Chapter 4. Differentiation

§1. Basic Properties of the Derivative

Let f be a real-valued function defined on an interval I in $\mathbb{R}$. The **derivative** of f at a point $a \in I$ is defined to be

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a},$$

if this limit exists as a real number. The derivative f at a is denoted by $f'(a)$. We say that f is **differentiable** at a if $f'(a)$ exists. We say that f is differentiable on I if $f'(x)$ exists at each $x \in I$. In this case, f' itself is a function from I to $\mathbb{R}$.

For example, let $n \in \mathbb{N}_0$ and let $f(x) = x^n$ for $x \in \mathbb{R}$. We show that $f'(x) = nx^{n-1}$ for all $x \in \mathbb{R}$. Indeed, for $n = 0$ we have $f(x) = 1$ for all $x \in \mathbb{R}$. Consequently, $f'(x) = 0$ for all $x \in \mathbb{R}$. Suppose $n \in \mathbb{N}$. For $a \in \mathbb{R}$ and $x \neq a$, we have

$$\frac{f(x)-f(a)}{x-a} = x^{n-1} + ax^{n-2} + \cdots + a^{n-2}x + a^{n-1}.$$

It follows that

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = a^{n-1} + aa^{n-2} + \cdots + a^{n-2}a + a^{n-1} = na^{n-1}.$$

If f is differentiable at a point a, then f is continuous at a. Indeed, for $x \neq a$ we have

$$f(x) = (x-a)\,\frac{f(x)-f(a)}{x-a} + f(a).$$

Hence, $\lim_{x\to a} f(x) = f(a)$.

Theorem 1.1. *Let f and g be two functions from an interval I to $\mathbb{R}$. Suppose that f and g are differentiable at a point $a \in I$.*

(1) For any $c \in \mathbb{R}$, the function cf is differentiable at a and $(cf)'(a) = cf'(a)$.

(2) The function $f+g$ is differentiable at a and $(f+g)'(a) = f'(a) + g'(a)$.

(3) The function fg is differentiable at a and $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

(4) If $g(a) \neq 0$, then the function f/g is differentiable at a and

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}.$$

Proof. (1) We have

$$(cf)'(a) = \lim_{x\to a} \frac{(cf)(x)-(cf)(a)}{x-a} = \lim_{x\to a} c\cdot \frac{f(x)-f(a)}{x-a} = cf'(a).$$

(2) This is true because the following identity holds for $x \neq a$:

$$\frac{(f+g)(x)-(f+g)(a)}{x-a} = \frac{f(x)-f(a)}{x-a} + \frac{g(x)-g(a)}{x-a}.$$

(3) For $x \in I \setminus \{a\}$, we have

$$\frac{(fg)(x)-(fg)(a)}{x-a} = f(x)\frac{g(x)-g(a)}{x-a} + g(a)\frac{f(x)-f(a)}{x-a}.$$

Taking the limit as $x \to a$ and noting that $\lim_{x\to a} f(x) = f(a)$, we obtain the product rule.

(4) Since $g(a) \neq 0$ and g is continuous at a, there exists an open interval J containing a such that $g(x) \neq 0$ for $x \in I \cap J$. For $x \in I \cap J$ we can write

$$\left(\frac{f}{g}\right)(x) - \left(\frac{f}{g}\right)(a) = \frac{g(a)f(x)-f(a)g(x)}{g(x)g(a)} = \frac{g(a)f(x)-g(a)f(a)+g(a)f(a)-f(a)g(x)}{g(x)g(a)}.$$

Hence, for $x \in I \cap J$ and $x \neq a$ we have

$$\frac{(f/g)(x)-(f/g)(a)}{x-a} = \left(g(a)\frac{f(x)-f(a)}{x-a} - f(a)\frac{g(x)-g(a)}{x-a}\right)\frac{1}{g(x)g(a)}.$$

Taking the limit as $x \to a$ and noting that $\lim_{x\to a} g(x) = g(a)$, we obtain the quotient rule. □

For example, let $p(x) := c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$ for $x \in \mathbb{R}$, where $c_0, c_1, c_2, \ldots, c_n$ are real numbers. Then p is a polynomial and

$$p'(x) = c_1 + 2c_2x + \cdots + nc_nx^{n-1}.$$

Suppose p and q are two polynomials. Let $Z_q := \{x \in \mathbb{R} : q(x) = 0\}$. Let h be the rational function given by $h(x) := p(x)/q(x)$ for $x \in \mathbb{R} \setminus Z_q$. If q is not identically 0, then Z_q is a finite set. In this case, by the quotient rule we obtain

$$h'(x) = \frac{p'(x)q(x)-p(x)q'(x)}{[q(x)]^2}, \quad x \in \mathbb{R} \setminus Z_q.$$

In particular, if $n \in \mathbb{N}$ and $h(x) := 1/x^n$ for $x \in \mathbb{R} \setminus \{0\}$, then

$$h'(x) = \frac{-nx^{n-1}}{x^{2n}} = -nx^{-n-1}, \quad x \in \mathbb{R} \setminus \{0\}.$$

Theorem 1.2. *(The Chain Rule) Let f be a function from an interval I to an interval J, and let g be a function from J to $\mathbb{R}$. If f is differentiable at a and g is differentiable at $f(a)$, then the composite function $g \circ f$ is differentiable at a and $(g \circ f)'(a) = g'(f(a))f'(a)$.*

Proof. Let h be the function from J to $\mathbb{R}$ given by

$$h(y) := \frac{g(y) - g(f(a))}{y - f(a)} \quad \text{for } y \in J \setminus \{f(a)\},$$

and $h(f(a)) := g'(f(a))$. Since g is differentiable at $f(a)$, we have

$$\lim_{y \to f(a)} h(y) = g'(f(a)) = h(f(a)).$$

Hence the function h is continuous at $f(a)$. Moreover,

$$g(y) - g(f(a)) = h(y)(y - f(a)) \quad \forall\, y \in J.$$

Consequently, for $x \in I \setminus \{a\}$ we have

$$\frac{g \circ f(x) - g \circ f(a)}{x - a} = h(f(x))\, \frac{f(x) - f(a)}{x - a}.$$

Taking the limit in the above equation as $x \to a$, we obtain $(g \circ f)'(a) = g'(f(a))f'(a)$. □

Theorem 1.3. *(Inverse Function Theorem) Let f be a real-valued function on an interval I in $\mathbb{R}$. If f is strictly monotone and continuous, then $J := f(I)$ is an interval in $\mathbb{R}$ and the inverse function g of f is continuous. If, in addition, f is differentiable at some point $c \in I$ and $f'(c) \neq 0$, then g is differentiable at $f(c)$ and*

$$g'(f(c)) = \frac{1}{f'(c)}.$$

Proof. It suffices to prove the theorem for the case that f is strictly increasing. The first part of the theorem was proved in Theorem 5.3 of Chapter 3. Suppose that f is differentiable at $c \in I$ and $f'(c) \neq 0$. To each $y \in J$ let $x = g(y)$. Then $y = f(x)$. Since g is continuous, we have

$$\lim_{y \to f(c)} x = \lim_{y \to f(c)} g(y) = g(f(c)) = c.$$

Moreover, $y \neq f(c)$ implies $x \neq c$. Hence,

$$\lim_{y \to f(c)} \frac{g(y) - g(f(c))}{y - f(c)} = \lim_{x \to c} \frac{x - c}{f(x) - f(c)} = \lim_{x \to c} \frac{1}{\frac{f(x) - f(c)}{x - c}} = \frac{1}{f'(c)}.$$

This shows $g'(f(c)) = 1/f'(c)$. □

Let us find the derivative of the root function $g : x \mapsto \sqrt[n]{x}$, $x \in (0, \infty)$, where n is a positive integer. It is the inverse of the power function $f : x \mapsto x^n$, $x \in (0, \infty)$. In particular, $f(\sqrt[n]{x}) = x$ for all $x \in (0, \infty)$. By Theorem 1.3, g is differentiable on $(0, \infty)$ and

$$g'(x) = \frac{1}{f'(\sqrt[n]{x})} = \frac{1}{n(\sqrt[n]{x})^{n-1}} = \frac{1}{n}\, x^{1/n-1}, \quad x \in (0, \infty).$$

Moreover, let $h(x) = x^r$ for $x > 0$, where $r = m/n$, $m \in \mathbb{Z}$ and $n \in \mathbb{N}$. Consequently, $h(x) = [x^{1/n}]^m$. By the chain rule, we have

$$h'(x) = m[x^{1/n}]^{m-1}\frac{1}{n}\, x^{1/n-1} = r\, x^{r-1}, \quad x > 0.$$

§2. The Derivative of the Exponential and Logarithmic Functions

In this section we will find the derivatives of the exponential and logarithmic functions.

Fix $a \in (0, 1) \cup (1, \infty)$. Let $f(x) := a^x$ for $x \in (-\infty, \infty)$ and $g(x) := \log_a x$ for $x \in (0, \infty)$. First, we find the derivative of the logarithmic function g. Suppose $x > 0$. For $|h| < x$ we have

$$\log_a(x + h) - \log_a x = \log_a \frac{x + h}{x} = \log_a\Big(1 + \frac{h}{x}\Big).$$

Set $y := h/x$. Then $h = xy$ and

$$\frac{\log_a(x + h) - \log_a x}{h} = \frac{1}{xy} \log_a(1 + y) = \frac{1}{x} \log_a(1 + y)^{1/y}.$$

Clearly, $\lim_{h\to 0} y = \lim_{h\to 0}(h/x) = 0$. Hence

$$\lim_{h\to 0} \frac{\log_a(x + h) - \log_a x}{h} = \lim_{y\to 0} \frac{1}{x} \log_a(1 + y)^{1/y}.$$

We assert that $\lim_{y\to 0}(1 + y)^{1/y}$ exists as a positive real number. Assuming that our assertion is valid and $e := \lim_{y\to 0}(1 + y)^{1/y}$, we infer that

$$g'(x) = \lim_{h\to 0} \frac{\log_a(x + h) - \log_a x}{h} = \frac{\log_a e}{x}.$$

We write $\ln x$ for $\log_e x$ and call it the **natural logarithm** of x. Let $u(x) := e^x$ for $x \in \mathbb{R}$ and $v(x) := \ln x$ for $x \in (0, \infty)$. By what has been proved, $v'(x) = 1/x$ for $x \in (0, \infty)$. By the Inverse Function Theorem, u is differentiable on $\mathbb{R}$ and

$$u'(x) = u'(v(e^x)) = \frac{1}{v'(e^x)} = \frac{1}{1/e^x} = e^x, \quad x \in \mathbb{R}.$$

Note that $f(x) = a^x = (e^{\ln a})^x = e^{x \ln a}$. By the chain rule we obtain

$$f'(x) = e^{x \ln a} \ln a = a^x \ln a, \quad x \in \mathbb{R}.$$

For $\mu \in \mathbb{R}$, let q be the function given by $q(x) := x^\mu$ for $x > 0$. Then $q(x) = e^{\mu \ln x}$. By the chain rule we get

$$q'(x) = e^{\mu \ln x} \frac{\mu}{x} = x^\mu \frac{\mu}{x} = \mu x^{\mu - 1}, \quad x > 0.$$

In order to prove that $\lim_{y \to 0}(1 + y)^{1/y}$ exists, we first consider $\lim_{n\to\infty} s_n$, where $s_n := (1 + 1/n)^n$ for $n \in \mathbb{N}$.

By the Binomial Theorem we have

$$s_n = \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} \left(\frac{1}{n}\right)^k = \sum_{k=0}^{n} c_{n,k},$$

where $c_{n,k} := \binom{n}{k}(1/n)^k$. Clearly, $c_{n,0} = c_{n,1} = 1$. For $n \ge k \ge 2$ we have

$$\begin{aligned} c_{n,k} &= \frac{n!}{k!(n-k)!} \frac{1}{n^k} = \frac{1}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} \\ &= \frac{1}{k!} \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n} = \frac{1}{k!} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right). \end{aligned}$$

It follows that

$$c_{n+1,k} = \frac{1}{k!} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n+1}\right) > \frac{1}{k!} \prod_{j=1}^{k-1} \left(1 - \frac{j}{n}\right) = c_{n,k},$$

because $1 - j/(n+1) > 1 - j/n$ for $j = 1, \ldots, k-1$. Hence

$$s_{n+1} = \sum_{k=0}^{n+1} c_{n+1,k} > \sum_{k=0}^{n} c_{n+1,k} > \sum_{k=0}^{n} c_{n,k} = s_n.$$

This shows that $(s_n)_{n=1,2,\ldots}$ is an increasing sequence.

Next, we demonstrate that the sequence $(s_n)_{n=1,2,\ldots}$ is bounded. We have $c_{n,k} \le 1/k!$ for $n \ge k \ge 2$. Consequently,

$$s_n = \sum_{k=0}^{n} c_{n,k} \le 1 + 1 + \sum_{k=2}^{n} \frac{1}{k!} =: t_n.$$

We can use mathematical induction to prove that $k! \geq 2^{k-1}$ for all $k \geq 2$. It follows that

$$t_n \leq 2 + \sum_{k=2}^{n} \frac{1}{2^{k-1}} < 3.$$

Therefore, $s_n < 3$ for all $n \in \mathbb{N}$. Thus $\lim_{n\to\infty} s_n$ exists as a real number. Let e denote the limit.

Fix an integer $n \geq 2$. For $m > n$ we have

$$s_m = 2 + \sum_{k=2}^{m} \frac{1}{k!} \prod_{j=1}^{k-1} \Big(1 - \frac{j}{m}\Big) > 2 + \sum_{k=2}^{n} \frac{1}{k!} \prod_{j=1}^{k-1} \Big(1 - \frac{j}{m}\Big).$$

Letting $m \to \infty$ in the above inequality, we obtain $e \geq t_n$. Thus, $s_n \leq t_n \leq e$ for $n \geq 2$. By the squeeze theorem for sequences we get

$$e = \lim_{n\to\infty} t_n = 2 + \sum_{k=2}^{\infty} \frac{1}{k!}.$$

An easy calculation gives $e \approx 2.718281828459045$.

Since $\lim_{n\to\infty}(1 + 1/n)^n = e$, we have

$$\lim_{n\to\infty} \Big(1 + \frac{1}{n+1}\Big)^n = e \quad \text{and} \quad \lim_{n\to\infty} \Big(1 + \frac{1}{n}\Big)^{n+1} = e.$$

For given $\varepsilon > 0$, there exists some positive integer N such that

$$e - \varepsilon < \Big(1 + \frac{1}{n+1}\Big)^n < \Big(1 + \frac{1}{n}\Big)^{n+1} < e + \varepsilon \quad \forall n \geq N.$$

Choose $\delta := 1/N$. Suppose $0 < y < \delta$. Then $1/y \geq N$. Let n be the integer such that $n \leq 1/y < n+1$. It follows that $1/(n+1) < y \leq 1/n$. Clearly, $n \geq N$. Hence we have

$$e - \varepsilon < \Big(1 + \frac{1}{n+1}\Big)^n < (1+y)^{1/y} < \Big(1 + \frac{1}{n}\Big)^{n+1} < e + \varepsilon.$$

This shows $\lim_{y\to 0^+}(1+y)^{1/y} = e$. It remains to prove $\lim_{y\to 0^-}(1+y)^{1/y} = e$. For $-1 < y < 0$, set $z := -y/(1+y)$. Then $z > 0$ and $\lim_{y\to 0^-} z = 0$. Moreover, $z = -y/(1+y)$ implies $z(1+y) = -y$. So $y = -z/(1+z)$. Consequently,

$$\lim_{y\to 0^-} (1+y)^{1/y} = \lim_{z\to 0^+} (1+z)^{1+1/z} = \lim_{z\to 0^+} (1+z)(1+z)^{1/z} = e.$$

This completes the proof for $\lim_{y\to 0}(1+y)^{1/y} = e$.

§3. The Mean Value Theorem

Let f be a function from an interval I to $\mathbb{R}$, and let c be an interior point of I. We say that f has a **local maximum** (**local minimum**) at c, if there exists some $\delta > 0$ such that $f(x) \le f(c)$ ($f(x) \ge f(c)$) for all $x \in I \cap (c-\delta, c+\delta)$.

Theorem 3.1. *If f has either a local maximum or a local minimum at an interior point c of $I = [a,b]$ and if f is differentiable at c, then $f'(c) = 0$.*

Proof. Suppose that f has a local minimum at c. Then there exists some $\delta > 0$ such that $(c-\delta, c+\delta) \subset I$ and $f(x) \ge f(c)$ for all $x \in (c-\delta, c+\delta)$. Consequently, we have

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

Hence, $f'(c) = 0$. If f has a local maximum at c, the proof is similar. □

Theorem 3.2. *(Rolle's Theorem) Suppose that f is continuous on $[a,b]$ and is differentiable on (a,b). Suppose further that $f(a) = f(b)$. Then there exists at least one point c in (a,b) such that $f'(c) = 0$.*

Proof. If $f(x) = f(a)$ for all $x \in [a,b]$. then $f'(x) = 0$ for all $x \in [a,b]$, and the theorem is proved. Otherwise, f must have either a maximum value or a minimum value at some point $c \in (a,b)$. By Theorem 2.1, it follows that $f'(c) = 0$. □

Theorem 3.3. *(The Mean Value Theorem) Suppose that f is continuous on $[a,b]$ and 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. The line joining the points $(a, f(a))$ and $(b, f(b))$ has equation $y = m(x-a) + f(a)$, $x \in \mathbb{R}$, where $m := [f(b) - f(a)]/(b-a)$. Let $g(x) := f(x) - [m(x-a) + f(a)]$, $a \le x \le b$. Then g is continuous on $[a,b]$ and g is differentiable on (a,b) with $g'(x) = f'(x) - m$. Note that $g(a) = g(b) = 0$. By Rolle's theorem, there exists some $c \in (a,b)$ such that $g'(c) = 0$. For this c we have $f'(c) = m = [f(b) - f(a)]/(b-a)$. □

Theorem 3.4. *(The Generalized Mean Value Theorem) Let f and g be two functions each of which is continuous on $[a,b]$ and differentiable on (a,b). Then there exists a point $c \in (a,b)$ such that*

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c).$$

Proof. Let h be the function given by

$$h(x) := [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x), \quad x \in [a, b].$$

Then h is continuous on $[a, b]$ and differentiable on (a, b). By Rolle's theorem, $h'(c) = 0$ for some $c \in (a, b)$. This completes the proof of the theorem. □

§4. Applications of the Mean Value Theorem

The following theorem is an application of the mean value theorem to the study of monotone functions. Given an interval I in $\mathbb{R}$, recall that I° is the set of all interior points of I.

Theorem 4.1. *Let f be a real-valued function on an interval I in $\mathbb{R}$. Suppose that f is continuous on I and differentiable on I°. Then the following statements are true:*
(1) *If $f'(x) > 0$ for all $x \in I^\circ$, then f is strictly increasing on I.*
(2) *If $f'(x) < 0$ for all $x \in I^\circ$, then f is strictly decreasing on I.*
(3) *If $f'(x) \geq 0$ for all $x \in I^\circ$, then f is increasing on I.*
(4) *If $f'(x) \leq 0$ for all $x \in I^\circ$, then f is decreasing on I.*
(5) *If $f'(x) = 0$ for all $x \in I^\circ$, then f is constant on I.*

Proof. Let us prove (1). Consider $x_1, x_2 \in I$ with $x_1 < x_2$. By the mean value theorem, there exists some $c \in (x_1, x_2)$ such that $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$. Clearly, $c \in I^\circ$ and hence $f'(c) > 0$ by the assumption. It follows that $f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0$. This shows that f is strictly increasing on I.

Parts (2), (3), and (4) can be proved by using similar arguments. Finally, (5) follows immediately from parts (3) and (4). □

The following example illustrates an application of Theorem 4.1.

Example 1. Let $f(x) := x^3/(1 - x^2)$ for $x \in \mathbb{R} \setminus \{-1, 1\}$. Determine the intervals where f is monotone.

Solution. For $x \in \mathbb{R} \setminus \{-1, 1\}$ we have

$$f'(x) = \frac{(3x^2)(1 - x^2) - x^3(-2x)}{(1 - x^2)^2} = \frac{x^2(3 - x^2)}{(1 - x^2)^2}.$$

Hence $f'(x) < 0$ for $|x| > \sqrt{3}$ and $f'(x) > 0$ for $x \in (-\sqrt{3}, \sqrt{3}) \setminus \{-1, 1\}$. Thus, the function is strictly decreasing on $(-\infty, -\sqrt{3}]$ and $[\sqrt{3}, \infty)$, and strictly increasing on $[-\sqrt{3}, -1)$, $(-1, 1)$ and $(1, \sqrt{3}]$.

The mean value theorem is useful for proving certain inequalities.

Example 2. Prove the following inequality:

$$\frac{x}{1+x} \le \ln(1+x) \le x \quad \text{for all } x > -1.$$

Proof. Let $f(x) := x - \ln(1+x)$, $x > -1$. We have

$$f'(x) = 1 - \frac{1}{1+x} = \frac{x}{1+x}, \quad x > -1.$$

Hence, $f'(x) > 0$ for $x > 0$ and $f'(x) < 0$ for $x < 0$. This shows that f is strictly decreasing on $(-1, 0)$ and is strictly increasing on $(0, \infty)$. Therefore, $f(x) \ge f(0) = 0$ for $x > -1$, that is, $\ln(1+x) \le x$ for $x > -1$.

Let $g(x) := \ln(1+x) - x/(1+x)$, $x > -1$. We have

$$g'(x) = \frac{1}{1+x} - \frac{1}{(1+x)^2} = \frac{x}{(1+x)^2}.$$

Hence, $g'(x) < 0$ for $x \in (-1, 0)$ and $g'(x) > 0$ for $x \in (0, \infty)$. This shows that g is strictly decreasing on $(-1, 0)$ and is strictly increasing on $(0, \infty)$. Therefore, $g(x) \ge g(0) = 0$ for $x > -1$, that is, $x/(1+x) \le \ln(1+x)$ for $x > -1$. □

The following example generalizes the Bernoulli inequality.

Example 3. Let $\mu > 1$. Prove that $(1+x)^\mu \ge 1 + \mu x$ for all $x > -1$.

Proof. Let $f(x) := (1+x)^\mu - (1 + \mu x)$ for $x > -1$. Then

$$f'(x) = \mu(1+x)^{\mu-1} - \mu = \mu[(1+x)^{\mu-1} - 1].$$

Since $\mu > 1$, $(1+x)^{\mu-1} < 1$ for $-1 < x < 0$ and $(1+x)^{\mu-1} > 1$ for $x > 0$. Thus, $f'(x) < 0$ for $-1 < x < 0$ and $f'(x) > 0$ for $x > 0$. This shows that f is decreasing on $(-1, 0]$ and increasing on $[0, \infty)$. Therefore, for all $x > -1$, $f(x) \ge f(0)$, that is, $(1+x)^\mu \ge 1 + \mu x$.

As an application of the generalized Bernoulli inequality, we study the following limit:

$$\lim_{x\to\infty} \frac{x^\alpha}{a^x},$$

where $a > 1$ and $\alpha \in \mathbb{R}$. First, consider the case $\alpha < 1$. Let $b := a - 1 > 0$. The Bernoulli's inequality tells us that $a^x = (1+b)^x \ge 1 + bx$ for $x > 1$. Hence

$$0 < \frac{x^\alpha}{a^x} \le \frac{x^\alpha}{bx} = \frac{1}{bx^{1-\alpha}}, \quad x > 1.$$

Since $1-\alpha>0$, we have $\lim_{x\to\infty} 1/(bx^{1-\alpha})=0$. By the squeeze theorem for limits, we get $\lim_{x\to\infty} x^\alpha/a^x=0$. Next, consider the case $\alpha\ge 1$. Choose a positive integer $m>\alpha$. Then

$$\frac{x^\alpha}{a^x}=\left[\frac{x^{\alpha/m}}{(a^{1/m})^x}\right]^m.$$

Now we have $\alpha/m<1$ and $a^{1/m}>1$. Therefore,

$$\lim_{x\to\infty}\frac{x^{\alpha/m}}{(a^{1/m})^x}=0 \quad\text{and}\quad \lim_{x\to\infty}\frac{x^\alpha}{a^x}=0.$$

Setting $x=\log_a y$ in the above limit, we obtain

$$\lim_{y\to\infty}\frac{(\log_a y)^\alpha}{y}=0,$$

provided that $a>1$ and $\alpha\in\mathbb{R}$.

Let f be a continuous function from an interval I to $\mathbb{R}$. If f is differentiable on I° and there is a constant M such that $|f'(x)|\le M$ for all $x\in I^\circ$, then the mean value theorem gives

$$|f(x_1)-f(x_2)|\le M|x_1-x_2| \quad \forall\, x_1,x_2\in I.$$

Thus f is a Lipschitz function on I. In particular, f is uniformly continuous on I.

Example 4. Let $f(x)=\ln x$, $x\in(0,\infty)$. For a fixed $a>0$, prove that f is uniformly continuous on $[a,\infty)$.

Proof. For $x\ge a$ we have

$$|f'(x)|=\left|\frac{1}{x}\right|\le\frac{1}{a}.$$

By the mean value theorem, f is uniformly continuous on $[a,\infty)$.

§5. Taylor's Theorem

Let f be a real-valued function defined on an interval I in $\mathbb{R}$. If f is differentiable on I, then the derivative $f': x\mapsto f'(x)$ is also a function on I. If $c\in I$ and f' is differentiable at c, then the derivative of f' at c, denoted by $f''(c)$ or $f^{(2)}(c)$, is called the **second derivative** of f at c, and f is said to be **twice differentiable** at c. More generally, for $n\in\mathbb{N}$, if $f^{(n-1)}$ exists on I, and if $f^{(n-1)}$ is differentiable at c, then the derivative of $f^{(n-1)}$ at c, denoted by $f^{(n)}(c)$, is called the nth derivative of f at c, and f is said to be n-times differentiable at c. If f is n-times differentiable at every point of I, then we say that f is n-times differentiable on I.

Example 1. Let f be the function on $\mathbb{R}$ given by $f(x) := (x-a)^n$ for $x \in \mathbb{R}$, where $n \in \mathbb{N}_0$ and $a \in \mathbb{R}$ is a constant. For $k \in \mathbb{N}$, find $f^{(k)}$ and $f^{(k)}(a)$.

Solution. For $n \geq 2$ we have

$$f'(x) = n(x-a)^{n-1} \quad \text{and} \quad f''(x) = n(n-1)(x-a)^{n-2}.$$

More generally, for $k \leq n$ we have

$$f^{(k)}(x) = n(n-1)\cdots(n-k+1)(x-a)^{n-k}, \quad x \in \mathbb{R}.$$

Note that $f^{(n)}(x) = n!$ for $x \in \mathbb{R}$. So $f^{(n)}$ is a constant. In particular, $f^{(n)}(a) = n!$. Moreover, for $k > n$ we have $f^{(k)} = 0$ and $f^{(k)}(a) = 0$. If $k < n$, then $n - k \geq 1$, and hence $(x-a)^{n-k}$ vanishes when $x = a$. Therefore, $f^{(k)}(a) = 0$ for $k < n$.

Example 2. Let g be a function from an interval I to $\mathbb{R}$. Suppose that g is n-times differentiable on I, and that $g^{(n)}$ is differentiable on the interior of I. Let a and b be two distinct points in I. If $g^{(k)}(a) = 0$ for $k = 0, 1, \ldots, n$ and $g(b) = 0$, then there exists some ξ between a and b such that $g^{(n+1)}(\xi) = 0$.

Proof. For $k \in \mathbb{N}$ let P_k be the statement "either $k > n+1$ or there exists some ξ between a and b such that $f^{(k)}(\xi) = 0$". We shall use mathematical induction to prove that P_k is true for all $k \in \mathbb{N}$. For $k = 1$, since $g(a) = g(b) = 0$, by Rolle's theorem there exists some ξ between a and b such that $g'(\xi) = 0$. This verifies the base case. For the induction step, assuming that P_k is true, we wish to prove that P_{k+1} is true. If $k > n$, then $k + 1 > n + 1$; hence P_{k+1} is true. Let us consider the case $k \leq n$. By the induction hypothesis, $g^{(k)}(\eta) = 0$ for some η between a and b. But $g^{(k)}(a) = 0$. Applying Rolle's theorem to the function $g^{(k)}$, we see that there exists some ξ between a and η such that $(g^{(k)})'(\xi) = 0$. In other words, $g^{(k+1)}(\xi) = 0$. Now η is between a and b, and ξ is between a and η. We infer that ξ is between a and b and thereby complete the induction step. Consequently, P_{n+1} is true. This is the desired result. □

Let f be a function from an interval I to $\mathbb{R}$. Suppose that f is n-times differentiable on I. Given an interior point a of I, we wish to find a polynomial of degree at most n such that $p(a) = f(a), p'(a) = f'(a), \ldots, p^{(n)}(a) = f^{(n)}(a)$. We may express p in the following form:

$$p(t) = \sum_{k=0}^{n} c_k (t-a)^k, \quad t \in \mathbb{R}.$$

By Example 1 we have $p^{(k)}(a) = c_k k!$. Thus, $p^{(k)}(a) = f^{(k)}(a)$ if and only if $c_k = f^{(k)}(a)/k!$, $k = 0, 1, \ldots, n$. We write

$$T_n(f,a)(t) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(t-a)^k, \quad t \in \mathbb{R},$$

and call $T_n(f,a)$ the nth **Taylor polynomial** of f at a.

Theorem 5.1. *Let f be a function from an interval I to $\mathbb{R}$. Suppose that f is n-times differentiable on I for some $n \in \mathbb{N}_0$, and that $f^{(n)}$ is differentiable on the interior of I. For an interior point a of I, let $p_n := T_n(f,a)$ be the nth Taylor polynomial of f at a. Then for each $x \in I$, there exists some ξ between a and x such that*

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

Proof. We have $f(a) = p_n(a)$. Hence, we may assume $x \neq a$ in what follows. Let

$$g(t) := f(t) - p_n(t) - r(t-a)^{n+1}, \quad t \in I,$$

where r is so chosen that $g(x) = 0$. In other words, $f(x) - p_n(x) = r(x-a)^{n+1}$. We observe that the derivatives $g^{(k)}$ exist on I for $k = 0, 1, \ldots, n$. Moreover, $g^{(k)}(a) = 0$ for $k = 0, 1, \ldots, n$. By Example 2, there exists some ξ between a and x such that $g^{(n+1)}(\xi) = 0$. On the other hand, $g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (n+1)!r$. Hence we have

$$f^{(n+1)}(\xi) - (n+1)!r = 0.$$

It follows that $r = f^{(n+1)}(\xi)/(n+1)!$. Therefore,

$$f(x) = p_n(x) + r(x-a)^{n+1} = p_n(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

This completes the proof. □

Let $R_n(f,a) := f - T_n(f,a)$. Then $R_n(f,a)$ is called the **remainder** between f and $T_n(f,a)$. The above theorem shows that there exists some ξ between a and x such that

$$R_n(f,a) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

This is called the **Lagrange form of the remainder**.

Example 3. Let f be the function given by $f(x) = \sqrt{1+x}$ for $x \in (-1, \infty)$. Find its second Taylor polynomial at $a = 0$ and the corresponding Lagrange form of the remainder.

Solution. We have

$$f'(x) = \frac{1}{2}(1+x)^{-1/2}, \quad f''(x) = -\frac{1}{4}(1+x)^{-3/2}, \quad f'''(x) = \frac{3}{8}(1+x)^{-5/2}.$$

It follows that $f(0) = 1$, $f'(0) = 1/2$, and $f''(0) = -1/4$. Hence

$$\sqrt{1+x} = T_2(f,0)(x) + R_2(f,0)(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + R_2(f,0)(x),$$

where

$$R_2(f,0)(x) = \frac{f'''(\xi)}{3!}x^3 = \frac{1}{16}(1+\xi)^{-5/2}x^3$$

for some ξ between 0 and x. □

Now let f be an infinitely differentiable real-valued function on an interval I. The series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k,$$

as a function of x on I, is called the **Taylor series** of f about a. This series converges to $f(x)$ if and only if $\lim_{n\to\infty} R_n(f,a)(x) = 0$.

Let $f(x) := e^x$ for $x \in \mathbb{R}$. Then $f^{(k)}(x) = e^x$ for all $k \in \mathbb{N}$ and $x \in \mathbb{R}$. Consequently,

$$T_n(f,0)(x) = \sum_{k=0}^{n} \frac{x^k}{k!}, \quad x \in \mathbb{R},$$

and

$$R_n(f,0)(x) = \frac{e^{\xi}}{(n+1)!}x^{n+1},$$

where ξ is a real number between 0 and x. Suppose $M > 0$. For $x \in [-M, M]$ we have

$$|R_n(f,0)(x)| \le e^M \frac{M^{n+1}}{(n+1)!} \quad \text{and} \quad \lim_{n\to\infty} \frac{M^{n+1}}{(n+1)!} = 0.$$

Hence, the sequence $(T_n(f,0)(x))_{n=1,2,\ldots}$ converges to $f(x)$ for each $x \in \mathbb{R}$. Consequently,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad x \in \mathbb{R}.$$

Example 4. Let g be the function on $\mathbb{R}$ given by $g(x) = e^{-1/x}$ for $x > 0$ and $g(x) = 0$ for $x \le 0$. Clearly g is infinitely differentiable at any point in $\mathbb{R} \setminus \{0\}$. Moreover, $g^{(n)}(0) = 0$ for all $n \in \mathbb{N}_0$. Hence the Taylor series of g about 0 is identically zero, so g does not agree with its Taylor series in any open interval containing 0.

§6. Power Series

A **power series** in x about a is an infinite series of the form

$$\sum_{n=0}^{\infty} c_n (x-a)^n,$$

where $a \in \mathbb{R}$ and $c_n \in \mathbb{R}$ for $n \in \mathbb{N}_0$. The main purpose of this section is to study convergence of the power series.

Suppose that the power series $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges for some $x_0 \neq a$. Then it converges absolutely for all $x \in \mathbb{R}$ satisfying $|x-a| < |x_0 - a|$. Let us verify this assertion. Since the series $\sum_{n=0}^{\infty} c_n (x_0-a)^n$ converges, the sequence $(c_n (x_0-a)^n)_{n=0,1,\ldots}$ converges to 0. So there is a positive number M such that $|c_n(x_0-a)^n| \le M$ for all $n \in \mathbb{N}_0$. Then we have

$$\left|c_n(x-a)^n\right| = \left|c_n(x_0-a)^n\right|\left|(x-a)^n/(x_0-a)^n\right| \le M r^n,$$

where $r := |x-a|/|x_0-a|$. Since $|x-a| < |x_0-a|$, we have $0 \le r < 1$, and hence the geometric series $\sum_{n=0}^{\infty} M r^n$ converges. By the comparison test for series we see that the series $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges absolutely.

Theorem 6.1. *Given a power series $\sum_{n=0}^{\infty} c_n (x-a)^n$, there is $R \in [0,\infty)$ or $R = \infty$ with the following properties: (1) the power series $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges for all $x \in \mathbb{R}$ with $|x-a| < R$; (2) the power series diverges for all $x \in \mathbb{R}$ with $|x-a| > R$.*

Proof. Let S be the set of those $x \in \mathbb{R}$ for which the power series $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges. Since $a \in S$, S is nonempty. Let $R := \sup\{|x-a| : x \in S\}$. If $R = 0$, then $\sum_{n=0}^{\infty} c_n (x-a)^n$ diverges whenever $x \neq a$. If $0 < R < \infty$, then $|x-a| > R$ implies $x \notin S$; hence $\sum_{n=0}^{\infty} c_n (x-a)^n$ diverges. Now suppose that $|x-a| < R$, where $0 < R \le \infty$. By the definition of R, there exists some $x_0 \in S$ such that $|x-a| < |x_0-a|$. Thus $\sum_{n=0}^{\infty} c_n (x_0-a)^n$ converges. Therefore the series $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges. □

The extended real number $R \in [0,\infty]$ in the above theorem is called the **radius of convergence** of the power series $\sum_{n=0}^{\infty} c_n (x-a)^n$. From the proof of the above theorem we see that $(a-R, a+R) \subseteq S \subseteq [a-R, a+R]$. Hence S is an interval. It is called the **interval of convergence** of the power series. If $R = 0$, the interval of convergence is the degenerated interval $\{a\}$. If $R = \infty$, the interval of convergence is $(-\infty, \infty)$.

Example 1. Consider the following three power series:

$$\sum_{n=0}^{\infty} n! x^n, \quad \sum_{n=0}^{\infty} x^n, \quad \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

By using the ratio test we see that the series $\sum_{n=0}^{\infty} n!x^n$ diverges for any $x \neq 0$. So its radius of convergence is $R = 0$. The series $\sum_{n=0}^{\infty} x^n$ is a geometric series. It converges if and only if $-1 < x < 1$; hence its radius of convergence is $R = 1$. Finally, the power series $\sum_{n=0}^{\infty} x^n/n!$ converges for all $x \in \mathbb{R}$ and its radius of convergence is $R = \infty$.

Example 2. Determine the interval of convergence of the following power series:

$$\sum_{n=0}^{\infty} \frac{1}{3^n(n+1)}(x+2)^n.$$

Solution. Let $u_n := (x+2)^n/(3^n(n+1))$ for $n \in \mathbb{N}_0$. For $x \neq -2$ we have

$$\left|\frac{u_{n+1}}{u_n}\right| = \frac{|x+2|^{n+1}}{3^{n+1}(n+2)} \frac{3^n(n+1)}{|x+2|^n} = \frac{|x+2|}{3} \frac{n+1}{n+2}.$$

It follows that

$$\lim_{n\to\infty}\left|\frac{u_{n+1}}{u_n}\right| = \frac{|x+2|}{3}.$$

By the ratio test, the power series converges if $|x+2| < 3$ and diverges if $|x+2| > 3$. So its radius of convergence is $R = 3$. We observe that $|x+2| < 3$ if and only if $-3 < x+2 < 3$, which is equivalent to $-5 < x < 1$. The end points of the interval $(-5, 1)$ are -5 and 1. If $x = -5$, the series

$$\sum_{n=0}^{\infty} \frac{1}{3^n(n+1)}(-5+2)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}$$

is convergent, by the alternating series test. If $x = 1$, the series

$$\sum_{n=0}^{\infty} \frac{1}{3^n(n+1)}(1+2)^n = \sum_{n=0}^{\infty} \frac{1}{n+1}$$

is the harmonic series. So it diverges. We conclude that the interval of convergence of the power series is $[-5, 1)$.

Term-by-term differentiation of a power series is valid inside its interval of convergence.

Theorem 6.2. *Suppose that the power series $\sum_{n=0}^{\infty} c_n(x-a)^n$ has radius of convergence $R > 0$. For $x \in (a-R, a+R)$, let $f(x)$ be the sum of the series. Then f is differentiable on $(a-R, a+R)$ and*

$$f'(x) = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1} \quad \forall\, x \in (a-R, a+R).$$

Proof. First, we prove that the power series $\sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$ converges absolutely for all $x \in (a-R, a+R)$. For this purpose we fix a real number $x \in (a-R, a+R)$. Choose x_0

such that $|x-a| < x_0 - a < R$. By our assumption, the series $\sum_{n=0}^{\infty} c_n(x_0-a)^n$ converges. Hence the sequence $(c_n(x_0-a)^n)_{n=0,1,\dots}$ converges to 0. So there is a positive number M such that $|c_n(x_0-a)^{n-1}| \le M$ for all $n \in \mathbb{N}$. It follows that

$$\left|nc_n(x-a)^{n-1}\right| = \left|c_n(x_0-a)^{n-1}\right| n|x-a|^{n-1}/(x_0-a)^{n-1} \le Mnr^{n-1},$$

where $r := |x-a|/|x_0-a| < 1$. Thus the series $\sum_{n=1}^{\infty} Mnr^{n-1}$ converges. So the series $\sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$ converges absolutely, by the comparison test. Applying term-by-term differentiation to the series $\sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$, we see that the power series $\sum_{n=2}^{\infty} n(n-1)c_n(x-a)^{n-2}$ converges absolutely for all $x \in (a-R, a+R)$.

Next, we show that $f'(x) = g(x)$ for $x \in (a-R, a+R)$, where $g(x)$ is the sum of the series $\sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$. Let $h := (x_0-a) - |x-a|$. Then $h > 0$. For $0 < |t| < h$ we have

$$\frac{f(x+t)-f(x)}{t} - g(x) = \sum_{n=1}^{\infty} c_n \left[\frac{(x-a+t)^n - (x-a)^n}{t} - n(x-a)^{n-1}\right].$$

Let $u_n(t) := (x-a+t)^n - (x-a)^n$, $t \in \mathbb{R}$. For $n=1$ we have $u_1(t) = t$. For $n \ge 2$, by the Taylor theorem we get $u_n(t) = u_n(0) + u_n'(0)t + u_n''(\xi)t^2$ for some ξ between 0 and t. Consequently,

$$\frac{(x-a+t)^n - (x-a)^n}{t} - n(x-a)^{n-1} = \frac{u_n(t)-u_n(0)}{t} - u_n'(0) = tn(n-1)(x-a+\xi)^{n-2}.$$

We have

$$|x-a+\xi| \le |x-a| + |\xi| \le |x-a| + |t| < |x_0-a|.$$

It follows that

$$\left|\frac{f(x+t)-f(x)}{t} - g(x)\right| \le |t| \sum_{n=2}^{\infty} |c_n| n(n-1)|x_0-a|^{n-2}.$$

But the series $\sum_{n=2}^{\infty} |c_n| n(n-1)|x_0-a|^{n-2}$ converges and its sum is a constant independent of t. Therefore,

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

This shows that $f'(x) = g(x) = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$ for $x \in (a-R, a+R)$. □

Example 3. The power series $\sum_{n=0}^{\infty} x^n$ is a geometric series. We have

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad -1 < x < 1.$$

Differentiating the above power series term-by-term, we obtain

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}, \quad -1 < x < 1.$$

Suppose that the power series $\sum_{n=0}^{\infty} c_n(x-a)^n$ has radius of convergence $R > 0$. For $x \in (a-R, a+R)$, let $f(x)$ be the sum of the series. For $k \in \mathbb{N}$, differentiating the power series term-by-term k times, we get

$$f^{(k)}(x) = \sum_{n=k}^{\infty} c_n n(n-1)\cdots(n-k+1)(x-a)^{n-k}, \quad x \in (a-R, a+R).$$

Substituting a for x in the above equation, we obtain $f^{(k)}(a) = c_k k!$. Therefore,

$$c_k = \frac{f^{(k)}(a)}{k!}, \quad k = 0, 1, 2, \ldots.$$

Thus, $\sum_{n=0}^{\infty} c_n(x-a)^n$ is the Taylor series of f about a.

Example 4. Let $f(x) = \ln(1+x)$ for $x > -1$. Find the Taylor series of f about 0.

Solution. Let $g(x) := f'(x) = 1/(1+x)$ for $x > -1$. We have

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

For $-1 < x < 1$, let $h(x)$ be the sum of the power series $\sum_{n=0}^{\infty}(-1)^n x^{n+1}/(n+1)$. By Theorem 6.2, $h'(x) = g(x)$ for $x \in (-1, 1)$. On the other hand, $f'(x) = g(x)$ for $x \in (-1, 1)$. Hence, $f'(x) - h'(x) = 0$ for all $x \in (-1, 1)$. Consequently, $f - h$ is a constant on $(-1, 1)$. But $f(0) = 0$ and $h(0) = 0$. Therefore, $f(x) = h(x)$ for all $x \in (-1, 1)$. This shows that

$$\ln(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n, \quad x \in (-1, 1).$$

Note that the convergence of interval of the above power series is $(-1, 1]$. But the convergence of interval of the power series $\sum_{n=0}^{\infty}(-1)^n x^n$ is $(-1, 1)$.

§7. Length of Curves

In this section we study lengths of curves in the Euclidean plane.

We use $\mathbb{R}^2$ to denote the set of ordered pairs (x_1, x_2) of real numbers. For two points $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in $\mathbb{R}^2$, define

$$\rho(x, y) := \sqrt{(x_1 - y_1)^2 + (y_1 - y_2)^2}.$$

Then $\rho(x, y)$ represents the distance between x and y. We call ρ a **metric** on $\mathbb{R}^2$. The Euclidean plane is the set $\mathbb{R}^2$ equipped with the metric ρ. The metric ρ satisfies the following properties for $x, y, z \in \mathbb{R}^2$:

(1) $\rho(x, y) \geq 0$, and $\rho(x, y) = 0$ if and only if $x = y$,

(2) $\rho(x, y) = \rho(y, x)$, and

(3) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

The third property is called the **triangle inequality**.

Let u be a mapping from an interval I in $\mathbb{R}$ to $\mathbb{R}^2$. We say that u is continuous on I, if for every $a \in I$,

$$\lim_{t \to a} \rho(f(t), f(a)) = 0.$$

A curve in the Euclidean plane $\mathbb{R}^2$ is represented by a continuous mapping u from a closed interval $[a, b]$ to $\mathbb{R}^2$. Suppose $u(t) = (u_1(t), u_2(t))$ for $t \in [a, b]$, where u_1 and u_2 are real-valued continuous functions on $[a, b]$. Then u is a continuous mapping from $[a, b]$ to $\mathbb{R}^2$.

By a **partition** P of $[a, b]$ we mean a finite ordered set $\{t_0, t_1, \ldots, t_n\}$ such that

$$a = t_0 < t_1 < \cdots < t_n = b.$$

Let $P := \{t_0, t_1, \ldots, t_n\}$ be a partition of $[a, b]$. For $j \in \{1, \ldots, n\}$, the length of the line segment connecting two points $u(t_{j-1})$ and $u(t_j)$ is

$$\sqrt{[u_1(t_j) - u_1(t_{j-1})]^2 + [u_2(t_j) - u_2(t_{j-1})]^2}.$$

Let $L(u, P)$ denote the sum of the lengths of the line segments connecting $u(t_{j-1})$ and $u(t_j)$ for $j = 1, \ldots, n$. Then

$$L(u, P) = \sum_{j=1}^{n} \sqrt{[u_1(t_j) - u_1(t_{j-1})]^2 + [u_2(t_j) - u_2(t_{j-1})]^2}.$$

The **length** of the curve u is defined to be

$$L(u) := \sup\{L(u, P) : P \text{ is a partition of } [a, b]\}.$$

If $L(u) < \infty$, then u is said to be **rectifiable**.

For $a \leq c \leq d \leq b$, we use $u|_{[c,d]}$ to denote the restriction of u to the interval $[c, d]$.

Theorem 7.1. *Let $u = (u_1, u_2)$ be a continuous mapping from $[a, b]$ to $\mathbb{R}^2$. If u_1' and u_2' are continuous on $[a, b]$, then u is rectifiable and the function s given by $s(t) := L(u|_{[a,t]})$ for $a \le t \le b$ has the following property:*

$$s'(t) = \sqrt{[u_1'(t)]^2 + [u_2'(t)]^2}, \quad t \in [a, b].$$

Proof. Suppose $a \le c < d \le b$. For $k = 1, 2$, let

$$m_k := \inf\{|u_k'(t)| : t \in [c, d]\} \quad \text{and} \quad M_k := \sup\{|u_k'(t)| : t \in [c, d]\}.$$

Let $P = \{t_0, t_1, \ldots, t_n\}$ be a partition of $[c, d]$. By the mean value theorem, for each $j \in \{1, \ldots, n\}$ there exist some ξ_j and η_j in (t_{j-1}, t_j) such that

$$u_1(t_j) - u_1(t_{j-1}) = u_1'(\xi_j)(t_j - t_{j-1}) \quad \text{and} \quad u_2(t_j) - u_2(t_{j-1}) = u_2'(\eta_j)(t_j - t_{j-1}).$$

It follows that

$$m_k(t_j - t_{j-1}) \le |u_k(t_j) - u_k(t_{j-1})| \le M_k(t_j - t_{j-1}), \quad k = 1, 2.$$

Consequently, with $m := \sqrt{m_1^2 + m_2^2}$ and $M := \sqrt{M_1^2 + M_2^2}$ we have

$$\sum_{j=1}^n m(t_j - t_{j-1}) \le \sum_{j=1}^n \sqrt{[u_1(t_j) - u_1(t_{j-1})]^2 + [u_2(t_j) - u_2(t_{j-1})]^2} \le \sum_{j=1}^n M(t_j - t_{j-1}).$$

Hence, $m(d - c) \le L(u|_{[c,d]}, P) \le M(d - c)$. This is true for every partition P of $[c, d]$. Therefore,

$$m(d - c) \le L(u|_{[c,d]}) \le M(d - c).$$

In particular, u is rectifiable.

Now suppose $t, t+h \in [a, b]$. For $k = 1, 2$, let $m_{k,h}$ ($M_{k,h}$) be the infimum (supremum) of the function $|u_k'|$ on the interval with t and $t + h$ as the end points. Let

$$m_h := \sqrt{m_{1,h}^2 + m_{2,h}^2} \quad \text{and} \quad M_h := \sqrt{M_{1,h}^2 + M_{2,h}^2}.$$

We have $s(t+h) - s(t) = L(u_{[t,t+h]})$ for $h > 0$ and $s(t+h) - s(t) = -L(u_{[t+h,t]})$ for $h < 0$. Thus, by what has been proved we obtain

$$m_h \le \frac{s(t + h) - s(t)}{h} \le M_h, \quad h \ne 0.$$

Since u_1' and u_2' are continuous on $[a, b]$,

$$\lim_{h\to 0} m_h = \lim_{h\to 0} M_h = \sqrt{[u_1'(t)]^2 + [u_2'(t)]^2}.$$

Consequently,

$$s'(t) = \lim_{h\to 0} \frac{s(t+h)-s(t)}{h} = \sqrt{[u_1'(t)]^2 + [u_2'(t)]^2}, \quad t \in [a,b].$$

This completes the proof of the theorem. □

Let us consider the following example: $\gamma(t) = (\gamma_1(t), \gamma_2(t))$, where

$$\gamma_1(t) := \frac{1-t^2}{1+t^2} \quad \text{and} \quad \gamma_2(t) = \frac{2t}{1+t^2}, \quad 0 \le t \le 1.$$

We have

$$\gamma_1'(t) = \frac{-4t}{(1+t^2)^2} \quad \text{and} \quad \gamma_2'(t) = \frac{2(1-t^2)}{(1+t^2)^2}, \quad 0 \le t \le 1.$$

Clearly, $\gamma_1'(t) < 0$ and $\gamma_2'(t) > 0$ for $0 < t < 1$. Hence, γ_1 is strictly decreasing and γ_2 is strictly increasing on $[0,1]$. Thus, γ is a one-to-one and onto mapping from $[0,1]$ to $\{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1, x_1 \ge 0, x_2 \ge 0\}$, which is the part of the unit circle in the first quadrant. For $0 \le t \le 1$, let $s(t) := L(\gamma|_{[0,t]})$. Then

$$s'(t) = \sqrt{[\gamma_1'(t)]^2 + [\gamma_2'(t)]^2} = \frac{2}{1+t^2}, \quad t \in [0,1].$$

For $0 \le t \le 1$ and $n \in \mathbb{N}$ we observe that

$$\frac{1}{1+t^2} = \sum_{k=0}^{n} (-t^2)^k + \frac{(-t^2)^{n+1}}{1+t^2}.$$

This motivates us to introduce the function

$$r_n(t) := s(t) - \sum_{k=0}^{n} \frac{2(-1)^k t^{2k+1}}{2k+1}, \quad 0 \le t \le 1.$$

Clearly, $r_n(0) = 0$. Moreover,

$$r_n'(t) = s'(t) - 2\sum_{k=0}^{n} (-1)^k t^{2k} = \frac{2(-t^2)^{n+1}}{1+t^2}, \quad 0 \le t \le 1.$$

By the mean value theorem we have

$$|r_n(t)| = |r_n(t) - r_n(0)| \le \sup\{|r_n'(\tau)| : \tau \in [0,t]\} \le 2t^{2n+2}, \quad 0 \le t \le 1.$$

It follows that $\lim_{n\to\infty} r_n(t) = 0$ for $0 \le t < 1$. Consequently,

$$s(t) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{2(-1)^k t^{2k+1}}{2k+1} = \sum_{k=0}^{\infty} \frac{2(-1)^k t^{2k+1}}{2k+1}, \quad 0 \le t < 1.$$

Furthermore,

$$\sum_{k=0}^{\infty} \frac{2(-1)^k}{2k+1} - s(t) = \sum_{k=0}^{\infty} \frac{2(-1)^k(1-t^{2k+1})}{2k+1} = 2(1-t)\sum_{k=0}^{\infty}(-1)^k b_k(t), \quad 0 \le t < 1$$

where $b_k(t) := (1+t+\cdots+t^{2k})/(2k+1)$. In particular, $b_0(t) = 1$. For $0 \le t < 1$, the series $\sum_{k=0}^{\infty}(-1)^k b_k(t)$ is an alternating series with $b_k(t) \ge b_{k+1}(t)$ for all $k \in \mathbb{N}_0$. Indeed $b_k(t) \ge b_{k+1}(t)$ holds if

$$(2k+3)(1+t+\cdots+t^{2k}) - (2k+1)(1+t+\cdots+t^{2k}+t^{2k+1}+t^{2k+2}) \ge 0.$$

Let w denote the left side of the above inequality. Then for $0 \le t < 1$ we have

$$w = 2(1+t+\cdots+t^{2k}) - (2k+1)(t^{2k+1}+t^{2k+2}) = \sum_{j=0}^{2k}\left[2t^j - (t^{2k+1}+t^{2k+2})\right] \ge 0.$$

This verifies $b_k(t) \ge b_{k+1}(t)$ for $0 \le t < 1$ and $k \in \mathbb{N}_0$. It follows that

$$\sum_{k=0}^{\infty}(-1)^k b_k(t) \le b_0(t) = 1.$$

Consequently,

$$0 \le \sum_{k=0}^{\infty} \frac{2(-1)^k}{2k+1} - s(t) \le 2(1-t), \quad 0 \le t < 1.$$

Finally, by the squeeze theorem for limits, we obtain

$$s(1) = \lim_{t \to 1^-} s(t) = \sum_{k=0}^{\infty} \frac{2(-1)^k}{2k+1}.$$

Let π be the perimeter of the half unit circle. Then

$$\frac{\pi}{4} = \frac{s(1)}{2} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

The above series gives $\pi \approx 3.141592653589793$.

§8. Trigonometric Functions

In this section we introduce trigonometric functions and investigate their properties.

Given two distinct points A and B in the Euclidean plane, the **ray** $\overrightarrow{AB}$ is the set consisting of A together with all points on the line AB that are on the same side of A as B. The point A is the origin of the ray. Let $\overrightarrow{AB}$ and $\overrightarrow{AC}$ be two rays originating at the same point A, not lying on the same line. Then the **angle** $\angle BAC$ is the union of the rays $\overrightarrow{AB}$ and $\overrightarrow{AC}$ and the set of those points P satisfying the following two properties: (1) the line segment PC does not intersect the line AB; (2) the line segment PB does not intersect the line AC. The point A is called the **vertex** of $\angle BAC$.

In the Euclidean plane $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$, the x-axis is the line $\{(x, 0) : x \in \mathbb{R}\}$, and the y-axis is the line $\{(0, y) : y \in \mathbb{R}\}$. Let C be the unit circle $\{(x, y) : x^2 + y^2 = 1\}$. The unit circle intersects the x-axis at two points $A(1, 0)$ and $B(-1, 0)$. The arc AB is the upper half unit circle $\{(x, y) : x^2 + y^2 = 1 \text{ and } y \geq 0\}$. If $P(x, y)$ is a point on the unit circle with $y > 0$, then the arc AP is the the intersection of the unit circle with $\angle POA$. If $P(x, y)$ is a point on the unit circle with $y < 0$, then the arc BP is the the intersection of the unit circle with $\angle POB$ and we define the arc AP to be the arc AB followed by the arc BP. For a point $P(x, y)$ on the unit circle, let $\sigma(x, y)$ be the length of the arc AP. If $(x, y) = (1, 0)$, we define $\sigma(1, 0) = 0$. Then σ is a one-to-one function from the unit circle C onto $[0, 2\pi)$. Given $\theta \in [0, 2\pi)$, there exists a unique point $P(x, y)$ on the unit circle such that $\sigma(x, y) = \theta$. We define

$$\cos\theta := x \quad \text{and} \quad \sin\theta := y.$$

It follows that $\cos 0 = 1$, $\sin 0 = 0$, $\cos(\pi/2) = 0$, $\sin(\pi/2) = 1$, $\cos\pi = -1$, $\sin\pi = 0$, $\cos(3\pi/2) = 0$ and $\sin(3\pi/2) = -1$. In general, any $\theta \in \mathbb{R}$ can be uniquely represented as $\theta = \theta_0 + 2k\pi$, where $\theta_0 \in [0, 2\pi)$ and $k \in \mathbb{Z}$. Then we define

$$\cos\theta := \cos\theta_0 \quad \text{and} \quad \sin\theta := \sin\theta_0.$$

Thus the cosine and sine functions are 2π-periodic. Since the point $(\cos\theta, \sin\theta)$ lies on the unit circle, we have

$$\cos^2\theta + \sin^2\theta = 1 \quad \forall\, \theta \in \mathbb{R}.$$

Let us find the derivatives of the sine and cosine functions. For this purpose, we consider the set $E := \{(x, y) : x^2 + y^2 = 1, x \geq 0 \text{ and } y \geq 0\}$, which is the part of the unit circle in the first quadrant. It has the following parametric equations:

$$x = u(t) = \frac{1 - t^2}{1 + t^2} \quad \text{and} \quad y = v(t) = \frac{2t}{1 + t^2}, \quad 0 \leq t \leq 1.$$

Given a point $P(u(t), v(t))$ for some $t \in [0,1]$, the length of the arc AP is $\theta = \sigma(u(t), v(t))$. Let $s(t) := \sigma(u(t), v(t))$ for $t \in [0,1]$. In the last section we proved that s is a strictly increasing continuous function from $[0,1]$ onto $[0, \pi/2]$. Moreover,

$$s'(t) = \frac{2}{1+t^2} \quad \forall\, t \in [0,1].$$

Since $x = u(t)$, $y = v(t)$ and $\theta = s(t)$, for $\theta \in [0, \pi/2]$ we have

$$\cos'(\theta) = \frac{dx}{d\theta} = \frac{\frac{dx}{dt}}{\frac{d\theta}{dt}} = \frac{\frac{-4t}{(1+t^2)^2}}{\frac{2}{1+t^2}} = \frac{-2t}{1+t^2} = -\sin\theta$$

and

$$\sin'(\theta) = \frac{dy}{d\theta} = \frac{\frac{dy}{dt}}{\frac{d\theta}{dt}} = \frac{\frac{2(1-t^2)}{(1+t^2)^2}}{\frac{2}{1+t^2}} = \frac{1-t^2}{1+t^2} = \cos\theta.$$

From the definitions of the cosine and sine functions we can deduce that

$$\cos(\theta + \pi/2) = -\sin\theta \quad \text{and} \quad \sin(\theta + \pi/2) = \cos\theta, \quad \theta \in [0, \pi/2].$$

Moreover,

$$\cos(\theta + \pi) = -\cos\theta \quad \text{and} \quad \sin(\theta + \pi) = -\sin\theta, \quad \theta \in [0, \pi].$$

Furthermore, the cosine and sine functions are 2π-periodic. Therefore we conclude that

$$\cos'(\theta) = -\sin\theta \quad \text{and} \quad \sin'(\theta) = \cos\theta \quad \forall\, \theta \in (-\infty, \infty).$$

By using differentiation we can derive the following addition formulas:

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \quad \text{and} \quad \cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta, \quad \alpha, \beta \in \mathbb{R}.$$

Indeed, to prove the first formula, we consider the function q given by

$$q(\theta) := \sin\theta\cos(\gamma - \theta) + \cos\theta\sin(\gamma - \theta), \quad \theta \in \mathbb{R},$$

where $\gamma := \alpha + \beta$ is fixed. For every $\theta \in \mathbb{R}$ we have

$$q'(\theta) = \cos\theta\cos(\gamma - \theta) + \sin\theta\sin(\gamma - \theta) - \sin\theta\sin(\gamma - \theta) - \cos\theta\cos(\gamma - \theta) = 0.$$

Hence, $q(\alpha) = q(0)$. This establishes the first formula. The second formula can be proved similarly.

Now let us find the Taylor series of the sine function about 0. Let $f(x) := \sin x$ for $x \in (-\infty, \infty)$. We have

$$f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f^{(3)}(x) = -\cos x, \quad f^{(4)}(x) = \sin x.$$

In general, for $j = 0, 1, 2, \ldots$,

$$f^{(4j)}(x) = \sin x, \quad f^{(4j+1)}(x) = \cos x, \quad f^{(4j+2)}(x) = -\sin x, \quad f^{(4j+3)}(x) = -\cos x.$$

It follows that

$$f^{(2k)}(0) = 0 \quad \text{and} \quad f^{(2k+1)}(0) = (-1)^k, \quad k = 0, 1, 2, \ldots.$$

By the Taylor theorem we obtain

$$f(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + R_{2n+1}(x),$$

where

$$R_{2n+1}(x) = \frac{f^{(2n+2)}(\xi)}{(2n+2)!} x^{2n+2} = (-1)^{n+1} \sin \xi \frac{x^{2n+2}}{(2n+2)!}$$

with ξ between 0 and x. It follows that

$$|R_{2n+1}(x)| \le \frac{|x|^{2n+2}}{(2n+2)!}.$$

Consequently,

$$\lim_{n \to \infty} |R_{2n+1}(x)| = 0.$$

Therefore

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \quad x \in (-\infty, \infty).$$

Term-by-term differentiation of the above power series gives the Taylor series of the cosine function ahbout 0:

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots, \quad x \in (-\infty, \infty).$$

The other trigonometric functions, tangent, cotangent, secant, and cosecant, are defined as follows:

$$\tan \theta := \frac{\sin \theta}{\cos \theta} \quad \text{and} \quad \sec \theta := \frac{1}{\cos \theta} \quad \text{for } \theta \in \mathbb{R} \setminus \{k\pi + \pi/2 : k \in \mathbb{Z}\},$$

and

$$\cot\theta := \frac{\cos\theta}{\sin\theta} \quad \text{and} \quad \csc\theta := \frac{1}{\sin\theta} \quad \text{for } \theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}.$$

The derivatives of these functions are found by using the quotient rule:

$$\tan'(\theta) = \sec^2\theta \quad \text{and} \quad \sec'(\theta) = \tan\theta\sec\theta \quad \text{for } \theta \in \mathbb{R} \setminus \{k\pi + \pi/2 : k \in \mathbb{Z}\},$$

and

$$\cot'(\theta) = -\csc^2\theta \quad \text{and} \quad \csc'(\theta) = -\cot\theta\csc\theta \quad \text{for } \theta \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}.$$

Finally, let us investigate the inverse trigonometric functions. Let $f(\theta) := \sin\theta$ for $-\pi/2 \le \theta \le \pi/2$. Since $f'(\theta) = \cos\theta > 0$ for $-\pi/2 < \theta < \pi/2$, f is strictly increasing on $[-\pi/2, \pi/2]$. Thus, f maps $[-\pi/2, \pi/2]$ one-to-one and onto $[-1, 1]$. Hence, the inverse function f^{-1} is continuous and strictly increasing on $[-1, 1]$ and its range is $[-\pi/2, \pi/2]$. We define

$$\arcsin x := f^{-1}(x), \quad x \in [-1, 1].$$

By the inverse function theorem, with $x = \sin\theta$ we obtain

$$\arcsin'(x) = \frac{1}{f'(\theta)} = \frac{1}{\cos\theta} = \frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1.$$

Let $g(\theta) := \cos\theta$ for $0 \le \theta \le \pi$. Since $g'(\theta) = -\sin\theta < 0$ for $0 < \theta < \pi$, g is strictly decreasing on $[0, \pi]$. Thus, g maps $[0, \pi]$ one-to-one and onto $[-1, 1]$. Hence, the inverse function g^{-1} is continuous and strictly decreasing on $[-1, 1]$ and its range is $[0, \pi]$. We define

$$\arccos x := g^{-1}(x), \quad x \in [-1, 1].$$

It is easily verified that

$$\arccos x = \frac{\pi}{2} - \arcsin x, \quad x \in [-1, 1].$$

Let $h(\theta) := \tan\theta$ for $-\pi/2 < \theta < \pi/2$. Since $h'(\theta) = \sec^2\theta > 0$ for $-\pi/2 < \theta < \pi/2$, h is strictly increasing on $(-\pi/2, \pi/2)$. Thus, h maps $(-\pi/2, \pi/2)$ one-to-one and onto $(-\infty, \infty)$. Hence, the inverse function h^{-1} is continuous and strictly increasing on $(-\infty, \infty)$ and its range is $(-\pi/2, \pi/2)$. We define

$$\arctan x := h^{-1}(x), \quad x \in (-\infty, \infty).$$

By the inverse function theorem, with $x = \tan\theta$ we obtain

$$\arctan'(x) = \frac{1}{h'(\theta)} = \frac{1}{\sec^2\theta} = \frac{1}{1+x^2}, \quad -\infty < x < \infty.$$