_{n\to\infty} p_n = 0$.

Then $\sum_{k=1}^{\infty} a_k p_k$ converges.

Proof. Let $\varepsilon > 0$ be given. Since $\{p_n\}_{n=1}^{\infty}$ is decreasing and $\lim_{n\to\infty} p_n = 0$, there exists $N > 0$ such that

$$0 \leqslant p_n < \frac{\varepsilon}{2M+1} \qquad \text{whenever} \quad n \geqslant N\,.$$

Define $S_n = \sum_{k=1}^{n} a_k$. Then if $n \geqslant N$ and $\ell \geqslant 0$,

$$\begin{aligned}
\Big|\sum_{k=n}^{n+\ell} a_k p_k\Big| &= \big|(S_n - S_{n-1})p_n + (S_{n+1} - S_n)p_{n+1} + (S_{n+2} - S_{n+1})p_{n+2} + \cdots \\
&\quad + (S_{n+\ell-1} - S_{n+\ell-2})p_{n+\ell-1} + (S_{n+\ell} - S_{n+\ell-1})p_{n+\ell}\big| \\
&= \big|-S_{n-1}p_n + S_n(p_n - p_{n+1}) + S_{n+1}(p_{n+1} - p_{n+2}) + \cdots + S_{n+\ell-1}(p_{n+\ell-1} - p_{n+\ell}) \\
&\quad + S_{n+\ell}p_{n+\ell}\big| \\
&\leqslant |S_{n-1}p_n| + |S_n(p_n - p_{n+1})| + |S_{n+1}(p_{n+1} - p_{n+2})| + \cdots + |S_{n+\ell}(p_{n+\ell-1} - p_{n+\ell})| \\
&\quad + |S_{n+\ell+1}p_{n+\ell}| \\
&\leqslant Mp_n + M(p_n - p_{n+1}) + M(p_{n+1} - p_{n+2}) + \cdots + M(p_{n+\ell-1} - p_{n+\ell}) + Mp_{n+\ell} \\
&= 2Mp_n < \frac{2M\varepsilon}{2M+1} < \varepsilon\,.
\end{aligned}$$

The convergence of $\sum_{k=1}^{\infty} a_k p_k$ then follows from the Cauchy criteria (Theorem 9.27). $\square$

Corollary 9.68

Let $\{p_n\}_{n=1}^{\infty}$ be a decreasing sequence of real numbers. If $\lim_{n\to\infty} p_n = 0$, then $\sum_{k=1}^{\infty} (-1)^k p_k$ and $\sum_{k=1}^{\infty} (-1)^{k+1} p_k$ converge.

Example 9.69. The series $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^p}$ converges conditionally for $0 < p \leqslant 1$ since

1. $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^p}$ converges due the fact that

$$\Big|\sum_{k=1}^{n} (-1)^{k+1}\Big| \leqslant 1 \quad \text{and} \quad \Big\{\frac{1}{n^p}\Big\}_{n=1}^{\infty} \text{ is decreasing and converges to } 0\,.$$

2. $\sum_{k=1}^{\infty} \Big|\frac{(-1)^{k+1}}{k^p}\Big|$ diverges for it is a p-series with $0 < p \leqslant 1$.

Similarly, $\sum_{k=1}^{\infty} \frac{(-1)^k}{\ln(k+1)}$ converges conditionally.

Example 9.70. The series $\sum_{k=1}^{\infty} \frac{\sin k}{k^p}$ converges for $p > 0$ since

1. $\sum\limits_{k=1}^{n} \sin k = \dfrac{\cos\frac{1}{2} - \cos\frac{2k+1}{2}}{2\sin\frac{1}{2}}$; $\Big($thus $\Big|\sum\limits_{k=1}^{n} \sin k\Big| \leqslant \dfrac{1}{\sin\frac{1}{2}}\Big)$.

2. $\big\{\dfrac{1}{n^p}\big\}_{n=1}^{\infty}$ is decreasing and $\lim\limits_{n\to\infty} \dfrac{1}{n^p} = 0$.

We remark here that $\sum\limits_{k=1}^{\infty} \dfrac{\sin k}{k} = \dfrac{\pi - 1}{2}$. In fact, $\sum\limits_{k=1}^{\infty} \dfrac{\sin(kx)}{k}$ is the Fourier series of the function $\dfrac{\pi - x}{2}$.

- **Alternating Series Remainder**

Theorem 9.71

Let $\{a_n\}_{n=1}^{\infty}$, $\{p_n\}_{n=1}^{\infty}$ be sequences of real numbers satisfying conditions in Theorem 9.67. If $\Big|\sum\limits_{k=1}^{n} a_k\Big| \leqslant M$ for all $n \in \mathbb{N}$, then

$$\Big|\sum_{k=1}^{\infty} a_k p_k - \sum_{k=1}^{n} a_k p_k\Big| = \Big|\sum_{k=n+1}^{\infty} a_k p_k\Big| \leqslant 2Mp_{n+1}\,.$$

Moreover, if $a_k = (-1)^k$, then

$$\Big|\sum_{k=1}^{\infty}(-1)^{k+1}p_k - \sum_{k=1}^{n}(-1)^{k+1}p_k\Big| \leqslant p_{n+1} \qquad \forall\, n \in \mathbb{N}\,.$$

Sketch of Proof. Let $S_n = \sum\limits_{k=1}^{n} a_k$. According to the proof of the Abel test, we have

$$\begin{aligned}\Big|\sum_{k=n}^{n+\ell} a_k p_k\Big| \leqslant |S_{n-1}|p_n + |S_n|(p_n - p_{n+1}) + |S_{n+1}|(p_{n+1} - p_{n+2}) + \cdots + |S_{n+\ell}|(p_{n+\ell-1} - p_{n+\ell}) \\ + |S_{n+\ell+1}|p_{n+\ell}\,. \end{aligned} \tag{9.6.1}$$

Note that for the general case, by the fact that $|S_n| \leqslant M$ for all $n \in \mathbb{N}$ and $\{p_n\}_{n=1}^{\infty}$ is decreasing, we conclude that for all $\ell \geqslant 0$,

$$\Big|\sum_{k=n}^{n+\ell} a_k p_k\Big| \leqslant 2Mp_n \qquad \forall\, n \in \mathbb{N}\,;$$

thus if $n \in \mathbb{N}$,

$$\Big|\sum_{k=1}^{\infty} a_k p_k - \sum_{k=1}^{n} a_k p_k\Big| = \lim_{\ell\to\infty}\Big|\sum_{k=1}^{n+1+\ell} a_k p_k - \sum_{k=1}^{n} a_k p_k\Big| = \lim_{\ell\to\infty}\Big|\sum_{k=n+1}^{n+1+\ell} a_k p_k\Big| \leqslant 2Mp_{n+1}\,.$$

For the case of alternating series, we note that terms of $\{S_n\}_{n=1}^{\infty}$ are $\{1,0,1,0,1,\cdots\}$; thus (9.6.1) implies that

$$\Big|\sum_{k=1}^{\infty}(-1)^{k+1}p_k - \sum_{k=1}^{n}(-1)^{k+1}p_k\Big| \leqslant p_{n+1} \qquad \forall\, n \in \mathbb{N}\,. \qquad \square$$

Example 9.72. Approximate the sum of the series $\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k!}$ by its first six terms, we obtain that

$$\sum_{k=1}^{6}(-1)^{k+1}\frac{1}{k!} = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} - \frac{1}{6!} \approx 0.63194\,.$$

Moreover, by Theorem 9.71, we find that

$$\Big|\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k!} - \sum_{k=1}^{6}(-1)^{k+1}\frac{1}{k!}\Big| \leqslant \frac{1}{7!} = \frac{1}{5040} \approx 0.0002\,.$$

Example 9.73. Determine the number of terms required to approximate the sum of the series $\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^4}$ with an error of less than 0.0001.

By Theorem 9.71,

$$\Big|\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^4} - \sum_{k=1}^{n}\frac{(-1)^{k+1}}{k^4}\Big| \leqslant \frac{1}{(n+1)^4}\,;$$

thus choosing n such that $\frac{1}{(n+1)^4} \leqslant 0.0001$ (that is, $n \geqslant 9$), we obtain that

$$\Big|\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k^4} - \sum_{k=1}^{n}\frac{(-1)^{k+1}}{k^4}\Big| \leqslant 0.001 \qquad \forall\, n \geqslant 9\,.$$

9.7 Taylor Polynomials and Approximations

Suppose that $f : (a,b) \to \mathbb{R}$ is $(n+1)$-times continuously differentiable; that is, $\frac{d^k f}{dx^k}$ is continuous on (a,b) for $1 \leqslant k \leqslant n+1$, then for $x \in (a,b)$, the Fundamental Theorem of

Calculus and integration-by-parts imply that

$$\begin{aligned}
f(x) - f(c) = \int_c^x f'(t)\,dt &= f'(t)(t-x)\Big|_{t=c}^{t=x} - \int_c^x f''(t)(t-x)\,dt \\
&= -f'(c)(c-x) - \int_c^x f''(t)(t-x)\,dt \\
&= f'(c)(x-c) - \Big[f''(t)\frac{(t-x)^2}{2}\Big|_{t=c}^{t=x} - \int_c^x f'''(t)\frac{(t-x)^2}{2}\,dt \Big] \\
&= f'(c)(x-c) - \Big[-\frac{f''(c)}{2}(c-x)^2 - \int_c^x f'''(t)\frac{(t-x)^2}{2}\,dt \Big] \\
&= f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \int_c^x f'''(t)\frac{(t-x)^2}{2}\,dt \\
&= \cdots\cdots \\
&= f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n \\
&\quad + (-1)^n \int_c^x f^{(n+1)}(t)\frac{(t-x)^n}{n!}\,dt\,,
\end{aligned}$$

where the last equality can be shown by induction. Therefore,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + (-1)^n \int_c^x f^{(n+1)}(t)\frac{(t-x)^n}{n!}\,dt\,. \tag{9.7.1}$$

Definition 9.74

If f has n derivatives at c, then the polynomial

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

is called the n-th (order) Taylor polynomial for f at c. The n-th Taylor polynomial for f at 0 is also called the n-th (order) Maclaurin polynomial for f.

Example 9.75. The n-th Maclaurin polynomial for the function $f(x) = e^x$ is

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k = \sum_{k=0}^{n} \frac{1}{k!}x^k = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}\,.$$

Example 9.76. The n-th Maclaurin polynomial for the function $f(x) = \ln(1+x)$ is given by

$$\begin{aligned}
P_n(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=1}^{n} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=1}^{n} \frac{(-1)^{k-1}(k-1)!}{k!} x^k = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k \\
&= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + \frac{(-1)^{n-1}}{n} x^n \,,
\end{aligned}$$

here we have used $g^{(k)}(x) = (-1)^{k-1}(k-1)!(x+1)^{-k}$ to compute $g^{(k)}(0)$.

The n-th Taylor polynomial for the function $g(x) = \ln x$ at 1 is given by

$$\begin{aligned}
Q_n(x) &= \sum_{k=0}^{n} \frac{g^{(k)}(1)}{k!} (x-1)^k = \sum_{k=1}^{n} \frac{g^{(k)}(1)}{k!} (x-1)^k = \sum_{k=1}^{n} \frac{(-1)^{k-1}(k-1)!}{k!} (x-1)^k \\
&= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} (x-1)^k \\
&= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots + \frac{(-1)^{n-1}}{n} (x-1)^n \,,
\end{aligned}$$

here we have used $g^{(k)}(x) = (-1)^{k-1}(k-1)!x^{-k}$ to compute $g^{(k)}(1)$. We note that $Q_n(x) = P_n(x-1)$ (and $g(x) = f(x-1)$).

Example 9.77. The $(2n)$-th Maclaurin polynomial for the function $f(x) = \cos x$ is given by

$$\begin{aligned}
P_{2n}(x) &= \sum_{k=0}^{2n} \frac{f^{(k)}(0)}{k!} x^k = 1 + \sum_{k=1}^{2n} \frac{f^{(k)}(0)}{k!} x^k = 1 + \sum_{k=1}^{n} \frac{f^{(2k-1)}(0)}{(2k-1)!} x^{2k-1} + \sum_{k=1}^{n} \frac{f^{(2k)}(0)}{(2k)!} x^{2k} \\
&= 1 + \sum_{k=1}^{n} \frac{f^{(2k)}(0)}{(2k)!} x^{2k} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + \frac{(-1)^n}{(2n)!} x^{2n} \,,
\end{aligned}$$

here we have used $f^{(k)}(x) = \cos\left(x + \frac{k\pi}{2}\right)$ to compute $f^{(k)}(0)$. We also note that $P_{2n}(x) = P_{2n+1}(x)$ for all $n \in \mathbb{N}$.

The $(2n-1)$-th Maclaurin polynomial for the function $g(x) = \sin x$ is given by

$$\begin{aligned}
Q_{2n-1}(x) &= \sum_{k=0}^{2n-1} \frac{g^{(k)}(0)}{k!} x^k = \sum_{k=1}^{2n-1} \frac{g^{(k)}(0)}{k!} x^k = \sum_{k=1}^{n} \frac{g^{(2k-1)}(0)}{(2k-1)!} x^{2k-1} + \sum_{k=1}^{n} \frac{g^{(2k)}(0)}{(2k)!} x^{2k} \\
&= \sum_{k=1}^{n} \frac{g^{(2k-1)}(0)}{(2k-1)!} x^{2k-1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{(-1)^{n-1}}{(2n-1)!} x^{2n-1} \,,
\end{aligned}$$

here we have used $g^{(k)}(x) = \sin\left(x + \frac{k\pi}{2}\right)$ to compute $g^{(k)}(0)$. We also note that $Q_{2n-1}(x) = Q_{2n}(x)$ for all $n \in \mathbb{N}$.

9.7.1 Remainder of Taylor Polynomials

To measure the accuracy of approximating a function value $f(x)$ by the Taylor polynomial, we look for the difference $R_n(x) \equiv f(x) - P_n(x)$, where P_n is the n-th Taylor polynomial for f (centered at a certain number c). The function R_n is called the remainder associated with the approximation P_n.

- **Integral form of the remainder**

By (9.7.1), we find that if P_n is the n-th Taylor polynomial for f at c, then

$$R_n(x) = (-1)^n \int_c^x f^{(n+1)}(t)\frac{(t-x)^n}{n!}\,dt\,. \tag{9.7.2}$$

Example 9.78. Consider the function $f(x) = \exp(x) = e^x$. If P_n is the n-th Maclaurin polynomial for f, the remainder R_n associated with P_n is given by

$$R_n(x) = (-1)^n \int_0^x f^{(n+1)}(t)\frac{(t-x)^n}{n!}\,dt = (-1)^n \int_0^x e^t \frac{(t-x)^n}{n!}\,dt\,.$$

Therefore, if $x > 0$,

$$\Big|e^x - \sum_{k=0}^n \frac{x^k}{k!}\Big| = \Big|\int_0^x e^t \frac{(t-x)^n}{n!}\,dt\Big| \leqslant \int_0^x e^t \frac{(x-t)^n}{n!}\,dt \leqslant \int_0^x e^x \frac{x^n}{n!}\,dt = \frac{e^x x^{n+1}}{n!}\,. \tag{9.7.3}$$

Note that for each $x > 0$, the series $\sum\limits_{k=0}^{\infty} e^x \dfrac{x^{n+1}}{n!}$ converges since

$$\lim_{n\to\infty} \frac{e^x \dfrac{x^{(n+1)+1}}{(n+1)!}}{e^x \dfrac{x^{n+1}}{n!}} = \lim_{n\to\infty} \frac{x}{n+1} = 0\,;$$

thus the n-th term test shows that $\lim\limits_{n\to\infty} e^x \dfrac{x^{n+1}}{n!} = 0$. Therefore, for each $x > 0$,

$$\lim_{n\to\infty} \Big|e^x - \sum_{k=0}^n \frac{x^k}{k!}\Big| = 0$$

or equivalently,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots\,.$$

In particular, if $x = 1$, (9.7.3) implies that

$$\Big|e - \sum_{k=0}^{n} \frac{1}{k!}\Big| \leqslant \frac{e}{n!}\,;$$

thus $\Big|e - \sum\limits_{k=0}^{17} \frac{1}{k!}\Big| < 10^{-8}$.

Example 9.79. Consider the function $f(x) = \cos x$ and its $(2n)$-th Maclaurin polynomial P_{2n} in Example 9.77. If $x > 0$,

$$\begin{aligned}\big|f(x) - P_{2n}(x)\big| = \big|f(x) - P_{2n+1}(x)\big| &\leqslant \Big|\int_0^x f^{(2n+2)}(t)\frac{(t-x)^{2n+1}}{(2n+1)!}\,dt\Big| \leqslant \int_0^x \frac{(x-t)^{2n+1}}{(2n+1)!}\,dt \\ &= \frac{-(x-t)^{2n+2}}{(2n+2)!}\Big|_{t=0}^{t=x} = \frac{x^{2n+2}}{(2n+2)!}\,,\end{aligned}$$

while if $x < 0$,

$$\begin{aligned}\big|f(x) - P_{2n}(x)\big| = \big|f(x) - P_{2n+1}(x)\big| &\leqslant \Big|\int_0^x f^{(2n+2)}(t)\frac{(t-x)^{2n+1}}{(2n+1)!}\,dt\Big| \leqslant \int_x^0 \frac{(t-x)^{2n+1}}{(2n+1)!}\,dt \\ &= \frac{(t-x)^{2n+2}}{(2n+2)!}\Big|_{t=0}^{t=x} = \frac{(-x)^{2n+2}}{(2n+2)!}\,.\end{aligned}$$

Therefore,

$$\Big|\cos x - \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!}x^{2k}\Big| \leqslant \frac{|x|^{2n+2}}{(2n+2)!} \qquad \forall\, x \in \mathbb{R}\,. \tag{9.7.4}$$

Similarly,

$$\Big|\sin x - \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!}x^{2k+1}\Big| \leqslant \frac{|x|^{2n+3}}{(2n+3)!} \qquad \forall\, x \in \mathbb{R}\,. \tag{9.7.5}$$

Moreover, by the fact that

$$\lim_{n\to\infty} \frac{\dfrac{|x|^{2(n+1)+2}}{[2(n+1)+2]!}}{\dfrac{|x|^{2n+2}}{(2n+2)!}} = \lim_{n\to\infty} \frac{x^2}{(2n+3)(2n+4)} = 0 < 1$$

and

$$\lim_{n\to\infty} \frac{\dfrac{|x|^{2(n+1)+3}}{[2(n+1)+3]!}}{\dfrac{|x|^{2n+3}}{(2n+3)!}} = \lim_{n\to\infty} \frac{x^2}{(2n+4)(2n+5)} = 0 < 1$$

the ratio test implies that $\sum\limits_{k=0}^{\infty} \frac{|x|^{2n+2}}{(2n+2)!}$ and $\sum\limits_{k=0}^{\infty} \frac{|x|^{2n+3}}{(2n+3)!}$ converge; thus for each $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{|x|^{2n+2}}{(2n+2)!} = \lim_{n\to\infty} \frac{|x|^{2n+3}}{(2n+3)!} = 0\,;$$

thus

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{(-1)^n}{(2n)!} x^{2n} + \cdots\,,$$

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^n}{(2n+1)!} x^{2n+1} + \cdots\,.$$

Using (9.7.4), we conclude that

$$\Big|\cos(0.1) - \sum_{k=0}^{3} \frac{(-1)^k}{(2k)!} (0.1)^{2k}\Big| \leqslant \frac{0.1^8}{8!}\,;$$

thus $\cos(0.1) \approx \sum\limits_{k=0}^{3} \frac{(-1)^k}{(2k)!} (0.1)^{2k} \approx 0.995004165$ which is accurate to nine decimal points.

Remark 9.80. By Example 9.78 and 9.79, conceptually we can explain why the Euler identity $e^{i\theta} = \cos\theta + i\sin\theta$ for all $\theta \in \mathbb{R}$. Recall that the $(2n)$-th Maclaurin polynomial for exp, cos, sin are

$$P^e_{2n}(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!}\,,$$

$$P^c_{2n}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{(-1)^n}{(2n)!} x^{2n}\,,$$

$$P^s_{2n}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^{n-1}}{(2n-1)!} x^{2n-1}\,.$$

Substitution $x = i\theta$, we find that

$$P^e_{2n}(i\theta) = P^c_{2n}(\theta) + iP^s_{2n}(\theta) \qquad \forall\, \theta \in \mathbb{R}\,.$$

Passing $n \to \infty$, by the fact that the remainders $R_n(x)$ for exp, sin and cos all converges to zero as $n \to \infty$ for each $x \in \mathbb{R}$ (and even $x \in \mathbb{C}$), we conclude that

$$e^{i\theta} = \cos\theta + i\sin\theta \qquad \forall\, \theta \in \mathbb{R}\,.$$

- **Lagrange form of the remainder**

Theorem 9.81: Taylor's Theorem

Let $f : (a, b) \to \mathbb{R}$ be $(n+1)$-times differentiable, and $c \in (a, b)$. Then for each $x \in (a, b)$, there exists ξ between x and c such that

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n + R_n(x)\,, \quad (9.7.6)$$

where Lagrange form of the remainder $R_n(x)$ is given by

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}\,.$$

Proof. We first show that if $h : (a, b) \to \mathbb{R}$ is m-times differentiable, and $c \in (a, b)$. Then for all $d \in (a, b)$ and $d \neq c$ there exists ξ between c and d such that

$$\frac{h(d) - \sum\limits_{k=0}^{m} \frac{h^{(k)}(c)}{k!}(d-c)^k}{(d-c)^{m+1}} = \frac{1}{m+1} \frac{h'(\xi) - \sum\limits_{k=0}^{m-1} \frac{(h')^{(k)}(c)}{k!}(\xi-c)^k}{(\xi-c)^m}\,. \qquad (9.7.7)$$

Let $F(x) = h(x) - \sum\limits_{k=0}^{m} \frac{h^{(k)}(c)}{k!}(x-c)^k$ and $G(x) = (x-c)^m$. Then F, G are continuous on $[c, d]$ (or $[d, c]$) and differentiable on (c, d) (or (d, c)), and $G'(x) \neq 0$ for all $x \neq c$. Therefore, the Cauchy Mean Value Theorem implies that there exists ξ between c and d such that

$$\frac{F(d) - F(c)}{G(d) - G(c)} = \frac{F'(\xi)}{G'(\xi)}\,,$$

and (9.7.7) is exactly the explicit form of the equality above.

Now we apply (9.7.7) successfully for $h = f, f', f'', \cdots$ and $f^{(n)}$ and find that

$$\begin{aligned}
\frac{f(d) - \sum\limits_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(d-c)^k}{(d-c)^{n+1}} &= \frac{1}{n+1} \frac{f'(d_1) - \sum\limits_{k=0}^{n-1} \frac{(f')^{(k)}(c)}{k!}(d_1-c)^k}{(d_1-c)^n} \\
&= \frac{1}{n+1} \cdot \frac{1}{n} \frac{f''(d_2) - \sum\limits_{k=0}^{n-2} \frac{(f'')^{(k)}(c)}{k!}(d_2-c)^k}{(d_2-c)^{n-1}} \\
&= \cdots\cdots \\
&= \frac{1}{(n+1)!} \frac{f^{(n)}(d_n) - f^{(n)}(c)}{d_n - c} = \frac{1}{(n+1)!} f^{(n+1)}(\xi)\,;
\end{aligned}$$

thus

$$f(d)-\sum_{k=0}^{n}\frac{f^{(k)}(c)}{k!}(d-c)^k=\frac{1}{(n+1)!}f^{(n+1)}(\xi)(d-c)^{n+1}\,.$$

(9.7.6) then follows from the equality above since $d \in (a,b)$ is given arbitrary. □

Example 9.82. In Example 9.76 we compute the Taylor polynomial Q_n for the function $y = \ln(1+x)$. Note that the Taylor Theorem implies that

$$\ln(1+x) = P_n(x) + R_n(x)\,,$$

where

$$R_n(x)=\frac{1}{(n+1)!}\Big(\frac{d^{n+1}}{dx^{n+1}}\Big|_{x=\xi}\ln(1+x)\Big)x^{n+1}=\frac{(-1)^n}{n+1}(1+\xi)^{-n-1}x^{n+1}$$

for some ξ between 0 and x.

1. If $-1<x<0$, then $R_n(x)=\dfrac{-1}{n+1}\Big(\dfrac{-x}{1+\xi}\Big)^{n+1}<0$; thus

$$\ln(1+x)\leqslant x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots+\frac{(-1)^n}{n}x^n \qquad \forall\, x\in(-1,0) \text{ and } n\in\mathbb{N}\,.$$

2. If $x > 0$, then

   (a) $R_n(x) < 0$ if n is odd; thus

$$\ln(1+x)\leqslant x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots+\frac{1}{2k+1}x^{2k+1} \qquad \forall\, x>0 \text{ and } k\in\mathbb{N}\,.$$

   (b) $R_n(x) > 0$ if n is even; thus

$$\ln(1+x)\geqslant x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\cdots+\frac{-1}{2k}x^{2k} \qquad \forall\, x>0 \text{ and } k\in\mathbb{N}\,.$$

Example 9.83. In this example we show that

$$\ln(1+x)=\sum_{k=1}^{\infty}\frac{(-1)^{k-1}x^k}{k}=x-\frac{x^2}{2}+\frac{x^3}{3}+\cdots+\frac{(-1)^{n-1}x^n}{n}+\cdots \qquad \forall\, x\in(0,1]\,. \tag{9.7.8}$$

Note that Taylor's Theorem implies that for all $x > -1$, there exists ξ between 0 and x such that the remainder associated with $P_n(x)=\sum\limits_{k=1}^{n}\dfrac{(-1)^{k-1}x^k}{k}$ is given by

$$R_n(x)=\frac{(-1)^n}{n+1}(1+\xi)^{-n-1}x^{n+1}\,.$$

Note that since ξ is between 0 and x, we always have

$$0 < \frac{x}{1+\xi} < 1 \qquad \forall\, x \in (0,1]\,;$$

thus $|R_n(x)| \leqslant \dfrac{1}{n+1}$ for all $x \in (-1,1]$ and (9.7.8) is concluded because

$$\lim_{n\to\infty} \big|R_n(x)\big| = 0\,.$$

Example 9.84. In this example we compute $\ln 2$. Note that using (9.7.8) we find that

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n-1}}{n} + R_n(1)\,,$$

where

$$R_n(1) = \frac{1}{(n+1)!}\Big(\frac{d^{n+1}}{dx^{n+1}}\Big|_{x=\xi} \ln(1+x)\Big) 1^{n+1} = \frac{(-1)^n}{n+1}(1+\xi)^{-(n+1)}$$

for some ξ between 0 and 1. Since ξ could be very closed to 0, in this case the best we can estimate $R_n(1)$ is

$$\big|R_n(1)\big| \leqslant \frac{1}{n+1}\,.$$

Therefore, to evaluate $\ln 2$ accurate to eight decimal point, it is required that $n = 10^8$.

Let $c = \dfrac{e}{2} \approx 1.359140914$. Then

$$\ln c = \ln\big(1 + (c-1)\big) = (c-1) - \frac{(c-1)^2}{2} + \cdots + \frac{(-1)^{n-1}}{n}(c-1)^n + R_n(c-1)\,,$$

where $R_n(c-1)$ is given by

$$R_n(c-1) = \frac{1}{(n+1)!}\Big(\frac{d^{n+1}}{dx^{n+1}}\Big|_{x=\xi} \ln(1+x)\Big)(c-1)^{n+1} = \frac{(-1)^n}{n+1}(1+\xi)^{-(n+1)}(c-1)^{n+1}$$

for some ξ between 0 and $c-1$. Note that

$$\big|R_n(c)\big| \leqslant \frac{(c-1)^{n+1}}{n+1}\,;$$

thus the value

$$(c-1) - \frac{(c-1)^2}{2} + \frac{(c-1)^3}{3} - \frac{(c-1)^4}{4} + \cdots + \frac{1}{17}(c-1)^{17}$$

to approximate $\ln c$ is accurate to eight decimal points $\big($since $\frac{1}{18}0.4^{18} < 10^{-8}\big)$. On the other hand, we have $\ln 2 = 1 - \ln c$, so the value

$$1 - (c-1) + \frac{(c-1)^2}{2} - \frac{(c-1)^3}{3} + \frac{(c-1)^4}{4} + \cdots - \frac{1}{17}(c-1)^{17}$$

to approximate $\ln 2$ is also accurate to eight decimal points.

9.8 Power Series

Recall that for all $x \in \mathbb{R}$, we have shown that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots,$$
$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{(-1)^n}{(2n)!} x^{2n} + \cdots,$$
$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{(-1)^n}{(2n+1)!} x^{2n+1} + \cdots.$$

The identities above show that the functions $y = \exp(x)$, $y = \cos x$, $y = \sin x$ can be defined using series whose terms are multiples of monomials of x. These kind of series are called power series. To be more precise, we have the following

Definition 9.85: Power Series

Let c be a real number. A power series (of one variable x) centered at c is an infinite series of the form

$$\sum_{k=0}^{\infty} a_k (x-c)^k = a_0 + a_1(x-c)^1 + a_2(x-c)^2 + \cdots,$$

where a_k is independent of x and represents the coefficient of the k-th term.

Theorem 9.86

Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of real numbers. If $\sum\limits_{k=0}^{\infty} a_k d^k$ converges, then $\sum\limits_{k=0}^{\infty} a_k (x-c)^k$ converges absolutely for all $x \in (c - |d|, c + |d|)$.

Proof. First we note that since $\sum\limits_{k=0}^{\infty} a_k d^k$ converges, $\lim\limits_{n\to\infty} a_n d^n = 0$; thus the boundedness of convergent sequence implies that there exists $M > 0$ such that

$$|a_n d^n| \leqslant M \qquad \forall\, n \in \mathbb{N}.$$

Suppose that $|x - c| < |d|$. Then there exists $\varepsilon > 0$ such that $|x - c| < |d| - \varepsilon$. Then

$$|a_n||x-c|^n = |a_n||d|^n \frac{|x-c|^n}{(|d|-\varepsilon)^n} \Big(\frac{|d|-\varepsilon}{|d|}\Big)^n \leqslant M\Big(\frac{|d|-\varepsilon}{|d|}\Big)^n.$$

Therefore, by the convergence of geometric series with ratio between -1 and 1, the direct comparison test implies that the series $\sum_{n=0}^{\infty} a_n(x-c)^n$ converges absolutely. $\square$

Corollary 9.87

For a power series centered at c, precisely one of the following is true.

1. The series converges only at c.
2. There exists $R > 0$ such that the series converges absolutely for $|x-c| < R$ and diverges for $|x-c| > R$.
3. The series converges absolutely for all x.

Definition 9.88: Radius of Convergence and Interval of Convergence

Let a power series centered at c be given. If the power series converges only at c, we say that the radius of convergence of the power series is 0. If the power series converges for $|x-c| < R$ but diverges for $|x-c| > R$, we say that the radius of convergence of the power series is R. If the power series converges for all x, we say that the radius of converges of the power series is ∞. The set of all values of x for which the power series converges is called the interval of convergence of the power series.

Remark 9.89. The radius of convergence of a power series centered at c is the greatest lower bound of the set

$$\left\{r > 0 \,\middle|\, \text{there exists } x \in (c-r, c+r) \text{ such that the power series diverges}\right\}.$$

Example 9.90. Consider the power series $\sum_{k=0}^{\infty} k!x^k$. Note that for each $x \neq 0$,

$$\lim_{k\to\infty} \frac{\left|(k+1)!x^{k+1}\right|}{\left|k!x^k\right|} = \lim_{k\to\infty} (k+1)|x| = \infty\,;$$

thus the ratio test implies that the power series $\sum_{k=0}^{\infty} k!x^k$ diverges for all $x \neq 0$. Therefore, the radius of convergence of $\sum_{k=0}^{\infty} k!x^k$ is 0, and the interval of convergence of $\sum_{k=0}^{\infty} k!x^k$ is $\{0\}$.

Example 9.91. Consider the power series $\sum\limits_{k=0}^{\infty} 3(x-2)^k$. Note that for each $x \in \mathbb{R}$,

$$\lim_{k\to\infty} \frac{3|x-2|^{k+1}}{3|x-2|^k} = \lim_{k\to\infty} |x-2| = |x-2|\,;$$

thus the ratio test implies that the power series $\sum\limits_{k=0}^{\infty} 3(x-2)^k$ converges absolutely if $|x-2| < 1$ and diverges if $|x-2| > 1$. Therefore, the radius of convergence is 1.

To see the interval of convergence, we still need to determine if the power series converges at end-point 1 or 3. However, the power series clearly does not converge at 1 and 3; thus the interval of convergence is $(1,3)$.

Example 9.92. Consider the power series $\sum\limits_{k=1}^{\infty} \frac{x^k}{k^2}$. Note that for each $x \in \mathbb{R}$,

$$\lim_{k\to\infty} \frac{\left|\dfrac{x^{k+1}}{(k+1)^2}\right|}{\left|\dfrac{x^k}{k^2}\right|} = \lim_{k\to\infty} \frac{k^2|x|}{(k+1)^2} = |x|\,;$$

thus the ratio test implies that the power series $\sum\limits_{k=0}^{\infty} \frac{x^k}{k^2}$ converges absolutely if $|x| < 1$ and diverges if $|x| > 1$. Therefore, the radius of convergence is 1.

To see the interval of convergence, we note that $\sum\limits_{k=1}^{\infty} \frac{1}{k^2}$ converges since it is a p-series with $p = 2$, and $\sum\limits_{k=1}^{\infty} \frac{(-1)^k}{k^2}$ converges since it converges absolutely (or simply because it is an alternating series). Therefore, the interval of convergence of the power series is $[-1,1]$.

Example 9.93. Consider the power series $\sum\limits_{k=1}^{\infty} \frac{x^k}{k}$. Note that for each $x \in \mathbb{R}$,

$$\lim_{k\to\infty} \frac{\left|\dfrac{x^{k+1}}{k+1}\right|}{\left|\dfrac{x^k}{k}\right|} = \lim_{k\to\infty} \frac{k|x|}{k+1} = |x|\,;$$

thus the ratio test implies that the power series $\sum\limits_{k=0}^{\infty} \frac{x^k}{k}$ converges absolutely if $|x| < 1$ and diverges if $|x| > 1$. Therefore, the radius of convergence is 1.

To see the interval of convergence, we note that $\sum\limits_{k=1}^{\infty} \frac{1}{k}$ diverges since it is a p-series with $p = 1$, and $\sum\limits_{k=1}^{\infty} \frac{(-1)^k}{k^2}$ converges since it is an alternating series. Therefore, the interval of convergence of the power series is $[-1,1)$.

Similarly, the power series $\sum_{k=1}^{\infty} \frac{(-1)^k x^k}{k}$ has interval of convergence $(-1, 1]$.

Example 9.94. Consider the power series $\sum_{k=1}^{\infty} \frac{x^k}{k^2}$. Note that for each $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{\left|\dfrac{x^{n+1}}{(n+1)^2}\right|}{\left|\dfrac{x^n}{n^2}\right|} = \lim_{n\to\infty} \frac{n^2|x|}{(n+1)^2} = |x|\,;$$

thus the ratio test implies that the power series $\sum_{k=1}^{\infty} \frac{x^k}{k^2}$ converges absolutely if $|x| < 1$ and diverges if $|x| > 1$. Therefore, the radius of convergence is 1.

To see the interval of convergence, we note that $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges since it is a p-series with $p = 2$, and $\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}$ also converges since it converges absolutely (or because of Dirichlet's test). Therefore, the interval of convergence of the power series is $[-1, 1]$.

Remark 9.95. Even though the examples above all has radius of convergence 1, it is not necessary that the radius of convergence of a power series is always 1. For example, the power series $\sum_{k=1}^{\infty} \frac{x^k}{2^k k}$ is obtained by replacing x by $\frac{x}{2}$ in Example 9.93; thus

$$\sum_{k=1}^{\infty} \frac{x^k}{2^k k} \text{ converges for } \frac{x}{2} \in [-1, 1)$$

or equivalent, the interval of convergence of $\sum_{k=1}^{\infty} \frac{x^k}{2^k k}$ is $[-2, 2)$; thus the radius of convergence of this power series is 2.

Example 9.96. The radius of convergence of the power series $\sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!}$ is ∞ since for all $x \in \mathbb{R}$,

$$\lim_{k\to\infty} \frac{\left|\dfrac{(-1)^{k+1} x^{2(k+1)+1}}{[2(k+1)+1]!}\right|}{\left|\dfrac{(-1)^k x^{2k+1}}{(2k+1)!}\right|} = \lim_{k\to\infty} \frac{\left|\dfrac{(-1)^{k+1} x^{2k+3}}{(2k+3)!}\right|}{\left|\dfrac{(-1)^k x^{2k+1}}{(2k+1)!}\right|} = \lim_{k\to\infty} \frac{x^2}{(2k+3)(2k+2)} = 0\,.$$

- **Differentiation and Integration of Power Series**

Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of real numbers and $c \in \mathbb{R}$. If the power series $\sum_{k=0}^{\infty} a_k (x-c)^k$ converges in an interval $(c-r, c+r)$, we can ask ourselves whether the function $f : (c-r, c+r)$

defined by $f(x) = \sum\limits_{k=0}^{\infty} a_k(x-c)^k$ is differentiable or not. We note that even though the power series is an infinite sum of differentiable functions (in fact, monomials), it is not clear if the limiting process $\dfrac{d}{dx}$ commutes with $\sum\limits_{k=0}^{\infty}$ since

$$\lim_{n\to\infty}\lim_{h\to 0} nh^2 = 0 \qquad \text{but} \qquad \lim_{h\to 0}\lim_{n\to\infty} nh^2 = \infty\,.$$

Theorem 9.97: Properties of Functions Defined by Power Series

If the function

$$f(x) = \sum_{k=0}^{\infty} a_k(x-c)^k = a_0 + a_1(x-c) + a_2(x-c)^2 + \cdots$$

has a radius of convergence of $R > 0$, then

1. f is differentiable on $(c-R, c+R)$ and

$$f'(x) = \sum_{k=1}^{\infty} ka_k(x-c)^{k-1} = a_1 + 2a_2(x-c) + 3a_3(x-c)^2 + \cdots\,.$$

2. an anti-derivative of f on $(c-R, c+R)$ is given by

$$\int f(x)\,dx = C + \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-c)^{k+1} = C + a_0(x-c) + \frac{a_1}{2}(x-c)^2 + \cdots\,.$$

The radius of convergence of the power series obtained by differentiating or integrating a power series term by term is the same as the original power series.

Remark 9.98. Theorem 9.97 states that, in many ways, a function defined by a power series behaves like a polynomial; that is, the derivative (or anti-derivative) of a power series can be obtained by term-by-term differentiation (or integration). However, it is not true for general functions defined by series of the form $\sum\limits_{k=0}^{\infty} b_k(x)$. For example, we have talked about (but did not prove) the series $\sum\limits_{k=1}^{\infty} \dfrac{\sin kx}{k}$ which is the same as $\dfrac{\pi - x}{2}$ on $(0, 2\pi)$; that is,

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k} = \frac{\pi - x}{2} \qquad \forall\, x \in (0, 2\pi)\,.$$

Then
$$-\frac{1}{2} = \frac{d}{dx}\sum_{k=1}^{\infty}\frac{\sin kx}{k} \qquad \forall\, x \in (0, 2\pi)$$
but
$$\frac{d}{dx}\sum_{k=1}^{\infty}\frac{\sin kx}{k} \neq \sum_{k=1}^{\infty}\frac{d}{dx}\frac{\sin kx}{k} = \sum_{k=1}^{\infty}\cos kx \qquad \forall\, x \in (0, 2\pi)$$
since the series $\sum\limits_{k=1}^{\infty}\cos kx$ does not converges for all $x \in (0, 2\pi)$.

Example 9.99. Consider the function f defined by power series
$$f(x) = \sum_{k=1}^{\infty}\frac{x^k}{k} = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots \qquad \forall\, x \in [-1, 1)\,.$$
Then the function
$$g(x) = \sum_{k=1}^{\infty}x^{k-1} = \sum_{k=0}^{\infty}x^k = 1 + x + x^2 + \cdots\,,$$
obtained by term-by-term differentiation, converges for $x \in (-1, 1)$, and the function
$$h(x) = \sum_{k=1}^{\infty}\frac{x^{k+1}}{k(k+1)} = \sum_{k=2}^{\infty}\frac{x^k}{k(k-1)} = \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{12} + \cdots$$
obtained by term-by-term differentiation, converges for $x \in [-1, 1]$.

Example 9.100. Suppose that x is a function of t satisfying
$$x''(t) + x(t) = 0\,, \qquad x(0) = x'(0) = 1\,.$$
Assume that $x(t) = \sum\limits_{k=0}^{\infty}a_k t^k$ for $t \in (-R, R)$ with some radius of convergence $R > 0$. Then Theorem 9.97 implies that
$$x''(t) = \sum_{k=2}^{\infty}k(k-1)a_k t^{k-2} = \sum_{k=0}^{\infty}(k+2)(k+1)a_{k+2}t^k \qquad \forall\, t \in (-R, R)\,;$$
thus if $t \in (-R, R)$,
$$\sum_{k=0}^{\infty}\big[(k+2)(k+1)a_{k+2} + a_k\big]t^k = \sum_{k=0}^{\infty}(k+2)(k+1)a_{k+2}t^k + \sum_{k=0}^{\infty}a_k t^k = x''(t) + x(t) = 0\,.$$
The equality above implies that
$$(k+2)(k+1)a_{k+2} + a_k = 0 \qquad \forall\, k \in \mathbb{N} \cup \{0\}\,.$$

Therefore,

$$a_{2k} = \frac{-1}{(2k)(2k-1)}a_{2k-2} = \frac{(-1)^2}{(2k)(2k-1)(2k-2)(2k-4)}a_{2k-4} = \cdots = \frac{(-1)^k}{(2k)!}a_0\,,$$
$$a_{2k+1} = \frac{-1}{(2k+1)(2k)}a_{2k-1} = \frac{(-1)^2}{(2k+1)(2k)(2k-1)(2k-2)}a_{2k-3} = \cdots = \frac{(-1)^k}{(2k+1)!}a_1\,.$$

Since $x(0) = x'(0) = 1$ implies $a_0 = a_1 = 1$, we have

$$x(t) = \sum_{k=0}^{\infty}\Big[\frac{(-1)^k}{(2k)!}t^{2k} + \frac{(-1)^k}{(2k+1)!}t^{2k+1}\Big] = \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}t^{2k} + \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}t^{2k+1} = \cos t + \sin t\,.$$

Corollary 9.101

For a function defined by power series

$$f(x) = \sum_{k=0}^{\infty} a_k(x-c)^k$$

(on a certain interval of convergence), the n-th Taylor polynomial for f at c is the n-th partial sum $\sum\limits_{k=0}^{n} a_k(x-c)^k$ of the power series.

9.9 Representation of Functions by Power Series

We have shown the following identities:

$$\exp(x) = \sum_{k=0}^{\infty}\frac{x^k}{k!} \qquad \forall\, x \in \mathbb{R}\,,$$
$$\sin x = \sum_{k=0}^{\infty}\frac{(-1)^k x^{2k+1}}{(2k+1)!} \qquad \forall\, x \in \mathbb{R}\,,$$
$$\cos x = \sum_{k=0}^{\infty}\frac{(-1)^k x^{2k}}{(2k)!} \qquad \forall\, x \in \mathbb{R}\,,$$
$$\ln(1+x) = \sum_{k=1}^{\infty}\frac{(-1)^{k-1} x^{k}}{k} \qquad \forall\, x \in (-1,1]\,.$$

In this section, we are interested in finding the power series representation (centered at c) of functions of the form

$$f(x) = \frac{1}{b-x}\,.$$

(without differentiating the function). In other words, for a given $c \in \mathbb{R}\backslash\{b\}$ we would like to find $\{a_k\}_{k=0}^{\infty}$ (which usually depends on c) such that $f(x)$ agrees with the power series

$$\sum_{k=0}^{\infty} a_k (x-c)^k$$

on a certain interval of convergence without differentiating f. For example, we know that

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \qquad \forall\, x \in (-1,1)\,;$$

thus to "expand the function about $\frac{1}{2}$"; that is, to write the function $y = \dfrac{1}{1-x}$ as a power series centered at $\frac{1}{2}$, we have

$$\frac{1}{1-x} = \frac{1}{\frac{1}{2} - \left(x - \frac{1}{2}\right)} = 2 \cdot \frac{1}{1 - 2\left(x - \frac{1}{2}\right)} = 2\sum_{k=0}^{\infty} \left[2\left(x - \tfrac{1}{2}\right)\right]^k \text{ if } x \text{ satisfying } 2\left|x - \tfrac{1}{2}\right| < 1\,.$$

In other words, we obtain

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} 2^{k+1}\left(x - \frac{1}{2}\right)^k \qquad \forall\, x \in (0,1)$$

without computing the derivatives of the function $y = \dfrac{1}{1-x}$ at $\dfrac{1}{2}$.

We emphasize that f is defined on $\mathbb{R}\backslash\{c\}$ and the power series $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ converges only on an interval; thus the function $y = f(x)$ is never the same as the function defined by power series.

- **Geometric Power Series**

Recall that the geometric series $\sum\limits_{k=0}^{\infty} r^k$ converges if and only if $|r| < 1$. The function $g(x) = \dfrac{1}{1-x}$ is defined on $\mathbb{R}\backslash\{1\}$, and by the fact that

$$\frac{1-x^{n+1}}{1-x} = 1 + x + x^2 + \cdots + x^n = \sum_{k=0}^{n} x^k \qquad \forall\, x \neq 1\,,$$

we find that if $|x| < 1$, then

$$\lim_{n\to\infty} \sum_{k=0}^{n} x^k = \lim_{n\to\infty} \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x}\,;$$

thus $\frac{1}{1-x} = \sum\limits_{k=0}^{\infty} x^k$ on $(-1,1)$. Therefore, for $c \neq b$,

$$\frac{1}{b-x} = \frac{1}{b-c} \cdot \frac{1}{1-\dfrac{x-c}{b-c}} = \frac{1}{b-c}\sum_{k=0}^{\infty}\Big(\frac{x-c}{b-c}\Big)^k \qquad \forall\, x \text{ satisfying } \Big|\frac{x-c}{b-c}\Big| < 1\,,$$

or equivalently,

$$\frac{1}{b-x} = \sum_{k=0}^{\infty} \frac{1}{(b-c)^{k+1}}(x-c)^k \qquad \forall\, x \in (c-|b-c|, c+|b-c|)\,.$$

Replacing x by $-x$, we find that

$$\frac{1}{b+x} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(b-c)^{k+1}}(x+c)^k \qquad \forall\, x \in (-c-|b-c|, -c+|b-c|)\,.$$

Example 9.102. Find a power series representation for $f(x) = \frac{1}{x}$, centered at 1.

To find the power series centered at 1, we rewrite $\frac{1}{x} = \frac{1}{1+(x-1)}$; thus

$$\frac{1}{x} = \frac{1}{1-(1-x)} = \sum_{k=0}^{\infty}(1-x)^k = \sum_{k=0}^{\infty}(-1)^k(x-1)^k \qquad \forall\, |x-1| < 1\,.$$

Example 9.103. Find a power series representation for $f(x) = \ln x$ centered at 1.

Note that $\frac{d}{dx}\ln x = \frac{1}{x}$; thus

$$\frac{d}{dx}\ln x = \sum_{k=0}^{\infty}(-1)^k(x-1)^k \qquad \forall\, x \in (0,2)\,.$$

Therefore, by Theorem 9.97,

$$\ln x = C + \sum_{k=0}^{\infty}\frac{(-1)^k}{k+1}(x-1)^{k+1} = C + \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}(x-1)^k \qquad \forall\, x \in (0,2)\,.$$

To determine the constant C, we let $x = 1$ and find that $\ln 1 = C$; thus $C = 0$ and we conclude that

$$\ln x = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}(x-1)^k \qquad \forall\, x \in (0,2)\,.$$

We note that the power series converges at $x = 2$, and Example 9.84 shows that

$$\ln 2 = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}\,.$$

In other words, the power series $\sum\limits_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}(x-1)^k$ is continuous at 2

• **Operations with Power Series**

Let $f(x) = \sum\limits_{k=0}^{\infty} a_k(x-c)^k$ have interval of convergence I_1 and $g(x) = \sum\limits_{k=0}^{\infty} b_k(x-c)^k$ have interval of convergence I_2.

1. $f(\alpha x) = \sum\limits_{k=0}^{\infty} a_k\alpha^k\big(x-\dfrac{c}{\alpha}\big)^k$ on $I \equiv \big\{x \in \mathbb{R} \,\big|\, \alpha x \in I_1\big\}$.

2. $f(x) + g(x) = \sum\limits_{k=0}^{\infty} (a_k + b_k)x^k$ on $I \equiv I_1 \cap I_2$.

3. If $c = 0$ and $N \in \mathbb{N}$, then $f(x^N) = \sum\limits_{k=0}^{\infty} a_k x^{Nk}$ on $I \equiv \big\{x \in \mathbb{R} \,\big|\, x^N \in I_1\big\}$.

4. $f(x)g(x) = \sum\limits_{k=0}^{\infty} d_k(x-c)^k$ on $I \equiv I_1 \cap I_2$, where $d_k = \sum\limits_{j=0}^{k} a_k b_{j-k}$.

Example 9.104. Find a power series for $f(x) = \arctan x$ centered at 0.

Note that $\dfrac{d}{dx}\arctan x = \dfrac{1}{1+x^2}$; thus

$$\frac{d}{dx}\arctan x = \frac{1}{1+x^2} = \sum_{k=0}^{\infty}(-1)^k x^{2k} \qquad \forall\, x \in (-1,1)\,.$$

By Theorem 9.97,

$$\arctan x = C + \sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}x^{2k+1} \qquad \forall\, x \in (-1,1)\,,$$

and the constant C is determined by applying the identity above at $x = 0$; thus $C = \arctan 0$ and

$$\arctan x = \sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}x^{2k+1} \qquad \forall\, x \in (-1,1)\,,$$

We note that the power series converges at $x = \pm 1$. Is it true that $\arctan 1 = 1 - \dfrac{1}{3} + \dfrac{1}{5} - \dfrac{1}{7} + \cdots$?

In general, suppose that the function f defined by power series $\sum\limits_{k=0}^{\infty} a_k(x-c)^k$ has a radius of convergence $R > 0$, and g is a continuous function defined on some interval I such that $f(x) = g(x)$ for all $x \in (c-R, c+R) \subsetneq I$. If f is also defined on $c+R$ (or $c-R$), by Theorem 9.97 it is not clear if $\lim\limits_{x\to c+R} f(x) = g(c+R)$ $\big($or $\lim\limits_{x\to c-R} f(x) = g(c-R)\big)$. The following theorem concerns with this issue.

Theorem 9.105: Continuity of Power Series at End-points

Let the radius of convergence of the power series $f(x) = \sum_{k=0}^{\infty} a_k(x-c)^k$ be r for some $r > 0$.

1. If $\sum_{k=0}^{\infty} a_k r^k$ converges, then f is continuous at $c + r$; that is,
$$\lim_{x \to (c+r)^-} f(x) = f(c+r)\,.$$
2. If $\sum_{k=0}^{\infty} a_k(-r)^k$ converges, then f is continuous at $c - r$; that is,
$$\lim_{x \to (c-r)^+} f(x) = f(c-r)\,.$$

Therefore, it is true that
$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \cdots + \frac{(-1)^n}{2n+1} + \cdots\,.$$

9.10 Taylor and Maclaurin Series

Definition 9.106

If a function f has derivatives of all orders at $x = c$, then the series
$$\sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k$$
is called the Taylor series for f at c. It is also called the Maclaurin series for f if $c = 0$.

Theorem 9.107: Convergence of Taylor Series

Let f be a function that has derivatives of all orders at $x = c$, and P_n be the n-th Taylor polynomial for f at c. If R_n, the remainder associated with P_n, has the property that
$$\lim_{n \to \infty} R_n(x) = 0 \qquad \forall\, x \in I$$
for some interval I, then the Taylor series for f converges and equals $f(x)$; that is,
$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k \qquad \forall\, x \in I\,.$$

Corollary 9.108

Let f be a function that has derivatives of all orders in an open interval I containing c. If there exists $M > 0$ such that $\left|f^{(k)}(x)\right| \leqslant M$ for all $x \in I$ and each $k \in \mathbb{N}$, then

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k \qquad \forall\, x \in I\,.$$

Proof. By the Taylor Theorem,

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + R_n(x)\,,$$

where

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

for some ξ between c and x. Since $\left|f^{(k)}(x)\right| \leqslant M$ for all $x \in I$ and $k \in \mathbb{N}$, we find that

$$\left|R_n(x)\right| \leqslant \frac{M}{(n+1)!}|x-c|^{n+1} \qquad \forall\, x \in I\,.$$

Therefore, by the fact that $\lim\limits_{n\to\infty} \dfrac{a^n}{n!} = 0$ for all $a \in \mathbb{R}$ (the same reasoning as in Example 9.79), the Squeeze Theorem implies that

$$\lim_{n\to\infty} R_n(x) = 0 \qquad \forall\, x \in I$$

and Theorem 9.107 further shows that $f(x) = \sum\limits_{k=0}^{\infty} \dfrac{f^{(k)}(c)}{k!}(x-c)^k$. $\square$

Example 9.109. Since the k-th derivatives of the sine function is bounded by 1; that is,

$$\left|\frac{d^k}{dx^k}\sin x\right| \leqslant 1 \qquad \forall\, x \in \mathbb{R} \text{ and } k \in \mathbb{N}\,,$$

Corollary 9.108 implies that for all $c \in \mathbb{R}$,

$$\sin x = \sum_{k=0}^{\infty} \frac{1}{k!}\sin\big(c + \frac{k\pi}{2}\big)(x-c)^k \qquad \forall\, x \in \mathbb{R}\,,$$

here we have used $\dfrac{d^k}{dx^k}\sin x = \sin\left(x + \dfrac{k\pi}{2}\right)$ to compute the k-th derivative of the sine function. In particular,

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \qquad \forall\, x \in \mathbb{R}\,.$$

Similarly, for all $c \in \mathbb{R}$,

$$\cos x = \sum_{k=0}^{\infty} \frac{1}{k!} \cos\left(c + \frac{k\pi}{2}\right)(x-c)^k \qquad \forall\, x \in \mathbb{R}\,.$$

Example 9.110. Consider the natural exponential function $y = \exp(x)$. Note that for all real numbers $R > 0$, we have

$$\left|\frac{d^k}{dx^k} e^x\right| = e^x \leqslant e^R \qquad \forall\, x \in (-R, R) \text{ and } k \in \mathbb{N}\,;$$

thus Corollary 9.108 implies that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \cdots \qquad \forall\, x \in (-R, R)\,.$$

Since the identity above holds for all $R > 0$, we conclude that

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \cdots \qquad \forall\, x \in \mathbb{R}\,.$$

Example 9.111 (Binomial Series)**.** In this example we consider the Maclaurin series, called the binomial series, of the function $f(x) = (1+x)^\alpha$, where $\alpha \in \mathbb{R}$ and $\alpha \neq \mathbb{N} \cup \{0\}$.

We compute the derivative of f and find that

$$\frac{d^k}{dx^k}(1+x)^\alpha = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)(1+x)^{\alpha-k}\,.$$

Therefore,

$$f^{(k)}(0) = \frac{d^k}{dx^k}\Big|_{x=0}(1+x)^\alpha = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)$$

and the Maclaurin series for f is

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

To see the radius of convergence of the Maclaurin series above, we use the ratio test and find that

$$\lim_{n\to\infty} \frac{\dfrac{\big|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(n+1)+1)\big|}{(n+1)!}|x|^{n+1}}{\dfrac{\big|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)\big|}{n!}|x|^n} = \lim_{n\to\infty} \frac{|\alpha-n|}{n+1}|x| = |x|\,;$$

thus the radius of convergence of the power series $\sum\limits_{k=0}^{\infty} \dfrac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}x^k$ is 1. Moreover, by Taylor's theorem, for each $x \in (-1,1)$ there exists ξ between 0 and x such that

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

where

$$R_n(x) = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n)}{(n+1)!}(1+\xi)^{\alpha-n-1}x^{n+1}\,.$$

Similar to Example 9.76, we have

$$\big|R_n(x)\big| \leqslant \frac{\big|\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n)\big|}{(n+1)!}x^\alpha \qquad \forall\, x \in (0,1);$$

thus (without detail reasoning) we find that

$$\lim_{n\to\infty} R_n(x) = 0 \qquad \forall\, x \in (0,1)\,.$$

Therefore,

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}x^k \qquad \forall\, x \in (0,1)\,.$$

In fact,

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}x^k \qquad \forall\, x \in (-1,1)\,.$$

9.11 Exercise

Problem 9.1. Determine whether the sequence $\{a_n\}_{n=1}^{\infty}$ converges or diverges. If it converges, find the limit.

(1) $a_n = \dfrac{\ln n}{\ln(2n)}$ (2) $a_n = \dfrac{(-1)^{n+1}n}{n+\sqrt{n}}$ (3) $a_n = n\sin\dfrac{1}{n}$ (4) $a_n = n - \sqrt{n+1}\sqrt{n+3}$

(5) $a_n = \sqrt[n]{n^2+n}$ (6) $a_n = (3^n+5^n)^{\frac{1}{n}}$ (7) $a_n = \dfrac{1}{\sqrt{n^2-1}-\sqrt{n^2+n}}$

(8) $a_n = \sqrt{n}\ln\left(1+\dfrac{1}{n}\right)$ (9) $a_n = \dfrac{1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)}{2^n n!}$ (10) $a_n = \dfrac{1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)}{2^n (n+1)!}$.

Problem 9.2. Determine whether the series $\sum\limits_{n=1}^{\infty} a_n$ is convergent or divergent. If it is convergent, find its sum.

(1) $a_n = \dfrac{1}{1+(\frac{2}{3})^n}$ (2) $a_n = \ln\left(\dfrac{n^2+1}{2n^2+1}\right)$ (3) $a_n = e^{-n} + \dfrac{1}{n(n+1)}$

(4) $a_n = \dfrac{1}{n^3-n}$ (5) $a_n = \dfrac{40n}{(2n-1)^2(2n+1)^2}$

Problem 9.3. Find values of x for which the following series converges.

(1) $\sum\limits_{n=1}^{\infty}(-4)^n(x-5)^n$ (2) $\sum\limits_{n=1}^{\infty}\dfrac{2^n}{x^n}$ (3) $\sum\limits_{n=1}^{\infty}\dfrac{\sin^n x}{3^n}$ (4) $\sum\limits_{n=1}^{\infty} e^{nx}$.

Problem 9.4. Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences of real numbers.

(1) Show that if $\lim\limits_{n\to\infty}(a_n+b_n)$ D.N.E. and $\lim\limits_{n\to\infty} b_n$ converges, then $\lim\limits_{n\to\infty} a_n$ D.N.E.

(2) Show that if $\sum\limits_{n=1}^{\infty}(a_n+b_n)$ diverges and $\sum\limits_{n=1}^{\infty} b_n$ converges, then $\sum\limits_{n=1}^{\infty} a_n$ diverges.

Problem 9.5. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers, and $\{\sigma_n\}_{n=1}^{\infty}$ be a sequence of real numbers defined by

$$\sigma_n = \frac{a_1+a_2+\cdots+a_n}{n} = \frac{1}{n}\sum_{k=1}^{n} a_k\,.$$

(1) Show that if $\lim\limits_{n\to\infty} a_n = a$ exists, then $\lim\limits_{n\to\infty}\sigma_n = a$.

(2) Suppose that $\lim\limits_{n\to\infty}\sigma_n = a$ exists, is it necessary that $\lim\limits_{n\to\infty} a_n = a$?

Problem 9.6. Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of real numbers defined recursively by

$$a_{n+1} = \sqrt{1+a_n} \quad \forall\, n\in\mathbb{N}\cup\{0\}\,, a_0 = 0\,.$$

Show that $\{a_n\}_{n=1}^{\infty}$ converges and find the limit.

Problem 9.7. Let $a_n = \left(1+\dfrac{1}{n}\right)^n$.

(1) Show that if $0 \leqslant a < b$, then

$$\frac{b^{n+1} - a^{n+1}}{b - a} < (n+1)b^n\,.$$

(2) Deduce that $b^n\big[(n+1)a - nb\big] < a^{n+1}$.

(3) Use $a = 1 + \dfrac{1}{n+1}$ and $b = 1 + \dfrac{1}{n}$ in (2) to show that $\{a_n\}_{n=1}^{\infty}$ is (strictly) increasing.

(4) Use $a = 1$ and $b = 1 + \dfrac{1}{2n}$ in (2) to show that $a_{2n} < 4$.

(5) Use (3) and (4) to show that $a_n < 4$.

(6) Deduce that $\{a_n\}_{n=1}^{\infty}$ converges.

Problem 9.8. Let a, b be positive real numbers, $a > b$. Let two sequence $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be given by the recursive relation

$$a_{n+1} = \frac{a_n + b_n}{2}\,,\ b_{n+1} = \sqrt{a_n b_n} \quad \forall\, n \in \mathbb{N}\,, \qquad a_1 = \frac{a+b}{2}\,,\ b_1 = \sqrt{ab}\,.$$

Complete the following.

(1) Show (by induction) that $a_n > a_{n+1} > b_{n+1} > b_n$ for all $n \in \mathbb{N}$.

(2) Deduce that $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ both converges.

(3) Show that $\lim\limits_{n\to\infty} a_n$ and $\lim\limits_{n\to\infty} b_n$ both exist and are identical.

Problem 9.9. Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of real number defined by the recursive relation

$$a_{n+1} = \frac{1}{2 + a_n} \quad \forall\, n \geqslant 0\,, \qquad a_0 = \frac{1}{2}\,.$$

Complete the following.

(1) Show that the sequence $\{a_{2n}\}_{n=0}^{\infty}$ is a decreasing sequence; that is, $a_{2n+2} \leqslant a_{2n}$ for all $n \in \mathbb{N} \cup \{0\}$.

(2) Show that the sequence $\{a_{2n+1}\}_{n=0}^{\infty}$ is an increasing sequence; that is, $a_{2n+3} \geqslant a_{2n+1}$ for all $n \in \mathbb{N} \cup \{0\}$.

(3) Show that $a_{2k+1} \leqslant a_{2\ell}$ for all $k, \ell \in \mathbb{N} \cup \{0\}$.

(4) Show that the two sequences $\{a_{2n}\}_{n=0}^{\infty}$ and $\{a_{2n+1}\}_{n=0}^{\infty}$ converges to the same limit.

(5) Show that $\{a_n\}_{n=0}^{\infty}$ converges.

Problem 9.10. The Fibonacci sequence $\{f_n\}_{n=1}^{\infty}$ is a sequence defined recursively by

$$f_1 = 1\,, \quad f_2 = 1 \qquad \text{and} \qquad f_{n+2} = f_{n+1} + f_n \quad \forall\, n \in \mathbb{N}\,.$$

Show the following.

(1) $\dfrac{1}{f_{n-1}f_{n+1}} = \dfrac{1}{f_{n-1}f_n} - \dfrac{1}{f_nf_{n+1}}$ for all $n \geqslant 2$.

(2) $\sum\limits_{n=2}^{\infty} \dfrac{1}{f_{n-1}f_{n+1}} = 1$.

(3) $\sum\limits_{n=2}^{\infty} \dfrac{f_n}{f_{n-1}f_{n+1}} = 2$.

Problem 9.11. Consider the series $\sum\limits_{n=1}^{\infty} \dfrac{n}{(n+1)!}$.

(1) Find the partial sum S_1, S_2, S_3 and S_4. Do you recognize the denominators? Use the pattern to guess a formula for S_n.

(2) Prove your guess by induction.

(3) Show that the given series is convergent, and find the sum.

Problem 9.12. Determine whether the series $\sum\limits_{n=1}^{\infty} a_n$ is convergent or divergent.

(1) $a_n = \dfrac{1}{n^{1+\frac{1}{n}}}$ (2) $a_n = \ln\left(1 + \dfrac{1}{n^2}\right)$ (3) $a_n = \dfrac{2^n + 3^n}{3^n + 4^n}$ (4) $a_n = \tan\dfrac{1}{n}$

(5) $a_n = \sin^n \dfrac{1}{\sqrt{n}}$ (6) $a_n = \dfrac{\arctan n}{n^{1.1}}$ (7) $a_n = \left[-\ln\left(e^2 + \dfrac{1}{n^2}\right)\right]^n$

(8) $a_n = \left(1 - \dfrac{1}{n}\right)^{n^2}$ (9) $a_n = \left(1 + \dfrac{1}{n}\right)^{-n^2}$ (10) $a_n = \dfrac{(n!)^2}{(2n)!}$ (11) $a_n = \dfrac{n!\ln n}{n(n+2)!}$

(12) $a_n = \dfrac{n!}{n^n}$ (13) $a_n = \dfrac{(-1)^n(3n)!}{n!(n+1)!(n+2)!}$ (14) $a_n = \dfrac{1\cdot 3\cdot 5\cdot\cdots\cdot(2n-1)}{2^n n!}$

(15) $a_n = (-1)^n\left(\sqrt{n+\sqrt{n}} - \sqrt{n}\right)$ (16) $a_n = (-1)^n\dfrac{(n!)^2 3^n}{(2n+1)!}$ (17) $a_n = \dfrac{(-1)^n(n!)^n}{n^{n^2}}$

Problem 9.13. Find all p and q such that $\sum\limits_{k=2}^{\infty} \dfrac{(\ln k)^q}{k^p}$ converges.

Problem 9.14. Show that if $\sum\limits_{k=1}^{\infty} a_k$ is a convergent series of positive terms, then $\sum\limits_{k=1}^{\infty} \sin a_k$ converges.

Problem 9.15. Let $S = \sum\limits_{k=1}^{\infty} \dfrac{1}{k^2}$. Euler found that $S = \dfrac{\pi^2}{6}$ in 1735 AD.

(1) Show that $S = 1 + \sum\limits_{k=1}^{\infty} \dfrac{1}{n^2(n+1)}$.

(2) Which of the sums $\sum\limits_{k=1}^{1000000} \dfrac{1}{k^2}$ or $1 + \sum\limits_{k=1}^{1000} \dfrac{1}{k^2(k+1)}$ should give a better approximation of S? Explain your answer.

Hint: (1) $\dfrac{1}{n^2(n+1)} = \dfrac{1}{n^2} - \dfrac{1}{n(n+1)}$.

Problem 9.16. Find all real numbers x such that $\sum\limits_{k=1}^{\infty} \dfrac{\cos(kx)}{\ln k}$ converges.

Problem 9.17. Show by example that $\sum\limits_{k=1}^{\infty} a_k b_k$ may diverge even if $\sum\limits_{k=1}^{\infty} a_k$ and $\sum\limits_{k=1}^{\infty} b_k$ both converge.

Problem 9.18. Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences of real numbers such that $a_n, b_n > 0$ for all $n \geqslant N$. Define

$$c_n = b_n - b_{n+1}\frac{a_{n+1}}{a_n} \qquad \forall\, n \in \mathbb{N}\,. \tag{9.11.1}$$

1. Show that if there exists a constant $r > 0$ such that $r < c_n$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ converges.

   Hint: Rewrite (9.11.1) as $b_n = c_n + \dfrac{a_{n+1}}{a_n} b_{n+1}$ and then obtain

$$\begin{aligned} b_N &= c_N + \frac{a_{N+1}}{a_N} b_{N+1} = c_N + \frac{a_{N+1}}{a_N}\Big(c_{N+1} + \frac{a_{N+2}}{a_{N+1}} b_{N+2}\Big) = c_N + \frac{a_{N+1}}{a_N} c_{N+1} + \frac{a_{N+2}}{a_N} b_{N+2} \\ &= c_N + \frac{a_{N+1}}{a_N} c_{N+1} + \frac{a_{N+2}}{a_N}\Big(c_{N+2} + \frac{a_{N+3}}{a_{N+2}} b_{N+3}\Big) = \cdots \\ &= c_N + \frac{a_{N+1}}{a_N} c_{N+1} + \frac{a_{N+2}}{a_N} c_{N+2} + \cdots + \frac{a_{N+n}}{a_N} c_{N+n} + \frac{a_{N+n+1}}{a_N} b_{N+n+1}\,. \end{aligned}$$

   Use the fact that $0 < r < c_n$ for all $n \geqslant N$ to conclude that

$$\sum_{k=N}^{N+n} a_k \leqslant \frac{a_N b_N}{r} \qquad \forall\, n \in \mathbb{N}\,.$$

Note that then the sequence of partial sum of $\sum\limits_{k=1}^{\infty} a_k$ then is bounded from above $\big($by $\sum\limits_{k=1}^{N-1} a_k + \dfrac{a_N b_N}{r}\big)$.

2. Show that if $\sum\limits_{k=1}^{\infty} \dfrac{1}{b_k}$ diverges and $c_n \leqslant 0$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.

   Hint: The fact that $c_n \leqslant 0$ for all $n \geqslant N$ implies that $b_n a_n \leqslant b_{n+1} a_{n+1}$ for all $n \geqslant N$. Use this fact to conclude that

$$\frac{a_N b_N}{b_n} \leqslant a_n \qquad \forall\, n \geqslant N$$

   and then apply the direct comparison test to conclude that $\sum\limits_{k=1}^{\infty} a_k$ diverges.

Problem 9.19. Let $\sum\limits_{k=1}^{\infty} a_k$ be a series with positive terms, and $\lim\limits_{n\to\infty} \dfrac{a_{n+1}}{a_n} = 1$. We know from class that the ratio test fails when this happens, but there are some refined results concerning this particular case.

1. **(Raabe's test)**:

   (a) If there exists a constant $\mu > 1$ such that $\dfrac{a_{n+1}}{a_n} < 1 - \dfrac{\mu}{n}$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ converges.

   (b) If there exists a constant $0 < \mu < 1$ such that $\dfrac{a_{n+1}}{a_n} > 1 - \dfrac{\mu}{n}$ for all $n \geqslant N$, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.

   Hint: Consider the sequence $\{b_n\}_{n=1}^{\infty}$ defined by $b_n = (n-1)a_n - na_{n+1}$. Then $\sum\limits_{k=1}^{\infty} b_k$ is a telescoping series. For case (a), show that $\big\{na_{n+1}\big\}_{n=N}^{\infty}$ is a positive decreasing sequence and then conclude that $\sum\limits_{k=1}^{\infty} b_k$ converges. Note that $b_n \geqslant (\mu-1)a_n$ for all $n \geqslant N$. For case (b), show that $\big\{na_{n+1}\big\}_{n=N}^{\infty}$ is a positive increasing sequence; thus $a_n \geqslant \dfrac{Na_{N+1}}{n-1}$ for all $n \geqslant N+1$ which implies that $\sum\limits_{k=1}^{\infty} a_k$ diverges.

   Remark: 注意到 (a) 說的是如果 $\{a_n\}_{n=1}^{\infty}$ 在某項之後「遞減得夠快」，那麼 $\sum\limits_{k=1}^{\infty} a_k$ 收斂。反之，如果 $\{a_n\}_{n=1}^{\infty}$「並非遞減得那麼快」，那麼 $\sum\limits_{k=1}^{\infty} a_k$ 發散。

2. **(Gauss's test)**: Suppose that there exist a positive constant $\epsilon > 0$, a constant μ, and a **bounded** sequence $\{R_n\}_{n=1}^{\infty}$ such that

$$\frac{a_{n+1}}{a_n} = 1 - \frac{\mu}{n} + \frac{R_n}{n^{1+\epsilon}} \qquad \text{for all } n \geqslant N\,.$$

(a) If $\mu > 1$, then $\sum\limits_{k=1}^{\infty} a_k$ converges. (b) If $\mu \leqslant 1$, then $\sum\limits_{k=1}^{\infty} a_k$ diverges.

Hint: Show that if $\mu > 1$ or $\mu < 1$, one can apply Raabe's test to conclude Gauss's test. For the case $\mu = 1$, let $b_n = (n-1)\ln(n-1)$ for $n \geqslant 2$. Using the second result of Problem 9.18 to show the divergence of $\sum\limits_{k=1}^{\infty} a_k$ (by showing that c_n defined by (9.11.1) is non-positive for all large enough n).

Problem 9.20. Complete the following.

1. Show that $\sum\limits_{k=1}^{\infty} \Big(1 - \dfrac{1}{\sqrt{k}}\Big)^k$ converges.
2. Show that $\sum\limits_{k=2}^{\infty} \dfrac{\log(k+1) - \log k}{(\log k)^2}$ converges.
3. Use Gauss's test to show that both the general harmonic series $\sum\limits_{k=1}^{\infty} \dfrac{1}{ak+b}$, where $a \neq 0$, and the series $\sum\limits_{k=1}^{\infty} \dfrac{1}{\sqrt{k}}$ diverge.
4. Show that $\sum\limits_{k=1}^{\infty} \dfrac{k!}{(\alpha+1)(\alpha+2)\cdots(\alpha+k)}$ converges if $\alpha > 1$ and diverges if $\alpha \leqslant 1$.
5. Test the following "hypergeometric" series for convergence or divergence:

   (a) $\sum\limits_{k=1}^{\infty} \dfrac{\alpha(\alpha+1)(\alpha+2)\cdots(\alpha+k-1)}{\beta(\beta+1)(\beta+2)\cdots(\beta+k-1)} = \dfrac{\alpha}{\beta} + \dfrac{\alpha(\alpha+1)}{\beta(\beta+1)} + \dfrac{\alpha(\alpha+1)(\alpha+2)}{\beta(\beta+1)(\beta+2)} + \cdots$.

   (b) $1 + \dfrac{\alpha\cdot\beta}{1\cdot\gamma} + \dfrac{\alpha(\alpha+1)\cdot\beta(\beta+1)}{1\cdot 2\gamma\cdot(\gamma+1)} + \dfrac{\alpha(\alpha+1)(\alpha+2)\cdot\beta(\beta+1)(\beta+2)}{1\cdot 2\cdot 3\cdot\gamma(\gamma+1)(\gamma+2)} + \cdots$.

Problem 9.21. Let $\sum\limits_{k=1}^{\infty} a_k$ be a conditionally convergent series. Show that $\sum\limits_{k=1}^{\infty} \big[1+\operatorname{sgn}(a_k)\big]a_k$ and $\sum\limits_{k=1}^{\infty} \big[1-\operatorname{sgn}(a_k)\big]a_k$ both diverge. Here the sign function sgn is defined by

$$\operatorname{sgn}(a) = \begin{cases} 1 & \text{if } a > 0\,, \\ 0 & \text{if } a = 0\,, \\ -1 & \text{if } a < 0\,. \end{cases}$$

Problem 9.22. A permutation of a non-empty set A is a one-to-one function from A onto A. Let $\pi : \mathbb{N} \to \mathbb{N}$ be a permutation of $\mathbb{N}$.

1. Suppose that $\{a_n\}_{n=1}^{\infty}$ be a convergent sequence of real numbers. Show that $\{a_{\pi(n)}\}_{n=1}^{\infty}$ is also convergent; that is, show that if $\{b_n\}_{n=1}^{\infty}$ is a sequence defined by $b_n = a_{\pi(n)}$, then $\{b_n\}_{n=1}^{\infty}$ also converges.

2. Suppose that $\sum\limits_{k=1}^{\infty} a_k$ is absolutely convergent. Show that $\sum\limits_{k=1}^{\infty} a_{\pi(k)}$ is also absolutely convergent, and
$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_{\pi(k)} \,.$$

3. Suppose that $\sum\limits_{k=1}^{\infty} a_k$ is conditionally convergent. Show that for each $r \in \mathbb{R}$, there exists a permutation $\pi : \mathbb{N} \to \mathbb{N}$ such that
$$\sum_{k=1}^{\infty} a_{\pi(k)} = r \,.$$

Problem 9.23. The second Taylor polynomial for a twice-differentiable function f at $x = c$ is called the quadratic approximation of f at $x = c$. Find the quadratic approximate of the following functions at $x = 0$.

(1) $f(x) = \ln \cos x$ (2) $f(x) = e^{\sin x}$ (3) $f(x) = \tan x$ (4) $f(x) = \dfrac{1}{\sqrt{1 - x^2}}$

(5) $f(x) = e^x \sin^2 x$ (6) $f(x) = e^x \ln(1 + x)$ (7) $f(x) = (\arctan x)^2$

Problem 9.24. Let f have derivatives through order n at $x = c$. Show that the n-th Taylor polynomial for f at c and its first n derivatives have the same values that f and its first n derivatives have at $x = c$.

Problem 9.25. Complete the following.

(1) Let $f, g : [a, b] \to \mathbb{R}$ be continuous and g is sign-definite; that is, $g(x) \geqslant 0$ for all $x \in [a, b]$ or $g(x) \leqslant 0$ for all $x \in [a, b]$. Show that there exists $c \in [a, b]$ such that
$$f(c) \int_a^b g(x)\, dx = \int_a^b f(x)g(x)\, dx \,. \tag{9.11.2}$$

(2) Let $f : [a, b] \to \mathbb{R}$ be a function, and $c \in [a, b]$. Prove (by induction) that if f is $(n+1)$-times continuously differentiable on $[a, b]$, then for all $x \in [a, b]$,

$$\begin{aligned} f(x) &= f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \cdots + \frac{f^{(n)}(c)}{n!}(x-c)^n \\ &\quad + (-1)^n \int_c^x f^{(n+1)}(t) \frac{(t-x)^n}{n!}\, dt \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k + (-1)^n \int_c^x f^{(n+1)}(t) \frac{(t-x)^n}{n!}\, dt\,. \end{aligned}$$

(3) Use (9.11.2) to show that if f is $(n+1)$-times continuously differentiable on $[a, b]$ and $c \in [a, b]$, then for all $x \in [a, b]$ there exists a point ξ between x and c such that

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

(4) Find and explain the difference between the conclusion above and Taylor's Theorem.

Problem 9.26. Suppose that f is differentiable on an interval centered at $x = c$ and that $g(x) = b_0 + b_1(x-c) + \cdots + b_n(x-c)^n$ is a polynomial of degree n with constant coefficients $b_0, b_1, \cdots, b_n$. Let $E(x) = f(x) - g(x)$. Show that if we impose on g the conditions

1. $E(c) = 0$ (which means "the approximation error is zero at $x = c$");
2. $\lim\limits_{x\to c} \dfrac{E(x)}{(x-c)^n} = 0$ (which means "the error is negligible when compared to $(x-c)^n$),

then g is the n-th Taylor polynomial for f at c. **Thus, the Taylor polynomial P_n is the only polynomial of degree less than or equal to n whose error is both zero at $x = c$ and negligible when compared with $(x-c)^n$.**

Problem 9.27. Show that if p is an polynomial of degree n, then

$$p(x+1) = \sum_{k=0}^{n} \frac{p^{(k)}(x)}{k!}\,.$$

Problem 9.28. In Chapter 3 we considered Newton's method for approximating a root/zero r of the equation $f(x) = 0$, and from an initial approximation x_1 we obtained successive approximations $x_2, x_3, \cdots$, where

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \forall\, n \geqslant 1\,.$$

Show that if f'' exists on an interval I containing r, x_n, and x_{n+1}, and $\left|f''(x)\right| \leqslant M$ and $\left|f'(x)\right| \geqslant K$ for all $x \in I$, then

$$|x_{n+1} - r| \leqslant \frac{M}{2K}|x_n - r|^2$$

This means that if x_n is accurate to d decimal places, then x_{n+1} is accurate to about $2d$ decimal places. More precisely, if the error at stage n is at most 10^{-m}, then the error at stage $n+1$ is at most $\frac{M}{2K}10^{-2m}$.

Hint: Apply Taylor's Theorem to write $f(r) = P_2(r) + R_2(r)$, where P_2 is the second Taylor polynomial for f at x_n.

Problem 9.29. Consider a function f with continuous first and second derivatives at $x = c$. Prove that if f has a relative maximum at $x = c$, then the second Taylor polynomial centered at $x = c$ also has a relative maximum at $x = c$.

Problem 9.30. Suppose that $f : [a, b] \to \mathbb{R}$ is three times continuously differentiable, $h = \frac{b-a}{2}$ and $c = \frac{a+b}{2}$. Show that there exists $\xi \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{2h} - \frac{h^2}{6} f^{(3)}(\xi)\,.$$

Hint: Find the difference $f(b) - f(a)$ by write f as the sum of its third Taylor polynomial about c and the corresponding remainder. Apply the Intermediate Value Theorem to deal with the sum of the remainders. We note that the identity above implies that

$$\left|f'(c) - \frac{f(c+h) - f(c-h)}{2h}\right| \leqslant \frac{h^2}{6} \max_{x\in[c-h,c+h]} \left|f^{(3)}(x)\right|.$$

Problem 9.31. Suppose that $f : [a, b] \to \mathbb{R}$ is four times continuously differentiable, $h = \frac{b-a}{2}$ and $c = \frac{a+b}{2}$. Show that there exists $\xi \in (a, b)$ such that

$$f''(c) = \frac{f(a) - 2f(c) + f(b)}{h^2} - \frac{f^{(4)}(\xi)}{12}h^2\,. \tag{9.11.3}$$

Hint: Find the sum $f(a) + f(b)$ by write f as the sum of its third Taylor polynomial about c and the corresponding remainder. Apply the Intermediate Value Theorem to deal with the sum of the remainders. We note that the identity above implies that

$$\left|f''(c) - \frac{f(c+h) - 2f(c) + f(c-h)}{h^2}\right| \leqslant \frac{h^2}{12} \max_{x\in[c-h,c+h]} \left|f^{(4)}(x)\right|.$$

Problem 9.32. Suppose that $f : [a, b] \to \mathbb{R}$ is four times continuously differentiable. Show that

$$\Big|\int_a^b f(x)\,dx - \frac{b-a}{6}\big[f(a) + 4f\big(\frac{a+b}{2}\big) + f(b)\big]\Big| \leqslant \frac{2h^5}{45}\max_{x\in[a,b]}\big|f^{(4)}(x)\big| \tag{9.11.4}$$

through the following steps.

1. Let $c = \dfrac{a+b}{2}$ and $h = \dfrac{b-a}{2}$. Write f as the sum of its third Taylor polynomial about c and the corresponding remainder and conclude that
$$\int_a^b f(x)\,dx = 2hf(c) + \frac{h^3}{3}f''(c) + \int_a^b R_3(x)\,dx\,.$$
2. Show (by Intermediate Value Theorem) that there exists $\xi \in (a, b)$ such that
$$\int_a^b R_3(x)\,dx = \frac{f^{(4)}(\xi)}{24}\int_a^b (x-c)^4\,dx = \frac{f^{(4)}(\xi)}{60}h^5\,. \tag{9.11.5}$$
3. Use (9.11.3) in (9.11.5) and conclude (9.11.4).

Problem 9.33. Find the interval of convergence of the following power series.

(1) $\sum\limits_{n=1}^{\infty}\big(1+\frac{1}{n}\big)^n x^n$ (2) $\sum\limits_{n=1}^{\infty}(\ln n)x^n$ (3) $\sum\limits_{n=1}^{\infty}\big(\sqrt{n+1}-\sqrt{n}\big)x^n$ (4) $\sum\limits_{n=1}^{\infty}\big(\frac{n}{n+1}\big)^{n^2}x^n$

(5) $\sum\limits_{n=1}^{\infty}\frac{n!}{(2n)!}x^n$ (6) $\sum\limits_{n=1}^{\infty}\frac{2\cdot4\cdot6\cdot\cdots\cdot(2n)}{3\cdot5\cdot7\cdot\cdots\cdot(2n+1)}x^{2n+1}$ (7) $\sum\limits_{n=1}^{\infty}\frac{(-1)^n 3\cdot7\cdot11\cdot\cdots\cdot(4n-1)}{4^n}x^n$

(8) $\sum\limits_{n=1}^{\infty}\frac{1}{2\cdot4\cdot6\cdot\cdots\cdot(2n)}x^n$ (10) $\sum\limits_{n=1}^{\infty}\frac{1}{1\cdot3\cdot5\cdot7\cdot\cdots\cdot(2n-1)}x^n$ (9) $\sum\limits_{n=1}^{\infty}\frac{n!}{3\cdot6\cdot9\cdot\cdots\cdot(3n)}x^n$

(10) $\sum\limits_{n=1}^{\infty}\frac{k(k+1)(k+2)\cdots(k+n-1)}{n!}x^n$, where k is a positive integer;

(11) $\sum\limits_{n=0}^{\infty}\frac{(n!)^k}{(kn)!}x^n$, where k is a positive integer; (12) $\sum\limits_{n=2}^{\infty}\frac{x^n}{n\ln n}$ (13) $\sum\limits_{n=2}^{\infty}\frac{x^n}{n(\ln n)^2}$

(14) $\sum\limits_{n=1}^{\infty}\big[2+(-1)^n\big](x+1)^{n-1}$

Problem 9.34. The function J_0 defined by

$$J_0(x) = \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{2^{2n}(n!)^2}$$

is called the Bessel function of the first kind of order 0. Find its domain (that is, the interval of convergence).

Problem 9.35. The function J_1 defined by

$$J_1(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!(n+1)!2^{2n+1}}$$

is called the Bessel function of the first kind of order 1. Find its domain (that is, the interval of convergence).

Problem 9.36. The function A defined by

$$A(x) = 1 + \frac{x^3}{2 \cdot 3} + \frac{x^6}{2 \cdot 3 \cdot 5 \cdot 6} + \frac{x^9}{2 \cdot 3 \cdot 5 \cdot 6 \cdot 8 \cdot 9} + \cdots$$

is called an Airy function after the English mathematician and astronomer Sir George Airy (1801–1892). Find the domain of the Airy function.

Problem 9.37. A function f is defined by

$$f(x) = 1 + 2x + x^2 + 2x^3 + x^4 + \cdots ;$$

that is, its coefficients are $c_{2n} = 1$ and $c_{2n+1} = 2$ for all $n \geqslant 0$. Find the interval of convergence of the series and find an explicit formula for $f(x)$.

Problem 9.38. Let $f : (-r, r) \to \mathbb{R}$ be n-times differentiable at 0, and $P_n(x)$ be the n-th Maclaurin polynomial for f.

1. Show that if $g(x) = x^\ell f(x^m)$ for some positive integers m and ℓ, then the $(mn+\ell)$-th Maclaurin polynomial for g is $x^\ell P_n(x^m)$.

2. Show that if $h(x) = x^\ell f(-x^m)$ for some positive integers m and ℓ, then the $(mn+\ell)$-th Maclaurin polynomial for h is $x^\ell P_n(-x^m)$.

3. Find the Maclaurin series for the following functions:

$$(1)\ y = \frac{1}{1+x^2} \quad (2)\ y = x^2 \arctan(x^3) \quad (3)\ y = \ln(1+x^4) \quad (4)\ y = x\sin(x^3)\cos(x^3).$$

Hint for (1) and (2): See Exercise 3 Problem 4.

Problem 9.39. To find the sum of the series $\sum\limits_{n=1}^{\infty} \frac{n^2}{2^n}$, express $\frac{1}{1-x}$ as a geometric series, differentiate both sides of the resulting equation with respect to x, multiply both sides of the result by x, differentiate again, multiply by x again, and set x equal to $\frac{1}{2}$. What do you get?

Problem 9.40. Complete the following.

(1) Use the power series of $y = \arctan x$ to show that

$$\pi = 2\sqrt{3}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)3^n}$$

(2) Using $x^3+1 = (x+1)(x^2-x+1)$, rewrite the integral $\displaystyle\int_0^{\frac{1}{2}}\frac{dx}{x^2-x+1}$ and then express $\dfrac{1}{1+x^3}$ as the sum of a power series to prove the following formula for π:

$$\pi = \frac{3\sqrt{3}}{4}\sum_{n=0}^{\infty}\frac{(-1)^n}{8^n}\Big(\frac{2}{3n+1}+\frac{1}{3n+2}\Big).$$

Problem 9.41. Show that the Bessel function of the first kind of order 0, denoted by J_0 and defined by

$$J_0(x) = \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{2^{2n}(n!)^2},$$

satisfies the differential equation

$$x^2y''(x) + xy'(x) + x^2y(x) = 0, \qquad y(0) = 1,\ y'(0) = 0.$$

Problem 9.42. Show that the Bessel function of the first kind of order 1, denoted by J_1 and defined by

$$J_1(x) = \sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{n!(n+1)!2^{2n+1}},$$

satisfies the differential equation

$$x^2y''(x) + xy'(x) + (x^2-1)y(x) = 0, \qquad y(0) = 0,\ y'(0) = \frac{1}{2}.$$

Problem 9.43. Suppose that $x_1(t)$ and $x_2(t)$ are functions of t satisfying the following equations

$$\begin{aligned} x_1''(t) - x_1(t) &= 0, \qquad x_1(0) = 1, \quad x_1'(0) = 0, \\ x_2''(t) - x_2(t) &= 0, \qquad x_2(0) = 0, \quad x_2'(0) = 1, \end{aligned}$$

where $'$ denotes the derivatives with respect to t.

1. Assume that the function $x_1(t)$ and $x_2(t)$ can be written as a power series (on a certain interval), that is, $x_1(t) = \sum\limits_{k=0}^{\infty} a_k t^k$ and $x_2(t) = \sum\limits_{k=0}^{\infty} b_k t^k$. Show that

$$(k+2)(k+1)a_{k+2} = a_k \quad \text{and} \quad (k+2)(k+1)b_{k+2} = b_k \qquad \forall\, k \geqslant 0\,.$$

2. Find a_k and b_k, and conclude that x_1 and x_2 are some functions that we have seen before.

3. Find a function $x(t)$ satisfying

$$x''(t) - x(t) = 0\,, \qquad x(0) = a\,, \quad x'(0) = b\,.$$

Note that x can be written as the linear combination of x_1 and x_2.

Problem 9.44. Find the series solution to the differential equation

$$y''(x) + x^2 y(x) = 0\,, \qquad y(0) = 1\,, \quad y'(0) = 0\,.$$

What is the radius of convergence of this series solution?

Problem 9.45. In this problem we try to establish the following theorem

Theorem 9.112

Let the radius of convergence of the power series $f(x) = \sum\limits_{k=0}^{\infty} a_k (x-c)^k$ be r for some $r > 0$.

1. If $\sum\limits_{k=0}^{\infty} a_k r^k$ converges, then f is continuous at $c+r$; that is $\lim\limits_{x\to(c+r)^-} f(x) = f(c+r)$.

2. If $\sum\limits_{k=0}^{\infty} a_k (-r)^k$ converges, then f is continuous at $c-r$; that is, $\lim\limits_{x\to(c-r)^+} f(x) = f(c-r)$.

Prove case 1 of the theorem above through the following steps.

(1) Let $A = \sum\limits_{k=0}^{\infty} a_k r^k$, and define

$$g(x) = f(rx+c) - A = -\sum_{k=1}^{\infty} a_k r^k + \sum_{k=1}^{\infty} a_k r^k x^k = \sum_{k=0}^{\infty} b_k x^k\,,$$

where $b_k = a_k r^k$ for each $k \in \mathbb{N}$ and $b_0 = -\sum_{k=1}^{\infty} a_k r^k$. Show that the radius of convergence of g is 1 and $\sum_{k=0}^{\infty} b_k = 0$. Moreover, show that f is continuous at $c + r$ if and only if g is continuous at 1.

(2) Let $s_n = b_0 + b_1 + \cdots + b_n$ and $S_n(x) = b_0 + b_1 x + \cdots + b_n x^n$. Show that

$$S_n(x) = (1 - x)(s_0 + s_1 x + \cdots + s_{n-1} x^{n-1}) + s_n x^n$$

and conclude that

$$g(x) = \lim_{n\to\infty} S_n(x) = (1 - x) \sum_{k=0}^{\infty} s_k x^k \,. \tag{9.11.6}$$

(3) Use (9.11.6) to show that g is continuous at 1. Note that you might need to use ε-δ argument.

Problem 9.46. Show that $\int_0^1 x^{-x}\, dx = \sum_{k=1}^{\infty} \frac{1}{k^k}$.

Hint: Write $x^{-x} = e^{-x \ln x}$ and use the Maclaurin series for exp to conclude that

$$\int_0^1 x^{-x}\, dx = \int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k (x \ln x)^k}{k!}\, dx \,.$$

Use the fact that $\int_0^1 \sum_{k=0}^{\infty} \frac{(-1)^k (x \ln x)^k}{k!}\, dx = \sum_{k=0}^{\infty} \int_0^1 \frac{(-1)^k (x \ln x)^k}{k!}\, dx$. You will also need to recall the Gamma function.

Problem 9.47. Show that $\int_0^1 \frac{\ln x \ln(1 + x)}{x}\, dx = \sum_{k=1}^{\infty} \frac{(-1)^k}{k^3}$.

Hint: Use (9.7.8) and rewrite the integral above as $\int_0^1 \ln x \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^{k-1}}{k}\, dx$. Assume that you know that

$$\int_0^1 \ln x \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^{k-1}}{k}\, dx = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \int_0^1 x^{k-1} \ln x\, dx \,.$$

Problem 9.48. Let $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ be sequence of real numbers such that the series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ both converge. Define $c_k = \sum_{j=0}^{k} a_j b_{k-j}$ and $C_n = \sum_{i=0}^{n} c_i$.

1. Show that if $\sum_{n=0}^{\infty} a_n$ converges absolutely, then

$$\lim_{n\to\infty} C_n = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big) \tag{9.11.7}$$

by completing the following.

   (a) Show that $C_n = \sum_{k=0}^{n} a_{n-k}B_k$, where $B_k = \sum_{i=0}^{k} b_i$ is the k-th partial sum of the series $\sum_{i=0}^{\infty} b_i$.

   (b) Let $A_k = \sum_{i=0}^{k} a_i$ be the k-th partial sum of the series $\sum_{i=0}^{\infty} a_i$, and $A = \lim_{n\to\infty} A_n$, $B = \sum_{n\to\infty} B_n$. Then

$$C_n - AB = \sum_{k=0}^{n} a_{n-k}(B_k - B) + (A_n - A)B\,.$$

   Use the ε-N argument to show that $\lim_{n\to\infty} C_n = AB$.

2. $\sum_{n=0}^{\infty} c_n$ is called the Cauchy product of the series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$. Show that (9.11.7) may fail if both $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ converges conditionally by looking at the example $a_n = b_n = \dfrac{(-1)^n}{\sqrt{n+1}}$ for all $n \in \mathbb{N}$.

Chapter 10

Vectors and the Geometry of Space

10.1 Preliminaries

In this section we review some of the materials from the high school (or even linear algebra). First we consider vectors in the plane. We let $\mathbf{i}$ (or $\mathbf{e}_1$) and $\mathbf{j}$ (or $\mathbf{e}_2$) denote the vectors $(1,0)$ and $(0,1)$, respectively. Any vectors $\mathbf{v}$ in the plane can be written as $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$. For two vectors $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$, the dot product of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} \cdot \mathbf{v}$, is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 = \sum_{j=1}^{2} u_j v_j \,.$$

Let θ denote the angle between two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$. The law of cosines then implies that

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta \,,$$

where $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2}$ and $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}$ denote the length of vectors $\mathbf{u}$ and $\mathbf{v}$, respectively.

Similar ideas can be extended for vectors in space. Let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ denote the vectors

$$\mathbf{i} = (1,0,0) \equiv \mathbf{e}_1 \,, \quad \mathbf{j} = (0,1,0) \equiv \mathbf{e}_2 \quad \text{and} \quad \mathbf{k} = (0,0,1) \equiv \mathbf{e}_3 \,.$$

The standard unit vector notation for a vector $\mathbf{v}$ in space is

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3 = \sum_{j=1}^{3} v_j \mathbf{e}_j \,.$$

For two vectors $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$, the dot product of $\mathbf{u}$ and $\mathbf{v}$, again denoted by $\mathbf{u} \cdot \mathbf{v}$, is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 = \sum_{j=1}^{3} u_jv_j\,.$$

If θ denote the angle between $\mathbf{u}$ and $\mathbf{v}$ when $\mathbf{u}, \mathbf{v}$ are non-zero vectors, then the law of cosines also implies that

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \cos\theta\,, \tag{10.1.1}$$

where $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2}$ and $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$ again denote the length of vectors $\mathbf{u}$ and $\mathbf{v}$, respectively.

10.2 The Cross Product of Two Vectors in Space

Definition 10.1

Let $\mathbf{u} = u_1\mathbf{i}+u_2\mathbf{j}+u_3\mathbf{k}$ and $\mathbf{v} = v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k}$ be vectors in space. The cross product of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u} \times \mathbf{v}$, is the vector

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}\,.$$

Remark 10.2. Using the notation of determinant, the cross product of $\mathbf{u}$ and $\mathbf{v}$ can be computed as

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}\,.$$

Remark 10.3. A sequence $(k_1, k_2, \cdots, k_n)$ of positive integers not exceeding n, with the property that no two of the k_i are equal, is called a ***permutation of degree*** n. The collection of all permutations of degree n is denoted by $\mathbb{P}(n)$. For $1 \leqslant i, j \leqslant n$ and $i \neq j$, the operator $\tau_{(i,j)}$ interchange the i-th and j-th elements of a sequence in $\mathbb{P}(n)$. For example, if n $= 3$, the permutation $(3, 1, 2)$ can be obtained by interchanging pairs of $(1, 2, 3)$ twice:

$$(1, 2, 3) \xrightarrow{\tau_{(1,3)}} (3, 2, 1) \xrightarrow{\tau_{(2,3)}} (3, 1, 2);$$

thus $(3, 1, 2)$ is called an even permutation of $(1, 2, 3)$. On the other hand, $(1, 3, 2)$ is obtained by interchanging pairs of $(1, 2, 3)$ once:

$$(1, 2, 3) \xrightarrow{\tau_{(2,3)}} (1, 3, 2);$$

thus $(1,3,2)$ is an odd permutation of $(1,2,3)$.

For $n=3$, the even and odd permutations can also be viewed as the orientation of the permutation (k_1,k_2,k_3). To be more precise, if $(1,2,3)$ is arranged in a counter-clockwise orientation (see Figure 10.1), then an even permutation of degree 3 is a permutation in the counter-clockwise orientation, while an odd permutation of degree 3 is a permutation in the clockwise orientation. From figure 10.1, it is easy to see that $(3,1,2)$ is an even permutation of degree 3 and $(1,3,2)$ is an odd permutation of degree 3.

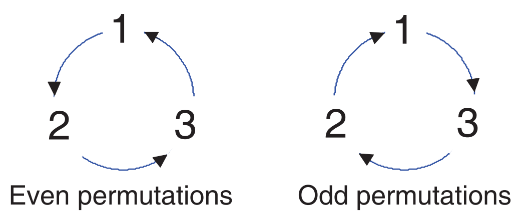


Figure 10.1: Even and odd permutations of degree 3

The permutation symbol is a function on $\mathbb{P}(n)$ defined by

$$\varepsilon_{k_1k_2\cdots k_n} = \begin{cases} 1 & \text{if } (k_1,k_2,\cdots,k_n) \text{ is an even permutation of } (1,2,\cdots,n)\,, \\ -1 & \text{if } (k_1,k_2,\cdots,k_n) \text{ is an odd permutation of } (1,2,\cdots,n)\,. \end{cases}$$

In general, one can define

$$\varepsilon_{k_1k_2\cdots k_n} = \begin{cases} 1 & \text{if } (k_1,k_2,\cdots,k_n) \text{ is an even permutation of } (1,2,\cdots,n)\,, \\ -1 & \text{if } (k_1,k_2,\cdots,k_n) \text{ is an odd permutation of } (1,2,\cdots,n)\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Using the permutation symbol, we have

$$\mathbf{u}\times\mathbf{v} = \sum_{i,j,k=1}^{3} \varepsilon_{ijk}u_jv_k\mathbf{e}_i = \sum_{i=1}^{3}\Big(\sum_{j,k=1}^{3}\varepsilon_{ijk}u_jv_k\Big)\mathbf{e}_i\,, \tag{10.2.1}$$

where $\mathbf{u}=(u_1,u_2,u_3)$ and $\mathbf{v}=(v_1,v_2,v_3)$. In other words, the i-th component of $\mathbf{u}\times\mathbf{v}$ is $\sum\limits_{j,k=1}^{3}\varepsilon_{ijk}u_jv_k$.

In the following, for simplicity we let $(\mathbf{u}\times\mathbf{v})_i$ denote the i-th component of the vector $\mathbf{u}\times\mathbf{v}$. In other words,

$$(\mathbf{u}\times\mathbf{v})_i = \sum_{j,k=1}^{3}\varepsilon_{ijk}u_jv_k\,.$$

Theorem 10.4

Let $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ be vectors in space, and c be a scalar.

(a) $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$.

(b) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$.

(c) $c(\mathbf{u} \times \mathbf{v}) = (c\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (c\mathbf{v})$.

(d) $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$.

(e) $\mathbf{u} \times \mathbf{u} = \mathbf{0}$.

(f) $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$.

We note that (b) and (c) can be simplified as

$$\mathbf{u} \times (c\mathbf{v} + d\mathbf{w}) = c(\mathbf{u} \times \mathbf{v}) + d(\mathbf{u} \times \mathbf{w}) \qquad \forall \text{ vectors in space } \mathbf{u}, \mathbf{v}, \mathbf{w} \text{ and scalars } c, d.$$

Proof of Theorem 10.4. We provide two proofs for (f), and the others are left as exercise.

1. Since $\mathbf{v} \times \mathbf{w} = (v_2w_3 - v_3w_2)\mathbf{i} + (v_3w_1 - v_1w_3)\mathbf{j} + (v_1w_2 - v_2w_1)\mathbf{k}$ and $\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}$, we find that
$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= u_1(v_2w_3 - v_3w_2) + u_2(v_3w_1 - v_1w_3) + u_3(v_1w_2 - v_2w_1) \\ &= w_1(u_2v_3 - u_3v_2) + w_2(u_3v_1 - u_1v_3) + w_3(u_1v_2 - u_2v_1) \\ &= \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}\,. \end{aligned}$$

2. Using (10.2.1) and the fact that $\varepsilon_{ijk} = \varepsilon_{kij}$,
$$\begin{aligned} \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) &= \sum_{i=1}^{3} u_i \sum_{j,k=1}^{3} \varepsilon_{ijk} v_j w_k = \sum_{i,j,k=1}^{3} \varepsilon_{ijk} u_i v_j w_k = \sum_{k=1}^{3} w_k \sum_{i,j=1}^{3} \varepsilon_{kij} u_i v_j \\ &= \sum_{k=1}^{3} w_k (\mathbf{u} \times \mathbf{v})_k = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v})\,. \end{aligned}$$ □

Lemma 10.5

Let δ_{ij} be the Kronecker delta defined by $\delta_{ij} = \begin{cases} 1 & \text{if } i = j\,, \\ 0 & \text{if } i \neq j\,. \end{cases}$ Then

$$\sum_{i=1}^{3} \varepsilon_{ijk} \varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\,. \tag{10.2.2}$$

Theorem 10.6: Geometric properties of the cross product

Let $\mathbf{u}$ and $\mathbf{v}$ be non-zero vectors in space, and let θ be the angle between $\mathbf{u}$ and $\mathbf{v}$.

(a) $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

(b) $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\| \sin\theta$.

(c) $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are scalar multiples of each other.

(d) $\|\mathbf{u} \times \mathbf{v}\|$ is the area of parallelogram having $\mathbf{u}$ and $\mathbf{v}$ as adjacent sides.

Proof. We only prove (b). Using (10.2.2),

$$\begin{aligned}
\|\mathbf{u} \times \mathbf{v}\|^2 = (\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{u} \times \mathbf{v}) &= \sum_{i=1}^{3} (\mathbf{u} \times \mathbf{v})_i (\mathbf{u} \times \mathbf{v})_i = \sum_{i=1}^{3} \Big(\sum_{j,k=1}^{3} \varepsilon_{ijk} u_j v_k \Big) \Big(\sum_{r,s=1}^{3} \varepsilon_{irs} u_r v_s \Big) \\
&= \sum_{i,j,k,r,s=1}^{3} \varepsilon_{ijk} \varepsilon_{irs} u_j v_k u_r v_s = \sum_{j,k,r,s=1}^{n} \Big(\sum_{i=1}^{3} \varepsilon_{ijk} \varepsilon_{irs} \Big) u_j v_k u_r v_s \\
&= \sum_{j,k,r,s=1}^{3} (\delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}) u_j v_k u_r v_s = \sum_{j,k=1}^{3} \big[u_j^2 v_k^2 - (u_j v_j)(u_k v_k) \big] \\
&= \Big(\sum_{j=1}^{3} u_j^2 \Big) \Big(\sum_{k=1}^{3} v_k^2 \Big) - \Big(\sum_{j=1}^{3} u_j v_j \Big) \Big(\sum_{k=1}^{3} u_k v_k \Big) = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - |\mathbf{u} \cdot \mathbf{v}|^2 .
\end{aligned}$$

Using (10.1.1), we find that

$$\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - |\mathbf{u} \cdot \mathbf{v}|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \cos^2\theta = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2\theta$$

which concludes (b). □

Definition 10.7: Triple Scalar Product

For vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ in space, the dot product of $\mathbf{u}$ and $\mathbf{v} \times \mathbf{w}$ is called the triple scalar product (of $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$).

Theorem 10.8

For $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$, the triple scalar product $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})$ is

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} .$$

Theorem 10.9

The volume V of a parallelepiped with vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ as adjacent edges is

$$V = \left|(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}\right| = \left|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})\right|.$$

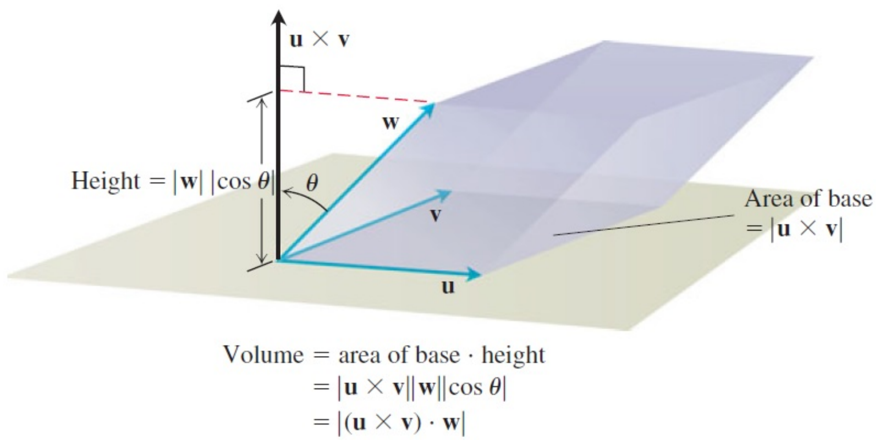


Figure 10.2: The number $|(\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}|$ is the volume of a parallelepiped.

10.2.1 Alternative definition of the cross product

We start with two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in space. If $\mathbf{u}$ and $\mathbf{v}$ are not parallel, they determine a plane. We select a unit vector $\boldsymbol{n}$ perpendicular to the plane by the right-hand rule; that is, the unit normal vector $\boldsymbol{n}$ points the way your right thumb points when your fingers curl through the angle θ from $\mathbf{u}$ to $\mathbf{v}$ (figure 10.3).

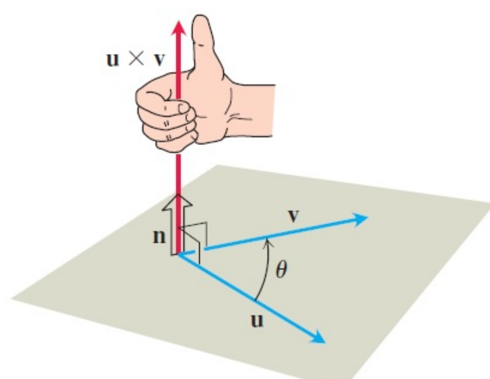


Figure 10.3: The construction of $\mathbf{u} \times \mathbf{v}$

Then we define a new vector as follows.

Definition 10.10

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in space, θ be the angle between $\mathbf{u}$ and $\mathbf{v}$, and $\boldsymbol{n}$ be a unit vector defined by the right-hand rule. The cross product $\mathbf{u} \times \mathbf{v}$ is the vector

$$\mathbf{u} \times \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}\| \sin\theta \, \boldsymbol{n} \, .$$

We note that if $\mathbf{u}$ and $\mathbf{v}$ are parallel, then $\boldsymbol{n}$ is not well-defined; however, in this case $\theta = 0$ or π so that $\sin\theta = 0$; thus the definition above suggests that $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ if $\mathbf{u}$ and $\mathbf{v}$ are parallel. This is indeed the case we should have in mind.

Using this definition of the cross product, properties (a)(c)(d)(e) in Theorem 10.4 clearly hold. For example, property (a) can be visualized by the following figure

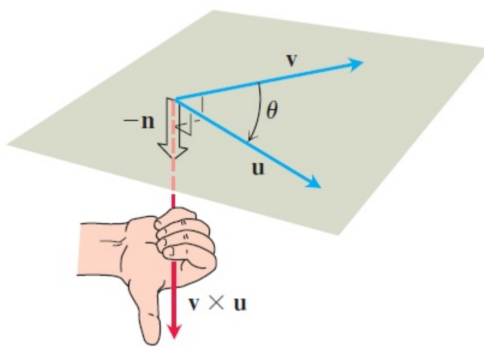


Figure 10.4: The construction of $\mathbf{u} \times \mathbf{v}$

In the following, we prove (b) in Theorem 10.4 under this alternative definition of cross product. To derive (b), we construct $\mathbf{u} \times \mathbf{v}$ in a new way (see Figure 10.5 for reference).

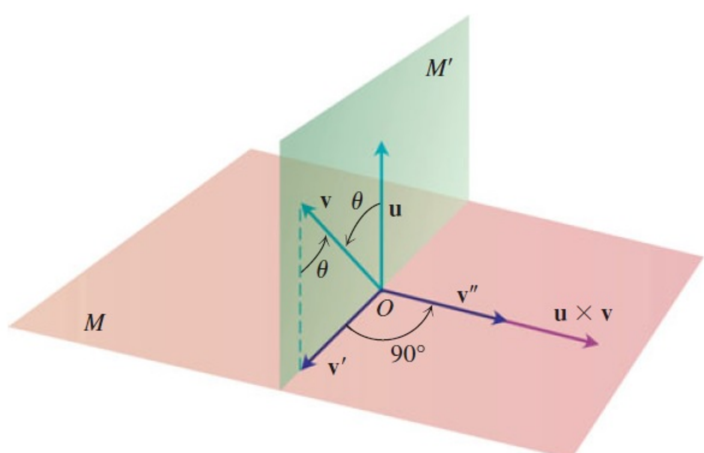


Figure 10.5: As explained in the text, $\mathbf{u} \times \mathbf{v} = \|\mathbf{u}\|\|\mathbf{v}''\|$. (The primes used here are purely notational and do not denote derivatives.)

We draw **u** and **v** from the common point O and construct a plane M perpendicular to **u** at O. We then project **v** orthogonally onto M, yielding a vector $\mathbf{v}'$ with length $\|\mathbf{v}\| \sin\theta$. We rotate $\mathbf{v}'$ $90°$ about **u** in the positive sense to produce a vector $\mathbf{v}''$. Finally, we multiply $\mathbf{v}''$ by the length of **u**. The resulting vector $\|\mathbf{u}\|\mathbf{v}''$ is equal to $\mathbf{u} \times \mathbf{v}$ since $\mathbf{v}''$ has the same direction as $\mathbf{u} \times \mathbf{v}$ by its construction and

$$\|\mathbf{u}\|\|\mathbf{v}''\| = \|\mathbf{u}\|\|\mathbf{v}'\| = \|\mathbf{u}\|\|\mathbf{v}\| \sin\theta = \|\mathbf{u} \times \mathbf{v}\| .$$

Now each of these three operations, namely,

1. projection onto M,
2. rotation about u through $90°$,
3. multiplication by the scalar $\|\mathbf{u}\|$,

when applied to a triangle whose plane is not parallel to **u**, will produce another triangle. If we start with the triangle whose sides are **v**, **w**, and $\mathbf{v} + \mathbf{w}$ (Figure 10.6) and apply these three steps, we successively obtain the following:

1. A triangle whose sides are $\mathbf{v}'$, $\mathbf{w}'$, and $(\mathbf{v} + \mathbf{w})'$ satisfying the vector equation
$$\mathbf{v}' + \mathbf{w}' = (\mathbf{v} + \mathbf{w})' .$$
2. A triangle whose sides are $\mathbf{v}''$, $\mathbf{w}''$, and $(\mathbf{v} + \mathbf{w})''$ satisfying the vector equation
$$\mathbf{v}'' + \mathbf{w}'' = (\mathbf{v} + \mathbf{w})'' .$$
3. A triangle whose sides are $\|\mathbf{u}\|\mathbf{v}''$, $\|\mathbf{u}\|\mathbf{w}''$, and $\|\mathbf{u}\|(\mathbf{v}+\mathbf{w})''$ satisfying the vector equation
$$\|\mathbf{u}\|\mathbf{v}'' + \|\mathbf{u}\|\mathbf{w}'' = \|\mathbf{u}\|(\mathbf{v} + \mathbf{w})'' .$$

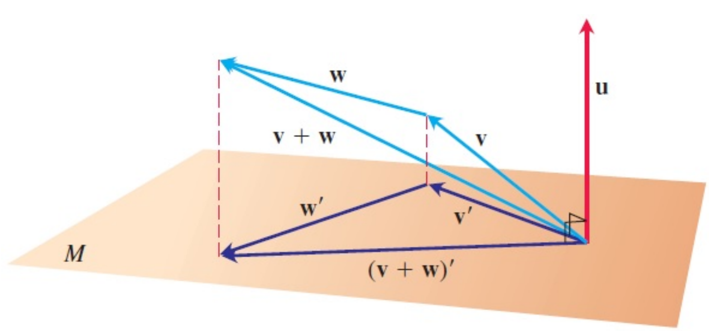


Figure 10.6: The vectors, **v**, **w**, $\mathbf{v} + \mathbf{w}$, and their projections onto a plane perpendicular to **u**.

Substituting $\|\mathbf{u}\|\mathbf{v}'' = \mathbf{u} \times \mathbf{v}$, $\|\mathbf{u}\|\mathbf{w}'' = \mathbf{u} \times \mathbf{w}$, and $\|\mathbf{u}\|(\mathbf{v} + \mathbf{w})'' = \mathbf{u} \times (\mathbf{v} + \mathbf{w})$ from our discussion above into this last equation gives $\mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w} = \mathbf{u} \times (\mathbf{v} + \mathbf{w})$, which is the law we wanted to establish.

When we apply the definition to calculate the pairwise cross products of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, we find that $\mathbf{i} \times \mathbf{j} = \mathbf{k}$, $\mathbf{j} \times \mathbf{k} = \mathbf{i}$ and $\mathbf{k} \times \mathbf{i} = \mathbf{j}$.

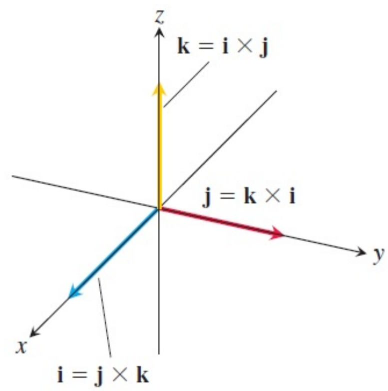


Figure 10.7: The pairwise cross products of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$.

Having establishing (b) in Theorem 10.4 under the alternative definition of cross product, we are able to derive the formula for cross product in Definition 10.1:

$$\begin{aligned}
\mathbf{u} \times \mathbf{v} &= (u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}) \times (v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) \\
&= u_1v_2(\mathbf{i} \times \mathbf{j}) + u_1v_3(\mathbf{i} \times \mathbf{k}) + u_2v_1(\mathbf{j} \times \mathbf{i}) + u_2v_3(\mathbf{j} \times \mathbf{k}) + u_3v_1(\mathbf{k} \times \mathbf{i}) + u_3v_2(\mathbf{k} \times \mathbf{j}) \\
&= (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}\,.
\end{aligned}$$

10.3 Polar Coordinate

In this section we review the polar coordinate (on the plane) that we introduction in Remark 0.8 and make some extensions. To form the polar coordinate system in the plane, fix a point O, called the pole (or origin), and construct from O an initial ray called the polar axis, as shown in Figure 10.8.

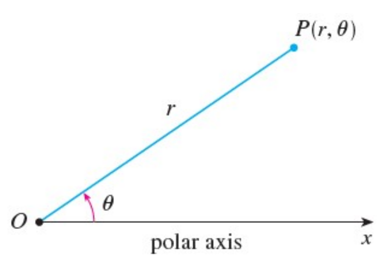


Figure 10.8: Polar coordinate

Then each point P in the plane can be assigned polar coordinates (r, θ), also called the polar representation of P, as follows.

$$\begin{aligned} r &= \text{distance from } O \text{ to } P, \\ \theta &= \text{angle (in } [0, 2\pi)) \text{ measured counterclockwise from polar axis to segment } \overline{OP}. \end{aligned}$$

Let the polar axis as the positive x-axis on the plane (that is, let $\mathbf{i}$ or $\mathbf{e}_1$ denote the unit vector pointing in the direction of the polar axis), and $\mathbf{j}$ or $\mathbf{e}_2$ be the unique unit vector in the plane obtained by rotating $\mathbf{i}$ counterclockwise by angle $\frac{\pi}{2}$. Then every point P in the plane can be expressed as an ordered pair (x, y) in the way that the vector $\overrightarrow{OP}$ can be expressed as $x\mathbf{e}_1 + y\mathbf{e}_2$. In other words, (x, y) is the Cartesian coordinate of P with $\mathbf{e}_1$ and $\mathbf{e}_2$ being the unit vectors on the x-axis and y-axis of the plane. If $\mathcal{P} \neq O$, and (x, y), (r, θ) are the Cartesian and polar coordinate of P, respectively, then we have

$$x = r\cos\theta\,, \qquad y = r\sin\theta\,, \tag{10.3.1a}$$

$$r = \sqrt{x^2 + y^2}\,, \qquad \theta = \begin{cases} \arctan\dfrac{y}{x} & \text{if } x > 0\,, \\ \dfrac{\pi}{2} & \text{if } x = 0 \text{ and } y > 0\,, \\ \pi + \arctan\dfrac{y}{x} & \text{if } x < 0\,, \\ \dfrac{3\pi}{2} & \text{if } x = 0 \text{ and } y < 0\,. \end{cases} \tag{10.3.1b}$$

(10.3.1a) is sometimes called the **polar-to-rectangular** and (10.3.1b) is sometimes called the **rectangular-to-polar** (coordinate) conversion. Note that the polar coordinate gives an one-to-one correspondence between the region $(0, \infty) \times [0, 2\pi)$ and the plane with the origin removed.

Remark 10.11. Often time we use the region $[0, \infty) \times [0, 2\pi]$ on the $r\theta$-plane to denote the set to which (r, θ) belongs. The segment $\{0\} \times [0, 2\pi]$ is treated as the origin (of the xy-plane), while the ray $[0, \infty) \times \{0\}$ and $[0, \infty) \times \{2\pi\}$ both represent x-axis.

Such as some rectangular regions can be easily represented using the Cartesian coordinate (for example, $[a, b] \times [c, d]$ represents a rectangle), some special regions in the plane can be easily represented using the polar coordinate.

Example 10.12. The sector enclosed by the circle with radius r_0 and two radii $\theta = \theta_0$ and $\theta = \theta_1$ can be expressed as $(r, \theta) \in [0, r_0] \times [\theta_0, \theta_1]$.

Curves in the region $[0, \infty) \times [0, 2\pi]$ of the $r\theta$-plane corresponds to curves in xy-plane through relation (10.3.1a). For example, the line segment $\{1\} \times [0, 2\pi]$ (or simply $r = 1$) corresponds to the unit circle centered at the origin, and the ray $[0, \infty) \times \{\theta_0\}$ (or simply $\theta = \theta_0$) corresponds to the ray to which the angle measured from the polar axis is θ_0.

Example 10.13. The curve $r = \cos\theta$ in the region $[0, \infty) \times [0, 2\pi]$ corresponds to the circle $x^2 + y^2 = x$ in the xy-plane.

As we did not distinguish the angle 0 and 2π, we should not distinguish any θ with all $\theta + 2k\pi$ ($k \in \mathbb{Z}$). In general, for a given point $P = (x, y)$ in Cartesian coordinate system, we should treat (r, θ) as the polar coordinate of P as long as (r, θ) satisfies (10.3.1a). This includes the possibility that r is negative since

$$(r\cos\theta, r\sin\theta) = \big((-r)\cos(\theta + \pi), (-r)\sin(\theta + \pi)\big)$$

which means if (r, θ) is a polar representation of P, then $(-r, \theta + \pi)$ is also a polar representation of P.

To be more precise, the polar coordinate (r, θ) of a point P satisfies

$r =$ "directed" distance from O to P,

$\theta =$ "directed" angle measured counterclockwise from polar axis to segment OP.

We note under this convention, each point have infinitely many polar representation.

Remark 10.14. 想像你身處原點，然後你的前方是 x 軸的正方向，而座標軸上有標記單位。現在在你前方放一面鏡子，而有另一個人出現在你的後方立於座標軸上的 -2 這個位置。你所看到的是，在你的「前方」有一個位置在 -2 的人，所以你很快速地標記他的極座標為 $(-2, 0)$。在此 -2 即為所謂的 directed distance 而 0 是 directed angle。directed distance 的正負號取決於你要不要在你觀察的那個 θ 方向加一面鏡子。

From now on, the polar coordinate, given the pole and the polar axis, refers to this non-unique polar representation of points in the plane.

Theorem 10.15

The polar coordinates (r, θ) of a point are relation to the Cartesian coordinates (x, y) of the point as follows.

Polar-to-Rectangular	**Rectangular-to-Polar**
$x = r\cos\theta$	$\tan\theta = \dfrac{y}{x}$
$y = r\sin\theta$	$r^2 = x^2 + y^2$

10.4 Cylindrical and Spherical Coordinates

10.4.1 The cylindrical coordinate

Definition 10.16

In a cylindrical coordinate system, a point P in space is presented by an ordered triple (r, θ, z) such that

1. (r, θ) is a polar representation of the projection of P in the xy-plane.
2. z is the directed distance from (r, θ) to P.

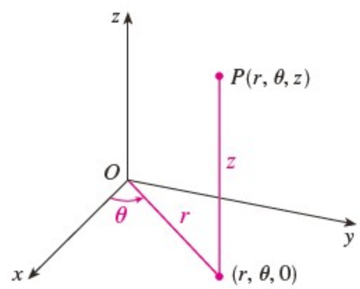


Figure 10.9: Cylindrical coordinate

The point $(0, 0, 0)$ is called the pole. Moreover, because the presentation of a point in the polar coordinate system is not unique, it follows that the representation in the cylindrical coordinate system is also not unique.

We have the coordinate conversion formula:

1. **Cylindrical to rectangular**: $x = r\cos\theta$, $y = r\sin\theta$, $z = z$.
2. **Rectangular to cylindrical**: $r^2 = x^2 + y^2$, $\tan\theta = \dfrac{y}{x}$, $z = z$.

10.4.2 The spherical coordinate

Definition 10.17

In a spherical coordinate system, a point P in space is represented by an ordered triple (ρ, θ, ϕ) such that

1. ρ is the distance between P and the origin (so $\rho \geqslant 0$).
2. θ is the same angle used in cylindrical coordinates for $r \geqslant 0$.
3. ϕ is the angle between the positive z-axis and the line segment $\overline{OP}$ (so $\phi \in [0, \pi]$).

Note that the first and third coordinates, ρ and ϕ, are nonnegative.

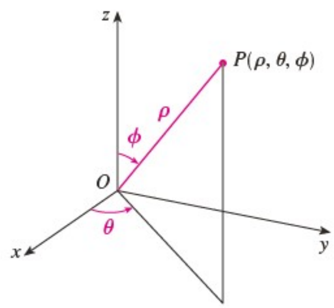


Figure 10.10: Spherical coordinate

The collection of all points whose "spherical representation" has the same $\rho > 0$ is the sphere center at the origin with radius ρ. Therefore, for fixed $\rho > 0$ the (θ, ϕ) coordinate system can be used to represent points on the sphere (centered at the origin with radius ρ) which is similar to the latitude-longitude system used to identify points on the surface of Earth. In fact, for $\rho = 6371$ kilometer (which is the radius of Earth), with the convention "north is positive and south is negative", "east is positive and west is negative", then θ is the latitude and $\frac{\pi}{2} - \phi$ is the longitude $\big($here $\theta = 0$ and $\theta = \pi$ correspond to the prime meridian (本初子午線) and the international date line (國際換日線), respectively, if $\theta \in (-\pi, \pi]\big)$.

We have the coordinate conversion formula:

1. **Spherical to rectangular**: $x = \rho\cos\theta\sin\phi$, $y = \rho\sin\theta\sin\phi$, $z = \rho\cos\phi$.
2. **Rectangular to spherical**: $\rho^2 = x^2 + y^2 + z^2$, $\tan\theta = \dfrac{y}{x}$, $\phi = \arccos\dfrac{z}{\sqrt{x^2 + y^2 + z^2}}$.

We can also convert the spherical coordinate to cylindrical coordinate and vice versa, by the following conversion formula:

1. **Spherical to cylindrical**: $r^2 = \rho^2\sin^2\phi$, $\theta = \theta$, $z = \rho\cos\phi$.
2. **Cylindrical to spherical**: $\rho = \sqrt{r^2 + z^2}$, $\theta = \theta$, $\phi = \arccos\dfrac{z}{\sqrt{r^2 + z^2}}$.

10.5 Exercise

Problem 10.1. In class we have introduced the permutation symbol ε_{ijk} and use it to define the cross product: for two given vectors $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k} = \sum\limits_{i=1}^{3} u_i\mathbf{e}_i$ and $\mathbf{v} =$

$v_1\mathbf{i}+v_2\mathbf{j}+v_3\mathbf{k}=\sum\limits_{i=1}^{3} v_i\mathbf{e}_i$, the cross product $\mathbf{u}\times\mathbf{v}$ is defined by

$$\mathbf{u}\times\mathbf{v}=\sum_{i=1}^{3}\Big(\sum_{j,k=1}^{3}\varepsilon_{ijk}u_jv_k\Big)\mathbf{e}_i=\sum_{i,j,k=1}^{3}\varepsilon_{ijk}u_jv_k\mathbf{e}_i\,.$$

Use the summation notation above without expanding the sum (不要展開成向量和的形式，直接用 Σ 操作) and the identity

$$\sum_{i=1}^{3}\varepsilon_{ijk}\varepsilon_{irs}=\delta_{jr}\delta_{ks}-\delta_{js}\delta_{kr}$$

to prove the following.

(1) $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=(\mathbf{u}\cdot\mathbf{w})\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\mathbf{w}$ for all vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ in space. (Is the associative law $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=(\mathbf{u}\times\mathbf{v})\times\mathbf{w}$ true?)

(2) $(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d})=\begin{vmatrix}\mathbf{a}\cdot\mathbf{c} & \mathbf{b}\cdot\mathbf{c}\\ \mathbf{a}\cdot\mathbf{d} & \mathbf{b}\cdot\mathbf{d}\end{vmatrix}$ for all vectors $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ in space.

Problem 10.2.

(1) Let $\mathbf{u},\mathbf{v}$ be vectors in space satisfying $\mathbf{u}\cdot\mathbf{v}=\sqrt{3}$ and $\mathbf{u}\times\mathbf{v}=\mathbf{i}+2\mathbf{j}+2\mathbf{k}$. Find the angle between $\mathbf{u}$ and $\mathbf{v}$.

(2) Let $\mathbf{u},\mathbf{v}$ be vectors in space. What can you conclude if $\mathbf{u}\times\mathbf{v}=\mathbf{0}$ and $\mathbf{u}\cdot\mathbf{v}=0$?

(3) Let $\mathbf{u},\mathbf{v},\boldsymbol{w}$ be vectors in space. Show that if $\mathbf{u}\cdot\mathbf{v}=\mathbf{u}\cdot\boldsymbol{w}$ and $\mathbf{u}\times\mathbf{v}=\mathbf{u}\times\boldsymbol{w}$, then $\mathbf{v}=\boldsymbol{w}$.

Problem 10.3.

(1) Let P be a point not on the line L that passes through the points Q and R. Show that the distance d from the point P to the line L is

$$d=\frac{\|\mathbf{a}\times\mathbf{b}\|}{\|\mathbf{a}\|}\,,$$

where $\mathbf{a}=\overrightarrow{QR}$ and $\mathbf{b}=\overrightarrow{QP}$.

(2) Let P be a point not on the plane that passes through the points Q, R, and S. Show that the distance d from P to the plane is

$$d = \frac{\left|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})\right|}{\|\mathbf{a} \times \mathbf{b}\|},$$

where $\mathbf{a} = \overrightarrow{QR}$, $\mathbf{b} = \overrightarrow{QS}$ and $\mathbf{c} = \overrightarrow{QP}$.

Problem 10.4. Show that the polar equation $r = a\sin\theta + b\cos\theta$, where $ab \neq 0$, represents a circle, and find its center and radius.

Problem 10.5. Replace the polar equations in the following questions with equivalent Cartesian equations.

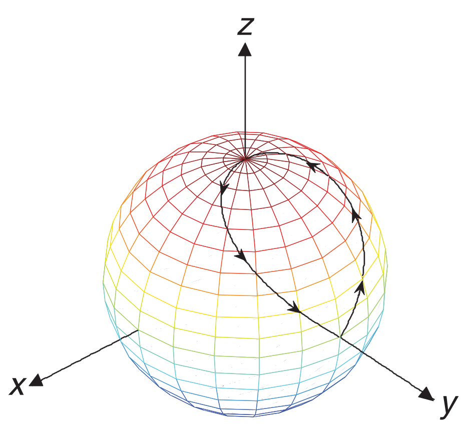
(1) $r^2 \sin 2\theta = 2$ (2) $r = 4\tan\theta\sec\theta$ (3) $r = \csc\theta e^{r\cos\theta}$ (4) $r\sin\theta = \ln r + \ln\cos\theta$.

Problem 10.6. Let C be a smooth curve parameterized by

$$\boldsymbol{r}(t) = (\cos t \sin t, \sin t \sin t, \cos t), \qquad t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

(1) Show that C is a closed curve on the unit sphere $\mathbb{S}^2$.

(2) Using the spherical coordinate, the curve C above corresponds to a curve on the $\theta\phi$-plane. Find the curve in the region $\left\{(\theta, \phi) \,\middle|\, 0 \leqslant \theta \leqslant 2\pi, 0 \leqslant \phi \leqslant \pi\right\}$.



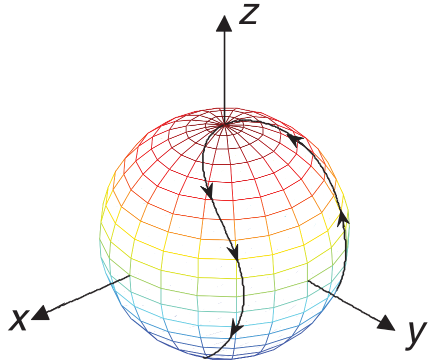
Remark: 想像球面是地球，有人開飛機飛行了 C 這個路線。這個路線在世界地圖上對應到另一個曲線，第二小題即是要求在世界地圖上這個曲線為何。

Problem 10.7. Let C be a smooth curve parameterized by

$$\boldsymbol{r}(t) = \big(\cos(\sin t)\sin t, \sin(\sin t)\sin t, \cos t\big), \qquad t \in [0, 2\pi].$$

(1) Show that C is a closed curve on the unit sphere $\mathbb{S}^2$.

(2) Using the spherical coordinate, the curve C above corresponds to a curve on the $\theta\phi$-plane. Find the curve in the region $\{(\theta, \phi) \,|\, 0 \leqslant \theta \leqslant 2\pi, 0 \leqslant \phi \leqslant \pi\}$.



Chapter 12

Vector-Valued Functions

12.1 Vector-Valued Functions of One Variable

Definition 12.1

A function of the form

$$\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} \qquad \text{or} \qquad \boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$$

is a vector-valued function of one variable, where the component function f, g and h are real-valued functions of the parameter t. Using the vector notation, vector-valued functions above are sometimes denoted by

$$\boldsymbol{r}(t) = \big(f(t), g(t)\big) \qquad \text{or} \qquad \boldsymbol{r}(t) = \big(f(t), g(t), h(t)\big) .$$

Remark 12.2. When $\boldsymbol{r}$ is a vector-valued function, we automatically assume that its components f, g (and h) have a common domain.

Definition 12.3: Limit of Vector-Valued Functions

1. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$, then
$$\lim_{t\to a} \boldsymbol{r}(t) = \Big(\lim_{t\to a} f(t)\Big)\mathbf{i} + \Big(\lim_{t\to a} g(t)\Big)\mathbf{j}$$
provided that the limits $\lim\limits_{t\to a} f(t)$ and $\lim\limits_{t\to a} g(t)$ exist.
2. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, then
$$\lim_{t\to a} \boldsymbol{r}(t) = \Big(\lim_{t\to a} f(t)\Big)\mathbf{i} + \Big(\lim_{t\to a} g(t)\Big)\mathbf{j} + \Big(\lim_{t\to a} h(t)\Big)\mathbf{k}$$
provided that the limits $\lim\limits_{t\to a} f(t)$, $\lim\limits_{t\to a} g(t)$ and $\lim\limits_{t\to a} h(t)$ exist.

Remark 12.4. Using the ϵ-δ language, the limit of a vector-valued function $\boldsymbol{r}$ is defined as follows: Let I be the domain of $\boldsymbol{r}$. The notation $\lim_{t \to a} \boldsymbol{r}(t) = \mathbf{L}$ means for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|\boldsymbol{r}(t) - \mathbf{L}\| < \varepsilon$ whenever $0 < |t - a| < \delta$ and $t \in I$.

Definition 12.5: Continuity of Vector-Valued Functions

A vector-valued function $\boldsymbol{r}$ is said to be continuous at a point a if the limit $\lim_{t \to a} \boldsymbol{r}(t)$ exists and $\lim_{t \to a} \boldsymbol{r}(t) = \boldsymbol{r}(a)$.

Definition 12.6: Differentiation of Vector-Valued Functions

The derivative of a vector-valued function $\boldsymbol{r}$ at a point a is

$$\boldsymbol{r}'(a) = \lim_{h \to 0} \frac{\boldsymbol{r}(a+h) - \boldsymbol{r}(a)}{h}$$

provided that the limit above exists. If $\boldsymbol{r}'(a)$ exists, then $\boldsymbol{r}$ is said to be differentiable at a and $\boldsymbol{r}'(a)$ is called the derivative of $\boldsymbol{r}$ at a. If $\boldsymbol{r}'(t)$ exists for all t in an interval I, then $\boldsymbol{r}$ is said to be differentiable on the interval I.

Theorem 12.7

1. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$, then
$$\boldsymbol{r}'(a) = f'(a)\mathbf{i} + g'(a)\mathbf{j}$$
provided that $f'(a)$ and $g'(a)$ exist.
2. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, then
$$\boldsymbol{r}'(a) = f'(a)\mathbf{i} + g'(a)\mathbf{j} + h'(a)\mathbf{k}$$
provided that $f'(a)$, $g'(a)$ and $h'(a)$ exist.

Theorem 12.8

Let $\boldsymbol{r}$ and $\boldsymbol{u}$ be differentiable vector-valued functions and f be a differentiable real-valued function.

(a) $\dfrac{d}{dt}(f\boldsymbol{r})(t) = f'(t)\boldsymbol{r}(t) + f\boldsymbol{r}'(t)$.

(b) $\dfrac{d}{dt}\big[\boldsymbol{r}(t) \pm \boldsymbol{u}(t)\big] = \boldsymbol{r}'(t) \pm \boldsymbol{u}'(t)$.

(c) $\dfrac{d}{dt}\big[\boldsymbol{r}(t) \star \boldsymbol{u}(t)\big] = \boldsymbol{r}'(t) \star \boldsymbol{u}(t) + \boldsymbol{r}(t) \star \boldsymbol{u}'(t)$, where $\star$ is the dot product or the cross product.

(d) $\dfrac{d}{dt}\boldsymbol{r}\big(f(t)\big) = f'(t)\boldsymbol{r}'\big(f(t)\big)$.

Proof. We only prove (c) for the case $\star$ being the cross product. Write $\boldsymbol{r}(t) = r_1(t)\mathbf{i} + r_2(t)\mathbf{j} + r_3(t)\mathbf{k}$ and $\boldsymbol{u}(t) = u_1(t)\mathbf{i} + u_2(t)\mathbf{j} + u_3(t)\mathbf{k}$. By the definition of the cross product, $\big[\boldsymbol{r}(t) \times \boldsymbol{u}(t)\big]_i$, the i-th component of $\boldsymbol{r}(t) \times \boldsymbol{u}(t)$, is given by $\sum\limits_{1 \leqslant j,k \leqslant 3} \varepsilon_{ijk} r_j(t) u_k(t)$. By the product rule,

$$\begin{aligned}\frac{d}{dt}\big[\boldsymbol{r}(t) \times \boldsymbol{u}(t)\big]_i &= \frac{d}{dt} \sum_{1 \leqslant j,k \leqslant 3} \varepsilon_{ijk} r_j(t) u_k(t) = \sum_{1 \leqslant j,k \leqslant 3} \varepsilon_{ijk} \frac{d}{dt}\big[r_j(t) u_k(t)\big] \\ &= \sum_{1 \leqslant j,k \leqslant 3} \varepsilon_{ijk}\big[r_j'(t) u_k(t) + r_j(t) u_j'(t)\big] = \boldsymbol{r}'(t) \times \boldsymbol{u}(t) + \boldsymbol{r}(t) \times \boldsymbol{u}'(t)\,,\end{aligned}$$

where we have used $\boldsymbol{r}'(t) = r_1'(t)\mathbf{i} + r_2'(t)\mathbf{j} + r_3'(t)\mathbf{k}$ and $\boldsymbol{u}'(t) = u_1'(t)\mathbf{i} + u_2'(t)\mathbf{j} + u_3'(t)\mathbf{k}$ to conclude the last equality. $\square$

Remark 12.9. The proof presented above in fact can be used to show that

$$\begin{aligned}&\frac{d}{dt}\begin{vmatrix} a_{11}(t) & a_{12}(t) & a_{13}(t) \\ a_{21}(t) & a_{22}(t) & a_{23}(t) \\ a_{31}(t) & a_{32}(t) & a_{33}(t) \end{vmatrix} \\ &\quad = \begin{vmatrix} a_{11}'(t) & a_{12}'(t) & a_{13}'(t) \\ a_{21}(t) & a_{22}(t) & a_{23}(t) \\ a_{31}(t) & a_{32}(t) & a_{33}(t) \end{vmatrix} + \begin{vmatrix} a_{11}(t) & a_{12}(t) & a_{13}(t) \\ a_{21}'(t) & a_{22}'(t) & a_{23}'(t) \\ a_{31}(t) & a_{32}(t) & a_{33}(t) \end{vmatrix} + \begin{vmatrix} a_{11}(t) & a_{12}(t) & a_{13}(t) \\ a_{21}(t) & a_{22}(t) & a_{23}(t) \\ a_{31}'(t) & a_{32}'(t) & a_{33}'(t) \end{vmatrix}\end{aligned}$$

since the determinant of $A = \big[a_{ij}(t)\big]_{1 \leqslant i,j \leqslant 3}$ is given by $\sum\limits_{1 \leqslant i,j,k \leqslant 3} \varepsilon_{ijk} a_{1i}(t) a_{2j}(t) a_{3k}(t)$. The formula above shows that the differentiation of determinants is obtained by differentiating row by row (or column by column).

• Integration of vector-valued functions of one variable

Similar to the differentiation of vector-valued functions which mimics the differentiation of real-valued functions, we can also define the Riemann integral of a vector-valued function $\boldsymbol{r}$ on $[a, b]$ as the "limit" of the Riemann sum

$$\sum_{k=1}^{n} \boldsymbol{r}(\xi_k)(t_k - t_{k-1})\,, \tag{12.1.1}$$

where $\{a = t_0 < t_1 < \cdots < t_n = b\}$ is a partition of $[a, b]$. To be more precise, a vector-valued function $\boldsymbol{r} : [a, b] \to \mathbb{R}^d$, where $d = 2$ or 3, is said to be Riemann integrable on $[a, b]$ if there exists a vector $\boldsymbol{A}$ such that for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ is a partition of $[a, b]$ satisfying $\|\mathcal{P}\| < \delta$, any Riemann sum of $\boldsymbol{r}$ for $\mathcal{P}$ (given by (12.1.1)) locates in $(\boldsymbol{A}-\varepsilon, \boldsymbol{A}+\varepsilon)$, where the vector $\boldsymbol{A} \pm \varepsilon$ is the vector obtained by adding or subtracting ε from each component of $\boldsymbol{A}$. The vector $\boldsymbol{A}$, if exists, is written as $\displaystyle\int_a^b \boldsymbol{r}(t)\,dt$. Since the limit of a vector-valued function can be computed componentwise, we have the following

Theorem 12.9

1. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$, then
$$\int_a^b \boldsymbol{r}(t)\,dt = \Big(\int_a^b f(t)dt\Big)\mathbf{i} + \Big(\int_a^b g(t)\,dt\Big)\mathbf{j}$$
provided that $\displaystyle\int_a^b f(t)\,dt$ and $\displaystyle\int_a^b g(t)\,dt$ exist.
2. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, then
$$\int_a^b \boldsymbol{r}(t)\,dt = \Big(\int_a^b f(t)dt\Big)\mathbf{i} + \Big(\int_a^b g(t)\,dt\Big)\mathbf{j} + \Big(\int_a^b h(t)\,dt\Big)\mathbf{k}$$
provided that $\displaystyle\int_a^b f(t)\,dt$, $\displaystyle\int_a^b g(t)\,dt$ and $\displaystyle\int_a^b h(t)\,dt$ exist.

The Fundamental Theorem of Calculus provides a way to compute the definite integral of vector-valued functions, and this enables us to define the indefinite integral of vector-valued functions as follows.

Definition 12.10

1. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i}+g(t)\mathbf{j}$, then the indefinite integral (anti-derivative) of $\boldsymbol{r}$ is
$$\int \boldsymbol{r}(t)\,dt = \Big(\int f(t)dt\Big)\mathbf{i} + \Big(\int g(t)\,dt\Big)\mathbf{j}$$
provided that $\displaystyle\int f(t)\,dt$ and $\displaystyle\int g(t)\,dt$ exist.
2. If $\boldsymbol{r}$ is a vector-valued function such that $\boldsymbol{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$, then the indefinite integral (anti-derivative) of $\boldsymbol{r}$ is
$$\int \boldsymbol{r}(t)\,dt = \Big(\int f(t)dt\Big)\mathbf{i} + \Big(\int g(t)\,dt\Big)\mathbf{j} + \Big(\int h(t)\,dt\Big)\mathbf{k}$$
provided that $\displaystyle\int f(t)\,dt$, $\displaystyle\int g(t)\,dt$ and $\displaystyle\int h(t)\,dt$ exist.

Having defined the indefinite integral of vector-valued functions, by the Fundamental Theorem of Calculus and Theorem 12.7 we have

$$\frac{d}{dt}\int \boldsymbol{r}(t)\,dt = \boldsymbol{r}(t)$$

as long as $\boldsymbol{r}$ is continuous.

12.2 Curves and Parametric Equations

Definition 12.11

A subset C in the plane (or space) is called a ***curve*** if C is the image of an interval $I \subseteq \mathbb{R}$ under a continuous vector-valued function $\boldsymbol{r}$. The continuous function $\boldsymbol{r} : I \to \mathbb{R}^2$ (or $\mathbb{R}^3$) is called a ***parametrization*** of the curve, and the equation

$$(x, y) = \boldsymbol{r}(t)\,,\ t \in I \qquad \big(\text{or } (x, y, z) = \boldsymbol{r}(t),\ \ t \in I\big)$$

is called a ***parametric equation*** of the curve. A curve C is called a ***plane curve*** if it is a subset in the plane.

Since a plane can be treated as a subset of space, in the following we always assume that the curve under discussion is a curve in space (so that the parametrization of the curve is given by $\boldsymbol{r} : I \to \mathbb{R}^3$).

Definition 12.12

A curve C is called ***simple*** if it has an injective parametrization; that is, there exists $\boldsymbol{r} : I \to \mathbb{R}^3$ such that $\boldsymbol{r}(I) = C$ and $\boldsymbol{r}(x) = \boldsymbol{r}(y)$ implies that $x = y$. A curve C with parametrization $\boldsymbol{r} : I \to \mathbb{R}^3$ is called ***closed*** if $I = [a, b]$ for some closed interval $[a, b] \subseteq \mathbb{R}$ and $\boldsymbol{r}(a) = \boldsymbol{r}(b)$. A ***simple closed*** curve C is a closed curve with parametrization $\boldsymbol{r} : [a, b] \to \mathbb{R}^3$ such that $\boldsymbol{r}$ is one-to-one on (a, b). A ***smooth*** curve C is a curve with differentiable parametrization $\boldsymbol{r} : I \to \mathbb{R}^3$ such that $\boldsymbol{r}'(t) \neq \mathbf{0}$ for all $t \in I$.

Example 12.13. The parabola $y = x^2 + 2$ on the plane is a simple smooth plane curve since $\boldsymbol{r} : \mathbb{R} \to \mathbb{R}^2$ given by $r(t) = t\mathbf{i} + (r^2 + 2)\mathbf{j}$ is an injective differentiable parametrization of this parabola. We note that $\widetilde{\boldsymbol{r}} : \big(-\frac{\pi}{2}, \frac{\pi}{2}\big) \to \mathbb{R}^2$ given by $\widetilde{\boldsymbol{r}}(t) = \tan t\,\mathbf{i} + (\sec^2 t + 1)\mathbf{j}$ is also an injective smooth parametrization of this parabola. In general, a curve usually has infinitely many parameterizations.

Example 12.14. Let $I \subseteq \mathbb{R}$ be an interval, and $\boldsymbol{r} : I \to \mathbb{R}^2$ be defined by $\boldsymbol{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j}$. Since $\boldsymbol{r}$ is continuous and the co-domain is $\mathbb{R}^2$, the image of I under $\boldsymbol{r}$, denoted by C, is a plane curve. We note that C is part of the unit circle centered at the origin. Moreover, C is a smooth curve since $\boldsymbol{r}'(t) \neq \mathbf{0}$ for all $t \in I$.

1. If $I = [a, b]$ and $|b - a| < 2\pi$, then C is a simple curve.
2. If $I = [0, 2\pi]$, then C is not a simple curve. However, C a simple closed curve.

Example 12.15. Let $\boldsymbol{r} : [0, 2\pi] \to \mathbb{R}^2$ be defined by $\boldsymbol{r}(t) = \sin t\mathbf{i} + \sin t \cos t\mathbf{j}$. The image $\boldsymbol{r}([0, 2\pi])$ is a curve called figure eight.

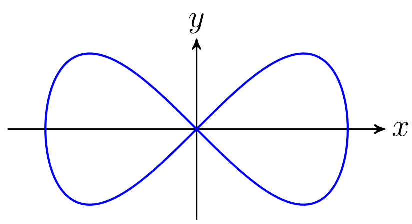


Figure 12.1: Figure eight

Example 12.16. Let $\boldsymbol{r} : \mathbb{R} \to \mathbb{R}^3$ be defined by $\boldsymbol{r}(t) = \cos t\mathbf{i} + \sin t\mathbf{j} + t\mathbf{k}$. Then the image $\boldsymbol{r}(\mathbb{R})$ is a simple smooth space curve. This curve is called a helix.

In the following, when a parametrization $\boldsymbol{r} : I \to \mathbb{R}^3$ of curves C is mentioned, we always assume that "there is no overlap"; that is, there are no intervals $[a, b], [c, d] \subseteq I$ satisfying that $\boldsymbol{r}([a, b]) = \boldsymbol{r}([c, d])$. If in addition

1. C is a simple curve, then $\boldsymbol{r}$ is injective, or
2. C is closed, then $I = [a, b]$ and $\boldsymbol{r}(a) = \boldsymbol{r}(b)$, or
3. C is simple closed, then $I = [a, b]$ and $\boldsymbol{r}$ is injective on $[a, b)$ and $\boldsymbol{r}(a) = \boldsymbol{r}(b)$.
4. C is smooth, then $\boldsymbol{r}$ is differentiable and $\boldsymbol{r}'(t) \neq \mathbf{0}$ for all $t \in I$.

12.2.1 Polar Graphs

In Example 10.13 we talk about one particular correspondence between a curve on the $r\theta$-plane and a curve on the xy-plane. The equation $r = \cos\theta$ is called a polar equation which means an equation in polar coordinate, and the corresponding curve given by the relation $(x, y) = (r\cos\theta, r\sin\theta)$ on the xy-plane is called the polar graph of this polar equation.

Definition 12.17

Let (r, θ) be the polar coordinate. A polar equation is an equation that r and θ satisfy. The polar graph of a polar equation is the collection of points $(r\cos\theta, r\sin\theta)$ in xy-plane with (r, θ) satisfying the given polar equation.

Remark 12.18. Usually, the polar equation under consideration is of the form

$$r = f(\theta) \qquad \text{or} \qquad \theta = g(r)$$

for some functions f and g. The polar graph of the polar equation $r = f(\theta)$ is the curve parameterized by the parametrization $\boldsymbol{r} : \mathbb{R} \to \mathbb{R}^2$ given by $\boldsymbol{r}(t) = f(t)\cos t\,\mathbf{i} + f(t)\sin t\,\mathbf{j}$ (where t is the role of θ), while the polar graph of the polar equation $\theta = g(r)$ is the curve parameterized by the parametrization $\boldsymbol{r} : \mathbb{R} \to \mathbb{R}^2$ given by $\boldsymbol{r}(t) = t\cos g(t)\mathbf{i} + t\sin g(t)\mathbf{j}$ (where t is the role of r).

Example 12.19. 1. The polar graph of the polar equation $r = r_0$, where $r_0 \neq 0$ is a constant, is the circle centered at the origin with radius $|r_0|$.

2. The polar graph of the polar equation $\theta = \theta_0$, where θ_0 is a constant, is the straight line with slope $\tan\theta_0$.

3. The polar graph of the polar equation $r = \sec\theta$ is $x = 1$ (in the xy-plane).

4. The polar graph of the polar equation $r = a\cos\theta$, where a is a constant, is the circle centered at $\left(\frac{a}{2}, 0\right)$ with radius $\frac{|a|}{2}$.

5. The polar graph of the polar equation $r = a\sin\theta$, where a is a constant, is the circle centered at $\left(0, \frac{a}{2}\right)$ with radius $\frac{|a|}{2}$.

Example 12.20. A conic section (圓錐曲線) can be defined purely in terms of plane geometry: it is the locus of all points P whose distance to a fixed point F (called the focus 焦點) is a constant multiple (called the eccentricity e 離心率) of the distance from P to a fixed line L (called the directrix 準線). For $0 < \mathrm{e} < 1$ we obtain an ellipse, for $\mathrm{e} = 1$ a parabola, and for $\mathrm{e} > 1$ a hyperbola.

Now we consider the polar equation whose polar graph represents a conic section. Let the focus be the pole of a polar coordinate, and the polar axis is perpendicular to the directrix but does not intersect the directrix. Then the eccentricity e is given by

$$\mathrm{e} = \frac{d(P, F)}{d(P, L)} \qquad \text{for all points } P \text{ on the conic section,} \tag{12.2.1}$$

where $d(P,F)$ is the distance between P and the focus F, and $d(P,L)$ is the distance between P and the directrix.

Let P denote the distance between the pole and the directrix, and the polar coordinate of points P on a conic section is (r,θ). Then (12.2.1) implies that

$$\mathrm{e} = \frac{r}{r\cos\theta + \mathrm{P}}\,.$$

Therefore, the polar equation of a conic section with eccentricity e is given by

$$r = \frac{\mathrm{eP}}{1-\mathrm{e}\cos\theta}\,.$$

In general, for a given conic section we let the principal ray denote the ray starting from the focus, perpendicular to the directrix without intersecting the directrix. Let the focus F be the pole of a polar coordinate and θ_0 be the directed angel from the polar axis to the principal ray. If (r,θ) is the polar representation of point P on the conic section, then (r,θ) satisfies

$$\mathrm{e} = \frac{r}{r\cos(\theta-\theta_0) + \mathrm{P}} \qquad \text{or equivalently,} \quad r = \frac{\mathrm{eP}}{1-\mathrm{e}\cos(\theta-\theta_0)}\,.$$

Example 12.21 (Limaçons - 蚶線)**.** The polar graph of the polar equation $r = a \pm b\cos\theta$ or $r = a \pm b\sin\theta$, where $a, b > 0$ are constants, is called a limaçon. A limaçons is also called a cardioid（心臟線）if $a = b$.

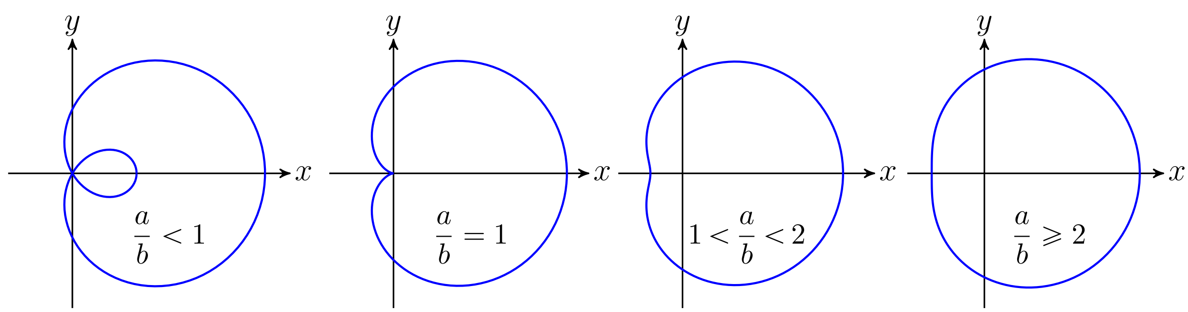


Figure 12.2: Limaçons $r = a \pm b\cos\theta$ with the ratio $\frac{a}{b}$ in different regions

(1) There is an inner loop when $\frac{a}{b} < 1$. (2) When $a = b$ it is also called the cardioid.
(3) When $1 < \frac{a}{b} < 2$, the region enclosed by the limaçon is not convex. This kind of limaçons is called dimpled limaçons. (4) When $\frac{a}{b} \geqslant 2$, it is called convex limaçons.

Example 12.22 (Rose curves)**.** The polar graph of the polar equation $r = a\cos n\theta$ or $r = a\sin n\theta$, where $a > 0$ is a given number and $n \geqslant 2$ is an integer, is called a rose curve.

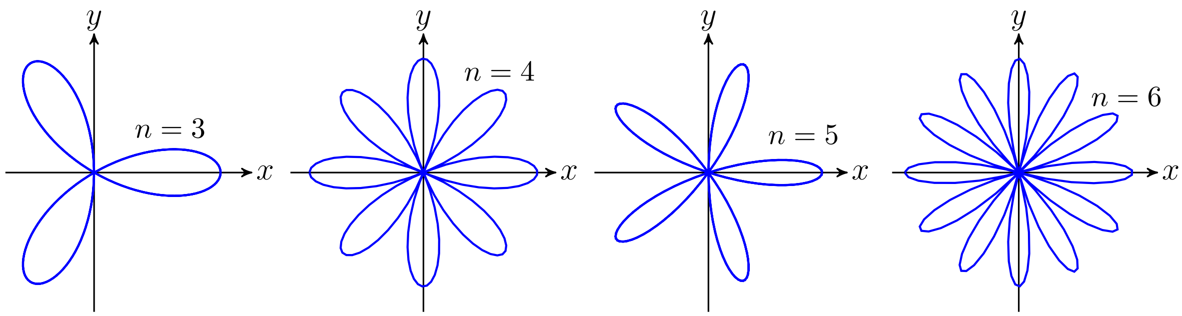


Figure 12.3: Rose curves $r = a\cos n\theta$: n petals when n is odd and $2n$ petals when n is even

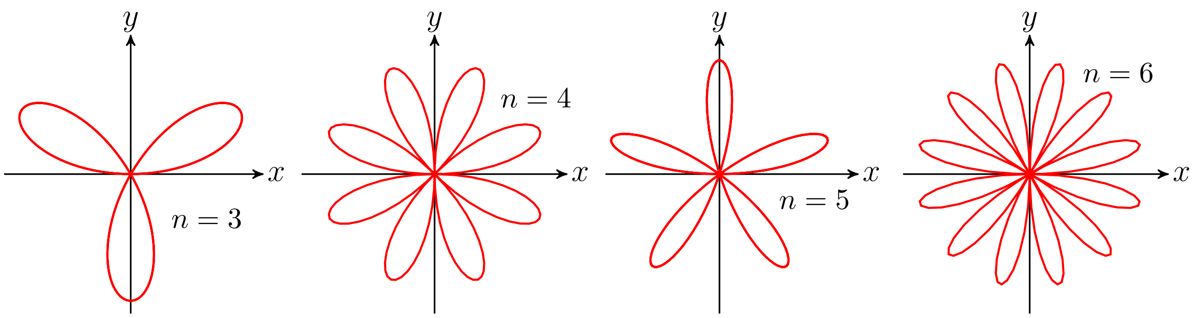


Figure 12.4: Rose curves $r = a\sin n\theta$: n petals when n is odd and $2n$ petals when n is even

Example 12.23 (Lemniscates - 雙紐線)**.** The polar graph of the polar equation $r^2 = a^2\sin 2\theta$ or $r^2 = a^2\cos 2\theta$ is called a lemniscate.

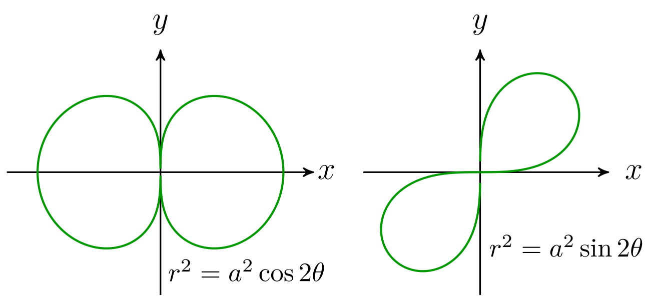


Figure 12.5: Lemniscate $r^2 = a^2\cos 2\theta$ or $r^2 = a^2\sin 2\theta$

12.3 Physical and Geometric Meanings of the Derivative of Vector-Valued Functions

Let $I \subseteq \mathbb{R}$ be an interval and $\boldsymbol{r} : I \to \mathbb{R}^3$ be a differentiable vector-valued function.

12.3.1 Physical meaning

Treat I as the time interval, and $\boldsymbol{r}(t)$ as the position of an object at time t. For $a, b \in I$ and

$a < b$, $\dfrac{\boldsymbol{r}(b) - \boldsymbol{r}(a)}{b - a}$ is the average velocity of the object in the time interval $[a, b]$. Therefore,

$$\boldsymbol{r}'(c) = \lim_{h \to 0} \frac{\boldsymbol{r}(c+h) - \boldsymbol{r}(c)}{h},$$

is the instantaneous velocity at $t = c$, and $\|\boldsymbol{r}'(c)\|$ is the instantaneous speed at $t = c$. If $\boldsymbol{r}$ is twice differentiable, then the derivative of the velocity vector $\boldsymbol{r}'$ is the acceleration.

Definition 12.24

Let $I \subseteq \mathbb{R}$ be the time interval and $\boldsymbol{r} : I \to \mathbb{R}^3$ be the position vector. The velocity vector, acceleration vector and the speed at time t are

$$\begin{aligned} \textbf{Velocity} &= \boldsymbol{v}(t) = \boldsymbol{r}'(t)\,, \\ \textbf{Acceleration} &= \boldsymbol{a}(t) = \boldsymbol{r}''(t)\,, \\ \textbf{Speed} &= \|\boldsymbol{v}(t)\| = \|\boldsymbol{r}'(t)\|\,. \end{aligned}$$

Example 12.25. Suppose a satellite is under uniform circular motion（等速率圓周運動）and the position of the satellite is given by

$$\boldsymbol{r}(t) = \big(R\cos(\omega t), R\sin(\omega t)\big)\,,$$

where R is the distance between the satellite and the center of Earth, and ω is the angular velocity. Then

$$\boldsymbol{r}'(t) = R\omega\big(-\sin(\omega t), \cos(\omega t)\big) \quad \text{and} \quad \boldsymbol{r}''(t) = -R\omega^2\big(\cos(\omega t), \sin(\omega t)\big)\,;$$

thus

$$\|\boldsymbol{a}(t)\| = \|\boldsymbol{r}''(t)\| = R\omega^2 = \frac{\|\boldsymbol{r}'(t)\|^2}{R} = \frac{\|\boldsymbol{v}(t)\|^2}{R}$$

which gives the famous formula for the centripetal acceleration（向心加速度）.

Example 12.26. In this example we consider the motion of a planet around a single sun. In the plane on which the planet moves, we introduce a polar coordinate system and a Cartesian coordinate system as follows:

1. Let the sun be the pole of the polar coordinate system, and fixed a polar axis on this plane.
2. Let $\mathbf{i}$ be the unit vector in the direction of the polar axis, and $\mathbf{j}$ be the corresponding unit vector obtained by rotating $\mathbf{i}$ counterclockwise by $\dfrac{\pi}{2}$.

Suppose the position of the planet on the plane at time $t \in I$ is given by $\boldsymbol{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$. For each $t \in I$, let $(r(t), \theta(t))$ be the polar representation of $(x(t), y(t))$ in the trajectory. We would like to determine the relation that $r(t)$ and $\theta(t)$ satisfy.

Define two vectors $\widehat{r}(t) = \cos\theta(t)\mathbf{i} + \sin\theta(t)\mathbf{j}$ and $\widehat{\theta}(t) = -\sin\theta(t)\mathbf{i} + \cos\theta(t)\mathbf{j}$. Then $\boldsymbol{r} = r\widehat{r}$. Moreover, let M and m be the mass of the sun and the planet, respectively. Then Newton's second law of motion implies that

$$-\frac{GMm}{r^2}\widehat{r} = m\boldsymbol{r}''. \tag{12.3.1}$$

By the fact that $\widehat{r}' = \theta'\widehat{\theta}$ and $\widehat{\theta}' = -\theta'\widehat{r}$, we find that

$$\begin{aligned}\boldsymbol{r}'' &= \frac{d}{dt}\big(r'\widehat{r} + r\theta'\widehat{\theta}\big) = r''\widehat{r} + r'\theta'\widehat{\theta} + r'\theta'\widehat{\theta} + r\theta''\widehat{\theta} - r(\theta')^2\widehat{r}\\ &= \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}.\end{aligned}$$

Therefore, (12.3.1) implies that

$$-\frac{GM}{r^2}\widehat{r} = \big[r'' - r(\theta')^2\big]\widehat{r} + \big[2r'\theta' + r\theta''\big]\widehat{\theta}.$$

Since $\widehat{r}$ and $\widehat{\theta}$ are linearly independent, we must have

$$-\frac{GM}{r^2} = r'' - r(\theta')^2, \tag{12.3.2a}$$

$$2r'\theta' + r\theta'' = 0. \tag{12.3.2b}$$

Note that (12.3.2b) implies that $(r^2\theta')' = 0$; thus $r^2\theta'$ is a constant. Since $mr^2\theta'$ is the angular momentum, (12.3.2b) implies that the angular momentum is a constant, so-called the conservation of angular momentum（角動量守恆）.

12.3.2 Geometric meaning

Suppose that the image $r(I)$ is a curve C. Since $\boldsymbol{r}(c+h) - \boldsymbol{r}(c)$ is the vector pointing from $\boldsymbol{r}(c)$ to $\boldsymbol{r}(c+h)$, we expect that $\boldsymbol{r}'(c)$, if it is not zero, is tangent to the curve at the point $\boldsymbol{r}(c)$. This motivates the following

Definition 12.27

Let C be a smooth curve represented by $\boldsymbol{r}$ on an interval I. The unit tangent vector $\mathbf{T}$ (associated with the parametrization $\boldsymbol{r}$) is defined as

$$\mathbf{T}(t) = \frac{\boldsymbol{r}'(t)}{\|\boldsymbol{r}'(t)\|}.$$

Remark 12.28. Since there are infinitely many parameterizations of a given smooth curve, different parameterizations of a smooth curve might provide different unit tangent vector. However, this is not the case and there are only two unit tangent vectors.

Theorem 12.29

Let $I \subseteq \mathbb{R}$ be an interval, and $\boldsymbol{r} : I \to \mathbb{R}^3$ be a differentiable vector-valued function. If $\|\boldsymbol{r}(t)\|$ is a constant function on I, then

$$\boldsymbol{r}(t) \cdot \boldsymbol{r}'(t) = 0 \qquad \forall\, t \in I\,.$$

Proof. Suppose that $\|\boldsymbol{r}(t)\| = C$ for some constant C. Since $\|\boldsymbol{r}(t)\|^2 = \boldsymbol{r}(t) \cdot \boldsymbol{r}(t)$,

$$\boldsymbol{r}(t) \cdot \boldsymbol{r}(t) = C^2 \qquad \forall\, t \in I\,;$$

thus by the fact that $\boldsymbol{r}$ is differentiable, Theorem 12.8 implies that

$$\boldsymbol{r}(t) \cdot \boldsymbol{r}'(t) = \frac{1}{2}\Big[\boldsymbol{r}(t) \cdot \boldsymbol{r}'(t) + \boldsymbol{r}'(t) \cdot \boldsymbol{r}(t)\Big] = \frac{1}{2}\frac{d}{dt}\big[\boldsymbol{r}(t) \cdot \boldsymbol{r}(t)\big] = 0 \qquad \forall\, t \in I\,. \qquad \square$$

Corollary 12.30

Let C be a smooth curve represented by $\boldsymbol{r}$ on an interval I, and $\mathbf{T}(t) = \dfrac{\boldsymbol{r}'(t)}{\|\boldsymbol{r}'(t)\|}$ be the unit tangent vector (associated with the parametrization r). If $\boldsymbol{T}$ is differentiable at t, then

$$\mathbf{T}(t) \cdot \mathbf{T}'(t) = 0 \qquad \forall\, t \in I\,.$$

Definition 12.31

Let C be a smooth curve represented by $\boldsymbol{r}$ on an interval I, and $\mathbf{T}(t) = \dfrac{\boldsymbol{r}'(t)}{\|\boldsymbol{r}'(t)\|}$ be the unit tangent vector (associated with r). If $\mathbf{T}'(t)$ exists and $\mathbf{T}'(t) \neq \mathbf{0}$, then the ***principal unit normal vector*** (associated with the parametrization $\boldsymbol{r}$) at t is defined as

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}\,.$$

Theorem 12.32

Let C be a smooth curve represented by $\boldsymbol{r}$ on an interval I, and the principal unit normal vector $\mathbf{N}$ exists, then the acceleration vector $\boldsymbol{a}$ lies in the plane determined by the unit tangent vector $\mathbf{T}$ and $\mathbf{N}$.

Proof. Let $\boldsymbol{v} = \boldsymbol{r}'$ be the velocity vector. Then

$$\boldsymbol{v} = \|\boldsymbol{v}\| \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = \|\boldsymbol{v}\| \frac{\boldsymbol{r}'}{\|\boldsymbol{r}'\|} = \|\boldsymbol{v}\|\mathbf{T}\,.$$

Therefore,

$$\boldsymbol{a} = \boldsymbol{v}' = \|\boldsymbol{v}\|'\mathbf{T} + \|\boldsymbol{v}\|\mathbf{T}' = \|\boldsymbol{v}\|'\mathbf{T} + \|\boldsymbol{v}\|\|\mathbf{T}'\|\mathbf{N}\,.$$

The equation above implies that $\boldsymbol{a}$ is written as a linear combination of $\mathbf{T}$ and $\mathbf{N}$, it follows that $\boldsymbol{a}$ lies in the plane determined by $\mathbf{T}$ and $\mathbf{N}$. $\square$

Remark 12.33. The coefficients of $\mathbf{T}$ and $\mathbf{N}$ in the proof above are called the ***tangential and normal components of acceleration*** and are denoted by

$$a_{\mathbf{T}} = \|\boldsymbol{v}\|' \qquad \text{and} \qquad a_{\mathbf{N}} = \|\boldsymbol{v}\|\|\mathbf{T}'\|$$

so that $\boldsymbol{a}(t) = a_{\mathbf{T}}(t)\mathbf{T}(t) + a_{\mathbf{N}}(t)\mathbf{N}(t)$. Moreover, we note that the formula for $a_{\mathbf{N}}$ above shows that $a_{\mathbf{N}} \geqslant 0$. The normal component of acceleration is also called the ***centripetal component of acceleration***.

The following theorem provides some convenient formulas for computing $a_{\mathbf{T}}$ and $a_{\mathbf{N}}$.

Theorem 12.34

Let C be a smooth curve represented by $\boldsymbol{r}$ on an interval I, and the principal unit normal vector $\mathbf{N}$ exists. Then the tangential and normal components of acceleration are given by

$$\begin{aligned} a_{\mathbf{T}} &= \|\boldsymbol{v}\|' = \boldsymbol{a}\cdot\mathbf{T} = \frac{\boldsymbol{v}\cdot\boldsymbol{a}}{\|\boldsymbol{v}\|}\,, \\ a_{\mathbf{N}} &= \|\boldsymbol{v}\|\|\mathbf{T}'\| = \boldsymbol{a}\cdot\mathbf{N} = \frac{\|\boldsymbol{v}\times\boldsymbol{a}\|}{\|\boldsymbol{v}\|} = \sqrt{\|\boldsymbol{a}\|^2 - a_{\mathbf{T}}^2}\,. \end{aligned}$$

Proof. It suffices to show the formula $a_{\mathbf{N}} = \dfrac{\|\boldsymbol{v}\times\boldsymbol{a}\|}{\|\boldsymbol{v}\|}$. Since $\boldsymbol{a} = a_{\mathbf{T}}\mathbf{T} + a_{\mathbf{N}}\mathbf{N}$, we find that

$$\boldsymbol{a}\times\mathbf{T} = a_{\mathbf{N}}(\mathbf{N}\times\mathbf{T})\,;$$

thus using the fact that $a_{\mathbf{N}} \geqslant 0$, by Theorem 10.6 we find that

$$a_{\mathbf{N}} = |a_{\mathbf{N}}| = \frac{\|a\times\mathbf{T}\|}{\|\mathbf{N}\times\mathbf{T}\|} = \frac{\|\boldsymbol{a}\times\mathbf{T}\|}{\|\mathbf{N}\|\|\mathbf{T}\|\sin\frac{\pi}{2}} = \|\boldsymbol{a}\times\mathbf{T}\| = \frac{\|\boldsymbol{v}\times\boldsymbol{a}\|}{\|\boldsymbol{v}\|}\,. \qquad \square$$

12.4 Arc Length and Area

12.4.1 Arc length

12.4.2 Area enclosed by simple closed curves

Let C be a simple closed curve in the plane parameterized by $\boldsymbol{r} : [a, b] \to \mathbb{R}^2$. Suppose that

1. $\boldsymbol{r}(t) = \big(x(t), y(t)\big)$ moves **counter-clockwise** (that is, the region enclosed by C is on the left-hand side when moving along C) as t increases.

2. There exists $c \in (a, b)$ such that x is strictly **increasing** on $[a, c]$ and is strictly **decreasing** on $[c, b]$ (this implies that every vertical line intersects with the curve C at at most two points)

3. $x'y$ is Riemann integrable on $[a, b]$ (for example, x is continuously differentiable on $[a, b]$).

Based on the assumption above, in the following we "prove" that

$$\text{the area of the region enclosed by } C \text{ is } -\int_a^b x'(t)y(t)\,dt\,. \tag{12.4.1}$$

We remark that condition 2 above implies that $\boldsymbol{r}(a)$ is the "leftmost" point of the curve, and $\boldsymbol{r}(c)$ is the "rightmost" point of the curve.

Since x is strictly increasing on $[a, c]$ and x is continuous, by the Intermediate Value Theorem (Theorem 1.58) we find that for each point $p \in \big[x(a), x(c)\big]$ there exists a unique $t \in [a, c]$ such that $x(t) = p$. Define $q = y(t)$. Then the map $p \mapsto q$ is a function. This implies that the curve $\boldsymbol{r}\big([a, c]\big)$ is the graph of a continuous function $f : \big[x(a), x(c)\big] \to \mathbb{R}$. Moreover, $y(t) = f(x(t))$ for all $t \in [a, c]$. Similarly, the curve $\boldsymbol{r}\big([c, b]\big)$, the "upper part of C", is the graph of a continuous function $g : \big[x(b), x(c)\big] \to \mathbb{R}$ and $y(t) = g(x(t))$ for all $t \in [c, b]$. Since $x(a) = x(b)$, the substitution of variable $x = x(t)$ implies that

$$\begin{aligned}
&\int_{x(a)}^{x(c)} \big[g(x) - f(x)\big]\,dx \\
&\quad = \int_{x(b)}^{x(c)} g(x)\,dx - \int_{x(a)}^{x(c)} f(x)\,dx = \int_b^c g(x(t))x'(t)\,dt - \int_a^c f(x(t))x'(t)\,dt \\
&\quad = \int_b^c y(t)x'(t)\,dt - \int_a^c y(t)x'(t)\,dt = -\int_a^b x'(t)y(t)\,dt\,;
\end{aligned}$$

thus (12.4.1) is concluded since the area of the region enclosed by C is given by the left-hand side of the equality above.

Similar argument can be applied to conclude that

$$\text{the area of the region enclosed by } C \text{ is } \int_a^b x(t)y'(t)\,dt\,. \tag{12.4.2}$$

if xy' is Riemann integrable on $[a,b]$ and every horizontal line intersects with the curve C at at most two points. Combining (12.4.1) and (12.4.2), we obtain that

$$\text{the area of the region enclosed by } C \text{ is } \frac{1}{2}\int_a^b \Big[x(t)y'(t) - x'(t)y(t)\Big]\,dt \tag{12.4.3}$$

provided that $x'y$ and xy' are Riemann integrable on $[a,b]$ and every vertical line and horizontal line intersects with the curve C at at most two points.

Remark 12.34. In general, **the restriction that every vertical line or horizontal line intersects with curve C at at most two points can be removed** from the condition for the use of (12.4.1), (12.4.2) and (12.4.3). We will see this later in Chapter **??** (but for now we will treat this as a fact for we have proved a special case).

Remark 12.35. Using the convention that $\mathbf{u}\times\mathbf{v} = u_1v_2 - u_2v_1$ when $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$ are vectors in the plane, (12.4.3) can be rewritten as

$$\text{the area of the region enclosed by } C \text{ is } \frac{1}{2}\int_a^b \boldsymbol{r}(t)\times\boldsymbol{r}'(t)\,dt\,. \tag{12.4.3'}$$

Without confusion, the area can also be written as $\frac{1}{2}\int_a^b \boldsymbol{r}(t)\times d\boldsymbol{r}(t)$.

Example 12.36. Let C be the curve parameterized by $\boldsymbol{r}(t) = (\cos t, \sin t)$, $t\in[0,2\pi]$. Then clearly $\boldsymbol{r}$ satisfies condition 1-3. Therefore, the area of the region enclosed by C can be computed by the following three ways:

1. Using (12.4.1),
$$-\int_0^{2\pi}\frac{d\cos t}{dt}\sin t\,dt = \int_0^{2\pi}\sin^2 t\,dt = \int_0^{2\pi}\frac{1-\cos(2t)}{2}\,dt = \frac{1}{2}\Big(t - \frac{\sin(2t)}{2}\Big)\Big|_{t=0}^{t=2\pi} = \pi\,.$$

2. Using (12.4.2),
$$\int_0^{2\pi}\cos t\,\frac{d\sin t}{dt}\,dt = \int_0^{2\pi}\cos^2 t\,dt = \int_0^{2\pi}\frac{1+\cos(2t)}{2}\,dt = \frac{1}{2}\Big(t + \frac{\sin(2t)}{2}\Big)\Big|_{t=0}^{t=2\pi} = \pi\,.$$

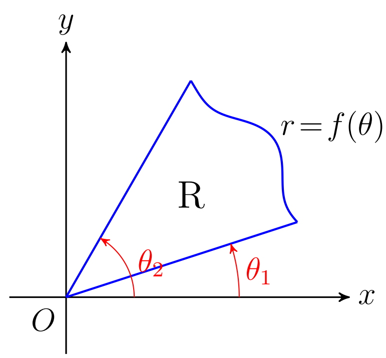
3. Using (12.4.3),
$$\frac{1}{2}\int_0^{2\pi}\Big(\cos t\frac{d\sin t}{dt} - \frac{d\cos t}{dt}\sin t\Big)\,dt = \frac{1}{2}\int_0^{2\pi}\big(\cos^2 t + \sin^2 t\big)\,dt = \frac{1}{2}\int_0^{2\pi}1dt = \pi\,.$$

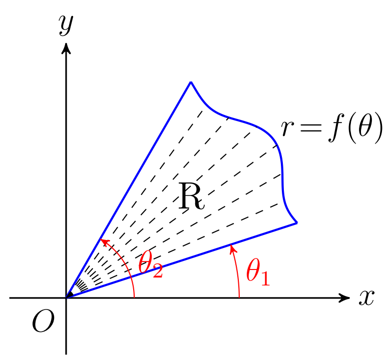
12.4.3 Area and arc length in polar coordinates

Now we consider the area of the region given by the polar representation

$$\big\{(r,\theta)\,\big|\,0\leqslant r\leqslant f(\theta),\theta_1\leqslant\theta\leqslant\theta_2\big\}\,, \tag{12.4.4}$$

where $f:[\theta_1,\theta_2]\to\mathbb{R}$ is non-negative and continuous.





Remark 12.37. Note that the region given in (12.4.4) is enclosed by the curve C parameterized by

$$\boldsymbol{r}(t)=\big(x(t),y(t)\big)=\begin{cases}\big(t-\theta_1+f(\theta_1)\big)\big(\cos\theta_1,\sin\theta_1\big) & \text{if } \theta_1-f(\theta_1)\leqslant t\leqslant\theta_1\,,\\ f(t)\big(\cos t,\sin t\big) & \text{if } \theta_1\leqslant t\leqslant\theta_2\,,\\ \big(\theta_2+f(\theta_2)-t\big)\big(\cos\theta_2,\sin\theta_2\big) & \text{if } \theta_2\leqslant t\leqslant\theta_2+f(\theta_2)\,.\end{cases}$$

Then

$$x(t)y'(t)-x'(t)y(t)=\big(x'(t),y'(t)\big)\cdot\big(-y(t),x(t)\big)=\begin{cases}0 & \text{if } \theta_1-f(\theta_1)\leqslant t\leqslant\theta_1\,,\\ f(t)^2 & \text{if } \theta_1\leqslant t\leqslant\theta_2\,,\\ 0 & \text{if } \theta_2\leqslant t\leqslant\theta_2+f(\theta_2)\,;\end{cases}$$

thus using (12.4.3) we find that

$$\text{the area given in (12.4.4) is } \frac{1}{2}\int_{\theta_1}^{\theta_2} f(\theta)^2\, d\theta\,.$$

Example 12.38 (Kepler's second law)**.**

12.5 Exercise

Problem 12.1. Let C be a curve parameterized by the vector-valued function $\boldsymbol{r} : [0,1] \to \mathbb{R}^2$,

$$\boldsymbol{r}(t) = \Big(\frac{e^t - e^{-t}}{e^t + e^{-t}}, \frac{2}{e^t + e^{-t}}\Big), \quad 0 \leqslant t \leqslant 1\,.$$

(1) Show that C is part of the unit circle centered at the origin.

(2) Plot the curve C. (The plot does not have to be very precise. You only need to specify the starting and end points as well as the orientation.)

(3) Find the length of the curve C.

Problem 12.2. Let C be the curve given by the parametric equations

$$x(t) = \frac{3+t^2}{1+t^2}\,, \qquad y(t) = \frac{2t}{1+t^2}$$

on the interval $t \in [0,1]$.

(1) In fact C is the graph of a function $y = f(x)$. Find f.

(2) Find the arc length of the curve C.

(3) Find the area of the surface formed by revolving the curve C about the y-axis.

Problem 12.3. In class we talked about how to find the total distance that you travel when you walk along a path according to the position vector $\boldsymbol{r} : [a,b] \to \mathbb{R}^2$. The total distance travelled can be computed by

$$\int_a^b \|\boldsymbol{r}'(t)\|\, dt$$

when $\boldsymbol{r}$ is continuously differentiable. Complete the following.

1. Let $\boldsymbol{r} : [0, 4\pi] \to \mathbb{R}^2$ be given by $\boldsymbol{r}(t) = 3\cos t\,\mathbf{i} + 3\sin t\,\mathbf{j}$. Find the image of $[0, 4\pi]$ under $\boldsymbol{r}$.
2. Compute the integral $\displaystyle\int_0^{4\pi} \|\boldsymbol{r}'(t)\|\,dt$. Does it agree with the length of the curve $C \equiv \boldsymbol{r}\big([0, 4\pi]\big)$?

Problem 12.4. To illustrate that the length of a smooth space curve does not depend on the parametrization you use to compute it, calculate the length of one turn of the helix in Example 1 with the following parametrizations.

1. $\boldsymbol{r}(t) = \cos(4t)\mathbf{i} + \sin(4t)\mathbf{j} + 4t\mathbf{k}$, $t \in \big[0, \dfrac{\pi}{2}\big]$.
2. $\boldsymbol{r}(t) = \cos\dfrac{t}{2}\mathbf{i} + \sin\dfrac{t}{2}\mathbf{j} + \dfrac{t}{2}\mathbf{k}$, $t \in [0, 4\pi]$.
3. $\boldsymbol{r}(t) = \cos t\,\mathbf{i} - \sin t\,\mathbf{j} - t\mathbf{k}$, $t \in [-2\pi, 0]$.

Problem 12.5. Parametrize the curve

$$\mathbf{r} = \mathbf{r}(t) = \arctan\frac{t}{\sqrt{1-t^2}}\mathbf{i} + \arcsin t\,\mathbf{j} + \arccos t\,\mathbf{k}, \quad t \in \Big[-1, 0.5\Big],$$

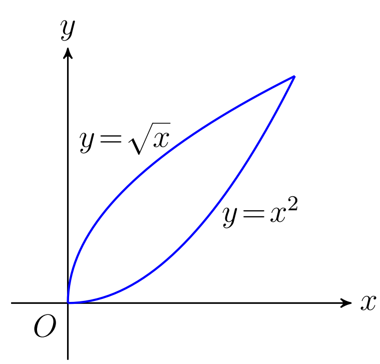
in the same orientation in terms of arc-length measured from the point where $t = 0$.

Problem 12.6. (15%) Parametrize the curve

$$\mathbf{r} = \mathbf{r}(t) = \arcsin\frac{t}{\sqrt{1+t^2}}\mathbf{i} + \arctan t\,\mathbf{j} + \arccos\frac{1}{\sqrt{1+t^2}}\mathbf{k}, \quad t \in [-1, 1],$$

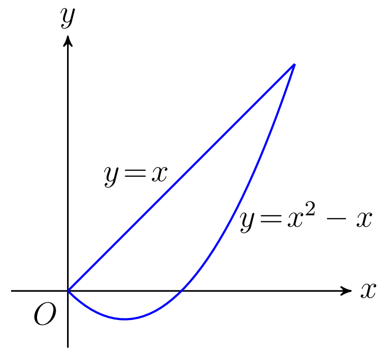
in the same orientation in terms of arc-length measured from the point where $t = 0$.

Problem 12.7. Give a parametrization of the simple closed curve C shown in the figure below



and find the area of the region enclosed by C using (12.4.1), (12.4.2) or (12.4.3).

Problem 12.8. Give a parametrization of the simple closed curve C shown in the figure below



and find the area of the region enclosed by C using (12.4.1), (12.4.2) or (12.4.3).

Problem 12.9. Let C_1 be the polar graph of the polar function $r = 1 + \cos\theta$ (which is a cardioid), and C_2 be the polar graph of the polar function $r = 3\cos\theta$ (which is a circle). See the following figure for reference.

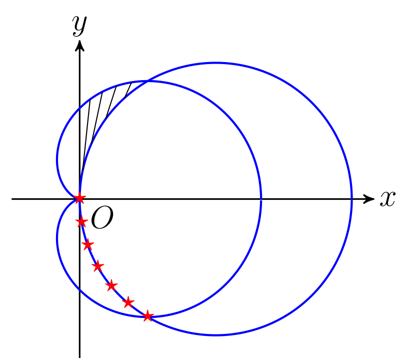


Figure 12.6: The polar graphs of the polar equations $r = 1 + \cos\theta$ and $r = 3\cos\theta$

(1) Find the intersection points of C_1 and C_2.

(2) Find the line L passing through the lowest intersection point and tangent to the curve C_2.

(3) Identify the curve marked by $\star$ on the θr-plane for $0 \leqslant \theta \leqslant 2\pi$.

(4) Find the area of the shaded region.

Problem 12.10. Let R be the region bounded by the lemniscate $r^2 = 2\cos 2\theta$ and is outside the circle $r = 1$ (see the shaded region in the graph).

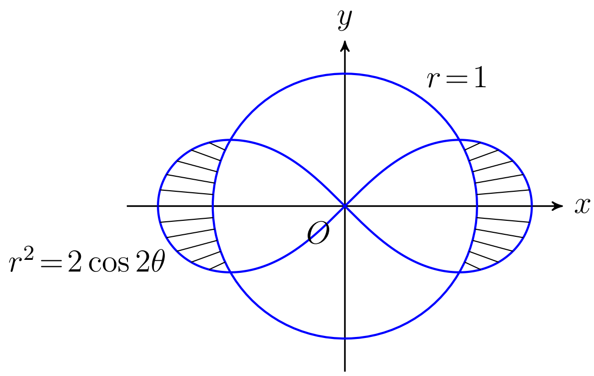


Figure 12.7: The polar graphs of the polar equations $r^2 = 2\cos 2\theta$ and $r = 1$

(1) Find the area of R.

(2) Find the slope of the tangent line passing thought the point on the lemniscate corresponding to $\theta = \dfrac{\pi}{6}$.

(3) Find the volume of the solid of revolution obtained by rotating R about the x-axis by complete the following:

(a) Suppose that (x, y) is on the lemniscate. Then (x, y) satisfies

$$y^4 + a(x)y^2 + b(x) = 0 \tag{12.5.1}$$

for some functions $a(x)$ and $b(x)$. Find $a(x)$ and $b(x)$.

(b) Solving (12.5.1), we find that $y^2 = c(x)$, where $c(x) = c_1x^2 + c_2 + c_3\sqrt{1+4x^2}$ for some constants c_1, c_2 and c_3. Then the volume of interests can be computed by

$$I = 2 \times \left[\pi \int_{\frac{\sqrt{3}}{2}}^{\sqrt{2}} c(x)dx - \pi \int_{\frac{\sqrt{3}}{2}}^{1} d(x)dx\right].$$

Compute $\displaystyle\int_{\frac{\sqrt{3}}{2}}^{1} \left[d(x) - (1-x^2)\right] dx$.

(c) Evaluate I by first computing the integral $\displaystyle\int_{\frac{\sqrt{3}}{2}}^{\sqrt{2}} \sqrt{1+4x^2}\, dx$, and then find I.

(4) Find the surface area of the surface of revolution obtained by rotating the boundary of R about the x-axis.

Problem 12.11. Let R be the region bounded by the circle $r = 1$ and outside the lemniscate $r^2 = -2\cos 2\theta$, and is located on the right half plane (see the shaded region in the graph).

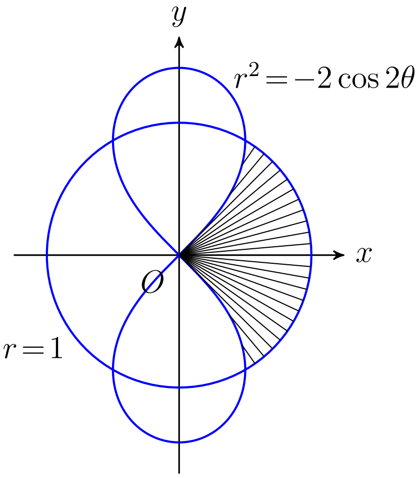


Figure 12.8: The polar graphs of the polar equations $r = 1$ and $r^2 = -2\cos 2\theta$

(1) Find the points of intersection of the circle $r = 1$ and the lemniscate $r^2 = -2\cos 2\theta$.

(2) Show that the straight line $x = \dfrac{1}{2}$ is tangent to the lemniscate at the points of intersection on the right half plane.

(3) Find the area of R.

(4) Find the volume of the solid of revolution obtained by rotating R about the x-axis by complete the following:

(a) Suppose that (x, y) is on the lemniscate. Then (x, y) satisfies

$$y^4 + a(x)y^2 + b(x) = 0 \tag{12.5.2}$$

for some functions $a(x)$ and $b(x)$. Find $a(x)$ and $b(x)$.

(b) Solving (12.5.2), we find that $y^2 = c(x)$, where $c(x) = c_1x^2 + c_2 + c_3\sqrt{1-4x^2}$ for some constants c_1, c_2 and c_3. Then the volume of interests can be computed by

$$I = \pi \int_0^{\frac{1}{2}} c(x)dx + \pi \int_{\frac{1}{2}}^1 d(x)dx.$$

Compute $\displaystyle\int_{\frac{1}{2}}^1 \Big[d(x) - (1 - x^2)\Big]dx$.

(c) Evaluate I by first computing the integral $\displaystyle\int_0^{\frac{1}{2}} \sqrt{1-4x^2}dx$, and then find I.

(5) Find the area of the surface of revolution obtained by rotating the boundary of R about the x-axis.

Problem 12.12. Let C_1, C_2 be the curves given by polar coordinate $r = 1 - 2\sin\theta$ and $r = 4 + 4\sin\theta$, respectively, and the graphs of C_1 and C_2 are given in Figure 12.9.

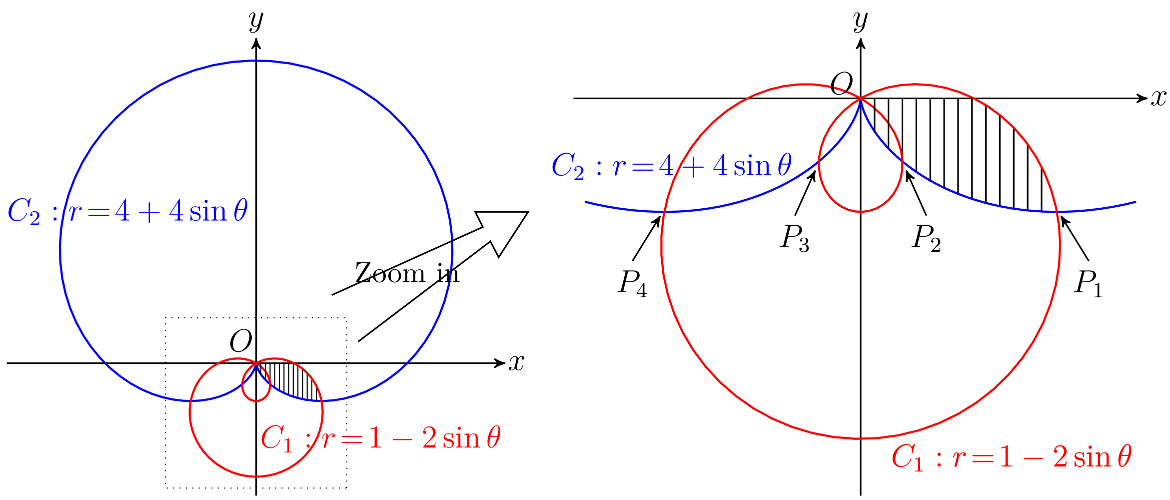


Figure 12.9: The polar graphs of the polar equations $r = 1 - 2\sin\theta$ and $r = 4 + 4\sin\theta$

(1) Let $P_1, \cdots, P_4$ be four points of intersection of curves C_1 and C_2 as shown in Figure 12.9 (the fifth one is the origin). What are the Cartesian coordinates of P_1 and P_2?

(2) Let L_1 and L_2 be two straight lines passing P_1 and tangent to C_1, C_2, respectively. Find the cosine value of the acute/smaller angle between L_1 and L_2.

(3) Compute the area of the shaded region.

Chapter 13

Functions of Several Variables

13.1 Introduction to Functions of Several Variables

Definition 13.1

Let D be a set of ordered pairs of real numbers. If to each ordered pair (x, y) in D there corresponds a unique real number $f(x, y)$, then f is a real-valued function of (two variables) x and y. The set D is the domain of f, and the corresponding set of values for $f(x, y)$ is the range of f. For the function $z = f(x, y)$, x and y are called the independent variables and z is called the dependent variable.

Definition 13.2

Let f, g be real-valued functions of two variables with domain D.

1. The sum of f and g, the difference of f and g and the product of f and g, denoted by $f + g$, $f - g$ and fg, are functions defined on D given by

$$\begin{aligned}(f+g)(x,y) &= f(x,y)+g(x,y) \quad \forall\,(x,y)\in D\,,\\ (f-g)(x,y) &= f(x,y)-g(x,y) \quad \forall\,(x,y)\in D\,,\\ (fg)(x,y) &= f(x,y)g(x,y) \qquad\quad \forall\,(x,y)\in D\,.\end{aligned}$$

2. The quotient of f and g, denoted by $\dfrac{f}{g}$, is a function defined on $D\backslash\big\{(x, y) \in D \,\big|\, g(x, y) = 0\big\}$ given by

$$\frac{f}{g}(x,y) = \frac{f(x,y)}{g(x,y)} \qquad \forall\,(x,y)\in D \text{ such that } g(x,y)\neq 0\,.$$

Remark 13.3. A function f of two variables should be given along with its domain. When the domain of a function is not specified, as before the domain should be treated as the collection of all (x, y) such that $f(x, y)$ is meaningful.

Definition 13.4

Let h be a real-valued function of two variables with domain D, and $g : I \to \mathbb{R}$ be a real-valued function (of one variable) on an interval I. The composite function of g and h, denoted by $g \circ h$, is a function defined on $D \cap \big\{(x, y) \in D \,\big|\, h(x, y) \in I\big\}$ given by

$$(g \circ h)(x, y) = g\big(h(x, y)\big) \qquad \forall\,(x, y) \in D \text{ such that } h(x, y) \in I\,.$$

Similar concepts such as real-valued functions of three variables, the sum, different, product, quotient and composition of functions of three variables can be defined accordingly.

Definition 13.5

Let D be a set of ordered pairs of real numbers, and $f : D \to \mathbb{R}$ be a real-valued function of two variables. The graph of f is the set of all points (x, y, z) for which $z = f(x, y)$ and $(x, y) \in D$.

Example 13.6. Let $r > 0$ be a real number. The graph of the function $z = f(x, y) = \sqrt{r^2 - x^2 - y^2}$ is the upper hemi-sphere of the sphere centered at the origin with radius r. On the other hand, the graph of the function $z = g(x, y) = -\sqrt{r^2 - x^2 - y^2}$ is the lower hemi-sphere of the sphere.

Definition 13.7: Level Curves

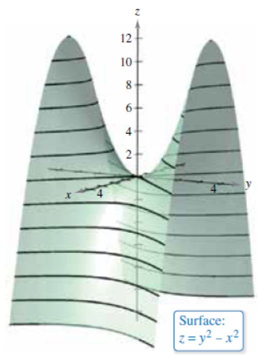
Let D be a set of ordered pairs of real numbers, and $f : D \to \mathbb{R}$ be a function of two variables. A level curve (or contour curve) of f is a collection of points (x, y) in D along which the value of $f(x, y)$ is constant.

Definition 13.8: Level Surfaces

Let D be a set of ordered pairs of real numbers, and $f : D \to \mathbb{R}$ be a function of three variables. A level surface of f is a collection of points (x, y, z) in D along which the value of $f(x, y, z)$ is constant.

Example 13.9. A level curve of the function $z = \sqrt{r^2 - x^2 - y^2}$ is a circle centered at the origin, and a level surface of the function $w = g(x, y, z) = x^2 + y^2 + z^2 - r^2$ is a sphere centered at the origin.

Example 13.10. The graph of $f(x, y) = y^2 - x^2$ is called a hyperbolic paraboloid. A level curve of a hyperbolic paraboloid is a hyperbola (or degenerated hyperbola), and each plane perpendicular to the xy-plane intersects the graph of $z = f(x, y)$ along a parabola (or degenerated parabola).



Hyperbolic paraboloid

13.2 Limits and Continuity

Definition 13.11

Let $\delta > 0$ be given. The δ-neighborhood about a point (x_0, y_0) in the plane is the open disk centered at (x_0, y_0) with radius δ given by

$$D\big((x_0, y_0), \delta\big) \equiv \Big\{(x, y) \,\Big|\, \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta\Big\}.$$

Definition 13.12

Let R be a collection of points in the plane. A point (x_0, y_0) (in R) is called an ***interior point*** of R if there exists $\delta > 0$ such that the δ-neighborhood about (x_0, y_0) lies entirely in R. If every point in R is an interior point of R, then R is called an open region. A point (x_0, y_0) is called a ***boundary point*** of R if every δ-neighborhood about (x_0, y_0) containing points inside R and point outsides R. In other words, (x_0, y_0) is a boundary point of R if

$$\forall\, \delta > 0, D\big((x_0, y_0), \delta\big) \cap R \neq \varnothing \quad \text{and} \quad D\big((x_0, y_0), \delta\big) \cap R^{\complement} \neq \varnothing\,.$$

If R contains all its boundary points, then R is called a closed region.

Remark 13.13. For $x \in \mathbb{R}$ and $\delta > 0$, let $D(x,\delta)$ denote the interval $(x-\delta, x+\delta)$ (and called the interval centered at x with radius r). Then for each $x \in (a,b)$, there exists $\delta > 0$ such that $D(x,r) \subseteq (a,b)$; thus (a,b) is called an open interval. The end-points a, b of the interval are boundary points of the interval, and $[a,b]$ is a closed interval since it contains all its boundary points.

Definition 13.14

Let f be a real-valued function of two variables defined, except possibly at (x_0, y_0), on an open disk centered at (x_0, y_0), and let L be a real number. Then

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L$$

if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big|f(x) - L\big| < \varepsilon \quad \text{whenever} \quad 0 < \sqrt{(x-x_0)^2 + (y-y_0)^2} < \delta\,.$$

Remark 13.15. If $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = L_1$ and $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = L_2$, then $L_1 = L_2$. In other words, the limit is unique when it exists.

The proof of the following is almost identical to the one of Theorem 1.14.

Theorem 13.16: Properties of Limits of Functions of Two Variables

Let $(a,b) \in \mathbb{R}^2$. Suppose that the limits

$$\lim_{(x,y)\to(a,b)} f(x,y) = L \qquad \text{and} \qquad \lim_{(x,y)\to(a,b)} g(x,y) = K\,.$$

both exist, and c is a constant.

1. $\lim\limits_{(x,y)\to(a,b)} c = c$, $\lim\limits_{(x,y)\to(a,b)} x = a$ and $\lim\limits_{(x,y)\to(a,b)} y = b$.

2. $\lim\limits_{(x,y)\to(a,b)} \big[f(x,y) \pm g(x,y)\big] = L + K$;

3. $\lim\limits_{(x,y)\to(a,b)} \big[f(x,y)g(x,y)\big] = LK$;

4. $\lim\limits_{(x,y)\to(a,b)} \dfrac{f(x,y}{g(x,y)} = \dfrac{L}{K}$ if $K \neq 0$.

Theorem 13.17: Squeeze

Let $(x_0, y_0) \in \mathbb{R}^2$. Suppose that f, g, h are functions of two variables such that

$$g(x, y) \leqslant f(x, y) \leqslant h(x, y)$$

except possible at (x_0, y_0), and $\lim\limits_{(x,y)\to(x_0,y_0)} g(x, y) = \lim\limits_{(x,y)\to(x_0,y_0)} h(x, y) = L$, then

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L\,.$$

Example 13.18. For $(a, b) \in \mathbb{R}^2$, find the limit $\lim\limits_{(x,y)\to(a,b)} \dfrac{5x^2y}{x^2 + y^2}$.

First we note that 1-3 of Theorem 13.16 implies that the function $f(x, y) = 5x^2y$ and $g(x, y) = x^2 + y^2$ has the property that

$$\lim_{(x,y)\to(a,b)} f(x, y) = 5a^2b \quad \text{and} \quad \lim_{(x,y)\to(a,b)} g(x, y) = a^2 + b^2\,.$$

Therefore, Theorem 13.16 again shows the following:

1. If $(a, b) \neq (0, 0)$, then 4 of Theorem 13.16 implies that

$$\lim_{(x,y)\to(a,b)} \frac{5x^2y}{x^2 + y^2} = \lim_{(x,y)\to(a,b)} \frac{f(x, y)}{g(x, y)} = \frac{5a^2b}{a^2 + b^2}\,.$$

2. If $(a, b) = (0, 0)$, then we cannot apply 4 of Theorem 13.16 to compute the limit. Nevertheless, since

$$\Big|\frac{5x^2y}{x^2 + y^2} - 0\Big| \leqslant 5|y| \qquad \forall\,(x, y) \neq (0, 0)\,,$$

the Squeeze Theorem implies that

$$\lim_{(x,y)\to(0,0)} \frac{5x^2y}{x^2 + y^2} = 0\,.$$

Example 13.19. Show that the limit $\lim\limits_{(x,y)\to(0,0)} \Big(\dfrac{x^2 - y^2}{x^2 + y^2}\Big)^2$ does not exist.

Let $f(x, y) = \Big(\dfrac{x^2 - y^2}{x^2 + y^2}\Big)^2$. By the definition of limits, if $\lim\limits_{(x,y)\to(0,0)} f(x, y) = L$ exists, then there exists $\delta > 0$ such that

$$\big|f(x, y) - L\big| < \frac{1}{2} \quad \text{whenever} \quad 0 < \sqrt{x^2 + y^2} < \delta$$

which implies that

$$L-\frac{1}{2}<f(x,y)<L+\frac{1}{2} \quad \text{whenever} \quad 0<\sqrt{x^2+y^2}<\delta\,. \tag{13.2.1}$$

However, when (x,y) satisfies $0<\sqrt{x^2+y^2}<\delta$ and $x=y$, then $f(x,y)=0$ while on the other hand, when (x,y) satisfies $0<\sqrt{x^2+y^2}<\delta$ and $y=0$, then $f(x,y)=1$. This is a contradiction because of (13.2.1).

• Another way of looking at this limit: When (x,y) approaches $(0,0)$ along the line $x=y$ (we use the notation $\lim\limits_{\substack{(x,y)\to(0,0)\\ x=y}}$ to denote this limit process), we find that

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

and when (x,y) approaches $(0,0)$ along the x-axis (we use the notation $\lim\limits_{\substack{(x,y)\to(0,0)\\ y=0}}$ to denote this limit process), we find that

$$\lim_{\substack{(x,y)\to(0,0)\\ y=0}} f(x,y)=1\,.$$

The uniqueness of the limit implies that the limit of f at $(0,0)$ does not exist.

13.2.1 Continuity of functions of two variables

Definition 13.20

A function f of two variables is ***continuous at a point*** (x_0,y_0) in an open region R if $f(x_0,y_0)$ is defined and is equal to the limit of $f(x,y)$ as (x,y) approaches (x_0,y_0); that is,

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)=f(x_0,y_0)\,.$$

In other words, f is continuous at (x_0,y_0) if for every $\varepsilon>0$ there exists $\delta>0$ such that

$$\big|f(x,y)-f(x_0,y_0)\big|<\varepsilon \quad \text{whenever} \quad \sqrt{(x-x_0)^2+(y-y_0)^2}<\delta\,.$$

The function f is ***continuous in the open region*** R if it is continuous at every point in R.

Remark 13.21. 1. Unlike the case that f does not have to be defined at (x_0,y_0) in order to consider the limit of f at (x_0,y_0), for f to be continuous at a point (x_0,y_0) f has to be defined at (x_0,y_0).

2. A point (x_0, y_0) is called a discontinuity of f if f is not continuous at (x_0, y_0). (x_0, y_0) is called a ***removable discontinuity*** of f if $\lim_{(x,y)\to(x_0,y_0)} f(x, y)$ exists.

Theorem 13.22

Let f and g be functions of two variables such that f and g are continuous at (x_0, y_0).

1. $f \pm g$ is continuous at (x_0, y_0).
2. fg is continuous at (x_0, y_0).
3. $\dfrac{f}{g}$ is continuous at (x_0, y_0) if $g(x_0, y_0) \neq 0$.

Theorem 13.23

If h is continuous at (x_0, y_0) and g is continuous at $h(x_0, y_0)$, then the composite function $g \circ h$ is continuous at (x_0, y_0); that is,

$$\lim_{(x,y)\to(x_0,y_0)} (g \circ h)(x, y) = g\big(h(x_0, y_0)\big) .$$

13.3 Partial Derivatives

Definition 13.24

Let f be a function of two variable. The first partial derivative of f with respect to x at (x_0, y_0), denoted by $f_x(x_0, y_0)$, is defined by

$$f_x(x_0, y_0) = \lim_{\Delta x\to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}$$

provided the limit exists. The first partial derivative of f with respect to y at (x_0, y_0), denoted by $f_y(x_0, y_0)$, is defined by

$$f_y(x_0, y_0) = \lim_{\Delta y\to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}$$

provided the limit exists. When f_x and f_y exist for all (x_0, y_0) (in a certain open region), f_x and f_y are simply called the first partial derivative of f with respect to x and y, respectively.

• **Notation**: For $z = f(x,y)$, the partial derivative f_x and f_y, can also be denoted by

$$\frac{\partial}{\partial x}f(x,y) = f_x(x,y) = z_x = \frac{\partial z}{\partial x} = \frac{\partial f}{\partial x}(x,y)$$

and

$$\frac{\partial}{\partial y}f(x,y) = f_y(x,y) = z_y = \frac{\partial z}{\partial y} = \frac{\partial f}{\partial y}(x,y)\,.$$

When evaluating the partial derivative at (x_0, y_0), we write

$$f_x(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0) = \frac{\partial}{\partial x}\Big|_{(x,y)=(x_0,y_0)} f(x,y)$$

and

$$f_y(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = \frac{\partial}{\partial y}\Big|_{(x,y)=(x_0,y_0)} f(x,y)\,.$$

Example 13.25. For $f(x,y) = xe^{x^2y}$, find f_x and f_y, and evaluate each at the point $(1, \ln 2)$.

Note that f_x is obtained by treating y as a constant and differentiate f with respect to x. Therefore, the product rule implies tat

$$f_x(x,y) = \Big(\frac{\partial}{\partial x}x\Big)e^{x^2y} + x\Big(\frac{\partial}{\partial x}e^{x^2y}\Big) = e^{x^2y} + x\cdot e^{x^2y}\cdot 2xy = (1+2x^2y)e^{x^2y}\,;$$

thus

$$f_x(1,\ln 2) = (1+2\ln 2)e^{\ln 2} = 2(1+2\ln 2)\,.$$

Similarly,

$$f_y(x,y) = \Big(\frac{\partial}{\partial y}x\Big)e^{x^2y} + x\Big(\frac{\partial}{\partial y}e^{x^2y}\Big) = x^3e^{x^2y}\,;$$

thus $f_y(1,\ln 2) = e^{\ln 2} = 2$.

Example 13.26. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

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

Then if $(x,y) \neq (0,0)$, we can apply the quotient rule (and product rule) to compute the partial derivatives and obtain that

$$\begin{aligned} f_x(x,y) &= \frac{(x^2+y^2)\frac{\partial}{\partial x}\big[xy(x^2-y^2)\big] - xy(x^2-y^2)\frac{\partial}{\partial x}(x^2+y^2)}{(x^2+y^2)^2} \\ &= \frac{(x^2+y^2)\big[y(x^2-y^2)+2x^2y\big] - xy(x^2-y^2)\cdot(2x)}{(x^2+y^2)^2} \\ &= \frac{x^4y+4x^2y^3-y^5}{(x^2+y^2)^2}\,. \end{aligned}$$

If $(x, y) = (0, 0)$, we cannot use the quotient rule to compute the derivative since the denominate is 0 (so that 4 of Theorem 13.16 cannot be applied), and we have to compute $f_x(0, 0)$ using the definition. By definition,

$$f_x(0,0) = \lim_{\Delta x \to 0} \frac{f(\Delta x, 0) - f(0,0)}{\Delta x} = 0\,.$$

Therefore,

$$f_x(x,y) = \begin{cases} \dfrac{x^4 y + 4x^2 y^3 - y^5}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

Similarly,

$$f_y(x,y) = \begin{cases} \dfrac{x^5 - 4x^3 y^2 - x y^4}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

• **Geometric meaning of partial derivatives**: Let $f(x, y)$ be a function of two variable, (x_0, y_0) be given, and $z_0 = f(x_0, y_0)$. Consider the graph of the function $z = f(x, y_0)$ (of one variable) on the xz-plane. If the graph $z = f(x, y_0)$ has a tangent line at (x_0, z_0), then the slope of the tangent line at (x_0, z_0) is given by

$$\lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

and this limit, if exists, is $f_x(x_0, y_0)$. This is called ***the slopes in the x-direction of the surface*** $z = f(x, y)$ ***at the point*** (x_0, y_0, z_0). Similarly, the slope of the tangent line of the graph of $z = f(x_0, y)$ at (y_0, z_0) is $f_y(x_0, y_0)$, and is called ***the slopes in the y-direction of the surface*** $z = f(x, y)$ ***at the point*** (x_0, y_0, z_0).

• **Partial derivatives of functions of three or more variables:**

The concept of partial derivatives can be extended to functions of three or more variables. For example, if $w = f(x, y, z)$, then

$$\begin{aligned} \frac{\partial w}{\partial x} &= f_x(x,y,z) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y, z) - f(x,y,z)}{\Delta x}\,, \\ \frac{\partial w}{\partial y} &= f_y(x,y,z) = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y, z) - f(x,y,z)}{\Delta y}\,, \\ \frac{\partial w}{\partial z} &= f_z(x,y,z) = \lim_{\Delta z \to 0} \frac{f(x, y, z+\Delta z) - f(x,y,z)}{\Delta z}\,. \end{aligned}$$

In general, if $w = f(x_1, x_2, \cdots, x_n)$, then there are n first partial derivatives denoted by

$$\frac{\partial w}{\partial x_k} = f_{x_k}(x_1, x_2, \cdots, x_n)\,, \quad k = 1, 2, \cdots, n\,.$$

• **Higher-order partial derivatives:**

We can also take higher-order partial derivatives of functions of several variables. For example, for $z = f(x, y)$,

1. Differentiate twice with respect to x:
$$\frac{\partial}{\partial x}\Big(\frac{\partial f}{\partial x}\Big) = \frac{\partial^2 f}{\partial x^2} = f_{xx} \,.$$
2. Differentiate twice with respect to y:
$$\frac{\partial}{\partial y}\Big(\frac{\partial f}{\partial y}\Big) = \frac{\partial^2 f}{\partial y^2} = f_{yy} \,.$$
3. Differentiate first with respect to x and then with respect to y:
$$\frac{\partial}{\partial y}\Big(\frac{\partial f}{\partial x}\Big) = \frac{\partial^2 f}{\partial y \partial x} = f_{xy} \,.$$
4. Differentiate first with respect to y and then with respect to x:
$$\frac{\partial}{\partial x}\Big(\frac{\partial f}{\partial y}\Big) = \frac{\partial^2 f}{\partial x \partial y} = f_{yx} \,.$$

The third and fourth cases are called ***mixed partial derivatives***.

Example 13.27. In this example we compute the second partial derivatives of the function given in 13.26. We have obtained that

$$f_x(x, y) = \begin{cases} \dfrac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0)\,, \\ 0 & \text{if } (x, y) = (0, 0)\,, \end{cases}$$

and

$$f_y(x, y) = \begin{cases} \dfrac{x^5 - 4x^3 y^2 - xy^4}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0)\,, \\ 0 & \text{if } (x, y) = (0, 0)\,. \end{cases}$$

If $(x, y) \neq (0, 0)$, the quotient rule, the product rule and the chain rule (for functions of one variable) together show that

$$\begin{aligned} f_{xx}(x, y) &= \frac{(x^2 + y^2)^2 \frac{\partial}{\partial x}(x^4 y + 4x^2 y^3 - y^5) - (x^4 y + 4x^2 y^3 - y^5) \frac{\partial}{\partial x}(x^2 + y^2)^2}{(x^2 + y^2)^4} \\ &= \frac{(x^2 + y^2)^2 (4x^3 y + 8xy^3) - (x^4 y + 4x^2 y^3 - y^5) \cdot \big[2(x^2 + y^2) \cdot (2x)\big]}{(x^2 + y^2)^3} \\ &= \frac{(x^2 + y^2)(4x^3 y + 8xy^3) - 4x(x^4 y + 4x^2 y^3 - y^5)}{(x^2 + y^2)^3} = \frac{-4x^3 y^3 + 12xy^5}{(x^2 + y^2)^3} \,. \end{aligned}$$

Similarly, if $(x,y) \neq (0,0)$,

$$
\begin{aligned}
f_{yy}(x,y) &= \frac{(x^2+y^2)^2(-8x^3y-4xy^3)-(x^5-4x^3y^2-xy^4)\cdot\left[2(x^2+y^2)\cdot(2y)\right]}{(x^2+y^2)^2} \\
&= \frac{-12x^5y+4x^3y^3}{(x^2+y^2)^3}\,, \\
f_{xy}(x,y) \quad &= \frac{(x^2+y^2)(x^4+12x^2y^2-5y^4)-4y(x^4y+4x^2y^3-y^5)}{(x^2+y^2)^3} \\
&= \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3}
\end{aligned}
$$

and

$$
\begin{aligned}
f_{yx}(x,y) &= \frac{(x^2+y^2)(5x^4-12x^2y^2-y^4)-4x(x^5-4x^3y^2-xy^4)}{(x^2+y^2)^3} \\
&= \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3}\,.
\end{aligned}
$$

We note that when $(x,y) \neq (0,0)$, $f_{xy}(x,y) = f_{yx}(x,y)$.

Since $f_x(x,0) = f_y(0,y) = 0$ for all $x \neq 0$, we find that

$$
f_{xx}(0,0) = \lim_{\Delta x \to 0} \frac{f_x(\Delta x,0)-f_x(0,0)}{\Delta x} = 0
$$

and

$$
f_{yy}(0,0) = \lim_{\Delta y \to 0} \frac{f_y(0,\Delta y)-f_y(0,0)}{\Delta y} = 0\,.
$$

Finally, we compute $f_{xy}(0,0)$ and $f_{yx}(0,0)$. By definition,

$$
f_{xy}(0,0) = \lim_{\Delta y \to 0} \frac{f_x(0,\Delta y)-f_x(0,0)}{\Delta y} = \lim_{\Delta y \to 0} \frac{\frac{-\Delta y^5}{\Delta y^4}}{\Delta y} = -1
$$

and

$$
f_{yx}(0,0) = \lim_{\Delta x \to 0} \frac{f_y(\Delta x,0)-f_y(0,0)}{\Delta x} = \lim_{\Delta y \to 0} \frac{\frac{\Delta x^5}{\Delta x^4}}{\Delta x} = 1\,.
$$

We note that $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Theorem 13.28: Clairaut's Theorem

If f is a function of x and y such that f_{xy} and f_{yx} are continuous on an open disk D, then

$$
f_{xy}(x,y) = f_{yx}(x,y) \qquad \forall\,(x,y) \in D\,.
$$

In the following, we prove the following more general version:

> If f is a function of x and y such that on an open disk D f_{xy} is continuous and f_{yx} exists, then $f_{xy}(x,y) = f_{yx}(x,y)$ for all $(x,y) \in D$.

Proof. Let $(a,b) \in D$ be given. Then

$$\begin{aligned} f_{yx}(a,b) = (f_y)_x(a,b) &= \lim_{h\to 0} \frac{f_y(a+h,b) - f_y(a,b)}{h} \\ &= \lim_{h\to 0} \frac{\displaystyle\lim_{k\to 0} \frac{f(a+h,b+k) - f(a+h,b)}{k} - \lim_{k\to 0} \frac{f(a,b+k) - f(a,b)}{k}}{h} \\ &= \lim_{h\to 0} \lim_{k\to 0} \frac{f(a+h,b+k) - f(a,b+k) - f(a+h,b) - f(a,b)}{hk} \,. \end{aligned}$$

Define

$$Q(h,k) \equiv \frac{f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b)}{hk} \,.$$

Then the computation above shows that

$$\lim_{h\to 0} \lim_{k\to 0} Q(h,k) = f_{yx}(a,b) \,. \tag{13.3.1}$$

For $h,k \neq 0$ such that $(a+h,b+k) \in D$, define $\varphi(x,y) = f(x,y+k) - f(x,y)$. Then $Q(h,k) = \dfrac{\varphi(a+h,b) - \varphi(a,b)}{hk}$. By the mean value theorem for functions of one variable (Theorem 3.9),

$$Q(h,k) = \frac{\varphi_x(a+\theta_1 h,b)h}{hk} = \frac{f_x(a+\theta_1 h,b+k) - f_x(a+\theta_1 h,b)}{k}$$

for some functions $\theta_1 = \theta_1(h)$ satisfying $0 < \theta_1 < 1$. Applying the mean value theorem again,

$$\begin{aligned} Q(h,k) &= \frac{f_x(a+\theta_1 h,b+k) - f_x(a+\theta_1 h,b)}{k} = \frac{f_{xy}(a+\theta_1 h,b+\theta_2 k)k}{k} \\ &= f_{xy}(a+\theta_1 h,b+\theta_2 k) \end{aligned}$$

for some functions $\theta_2 = \theta_2(h,k)$ satisfying $0 < \theta_2 < 1$. Therefore, we establish that there exist functions $\theta_1 = \theta_1(h)$ and $\theta_2 = \theta_2(h,k)$ such that

$$Q(h,k) = f_{xy}(x+\theta_1 h, y+\theta_2 k) \,.$$

Passing to the limit as $k \to 0$ first then $h \to 0$, using (13.3.1) and the continuity of f_{xy} we conclude that $f_{xy}(a,b) = f_{yx}(a,b)$. $\square$

Example 13.29. Let $f(x, y, z) = ye^x + x \ln z$. Then $f_x(x, y, z) = ye^x + \ln z$, $f_y(x, y, z) = e^x$ and $f_z(x, y, z) = \dfrac{x}{z}$. Therefore,

$$\begin{aligned}
f_{xy}(x, y, z) &= e^x = f_{yx}(x, y, z)\,, \\
f_{xz}(x, y, z) &= \frac{1}{z} = f_{zx}(x, y, z) \quad \forall\, z \neq 0\,, \\
f_{yz}(x, y, z) &= 0 = f_{zy}(x, y, z)\,.
\end{aligned}$$

13.4 Differentiability of Functions of Several Variables

Recall that a function $f : (a, b) \to \mathbb{R}$ is said to be differentiable at a point $c \in (a, b)$ if the limit

$$\lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x}$$

exists. The differentiability of f at c can be rephrased as follows:

A function $f : (a, b) \to \mathbb{R}$ is said to be differentiable at $c \in (a, b)$ if there exists $m \in \mathbb{R}$ such that

$$\lim_{\Delta x \to 0} \left| \frac{f(c + \Delta x) - f(c) - m\Delta x}{\Delta x} \right| = 0\,.$$

or equivalently,

$$\lim_{x \to c} \left| \frac{f(x) - f(c) - m(x - c)}{x - c} \right| = 0\,.$$

This equivalent way of defining differentiability of functions of one variable motivate the following

Definition 13.30

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. For $(x_0, y_0) \in R$, f is said to be differentiable at (x_0, y_0) if there exist real numbers A, B such that

$$\lim_{(x,y) \to (x_0,y_0)} \frac{\big|f(x, y) - f(x_0, y_0) - (A, B) \cdot (x - x_0, y - y_0)\big|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0\,.$$

Suppose that f is differentiable at (x_0, y_0). When (x, y) approaches (x_0, y_0) along the line $y = y_0$, we find that

$$\begin{aligned}
0 &= \lim_{\substack{(x,y)\to(x_0,y_0)\\ y=y_0}} \frac{\big|f(x,y) - f(x_0,y_0) - A(x-x_0) - B(y-y_0)\big|}{\sqrt{(x-x_0)^2+(y-y_0)^2}} \\
&= \lim_{x\to x_0} \frac{\big|f(x,y_0) - f(x_0,y_0) - A(x-x_0)\big|}{|x-x_0|} = \lim_{x\to x_0} \Big|\frac{f(x,y_0)-f(x_0,y_0)}{x-x_0} - A\Big|
\end{aligned}$$

which implies that the number A must be $f_x(x_0,y_0)$. Similarly, $B = f_y(x_0,y_0)$, and we have the following alternative

Definition 13.31

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. For $(x_0,y_0) \in R$, f is said to be differentiable at (x_0,y_0) if ($f_x(x_0,y_0)$, $f_y(x_0,y_0)$ both exist and)

$$\lim_{(x,y)\to(x_0,y_0)} \frac{\big|f(x,y) - f(x_0,y_0) - f_x(x_0,y_0)(x-x_0) - f_y(x_0,y_0)(y-y_0)\big|}{\sqrt{(x-x_0)^2+(y-y_0)^2}} = 0\,.$$

Remark 13.32. The ordered pair $\big(f_x(x_0,y_0), f_y(x_0,y_0)\big)$ if called the derivative of f at (x_0,y_0) if f is differentiable at (x_0,y_0) and is usually denoted by $(Df)(x_0,y_0)$.

2. Using ε-δ notation, we find that f is differentiable at (x_0,y_0) if for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\begin{aligned}
&\big|f(x,y) - f(x_0,y_0) - f_x(x_0,y_0)(x-x_0) - f_y(x_0,y_0)(y-y_0)\big| \\
&\qquad \leqslant \varepsilon\sqrt{(x-x_0)^2+(y-y_0)^2} \qquad \text{whenever} \quad \sqrt{(x-x_0)^2+(y-y_0)^2} < \delta\,.
\end{aligned}$$

Now suppose that f is a function of two variables such that $f_x(x_0,y_0)$ and $f_y(x_0,y_0)$ exist. Define

$$\varepsilon(x,y) = \begin{cases} \dfrac{f(x,y) - f(x_0,y_0) - f_x(x_0,y_0)(x-x_0) - f_y(x_0,y_0)(y-y_0)}{\sqrt{(x-x_0)^2+(y-y_0)^2}} & \text{if } (x,y) \neq (x_0,y_0)\,, \\ 0 & \text{if } (x,y) = (x_0,y_0)\,. \end{cases}$$

Let $\Delta x = x - x_0$, $\Delta y = y - y_0$ and $\Delta z = f(x,y) - f(x_0,y_0)$. Then

$$\Delta z = f_x(x_0,y_0)\Delta x + f_y(x_0,y_0)\Delta y + \varepsilon(x,y)\sqrt{\Delta x^2 + \Delta y^2}\,,$$

and f is differentiable at (x_0,y_0) if and only if $\lim\limits_{(x,y)\to(x_0,y_0)} \varepsilon(x,y) = 0$.

Finally, define

$$\varepsilon_1(x,y) = \begin{cases} \dfrac{\varepsilon(x,y)\Delta x}{\sqrt{\Delta x^2 + \Delta y^2}} & \text{if } (x,y) \neq (x_0,y_0)\,, \\ 0 & \text{if } (x,y) \neq (x_0,y_0)\,, \end{cases}$$
$$\varepsilon_2(x,y) = \begin{cases} \dfrac{\varepsilon(x,y)\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}} & \text{if } (x,y) \neq (x_0,y_0)\,, \\ 0 & \text{if } (x,y) \neq (x_0,y_0)\,, \end{cases} ,$$

then

$$0 \leqslant |\varepsilon_1(x,y)|, |\varepsilon_2(x,y)| \leqslant |\varepsilon(x,y)| = \sqrt{\varepsilon_1(x,y)^2 + \varepsilon_2(x,y)^2}$$

thus the Squeeze Theorem shows that

$$\lim_{(x,y)\to(x_0,y_0)} \varepsilon(x,y) = 0 \quad \text{if and only if} \quad \lim_{(x,y)\to(x_0,y_0)} \varepsilon_1(x,y) = \lim_{(x,y)\to(x_0,y_0)} \varepsilon_2(x,y) = 0\,.$$

By the fact that $\varepsilon(x,y)\sqrt{\Delta x^2 + \Delta y^2} = \varepsilon_1(x,y)\Delta x + \varepsilon_2(x,y)\Delta y$, the alternative definition above can be rewritten as

> Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. For $(x_0, y_0) \in R$, f is said to be differentiable at (x_0, y_0) if ($f_x(x_0, y_0)$, $f_y(x_0, y_0)$ both exist and) there exist functions ε_1 and ε_2 such that
>
> $$\Delta z = f_x(x_0,y_0)\Delta x + f_y(x_0,y_0)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y\,,$$
>
> where both ε_1 and ε_2 approaches 0 as $(x, y) \to (x_0, y_0)$.

Example 13.33. Show that the function $f(x,y) = x^2 + 3y$ is differentiable at every point in the plane.

Let $(a,b) \in \mathbb{R}^2$ be given. Then $f_x(a,b) = 2a$ and $f_y(a,b) = 3$. Therefore,

$$\begin{aligned} \Delta z - f_x(a,b)\Delta x - f_y(a,b)\Delta y &= x^2 + 3y - a^2 - 3b - 2a(x-a) - 3(y-b) \\ &= (x-a)^2 = \varepsilon_1(x,y)\Delta x + \varepsilon_2(x,y)\Delta y\,, \end{aligned}$$

where $\varepsilon_1(x,y) = x - a$ and $\varepsilon_2(x,y) = 0$. Since

$$\lim_{(x,y)\to(a,b)} \varepsilon_1(x,y) = 0 \qquad \text{and} \qquad \lim_{(x,y)\to(a,b)} \varepsilon_2(x,y) = 0\,,$$

by the definition we find that f is differentiable at (a,b).

Example 13.34. The function f given in Example 13.26 is differentiable at $(0,0)$ since if $(x,y) \neq (0,0)$,

$$\frac{\big|f(x,y)-f(0,0)-f_x(0,0)x-f_y(0,0)y\big|}{\sqrt{x^2+y^2}} = \frac{\big|xy(x^2-y^2)\big|}{(x^2+y^2)^{\frac{3}{2}}} \leqslant \frac{|x^2-y^2|}{\sqrt{x^2+y^2}} \leqslant |x|+|y|$$

and the Squeeze Theorem shows that

$$\lim_{(x,y)\to(0,0)} \frac{\big|f(x,y)-f(0,0)-f_x(0,0)(x-0)-f_y(0,0)(y-0)\big|}{\sqrt{x^2+y^2}} = 0\,.$$

• Differentiability of functions of several variables

A real-valued function f of n variables is differentiable at $(a_1, a_2, \cdots, a_n)$ if there exist n real numbers $A_1, A_2, \cdots, A_n$ such that

$$\lim_{(x_1,\cdots,x_n)\to(a_1,\cdots,a_n)} \frac{\big|f(x_1,\cdots,x_n)-f(a_1,\cdots,a_n)-(A_1,\cdots,A_n)\cdot(x_1-a_1,\cdots,x_n-a_n)\big|}{\sqrt{(x_1-a_1)^2+\cdots+(x_n-a_n)^2}} = 0\,.$$

We also note that when f is differentiable at $(a_1, \cdots, a_n)$, then these numbers $A_1, A_2, \cdots, A_n$ must be $f_{x_1}(a_1, \cdots, a_n)$, $f_{x_2}(a_1, \cdots, a_n)$, $\cdots$, $f_{x_n}(a_1, \cdots, a_n)$, respectively.

It is usually easier to compute the partial derivatives of a function of several variables than determine the differentiability of that function. Is there any connection between some specific properties of partial derivatives and the differentiability? We have the following

Theorem 13.35

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. If f_x and f_y are continuous in a neighborhood of $(x_0, y_0) \in R$, then f is differentiable at (x_0, y_0). In particular, if f_x and f_y are continuous on R, then f is differentiable on R; that is, f is said to be differentiable at every point in R.

Therefore, the differentiability of f in Example 13.26 at any point $(x_0, y_0) \neq (0,0)$ can be guaranteed since f_x and f_y are continuous on $\mathbb{R}^2\backslash\{(0,0)\}$.

Theorem 13.36

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. If f is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0).

Proof. By the definition of differentiability, if f is differentiable at (x_0, y_0), then there exists function ε_1 and ε_2 such that

$$\lim_{(x,y)\to(x_0,y_0)} \varepsilon_1(x,y) = \lim_{(x,y)\to(x_0,y_0)} \varepsilon_w(x,y) = 0$$

and

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &+ \varepsilon_1(x,y)(x-x_0) + \varepsilon_2(x,y)(y-y_0)\,. \end{aligned}$$

Then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$. □

Example 13.37. Consider the function

$$f(x,y) = \begin{cases} \dfrac{-3xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

Then f is not continuous at $(0,0)$ since

$$\lim_{\substack{(x,y)\to(0,0) \\ y=0}} f(x,y) = 0 \qquad \text{but} \qquad \lim_{\substack{(x,y)\to(0,0) \\ x=y}} f(x,y) = -\frac{3}{2}\,.$$

However, we note that

$$f_x(0,0) = \lim_{\Delta x\to 0} \frac{f(\Delta x,0) - f(0,0)}{\Delta x} = 0 \quad \text{and} \quad f_y(0,0) = \lim_{\Delta y\to 0} \frac{f(0,\Delta y) - f(0,0)}{\Delta y} = 0\,.$$

Therefore, the existence of partial derivatives at a point in all directions does **not** even imply the continuity.

13.5 Chain Rules for Functions of Several Variables

Recall the chain rule for functions of one variable:

> Let I, J be open intervals, $f : J \to \mathbb{R}$, $g : I \to \mathbb{R}$ be real-valued functions, and the range of g is contained in J. If g is differentiable at $c \in I$ and f is differentiable at $g(c)$, then $f \circ g$ is differentiable at c and
>
> $$\frac{d}{dx}\Big|_{x=c} (f\circ g)(x) = f'(g(c))g'(c)\,.$$

For functions of two variables, we have the following

Theorem 13.37

Let $z = f(x, y)$ be a differentiable function (of x and y). If $x = g(t)$ and $y = h(t)$ are differentiable functions (of t), then $z(t) = f\big(x(t), y(t)\big)$ is differentiable and

$$z'(t) = f_x\big(x(t), y(t)\big)x'(t) + f_y\big(x(t), y(t)\big)y'(t)\,.$$

Let $\gamma(t) = \big(x(t), y(t)\big)$. Then $\gamma'(t) = \big(x'(t), y'(t)\big)$, and the chain rule above can be written as

$$\frac{d}{dt}(f \circ \gamma)(t) = (Df)(\gamma(t)) \cdot \gamma'(t)\,.$$

A short-hand notation of the identity above

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = (f_x, f_y) \cdot (x', y')\,.$$

Corollary 13.38

Let $z = f(x, y)$ be a differentiable function (of x and y).

1. If $x = u(s, t)$ and $y = v(s, t)$ are such that $\dfrac{\partial u}{\partial s}$ and $\dfrac{\partial v}{\partial s}$ exist, then the first partial derivative $\dfrac{\partial z}{\partial s}$ of the function $z(s, t) = f\big(u(s, t), v(s, t)\big)$ exists and

$$z_s(s, t) = f_x\big(u(s, t), v(s, t)\big)u_s(s, t) + f_y\big(u(s, t), v(s, t)\big)v_s(s, t)\,.$$

2. If $x = u(s, t)$ and $y = v(s, t)$ are such that $\dfrac{\partial u}{\partial t}$ and $\dfrac{\partial v}{\partial t}$ exist, then the first partial derivative $\dfrac{\partial z}{\partial t}$ of the function $z(s, t) = f\big(u(s, t), v(s, t)\big)$ exists and

$$z_t(s, t) = f_x\big(u(s, t), v(s, t)\big)u_t(s, t) + f_y\big(u(s, t), v(s, t)\big)v_t(s, t)\,.$$

Example 13.39. Let $f(x, y) = x^2y - y^2$. Find $\dfrac{dz}{dt}$, where $z(t) = f(\sin t, e^t)$.

1. Since $z(t) = e^t \sin^2 t - e^{2t}$, by the product rule and the chain rule for functions of one variable, we find that

$$z'(t) = \frac{de^t}{dt}\sin^2 t + e^t\frac{d\sin^2 t}{dt} - 2e^{2t} = e^t \sin^2 t + 2e^t \sin t \cos t - 2e^{2t}\,.$$

2. By the chain rule for functions of two variables,

$$\begin{aligned} z'(t) &= \big(f_x(\sin t, e^t), f_y(\sin t, e^t)\big) \cdot \frac{d}{dt}(\sin t, e^t) \\ &= (2xy, x^2 - 2y)\Big|_{(x,y)=(\sin t, e^t)} \cdot (\cos t, e^t) \\ &= (2e^t \sin t, \sin^2 t - 2e^t) \cdot (\cos t, e^t) \\ &= 2e^t \sin t \cos t + e^t \sin^2 t - 2e^{2t}\,. \end{aligned}$$

Example 13.40. Let $f(x,y) = 2xy$. Find $\dfrac{\partial z}{\partial s}$ and $\dfrac{\partial z}{\partial t}$, where $z(s,t) = f\big(s^2+t^2, \frac{s}{t}\big)$.

1. Since $z(s,t) = 2(s^2+t^2)\dfrac{s}{t} = \dfrac{2s^3}{t} + 2st$, by the product rule we find that

$$\frac{\partial z}{\partial s}(s,t) = \frac{6s^2}{t} + 2t \qquad \text{and} \qquad \frac{\partial z}{\partial t}(s,t) = -\frac{2s^3}{t^2} + 2s\,.$$

2. By the chain rule for functions of two variables,

$$\begin{aligned} \frac{\partial z}{\partial s}(s,t) &= \big(f_x(s^2+t^2, s/t), f_y(s^2+t^2, s/t)\big) \cdot \frac{\partial}{\partial s}\big(s^2+t^2, \frac{s}{t}\big) \\ &= \big(\frac{2s}{t}, 2(s^2+t^2)\big) \cdot \big(2s, \frac{1}{t}\big) = \frac{4s^2}{t} + \frac{2s^2+2t^2}{t} = \frac{6s^2}{t} + 2t \end{aligned}$$

and

$$\begin{aligned} \frac{\partial z}{\partial t}(s,t) &= \big(f_x(s^2+t^2, s/t), f_y(s^2+t^2, s/t)\big) \cdot \frac{\partial}{\partial t}\big(s^2+t^2, \frac{s}{t}\big) \\ &= \big(\frac{2s}{t}, 2(s^2+t^2)\big) \cdot \big(2t, -\frac{s}{t^2}\big) = 4s - \frac{2s^3+2st^2}{t^2} = -\frac{2s^3}{t^2} + 2s\,. \end{aligned}$$

- **The chain rule for functions of several variables**

Suppose that $w = f(x_1, x_2, \cdots, x_n)$ be a differentiable function (of $x_1, x_2, \cdots, x_n$). If each x_i is a differentiable function of m variables $t_1, t_2, \cdots, t_m$, then

$$\begin{aligned} \frac{\partial w}{\partial t_1} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_1} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_1} = \sum_{j=1}^{n} \frac{\partial w}{\partial x_j}\frac{\partial x_j}{\partial t_1}\,, \\ \frac{\partial w}{\partial t_2} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_2} = \sum_{j=1}^{n} \frac{\partial w}{\partial x_j}\frac{\partial x_j}{\partial t_2}\,, \\ &\vdots \\ \frac{\partial w}{\partial t_m} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_m} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_m} + \cdots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_m} = \sum_{j=1}^{n} \frac{\partial w}{\partial x_j}\frac{\partial x_j}{\partial t_m}\,. \end{aligned}$$

Using the notation of the matrix multiplication,

$$\begin{bmatrix} \dfrac{\partial w}{\partial t_1} & \dfrac{\partial w}{\partial t_2} & \cdots & \dfrac{\partial w}{\partial t_m} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} & \cdots & \dfrac{\partial f}{\partial x_n} \end{bmatrix} \begin{bmatrix} \dfrac{\partial x_1}{\partial t_1} & \dfrac{\partial x_1}{\partial t_2} & \cdots & \dfrac{\partial x_1}{\partial t_m} \\ \dfrac{\partial x_2}{\partial t_1} & \dfrac{\partial x_2}{\partial t_2} & \cdots & \dfrac{\partial x_2}{\partial t_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial x_n}{\partial t_1} & \dfrac{\partial x_n}{\partial t_2} & \cdots & \dfrac{\partial x_n}{\partial t_m} \end{bmatrix}.$$

- **Differentiation of determinant functions**

For an $n \times n$ matrix A, let $\text{Cof}(A)$ denote the cofactor matrix of A; that is, the (i,j)-th entry of $\text{Cof}(A)$ is the determinant of the matrix obtained by deleting the i-th row and j-th column of A or

$$\big[\text{Cof}(A)\big]_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1\text{n}} \\ \vdots & \ddots & & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & a_{(i-1)2} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)\text{n}} \\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)\text{n}} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{\text{n}1} & a_{\text{n}2} & \cdots & a_{\text{n}(j-1)} & a_{\text{n}(j+1)} & \cdots & a_{\text{nn}} \end{vmatrix}.$$

Then the determinant of A, using the reductive algorithm, can be computed by

$$\det(A) = \sum_{k=1}^{n} a_{ik}\big[\text{Cof}(A)\big]_{ik} \qquad \forall\, 1 \leqslant i \leqslant n\,. \tag{13.5.1}$$

On the other hand, the determinant of an $n \times n$ matrix $A = [a_{ij}]_{1 \leqslant i,j \leqslant n}$ can be viewed as a real-valued function of n^2 variable:

$$f(a_{11}, a_{12}, \cdots, a_{1n}, a_{21}, a_{22}, \cdots, a_{2n}, a_{31} \cdots, a_{nn}) = \det([\text{a}_{\text{ij}}])\,.$$

Since for each $1 \leqslant i \leqslant n$ the (i,k)-th entry of the cofactor matrix $\text{Cof}(A)_{ik}$ is independent of a_{ij} for all $1 \leqslant j,k \leqslant n$, we have $\dfrac{\partial f}{\partial a_{ij}} = \big[\text{Cof}(A)\big]_{ij}$; thus if $a_{ij} = a_{ij}(t)$ is a function of t for all $1 \leqslant i,j \leqslant n$, with $A = A(t) = \big[a_{ij}(t)\big]_{1 \leqslant i,j \leqslant n}$ in mind the chain rule implies that

$$\frac{d}{dt} f\big(a_{11}(t), a_{12}(t), \cdots, a_{nn}(t)\big) = \sum_{i,j=1}^{n} \big[\text{Cof}(A)\big]_{ij} \frac{da_{ij}(t)}{dt}\,. \tag{13.5.2}$$

Let $\mathrm{Adj}(A)$ be the transpose of the cofactor matrix, called the adjoint matrix, of A, then (13.5.2) implies that

$$\frac{d}{dt}\det(A) = \sum_{i,j=1}^{n} \big[\mathrm{Adj}(A)\big]_{ji} \frac{da_{ij}}{dt} = \mathrm{tr}\Big(\mathrm{Adj}(A)\frac{dA}{dt}\Big), \tag{13.5.3}$$

where $\mathrm{tr}(M)$ denotes the trace of a square matrix M and $\dfrac{dA}{dt} = \Big[\dfrac{da_{ij}}{dt}\Big]_{1\leqslant i,j\leqslant n}$. In particular, if A is invertible, then $A^{-1} = \dfrac{1}{\det(A)}\mathrm{Adj}(A)$; thus for invertible matrix $A = \big[a_{ij}(t)\big]$, we have

$$\frac{d}{dt}\det(A) = \mathrm{tr}\Big(\det(A)A^{-1}\frac{dA}{dt}\Big) = \det(A)\mathrm{tr}\Big(A^{-1}\frac{dA}{dt}\Big) \tag{13.5.4}$$

or

$$\frac{d}{dt}\ln\big|\det(A)\big| = \mathrm{tr}\Big(A^{-1}\frac{dA}{dt}\Big).$$

Example 13.41. Let $A(t) = \begin{bmatrix} f(t) & g(t) \\ h(t) & k(t) \end{bmatrix}$. Then

$$\begin{aligned}\frac{d}{dt}\det(A) &= \mathrm{tr}\Big(\begin{bmatrix} k & -g \\ -h & f\end{bmatrix}\begin{bmatrix} f' & g' \\ h' & k'\end{bmatrix}\Big) = \mathrm{tr}\Big(\begin{bmatrix} kf'-gh' & kg'-gk' \\ -hf'+fh' & -hg'+fk'\end{bmatrix}\Big) \\ &= kf'-gh'-hg'+fk' = (fk-gh)'.\end{aligned}$$

• Taylor's theorem for functions of two variables

Let $R \subseteq \mathbb{R}^2$ be an open region, and $f : R \to \mathbb{R}$ be a function of two variables. For $(x, y), (a, b) \in R$, define $g(t) = f\big(a+t(x-a), b+t(y-b)\big)$. Suppose that all the k-th partial derivatives of f are continuous for $0 \leqslant k \leqslant n+1$ (which, by Theorem 13.35, implies that g is $(n+1)$-times differentiable), then Taylor's Theorem implies that there exists $\xi \in (0,1)$ such that

$$g(1) = \sum_{k=0}^{n} \frac{g^{(k)}(0)}{k!} + \frac{g^{(n+1)}(\xi)}{(n+1)!}.$$

Now we compute $g^{(k)}(0)$. First by the chain rule,

$$\begin{aligned}g'(t) &= \frac{d}{dt} f\big(a+t(x-a), b+t(y-b)\big) \\ &= f_x\big(a+t(x-a), b+t(y-b)\big)(x-a) + f_y\big(tx+(1-t)a, ty+(1-t)b\big)(y-b);\end{aligned}$$

thus $g'(0) = f_x(a, b)(x-a) + f_y(a, b)(y-b)$. In general, we can prove by induction that

$$g^{(k)}(t) = \sum_{j=0}^{k} C^k_j \frac{\partial^k f}{\partial x^{k-j}\partial y^j}\big(a+t(x-a), b+t(y-b)\big)(x-a)^{k-j}(y-b)^j \tag{13.5.5}$$

under the assumption that the k-th partial derivatives are continuous (on an open region containing the line segment connecting (x,y) and (a,b)). To see this, we first simplify the notation by letting $\gamma(t) = \big(a+t(x-a), b+t(y-b)\big)$. We note that (13.5.5) holds for $k=1$. Suppose that (13.5.5) holds for $k=\ell$. Then by the chain rule and Theorem 13.28, we find that

$$\begin{aligned}
g^{(\ell+1)}(t) &= \frac{d}{dt} g^{(\ell)}(t) = \frac{d}{dt} \sum_{j=0}^{\ell} C^\ell_j \frac{\partial^\ell f}{\partial x^{\ell-j}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell-j}(y-b)^j \\
&= \sum_{j=0}^{\ell} C^\ell_j \Big[\frac{\partial^{\ell+1} f}{\partial x^{\ell-j+1}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell-j+1}(y-b)^j \\
&\qquad\qquad + \frac{\partial^{\ell+1} f}{\partial x^{\ell-j}\partial y^{j+1}}\big(\gamma(t)\big)(x-a)^{\ell-j}(y-b)^{j+1}\Big] \\
&= \sum_{j=0}^{\ell} C^\ell_j \frac{\partial^{\ell+1} f}{\partial x^{\ell+1-j}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell+1-j}(y-b)^j \\
&\quad + \sum_{j=1}^{\ell+1} C^\ell_{j-1} \frac{\partial^{\ell+1} f}{\partial x^{\ell+1-j}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell+1-j}(y-b)^j \\
&= \frac{\partial^{\ell+1} f}{\partial x^{\ell+1}}\big(\gamma(t)\big)(x-a)^{\ell+1} + \frac{\partial^{\ell+1} f}{\partial y^{\ell+1}}\big(\gamma(t)\big)(y-b)^{\ell+1} \\
&\quad + \sum_{j=1}^{\ell} (C^\ell_j + C^\ell_{j-1}) \frac{\partial^{\ell+1} f}{\partial x^{\ell+1-j}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell+1-j}(y-b)^j .
\end{aligned}$$

By Pascal's Theorem,

$$\begin{aligned}
g^{(\ell+1)}(t) &= \frac{\partial^{\ell+1} f}{\partial x^{\ell+1}}\big(\gamma(t)\big)(x-a)^{\ell+1} + \frac{\partial^{\ell+1} f}{\partial y^{\ell+1}}\big(\gamma(t)\big)(y-b)^{\ell+1} \\
&\quad + \sum_{j=1}^{\ell} C^{\ell+1}_j \frac{\partial^{\ell+1} f}{\partial x^{\ell+1-j}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell+1-j}(y-b)^j \\
&= \sum_{j=0}^{\ell+1} C^{\ell+1}_j \frac{\partial^{\ell+1} f}{\partial x^{\ell+1-j}\partial y^j}\big(\gamma(t)\big)(x-a)^{\ell+1-j}(y-b)^j ;
\end{aligned}$$

thus we establish (13.5.5) by induction. Therefore, by the fact that $g(1) = f(x,y)$ and $g(0) = f(a,b)$,

$$f(x,y) = \sum_{k=0}^{n} \frac{1}{k!} \sum_{j=0}^{k} C^k_j \frac{\partial^k f}{\partial x^{k-j}\partial y^j}(a,b)(x-a)^{k-j}(y-b)^j + R_n(x,y) , \tag{13.5.6}$$

where

$$R_n(x,y) = \frac{1}{(n+1)!} \sum_{j=0}^{n+1} C_j^{n+1} \frac{\partial^k f}{\partial x^{k-j} \partial y^j} \big(a + \xi(x-a), b + \xi(y-b)\big)(x-a)^{n+1-j}(y-b)^j \,.$$

The "polynomial" of two variables

$$P_n(x,y) = \sum_{k=0}^{n} \frac{1}{k!} \sum_{j=0}^{k} C_j^k \frac{\partial^k f}{\partial x^{k-j} \partial y^j}(a,b)(x-a)^{k-j}(y-b)^j$$

is called the n-th Taylor polynomial for f centered at (a,b), and the function R_n is the remainder associated with P_n.

Expanding the sum, we find that

$$\begin{aligned}
P_n(x,y) = {} & f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) \\
& + \frac{1}{2!}\Big[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\Big] \\
& + \frac{1}{3!}\Big[f_{xxx}(a,b)(x-a)^3 + 3f_{xxy}(a,b)(x-a)^2(y-b) + 3f_{xyy}(x-a)(y-b)^2 \\
& \qquad + f_{yyy}(a,b)(y-b)^3\Big] + \cdots + \\
& + \frac{1}{n!}\Big[\frac{\partial^n f}{\partial x^n}(a,b)(x-a)^n + C_1^n \frac{\partial^n f}{\partial x^{n-1}\partial y}(a,b)(x-a)^{n-1}(y-b) + \cdots\cdots + \\
& \qquad + C_{n-1}^n \frac{\partial^n f}{\partial x \partial y^{n-1}}(a,b)(x-a)(y-b)^{n-1} + \frac{\partial^n f}{\partial y^n}(a,b)(y-b)^n\Big] \,.
\end{aligned}$$

Example 13.42. Find the third Taylor polynomial for the function $f(x,y) = \sin(xy)$ centered at $(0,0)$.

We compute the first, the second and the third partial derivatives of f as follows:

$$\begin{aligned}
& f_x(x,y) = y\cos(xy)\,, \quad f_y(x,y) = x\cos(xy)\,, \\
& f_{xx}(x,y) = -y^2\sin(xy)\,, \quad f_{xy}(x,y) = \cos(xy) - xy\sin(xy)\,, \quad f_{yy}(x,y) = -x^2\sin(xy)\,, \\
& f_{xxx}(x,y) = -y^3\cos(xy)\,, \quad f_{xxy}(x,y) = -2y\sin(xy) - xy^2\cos(xy)\,, \\
& f_{xyy}(x,y) = -2x\sin(xy) - x^2y\cos(xy)\,, \quad f_{yyy}(x,y) = -x^3\cos(xy)\,.
\end{aligned}$$

Therefore, the only non-vanishing term, when plugging $(x,y) = (0,0)$, is $f_{xy}(0,0) = 1$; thus

$$P_3(x,y) = \frac{1}{2!} \cdot 2f_{xy}(0,0)(x-0)(y-0) = xy\,.$$

Example 13.43. Find the second Taylor polynomial for the function $f(x, y) = \exp(x^2+2y)$ centered at $(0, 0)$.

We compute the first and the second partial derivatives of f as follows:

$$\begin{aligned} f_x(x, y) &= 2x \exp(x^2 + 2y)\,, \quad f_y(x, y) = 2\exp(x^2 + 2y)\,, \\ f_{xx}(x, y) &= (2 + 4x^2)\exp(x^2 + 2y)\,, \quad f_{xy}(x, y) = 4x\exp(x^2 + 2y)\,, \\ f_{yy}(x, y) &= 4\exp(x^2 + 2y)\,. \end{aligned}$$

Therefore, $f_x(0, 0) = f_{xy}(0, 0) = 0$, $f_y(0, 0) = f_{xx}(0, 0) = 2$, $f_{yy}(0, 0) = 4$; thus

$$\begin{aligned} P_2(x, y) &= f(0, 0) + f_x(0, 0)x + f_y(0, 0)y + \frac{1}{2!}\big[f_{xx}(0, 0)x^2 + 2f_{xy}(0, 0)xy + f_{yy}(0, 0)y^2\big] \\ &= 1 + 2y + x^2 + 2y^2\,. \end{aligned}$$

- **Implicit partial differentiation**

In Section 2.4 we have talked about finding derivatives of a function $y = f(x)$ which is defined implicitly by $F(x, y) = 0$ (when F is giving explicitly). Now suppose that $z = F(x, y)$ is a differentiable function and the relation $F(x, y) = 0$ defines a differentiable function $y = f(x)$ implicitly (so that $F(x, f(x)) = 0$). By the chain rule,

$$0 = \frac{d}{dx}F(x, f(x)) = F_x(x, f(x)) + F_y(x, f(x))f'(x)$$

which implies that

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))} \quad \text{if} \quad F_y(x, f(x)) \neq 0\,.$$

Since f is in general unknown (but exists), we usually write the identity above as

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)} \quad \text{if } F(x, y) = 0 \text{ and } F_y(x, y) \neq 0\,.$$

In fact, when F_x and F_y are continuous in an open region R, and $F(a, b) = 0$ and $F_y(a, b) \neq 0$ at some point $(a, b) \in R$, the relation $F(x, y) = 0$ defines a function $y = f(x)$ implicitly near (a, b) and f is continuously differentiable near $x = a$. This is the Implicit Function Theorem and the precise statement is stated as follows.

Theorem 13.44: Implicit Function Theorem (Special case)

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $F : R \to \mathbb{R}$ be a function of two variables such that F_x and F_y are continuous in a neighborhood of $(a, b) \in R$. If $F(a, b) = 0$ and $F_y(a, b) = 0$, then there exists $\delta > 0$ and a unique function $f : (a - \delta, a + \delta) \to \mathbb{R}$ satisfying $F(x, f(x)) = 0$ for all $x \in (a - \delta, a + \delta)$, and $b = f(a)$. Moreover, f is differentiable on $(a - \delta, a + \delta)$, and

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))} \qquad \forall\, x \in (a - \delta, a + \delta)\,.$$

In general, if F is a function of n variables $(x_1, x_2, \cdots, x_n)$ such that F_{x_1}, F_{x_2}, $\cdots$, F_{x_n} are continuous in a neighborhood of $(a_1, a_2, \cdots, a_n.$ If $F(a_1, a_2, \cdots, a_n) = 0$ and $F_{x_n}(a_1, a_2, \cdots, a_n) \neq 0$, then locally there exists a unique function f satisfying $F(x_1, \cdots, x_{n-1}, f(x_1, \cdots, x_{n-1})) = 0$ and $a_n = f(a_1, \cdots, a_{n-1})$. Moreover, for $1 \leqslant j \leqslant n - 1$,

$$\frac{\partial f}{\partial x_j}(x_1, \cdots, x_{n-1}) = -\frac{F_{x_j}(x_1, \cdots, x_{n-1}, f(x_1, \cdots, x_{n-1}))}{F_{x_n}(x_1, \cdots, x_{n-1}, f(x_1, \cdots, x_{n-1}))}\,.$$

Example 13.45. Find $\dfrac{dy}{dx}$ if (x, y) satisfies $y^3 + y^2 - 5y - x^2 + 4 = 0$.

Let $F(x, y) = y^3 + y^2 - 5y - x^2 + 4$. Then $F_x(x, y) = -2x$ and $F_y(x, y) = 3y^2 + 2y - 5$. Therefore,

$$\frac{dy}{dx} = -\frac{F_x(x, y)}{F_y(x, y)} = \frac{2x}{3y^2 + 2y - 5}\,.$$

Example 13.46. Find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ if (x, y, z) satisfies $3x^2 z - x^2 y^2 + 2z^3 + 3yz - 5 = 0$.

Let $F(x, y, z) = 3x^2 z - x^2 y^2 + 2z^3 + 3yz - 5$. Then $F_x(x, y, z) = 6xz - 2xy^2$, $F_y(x, y, z) = -2x^2 y + 3z$ and $F_z(x, y, z) = 3x^2 + 6z^2 + 3y$. Therefore,

$$\frac{\partial z}{\partial x} = -\frac{F_x(x, y, z)}{F_z(x, y, z)} = \frac{2xy^2 - 6xz}{3x^2 + 6z^2 + 3y}$$

and

$$\frac{\partial z}{\partial y} = -\frac{F_y(x, y, z)}{F_z(x, y, z)} = \frac{2x^2 y - 3z}{3x^2 + 6z^2 + 3y}\,.$$

13.6 Directional Derivatives and Gradients

Let f be a function of two variables. From the discussion above we know that the existence of f_x and f_y does not guarantee the differentiability of f. Since f_x and f_y are the rate of change of the function f in two special directions $(1,0)$ and $(0,1)$, we can ask ourselves whether f is differentiable if the rate of change of f exist in all direction.

Definition 13.47

Let f be a function of two variables x and y, and let $\boldsymbol{u} = \cos\theta\mathbf{i}+\sin\theta\mathbf{j}$, where $\mathbf{i} = (1,0)$ and $\mathbf{j} = (0,1)$, be a unit vector. The directional derivative of f in the direction of $\boldsymbol{u}$ at (a,b), denoted by $D_{\boldsymbol{u}}f(a,b)$, is the limit

$$D_{\boldsymbol{u}}f(a,b) = \lim_{h\to 0}\frac{f(a+h\cos\theta, b+h\sin\theta)-f(a,b)}{h}$$

provided this limit exists.

Example 13.48. Find the direction derivative of $f(x,y) = x^2\sin 2y$ at $\left(1, \frac{\pi}{2}\right)$ in the direction of $\boldsymbol{v} = 3\mathbf{i} - 4\mathbf{j}$.

We first normalize the vector $\boldsymbol{v}$ and find that $\boldsymbol{u} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}$ is in the same direction of $\boldsymbol{v}$ and has unit length. Therefore, for $h \neq 0$,

$$\frac{f\left(1+\frac{3h}{5}, \frac{\pi}{2}-\frac{4h}{5}\right)-f\left(1,\frac{\pi}{2}\right)}{h} = \frac{\left(1+\frac{3h}{5}\right)^2\sin\left(\pi-\frac{8h}{5}\right)-1^2\sin\pi}{h} = \left(1+\frac{3h}{5}\right)^2\frac{\sin\frac{8h}{5}}{h};$$

thus by the fact that $\lim\limits_{h\to 0}\frac{\sin h}{h} = 1$, we find that

$$\lim_{h\to 0}\frac{f\left(1+\frac{3h}{5}, \frac{\pi}{2}-\frac{4h}{5}\right)-f\left(1,\frac{\pi}{2}\right)}{h} = \lim_{h\to 0}\left(1+\frac{3h}{5}\right)^2\frac{\sin\frac{8h}{5}}{h} = \frac{8}{5}.$$

Theorem 13.49

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. If f is differentiable at $(x_0, y_0) \in R$, then for all unit vector $\boldsymbol{v} = \cos\theta\mathbf{i}+\sin\theta\mathbf{j}$,

$$(D_{\boldsymbol{u}}f)(x_0,y_0) = f_x(x_0,y_0)\cos\theta + f_y(x_0,y_0)\sin\theta = (Df)(x_0,y_0)\cdot\boldsymbol{u}.$$

Proof. Let $g(t) = f(x_0 + t\cos\theta, y_0 + t\sin\theta)$. Then by the chain rule for functions of two variables,

$$(D_{\boldsymbol{u}}f)(x_0, y_0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h} = g'(0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta\,. \qquad \square$$

Example 13.50. In this example we re-compute of the direction derivative in Example 13.48 using Theorem 13.49. Note that $f(x, y) = x^2 \sin 2y$ is differentiable on $\mathbb{R}^2$ since $f_x(x, y) = 2x \sin 2y$ and $f_y(x, y) = 2x^2 \cos 2y$ are continuous (so that Theorem 13.35) guarantees the differentiability of f). Therefore, Theorem 13.49 implies that

$$(D_{\boldsymbol{u}}f)\big(1, \frac{\pi}{2}\big) = \frac{3}{5} f_x\big(1, \frac{\pi}{2}\big) - \frac{4}{5} f_y\big(1, \frac{\pi}{2}\big) = \frac{3}{5}\cdot 2 \cdot \sin\pi - \frac{4}{5}\cdot 2\cdot 1^2 \cdot \cos\pi = \frac{8}{5}\,.$$

Unfortunately, the existence of directional derivative of f in all directions does not imply the differentiability of f.

Example 13.51. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$f(x, y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4} & \text{if } (x, y) \neq (0, 0)\,, \\ 0 & \text{if } (x, y) = (0, 0)\,, \end{cases}$$

and $\boldsymbol{u} = (\cos\theta, \sin\theta) \in \mathbb{R}^2$ be a unit vector. Then if $\cos\theta \neq 0$ $\big($or equivalently, $\theta \neq \dfrac{\pi}{2}, \dfrac{3\pi}{2}\big)$,

$$(D_{\boldsymbol{u}}f)(0, 0) = \lim_{h\to 0} \frac{f(h\cos\theta, h\sin\theta) - f(0, 0)}{h} = \lim_{t\to 0} \frac{h^3\cos\theta\sin\theta^2}{h(h^2\cos\theta^2 + h^4\sin\theta^4)} = \frac{\sin\theta^2}{\cos\theta}$$

while if $\cos\theta = 0$,

$$(D_{\boldsymbol{u}}f)(0, 0) = \lim_{h\to 0} \frac{f(h\cos\theta, h\sin\theta) - f(0, 0)}{h} = 0\,.$$

Therefore, the directional derivative of f at $(0, 0)$ exist in all directions. However, f is not continuous at $(0, 0)$ since if (x, y) approaches $(0, 0)$ along the curve $x = my^2$ with $m \neq 0$, we have

$$\lim_{\substack{(x,y)\to(0,0)\\ x=my^2}} f(x, y) = \lim_{y\to 0} f(my^2, y) = \lim_{y\to 0} \frac{my^4}{m^2y^4 + y^4} = \frac{m}{m^2 + 1}$$

which depends on m. Therefore, f is not continuous at $(0, 0)$.

Definition 13.52

Let $z = f(x, y)$ be a function of x and y such that $f_x(a, b)$ and $f_y(a, b)$ exists. Then the gradient of f at (a, b), denoted by $(\nabla f)(a, b)$ or $(\mathbf{grad}f)(a, b)$, is the vector $(f_x(a, b), f_y(a, b))$; that is,

$$(\nabla f)(a, b) = \big(f_x(a, b), f_y(a, b)\big) = f_x(a, b)\mathbf{i} + f_y(a, b)\mathbf{j}\,.$$

- **Functions of several variables**

Definition 13.53

Let f be a function of n variables. The directional derivative of f at $(a_1, a_2, \cdots, a_n)$ in the direction $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)$, where $u_1^2 + u_2^2 + \cdots + u_n^2 = 1$, is the limit

$$(D_{\boldsymbol{u}} f)(a_1, a_2, \cdots, a_n) = \lim_{h \to 0} \frac{f(a_1 + hu_1, a_2 + hu_2, \cdots, a_n + hu_n) - f(a_1, a_2, \cdots, a_n)}{h}$$

provided that the limit exists. The gradient of f at $(a_1, a_2, \cdots, a_n)$, denoted by $(\nabla f)(a_1, a_2, \cdots, a_n)$, is the vector

$$(\nabla f)(a_1, a_2, \cdots, a_n) = \big(f_{x_1}(a_1, \cdots, a_n), f_{x_2}(a_1, \cdots, a_n), \cdots, f_{x_n}(a_1, \cdots, a_n)\big).$$

Theorem 13.54

Let f be a function of n variables. If f is differentiable at $(a_1, a_2, \cdots, a_n)$ and $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)$ is a unit vector, then

$$(D_{\boldsymbol{u}} f)(a_1, a_2, \cdots, a_n) = (\nabla f)(a_1, \cdots, a_n) \cdot \boldsymbol{u}.$$

- **Properties of the gradient**

Theorem 13.55

Let f be a function of two variables. If f has continuous first partial derivatives f_x and f_y in a neighborhood of (x_0, y_0) and $(\nabla f)(x_0, y_0) \neq \mathbf{0}$, then $(\nabla f)(x_0, y_0)$ is perpendicular/normal to the level curve $f(x, y) = f(x_0, y_0)$ at (x_0, y_0). Moreover, the value of f at (x_0, y_0) increase most rapidly in the direction $\dfrac{(\nabla f)(x_0, y_0)}{\|(\nabla f)(x_0, y_0)\|}$ and decreases most rapidly in the direction $-\dfrac{(\nabla f)(x_0, y_0)}{\|(\nabla f)(x_0, y_0)\|}$, where $\|\cdot\|$ denotes the length of the vector.

Remark 13.56. 1. Let $f : (a, b) \to \mathbb{R}$ be differentiable. The graph of the function $y = f(x)$ can be view as the level set $F(x, y) = y - f(x)$ through point $(c, f(c))$ (that is, $F(x, y) = F(c, f(c))$). We note that at the slope of the tangent line $(c, f(c))$ if $f'(c)$ (so that $(1, f'(c))$ is a tangent vector at $(c, f(c))$); thus the vector $(-f'(c), 1)$ is perpendicular to the graph of f at $(c, f(c))$. The theorem above generalizes this result.

2. The terminology "the value of f at (x_0, y_0) increase most rapidly in the direction $\boldsymbol{u}$", where $\boldsymbol{u}$ is a unit vector, means that the directional derivative $(D_{\boldsymbol{v}}f)(x_0, y_0)$, treated as a function of $\boldsymbol{v}$, attains its maximum at $\boldsymbol{v} = \boldsymbol{u}$.

Example 13.57. Let $f(x, y) = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$. Then the level curve $f(x, y) = 1$ is an ellipse and the normal vector of this level curve at point $(a\cos\theta, b\sin\theta)$ is given by

$$\big(f_x(a\cos\theta, b\sin\theta), f_y(a\cos\theta, b\sin\theta)\big) = \big(\frac{2\cos\theta}{a}, \frac{2\sin\theta}{b}\big)\,.$$

Example 13.58. A heat-seeking particle is located at the point $(2, -3)$ on a metal plate whose temperature at (x, y) is $T(x, y) = 20 - 4x^2 - y^2$. Find the path of the particle as it continuously moves in the direction of maximum temperature increase.

Suppose the path of the particle is given by $(x(t), y(t))$. Then

$$(x'(t), y'(t)) /\!/ \ (\nabla T)(x(t), y(t)) = \big(-8x(t), -2y(t)\big)\,.$$

Therefore, there exists a function $k(t)$ such that $-8x = k\dfrac{dx}{dt}$ and $-2y = k\dfrac{dy}{dt}$; thus

$$\frac{d}{dt}\big(\ln|x| - 4\ln|y|\big) = 0\,.$$

Then $|x||y|^{-4} = C$. Since $(x(t), y(t))$ passes through $(2, -3)$, we find that $C = \dfrac{2}{81}$; thus (x, y) satisfies $x = \dfrac{2}{81}y^4$.

Theorem 13.59

Let f be a function of three variables. If f has continuous first partial derivatives f_x, f_y, f_z in a neighborhood of (x_0, y_0, z_0) and $(\nabla f)(x_0, y_0, z_0) \neq \mathbf{0}$, then $(\nabla f)(x_0, y_0, z_0)$ is perpendicular/normal to the level surface $f(x, y, z) = f(x_0, y_0, z_0)$ at (x_0, y_0, z_0). Moreover, the value of f at (x_0, y_0, z_0) increase most rapidly in the direction $\dfrac{(\nabla f)(x_0, y_0, z_0)}{\|(\nabla f)(x_0, y_0, z_0)\|}$ and decreases most rapidly in the direction $-\dfrac{(\nabla f)(x_0, y_0, z_0)}{\|(\nabla f)(x_0, y_0, z_0)\|}$, where $\|\cdot\|$ denotes the length of the vector.

Proof. We have shown that $(\nabla F)(x_0, y_0, z_0)$ is perpendicular to the level surface $F(x, y, z) = F(x_0, y_0, z_0)$ in Theorem 13.63, so it suffices to show that $(D_{\boldsymbol{v}}F)(x_0, y_0, z_0)$ attains its maximum at $\boldsymbol{v} = \boldsymbol{u}$. Nevertheless, by Theorem 13.54, we find that

$$(D_{\boldsymbol{v}}F)(x_0, y_0, z_0) = (\nabla F)(x_0, y_0, z_0) \cdot \boldsymbol{v} = \|(\nabla F)(x_0, y_0, z_0)\|\cos\theta\,,$$

where θ is the angle between $(\nabla F)(x_0, y_0, z_0)$ and $\boldsymbol{v}$. Clearly $(D_{\boldsymbol{v}}F)(x_0, y_0, z_0)$ attains its maximum when $\theta = 0$ which shows that $(D_{\boldsymbol{v}}F)(x_0, y_0, z_0)$ attains its maximum at $\boldsymbol{v} = \dfrac{(\nabla F)(x_0, y_0, z_0)}{\|(\nabla F)(x_0, y_0, z_0)\|}$. □

Example 13.60 (Gradient method of finding local minimum of a function)**.** Suppose that you are looking for the minimum of a function $f : \mathbb{R}^2 \to \mathbb{R}$. You do not know where the minimum point of f is, so you start with (conjecturing a possible) point (a, b) and hope to find a curve C that connects (a, b) and the minimum point. Suppose that C is parameterized by $\boldsymbol{r} : [a, b] \to \mathbb{R}^2$. By the fact that $-(\nabla f)(\boldsymbol{x})$ points to the direction to which f decreases most rapidly, we expected that

$$\boldsymbol{r}'(t) \mathbin{/\!/} -(\nabla f)(\boldsymbol{r}(t))\,.$$

In particular, we choose $\boldsymbol{r}'(t) = -(\nabla f)(\boldsymbol{r}(t))$ and hope that we can find $\boldsymbol{r}$ (so that we can find C). We note that we can also choose $\boldsymbol{r}'(t) = -\dfrac{(\nabla f)(\boldsymbol{r}(t))}{\|(\nabla f)(\boldsymbol{r}(t))\|}$ which implies that $\boldsymbol{r}'$ never vanishes so that the tangent direction indeed points to the direction $-(\nabla f)(\boldsymbol{r}(t))$.

Sometimes it is very hard to find the solution $\boldsymbol{r}$ to the differential equationt, so instead we choose a different strategy. Starting at the point (a, b), we move forward in the direction $-(\nabla f)(a, b)$ and stop temporally at $(a_1, b_1) \equiv (a, b) - t_0(\nabla f)(a, b)$ for some $t > 0$. Then we move forward in the direction $-(\nabla f)(a_1, b_1)$ and stop temporally at $(a_2, b_2) \equiv (a_1, b_1) - t_1(\nabla)(a_1, b_1)$. Continue this process, we obtain a sequence of stops $\{(a_k, b_k)\}_{k=1}^{\infty}$ given by

$$(a_{k+1}, b_{k+1}) = (a_k, b_k) - t_k(\nabla f)(a_k, b_k) \tag{13.6.1}$$

for some sequence $\{t_k\}_{k=0}^{\infty}$ of non-negative numbers to be chosen. One way of choosing the step-size t_k, called the method of exact line search, is to choose t_k so that

$$f\big((a_k, b_k) - t_k(\nabla f)(a_k, b_k)\big) = \min_{t>0} f\big((a_k, b_k) - t(\nabla f)(a_k, b_k)\big)\,.$$

Such t_k must satisfy that

$$\frac{d}{dt}\Big|_{t=t_k} f\big((a_k, b_k) - t(\nabla f)(a_k, b_k)\big) = 0$$

which implies that t_k satisfies that $(\nabla f)\big((a_k, b_k) - t_k(\nabla f)(a_k, b_k)\big) \cdot (\nabla f)(a_k, b_k) = 0$. Therefore, (13.6.1) implies that

$$(\nabla f)(a_{k+1}, b_{k+1}) \cdot (\nabla f)(a_k, b_k) = 0 \qquad \forall\, k \in \mathbb{N} \cup \{0\}$$

which shows that the exact line search algorithm of constructing minimizing sequence produces a zigzag path connecting the starting point and the minimum point.

13.7 Tangent Planes and Normal Lines

- **The tangent plane of surfaces**

Any three points in space that are not collinear defines a plane. Suppose that $\mathcal{S}$ is a "surface" (which we have not define yet, but please use the common sense to think about it), and $P_0 = (x_0, y_0, z_0)$ is a point on the plane. Given another two point $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$ on the surface such that P_0, P_1, P_2 are not collinear, let $T_{P_1P_2}$ denote the plane determined by P_0, P_1 and P_2. If the plane "approaches" a certain plane as P_1, P_2 approaches P_0, the "limit" is called the tangent plane of $\mathcal{S}$ at P_0.

Now suppose that the surface $\mathcal{S}$ is the graph of a function of two variables $z = f(x, y)$. Consider the tangent plane of $\mathcal{S}$ at $P_0 = (x_0, y_0, z_0)$, where $z_0 = f(x_0, y_0)$. The plane $T_{P_1P_2}$, where $P_1 = (x_0 + h, y_0, f(x_0 + h, y_0))$ and $P_2 = (x_0, y_0 + k, f(x_0, y_0 + k))$, is given by

$$\Big[\big(h, 0, f(x_0+h, y_0) - f(x_0, y_0)\big) \times \big(0, k, f(x_0, y_0+k) - f(x_0, y_0)\big)\Big] \cdot (x - x_0, y - y_0, z - z_0) = 0\,,$$

where $\boldsymbol{u} \cdot \boldsymbol{v}$ and $\boldsymbol{u} \times \boldsymbol{v}$ are the inner product and the cross product of $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively. For $(h, k) \neq (0, 0)$, divide both sides by hk and pass to the limit as $(h, k) \to (0, 0)$, we find that the limit is

$$\Big[\big(1, 0, f_x(x_0, y_0)\big) \times \big(0, 1, f_y(x_0, y_0)\big)\Big] \cdot (x - x_0, y - y_0, z - z_0) = 0\,,$$

provided that $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exists. Computing the cross product, we find that

$$\big(1, 0, f_x(x_0, y_0)\big) \times \big(0, 1, f_y(x_0, y_0)\big) = \big(-f_x(x_0, y_0), -f_y(x_0, y_0), 1\big)\,;$$

thus if the tangent plane exists at (x_0, y_0, z_0), the tangent plane must be

$$\big(-f_x(x_0, y_0), -f_y(x_0, y_0), 1\big) \cdot \big(x - x_0, y - y_0, z - f(x_0, y_0)\big) = 0$$

or equivalently,

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)\,.$$

On the other hand, if f is differentiable at (x_0, y_0), then

$$\begin{aligned} f(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ + \varepsilon_1(x, y)(x - x_0) + \varepsilon_2(x, y)(y - y_0) \end{aligned}$$

for some functions ε_1, ε_2 satisfying $\lim\limits_{(x,y)\to(x_0,y_0)} \varepsilon_1(x,y) = \lim\limits_{(x,y)\to(x_0,y_0)} \varepsilon_2(x,y) = 0$. This shows that the rate of convergence of the quantity

$$\big|f(x,y) - f(x_0,y_0) - f_x(x_0,y_0)(x-x_0) - f_y(x_0,y_0)(y-y_0)\big|\,,$$

as (x,y) approaches (x_0,y_0), is "faster than linear" and this is exactly what we have in mind when talking about tangent planes. Therefore, we conclude that

Theorem 13.61

Let $R \subseteq \mathbb{R}^2$ be an open region in the plane, and $f : R \to \mathbb{R}$ be a function of two variables. If f is differentiable at $(x_0,y_0) \in R$, the tangent plane of the graph of f at $(x_0,y_0,f(x_0,y_0))$ is given by

$$z = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)\,,$$

and the vector $(f_x(x_0,y_0), f_y(x_0,y_0), -1)$ is a normal vector to the graph of f at $(x_0,y_0,f(x_0,y_0))$.

Example 13.62. Find the equation of the normal line to the surface $xyz = 12$ at the point $(2,-2,-3)$.

Let $F(x,y,z) = xyz - 12$. Then $(F_x,F_y,F_z)(2,-2,-3) = (6,-6,-4)$. Therefore, the vector $(6,-6,-4)$ is normal to the surface $xyz = 12$ at $(2,-2,-3)$ and the normal line passing through $(2,-2,-3)$ is

$$\frac{x-2}{6} = \frac{y+2}{-6} = \frac{z+3}{-4}\,.$$

Now suppose that the function of three variables $w = F(x,y,z)$ is continuously differentiable; that is, F_x, F_y, F_z are continuous. Suppose that for some (x_0,y_0,z_0) in the domain, $(F_x(x_0,y_0,z_0), F_y(x_0,y_0,z_0), F_z(x_0,y_0,z_0)) \neq \mathbf{0}$. W.L.O.G., we assume that $F_z(x_0,y_0,z_0) \neq 0$. Then the Implicit Function Theorem (Theorem 13.44) implies that there exists a unique differentiable function $z = f(x,y)$ such that

$$F(x,y,f(x,y)) = 0 \qquad \text{and} \qquad z_0 = f(x_0,y_0)\,.$$

By the discussion above, the tangent plane of the graph of f at (x_0,y_0,z_0) is given by

$$z = z_0 + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$$

and the implicit partial differentiation further shows that the tangent plane above can be rewritten as

$$z = z_0 - \frac{F_x(x_0, y_0, z_0)}{F_z(x_0, y_0, z_0)}(x - x_0) - \frac{F_y(x_0, y_0, z_0)}{F_z(x_0, y_0, z_0)}(y - y_0)\,.$$

Therefore, the tangent plane of the graph of f at (x_0, y_0, z_0) is given by

$$\big(F_x(x_0, y_0, z_0), F_y(x_0, y_0, z_0), F_z(x_0, y_0, z_0)\big) \cdot (x - x_0, y - y_0, z - z_0) = 0\,.$$

On the other hand, note that the graph of f is the same as the level surface $F(x, y, z) = F(x_0, y_0, z_0)$; thus we conclude that

Theorem 13.63

Let $w = F(x, y, z)$ be a function of three variables such that F_x, F_y and F_z are continuous. If $\big(F_x(x_0, y_0, z_0), F_y(x_0, y_0, z_0), F_z(x_0, y_0, z_0)\big) \neq \mathbf{0}$, then the tangent plane of the level surface $F(x, y, z) = F(x_0, y_0, z_0)$ at (x_0, y_0, z_0) is given by

$$\big(F_x(x_0, y_0, z_0), F_y(x_0, y_0, z_0), F_z(x_0, y_0, z_0)\big) \cdot (x - x_0, y - y_0, z - z_0) = 0\,,$$

and the vector $\big(F_x(x_0, y_0, z_0), F_y(x_0, y_0, z_0), F_z(x_0, y_0, z_0)\big)$ is a normal vector to the level surface $F(x, y, z) = F(x_0, y_0, z_0)$.

Example 13.64. Find an equation of the normal line and the tangent plane to the paraboloid

$$z = 1 - \frac{1}{10}(x^2 + 4y^2)$$

at the point $\big(1, 1, \frac{1}{2}\big)$.

Let $F(x, y, z) = z - 1 + \frac{1}{10}(x^2 + 4y^2)$. Then $F_z\big(1, 1, \frac{1}{2}\big) \equiv \big(\frac{1}{5}, \frac{4}{5}, 1\big) \neq \mathbf{0}$; thus Theorem 13.63 implies that the tangent plane of the given paraboloid at $\big(1, 1, \frac{1}{2}\big)$ is

$$z = \frac{1}{2} - \frac{1}{5}(x - 1) - \frac{4}{5}(y - 1) = \frac{3}{2} - \frac{1}{5}x - \frac{4}{5}y\,.$$

An equation of the normal line at $\big(1, 1, \frac{1}{2}\big)$ is given by

$$\frac{x - 1}{1/5} = \frac{y - 1}{4/5} = \frac{z - 1/2}{1}\,.$$

13.8 Extrema of Functions of Several Variables

13.8.1 Absolute extrema and relative extrema

Theorem 13.65: Extreme Value Theorem

Let f be a continuous function of two variables x and y defined on a closed bounded region R in the plane.

1. There is at least one point in R at which f takes on a minimum value.
2. There is at least one point in R at which f takes on a maximum value.

A minimum is also called an absolute minimum and a maximum is also called an absolute maximum. As in the case of functions of one variable, there are relative extrema defined as follows.

Definition 13.66: Relative Extrema

Let f be a function defined on a region R containing (x_0, y_0).

1. The function f has a relative minimum at (x_0, y_0) if $f(x, y) \geqslant f(x_0, y_0)$ for all (x, y) in an open disk containing (x_0, y_0).
2. The function f has a relative maximum at (x_0, y_0) if $f(x, y) \leqslant f(x_0, y_0)$ for all (x, y) in an open disk containing (x_0, y_0).

Similar to the critical points for functions of one variable defined in Definition 3.4,we have the following

Definition 13.67: Critical Points

Let f be defined on an open region R containing (x_0, y_0). The point (x_0, y_0) is a critical point of f if one of the following is true.

1. $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$;
2. $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ does not exist.

Similar to Theorem 3.5, we have the following necessary condition for points where f attains its relative extrema.

Theorem 13.68

Let R be an open region in the plane, and $f : R \to \mathbb{R}$ be continuous. If f has a relative extremum at (x_0, y_0) on an open region R, then (x_0, y_0) is a critical point of f.

Example 13.69. Determine the relative extrema of the function

$$f(x,y) = -x^3 + 4xy - 2y^2 + 1\,.$$

First we find the critical points of f. Since f is differentiable, the critical points are those points at which the gradient of f is the zero vector. Since $f_x(x,y) = -3x^2 + 4y$ and $f_y(x,y) = 4x - 4y$, if (a,b) is a critical point of f, then $-3a^2 + 4b = 4a - 4b = 0$. Therefore, $(0,0)$ and $\big(\frac{4}{3}, \frac{4}{3}\big)$ are the only critical points of f.

Note that $(0,0)$ is not a relative extremum of f since $f(x,0)$ does not attain its extremum at $x = 0$. Near $\big(\frac{4}{3}, \frac{4}{3}\big)$, we find that if $|h|, |k| \ll 1$,

$$\begin{aligned}
f\Big(\frac{4}{3}+h, \frac{4}{3}+k\Big) &= -\Big(h+\frac{4}{3}\Big)^3 + 4\Big(\frac{4}{3}+h\Big)\Big(\frac{4}{3}+k\Big) - 2\Big(k+\frac{4}{3}\Big)^2 + 1 \\
&= -h^3 - 4h^2 - \frac{16h}{3} - \frac{64}{27} + 4\Big(\frac{16}{9} + \frac{4}{3}h + \frac{4}{3}k + hk\Big) - 2\Big(k^2 + \frac{8}{3}k + \frac{16}{9}\Big) + 1 \\
&= -h^3 - 4h^2 + 4hk - 2k^2 + f\Big(\frac{4}{3}, \frac{4}{3}\Big) \\
&= f\Big(\frac{4}{3}, \frac{4}{3}\Big) - 2(k-h)^2 - h^2(2+h) \leqslant f\Big(\frac{4}{3}, \frac{4}{3}\Big)\,.
\end{aligned}$$

Therefore, f has a relative maximum at $\big(\frac{4}{3}, \frac{4}{3}\big)$.

13.8.2 The second partials test

A critical point of a function of two variables do not always yield relative maxima or minima.

Definition 13.70

Let f be a function of two variables. A point (x_0, y_0) is a saddle point of f if (x_0, y_0) is a critical point of f but f does not attain its extrema at (x_0, y_0).

Theorem 13.71

Suppose that a function f of two variables has continuous second partial derivatives on an open region containing a point (a, b) for which $f_x(a, b) = f_y(a, b) = 0$. Let

$$D = f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2 = \begin{vmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{vmatrix} .$$

1. If $D > 0$ and $f_{xx}(a, b) > 0$, then f has a relative minimum at (a, b).
2. If $D > 0$ and $f_{xx}(a, b) < 0$, then f has a relative maximum at (a, b).
3. If $D < 0$, then $(a, b, f(a, b))$ is a saddle point.
4. The test is inconclusive if $D = 0$.

Example 13.72. Consider the relative extrema of the function given in Example 13.69. We have computed that $(0, 0)$ and $\left(\frac{4}{3}, \frac{4}{3}\right)$ are the only critical points of f.

1. The point $(0, 0)$: we compute the second partial derivatives and obtain that

$$f_{xx}(0,0) = 0\,, \quad f_{xy}(0,0) = 4 \quad \text{and} \quad f_{yy}(0,0) = -4\,.$$

   Therefore, $D = -16 < 0$ which implies that $(0, 0)$ is a saddle point.

2. The point $\left(\frac{4}{3}, \frac{4}{3}\right)$: we compute the second partial derivatives and obtain that

$$f_{xx}\left(\frac{4}{3}, \frac{4}{3}\right) = -8\,, \quad f_{xy}\left(\frac{4}{3}, \frac{4}{3}\right) = 4 \quad \text{and} \quad f_{yy}\left(\frac{4}{3}, \frac{4}{3}\right) = -4\,.$$

   Therefore, $D = 16 > 0$. Since $f_{xx}\left(\frac{4}{3}, \frac{4}{3}\right) < 0$, f has a relative maximum at $\left(\frac{4}{3}, \frac{4}{3}\right)$.

Example 13.73. Find the absolute extrema of the function $f(x, y) = \sin(xy)$ on the closed region given by $0 \leqslant x \leqslant \pi$ and $0 \leqslant y \leqslant 1$.

From the partial derivatives

$$f_x(x,y) = y\cos(xy) \quad \text{and} \quad f_y(x,y) = x\cos(xy)\,,$$

we find that each point on the hyperbola $xy = \frac{\pi}{2}$ is a critical point of f. The value of f at each of these points is $\sin\frac{\pi}{2} = 1$ which is the maximum of the sine function. Therefore, the maximum of f is 1.

The minimum of f occurs at the boundary of the region.

1. $x = 0$ and $0 \leqslant y \leqslant 1$: then $f(x, y) = 0$.

2. $x = \pi$ and $0 \leqslant y \leqslant 1$: then $f(x, y) = \sin(\pi y)$. The critical points of the function $g(y) = \sin(\pi y)$ occurs at $y = \frac{1}{2}$ since $g'\big(\frac{1}{2}\big) = \pi \cos\big(\frac{\pi}{2}\big) = 0$. Since $g\big(\frac{1}{2}\big) = 1$ and $g(0) = g(1) = 0$, we find that the minimum of g is 0.

3. $y = 0$ and $0 \leqslant x \leqslant \pi$: then $f(x, y) = 0$.

4. $y = 1$ and $0 \leqslant x \leqslant \pi$: then $f(x, y) = \sin x$ whose minimum on $[0, \pi]$ is 0.

Therefore, the minimum of f is 0.

The concepts of relative extrema and critical points can be extended to functions of three or more variables. On the other hand, the second derivative test for functions of three or more variables are more tricky, and we will not talk about this until the course of Advance Calculus.

13.9 Applications of Extrema

Theorem 13.74

The least squares regression line for n points $\big\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\big\}$ is given by $y = ax + b$, where

$$a = \frac{n \sum\limits_{i=1}^{n} x_i y_i - \Big(\sum\limits_{i=1}^{n} x_i\Big)\Big(\sum\limits_{i=1}^{n} y_i\Big)}{n \sum\limits_{i=1}^{n} x_i^2 - \Big(\sum\limits_{i=1}^{n} x_i\Big)^2} \quad \text{and} \quad b = \frac{1}{n}\Big(\sum_{i=1}^{n} y_i - a \sum_{i=1}^{n} x_i\Big). \tag{13.9.1}$$

Proof. For $a, b \in \mathbb{R}$, define $S(a, b) = \sum\limits_{i=1}^{n} (ax_i + b - y_i)^2$. Then

$$\frac{\partial S}{\partial a}(a, b) = 2 \sum_{i=1}^{n} (ax_i + b - y_i) x_i\,,$$
$$\frac{\partial S}{\partial b}(a, b) = 2 \sum_{i=1}^{n} (ax_i + b - y_i)\,.$$

The critical points (a,b) of S satisfies

$$a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i\,, \tag{13.9.2a}$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} 1 = \sum_{i=1}^{n} y_i \tag{13.9.2b}$$

which implies that (a,b) are given by (13.9.1). Clearly such (a,b) minimizes S. $\square$

Remark 13.75. An easy way to memorize the equations (a,b) satisfies is given in this remark. We assume (even though in general it is a false assumption) that the line $y = ax+b$ passes through (x_1,y_1), (x_2,y_2), $\cdots$, (x_n,y_n). Then $y_i = ax_i + b$ for all $1 \leqslant i \leqslant n$; thus in matrix form, we have

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 & 1 \\ y_2 & 1 \\ \vdots & \vdots \\ y_n & 1 \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 & 1 \\ y_2 & 1 \\ \vdots & \vdots \\ y_n & 1 \end{bmatrix}$$

which implies (13.9.2).

13.10 Lagrange Multipliers

The concept of this section is to find the extrema of a function of several variables subject to certain constraints:

> Find extrema of the function $w = f(x_1, x_2, \cdots, x_n)$ when $(x_1, x_2, \cdots, x_n)$ satisfies $g_1(x_1, \cdots, x_n) = g_2(x_1, \cdots, x_n) = \cdots = g_m(x_1, \cdots, x_n) = 0$.

Theorem 13.76: Lagrange Multiplier Theorem

Let f and g be continuously differentiable functions of two variables. Suppose that on the level curve $g(x, y) = c$ the function f attains its extrema at (x_0, y_0). If $(\nabla g)(x_0, y_0) \neq \mathbf{0}$, then there is a real value λ such that

$$(\nabla f)(x_0, y_0) = \lambda(\nabla g)(x_0, y_0)\,.$$

Proof. First we note that (x_0, y_0) is on the level curve $g(x, y) = c$; thus $c = g(x_0, y_0)$.

Define $F(x, y) = g(x, y) - g(x_0, y_0)$. Then F has continuous first partial derivatives, and $(\nabla F)(x_0, y_0) = (\nabla g)(x_0, y_0) \neq \mathbf{0}$. Then either $F_x(x_0, y_0) \neq 0$ or $F_y(x_0, y_0) \neq 0$. Suppose that $F_y(x_0, y_0) \neq 0$. Then the Implicit Function Theorem implies that there exists $\delta > 0$ a unique differentiable function $h : (x_0 - \delta, x_0 + \delta) \to \mathbb{R}$ such that

$$F(x, h(x)) = 0 \qquad \text{and} \qquad y_0 = h(x_0)\,.$$

In other words, the set $\big\{(x, h(x)) \,\big|\, x_0 - \delta < x < x_0 + \delta\big\}$ is a subset of the level curve $g(x, y) = g(x_0, y_0)$. Therefore, the function $G : (x_0 - \delta, x_0 + \delta) \to \mathbb{R}$ defined by $G(x) = f(x, h(x))$ attains its extrema at (an interior point) x_0; thus

$$G'(x_0) = f_x(x_0, y_0) + f_y(x_0, y_0)h'(x_0) = 0\,.$$

Since the implicit differentiation shows that

$$h'(x_0) = -\frac{F_x(x_0, h(x_0))}{F_y(x_0, h(x_0))} = -\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)}\,,$$

we conclude that

$$f_x(x_0, y_0) - f_y(x_0, y_0)\frac{g_x(x_0, y_0)}{g_y(x_0, y_0)} = 0\,.$$

If $f_y(x_0, y_0) = 0$, then $f_x(x_0, y_0) = 0$ which implies that $(\nabla f)(x_0, y_0) = \mathbf{0} = 0 \cdot (\nabla g)(x_0, y_0)$. If $f_y(x_0, y_0) \neq 0$, then

$$\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)} = \frac{g_x(x_0, y_0)}{g_y(x_0, y_0)}$$

which implies that $(\nabla f)(x_0, y_0) /\!/ (\nabla g)(x_0, y_0)$; thus there exists λ such that

$$(\nabla f)(x_0, y_0) = \lambda(\nabla g)(x_0, y_0)\,.$$

Similar argument can be applied to the case $F_x(x_0, y_0) \neq 0$, and we omit the proof for this case. $\square$

Remark 13.77. The scalar λ in the theorem above is called a Lagrange multiplier.

Example 13.78. Find the extreme value of $f(x, y) = 4xy$ subject to the constraint

$$\frac{x^2}{9} + \frac{y^2}{16} = 1\,.$$

Let $g(x, y) = \frac{x^2}{9} + \frac{y^2}{16} - 1$. Suppose that on the level curve $g(x, y) = 0$ the function f attains its extrema at (x_0, y_0). Note that then $(\nabla g)(x_0, y_0) \neq \mathbf{0}$ (since $(x_0, y_0) \neq (0, 0)$); thus the Lagrange Multiplier Theorem implies that there exists $\lambda \in \mathbb{R}$ such that

$$(4y_0, 4x_0) = (\nabla f)(x_0, y_0) = \lambda(\nabla g)(x_0, y_0) = \lambda\Big(\frac{2x_0}{9}, \frac{y_0}{8}\Big)\,.$$

Therefore, (x_0, y_0) satisfies $4y_0 = \frac{2\lambda x_0}{9}$ and $4x_0 = \frac{\lambda y_0}{8}$, as well as $\frac{x_0^2}{9} + \frac{y_0^2}{16} = 1$. Therefore, $\lambda \neq 0$, and

$$4x_0 = \frac{\lambda y_0}{8} = \frac{\lambda}{8} \cdot \frac{\lambda x_0}{18} = \frac{\lambda^2 x_0}{144}\,.$$

The identity above implies that $x_0 = 0$ or $\lambda = \pm 24$.

1. If $x_0 = 0$, then $y_0 = \pm 4$ which shows that $\lambda = 0$, a contradiction.

2. If $\lambda = \pm 24$, then $x_0 = \pm\frac{3y_0}{4}$; thus

$$1 = \frac{1}{9} \cdot \frac{9y_0^2}{16} + \frac{y_0^2}{16} = \frac{y_0^2}{8}\,.$$

   Therefore, $y_0 = \pm 2\sqrt{2}$ which implies that $x_0 = \pm\frac{3\sqrt{2}}{2}$. At these (x_0, y_0), $f(x_0, y_0) = \pm 24$. Therefore, on the ellipse $\frac{x^2}{9} + \frac{y^2}{16} = 1$ the maximum of f is 24 $\Big($at $(x_0, y_0) = \big(\pm 2\sqrt{2}, \pm\frac{3\sqrt{2}}{2}\big)\Big)$ and the minimum of f is -24 $\Big($at $(x_0, y_0) = \big(\pm 2\sqrt{2}, \mp\frac{3\sqrt{2}}{2}\big)\Big)$.

Example 13.79. Find the extreme value of $f(x, y) = 4xy$, where $x > 0$ and $y > 0$, subject to the constraint $\frac{x^2}{9} + \frac{y^2}{16} = 1$. From the previous example we find that the maximum of f is 24 $\Big($at $(x_0, y_0) = \big(2\sqrt{2}, \frac{3\sqrt{2}}{2}\big)\Big)$. The minimum of f occurs at the end-points $(0, 4)$ or $(3, 0)$. In either points, the value of f is 0; thus the minimum of f is 0.

Example 13.80. Find the extreme value of $f(x,y) = 4xy$, where (x,y) satisfies $\dfrac{x^2}{9}+\dfrac{y^2}{16} \leqslant 1$.

We have find the extreme value of f, under the constraint $\dfrac{x^2}{9}+\dfrac{y^2}{16} = 1$, is ± 24. Therefore, it suffices to consider the extreme value of f in the interior $\dfrac{x^2}{9}+\dfrac{y^2}{16} < 1$.

Assume that f attains its extreme value at an interior point (x_0, y_0). Then (x_0, y_0) is a critical point of f; thus

$$f_x(x_0, y_0) = f_y(x_0, y_0) = 0$$

which implies that $(x_0, y_0) = (0, 0)$. Since $f(0,0) = 0$, $f(0,0)$ is not an extreme value of f. Therefore, the extreme value of f on the region $\dfrac{x^2}{9}+\dfrac{y^2}{16} \leqslant 1$ is ± 24.

We note that $(0,0)$ in fact is a saddle point of f since $f_{xx}(0,0)f_{yy}(0,0) - f_{xy}(0,0)^2 = -16 < 0$.

Example 13.81. Find the extreme value of $f(x,y) = x^2 + 6(y^2+y+1)^2$ subject to the constraint $x^2 + (y^3-1)^2 = 1$ (using the method of Lagrange multipliers).

Let $g(x,y) = x^2 + (y^3-1)^2$. We first compute the gradient of f and g as follows:

$$(\nabla f)(x,y) = \big(2x, 12(2y+1)(y^2+y+1)\big) \quad \text{and} \quad (\nabla g)(x,y) = \big(2x, 6y^2(y^3-1)\big)\,.$$

Assume that f, under the constraint $g = 1$, attains its extrema at (x_0, y_0). Then

1. If $(\nabla g)(x_0, y_0) \neq \mathbf{0}$, then the Lagrange multiplier theorem implies that there exists $\lambda \in \mathbb{R}$ such that

$$\big(2x_0, 12(2y_0+1)(y_0^2+y_0+1)\big) = \lambda\big(2x_0, 6y_0^2(y_0^3-1)\big)\,. \tag{13.10.1}$$

Therefore, $x_0(\lambda-1) = 0$ and $2(2y_0+1) = \lambda y_0^2(y_0-1)$.

(a) $x_0 = 0$, then $g(x_0, y_0) = 1$ implies that $y_0 = \sqrt[3]{2}$ ($y_0 = 0$ cannot be true because no λ will verify (13.10.1)); thus $f(x_0, y_0) = 6(\sqrt[3]{4}+\sqrt[3]{2}+1)^2$.

(b) $\lambda = 1$, then $4y_0 + 2 = y_0^2(y_0-1)$ or equivalently, $y_0^3 - y_0^2 - 2(2y_0+1) = 0$. Note that

$$y_0^3 - y_0^2 - 4y_0 - 2 = (y_0+1)(y_0^2 - 2y_0 - 2)\,;$$

thus $y_0 = -1$ (impossible since $g(x_0, -1) \neq 1$) or $y_0 = 1 \pm \sqrt{3}$ (both are impossible since $g(x_0, 1\pm\sqrt{3}) \neq 1$).

2. If $(\nabla g)(x_0, y_0) = \mathbf{0}$, then $(x_0, y_0) = (0, 0)$; thus $f(x_0, y_0) = 1$.

Therefore, the maximum of f, under the constraint $g = 1$, is $f(0, \sqrt[3]{2}) = 6(\sqrt[3]{4} + \sqrt[3]{2} + 1)^2$ and the minimum of f, under the constraint $g = 1$, is $f(0, 0) = 1$.

Similar argument of proving Theorem 13.76 can be used to show the following

Theorem 13.82

Let f and g be continuously differentiable functions of n variables. Suppose that on the level curve $g(x_1, \cdots, x_n) = c$ the function f attains its extrema at $(a_1, \cdots, a_n)$. If $(\nabla g)(a_1, \cdots, a_n) \neq \mathbf{0}$, then there is a real value λ such that

$$(\nabla f)(a_1, \cdots, a_n) = \lambda(\nabla g)(a_1, \cdots, a_n) .$$

Example 13.83. Find the minimum value of $f(x, y, z) = 2x^2 + y^2 + 3z^2$ subject to the constraint $2x - 3y - 4z = 49$.

Let $g(x, y, z) = 2x - 3y - 4z - 49$. Then $(\nabla g) \neq \mathbf{0}$; thus if f attains its relative extrema at (x_0, y_0, z_0), there exists $\lambda \in \mathbb{R}$ such that $(\nabla f)(x_0, y_0, z_0) = \lambda(\nabla g)(x_0, y_0, z_0)$. Therefore,

$$(4x_0, 2y_0, 6z_0) = \lambda(2, -3, -4)$$

or equivalently, $\lambda = 2x_0 = -\dfrac{2}{3}y_0 = -\dfrac{3}{2}z_0$. Since $2x_0 - 3y_0 - 4z_0 = 49$, we find that $\lambda = 6$ which implies that

$$(x_0, y_0, z_0) = (3, -9, -4) .$$

Since f grows beyond any bound as $\sqrt{x^2 + y^2 + z^2}$ approaches ∞, we find that $f(3, -9, -4) = 147$ is the minimum of f.

Next, we consider the optimization problem of finding the extreme value of a function of three variables $w = f(x, y, z)$ subject to two constraints $g(x, y, z) = h(x, y, z) = 0$.

Theorem 13.84: Lagrange Multiplier Theorem - More General Version

Let f, g and h be continuously differentiable functions of three variables. Suppose that subject to the constraints $g(x, y, z) = h(x, y, z) = c$ the function f attains its extrema at (x_0, y_0, z_0). If $(\nabla g)(x_0, y_0, z_0) \times (\nabla h)(x_0, y_0, z_0) \neq \mathbf{0}$, then there are real numbers λ and μ such that

$$(\nabla f)(x_0, y_0, z_0) = \lambda(\nabla g)(x_0, y_0, z_0) + \mu(\nabla h)(x_0, y_0, z_0) .$$

Example 13.85. Find the extreme value of the function $f(x,y,z) = 20 + 2x + 2y + z^2$ subject to two constraints $x^2 + y^2 + z^2 = 11$ and $x + y + z = 3$.

Let $g(x,y,z) = x^2 + y^2 + z^2 - 11$ and $h(x,y,z) = x + y + z - 3$. We first note that if (x,y,z) satisfies $g(x,y,z) = h(x,y,z) = 0$, then $(\nabla g)(x,y,z) \times (\nabla h)(x,y,z) \neq \mathbf{0}$. Moreover, f attains its extrema on the intersection of the level surface $g(x,y,z) = 0$ and $h(x,y,z) = 0$ (since the intersection is closed and bounded). Suppose that f attains its extrema at (x_0,y_0,z_0). Then there exists $\lambda, \mu \in \mathbb{R}$ such that

$$\begin{aligned}(\nabla f)(x_0,y_0,z_0) &= \lambda(\nabla g)(x_0,y_0,z_0) + \mu(\nabla h)(x_0,y_0,z_0)\,,\\ g(x_0,y_0,z_0) &= h(x_0,y_0,z_0) = 0\,.\end{aligned}$$

Therefore,

$$2\lambda x_0 + \mu = 2\,, \tag{13.10.2a}$$

$$2\lambda y_0 + \mu = 2\,, \tag{13.10.2b}$$

$$2(\lambda - 1)z_0 + \mu = 0\,, \tag{13.10.2c}$$

$$x_0^2 + y_0^2 + z_0^2 = 11\,, \tag{13.10.2d}$$

$$x_0 + y_0 + z_0 = 3\,. \tag{13.10.2e}$$

(13.10.2a,b) implies that $\lambda(x_0 - y_0) = 0$; thus $\lambda = 0$ or $x_0 = y_0$.

1. If $\lambda = 0$, then (13.10.2a) implies $\mu = 2$ and (13.10.2c) implies $\mu = 2z_0$. Therefore, $z_0 = 1$ which further shows $x_0^2 + y_0^2 = 10$ and $x_0 + y_0 = 2$. Then $(x_0,y_0) = (3,-1)$ or $(-1,3)$. Therefore, when $\lambda = 0$,

$$(x_0,y_0,z_0) = (3,-1,1) \quad \text{or} \quad (x_0,y_0,z_0) = (-1,3,1)\,.$$

2. If $x_0 = y_0$, then (13.10.2d,e) implies that $2x_0^2 + z_0^2 = 11$ and $2x_0 + z_0 = 3$. Therefore, $x_0 = y_0 = \dfrac{3 \pm 2\sqrt{3}}{3}$, $z_0 = \dfrac{3 \mp 4\sqrt{3}}{3}$.

Since $f(3,-1,1) = f(-1,3,1) = 25$ and

$$f\Big(\frac{3+2\sqrt{3}}{3}, \frac{3+2\sqrt{3}}{3}, \frac{3-4\sqrt{3}}{3}\Big) = f\Big(\frac{3-2\sqrt{3}}{3}, \frac{3-2\sqrt{3}}{3}, \frac{3+4\sqrt{3}}{3}\Big) = \frac{91}{3}\,,$$

we conclude that the maximum and minimum value of f subject to $g = h = 0$ are $\dfrac{91}{3}$ and 25, respectively.

Example 13.86. Find the extreme value of $f(x,y,z)=z$ subject to the constraints $x^4+y^4-z^3=0$ and $y=z$.

Let $g(x,y,z)=x^4+y^4-z^3$ and $h(x,y,z)=y-z$. Then

$$(\nabla g)(x,y,z)=(4x^3,4y^3,-3z^2) \qquad \text{and} \qquad (\nabla h)(x,y,z)=(0,1,-1)$$

which implies that

$$(\nabla g)(x,y,z)\times(\nabla h)(x,y,z)=(3z^2-4y^3,4x^3,4x^3)\,.$$

Suppose the extreme value of f, under the constraints $g=h=0$, occurs at (x_0,y_0,z_0).

1. If $(\nabla g)(x_0,y_0,z_0)\times(\nabla h)(x_0,y_0,z_0)=\mathbf{0}$, then $(x_0,y_0,z_0)=(0,0,0)$ and $f(0,0,0)=0$.

2. If $(\nabla g)(x_0,y_0,z_0)\times(\nabla h)(x_0,y_0,z_0)\neq\mathbf{0}$, then the Lagrange Multiplier Theorem implies that there exist $\lambda,\mu\in\mathbb{R}$ such that

$$(\nabla f)(x_0,y_0,z_0)=\lambda(\nabla g)(x_0,y_0,z_0)+\mu(\nabla h)(x_0,y_0,z_0)\,.$$

Therefore, (x_0,y_0,z_0) satisfies that

$$4\lambda x_0^3=0\,, \tag{13.10.3a}$$
$$4\lambda y_0^3+\mu=0\,, \tag{13.10.3b}$$
$$-3\lambda z_0^2-\mu=1\,, \tag{13.10.3c}$$
$$x_0^4+y_0^4-z_0^3=0\,, \tag{13.10.3d}$$
$$y_0-z_0=0\,. \tag{13.10.3e}$$

Then (13.10.3a) implies that $\lambda=0$ or $x_0=0$.

(a) If $\lambda=0$, then (13.10.3b) shows $\mu=0$; thus using (13.10.3c), we obtain a contradiction $0=-1$. Therefore, $\lambda\neq 0$.

(b) If $x_0=0$ (and $\lambda\neq0$), then (13.10.3d) implies that $y_0^4-z_0^3=0$. Together with (13.10.3e), we find that $y_0=0$ or $y_0=1$. However, if $y_0=0$, then (13.10.3b) shows that $\mu=0$ which again implies a contradiction $0=1$ from (13.10.3c). Therefore, $y_0=z_0=1$ (and there are λ,μ satisfying (13.10.3b,c) for $y_0=z_0=1$ but the values of λ and μ are not important).

Therefore, the Lagrange Multiplier Theorem only provides one possible $(x_0, y_0, z_0) = (0, 1, 1)$ where f attains its extreme value.

Since the intersection of the level surface $g = 0$ and $h = 0$ is closed and bounded, f must attains its maximum and minimum subject to the constraints $g = h = 0$. Since $(0, 0, 0)$ and $(0, 1, 1)$ are the only possible points where f attains its extrema, the maximum and minimum of f, subject to the constraint $g = h = 0$, is $f(0, 1, 1) = 1$ and $f(0, 0, 0) = 0$, respectively.

13.11 Exercise

Problem 13.1. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function such that

$$f(x, y) + f(y, z) + f(z, x) = 0 \qquad \forall\, x, y, z \in \mathbb{R}\,.$$

Show that there exists $g : \mathbb{R} \to \mathbb{R}$ such that

$$f(x, y) = g(x) - g(y) \qquad \forall\, x, y \in \mathbb{R}\,.$$

Problem 13.2. In the following sub-problems, find the limit if it exists or explain why it does not exist.

(1) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x+y}{x^2+y}$ (2) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x}{x^2-y^2}$ (3) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^4+y^2}$

(4) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$ (5) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3-y^3}{x^2+y^2}$ (6) $\displaystyle\lim_{(x,y)\to(0,0)} (x^2+y^2)\ln(x^2+y^2)$

(7) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^4+y^4}$ (8) $\displaystyle\lim_{(x,y)\to(0,0)} y\sin\frac{1}{x}$ (9) $\displaystyle\lim_{(x,y)\to(0,0)} x\cos\frac{1}{y}$

(10) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{\sqrt{x^2+y^2+1}-1}$ (11) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{xy+yz+zx}{x^2+y^2+z^2}$

(12) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{xy+yz^2+xz^2}{x^2+y^2+z^2}$ 13) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \arctan\frac{1}{x^2+y^2+z^2}$

Problem 13.3. Discuss the continuity of the functions given below.

1. $f(x, y) = \begin{cases} \dfrac{\sin(xy)}{xy} & \text{if } xy \neq 0\,, \\ 1 & \text{if } xy = 0\,. \end{cases}$

2. $f(x,y) = \begin{cases} \dfrac{e^{-x^2-y^2}-1}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 1 & \text{if } (x,y) = (0,0)\,. \end{cases}$

3. $f(x,y) = \begin{cases} \dfrac{\sin(x^3+y^4)}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$

Problem 13.4. Let $f(x,y) = \begin{cases} 0 & \text{if } y \leqslant 0 \text{ or } y \geqslant x^4\,, \\ 1 & \text{if } 0 < y < x^4\,. \end{cases}$

1. Show that $f(x,y) \to 0$ as $(x,y) \to (0,0)$ along any path through $(0,0)$ of the form $y = mx^\alpha$ with $0 < \alpha < 4$.

2. Show that f is discontinuous on two entire curves.

Problem 13.5. Find $\dfrac{\partial}{\partial x}\Big|_{(x,y,z)=(\ln 4, \ln 9, 2)} \sum\limits_{n=0}^{\infty} \dfrac{(x+y)^n}{n!z^n}$. Do not write the answer in terms of an infinite sum.

Problem 13.6. Let $f(x,y) = (x^2+y^2)^{\frac{2}{3}}$. Find the partial derivative $\dfrac{\partial f}{\partial x}$.

Problem 13.7. Let $f(x,y,z) = xy^2z^3 + \arcsin(x\sqrt{z})$. Find f_{xzy} in the region $\big\{(x,y,z) \,\big|\, |x^2z| < 1\big\}$.

Problem 13.8. Let $\vec{\boldsymbol{a}} = (a_1, a_2, \cdots, a_n)$ be a unit vector, $\vec{\boldsymbol{x}} = (x_1, x_2, \cdots, x_n)$, and $f(x_1, x_2, \cdots, x_n) = \exp(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{x}})$. Show that

$$\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} = f\,.$$

Problem 13.9. Let $f(x,y) = x(x^2+y^2)^{-\frac{3}{2}} e^{\sin(x^2y)}$. Find $f_x(1,0)$.

Problem 13.10. Let $f(x,y) = \displaystyle\int_1^y \frac{dt}{\sqrt{1-x^3t^3}}$. Show that

$$f_x(x,y) = \int_1^y \Big(\frac{\partial}{\partial x} \frac{1}{\sqrt{1-x^3t^3}}\Big)\, dt$$

in the region $\big\{(x,y) \,\big|\, x < 1, y > 1 \text{ and } xy < 1\big\}$.

Problem 13.11. The gas law for a fixed mass m of an ideal gas at absolute temperature T, pressure P, and volume V is $PV = mRT$, where R is the gas constant. Show that

$$\frac{\partial P}{\partial V}\frac{\partial V}{\partial T}\frac{\partial T}{\partial P} = -1\,.$$

Problem 13.12. The total resistance R produced by three conductors with resistances R_1, R_2, R_3 connected in a parallel electrical circuit is given by the formula

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\,.$$

Find $\dfrac{\partial R}{\partial R_1}$ by directly taking the partial derivative of the equation above.

Problem 13.13. Find the value of $\dfrac{\partial z}{\partial x}$ at the point $(1, 1, 1)$ if the equation

$$xy + z^3 x - 2yz = 0$$

defines z as a function of the two independent variables x and y and the partial derivative exists.

Problem 13.14. Find the value of $\dfrac{\partial x}{\partial z}$ at the point $(1, -1, -3)$ if the equation

$$xz + y \ln x - x^2 + 4 = 0$$

defines x as a function of the two independent variables y and z and the partial derivative exists.

Problem 13.15. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function such that $f_x(a, b)$ and $f_y(a, b)$ exists. Suppose that $c = f(a, b)$.

1. Using the geometric meaning of partial derivatives, explain what the vectors $(1, 0, f_x(a, b))$ and $(0, 1, f_y(a, b))$ mean.

2. Suppose that you know that there is a tangent plane (which we have not talked about, but you can roughly imagine what it is) of the graph of f at (a, b, c). What should the equation of the tangent plane be?

Problem 13.16. Define

$$f(x, y) = \begin{cases} x^2 \arctan \dfrac{y}{x} - y^2 \arctan \dfrac{x}{y} & \text{if } x, y \neq 0\,, \\ 0 & \text{if } x = 0 \text{ or } y = 0\,. \end{cases}$$

Find $f_{xy}(0, 0)$ and $f_{yx}(0, 0)$.

Problem 13.17. Show that each of the following functions is not differentiable at the origin.

(1) $f(x,y) = \sqrt[3]{x}\cos y$ (2) $f(x,y) = \sqrt{|xy|}$

Problem 13.18. In the following, show that both $f_x(0,0)$ and $f_y(0,0)$ both exist but that f is not differentiable at $(0,0)$.

(1) $f(x,y) = \begin{cases} \dfrac{5x^2y}{x^3+y^3} & \text{if } x^3+y^3 \neq 0\,, \\ 0 & \text{if } x^3+y^3 = 0\,. \end{cases}$

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

(3) $f(x,y) = \begin{cases} \dfrac{3x^2y}{x^4+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$

(4) $f(x,y) = \begin{cases} \dfrac{\sin(x^3+y^4)}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$

Problem 13.19. Let $f, g : (a,b) \to \mathbb{R}$ be real-valued function, $h(x,y) = f(x)g(y)$, and $c, d \in (a,b)$. Show that if f is differentiable at c and g is differentiable at d, then h is differentiable at (c,d).

Problem 13.20. Show that the function $f(x,y) = \sqrt{x^2+y^2}\sin\sqrt{x^2+y^2}$ is differentiable at $(0,0)$.

Problem 13.21. Investigate the differentiability of the following functions at the point $(0,0)$.

(1) $f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ (2) $f(x,y) = \begin{cases} \dfrac{xy}{x+y^2} & \text{if } x+y^2 \neq 0\,, \\ 0 & \text{if } x+y^2 = 0 \end{cases}$

(3) $f(x,y) = \begin{cases} (x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$

Problem 13.22. Use the chain rule for functions of several variables to compute $\dfrac{dz}{dt}$ or $\dfrac{dw}{dt}$.

(1) $z = \sqrt{1+xy}$, $x = \tan t$, $y = \arctan t$.

(2) $w = x\exp\left(\frac{y}{z}\right)$, $x = t^2$, $y = 1 - t$, $z = 1 + 2t$.

(3) $w = \ln\sqrt{x^2 + y^2 + z^2}$, $x = \sin t$, $y = \cos t$, $z = \tan t$.

(4) $w = xy\cos z$, $x = t$, $y = t^2$, $z = \arccos t$.

(5) $w = 2ye^x - \ln z$, $x = \ln(t^2 + 1)$, $y = \arctan t$, $z = e^t$.

Problem 13.23. Use the chain rule for functions of several variables to compute $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

(1) $z = \arctan(x^2 + y^2)$, $x = s\ln t$, $y = te^s$.

(2) $z = \arctan\frac{x}{y}$, $x = s\cos t$, $y = s\sin t$.

(3) $z = e^x\cos y$, $x = st$, $y = s^2 + t^2$.

Problem 13.24. Assume that $z = f\left(ts^2, \frac{s}{t}\right)$, $\frac{\partial f}{\partial x}(x, y) = xy$, $\frac{\partial f}{\partial y}(x, y) = \frac{x^2}{2}$. Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

Problem 13.25. Find the partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ at given points.

(1) $\sin(x + y) + \sin(y + z) + \sin(x + z) = 0$, $(x, y, z) = (\pi, \pi, \pi)$.

(2) $xe^y + ye^z + 2\ln x - 2 - 3\ln 2 = 0$, $(x, y, z) = (1, \ln 2, \ln 3)$.

(3) $z = e^x\cos(y + z)$, $(x, y, z) = (0, -1, 1)$.

Problem 13.26. Let f be differentiable, and $z = \frac{1}{y}\big[f(ax + y) + g(ax - y)\big]$. Show that

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

Problem 13.27. Suppose that we substitute polar coordinates $x = r\cos\theta$ and $y = r\sin\theta$ in a differentiable function $z = f(x, y)$.

(1) Show that $\frac{\partial z}{\partial r} = f_x\cos\theta + f_y\sin\theta$ and $\frac{1}{r}\frac{\partial r}{\partial\theta} = -f_x\sin\theta + f_y\cos\theta$.

(2) Solve the equations in part (1) to express f_x and f_y in terms of $\frac{\partial z}{\partial r}$ and $\frac{\partial z}{\partial\theta}$.

(3) Show that $(f_x)^2 + (f_y)^2 = \Big(\dfrac{\partial z}{\partial r}\Big)^2 + \dfrac{1}{r^2}\Big(\dfrac{\partial z}{\partial \theta}\Big)^2$.

(4) Suppose in addition that f_x and f_y are differentiable. Show that

$$f_{xx} + f_{yy} = \frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2}\,.$$

Problem 13.28. Let R be an open region in $\mathbb{R}^2$ and $f : R \to \mathbb{R}$ be a real-valued function. In class we have talked about the differentiability of f. For $k \geqslant 2$, the k-times differentiability of f is defined inductively: for $k \in \mathbb{N}$, f is said to be $(k+1)$-times differentiable at (a,b) if the k-th partial derivative $\dfrac{\partial^k f}{\partial x^{k-j}\partial y^j}$ is differentiable at (a,b) for all $0 \leqslant j \leqslant k$ $\Big($note that in order to achieve this, $\dfrac{\partial^k f}{\partial x^{k-j}\partial y^j}$ has to be defined in a neighborhood of (a,b) for all $0 \leqslant j \leqslant k\Big)$. f is said to be k-times differentiable on R if f is k-times differentiable at (a,b) for all $(a,b) \in R$. f is said to be k-times continuously differentiable on R if the k-th partial derivative $\dfrac{\partial^k f}{\partial x^{k-j}\partial y^j}$ is continuous at (a,b) for all $0 \leqslant j \leqslant k$.

(1) Show that if f is $(k+1)$-times differentiable on R, then f is k-times continuously differentiable on R.

(2) Show that if f is k-times continuously differentiable on R, then f is k-times differentiable on R.

Hint: In this problem Theorem 13.35 is used (without proving yet).

Problem 13.29. Let $f(x,) = \sqrt[3]{xy}$.

(1) Show that f is continuous at $(0,0)$.

(2) Show that f_x and f_y exist at the origin but that the directional derivatives at the origin in all other directions do not exist.

Problem 13.30. Let

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^4 + y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

(1) Show that the directional derivative of f at the origin exists in all directions $\boldsymbol{u}$, and

$$(D_{\boldsymbol{u}} f)(0,0) = \Big(\frac{\partial f}{\partial x}(0,0), \frac{\partial f}{\partial y}(0,0)\Big) \cdot \boldsymbol{u}\,.$$

(2) Determine whether f is differentiable at $(0, 0)$ or not.

Problem 13.31. Let $\boldsymbol{u} = (a, b)$ be a unit vector and f be twice continuously differentiable. Show that

$$D^2_{\boldsymbol{u}} f = f_{xx} a^2 + 2 f_{xy} ab + f_{yy} b^2 \,,$$

where $D^2_{\boldsymbol{u}} f = D_{\boldsymbol{u}}(D_{\boldsymbol{u}} f)$.

Problem 13.32. Show that the operation of taking the gradient of a function has the given property. Assume that u and v are differentiable functions of x and y and that a, b are constants.

(1) $\nabla(au + bv) = a\nabla u + b\nabla v$.

(2) $\nabla(uv) = u\nabla v + v\nabla u$.

(3) $\nabla\big(\frac{u}{v}\big) = \dfrac{v\nabla u - u\nabla v}{v^2}$.

(4) $\nabla(u^n) = nu^{n-1}\nabla u$.

Problem 13.33. Show that the equation of the tangent plane to the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ at the point (x_0, y_0, z_0) can be written as

$$\frac{xx_0}{a^2} + \frac{yy_0}{b^2} + \frac{zz_0}{c^2} = 1 \,.$$

Problem 13.34. Show that the equation of the tangent plane to the elliptic paraboloid $\dfrac{z}{c} = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$ at the point (x_0, y_0, z_0) can be written as

$$\frac{2xx_0}{a^2} + \frac{2yy_0}{b^2} = \frac{z + z_0}{c} \,.$$

Problem 13.35. Let f be a differentiable function and consider the surface $z = xf\big(\frac{y}{x}\big)$. Show that the tangent plane at any point (x_0, y_0, z_0) on the surface passes through the origin.

Problem 13.36. Prove that the angle of inclination θ of the tangent plane to the surface $z = f(x, y)$ at the point (x_0, y_0, z_0) satisfies

$$\cos\theta = \frac{1}{\sqrt{f_x(x_0, y_0)^2 + f_y(x_0, y_0)^2 + 1}} \,.$$

Problem 13.37. In the following problems, find all relative extrema and saddle points of the function. Use the Second Partials Test when applicable.

(1) $f(x, y) = x^2 - xy - y^2 - 3x - y$ (2) $f(x, y) = 2xy - \frac{1}{2}(x^4 + y^4) + 1$

(3) $f(x, y) = xy - 2x - 2y - x^2 - y^2$ (4) $f(x, y) = x^3 + y^3 - 3x^2 - 3y^2 - 9x$

(5) $f(x, y) = \sqrt{56x^2 - 8y^2 - 16x - 31} + 1 - 8x$ (6) $f(x, y) = \frac{1}{x} + xy + \frac{1}{y}$

(7) $f(x, y) = \ln(x + y) + x^2 - y$ (8) $f(x, y) = 2\ln x + \ln y - 4x - y$

(9) $f(x, y) = xy \exp\big(-\frac{x^2 + y^2}{2}\big)$ (10) $f(x, y) = xy + e^{-xy}$

(11) $f(x, y) = (x^2 + y^2)e^{-x}$ (12) $f(x, y) = \Big(\frac{1}{2} - x^2 + y^2\Big)\exp(1 - x^2 - y^2)$

Problem 13.38. In the following problems, find the absolute extrema of the function over the region R (which contains boundaries).

(1) $f(x, y) = x^2 + xy$, and $R = \big\{(x, y)\,\big|\,|x| \leqslant 2, |y| \leqslant 1\big\}$

(2) $f(x, y) = 2x - 2xy + y^2$, and R is the region in the xy-plane bounded by the graphs of $y = x^2$ and $y = 1$.

(3) $f(x, y) = \dfrac{4xy}{(x^2 + 1)(y^2 + 1)}$, and $R = \big\{(x, y)\,\big|\,0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\big\}$.

(4) $f(x, y) = xy^2$, and $R = \big\{(x, y)\,\big|\,x \geqslant 0, y \geqslant 0, x^2 + y^2 \leqslant 3\big\}$.

(5) $f(x, y) = 2x^3 + y^4$, and $R = \big\{(x, y)\,\big|\,x^2 + y^2 \leqslant 1\big\}$.

Problem 13.39. Show that $f(x, y) = x^2 + 4y^2 - 4xy + 2$ has an infinite number of critical points and that the discriminant $f_{xx}f_{yy} - f_{xy}^2 = 0$ at each one. Then show that f has a local (and absolute) minimum at each critical point

Problem 13.40. Show that $f(x, y) = x^2ye^{-x^2-y^2}$ has maximum values at $\big(\pm 1, \frac{1}{\sqrt{2}}\big)$ and minimum values at $\big(\pm 1, -\frac{1}{\sqrt{2}}\big)$. Show also that f has infinitely many other critical points and the discriminant $f_{xx}f_{yy} - f_{xy}^2 = 0$ at each of them. Which of them give rise to maximum values? Minimum values? Saddle points?

Problem 13.41. Find two numbers a and b with $a \leqslant b$ such that

$$\int_a^b \sqrt[3]{24 - 2x - x^2}\, dx$$

has its largest value.

Problem 13.42. Let $m > n$ be natural numbers, and A be an $m \times n$ real matrix, $\boldsymbol{b} \in \mathbb{R}^m$ be a vector.

(1) Show that if the minimum of the function $f(x_1, \cdots, x_n) = \|A\boldsymbol{x} - \boldsymbol{b}\|$ occurs at the point $\boldsymbol{c} = (c_1, \cdots, c_n)$, then $\boldsymbol{c}$ satisfies $A^{\mathrm{T}} A\boldsymbol{c} = A^{\mathrm{T}}\boldsymbol{b}$.

(2) Find the relation between the linear regression and (1).

Problem 13.43. Let $\big\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\big\}$ be n points with $x_i \neq x_j$ if $i \neq j$. Use the Second Partials Test to verify that the formulas for a and b given by

$$a = \frac{n \sum\limits_{i=1}^{n} x_i y_i - \Big(\sum\limits_{i=1}^{n} x_i\Big)\Big(\sum\limits_{i=1}^{n} y_i\Big)}{n \sum\limits_{i=1}^{n} x_i^2 - \Big(\sum\limits_{i=1}^{n} x_i\Big)^2} \quad \text{and} \quad b = \frac{1}{n}\Big(\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i\Big)$$

indeed minimize the function $S(a, b) = \sum\limits_{i=1}^{n} (ax_i + b - y_i)^2$.

Problem 13.44. The Shannon index (sometimes called the Shannon-Wiener index or Shannon-Weaver index) is a measure of diversity in an ecosystem. For the case of three species, it is defined as

$$H = -p_1 \ln p_1 - p_2 \ln p_2 - p_3 \ln p_3\,,$$

where p_i is the proportion of species i in the ecosystem.

(1) Express H as a function of two variables using the fact that $p_1 + p_2 + p_3 = 1$.

(2) What is the domain of H?

(3) Find the maximum value of H. For what values of p_1, p_2, p_3 does it occur?

Problem 13.45. Three alleles (alternative versions of a gene) A, B, and O determine the four blood types A (AA or AO), B (BB or BO), O (OO), and AB. The Hardy-Weinberg

Law states that the proportion of individuals in a population who carry two different alleles is

$$P = 2pq + 2pr + 2rq\,,$$

where p, q, and r are the proportions of A, B, and O in the population. Use the fact that $p+q+r=1$ to show that P is at most $\frac{2}{3}$.

Problem 13.46. Find an equation of the plane that passes through the point $(1,2,3)$ and cuts off the smallest volume in the first octant.

Problem 13.47. Use the method of Lagrange multipliers to complete the following.

(1) Maximize $f(x,y) = \sqrt{6-x^2-y^2}$ subject to the constraint $x+y-2=0$.

(2) Minimize $f(x,y) = 3x^2 - y^2$ subject to the constraint $2x-2y+5=0$.

(3) Minimize $f(x,y) = x^2+y^2$ subject to the constraint $xy^2 = 54$.

(4) Maximize $f(x,y,z) = e^{xyz}$ subject to the constraint $2x^2+y^2+z^2 = 24$.

(5) Maximize $f(x,y,z) = \ln(x^2+1) + \ln(y^2+1) + \ln(z^2+1)$ subject to the constraint $x^2+y^2+z^2 = 12$.

(6) Maximize $f(x,y,z) = x+y+z$ subject to the constraint $x^2+y^2+z^2=1$.

(7) Maximize $f(x,y,z,t) = x+y+z+t$ subject to the constraint $x^2+y^2+z^2+t^2=1$.

(8) Minimize $f(x,y,z) = x^2+y^2+z^2$ subject to the constraints $x+2z=6$ and $x+y=12$.

(9) Maximize $f(x,y,z) = z$ subject to the constraints $x^2+y^2+z^2=36$ and $2x+y-z=2$.

(10) Maximize $f(x,y,z) = yz+xy$ subject to the constraint $xy=1$ and $y^2+z^2=1$.

Problem 13.48. Use the method of Lagrange multipliers to find the extreme values of the function $f(x_1,x_2,\cdots,x_n) = x_1+x_2+\cdots+x_n$ subject to the constraint $x_1^2+x_2^2+\cdots+x_n^2=1$.

Problem 13.49. (1) Use the method of Lagrange multipliers to show that the product of three positive numbers x, y, and z, whose sum has the constant value S, is a maximum when the three numbers are equal. Use this result to show that

$$\frac{x+y+z}{3} \geqslant \sqrt[3]{xyz} \qquad \forall\, x,y,z>0\,.$$

(2) Generalize the result of part (1) to prove that the product $x_1x_2x_3\cdots x_n$ is maximized, under the constraint that $\sum\limits_{i=1}^{n} x_i = S$ and $x_i \geqslant 0$ for all $1 \leqslant i \leqslant n$, when

$$x_1 = x_2 = x_3 = \cdots = x_n\,.$$

Then prove that

$$\sqrt[n]{x_1x_2\cdots x_n} \leqslant \frac{x_1 + x_2 + \cdots + x_n}{n} \qquad \forall\, x_1, x_2, \cdots, x_n \geqslant 0\,.$$

Problem 13.50. (1) Maximize $\sum\limits_{i=1}^{n} x_iy_i$ subject to the constraints $\sum\limits_{i=1}^{n} x_i^2 = 1$ and $\sum\limits_{i=1}^{n} y_i^2 = 1$.

(2) Put $x_i = \dfrac{a_i}{\sqrt{\sum\limits_{j=1}^{n} a_j^2}}$ and $y_i = \dfrac{b_i}{\sqrt{\sum\limits_{j=1}^{n} b_j^2}}$ to show that

$$\sum_{i=1}^{n} a_ib_i \leqslant \sqrt{\sum_{j=1}^{n} a_j^2}\sqrt{\sum_{j=1}^{n} b_j^2}$$

for any numbers $a_1, a_2, \cdots, a_n$, $b_1, b_2, \cdots, b_n$. This inequality is known as the Cauchy-Schwarz Inequality.

Problem 13.51. Find the points on the curve $x^2 + xy + y^2 = 1$ in the xy-plane that are nearest to and farthest from the origin.

Problem 13.52. If the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ is to enclose the circle $x^2 + y^2 = 2y$, what values of a and b minimize the area of the ellipse?

Problem 13.53. (1) Use the method of Lagrange multipliers to prove that the rectangle with maximum area that has a given perimeter p is a square.

(2) Use the method of Lagrange multipliers to prove that the triangle with maximum area that has a given perimeter p is equilateral.

Hint: Use Heron's formula for the area:

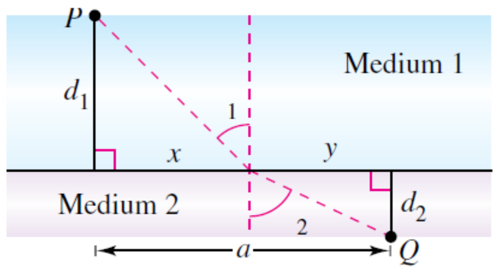
$$A = \sqrt{s(s-x)(s-y)(s-z)}\,,$$

where $s = \dfrac{p}{2}$ and x, y, z are the lengths of the sides.

Problem 13.54. When light waves traveling in a transparent medium strike the surface of a second transparent medium, they tend to "bend" in order to follow the path of minimum time. This tendency is called refraction and is described by Snell' s Law of Refraction,

$$\frac{\sin\theta_1}{\mathrm{v}_1} = \frac{\sin\theta_2}{\mathrm{v}_2}\,,$$

where θ_1 and θ_2 are the magnitudes of the angles shown in the figure, and v_1 and v_2 are the velocities of light in the two media. Use the method of Lagrange multipliers to derive this law using $x + y = a$.



Problem 13.55. A set $C \subseteq \mathbb{R}^n$ is said to be convex if

$$t\boldsymbol{x} + (1-t)\boldsymbol{y} \in C \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in C \text{ and } t \in [0,1]\,.$$

(一個 $\mathbb{R}^n$ 中的集合 C 被稱為凸集合如果 C 中任兩點 $\boldsymbol{x}$ 與 $\boldsymbol{y}$ 之連線所形成的線段也在 C 中)。

Suppose that $C \subseteq \mathbb{R}^n$ is a convex set, and $f : C \to \mathbb{R}$ be a differentiable real-valued function. Show that if f on C attains its minimum at a point $\boldsymbol{x}^*$, then

$$(\nabla f)(\boldsymbol{x}^*) \cdot (\boldsymbol{x} - \boldsymbol{x}^*) \geqslant 0 \qquad \forall\, \boldsymbol{x} \in C\,. \tag{$\star$}$$

Hint: Recall that $(\nabla f)(\boldsymbol{x}^*) \cdot (\boldsymbol{x} - \boldsymbol{x}^*)$, when f is differentiable at x^*, is the directional derivative of f at $\boldsymbol{x}^*$ in the "direction" $(\boldsymbol{x} - \boldsymbol{x}^*)$.
Remark: A point $\boldsymbol{x}^*$ satisfying ($\star$) is sometimes called a ***stationary point*** of f in C.

Problem 13.56. Let B be the unit closed ball centered at the origin given by

$$B = \left\{\boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \,\middle|\, \|\boldsymbol{x}\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2 \leqslant 1\right\},$$

and $f : B \to \mathbb{R}$ be a differentiable real-valued function. Consider the minimization problem $\min\limits_{\boldsymbol{x}\in B} f(\boldsymbol{x})$.

(1) Show that if f attains its minimum at $\boldsymbol{x}^* \in B$, then there exists $\lambda \leqslant 0$ such that

$$(\nabla f)(\boldsymbol{x}^*) = \lambda \boldsymbol{x}^* .$$

(2) Find the minimum of the function $f(x, y) = x^2 + 2y^2 - x$ on the unit closed disk centered at the origin $\{(x, y) \,|\, x^2 + y^2 \leqslant 1\}$ using (1).

Problem 13.57. Let $\boldsymbol{a} \in \mathbb{R}^3$ be a vector, $b \in \mathbb{R}$, and C be a half plane given by

$$C = \{\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, \boldsymbol{a} \cdot \boldsymbol{x} \leqslant b\} ,$$

and $f : C \to \mathbb{R}$ be a differentiable real-valued function. Consider the minimization problem $\min\limits_{\boldsymbol{x} \in C} f(\boldsymbol{x})$. Show that if f attains its minimum at $\boldsymbol{x}^* \in C$, then there exists $\lambda \leqslant 0$ such that

$$(\nabla f)(\boldsymbol{x}^*) = \lambda \boldsymbol{a} .$$

Chapter 14

Multiple Integration

14.1 Double Integrals and Volume

Let R be a closed and bounded region in the plane, and $f : R \to \mathbb{R}$ be a non-negative continuous function. We are interested in the volume of the solid in space

$$\mathrm{D} = \big\{(x, y, z) \,\big|\, (x, y) \in R\,, 0 \leqslant z \leqslant f(x, y)\big\}\,.$$

First we assume that $R = [a, b] \times [c, b] = \big\{(x, y) \,\big|\, a \leqslant x \leqslant b, c \leqslant y \leqslant d\big\}$ be a rectangle. Let $\mathcal{P}_x = \{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$ and $\mathcal{P}_y = \{c = y_0 < y_1 < \cdots < y_m = d\}$ be partitions of $[a, b]$ and $[c, d]$, respectively, R_{ij} denote the rectangle $[x_{i-1}, x_i] \times [y_{j-1}, y_j]$, and $\big\{(\alpha_i, \beta_j)\big\}_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m}$ be a collection of points such that $\alpha_i \in [x_{i-1}, x_i]$ and $\beta_j \in [y_{j-1}, y_j]$. Then as before, we consider an approximation of the volume of D given by

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f(\alpha_i, \beta_j)(x_i - x_{i-1})(y_j - y_{j-1})\,.$$

Then the limit of the sum above, as $\|\mathcal{P}_x\|$, $\|\mathcal{P}_y\|$ approaches zero, is the volume of D. The collection of rectangles $\mathcal{P} = \{R_{ij}\}_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m}$ is called a partition of R.

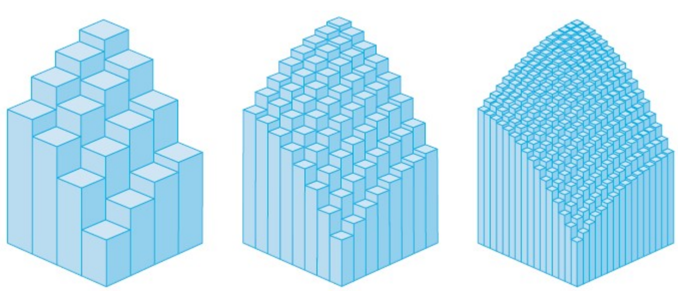


Figure 14.1: The volume of D can be obtained by making $\|\mathcal{P}_x\|, \|P_y\| \to 0$.

In general, by relabeling the rectangles as $R_1, R_2, \cdots, R_{nm}$ (thus $\mathcal{P} = \{R_k \,|\, 1 \leqslant k \leqslant nm\}$), and letting $\big\{(\xi_k, \eta_k)\big\}_{k=1}^{nm}$ be a collection of point in R such that $(\xi_k, \eta_k) \in R_k$, we can consider an approximation of the volume of the solid given by

$$\sum_{k=1}^{nm} f(\xi_k, \eta_k) A_k \,,$$

where A_k is the area of the rectangle R_k. The sum above is called a ***Riemann sum of*** f ***for partition*** $\mathcal{P}$. With $\|\mathcal{P}\|$, called the norm of $\mathcal{P}$, denoting the maximum length of the diagonal of R_k; that is,

$$\|\mathcal{P}\| = \max\big\{\ell_k \,\big|\, \ell_k \text{ is the length of the diagonal of } R_k, 1 \leqslant k \leqslant nm\big\}\,,$$

then the volume of D is the "limit"

$$\lim_{\|\mathcal{P}\| \to 0} \sum_{k=1}^{nm} f(\xi_k, \eta_k) A_k$$

as long as "the limit exists". Similar to the discussion of the limit of Riemann sums in the case of functions of one variable, we can remove the restrictions that f is continuous and non-negative on R and still consider the limit of the Riemann sums. We have the following

Definition 14.1

Let $R = [a, b] \times [c, d]$ be a rectangle in the plane, and $f : R \to \mathbb{R}$ be a function. f is said to be Riemann integrable on R if there exists a real number V such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is partition of R satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sums of f for the partition $\mathcal{P}$ belongs to the interval $(V - \varepsilon, V + \varepsilon)$. Such a number V (is unique if it exists and) is called the ***Riemann integral*** or ***double integral of*** f ***on*** R and is denoted by $\displaystyle\iint_R f(x, y)\, dA$ or simply $\displaystyle\int_R f(x, y)\, d(x, y)$.

How about the case that the base R of the solid is not a closed and bounded rectangle? In this case we choose $r > 0$ large enough such that $R \subseteq [-r, r]^2 \equiv [-r, r] \times [-r, r]$ and then for a function $f : R \to \mathbb{R}$, define $\widetilde{f} : [-r, r]^2 \to \mathbb{R}$ by

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{if } x \in R\,, \\ 0 & \text{if } x \notin R\,. \end{cases}$$

We define $\displaystyle\iint_R f(x, y)\, dA$ as $\displaystyle\iint_{[-r,r]^2} \widetilde{f}(x, y)\, dA$ (when the latter double integral exists).

Before proceeding, let us talk about a special class of regions.

Definition 14.2

A region R is said to be have area if the constant function 1 is Riemann integrable on R. If R has area, then the area of R is defined as the integral $\iint\limits_R 1\,dA$.

The following theorem is an analogy of Theorem 4.10.

Theorem 14.3

Let R be a closed and bounded region in the plane, and $f : R \to \mathbb{R}$ be a function. If R has area and f is continuous on R, then f is Riemann integrable on R.

Similar to the properties for integrals of functions of one variable, we have the following

Theorem 14.4: Properties of double integrals

Let R be a closed and bounded region in the plane, $f, g : R \to \mathbb{R}$ be functions that are Riemann integrable on R, and c be a real number.

1. cf is Riemann integrable on R, and
$$\iint\limits_R (cf)(x,y)\,dA = c\iint\limits_R f(x,y)\,dA\,.$$
2. $f \pm g$ are Riemann integrable on R, and
$$\iint\limits_R (f \pm g)(x,y)\,dA = \iint\limits_R f(x,y)\,dA \pm \iint\limits_R g(x,y)\,dA\,.$$
3. If $f(x,y) \geqslant g(x,y)$ for all $(x,y) \in R$, then
$$\iint\limits_R f(x,y)\,dA \geqslant \iint\limits_R g(x,y)\,dA\,.$$
4. $|f|$ is Riemann integrable, and
$$\Big|\iint\limits_R f(x,y)\,dA\Big| \leqslant \iint\limits_R \big|f(x,y)\big|\,dA\,.$$

Definition 14.5

Two bounded regions R_1 and R_2 in the plane are said to be non-overlapping if $R_1 \cap R_2$ has zero area.

Theorem 14.6

Let R_1 and R_2 be non-overlapping regions in the plane, $R = R_1 \cup R_2$, and $f : R \to \mathbb{R}$ be such that f is Riemann integrable on R_1 and R_2. Then f is Riemann integrable on R and

$$\iint\limits_{R} f(x,y)\,dA = \iint\limits_{R_1} f(x,y)\,dA + \iint\limits_{R_2} f(x,y)\,dA\,.$$

14.2 The Iterated Integrals and Fubini's Theorem

Let R be a bounded region with area, and $f : R \to \mathbb{R}$ be a non-negative continuous function. As explained in the previous section, the volume of the solid

$$\mathrm{D} = \big\{(x,y,z)\,\big|\,(x,y) \in R, 0 \leqslant z \leqslant f(x,y)\big\}$$

is given by $\displaystyle\iint\limits_{R} f(x,y)\,dA$. We are concerned with computing this double integral in this section.

Recall from Section 7.2 that if D is a solid lies between two planes $x = a$ and $x = b$ $(a < b)$, and the area of the cross section of D taken perpendicular to the x-axis is $A(x)$, then

$$\text{the volume of D} = \int_a^b A(x)\,dx\,.$$

Therefore, if the region R is given by

$$R = \big\{(x,y)\,\big|\,a \leqslant x \leqslant b, g_1(x) \leqslant y \leqslant g_2(x)\big\}$$

for some continuous functions $g_1, g_2 : [a,b] \to \mathbb{R}$, then the area of the cross section of D taken perpendicular to the x axis is

$$A(x) = \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy$$

which shows that the volume of D is given by $\displaystyle\int_a^b \Big(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy \Big)\,dx$. Therefore, in this special case we find that

$$\iint_R f(x,y)\,dA = \int_a^b \Big(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy \Big)\,dx\,. \tag{14.2.1}$$

Similarly, recall that if D lies between $y = c$ and $y = d$ ($c < d$), and the area of the cross section of D taken perpendicular to the y-axis is $A(y)$, then

$$\text{the volume of D} = \int_c^d A(y)\,dy\,;$$

thus similar argument shows that

$$\iint_R f(x,y)\,dA = \int_c^d \Big(\int_{h_1(y)}^{h_2(y)} f(x,y)\,dx \Big)\,dy\,. \tag{14.2.2}$$

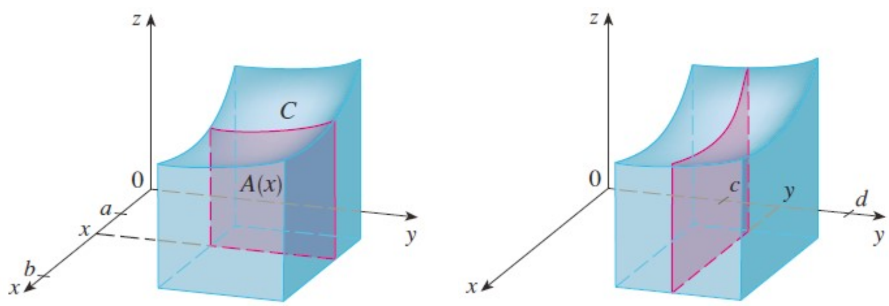


Figure 14.2: Finding the volume of D using the method of cross section

We note that in formulas (14.2.1), we have to compute the integral $\displaystyle\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy$ for each fixed $x \in [a,b]$ which gives the area of the cross section $A(x)$, then compute the integral $\displaystyle\int_a^b A(x)\,dx$ to obtain the volume of D. This way of computing double integrals is called ***iterated integrals***, and sometime we omit the parentheses and write it as

$$\iint_R f(x,y)\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dydx\,.$$

Similarly, the iterated integral appearing in (14.2.2) can also be written as

$$\iint_R f(x,y)\,dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dxdy\,.$$

The evaluation of the double integral $\displaystyle\iint\limits_R f(x,y)\,dA$ can be generalized for a more general class of functions, and it is called the Fubini Theorem.

Theorem 14.7: Fubini's Theorem

Let R be a region in the plane, and $f : R \to \mathbb{R}$ be continuous (but no necessary non-negative).

1. If R is given by $R = \big\{(x,y)\,\big|\, a \leqslant x \leqslant b, g_1(x) \leqslant y \leqslant g_2(x)\big\}$, then
$$\iint\limits_R f(x,y)\,dA = \int_a^b \Big(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\Big)\,dx\,.$$
2. If R is given by $R = \big\{(x,y)\,\big|\, c \leqslant y \leqslant d, g_1(x) \leqslant y \leqslant g_2(x)\big\}$, then
$$\iint\limits_R f(x,y)\,dA = \int_c^d \Big(\int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\Big)\,dy\,.$$

Example 14.8. Find the volume of the solid region bounded by the paraboloid $z = 4 - x^2 - 2y^2$ and the xy-plane. By the definition of double integrals, the volume of this solid is given by $\displaystyle\iint\limits_R (4 - x^2 - 2y^2)\,dA$, where R is the region $\big\{(x,y)\,\big|\, x^2 + 2y^2 \leqslant 4\big\}$. Writing R as
$$R = \Big\{(x,y)\,\Big|\, -2 \leqslant x \leqslant 2\,, -\sqrt{\frac{4-x^2}{2}} \leqslant y \leqslant \sqrt{\frac{4-x^2}{2}}\Big\}$$
or
$$R = \big\{(x,y)\,\big|\, -\sqrt{2} \leqslant y \leqslant \sqrt{2}\,, -\sqrt{4-2y^2} \leqslant x \leqslant \sqrt{4-2y^2}\big\}\,,$$
the Fubini Theorem then implies that
$$\begin{aligned}\iint\limits_R (4 - x^2 - 2y^2)\,dA &= \int_{-2}^{2} \Big(\int_{-\sqrt{\frac{4-x^2}{2}}}^{\sqrt{\frac{4-x^2}{2}}} (4 - x^2 - 2y^2)\,dy\Big)\,dx \\ &= \int_{-\sqrt{2}}^{\sqrt{2}} \Big(\int_{-\sqrt{4-2y^2}}^{\sqrt{4-2y^2}} (4 - x^2 - 2y^2)\,dx\Big)\,dy\,.\end{aligned}$$

1. Integrating in y first then integrating in x: for fixed $x \in [-2,2]$,

$$\int_{-\sqrt{\frac{4-x^2}{2}}}^{\sqrt{\frac{4-x^2}{2}}}(4-x^2-2y^2)\,dy = \int_{-\sqrt{\frac{4-x^2}{2}}}^{\sqrt{\frac{4-x^2}{2}}}(4-x^2)\,dy - 2\int_{-\sqrt{\frac{4-x^2}{2}}}^{\sqrt{\frac{4-x^2}{2}}} y^2\,dy$$
$$= \sqrt{2}(4-x^2)^{\frac{3}{2}} - \frac{4}{3}\Big(\sqrt{\frac{4-x^2}{2}}\Big)^3 = \frac{2\sqrt{2}}{3}(4-x^2)^{\frac{3}{2}}\,.$$

Therefore, by the substitution $x = 2\sin\theta$ (so that $dx = 2\cos\theta\,d\theta$),

$$\iint\limits_R (4-x^2-2y^2)\,dA = \frac{2\sqrt{2}}{3}\int_{-2}^{2}(4-x^2)^{\frac{3}{2}}\,dx = \frac{2\sqrt{2}}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 8\cos^3\theta\cdot 2\cos\theta d\theta$$
$$= \frac{32\sqrt{2}}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^4\theta\,d\theta = \frac{64\sqrt{2}}{3}\int_{0}^{\frac{\pi}{2}}\cos^4\theta\,d\theta$$
$$= \frac{64\sqrt{2}}{3}\int_{0}^{\frac{\pi}{2}}\Big(\frac{1+\cos 2\theta}{2}\Big)^2\,d\theta$$
$$= \frac{16\sqrt{2}}{3}\int_{0}^{\frac{\pi}{2}}\Big(1+2\cos 2\theta+\frac{1+\cos 4\theta}{2}\Big)\,d\theta$$
$$= \frac{16\sqrt{2}}{3}\Big[\frac{3}{2}\cdot\frac{\pi}{2}+\sin\Big(2\cdot\frac{\pi}{2}\Big)+\frac{1}{8}\sin\Big(4\cdot\frac{\pi}{2}\Big)\Big] = 4\sqrt{2}\pi\,.$$

2. Integrating in x first then integrating in y: for fixed $y \in [-\sqrt{2}, \sqrt{2}]$,

$$\int_{-\sqrt{4-2y^2}}^{\sqrt{4-2y^2}}(4-x^2-2y^2)\,dx = \int_{-\sqrt{4-2y^2}}^{\sqrt{4-2y^2}}(4-2y^2)\,dx - \int_{-\sqrt{4-2y^2}}^{\sqrt{4-2y^2}} x^2\,dx$$
$$= 2(4-2y^2)^{\frac{3}{2}} - \frac{2}{3}(4-2y^2)^{\frac{3}{2}} = \frac{4}{3}(4-2y^2)^{\frac{3}{2}}\,;$$

thus by the substitution of variable $y = \sqrt{2}\sin\theta$ (so that $dy = \sqrt{2}\cos\theta\,d\theta$),

$$\iint\limits_R (4-x^2-2y^2)\,dA = \frac{4}{3}\int_{-\sqrt{2}}^{\sqrt{2}}(4-2y^2)^{\frac{3}{2}}\,dy = \frac{4}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 8\cos^3\theta\cdot\sqrt{2}\cos\theta\,d\theta$$
$$= \frac{32\sqrt{2}}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos^4\theta\,d\theta = \frac{64\sqrt{2}}{3}\int_{0}^{\frac{\pi}{2}}\cos^4\theta\,d\theta = 4\sqrt{2}\pi\,.$$

Example 14.9. Find the volume of the solid region bounded above by the paraboloid $z = 1-x^2-y^2$ and below by the plane $z = 1-y$.

Let R be the region in the plane whose boundary points (x, y) satisfies $1-x^2-y^2 = 1-y$ or equivalently, $x^2+y^2-y = 0$. Then the volume of the solid described above is given by

$\iint\limits_R \big[(1-x^2-y^2)-(1-y)\big]\, dA$. Note that the region R is a disk centered at $\big(0,\frac{1}{2}\big)$ with radius $\frac{1}{2}$ and can be written as

$$R = \big\{(x,y)\,\big|\, 0 \leqslant y \leqslant 1\,, -\sqrt{y-y^2} \leqslant x \leqslant \sqrt{y-y^2}\big\}\,.$$

Therefore,

$$\begin{aligned}\iint\limits_R \big[(1-x^2-y^2)-(1-y)\big]\, dA &= \int_0^1 \Big(\int_{-\sqrt{y-y^2}}^{\sqrt{y-y^2}} (y-x^2-y^2)\, dx\Big) dy \\ &= \int_0^1 \Big(2(y-y^2)^{\frac{3}{2}} - \frac{2}{3}(y-y^2)^{\frac{3}{2}}\Big)\, dy = \frac{4}{3}\int_0^1 (y-y^2)^{\frac{3}{2}}\, dy = \frac{4}{3}\int_0^1 \Big[\frac{1}{4} - \big(y-\frac{1}{2}\big)^2\Big]^{\frac{3}{2}} dy\,.\end{aligned}$$

Making the substitution of variable $y - \frac{1}{2} = \frac{1}{2}\sin\theta$ $\big($so that $dy = \frac{1}{2}\cos\theta\, d\theta\big)$,

$$\iint\limits_R \big[(1-x^2-y^2)-(1-y)\big]\, dA = \frac{4}{3}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\cos^3\theta}{8}\cdot\frac{1}{2}\cos\theta\, d\theta = \frac{1}{6}\int_0^{\frac{\pi}{2}} \cos^4\theta\, d\theta = \frac{\pi}{32}\,.$$

Example 14.10. Find the iterated integral $\displaystyle\int_0^1 \Big(\int_y^1 e^{-x^2}\, dx\Big) dy$.

Let $R = \big\{(x,y)\,\big|\, 0 \leqslant y \leqslant 1\,, y \leqslant x \leqslant 1\big\}$. Since R can also be expressed as $R = \big\{(x,y)\,\big|\, 0 \leqslant x \leqslant 1\,, 0 \leqslant y \leqslant x\big\}$, by the Fubini Theorem we find that

$$\begin{aligned}\int_0^1 \Big(\int_y^1 e^{-x^2}\, dx\Big)\, dy &= \iint\limits_R e^{-x^2}\, dA = \int_0^1 \Big(\int_0^x e^{-x^2}\, dy\Big) dx = \int_0^1 x e^{-x^2}\, dx \\ &= -\frac{1}{2}e^{-x^2}\Big|_{x=0}^{x=1} = \frac{1}{2}\big(1-e^{-1}\big)\,.\end{aligned}$$

14.3 Surface Area

14.3.1 Surface area of graph of functions

Let $R = [a,b]\times[c,d]$ be a rectangle in the plane, and $f : R \to \mathbb{R}$ be a continuously differentiable function. We are interested in the area of the surface

$$S = \big\{(x,y,z)\,\big|\,(x,y)\in R\,, z = f(x,y)\big\}\,.$$

Let $\mathcal{P} = \big\{R_{ij}\,\big|\, 1 \leqslant i \leqslant n\,, 1 \leqslant j \leqslant m\big\}$ be a partition of R. Partition each rectangle $R_{ij} = [x_{i-1},x_i]\times[y_{j-1},y_j]$ into two triangles Δ_{ij}^1 and Δ_{ij}^2, where Δ_{ij}^1 has vertices (x_{i-1},y_{j-1}),

(x_i, y_{j-1}), (x_{i-1}, y_j) and Δ^2_{ij} has vertices (x_i, y_j), (x_{i-1}, y_j), (x_i, y_{j-1}). Then intuitively, the area of the surface $f(\Delta^1_{ij})$ can be approximated by the area of the triangle T^1_{ij} with vertices $\big(x_{i-1}, y_{j-1}, f(x_{i-1}, y_{j-1})\big)$, $\big(x_i, y_{j-1}, f(x_i, y_{j-1})\big)$ and $\big(x_i, y_j, f(x_i, y_j)\big)$, while the area of the surface $f(\Delta^2_{ij})$ can be approximated by the area of the triangle T^2_{ij} with vertices $\big(x_i, y_j, f(x_i, y_j)\big)$, $\big(x_{i-1}, y_j, f(x_{i-1}, y_j)\big)$ and $\big(x_i, y_{j-1}, f(x_i, y_{j-1})\big)$. Therefore, the area of the surface $f(R_{ij})$ can be approximated by the sum of area of triangles T^1_{ij} and T^2_{ij}, and the area of the surface S can be approximated by the sum of the area of the triangles T^1_{ij} and T^2_{ij}, where is sum is taken over all $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$.

Now we compute the area of the triangles T^1_{ij} and T^2_{ij}. We remark that for a triangle T with vertices P_1, P_2, P_3, letting $\boldsymbol{u} = \overrightarrow{P_1P_2} = P_2 - P_1$ and $\boldsymbol{v} = \overrightarrow{P_1P_3} = P_3 - P_1$, the area of T can be computed by $\dfrac{1}{2}\|\boldsymbol{u} \times \boldsymbol{v}\|$. Therefore, the area of T^1_{ij} is given by

$$\begin{aligned} |T^1_{ij}| = \frac{1}{2}\Big\| &\big(x_i - x_{i-1}, 0, f(x_i, y_{j-1}) - f(x_{i-1}, y_{j-1})\big) \times \\ &\times \big(0, y_j - y_{j-1}, f(x_{i-1}, y_j) - f(x_{i-1}, y_{j-1})\big)\Big\| . \end{aligned}$$

By the mean value theorem, there exist $\xi^*_i \in (x_{i-1}, x_i)$ and $\eta^*_j \in (y_{j-1}, y_j)$ such that

$$\begin{aligned} f(x_i, y_{j-1}) - f(x_{i-1}, y_{j-1}) &= f_x(\xi^*_i, y_{j-1})(x_i - x_{i-1}) , \\ f(x_{i-1}, y_j) - f(x_{i-1}, y_{j-1}) &= f_y(x_{i-1}, \eta^*_j)(y_j - y_{j-1}) ; \end{aligned}$$

thus we obtain that

$$\begin{aligned} |T^1_{ij}| &= \frac{1}{2}\big\|\big(1, 0, f_x(\xi^*_i, y_{j-1})\big) \times \big(0, 1, f_y(x_{i-1}, \eta^*_j)\big)\big\| \\ &= \frac{1}{2}\big\|\big(-f_x(\xi^*_i, y_{j-1}), -f_y(x_{i-1}, \eta^*_j), 1\big)\big\|(x_i - x_{i-1})(y_j - y_{j-1}) \\ &= \frac{1}{2}\sqrt{1 + f_x(\xi^*_i, y_{j-1})^2 + f_y(x_{i-1}, \eta^*_j)^2}(x_i - x_{i-1})(y_j - y_{j-1}) . \end{aligned}$$

Similarly, there exist $\xi^{**}_i \in (x_{i-1}, x_i)$ and $\eta^{**}_j \in (y_{j-1}, y_j)$ such that the area of the triangle T^2_{ij} is given by

$$|T^2_{ij}| = \frac{1}{2}\sqrt{1 + f_x(\xi^{**}_i, y_j)^2 + f_y(x_i, \eta^{**}_j)^2}(x_i - x_{i-1})(y_j - y_{j-1}) .$$

Let $M = \max\limits_{(x,y)\in R}\big(|f_x(x, y)| + |f_y(x, y)|\big)$, $|R| = (b - a)(d - c)$, and $\varepsilon > 0$ be a given (but arbitrary) number. Suppose that

$$\big|f_x(\alpha, \beta) - f_x(\xi, \eta)\big| + \big|f_y(\alpha, \beta) - f_y(\xi, \eta)\big| < \frac{\varepsilon}{2|R|(1 + M)} \qquad \forall\, (\alpha, \beta), (\xi, \eta) \in R_{ij} . \quad (14.3.1)$$

Then

$$
\begin{aligned}
&\Big|\sqrt{1+f_x(\alpha,\beta)^2+f_y(\alpha^*,\beta^*)^2}-\sqrt{1+f_x(\xi,\eta)^2+f_y(\xi,\eta)^2}\Big| \\
&\quad=\left|\frac{f_x(\alpha,\beta)^2+f_y(\alpha^*,\beta^*)^2-f_x(\xi,\eta)^2-f_y(\xi,\eta)^2}{\sqrt{1+f_x(\alpha,\beta)^2+f_y(\alpha^*,\beta^*)^2}+\sqrt{1+f_x(\xi,\eta)^2+f_y(\xi,\eta)^2}}\right| \\
&\quad\leqslant\frac{1}{2}\Big[\big|f_x(\alpha,\beta)-f_x(\xi,\eta)\big|\big|f_x(\alpha,\beta)+f_x(\xi,\eta)\big| \\
&\qquad\quad+\big|f_y(\alpha^*,\beta^*)-f_y(\xi,\eta)\big|\big|f_y(\alpha^*,\beta^*)+f_y(\xi,\eta)\big|\Big] \\
&\quad\leqslant\frac{2M}{2}\Big[\big|f_x(\alpha,\beta)-f_x(\xi,\eta)\big|+\big|f_y(\alpha^*,\beta^*)-f_y(\xi,\eta)\big|\Big]\leqslant\frac{M\varepsilon}{2|R|(1+M)}<\frac{\varepsilon}{2|R|}\,.
\end{aligned}
$$

Therefore, if (14.3.1) holds for all $1\leqslant i\leqslant n$ and $1\leqslant j\leqslant m$, then for $(\xi_{ij},\eta_{ij})\in R_{ij}$, we have

$$
\begin{aligned}
&\Big||T_{ij}^1|+|T_{ij}^2|-\sqrt{1+f_x(\xi_{ij},\eta_{ij})^2+f_y(\xi_{ij},\eta_{ij})^2}(x_i-x_{i-1})(y_j-y_{j-1})\Big| \\
&\quad\leqslant\Big|\frac{1}{2}\sqrt{1+f_x(\xi_i^*,y_{j-1})^2+f_y(x_{i-1},\eta_j^*)^2}+\frac{1}{2}\sqrt{1+f_x(\xi_i^{**},y_j)^2+f_y(x_i,\eta_j^{**})^2} \\
&\qquad-\sqrt{1+f_x(\xi_{ij},\eta_{ij})^2+f_y(\xi_{ij},\eta_{ij})^2}\Big|(x_i-x_{i-1})(y_j-y_{j-1}) \\
&\quad\leqslant\frac{\varepsilon}{2|R|}(x_i-x_{i-1})(y_j-y_{j-1})\,;
\end{aligned}
$$

thus if (14.3.1) holds for all $1\leqslant i\leqslant n$ and $1\leqslant j\leqslant m$, then for $(\xi_{ij},\eta_{ij})\in R_{ij}$,

$$
\begin{aligned}
&\left|\sum_{i=1}^{n}\sum_{j=1}^{m}\big(|T_{ij}^1|+|T_{ij}^2|\big)-\sum_{i=1}^{n}\sum_{j=1}^{m}\sqrt{1+f_x(\xi_{ij},\eta_{ij})^2+f_y(\xi_{ij},\eta_{ij})^2}(x_i-x_{i-1})(y_j-y_{j-1})\right| \\
&\quad\leqslant\sum_{i=1}^{n}\sum_{j=1}^{m}\Big||T_{ij}^1|+|T_{ij}^2|-\sqrt{1+f_x(\xi_{ij},\eta_{ij})^2+f_y(\xi_{ij},\eta_{ij})^2}(x_i-x_{i-1})(y_j-y_{j-1})\Big| \\
&\quad\leqslant\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{\varepsilon}{2|R|}(x_i-x_{i-1})(y_j-y_{j-1})=\frac{\varepsilon}{2}\,.
\end{aligned}
$$

Finally, we state as a fact that there exists $\delta_1>0$ such that (14.3.1) holds as long as $\|\mathcal{P}\|<\delta_1$. This property is called the ***uniform continuity*** of continuous functions on closed and bounded sets.

On the other hand, since the function $z=\sqrt{1+f_x(x,y)^2+f_y(x,y)^2}$ is continuous on R (and R has area), it is Riemann integrable on R. Let

$$
\mathrm{I}=\iint\limits_{R}\sqrt{1+f_x(x,y)^2+f_y(x,y)^2}\,dA\,.
$$

Then there exists $\delta_2 > 0$ such that if $\mathcal{P}$ is a partition of R satisfying $\|\mathcal{P}\| < \delta_2$, then any Riemann sum of f for the partition $\mathcal{P}$ belongs to $\big(\mathrm{I} - \frac{\varepsilon}{2}, \mathrm{I} + \frac{\varepsilon}{2}\big)$. Therefore,

$$\Big|\sum_{i=1}^{n}\sum_{j=1}^{m}\sqrt{1 + f_x(\xi_{ij}, \eta_{ij})^2 + f_y(\xi_{ij}, \eta_{ij})^2}(x_i - x_{i-1})(y_j - y_{j-1}) - \mathrm{I}\Big| < \frac{\varepsilon}{2}\,.$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Then $\delta > 0$, and if $\mathcal{P} = \big\{R_{ij}\,\big|\,R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j], 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\big\}$ is a partition of R satisfying $\|\mathcal{P}\| < \delta$, then by choosing a collection of points $\big\{(\xi_{ij}, \eta_{ij})\big\}_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m}$ such that $(\xi_{ij}, \eta_{ij}) \in R_{ij}$, we conclude that

$$\begin{aligned}
&\Big|\sum_{i=1}^{n}\sum_{j=1}^{m}\big(|T_{ij}^1| + |T_{ij}^2|\big) - \mathrm{I}\Big| \\
&\quad\leqslant \Big|\sum_{i=1}^{n}\sum_{j=1}^{m}\big(|T_{ij}^1| + |T_{ij}^2|\big) - \sum_{i=1}^{n}\sum_{j=1}^{m}\sqrt{1 + f_x(\xi_{ij}, \eta_{ij})^2 + f_y(\xi_{ij}, \eta_{ij})^2}(x_i - x_{i-1})(y_j - y_{j-1})\Big| \\
&\qquad + \Big|\sum_{i=1}^{n}\sum_{j=1}^{m}\sqrt{1 + f_x(\xi_{ij}, \eta_{ij})^2 + f_y(\xi_{ij}, \eta_{ij})^2}(x_i - x_{i-1})(y_j - y_{j-1}) - \mathrm{I}\Big| < \varepsilon\,.
\end{aligned}$$

This means that the approximation of the area of the surface S can be made arbitrarily closed to I; thus the area of the surface S must be I. In general, we have the following

Theorem 14.11

Let R be a closed region in the plane, and $f : R \to \mathbb{R}$ be a continuously differentiable function. Then the area of the surface $S = \big\{(x, y, z)\,\big|\,(x, y) \in R\,, z = f(x, y)\big\}$ is given by

$$\iint\limits_R \sqrt{1 + \|(\nabla f)(x, y)\|^2}\,dA = \iint\limits_R \sqrt{1 + f_x(x, y)^2 + f_y(x, y)^2}\,dA\,.$$

Example 14.12. Find the surface area of the sphere with radius r.

Let $f(x, y) = \sqrt{r^2 - x^2 - y^2}$ and $R = \big\{(x, y)\,\big|\,x^2 + y^2 \leqslant r^2\big\}$. Then the surface area of the sphere with radius r is given by

$$2\iint\limits_R \sqrt{1 + f_x(x, y)^2 + f_y(x, y)^2}\,dA = 2r\iint\limits_R \frac{1}{\sqrt{r^2 - x^2 - y^2}}\,dA\,.$$

Since R can also be expressed as $R = \big\{(x, y)\,\big|\,-r \leqslant x \leqslant r, -\sqrt{r^2 - x^2} \leqslant y \leqslant \sqrt{r^2 - x^2}\big\}$, the Fubini Theorem then implies that

$$\iint\limits_R \frac{1}{\sqrt{r^2 - x^2 - y^2}}\,dA = \int_{-r}^{r}\Big(\int_{-\sqrt{r^2 - x^2}}^{\sqrt{r^2 - x^2}} \frac{1}{\sqrt{r^2 - x^2 - y^2}}\,dy\Big)dx\,.$$

By Theorem 5.63, we find that for each $-r < x < r$,

$$\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\sqrt{r^2-x^2-y^2}}\,dy = \arcsin \frac{y}{\sqrt{r^2-x^2}}\Big|_{y=-\sqrt{r^2-x^2}}^{y=\sqrt{r^2-x^2}} = \arcsin 1 - \arcsin(-1) = \pi\,.$$

Therefore,

$$\int_{-r}^{r} \Big(\int_{-\sqrt{r^2-x^2}}^{\sqrt{r^2-x^2}} \frac{1}{\sqrt{r^2-x^2-y^2}}\,dy\Big)dx = \int_{-r}^{r} \pi\,dx = 2\pi r$$

which implies that the surface area of a sphere with radius r is $4\pi r^2$.

Example 14.13. In this example we consider the surface area of the upper hemi-sphere $z = \sqrt{r^2-x^2-y^2}$ that lies above the disk $R = \big\{(x,y)\,\big|\,x^2+y^2 \leqslant \sigma^2\big\}$, where $0 < \sigma < r$. Let $f(x,y) = \sqrt{r^2-x^2-y^2}$. Since R can also be expressed by

$$R = \big\{(x,y)\,\big|\, -r\sigma \leqslant x \leqslant \sigma, -\sqrt{\sigma^2-x^2} \leqslant y \leqslant \sqrt{\sigma^2-x^2}\big\}\,,$$

the Fubini Theorem implies that the surface area of interest is given by

$$\begin{aligned}
&\iint\limits_R \sqrt{1+f_x(x,y)^2+f_y(x,y)^2}\,dA \\
&\quad= \iint\limits_R \frac{r}{\sqrt{r^2-x^2-y^2}}\,dA = r\int_{-\sigma}^{\sigma}\Big(\int_{-\sqrt{\sigma^2-x^2}}^{\sqrt{\sigma^2-x^2}} \frac{1}{\sqrt{r^2-x^2-y^2}}\,dy\Big)dx\,.
\end{aligned}$$

By Theorem 5.63, we find that

$$\begin{aligned}
&\int_{-\sigma}^{\sigma}\Big(\int_{-\sqrt{\sigma^2-x^2}}^{\sqrt{\sigma^2-x^2}} \frac{1}{\sqrt{r^2-x^2-y^2}}\,dy\Big)dx = \int_{-\sigma}^{\sigma}\Big(\arcsin\frac{y}{\sqrt{r^2-x^2}}\Big|_{y=-\sqrt{\sigma^2-x^2}}^{y=\sqrt{\sigma^2-x^2}}\Big)dx \\
&\quad= 2\int_{-\sigma}^{\sigma} \arcsin\frac{\sqrt{\sigma^2-x^2}}{\sqrt{r^2-x^2}}\,dx = 2\int_{-\sigma}^{\sigma} \arctan\frac{\sqrt{\sigma^2-x^2}}{\sqrt{r^2-\sigma^2}}\,dx \\
&\quad= 2\Big[x\arctan\frac{\sqrt{\sigma^2-x^2}}{\sqrt{r^2-\sigma^2}}\Big|_{x=-\sigma}^{x=\sigma} - \int_{-\sigma}^{\sigma} x\frac{d}{dx}\arctan\frac{\sqrt{\sigma^2-x^2}}{\sqrt{r^2-\sigma^2}}\,dx\Big] \\
&\quad= -2\int_{-\sigma}^{\sigma}\frac{x\cdot\frac{1}{\sqrt{r^2-\sigma^2}}\frac{-x}{\sqrt{\sigma^2-x^2}}}{1+\frac{\sigma^2-x^2}{r^2-\sigma^2}}\,dx = 2\sqrt{r^2-\sigma^2}\int_{-\sigma}^{\sigma}\frac{x^2-r^2+r^2}{(r^2-x^2)\sqrt{\sigma^2-x^2}}\,dx \\
&\quad= -2\sqrt{r^2-\sigma^2}\pi + 2\sqrt{r^2-\sigma^2}\int_{-\sigma}^{\sigma}\frac{r^2}{(r^2-x^2)\sqrt{\sigma^2-x^2}}\,dx\,.
\end{aligned}$$

Using the substitution $x = \sigma\sin\dfrac{\theta}{2}$, we find that

$$\begin{aligned}
\int_{-\sigma}^{\sigma}\frac{r^2}{(r^2-x^2)\sqrt{\sigma^2-x^2}}\,dx &= \int_{-\pi}^{\pi}\frac{r^2}{2(r^2-\sigma^2\sin^2\frac{\theta}{2})}\,d\theta = \int_{-\pi}^{\pi}\frac{r^2}{2r^2-\sigma^2(1-\cos\theta)}\,d\theta \\
&= r^2\int_{-\pi}^{\pi}\frac{1}{(2r^2-\sigma^2)+\sigma^2\cos\theta}\,d\theta\,.
\end{aligned}$$

and further substitution $\tan\dfrac{\theta}{2} = t$ implies that

$$\begin{aligned}
\int_{-\sigma}^{\sigma} \frac{r^2}{(r^2-x^2)\sqrt{\sigma^2-x^2}}\,dx &= \int_{-\infty}^{\infty} \frac{r^2}{(2r^2-\sigma^2)+\sigma^2\frac{1-t^2}{1+t^2}}\,\frac{2dt}{1+t^2}\\
&= \int_{-\infty}^{\infty} \frac{2r^2}{2r^2(1+t^2)-\sigma^2(1+t^2)+\sigma^2(1-t^2)}\,dt\\
&= \int_{-\infty}^{\infty} \frac{r^2}{r^2+(r^2-\sigma^2)t^2}\,dt\\
&= \frac{r}{\sqrt{r^2-\sigma^2}}\arctan\Big(\frac{\sqrt{r^2-\sigma^2}}{r}t\Big)\Big|_{t=-\infty}^{\infty} = \frac{\pi r}{\sqrt{r^2-\sigma^2}}\,.
\end{aligned}$$

Therefore, the surface area of interest is given by

$$\iint\limits_R \frac{r}{\sqrt{r^2-x^2-y^2}}\,dA = 2r\sqrt{r^2-\sigma^2}\Big[-\pi+\frac{\pi r}{\sqrt{r^2-\sigma^2}}\Big] = 2\pi r\big(r-\sqrt{r^2-\sigma^2}\big)\,.$$

Example 14.14. Find the surface area of the paraboloid $z = 1+x^2+y^2$ that lies above the unit disk.

Let $f(x,y) = 1+x^2+y^2$ and $R = \big\{(x,y)\,\big|\,-1\leqslant x\leqslant 1, -\sqrt{1-x^2}\leqslant y\leqslant\sqrt{1-x^2}\big\}$, the Fubini Theorem implies that the surface area of interest is given by

$$\iint\limits_R \sqrt{1+f_x(x,y)^2+f_y(x,y)^2}\,dA = \int_{-1}^{1}\Big(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\sqrt{1+4x^2+4y^2}\,dy\Big)dx\,.$$

Using (8.3.1), we find that $\displaystyle\int\sqrt{a^2+b^2u^2}\,du = \frac{a^2}{2b}\Big[\frac{bu\sqrt{a^2+b^2u^2}}{a^2}+\ln\big(bu+\sqrt{a^2+b^2u^2}\big)\Big]+C$ if $a,b>0$; thus

$$\begin{aligned}
&\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\sqrt{1+4x^2+4y^2}\,dy = 2\int_0^{\sqrt{1-x^2}}\sqrt{1+4x^2+4y^2}\,dy\\
&\quad= \frac{1+4x^2}{2}\Big[\frac{2y\sqrt{1+4x^2+4y^2}}{1+4x^2}+\ln\big(2y+\sqrt{1+4x^2+4y^2}\big)\Big]\Big|_{y=0}^{y=\sqrt{1-x^2}}\\
&\quad= \sqrt{5}\sqrt{1-x^2}+\frac{1+4x^2}{2}\ln\frac{\sqrt{5}+2\sqrt{1-x^2}}{\sqrt{1+4x^2}}\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\iint\limits_R \sqrt{1+f_x(x,y)^2+f_y(x,y)^2}\,dA &= \int_{-1}^{1}\Big[\sqrt{5}\sqrt{1-x^2}+\frac{1+4x^2}{2}\ln\frac{\sqrt{5}+2\sqrt{1-x^2}}{\sqrt{1+4x^2}}\Big]\,dx\\
&= \frac{\sqrt{5}}{2}\pi+\frac{1}{2}\int_{-1}^{1}(1+4x^2)\ln\frac{\sqrt{5}+2\sqrt{1-x^2}}{\sqrt{1+4x^2}}\,dx\,.
\end{aligned}$$

Integrating by parts,

$$
\begin{aligned}
&\int_{-1}^{1}(1+4x^2)\ln\frac{\sqrt{5}+2\sqrt{1-x^2}}{\sqrt{1+4x^2}}\,dx\\
&\quad=\Big(x+\frac{4}{3}x^3\Big)\ln\frac{\sqrt{5}+2\sqrt{1-x^2}}{\sqrt{1+4x^2}}\bigg|_{x=-1}^{x=1}-\int_{-1}^{1}\Big(x+\frac{4}{3}x^3\Big)\frac{d}{dx}\ln\frac{\sqrt{5}+2\sqrt{1-x^2}}{\sqrt{1+4x^2}}\,dx\\
&\quad=-\int_{-1}^{1}\Big(x+\frac{4}{3}x^3\Big)\frac{\sqrt{1+4x^2}}{\sqrt{5}+2\sqrt{1-x^2}}\,\frac{\dfrac{-2x}{\sqrt{1-x^2}}\sqrt{1+4x^2}-\dfrac{4x}{\sqrt{1+4x^2}}(\sqrt{5}+2\sqrt{1-x^2})}{1+4x^2}\,dx\\
&\quad=-\int_{-1}^{1}\Big(x+\frac{4}{3}x^3\Big)\frac{-2x}{\sqrt{5}+2\sqrt{1-x^2}}\,\frac{5+2\sqrt{5}\sqrt{1-x^2}}{(1+4x^2)\sqrt{1-x^2}}\,dx\\
&\quad=\frac{\sqrt{5}}{3}\int_{-1}^{1}\frac{2x(3x+4x^3)}{(1+4x^2)\sqrt{1-x^2}}\,dx=\frac{\sqrt{5}}{3}\int_{-1}^{1}\frac{-1+3(1+4x^2)-2(1-x^2)(1+4x^2)}{(1+4x^2)\sqrt{1-x^2}}\,dx\\
&\quad=\frac{-\sqrt{5}}{3}\int_{-1}^{1}\frac{1}{(1+4x^2)\sqrt{1-x^2}}\,dx+\sqrt{5}\int_{-1}^{1}\frac{1}{\sqrt{1-x^2}}\,dx-\frac{2\sqrt{5}}{3}\int_{-1}^{1}\sqrt{1-x^2}\,dx\\
&\quad=\frac{-\sqrt{5}}{3}\int_{-1}^{1}\frac{1}{(1+4x^2)\sqrt{1-x^2}}\,dx+\frac{2\sqrt{5}}{3}\pi\,.
\end{aligned}
$$

By the substitution of variable $x=\sin\theta$, we find that

$$
\begin{aligned}
\int_{-1}^{1}\frac{1}{(1+4x^2)\sqrt{1-x^2}}\,dx&=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{1}{1+4\sin^2\theta}\,d\theta=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{1}{1+2(1-\cos 2\theta)}\,d\theta\\
&=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{1}{3-2\cos 2\theta}\,d\theta=\frac{1}{2}\int_{-\pi}^{\pi}\frac{1}{3-2\cos\phi}\,d\phi\,.
\end{aligned}
$$

By the substitution of variable $\tan\dfrac{\phi}{2}=t$, we further obtain that

$$
\begin{aligned}
\int_{-1}^{1}\frac{1}{(1+4x^2)\sqrt{1-x^2}}\,dx&=\frac{1}{2}\int_{-\infty}^{\infty}\frac{1}{3-2\frac{1-t^2}{1+t^2}}\,\frac{2dt}{1+t^2}=\int_{-\infty}^{\infty}\frac{1}{1+5t^2}\,dt\\
&=\frac{1}{\sqrt{5}}\arctan(\sqrt{5}t)\Big|_{t=-\infty}^{t=\infty}=\frac{\pi}{\sqrt{5}}\,.
\end{aligned}
$$

Therefore,

$$
\iint\limits_{R}\sqrt{1+f_x(x,y)^2+f_y(x,y)^2}\,dA=\frac{\sqrt{5}}{2}\pi+\frac{1}{2}\Big[-\frac{\sqrt{5}}{3}\cdot\frac{\pi}{\sqrt{5}}+\frac{2\sqrt{5}\pi}{3}\Big]=\frac{\pi}{6}(5\sqrt{5}-1)\,.
$$

14.3.2 Surface area of parametric surfaces

Definition 14.15: Parametric Surfaces

Let X, Y and Z be functions of u and v that are continuous on a domain D in the uv-plane. The collection of points

$$\Sigma \equiv \left\{ \boldsymbol{r} \in \mathbb{R}^3 \,\middle|\, \boldsymbol{r} = X(u,v)\mathbf{i} + Y(u,v)\mathbf{j} + Z(u,v)\mathbf{k} \text{ for some } (u,v) \in D \right\}$$

is called a parametric surface. The equations $x = X(u,v)$, $y = Y(u,v)$, and $z = Z(u,v)$ are the parametric equations for the surface, and $\boldsymbol{r} : D \to \mathbb{R}^3$ given by $\boldsymbol{r}(u,v) = X(u,v)\mathbf{i} + Y(u,v)\mathbf{j} + Z(u,v)\mathbf{k}$ is called a parametrization of Σ.

Example 14.16. Let R be an open region in the plane, and $f : R \to \mathbb{R}$ be a continuous function. Then the graph of f is a parametric surface. In fact,

$$\text{the graph of } f = \left\{ \boldsymbol{r} \in \mathbb{R}^3 \,\middle|\, \boldsymbol{r} = \big(x, y, f(x,y)\big) \text{ for some } (x,y) \in R \right\}.$$

Therefore, a parametric surface can be viewed as a generalization of surfaces being graphs of functions.

Example 14.17. Let $\mathbb{S}^2 = \left\{ (x,y,z) \in \mathbb{R}^3 \,\middle|\, x^2 + y^2 + z^2 = 1 \right\}$ be the unit sphere in $\mathbb{R}^3$. Consider

$$\boldsymbol{r}(\theta, \phi) = \big(\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi\big), \quad (\theta, \phi) \in D = [0, 2\pi] \times [0, \pi].$$

Then $\boldsymbol{r} : D \to \mathbb{S}^2$ is a continuous bijection; thus $\mathbb{S}^2$ is a parametric surface.

Example 14.18. Consider the torus shown below

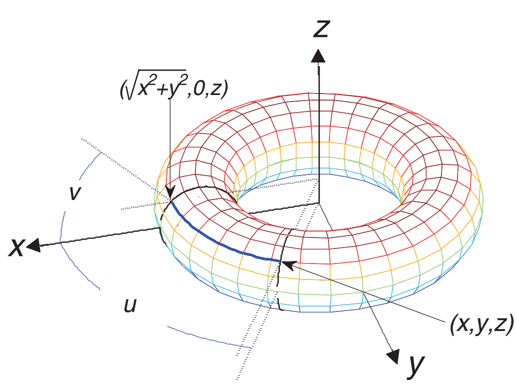


Figure 14.3: Torus with parametrization $\boldsymbol{r}(u,v)$. (temporary picture)

Note that the torus has a parametrization

$$\boldsymbol{r}(u,v) = \big((a+b\cos v)\cos u, (a+b\cos v)\sin u, b\sin v\big), \quad (u,v) \in [0,2\pi]\times[0,2\pi] .$$

Therefore, the torus is a parametric surface.

Remark 14.19. Similar to the case of curves, it is not required that the parametrization $\boldsymbol{r}$ is one-to-one; thus self-intersection of surface is allowed for defining parametric surface. However, we always assume that the "area" of the part of intersection is zero. This requirement is similar to the case that the parametrization of a curve that we discussed in Chapter 12 has non-overlapping property (see page 281).

Definition 14.20

A parametric surface

$$\Sigma \equiv \Big\{\boldsymbol{r} \in \mathbb{R}^3 \,\Big|\, \boldsymbol{r} = X(u,v)\mathbf{i} + Y(u,v)\mathbf{j} + Z(u,v)\mathbf{k} \text{ for some } (u,v) \in D\Big\} .$$

is said to be regular if X, Y, Z are differentiable funcitons and

$$\boldsymbol{r}_u(u,v) \times \boldsymbol{r}_v(u,v) \neq \mathbf{0} \qquad \forall\,(u,v) \in D\,,$$

where $\boldsymbol{r}_u \equiv X_u\mathbf{i} + Y_u\mathbf{j} + Z_u\mathbf{k}$ and $\boldsymbol{r}_v \equiv X_v\mathbf{i} + Y_v\mathbf{j} + Z_v\mathbf{k}$.

Remark 14.21. Let $\mathcal{V}$ be an open region in the plane. A vector-valued function $\psi : \mathcal{V} \to \mathbb{R}^3$ is differentiable if each component of ψ is differentiable, and the derivative of ψ, denoted by $D\psi$, is defined by

$$\big[D\psi(u,v)\big] = \begin{bmatrix} \dfrac{\partial \psi_1}{\partial u}(u,v) & \dfrac{\partial \psi_1}{\partial v}(u,v) \\ \dfrac{\partial \psi_2}{\partial u}(u,v) & \dfrac{\partial \psi_2}{\partial v}(u,v) \\ \dfrac{\partial \psi_3}{\partial u}(u,v) & \dfrac{\partial \psi_3}{\partial v}(u,v) \end{bmatrix} .$$

Therefore, a parametric surface

$$\Sigma \equiv \Big\{\boldsymbol{r} \in \mathbb{R}^3 \,\Big|\, \boldsymbol{r} = X(u,v)\mathbf{i} + Y(u,v)\mathbf{j} + Z(u,v)\mathbf{k} \text{ for some } (u,v) \in D\Big\} .$$

is regular if for each $(u,v) \in D$ the derivative $\big[D\psi(u,v)\big]$ has full rank.

Question: What does it mean by that a parametric surface is regular?

Suppose that

$$\Sigma \equiv \Big\{ \boldsymbol{r} \in \mathbb{R}^3 \,\Big|\, \boldsymbol{r} = X(u,v)\mathbf{i} + Y(u,v)\mathbf{j} + Z(u,v)\mathbf{k} \text{ for some } (u,v) \in D \Big\}\,.$$

is regular. Then at each point $p = \boldsymbol{r}(u_0, v_0)$, $\boldsymbol{r}_u(u_0, v_0)$ and $\boldsymbol{r}_v(u_0, v_0)$ are tangent vectors to Σ so that $\boldsymbol{r}_u(u_0, v_0) \times \boldsymbol{r}_v(u_0, v_0)$ is normal to the tangent plane of Σ at p. In other words, a parametric surface is regular if every point $p \in \Sigma$ has a tangent plane (denoted by $T_p\Sigma$).

Example 14.22. Let $\mathbb{S}^2$ be the unit sphere given in Example 14.17. Then

$$\begin{aligned}
\boldsymbol{r}_\theta(\theta, \phi) &= \big(-\sin\theta \sin\phi, \cos\theta \sin\phi, 0\big)\,,\\
\boldsymbol{r}_\phi(\theta, \phi) &= \big(\cos\theta \cos\phi, \sin\theta \cos\phi, -\sin\phi\big)
\end{aligned}$$

so that

$$\begin{aligned}
(\boldsymbol{r}_u \times \boldsymbol{r}_v)(\theta, \phi) &= \big(-\cos\theta \sin^2\phi, -\sin\theta \sin^2\phi, -\sin\phi \cos\phi\big)\\
&= -\sin\phi\big(\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi\big)
\end{aligned}$$

which is non-zero if $\phi \neq 0$ and π. Therefore, $\mathbb{S}^2\backslash\{$the north and the south poles$\}$ is a regular parametric surface (with the same parametrization except that the domain becomes $[0, 2\pi] \times (0, \pi)$).

Example 14.23. Let the torus be given in Example 14.18. Then

$$\begin{aligned}
\boldsymbol{r}_u(u, v) &= \big(-(a + b\cos v)\sin u, (a + b\cos v)\cos u, 0\big)\,,\\
\boldsymbol{r}_v(u, v) &= \big(-b\sin v \cos u, -b\sin v \sin u, b\cos v\big)
\end{aligned}$$

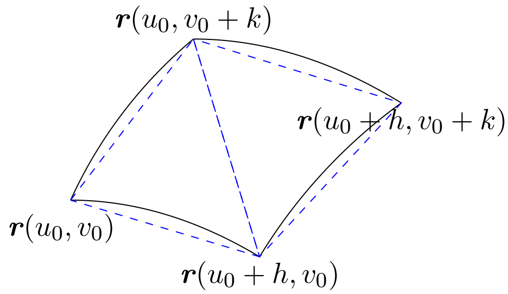
so that

$$\begin{aligned}
(\boldsymbol{r}_u \times \boldsymbol{r}_v)(u, v) &= \big(b(a + b\cos v)\cos u \cos v, b(a + b\cos v)\cos v \sin u, b(a + b\cos v)\sin v\big)\\
&= b(a + b\cos v)\big(\cos u \cos v, \sin u \cos v, \sin v\big)\,.
\end{aligned}$$

Since $\boldsymbol{r}_u \times \boldsymbol{r}_v \neq \mathbf{0}$, we find that the torus is a regular parametric surface.

Question: How to compute the surface area of a regular parametric surface?

Let $p = \boldsymbol{r}(u_0, v_0)$ be a point in Σ, and we consider the surface area of the region $\boldsymbol{r}\big([u_0, u_0 + h] \times [v_0, v_0 + k]\big)$, where h, k are very small. This area can be approximated by the sum of the area of two triangles, one with vertices $\boldsymbol{r}(u_0, v_0)$, $\boldsymbol{r}(u_0 + h, v_0)$, $\boldsymbol{r}(u_0, v_0 + k)$ and the other with vertices $\boldsymbol{r}(u_0 + h, v_0)$, $\boldsymbol{r}(u_0, v_0 + k)$, $\boldsymbol{r}(u_0 + h, v_0 + k)$.



The area of the triangle with vertices $\boldsymbol{r}(u_0, v_0)$, $\boldsymbol{r}(u_0+h, v_0)$, $\boldsymbol{r}(u_0, v_0+k)$ is

$$A_1 = \frac{1}{2}\big\|\big(\boldsymbol{r}(u_0+h, v_0) - \boldsymbol{r}(u_0, v_0)\big) \times \big(\boldsymbol{r}(u_0, v_0+k) - \boldsymbol{r}(u_0, v_0)\big)\big\|_{\mathbb{R}^3} .$$

By the mean value theorem,

$$\begin{aligned}
&\boldsymbol{r}(u_0+h, v_0) - \boldsymbol{r}(u_0, v_0) \\
&\qquad = \big[X(u_0+h, v_0) - X(u_0, v_0)\big]\mathbf{i} + \big[Y(u_0+h, v_0) - Y(u_0, v_0)\big]\mathbf{j} \\
&\qquad\quad + \big[Z(u_0+h, v_0) - Z(u_0, v_0)\big]\mathbf{k} \\
&\qquad = h\big[X_u(u_0+\theta_1 h, v_0)\mathbf{i} + Y_u(u_0+\theta_2 h, v_0)\mathbf{j} + Z_u(u_0+\theta_3 h, v_0)\mathbf{k}\big]
\end{aligned}$$

for some $\theta_1, \theta_2, \theta_3 \in (0,1)$. Suppose that $\boldsymbol{r}$ is continuously differentiable; that is, X, Y, Z are continuously differentiable, then

$$\begin{aligned}
X_u(u_0+\theta_1 h, v_0) &= X_u(u_0, v_0) + E_1(u_0, v_0, h)\,, \\
Y_u(u_0+\theta_2 h, v_0) &= Y_u(u_0, v_0) + E_2(u_0, v_0, h)\,, \\
Z_u(u_0+\theta_3 h, v_0) &= Z_u(u_0, v_0) + E_3(u_0, v_0, h)\,,
\end{aligned}$$

where E_1, E_2, E_2 approach zero as $h \to 0$. Therefore,

$$\boldsymbol{r}(u_0+h, v_0) - \boldsymbol{r}(u_0, v_0) = h\big[\boldsymbol{r}_u(u_0, v_0) + \boldsymbol{E}_1(u_0, v_0, h)\big]\,,$$

where $\boldsymbol{E}_1 = E_1\mathbf{i} + E_2\mathbf{j} + E_3\mathbf{k}$ satisfying that $\lim\limits_{h\to 0} \boldsymbol{E}_1(u_0, v_0; h) = 0$. Similarly,

$$\boldsymbol{r}(u_0, v_0+k) - \boldsymbol{r}(u_0, v_0) = k\big[\boldsymbol{r}_u(u_0, v_0) + \boldsymbol{E}_2(u_0, v_0, h)\big]\,,$$

where $\lim\limits_{h\to 0} \boldsymbol{E}_1(u_0, v_0; h) = 0$. The discussion above shows that

$$\lim_{(h,k)\to(0,0)} \frac{\big(\boldsymbol{r}(u_0+h, v_0) - \boldsymbol{r}(u_0, v_0)\big) \times \big(\boldsymbol{r}(u_0, v_0+k) - \boldsymbol{r}(u_0, v_0)\big)}{hk} - \boldsymbol{r}_u(u_0, v_0) \times \boldsymbol{r}_v(u_0, v_0) = \mathbf{0}$$

which further implies that

$$A_1 = \frac{1}{2}\big\|\boldsymbol{r}_u(u_0,v_0)\times\boldsymbol{r}_v(u_0,v_0)\big\|hk + \mathcal{E}_1(u_0,v_0,h,k)hk$$

for some function $\mathcal{E}_1$ which is bounded and converges to 0 as $(h,k)\to(0,0)$. Similarly, the area of the triangle with vertices $\boldsymbol{r}(u_0+h,v_0)$, $\boldsymbol{r}(u_0,v_0+k)$, $\boldsymbol{r}(u_0+h,v_0+k)$ is

$$A_2 = \frac{1}{2}\big\|\boldsymbol{r}_u(u_0,v_0)\times\boldsymbol{r}_v(u_0,v_0)\big\|hk + \mathcal{E}_2(u_0,v_0,h,k)hk$$

for some function $\mathcal{E}_2$ which is bounded and converges to 0 as $(h,k)\to(0,0)$. The two formulas for A_1 and A_2 shows that

$$\begin{aligned}&\text{the surface area of } \boldsymbol{r}\big([u_0,u_0+h]\times[v_0,v_0+k]\big)\\ &\qquad = \big\|\boldsymbol{r}_u(u_0,v_0)\times\boldsymbol{r}_v(u_0,v_0)\big\|hk + \mathcal{E}(u_0,v_0,h,k)hk\end{aligned} \tag{14.3.2}$$

for some bounded function $\mathcal{E}$ which converges to 0 as the last two variables h,k approach 0.

Now consider the surface area of $\boldsymbol{r}([a,a+L]\times[b,b+W])$. Let $\varepsilon>0$ be given. Choose $N>0$ such that

$$\big|\mathcal{E}(u,v;h,k)\big| < \frac{\varepsilon}{2LW} \quad \forall\, 0<h<\frac{L}{N}, 0<k<\frac{W}{N} \text{ and } (u,v)\in[a,a+L]\times[b,b+W]\,.$$

Denote $\|\boldsymbol{r}_u\times\boldsymbol{r}_v\|$ by $\sqrt{\mathrm{g}}$. Then

$$\Big|\sum_{j=1}^{m}\sum_{i=1}^{n}\sqrt{\mathrm{g}\big(a+\frac{i-1}{n}L, b+\frac{j-1}{m}M\big)}\frac{L}{n}\frac{W}{m} - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big| < \frac{\varepsilon}{2} \quad \text{if } n,m\geqslant N\,.$$

Then for $n,m\geqslant N$, with (h,k) denoting $\big(\frac{L}{n},\frac{W}{m}\big)$ (14.3.2) implies that

$$\begin{aligned}&\Big|\text{the surface area of } \boldsymbol{r}([a,a+L]\times[b,b+W]) - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big|\\ &\quad = \Big|\sum_{j=1}^{m}\sum_{i=1}^{n}\text{the surface area of } \boldsymbol{r}([a+(i-1)h,a+ih]\times[b+(j-1)k,b+jk])\\ &\qquad - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big|\\ &\quad \leqslant \Big|\sum_{j=1}^{m}\sum_{i=1}^{n}\sqrt{\mathrm{g}\big(a+(i-1)h,b+(j-1)k\big)}hk - \int_{[a,a+L]\times[b,b+W]}\sqrt{\mathrm{g}}\,d\mathbb{A}\Big|\\ &\qquad + \Big|\sum_{j=1}^{m}\sum_{i=1}^{n}f(a+(i-1)h,b+(j-1)k;h,k)hk\Big|\\ &\quad < \frac{\varepsilon}{2}+\frac{\varepsilon}{2LW}\sum_{j=1}^{m}\sum_{i=1}^{n}hk = \varepsilon\,.\end{aligned}$$

The discussion above verifies the following

Theorem 14.24

Let D be an open region in the plane, and

$$\Sigma \equiv \left\{ \boldsymbol{r} \in \mathbb{R}^3 \,\middle|\, \boldsymbol{r} = X(u,v)\mathbf{i} + Y(u,v)\mathbf{j} + Z(u,v)\mathbf{k} \ \text{ for some } (u,v) \in D \right\}.$$

be a regular parametric surface so that r is continuously differentiable; that is, $X_u, X_v, Y_u, Y_v, Z_u, Z_v$ are continuous. Then

$$\text{the surface area of } \Sigma = \iint_D \|\boldsymbol{r}_u(u,v) \times \boldsymbol{r}_v(u,v)\| \, d(u,v)\,.$$

Example 14.25. Let R be an open region in the plane, and $f : R \to \mathbb{R}$ is continuously differentiable. Then Theorem 14.24 implies that the surface area of the graph of f is given by

$$\iint_R \|(\boldsymbol{r}_x \times \boldsymbol{r}_y)(x,y)\| \, d\mathbb{A}\,,$$

where the parametrization $\boldsymbol{r}$ is given by $\boldsymbol{r}(x,y) = \big(x, y, f(x,y)\big)$, $(x,y) \in R$. This formula agrees with what Theorem 14.11 provides.

Example 14.26. With the parametrization of the unit sphere $\mathbb{S}^2$ given in Example 14.22, by Theorem 14.24 the surface area of $\mathbb{S}^2$ is given by

$$\iint_{[0,2\pi]\times[0,\pi]} \|(\boldsymbol{r}_\theta \times \boldsymbol{r}_\phi)(\theta,\phi)\| d(\theta,\phi) = \int_0^\pi \Big(\int_0^{2\pi} \sin\phi \, d\theta \Big) d\phi = 4\pi\,.$$

Example 14.27. With the parametrization of the torus given in Example 14.23, by Theorem 14.24 the surface area of the torus is given by

$$\iint_{[0,2\pi]\times[0,2\pi]} b(a + b\cos v)\, d(u,v) = \int_0^{2\pi} \Big(\int_0^{2\pi} (ab + b^2 \cos \boldsymbol{v})\, du \Big) dv = 4\pi^2 ab\,.$$

14.4 Triple Integrals and Applications

Let Q be a bounded region in space, and $f : Q \to \mathbb{R}$ be a non-negative function which described the point density of the region. We are interested in the mass of Q.

We start with the simple case that $Q = [a, b] \times [c, d] \times [r, s]$ is a cube. Let

$$\begin{aligned}\mathcal{P}_x &= \{a = x_0 < x_1 < \cdots < x_n = b\}\,,\\ \mathcal{P}_y &= \{c = y_0 < y_1 < \cdots < y_m = d\}\,,\\ \mathcal{P}_z &= \{r = z_0 < z_1 < \cdots < z_p = s\}\,,\end{aligned}$$

be partitions of $[a, b]$, $[c, d]$, $[r, s]$, respectively, and $\mathcal{P}$ be a collection of non-overlapping cubes given by

$$\mathcal{P} = \big\{R_{ijk} \,\big|\, R_{ijk} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k], 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m, 1 \leqslant k \leqslant p\big\}\,.$$

Such a collection $\mathcal{P}$ is called a partition of Q, and the norm of $\mathcal{P}$ is the maximum of the length of the diagonals of all R_{ijk}; that is

$$\|\mathcal{P}\| = \max\Big\{\sqrt{(x_i - x_{i-1})^2 + (y_j - y_{j-1})^2 + (z_k - z_{k-1})^2}\,\Big|\, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m, 1 \leqslant k \leqslant p\Big\}\,.$$

A Riemann sum of f for this partition $\mathcal{P}$ is given by

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{p} f(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})\,.$$

The mass of Q then should be the "limit" of Riemann sums as $\|\mathcal{P}\|$ approaches zero. In general, we can remove the restrictions that f is non-negative on R and still consider the limit of the Riemann sums. We have the following

Theorem 14.28

Let $Q = [a, b] \times [c, d] \times [r, s]$ be a cube in space, and $f : Q \to \mathbb{R}$ be a function. f is said to be Riemann integrable on Q if there exists a real number I such that for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of Q satisfying $\|\mathcal{P}\| < \delta$, then any Riemann sum of f for $\mathcal{P}$ belongs to $(I - \varepsilon, I + \epsilon)$. Such a number I (is unique if it exists and) is called the ***Riemann integral*** or ***triple integral of*** f ***on*** Q and is denoted by $\displaystyle\iiint_Q f(x, y, z)\, dV$.

For general bounded region Q in space, let $r > 0$ be such that $Q \subseteq [-r, r]^3$, and we define $\displaystyle\iiint_Q f(x, y, z)\, dV$ as $\displaystyle\iiint_{[-r,r]^3} \widetilde{f}(x, y, z)\, dV$, where $\widetilde{f}$ is the zero extension of f given by

$$\widetilde{f}(x, y, z) = \begin{cases} f(x, y, z) & \text{if } (x, y, z) \in R\,,\\ 0 & \text{if } (x, y, z) \notin R\,.\end{cases}$$

Some of the properties of double integrals in Theorem 14.4 can be restated in terms of triple integrals.

1. $\displaystyle\iiint\limits_{Q} (cf)(x,y,z)\,dV = c\iiint\limits_{Q} f(x,y,z)\,dV$.

2. $\displaystyle\iiint\limits_{Q} (f+g)(x,y,z)\,dV = \iiint\limits_{Q} f(x,y,z)\,dV + \iiint\limits_{Q} g(x,y,z)\,dV$.

3. $\displaystyle\iiint\limits_{Q_1\cup Q_2} f(x,y,z)\,dV = \iiint\limits_{Q_1} f(x,y,z)\,dV + \iiint\limits_{Q_2} f(x,y,z)\,dV$ for all "non-overlapping" solid regions Q_1 and Q_2.

Similar to Fubini's Theorem for the evaluation of double integrals, we have the following

Theorem 14.29: Fubini's Theorem

Let Q be a region in space, and $f : Q \to \mathbb{R}$ be continuous. If Q is given by $Q = \big\{(x,y,z)\,\big|\,(x,y) \in R, g_1(x,y) \leqslant z \leqslant g_2(x,y)\big\}$ for some region R in the xy-plane, then

$$\iiint\limits_{Q} f(x,y,z)\,dV = \iint\limits_{R} \Big(\int_{g_1(x,y)}^{g_2(x,y)} f(x,y,z)\,dz\Big)\,dA\,.$$

In particular, if R is expressed by $R = \big\{(x,y)\,\big|\,a \leqslant x \leqslant b, h_1(x) \leqslant y \leqslant h_2(y)\big\}$, then

$$\iiint\limits_{Q} f(x,y,z)\,dV = \int_a^b \Big[\int_{h_1(x)}^{h_2(x)} \Big(\int_{g_1(x,y)}^{g_2(x,y)} f(x,y,z)\,dz\Big)dy\Big]\,dx\,.$$

Example 14.30. Find the volume of the region Q bounded below by the paraboloid $z = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 6$.

Suppose Q is a solid region in space with uniform density 1 (or say, this region is occupied by water). Then the volume of Q is identical to the mass (in terms of its numerical value); thus we find that the volume of Q is given by $\displaystyle\iiint\limits_{Q} 1\,dV$. To apply the Fubini Theorem, we need to express Q as $\big\{(x,y,z)\,\big|\,(x,y) \in R, g_1(x,y) \leqslant z \leqslant g_2(x,y)\big\}$. Nevertheless, if R is the bounded region in the plane enclosed by the curve $(x^2+y^2)^2 + x^2 + y^2 = 6$ (which in fact gives $x^2 + y^2 = 2$), then

$$Q = \big\{(x,y,z)\,\big|\,(x,y) \in R, x^2 + y^2 \leqslant z \leqslant \sqrt{6 - x^2 - y^2}\big\}$$

and the Fubini Theorem implies that

$$\text{the volume of } Q = \int_R \Big(\int_{x^2+y^2}^{\sqrt{6-x^2-y^2}} 1\, dz \Big)\, dA\,.$$

Solving for R, we find that $R = \big\{(x,y)\,\big|\, -\sqrt{2} \leqslant x \leqslant \sqrt{2}, -\sqrt{2-x^2} \leqslant y \leqslant \sqrt{2-x^2}\big\}$; thus by the Fubini Theorem we find that

$$\text{the volume of } Q = \int_{-\sqrt{2}}^{\sqrt{2}} \Big[\int_{-\sqrt{2-x^2}}^{\sqrt{2-x^2}} \Big(\int_{x^2+y^2}^{\sqrt{6-x^2-y^2}} 1\, dz \Big) dy \Big] dx\,.$$

Example 14.31. Evaluate $\displaystyle\int_0^{\sqrt{\pi/2}} \Big[\int_x^{\sqrt{\pi/2}} \Big(\int_1^3 \sin(y^2)\, dz \Big) dy \Big] dx$.

Let $R = \big\{(x,y)\,\big|\, 0 \leqslant x \leqslant \sqrt{\pi/2}, x \leqslant y \leqslant \sqrt{\pi/2}\big\}$, then the domain of integration is given by

$$Q = \big\{(x,y,z)\,\big|\, 0 \leqslant x \leqslant \sqrt{\pi/2}, x \leqslant y \leqslant \sqrt{\pi/2}, 1 \leqslant z \leqslant 3\big\}$$

and the iterated integral given above is the triple integral $\displaystyle\iiint\limits_Q \sin(y^2)\, dV$.

Since R can also be expressed as $R = \big\{(x,y)\,\big|\, 0 \leqslant y \leqslant \sqrt{\pi/2}, 0 \leqslant x \leqslant y\big\}$, by the Fubini Theorem we find that

$$\begin{aligned}
&\int_0^{\sqrt{\pi/2}} \Big[\int_x^{\sqrt{\pi/2}} \Big(\int_1^3 \sin(y^2)\, dz \Big) dy \Big] dx = \iiint\limits_Q \sin(y^2)\, dV \\
&\quad = \int_0^{\sqrt{\pi/2}} \Big[\int_0^y \Big(\int_1^3 \sin(y^2)\, dz \Big) dx \Big] dy = \int_0^{\sqrt{\pi/2}} 2y\sin(y^2)\, dy = -\cos(y^2)\Big|_{y=0}^{y=\sqrt{\pi/2}} = 1\,.
\end{aligned}$$

Example 14.32. Compute the iterated integrals

$$\int_0^6 \Big[\int_{\frac{z}{2}}^3 \Big(\int_{\frac{z}{2}}^y dx \Big) dy \Big] dz + \int_0^6 \Big[\int_3^{\frac{12-z}{2}} \Big(\int_{\frac{z}{2}}^{6-y} dx \Big) dy \Big] dz\,,$$

then write the sum above as a single iterated integral in the order $dydzdx$ and $dzdydx$.

We compute the two integrals above as follows:

$$\begin{aligned}
\int_0^6 \Big[\int_{\frac{z}{2}}^3 \Big(\int_{\frac{z}{2}}^y dx \Big) dy \Big] dz &= \int_0^6 \Big[\int_{\frac{z}{2}}^3 \Big(y - \frac{z}{2}\Big)\, dy \Big] dz = \int_0^6 \Big(\frac{y^2 - yz}{2}\Big|_{y=\frac{z}{2}}^{y=3} \Big)\, dz \\
&= \frac{1}{2}\int_0^6 \Big(9 - 3z + \frac{z^2}{4}\Big)\, dz = \frac{1}{2}\Big(9z - \frac{3z^2}{2} + \frac{z^3}{12}\Big)\Big|_{z=0}^{z=6} = 9\,,
\end{aligned}$$

and

$$\int_0^6 \Big[\int_3^{\frac{12-z}{2}} \Big(\int_{\frac{z}{2}}^{6-y} dx\Big) dy\Big] dz = \int_0^6 \Big[\int_3^{\frac{12-z}{2}} \big(6 - y - \frac{z}{2}\big)\, dy\Big] dz$$

$$= \frac{1}{2}\int_0^6 \big(12y - y^2 - yz\big)\Big|_{y=3}^{y=\frac{12-z}{2}} dz$$

$$= \frac{1}{2}\int_0^6 \Big[6(12-z) - \frac{144 - 24z + z^2}{4} - \frac{(12-z)z}{2} - 36 + 9 + 3z\Big)\, dz$$

$$= \frac{1}{2}\int_0^6 \Big(72 - 6z - 36 + 6z - \frac{z^2}{4} - 6z + \frac{z^2}{2} - 27 + 3z\Big)\, dz$$

$$= \frac{1}{2}\int_0^6 \Big(9 - 3z + \frac{z^2}{4}\Big)\, dz = \frac{1}{2}\Big(9z - \frac{3z^2}{2} + \frac{z^3}{12}\Big)\Big|_{z=0}^{z=6} = 9\,.$$

Therefore, the sum of the two integrals is 18.

Let

$$Q_1 = \Big\{(x, y, z)\,\Big|\, 0 \leqslant z \leqslant 6, \frac{z}{2} \leqslant y \leqslant 3, \frac{z}{2} \leqslant x \leqslant y\Big\}\,,$$

$$Q_2 = \Big\{(x, y, z)\,\Big|\, 0 \leqslant z \leqslant 6, 3 \leqslant y \leqslant \frac{12-z}{2}, \frac{z}{2} \leqslant x \leqslant 6 - y\Big\}\,.$$

Then the Fubini Theorem implies that

$$\int_0^6 \Big[\int_{\frac{z}{2}}^3 \Big(\int_{\frac{z}{2}}^{y} dx\Big) dy\Big] dz = \iiint\limits_{Q_1} dV\,, \qquad \int_0^6 \Big[\int_3^{\frac{12-z}{2}} \Big(\int_{\frac{z}{2}}^{6-y} dx\Big) dy\Big] dz = \iiint\limits_{Q_2} dV\,.$$

Let $Q = Q_1 \cup Q_2$. Since Q_1 and Q_2 are non-overlapping solid regions (their intersection is a subset of the plane $y = 3$). Then

$$\iiint\limits_{Q_1} dV + \iiint\limits_{Q_2} dV = \iiint\limits_{Q} dV\,.$$

1. Let R be the projection of Q onto the xz-plane. Then $R = \big\{(x, z)\,\big|\, 0 \leqslant x \leqslant 3, 0 \leqslant z \leqslant 2x\big\}$ (where $z = 2x$ is the projection of the plane $x = \dfrac{z}{2}$ onto the xz-plane), and Q can also be expressed as

   $$Q = \big\{(x, y, z)\,\big|\,(x, z) \in R, x \leqslant y \leqslant 6 - x\big\}\,.$$

   Therefore, the volume of Q is given by

   $$\int_0^3 \Big[\int_0^{2x} \Big(\int_x^{6-x} dy\Big) dz\Big] dx = \int_0^3 \Big[\int_0^{2x} (6 - 2x)\, dz\Big] dx$$

   $$= \int_0^3 2x(6 - 2x)\, dx = \Big(6x^2 - \frac{4x^3}{3}\Big)\Big|_{x=0}^{x=3} = 54 - 36 = 18\,.$$

2. Let S be the projection of Q onto the xy-plane. Then $S = \big\{(x, y) \,\big|\, 0 \leqslant x \leqslant 3, x \leqslant y \leqslant 6 - x\big\}$, and Q can also be expressed as

$$Q = \big\{(x, y, z) \,\big|\, (x, y) \in S, 0 \leqslant z \leqslant 2x\big\}.$$

Therefore, the volume of Q is given by

$$\int_0^3 \Big[\int_x^{6-x} \Big(\int_0^{2x} dz\Big) dy\Big] dx = \int_0^3 \Big[\int_x^{6-x} 2x\, dy\Big] dx = \int_0^3 2x(6-2x)\, dx = 18\,.$$

14.5 Change of Variables Formula

In this section, we consider the version of substitution of variables in multiple integrals. We have used the technique of substitution of variable to evaluate the iterated integrals in, for example, Example 14.13 and 14.14; however, these substitutions of variable always assume that other variables are independent of the new variable introduced by the substitution of variable. We would like to investigate the effect of making a change of variables such as $x = r\cos\theta$, $y = r\sin\theta$ in computing the double integrals.

14.5.1 Double integrals in polar coordinates

We start our discussion with double integrals in polar coordinates. Suppose that R is the shaded region shown in Figure 14.4 and $f : R \to \mathbb{R}$ is continuous.

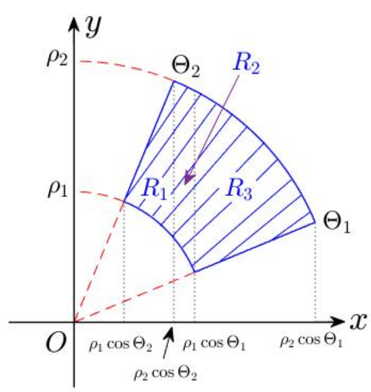


Figure 14.3: Rectangle in polar coordinates

Then to compute the double integral $\displaystyle\iint\limits_R f(x, y)\, dA$ using the Fubini theorem directly, we need to divide R into three sub-regions R_1, R_2, R_3 given by

$$
\begin{aligned}
R_1 &= \left\{(x,y)\,\middle|\,\rho_1\cos\Theta_2 \leqslant x \leqslant \rho_2\cos\Theta_2, \sqrt{\rho_1^2 - x^2} \leqslant y \leqslant x\tan\Theta_2\right\},\\
R_2 &= \left\{(x,y)\,\middle|\,\rho_2\cos\Theta_2 \leqslant x \leqslant \rho_1\cos\Theta_1, \sqrt{\rho_2^2 - x^2} \leqslant y \leqslant \sqrt{\rho_1^2 - x^2}\right\},\\
R_3 &= \left\{(x,y)\,\middle|\,\rho_1\cos\Theta_1 \leqslant x \leqslant \rho_2\Theta_2, x\tan\Theta_1 \leqslant y \leqslant \sqrt{\rho_2^2 - x^2}\right\},
\end{aligned}
$$

and write

$$
\iint\limits_R f(x,y)\,dA = \iint\limits_{R_1} f(x,y)\,dA + \iint\limits_{R_2} f(x,y)\,dA + \iint\limits_{R_3} f(x,y)\,dA\,.
$$

However, we know that the region R above is a rectangle in $r\theta$-plane, where (r,θ) is the polar coordinates on the plane. To be more precise, in polar coordinate the region R can be expressed as $R' \equiv \{(r,\theta)\,|\,\rho_1 \leqslant r \leqslant \rho_2, \Theta_1 \leqslant \theta \leqslant \Theta_2\}$, which means that every point (x,y) in R can be written as $(r\cos\theta, r\sin\theta)$ for $(r,\theta) \in R'$, and vice versa. One should expect that it should be easier to write down the iterated integral for computing $\displaystyle\iint\limits_R f(x,y)\,dA$.

Let $\mathcal{P}_r = \{\rho_1 = r_0 < r_1 < \cdots < r_n = \rho_2\}$ and $\mathcal{P}_\theta = \{\Theta_1 = \theta_0 < \theta_1 < \cdots < \theta_m = \Theta_2\}$ be partitions of $[\rho_1,\rho_2]$ and $[\Theta_1,\Theta_2]$, respectively, $R_{ij} = [r_{i-1}, r_i] \times [\theta_{j-1}, \theta_j]$ be rectangles in the $r\theta$-plane, S_{ij} be the sub-region in the xy-plane corresponds to R_{ij} under the polar coordinate; that is,

$$
S_{ij} = \{(r\cos\theta, r\sin\theta)\,|\,r \in [r_{i-1}, r_i], \theta \in [\theta_{j-1}, \theta_j]\}\,.
$$

The collection $\mathcal{P} = \{S_{ij}\,|\,1 \leqslant i \leqslant n, 1 \leqslant j \leqslant m\}$ is called a partition of rectangles in polar coordinates, and the norm of $\mathcal{P}$, denoted by $\|\mathcal{P}\|$, is the maximum diameter of S_{ij}.

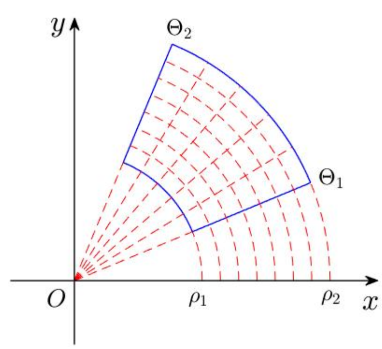


Figure 14.4: Rectangle in polar coordinates

A Riemann sum of f for partition $\mathcal{P}$ is of the form $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} f(\xi_{ij},\eta_{ij})|S_{ij}|$, where $|S_{ij}|$ is the area of S_{ij} and $\big\{(\xi_{ij},\eta_{ij})\big\}_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$ be collection of points satisfying $(\xi_{ij},\eta_{ij}) \in S_{ij}$. Then intuitively $\iint\limits_{R} f(x,y)\,dA$ is the limit of Riemann sums of f for $\mathcal{P}$ as $\|\mathcal{P}\|$ approaches zero.

To see the limit of Riemann sums, we choose a particular partition $\mathcal{P}$ and collection $\big\{(\xi_{ij},\eta_{ij})\big\}_{1\leqslant i\leqslant n, 1\leqslant j\leqslant m}$. We equally partition $[\rho_1,\rho_2]$ and $[\Theta_1,\Theta_2]$ into n and m sub-intervals. Let $\Delta r = \dfrac{\rho_2-\rho_1}{n}$ and $\Delta\theta = \dfrac{\Theta_2-\Theta_1}{m}$, and $r_i = \rho_1 + i\Delta r$ and $\theta_j = \Theta_1 + j\Delta\theta$, and $\xi_{ij} = r_i\cos\theta_j$ and $\eta_{ij} = r_i\sin\theta_j$. Noting that

$$|S_{ij}| = \frac{1}{2}(r_i^2 - r_{i-1}^2)(\theta_j - \theta_{j-1}) = \frac{1}{2}(r_i + r_{i-1})\Delta r\Delta\theta = r_i\Delta r\Delta\theta - \frac{1}{2}\Delta r^2\Delta\theta\,,$$

we find that

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{m} f(\xi_{ij},\eta_{ij})|S_{ij}| &= \sum_{i=1}^{n}\sum_{j=1}^{m} f(r_i\cos\theta_j, r_i\sin\theta_j)r_i\Delta r\Delta\theta \\ &\quad - \frac{\Delta r}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} f(r_i\cos\theta_j, r_i\sin\theta_j)\Delta r\Delta\theta\,.\end{aligned}$$

Let $g(r,\theta) = rf(r\cos\theta, r\sin\theta)$ and $h(r,\theta) = f(r\cos\theta, r\sin\theta)$, then

$$\sum_{i=1}^{n}\sum_{j=1}^{m} f(\xi_{ij},\eta_{ij})|S_{ij}| = \sum_{i=1}^{n}\sum_{j=1}^{m} g(r_i,\theta_j)\Delta r\Delta\theta - \frac{\Delta r}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} h(r_i,\theta_j)\Delta r\Delta\theta\,.$$

As n,m approach ∞, we find that

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{m} g(r_i,\theta_j)\Delta r\Delta\theta &\to \iint\limits_{R'} g(r,\theta)\,d(r,\theta) = \iint\limits_{R'} f(r\cos\theta, r\sin\theta)r\,d(r,\theta)\,,\\ \sum_{i=1}^{n}\sum_{j=1}^{m} h(r_i,\theta_j)\Delta r\Delta\theta &\to \iint\limits_{R'} h(r,\theta)\,d(r,\theta) = \iint\limits_{R'} f(r\cos\theta, r\sin\theta)\,d(r,\theta)\,,\end{aligned}$$

where the right-hand side integrals denotes the double integrals on the rectangle R'. Therefore, the limit of Riemann sums of f for $\mathcal{P}$ as $\|\mathcal{P}\|$ approaches zero is

$$\iint\limits_{R'} f(r\cos\theta, r\sin\theta)r\,d(r,\theta)\,;$$

thus

$$\iint\limits_{R} f(x,y)\,d(x,y) = \iint\limits_{R'} f(r\cos\theta, r\sin\theta)r\,d(r,\theta)\,. \tag{14.5.1}$$

14.5.2 Jacobian

Recall the substitution of variables formula for the integral of functions of one variable:

$$\int_a^b f\big(g(x)\big)g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du\,.$$

Suppose that $g:[a,b]\to\mathbb{R}$ is one-to-one. If g is increasing, then $g'\geqslant 0$ and $g([a,b]) = \big[g(a),g(b)\big]$; thus the formula above can be rewritten as

$$\int_{g([a,b])} f(u)\,du = \int_{[a,b]} f(g(x))g'(x)\,dx = \int_{[a,b]} f(g(x))\big|g'(x)\big|\,dx\,.$$

If g is decreasing, then $g'\leqslant 0$ and $g([a,b]) = \big[g(b),g(a)\big]$; thus the formula above can be written as

$$\int_{g([a,b])} f(u)\,du = -\int_{[a,b]} f(g(x))g'(x)\,dx = \int_{[a,b]} f(g(x))\big|g'(x)\big|\,dx\,.$$

Therefore, in either cases we have a rewritten version of the substitution of variable formula

$$\int_{g([a,b])} f(u)\,du = \int_{[a,b]} f(g(x))\big|g'(x)\big|\,dx\,.$$

In this section, we are concerned with the substitution of variable formula (usually called the change of variables formula in the case of multiple integrals) for double and triple integrals, here the substitution of variables is usually given by $x = x(u,v), y = y(u,v)$ for the case of double integrals and $x = x(u,v,w)$, $y = y(u,v,w)$, $z = z(u,v,w)$ for the case of triple integrals.

Consider the double integral $\displaystyle\iint\limits_R f(x,y)\,dA$. Suppose that we have the change of variables $x = x(u,v)$ and $y = y(u,v)$, and the Fubini Theorem implies that the double integral can be written as $\displaystyle\int\Big(\int f(x,y)\,dy\Big)dx$, here we do not write the upper limit and lower limit explicitly. Note the inner integral in the iterated integral is computed by assuming that x is fixed. When x is a fixed constant, the relation $x = x(u,v)$ gives a relation between u and v, and the implicit differentiation provides that

$$\frac{du}{dv} = -\frac{x_v(u,v)}{x_u(u,v)}$$

if $x_u \neq 0$. Making the substitution of the variable $y = y(u, v)$ with u, v satisfying the relation $x = x(u, v)$, we find that

$$\begin{aligned} dy &= y_u(u,v)du + y_v(u,v)dv = y_u(u,v)\frac{du}{dv}dv + y_v(u,v)dv \\ &= \frac{x_u(u,v)y_v(u,v) - x_v(u,v)y_u(u,v)}{x_u(u,v)}\,dv\,; \end{aligned}$$

thus

$$\int f(x,y)\,dy = \int f(x(u,v),y(u,v))\Big|\frac{x_u(u,v)y_v(u,v) - x_v(u,v)y_u(u,v)}{x_u(u,v)}\Big|\,dv\,.$$

Therefore, the substitution of variable $x = x(u, v)$, where "v is treated as a constant since it has been integrated", is

$$\begin{aligned} &\int \Big(\int f(x,y)\,dy\Big)dx \\ &\quad = \int \Big(\int f(x(u,v),y(u,v))\Big|\frac{x_u(u,v)y_v(u,v) - x_v(u,v)y_u(u,v)}{x_u(u,v)}\Big|\,dv\Big)\big|x_u(u,v)\big|\,du \\ &\quad = \int \Big(\int f(x(u,v),y(u,v))\big|x_u(u,v)y_v(u,v) - x_v(u,v)y_u(u,v)\big|\,dv\Big)du\,. \end{aligned} \tag{14.5.2}$$

Example 14.33. Consider the change of variables using polar coordinate $x = r\cos\theta$, $y = r\sin\theta$ (treat r, θ as the u, v variables, respectively). Then

$$|x_u y_v - x_v y_u| = |\cos\theta\cdot r\cos\theta - (-r\sin\theta)\cdot\sin\theta| = |r| = r\,;$$

thus (14.5.2) implies the change of variables formula for polar coordinates (14.5.1).

Now we consider the possible change of variables formula for triple integrals. Suppose that by the Fubini Theorem,

$$\iiint\limits_{Q} f(x,y,z)\,dV = \int\Big[\int\Big(\int f(x,y,z)\,dz\Big)dy\Big]dx\,,$$

where again we do not state explicitly the upper and the lower limit of each integral. For a given change of variables $x = x(u, v, w)$, $y = y(u, v, w)$ and $z = z(u, v, w)$, the first integral that we need to evaluate is $\int f(x, y, z)\,dz$, and this integral is computed by assuming that x, y are fixed constants. When x and y are fixed constants, the relations $x = x(u, v, w)$ and $y = y(u, v, w)$ give a relation among u, v, w. Suppose that these relations imply that u and v

are differentiable functions of w, then the implicit differentiation (when applicable) provides that

$$\begin{aligned} 0 &= x_u(u,v,w)\frac{du}{dw} + x_v(u,v,w)\frac{dv}{dw} + x_w(u,v,w)\,, \\ 0 &= y_u(u,v,w)\frac{du}{dw} + y_v(u,v,w)\frac{dv}{dw} + y_w(u,v,w)\,; \end{aligned}$$

thus if $x_u y_v - x_v y_u \neq 0$, we have

$$\begin{aligned} \frac{du}{dw} &= \frac{x_v(u,v,w)y_w(u,v,w) - x_w(u,v,w)y_v(u,v,w)}{x_u(u,v,w)y_v(u,v,w) - x_v(u,v,w)y_u(u,v,w)}\,, \\ \frac{dv}{dw} &= \frac{x_w(u,v,w)y_u(u,v,w) - x_u(u,v,w)y_w(u,v,w)}{x_u(u,v,w)y_v(u,v,w) - x_v(u,v,w)y_u(u,v,w)}\,, \end{aligned}$$

and these identities further imply that

$$\begin{aligned} dz &= z_u(u,v,w)du + z_v(u,v,w)dv + z_w(u,v,w)dw \\ &= \Big[z_u\frac{x_v y_w - x_w y_v}{x_u y_v - x_v y_u} + z_v\frac{x_w y_u - x_u y_w}{x_u y_v - x_v y_u} + z_w\Big](u,v,w)dw \\ &= \Big[\frac{x_v y_w z_u - x_w y_v z_u + x_w y_u z_v - x_u y_w z_v + x_u y_v z_w - x_v y_u z_w}{x_u y_v - x_v y_u}\Big](u,v,w)dw\,. \end{aligned}$$

Therefore,

$$\begin{aligned} \int f(x,y,z)\,dz = \int & f(x(u,v,w),y(u,v,w),z(u,v,w))\times \\ & \times\Big|\frac{x_v y_w z_u - x_w y_v z_u + x_w y_u z_v - x_u y_w z_v + x_u y_v z_w - x_v y_u z_w}{x_u y_v - x_v y_u}\Big|(u,v,w)\,dw\,, \end{aligned}$$

and (14.5.2), by treating w as a constant since it has been integrated, implies that

$$\begin{aligned} &\int\Big[\int\Big(\int f(x,y,z)\,dz\Big)dy\Big]dx \\ &\quad= \int\Big[\int\Big(\int f(x(u,v,w),y(u,v,w),z(u,v,w))\times \\ &\qquad\qquad \times\Big|\frac{x_v y_w z_u - x_w y_v z_u + x_w y_u z_v - x_u y_w z_v + x_u y_v z_w - x_v y_u z_w}{x_u y_v - x_v y_u}\Big|(u,v,w)\,dw\Big)\times \\ &\qquad\qquad \times\big|x_u(u,v,w)y_v(u,v,w) - x_v(u,v,w)y_u(u,v,w)\big|\,dv\Big]du \\ &\quad= \int\Big[\int\Big(\int f(x(u,v,w),y(u,v,w),z(u,v,w))\times \\ &\qquad\qquad \times\big|x_v y_w z_u - x_w y_v z_u + x_w y_u z_v - x_u y_w z_v + x_u y_v z_w - x_v y_u z_w\big|(u,v,w)\,dw\Big)dv\Big]du\,. \end{aligned}$$

The naive (but **wrong**) computations above motivate the following

Definition 14.34

If $x = x(u, v)$ and $y = y(u, v)$, the ***Jacobian*** of x and y with respect to u and v, denoted by $\dfrac{\partial(x, y)}{\partial(u, v)}$, is

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

If $x = x(u, v, w)$, $y = y(u, v, w)$ and $z = z(u, v, w)$, the ***Jacobian*** of x, y and z with respect to u, v and w, denoted by $\dfrac{\partial(x, y, z)}{\partial(u, v, w)}$, is

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{vmatrix} = x_u y_v z_w + x_w y_u z_v + x_v y_w z_u - x_w y_v z_u - x_v y_u z_w - x_u y_w z_v \,.$$

In general, if $g_1, g_2, \cdots, g_n$ are functions of n-variables (whose variables are denoted by $u_1, u_2, \cdots, u_n$), then the Jacobian of $g_1, g_2, \cdots, g_n$ (with respect to $u_1, u_2, \cdots, u_n$), denoted by $\dfrac{\partial(g_1, \cdots, g_n)}{\partial(u_1, \cdots, u_n)}$, is

$$\frac{\partial(g_1, \cdots, g_n)}{\partial(u_1, \cdots, u_n)} = \begin{vmatrix} \dfrac{\partial g_1}{\partial u_1} & \dfrac{\partial g_1}{\partial u_2} & \cdots & \dfrac{\partial g_1}{\partial u_n} \\ \dfrac{\partial g_2}{\partial u_1} & \dfrac{\partial g_2}{\partial u_2} & \cdots & \dfrac{\partial g_2}{\partial u_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial g_n}{\partial u_1} & \dfrac{\partial g_n}{\partial u_2} & \cdots & \dfrac{\partial g_n}{\partial u_n} \end{vmatrix} .$$

Example 14.35. The Jacobian of the change of variables given by the polar coordinate $x = a + r\cos\theta$, $y = b + r\sin\theta$ is

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r \,.$$

The Jacobian of the change of variables given by the spherical coordinate $x = \rho\cos\theta\sin\phi$, $y = \rho\sin\theta\sin\phi$, $z = \rho\cos\phi$ is

$$\begin{aligned} \frac{\partial(x, y, z)}{\partial(\rho, \theta, \phi)} &= \begin{vmatrix} \cos\theta\sin\phi & -\rho\sin\theta\sin\phi & \rho\cos\theta\cos\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\phi & 0 & -\rho\sin\phi \end{vmatrix} \\ &= -\rho^2\cos^2\theta\sin^3\phi - \rho^2\sin^2\theta\sin\phi\cos^2\phi - \rho^2\cos^2\theta\sin\phi\cos^2\phi - \rho^2\sin^2\theta\sin^3\phi \\ &= -\rho^2\cos^2\theta\sin\phi - \rho^2\sin^2\theta\sin\phi = -\rho^2\sin\phi \,. \end{aligned}$$

The Jacobian of the change of variables given by the cylindrical coordinate $x = r\cos\theta$, $y = r\sin\theta$, $z = z$ is

$$\frac{\partial(x,y,z)}{\partial(r,\theta,z)} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r\,.$$

Even though the derivation of the change of variables is wrong; however, the conclusion is in fact correct, and we have the following

Theorem 14.36

Let $\mathrm{O} \subseteq \mathbb{R}^2$ be an open set that has area, and $g = (g_1, g_2) : \mathrm{O} \to \mathbb{R}^2$ be an one-to-one continuously differentiable function such that g^{-1} is also continuously differentiable. Assume that the Jacobian of g_1, g_2 (with respective to their variables) does not vanish in O. If $f : g(\mathrm{O}) \to \mathbb{R}$ is integrable (on $g(\mathrm{O})$), then

$$\iint\limits_{g(\mathrm{O})} f(x,y)\,dA = \iint\limits_{\mathrm{O}} f\big(g_1(u,v), g_2(u,v)\big)\Big|\frac{\partial(g_1,g_2)}{\partial(u,v)}\Big|\,dA'\,,$$

where the integral on the right-hand side is the double integral of the function $f\big(g_1(u,v), g_2(u,v)\big)\Big|\frac{\partial(g_1,g_2)}{\partial(u,v)}\Big|$ (with variables u, v) on O.

Theorem 14.37

Let $\mathrm{O} \subseteq \mathbb{R}^3$ be an open set that has volume (that is, the constant function is Riemann integrable on O), and $g = (g_1, g_2, g_3) : \mathrm{O} \to \mathbb{R}^3$ be an one-to-one continuously differentiable function such that g^{-1} is also continuously differentiable. Assume that the Jacobian of g_1, g_2, g_3 (with respective to their variables) does not vanish in O. If $f : g(\mathrm{O}) \to \mathbb{R}$ is integrable (on $g(\mathrm{O})$), then

$$\iiint\limits_{g(\mathrm{O})} f(x,y,z)\,dV = \iiint\limits_{\mathrm{O}} f\big(g_1(u,v,w), g_2(u,v,w), g_3(u,v,w)\big)\Big|\frac{\partial(g_1,g_2,g_3)}{\partial(u,v,w)}\Big|\,dV'\,,$$

where the integral on the right-hand side is the triple integral of the function $f\big(g_1(u,v,w), g_2(u,v,w), g_3(u,v,w)\big)\Big|\frac{\partial(g_1,g_2,g_3)}{\partial(u,v,w)}\Big|$ (with variables u, v, w) on O.

Remark 14.38. Suppose that O is an open set in the plane such that the boundary of O, denoted by $\partial\mathrm{O}$, has zero area. Under suitable assumptions (for example, if the set of discontinuities of f has zero area and f is bounded above or below by a constant), we have

$$\iint\limits_{\mathrm{O}} f(x,y)\,dA = \iint\limits_{\bar{\mathrm{O}}} f(x,y)\,dA\,. \tag{14.5.3}$$

Example 14.39. Let $B = \{(x,y)\,|\,x^2+y^2<R^2\} - [0,1)\times\{0\}$. Then the polar coordinate $x = x(r,\theta) = r\cos\theta$ and $y = y(r,\theta) = r\cos\theta$ is an one-to-one continuously differentiable function from $\mathrm{O} \equiv (0,R)\times(0,2\pi) \to \mathbb{R}^2$ and the inverse function $r = r(x,y) = \sqrt{x^2+y^2}$ and

$$\theta = \theta(x,y) = \begin{cases} \arccos\dfrac{x}{\sqrt{x^2+y^2}} & \text{if } y>0\,,\\ \pi & \text{if } y=0\,,\\ 2\pi - \arccos\dfrac{x}{\sqrt{x^2+y^2}} & \text{if } y<0\,, \end{cases}$$

is also continuously differentiable (which you proved in Quiz). Therefore, the change of variables formula implies that

$$\iint\limits_{B} f(x,y)dA = \iint\limits_{(0,R)\times(0,2\pi)} f(r\cos\theta, r\sin\theta)r\,dA'\,.$$

Let $D(R) = \{(x,y)\,|\,x^2+y^2\leqslant R^2\}$. Then $D = B\cup\partial B$ and $[0,R]\times[0,2\pi] = (0,R)\times(0,2\pi)\cup\partial\big[(0,R)\times(0,2\pi)\big]$; thus (14.5.3) further implies that

$$\iint\limits_{D(R)} f(x,y)dA = \iint\limits_{[0,R]\times[0,2\pi]} f(r\cos\theta, r\sin\theta)r\,dA'\,.$$

In general, if a region R, in polar coordinate, can be expressed as

$$R = \{(r,\theta)\,|\,a\leqslant\theta\leqslant b, g_1(\theta)\leqslant r\leqslant g_2(\theta)\}\,,$$

then

$$\iint\limits_{R} f(x,y)\,dA = \int_a^b\Big(\int_{g_1(\theta)}^{g_2(\theta)} f(r\cos\theta, r\sin\theta)r\,dr\Big)\,d\theta\,,$$

while if R, in polar coordinate, can be expressed as

$$R = \{(r,\theta)\,|\,c\leqslant r\leqslant d, h_1(r)\leqslant\theta\leqslant h_2(r)\}\,,$$

then

$$\iint\limits_{R} f(x,y)\,dA = \int_c^d\Big(\int_{h_1(r)}^{h_2(r)} f(r\cos\theta, r\sin\theta)r\,d\theta\Big)\,dr\,.$$

Example 14.40. In this example we compute the double integral $\iint\limits_R \sqrt{1+4x^2+4y^2}\,dA$ that appears in Example 14.14, where $R = \{(x,y) \,|\, x^2+y^2 \leqslant 1\}$.

Using the polar coordinate, $R = \{(r,\theta) \,|\, 0 \leqslant r \leqslant 1, 0 \leqslant \theta \leqslant 2\pi\}$; thus

$$\iint\limits_R \sqrt{1+4x^2+4y^2}\,dA = \int_0^{2\pi}\Big(\int_0^1 \sqrt{1+4r^2}\cdot r\,dr\Big)d\theta = \int_0^{2\pi}\Big[\frac{1}{12}(1+4r^2)^{\frac{3}{2}}\Big]\Big|_{r=0}^{r=1} d\theta$$
$$= \int_0^{2\pi}(5\sqrt{5}-1)\,d\theta = 2\pi(5\sqrt{5}-1)\,.$$

Example 14.41. In this example we compute the double integral $\iint\limits_R \frac{r}{\sqrt{r^2-x^2-y^2}}\,dA$ that appears in Example 14.13, where $R = \{(x,y) \,|\, x^2+y^2 \leqslant \sigma^2\}$ with $0 < \sigma < r$.

Using the polar coordinate (here we let ρ be the radial variable instead of r since r in this integral is a fixed constant), $R = \{(\rho,\theta) \,|\, 0 \leqslant \rho \leqslant \sigma, 0 \leqslant \theta \leqslant 2\pi\}$; thus

$$\iint\limits_R \frac{r}{\sqrt{r^2-x^2-y^2}}\,dA = \int_0^{2\pi}\Big(\int_0^{\sigma} \frac{r}{\sqrt{r^2-\rho^2}}\cdot\rho\,d\rho\Big)d\theta = \int_0^{2\pi}\Big(-r\sqrt{r^2-\rho^2}\Big)\Big|_{\rho=0}^{\rho=\sigma} d\theta$$
$$= \int_0^{2\pi}\big(r^2-r\sqrt{r^2-\sigma^2}\big)\,d\theta = 2\pi\big(r^2-r\sqrt{r^2-\sigma^2}\big)\,.$$

Example 14.42. Let S be the subset of the upper hemisphere $z = \sqrt{1-x^2-y^2}$ enclosed by the curve C shown in the figure below

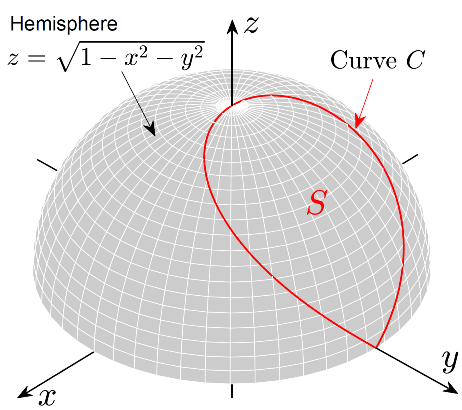


Figure 14.5: Curve S on the upper hemisphere

where each point of C corresponds to some point $(\cos t\sin t, \sin^2 t, \cos t)$ with $t \in \big[-\frac{\pi}{2}, \frac{\pi}{2}\big]$. Find the surface of S.

Let (x,y) be a boundary point of R. The $(x,y) = (\cos t \sin t, \sin^2 t)$ for some $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$; thus

$$x^2 + y^2 = \cos^2 t \sin^2 t + \sin^4 t = (\cos^2 t + \sin^2 t)\sin^2 t = \sin^2 t = y\,.$$

Therefore, the boundary of R consists of points (x,y) satisfying $x^2+y^2=y$ which shows that R is a disk centered at $\left(0,\frac{1}{2}\right)$ with radius $\frac{1}{2}$. Therefore,

$$R = \left\{(x,y)\,\middle|\,0 \leqslant y \leqslant 1, -\sqrt{y-y^2} \leqslant x \leqslant \sqrt{y-y^2}\right\},$$

and by Theorem 14.11 the surface area of S is given by $\displaystyle\iint\limits_R \frac{1}{\sqrt{1-x^2-y^2}}\,dA$.

Now we apply the change of variables using the polar coordinates to compute this double integral. Since we have found the Jacobian of this change of variables, we only need to find the corresponding region R' of R in the $r\theta$-plane and the change of variables formula shows that the surface area of S is $\displaystyle\iint\limits_{R'} \frac{r}{\sqrt{1-r^2}}\,dA'$.

By the fact that the boundary of R' maps to the boundary of R under the change of variables $x = r\cos\theta$ and $y = r\sin\theta$, we find that if (r,θ) is a boundary point of R', then (r,θ) satisfies $r^2 = r\sin\theta$; thus the boundary of R' consists of points (r,θ) satisfying $r = \sin\theta$ or $r = 0$ in the $r\theta$-plane. Since R locates on the upper half plane, $0 \leqslant \theta \leqslant \pi$, and the center of the disk R corresponds to point $\left(\frac{1}{2},\frac{\pi}{2}\right)$ in the $r\theta$-plane, we conclude that

$$R' = \left\{(r,\theta)\,\middle|\,0 \leqslant \theta \leqslant \pi, 0 \leqslant r \leqslant \sin\theta\right\}.$$

Therefore,

$$\begin{aligned}\iint\limits_{R'} \frac{r}{\sqrt{1-r^2}}\,dA' &= \int_0^\pi \left(\int_0^{\sin\theta} \frac{1}{\sqrt{1-r^2}}\,r\,dr\right)d\theta = \int_0^\pi \left[\left(-\sqrt{1-r^2}\right)\Big|_{r=0}^{r=\sin\theta}\right]d\theta \\ &= \int_0^\pi \left(1 - |\cos\theta|\right)d\theta = \pi - 2\int_0^{\frac{\pi}{2}} \cos\theta\,d\theta = \pi - 2\left(\sin\theta\,\Big|_{\theta=0}^{\theta=\frac{\pi}{2}}\right) = \pi - 2\,.\end{aligned}$$

Example 14.43. In this example we compute the improper integral $\displaystyle\int_0^\infty e^{-x^2}\,dx$. First we note that this improper integral converges since $0 \leqslant e^{-x^2} \leqslant e^{-x}$ for all $x \geqslant 1$ and $\displaystyle\int_1^\infty e^{-x}\,dx = e^{-1} < \infty$, the comparison test implies that $\displaystyle\int_1^\infty e^{-x^2}\,dx$ converges.

Let $I = \displaystyle\int_0^{\infty} e^{-x^2}\,dx$. Then $I = \displaystyle\int_0^{\infty} e^{-y^2}\,dy$; thus

$$\begin{aligned} I^2 &= \Big(\int_0^{\infty} e^{-x^2}\,dx\Big)\Big(\int_0^{\infty} e^{-y^2}\,dy\Big) = \int_0^{\infty}\Big(\int_0^{\infty} e^{-y^2}\,dy\Big)e^{-x^2}\,dx \\ &= \int_0^{\infty}\Big(\int_0^{\infty} e^{-x^2}e^{-y^2}\,dy\Big)\,dx = \int_0^{\infty}\Big(\int_0^{\infty} e^{-(x^2+y^2)}\,dy\Big)\,dx = \iint\limits_R e^{-(x^2+y^2)}\,dA\,, \end{aligned}$$

where R is the first quadrant of the plane. In polar coordinate, the first quadrant can be expressed as $0 < r < \infty$ and $0 < \theta < \frac{\pi}{2}$; thus using the polar coordinate we find that

$$I^2 = \int_0^{\frac{\pi}{2}}\Big(\int_0^{\infty} e^{-r^2} r\,dr\Big)d\theta = \int_0^{\frac{\pi}{2}}\Big(-\frac{1}{2}e^{-r^2}\Big)\Big|_{r=0}^{r=\infty}\,d\theta = \frac{\pi}{4}\,.$$

By the fact that $I \geqslant 0$, we conclude that $I = \dfrac{\sqrt{\pi}}{2}$.

Example 14.44. The Jacobian in the change of variable using spherical coordinate is $\rho^2\sin\phi$ Let Q be a solid region in space, and $f : Q \to \mathbb{R}$ be continuous. Suppose that Q, in spherical coordinate, can be expressed as

$$\big\{(\rho,\theta,\phi)\,\big|\,a \leqslant \phi \leqslant b, g_1(\phi) \leqslant \theta \leqslant g_2(\phi),$$

Example 14.45. In this example we reconsider the volume of Q in Example 14.30, where

$$Q = \big\{(x,y,z)\,\big|\,(x,y) \in R, x^2+y^2 \leqslant z \leqslant \sqrt{6-x^2-y^2}\big\}\,,$$

and R is a disk centered at the origin with radius $\sqrt{2}$.

Using the cylindrical coordinate, the region Q can be expressed as

$$\big\{(r,\theta,z)\,\big|\,0 \leqslant r \leqslant \sqrt{2}, 0 \leqslant \theta \leqslant 2\pi, r^2 \leqslant z \leqslant \sqrt{6-r^2}\big\}\,.$$

Therefore, the volume of Q is given by

$$\begin{aligned} \iiint\limits_Q dV &= \int_0^{2\pi}\Big[\int_0^{\sqrt{2}}\Big(\int_{r^2}^{\sqrt{6-r^2}} r\,dz\Big)dr\Big]d\theta = \int_0^{2\pi}\Big[\int_0^{\sqrt{2}} r\big(\sqrt{6-r^2}-r^2\big)\,dr\Big]d\theta \\ &= \int_0^{2\pi}\Big[-\frac{1}{3}(6-r^2)^{\frac{3}{2}} - \frac{1}{4}r^4\Big]\Big|_{r=0}^{r=\sqrt{2}}\,d\theta = \int_0^{2\pi}\big(-\frac{8}{3} - 1 + 2\sqrt{6}\big)\,d\theta = 2\pi\big(2\sqrt{6} - \frac{11}{3}\big)\,. \end{aligned}$$

Example 14.46. Find the volume of the solid region Q bounded below by the cone $z = \sqrt{x^2+y^2}$ and above by the sphere $x^2+y^2+z^2=9$.

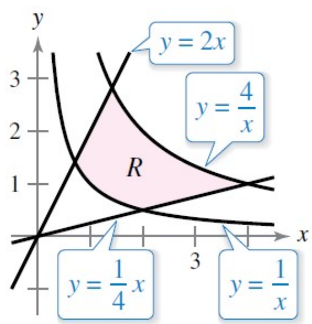
Using spherical coordinate, Q can be expressed as

$$\left\{(\rho,\theta,\phi)\,\middle|\, 0\leqslant\rho\leqslant 3, 0\leqslant\theta\leqslant 2\pi, 0\leqslant\phi\leqslant\frac{\pi}{4}\right\}.$$

Therefore, the volume of Q is given by

$$\iiint\limits_Q dV = \int_0^{\frac{\pi}{4}}\Big[\int_0^{2\pi}\Big(\int_0^3 \rho^2\sin\phi\,d\rho\Big)d\theta\Big]d\phi = 18\pi\int_0^{\frac{\pi}{4}}\sin\phi\,d\phi = 18\pi\big(1-\frac{\sqrt{2}}{2}\big).$$

Example 14.47. Find the double integral $\displaystyle\iint\limits_R e^{-\frac{xy}{2}}\,dA$, where R is the region given in the following figure.



Consider the following change of variables: $x=\sqrt{\dfrac{v}{u}}$ and $y=\sqrt{uv}$. In order to apply the change of variables formula to find the double integral, we need to know

1. What is the Jacobian of this change of variable?

2. What is the corresponding region of integration in the uv-plane?

We first note that for the change of variables to make sense, u, v have the same sign. W.L.O.G., we assume that the corresponding region in the uv-plane lies in the first quadrant. We compute the Jacobian and find that

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{1}{2}\sqrt{\dfrac{u}{v}}\cdot\dfrac{-v}{u^2} & \dfrac{1}{2}\sqrt{\dfrac{u}{v}}\cdot\dfrac{1}{u} \\ \dfrac{1}{2}\dfrac{v}{\sqrt{uv}} & \dfrac{1}{2}\dfrac{u}{\sqrt{uv}} \end{vmatrix} = \frac{1}{4}\cdot\frac{-1}{u}-\frac{1}{4}\cdot\frac{1}{u} = -\frac{1}{2u}.$$

Now we find the corresponding region R' in the uv-plane. The rule of thumb is that a one-to-one continuously differentiable function whose Jacobian does not vanish maps the boundary of a region to the boundary of its image. Therefore, the boundary of R' is given by $u = \frac{1}{2}$, $u = 2$ and $v = 1$, $v = 4$. Since the point (x, y) satisfying $xy = 2$ and $\frac{y}{x} = 1$ corresponds to $u = 1$ and $v = 2$, we find that $R' = \left[\frac{1}{2}, 2\right] \times [1, 4]$. Therefore, the change of variable formula implies that

$$\iint\limits_R e^{-\frac{xy}{2}}\,dA = \iint\limits_{[\frac{1}{2},2]\times[1,4]} e^{-\frac{v}{2}}\frac{1}{2u}\,dA' = \int_{\frac{1}{2}}^{2}\Big(\int_1^4 \frac{e^{-\frac{v}{2}}}{2u}\,dv\Big)du$$

$$= \int_{\frac{1}{2}}^{2}\Big[\Big(\frac{-e^{-\frac{v}{2}}}{u}\Big)\Big|_{v=1}^{v=4}\Big]\,du = \big(e^{-\frac{1}{2}} - e^{-2}\big)\int_{\frac{1}{2}}^{2}\frac{1}{u}\,du = 3\ln 2\big(e^{-\frac{1}{2}} - e^{-2}\big)\,.$$

A more fundamental question is: why do we choose this change of coordinate? The general philosophy is to "straighten" the boundary so that in the new coordinate system the corresponding region becomes a region bounded by straight lines. Observing that the boundaries of the region R consists of four curves $\frac{y}{x} = \frac{1}{4}$, $\frac{y}{x} = 2$, $xy = 1$ and $xy = 4$, it is quite intuitive that we choose $u = \frac{y}{x}$ and $v = xy$ as our change of variables (in a reverse order). Solving for x, y in terms of u, v, we find that $x = \sqrt{\frac{v}{u}}$ and $y = \sqrt{uv}$.

14.6 Exercise

Problem 14.1. Evaluate the following iterated integrals.

(1) $\int_{-1}^{1}\Big(\int_0^1 ye^{x^2+y^2}\,dx\Big)dy$ (2) $\int_0^2\Big(\int_y^{\sqrt{8-y^2}}\sqrt{x^2+y^2}\,dx\Big)dy$ (3) $\int_0^1\Big(\int_{\sqrt{y}}^1 e^{x^3}\,dx\Big)dy$

(4) $\int_0^1\Big(\int_y^1 \frac{1}{1+x^4}\,dx\Big)dy$ (5) $\int_0^4\Big(\int_{\frac{x}{2}}^2 \sin(y^2)\,dy\Big)dx$ (6) $\int_0^4\Big(\int_{\sqrt{x}}^2 \frac{1}{y^3+1}\,dy\Big)\,dx$

(7) $\int_0^2\Big(\int_x^2 x\sqrt{1+y^3}\,dy\Big)dx$ (8) $\int_0^2\Big(\int_{\frac{y}{2}}^1 \exp(x^2)\,dx\Big)dy$ (9) $\int_0^1\Big(\int_0^1 \frac{y}{1+x^2y^2}\,dx\Big)\,dy$

(10) $\int_0^{\frac{\pi}{2}}\Big(\int_x^{\frac{\pi}{2}} \frac{\sin y}{y}\,dy\Big)dx$ (11) $\int_0^2\Big(\int_{y^2}^4 \sqrt{x}\sin x\,dx\Big)dy$ (12) $\int_0^2\Big(\int_0^{4-x^2} \frac{xe^{2y}}{4-y}dy\Big)dx$

(13) $\int_0^1\Big(\int_{\arcsin y}^{\frac{\pi}{2}} \cos x\sqrt{1+\cos^2 x}\,dx\Big)dy$ (14) $\int_{-5}^{5}\Big[\int_0^{\sqrt{25-x^2}}\Big(\int_0^{\frac{1}{x^2+y^2}} \sqrt{x^2+y^2}\,dz\Big)dy\Big]dx$

(15) $\int_0^4\Big[\int_0^1\Big(\int_{2y}^{y} \frac{2\cos(x^2)}{\sqrt{z}}\,dx\Big)dy\Big]dz$ (16) $\int_0^1\Big[\int_0^1\Big(\int_{x^2}^1 xz\exp(zy^2)\,dy\Big)dx\Big]dz$

(17) $\displaystyle\int_0^1 \left[\int_{\sqrt[3]{z}}^1 \left(\int_0^{\ln 3} \frac{\pi e^{2x} \sin(\pi y^2)}{y^2}\, dx \right) dy \right] dz$ (18) $\displaystyle\int_0^2 \left[\int_0^{4-x^2} \left(\int_0^x \frac{\sin(2z)}{4-z}\, dy \right) dz \right] dx$

Problem 14.2. Evaluate the double integral $\displaystyle\iint\limits_R f(x,y)\, dA$ with the following f and R.

(1) $f(x,y) = y^2 e^{xy}$, and R is the region bounded by $y = x$, $y = 4$ and $x = 0$.

(2) $f(x,y) = xy$, and R is the region bounded by the line $y = x - 1$ and parabola $y^2 = 2x + 6$.

(3) $f(x,y) = \sin^4(x+y)$, and R is the triangle enclosed by the lines $y = 0$, $y = 2x$, and $x = 1$.

(4) $f(x,y) = x^2 + x^2y^3 - y^2 \sin x$, and $R = \{(x,y) \,|\, |x| + |y| \leqslant 1\}$.

(5) $f(x,y) = |x| + |y|$, and $R = \{(x,y) \,|\, |x| + |y| \leqslant 1\}$.

(6) $f(x,y) = xy$, and R is the region in the first quadrant bounded by curves $x^2 + y^2 = 4$, $x^2 + y^2 = 9$, $x^2 - y^2 = 1$ and $x^2 - y^2 = 4$.

(7) $f(x,y) = x$, and R is the region in the first quadrant bounded by curves $4x^2 - y^2 = 4$, $4x^2 - y^2 = 16$, $y = x$ and the x-axis.

(8) $f(x,y) = \exp(-x^2 - 4y^2)$, and $R = \{(x,y) \,|\, x^2 + 4y^2 \leqslant 1\}$.

(9) $f(x,y) = \exp\left(\dfrac{2y-x}{2x+y}\right)$, and R is the trapezoid with vertices $(0,2)$, $(1,0)$, $(4,0)$ and $(0,8)$.

Problem 14.3. Evaluate the triple integral $\displaystyle\iiint\limits_D f(x,y,z)\, dV$ with the following f and D.

(1) $f(x,y,z) = x - y + z^2$, and D is the solid region bounded above by $z = 1 + x^2 + y^2$, bounded below by $z = 0$, and inside $x^2 + y^2 = 4$.

(2) $f(x,y,z) = 1$, and D is the solid region bounded by $z = x^2 + y^2$, $x^2 + y^2 = 4$ and $z = 0$.

(3) $f(x,y,z) = 1$, and $D = \left\{(x,y,z) \in \mathbb{R}^3 \,\middle|\, \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} \leqslant 1\right\}$, where $a, b, c > 0$.

Problem 14.4. Evaluate the integral $\displaystyle\int_0^2 \big[\arctan(\pi x) - \arctan x\big]\,dx$ by converting the integral into a double integral and evaluating the double integral by changing the order of integration.

Problem 14.5. Let a, b be positive constants. Evaluate the integral

$$\int_0^a \Big(\int_0^b \exp\big(\max\{b^2x^2, a^2y^2\}\big)\,dy \Big)\,dx\,.$$

Problem 14.6. Show that if $\lambda > \dfrac{1}{2}$, there does not exist a real-valued continuous function u such that for all x in the closed interval $[0, 1]$,

$$u(x) = 1 + \lambda \int_x^1 u(y)u(y-x)\,dy\,.$$

Hint: Assume the contrary that there exists such a function u. Integrate the equation above on the interval $[0, 1]$.

Problem 14.7. Find the surface area for the portion of the surface $z = xy$ that is inside the cylinder $x^2 + y^2 = 1$.

Problem 14.8. Let Σ be a parametric surface parameterized by

$$\boldsymbol{r}(u, v) = X(u, v)\mathbf{i} + Y(u, v)\mathbf{j} + Z(u, v)\mathbf{k}\,, \quad (u, v) \in R\,.$$

Define $E = \boldsymbol{r}_u \cdot \boldsymbol{r}_u$, $F = \boldsymbol{r}_u \cdot \boldsymbol{r}_v$ and $G = \boldsymbol{r}_v \cdot \boldsymbol{r}_v$. Show that

$$\|\boldsymbol{r}_u \times \boldsymbol{r}_v\|^2 = EG - F^2\,.$$

Hint: You can try to make use of ε_{ijk}, the permutation symbol.
Remark: This quantity $EG - F^2$ is called the first fundamental form (associated with the parametrization $\boldsymbol{r}$).

Problem 14.9. Let $k > 0$ be a constant. Show that the surface area of the cone $z = k\sqrt{x^2 + y^2}$ that lies above the circular region $x^2 + y^2 \leqslant r^2$ in the xy-plane is $\pi r^2\sqrt{k^2 + 1}$ by the following methods:

1. Use the formula $\displaystyle\iint_R \sqrt{1 + \|(\nabla f)(x, y)\|^2}\,dA$ directly.

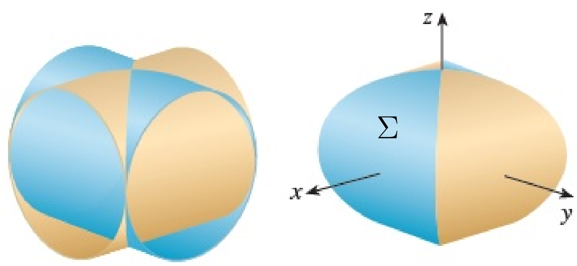
2. Find a parametrization of the cone above using r, θ (from the polar coordinate) as the parameters and make use of the formula $\displaystyle\iint_D \|(\boldsymbol{r}_r \times \boldsymbol{r}_\theta)(r,\theta)\|\, d(r,\theta)$.

Problem 14.10. Let Σ be the surface formed by rotating the curve

$$C = \left\{(x,y,z) \in \mathbb{R}^3 \,\middle|\, x = \cos z, y = 0, -\frac{\pi}{2} \leqslant z \leqslant \frac{\pi}{2}\right\}$$

about the z-axis. Find a parametrization for Σ and compute its surface area.

Problem 14.11. The figure below shows the surface created when the cylinder $y^2 + z^2 = 1$ intersects the cylinder $x^2 + z^2 = 1$. Let Σ be the part shown in the figure.



(1) Find the area of Σ using the formula $\displaystyle\iint_R \sqrt{1 + \|(\nabla f)(x,y)\|^2}\, dA$.

(2) Parameterize Σ using θ, z as parameters (from the cylindrical coordinate) and find the area of this surface using the formula $\displaystyle\iint_D \|(\boldsymbol{r}_\theta \times \boldsymbol{r}_z)(\theta,z)\|\, d(r,\theta)$.

(3) Parameterize Σ using θ, ϕ as parameters (from the spherical coordinate) and find the area of this surface using the formula $\displaystyle\iint_D \|(\boldsymbol{r}_\theta \times \boldsymbol{r}_\phi)(\theta,z)\|\, d(r,\theta)$.

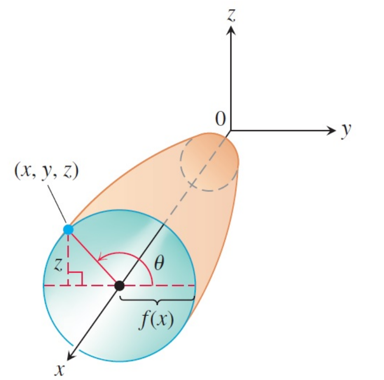
(3) Find the volume of this intersection using triple integrals.

Problem 14.12. Let Σ be the surface obtained by rotating the smooth curve $y = f(x)$, $a \leqslant x \leqslant b$ about the x-axis, where $f(x) > 0$.

1. Show that

$$\boldsymbol{r}(x,\theta) = x\mathbf{i} + f(x)\cos\theta\,\mathbf{j} + f(x)\sin\theta\,\mathbf{k}, \quad (x,\theta) \in [a,b] \times [0,2\pi],$$

is a parametrization of Σ, where θ is the angle of rotation about the x-axis (see the accompanying figure).

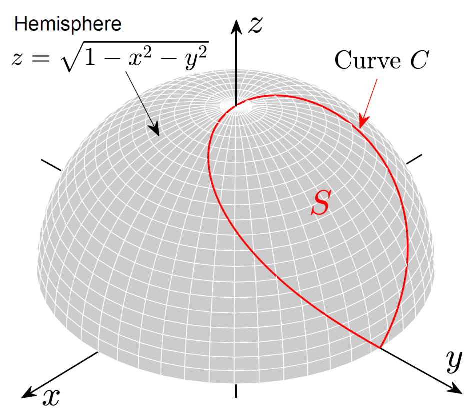


2. Show that the surface area of Σ is

$$\int_a^b 2\pi f(x)\sqrt{1+f'(x)^2}\,dx$$

using the formula $\displaystyle\iint_D \|(\boldsymbol{r}_r \times \boldsymbol{r}_\theta)(r,\theta)\|\,d(r,\theta)$.

Problem 14.13. Let S be the subset of the upper hemisphere $z=\sqrt{1-x^2-y^2}$ enclosed by the curve C shown in the figure below



where each point of C corresponds to some point $(\cos t\sin t, \sin^2 t, \cos t)$ with $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Find the surface of S via the following steps:

(1) Let R be the region obtained by projecting S onto the xy-plane along the z-axis. Suppose that R can be expressed as $R=\left\{(x,y)\,\middle|\, c\leqslant y\leqslant d, g_1(y)\leqslant x\leqslant g_2(y)\right\}$. Find c, d and g_1, g_2, and find the surface area of S using the formula $\displaystyle\iint_R \sqrt{1+\|(\nabla f)(x,y)\|^2}\,dA$.

(2) The surface S is a parametric surface parameterized by

$$S = \left\{ \boldsymbol{r} \,\middle|\, \boldsymbol{r} = \cos\theta\sin\phi\,\mathbf{i} + \sin\theta\sin\phi\,\mathbf{j} + \cos\phi\,\mathbf{k} \text{ for some } (\theta,\phi) \in D \right\}.$$

Find the domain D inside the rectangle $[0, 2\pi] \times [0, \pi]$, and find the surface area of S using the formula $\displaystyle\iint\limits_D \|(\boldsymbol{r}_\theta \times \boldsymbol{r}_\phi)(\theta,\phi)\|\, d(\theta,\phi)$.

Problem 14.14. Rewrite the following iterated integrals as an equivalent iterated integral in the five other orders.

(1) $\displaystyle\int_0^1 \Big[\int_y^1 \Big(\int_0^y f(x,y,z)\,dz\Big)dx\Big]dy$ (2) $\displaystyle\int_0^1 \Big[\int_y^1 \Big(\int_0^z f(x,y,z)\,dx\Big)dz\Big]dy$

(3) $\displaystyle\int_0^1 \Big[\int_0^{1-x^2} \Big(\int_0^{1-x} f(x,y,z)\,dy\Big)dz\Big]dx$ (4) $\displaystyle\int_0^3 \Big[\int_0^x \Big(\int_0^{9-x^2} f(x,y,z)\,dz\Big)dy\Big]dx$

(5) $\displaystyle\int_0^1 \Big[\int_{\sqrt{x}}^1 \Big(\int_0^{1-y} f(x,y,z)\,dz\Big)dy\Big]dx$ (6) $\displaystyle\int_{-1}^1 \Big[\int_{x^2}^1 \Big(\int_0^{1-y} f(x,y,z)\,dz\Big)dy\Big]dx$

Problem 14.15. Find volume of the solid that lies under $z = x^2 + y^2$ and above the region R in the xy-plane bounded by the line $y = 2x$ and parabola $y = x^2$.

Problem 14.16. Evaluate the triple integral $\displaystyle\iiint\limits_D dV$, where D is bounded by $z = x^2 + y^2$, $x^2 + y^2 = 4$ and $z = 0$.

Problem 14.17. Evaluate the double integral $\displaystyle\iint\limits_R \arctan\frac{y}{x}\,dA$ using the polar coordinate, where

$$R = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, 1 \leqslant x^2 + y^2 \leqslant 4, 0 \leqslant y \leqslant x\right\}.$$

Problem 14.18. Evaluate the triple integral $\displaystyle\iiint\limits_D x\exp(x^2 + y^2 + z^2)\,dV$, where D is the portion of the unit ball $x^2 + y^2 + z^2 \leqslant 1$ that lies in the first octant.

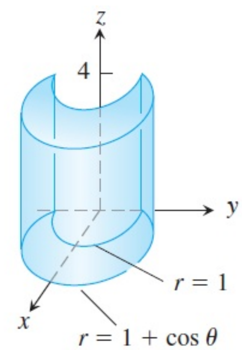
Problem 14.19. Evaluate the triple integral $\displaystyle\iiint\limits_D \sqrt{x^2 + y^2 + z^2}\,dV$, where D is the region lying above the cone $z = \sqrt{x^2 + y^2}$ and between the spheres $x^2+y^2+z^2 = 1$ and $x^2+y^2+z^2 = 4$.

Problem 14.20. Use the cylinder coordinate to find the volume of the ball $x^2+y^2+z^2 = a^2$.

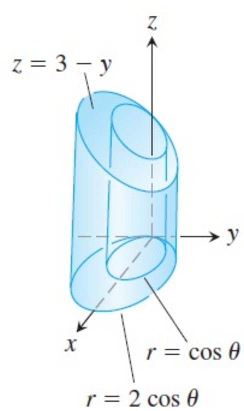
Problem 14.21. Use the spherical coordinate to find the volume of the cylindricality $x^2 + y^2 = r^2$, where $0 \leqslant z \leqslant h$.

Problem 14.22. Compute the volume of D given below using triple integrals in cylindrical coordinates.

(1) D is the solid right cylinder whose base is the region in the xy-plane that lies inside the cardioid $r = 1 + \cos\theta$ and outside the circle $r = 1$ and whose top lies in the plane $z = 4$.

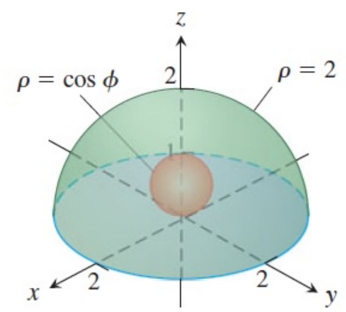


(2) D is the solid right cylinder whose base is the region between the circles $r = \cos\theta$ and $r = 2\cos\theta$ and whose top lies in the plane $z = 3 - y$.

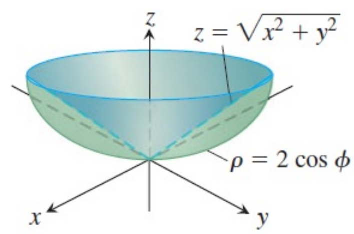


Problem 14.23. Compute the volume of D given below using triple integrals in spherical coordinates.

(1) D is the solid between the sphere $\rho = \cos\phi$ and the hemisphere $\rho = 2$, $z \geqslant 0$.



(2) D is the solid bounded below by the sphere $\rho = 2\cos\phi$ and above by the cone $z = \sqrt{x^2+y^2}$.



Problem 14.24. Convert the integral

$$\int_{-1}^{1}\Big[\int_{0}^{\sqrt{1-y^2}}\Big(\int_{0}^{x}(x^2+y^2)\,dz\Big)dx\Big]dy$$

to an equivalent integral in cylindrical coordinates and evaluate the result.

Problem 14.25. Find the integrals given below with specific change of variables.

(1) Find $\int_0^2\Big(\int_{\frac{y}{2}}^{\frac{y+4}{2}} y^3(2x-y)e^{(2x-y)^2}\,dx\Big)dy$ using change of variables $x = u + \frac{1}{2}v$, $y = v$.

(2) Find $\displaystyle\iint_{[0,1]\times[0,1]} \frac{1}{(1+xy)\ln(xy)}\,dA$ by making the change of variables $u = xy$ and $v = y$.

(3) Find $\int_1^2\Big(\int_{\frac{1}{y}}^{y}(x^2+y^2)\,dx\Big)dy + \int_2^4\Big(\int_{\frac{y}{4}}^{\frac{4}{y}}(x^2+y^2)\,dx\Big)dy$ using change of variables $x = \frac{u}{v}$, $y = uv$.

(4) Find $\int_0^1\Big(\int_0^{2\sqrt{1-x}}\sqrt{x^2+y^2}\,dy\Big)dx$ using change of variables $x = u^2 - v^2$, $y = 2uv$.

(5) Let R be the region in the first quadrant of the xy-plane bounded by the hyperbolas $xy = 1$, $xy = 9$ and the lines $y = x$, $y = 4x$. Find $\displaystyle\iint_R \Big(\sqrt{\frac{y}{x}} + \sqrt{xy}\Big)\,dA$ using the change of variables $x = \frac{u}{v}$, $y = uv$.

(6) Let D be the solid region in xyz-space defined by

$$D = \left\{(x, y, z) \,\middle|\, 1 \leqslant x \leqslant 2, 0 \leqslant xy \leqslant 2, 0 \leqslant z \leqslant 1\right\}.$$

Find $\displaystyle\iiint\limits_D (x^2y + 3xyz)\, dV$ using change of variables $u = x$, $v = xy$, $w = 3z$.

Problem 14.26. Evaluate the double integral $\displaystyle\iint\limits_R (x+y)e^{x^2-y^2}\, dA$, where R is rectangle enclosed by the lines $x - y = 0$, $x - y = 2$, $x + y = 0$, and $x + y = 3$.

Problem 14.27. Let f be continuous on $[0,1]$ and let R be the triangular region with vertices $(0,0)$, $(1,0)$, and $(0,1)$. Show that

$$\iint\limits_R f(x+y)\, dA = \int_0^1 uf(u)\, du\,.$$

Problem 14.28. Let A be the area of the region in the first quadrant bounded by the line $y = \dfrac{1}{2}x$, the x-axis, and the ellipse $\dfrac{1}{9}x^2 + y^2 = 1$. Find the positive number m such that A is equal to the area of the region in the first quadrant bounded by the line $y = mx$, the y-axis, and the ellipse $\dfrac{1}{9}x^2 + y^2 = 1$.
Hint: Try to make change of variables so that the computation of the area of the region in the first quadrant bounded by the line $y = mx$, the y-axis, and the ellipse $\dfrac{1}{9}x^2 + y^2 = 1$ looks the same as the former one.