微积分

$W \times F$

目录

第 1 章　从自然数到实数

温故而知新，可以为师矣。

孔子

我不认为自己知道那些我不知道的事情。

苏格拉底

以“你识数吗？”作为这一章的副标题应该是恰如其分的。微积分的研究对象是函数，其自变量和函数值都是实数，要研究函数的性质就必须先了解实数的性质。

人们对于数的认知开始于幼儿时期的数(shǔ，counting)数(shù，numbers)，到小学和中学阶段，逐步认识了自然数、分数（也就是有理数）和实数，并且学习了它们的一些基本运算（主要是代数运算，加减乘除，乘方和开方）。我们知道有理数都可以表示为整数的比，从而有理数的四则运算可以通过分数形式转化为整数的四则运算。此外，我们还知道利用阿拉伯数字可以把实数表示为十进制小数，有理数可以表示为无限循环小数，而像 $\sqrt{2}, \pi$ 这样的数只能写成无限不循环小数，它们是无理数。而对两个十进制无限小数，确定它们的和差积商的十进制小数表达也并非是轻而易举的事情。

在这一章中，我们要回答的主要问题是：为什么要引进实数？为什么有理数无法满足微积分的需要？实数和有理数的根本区别在哪里？这需要我们对实数有一个更深入的理解。

我们假定读者熟悉基本逻辑规则以及集合的语言，在本章最后一节的附录中我们列举了一些常用记号供读者对照查看。

1.1　自然数集

定义 1.1.1. **自然数集** $\mathbb{N}$ 是这样一个集合，存在元素 $0 \in \mathbb{N}$ 以及一个单射 $S : \mathbb{N} \to \mathbb{N} \setminus \{0\} =: \mathbb{N}^{*}$[1]，使得 $\mathbb{N}$ 满足**数学归纳法原理**，即：对 $\mathbb{N}$ 的任意子集 A，只

要 (1) $0 \in A$ 且 (2) $n \in A \Rightarrow S(n) \in A$，就有 $A = \mathbb{N}$。

称 $S(n)$ 为 n 的**后继**，记 $1 = S(0)$，$2 = S(1)$，$3 = S(2)$ 等等。

注 1.1.2. (1) 上述定义中所称“单射”是这样的映射 $f : A \to B$，它满足 $f(x) = f(y) \Rightarrow x = y$，即任何一个函数值只能被自变量的一个值取得。
(2) 由数学归纳法原理可以证明上述单射 $S : \mathbb{N} \to \mathbb{N}^*$ 是个满射，从而是个一一对应：令 $A = \{0\} \cup S(\mathbb{N})$。则 $0 \in A$，并且若 $n \in A$，则 $S(n) \in S(\mathbb{N}) \subset A$。因此 $A = \mathbb{N}$。所以 $S(\mathbb{N}) = \mathbb{N}^*$。

定义 1.1.3. 自然数集 $\mathbb{N}$ 上的**加法**和**乘法**。对任意 $m, n \in \mathbb{N}$，定义

$$m + n = \begin{cases} m, & n = 0; \\ S(m + n'), & n = S(n'). \end{cases}$$

$$m \cdot n = \begin{cases} 0, & n = 0; \\ m \cdot n' + m, & n = S(n'). \end{cases}$$

由上述定义知

$$S(m) = S(m + 0) = m + S(0) = m + 1,$$
$$m + (n + 1) = m + S(n) = S(m + n) = (m + n) + 1.$$

定理 1.1.4. $\mathbb{N}$ 上的加法满足**结合律**：即对任意 $m, n, k \in \mathbb{N}$，

$$m + (n + k) = (m + n) + k.$$

证明. 令 $A = \{\, k \in \mathbb{N} | \forall m, n \in \mathbb{N}, m + (n + k) = (m + n) + k\}$。

由加法定义知 $0, 1 \in A$。

若 $k \in A$，则

$$\begin{aligned} m + (n + (k + 1)) &= m + ((n + k) + 1) = (m + (n + k)) + 1 && (1 \in A) \\ &= ((m + n) + k) + 1 && (k \in A) \\ &= (m + n) + (k + 1). && (1 \in A) \end{aligned}$$

因此 $k + 1 \in A$。从而 $A = \mathbb{N}$。 $\square$

定理 1.1.5. 0 是 $\mathbb{N}$ 的加法**单位元**：即对任意 $m \in \mathbb{N}$，$m + 0 = 0 + m = m$。

证明. 令 $A=\{m\in\mathbb{N}|\, m+0=0+m=m\}$。

因为 $0+0=0$，所以 $0\in A$。

设 $m\in A$。则

$$(m+1)+0=m+1=(0+m)+1=0+(m+1),$$

所以 $m+1\in A$。

因此 $A=\mathbb{N}$。 □

定理 1.1.6. $\mathbb{N}$ 上的加法满足**交换律**：即对任意 $m,n\in\mathbb{N}$，$m+n=n+m$。

证明. 令 $A=\{n\in\mathbb{N}|\,\forall m\in\mathbb{N}, m+n=n+m\}$。

由于 0 是加法单位元，所以 $0\in A$。

令 $B=\{m\in\mathbb{N}|\, m+1=1+m\}$。则 $0\in B$。对任意 $m\in B$，

$$1+(m+1)=(1+m)+1=(m+1)+1,$$

所以 $m+1\in B$，因此 $B=\mathbb{N}$。从而 $1\in A$。

设 $n\in A$。则

$$\begin{aligned} m+(n+1) &= (m+n)+1 && \text{(加法结合律)}\\ &= (n+m)+1 && (n\in A)\\ &= n+(m+1) && \text{(加法结合律)}\\ &= n+(1+m) && (1\in A)\\ &= (n+1)+m, && \text{(加法结合律)} \end{aligned}$$

因此 $n+1\in A$。从而 $A=\mathbb{N}$。 □

定理 1.1.7. $\mathbb{N}$ 上的加法对乘法满足**分配律**：即对任意 $m,n,k\in\mathbb{N}$，

$$(m+n)\cdot k=m\cdot k+n\cdot k.$$

定理 1.1.8. 1 是$\mathbb{N}$ 上乘法的单位元：即对任意 $m\in\mathbb{N}$，

$$m\cdot 1=1\cdot m=m.$$

定理 1.1.9. $\mathbb{N}$ 上的乘法满足交换律：即对任意 $m,n\in\mathbb{N}$，$m\cdot n=n\cdot m$。

定理 1.1.10. $\mathbb{N}$ 上的乘法满足结合律：即对任意 $m,n,k\in\mathbb{N}$，

$$m\cdot(n\cdot k)=(m\cdot n)\cdot k.$$

习题1.1

1. 证明定理1.1.7，1.1.9和1.1.10。

2. 证明对任意 $n \in \mathbb{N}$，

$$2 \cdot (0 + 1 + 2 + 3 + \cdots + n) = n(n + 1),$$

$$1 + 3 + 5 + \cdots + (2n + 1) = (n + 1) \cdot (n + 1).$$

3. 设 $m, n, k \in \mathbb{N}$，证明
(1) （加法消去律）：若 $m + k = n + k$，则 $m = n$。
(2) （乘法消去律）：若 $k \neq 0$ 且 $m \cdot k = n \cdot k$，则 $m = n$。

4. (1) 证明：对任意 $m, n \in \mathbb{N}$，存在 $k \in \mathbb{N}$ 使得 $m = n + k$ 或者 $n = m + k$。对前者我们记 $n \leq m$。
(2) 证明对 $\mathbb{N}$ 的任意非空子集 A，存在 $m \in A$ 使得对任意 $n \in A$ 都有 $m \leq n$。即 $\mathbb{N}$ 的任意非空子集都有最小值。

5. 设 $A \subseteq \mathbb{N}$ 满足：对任意 $k \in \mathbb{N}$，若 $S(k) \in A$，则 $k \in A$。证明以下三种情况之一恰有一个成立：
(1) $A = \emptyset$；
(2) 存在 $n \in \mathbb{N}$ 使得 $n \in A$ 但 $S(n) \notin A$；
(3) $A = \mathbb{N}$。

1.2 整数集

定义 1.2.1. 称 $\mathbb{Z} = \{m - n | m, n \in \mathbb{N}\}$ 为**整数集**，称其元素为**整数**。约定：对 $m, n, p, q \in \mathbb{N}$，

$$m - n = p - q, \quad 若\ m + q = p + n.$$

定义 $\mathbb{Z}$上的加法和乘法为：

$$(m - n) + (p - q) = (m + p) - (n + q),$$
$$(m - n) \cdot (p - q) = (m \cdot p + n \cdot q) - (n \cdot p + m \cdot q)。$$

定理 1.2.2. $\mathbb{Z}$ 上的加法和乘法是良好定义的，即：若

$$m_1 - n_1 = m_2 - n_2, \quad p_1 - q_1 = p_2 - q_2,$$

则

$$(m_1 - n_1) + (p_1 - q_1) = (m_2 - n_2) + (p_2 - q_2),$$

$$(m_1 - n_1) \cdot (p_1 - q_1) = (m_2 - n_2) \cdot (p_2 - q_2).$$

定理 1.2.3. *1.* $\mathbb{Z}$ 上的加法和乘法满足交换律、结合律、分配律。

2. 记 $0 = 0-0$。则 0 是 $\mathbb{Z}$ 上的**加法单位元**：对任意 $z \in \mathbb{Z}$，$z+0 = 0+z = z$。

3. 对任意 $z \in \mathbb{Z}$，$z = m - n$（$m, n \in \mathbb{N}$），令 $-z = n - m$，则 $-z \in \mathbb{Z}$ 满足 $z + (-z) = (-z) + z = 0$。即任何整数 z 有**加法逆元** $-z$。

4. $1 = 1 - 0$ 是 $\mathbb{Z}$ 上的**乘法单位元**：对任意 $z \in \mathbb{Z}$，$z \cdot 1 = 1 \cdot z = z$。

上述定理证明留作练习。

定理 1.2.4. $f : \mathbb{N} \to \mathbb{Z}$， $f(n) = n - 0$*(* $n \in \mathbb{N}$*)*是 $\mathbb{N}$ 到 $f(\mathbb{N})$ 的一个一一对应，且 $f(0) = 0$，$f(S(n)) = f(n) + 1$。

证明. 若 $f(n) = f(m)$，则 $n - 0 = m - 0$，从而按整数的定义，$n + 0 = m + 0$，所以 $n = m$。所以 $f : \mathbb{N} \to f(\mathbb{N})$ 是一一对应。 $f(0) = 0 - 0 = 0$，

$$
\begin{aligned}
f(S(n)) &= S(n) - 0 && (\ f \text{ 的定义}) \\
&= (n+1) - 0 && (\text{自然数加法的定义}) \\
&= (n-0) + (1-0) && (\text{整数加法的定义}) \\
&= f(n) + 1. && (\ f \text{ 和整数 } 1 \text{ 的定义})
\end{aligned}
$$

□

根据这个定理，我们可视 $f(\mathbb{N}) = \mathbb{N}$，从而 $\mathbb{N} \subseteq \mathbb{Z}$。记 $\mathbb{Z}^+ = f(\mathbb{N}^*)$，

定理 1.2.5. 对任意 $z \in \mathbb{Z}$， $z \in \mathbb{Z}^+$、$z = 0$ 和 $-z \in \mathbb{Z}^+$ 三者中恰有一个成立。

证明. 我们证明对任何 $z \in \mathbb{Z}$，$z \in \mathbb{Z}^+$ 与 $-z \in \mathbb{Z}^+$中至多只有一个成立。

若 $z = m - n \in \mathbb{Z}^+$ 满足 $-z \in \mathbb{Z}^+$，则存在 $k, l \in \mathbb{N}$ 使得

$$
m - n = S(k) - 0, \quad n - m = S(l) - 0.
$$

二者相加得到

$$
0 - 0 = (m+n) - (n+m) = (S(k) + S(l)) - 0 = S(k + S(l)) - 0,
$$

从而 $0 = S(k + S(l))$，但这与 $0 \notin S(\mathbb{N})$ 矛盾。所以对任何 $z \in \mathbb{Z}$，$z \in \mathbb{Z}^+$ 与 $-z \in \mathbb{Z}^+$ 中至多只有一个成立。

又因为 $0 = -0$，所以 $0 \notin \mathbb{Z}^+$，$-0 \notin \mathbb{Z}^+$。因此对任意 $z \in \mathbb{Z}$，$z = 0$、$z \in \mathbb{Z}^+$ 与 $-z \in \mathbb{Z}^+$中至多只有一个成立。

设 $z = m - n$ （ $m, n \in \mathbb{N}$ ）满足 $z \neq 0$。则 $m \neq n$ 。由习题1.1中的结论可知，存在 $k \in \mathbb{N}^* = S(\mathbb{N})$ 使得 $m = n + k$ 或者 $n = m + k$ 。对前者，我们有 $z \in \mathbb{Z}^+$，对后者我们有 $-z = n - m \in \mathbb{Z}^+$。 □

习题1.2

1. 证明定理 1.2.2 和 1.2.3。

1.3 有理数集

定义 1.3.1. 称 $\mathbb{Q} := \left\{ \frac{z_1}{z_2} \middle| z_1 \in \mathbb{Z}, z_2 \in \mathbb{Z}\backslash\{0\} \right\}$ 为**有理数集**，称其元素为**有理数**。约定：对 $z_1, u_1 \in \mathbb{Z}$ 和 $z_2, u_2 \in \mathbb{Z}\backslash\{0\}$：

$$\frac{z_1}{z_2} = \frac{u_1}{u_2}, \quad 若 z_1 \cdot u_2 = z_2 \cdot u_1.$$

定义 $\mathbb{Q}$ 上的加法和乘法为

$$\frac{z_1}{z_2} + \frac{u_1}{u_2} = \frac{z_1 \cdot u_2 + z_2 \cdot u_1}{z_2 \cdot u_2},$$
$$\frac{z_1}{z_2} \cdot \frac{u_1}{u_2} = \frac{z_1 \cdot u_1}{z_2 \cdot u_2},$$

并记

$$0 := \frac{0}{1}, \quad 1 := \frac{1}{1},$$
$$-\frac{z_1}{z_2} := \frac{-z_1}{z_2},$$
$$\left(\frac{z_1}{z_2}\right)^{-1} := \frac{z_2}{z_1}, \quad 若 z_1, z_2 \in \mathbb{Z}\backslash\{0\}.$$

定理 1.3.2. $\mathbb{Q}$ 是一个域，即 $\mathbb{Q}$ 上的加法和乘法是良好定义的，且满足交换律、结合律和分配律；$0, 1$ 分别是加法和乘法的单位元，且$1 \neq 0$；任何 $r \in \mathbb{Q}$ 有加法逆元；任何 $r \in \mathbb{Q}\backslash\{0\}$ 有乘法逆元 r^{-1}，即 $r \cdot r^{-1} = r^{-1} \cdot r = 1$。

对域中元素 a 和 $b \neq 0$ 以及正整数 n，记

$$\frac{a}{b} := ab^{-1}, \qquad a^n := \underbrace{a \cdot \cdots \cdot a}_{n个a}.$$

定理 1.3.3. $f : \mathbb{Z} \to \mathbb{Q}, f(z) = \frac{z}{1}$ 是一个单射，并且满足对任意 $z_1, z_2 \in \mathbb{Z}$，$f(z_1 + z_2) = f(z_1) + f(z_2)$，$f(z_1 \cdot z_2) = f(z_1) \cdot f(z_2)$，$f(0) = 0$，$f(1) = 1$。因此可视 $\mathbb{Z} \subseteq \mathbb{Q}$。

定理 1.3.4. 记 $\mathbb{Q}^+ = \left\{ \frac{z_1}{z_2} \in \mathbb{Q} \middle| z_1, z_2 \in \mathbb{Z}^+ \right\}$，则 $\forall r \in \mathbb{Q}$，$r \in \mathbb{Q}^+$、$r = 0$ 和 $-r \in \mathbb{Q}^+$ 三者中恰有一个成立。

定义 1.3.5. $\forall r_1, r_2 \in \mathbb{Q}$，定义

$$r_1 < r_2（也记作 r_2 > r_1），若 r_2 - r_1 \in \mathbb{Q}^+;$$
$$r_1 \leq r_2（也记作 r_2 \geq r_1），若 r_1 < r_2 或者 r_1 = r_2。$$

推论 1.3.6. $\forall r_1, r_2 \in \mathbb{Q}$，$r_1 < r_2$、$r_1 = r_2$ 和 $r_1 > r_2$ 三者中恰有一个成立。

定理 1.3.7. $\forall r_1, r_2, r_3 \in \mathbb{Q}$，

$$r_1 < r_2 \Rightarrow r_1 + r_3 < r_2 + r_3; \qquad r_1 < r_2 \text{且 } r_3 > 0 \Rightarrow r_1 \cdot r_3 < r_2 \cdot r_3.$$

推论 1.3.8. $\forall r_1, r_2, r_3 \in \mathbb{Q}$，

$$r_1 \le r_2 \text{ 且 } r_2 \le r_3 \Rightarrow r_1 \le r_3;\text{（传递）}$$
$$r_1 \le r_1;\text{（自反）}$$
$$r_1 \le r_2 \text{ 且 } r_2 \le r_1 \Rightarrow r_1 = r_2.\text{（反对称）}$$

上述定理和推论的证明留作练习。

定义 1.3.9. 设 $\mathbb{F}$ 是一个域，$\mathbb{F}^+ \subseteq \mathbb{F}$ 满足：(1) $\forall x \in \mathbb{F}$，$x \in \mathbb{F}^+$、$x = 0$ 和 $-x \in \mathbb{F}^+$ 三者中恰有一个成立；(2) $\forall x, y \in \mathbb{F}^+$，$x + y, xy \in \mathbb{F}^+$。则称 $(\mathbb{F}, \mathbb{F}^+)$ 是一个**序域**（也简称 $\mathbb{F}$ 是一个序域）。

总结：$(\mathbb{Q}, \mathbb{Q}^+)$ 是一个序域。

习题1.3

1. 证明本节中未证明的定理和推论。

2. (a) 证明对于任何有理数 $r \in \mathbb{Q}$，存在唯一的整数 $z \in \mathbb{Z}$ 使得 $z \le r < z + 1$。我们记 $z = [r]$，称它为 r 的**整数部分**。

   (b) （正整数的进位制表示法）任意给定正整数 $p > 1$，证明对于任意正整数 N，N 可以唯一写成如下形式

   $$N = a_n p^n + a_{n-1} p^{n-1} + \cdots + a_1 p + a_0,$$

   其中 a_n, a_{n-1}, a_1, a_0 是小于 p 的非负整数，$a_n > 0$。

3. 重要的不等式

   (a) 证明在任何序域 $\mathbb{F}$ 中 $x^2 \ge 0 (\forall x \in \mathbb{F})$，并且 $x^2 = 0$ 当且仅当 $x = 0$。

   (b) (**Cauchy-Schwarz 不等式**) 证明对任意实数 $a_1, \ldots, a_n, b_1, \ldots, b_n$，

   $$(a_1 b_1 + \cdots + a_n b_n)^2 \le (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2).$$

   等式成立当且仅当存在实数 λ 使得

   $$a_k = \lambda b_k, \quad \forall 1 \le k \le n,$$

或者

$$b_k = \lambda a_k, \quad \forall 1 \le k \le n.$$

事实上，成立等式

$$(a_1^2+\cdots+a_n^2)(b_1^2+\cdots+b_n^2)-(a_1b_1+\cdots+a_nb_n)^2 = \sum_{1\le i<j\le n}(a_ib_j-a_jb_i)^2.$$

这个等式的几何意义是 $\mathbb{R}^n$ 中由向量 $\mathbf{a} = (a_1, a_2, \ldots, a_n)^T$ 和 $\mathbf{b} = (b_1, b_2, \ldots, b_n)^T$ 所形成的平行四边形的面积与它在各 2 维坐标平面中的投影平行四边形面积之间的关系。

(c) （**Bernoulli 不等式**）设 $x_1, \ldots, x_n > -1$，且 $x_ix_j \ge 0(\forall i, j \in \{1, 2, \ldots, n\})$。证明

$$(1+x_1)\cdot \cdots \cdot (1+x_n) \ge 1 + (x_1 + \cdots + x_n),$$

其中等号成立当且仅当 $n = 1$ 或者 $x_1, \ldots, x_n$ 中至多有一个非零。经典的 Bernoulli 不等式：

$$(1+x)^n \ge 1 + nx, \quad \forall n \in \mathbb{N}, \forall x > -1.$$

(d) 利用 Bernoulli 不等式证明对任何正整数 n 以及任何正数 a, b ，都有

$$ab^n \le \left(\frac{a+nb}{n+1}\right)^{n+1},$$

且等号成立当且仅当 $a = b$。并利用这个不等式证明对任何正整数 n，都有

$$\left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n+1}\right)^{n+1} < \left(1+\frac{1}{n+1}\right)^{n+2} < \left(1+\frac{1}{n}\right)^{n+1}.$$

(e) （**算术-几何平均值不等式**）利用(d)中不等式证明：在任何序域 $\mathbb{F}$ 中，对任意正整数 n 和非负数 $x_1, x_2, \ldots, x_n \in \mathbb{F}$，都成立

$$x_1x_2\cdot \cdots \cdot x_n \le \left(\frac{x_1+x_2+\cdots+x_n}{n}\right)^n.$$

其中等号成立当且仅当 $x_1 = x_2 = \cdots = x_n$。（提示：不是每个序域中都可以使用开方运算）

(f) （**广义的算术-几何平均值不等式**）证明对任何非负数 $x_1, x_2, \ldots, x_n \in \mathbb{F}$ 和任何正整数 $p_1, p_2, \ldots, p_n$ 都成立

$$x_1^{p_1}x_2^{p_2}\cdot \cdots \cdot x_n^{p_n} \le \left(\frac{p_1x_1+p_2x_2+\cdots+p_nx_n}{p_1+p_2+\cdots+p_n}\right)^{p_1+p_2+\cdots+p_n}.$$

其中等号成立当且仅当 $x_1 = x_2 = \cdots = x_n$。

4. 证明 $(\mathbb{Q},\mathbb{Q}^+)$ 是最小的序域。即对任何序域 $(\mathbb{F},\mathbb{F}^+)$，$\mathbb{Q}\subseteq\mathbb{F}$，$\mathbb{Q}^+\subseteq\mathbb{F}^+$。

5. $\mathbb{Z}_2=\{0,1\}$ 按如下加法和乘法构成一个域，它是最小的域，但不是序域。

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

6. 令 $\mathbb{F}=\{x+\mathrm{i}y|x,y\in\mathbb{Q}\}$，对 $x_1+\mathrm{i}y_1,x_2+\mathrm{i}y_2\in\mathbb{F}$，定义加法和乘法

$$(x_1+\mathrm{i}y_1)+(x_2+\mathrm{i}y_2)=(x_1+x_2)+\mathrm{i}(y_1+y_2),$$
$$(x_1+\mathrm{i}y_1)\cdot(x_2+\mathrm{i}y_2)=(x_1x_2-y_1y_2)+\mathrm{i}(x_1y_2+x_2y_1).$$

证明

(a) $\mathbb{F}$ 是一个域；

(b) $\mathrm{i}^2=-1$；

(c) $\mathbb{F}$ 不是一个序域。

1.4 实数集

设 $(\mathbb{F},\mathbb{F}^+)$ 是一个序域。

定义 1.4.1. 对 $x\in\mathbb{F}$，记

$$|x|:=\begin{cases}x, & 若\ x\in\mathbb{F}^+,\\ 0, & 若\ x=0,\\ -x, & 若\ -x\in\mathbb{F}^+.\end{cases}$$

定理 1.4.2. *1.* $\forall x\in\mathbb{F}$，$|x|\geq 0$；$|x|=0$ 当且仅当 $x=0$。

2. $\forall x,y\in\mathbb{F}$，$|xy|=|x|\cdot|y|$，$|x+y|\leq|x|+|y|$。

3. 于是 $d(x,y)=|x-y|$ 定义了 $\mathbb{F}$ 上一个距离：

$$d(x,y)=d(y,x)\geq 0;$$
$$d(x,y)=0\ 当且仅当\ x=y;$$
$$d(x,y)\leq d(x,z)+d(z,y).$$

证明. 证明作为练习留给读者完成。 □

有了距离，就可以定量研究逼近的误差。因为有序，所以就有所谓不足近似值和过剩近似值。这都离不开下面这些概念。

定义 1.4.3. 设 $A \subseteq \mathbb{F}$。

称 A 在域 $\mathbb{F}$ 中**有界**，若存在 $M \in \mathbb{F}$ 使得：$x \in A \Rightarrow |x| \le M$。此时称这样的 M 为 A（在域 $\mathbb{F}$ 中）的一个**界**。

称 A 在域 $\mathbb{F}$ 中**有上界**（相应地，**有下界**），若存在 $b \in \mathbb{F}$ 使得：$x \in A \Rightarrow x \le b$（相应地，$x \in A \Rightarrow x \ge b$）。此时称这样的 b 为 A在域 $\mathbb{F}$ 中的一个**上界**（相应地，**下界**）。

如果 $M \in A$ 是 A 的上界，则称 M 为 A 的**最大值**，记 $M =: \max A =: \max\limits_{x\in A} x$。如果 $m \in A$ 是 A 的下界，则称 m 为 A 的**最小值**，记 $m =: \min A =: \min\limits_{x\in A} x$。

如果 A 在域 $\mathbb{F}$ 中有最小上界，则称这个最小上界为 A 在域 $\mathbb{F}$ 中的**上确界**，记为 $\sup_{\mathbb{F}} A$ 或 $\sup\limits_{x\in A} x$。如果 A 在域 $\mathbb{F}$ 中有最大下界，则称这个最大下界为 A 在域 $\mathbb{F}$ 中的**下确界**，记为 $\inf_{\mathbb{F}} A$ 或 $\inf\limits_{x\in A} x$。

在不引起歧义的情况下，我们简记 $\sup A := \sup_{\mathbb{F}} A$， $\inf A := \inf_{\mathbb{F}} A$。

不难发现：对序域 $\mathbb{F}$ 的子集 A，

1. A 有界当且仅当 A 既有上界又有下界。
2. 若 A 非空且只有有限多个元素，则 A 必然有最大值和最小值。
3. 如果 A 有最大值，则 $\max A = \sup A$；如果 A 有最小值，则 $\min A = \inf A$。

例 1.4.4. 对 $A = \{x \in \mathbb{F} | 0 \le x < 1\}$，$A$ 有最小值 0，从而 $\inf A = \min A = 0$。A 有上界，但没有最大值：$\forall x \in A$，$x < \frac{x+1}{2} < 1$，所以 $\frac{x+1}{2} \in A$。1 是 A 的最小上界，$\sup A = 1$：1 是 A 的一个上界；$\forall x < 1$，$\frac{\max\{x,0\}+1}{2} \in A$，$\frac{\max\{x,0\}+1}{2} > x$，所以 x 不是 A 的上界。 □

例 1.4.5. 空集 $\emptyset$ 是有界集，任何 $M \in \mathbb{F}$ 既是 $\emptyset$ 的上界也是 $\emptyset$ 的下界。$\emptyset$ 既无上确界，也无下确界。 □

例 1.4.6. 自然数集 $\mathbb{N}^*$ 在有理数集 $\mathbb{Q}$ 中无上界。事实上，对任何有理数 $\frac{m}{n} \in \mathbb{Q}$，$\frac{m}{n} \le \left|\frac{m}{n}\right| = \frac{|m|}{|n|} \le |m| < |m| + 1$，其中 $|m| + 1 \in \mathbb{N}^*$。所以 $\frac{m}{n}$ 不是 $\mathbb{N}^*$ 的上界。 □

例 1.4.7. 集合 $A = \left\{x \in \mathbb{Q} | x > 0, x^2 < 2\right\}$ 非空，在 $\mathbb{Q}$ 中有上界，但在 $\mathbb{Q}$ 中没有上确界。

证明. A 不是空集，因为 $1 \in A$。 A 在 $\mathbb{Q}$ 中有上界，因为从如下断言知 2 是 A 的一个上界。

断言： *若 $y > 0$ 满足 $y^2 \geq 2$，则 y 是 A 的一个上界。*

断言的证明： 若 y 不是 A 的一个上界，则存在 $x \in A$ 使得 $x > y > 0$，于是 $x^2 > x \cdot y > y^2 \geq 2$，这与 $x \in A$ 矛盾。**断言证毕。**

设 $b \in \mathbb{Q}$ 满足 $b > 0$。

若 $b^2 < 2$，则 $\frac{1}{\frac{2}{b^2}-1} \in \mathbb{Q}$，因 $\mathbb{N}^*$ 在 $\mathbb{Q}$ 中无上界，故可取正整数 $N > \frac{1}{\frac{2}{b^2}-1}$，于是

$$\left(b + \frac{b}{3N}\right)^2 = b^2\left(1 + \frac{2}{3N} + \frac{1}{9N^2}\right) \leq b^2\left(1 + \frac{7}{9N}\right) < b^2\left(1 + \frac{1}{N}\right) < 2.$$

因此 $b + \frac{b}{3N} \in A$，从而 b 不是 A 的上界。同时这还表明 A 没有最大值。

若 $b^2 > 2$，则取正整数 $N > \frac{1}{1-\frac{2}{b^2}}$，则

$$\left(b - \frac{b}{2N}\right)^2 = b^2\left(1 - \frac{1}{N} + \frac{1}{4N^2}\right) > b^2\left(1 - \frac{1}{N}\right) > 2.$$

而 $0 < b - \frac{b}{2N} < b$，所以由断言知 $b - \frac{b}{2N}$ 是 A 的上界，所以 b 不是 A 的最小上界。

若 $b^2 = 2$，因 $b \in \mathbb{Q}$，所以存在 $m, n \in \mathbb{N}^*$ 使得 $b = \frac{m}{n}$，$m^2 = 2n^2$。由 $n^2 < m^2 = 2n^2 < (2n)^2$ 知 $n < m < 2n$。取 $n' = m - n$，$m' = 2n - m$，则 $m', n' \in \mathbb{N}^*$，$n' < n$，$m' < m$，

$$m'^2 - 2n'^2 = (2n - m)^2 - 2(m - n)^2 = 2n^2 - m^2 = 0,$$

所以 $b = \frac{m'}{n'}$。这个过程可以不断重复，但这与不超过 m, n 的自然数只有有限多个是矛盾的。因此 $b^2 \neq 2$。

因此 A 在 $\mathbb{Q}$ 中没有上确界。 □

另一方面，根据勾股定理，边长为 1 的正方形的对角线长度 x 满足 $x^2 = 1^2 + 1^2 = 2$。

就像上面例子表明的那样，在很多问题中有理数无法满足我们的要求的，这就需要在有理数的基础上进一步扩展数的概念。Dedekind 通过对有理数进行分割的方式构造了实数。

定义 1.4.8. 设 A 是有理数集 $\mathbb{Q}$ 的一个非空子集。称 A 是一个**实数**，如果

1. A 在 $\mathbb{Q}$ 中有上界；
2. $\forall x \in A$，$y \in \mathbb{Q}$ 且 $y < x \Rightarrow y \in A$；
3. 若 A 在 $\mathbb{Q}$ 中有上确界，则 $\sup_{\mathbb{Q}} A \in A$。

记 $\mathbb{R}$ 为所有实数组成的集合，称为**实数集**。

定义 1.4.9. 在 $\mathbb{R}$ 上定义序：对 $A, B \in \mathbb{R}$，$A \leq B$ 当且仅当它们作为有理数集 $\mathbb{Q}$ 的子集满足 $A \subseteq B$。

定理 1.4.10. $\mathbb{R}$ 上的序 $\leq$ 满足传递、自反和反对称，并且 $\forall A, B \in \mathbb{R}$，$A \leq B$ 和 $B \leq A$ 中至少有一个成立。

定理 1.4.11. 定义 $f : \mathbb{Q} \to \mathbb{R}$，$f(r) = A_r = \{x \in \mathbb{Q} | x \leq r\}$。则 $f : \mathbb{Q} \to \mathbb{R}$ 是严格增函数，即 $\forall r, s \in \mathbb{Q}$，$r < s \Rightarrow f(r) < f(s)$。但 f 不是满射。

证明. 设 $r, s \in \mathbb{Q}, r < s$。则 $s \in A_s \backslash A_r$，所以 $A_r \subset A_s$。因此 $f : \mathbb{Q} \to \mathbb{R}$ 是严格增函数。令 $A = \left\{x \in \mathbb{Q} | x < 0 \text{ 或 } x^2 < 2\right\}$。由例1.4.7知，$A$ 非空，在 $\mathbb{Q}$ 中有上界但无上确界。设 $x \in A$，$y \in \mathbb{Q}$ 满足 $y < x$。若 $y \leq 0$，则 $y \in A$；若 $y > 0$，则 $x > y > 0$，从而 $y^2 < x^2 < 2$，所以 $y \in A$。因此 A 是个实数。但 A 无最大值，所以 $\forall r \in \mathbb{Q}$，$A \neq A_r$。即 $A \notin f(\mathbb{Q})$。故 $f : \mathbb{Q} \to \mathbb{R}$ 不是满射。 □

因此可视 $\mathbb{Q} \subset \mathbb{R}$。

定理 1.4.12. $\mathbb{Q}$ 在 $\mathbb{R}$ 中**稠密**，即 $\forall A, B \in \mathbb{R}, A < B$，存在 $r \in \mathbb{Q}$ 使得 $A < A_r < B$。

证明. 任取 $A, B \in \mathbb{R}, A < B$。则 $A \subset B$，因此存在 $b \in \mathbb{Q}$ 使得 $b \in B \backslash A$。

因为 $B \in \mathbb{R}$ 且 $b \in B$，所以 $A_b \subseteq B$。因为 $b \notin A$，所以 b 是 A 的上界（假设存在 $x \in A$ 使得 $b < x$，则由 $A \in \mathbb{R}$ 知 $b \in A$，矛盾），但不是 A 的上确界。

因此存在 $c \in \mathbb{Q}$ 使得 $c < b$ 且 c 是 A 的上界。故 $r = \frac{b+c}{2} \in \mathbb{Q}$ 满足 $A \subseteq A_c \subset A_r \subset A_b \subseteq B$。 □

定理 1.4.13. $\mathbb{R}$ 中任何非空有上界的集合都有上确界。

证明. 设 $\mathcal{A} \subseteq \mathbb{R}$ 非空有上界。令 $A^* = \bigcup\limits_{A \in \mathcal{A}} A$。则 $A^* \subseteq \mathbb{Q}$。

对 $\mathcal{A}$ 的任何上界 $B \in \mathbb{R}$，对任意 $A \in \mathcal{A}$，$A \leq B$。取 $\tilde{A} \in \mathcal{A}$，则 $\tilde{A} \subseteq A^* \subseteq B$。所以 $A^* \neq \emptyset$。因为 B 在 $\mathbb{Q}$ 中有上界，所以 A^* 在 $\mathbb{Q}$ 中有上界。

设 $x \in A^*$，$y \in \mathbb{Q}$ 满足 $y < x$。则存在 $A(x) \in \mathcal{A}$，$x \in A(x)$，因为 $A(x) \in \mathbb{R}$，所以 $y \in A(x) \subseteq A^*$。

(1) 如果 A^* 在 $\mathbb{Q}$ 中没有上确界，则令 $\bar{A} = A^*$，于是 $\bar{A} \in \mathbb{R}$。$\bar{A}$ 是 $\mathcal{A}$ 的上界。对 $\mathcal{A}$ 的任何上界 $B \in \mathbb{R}$，$\bar{A} = A^* \subseteq B$，所以 $\bar{A} \leq B$，所以 $\bar{A}$ 是 $\mathcal{A}$ 的最小上界，即上确界。

(2) 如果 A^* 在 $\mathbb{Q}$ 中有上确界 $b^* \in \mathbb{Q}$，则令 $\bar{A} = A^* \bigcup \{b^*\}$。易见 b^* 是 $\bar{A}$ 的最大值。对任意 $y \in \mathbb{Q}$，若 $y < b^*$，则 y 不是 A^* 在 $\mathbb{Q}$ 中的上界，所以存在 $x \in A^*$ 使得 $y < x$。于是 $y \in A^* \subseteq \bar{A}$。所以 $\bar{A} \in \mathbb{R}$ 且 $\bar{A}$ 是 $\mathcal{A}$ 的上界。接下来，我们证明 $\bar{A}$ 是 $\mathcal{A}$ 的上确界。

假设 $\bar{A}$ 不是 $\mathcal{A}$ 的上确界（即最小上界），则存在 $\mathcal{A}$ 的上界 B，使得 $B < \bar{A}$。由有理数集在实数集中的稠密性（定理1.4.11），存在有理数 $r \in \mathbb{Q}$ 使得 $B < A_r < \bar{A}$。所以 $r \in \bar{A}$ 且 $r < b^*$，因此 $r \in A^*$。从而存在 $A(r) \in \mathcal{A}$ 使得 $r \in A(r)$。而 $A(r) \le B$，所以 $r \in B$，因此 $A_r \subseteq B$，这与 $B < A_r$ 矛盾。所以 $\bar{A}$ 是 $\mathcal{A}$ 的上确界。

总之，$\mathcal{A}$ 在 $\mathbb{R}$ 中有上确界。 □

下面考虑 $\mathbb{R}$ 上的运算。

定义 1.4.14. 设 $A, B \in \mathbb{R}$，定义

$$\begin{aligned} A + B &= \{r \in \mathbb{Q} | \forall a \in \mathbb{Q}\backslash A, \forall b \in \mathbb{Q}\backslash B, r \le a + b\}, \\ -A &= \{r \in \mathbb{Q} | \forall a \in A, r \le -a\}, \\ 0 &= \{r \in \mathbb{Q} | r \le 0\}, \\ 1 &= \{r \in \mathbb{Q} | r \le 1\}. \end{aligned}$$

对 $A, B \in \mathbb{R}$，如果 $A \ge 0$ 且 $B \ge 0$，则定义

$$A \cdot B = \{r \in \mathbb{Q} | \forall a \in \mathbb{Q}\backslash A, \forall b \in \mathbb{Q}\backslash B, r \le a \cdot b\};$$

如果 $A \ge 0$ 且 $B < 0$，则定义 $A \cdot B = -(A \cdot (-B))$；如果 $A < 0$ 且 $B \ge 0$，则定义 $A \cdot B = -((-A) \cdot B)$；如果 $A < 0$ 且 $B < 0$，则定义 $A \cdot B = (-A) \cdot (-B)$。

定理 1.4.15. *$(\mathbb{R}, \mathbb{R}^+)$ 是一个序域，其中 $\mathbb{R}^+ = \{A \in \mathbb{R} | A > 0\}$。*

定理 1.4.16 (阿基米德性质). *$\mathbb{N}$ 在 $\mathbb{R}$ 中没有上界。*

证明. 假设 $\mathbb{N}$ 在 $\mathbb{R}$ 中有上界。因为 $\mathbb{N}$ 非空（$1 \in \mathbb{N}$），所以由定理 1.4.13 知 $\mathbb{N}$ 在 $\mathbb{R}$ 中有上确界 $b \in \mathbb{R}$。从而 $b - 1$ 不是 $\mathbb{N}$ 的上界，因此存在 $n \in \mathbb{N}$ 使得 $b - 1 < n$，于是 $b < n + 1$，而 $n + 1 \in \mathbb{N}$，所以 b 不是 $\mathbb{N}$ 的上界，但这与 b 是 $\mathbb{N}$ 的上确界矛盾。 □

最后，我们可以总结一下这一章中最重要的结论：实数集 $\mathbb{R}$ 是唯一一个具有确界性质的序域。

定理 1.4.17. *若 $(\mathbb{F}_1, \mathbb{F}_1^+)$ 和 $(\mathbb{F}_2, \mathbb{F}_2^+)$ 是两个序域，都具有确界性质，则存在唯一的一一对应 $f : \mathbb{F}_1 \to \mathbb{F}_2$ 使得 f 保持加法、乘法和序，即*

$$f(x + y) = f(x) + f(y), \quad f(x \cdot y) = f(x) \cdot f(y), \quad \forall x, y \in \mathbb{F}_1,$$

且 $f(\mathbb{F}_1^+) = \mathbb{F}_2^+$。

证明. 证明留给读者完成。 □

习题1.4

1. 证明本节中未经证明的定理和推论。
2. 讨论集合 $\{q^n|n\in\mathbb{N}\}$ 的有界性，其中 $q\in\mathbb{R}$。
3. 设 $\mathbb{F}$ 是一个阿基米德序域（即 $\mathbb{F}$ 是一个序域，且 $\mathbb{N}$ 在 $\mathbb{F}$ 中无上界）。
   (a) 对 $a,b\in\mathbb{F}$，证明 $a\le b$ 当且仅当对任意 $n\in\mathbb{N}^*$， $a<b+\frac{1}{n}$。
   (b) 证明 $\mathbb{Q}$ 在 $\mathbb{F}$ 中稠密。
4. 称 $I\subseteq\mathbb{R}$ 是一个**区间**，如果对任意 $x,y\in I$ 以及任意 $z\in\mathbb{R}$，$x<z<y\Rightarrow z\in I$。证明区间必然是如下九种集合之一：

$$\begin{aligned}
&(-\infty,+\infty)=\mathbb{R},\\
&(-\infty,b)=\{x\in\mathbb{R}|x<b\},\qquad (-\infty,b]=\{x\in\mathbb{R}|x\le b\},\\
&[a,b]=\{x\in\mathbb{R}|a\le x\le b\},\qquad [a,b)=\{x\in\mathbb{R}|a\le x<b\},\\
&(a,b]=\{x\in\mathbb{R}|a<x\le b\},\qquad (a,b)=\{x\in\mathbb{R}|a<x<b\},\\
&[a,+\infty)=\{x\in\mathbb{R}|a\le x\},\qquad (a,+\infty)=\{x\in\mathbb{R}|a<x\}.
\end{aligned}$$

5. (a) 证明：对任意实数 x，存在唯一的整数 $[x]$ 使得 $[x]\le x<[x]+1$；
   (b) （实数的 p 进制表示）设 $p>1$ 是正整数，证明对任意实数 $x>0$，存在唯一的整数 k 和整数 a_k 使得 $1\le a_k\le p-1$ 且 $0\le x-a_kp^k<p^k$。
6. **从乘方到开方**：关于有理数指数的幂函数
   (a) (Bernoulli 不等式)证明：对任意整数 n 和任何实数 $x>-1$，

$$(1+x)^n\ge 1+nx,$$

   其中等号成立当且仅当 $n=0$ 或 $n=1$ 或 $x=0$。
   (b) 对任意 $m\in\mathbb{Z},n\in\mathbb{N}^*$ 和任何正数 x。证明：
   (1) 集合 $A_{x,m,n}=\{r\in\mathbb{Q}|r>0,r^n<x^m\}$ 有上确界；
   (2) $A_{x,m,n}$ 的上确界是满足 $y^n=x^m$ 的唯一正实数 y，记 $y=x^{\frac{m}{n}}$。
   (c) 证明：对任何 $x>0$ 及任何有理数 r，x^r 有唯一定义且与分数表示 $r=\frac{m}{n}$ 的选取无关。
   (d) 证明：对正有理数 r，函数 $f_r:(0,+\infty)\to(0,+\infty)$，$f(x)=x^r$ 是严格增满射；对负有理数 r，$f:(0,+\infty)\to(0,+\infty)$，$f(x)=x^r$ 是严格减满射。

(e) 对 $x>0$，证明 $x^0=1$，并且对任意 $r,s\in\mathbb{Q}$ 以及任意 $x,y>0$，$x^r\cdot x^s=x^{r+s}$，$(x^r)^s=x^{r\cdot s}$，$x^r\cdot y^r=(x\cdot y)^r$。

7. 关于有理数集上的指数函数。设 $a>1$。证明：函数 $f:\mathbb{Q}\to\mathbb{R}$，$f(x)=a^x$ 是严格增函数。

8. **从乘方到对数** 设 $a>1$。

(a) 证明：对任何正数 x ，存在整数 m 使得 $a^m\le x<a^{m+1}$。

(b) 对任何 $x>0$，记集合

$$A_x:=\left\{\frac{m}{n}\middle| a^m\le x^n, m\in\mathbb{Z}, n\in\mathbb{N}^*\right\}.$$

证明：如果 $a^m>x^n$ ($m\in\mathbb{Z},n\in\mathbb{N}^*$)，则 $\frac{m}{n}$ 是 A_x 的一个上界。

(c) 证明：对任何 $x>0$，A_x 非空、有上界。

(d) 证明：$f:\mathbb{R}^+\to\mathbb{R}$，$f(x)=\sup A_x$ 是严格增函数。

(e) 证明：$f(a)=1$ 且 $f(xy)=f(x)+f(y),\quad\forall x,y\in\mathbb{R}^+$。

(f) 证明：若 $b>0$ 且 $b\ne 1$，则存在唯一的单调函数 $g:\mathbb{R}^+\to\mathbb{R}$ 满足 $g(b)=1$ 且 $g(xy)=g(x)+g(y),\quad\forall x,y\in\mathbb{R}^+$。记 g 为 $\log_b$，称为以 b 为底的**对数函数**。

(g) 证明：若 $a,b>0$ 且 $a,b\ne 1$，则 $\log_b x=\frac{\log_a x}{\log_a b},\quad x>0$。

9. 证明 $\{\sin n|n\in\mathbb{N}\}$ 在区间 $[-1,1]$ 中稠密。

10. $\mathbb{Q}$ 是最小的阿基米德序域。$\mathbb{R}$ 是否为最大的阿基米德序域？

1.5 附：集合与逻辑

数学推理依靠逻辑，而集合是现代数学的语言。

设 p 是一个陈述。则 $\neg p$ 也是一个陈述，称为“**非** p”，它们的真值关系为

p	真	假
$\neg p$	假	真

设 p,q 是两个陈述。

$p\wedge q$ 和 $p\vee q$ 是两个陈述，分别称为“ p **且** q ”和“ p **或** q”，它们的真值为

$p\wedge q$		q 真	q 假
p	真	真	假
	假	假	假

$p\vee q$		q 真	q 假
p	真	真	真
	假	真	假

即当且仅当 p, q 都是真时，$p \wedge q$ 为真；当且仅当 p, q 都是假时，$p \vee q$ 为假。

陈述 $p \Rightarrow q$ 表示 $(\neg p) \wedge q$，称为“ p **蕴涵** q ”或者“**若** p**，则** q ”，其中 p 称为**前件**，q 称为**后件**。$p \Rightarrow q$ 的真值为

$p \Rightarrow q$		q 真	q 假
p	真	真	假
	假	真	真

即当且仅当 p 真且 q 假时，$p \Rightarrow q$ 为假。$p \Rightarrow q$ 也可以写成 $q \Leftarrow p$。

记 $p \Leftrightarrow q$ 表示“ $p \Rightarrow q$ 且 $q \Rightarrow p$”，称为 p **和** q **等价**。

在日常语言中，$p \Rightarrow q$ 有以下不同的表述形式：“若 p，则 q ”，“如果 p，那么 q ”，“如果 p，就有 q ”，“当 p 成立时，q 成立”，“只要 p 成立，q 就成立”，“只有 q 成立，p 才成立”等等。

当 $p \Rightarrow q$ 为真时，p 称为 q 的**充分条件**，q 称为 p 的**必要条件**。当 $p \Leftrightarrow q$ 为真时，称 p 和 q 为对方的**充分必要条件**。$p \Leftrightarrow q$ 为真当且仅当 p, q 都是真或都是假。

易见以下陈述为真

$$p \vee (\neg p), \quad \neg(\neg p) \Leftrightarrow p,$$
$$\neg(p \wedge q) \Leftrightarrow (\neg p) \vee (\neg q), \qquad \neg(p \vee q) \Leftrightarrow (\neg p) \wedge (\neg q),$$

另外，

$$(p \wedge p) \Leftrightarrow p, \quad (p \vee p) \Leftrightarrow p,$$
$$p \vee (q \vee r) \Leftrightarrow (p \vee q) \vee r, \qquad p \wedge (q \wedge r) \Leftrightarrow (p \wedge q) \wedge r,$$
$$p \vee (q \wedge r) \Leftrightarrow (p \wedge q) \vee (p \wedge r), \qquad p \wedge (q \vee r) \Leftrightarrow (p \vee q) \wedge (p \vee r).$$

特别要提醒读者注意的是：p，q 和 $p \Rightarrow q$ 是三个不同的陈述；并且当条件 p 为假时，无论后件 q 是真还是假，蕴涵关系 $p \Rightarrow q$ 都是真的。另外，

$$(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p),$$

即逆否命题与原命题具有相同的真值，这为反证法提供了逻辑依据。

设 A 是一个集合。记号 $x \in A$ 表示对象 x 是集合 A 的一个成员，称 x 是 A 的一个**元素**。$x \notin A$ 表示 $\neg(x \in A)$。

设 A, B 是两个集合。如果

$$x \in A \Rightarrow x \in B,$$

则称 A 是 B 的一个**子集**，记为 $A \subseteq B$。如果进一步有$A \neq B$，则称 A 是 B 的一个**真子集**，记为 $A \subset B$。根据蕴涵关系的逻辑可以证明空集 $\emptyset$ 是任何集合的子集。

设 X 是一个集合，$p(x)$ 是一个包括变元 x 的陈述。则记

$$\{x \in X | p(x)\}$$

为 X 中所有使得 $p(x)$ 为真的那些元素 x 组成的集合。

一般而言，$\{x|p(x)\}$ 可能并不是一个集合。例如 $A = \{x|x \notin x\}$，如果 A 是个集合，那么 $A \in A$ 和 $A \notin A$ 中必有一个成立。但无论它们中的哪一个成立都会和 A 的定义矛盾。为了避免产生类似的矛盾，所以数学家对“集合”这个看似朴素的概念做了必要的局限，这导致了公理化集合论的产生。在公理化集合论中，对于已有的集合 X，$\{x \in X|p(x)\}$ 是一个集合。

陈述“ $\forall x \in A, p(x)$ ” 为真当且仅当对 A 的所有元素 x，$p(x)$ 为真。陈述“ $\exists x \in A, p(x)$ ”为真当且仅当存在 A 的元素 x 使得 $p(x)$ 为真。

陈述“ $\forall x \in A, p(x)$ ” 等价于 $x \in A \Rightarrow p(x)$。在日常语言中，这个陈述可以有不同形式的表述，例如“凡 A 中成员 x，都 $p(x)$”，“对 A 中任一/任意/每个/所有 x，都 $p(x)$”等。

易见

$$\neg(\forall x \in A, p(x)) = \exists x \in A, \neg(p(x));$$
$$\neg(\exists x \in A, p(x)) = \forall x \in A, \neg(p(x)).$$

但是当 $\forall$ （对所有，对每一个，对任意）和 $\exists$ （存在，至少有一个）同时出现在一个陈述中时，它们的前后顺序不能随意改变，比如：“对每个男人 A，都存在一个男人 B 使得 A 是 B 的儿子”和“存在一个男人 B 使得每一个男人 A 都是 B 的儿子”的意思是完全不同的。

设 Λ 是一个集合，对每个 $\lambda \in \Lambda$，A_λ 是 X 的一个子集，记

$$\bigcap_{\lambda \in \Lambda} A_\lambda = \{x \in X | \forall \lambda \in \Lambda, x \in A_\lambda\},$$
$$\bigcup_{\lambda \in \Lambda} A_\lambda = \{x \in X | \exists \lambda \in \Lambda, x \in A_\lambda\},$$

它们分别称为集合 A_λ 们的**交集**和**并集**。对 X 的子集 A，

$$A^c = \{x \in X | x \notin A\}$$

称为 A 的**补集**。

易见

$$\left(\bigcap_{\lambda \in \Lambda} A_\lambda\right)^c = \bigcup_{\lambda \in \Lambda} A_\lambda^c,$$
$$\left(\bigcup_{\lambda \in \Lambda} A_\lambda\right)^c = \bigcap_{\lambda \in \Lambda} A_\lambda^c.$$

记

$$\emptyset_X = \{x \in X | x \neq x\},$$

称它为 X 中的**空集**。易见对于 X 的任何子集 A，都有 $\emptyset_X \subseteq A$。

第 2 章　连续与极限

一个物理量 y 的值由另一个物理量 x 决定，x 的值可以直接观测，而 y 的值只能通过 $y = f(x)$ 间接得到。由于任何测量都有误差，所以观测值只可能是近似值。一个自然的问题是：如果 x 是 x^* 的近似值，那么 $f(x)$ 是否也是 $f(x^*)$ 的近似值呢？或者更严格地说，我们能否通过控制 x 和 x^* 之间的误差，来减小 $f(x)$ 与 $f(x^*)$ 之间的误差呢？

本章研究函数在一点附近或无穷远附近的渐近行为，其中在一点附近的渐近行为涉及函数的连续性和极限。我们把数列看成自变量取正整数值的函数，因此把数列极限作为函数极限的一种特殊情况。对无穷小量和无穷大量的概念以及大 O 和小 o 的语言，我们做了一个简明扼要又不失严谨的介绍。

2.1　函数的连续性

在上一章的习题里，我们已经对所有有理数 r 确定了幂 2^r 的意义。如果要试图对所有实数（特别是无理数）x 定义 2^x，使其适合已经确定了的 2^r，就必须研究函数 $f : \mathbb{Q} \to \mathbb{R}$，$f(r) = 2^r$ 的性质，也就是它的连续性。

有鉴于此，在本章中函数的定义域可以是任何实数子集，而不必是区间，这是要特别提醒读者注意的。

定义 2.1.1. 称函数 $f : I \to \mathbb{R}$ 在 $a \in I$ 处**连续**，如果对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得对任意 $x \in I$，只要 $|x - a| < \delta$，就成立 $|f(x) - f(a)| < \varepsilon$。即

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in I \big(|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon \big).$$

换句话说，凡 I 中 δ-接近于 a 的 x，其函数值 $f(x)$ 都 ε-接近于 $f(a)$。

称 $f : I \to \mathbb{R}$ 是**连续函数**，如果 f 在每个 $a \in I$ 处连续。

上述定义中，$|f(x) - f(a)|$ 和 $|x - a|$ 是距离，也可以理解为误差。ε 和 δ 出现在不等式较大的一侧，是很小的正数，用来控制距离和误差。$|f(x) - f(a)| < \varepsilon$

是目的，所以 ε 可以任意变小。$|x-a|<\delta$ 是为了实现这个目的而采取的手段，δ 会随 ε 的改变而做出适当调整，因此通常把 δ 写成 δ_ϵ 的形式，来体现 δ_ε 依赖于 ε。可以任意变小的 ε 和与之相应变小的 δ_ε 体现了上述定义本质上是一个动态过程。

本书中我们有时也讨论变量取值为复数的函数，易见上述定义也适用于自变量或函数值为复数的函数。

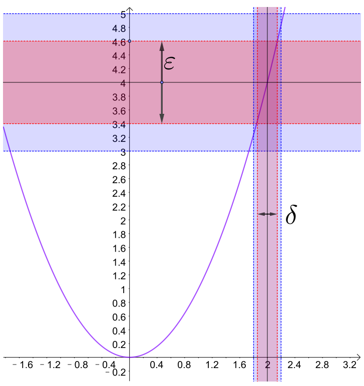
例 2.1.2. 任何常值函数都是连续函数。 □

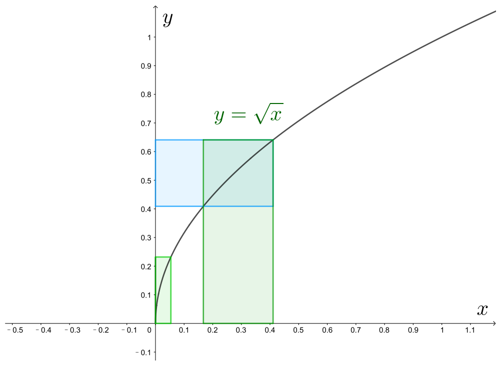
例 2.1.3. x^2 是连续函数。

证明. 对任意 $a\in\mathbb{R}$，任意给定 $\varepsilon>0$，对任意 x，记 $h=x-a$，则

$$\begin{aligned}|x^2-a^2| &= |(a+h)^2-a^2| = |2ah+h^2| \\ &\le (2|a|+|h|)\,|h| \\ &\le (2|a|+1)\,|h| && (\text{若 } |h|<1) \\ &< \varepsilon, && (\text{若 } |h|<\frac{\varepsilon}{1+2|a|})\end{aligned}$$

因此当 $|x-a|<\frac{\varepsilon}{1+2|a|+\varepsilon}$ 时，$|x^2-a^2|<\varepsilon$，x^2 在 a 处连续，x^2 是连续函数。 □





例 2.1.4. $\sqrt{x}$ 是区间 $[0,+\infty)$ 上的连续函数。

证明. 任意给定 $a\ge 0$，任意给定 $\varepsilon>0$，则对满足 $|x-a|<\varepsilon^2$ 的任意非负实数 x，都有

$$|\sqrt{x}-\sqrt{a}|^2 = |\sqrt{x}-\sqrt{a}||\sqrt{x}-\sqrt{a}| \le |\sqrt{x}-\sqrt{a}||\sqrt{x}+\sqrt{a}| = |x-a| < \varepsilon^2.$$

所以 $|\sqrt{x}-\sqrt{a}|<\varepsilon$，从而 $\sqrt{x}$ 在 a 处连续。从而 $\sqrt{x}$ 是连续函数。 □

注 2.1.5. (1) 按定义，验证连续就是验证以下逻辑命题

$$\forall\varepsilon>0\left[\exists\delta>0\left(\forall x\in I\big(|x-a|<\delta\Rightarrow|f(x)-f(a)|<\varepsilon\big)\right)\right].$$

注意最后这个蕴涵关系只是充分条件，不是充分必要条件，所以验证连续性并不是求解不等式 $|f(x)-f(a)|<\varepsilon$，而只是找到一个使这个不等式成立的形如 $|x-a|<\delta$ 的充分条件。

(2) 上述两个例子表现出来的连续性质有所不同。

对函数 $\sqrt{x}$ 而言，

$$\delta(\varepsilon)=\varepsilon^2$$

只与 ε 有关，可以适用于所有 $a\in[0,+\infty)$。

对函数 x^2 而言，

$$\delta(\varepsilon,a)=\frac{\varepsilon}{1+\varepsilon+2|a|}$$

不仅与 ε 有关，也与 a 有关。对任何 $\varepsilon,\delta>0$，当 $n>\frac{\varepsilon}{2}+\frac{1}{\delta}$ 时，对于 $a=n^2$，$|n^2+\frac{1}{n}-n^2|=\frac{1}{n}<\delta$，但是

$$\left|\left(n^2+\frac{1}{n}\right)^2-(n^2)^2\right|=\left|2n+\frac{1}{n^2}\right|>2n>\varepsilon.$$

稍后我们会更细致地研究这两种不同的连续性质。

例 2.1.6. $\frac{1}{x}$ 是 $\mathbb{R}\backslash\{0\}$ 上的连续函数。

证明. 对任意 $a\neq0$，只要 $|x-a|<\delta$，就有

$$|x|\geq|a|-|x-a|>|a|-\delta,$$

当 $\delta<|a|$时，$|x|>0$，从而 $x\neq0$，且

$$\left|\frac{1}{x}-\frac{1}{a}\right|=\frac{|x-a|}{|a||x|}<\frac{\delta}{|a|(|a|-\delta)}.$$

对任意 $\varepsilon>0$，由

$$0<\frac{\delta}{|a|(|a|-\delta)}<\varepsilon$$

解得

$$0<\delta<\frac{|a|^2\varepsilon}{1+|a|\varepsilon}.$$

所以当$|x-a|<\frac{a^2\varepsilon}{1+|a|\varepsilon}$时，$\left|\frac{1}{x}-\frac{1}{a}\right|<\varepsilon$。从而 $\frac{1}{x}$ 在任何 $a\in\mathbb{R}\backslash\{0\}$ 处连续。 □

以下定理确保在复合以及四则运算下可以保持函数的连续性，它使得我们可以通过函数的代数结构把一些复杂函数的连续性归结到一些“基础构件”的连续性，从而避免陷入繁琐的不等式论证。但是，如果我们希望对误差进行细致的分析，不等式的讨论仍然是至关重要的，这也是我们给出详细证明的目的。

定理 2.1.7 (连续函数的复合). 设函数 f 在 a 处连续，g 在 $b=f(a)$ 处连续。则 f 与 g 的**复合函数** $g\circ f$ 在 a 处连续，其中 $(g\circ f)(x)=g(f(x))$。 □

证明. 对任意 $\varepsilon>0$，存在 $\delta>0$ 使得对 g 定义域中任意 y，只要 $|y-b|<\delta$，就有 $|g(y)-g(b)|<\varepsilon$。

对应于这个 δ，存在 $\sigma>0$ 使得对 f 定义域中任意 x，只要 $|x-a|<\sigma$，就有 $|f(x)-f(a)|<\delta$。

于是

$$|g(f(x))-g(f(a))|=|g(f(x))-g(b)|<\varepsilon.$$

因此 $g\circ f$ 在 a 处连续。 □

利用数学归纳法，这个定理可以推广到任意有限多个函数的情形。

定理 2.1.8 (连续函数四则运算). 设函数 f_1 和 f_2 都在 a 处连续。则 f_1+f_2，$f_1\cdot f_2$，$\frac{f_1}{f_2}$（若 $f_2(a)\neq 0$）都在 a 处连续。

证明. (1) 加法和数乘的连续性。

对任意 $\varepsilon>0$，存在正数 δ_1,δ_2 使得：对任意 $x\in I$，只要 $|x-a|<\delta_k$ ，就有 $|f_k(x)-f_k(a)|<\frac{\varepsilon}{2}$ ($k=1,2$)。因此当 $|x-a|<\min\{\delta_1,\delta_2\}$ 时，

$$|f_1(x)+f_2(x)-f_1(a)-f_2(a)|\leq\sum_{k=1}^{2}|f_k(x)-f_k(a)|<\varepsilon.$$

因此 f_1+f_2 在 a 连续。

(2) 乘法的连续性。

先证明 $f_1(a)=f_2(a)=0$ 时的情形。此时，对任意 $\varepsilon>0$，存在正数 δ 使得：对任意 $x\in I$，只要 $|x-a|<\delta$ ，就有

$$|f_k(x)|<\min\{\varepsilon,1\},\quad k=1,2.$$

于是

$$|f_1(x)f_2(x)|=|f_1(x)||f_2(x)|<1\cdot\varepsilon=\varepsilon.$$

因此 $f_1\cdot f_2$ 在 a 处连续。

再证明 f_1 是常数的情形。此时，对任意 $\varepsilon>0$，存在正数 δ 使得：对任意 $x\in I$，只要 $|x-a|<\delta$ ，就有

$$|f_2(x)|<\frac{\varepsilon}{|f_1(a)|+1}.$$

于是

$$|f_1(x)f_2(x)-f_1(a)f_2(a)|=|f_1(a)||f_2(x)-f_2(a)|\leq|f_1(a)|\frac{\varepsilon}{|f_1(a)|+1}<\varepsilon.$$

因此 $f_1\cdot f_2$ 在 a 处连续。

对一般情形，

$$f_1(x)f_2(x) = (f_1(x) - f_1(a))(f_2(x) - f_2(a)) + f_2(a)(f_1(x) - f_1(a)) + f_1(a)f_2(x),$$

由加法的连续性以及上述两个特殊情形下的连续性知，上式右端是连续函数。因此 $f_1 \cdot f_2$ 在 a 连续。

乘法的连续性也可以利用如下等式

$$f(x)g(x) = \left(\frac{f(x)+g(x)}{2}\right)^2 - \left(\frac{f(x)-g(x)}{2}\right)^2$$

以及加法的连续性、x^2 的连续性、复合函数的连续性以及数乘运算的连续性得到。

(3) 除法的连续性。

由 $\frac{1}{x}$ 的连续性和复合函数的连续性定理知 $\frac{1}{f_2}$ 在 a 处连续。再由乘法的连续性结论知 $\frac{f_1}{f_2}$ 在 a 处连续。 □

推论 2.1.9. *代数函数(多项式，有理函数——多项式的商)都是连续函数。* □

注 2.1.10. 本节中我们使用了很多不等式论证。这种不等式论证方法是误差估计的基本方法，在实验科学和工程技术中会经常遇到。

习题2.1

1. 如何仅用十进制有限小数加减乘运算得到 $\sqrt{2}$ 的给定精度的近似值？

2. 证明以下函数的连续性

   (1) $f(x) = |x|$；

   (2) $g(x) = \max\{f_1(x), \ldots, f_n(x)\}$ (其中 $f_1, \ldots, f_n$ 都是 I 上的连续函数)

3. **乘方和开方的连续性**。用定义证明对任何正整数 n，x^n 和 $x^{\frac{1}{n}}$ 都是区间 $(0, +\infty)$ 上的连续函数。分别对 $f(x) = x^n$ 和 $f(x) = x^{\frac{1}{n}}$ 判断是否对任意 $\varepsilon > 0$ 都存在与 $x_0 > 0$ 无关的正数 δ 使得只要正数 x 满足 $|x - x_0| < \delta$ 就有 $|f(x) - f(x_0)| < \varepsilon$。

4. **对数函数的连续性**。证明：对 $a > 1$，对数函数 $\log_a : \mathbb{R}^+ \to \mathbb{R}$ （定义见习题1.4）是连续函数。

5. 以下论证希望表明任何严格增函数都是连续函数。

   证明. 设 $f : (a, b) \to (c, d)$ 是一个严格增函数。任取 $x_0 \in (a, b)$。对任意 $\varepsilon > 0$，取 $x_\pm$ 使得 $f(x_\pm) = f(x_0) \pm \varepsilon$ 以及 $\delta = \min\{x_+ - x_0, x_0 - x_-\}$。则 $x_- < x_0 < x_+$，$\delta > 0$，且对任意 $x \in (x_0 - \delta, x_0 + \delta)$，

   $$f(x_0) - \varepsilon = f(x_-) < f(x) < f(x_+) = f(x_0) + \varepsilon.$$

因此 f 在 x_0 处连续。 □

请问上述结论对吗？这个证明对吗？如果你发现这个证明有问题，你能够把它修改成一个正确的证明吗？

6. 以下论证希望表明 x^2 是连续函数。

   对于 $a \in \mathbb{R}$，任取正数 ε ($0 < \varepsilon < a^2$)，当

$$|x - a| < \min\left\{\sqrt{a^2+\varepsilon} - a, a - \sqrt{a^2-\varepsilon}\right\}$$

   时，

$$\sqrt{a^2-\varepsilon} - a < x - a < \sqrt{a^2+\varepsilon} - a,$$

   从而

$$\sqrt{a^2-\varepsilon} < x < \sqrt{a^2+\varepsilon},$$

   因此

$$a^2 - \varepsilon < x^2 < a^2 + \varepsilon,$$

   于是 $|x^2 - a^2| < \varepsilon$，所以 x^2 在 $x = a$ 处连续。 □

   请问：这个论证有问题吗？

7. 设 $f : \mathbb{R} \to \mathbb{R}$ 连续，且满足对任意 $x, y \in \mathbb{R}$，

$$f(x+y) = f(x) + f(y).$$

   证明对任意 $x \in \mathbb{R}$，$f(x) = f(1)x$。

8. 设 $f : \mathbb{R} \to \mathbb{R}$ 连续，且满足对任意 $x, y \in \mathbb{R}$，

$$f(\sqrt{x^2+y^2}) = f(x)f(y).$$

   证明对任意 $x \in \mathbb{R}$，$f(x) = (f(1))^{x^2}$。

9. 设 $f : \mathbb{R} \to \mathbb{R}$ 在任何有界闭区间上有界，且满足对任意 $x, y \in \mathbb{R}$，

$$f\left(\frac{x+y}{2}\right) \le \frac{f(x)+f(y)}{2}.$$

   证明 f 是连续函数。

2.2 函数在一点处的极限，函数的可去间断点

极限 $\lim\limits_{x\to a} f(x)$ 与 f 由靠近 a 但不是 a 的那些 x 所对应的函数值所决定，它反映了函数 f 在 a 附近的性质，它可以帮助我们更好地了解函数的连续性。

定义 2.2.1. 称 $\{x \in \mathbb{R} | 0 < |x-a| < \varepsilon\}$ 为 a 的 **ε-去心邻域**。称 $\{x \in I | 0 < |x-a| < \varepsilon\}$ 为 a **在 I 中的 ε-去心邻域**。

定义 2.2.2. 称 a 是 I 的一个**孤立点**，如果 $a \in I$ 且 a 的某个去心邻域中没有 I 的成员，即存在 $\delta_0 > 0$ 使得

$$x \in I \text{ 且 } |x-a| < \delta_0 \Rightarrow x = a.$$

称 $a \in \mathbb{R}$ 是集合 I 的一个**聚点**，如果对任意 $\varepsilon > 0$，存在 $x \in I$ 使得 $0 < |x-a| < \varepsilon$。即在 a 的任意去心邻域中都有 I 的成员。

简言之，I 的孤立点必然是 I 的成员，但不能被 I 的其他成员逼近；I 的聚点未必是 I 的成员，但可以被 I 的（其他）成员逼近。

例 2.2.3. 1. 区间 $(0,1)$ 的所有聚点组成的集合为 $[0,1]$。

2. 0 是集合 $\left\{ \frac{1}{n} \middle| n \in \mathbb{N}^* \right\}$ 的唯一聚点。

3. 任何实数都是 $\mathbb{Q}$ 的聚点。

4. $\mathbb{Z}$ 没有聚点。

5. 对非空集合 I，I 中的每个数要么是 I 的孤立点，要么是 I 的聚点。

由定义知，如果 a 是 I 的一个孤立点，则 $f: I \to \mathbb{R}$ 在 a 处连续。这是连续性的一种平凡情形。

连续性的非平凡情形与极限有关。函数 f 在其自变量 x 趋于 a 时的极限，由 f 的定义域中那些足够靠近 a 却不同于 a 的 x 所对应的函数值 $f(x)$ 决定。

定义 2.2.4. 设 $a \in \mathbb{R}$ 是集合 I 的一个聚点，称**当 x 趋于 a 时函数 $f: I \to \mathbb{R}$ 存在极限**（也称**收敛**），如果存在 $A \in \mathbb{R}$ 使得：对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得对任意 $x \in I$，只要 $0 < |x-a| < \delta$，就成立 $|f(x) - A| < \varepsilon$。此时称 A 为**当 x 趋于 a 时 $f: I \to \mathbb{R}$ 的一个极限**，记为 $\lim\limits_{x \to a} f(x) = A$。即

$$\lim_{x \to a} f(x) = A: \quad \forall \varepsilon > 0, \exists \delta > 0, \forall x \in I \big(0 < |x-a| < \delta \Rightarrow |f(x) - A| < \varepsilon\big).$$

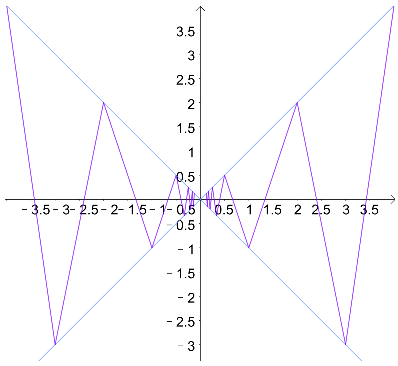
注意在上述定义中出现的"任意"、"存在"是不能随意交换先后顺序的。

定理 2.2.5. *若当 x 趋于 a 时函数 $f: I \to \mathbb{R}$ 存在极限，则极限唯一。*

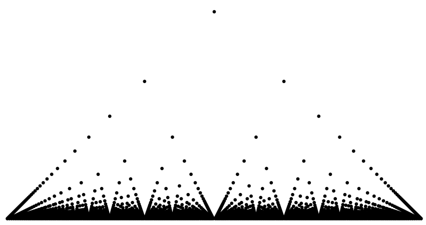
证明. (反证法) 设 A, B 都是当 x 趋于 a 时 f 的极限。假设 $A \neq B$。则存在 $\delta > 0$ 使得对任意 $x \in I$，只要 $0 < |x-a| < \delta$，就成立 $|f(x)-A| < \frac{|A-B|}{4}$，$|f(x)-B| < \frac{|A-B|}{4}$。于是

$$|A-B| \le |f(x)-A| + |f(x)-B| < \frac{|A-B|}{2},$$

矛盾。所以 $A = B$。 □



(a)



(b)

图 2.1:

下面这个例子表明，极限 $\lim_{x\to a} f(x)$ 与 f 是否在 a 有定义或 $f(a)$ 的值无关。

例 2.2.6. 证明：$\lim_{x\to 0} xg(x) = 0$，其中函数 $g: \mathbb{R}\backslash\{0\} \to \mathbb{R}$ 是一个奇函数，并且满足：对任意正整数 n，在区间 $[\frac{1}{n+1}, \frac{1}{n}]$ 上 $y = g(x)$ 是连接点 $(\frac{1}{n}, (-1)^n)$ 到点 $(\frac{1}{n+1}, (-1)^{n+1})$ 的直线段，在区间 $[n, n+1]$ 上 $y = g(x)$ 是连接点 $(n, (-1)^n)$ 到点 $(n+1, (-1)^{n+1})$ 的直线段。

证明. $xg(x)$ 的定义域为 $I = (-\infty, 0) \bigcup (0, +\infty)$，0 是 I 的聚点。对任意 $\varepsilon > 0$，取 $\delta = \varepsilon$，则 $0 < |x-0| < \varepsilon$ 时

$$|xg(x) - 0| = |x|\,|g(x)| \le |x| < \varepsilon,$$

所以 $\lim_{x\to 0} xg(x) = 0$。见图2.1(a)。 □

例 2.2.7. 对 Riemann 函数（见图2.1(b)）

$$R(x) = \begin{cases} \frac{1}{p}, & \text{若 } x = \frac{q}{p} \text{ 是既约分数;} \\ 0, & \text{若 } x \text{ 是无理数,} \end{cases}$$

这里称 $\frac{q}{p}$ 是既约分数是指，$p \in \mathbb{N}^*, q \in \mathbb{Z}$，$p, q$ 的最大的正整数公因子为 1。证明对任意 $a \in \mathbb{R}$，$\lim_{x\to a} R(x) = 0$。

证明. 对任意 $\varepsilon > 0$，取正整数 $N > \frac{1}{\varepsilon}$。

对分母不大于 N 的任意两个不同的分数 $\frac{q_1}{p_1}, \frac{q_2}{p_2}$，

$$\left|\frac{q_1}{p_1} - \frac{q_2}{p_2}\right| \geq \frac{1}{p_1 p_2} \geq \frac{1}{N^2}.$$

因为区间 $(a - \frac{1}{3N^2}, a + \frac{1}{3N^2})$ 长度小于 $\frac{1}{N^2}$，所以其中分母不超过 N 的既约分数最多只有一个，所以存在 $\delta_\varepsilon > 0$ 使得当 $0 < |x-a| < \delta_\epsilon$ 时，x 要么是一个无理数，要么是一个分母大于 N 的既约分数，无论哪种情况都有

$$0 \leq R(x) < \frac{1}{N} < \varepsilon.$$

所以 $\lim\limits_{x\to a} R(x) = 0$。 □

下述定理说明了极限与连续性之间的关系。

定理 2.2.8. *设 a 是 I 的聚点。则*

1. *若在 a 的一个去心邻域中 $f(x) = g(x)$，$\lim\limits_{x\to a} f(x) = A$，则 $\lim\limits_{x\to a} g(x) = A$。*
2. *$\lim\limits_{x\to a} f(x) = A$ 的充分必要条件是*

$$\tilde{f}(x) = \begin{cases} f(x), & x \in I\backslash\{a\}, \\ A, & x = a \end{cases}$$

   在 $a \in I$ 处连续。

3. *f 在 a 处连续当且仅当 $\lim\limits_{x\to a} f(x) = f(a)$。*

□

例 2.2.9. $\lim\limits_{x\to 1} \dfrac{x^2-3x+2}{x^2-x} = -1$。

证明. 对任何实数 x，只要 $0 < |x-1| < 1$，就有 $x \neq 0$ 且 $x \neq 1$，

$$\frac{x^2-3x+2}{x^2-x} = \frac{(x-1)(x-2)}{x(x-1)} = \frac{x-2}{x},$$

等式最右端在 $x = 1$ 处连续，故 $\lim\limits_{x\to 1} \frac{x^2-3x+2}{x^2-x} = \lim\limits_{x\to 1} \frac{x-2}{x} = \frac{1-2}{1} = -1$。见图2.2。

□

例 2.2.10. 设 $a > 0$，求 $\lim\limits_{x\to a} \dfrac{\sqrt{x}-\sqrt{a}}{x-a}$。

解： 对任何正数 $x \neq a$，$\frac{\sqrt{x}-\sqrt{a}}{x-a} = \frac{1}{\sqrt{x}+\sqrt{a}}$。等式右端在 $x = a$ 处连续，所以

$$\lim_{x\to a} \frac{\sqrt{x}-\sqrt{a}}{x-a} = \lim_{x\to a} \frac{1}{\sqrt{x}+\sqrt{a}} = \frac{1}{2\sqrt{a}}.$$

□

利用连续函数的四则运算和复合性质，我们得到如下结论，它们把复杂情况分解为一些较简单的情形。

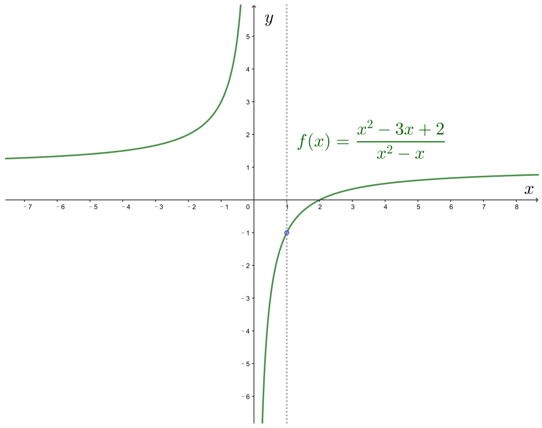


图 2.2:

定理 2.2.11 **(复合函数的极限).** 设 $\lim_{x\to a} f(x) = b$，g 在点 b 处连续。则 $\lim_{x\to a} g(f(x)) = g(b)$。 □

注 2.2.12. (1) 复合函数极限定理就是在极限计算中可以采取适当的换元。

(2) 上述定理能否把 g 的连续性条件改为 $\lim_{y\to b} g(y) = B$，并把结论改为 $\lim_{x\to a} g(f(x)) = B$？为什么？

定理 2.2.13. 设 $\lim_{x\to a} f(x) = A$，$\lim_{x\to a} g(x) = B$。则

1. $\lim_{x\to a} [f(x)+g(x)]$ 存在，且 $\lim_{x\to a} [f(x)+g(x)] = A+B$。
2. $\lim_{x\to a} [f(x)\cdot g(x)]$ 存在，且 $\lim_{x\to a} [f(x)\cdot g(x)] = A\cdot B$。
3. 若 $B \neq 0$，则 $\lim_{x\to a} \frac{f(x)}{g(x)}$ 存在，且 $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{A}{B}$。 □

利用函数之间的大小关系，我们也可以通过适当放缩，把把复杂函数极限的计算转化为简单函数的极限。

定理 2.2.14 **(夹挤定理).** 设 a 是 I 的聚点。若在 a 的一个去心邻域中总有

$$l(x) \leq f(x) \leq u(x),$$

且 $\lim_{x\to a} l(x) = \lim_{x\to a} u(x) = A$。则 $\lim_{x\to a} f(x)$ 存在且 $\lim_{x\to a} f(x) = A$。

证明. 设 $\delta_0 > 0$ 满足对任意 $x \in I$ 且 $0 < |x-a| < \delta_0$，都有

$$l(x) \leq f(x) \leq u(x).$$

对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得对任意 $x \in I$，只要 $0 < |x-a| < \delta$，就成立 $|l(x)-A| < \varepsilon$,$|u(x)-A| < \varepsilon$。则当 $x \in I$ 且 $0 < |x-a| < \min\{\delta, \delta_0\}$ 时，

$$-\varepsilon < l(x) - A \leq f(x) - A \leq u(x) - A < \varepsilon,$$

于是 $|f(x)-A|<\varepsilon$，因此 $\lim\limits_{x\to a} f(x)=A$。 □

在下面这个例子中，我们将得到三角函数的一些重要极限，利用它们我们可以得到三角函数的连续性和可微性。

例 2.2.15. 三角函数的连续性和重要极限

$$\lim_{x\to 0}\sin x=0,\quad \lim_{x\to 0}\cos x=1,\quad \lim_{x\to 0}\frac{\sin x}{x}=\lim_{x\to 0}\frac{\tan x}{x}=\lim_{x\to 0}\frac{1-\cos x}{x^2/2}=1.$$

解： 按三角函数的几何定义，在平面直角坐标系中，点 $(1,0)$ 沿半径长为 1 的

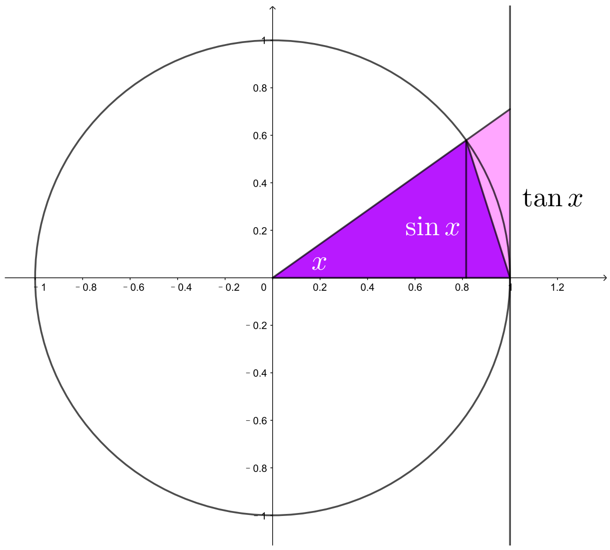


图 2.3:

圆周逆时针方向移动长度 x （当 $x<0$ 时，顺时针移动长度 $|x|$ ）时，处于坐标为 $(\cos x,\sin x)$ 的点处。半径长为 1 的圆周周长为 2π。

当 $0<x<\frac{\pi}{2}$ 时，由图2.3中等腰三角形、扇形、直角三角形所确定的三个区域的包含关系，以及相应面积的大小顺序，得到

$$0<\frac{\sin x}{2}<\frac{x}{2}<\frac{\tan x}{2}.$$

再由上述函数的奇偶性知，当 $0<|x|<\dfrac{\pi}{2}$ 时，

$$0<|\sin x|<|x|<|\tan x|,$$

而 $\lim\limits_{x\to 0}x=0$，由夹挤定理，$\lim\limits_{x\to 0}\sin x=0$。由 $\cos x=\sqrt{1-\sin^2 x}$ 以及极限的四则运算和复合函数极限性质，

$$\lim_{x\to 0}\cos x=\lim_{x\to 0}\sqrt{1-\sin^2 x}=\sqrt{1-\left(\lim_{x\to 0}\sin x\right)^2}=1.$$

由
$$0<\frac{\sin x}{x}<1<\frac{\tan x}{x}$$
得
$$\cos x<\frac{\sin x}{x}<1,$$
再由夹挤定理得 $\lim\limits_{x\to 0}\frac{\sin x}{x}=1$。由极限的四则运算性质以及复合函数极限性质得
$$\lim_{x\to 0}\frac{\tan x}{x}=\frac{\lim\limits_{x\to 0}\frac{\sin x}{x}}{\lim\limits_{x\to 0}\cos x}=1.$$
$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\lim_{x\to 0}\frac{2\sin^2\frac{x}{2}}{4\left(\frac{x}{2}\right)^2}=\frac{1}{2}\left(\lim_{x\to 0}\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2=\frac{1}{2}.$$
□

利用实数集的序结构，我们可以得到极限与序的关系。

定理 2.2.16 (**保序性和有界性**)**.** *设* $\lim\limits_{x\to a}f(x)=A$，$\lim\limits_{x\to a}g(x)=B$。*则*

1. *若* $A<B$。*则在* a *的某个去心邻域中总成立* $f(x)<g(x)$。
2. f *在* a *的某个去心邻域中有界。*
3. *若在* a *的某个去心邻域中总成立* $f(x)\le g(x)$，*则* $A\le B$。

证明. (1) 若 $A<B$，则存在 $\delta_0>0$ 使得对任意 $x\in I$ 且 $0<|x-a|<\delta_0$，
$$|f(x)-A|<\frac{B-A}{2},\qquad |g(x)-B|<\frac{B-A}{2},$$
所以
$$f(x)\le|f(x)-A|+A<\frac{A+B}{2}<B-|g(x)-B|<g(x).$$
(2) $A-1<\lim\limits_{x\to a}f(x)<A+1$，所以由(1)，存在 $\delta_0>0$ 使得对任意 $x\in I$ 且 $0<|x-a|<\delta_0$，$A-1<f(x)<A+1$。
(3) 由(1)和反证法知结论成立。 □

例 2.2.17. $\lim\limits_{x\to 0}\frac{1}{x}\sin\frac{1}{x}$ 不存在。

证明. 记 $f(x)=\frac{1}{x}\sin\frac{1}{x}$。对任意 $M>0$ 和任意 $\delta>0$，由阿基米德性质，存在正整数 n 使得 $2n\pi+\frac{\pi}{2}>M+\frac{1}{\delta}$。取 $x_n=\left(2n\pi+\frac{\pi}{2}\right)^{-1}$，所以 $0<x_n<\delta$，$f(x_n)=2n\pi+\frac{\pi}{2}>M$。所以 f 在 $x=0$ 的任意近旁无界。所以 $\lim\limits_{x\to 0}f(x)$ 不存在。 □

函数在聚点附近有界是存在极限的必要条件，但不是充分条件。

例 2.2.18. $\lim\limits_{x\to 0}\sin\frac{1}{x}$ 不存在。

证明. 假设 $A=\lim\limits_{x\to 0}\sin\frac{1}{x}$ 存在。则存在 $n\in\mathbb{N}^*$ 使得当 $0<|x|<\frac{1}{n}$ 时，$\sin\frac{1}{x}\in(A-\frac{1}{4},A+\frac{1}{4})$。令

$$x_n=\frac{2}{4n\pi+\pi},\qquad y_n=\frac{1}{2n\pi}.$$

则 $0<x_n<y_n<\frac{1}{n}$。所以

$$1=\left|\sin\frac{1}{x_n}-\sin\frac{1}{y_n}\right|<\left(A+\frac{1}{4}\right)-\left(A-\frac{1}{4}\right)=\frac{1}{2},$$

矛盾。因此 $\lim\limits_{x\to 0}\sin\frac{1}{x}$ 不存在。 □

习题2.2

1. 为观察极限 $\lim\limits_{x\to 1}\frac{x^2-3x+2}{x^2-x}$，请用Excel做数值计算，观察并思考其中出现的问题。

2. （有理分式与根式，换元及四则运算）求极限

   (1) $\lim\limits_{x\to 2}\frac{x^2-3x+2}{x^3+x^2-6x}$ (提示令 $h=x-2$)；

   (2) $\lim\limits_{x\to a}\frac{x^n-a^n}{x-a}$ (n 是正整数，提示令 $h=x-a$)；

   (3) $\lim\limits_{x\to a}\frac{x^{m/n}-a^{m/n}}{x-a}$ (m,n 是正整数，提示令 $y=\sqrt[n]{x}$)；

   (4) $\lim\limits_{x\to 0}\frac{\sqrt{1+x}-\sqrt[3]{1-x}}{x}$；

   (5) $\lim\limits_{x\to 4}\frac{\sqrt{1+2x}-3}{\sqrt{x}-2}$。

3. （三角函数，换元及四则运算）求极限

   (1) $\lim\limits_{x\to a}\sin x$，$\lim\limits_{x\to a}\cos x$ (提示令 $h=x-a$)；

   (2) $\lim\limits_{x\to a}\frac{\sin x-\sin a}{x-a}$ (提示令 $h=x-a$)；

   (3) $\lim\limits_{x\to 0}x\cot x$；

   (4) $\lim\limits_{x\to 0}\frac{\sqrt{\cos x}-\sqrt[3]{\cos x}}{\sin x^2}$；

   (5) $\lim\limits_{x\to 0}\frac{1-\cos x\sqrt{\cos 2x}\sqrt[3]{\cos 3x}}{x^2}$；

   (6) $\lim\limits_{x\to\frac{\pi}{3}}\frac{\tan^3 x-3\tan x}{\cos\left(x+\frac{\pi}{6}\right)}$ (提示令 $h=x-\frac{\pi}{3}$)。

4. 求极限

   (1) $\lim\limits_{x\to 0}x\left[\frac{1}{x}\right]$ ($[t]$ 表示不大于 t 的最大整数)；

   (2) $\lim\limits_{x\to 0}x\cos\frac{1}{x}$。

5. 设 $f:\mathbb{R}\to\mathbb{R}$ 是连续函数，$c>0$。证明

$$g:\mathbb{R}\to\mathbb{R},\quad g(x)=\begin{cases}c, & f(x)>c;\\ f(x), & |f(x)|\le c;\\ -c, & f(x)<-c\end{cases}$$

   是连续函数。

6. **三角函数的基本性质及其推论** 正弦和余弦 $\sin, \cos$ 是两个定义在整个 $\mathbb{R}$ 上的函数，满足：存在正数 π 使得

(a) $\cos 0 = \sin \frac{\pi}{2} = 1$，$\cos \pi = -1$；

(b) 对任意实数 x, y，

$$\cos(x-y) = \cos x \cos y + \sin x \sin y;$$

(c) 对任意 $0 < x < \frac{\pi}{2}$，

$$0 < \cos x < \frac{\sin x}{x} < \frac{1}{\cos x}.$$

利用性质(a)和(b)证明：

(1) 对任意实数 x，$\cos^2 x + \sin^2 x = 1$；

(2) $\sin 0 = \sin \pi = \cos \frac{\pi}{2} = 0$；

(3) 对任意实数 x，

$$\cos(-x) = \cos x, \quad \cos\left(\frac{\pi}{2} - x\right) = \sin x, \quad \cos(\pi - x) = -\cos x;$$

(4) 对任意实数 x，

$$\begin{aligned}
\sin(-x) &= -\sin x, \\
\sin\left(\frac{\pi}{2} + x\right) &= \cos x, \qquad \cos\left(\frac{\pi}{2} + x\right) = -\sin x, \\
\sin(\pi + x) &= -\sin x, \qquad \cos(\pi + x) = -\cos x, \\
\sin(2\pi + x) &= \sin x, \qquad \cos(2\pi + x) = \cos x;
\end{aligned}$$

(5) 对任意实数 x, y，

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

并且

$$\begin{aligned}
\sin x - \sin y &= 2 \sin \frac{x-y}{2} \cos \frac{x+y}{2}, \\
\cos x - \cos y &= -2 \sin \frac{x-y}{2} \sin \frac{x+y}{2};
\end{aligned}$$

(6) 对任意实数 x，

$$\begin{aligned}
\sin 2x &= 2 \sin x \cos x, \\
\cos 2x &= \cos^2 x - \sin^2 x \\
&= 2\cos^2 x - 1 = 1 - 2\sin^2 x;
\end{aligned}$$

利用(c)和(6)证明

(7) 对任意实数 $0 < x < \frac{\pi}{2}$，

$$\cos x < \frac{\sin x}{x} < 1;$$

利用(5)和(7)证明

(8) $\sin, \cos$ 是连续函数，并且 $\lim_{x\to 0} \frac{\sin x}{x} = 1$。

(9) 证明 $\cos$ 在区间 $[0,\pi]$ 上是减函数， $\sin$ 在区间 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 上是增函数。

2.3 单侧极限与间断，单调性与极限

因为实数集具有序结构，所以对于一元函数，其自变量与函数值的大小变化都只有两种可能（从小变大或从大变小），自变量与函数值之间大小关系的协调性体现为函数的单调性。这使得一元函数尤其是单调函数在连续性和极限行为方面有一些独特的性质。在这一节中，我们讨论函数的单侧极限与单侧连续性，特别对单调函数的连续性做深入讨论，最终给出指数函数、对数函数和幂函数的严格定义及连续性的证明。

定义 2.3.1 (函数的单侧极限与单侧连续性). 设函数 $f: I \to \mathbb{R}$ 在 $a \in \mathbb{R}$ 左侧任意近旁有定义，即对任意 $\delta > 0$， $I \cap (a-\delta, a) \neq \emptyset$。称 $A \in \mathbb{R}$ 为 f **在** a **左侧的极限**（记为 $\lim_{x\to a^-} f(x) = A$），如果对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得只要 $x \in I \cap (a-\delta, a)$，就有 $|f(x) - A| < \varepsilon$。

称 f 在 $a \in \mathbb{R}$ **左连续**，如果 $\lim_{x\to a^-} f(x) = f(a)$。

类似定义 f **在** a **右侧的极限**（记为 $\lim_{x\to a^+} f(x) = A$）以及 f 在 $a \in \mathbb{R}$ **右连续**。

左右极限也可以分别记为 $f(a-)$ 和 $f(a+)$。

易见对单侧极限，上述（双侧）极限所有的性质（唯一性、四则运算、复合函数、夹挤定理、保序性和有界性）都成立，读者可以自己陈述相应的结论并给出证明。

定理 2.3.2. *设 f 在 a 左右两侧任意近旁有定义。则*

1. *$\lim_{x\to a} f(x)$ 存在当且仅当单侧极限 $\lim_{x\to a^-} f(x)$ 和 $\lim_{x\to a^+} f(x)$ 都存在且相等。此时 $\lim_{x\to a} f(x) = \lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x)$。*

2. *f 在 a 处连续当且仅当 f 在 a 处既右连续又左连续。*

定义 2.3.3 (**函数的间断点**)**.** 设 $a \in \mathbb{R}$ 是 I 的聚点。称 $a \in \mathbb{R}$ 是函数 $f : I \to \mathbb{R}$ 的一个**间断点**，若 f 在 a 处无定义或不连续。

进一步，若极限 $\lim\limits_{x\to a} f(x)$ 存在，但 f 在 a 处无定义或者 $\lim\limits_{x\to a} f(x) \neq f(a)$，则称 $a \in \mathbb{R}$ 是函数 f 的一个**可去间断点**。

若单侧极限 $\lim\limits_{x\to a^-} f(x)$ 和 $\lim\limits_{x\to a^+} f(x)$ 都存在但不相等，则称 $a \in \mathbb{R}$ 是函数 f 的一个**跳跃间断点**。

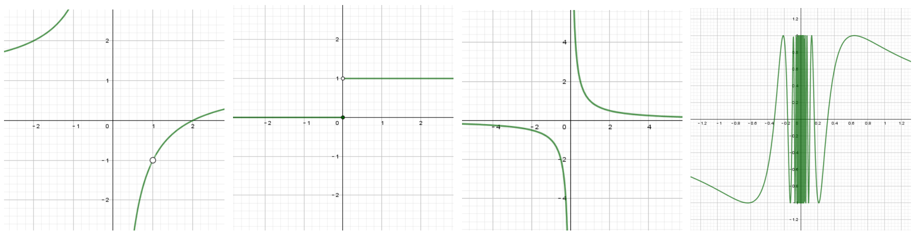
可去间断点和跳跃间断点统称为**第一类间断点**。

第一类间断点以外的间断点称为**第二类间断点**。

注 2.3.4. (1) f 在它的可去间断点 a 处不连续，是由于 f 在 a 处未定义或 $f(a)$ 的值不恰当所致，只要把 f 在 a 的值改为 $\lim\limits_{x\to a} f(x)$，即可让函数在 a 处变为连续。

(2) 如果 a 是 f 的跳跃间断点，则可以把 f 在 a 的值改为某一侧的单侧极限值，从而使函数在那一侧是连续的，但无论如何不可能让函数在 x_0 处连续。

(3) 如果 a 是 f 的第二类间断点，则在 a 处某一侧，无论怎样修改 f 在 a 的值都无法使函数 f 在这一侧连续。



例 2.3.5. (1) $x = 1$ 是 $\frac{x^2-3x+2}{x^2-x}$ 的可去间断点。

(2) $x = 0$ 是函数 $h(x) = \begin{cases} 1, & x > 0; \\ 0, & x \leq 0 \end{cases}$ 的跳跃间断点。

(3) $x = 0$ 是函数 $\frac{1}{x}$ 和 $\sin\frac{1}{x}$ 的第二类间断点。

细心的读者可能会注意到，到目前为止出现的所有涉及极限的结论，都是在某些极限存在的前提下得到的。下面这个结论与众不同，它给出了极限存在的一个充分条件，它是实数连续性的一个重要体现。

定理 2.3.6 (**单调有界 $\Rightarrow$ 单侧极限存在**)**.** *1. 设 f 在 a 左侧任意近旁有定义，且存在 $\delta_0 > 0$ 使得 f 在 $I \bigcap (a-\delta_0, a)$ 上单调不减且有上界（或单调不增且有下界），则 $\lim\limits_{x\to a^-} f(x)$ 存在。*

2. 设 f 在 a 右侧任意近旁有定义，且存在 $\delta_0 > 0$ 使得 f 在 $I \bigcap (a, a+\delta_0)$ 上单调不减且有下界（或单调不增且有上界），则 $\lim\limits_{x\to a^+} f(x)$ 存在。

3. 设 $f: I \to \mathbb{R}$ 是单调函数。若 f 在 a 左右两侧任意近旁有定义，则单侧极限 $\lim\limits_{x\to a^-} f(x)$ 和 $\lim\limits_{x\to a^+} f(x)$ 都存在，并且或者 f 在 a 处连续，或者 a 是 f 的跳跃间断点，或者 a 是 f 的可去间断点且 f 在 a 无定义。

证明. (1) 由确界公理，上确界 $A = \sup\limits_{x\in I\bigcap(a-\delta_0,a)} f(x)$ 存在。对任意 $\varepsilon > 0$，因为 $A-\varepsilon$ 不是 $\{f(x)|\, x\in I\bigcap(a-\delta_0,a)\}$ 的上界，所以存在 $x_1 \in I\bigcap(a-\delta_0,a)$ 使得 $f(x_1) > A-\varepsilon$。由 f 的单调性知，对任意 $x\in I\bigcap(x_1,a)$，$A-\varepsilon < f(x_1) \le f(x) \le A$。所以 $\lim\limits_{x\to a^-} f(x) = A$。

(2) 与 (1) 类似，留作练习。

(3) 不妨设 f 单调不减。任取 $x_1 < a < x_2$ 使得 $f(x_1), f(x_2)$ 都有定义。则对任意 $x \in I\bigcap(x_1,a)$ 和任意 $y \in I\bigcap(a,x_2)$，$f(x_1) \le f(x) \le f(y) \le f(x_2)$。于是由(1)和(2)，$\lim\limits_{x\to a^-} f(x)$ 和 $\lim\limits_{x\to a^+} f(x)$ 都存在。且由保序性知 $\lim\limits_{x\to a^-} f(x) \le \lim\limits_{x\to a^+} f(x)$。

如果 $\lim\limits_{x\to a^-} f(x) < \lim\limits_{x\to a^+} f(x)$，则 a 是 f 的跳跃间断点。

如果 $\lim\limits_{x\to a^-} f(x) = \lim\limits_{x\to a^+} f(x)$ 且 f 在 a 无定义，则 a 是 f 的可去间断点。

如果 $\lim\limits_{x\to a^-} f(x) = \lim\limits_{x\to a^+} f(x)$ 且 f 在 a 有定义，则对任意 $x \in I\bigcap(x_1,a)$ 和任意 $y \in I\bigcap(a,x_2)$，

$$f(x_1) \le f(x) \le f(a) \le f(y) \le f(x_2),$$

所以 $\lim\limits_{x\to a^-} f(x) \le f(a) \le \lim\limits_{x\to x_0^+} f(x)$。若 $\lim\limits_{x\to a^-} f(x) = \lim\limits_{x\to a^+} f(x)$，则

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

从而 f 在 a 处连续。 □

从 $\sin\frac{1}{x}$ 在 $x\to 0^+$ 和 $\frac{1}{x}$ 在 $x \to 0^+$ 的情况我们知道，仅靠有界性或仅靠单调性都不足以保证极限存在。

定义 2.3.7. 如果函数 $f: A\to B$ 是一一对应，则对任意 $y\in B$，存在唯一的 $x \in A$ 使得 $f(x) = y$，此时记 $x = f^{-1}(y)$，称 $f^{-1}: B \to A$ 为 $f: A\to B$ 的**反函数**。

定理 2.3.8 (**单调函数的连续性、反函数的连续性**)**.** 设 $I\subset\mathbb{R}$ 是区间，$f: I\to\mathbb{R}$ 是单调函数。则

1. f 连续当且仅当 $f(I)$ 是区间。

2. 若 f 是严格单调的连续函数，则 $f^{-1}: f(I)\to I$ 连续。

证明. 不妨设 f 单调不减。

(1) f 连续 $\Rightarrow f(I)$ 是区间：

对任意 $f(x_1), f(x_2) \in f(I)$ ($x_1, x_2 \in I$)以及任意 $y \in \big(f(x_1), f(x_2)\big)$，令 $A := \{x \in I | f(x) < y\}$。则 $x_1 \in A$，x_2 是 A 的上界。从而 A 有上确界 x_0。于是 $x_1 \leq y \leq x_2$，因为 I 是区间，所以 $x_0 \in I$。

对任意 $x \in I \cap (-\infty, x_0)$ 和任意 $z \in I \cap (x_0, +\infty)$，存在 $x_3 \in A$ 使得 $x < x_3 < x_0$，同时，$z \notin A$。于是

$$f(x) \leq f(x_3) < y \leq f(z), \quad f(x) \leq f(x_3) \leq f(x_0) \leq f(z),$$

从而 $\lim\limits_{x \to x_0^-} f(x) \leq \min\{y, f(x_0)\}$ 并且 $\lim\limits_{x \to x_0^+} f(x) \geq \max\{y, f(x_0)\}$。

因为 f 在 x_0 处连续，所以 $\lim\limits_{x \to x_0^-} f(x) = \lim\limits_{x \to x_0^+} f(x)$。于是 $\max\{y, f(x_0)\} \leq \min\{y, f(x_0)\}$，所以 $y = f(x_0) \in f(I)$。因此 $f(I)$ 是区间。

$f(I)$ 是区间 $\Rightarrow f$ 连续：

假设 f 在 $x_0 \in I$ 处不连续。则要么 $f(x_0) < \lim\limits_{x \to x_0^+} f(x)$，要么 $f(x_0) > \lim\limits_{x \to x_0^-} f(x)$。

若 $f(x_0) < \lim\limits_{x \to x_0^+} f(x) =: \alpha$，则对任意 $x_1, x_2 \in I, x_1 \leq x_0 < x_2$，成立不等式 $f(x_1) \leq f(x_0) < \alpha \leq f(x_2)$。从而 $\frac{f(x_0)+\alpha}{2} \notin f(I)$，但这与 $f(I)$ 是区间矛盾。

类似可证 $f(x_0) > \lim\limits_{x \to x_0^-} f(x)$ 也与 $f(I)$ 是区间矛盾。

所以 $\lim\limits_{x \to x_0^-} f(x) = f(x_0) = \lim\limits_{x \to x_0^+} f(x)$（当 x_0 是 区间 I 的端点时，这里有一个单侧极限不能定义），从而 f 在 x_0 处连续。

(2) 若 f 严格单调且连续，则 $f(I)$ 是区间且 $f^{-1} : f(I) \to I$ 严格单调。又 $f^{-1}(f(I)) = I$ 是区间，所以根据(1)，f^{-1} 连续。 $\square$

例 2.3.9 (乘方与开方的连续性). 对任何正整数 n，我们已知 x^n 是区间 $(0, +\infty)$ 到 $(0, +\infty)$ 的严格增的连续函数，因此其值域是一个区间。

对任意正数 y，取正整数 $m > y + \frac{1}{y}$，则 $m^n \geq m > y > \frac{1}{m} \geq \left(\frac{1}{m}\right)^n$，所以 y 在 x^n 的值域中。所以 x^n 的值域为 $(0, +\infty)$。

因此 x^n 的反函数 $x^{\frac{1}{n}}$ 是区间 $(0, +\infty)$ 到 $(0, +\infty)$ 上的严格增的连续函数。 $\square$

例 2.3.10 (对数函数的连续性，以及指数函数——作为对数函数的反函数——的连续性). 对任何正数 $a > 1$，我们已知对数函数 $\log_a : (0, +\infty) \to \mathbb{R}$ 是严格增的连续函数（见习题1.4和习题2.1），因此其值域是一个区间。

对任意实数 y，取正整数 $m > |y|$，则 $\log_a(a^m) = m > y > -m = \log_a(a^{-m})$，所以 y 在 $\log_a$ 的值域中。所以 $\log_a$ 的值域为 $\mathbb{R}$。

因此对数函数 $\log_a$ 的反函数——指数函数 a^x 是 $\mathbb{R}$ 到区间 $(0,+\infty)$ 上严格增的连续函数。

当 $0 < a < 1$ 时，对数函数 $\log_a = -\log_{1/a}$ 是从区间 $(0,+\infty)$ 到 $\mathbb{R}$ 上的严格减的连续函数，其反函数 a^x 是 $\mathbb{R}$ 到区间 $(0,+\infty)$ 上严格减的连续函数。 □

例 2.3.11 (**幂函数的连续性**). 对任何正数 x 和实数 μ，定义

$$x^\mu = 2^{\mu \log_2 x},$$

则 x^μ 是区间 $(0,+\infty)$ 上的连续函数。

当 $\mu > 0$ 时，x^μ 严格增，值域为 $(0,+\infty)$。

当 $\mu < 0$ 时，x^μ 严格减，值域为 $(0,+\infty)$。

当 $\mu = 0$ 时，$x^0 = 1$。 □

下面我们不借助对数函数，直接利用从有理数指数逼近的办法定义任意实数指数的指数函数，并证明它的连续性。

例 2.3.12 (**指数函数的另一定义以及连续性**). 设 $a > 1$，考虑 $f : \mathbb{Q} \to \mathbb{R}$，$f(x) = a^x$。证明

1. $f : \mathbb{Q} \to \mathbb{R}$ 是严格增函数；
2. 对任意 $x_0 \in \mathbb{R}$，极限 $\lim\limits_{x \to x_0} f(x)$ 都存在，记为 $\hat{f}(x_0)$；
3. f 在任何 $x_0 \in \mathbb{Q}$ 处连续，$\hat{f}(x_0) = f(x_0)$；
4. $\hat{f} : \mathbb{R} \to \mathbb{R}$ 是严格增函数；
5. $\hat{f} : \mathbb{R} \to \mathbb{R}$ 是连续函数。

因此可记 $a^x = \lim\limits_{r \to x} a^r$。

证明. (1) 对任意两个有理数 $r < s$，存在正整数 N 和整数 m, n 使得 $r = \frac{m}{N}, s = \frac{n}{N}$。从而 $n - m \in \mathbb{N}^*$。因为 x^N 严格增，$\left(a^{\frac{1}{N}}\right)^N = a > 1 = 1^N$，所以 $a^{\frac{1}{N}} > 1$。故

$$\frac{a^s}{a^r} = a^{\frac{n-m}{N}} = \left(a^{\frac{1}{N}}\right)^{n-m} > 1.$$

所以 $a^r < a^s$。所以 $f(x) = a^x$ 是 $\mathbb{Q}$ 上的严格增函数。

(2) 根据定理2.3.6，对任意 $x_0 \in \mathbb{R}$，$A = \lim\limits_{x \to x_0^-} f(x)$ 和 $B = \lim\limits_{x \to x_0^+} f(x)$ 都存在。因 f 严格增，所以 $A \le B$。对任意正整数 N，由有理数的稠密性，存在有理数 r, s 满足

$$x_0 - \frac{1}{N} < r < x_0.$$

于是 $x_0 < r + \frac{1}{N} < x_0 + \frac{1}{N}$，$a^r \le A \le B \le a^{r+\frac{1}{N}}$，所以 $1 \le \frac{B}{A} < \dfrac{a^{r+\frac{1}{N}}}{a^r} = a^{\frac{1}{N}}$，从而 $\left(\frac{B}{A}\right)^N < a$。根据 Bernoulli 不等式，

$$a > \left(\frac{B}{A}\right)^N > N\left(\frac{B}{A} - 1\right).$$

若 $A < B$，则 $N < \frac{a}{\frac{B}{A}-1}$，这与阿基米德性质矛盾。故 $A = B$，$\lim\limits_{x\to x_0} f(x)$ 存在。

(3) 由(2)及定理2.3.6(3)，f 在任何 $x_0 \in \mathbb{Q}$ 处连续。

(4) 设 $x, y \in \mathbb{R}$ 满足 $x < y$。由有理数的稠密性，存在有理数 r_0, s_0 使得 $x < r_0 < s_0 < y$。任取有理数 r, s，满足 $x < r < r_0 < s_0 < s < y$。则

$$a^r < a^{r_0} < a^{s_0} < a^s,$$

从而 $\hat{f}(x) = \lim\limits_{r\to x} a^r \le a^{r_0} < a^{s_0} \le \lim\limits_{s\to y} a^s = \hat{f}(y)$，$\hat{f}: \mathbb{R} \to \mathbb{R}$ 是严格增函数。

(5) 任取 $x_0 \in \mathbb{R}$。由(4)和定理2.3.6，$A = \lim\limits_{x\to x_0^-} \hat{f}(x)$ 和 $B = \lim\limits_{x\to x_0^+} \hat{f}(x)$ 都存在，且 $A \le B$。

因为 $a^r (r \in \mathbb{Q})$ 连续，所以对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得当 $x \in \mathbb{Q}$ 满足 $0 < x < \delta$ 时，$1 = a^0 < a^x < 1 + \varepsilon$。

取正整数 $N > \frac{1}{\delta}$，取有理数 r 满足 $x_0 - \frac{1}{N} < r < x_0$，于是

$$a^r = \hat{f}(r) < A \le B < \hat{f}(r + \frac{1}{N}) = a^{r+\frac{1}{N}},$$

从而

$$1 \le \frac{B}{A} < \frac{a^{r+\frac{1}{N}}}{a^r} = a^{\frac{1}{N}} < 1 + \varepsilon.$$

所以 $A = B$，再由定理2.3.6(3)知 $\hat{f}$ 在 x_0 处连续。 $\square$

对 $0 < a < 1$，可以类似定义 a^x 并得到相应的性质。

习题2.3

1. 指数函数的运算性质。证明对任意 $a, b > 0$ 和 $x, y \in \mathbb{R}$，$a^x \cdot a^y = a^{x+y}$，$(a^x)^y = a^{xy}$，$a^x \cdot b^x = (ab)^x$。

2. 幂函数的运算性质。证明对任意 $x, y > 0$ 以及任意 $\lambda, \mu \in \mathbb{R}$，

$$x^\mu y^\mu = (xy)^\mu, \quad x^\lambda x^\mu = x^{\lambda+\mu}, \quad (x^\mu)^\lambda = x^{\lambda\mu}.$$

3. 是否每个函数都可以写成一个单调不增函数与一个单调不减函数的和？

4. 设 A 是 I 的一个稠密子集。是否每个严格单调的连续函数 $f : A \to \mathbb{R}$ 都可以扩充为一个连续函数 $\hat{f} : I \to \mathbb{R}$，即 $\forall x \in A, \hat{f}(x) = f(x)$？

5. 设 $f:(a,b)\to\mathbb{R}$ 是一个有界函数，

$$g:(a,b)\to\mathbb{R},\quad g(x)=\sup_{a<t\leq x}f(t).$$

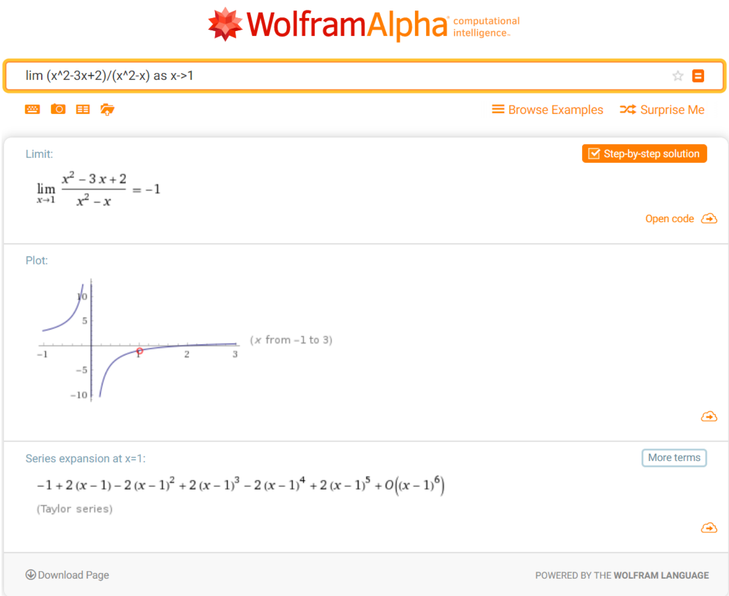
(1) 证明：若 f 是连续函数，则 g 是连续函数。

(2) 如果不假定 f 是连续函数，那么 g 是否为连续函数？是否具有单侧连续性？

6. 你能构造出一个单调有界的函数，使其具有尽可能多的间断点吗？

用WolframAlpha进行计算

访问www.wolframalpha.com网页，在对话框中输入“`lim (x^2-3x+2)/(x^2-x)`



`as x->1`”，然后按后面的“=”按钮或直接回车，就能得到 $\lim\limits_{x\to1}\frac{x^2-3x+2}{x^2-x}$ 的值以及函数 $\frac{x^2-3x+2}{x^2-x}$ 的图像和它在 $x\to1$ 时的渐近展开式。

2.4 无穷远、无穷小与无穷大，数列的极限

无穷远是从自变量的角度说的。

定义 2.4.1. 设 $A\in\mathbb{R}$，函数 $f:I\to\mathbb{R}$ 满足对任意 $N>0$，$I\bigcap[N,+\infty)\neq\emptyset$。$\lim\limits_{x\to+\infty}f(x)=A$ 表示：对任意 $\varepsilon>0$，存在 $N>0$ 使得对任意 $x\in I$，只要 $x>N$，就成立 $|f(x)-A|<\varepsilon$。这时我们称当 $x\to+\infty$ 时，$f(x)$ **收敛**。

类似定义 $\lim\limits_{x\to-\infty}f(x)=A$，请读者给出相应的陈述。

定义 2.4.2. 设函数 $f : I \to \mathbb{R}$ 满足对任意 $N > 0$，$I \bigcap (-\infty, -N] \neq \emptyset$ 且 $I \bigcap [N, +\infty) \neq \emptyset$。 $\lim\limits_{x\to\infty} f(x) = A$ 表示：对任意 $\varepsilon > 0$，存在 $N > 0$ 使得对任意 $x \in I$，只要 $|x| > N$，就成立 $|f(x) - A| < \varepsilon$。

无穷小和无穷大是从函数值的角度说的。

定义 2.4.3. 设 c 是一个实数或者为 $+\infty, -\infty, \infty$ 之一。

称当 $x \to c$ 时函数 f 是一个**无穷小**，如果 $\lim\limits_{x\to c} f(x) = 0$。

称当 $x \to c$ 时函数 f 是一个**无穷大**，如果 $\lim\limits_{x\to c} \frac{1}{f(x)} = 0$，此时记 $\lim\limits_{x\to c} f(x) = \infty$。

称当 $x \to c$ 时函数 f 是一个**正（负）无穷大**，如果当 $x \to c$ 时 f 是无穷大，且 $f(x) > 0$（ $f(x) < 0$），此时记 $\lim\limits_{x\to c} f(x) = +\infty$（ $\lim\limits_{x\to c} f(x) = -\infty$）。

数列 $\{a_n\}_{n\geq 1}$ 实际上是一个定义在正整数集 $\mathbb{N}^*$ 上的函数 $f : \mathbb{N}^* \to \mathbb{R}$，$f(n) = a_n$。所以上述定义适用于数列，但为了方便读者，我们还是把定义再叙述一遍。

定义 2.4.4. 称数列 $\{a_n\}$ **收敛**，如果存在实数 A 使得 $\lim\limits_{n\to+\infty} a_n = A$，即对任意 $\varepsilon > 0$，存在 $N > 0$ 使得对任意 $n > N$，就成立 $|a_n - A| < \varepsilon$。

称数列 $\{a_n\}$ 是一个**无穷小数列**，如果 $\lim\limits_{n\to+\infty} a_n = 0$。

称数列 $\{a_n\}$ 是一个**无穷大数列**，如果 $\lim\limits_{n\to+\infty} a_n = \infty$。

称数列 $\{a_n\}$ 是一个**正（负）无穷大数列**，如果 $\lim\limits_{n\to+\infty} a_n = +\infty$（ $\lim\limits_{n\to+\infty} a_n = -\infty$）。

现在，在符号 $\lim\limits_{x\to c} f(x) = A$ 中，c 和 A 都可以是一个实数，也可以是 $+\infty, -\infty, \infty$ 中的任何一个，c 还可以是某个实数的左侧或右侧。这样这个符号有二十多种组合情形。为了便于统一讨论，我们引入“邻域”和“去心邻域”的概念。

定义 2.4.5. 对 $a \in \mathbb{R}$，称 $V \subset \mathbb{R}$ 为 a 的一个**邻域**，如果 V 包含一个含 a 的开区间；此时称 $V \backslash \{a\}$ 为 a 的一个**去心邻域**。

任何以 a 为右端点（左端点）的开区间称为 a 的一个**左侧去心邻域**（**右侧去心邻域**）。

称任何形如$(N,+\infty)$（$N\in\mathbb{R}$）的开区间为 $+\infty$ 的一个**邻域**。

称任何形如$(-\infty,N)$（$N\in\mathbb{R}$）的开区间为 $-\infty$ 的一个**邻域**。

称任何形如$(-\infty,-N)\bigcup(N,+\infty)$（$N>0$）的开区间为 ∞ 的一个**邻域**。$+\infty,-\infty,\infty$ 的邻域也称为它们的**去心邻域**。

注 2.4.6.

1. 有限多个邻域的交集仍是邻域，有限多个去心邻域的交集仍是去心邻域。
2. $\lim\limits_{x\to c}f(x)=A$，即对 A 的任何邻域 W，存在 c 的去心邻域 V 使得 $f(V)\subseteq W$。
3. f 在 x_0 处连续，即对 $f(x_0)$ 的任何邻域 W，存在 x_0 的邻域 V 使得 $f(V)\subseteq W$。

例 2.4.7. $\lim\limits_{x\to c}f(x)=\infty$ 当且仅当对任意 $M>0$，存在 c 的一个去心邻域 V 使得对任意 $x\in V$，都有 $|f(x)|>M$。

证明. $\lim\limits_{x\to c}f(x)=\infty$ 当且仅当 $\lim\limits_{x\to c}\frac{1}{f(x)}=0$，后者当且仅当对任意 $\varepsilon>0$，存在 c 的一个去心邻域 V 使得对任意 $x\in V$，都有 $\left|\frac{1}{f(x)}\right|<\varepsilon$，即 $|f(x)|>\frac{1}{\varepsilon}$。

对任意 $M>0$，取 $\varepsilon=1/M$，于是对任意 $x\in V$，都有 $|f(x)|>M$。

反之，对任意 $\varepsilon>0$，取 $M=1/\varepsilon$，于是对任意 $x\in V$，都有 $\left|\frac{1}{f(x)}\right|<\varepsilon$。

$W=\{y\in\mathbb{R}||y|>M\}$ 是 ∞ 的一个邻域。 □

例 2.4.8. 设 $\alpha>0$。则 $\lim\limits_{x\to+\infty}x^\alpha=+\infty$，$\lim\limits_{x\to0^+}x^\alpha=0$。

证明. 取正整数 n 使得 $0<\frac{1}{n}<\alpha$。对任意 $N\geq1$，当 $x>N^n$ 时，利用指数函数与幂函数的单调性，$x^\alpha>x^{\frac{1}{n}}>(N^n)^{\frac{1}{n}}=N$，所以 $\lim\limits_{x\to+\infty}x^\alpha=+\infty$。

当 $0<x<\frac{1}{N^n}$ 时，$\left(\frac{1}{x}\right)^\alpha>N$，$0<x^\alpha<\frac{1}{N}$。所以 $\lim\limits_{x\to0^+}x^\alpha=0$。 □

例 2.4.9. 设 $a>1$。证明 $\lim\limits_{x\to-\infty}a^x=0$，$\lim\limits_{x\to+\infty}a^x=+\infty$。

证明. 对任意 $\varepsilon>0$，取正整数 $N>\frac{1}{\varepsilon(a-1)}$。由 Bernoulli 不等式，

$$a^N=(1+a-1)^N>N(a-1)>\frac{1}{\varepsilon}.$$

所以对任意 $x<-N$，$0\leq a^x<a^{-N}=\frac{1}{a^N}<\varepsilon$。所以 $\lim\limits_{x\to-\infty}a^x=0$。

对任意 $x>N$，$a^x>a^N$，所以 $0<\frac{1}{a(x)}<\frac{1}{a^N}<\varepsilon$，因此 $\lim\limits_{x\to+\infty}a^x=+\infty$。

□

当极限值为实数时，上节中极限的唯一性、四则运算、夹挤定理、保序性和有界性、单调有界收敛的结论对在无穷远处的极限（包括数列极限）也都成立。请读者给出相应的陈述和证明。

相比于连续变量的函数，数列的单调性往往可以用初等的办法得到。

例 2.4.10. 证明数列 $\left(1+\frac{1}{n}\right)^n$ 和 $\sum\limits_{k=0}^{n}\frac{1}{k!}$ 都有极限，且极限值相等。记 e 为它们共同的极限。

证明. 记 $b_n=\sum\limits_{k=0}^{n}\frac{1}{k!}$。易见 b_n 单调增，且对 $n\geq 2$，

$$b_n=1+1+\sum_{k=2}^{n}\frac{1}{k!}<1+1+\sum_{k=2}^{n}\frac{1}{(k-1)k}=3-\frac{1}{n}<3.$$

因此 $\lim\limits_{n\to+\infty}b_n$ 存在，记 $\mathrm{e}=\lim\limits_{n\to+\infty}b_n$。

记 $a_n=\left(1+\frac{1}{n}\right)^n$。由习题1.3-3(d)的结论知 $\{a_n\}_{n\geq 1}$ 单调增。

对任意正整数 n，利用二项式展开得到

$$a_n=\left(1+\frac{1}{n}\right)^n=1+\sum_{k=1}^{n}\frac{n\cdots(n-k+1)}{k!}\frac{1}{n^k}\leq 1+\sum_{k=1}^{n}\frac{1}{k!}=b_n<\mathrm{e},$$

所以 $\{a_n\}_{n\geq 1}$ 有上界。从而 $\alpha=\lim\limits_{n\to+\infty}a_n$ 存在，并且 $\alpha\leq\mathrm{e}$。

对任意正整数 N 以及任意正整数 $n\geq N$，

$$a_n=1+\sum_{k=1}^{n}\frac{n\cdots(n-k+1)}{k!}\frac{1}{n^k}\geq 1+\sum_{k=1}^{N}\frac{n\cdots(n-k+1)}{n^k}\frac{1}{k!},$$

对上式最左和最右两端同时让 $n\to+\infty$ 取极限，由极限的保序性和四则运算性质（这时 N 是固定的）得到

$$\alpha\geq 1+\sum_{k=1}^{N}\frac{1}{k!}=b_N.$$

再让 $N\to+\infty$，得到 $\alpha\geq\mathrm{e}$。因此 $\lim\limits_{n\to+\infty}a_n=\alpha=\mathrm{e}$。 □

请读者思考，上述证明最后一段为什么要分两次取极限。

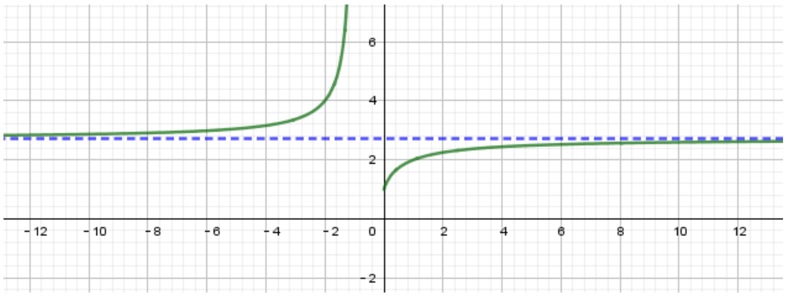
例 2.4.11. 证明 $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x=\mathrm{e}$。

证明. 因为 $\lim\limits_{n\to+\infty}\left(1+\frac{1}{n}\right)^n=\mathrm{e}$，所以

$$\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^{n+1}=\mathrm{e},\quad \lim_{n\to+\infty}\left(1+\frac{1}{n+1}\right)^{n}=\mathrm{e}.$$

对任意 $\varepsilon>0$，存在正整数 N 使得对任意正整数 $n\geq N$，

$$\left(1+\frac{1}{n}\right)^{n+1}<\mathrm{e}+\varepsilon,\quad \left(1+\frac{1}{n+1}\right)^{n}>\mathrm{e}-\varepsilon.$$



当 $x \geq N$ 时，存在正整数 $n \geq N$ 使得 $n \leq |x| < n+1$，因此

$$\mathrm{e} - \varepsilon < \left(1 + \frac{1}{n+1}\right)^n \leq \left(1 + \frac{1}{x}\right)^x \leq \left(1 + \frac{1}{n}\right)^{n+1} < \mathrm{e} + \varepsilon.$$

因此 $\lim\limits_{x \to +\infty} \left(1 + \frac{1}{x}\right)^x = \mathrm{e}$。

对 $y < 0$，

$$\lim_{y \to -\infty} \left(1 + \frac{1}{y}\right)^y \overset{x=-y}{=} \lim_{x \to +\infty} \left(1 - \frac{1}{x}\right)^{-x} = \lim_{x \to +\infty} \left(1 + \frac{1}{x-1}\right)^x = \mathrm{e}.$$

因此 $\lim\limits_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = \mathrm{e}$。 □

注 2.4.12. 极限 $\lim\limits_{x \to \infty} \left(1 + \frac{1}{x}\right)^x$ 对进一步研究指数函数和对数函数特别重要，但其存在性却不容易直接得到。因此我们先考虑它的一个特殊情况，即 x 是正整数的情况。这时，数列 $\left(1 + \frac{1}{n}\right)^n$ 的单调性可以用上面的初等方法得到，该数列的有界性可借助另一个数列 $b_n = \sum\limits_{k=0}^{n} \frac{1}{k!}$ 得到证明。这两个数列具有相同的极限值，在实际计算中，前者要重新计算乘方，后者只需进行累乘和累加，从运算复杂性和控制误差两个方面看，后者在计算 e 的近似值时都有着明显的优势。

上例证明的最后一步用到了极限换元，也就是复合函数的极限，其一般结论如下。

定理 2.4.13 (复合函数的极限). *设 $\lim\limits_{x \to c} f(x) = A$，$\lim\limits_{y \to A} g(y) = B$，其中 c, A, B 可以是某个实数或无穷大 $\infty, +\infty, -\infty$，或某个实数的某一侧。则在下述三个条件之一成立时，都有 $\lim\limits_{x \to c} g(f(x)) = B$：*

1. *$A = \infty, +\infty, -\infty$；*
2. *A 是实数或实数的某侧，且 $f(x)$ 总在 A 的去心邻域中；*
3. *A 是实数或实数的某侧，g 在 A 处连续。*

证明. 我们只证明前两种情形，第三种情形的证明留给读者自己完成。

对前两种情形，都存在 c 的去心邻域 W_0 和 A 的去心邻域 V_0，使得对任意 $x \in W_0$，$f(x) \in V_0$。

由 $\lim\limits_{y\to A} g(y) = B$ 知，对 B 的任意邻域 U，存在 A 的去心邻域 V 使得 $g(V) \subseteq U$。

若 A 是无穷远，则取 $V_1 = V$；若 A 是有限实数，则取 $V_1 = V \cup \{A\}$。于是 V_1 是 A 的邻域。

由 $\lim\limits_{x\to c} f(x) = A$ 知，对 A 的邻域 V_1，存在 c 的去心邻域 W 使得 $f(W) \subseteq V_1$。

于是对任意 $x \in W_0 \cap W$，$f(x) \in V_0 \cap V_1 \subseteq V$，从而 $g(f(x)) \in U$。 □

例 2.4.14 (与指数函数、对数函数有关的重要极限).

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}} = \mathrm{e}, \quad \lim_{x\to 0}\frac{\ln(1+x)}{x} = 1, \quad \lim_{x\to 0}\frac{\mathrm{e}^x - 1}{x} = 1.$$

证明. 由复合函数极限定理知，

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}} \overset{t=\frac{1}{x}}{=} \lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t = \mathrm{e}.$$

由复合函数极限定理以及对数函数 $\ln = \log_{\mathrm{e}}$ 的连续性知

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = \lim_{x\to 0}\ln(1+x)^{\frac{1}{x}} = \ln\mathrm{e} = 1.$$

由复合函数极限定理以及极限的四则运算性质知

$$\lim_{x\to 0}\frac{\mathrm{e}^x-1}{x} \overset{y=\mathrm{e}^x-1}{=} \lim_{y\to 0}\frac{y}{\ln(1+y)} = \frac{1}{\lim\limits_{y\to 0}\dfrac{\ln(1+y)}{y}} = 1.$$

□

我们可以通过换元，把无穷远处的极限变为有限点处的极限。

例 2.4.15. 证明 $\lim\limits_{x\to\infty}\sqrt{\frac{x^2+1}{x^2-1}} = 1$。

证明. 不难用定义证明 $\lim\limits_{x\to\infty}\frac{1}{x^2} = 0$。因为

$$\sqrt{\frac{x^2+1}{x^2-1}} = \sqrt{\frac{1+t}{1-t}}, \quad t = \frac{1}{x^2},$$

而且 $\sqrt{\frac{1+t}{1-t}}$ 在 $t=0$ 处连续，所以 $\lim\limits_{x\to\infty}\sqrt{\frac{x^2+1}{x^2-1}} = 1$。 □

上述定理有以下推论。

推论 2.4.16. 设 $\lim\limits_{n\to+\infty} a_n = A$，则对任意单射 $f:\mathbb{N}^*\to\mathbb{N}^*$，都有 $\lim\limits_{k\to+\infty} a_{f(k)} = A$。

当 f 严格增时，称数列 $\{a_{f(k)}\}_{k\geq 1}$ 为 $\{a_n\}_{n\geq 1}$ 的**子列**。因此：若 $\lim\limits_{n\to+\infty} a_n = A$，则 $\{a_n\}_{n\geq 1}$ 的任何子列也满足 $\lim\limits_{k\to+\infty} a_{f(k)} = A$。

证明. 对任意正整数 M，集合 $f^{-1}\{1,2,\ldots,M\}$ 是有限集，取

$$N_M := \min \mathbb{N}^*\backslash f^{-1}\{1,2,\ldots,M\},$$

于是对任意 $n \geq N_M$，$f(n) > M$。因此 $\lim\limits_{n\to+\infty} f(n) = +\infty$。

所以由复合函数极限有 $\lim\limits_{k\to+\infty} a_{f(k)} = A$。 □

定理 2.4.17 (Heine). *函数 g 在y_0 处连续当且仅当对 g 的定义域中任意满足 $\lim\limits_{n\to+\infty} y_n = y_0$ 的数列$\{y_n\}_{n\geq1}$，都有 $\lim\limits_{n\to+\infty} g(y_n) = g(y_0)$。*

证明. 必要性部分是前述定理的推论。下证充分性部分：设任意满足 $\lim\limits_{n\to+\infty} y_n = y_0$ 的数列 $\{y_n\}_{n\geq1}$，都有 $\lim\limits_{x\to+\infty} g(y_n) = g(y_0)$。假设 g 在 y_0 不连续，则存在 $\varepsilon_0 > 0$ 使得对任意正整数 n 都存在 y_n 使得 $|y_n - y_0| < \frac{1}{n}$，$|g(y_n) - g(y_0)| \geq \varepsilon_0$。于是 $\lim\limits_{n\to+\infty} y_n = y_0$，但是 $\lim\limits_{n\to+\infty} g(y_n) = g(y_0)$ 不成立。这与已知矛盾。 □

例 2.4.18. 证明 Dirichlet 函数

$$D(x) = \begin{cases} 1, & \text{若 } x \text{ 是有理数;} \\ 0, & \text{若 } x \text{ 是无理数} \end{cases}$$

处处间断。

证明. 对任意 $a \in \mathbb{R}$，由于有理数的稠密性，所以对于任何正整数 n，存在有理数 x_n 使得 $\frac{1}{n+1} < |x_n - a| < \frac{1}{n}$。$\lim\limits_{n\to+\infty} x_n = a$，$\lim\limits_{n\to+\infty} D(x_n) = 1$。

另一方面，取 $y_n = x_n + \frac{\sqrt{2}}{4n+4}$，则 y_n 都是无理数，满足 $|y_n - a| \leq |x_n - a| + |x_n - y_n| \leq \frac{3}{n}$，$\lim\limits_{n\to+\infty} y_n = a$，$\lim\limits_{n\to+\infty} D(y_n) = 0$。

所以 D 在 a 处间断。

同时

$$|y_n - a| \geq |x_n - a| - |x_n - y_n| > \frac{1}{n+1} - \frac{\sqrt{2}}{4n+4} > \frac{1}{2n+2} > |x_n - y_n|.$$

易见我们可以取 x_n 位于 a 的某一侧，则 y_n 与 x_n 位于 a 的同侧。因此 $\lim\limits_{x\to a^+} D(x)$ 和 $\lim\limits_{x\to a^-} D(x)$ 都不存在，所以 a 是 D 的第二类间断点。 □

对于无穷大量，有以下结论

定理 2.4.19. *设 $\lim\limits_{x\to c} f(x) = (\pm)\infty$。*

1. 若 $|g(x)|$ 有正下界，则

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

其正负由 f 和 g 的正负决定。

2. 若 g 是一个有界函数，则

$$\lim_{x\to c}\big(f(x)+g(x)\big)=(\pm)\infty.$$

3. 若 $\lim\limits_{x\to c} f(x)=+\infty$，且 $f(x)\le g(x)$，则 $\lim\limits_{x\to c} g(x)=+\infty$。

证明. (1) 设 $|g(x)|\ge m>0$。则

$$\left|\frac{1}{f(x)g(x)}\right|\le\frac{1}{m}\left|\frac{1}{f(x)}\right|\to 0,\quad x\to c,$$

所以

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

(2) 设 $|g(x)|\le M$。于是

$$\left|\frac{g(x)}{f(x)}\right|\le M\cdot\left|\frac{1}{f(x)}\right|\to 0,\quad x\to c,$$

所以

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

因此

$$\frac{1}{f(x)+g(x)}=\frac{1}{f(x)}\cdot\frac{1}{1+\frac{g(x)}{f(x)}}\to 0,\quad x\to c,$$

故 $\lim\limits_{x\to c}\frac{1}{f(x)+g(x)}=0$， $\lim\limits_{x\to c}\big(f(x)+g(x)\big)=\infty$，且 $f(x)+g(x)$ 与 $f(x)$ 具有相同的正负号。

(3) 读者可以自行给出证明。 □

定义 2.4.20. 如果 $a\in\mathbb{R}$ 并且 $\lim\limits_{x\to a} f(x)=\infty$，则称直线 $x=a$ 为函数图像 $y=f(x)$ 在 a 处的**竖直渐近线**。

定义 2.4.21. 如果 $\lim\limits_{x\to\pm\infty}(f(x)-(kx+b))=0$，则称直线 $y=kx+b$ 为函数图像 $y=f(x)$ 在 $x\to\pm\infty$ 时的**斜渐近线**。若其中 $k=0$，这渐近线又称为**水平渐近线**。

易见，$y=kx+b$ 为函数图像 $y=f(x)$ 在 $x\to\pm\infty$ 时的渐近线当且仅当

$$k=\lim_{x\to\pm\infty}\frac{f(x)}{x},\quad b=\lim_{x\to\pm\infty}\big(f(x)-kx\big).$$

例 2.4.22. 设 $f(x)=\sqrt{\frac{x^3}{x-1}}$。求 $y=f(x)$ 的渐近线。

解： f 的定义域为 $(-\infty,0]\bigcup(1,+\infty)$，$f:(-\infty,0]\bigcup(1,+\infty)\to\mathbb{R}$ 为连续函数。

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

所以 $x=1$ 是 $y=f(x)$ 在 $x\to 1^+$ 时的竖直渐近线。

当 $x>1$ 时，

$$\sqrt{\frac{x^3}{x-1}}=x\sqrt{\frac{1}{1-\frac{1}{x}}},\quad \lim_{x\to+\infty}\sqrt{\frac{1}{1-\frac{1}{x}}}=1,$$

$$\sqrt{\frac{x^3}{x-1}}-x=x\left[\sqrt{\frac{1}{1-\frac{1}{x}}}-1\right]=\frac{x\left(1-\left(1-\frac{1}{x}\right)\right)}{\sqrt{1-\frac{1}{x}}\left(1+\sqrt{1-\frac{1}{x}}\right)}\to\frac{1}{2},\quad x\to+\infty,$$

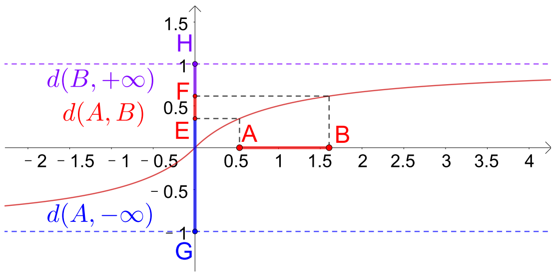
所以 $y=x+\frac{1}{2}$ 是函数图象 $y=f(x)$ 在 $x\to+\infty$ 时的斜渐近线。

类似可证 $\lim\limits_{x\to+\infty}\left(\sqrt{\frac{x^3}{x-1}}+x\right)=-\frac{1}{2}$，所以 $y=-x-\frac{1}{2}$ 是函数图象 $y=f(x)$ 在 $x\to-\infty$ 时的斜渐近线。 □

习题2.4

1. 记 $\overline{\mathbb{R}}$ 为实数集 $\mathbb{R}$ 添加两个元素 $-\infty,+\infty$ 所得到的集合。考虑 $\overline{\mathbb{R}}$ 上的函数

$$h(x)=\begin{cases}\frac{x}{1+|x|}, & x\in\mathbb{R};\\ 1, & x=+\infty;\\ -1, & x=-\infty.\end{cases}$$



(a) 证明：$h|_{\mathbb{R}}$ 是 $\mathbb{R}$ 到区间 $(-1,1)$ 上的严格增的连续函数。

(b) 证明：

$$d(x,y)=|h(x)-h(y)|,\quad x,y\in\overline{\mathbb{R}}$$

是 $\overline{\mathbb{R}}$ 上的一个距离。

(c) 证明：对任意 $x,y\in\mathbb{R}$，$d(x,y)\le 2|x-y|$。

(d) 设 $x_0 \in \overline{\mathbb{R}}$，$A \in \mathbb{R}$。证明：$\lim\limits_{x \to x_0} f(x) = A$（按通常含义）当且仅当对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得对任意 $x \in \mathbb{R}$，只要 $0 < d(x, x_0) < \delta$，就有 $|f(x) - A| < \varepsilon$。后者我们记为 $\lim\limits_{d(x,x_0) \to 0} f(x) = A$。

(e) 对任意 $x_0 \in \overline{\mathbb{R}}$，$x_0$ 在 $\mathbb{R}$ 中的去心邻域就是 x_0 在 $\overline{\mathbb{R}}$ 中按距离 d 给出的去心邻域。

(f) 对任意 $x_0, A \in \overline{\mathbb{R}}$，$\lim\limits_{x \to x_0} f(x_0) = A$ 当且仅当对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得只要 $0 < d(x, x_0) < \delta$，就有 $d(f(x), A) < \varepsilon$。

2. 关于极限 $\lim\limits_{x \to \infty} \sqrt{\frac{x^2+1}{x^2-1}}$，有人认为：当 x 充分大时，$x^2 + 1 \approx x^2$，$x^2 - 1 \approx x^2$，所以 $\sqrt{\frac{x^2+1}{x^2-1}} \approx 1$，从而 $\lim\limits_{x \to \infty} \sqrt{\frac{x^2+1}{x^2-1}} = 1$。请问这样的说法成立吗？为什么？作为一个对照，请讨论极限 $\lim\limits_{x \to +\infty} \left(\sqrt{x^2 + x^\alpha} - x\right)$，其中 $0 < \alpha < 2$。

3. 设 $0 < a < b$，$f(x) = \left(\frac{a^x + b^x}{2}\right)^{\frac{1}{x}}$。求极限 $\lim\limits_{x \to -\infty} f(x)$ 和 $\lim\limits_{x \to +\infty} f(x)$。

4. 叙述并证明数列情形的单调有界收敛定理。

5. (p **进制小数**) 设 $p > 1$ 是正整数。对任何实数 x，令

$$a_0 = \lfloor x \rfloor, \quad x_0 = a_0,$$

对正整数 n，记

$$a_n = \lfloor p^n (x - x_{n-1}) \rfloor, \quad x_n = x_{n-1} + \frac{a_n}{p^n}.$$

证明对每个正整数 n，a_n 是小于 p 的非负整数，且

$$x_n \le x < x_n + \frac{1}{p^n}.$$

从而 $\lim\limits_{n \to +\infty} x_n = x$。这样，每个实数 x 与唯一一个整数数列 $\{a_n\}_{n \ge 0}$ 对应，后者满足 $a_0 \in \mathbb{Z}$，$a_n \in \{0, 1, \ldots, p-1\} (n \ge 1)$，且

$$a_0 + \frac{a_1}{p} + \frac{a_2}{p^2} + \cdots + \frac{a_n}{p^n} \le x < a_0 + \frac{a_1}{p} + \frac{a_2}{p^2} + \cdots + \frac{a_n + 1}{p^n}.$$

对非负实数 x，$a_0.a_1a_2 \cdots a_n \cdots$ 就是 x 的 p 进制小数展开。

6. 设 $f : (0, +\infty) \to (0, +\infty)$ 满足：若 $a_0.a_1a_2a_3 \cdots$ 是 x 的二进制小数展开，则

$$f(x) = a_0.0a_10a_20a_3 \cdots.$$

求 f 的所有间断点。

7. 设正整数 $m \geq 3$，a_n 和 b_n 分别是半径为 1 的圆的内接和外切正 $2^n m$ 边形的边长，p_n 和 q_n 是相应的周长。

(1) 证明

$$a_{n+1} = \sqrt{2 - \sqrt{4 - a_n^2}}, \quad p_{n+1} = 2^{n+1} m \sqrt{2 - \sqrt{4 - \frac{p_n^2}{2^{2n} m^2}}},$$

并计算圆的周长的近似值。

(2) 证明

$$a_{n+1} = \frac{a_n}{2\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{a_n^2}{4}}}}, \quad p_{n+1} = \frac{p_n}{\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{p_n^2}{2^{2n+2} m^2}}}},$$

并计算圆的周长的近似值。

(3) 证明

$$b_{n+1} = \frac{b_n}{1 + \sqrt{1 + \frac{b_n^2}{4}}}, \quad q_{n+1} = \frac{2q_n}{1 + \sqrt{1 + \frac{q_n^2}{2^{2n+2} m^2}}},$$

并计算圆的周长的近似值。

(4) 证明

$$\frac{a_n}{2} = \sin\frac{\theta_n}{2}, \quad \frac{b_n}{2} = \tan\frac{\theta_n}{2}, \quad \theta_n = \frac{2\pi}{2^n m},$$

$$a_{n+1} = \sqrt{\frac{a_n b_{n+1}}{2}}, \quad b_{n+1} = \frac{a_n b_n}{a_n + b_n},$$

$$p_{n+1} = \sqrt{p_n q_{n+1}}, \quad q_{n+1} = \frac{2 p_n q_n}{p_n + q_n}.$$

并计算圆的周长的近似值。

(5) 通过数值计算，观察并比较以上算法的逼近效果，解释其中出现的现象。

(6) 证明 $p_n < p_{n+1} < q_{n+1} < q_n$，从而极限 $\lim\limits_{n\to+\infty} p_n$，$\lim\limits_{n\to+\infty} q_n$ 都存在。

(7) 证明 $\lim\limits_{n\to+\infty} p_n = \lim\limits_{n\to+\infty} q_n$。

8. 用 Excel 计算以下数列 $\left(1+\frac{1}{n}\right)^n$，$\left(1+\frac{1}{n}\right)^{n+1}$，$\left(1+\frac{1}{n}\right)^{n+\lambda}$($0 < \lambda < 1$)，$\sum\limits_{k=0}^{n} \frac{1}{k!}$，观察它们的单调性并比较它们的收敛速度。

9. (1) 证明数列 $c_n = 1 + \frac{1}{1!} + \cdots + \frac{1}{n!} + \frac{1}{n!n}$ 单调减，且 $\lim\limits_{n\to+\infty} c_n = \mathrm{e}$。

(2) 证明对任意正整数 n，

$$\frac{1}{(n+1)!} < \mathrm{e} - \left(1 + \frac{1}{1!} + \cdots + \frac{1}{n!}\right) < \frac{1}{n!n}.$$

(3) $\lim\limits_{n\to+\infty} n \sin(2\pi n!\mathrm{e}) = 2\pi$。

(4) 证明 e 是无理数。

10. 证明

(1) $\left(1+\frac{1}{n}\right)^n < \mathrm{e} < \left(1+\frac{1}{n}\right)^{n+1}$；

(2) $\frac{1}{n+1} < \ln(n+1) - \ln n < \frac{1}{n}$；

(3) 数列

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln n.$$

单调递减；

(4) 数列

$$b_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln(n+1).$$

单调递增；

(5) 极限 $\lim\limits_{n\to+\infty} a_n$ 和 $\lim\limits_{n\to+\infty} b_n$ 存在且相等，这个共同的极限值称为**欧拉常数**；

(6) $\lim\limits_{n\to+\infty}\left(1-\frac{1}{2}+\frac{1}{3}-\cdots+\frac{(-1)^{n-1}}{n}\right) = \ln 2$；

(7) $\lim\limits_{n\to+\infty}\left(\frac{1}{n}+\frac{1}{n+1}+\frac{1}{n+2}\cdots+\frac{1}{2n}\right) = \ln 2$。

11. 设 $a_1 > 0, b_1 > 0$，$a_{n+1} = \frac{a_n+b_n}{2}$， $b_{n+1} = \sqrt{a_n b_n}$。证明 $\{a_n\}_{n\geq1}$ 和 $\{b_n\}_{n\geq1}$ 收敛到相同极限。

12. **对数函数的另一种定义方式**。设 $x > 0$ 。对正整数 n ，记

$$a_n(x) = 2^n\left(x^{\frac{1}{2^n}} - 1\right).$$

证明 $\{a_n(x)\}_{n\geq1}$ 单调有界。定义

$$\ln x = \lim_{n\to+\infty} 2^n\left(x^{\frac{1}{2^n}} - 1\right),$$

证明对任意正数 x, y， $\ln(xy) = \ln x + \ln y$。

13. （几何级数） (1) 对任意 $x \in \mathbb{R}$ 及任意正整数 n，记 $S_n = 1 + x + x^2 + \cdots + x^n$。证明：数列 $\{S_n\}_{n\geq1}$ 收敛当且仅当 $|x| < 1$。并求 $\lim\limits_{n\to+\infty} S_n$ 的值。

(2) 设 $p > 1$ 是正整数。证明 $x \in \mathbb{R}$ 是有理数当且仅当在 p 进制下 x 是一个有限小数或一个无限循环小数。

14. 设 $a_n > 0$，$\lim\limits_{n\to+\infty} \frac{a_{n+1}}{a_n} = \alpha$。

(1) 设 $0 \leq A < \alpha < B$。证明存在 N 使得当 $n \geq N$ 时 $A < \frac{a_{n+1}}{a_n} < B$。从而 $\frac{a_N}{A^N}A^n < a_n < \frac{a_N}{B^N}B^n$。

(2) 设 $\alpha < 1$，证明 $\lim\limits_{n\to+\infty} a_n = 0$。

(3) 利用(1)证明 $\lim\limits_{n\to+\infty} \sqrt[n]{a_n} = \alpha$。

15. 设 $a > 1$ 及 $k > 0$。求 $\lim\limits_{n\to+\infty} \frac{a^n}{n!}$， $\lim\limits_{n\to+\infty} \frac{n^k}{a^n}$， $\lim\limits_{n\to+\infty} \frac{n!}{n^n}$。(提示：用第1(a)xiv题结论)

16. 设 $a_n > 0$，$\lim\limits_{n\to+\infty} a_n = a$。证明 $\lim\limits_{n\to+\infty} \sqrt[n]{a_1 a_2 \ldots a_n} = a$。(提示：用第1(a)xiv题结论)

17. 求 $\lim\limits_{n\to+\infty} \sqrt[n]{a}$，$\lim\limits_{n\to+\infty} \sqrt[n]{n}$，$\lim\limits_{n\to+\infty} \frac{\ln n}{n}$，$\lim\limits_{n\to+\infty} \frac{n}{\sqrt[n]{n!}}$，$\lim\limits_{n\to+\infty} \sqrt[n]{\frac{1\cdot3\cdot5\cdots(2n-1)}{2\cdot4\cdot6\cdots(2n)}}$。(提示：用第1(a)xiv题结论)

18. 设 $a > 1$ 及 $k > 0$。求 $\lim\limits_{x\to+\infty} x^{\frac{1}{x}}$，$\lim\limits_{x\to+\infty} \frac{\ln x}{x^k}$，$\lim\limits_{x\to+\infty} \frac{x^k}{a^x}$，$\lim\limits_{x\to0^+} x^k \ln x$。(提示：用第**??**题结论)

19. 设 $\lim\limits_{n\to+\infty} a_n = \alpha$。证明 $\lim\limits_{n\to+\infty} \frac{a_1+a_2+\cdots+a_n}{n} = \alpha$。

20. 求 $\lim\limits_{n\to+\infty} \frac{1\cdot3\cdot5\cdots(2n-1)}{2\cdot4\cdot6\cdots(2n)}$。(提示：用不等式 $\sqrt{ab} \le \frac{a+b}{2}$)

21. 写出一个数列 $\{a_n\}$ 使得对于任意 $A \in \mathbb{R}$ 或 $A = +\infty, -\infty, \infty$，都存在子列 $\{a_{n_k}\}$ 使得 $\lim\limits_{k\to+\infty} a_{n_k} = A$。

22. 证明：任何一个实数数列都含有一个单调子列。

23. 证明关于实数集 $\mathbb{R}$ 的以下两个陈述是等价的：
(1) $\mathbb{R}$ 是一个序域，且满足确界公理，即任何非空有上界的实数集合 $X \subset \mathbb{R}$ 都在 $\mathbb{R}$ 中有上确界。
(2) $\mathbb{R}$ 是一个序域，且满足：任何单调不减且有上界的实数数列 $\{x_n\}_{n\ge1}$ 都在 $\mathbb{R}$ 中有极限。

用WolframAlpha进行计算
在www.wolframalpha.com网页的对话框中输入

```
lim sqrt((x^2+1)/(x^2-1)) as x->infinity
```

检查计算结果。

用WolframAlpha进行计算
在www.wolframalpha.com网页的对话框中输入

```
sqrt(x^3/(x-1)) as x goes to -infinity
```

得到 $\sqrt{\frac{x^3}{x-1}}$ 在 $x \to -\infty$ 时的渐近展开式，从中可以得到 $y = \sqrt{\frac{x^3}{x-1}}$ 在 $x \to -\infty$ 处的渐近线。在对话框中输入

```
asymptotes | sqrt(x^3/(x - 1))
```

会得到 $\sqrt{\frac{x^3}{x-1}}$ 在 $x \to 1$ 和 $x \to +\infty$ 时的渐近线。

2.5 大 O 与小 o，函数的主项，阶的比较

在解决复杂问题时，我们往往采取突出主要矛盾忽略次要矛盾的办法，使得在不改变问题本质的情况下，问题得到适当的简化。

把一个函数分解成具有简单形式的主要项（简称主项）和次要项是微积分中常用的一个办法。例如对多项式 $P(x) = 3x - 2x^2$ 而言，当 $x \to 0$ 时，$-2x^2$ 的绝对值要远比 $3x$ 的绝对值更小，此时是 $P(x)$ 的两项中 $3x$ 对 $P(x)$ 的贡献是主要的，它是 $P(x)$ 主项；而当 $x \to +\infty$ 时，$-2x^2$ 和 $3x$ 都是无穷大量，但$-2x^2$ 的绝对值要远比 $3x$ 的绝对值更大，此时 $-2x^2$ 是 $P(x)$ 的主项，相比之下 $3x$ 对 $P(x)$ 的贡献可以被忽略。

在一个极限过程中，比较不同函数大小变化的相对快慢，涉及到阶的概念，而大 O 和小 o 提供了一个描述阶的一种方式。

定义 2.5.1. "$f(x) = O(g(x)),\ \ x \to a$"表示：存在常数 $M > 0$ 以及 a 的去心邻域 V 使得对任意 $x \in V$，$|f(x)| \le M\,|g(x)|$。

"$f(x) = o(g(x)),\ \ x \to a$"表示：对任意 $\varepsilon > 0$，都存在 a 的去心邻域 V_ε 使得对任意 $x \in V_\varepsilon$，$|f(x)| \le \varepsilon\,|g(x)|$。

称当 $x \to a$ 时，函数 f 是**有界量**，如果 $f(x) = O(1),\ \ x \to a$。即存在常数 $M > 0$ 以及 a 的去心邻域 V 使得对任意 $x \in V$，$|f(x)| \le M$。

因此，
$x \to a$ 时，f 是无穷小量，当且仅当 $f(x) = o(1),\ \ x \to a$；
$x \to a$ 时，f 是无穷大量，当且仅当 $\frac{1}{f(x)} = o(1),\ \ x \to a$；
$\lim\limits_{x\to a} f(x) = A \in \mathbb{R}$ 当且仅当 $f(x) - A = o(1),\ \ x \to a$。

例 2.5.2. 当 $x \to 0$ 时，$1 - \cos x = O(x^2)$，$\sin(x^2) = O(x^2)$，$x^2 = O(x^2)$。但这三个函数并不恒等。

同样，$x \to 0$ 时，$\mathrm{e}^x - 1 - x = o(x)$，$\sin x - x = o(x)$，$\ln(1+x) - x = o(x)$。这三个函数也不恒等。 □

注 2.5.3. 上述例子表明 $O(g(x))$ 或 $o(g(x))$ 都不是一个函数，而是一个由函数组成的集合，也就是一类函数，这类函数具有上述定义所描述的那样的共同性质。$f(x) = O(g(x))$ 的确切含义应该是 $f \in O(g)$。但后者这个表达虽然准确，但不适合计算，所以习惯上仍使用 $f(x) = O(g(x))$ 的表达形式。

O 和 o 满足一些运算性质，例如

- $f = O(f)$。
- 若 $f = o(g)$，则 $f = O(g)$。

- 若 $f = O(p)$，则 $fo(g) = o(pg)$，$fO(g) = O(pg)$。特别地，因为 $f = O(f)$，所以 $fo(g) = o(fg)$，$fO(g) = O(fg)$。若 f 是有界函数，则$fO(g) = O(g)$，$fo(g) = o(g)$。
- $o(h)$ 和 $O(h)$ 都是线性空间。即，若 $x \to c$ 时，$f = o(h), g = o(h)$，则对任意 $\lambda, \mu \in \mathbb{R}$，$\lambda f + \mu g = o(h)$。对 O 成立类似结论。

定义 2.5.4. (1) 设 $x \to a$ 时，f, g 都是无穷小量。

称 f 是**不比** g **更低阶**的无穷小量，如果 $f(x) = O(g(x))$， $x \to a$；

称 f 是**比** g **更高阶**的无穷小量，如果 $f(x) = o(g(x))$， $x \to a$。

(2) 设 $x \to a$ 时，f, g 都是无穷大量。

称 f 是**不比** g **更低阶**的无穷大量，如果 $\frac{1}{f}$ 是不比 $\frac{1}{g}$ 更高阶的无穷大量，即 $g(x) = O(f(x))$， $x \to a$；

称 f 是比 g **更高阶**的无穷大量，如果当 $x \to a$ 时，$\frac{1}{f}$ 是比 $\frac{1}{g}$ 更高阶的无穷小量，即 $g(x) = o(f(x))$， $x \to a$。

例 2.5.5. (1) $x \to 0$ 时，x^β 是比 x^α 更高阶的无穷小量当且仅当 $0 < \alpha < \beta$，x^ρ 是比 x^τ 更高阶的无穷大量当且仅当 $\rho < \tau < 0$。

(2) $x \to +\infty$ 时，x^β 是比 x^α 更高阶的无穷大量当且仅当 $0 < \alpha < \beta$，x^ρ 是比 x^τ 更高阶的无穷小量当且仅当 $\rho < \tau < 0$。

(3) 对 $\alpha \le 1$，当 $0 < |x| < 1$ 时，

$$|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x} \le |x|^\alpha |x|^{1-\alpha} \left(\sin^2 \frac{1}{x} + \cos^2 \frac{1}{x} \right) = |x|^\alpha |x|^{1-\alpha},$$

所以 $|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x} = O(|x|)$；对 $\alpha < 1$，$|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x} = o(|x|^\alpha)$。

对 $\alpha > 1$，取 $x_n = \frac{1}{2n\pi + \frac{\pi}{2}}$，则

$$\frac{|x_n| \sin^2 \frac{1}{x_n} + x_n^2 \cos^2 \frac{1}{x_n}}{|x_n|^\alpha} = \frac{1}{x_n^{\alpha-1}} \to +\infty, \quad n \to +\infty,$$

所以 $|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x} \ne O(|x|^\alpha)$。

对 $\beta \ge 2$，

$$|x|^\beta = |x|^\beta \sin^2 \frac{1}{x} + |x|^\beta \cos^2 \frac{1}{x} \le |x|^{\beta-2} \left(|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x} \right).$$

所以 $x^2 = O\left(|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x}\right)$；对 $\beta > 2$，$|x|^\beta = o\left(|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x}\right)$。

对 $\beta < 2$，取 $x_n = \frac{1}{n\pi}$，则

$$\frac{|x_n|^\beta}{|x_n| \sin^2 \frac{1}{x_n} + x_n^2 \cos^2 \frac{1}{x_n}} = \frac{1}{x_n^{2-\beta}} \to +\infty, \quad n \to +\infty,$$

所以 $|x|^\beta \ne O\left(|x| \sin^2 \frac{1}{x} + x^2 \cos^2 \frac{1}{x}\right)$。

“不比……更低阶”和“比……更高阶”满足传递性，但不是任何两个无穷小量都可以比较高阶，例如，$x\sin^2\frac{1}{x}+x^2\cos^2\frac{1}{x}$ 是无穷小量，但是它无法和 $|x|^{3/2}$ 比较哪个是更高阶的无穷小量。

定义 2.5.6. (1) 称 $x\to a$ 时 f,g **同阶**，如果当 $x\to a$ 时，$f(x)=O(g(x))$ 且 $g(x)=O(f(x))$，即存在常数 $M>1$ 以及 a 的去心邻域 V 使得对任意 $x\in V$，

$$\frac{1}{M}|g(x)|\le|f(x)|\le M|g(x)|.$$

(2) 称 $x\to a$ 时 f,g **等价**，如果 $f(x)=g(x)+o(g(x)),\quad x\to a$。

如果 $g(x)\neq 0$，则 $f(x)=g(x)+o(g(x))$，$x\to a$ 当且仅当 $\lim\limits_{x\to a}\frac{f(x)}{g(x)}=1$。于是我们此前得到的一些极限现在可以写成如下形式。

例 2.5.7. 当 $x\to 0$ 时，

$$\sin x=x+o(x),\qquad \cos x=1-\frac{x^2}{2}+o(x^2),\qquad \tan x=x+o(x),$$
$$\ln(1+x)=x+o(x),\qquad \mathrm{e}^x=1+x+o(x),\qquad (1+x)^r=1+rx+o(x).$$

这些等式的作用是把超越函数渐近展开为主项是多项式的形式。

证明. 我们只证明最后一个。

$$\begin{aligned}(1+x)^r&=\mathrm{e}^{r\ln(1+x)}\\&=\mathrm{e}^{rx+o(x)}\\&=1+rx+o(x)+o(rx+o(x))\\&=1+rx+o(x).\end{aligned}$$

上述过程中使用了如下事实

$$ro(x)=o(x);$$
$$\text{若 } f(x)=O(g(x))\text{，则 } o(f(x))=o(g(x));$$
$$o(x)+o(x)=o(x).$$

类似的关于 O,o 的运算性质我们留作习题。 □

例 2.5.8. 求 $\lim\limits_{x\to 0}\dfrac{\sqrt{1+2x^4}-\sqrt[3]{1-x^4}}{\sin^2 x\,(1-\cos x)}$.

解：

$$
\begin{aligned}
&\lim_{x\to 0}\frac{\sqrt{1+2x^4}-\sqrt[3]{1-x^4}}{\sin^2 x\,(1-\cos x)}\\
=&\lim_{x\to 0}\frac{1+\frac{1}{2}(2x^4)+o(2x^4)-\left(1+\frac{1}{3}(-x^4)+o(-x^4)\right)}{[x+o(x)]^2\left[1-\left(1-\frac{x^2}{2}+o(x^2)\right)\right]}\\
=&\lim_{x\to 0}\frac{\frac{4}{3}x^4+o(x^4)}{\frac{x^4}{2}\,[1+o(1)]^2\,[1+o(1)]}\\
=&\lim_{x\to 0}\frac{\frac{4x^4}{3}[1+o(1)]}{\frac{x^4}{2}\,[1+o(1)]}\\
=&\frac{8}{3}
\end{aligned}
$$

□

例 2.5.9. 求 $\lim\limits_{x\to 0}\dfrac{\tan x-\sin x}{x^2\ln(1+x)}$。

解：

$$\frac{\tan x-\sin x}{x^2\ln(1+x)}=\frac{x+o(x)-[x+o(x)]}{x^2(x+o(x))}=\frac{o(x)}{x^3+o(x^3)},$$

这样展开是无法确定最后的极限的，因为分子 $o(x)$ 无法与分母中的主项 x^3 比较。

$$
\begin{aligned}
\frac{\tan x-\sin x}{x^2\ln(1+x)}&=\frac{\tan x(1-\cos x)}{x^2(x+o(x))}=\frac{[x+o(x)]\left[\frac{x^2}{2}+o(x^2)\right]}{x^3+o(x^3)}=\frac{\frac{x^3}{2}+o(x^3)}{x^3+o(x^3)}\\
&=\frac{\frac{1}{2}+\frac{o(x^3)}{x^3}}{1+\frac{o(x^3)}{x^3}}=\frac{\frac{1}{2}+o(1)}{1+o(1)},
\end{aligned}
$$

所以 $\lim\limits_{x\to 0}\dfrac{\tan x-\sin x}{x^2\ln(1+x)}=\dfrac{1}{2}$。 □

注 2.5.10. 有些教材用 “$f(x)\sim g(x), x\to a$” 来表示 f 与 g 等价，并提出了无穷小等价替换的做法。例如

$$\sin x\sim x,\quad \tan x\sim x,\quad \ln(1+x)\sim x,\qquad x\to 0,$$

当 $x\to 0$ 时用 x 替换 $\sin x$、$\tan x$、$\ln x$ 等。初学者常常会据此得到

$$\lim_{x\to 0}\frac{\tan x-\sin x}{x^2\ln(1+x)}=\lim_{x\to 0}\frac{x-x}{x^2x}=0.$$

这明显是错误的，但初学者却不明就里。所以我们不建议读者使用 $f(x) \sim g(x)$，$x \to a$ 这种看似简单的表达形式。

以下定理告诉我们上面定义的“等价”本质上是一个对称的关系。它同时也给我们提供了一个重要的工具，我们会多次看到它的应用，这里我们介绍一下 Newton 是如何利用这个想法发现广义二项式展开的。

定理 2.5.11. *若* $f(x)+o(f(x))=g(x)+o(g(x))$, $x\to a$，*则* $g(x)=f(x)+o(f(x))$, $x\to a$。

证明. $f(x)+\alpha(x)=g(x)+\beta(x)$，其中 $\alpha=o(f)$, $\beta=o(g)$, $x\to a$。于是对任意 $0<\varepsilon<1$，存在 a 的去心邻域 V_ε 使得对任意 $x\in V_\varepsilon$，

$$|\alpha(x)|\le\frac{\varepsilon}{3}\left|f(x)\right|,\quad |\beta(x)|\le\frac{\varepsilon}{3}\left|g(x)\right|.$$

从而

$$|g(x)|=|f(x)+\alpha(x)-\beta(x)|\le|f(x)|+|\alpha(x)|+|\beta(x)|\le\frac{4}{3}\left|f(x)\right|+\frac{1}{3}\left|g(x)\right|,$$

所以 $|g(x)|\le 2\left|f(x)\right|$。因此

$$|\beta(x)|\le\frac{\varepsilon}{3}\left|g(x)\right|\le\frac{2\varepsilon}{3}\left|f(x)\right|.$$

从而

$$|g(x)-f(x)|=|\alpha(x)-\beta(x)|\le|\alpha(x)|+|\beta(x)|\le\varepsilon\left|f(x)\right|.$$

所以 $g(x)=f(x)+o(f(x))$, $x\to a$。若 $f(x)=g(x)+o(g(x))$，则取 $\alpha(x)=0$，则由上述结论知 $g(x)=f(x)+o(f(x))$, $x\to a$。 □

例 2.5.12. 对正有理数 r， $(1+x)^r=1+rx+\dfrac{r(r-1)}{2!}x^2+o(x^2)$, $x\to 0$。

证明. 设 $r=\frac{m}{n}$，$(1+x)^{\frac{m}{n}}=1+\alpha(x)$。则 $\lim\limits_{x\to 0}\alpha(x)=0$。

于是 $(1+x)^m=[1+\alpha(x)]^n$，从而

$$1+mx+o(x)=1+n\alpha(x)+o(\alpha(x)),\quad x\to 0.$$

根据定理2.5.11，$\alpha(x)=\frac{m}{n}x+o(x)=rx+o(x)$，$x\to 0$。再令 $(1+x)^{\frac{m}{n}}=1+rx+\beta(x)$。则 $\lim\limits_{x\to 0}\beta(x)=0$，于是 $(1+x)^m=[1+rx+\beta(x)]^n$，从而

$$\begin{aligned}&1+mx+\frac{m(m-1)}{2}x^2+o(x^2)\\=&\left(1+rx\right)^n+n\left(1+rx\right)^{n-1}\beta(x)+o(\beta(x))\\=&1+nrx+\frac{r^2n(n-1)}{2}x^2+o(x^2)+n\beta(x)+o(\beta(x))\end{aligned}$$

所以根据定理2.6.3，

$$\beta(x)=\frac{1}{2}\left[r(m-1)-r^2(n-1)\right]x^2+o(x^2)=\frac{r(r-1)}{2}x^2+o(x^2).$$

所以 $(1+x)^r=1+rx+\frac{r(r-1)}{2!}x^2+o(x^2)$, $x\to 0$ □

历史上 Newton 曾用类似的办法得到：对任意有理数 r，

$$(1+x)^r = 1 + rx + \frac{r(r-1)}{2!}x^2 + \cdots + \frac{r(r-1)\cdots(r-k+1)}{k!}x^k + \cdots, \quad x \to 0.$$

当 r 是正整数时，这是人们早已熟知的二项式展开，此时上式右端只有有限多项，到 $k = r$ 时结束。Newton 发现这个展开对有理数也成立，它被称为**广义 Newton 二项式**[1]。 Newton 从这样一个基本发现出发得到了一系列重要的结果。广义二项式这个结论我们将在后面将用 Taylor 展开的办法得到，那时指数 r 可以是任意实数。

习题2.5

1. 证明 $o(f) \pm o(f) = o(f)$，$O(f) \pm O(f) = O(f)$，$o(f) \cdot O(g) = o(fg)$，$O(f) \cdot O(g) = O(fg)$；如果 $g \neq 0$，则 $f = O(g)$ 当且仅当 $\frac{f}{g} = O(1)$，$f = o(g)$ 当且仅当 $\frac{f}{g} = o(1)$。

2. 设 $0 < a < b$，$f(x) = \left(\frac{a^x + b^x}{2}\right)^{\frac{1}{x}}$。求 $\lim\limits_{x\to 0} f(x)$，$\lim\limits_{x\to -\infty} f(x)$，$\lim\limits_{x\to +\infty} f(x)$。

3. 问

(a)

$$\lim_{x\to 0} \frac{\sin\left(x^2 \sin\frac{1}{x}\right)}{x^2 \sin\frac{1}{x}} = 1$$

成立吗？为什么？

(b)

$$\lim_{x\to 0} \frac{\sin\left(x^2 \sin\frac{1}{x}\right)}{x} = \lim_{x\to 0} \frac{\sin\left(x^2 \sin\frac{1}{x}\right)}{x^2 \sin\frac{1}{x}} \cdot \frac{x^2 \sin\frac{1}{x}}{x}$$

成立吗？为什么？

(c)

$$\lim_{x\to 0} \frac{\sin\left(x^2 \sin\frac{1}{x}\right)}{x} = \lim_{x\to 0} \frac{x^2 \sin\frac{1}{x}}{x}$$

成立吗？为什么？

4. (1) 设 f 在 $x = 0$ 处连续，$f(0) = 0$，并且 $\lim\limits_{x\to 0} \dfrac{f(2x) - f(x)}{x} = \lambda$。证明

$$f(x) = \lambda x + o(x), \qquad x \to 0.$$

(2) 利用上述结果证明

$$\mathrm{e}^x = 1 + x + \frac{x^2}{2} + o(x^2), \qquad x \to 0.$$

---

[1]在1676年9月13日通过皇家学会秘书 Henry Oldenburg 写给 Leibniz 的信里 Newton 解释了这个结果。https://cudl.lib.cam.ac.uk/view/MS-ADD-03977/63

你能用这个办法得到 e^x 在 $x \to 0$ 时的更高阶的展开形式吗？(提示：考虑 $e^x = 1 + x + xf(x)$ 并利用 $e^{2x} = e^x e^x$)

(3) 你能得到 $\sin x$ 在 $x \to 0$ 时的更高阶展开吗？

5. **Aitken 加速法**是由新西兰数学家 Alexander Aitken（1895 年 4 月 1 日–1967 年 11 月 3 日）于 1926 年提出的一个加速数列收敛的方法。如果数列 $\{x_n\}$ 收敛于 x^*，那么在坐标平面中过点 (x_n, x_{n+1}) 和点 (x_{n+1}, x_{n+2}) 做一条直线，认为这条直线与对角线 $y = x$ 的交点为 (x^*, x^*) 的近似值，这样得到 x^* 的近似值

$$y_n = \frac{x_{n+2}x_n - x_{n+1}^2}{x_{n+2} + x_n - 2x_{n+1}} = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} + x_n - 2x_{n+1}}.$$

(1) 如果 $\lim\limits_{n\to+\infty} \frac{x_{n+1}-x^*}{x_n-x^*} = \mu \in (0,1)$，证明 $\lim\limits_{n\to+\infty} \frac{|y_n-x^*|}{|x_n-x^*|} = 0$。

(2) 如果 $x_n = x^* + \frac{\alpha}{n^p} + o\left(\frac{1}{n^p}\right)$ ($n \to +\infty$)，求 $|y_n - x^*|$ 的阶。

习题讨论课1

1. 求 $\lim\limits_{x\to 1}\dfrac{x^3-3x+2}{x^4-4x+3}$。

2. 求 $\lim\limits_{x\to 0}\dfrac{(1+x)^\mu-1}{x}$。

3. 求 $\lim\limits_{x\to -8}\dfrac{\sqrt{1-x}-3}{2+\sqrt[3]{x}}$。

4. 求 $\lim\limits_{x\to a}\dfrac{x^\mu-a^\mu}{x-a}$。

5. 求 $\lim\limits_{x\to \infty}\dfrac{(x-1)(x-2)(x-3)(x-5)}{(5x-1)^5}$。

6. 求 $\lim\limits_{x\to\infty}\left(\sqrt[3]{x^3+x^2+1}-\sqrt[3]{x^3-x^2+1}\right)$。

7. 求 $\lim\limits_{x\to 1}\arctan\dfrac{1}{1-x}$。

8. 设 $\lim\limits_{n\to+\infty}a_n=A$，其中 $A\in\mathbb{R}$ 或 $A=-\infty$ 或 $A=+\infty$ 或 $A=\infty$。请判断
$$\lim_{n\to+\infty}\frac{a_1+a_2+\cdots+a_n}{n}=A$$
是否成立。若成立，请用极限的定义给出证明；若不成立，请给出反例。

9. 设 $a_n>0$ 满足 $\lim\limits_{n\to+\infty}a_n=A$，其中 $A\geq 0$ 或 $A=+\infty$。请判断
$$\lim_{n\to+\infty}\sqrt[n]{a_1a_2\cdots a_n}=A$$
是否成立。若成立，请给出证明；若不成立，请给出反例。

10. 设 $a>1,\alpha>0,\beta>0$。证明
$$\lim_{n\to+\infty}\sqrt[n]{n}=1,\quad \lim_{n\to+\infty}\frac{\ln n}{n}=0,\quad \lim_{x\to+\infty}\frac{\ln x}{x^\alpha}=0,\quad \lim_{x\to+\infty}\frac{x^\beta}{a^x}=0.$$

11. 求 $\lim\limits_{n\to+\infty}\dfrac{a^n}{n!}$，$\lim\limits_{n\to+\infty}n!\left(\dfrac{a}{n}\right)^n$。

12. 设 $x_n>0(\forall n\geq 1)$。证明数列 $\left\{\sum\limits_{k=1}^n x_k\right\}$ 收敛当且仅当数列 $\left\{\prod\limits_{k=1}^n(1+x_k)\right\}_{n\geq 1}$ 收敛。

13. 设 $f:\mathbb{R}\to\mathbb{R}$ 满足
$$f\left(\frac{x+y}{2}\right)\leq\frac{f(x)+f(y)}{2},\quad \forall x,y\in\mathbb{R}.$$

证明

(1) 对任何有理数 $r \in [0,1]$，

$$f((1-r)x+ry) \le (1-r)f(x)+rf(y), \quad \forall x,y \in \mathbb{R}.$$

(2) 如果 f 连续，则对任意 $t \in [0,1]$，

$$f((1-t)x+ty) \le (1-t)f(x)+tf(y), \quad \forall x,y \in \mathbb{R}.$$

(3) 如果 f 连续，则对任意 $t_1,t_2,\ldots,t_n \in [0,1]$，只要 $t_1+t_2+\cdots+t_n=1$，就有对任意 $x_1,x_2,\ldots,x_n \in \mathbb{R}$，

$$f(t_1x_1+t_2x_2+\cdots+t_nx_n) \le t_1f(x_1)+t_2f(x_2)+\cdots+t_nf(x_n).$$

(4) 如果 f 在任何有界闭区间中有界，则 f 连续。

(5) 如果 f 没有第二类间断点，则 f 连续。

(6) 如果 f 单调，则 f 连续。

14. 设 $\{x_n\}_{n\ge1}$ 满足对任意正整数 n,m，$0 \le x_{n+m} \le x_n+x_m$。证明 $\lim\limits_{n\to+\infty}\dfrac{x_n}{n} = \inf\limits_{n\ge1}\dfrac{x_n}{n}$。

第 3 章　深入了解连续性

在前面两章中，我们重新认识了实数，特别是实数的确界性质，了解了连续函数和极限的概念和基本性质。在这一章里，我们将要对实数和函数的连续性进一步研究，为后续内容的学习打好基础。

3.1　实数的连续性

为了寻找一个问题的答案，我们往往先大致预估一下答案所在的范围，然后再逐步缩小这个范围。以下“有界闭区间套定理”正是这样一个想法，利用它我们能够通过一系列筛选最终找到问题的答案。

定义 3.1.1 (闭集). 称 $I \subseteq \mathbb{R}$ 是**闭集**，如果 I 满足：

$$\text{数列 } \{x_n\}_{n\geq 1} \text{ 满足 } x_n \in I\ (\forall n \geq 1) \text{ 且 } \lim_{n\to+\infty} x_n = x_0 \Rightarrow x_0 \in I\text{。}$$

如果一个区间是闭集，那么称它是一个**闭区间**。闭区间有如下几种形式：

$$[a,b], \quad [a,+\infty), \quad (-\infty,b], \quad (-\infty,+\infty), \quad \emptyset,$$

其中只有第一种是非空的有界闭区间。

定理 3.1.2 (有界闭区间套定理). 设一列非空有界闭区间 $[a_n,b_n]$ 构成一个区间套，即对任意正整数 n，$[a_{n+1},b_{n+1}] \subseteq [a_n,b_n]$。则 $\bigcap\limits_{n\geq 1}[a_n,b_n] \neq \emptyset$。若进一步 $\lim\limits_{n\to+\infty}(b_n-a_n)=0$，则存在唯一的实数 A 使得 $\bigcap\limits_{n\geq 1}[a_n,b_n]=\{A\}$。

证明. 由 $[a_{n+1},b_{n+1}] \subseteq [a_n,b_n]$ 知

$$a_1 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots \leq b_{n+1} \leq b_n \leq \cdots \leq b_1,$$

所以数列 $\{a_n\}_{n\geq 1}$ 单调不减有上界，数列 $\{b_n\}_{n\geq 1}$ 单调不增有下界。因此确界 $A = \sup\limits_{n\geq 1} a_n$，　$B = \inf\limits_{n\geq 1} b_n$ 存在，而且 $A \leq B$。

$c \in \bigcap\limits_{n\geq 1}[a_n,b_n]$ 当且仅当 c 是 $\{a_n\}_{n\geq 1}$ 上界且是 $\{b_n\}_{n\geq 1}$ 下界，当且仅当 $A \leq c \leq B$。因此 $\bigcap\limits_{n\geq 1}[a_n,b_n] = [A,B]$。若 $\lim\limits_{n\to+\infty}(b_n-a_n) = 0$，则由 $0 \leq B-A \leq b_n-a_n$ 知 $B-A=0$，所以 $A=B$，$\bigcap\limits_{n\geq 1}[a_n,b_n] = \{A\}$。 □

在上述定理中，“有界”、“闭”、“套”都是不可或缺的重要条件。

以下的“列紧性”定理可以帮助我们从一堆看似杂乱无章的近似解中找到逼近最终解的线索。同时，从它的证明过程中我们可以体会有界闭区间套这个重要工具是如何帮助我们解决具体问题的。

为了找到数列 $\{x_n\}$ 的一个收敛子列 $\{x_{n_k}\}$ 以及相应的极限 A，我们只需找到这样的实数 A：在 A 的任意近旁，都聚集着数列 $\{x_n\}$ 的无穷多项。

定理 3.1.3 (列紧性). *任何有界的实数列必含有收敛的子列。*

证明. 设 $\{x_n\}_{n\geq 1}$ 有界。于是存在实数 a_1, b_1 使得对任意 $n \geq 1$，$x_n \in [a_1, b_1]$。取 $n_1 = 1$。

如果区间 $\left[a_1, \frac{a_1+b_1}{2}\right]$ 中含有数列 $\{x_n\}_{n\geq 1}$ 的无穷多项，则取 $a_2 = a_1$， $b_2 = \frac{a_1+b_1}{2}$；否则取 $a_2 = \frac{a_1+b_1}{2}$， $b_2 = b_1$。因此 $[a_2, b_2]$ 中含有数列 $\{x_n\}_{n\geq 1}$ 的无穷多项，取 $n_2 > n_1$ 使 $x_{n_2} \in [a_2, b_2]$。

按照这个办法可以得到有界闭区间套$[a_k, b_k]$，以及 $\{x_n\}_{n\geq 1}$ 的子列 $\{x_{n_k}\}_{k\geq 1}$ 使得对任意 $k \geq 1$，$x_{n_k} \in [a_k, b_k]$。因为 $\lim\limits_{k\to+\infty}(b_k - a_k) = \lim\limits_{k\to+\infty}\frac{b_1-a_1}{2^{k-1}} = 0$，所以由闭区间套定理（定理3.1.2），存在 $A \in \mathbb{R}$ 使得 $\bigcap\limits_{k\geq 1}[a_k, b_k] = \{A\}$。因为 $A, x_{n_k} \in [a_k, b_k]$，所以

$$|x_{n_k} - A| \leq |b_k - a_k| = \frac{b_1 - a_1}{2^{k-1}},$$

从而 $\lim\limits_{k\to+\infty} x_{n_k} = A$。 □

根据定义，要验证一个数列 $\{a_n\}_{n\geq 1}$ 收敛，就是要找到这样的实数 A，使得 $|a_n - A|$ 随着 n 变大而越来越接近于零。我们把 a_n 当成 A 的近似值，如果无法得到 A 的精确值，那么我们就无法知道误差 $|a_n - A|$ 的大小。在这种情况下，就无法利用极限的定义来说明数列 $\{a_n\}_{n\geq 1}$ 收敛到 A。以下的“Cauchy准则”告诉我们可以通过测量值之间的误差 $|a_m - a_n|$ 来验证数列 $\{a_n\}_{n\geq 1}$ 是否收敛。

定理 3.1.4 (数列收敛的 Cauchy 准则). *数列 $\{a_n\}_{n\geq 1}$ 收敛当且仅当它是一个* **Cauchy 数列**，*即对任意 $\varepsilon > 0$，存在 $N_\varepsilon \geq 1$ 使得对任意 $m, n \geq N_\varepsilon$，$|a_m - a_n| < \varepsilon$。*

证明. 必要性(即收敛数列是 Cauchy 数列)的证明留给读者完成。下证充分性。

设 $\{a_n\}_{n\geq 1}$ 是 Cauchy 数列。则存在 n_1 使得对任意 $m, n \geq n_1$，$|a_m - a_n| < 1$。从而对任意正整数 n，

$$|a_n| \leq |a_{n_1}| + |a_n - a_{n_1}| \leq |a_{n_1}| + 1 + \max_{1\leq k\leq n_1} |a_k - a_{n_1}|.$$

所以 $\{a_n\}_{n\geq 1}$ 是有界数列，从而根据定理3.1.3，$\{a_n\}_{n\geq 1}$ 有收敛子列。设 $A = \lim\limits_{k\to+\infty} a_{n_k}$。存在 K 使得 $n_K \geq N_\varepsilon$ 且 $|a_{n_K} - A| < \varepsilon$。于是对任意 $n \geq N_\varepsilon$，

$$|a_n - A| \leq |a_n - a_{n_K}| + |a_{n_K} - A| < 2\varepsilon.$$

因此 $\lim\limits_{n\to+\infty} a_n = A$。 □

我们再给一个证明，可以进一步体会如何使用有界闭区间套定理这个工具。

充分性的另一个证明. 我们用有界闭区间套定理给充分性一个另外的证明。跟上面证明一样，我们知道 Cauchy 数列是有界的，设 x_1, y_1 使得数列 $\{a_n\}$ 中的所有项都在有界闭区间 $[x_1, y_1]$ 中。令

$$u_1 = \frac{2x_1 + y_1}{3}, \quad v_1 = \frac{x_1 + 2y_1}{3}.$$

则由 Cauchy 数列的定义知，$[x_1, u_1]$ 和 $[v_1, y_1]$ 这两个闭区间中至少有一个不能包含数列 $\{a_n\}$ 中的无穷多项。

如果 $[x_1, u_1]$ 不包含数列 $\{a_n\}$ 中的无穷多项，则取 $x_2 = u_1, y_2 = y_1$；否则取 $x_2 = x_1, y_2 = v_1$。总之，有界闭区间 $[x_2, y_2]$ 之外最多只有数列 $\{a_n\}$ 中的有限多项，即存在正整数 N_1 使得对任意 $n \geq N_1$，$a_n \in [x_2, y_2]$。

如此得到正整数列 $N_1 \leq N_2 \leq N_3 \leq \cdots$ 以及有界闭区间套 $[x_k, y_k]$，使得对任意 $n \geq N_k$，$a_n \in [x_k, y_k]$；$|x_k - y_k| = \left(\frac{2}{3}\right)^{k-1} |x_1 - y_1|$。

由有界闭区间套定理知存在实数 A 使得 $\{A\} = \bigcap\limits_{k\geq 1} [x_k, y_k]$。从而对任意 $k \geq 1$，对任意 $n \geq N_k$，

$$|a_n - A| \leq |x_k - y_k| = \left(\frac{2}{3}\right)^{k-1} |x_1 - y_1|,$$

所以 $\lim\limits_{n\to+\infty} a_n = A$。 □

习题3.1

1. 设 I 是一个有界闭区间，则 $I = [\alpha, \beta] = \{x \in \mathbb{R} | \alpha \leq x \leq \beta\}$，其中 $\alpha = \inf I$，$\beta = \sup I$。

2. 设 $f : \mathbb{R} \to \mathbb{R}$ 是连续函数。证明对任何实数 b，$\{x \in \mathbb{R} | f(x) \leq b\}$ 是闭集；一般地，对任何闭集 $F \subseteq \mathbb{R}$，$f^{-1}(F) = \{x \in \mathbb{R} | f(x) \in F\}$ 是闭集。

3. 设 $\{F_n\}$ 是一列有界的非空闭集，满足 $F_{n+1} \subset F_n$。

(a) 证明 $\bigcap\limits_{n\geq 1} F_n \neq \emptyset$。

(b) 如果进一步，$\lim\limits_{n\to+\infty} \sup\limits_{x,y\in F_n} |x-y| = 0$，则 $\bigcap\limits_{n\geq 1} F_n \neq \emptyset$ 是单点集。

4. 证明不含任何收敛子列的数列一定是无穷大数列。

5. 设 F 是一个序域。证明以下结论都等价于 $F = \mathbb{R}$。

(a) F 满足确界公理。

(b) F 满足任何单调不减且有上界的序列都收敛。

(c) F 满足阿基米德性质以及闭区间套性质。

(d) F 中任何有界数列都有收敛子列。

(e) F 满足阿基米德性质以及 Cauchy 准则。

3.2 应用：迭代与不动点

把计算器设为弧度制，然后随便输入一个数字，不断地按 cos 按钮，你会看到什么现象？换一个数字，重复上述过程，你会有什么发现？

很多时候，一个数列是靠递推关系 $x_{n+1} = f(x_n)$ 产生的，或者说这个数列由初值 x_0 和 f 的不断迭代产生。其中一个简单情形是 x_0 是 f 的一个不动点，$f(x_0) = x_0$，这时产生的数列是常数数列。一个稍微复杂一点的情形是 x_0 经 f 的迭代收敛到一个不动点，这个不动点是数列 x_n 的极限。很多数学问题可以转化为解方程，而后者又可以改写成不动点形式。同时，很多科学中的复杂问题要用算法才能解决，而迭代算法是其中很重要的一类方法。所以不动点问题无论在理论研究中还是在实际应用中都有重要意义。在这一节中，我们要研究在什么条件下一个函数产生的迭代数列可以收敛到一个不动点。

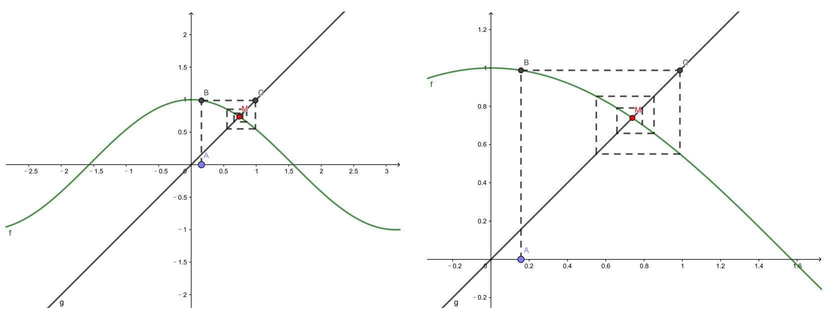


图 3.1:

定理 3.2.1 (压缩不动点定理). 设 I 是闭集，$f: I \to \mathbb{R}$ 满足 $f(I) \subseteq I$，且存在常数 $0 < \lambda < 1$ 使得对任意 $x, y \in I$，$|f(x) - f(y)| \leq \lambda |x-y|$（此时称 f 为集

合 I 上的一个**压缩映射**)。则存在唯一的 $x^* \in I$ 使得 $f(x^*) = x^*$，并且对任意 $x \in I$，$\lim\limits_{n\to+\infty} f^n(x) = x^*$。其中 $f^n = \underbrace{f \circ \cdots \circ f}_{n\text{个}f}$ 是 f 的 n 次迭代。

证明. 任取 $x \in I$，记 $x_n = f^n(x)(\,n = 0, 1, 2, \dots\,)$。则

$$|x_{n+2} - x_{n+1}| = |f(x_{n+1}) - f(x_n)| \le \lambda\,|x_{n+1} - x_n| \le \lambda^{n+1}\,|x_1 - x_0|\,,$$

从而

$$|x_{n+p} - x_n| \le \sum_{k=1}^{p} |x_{n+k} - x_{n+k-1}| \le \sum_{k=1}^{p} \lambda^{n+k-1}\,|x_1 - x_0| \le \frac{\lambda^n}{1-\lambda}\,|x_1 - x_0|\,,$$

所以 $\{x_n\}_{n\ge1}$ 是 Cauchy 数列。根据定理3.1.4，极限 $\lim\limits_{n\to+\infty} x_n = x^*$ 存在。在上述不等式中令 $p \to +\infty$，得到

$$|x^* - x_n| \le \frac{\lambda^n}{1-\lambda}\,|x_1 - x_0|\,.$$

于是

$$\begin{aligned}|f(x^*) - x^*| &\le |f(x^*) - f(x_n)| + |f(x_n) - x^*| \\ &\le \lambda\,|x^* - x_n| + |x_{n+1} - x^*| \le \frac{2\lambda^{n+1}}{1-\lambda}\,|x_1 - x_0|\,,\end{aligned}$$

令 $n \to +\infty$，得到 $|f(x^*) - x^*| \le 0$，从而 $f(x^*) = x^*$。对任意 $y \in I$，

$$|f^n(y) - x^*| = |f^n(y) - f^n(x^*)| \le \lambda^n\,|y - x^*|\,,$$

所以 $\lim\limits_{n\to+\infty} f^n(y) = x^*$。若 $\tilde{x} \in I$ 满足 $f(\tilde{x}) = \tilde{x}$，则 $\tilde{x} = \lim\limits_{n\to+\infty} f^n(\tilde{x}) = x^*$。
□

例 3.2.2. 考虑 $f : \mathbb{R} \to \mathbb{R}, f(x) = \cos x$，$x_n = f^n(x_0)$。

对任意 $x_0 \in \mathbb{R}$，$x_1 = f(x_0) \in [-1, 1] \subset \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$，所以 $x_2 \in [0, 1]$。因此对任意 $n \ge 2$，$x_n \in [0, 1]$。

对任意 $x, y \in [0, 1]$，

$$|f(x) - f(y)| = |\cos x - \cos y| = \left|2\sin\frac{x+y}{2}\sin\frac{y-x}{2}\right| \le \sin 1 \cdot |x - y|.$$

所以 f 是映闭区间 $[0, 1]$ 到其自身的压缩映射，因此 f 有唯一的不动点 x^*，对任意 $x_0 \in \mathbb{R}$，$f^n(x_0)$ 收敛到 x^*。 □

一般而言，压缩条件通常只是一个在不动点附近的局部条件。在实际应用中，会出现非压缩情形、大范围情形以及高维情形，函数迭代会展现出更丰富也更复杂的行为，其中的奥秘值得去探索。

习题3.2

1. 把 Kepler 方程 $x-\varepsilon\sin x=y$ 改写为不动点形式 $x=\varepsilon\sin x+y$。证明
(1) 若 $0<\varepsilon<1$，则对任意 $y\in\mathbb{R}$， Kepler 方程有唯一解 $x(y)$，讨论 $x(y)$ 关于 y 的连续性。
(2) 若 $y\in\mathbb{R}$，则对任意 $\varepsilon\in(0,1)$， Kepler 方程有唯一解 $x(\varepsilon)$，讨论 $x(\varepsilon)$ 关于 ε 的连续性。

2. 为了求解 $x^2+px+q=0(\ q\neq 0\ )$，把该方程改写为 $x=-\frac{q}{x}-p$。并利用迭代 $x_{n+1}=-\frac{q}{x_n}-p$ 计算上述方程的近似解。这个算法的优势在于它只用简单的算术运算，而不必求平方根。
(1) 问对怎样的初值 x_1，这个迭代会收敛到原二次方程的零点？
(2) 我们知道方程 $x^2-x-1=0$ 有两个不同的实数解，请利用数值计算，看看能否通过选择不同的初值得到所有的解，观察计算结果，发现其中的问题并给予解释。

3. 设 $c>0$，$a_{n+1}=\sqrt{c+a_n}$。求 a_1 的范围使得数列 $\{a_n\}_{n\geq 1}$ 有意义且收敛。提示：考虑函数 $f(x)=\sqrt{x+c}$。

4. 证明 $x_{n+1}=\frac{ax_n+b}{\alpha x_n+\beta}$ 可以通过适当的形如 $y_n=cx_n+d$ 的变换变为以下形式之一：
$$(1)\ y_{n+1}=y_n+A;\quad (2)\ y_{n+1}=Ay_n;\quad (3)\ y_{n+1}=\frac{A}{y_n};\quad (4)\ y_{n+1}=1+\frac{A}{y_n}.$$
然后分别讨论它们的极限行为。

3.3 连续函数的介值性质，反函数的连续性

定理 3.3.1 (Bolzano 1817). 设 $I\subseteq\mathbb{R}$ 是区间，$f:I\to\mathbb{R}$ 连续。则 $f(I)=\{f(x)|x\in I\}$ 也是区间。等价说法是，若 $x_1,x_2\in I$，$f(x_1)<f(x_2)$，则对任意 $y\in(f(x_1),f(x_2))$，存在介于 x_1,x_2 之间的 x（从而 $x\in I$）使得 $f(x)=y$。

证明. (**二分法**) 不妨设 $x_1<x_2$。假设对任意 $x\in(x_1,x_2)$，$f(x)\neq y$。

取 $a_1=x_1$，$b_1=x_2$。若 $f\left(\frac{a_1+b_1}{2}\right)>y$ ，则取 $a_2=a_1$，$b_2=\frac{a_1+b_1}{2}$；若 $f\left(\frac{a_1+b_1}{2}\right)<y$ ，则取 $a_2=\frac{a_1+b_1}{2}$，$b_2=b_1$。

如此不断进行下去得到闭区间套$[a_{n+1},b_{n+1}]\subseteq[a_n,b_n]$，对任意 n，$f(a_n)<y<f(b_n)$。由闭区间套定理（定理3.1.2），存在实数 ξ 使得 $\bigcap\limits_{n\geq 1}[a_n,b_n]=\{\xi\}$。则 $\xi\in I$，$\lim\limits_{n\to+\infty}a_n=\lim\limits_{n\to+\infty}b_n=\xi$。

因为 f 在 ξ 连续，所以 $\lim\limits_{n\to+\infty}f(a_n)=\lim\limits_{n\to+\infty}f(b_n)=f(\xi)$。由数列的保序性，$f(\xi)\leq y\leq f(\xi)$，所以 $f(\xi)=y$。但这与假设对任意 $x\in(x_1,x_2)$，$f(x)\neq y$ 矛盾。所以存在 $x\in(x_1,x_2)$ 使得 $f(x)=y$。 □

推论 3.3.2. 设 $I\subseteq\mathbb{R}$ 是区间，$f:I\to\mathbb{R}$ 是连续单射。则 f 是严格单调函数。

证明. 断言：若 $x_1, x_2 \in I$ 满足 $x_1 < x_2$ 且 $f(x_1) < f(x_2)$，则 f 在区间 $[x_1, x_2]$ 上是严格增函数。

断言的证明：任取 $x, y \in (x_1, x_2)$ 满足 $x < y$。

若 $f(x) < f(x_1)$，则 $f(x) < f(x_1) < f(x_2)$，于是由介值性质知存在 $\xi \in (x, x_2)$ 使得 $f(\xi) = f(x_1)$，这与 f 是单射矛盾。因此 $f(x) > f(x_1)$。

同理可证 $f(x) < f(x_2)$。所以 $f(x_1) < f(x) < f(x_2)$。

对 x, y, x_2 重复上面对 x_1, x, x_2 的讨论知，$f(x) < f(y) < f(x_2)$。

所以 $f(x_1) < f(x) < f(y) < f(x_2)$。因此 f 在区间 $[x_1, x_2]$ 上是严格增函数。

由断言知 f 在区间 I 的任何有界闭子区间上都是严格单调的。所以 f 在 I 上是严格单调的。 □

推论 3.3.3. 设 $I \subseteq \mathbb{R}$ 是区间，$f : I \to \mathbb{R}$ 是连续单射。则 $f^{-1} : f(I) \to I$ 是连续函数。

证明. 由前一个推论知 $f : I \to \mathbb{R}$ 是严格单调函数。由介值性质知 $f(I)$ 是区间。于是 $f^{-1} : f(I) \to I$ 是严格单调的，并且它把区间 $f(I)$ 映为区间 I，所以 f^{-1} 连续。 □

例 3.3.4. $\sin : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to (-1, 1)$，$\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$，$\cos : (0, \pi) \to (-1, 1)$，$\cot : (0, \pi) \to \mathbb{R}$ 都是连续的严格单调函数，它们的反函数 $\arcsin : (-1, 1) \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$，$\arctan : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$，$\arccos : (-1, 1) \to (0, \pi)$，$\operatorname{arccot} : \mathbb{R} \to (0, \pi)$ 也都是严格单调的连续函数。 □

定义 3.3.5. 多项式、幂函数、指数函数、对数函数、三角函数和反三角函数称为**基本初等函数**。有限多个基本初等函数通过有限多次四则运算和复合得到的函数称为**初等函数**。

初等函数都是连续函数。

习题3.3

1. 求正数 a 的平方根$\sqrt{a}$ 就是解方程 $x^2 = a$。古巴比伦人给出一个求 $\sqrt{a}$ 近似值的方法：$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$——对 a 的任何近似值 x_n，x_n 与 $\frac{a}{x_n}$ 必然一个是 $\sqrt{a}$ 的不足近似值，另一个是过剩近似值。作为 $\sqrt{a}$ 的近似值，$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ 总是比 x_n 与 $\frac{a}{x_n}$ 中较差的一个近似值更好。
   (1) 请对这个说法给予讨论。
   (2) 分别利用二分法和古巴比伦人的方法计算 $\sqrt{2}$ 的近似值，并对计算结果进行比较，你能解释你看到的现象吗？
   (3) 对不同的 $0 < \lambda < 1$，比较 $x_{n+1} = \lambda x_n + (1 - \lambda)\frac{a}{x_n}$ 的收敛效果。

2. 设 n 是正整数，连续函数 $f:\mathbb{R}\to\mathbb{R}$ 满足 $f(x)=x^{2n-1}+o(x^{2n-1}),\quad x\to\infty$。证明 f 是满射。问：是否存在 $N>0$ 使得对任意 $|y|>N$，$f(x)=y$ 有唯一解？若连续函数 $g:\mathbb{R}\to\mathbb{R}$ 满足 $g(x)=x^{2n}+o(x^{2n})$，$x\to\infty$，问 g 是否为满射？

3. 设 $f:I\to\mathbb{R}$ 是区间 I 上的一个单射。证明如果 f 连续，则 f 是严格单调的，从而 $f^{-1}:f(I)\to I$ 连续。

4. 设连续函数 $f:[a,b]\to\mathbb{R}$ 满足 $f[a,b]\subseteq[a,b]$ 或者 $[a,b]\subseteq f[a,b]$。证明 f 至少有一个不动点。如果把 $[a,b]$ 改成任意区间，结论是否还成立？

5. 用确界性质证明连续函数介值定理。设 $f:[a,b]\to\mathbb{R}$ 连续，$f(a)<c<f(b)$。记 $A=\{x\in[a,b]|f(x)<c\}$，证明 $\xi=\sup A$ 满足 $f(\xi)=c$。[1]

6. 证明例3.3.4中的结论。

3.4 有界闭集上的连续函数

引理 3.4.1. *$I\subset\mathbb{R}$ 是有界闭集当且仅当 I 中任何数列都含有在 I 中收敛的子列。*

证明. （必要性）设 I 是有界闭集，$x_n\in I(n\geq 1)$。则 $\{x_n\}_{n\geq 1}$ 是有界数列，从而有收敛子列 $\{x_{n_k}\}_{k\geq 1}$，$x_0=\lim\limits_{k\to+\infty}x_{n_k}$。因为 I 是闭集，所以 $x_0\in I$。

（充分性）设 I 中任何数列都含有在 I 中收敛的子列。

假若 I 无界，则存在 $x_n\in I$ 使得 $|x_n|>n$。于是 $\{x_n\}_{n\geq 1}$ 的任何子列都无界，从而没有收敛子列。这与已知矛盾。因此 I 有界。

设 $x_n\in I$ 且存在极限 $x_0=\lim\limits_{n\to+\infty}x_n$。根据已知，$\{x_n\}_{n\geq 1}$ 有在 I 中收敛的子列 $\{x_{n_k}\}_{k\geq 1}$。因为 $x_0=\lim\limits_{k\to+\infty}x_{n_k}$，所以 $x_0\in I$。因此 I 是闭集。 □

定理 3.4.2. *设 $I\subseteq\mathbb{R}$ 是有界闭集，$f:I\to\mathbb{R}$ 连续。则 $f(I)=\{f(x)|x\in I\}$ 是有界闭集。*

证明. 任取 $f(x_n)\in f(I)$，其中 $x_n\in I$。

因为 I 是有界闭集，所以根据引理，$\{x_n\}_{n\geq 1}$ 有收敛子列 $\{x_{n_k}\}_{k\geq 1}$，满足 $x_0=\lim\limits_{k\to+\infty}x_{n_k}\in I$。由 f 的连续性知 $\lim\limits_{k\to+\infty}f(x_{n_k})=f(x_0)\in f(I)$，所以根据引理知 $f(I)$ 是有界闭集。 □

推论 3.4.3. *设 $I\subseteq\mathbb{R}$ 是有界闭集，$f:I\to\mathbb{R}$ 连续。则 $f(I)=\{f(x)|x\in I\}$ 有最大值和最小值*

[1]这是 Bolzano 给出的证明方法，见 Steve Ross 编著的 The Mathematical Works by Bernard Bolzano, Oxford University Press, 2006

证明. 由定理3.4.1知 $f(I) \subset \mathbb{R}$ 是有界闭集。因此 $M = \sup f(I), m = \inf f(I)$ 存在，并且 $M, m \in f(I)$。从而 M 和 m 分别是 f 在 I 上的最大值和最小值。 □

多项式方程求解是数学史中的重要问题之一。虽然只有四次和四次以下的多项式存在求根公式，即可以用多项式的次数通过有限多次四则运算和开方运算得到根的表达式，但是以下的代数学基本定理确保了多项式的根的存在性。这个定理有很多不同的证明，我们这里给出一个比较初等的证明。为此我们把之前的一些概念和结果推广到复数情形。

复变量函数 $f : \mathbb{C} \to \mathbb{C}$ 的连续性：对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得当 $|z - z_0| < \delta$ 时，$|f(z) - f(z_0)| < \varepsilon$。这里对复数 $z = x + \mathrm{i}y$（$x, y \in \mathbb{R}$ 是 z 的实部和虚部）$|z| = \sqrt{x^2 + y^2}$。

集合 $K \subseteq \mathbb{C}$ 的有界性、K 是闭集以及复数列的极限概念都与实数情形的定义形式相同，不再赘述。

集合 $K \subseteq \mathbb{C}$ 有界当且仅当 K 的实部集合 $\{x \in \mathbb{R} | \exists z \in K,\ x \text{ 是 } z \text{ 的实部}\}$ 与 K 的虚部集合 $\{y \in \mathbb{R} | \exists z \in K,\ y \text{ 是 } z \text{ 的虚部}\}$ 都是有界集。因此 $\mathbb{C}$ 中任何有界数列必有收敛子列。

与定理3.4.2和推论3.4.3的证明相同，可以得到：对任何连续函数 $f : \mathbb{C} \to \mathbb{C}$ 以及任何非空有界闭集 $K \subseteq \mathbb{C}$，$|f(z)|$ 在 K 上有最小值。

例 3.4.4 (代数学基本定理). 任何复系数非常值多项式至少有一个复数根。

证明. 设 $P(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$，$n \geq 1$。

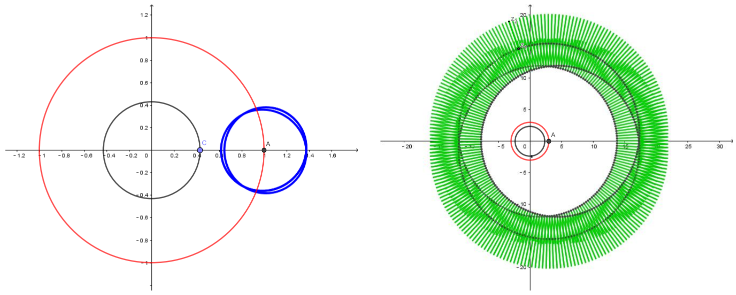


图 3.2: 左：存在 z 使得 $|Q(z)| < 1$；右：$|z| > R \Rightarrow |P(z)| > |P(0)|$

第一步：证明 $P(z)$ 没有离原点最近的非零值。

若 $P(z_0) \neq 0$，则令 $Q(z) = \frac{P(z+z_0)}{P(z_0)}$。于是 $Q(z)$ 是 n 次多项式满足 $Q(z) = 1 + b_m z^m + \cdots + b_n z^n$，其中正整数 $m \leq n$ 且 $b_m \neq 0$。

记 $\tilde{Q}(z) = 1 + b_m z^m$。设 $b_m = r\mathrm{e}^{\mathrm{i}\theta}$。取 $z^* = \varepsilon \mathrm{e}^{\mathrm{i}\frac{\pi-\theta}{m}}$，其中

$$0 < \varepsilon < \min\left\{1, \frac{1}{\sqrt[m]{r}}, \frac{r}{|b_{m+1}| + \cdots + |b_n|}\right\},$$

则 $\tilde{Q}(z^*) = 1 - r\varepsilon^m \in (0,1)$，

$$\begin{aligned}|Q(z^*)| &\le \left|Q(z^*) - \tilde{Q}(z^*)\right| + \left|\tilde{Q}(z^*)\right| \\ &\le (|b_{m+1}| + \cdots + |b_n|)\,\varepsilon^{m+1} + (1 - r\varepsilon^m) < 1,\end{aligned}$$

从而 $|P(z^* + z_0)| < |P(z_0)|$。即 $P(z_0)$ 不是 $P(z)$ 值域中离原点最近的点。

第二步：证明 $|P(z)|$ 有最小值 $|P(z_0)|$。

取 $R = 1 + 2\,(|a_{n-1}| + \cdots + |a_1| + |a_0|)$，则当 $|z| \ge R$ 时，

$$\begin{aligned}|P(z)| &= \left|z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0\right| \\ &\ge |z|^n - |a_{n-1}|\,|z|^{n-1} - \cdots - |a_1|\,|z| - |a_0| \\ &\ge |z|^n - (|a_{n-1}| + \cdots + |a_1| + |a_0|)\,|z|^{n-1} \\ &= |z|^n \left(1 - \frac{|a_{n-1}| + \cdots + |a_1| + |a_0|}{|z|}\right) \\ &\ge \frac{|z|^n}{2} \ge \frac{|z|}{2} > |a_0| = |P(0)|.\end{aligned}$$

因为 $P : \mathbb{C} \to \mathbb{C}$ 是连续函数，所以 $|P(z)|$ 在有界闭集 $K = \{z \in \mathbb{C}|\,|z| \le R\}$ 上有最小值 $|P(z_0)|$。从而对任意 $z \in \mathbb{C}$，

$$|P(z)| \ge \min\{|P(z)|\,, |P(0)|\} \ge |P(z_0)|\,.$$

所以 $|P(z_0)|$ 是 $|P(z)|$ 在 $\mathbb{C}$ 上的最小值。

再由第一步结论知 $P(z_0) = 0$。 □

习题3.4

1. 设 $I \subset \mathbb{R}$。证明 I 是有界闭集当且仅当 I 中每个数列 $\{x_n\}_{n\ge 1}$ 都有在 I 中收敛的子列。

2. 设 n 是正整数，连续函数 $f : \mathbb{R} \to \mathbb{R}$ 满足 $f(x) = x^{2n} + o(x^{2n}),\quad x \to \infty$。证明 f 有最小值。

3. (a) 用确界性质证明：设 $f : [a,b] \to \mathbb{R}$ 连续，则 f 在 $[a,b]$ 上有上界。（提示：考虑 $A = \{x \in [a,b] | f$ 在区间 $[a,x]$ 上有上界$\}$）

   (b) 证明连续函数 $f : [a,b] \to \mathbb{R}$ 有最大值。（提示：记 $M = \sup\{f(x) | a \le x \le b\}$，用反证法证明 M 是 f 的函数值，考虑 $g(x) = \frac{1}{M - f(x)}$）

4. 称 $f : I \to \mathbb{R}$ 是一个**下半连续函数**，如果对任意$x_0 \in I$ 以及任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得

$$x \in I, |x - x_0| < \delta \Longrightarrow f(x) > f(x_0) - \varepsilon.$$

   证明：如果 $I \subset \mathbb{R}$ 是有界闭集，则 f 在 I 上取得最小值。

3.5 函数的一致连续性

在一些实际问题中，一些物理量 y 可以从一些容易观测的物理量 x 通过特定的法则 $y=f(x)$ 计算得到。实际测量总会有误差，所以我们很难得到 x 的精确值，进而也就无法得到 y 的精确值。但是我们总还是可以得到 x 的一些近似值 $\hat{x}$，我们希望近似值 $\hat{y}=f(\hat{x})$ 不要明显偏离精确值 y。这就需要转化方式(即函数 f)是连续的。

但在实际问题中 f 仅仅连续是不够的。因为我们既然无从知道 x 的精确值，也就无从知道 $y=f(x)$ 的精确值，所以我们既无法直接知道误差 $|\hat{x}-x|$ 的大小，也无法知道误差 $|f(\hat{x})-f(x)|$ 的大小。一个现实的问题时，当我们不知道 x 的真值时，如果 x 的两个近似值（通常取一对不足近似值和过剩近似值） x_1,x_2 彼此足够接近时，是否能保证 $f(x_1),f(x_2)$ 也足够接近呢？或者说，$|x_1-x_2|$ 小到什么程度可以使误差 $|f(x_1)-f(x_2)|$ 满足给定的精度要求？这就要求 f 是一致连续函数。

定义 3.5.1. 称函数 $f:I\to\mathbb{R}$ 在 $K\subseteq I$ 上是**一致连续的**，如果对任意 $\varepsilon>0$，存在 $\delta_\varepsilon>0$ 使得对任意 $x_0\in K$ 以及任意 $x\in I$，只要 $|x-x_0|<\delta_\varepsilon$，就有 $|f(x)-f(x_0)|<\varepsilon$。这里 δ_ε 只与 ε 有关，与 $x_0\in K$ 和 $x\in I$ 无关。

例 3.5.2. 称函数 $f:I\to\mathbb{R}$ 是一个 **Lipschitz 函数**，如果存在常数 $L>0$ 使得对任意 $x,y\in I$，$|f(x)-f(y)|\le L|x-y|$。Lipschitz 函数都是一致连续的。

例 3.5.3. 由例2.1.3的证明知函数 $\sqrt{x}$ 在 $[0,+\infty)$ 上是一致连续的。

例 3.5.4. 函数 x^2 在 $\mathbb{R}$ 上不是一致连续的，但在任何有界闭区间 $[-N,N]$ 上是一致连续的。

证明. 对任意 $\varepsilon>0$，取 $\delta_\varepsilon=\frac{\varepsilon}{2N+1+\varepsilon}$，则对任意 $x_0\in[-N,N]$ 以及 $|x-x_0|<\delta_\varepsilon$，$|x|\le|x_0|+\delta_\varepsilon<N+1$，$\left|x^2-x_0^2\right|=|x-x_0|\,|x+x_0|\le|x-x_0|\,(|x|+|x_0|)\le(2N+1)\,|x-x_0|<\varepsilon$。所以在任何有界闭区间 $[-N,N]$ 上，x^2 是一致连续的。

假设 x^2 在 $\mathbb{R}$ 上是一致连续的。因为 $\left|\left(N+\frac{1}{N}\right)^2-N^2\right|=\left|2+\frac{1}{N^2}\right|\ge 2$，所以对 $0<\varepsilon<2$，相应的 δ_ε 必须满足 $0<\delta_\varepsilon<\frac{1}{N}$。而 N 是任意的，所以不存在定义中所说的只依赖与 ε 的正数 δ_ε。 □

例 3.5.5. 证明 $\lim\limits_{x\to+\infty}\left(\sin\sqrt{x^2+1}-\sin x\right)=0$。

证明. 因为

$$|\sin x-\sin y|=\left|2\sin\frac{x-y}{2}\cos\frac{x+y}{2}\right|\le|x-y|$$

所以 $\sin:\mathbb{R}\to\mathbb{R}$ 是一致连续函数。而当 $x\to+\infty$ 时，

$$\left|\sqrt{x^2+1}-x\right| = x\left|\sqrt{1+\frac{1}{x^2}}-1\right| = x\left(\frac{1}{2x^2}+o\left(\frac{1}{x^2}\right)\right) = \frac{1}{x}+o\left(\frac{1}{x}\right)\to 0,$$

所以 $\lim\limits_{x\to+\infty}\left(\sin\sqrt{x^2+1}-\sin x\right)=0$。 □

定理 3.5.6. 设 $f:I\to\mathbb{R}$ 连续，$K\subseteq I$ 是有界闭集。则 f 在 K 上是一致连续的。

证明. 假设 f 在 K 上不是一致连续的。则存在 $\varepsilon>0$ 以及数列 $x_n\in K$ 和 $y_n\in I$ 使得 $|x_n-y_n|<\frac{1}{n}$ 但 $|f(x_n)-f(y_n)|\geq\varepsilon$。因为 $x_n\in K$ 有界，所以 $\{x_n\}_{n\geq 1}$ 有收敛子列。不妨设 $\lim\limits_{n\to+\infty}x_n=x_0$。则 $x_0\in K$ 且 $\lim\limits_{n\to+\infty}y_n=x_0$。从而 $\lim\limits_{n\to+\infty}f(x_n)=f(x_0)=\lim\limits_{n\to+\infty}f(y_n)$。但这与 $|f(x_n)-f(y_n)|\geq\varepsilon$ 矛盾。所以 f 在 K 上是一致连续的。 □

习题3.5

1. $\sin(x^2)$ 是 $\mathbb{R}$ 上的一致连续函数吗？为什么？
2. 讨论 x^α 在区间 $(0,+\infty)$ 上的一致连续性。
3. 证明 $\mathbb{R}$ 上任何连续的周期函数都是一致连续的。
4. 讨论一致连续性与函数四则运算、复合、反函数之间的关系。
5. 用确界性质证明：设 $f:[a,b]\to\mathbb{R}$ 连续，则 f 在 $[a,b]$ 上一致连续。（提示：考虑 $A=\{x\in[a,b]|f$ 在区间 $[a,x]$ 上一致连续$\}$）

习题讨论课2

1. 以下陈述哪个与“$\{a_n\}_{n\geq 1}$ 是 Cauchy 列”等价？

   (a) 对任意 $\varepsilon > 0$，存在正整数 N 使得对任意正整数 $n \geq N$，$|a_n - a_N| < \varepsilon$。

   (b) 对任意 $\varepsilon > 0$，对任意正整数 p，存在正整数 N 使得对任意正整数 $n \geq N$，$|a_{n+p} - a_n| < \varepsilon$。

   (c) 对任意正整数 p，$\lim\limits_{n\to+\infty} |a_{n+p} - a_n| = 0$。

2. 关于闭区间套定理

   (a) 如果区间套不是闭区间，请问该定理结论是否成立？

   (b) 如果区间套不是有界区间，请问该定理结论是否成立？

   (c) 如果用有界闭集代替有界闭区间，请问该定理结论是否成立？

   (d) 请给出 $\mathbb{R}^n$ 中闭集的定义。在 $\mathbb{R}^n$ 中是否成立有界闭集套定理？

   (e) 用闭区间套性质和阿基米德性质证明任何 Cauchy 实数数列都收敛。

3. 如果实数数列 $\{a_n\}_{n\geq 1}$ 不含任何收敛子列，证明 $\lim\limits_{n\to+\infty} a_n = \infty$。

4. 设 $0 < \lambda < 1$，$x_1 = a$，$x_2 = b$，$x_n = \lambda x_{n-1} + (1-\lambda)x_{n-2}$。证明 $\{x_n\}_{n\geq 1}$ 收敛，并求 $\lim\limits_{n\to\infty} x_n$ 的值。

5. 设 $\lambda_1, \ldots, \lambda_k > 0$ 满足 $\lambda_1 + \cdots + \lambda_k = 1$。$x_1 = a_1, \ldots, x_k = a_k$，

   $$x_n = \lambda_1 x_{n-1} + \cdots + \lambda_k x_{n-k}.$$

   证明 $\{x_n\}_{n\geq 1}$ 收敛，并求 $\lim\limits_{n\to\infty} x_n$ 的值。。

6. 证明任给 $(x_1, y_1) \in (-\sqrt{2}, 2+\sqrt{2}) \times (-1-\sqrt{2}, 1+\sqrt{2})$，由

   $$\begin{cases} x_{n+1} = \dfrac{1+y_n^2}{2}, \\ y_{n+1} = x_n - \dfrac{x_n^2}{2} \end{cases}$$

   产生的数列 $\{x_n\}_{n\geq 1}, \{y_n\}_{n\geq 1}$ 都收敛。求或估计 $\lim\limits_{n\to\infty} x_n$，$\lim\limits_{n\to\infty} y_n$ 的值。

7. 设 $f : [a,b] \to \mathbb{R}$ 是连续函数，满足 $f[a,b] \subseteq [a,b]$。

   (a) 证明：存在 $\xi \in [a,b]$ 使得 $f(\xi) = \xi$。

   (b) 证明：如果 f 是单调函数，则对任何 $x_0 \in [a,b]$，要么 $f^n(x_0)$ 收敛于 f 的一个不动点，要么 $\{f^n(x_0)\}$ 收敛于 f 的一个周期为 2 的周期轨道：即存在 $\eta \in [a,b]$ 满足 $f^2(\eta) = \eta$，但 $f(\eta) \neq \eta$，使得

   $$\lim_{n\to\infty} \min\left\{|f^n(x_0) - \eta|, |f^n(x_0) - f(\eta)|\right\} = 0.$$

(c) 如果去掉 f 的单调性，(b)中结论是否成立？

8. 设 $f:[a,b]\to\mathbb{R}$ 是连续函数，满足 $[a,b]\subseteq f[a,b]$。证明：存在 $\xi\in[a,b]$ 使得 $f(\xi)=\xi$。

9. 设 $f:[a,+\infty)\to\mathbb{R}$ 连续，令

$$g(x)=\sup_{a\le t\le x} f(t).$$

(a) 证明 g 在 $[a,+\infty)$ 上连续。

(b) 若 f 是一致连续的，问 g 是否一致连续的？

10. 设 $I\subseteq\mathbb{R}$ 是区间，$f:I\to\mathbb{R}$ 是连续单射，证明：$f^{-1}:f(I)\to I$ 连续。

11. 设 $f:[a,+\infty)\to\mathbb{R}$ 一致连续，$g:[a,+\infty)\to\mathbb{R}$ 连续，满足 $\lim\limits_{x\to+\infty}(f(x)-g(x))=0$。证明 g 在 $[a,+\infty)$ 上一致连续。

12. 证明

(a) 连续的周期函数是一致连续的。

(b) 连续的非常值的周期函数的周期集合没有聚点。

13. 对 $t>0$，考虑 $f_t(x)=(1-x)^t-x$，$0\le x\le 1$。

(a) 证明函数 $f_t:[0,1]\to\mathbb{R}$ 连续且严格减。

(b) 证明对任意 $t>0$，$f_t(x)=0$ 有唯一解 $x_t\in(0,1)$。

(c) 证明 x_t 关于 $t>0$ 严格减，并且 $\lim\limits_{t\to+\infty}x_t=0$。

(d) 证明当 $t\to+\infty$ 时，$x_t=\frac{\ln t}{t}+o\left(\frac{\ln t}{t}\right)$。

14. 1911年德国数学家 Otto Toeplitz 提出了如下猜想：任何一条平面简单封闭曲线上必存在四个点，它们是一个正方形的四个顶点。这也称为“内接正方形问题”或“正方形桩子问题”。这里考虑一种特殊情况。设 $f:[0,1]\to\mathbb{R}$ 是一个连续函数，满足 $f(0)=f(1)=0<f(x)$ ($\forall 0<x<1$)。曲线 γ 由区间 $[0,1]$ 对应的线段和 f 的图像组成。证明：

(a) γ 有一个内接正方形。

(b) 对任意 $\lambda>0$，γ 有一个内接矩形，其相邻两边的长度比为 λ。

关于这个问题可以参见 https://www.webpages.uidaho.edu/ markn/squares/ 和 https://arxiv.org/pdf/2005.09193.pdf

第 4 章　导数与微分

> 事实上，这些运算的基础足够明显，但是因为我现在无法继续解释它，所以我宁愿把它隐藏成这样：*6accdae13eff7i3l9n4o4qrr4s8t12ux*。我在这个基础上，已经尝试过简化曲线积分的理论，并且得到了某些一般定理。
>
> 牛顿写给莱布尼茨的第2封信(1676年10月24日)

4.1　导数与微分的概念，函数的局部线性近似

Newton 为了研究运动的一般规律而创造了微积分。他的伟大发现是：运动的整体和长久的行为(比如由 Kepler 的三个行星运动定律所刻画的现象)可以由局部的、瞬时的运动规律（例如万有引力的平方反比关系）决定。而刻画局部、瞬时的规律就需要引入导数和微分的概念。

运动学中把位移视为时间的函数 $S(t)$，在从 t_0 到 t_1 的时段内的平均速度为

$$\bar{v} = \frac{S(t_1) - S(t_0)}{t_1 - t_0},$$

引入平均速度，可以把从 t_0 到 t_1 时物体从 $S(t_0)$ 移动到 $S(t_1)$ 的运动等效成以 $\bar{v}$ 的匀速直线运动。

一般地，变量 y 是自变量 x 的函数 $y = f(x)$，当自变量从 a 到 x 的过程中函数值 y 的平均变化率为

$$\frac{\Delta y}{\Delta x} = \frac{f(x) - f(a)}{x - a},$$

从 a 到 x 的过程中，y 的最终改变量可以用以 $\frac{f(x)-f(a)}{x-a}$ 为变化率的均匀变化来看待。

如果平均变化率为常数 λ，则 $\frac{\Delta y}{\Delta x} = \frac{f(x)-f(a)}{x-a} = \lambda$，从而 $f(x) = f(a) + \lambda(x - a)$，在 (x, y) 坐标平面上，函数 f 的图像为一条直线，λ 是这条直线的斜

率。一般情形下，$\frac{\Delta y}{\Delta x}=\frac{f(x)-f(a)}{x-a}$ 是函数 f 的图像过 $(a,f(a))$ 和 $(x,f(x))$ 两点的直线（称为 $y=f(x)$ 的割线）的斜率。

如果极限 $\lim\limits_{x\to a}\frac{f(x)-f(a)}{x-a}=\lambda$，则 $f(x)=f(a)+\lambda(x-a)+o(x-a)$，$x\to a$。即对充分接近于 a 的所有 x，y 的改变量 $f(x)-f(a)$ 的主要部分 $\lambda(x-a)$ 与 x 的改变量 $x-a$ 成比例（比例系数为 λ），我们称之为 $f(x)-f(a)$ 的线性主要部分。微分学的核心想法就是当自变量发生微小变化时找到相应的函数值变化的线性主要部分。它使得我们可以通过局部上把函数看成简单的一次函数 $y=kx+b$ 的方式对复杂函数进行研究，也使得我们可以通过研究曲线的切线而获得曲线自身的性质。我们希望提请读者注意的是，作为比例系数或切线斜率的导数只是一个便于计算的量，而作为线性函数或几何中切线的微分才是微分学中真正核心的对象。

为了保证自变量可以在一定范围内自由变化，我们引入内点的概念。

定义 4.1.1. 称 a 是集合 $I\subseteq\mathbb{R}$ 的一个**内点**，如果存在正数 $\delta>0$ 使得

$$\{x\in\mathbb{R}|\,|x-a|<\delta\}\subseteq I.$$

即 a 及其附近所有的点都在 I 中。

称 $I\subseteq\mathbb{R}$ 是一个**开集**，如果任何 $a\in I$ 都是 I 的内点。

定义 4.1.2. 称 $L:\mathbb{R}\to\mathbb{R}$ 是一个**线性函数**，如果它满足对任意 $x,y\in\mathbb{R}$ 以及任意 $\lambda\in\mathbb{R}$ 都有，$L(x+y)=L(x)+L(y)$，$L(\lambda x)=\lambda L(x)$。

取 $x=1$ 得到 $L(\lambda)=L(1)\lambda$，所以线性函数 L 就是以 $L(1)$ 为系数的正比例函数。

定义 4.1.3. 设 a 是集合 $I\subseteq\mathbb{R}$ 的一个内点，称 $f:I\to\mathbb{R}$ 在 a 处**可微**，如果存在常数 A 使得

$$f(a+h)=f(a)+Ah+o(h),\quad h\to 0,$$

此时称以 A 为系数的线性函数 $h\mapsto Ah$ 为 f 在 a 处的**微分**，记为

$$\mathrm{d}f(a):\mathbb{R}\to\mathbb{R},\quad \mathrm{d}f(a)(h)=Ah.$$

易见

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

称上式等号右侧的极限为 f 在 a 处的**导数**，记为 $f'(a) = \dfrac{\mathrm{d}f}{\mathrm{d}x}(a)$。所以，

$$\mathrm{d}f(a)(h) = f'(a)h.$$

传统上，人们习惯用变量代替函数，对 $y = f(x)$，导数 $f'(a)$ 也用 Leibniz 符号写成 $\left.\dfrac{\mathrm{d}y}{\mathrm{d}x}\right|_a$。

如果 f 在任意 $x \in I$ 处都可微，则称 **f 在 I 上可微**。此时，称 $f' : I \to \mathbb{R}$，$x \mapsto f'(x)$ 为 f 在 I 上的**导函数**。

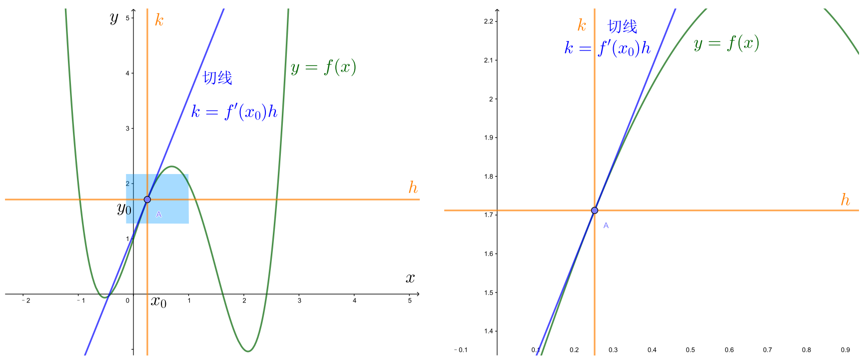


图 4.1: 微分：函数图像的切线/函数的局部线性近似

注 4.1.4. （**微分的几何含义**）上述定义表明，若 f 在 a 处可微，则

$$f(x) = f(a) + f'(a)(x-a) + o(x-a), \quad x \to a,$$

即在 a 附近 f 可以近似写为

$$y = f(a) + f'(a)(x-a)$$

的形式，几何上它是 (x, y) 坐标平面中过点 $(a, f(a))$ 并以 $f'(a)$ 为斜率的直线，我们把这条直线称为函数图像 $y = f(x)$ 在点 $(a, f(a))$ 处的**切线**。

在点 $P(a, f(a))$ 附近，引入新坐标系

$$X = x - a, \quad Y = y - f(a).$$

则 P 是 (X, Y) 坐标系中的原点，函数 f 的图像写成 $Y = f(X+a) - f(a)$，它在点 P 处的切线方程写为 $Y = f'(a)X$，这切线就是 f 在 a 处的微分 $\mathrm{d}f(a)$ 的图像。

注 4.1.5. （ **微分的符号表达** ）由于历史和习惯的原因，人们用变量来表示函数，而且不区分函数与函数值，函数 f 常被写成 $y = f(x)$ 的形式（例如 $y = x^2$），此时 $y, f(x), f$ 被视为都表示函数 f。因而函数 f 在 x 处的微分 $\mathrm{d}f(x)$ 也被写成 $\mathrm{d}y$。

当 x 是自变量时，恒同映射 $\iota : \mathbb{R} \to \mathbb{R}$ $\iota(x) = x$。此时 x 也是函数值，把它视同于恒同映射 ι。因为

$$\iota(x+h) = x + h = \iota(x) + h,$$

所以 ι 可微，且 $\mathrm{d}\iota(x)(h) = h$，因此 $\mathrm{d}x(h) = h$。从而对任意可微函数 f，

$$\mathrm{d}f(x)(h) = f'(x)h = f'(x)\mathrm{d}x(h), \quad \forall h \in \mathbb{R}.$$

所以作为函数，

$$\mathrm{d}f(x) = f'(x)\mathrm{d}x.$$

需要指出的是，这里 $\mathrm{d}f(x)$ 和 $f'(x)$ 中的 x 是相同的，分别指在 x 处计算微分和导数；$\mathrm{d}x$ 中的 x 是函数，$\mathrm{d}x$ 本身就是恒同映射。

定理 4.1.6. *若 f 在 a 可微，则对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得对任意 $x \in I$ 且 $|x - a| < \delta$，都有*

$$|f(x) - f(a) - f'(a)(x-a)| \leq \varepsilon |x - a|.$$

从而 f 在 a 连续。 □

例 4.1.7. 任何线性函数 $L(x) = \lambda x$ 是可微的：$L(x+h) - L(x) = L(h)$，而且对于任何 $a \in \mathbb{R}$，$\mathrm{d}L(a)(h) = L(h)$，从而$\mathrm{d}L(a) = L$。

例 4.1.8. 求 $f(x) = x^2$ 的导数与微分。

解：

$$(x+h)^2 - x^2 = 2xh + h^2 = 2xh + o(h), \quad h \to 0,$$

其中 $2xh$ 关于变量 h 是线性的，所以 $\mathrm{d}(x^2)(h) = 2xh$，$(x^2)' = 2x$。 □

例 4.1.9. 求 $f(x) = \dfrac{1}{x}$ 的导数与微分。

解： 对 $x \neq 0$，

$$\frac{1}{x+h} - \frac{1}{x} = \frac{-h}{x(x+h)} = -\frac{h}{x^2} + \frac{h^2}{x^2(x+h)} = -\frac{h}{x^2} + o(h), \quad h \to 0,$$

其中 $-\frac{h}{x^2}$ 关于变量 h 是线性的，所以 $\mathrm{d}\left(\frac{1}{x}\right)(h) = -\frac{h}{x^2}$，$\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$。 □

例 4.1.10. 求 $\exp(x) = \mathrm{e}^x$ 的导数与微分。

解：

$$e^{x+h} - e^x = e^x(e^h - 1) = e^x(h + o(h)) = e^x h + o(h), \quad h \to 0,$$

其中 $e^x h$ 关于变量 h 是线性的，所以 $d(e^x)(h) = e^x h$，$(e^x)' = e^x$。 □

例 4.1.11. 求 $\sin x$ 的导数与微分。
解：

$$\begin{aligned}\sin(x+h) - \sin x &= \cos x \sin h + \sin x(\cos h - 1)\\ &= \cos x \cdot (h + o(h)) + \sin x \cdot o(h)\\ &= \cos x \cdot h + o(h), \quad h \to 0\end{aligned}$$

其中 $\cos x \cdot h$ 关于变量 h 是线性的，所以 $d(\sin x)(h) = \cos x \cdot h$，$(\sin x)' = \cos x$。

□

习题4.1

1. 证明： f 在 x 处可微当且仅当存在 $A \in \mathbb{R}$ 使得对任意 $B \in \mathbb{R}\backslash\{A\}$，

   $$f(x+h) - f(x) - Ah = o\big(f(x+h) - f(x) - Bh\big), \quad h \to 0.$$

   此时 $A = f'(x)$。因此函数在一点处可微当且仅当它有唯一的局部最佳线性近似，这个最佳线性近似就是函数的微分。

2. 有人说，可微函数 f 在点 x 处的微分 $f'(x)h$ 是函数改变量 $f(x+h) - f(x)$ 的主要部分。你是否同意这个说法？为什么？

3. 对一元实数变量函数 f，我们可以定义 f 在点 x 处的左侧或右侧的导数，分别记为 $f'_-(x)$ 和 $f'_+(x)$。

   (a) 请给出 $f'_-(x)$ 和 $f'_+(x)$ 的具体定义陈述。

   (b) 证明：f 在 x 处可导当且仅当 f 在点 x 处的两侧都存在单侧导数，且 $f'(x) = f'_-(x) = f'_+(x)$。

   (c) 请分别举例说明存在这样的函数 f ，它在 x 处

      i. f 不可导，但在 x 两侧都可导；

      ii. 仅在 x 的一侧可导；

      iii. 两侧都不可导。

4.2 导数与微分的运算法则

两个线性函数 $u = 2y$ 和 $y = -3x$ 的复合 $u = 2(-3x) = -6x$ 仍然是线性函数，它们的系数满足 $-6 = 2 \times (-3)$。微分是函数的局部线性近似，复合函数微分满足如下结论。

定理 4.2.1 (复合函数导数的链索法则). 设 f 在 a 可微，g 在 $b=f(a)$ 处可微，则 $g\circ f$ 在 a 处可微，且

$$\mathrm{d}(g\circ f)(a)=\mathrm{d}g(b)\circ\mathrm{d}f(a)=\mathrm{d}g(f(a))\circ\mathrm{d}f(a),\quad (\text{复合的微分}=\text{微分的复合})$$

即 $(g\circ f)'(a)=g'(b)f'(a)=g'(f(a))f'(a)$。

如果记 $u=g(y)$，$y=f(x)$，使用 *Leibniz* 的符号，则上述等式写成

$$\left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_a=\left.\frac{\mathrm{d}u}{\mathrm{d}y}\right|_b\cdot\left.\frac{\mathrm{d}y}{\mathrm{d}x}\right|_a.$$

证明. 记 $L=\mathrm{d}f(a)$，$K=\mathrm{d}g(b)$。由 $f(a+h)=f(a)+L(h)+\alpha(h)$，$g(b+w)=g(b)+K(w)+\beta(w)$，得到

$$\begin{aligned}g(f(a+h))&=g(f(a))+K(f(a+h)-f(a))+\beta(f(a+h)-f(a))\\&=g(f(a))+K(L(h)+\alpha(h))+\beta(L(h)+\alpha(h))\\&=g(f(a))+K(L(h))+K(\alpha(h))+\beta(L(h)+\alpha(h))\end{aligned}$$

当 $0<|h|<\delta$ 时，$|\alpha(h)|\le\varepsilon|h|$。当 $0<|w|<\delta$ 时，$|\beta(w)|\le\varepsilon|w|$。因为 $\alpha(0)=\beta(0)$，所以这两个不等式对 $h=0$ 和 $w=0$ 也成立。所以

$$|K(\alpha(h))|=|g'(b)|\,|\alpha(h)|\le\varepsilon\,|g'(b)|\,|h|.$$

当 $|h|<\dfrac{\delta}{1+|f'(a)|+\varepsilon}$ 时，

$$|L(h)+\alpha(h)|\le|f'(a)|\,|h|+\varepsilon\,|h|\le(|f'(a)|+\varepsilon)\,|h|<\delta,$$

所以 $|\beta(L(h)+\alpha(h))|\le\varepsilon\,|L(h)+\alpha(h)|\le\varepsilon\,(|f'(a)|+\varepsilon)\,|h|$，从而

$$\begin{aligned}|K(\alpha(h))+\beta(L(h)+\alpha(h))|&\le\varepsilon\,|g'(b)|\,|h|+\varepsilon\,(|f'(a)|+\varepsilon)\,|h|\\&=\varepsilon\,(|f'(a)|+|g'(b)|+\varepsilon)\,|h|,\end{aligned}$$

因此 $K(\alpha(h))+\beta(L(h)+\alpha(h))=o(h),\quad h\to0$，

$$g(f(a+h))=g(f(a))+K\circ L(h)+o(h),\quad h\to0,$$

从而 $g\circ f$ 在 a 处可微，且

$$\mathrm{d}(g\circ f)(a)=K\circ L,\qquad (g\circ f)'(a)=g'(b)f'(a).$$

□

注 4.2.2. 设 $z=z(y)$ ，$y=f(x)$，以及 $x=x(u)$ 都可微，则对于 $z=g(u)=z(f(x(u)))$，

$$\mathrm{d}z=\frac{\mathrm{d}z}{\mathrm{d}y}\mathrm{d}y,\qquad \mathrm{d}y=\frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x,\qquad \mathrm{d}x=\frac{\mathrm{d}x}{\mathrm{d}u}\mathrm{d}u.$$

由链索法则

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

所以 $\mathrm{d}z=\frac{\mathrm{d}z}{\mathrm{d}y}\mathrm{d}y=\frac{\mathrm{d}z}{\mathrm{d}x}\mathrm{d}x=\frac{\mathrm{d}z}{\mathrm{d}u}\mathrm{d}u$，这被称为“一阶微分的形式不变性”，其本质是曲线之间的相切关系与坐标系的选取无关。

定理 4.2.3. 设 f,g 在 a 可微，则 $\alpha f+\beta g, fg(\alpha,\beta\in\mathbb{R})$在 a 处可微，且

1. *(线性)*

$$\mathrm{d}(\alpha f+\beta g)(a)=\alpha\mathrm{d}f(a)+\beta\mathrm{d}g(a),\quad (\alpha f+\beta g)'(a)=\alpha f'(a)+\beta g'(a);$$

2. *(Leibniz)*

$$\mathrm{d}(fg)(a)=g(a)\mathrm{d}f(a)+f(a)\mathrm{d}g(a),\quad (fg)'(a)=g(a)f'(a)+f(a)g'(a);$$

3. 如果 $g(a)\neq 0$，则 $\frac{f}{g}$ 在 a 处可微，且

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

证明. 线性性质留给读者作为练习，除法结论可由乘法和链索法则以及例4.1.9的结论得到。下面只证明乘法的结论。

由 $f(a+h)=f(a)+\mathrm{d}f(a)(h)+o(h)$，$g(a+h)=g(a)+\mathrm{d}g(a)(h)+o(h)$，得到

$$\begin{aligned}&f(a+h)g(a+h)-f(a)g(a)\\ =&[f(a)+\mathrm{d}f(a)(h)+o(h)]\cdot[g(a)+\mathrm{d}g(a)(h)+o(h)]-f(a)g(a)\\ =&g(a)\mathrm{d}f(a)(h)+f(a)\mathrm{d}g(a)(h)+\mathrm{d}f(a)(h)\cdot\mathrm{d}g(a)(h)+o(h)\\ =&g(a)\mathrm{d}f(a)(h)+f(a)\mathrm{d}g(a)(h)+o(h)\end{aligned}$$

所以 $\mathrm{d}(fg)(a)=g(a)\mathrm{d}f(a)+f(a)\mathrm{d}g(a)$。 □

例 4.2.4. $(a^x)'=\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{x\ln a}\right)=\frac{\mathrm{d}}{\mathrm{d}y}\mathrm{e}^y\Big|_{y=x\ln a}\cdot\frac{\mathrm{d}}{\mathrm{d}x}(x\ln a)=\mathrm{e}^{x\ln a}\ln a=a^x\ln a$。

例 4.2.5. 求 $(\cos x)'$ 和 $(\tan x)'$。

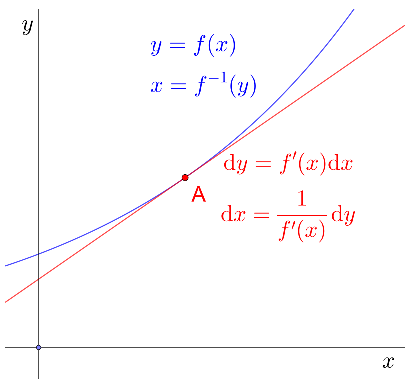
解：

$$\begin{aligned}(\cos x)'&=(\sin(\pi/2-x))'=\cos(\pi/2-x)(\pi/2-x)'\\ &=\sin x\cdot(-1)=-\sin x.\\ (\tan x)'&=\left(\frac{\sin x}{\cos x}\right)'=\frac{\cos x(\sin x)'-\sin x(\cos x)'}{\cos^2 x}\\ &=\frac{\cos^2 x-\sin x(-\sin x)}{\cos^2 x}=\frac{1}{\cos^2 x}=1+\tan^2 x.\end{aligned}$$

□

当 $\lambda \neq 0$ 时，线性函数 $y=\lambda x$ 有反函数 $x=\lambda^{-1}y$，对于微分成立如下结论。

定理 4.2.6. *设函数 $f:(a,b)\to(\alpha,\beta)$ 是严格单调的连续函数，f 在 x_0 可微，且 $f'(x_0)\neq 0$。则 f 的反函数 f^{-1} 在 $y_0=f(x_0)$ 可微，且 $\mathrm{d}(f^{-1})(y_0)=(\mathrm{d}f(x_0))^{-1}$ (反函数的微分是微分的反函数)，$\left(f^{-1}\right)'(y_0)=\dfrac{1}{f'(x_0)}$。*



证明. $y=y_0+f'(x_0)(x-x_0)+o(x-x_0)$，$x\to x_0$。

因为 $f'(x_0)\neq 0$，所以

$$y-y_0=f'(x_0)(x-x_0)+o\left(f'(x_0)(x-x_0)\right),\quad x\to x_0.$$

从而

$$y-y_0+o(y-y_0)=f'(x_0)(x-x_0),\quad x\to x_0.$$

$x\to x_0$ 时 $y\to y_0$。又反函数连续，所以 $y\to y_0$ 时，$x=f^{-1}(y)\to x_0=f^{-1}(y_0)$，所以

$$f^{-1}(y)-f^{-1}(y_0)=x-x_0=\frac{y-y_0}{f'(x_0)}+o(y-y_0),\quad y\to y_0.$$

从而 f^{-1} 在 $y_0=f(x_0)$ 可微，且 $\left(f^{-1}\right)'(y_0)=\dfrac{1}{f'(x_0)}$。 □

例 4.2.7.

$$(\ln x)'=\frac{1}{\frac{\mathrm{d}}{\mathrm{d}y}(\mathrm{e}^y)}=\left.\frac{1}{\mathrm{e}^y}\right|_{y=\ln x}=\frac{1}{x},$$

$$(x^\alpha)'=\left(\mathrm{e}^{\alpha\ln x}\right)'=\mathrm{e}^{\alpha\ln x}\cdot\frac{\alpha}{x}=x^\alpha\cdot\frac{\alpha}{x}=\alpha x^{\alpha-1},$$

$$(x^x)'=\left(\mathrm{e}^{x\ln x}\right)'=\mathrm{e}^{x\ln x}\cdot\left(\ln x+x\cdot\frac{1}{x}\right)=x^x(\ln x+1),$$

$$(\arcsin x)' = \frac{1}{\frac{\mathrm{d}}{\mathrm{d}y}(\sin y)\Big|_{y=\arcsin x}} = \frac{1}{\cos y}\Big|_{y=\arcsin x} = \frac{1}{\sqrt{1-x^2}},$$

$$(\arccos x)' = \frac{1}{\frac{\mathrm{d}}{\mathrm{d}y}(\cos y)\Big|_{y=\arccos x}} = \frac{1}{-\sin y}\Big|_{y=\arccos x} = -\frac{1}{\sqrt{1-x^2}},$$

$$(\arctan x)' = \frac{1}{\frac{\mathrm{d}}{\mathrm{d}y}(\tan y)\Big|_{y=\arctan x}} = \frac{1}{1+\tan^2 y}\Big|_{y=\arctan x} = \frac{1}{1+x^2}.$$

□

例 4.2.8. 当 $0<\varepsilon<1$ 时，$f(x)=x-\varepsilon\sin x$ 有连续的反函数(习题3.2第1题)。$f'(x)=1-\varepsilon\cos x>0$，所以 f 的反函数 f^{-1} 可微，并且对等式

$$f^{-1}(y)-\varepsilon\sin f^{-1}(y)=y$$

两边求导得到

$$\frac{\mathrm{d}}{\mathrm{d}y}f^{-1}(y)-\varepsilon\cos f^{-1}(y)\frac{\mathrm{d}}{\mathrm{d}y}f^{-1}(y)=1,$$

从而

$$\frac{\mathrm{d}}{\mathrm{d}y}f^{-1}(y)=\frac{1}{1-\varepsilon\cos f^{-1}(y)}.$$

□

习题4.2

1. 称两个函数 f,g 在 x 处**相切**，如果 f,g 在 x_0 处连续，并且

$$f(x)-g(x)=o(x-x_0),\quad x\to x_0.$$

证明

(a) “相切”是一个等价关系；

(b) f 在 x_0 处可微当且仅当 f 与某个一次函数 $A+B(x-x_0)$ 在 x_0 处相切，此时 $A=f(x_0)$，$B=f'(x_0)$。

(c) 若 f 在 x_0 处可微，且与 g 在 x_0 处相切，则 $f(x_0)=g(x_0)$ 且 $f'(x_0)=g'(x_0)$。

(d) 若 f,g 在 x_0 处相切，函数 u 满足在 x_0 的一个邻域中

$$f(x)\le u(x)\le g(x),$$

则 u,f 在 x_0 处相切。

2. 记 $f(x)=x$，$g(x)=x+x^2$。取

$$x_n=\frac{1}{2^n},\quad y_n=\frac{1}{2^n}+\frac{1}{5^n},\quad n\in\mathbb{N}^*.$$

(a) 证明：对任意正整数n，都有 $x_{n+1} < y_{n+1} < x_n$；

(b) 令函数u满足：在区间 $[y_{n+1}, x_n]$ 上 $y = u(x)$ 是连接点 $B_{n+1}(y_{n+1}, y_{n+1})$ 和点 $A_n(x_n, g(x_n))$ 的直线段；在区间 $[x_{n+1}, y_{n+1}]$ 上 $y = u(x)$ 是连接点 A_{n+1} 和点 B_{n+1} 的直线段；在区间 $(-\infty, 0]$ 上，$u(x) = x$。证明：u 是连续函数，在 $x = 0$ 处可微，并且 $u'(0) = 1$，但在任何区间 $[0, \delta)$ 内，u 都不是单调函数。

4.3 高阶导数

定义 4.3.1. 若 $f : I \to \mathbb{R}$ 在每个 $x \in I$ 处都可微，则称 $f' : I \to \mathbb{R}$ 为 f 的**一阶导函数**。若 f' 是连续函数，则称 f 是 $\mathscr{C}^1$ **函数**。

如果 f' 在 $x_0 \in I$ 处可微，则称 f 在 x_0 处**二阶可微**，记

$$f''(x_0) = (f')'(x_0), \quad \left.\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right|_{x_0} = \left.\frac{\mathrm{d}}{\mathrm{d}x}\left(\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_x\right)\right|_{x_0},$$

称为 f 在 x_0 处的**二阶导数**。

一般地，若 f 的 n 阶导函数 $f^{(n)}$ 在 $x_0 \in I$ 处可微，则称 f 在 x_0 处 $n+1$ **阶可微**，记

$$f^{(n+1)}(x_0) = \left(f^{(n)}\right)'(x_0), \quad \left.\frac{\mathrm{d}^{n+1} f}{\mathrm{d}x^{n+1}}\right|_{x_0} = \left.\frac{\mathrm{d}}{\mathrm{d}x}\left(\left.\frac{\mathrm{d}^n f}{\mathrm{d}x^n}\right|_x\right)\right|_{x_0},$$

称为 f 在 x_0 处的$n+1$ **阶导数**。

如果 n 阶导函数 $f^{(n)}$ 连续，则称 f 是C^n **函数**。

称 f 是$\mathscr{C}^\infty$ **函数**，如果对任意正整数 n，f 是 C^n 函数。

用 Leibniz 的符号，n 阶求导运算 $\dfrac{\mathrm{d}^n}{\mathrm{d}x^n}$ 可以看成算子(一个操作)，它是把 $\dfrac{\mathrm{d}}{\mathrm{d}x}$ 连续作用 n 次：

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} = \underbrace{\frac{\mathrm{d}}{\mathrm{d}x} \cdots \frac{\mathrm{d}}{\mathrm{d}x}}_{n\text{个}}.$$

例 4.3.2. Newton 的运动第二定律说，物体运动的加速度与物体所受的外力成正比，即

$$F = ma = m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}.$$

其中 $x = x(t)$ 是位移函数，$a = \dfrac{\mathrm{d}^2 x}{\mathrm{d}t^2}$ 是加速度，F 是外力，t 是时间。

例如，在弹簧振子中，按 Hook 定律，弹簧作用在振子上的力大小与振子偏离弹簧平衡位置的距离成正比，方向相反。于是

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx.$$

这是一个关于 $x(t)$ 的二阶微分方程。

现在让我们分头来计算下面这个问题。

例 4.3.3. 对 $y = \mathrm{e}^{\sin(x^2)}$，求 y''。

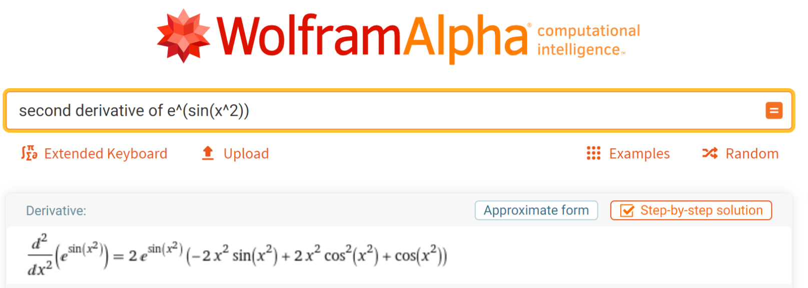
解：

$$y' = \mathrm{e}^{\sin(x^2)}\left(\sin(x^2)\right)' = \mathrm{e}^{\sin(x^2)}\cos(x^2)\cdot 2x.$$

$$\begin{aligned} y'' &= \mathrm{e}^{\sin(x^2)}\left(\cos(x^2)\cdot 2x\right)^2 + \mathrm{e}^{\sin(x^2)}\left[-\sin(x^2)\cdot(2x)^2 + \cos(x^2)\cdot 2\right] \\ &= \mathrm{e}^{\sin(x^2)}\left[4x^2\cos^2(x^2) - 4x^2\sin(x^2) + 2\cos(x^2)\right]. \end{aligned}$$

$\square$

你算对了吗？我相信你根本没有兴趣去计算三阶导数 $y^{(3)}$。一方面，这些计算虽然繁琐，但完全是根据确定的规则按部就班进行，现在的计算机软件完全能够胜任。有了得心应手的工具，这样的事为什么不交给机器自动完成呢？人的智慧不应该用于处理更有价值更有挑战的问题吗？另一方面，也是更重要的，你可能会问：为什么要去算这些高阶导数呢？这个问题等我们在后面介绍了 Taylor 公式你就会明白，高阶导数可以帮助我们用多项式来近似更复杂的函数。当然还有一点，很多实际问题的数学模型是用高阶微分方程刻画的。



通常高阶导数计算是比较复杂的，对某些简单运算有一些规律性的结果。

定理 4.3.4. 设 f, g 在 x_0 处 n 阶可微，则 $\alpha f + \beta g, fg$ 都在 x_0 处 n 阶可微，且

$$(\alpha f + \beta g)^{(n)} = \alpha f^{(n)} + \beta g^{(n)}, \qquad \text{(线性)}$$

$$(f\cdot g)^{(n)} = \sum_{k=0}^{n}\binom{n}{k} f^{(k)} g^{(n-k)}. \qquad \text{(Leibniz)}$$

证明. 对 n 做归纳。线性留给读者自己验证。我们这里只证明乘积的高阶可微性和 Leibniz 公式。

$n=1$ 时，结论成立，

$$(f\cdot g)'=f'\cdot g+f\cdot g'.$$

假设 n 时结论成立。设 f,g 在 x_0 处 $n+1$ 阶可微。则 f',g' 在 x_0 处 n 阶可微。于是根据归纳假设

$$\begin{aligned}(f'\cdot g)^{(n)}&=\sum_{k=0}^{n}\binom{n}{k}f^{(k+1)}g^{(n-k)}=\sum_{j=1}^{n+1}\binom{n}{j-1}f^{(j)}g^{(n+1-j)},\\(f\cdot g')^{(n)}&=\sum_{j=0}^{n}\binom{n}{j}f^{(j)}g^{(n-j+1)}.\end{aligned}$$

因此 $(f\cdot g)'$ 在 x_0 处 n 阶可微，从而 $f\cdot g$ 在 x_0 处 $n+1$ 阶可微，并且

$$\begin{aligned}(f\cdot g)^{(n+1)}&=(f'\cdot g+f\cdot g')^{(n)}\\&=fg^{(n+1)}+f^{(n+1)}g+\sum_{j=1}^{n}\left[\binom{n}{j-1}+\binom{n}{j}\right]f^{(j)}g^{(n+1-j)}\\&=\sum_{j=0}^{n+1}\binom{n+1}{j}f^{(j)}g^{(n+1-j)}.\end{aligned}$$

因此 $n+1$ 时结论成立。从而结论对任意正整数 n 成立。 □

定理 4.3.5. [1] *设 f 在 x 处 n 阶可微，g 在 $y=f(x)$ 处 n 阶可微。则*

1. *$g\circ f$ 在 x_0 处 n 阶可微。*
2. *$(g(ax+b))^{(n)}=a^n g^{(n)}(ax+b)$。*
3. *若 $f(x_0)\neq 0$，则 $\frac{1}{f}$ 在 x_0 处 n 阶可微。*
4. *若 f 可逆且 $f'(x_0)\neq 0$，则 f^{-1} 在 $y_0=f(x_0)$ 处 n 阶可微。*

证明. (1) $n=1$ 时，结论成立且

$$(g\circ f)'=(g'\circ f)\cdot f'.$$

假设 n 时结论成立。设 f,g 在 x_0 处 $n+1$ 阶可微。则 f,g,f',g' 在 x_0 处 n 阶可微。于是根据归纳假设 $g'\circ f$ 在 x_0 处 n 阶可微，从而 $(g'\circ f)\cdot f'$ 在 x_0 处 n 阶可微，因此 $(g\circ f)'$ 在 x_0 处 n 阶可微，所以 $g\circ f$ 在 x_0 处 $n+1$ 阶可微。

(2) 证明留给读者。

(3) 因为 $f(x_0)\neq 0$ 且 f 连续，所以在 x_0 的一个邻域内 $f(x)\neq 0$ 。因为 $g(y)=\frac{1}{y}$ 是任意阶可微的，所以 $g\circ f=\frac{1}{f}$ 是 n 阶可微的。

[1]这些结论的另外一种证明可以参见小平邦彦所著《微积分入门》一书。

(4) $n=1$ 时，结论成立且

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

假设 n 时结论成立。设 f 在 x_0 处 $n+1$ 阶可微。则根据归纳假设 f', f^{-1} 在 x_0 处 n 阶可微。从而 $\frac{1}{f' \circ f^{-1}}$ 在 x_0 处 n 阶可微，从而 $(f^{-1})'$ 在 x_0 处 n 阶可微，因此 f^{-1} 在 x_0 处 $n+1$ 阶可微。 □

推论 4.3.6. *设 f, g 在 x_0 处 n 阶可微，且 $g(x_0) \neq 0$，则 $\frac{f}{g}$ 在 x_0 处 n 阶可微。* □

推论 4.3.7. *初等函数都是 $\mathscr{C}^\infty$ 函数。* □

例 4.3.8.

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

□

例 4.3.9.

$$\begin{pmatrix} \cos x \\ \sin x \end{pmatrix}' = \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \cos x \\ \sin x \end{pmatrix},$$

于是

$$\begin{pmatrix} \cos x \\ \sin x \end{pmatrix}^{(n)} = J^n \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad J^2 = -I, J^3 = -J, J^4 = I.$$

另外，如果记

$$\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i} \sin x,$$

则

$$(\mathrm{e}^{\mathrm{i}x})' = (\cos x)' + \mathrm{i}(\sin x)' = -\sin x + \mathrm{i}\cos x = \mathrm{i}\mathrm{e}^{\mathrm{i}x}.$$

因此

$$(\cos x)^{(n)} + \mathrm{i}(\sin x)^{(n)} = (\mathrm{e}^{\mathrm{i}x})^{(n)} = \mathrm{i}^n \mathrm{e}^{\mathrm{i}x} = \mathrm{i}^n \cos x + \mathrm{i}^{n+1} \sin x.$$

所以

$$(\cos x)^{(n)} = \operatorname{Re}\left(\mathrm{i}^n \cos x + \mathrm{i}^{n+1} \sin x\right) = \begin{cases} (-1)^k \cos x, & n = 2k; \\ (-1)^k \sin x, & n = 2k-1. \end{cases}$$

$$(\sin x)^{(n)} = \operatorname{Im}\left(\mathrm{i}^n \cos x + \mathrm{i}^{n+1} \sin x\right) = \begin{cases} (-1)^k \sin x, & n = 2k; \\ (-1)^{k+1} \cos x, & n = 2k-1. \end{cases}$$

□

例 4.3.10. 求 $\dfrac{1}{x^2-1}$ 的 n 阶导数。

解：

$$\begin{aligned} \left(\frac{1}{x^2-1}\right)^{(n)} &= \left(\frac{1}{2}\left(\frac{1}{x-1} - \frac{1}{x+1}\right)\right)^{(n)} \\ &= \frac{1}{2}\left(\frac{1}{x-1}\right)^{(n)} - \frac{1}{2}\left(\frac{1}{x+1}\right)^{(n)} \\ &= \frac{(-1)^{n-1}(n-1)!}{2}\left(\frac{1}{(x-1)^n} - \frac{1}{(x+1)^n}\right). \end{aligned}$$

□

例 4.3.11. 求 $\tan^{(3)}(0)$。

解： 记 $f(x) = \tan x$。则

$$\cos x \cdot f(x) = \sin x.$$

于是利用乘积高阶导数的 Leibniz 公式，得到

$$\begin{gathered} \cos 0 \cdot f'(0) - \sin 0 \cdot f(0) = \cos 0, \\ \cos 0 \cdot f''(0) - 2\sin 0 \cdot f'(0) - \cos 0 \cdot f(0) = -\sin 0, \\ \cos 0 \cdot f^{(3)}(0) - 3\sin 0 \cdot f''(0) - 3\cos 0 \cdot f'(0) + \sin 0 \cdot f(0) = -\cos 0. \end{gathered}$$

所以

$$f'(0) = 1, \quad f''(0) = 0, \quad f^{(3)}(0) = 2.$$

□

例 4.3.12. 设 $0 < \varepsilon < 1$。证明 $f(x) = x - \varepsilon \sin x$ 有 $\mathscr{C}^\infty$ 的反函数，并求 $(f^{-1})''(0)$ 和 $(f^{-1})^{(3)}(0)$ 的值。

解： $f \in \mathscr{C}^\infty(\mathbb{R})$，$f'(x) = 1 - \varepsilon \cos x > 0$，所以 $f^{-1} \in \mathscr{C}^\infty(\mathbb{R})$。设 $y = f(x)$，在恒等式

$$y = x - \varepsilon \sin x, \qquad (x = f^{-1}(y))$$

两边对 y 逐次求导，依次得到

$$
\begin{aligned}
1 &= \frac{\mathrm{d}x}{\mathrm{d}y} - \varepsilon\cos x \cdot \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\mathrm{d}x}{\mathrm{d}y}(1-\varepsilon\cos x),\\
0 &= \frac{\mathrm{d}^2x}{\mathrm{d}y^2}(1-\varepsilon\cos x) + \varepsilon\sin x\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2,\\
0 &= \frac{\mathrm{d}^3x}{\mathrm{d}y^3}(1-\varepsilon\cos x) + \frac{\mathrm{d}^2x}{\mathrm{d}y^2}\varepsilon\sin x\frac{\mathrm{d}x}{\mathrm{d}y} + \varepsilon\cos x\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^3 + 2\varepsilon\sin x\frac{\mathrm{d}x}{\mathrm{d}y}\frac{\mathrm{d}^2x}{\mathrm{d}y^2}.
\end{aligned}
$$

因为 $f(0)=0$，所以 $0=f^{-1}(0)$，代入上式得到

$$
\begin{aligned}
1 &= \left.\frac{\mathrm{d}x}{\mathrm{d}y}\right|_{y=0}(1-\varepsilon) \Rightarrow \left.\frac{\mathrm{d}x}{\mathrm{d}y}\right|_{y=0} = \frac{1}{1-\varepsilon},\\
0 &= \left.\frac{\mathrm{d}^2x}{\mathrm{d}y^2}\right|_{y=0}(1-\varepsilon) \Rightarrow \left.\frac{\mathrm{d}^2x}{\mathrm{d}y^2}\right|_{y=0} = 0,\\
0 &= \left.\frac{\mathrm{d}^3x}{\mathrm{d}y^3}\right|_{y=0}(1-\varepsilon) + \varepsilon\left(\left.\frac{\mathrm{d}x}{\mathrm{d}y}\right|_{y=0}\right)^3 \Rightarrow \left.\frac{\mathrm{d}^3x}{\mathrm{d}y^3}\right|_{y=0} = \frac{-\varepsilon}{(1-\varepsilon)^4}.
\end{aligned}
$$

在后面我们学习了 Taylor 公式后，我们会知道在 $x=0$ 附近，$f^{-1}(y)$ 可以近似表示为

$$
x = \frac{y}{1-\varepsilon} - \frac{\varepsilon y^3}{6(1-\varepsilon)^4}.
$$

□

习题4.3

1. 证明定理 4.3.4 和定理 4.3.5。

2. 设 f,g 都是 n 阶可微函数。证明

$$
(g\circ f)^{(n)} = \sum_{j=0}^{n-1}\binom{n-1}{j}(g'\circ f)^{(j)}\cdot f^{(n-j)}.
$$

3. 设 f 是 n 阶可微函数，$f(x)\neq 0$。证明

$$
\left(\frac{1}{f}\right)^{(n)} = -\sum_{0\le i\le k<n}\frac{(n-1)!}{i!(k-i)!(n-1-k)!}\left(\frac{1}{f}\right)^{(i)}\cdot\left(\frac{1}{f}\right)^{(k-i)}\cdot f^{n-k}.
$$

4. 求 $(\arctan x)^{(n)}$。

5. 对 $x\neq 0$ 及正整数 n，求 $\left(\frac{\mathrm{e}^x-1}{x}\right)^{(n)}$。

6. 证明 $x(t)=\sin\left(\sqrt{\frac{k}{m}}t\right)$ 是方程

$$
m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx
$$

的一个解。你能写出更多的解吗？你能写出所有解吗？

7. 设 $A, \alpha, \beta \in \mathbb{R}$，讨论函数

$$f(x) = \begin{cases} |x|^\alpha \ln|x|, & x \neq 0; \\ A, & x = 0 \end{cases}$$

和函数

$$g(x) = \begin{cases} |x|^\alpha \sin\frac{1}{|x|^\beta}, & x \neq 0; \\ A, & x = 0 \end{cases}$$

的连续性和可微性以及高阶可微性。

8. 证明 $f(x) = \begin{cases} \mathrm{e}^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0. \end{cases}$ 是 $\mathscr{C}^\infty$ 函数，且对所有正整数 n，$f^{(n)}(0) = 0$。

9. 证明 $f : \mathbb{R} \to \mathbb{R}, f(x) = x - \varepsilon \sin x (\ 0 < \varepsilon < 1\ )$的反函数 f^{-1} 是 $\mathscr{C}^\infty$ 函数，并求 $(f^{-1})^{(2)}(0), (f^{-1})^{(3)}(0)$。

10. （常系数线性微分方程和拟多项式）对多项式

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

证明

(a)

$$P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) e^{\lambda x} = \left[a_n \frac{\mathrm{d}^n}{\mathrm{d}x^n} + a_{n-1} \frac{\mathrm{d}^{n-1}}{\mathrm{d}x^{n-1}} + \cdots + a_1 \frac{\mathrm{d}}{\mathrm{d}x} + a_0\right] \mathrm{e}^{\lambda x} = P(\lambda)\mathrm{e}^{\lambda x}.$$

(b)

$$P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\left(\frac{x^m}{m!} e^{\lambda x}\right) = \sum_{k=0}^{m} \frac{P^{(m-k)}(\lambda)}{(m-k)!} \frac{x^k}{k!} \mathrm{e}^{\lambda x}.$$

(c) $y(x) = \mathrm{e}^{\lambda x}$ 是 微分方程 $P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) y = 0$ 的解当且仅当 λ 是多项式 P 的根。

(d) 若 λ 是多项式 P 的 m 重根，则 $\mathrm{e}^{\lambda x}, x\mathrm{e}^{\lambda x}, \ldots, x^{m-1}\mathrm{e}^{\lambda x}$ 都是微分方程 $P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) y = 0$ 的解。

(e) 对任意 λ 和任意正整数 m，记

$$V_m(\lambda) = \{b_0 \mathrm{e}^{\lambda x} + b_1 x \mathrm{e}^{\lambda x} + \cdots + b_{m-1} x^{m-1} \mathrm{e}^{\lambda x}\},$$

则 $P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)$ 是线性空间 $V_m(\lambda)$ 到其自身的一个线性映射，按照基底 $1, x\mathrm{e}^{\lambda x}, \ldots, \frac{x^{m-1}}{(m-1)!}\mathrm{e}^{\lambda x}$ 的顺序，$P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)$ 的表示矩阵为

$$\begin{pmatrix} P(\lambda) & P'(\lambda) & \frac{P''(\lambda)}{2!} & \frac{P^{(3)}(\lambda)}{3!} & \cdots & \frac{P^{(m)}(\lambda)}{m!} \\ 0 & P(\lambda) & P'(\lambda) & \frac{P''(\lambda)}{2!} & \cdots & \frac{P^{(m-1)}(\lambda)}{(m-1)!} \\ 0 & 0 & P(\lambda) & P'(\lambda) & \cdots & \frac{P^{(m-2)}(\lambda)}{(m-2)!} \\ 0 & 0 & 0 & P(\lambda) & \cdots & \frac{P^{(m-3)}(\lambda)}{(m-3)!} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & P(\lambda) \end{pmatrix}$$

(f) 证明：如果 λ 是多项式 P 的 k 重根，则对任意非负整数 m，都有

$$P\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) : V_{m+k}(\lambda) \to V_m(\lambda)$$

是满射，事实上，它是 $x^k V_m(\lambda) \to V_m(\lambda)$ 线性同构。

11. 对 $\mu \in \mathbb{R}$，记 $P_\mu(t) = t - \mu$ 为关于未定元 t 的多项式。于是

$$(P_\mu(t))^n = (t-\mu)^n = \underbrace{(t-\mu)\cdots(t-\mu)}_{n\text{个}}.$$

这里称 t 是未定元，从而 t 可以是一个数，也可以是一个操作。以下取 $t = \frac{\mathrm{d}}{\mathrm{d}x}$，对可微函数 $f(x)$，定义

$$P_\mu\left(\frac{\mathrm{d}}{\mathrm{d}x}\right) f = \left(\frac{\mathrm{d}}{\mathrm{d}x} - \mu\right) f = \frac{\mathrm{d}f(x)}{\mathrm{d}x} - \mu f(x),$$

$$\left(P_\mu\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\right)^n f = \left(\frac{\mathrm{d}}{\mathrm{d}x} - \mu\right)^n f = \underbrace{\left(\frac{\mathrm{d}}{\mathrm{d}x} - \mu\right)\cdots\left(\frac{\mathrm{d}}{\mathrm{d}x} - \mu\right)}_{n\text{个}} f.$$

计算 $P_\mu\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\mathrm{e}^{\lambda x}$，$\left(P_\mu\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\right)^n \left(\frac{x^k}{k!}\mathrm{e}^{\lambda x}\right)$。

12. (1) 设 $f(t)$ 是 $\mathscr{C}^\infty$ 函数，记 $y(x) = f(\ln x)$。证明对任意正整数 n，

$$x^n \frac{\mathrm{d}^n y}{\mathrm{d}x^n} = \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} - (n-1)\right)\cdots\left(\frac{\mathrm{d}}{\mathrm{d}t} - 1\right)\frac{\mathrm{d}}{\mathrm{d}t} f(t)\right]\Bigg|_{t=\ln x}.$$

(2) 形如

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

的微分方程称为 **Euler 方程**。证明经过换元 $x = \mathrm{e}^t$，Euler 方程可以变为常系数线性微分方程

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + b_{n-1}\frac{\mathrm{d}^{(n-1)} y}{\mathrm{d}t^{n-1}} + \cdots + b_1 \frac{\mathrm{d}y}{\mathrm{d}t} + b_0 y = f(\mathrm{e}^t).$$

4.4 应用：曲线的切线、法线和曲率

定义 4.4.1. 设 $(x(t), y(t))$ 满足 $x, y : (a, b) \to \mathbb{R}$ 都是 $\mathscr{C}^1$ 函数，且对任意 $t \in (a, b)$，$x'(t)^2 + y'(t)^2 > 0$。则称 $\gamma : (x(t), y(t))(\ t \in (a, b) \ )$是一条**正则曲线**。称直线

$$\begin{cases} x = x(t_0) + x'(t_0)(t - t_0), \\ y = y(t_0) + y'(t_0)(t - t_0), \end{cases} \quad t \in \mathbb{R}$$

为曲线 γ 在点 $(x(t_0), y(t_0))$ (或 $t = t_0$)处的**切线**。称过点 $(x(t_0), y(t_0))$ 且与曲线 γ 在点 $(x(t_0), y(t_0))$ 处的切线垂直的直线为曲线 γ 在点 $(x(t_0), y(t_0))$ 处的**法线**。

例 4.4.2. 求抛物线 $x = y^2$ 在 (a^2, a) 处的切线与法线。

解： 切线：$\begin{cases} x = a^2 + 2a(t - a), \\ y = a + (t - a) \end{cases}$，即 $x = a^2 + 2a(y - a)$。

法线：$x - a^2 = -\frac{1}{2a}(y - a)$，即 $y = a - 2a(x - a^2)$。 □

例 4.4.3. 证明平行于抛物线对称轴的入射光线经抛物线反射后汇聚于焦点。

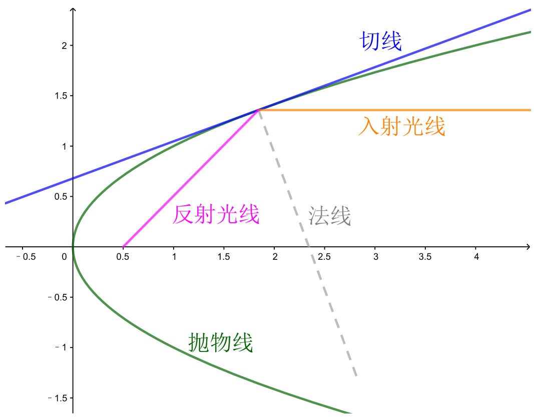


图 4.2:

解： 水平入射光线在点 $P(t, f(t))$ 处经曲线 $y = f(x)$ 反射汇聚于点 $F(c, 0)$ 处，当且仅当直线 FP 的倾角 β 是曲线 $y = f(x)$ 在点 $P(t, f(t))$ 处的切线的倾角 α 的 2 倍，即

$$\frac{f(t)}{t - c} = \tan 2\alpha = \frac{2f'(t)}{1 - f'(t)^2},$$

也就是

$$yy'^2 + 2(x-c)y' - y = 0,$$

解得

$$y' = \frac{-(x-c) \pm \sqrt{(x-c)^2 + y^2}}{y}.$$

记 $u(x) = \sqrt{(x-c)^2 + y(x)^2}$，则上述微分方程等价于 $u' = \pm 1$。

对于 $y(x) = \pm\sqrt{4cx}$，

$$u(x) = \sqrt{(x-c)^2 + y(x)^2} = \pm(x+c),$$

满足 $u' = \pm 1$ 。因此水平光线经抛物线 $4cx = y^2$ 反射后汇聚于焦点 $(c, 0)$。 □

例 4.4.4. 求当 $t \to a$ 时抛物线 $x = y^2$ 在点 (a^2, a) 处的法线与点在点 (t^2, t) 处的法线的交点的极限。

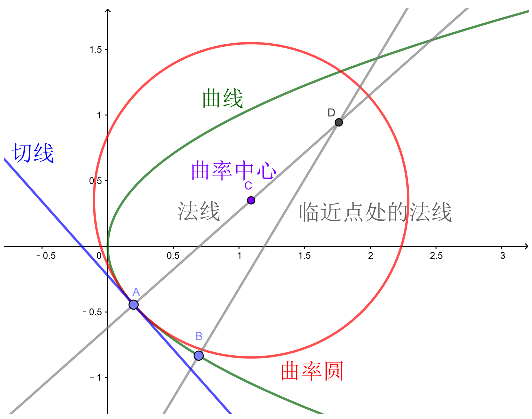


图 4.3:

解： 抛物线在点 (a^2, a) 处的法线：$y = a - 2a(x - a^2)$。在点 (t^2, t) 处的法线：$y = t - 2t(x - t^2)$。它们的交点

$$\begin{cases} x = \dfrac{t + 2t^3 - a - 2a^3}{2(t-a)}, \\ y = a - 2a(x - a^2) \end{cases}$$

让 $t \to a$ 得到

$$\begin{cases} x = \dfrac{1}{2}(1 + 6a^2) = \frac{1}{2} + 3a^2, \\ y = a - 2a\left(\dfrac{1}{2} + 3a^2 - a^2\right) = -4a^3 \end{cases}$$

所以两条法线的交点的极限为 $(\frac{1}{2} + 3a^2, -4a^3)$。 □

定理 4.4.5. 设 $\gamma:(x(t),y(t))$($t\in(a,b)$)是一条正则曲线，满足 $x(t),y(t)$ 都是 $\mathscr{C}^2$ 函数。则当 $t\to t_0$ 时，γ 在点 $(x(t_0),y(t_0))$ 处的法线与点 $(x(t),y(t))$ 处的法线的交点有极限，这个极限 C称为 γ 在点 $(x(t_0),y(t_0))$ 处的**曲率圆中心**，C 到点 $(x(t_0),y(t_0))$ 的距离 R 称为**曲率半径**，称以 C 为圆心以 R 为半径的圆为 γ 在点 $(x(t_0),y(t_0))$ 处的**曲率圆**，称曲率半径的倒数 $\kappa=\frac{1}{R}$ 为 γ 在点 $(x(t_0),y(t_0))$ 处的**曲率**。

例 4.4.6. 求抛物线 $x=y^2$ 的曲率。

解： 由例4.4.4知抛物线在点 (a^2,a) 处的曲率中心为 $(\frac{1}{2}+3a^2,-4a^3)$。

所以曲率半径

$$R=\sqrt{\left(\frac{1}{2}+3a^2-a^2\right)^2+(-4a^3-a)^2}=\sqrt{\frac{1}{4}+3a^2+12a^4+16a^6},$$

曲率

$$\kappa=\frac{1}{R}=\frac{1}{\sqrt{\frac{1}{4}+3a^2+12a^4+16a^6}}.$$

曲率最大值为 2，在 $a=0$ 时即 $(0,0)$ 点处取得。

在 $(0,0)$ 附近，抛物线 $x=y^2$ 上的点到 $(0,0)$ 处法线上的点 $(c,0)$ 的距离

$$\begin{aligned}\sqrt{(x-c)^2+y^2}&=\sqrt{(y^2-c)^2+y^2}=\sqrt{c^2-(2c-1)y^2+y^4}\\&=\begin{cases}|c|\left(1-\frac{2c-1}{2c^2}y^2+o(y^2)\right), & c\neq 0;\\ |y|\left(1+o(1)\right), & c=0.\end{cases}\end{aligned}$$

当且仅当 $c=\frac{1}{2}$ 时，$\sqrt{(x-c)^2+y^2}-c=o(y^2)$。这表明，虽然所有圆 $(x-c)^2+y^2=c^2$ 都与抛物线 $x=y^2$ 在 $(0,0)$ 处相切（它们有共同切线），但唯有圆 $\left(x-\frac{1}{2}\right)^2+y^2=\frac{1}{4}$ 与抛物线才有更高程度的相切关系。 □

习题4.4

1. 证明从椭圆的一个焦点射出的光线经椭圆反射后汇聚于另一焦点。

2. 曲率与曲率圆

   (a) 证明定理4.4.5，并推导曲线的曲率公式。

   (b) 证明在 $\mathscr{C}^2$ 曲线任意一点处，与曲线相切的所有圆中只有曲率圆与曲线是二阶相切的。

   (c) 证明曲线的曲率与曲线的参数表示无关，即若 $t=t(s)$ 是可逆映射，且 $t(s)$ 和它的反函数 $s(t)$ 都是 $\mathscr{C}^2$ 的，则 $(x(t(s)),y(t(s)))$ 在 s_0 处的曲率与 $(x(t),y(t))$ 在 $t_0=t(s_0)$ 处的曲率相等。

3. 称 $\mathbb{R}^n$ 中的 $\mathscr{C}^2$ 曲线 $\gamma(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ 是**正则曲线**，如果

$$\gamma'(t) = (x_1'(t), \ldots, x_n'(t)) \neq 0, \quad \forall t \in (a, b).$$

记

$$T(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|} = \frac{(x_1'(t), x_2'(t), \ldots, x_n'(t))}{\sqrt{x_1'(t)^2 + x_2'(t)^2 + \cdots x_n'(t)^2}}.$$

即 $T(t)$ 为曲线在 $\gamma(t)$ 处的**单位切向量**。证明：

(a) 对任意 $t \in (a, b)$，$T'(t)$ 与 $T(t)$ 正交；

(b) $\dfrac{T'(t)}{\|\gamma'(t)\|}$ 只与曲线方向有关，与曲线参数表示无关，即：如果 $t = t(s)$ 满足 $t'(s) > 0(\ \forall s\ )$，则

$$\frac{T'(t)}{\|\gamma'(t)\|} = \frac{\tilde{T}'(s)}{\|\tilde{\gamma}'(s)\|},$$

如果 $t = t(s)$ 满足 $t'(s) < 0(\ \forall s\ )$，则

$$\frac{T'(t)}{\|\gamma'(t)\|} = -\frac{\tilde{T}'(s)}{\|\tilde{\gamma}'(s)\|},$$

其中 $\tilde{\gamma} = \gamma(t(s))$，$\tilde{T}(s) = \dfrac{\tilde{\gamma}'(s)}{\|\tilde{\gamma}'(s)\|}$。因此 $\dfrac{\|T'(t)\|}{\|\gamma'(t)\|}$ 与曲线参数表示无关，它反映了曲线的一个几何性质。

(c) $n = 2$ 时曲率

$$\kappa = \frac{\|T'(t)\|}{\|\gamma'(t)\|} = \frac{\left\|\gamma''(t) - \langle \gamma''(t), \gamma'(t) \rangle \dfrac{\gamma'(t)}{\|\gamma'(t)\|^2}\right\|}{\|\gamma'(t)\|^2}.$$

4.5 应用：Newton-Raphson 方法

我们以计算 $\sqrt{10}$ 的近似值为例，介绍 Newton 当初的想法。

首先，因为 $3 < \sqrt{10} < 4$，所以可设 $\sqrt{10} = 3 + x$，于是

$$10 = (3 + x)^2 = 9 + 6x + x^2,$$

因为 $0 < x < 1$，所以忽略比 $6x$ 更小的项 x^2，上述方程变成

$$10 = 9 + 6x,$$

解得近似值 $x = \frac{1}{6} \approx 0.17$。于是 $\sqrt{10} \approx 3.17$。

再设 $\sqrt{10} = 3.17 + x$，于是

$$10 = 10.0489 + 6.34x + x^2,$$

舍去 x^2，求得 x 的近似值 $x = -0.00771$，于是$\sqrt{10} \approx 3.17 - 0.00771 = 3.16229$。再次使用上述方法，得到 $\sqrt{10} \approx 3.1622776602$。

Newton 不断求解由近似值导致的新方程从而得到更进一步的近似值，他的这个想法被 Raphson 改写成了如下迭代形式，这个方法被称为 Newton-Raphson 方法，是数值计算求解方程最高效的算法。

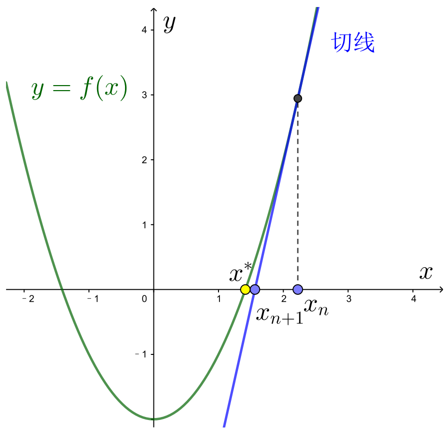


图 4.4:

设 $f : (a, b) \to \mathbb{R}$ 是可微映射。设 x^* 是方程 $f(x) = 0$ 的解。假设 x_n 是 x^* 的一个近似解。在 x_n 处把函数 $f(x)$ 局部线性化：

$$y = f(x_n) + f'(x_n)(x - x_n).$$

由线性方程

$$0 = f(x_n) + f'(x_n)(x - x_n)$$

解得

$$x = x_n - \frac{f(x_n)}{f'(x_n)}.$$

这样得到 Newton 迭代

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

这样就得到了 x^* 的一个新的近似值 x_{n+1} 。

例 4.5.1. 用 Newton-Raphson 方法解方程 $x^2 = a$。

解： 曲线 $y = x^2 - a$ 在 $(x_n, x_n^2 - a)$ 处的切线

$$y = x_n^2 - a + 2x_n(x - x_n) = 2x_n x - x_n^2 - a.$$

解线性方程 $0 = 2x_n x - x_n^2 - a$ 得到

$$x_{n+1} = \frac{x_n^2 + a}{2x_n} = \frac{x_n}{2} + \frac{a}{2x_n}.$$

最后这个公式正是古代巴比伦人求 $\sqrt{a}$ 的方法。

记

$$F(x)=\frac{x}{2}+\frac{a}{2x}.$$

对 $x>0$，

$$F(x)=\frac{x}{2}+\frac{a}{2x}\geq 2\sqrt{\frac{x}{2}\cdot\frac{a}{2x}}=\sqrt{a};$$

对任意 $x,y\in[\sqrt{a},+\infty)$，

$$|F(x)-F(y)|=\left|\frac{x}{2}+\frac{a}{2x}-\frac{y}{2}-\frac{a}{2y}\right|=\left(\frac{1}{2}-\frac{a}{2xy}\right)|x-y|\leq\frac{1}{2}|x-y|,$$

从而 F 是区间 $(0,+\infty)$ 到闭区间$[\sqrt{a},+\infty)$ 中的压缩映射。因此对任意 $x_1>0$，因此 $\lim\limits_{n\to+\infty}x_n$ 存在且为 F 的唯一不动点，易见这不动点为 $\sqrt{a}$。

对 $n\geq 2$，$x_n=F(x_{n-1})\geq\sqrt{a}$，从而

$$\left|x_{n+1}-\sqrt{a}\right|=\left|\frac{x_n-\sqrt{a}}{2x_n}\right|\left|x_n-\sqrt{a}\right|\leq\frac{1}{2\sqrt{a}}\left|x_n-\sqrt{a}\right|^2,$$

所以这个方法是二次收敛的。

表 4.1: 用 Newton-Raphson 方法计算 $\sqrt{10}$ 的近似值

迭代次数 n	x_n	x_n^2-10
0	3.000000000000000	-1.000000000000000
1	3.166666666666670	2.77778×10^{-2}
2	3.162280701754390	1.92367×10^{-5}
3	3.162277660169840	9.24949×10^{-12}
4	3.162277660168380	0

以 $a=10$，$x_1=3$ 为例。

$$|x_1-\sqrt{10}|<\frac{1}{10}.$$

而

$$\left|x_{n+1}-\sqrt{10}\right|\leq\frac{1}{2\sqrt{10}}\left|x_n-\sqrt{10}\right|^2<\frac{2}{10}\left|x_n-\sqrt{10}\right|^2,$$

因此

$$\left|x_n-\sqrt{10}\right|<\frac{2^{2^{n-1}-1}}{10^{2^{n-1}-1}}\left(\frac{1}{10}\right)^{2^{n-1}},\quad n\geq 2.$$

通过上面的表格可以看到如果以 3 为初值计算 $\sqrt{10}$ 的近似值，只需 4 次迭代就可以得到 15 位有效小数位数的近似值了。 □

Newton 法虽然是求解方程的一个好方法（ Newton 法的收敛性我们稍后证明），但 Newton 迭代中每次都需要重新计算导数值 $f'(x_n)$。如果 f 的导数 f'

连续，则当 x_n 接近 x^* 时，$f'(x_n) \approx f'(x^*)$。于是我们可以考虑一个简化了的 Newton 迭代：

$$x_{n+1} = x_n - \frac{f(x_n)}{\lambda}.$$

而

$$f(x_n) = f(x^*) + f'(x^*)(x_n - x^*) + \alpha(x_n) = f'(x^*)(x_n - x^*) + \alpha(x_n),$$

其中 $\alpha(x) = o(x - x^*)$（ $x \to x^*$）。

取 λ 使得

$$0 < \frac{f'(x^*)}{\lambda} < 2.$$

当 x_n 足够靠近 x^* 时，

$$\left|\frac{\alpha(x_n)}{\lambda(x_n - x^*)}\right| < \frac{f'(x^*)}{2\lambda} < 1,$$

从而

$$\begin{aligned}
|x_{n+1} - x^*| &= \left|x_n - x^* - \frac{f'(x^*)(x_n - x^*) + \alpha(x_n)}{\lambda}\right| \\
&\le |x_n - x^*| \left(1 - \frac{f'(x^*)}{\lambda} + \left|\frac{\alpha(x_n)}{\lambda(x_n - x^*)}\right|\right) \\
&\le |x_n - x^*| \left(1 - \frac{f'(x^*)}{2\lambda}\right).
\end{aligned}$$

于是 $\lim\limits_{n\to+\infty} x_n = x^*$。

但是对实际应用而言，这样的结果意义有限，因为算法必须在有限步之内完成，由于我们不知道 x^* 的值，所以无法判定 x_n 是否已经在给定的误差范围内靠近了 x^*。关于 Newton-Raphson 方法的进一步的讨论我们留在下一章里完成。

习题4.5

1. 用 Newton-Raphson 方法求 $f(x) = \cos x$ 的不动点。比较 Newton-Raphson 迭代与 $x_{n+1} = \cos x_n$ 的收敛效率以及不同初值对迭代收敛性的影响。

2. 用 Newton-Raphson 方法给出计算非零实数 a 的倒数 $\frac{1}{a}$ 的算法，使得计算过程中只使用加减法和乘法运算。这个算法可以提高除法的计算效率，再结合Karatsuba(1960)提出的快速乘法，可以使除法效率进一步提高。如果你希望更深入地了解除法算法，建议你阅读英文版维基百科 division algorithm 词条。

3. 在电子游戏《雷神之锤III竞技场》的源代码中有一个“平方根倒数快速算法”（ FastInvSqrt()）用于快速计算 $\frac{1}{\sqrt{x}}$，这在数字信号处理以及计算机

图形学中可以提高运算效率。在这个这个算法中使用了 Newton-Raphson 方法，并且利用了一个“神奇的”初值，使得 Newton 迭代只需一步就可完成。关于这个算法以及它的历史见维基百科 Fast inverse square root 词条。

4.6 应用：用微分方程建立数学模型

例 4.6.1. 放射性物质衰变。某放射性物质的质量 $y(t)$ 关于时间 t 连续，并且随着时间 t 的增加而减少，相同质量的该物质经过相同时间后的剩余质量相同。则

$$\frac{y(t+s)}{y(t)}=\frac{y(s)}{y(0)}.$$

记 $Y(t)=\frac{y(t)}{y(0)}$，则 $Y(t)$ 连续，$Y(0)=1$，且

$$Y(t+s)=Y(t)Y(s).$$

于是

$$\frac{Y(t+s)-Y(t)}{s}=Y(t)\frac{Y(s)-1}{s}=Y(t)\frac{Y(s)-Y(0)}{s}.$$

因此，若 Y 在 $t=0$ 可微，则 Y 在任意 t 可微，且

$$Y'(t)=Y'(0)Y(t).$$

记 $\lambda=-Y'(0)$，则 $\lambda>0$，

$$y'(t)=-\lambda y(t).$$

不难验证 $y(t)=Ce^{-\lambda t}$ 是该方程的解，将来我们会知道这是方程的所有的解。

由 $\frac{y(T)}{y(0)}=\frac{1}{2}$ 解得 $T=\frac{\ln 2}{\lambda}$，这是该放射性物质的**半衰期**。

古生物学家和考古学家通过分析化石样本中某些放射性物质的量估计古生物或文物的年代。 □

例 4.6.2. Malthus 人口模型。假设人口数量 $y(t)$ 随时间 t 变化，人口自然增长率是常数 λ，

$$y'(t)=\lambda y(t).$$

如果 $\lambda>0$，则会出现人口爆炸性增长；如果 $\lambda<0$，则会出现人口指数性衰减；如果 $\lambda=0$，则人口数量将一成不变。所以 Malthus 的人口模型对预测长期的人口变化是不尽合理的。 □

例 4.6.3. 修正的人口模型，Logistic 模型。假设人口数量 $y(t)$ 随时间 t 变化，人口自然增长率 λ 与当前人口数量 $y(t)$ 有关，

$$\lambda=a-by(t),$$

当 $y(t)$ 较小时，$\lambda \approx a > 0$，人口近似指数增长；但当 $y(t)$ 较大时，由于受到环境因素制约，人口增长率降低。

$$y'(t) = (a - by(t))y(t).$$

记 $Y(t) = \dfrac{b}{a}y(t)$，则

$$Y'(t) = a(1 - Y(t))Y(t).$$

最后这个方程称为 Logistic 方程。 □

例 4.6.4. 捕猎模型。在Logistic模型下考虑捕猎因素，动物数量为 $x(t)$ 满足

$$x' = (a - bx)x - c,$$

其中 $c > 0$ 是单位时间的捕获量。 □

例 4.6.5. 捕食者-饵模型，Lotka-Volterra 模型。两种生物，其中一种生物(称为捕食者)捕食另一种生物(称为饵)，捕食者的数量为 $x(t)$，饵的数量为 $y(t)$。

两种生物各自有自然增长率，但实际增长率也受彼此数量的影响：

捕食者的增长率随饵的数量变大而变大，当 $y(t)$ 比较大时，饵充足，捕食者数量正增长；当 $y(t)$ 比较小时，因为食物稀缺，捕食者数量的增长率降低，甚至可能出现数量负增长。

$$x'(t) = (ay(t) - b)x(t).$$

饵的增长率随捕食者数量的变大而变小，当 $x(t)$ 比较大时，因天敌数量大，饵数量出现负增长；当 $x(t)$ 比较小时，因为缺少天敌，饵的数量是正增长。

$$y'(t) = (p - qx(t))y(t).$$

这两个方程所组成的微分方程组称为 Lotka-Volterra 模型：

$$\begin{cases} x' = (ay - b)x, \\ y' = (p - qx)y. \end{cases}$$

□

例 4.6.6. 流感模型。这是一个传染病模型，适用于普通感冒和流感这样的疾病，其特点是人群分为易感人群和感染者两类，数量分别是 $S(t), I(t)$，感染者不会免疫。

$$\begin{cases} S' = -\frac{\beta SI}{N} + \gamma I, \\ I' = \frac{\beta SI}{N} - \gamma I, \end{cases}$$

其中 $N = S(t) + I(t)$ 是人口总数，β 是感染率，γ 是康复系数。$\frac{SI}{N}$ 的分子反映了易感人群与感染者的接触程度，分母 N 是为了平衡系数 β 的量纲。不难发现这个微分方程组是 Lotka-Volterra 模型的一个特例。

利用 $N = S(t) + I(t)$ ，可以得到

$$I' = \left(\beta - \gamma - \frac{\beta}{N} I\right) I,$$

这是 Logistic 模型。 □

以上这些问题最终都归结为一个（或多个）函数和它（它们的）的导数满足的等式关系，这样的等式称为**微分方程（组）**。在上述问题中，微分方程中出现的导数都是一阶导数，这样的微分方程（组）叫做**一阶微分方程（组）**。

在动力学中，Newton第二定律断言运动力的作用而改变，具体而言，运动的加速度与力成正比。加速度是位移函数的二阶导数，所以动力学方程通常是**二阶微分方程**。

一个微分方程（组）的**阶**，就是这个方程（组）中出现的未知函数的最高阶导数的阶数。

例 4.6.7. 弹簧振子-胡克定律。胡克定律：弹簧振子偏离其平衡位置时受到弹簧的作用力 F，F 的大小与偏离平衡位置的位移 x 的大小成比例，F 的方向位移 x 的方向相反。于是根据 Newton 定律和胡克定律得到微分方程：

$$mx'' = F = -kx.$$

这是个二阶微分方程，如果令 $y = x'$，则上述方程可以写成

$$\begin{cases} x' = y, \\ y' = -\frac{k}{m}x. \end{cases}$$

□

例 4.6.8. 带有阻尼的弹簧振子。弹簧振子受到阻力作用，阻力的大小与速度 x' 的大小成比例，方向相反。

$$mx'' = -kx - ax'.$$

这是个二阶微分方程，如果令 $y = x'$，则上述方程可以写成

$$\begin{cases} x' = y, \\ y' = -\frac{k}{m}x - \frac{a}{m}y. \end{cases}$$

□

例 4.6.9. 单摆。一根细杆长度为 l，质量忽略不计，其一端固定，另一端有一个质量为 m 的小球。在小球所受重力作用下，细杆绕着其固定端在竖直平面中摆动，细杆垂直向下为平衡位置，θ 为细杆偏离垂直向下的角度。由 Newton 运动定律可得

$$ml\theta'' = -mg\sin\theta.$$

这是个二阶微分方程，如果令 $y = x'$，则上述方程可以写成

$$\begin{cases} \theta' = y, \\ y' = -\frac{g}{l}\sin\theta. \end{cases}$$

□

虽然我们现在还不知道如何求解这些微分方程，但是我们可以通过几何的办法直观地了解微分方程，从而预见到解的一些性质。

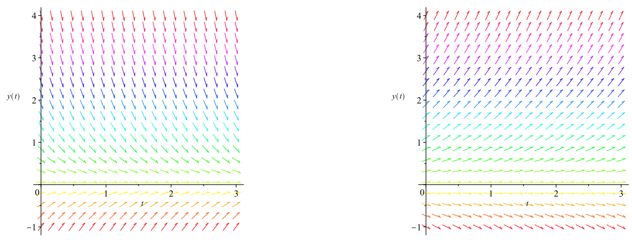


图 4.5: 放射性衰变 / Malthus人口模型

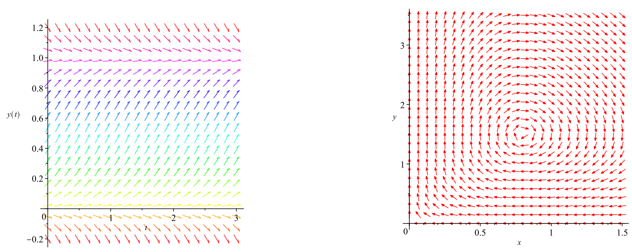


图 4.6: Logistic人口模型 / Lotka-Volterra模型

对微分方程
$$\frac{\mathrm{d}y}{\mathrm{d}t}=f(t,y),$$
我们可以在 (t,y)坐标系的平面中的每个点 (t,y) 处画一条斜率为 $f(t,y)$ 的短直线段，我们称 $f(t,y)$ 是 (t,y)坐标平面上的一个**斜率场**或**方向场**。如果 $y(t)$ 是上述微分方程的**解**(即 $y'(t)=f(t,y(t))$)，则曲线 $y=y(t)$ 在其所经过的每个点处都与方向场 $f(t,y)$ 相切，称曲线 $y=y(t)$ 是方向场 $f(t,y)$ 的一条**积分曲线**。

对微分方程组
$$\begin{cases}x'=f(x,y),\\ y'=g(x,y).\end{cases}$$
我们可以在 (x,y) 坐标系的平面中的每个点 (x,y) 处画一个向量 $(f(x,y),g(x,y))$，我们称 $(f(x,y),g(x,y))$ 是 (x,y) 坐标平面上的一个**向量场**。如果 $(x(t),y(t))$ 是上述微分方程的**解**，则曲线 $y=y(t)$ 在其所经过的每个点处都与方向场 $f(t,y)$ 相切，称曲线 $x=x(t),y=y(t)$ 是向量场 $(f(x,y),g(x,y))$ 的一条**积分曲线**。

从上面这些方程所确定的方向场或向量场的图像可以看出，当我们画出足够密集的方向线段或向量时，我们不难从中发现相应的积分曲线的形状，从而了解微分方程解的一些性质。

习题4.6

1. 设连续函数 $Y:\mathbb{R}\to\mathbb{R}$ 满足 $Y(0)=1$，$Y(t+s)=Y(t)Y(s)(\forall t,s\in\mathbb{R})$。证明 $Y(t)=a^t(\forall t\in\mathbb{R})$，其中 $a=Y(1)$。

2. 四只小狗逆时针排列在一个正方形的四个顶点处，它们同时以大小相同且恒定不变的速度追向位于自己前方的那只小狗。建立适当的坐标系，并写出每只小狗的位置坐标 $(x(t),y(t))$ 所满足的微分方程。

3. 一只兔子沿 (x,y)坐标平面中的 y 轴奔跑，速度为 v_h。当兔子经过坐标原点时，被一只位于 x 轴上点 $(a,0)$ 处的狼所发现。狼以速度 v_w 朝兔子追去。写出狼的运动轨迹 $(x(t),y(t))$ 所满足的微分方程。

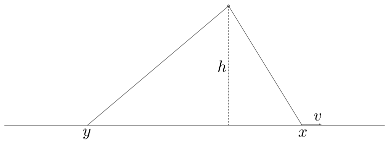


图 4.7:

习题讨论课3

1. 求 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 和 $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$，其中
   (a) $y=\sin^6 x+\cos^6 x$，
   (b) $y=\frac{x^2-3}{(x-1)^2(x^2+1)}$，
   (c) $y=\left(1+\frac{1}{x}\right)^x$，
   (d) $y=g(h(x))$，其中 g,h 是二阶可微函数，
   (e) $x=g(y)$，其中 g 是二阶可微的严格增函数，
   (f) $x=x(t),y=y(t)$ 是二阶可微函数，$x(t)$ 是严格增函数。
2. 称曲线 $y=f(x)$ 与 $y=g(x)$ 在 $x=a$ 处相切，如果 $f(a)=g(a)$ 且 $f'(a)=g'(a)$。证明：若曲线 $y=f(x)$ 与 $y=g(x)$ 在 $x=a$ 处，$x=x(t)$，$z=z(y)$ 是可微函数，$x(t_0)=a$，则曲线 $z=z(f(x(t)))$ 与 $z=z(g(x(t)))$ 在 $t=t_0$ 处相切。
3. 设 $m\geq 2$ 是正整数，$a>0$。写出求解方程 $x^m=a$ 的 Newton 迭代，并讨论这个算法的收敛性。
4. 求曲线 $y=\frac{1}{x}$ 的曲率，并证明 $y=\frac{1}{x}$ 和它的曲率圆二阶相切。
5. 如图??所示，两个铁块由长度为 L 的绳子相连，悬挂在距地面高度为 h 处的一个定滑轮处，当右侧的铁块位于 x 处以速度 v 向右移动时，求左侧的铁块的移动速度。
6. 证明 Legendre 多项式 $P(x)=\frac{\{(x^2-1)^n\}^{(n)}}{2^n n!}$ 满足
$$(1-x^2)P''(x)-2xP'(x)+n(n+1)P(x)=0.$$
7. 四只小狗逆时针排列在一个正方形的四个顶点处，它们同时以大小相同且恒定不变的速度追向位于自己前方的那只小狗。建立适当的坐标系，并写出每只小狗的位置坐标 $(x(t),y(t))$ 所满足的微分方程。
8. 一只兔子沿 (x,y)坐标平面中的 y 轴奔跑，速度为 v_h。当兔子经过坐标原点时，被一只位于 x 轴上点 $(a,0)$ 处的狼所发现。狼以速度 v_w朝兔子追去。写出狼的运动轨迹 $(x(t),y(t))$ 所满足的微分方程。

第 5 章　用导数研究函数性质

5.1　微分中值定理

函数在一点处的最佳线性逼近由其微分给出，函数在该点处的导数是微分中的比例系数。然而，这个线性函数究竟在多大程度上反映了原来函数的性质？由一点处的导数究竟能得到函数哪些性质？若 $\alpha < f'(x_0) < \beta$，则存在 $\delta > 0$ 使得

$$\alpha < \frac{f(x)-f(x_0)}{x-x_0} < \beta, \quad \forall 0 < |x-x_0| < \delta,$$

从而

$$\begin{aligned} &\forall x \in (x_0, x_0+\delta), \quad f(x_0)+\alpha(x-x_0) < f(x) < f(x_0)+\beta(x-x_0), \\ &\forall x \in (x_0-\delta, x_0), \quad f(x_0)+\beta(x-x_0) < f(x) < f(x_0)+\alpha(x-x_0). \end{aligned}$$

即任给一个包含切线 $y = f(x_0)+f'(x_0)(x-x_0)$ 在内的锥形区域，f 在 x_0 附近的函数图像总在这个锥形区域中。

特别地，如果 $f'(x_0) > 0$，即 $\alpha = 0$，则在 x_0 的一个邻域内，

$$x < x_0 \Rightarrow f(x) < f(x_0); \qquad x > x_0 \Rightarrow f(x) > f(x_0).$$

因此当 $f'(x_0)$ 非零时，微分 $\mathrm{d}f(x_0)$ 的单调性决定了 x_0 附近点的函数值与 $f(x_0)$ 的大小关系。但是以下例子表明，仅由 x_0 处微分 $\mathrm{d}f(x_0)$ 的单调性无法确保 f 在 x_0 的任意邻域内具有与 $\mathrm{d}f(x_0)$ 同样的单调性。

例 5.1.1. $f(x) = \begin{cases} x + Ax^2 \sin\frac{1}{x}, & x \neq 0, \\ 0, & x = 0 \end{cases}$，其中 $A > \frac{\pi}{2}$。则 $f'(0) = 1$，但在 $x = 0$ 的任何邻域内 f 都不是单调的。

证明. 由 $f'(0) = 1 > 0$ 可知在 $x = 0$ 的任何邻域内 f 都不是单调不增的。另一方面，取 $x_k = \left(2k\pi + \frac{\pi}{2}\right)^{-1}$，$y_k = \left(2k\pi - \frac{\pi}{2}\right)^{-1}$，则对任意 $k \geq 1$，$0 < x_k < y_k < \frac{1}{k}$，

$$f(x_k)-f(y_k) = A\left(2k\pi+\frac{\pi}{2}\right)^{-2} + A\left(2k\pi-\frac{\pi}{2}\right)^{-2} - \pi\left(2k\pi+\frac{\pi}{2}\right)^{-1}\left(2k\pi-\frac{\pi}{2}\right)^{-1},$$

注意到
$$\lim_{k\to+\infty}\frac{A\left(2k\pi+\frac{\pi}{2}\right)^{-2}+A\left(2k\pi-\frac{\pi}{2}\right)^{-2}}{\pi\left(2k\pi+\frac{\pi}{2}\right)^{-1}\left(2k\pi-\frac{\pi}{2}\right)^{-1}}=\frac{2A}{\pi}>1,$$
故存在 $K\in\mathbb{N}^*$ 使得 $\forall k\geq K$，$f(x_k)>f(y_k)$。故在 $x=0$ 的任何邻域内 f 都不是单调不减的。 □

上述例子还表明：f 在一点 x_0 处导数非零（即 $\mathrm{d}f(x_0)$ 是单射）不足以保证 f 在 x_0 的某个邻域内是单射。

定义 5.1.2. 设函数 f 在 x_0 的一个邻域 U 内有定义，且 $\forall x\in U$，$f(x_0)\geq f(x)$。则称 x_0 为 f 的一个**极大值点**，称 $f(x_0)$ 为 f 的一个**极大值**。类似定义 f 的**极小值点**和**极小值**。f 的极大值（点）和极小值（点）统称为 f 的**极值（点）**。

注 5.1.3. 若 f 在 x_0 的邻域内有定义，并且在 x_0 达到最大（小）值，则 x_0 是 f 的一个极大（小）值点。

定义 5.1.4. 若 $f'(x_0)=0$，则称 x_0 为 f 的一个**临界点**。

定理 5.1.5 (Fermat 引理). *若 f 在极值点 x_0 处可微，则 x_0 是 f 的临界点。*

定理 5.1.6 (Rolle 定理). *设函数 f 在 (a,b) 内可微($-\infty\leq a<b\leq+\infty$)，并且*
$$\lim_{x\to a^+}f(x)=\lim_{x\to b^-}f(x)=A\in\mathbb{R}\bigcup\{-\infty,+\infty\},$$
则存在 $\xi\in(a,b)$ 使得 $f'(\xi)=0$。

证明. 函数 f 在 (a,b) 内可微，从而在 (a,b) 内连续。如果 f 是常值函数，则结论自然成立。下设 f 不是常值函数，不妨设存在 $c\in(a,b)$ 使得 $f(c)>A$。于是存在 a_1,b_1 使得 $a<a_1<c<b_1<b$，并且 $\forall x\in(a,a_1]\bigcup[b_1,b)$，$f(x)<f(c)$。$f$ 连续，在有界闭区间 $[a_1,b_1]$ 上有最大值点 ξ，从而 $\xi\in(a,b)$ 是 f 在区间 (a,b) 中的最大值点，由 Fermat 引理知 $f'(\xi)=0$。 □

注 5.1.7. 若 a,b 是有限数，则上述定理结论可写成：存在 $\theta\in(0,1)$ 使得 $f'(\theta a+(1-\theta)b)=0$。

定理 5.1.8 (Cauchy 中值定理). *设 $-\infty\leq a<b\leq+\infty$，函数 $x(t),y(t)$ 在区间 (a,b) 内可微，满足极限 $x(a^+),x(b^-),y(a^+),y(b^-)$ 都收敛，则存在 $\xi\in(a,b)$ 使得*
$$x'(\xi)\left(y(b^-)-y(a^+)\right)=y'(\xi)\left(x(b^-)-x(a^+)\right).$$

Cauchy 中值定理的几何解释：平面曲线 $\Gamma:(x(t),y(t))$ 在某点 $P(x(\xi),y(\xi))$ 处的切向量 $(x'(\xi),y'(\xi))$ 与点 $A(x(a^+),y(a^+))$ 和点 $B(x(b^-),y(b^-))$ 所连的直线 l_{AB} 平行。

证明. 记

$$f(t)=\left(x(t)-x(a^+)\right)\left(y(b^-)-y(a^+)\right)-\left(x(b^-)-x(a^+)\right)\left(y(t)-y(a^+)\right),$$

则 f 在区间 (a,b) 上可微，并且 $f(a^+)=0=f(b^-)$。因此由Rolle定理，存在 $\xi\in(a,b)$ 使得 $f'(\xi)=0$，于是

$$x'(\xi)\left(y(b^-)-y(a^+)\right)=y'(\xi)\left(x(b^-)-x(a^+)\right).$$

□

注 5.1.9. 当 a,b 是有限实数时，条件“$g(a^+),g(b^-)\in\mathbb{R}$，$f(a^+),f(b^-)\in\mathbb{R}$”可以用“$f,g$ 在 $[a,b]$ 上连续”代替。

下面这个例子表明， Cauchy 微分中值定理对二维以上空间中的曲线是不成立的!

例 5.1.10. 三维空间中的光滑曲线，$x(t)=t,y(t)=\cos t,z(t)=\sin t,0\le t\le 2\pi$。如图??所示，这是一条缠绕圆柱面的曲线。

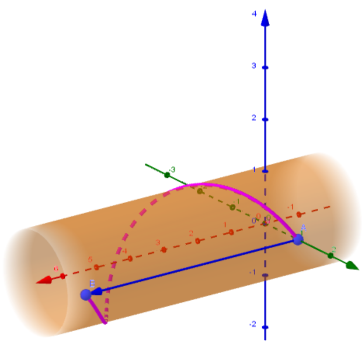


图 5.1:

$$(x(2\pi)-x(0),y(2\pi)-y(0),z(2\pi)-z(0))=(2\pi,0,0),$$

但 $(x'(t),y'(t),z'(t))=(1,-\sin t,\cos t)$ 从不与 $(2\pi,0,0)$ 平行。 □

例 5.1.11. 考虑由光滑的参数方程 $x(t)=t^3|t|$，$y(t)=t^2$，$-1\le t\le 1$刻画的平面曲线，该曲线的图形如图??所示。

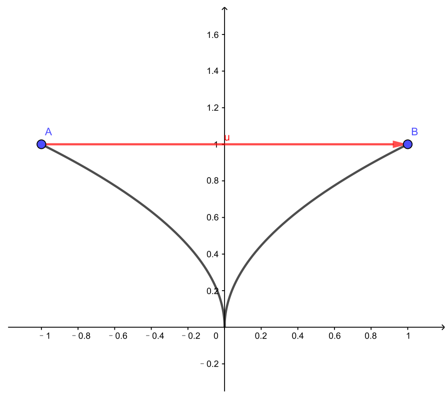


图 5.2:

$$(x(1) - x(-1), y(1) - y(-1)) = (2, 0).$$

曲线上哪里的切线与 $(2,0)$ 平行？这违背 Cauchy 微分中值定理吗？

实际上， Cauchy 微分中值定理断言，如果质点运动是可微的(即任何时刻都存在有限的瞬时速度)，那么在任何一个时段中，总有那样的时刻，质点当时的瞬时速度与质点在该段时间内发生的位移向量平行。在上面这个例子中，$(x'(0), y'(0)) = \mathbf{0}$。 □

定理 5.1.12 (Lagrange 中值定理). *设函数 f 在有界闭区间 $[a,b]$ 上连续，在开区间 (a,b) 内可微，则存在 $\xi \in (a,b)$ 使得 $f'(\xi) = \dfrac{f(b)-f(a)}{b-a}$。*

这实际上是 Cauchy 微分中值定理的特例。但下面给出一个单独的证明。

证明. 对函数 $F(x) = f(x) - f(a) - \dfrac{f(b)-f(a)}{b-a}(x-a)$ 用 Rolle 定理。 □

上述定理结论可写成：$\exists \theta \in (0,1)$ 使得 $f(b) = f(a) + f'(\theta a + (1-\theta)b)(b-a)$。

例 5.1.13. 设函数 $f : I \to \mathbb{R}$ 是可微函数，并且存在 $M > 0$ 使得 $\forall x \in I$，$|f'(x)| \le M$，则对任意 $x, y \in I$，$|f(x) - f(y)| \le M|x-y|$。因此导数有界的函数都是 Lipschitz 函数，从而是一致连续的。

证明. 由 Lagrange 中值定理，对任意 $x, y \in I$，存在介于 x, y 之间的 ξ (从而 $\xi \in I$)使得 $f(x) - f(y) = f'(\xi)(x-y)$，从而

$$|f(x) - f(y)| = |f'(\xi)|\,|x-y| \le M\,|x-y|\,.$$

□

由上述结论知，对一切 $x, y \in \mathbb{R}$，$|\sin x - \sin y| \le |x-y|$。

推论 5.1.14. *在区间 I 上，f 为常数当且仅当 $f'=0$。*

证明. 充分性：上例中取 $M=0$。必要性：用导数定义。 □

例 5.1.15. 求微分方程 $y'=\alpha y$ 的所有的解。

解： 不难发现 $y(x)=\mathrm{e}^{\lambda x}$ 是上述微分方程的解当且仅当 $\lambda=\alpha$。由于上述微分方程是线性方程，所以对任意常数 $C\in\mathbb{R}$，$C\mathrm{e}^{\alpha x}$ 也是该方程的解。

为了求上述微分方程的所有的解，我们使用**常数变易法**：任何函数可以表示为 $y(x)=C(x)\mathrm{e}^{\alpha x}$ 的形式。于是 $y'(x)=\mathrm{e}^{\alpha x}(C'(x)+\alpha C(x))$，于是 $y(x)=C(x)\mathrm{e}^{\alpha x}$ 是原微分方程的解当且仅当 $C'(x)\mathrm{e}^{\alpha x}=0$，即 $C'(x)=0$，即 $C(x)$ 为常数。

所以 $C\mathrm{e}^{\alpha x}$ (C 是任意常数)是原微分方程的所有的解。 □

例 5.1.16. 求微分方程 $y''-3y'+2y=0$ 的所有的解。

解： $y(x)=\mathrm{e}^{\lambda x}$ 是上述微分方程的一个解当且仅当 $\mathrm{e}^{\lambda x}(\lambda^2-3\lambda+2)=0$。$\lambda^2-3\lambda+2$ 称为该微分方程的**特征多项式**，它的零点 $\lambda=1$ 和 $\lambda=2$ 称为该微分方程的**特征指数**。因此 $y(x)=\mathrm{e}^{2x}$ 是原微分方程的一个非零解。

为了求上述微分方程的所有的解，我们使用**常数变易法**：任何函数可以表示为 $y(x)=C(x)\mathrm{e}^{2x}$ 的形式。

$y'(x)=\mathrm{e}^{2x}[C'(x)+2C(x)]$，$y''(x)=\mathrm{e}^{2x}[C''(x)+4C'(x)+4C(x)]$，于是 $y(x)=C(x)\mathrm{e}^{2x}$ 是原微分方程的解当且仅当 $C''(x)+C'(x)=0$。后者是关于 C' 的一个一阶微分方程——常数变易法使得方程的阶数降低了。从这个一阶方程解得 $C'(x)=A\mathrm{e}^{-x}$，于是 $[C(x)+A\mathrm{e}^{-x}]'=0$，因此 $C(x)=-A\mathrm{e}^{-x}+B$，其中 A,B 为任意常数。

所以 $y(x)=C_1\mathrm{e}^{x}+C_2\mathrm{e}^{2x}$ (C_1,C_2 是任意常数)是原微分方程所有的解。 □

例 5.1.17. 证明对区间 $[\sqrt{2}-1,\sqrt{2}+1]$ 中的任何有理数 $\frac{p}{q}$，都有

$$\left|\frac{p}{q}-\sqrt{2}\right|\geq\frac{1}{2(1+\sqrt{2})q^2}.$$

证明. 考虑 $P(x)=x^2-2$。则 $P'(x)=2x$，

$$\left|\frac{P\left(\frac{p}{q}\right)-P(\sqrt{2})}{\frac{p}{q}-\sqrt{2}}\right|=|P'(\xi)|=2\xi\leq 2(1+\sqrt{2}).$$

从而

$$\left|\frac{p}{q}-\sqrt{2}\right|\geq\frac{1}{2(1+\sqrt{2})}\left|P\left(\frac{p}{q}\right)\right|=\frac{1}{2(1+\sqrt{2})}\frac{|p^2-2q^2|}{q^2}\geq\frac{1}{2(1+\sqrt{2})q^2},$$

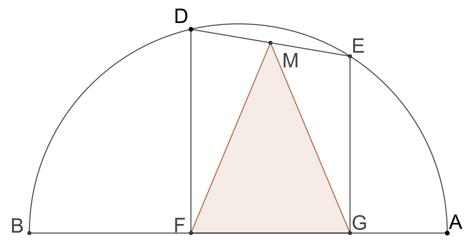
最后这个不等式成立是因为 $|p^2-2q^2|$ 是一个正整数。 □

习题5.1

1. 说明交管部门依据高速公路上区间测速结果判定机动车超速行驶的理论依据。

2. (Darboux定理)设函数 f 在有界区间 (a,b) 内可导，且 $f'_+(a)$ 和 $f'_-(b)$ 存在，且 $f'_+(a) < f'_-(b)$。则对任何 $c \in (f'_+(a), f'_-(b))$，存在 $\xi \in (a,b)$ 使得 $f'(\xi) = c$。由此可知 f 的导函数 f' 的间断点只可能是第二类间断的。

3. 如图，在以线段 AB 为直径的半圆周上，DE 是沿这个半圆周移动的一条弦，其长度是定值。M 是这条弦的中点，F, G 是 D, E 在直径 AB 上的垂直投影。 证明三角形 MFG 是一个形状固定的等腰三角形。（提示：只需



证明 $\angle MFG$ 是定值）

4. 线性微分方程与常数变易法
(1) $\mathrm{e}^{\lambda x}$ 是 n 阶常系数线性常微分方程

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$$

的解当且仅当 λ 是多项式 $\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ 的根。这个多项式称为该微分方程的**特征多项式**，它的根称为该微分方程的**特征值**或**特征指数**。
(2) [**常数变易法**] 设 λ 是微分方程

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$$

的特征值，证明 $y(x) = C(x)\mathrm{e}^{\lambda x}$ 是微分方程

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = f(x)$$

的解当且仅当 $C'(x)$ 是一个 $n-1$ 阶常系数线性微分方程的解，并求这个 $n-1$ 阶常系数线性微分方程。
(3) 证明：$y(x)$ 是微分方程 $y' - \lambda y = 0$ 的解当且仅当存在常数 C 使得 $y(x) = C\mathrm{e}^{\lambda x}$。

(4) 求微分方程 $y''-3y'+2y=0$ 的所有的解。一般地，若 $a^2>4b$，求微分方程 $y''+ay'+by=0$ 的所有的解。
(5) 求微分方程 $y''-2ay'+a^2y=0$ 的所有的解。

5. 设函数 f 在区间 $[a,b]$ 上连续，在开区间 (a,b) 中可微，$f(a)=f(b)=0$。证明对任意实数 λ，存在 $\xi\in(a,b)$ 使得 $f'(\xi)+\lambda f(\xi)=0$。

6. 设函数 f 满足 $f'(x)\le f(x)$，$f(0)=0$。证明对所有实数 x，$xf(x)\le 0$。

7. (1) 设 $y(x)$ 是微分方程 $\dfrac{\mathrm{d}^2y}{\mathrm{d}x^2}+\dfrac{\mathrm{d}F(y)}{\mathrm{d}y}=0$ 的解。证明 $\dfrac{1}{2}\left(\dfrac{\mathrm{d}y}{\mathrm{d}x}\right)^2+F(y)$ 为常数。
(2) 证明对任意 $A,B\in\mathbb{R}$，初值问题 $\begin{cases} y''+y=0\\ y(x_0)=A,\\ y'(x_0)=B\end{cases}$ 有唯一解。

8. 设 $\lambda\in\mathbb{R}$。讨论方程 $(1+x^2)\arctan x=\lambda x$ 的解的个数。

9. (Liouville，1851) 设无理数 x_0 是一个 n 阶代数数，即 x_0 是一个 n 次整系数多项式 $P(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 的根，但不是任何次数低于 n 的多项式的根。则存在常数 $A>0$ 使得对区间 $[x_0-1,x_0+1]$ 中的任意有理数 $\frac{p}{q}$，都有
$$\left|\frac{p}{q}-x_0\right|\ge\frac{1}{Aq^n}.$$

10. (Lagrange 插值多项式) 设函数 f 在区间 $[a,b]$ 上有定义，并且满足 $|f^{n+1}(x)|\le M(\forall x\in[a,b])$。设
$$a\le x_0<x_1<\cdots<x_n\le b,$$
f 的 n 次插值多项式为
$$P_n(x)=\sum_{k=0}^{n}\left(f(x_k)\prod_{0\le j\le n,j\ne k}\frac{x-x_j}{x_k-x_j}\right).$$
考虑
$$F(t):=f(t)-P_n(t)-\frac{(t-x_0)(t-x_1)\cdots(t-x_n)}{(x-x_0)(x-x_1)\cdots(x-x_n)}[f(x)-P_n(x)].$$
(a) 证明
$$|f(x)-P_n(x)|\le\frac{(b-a)^{n+1}M}{(n+1)!},\quad \forall x\in[a,b].$$
(b) 如果 $n=2$ 且 $x_1=x_0+h$，$x_2=x_0+2h$，证明
$$|f(x)-P_2(x)|\le\frac{h^3M}{9\sqrt{3}},\quad \forall x\in[a,b].$$

5.2 函数的单调性与极值

定理 5.2.1. 设 f 在区间 I 上可微。
(1) 如果对任何 $x \in I$，$f'(x) > 0$($f'(x) < 0$)，则 f 在区间 I 上严格增(相应地，严格减)；
(2) f 在区间 I 上单调不减(相应地，单调不增)当且仅当对任何 $x \in I$，$f'(x) \geq 0$($f'(x) \leq 0$)。

证明. 对任何 $x_1, x_2 \in I$，存在 $\xi \in I$ 使得 $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$。 □

例 5.2.2. 讨论函数 $f(x) = \left(1 + \dfrac{1}{x}\right)^x$ 在区间 $(0, +\infty)$ 上的单调性。

解： 因为 ln 是严格增函数，所以 f 与 $g(x) = \ln f(x) = x \ln\left(1 + \frac{1}{x}\right)$具有相同的单调性。

$$g'(x) = \ln(x+1) - \ln x + x\left(\frac{1}{x+1} - \frac{1}{x}\right).$$

$$g''(x) = 2\left(\frac{1}{x+1} - \frac{1}{x}\right) + x\left(\frac{-1}{(x+1)^2} + \frac{1}{x^2}\right) = \frac{-1}{(x+1)^2 x} < 0, \qquad \forall x > 0.$$

所以 g' 是减函数。因为

$$\lim_{x \to +\infty} g'(x) = \lim_{x \to +\infty}\left[\ln\left(1 + \frac{1}{x}\right) - \frac{1}{x+1}\right] = 0,$$

所以对任意 $x > 0$，$g'(x) > 0$，因此 g 在区间 $(0, +\infty)$ 上是增函数。从而 $f(x) = \left(1 + \dfrac{1}{x}\right)^x$ 在区间 $(0, +\infty)$ 上是增函数。 □

定理 5.2.3. 设存在 $\delta > 0$ 使得 f 在 $(x_0 - \delta, x_0 + \delta)$ 上连续。

1. 若对 $x \in (x_0 - \delta, x_0)$，$f'(x) \leq 0$；对 $x \in (x_0, x_0 + \delta)$，$f'(x) \geq 0$，则 x_0 是 f 在区间 $(x_0 - \delta, x_0 + \delta)$ 中的最小值点。

2. 若对 $x \in (x_0 - \delta, x_0)$，$f'(x) \geq 0$；对 $x \in (x_0, x_0 + \delta)$，$f'(x) \leq 0$，则 x_0 是 f 在区间 $(x_0 - \delta, x_0 + \delta)$ 中的最大值点。

3. 若 $f'(x_0) = 0$ 且 $f''(x_0) > 0$，则 x_0 是 f 的极小值点。

4. 若 $f'(x_0) = 0$ 且 $f''(x_0) < 0$，则 x_0 是 f 的极大值点。

5. 若 $f'(x_0) = f''(x_0) = \cdots = f^{(2n-1)}(x_0) = 0$，$f^{(2n)}(x_0) > 0$（ $f^{(2n)}(x_0) > 0$ ），则 x_0 是 f 的极小（大）值点。

6. 若 $f'(x_0) = f''(x_0) = \cdots = f^{(2n)}(x_0) = 0$，$f^{(2n+1)}(x_0) > 0$（ $f^{(2n+1)}(x_0) > 0$ ），则 f 在 x_0 的一个邻域内单调增（单调减）。

证明. (1)和(2)是定理5.2.1的推论。对 $-f$用(3)得到(4)。下面证明(3)。

由 $f'(x_0)=0$ 且 $f''(x_0)>0$知存在 $\delta_1>0$ 使得当 $x\in(x_0-\delta_1,x_0)$ 时 $f'(x)<f'(x_0)=0$，当 $x\in(x_0,x_0+\delta_1)$ 时 $f'(x)>f'(x_0)=0$，所以由(1)知 x_0 是 f 的极小值点。

(5)和(6)的证明留作练习。 □

例 5.2.4. 海滩上的救生员看到海中有人溺水。救生员在沙滩上的奔跑速度为 v_1，在水中的游泳速度为 v_2。求救生员的最佳救援路线。

解： 假设海岸线为直线。建立坐标系使海岸线为 x 轴，救生员起始位置在点 $A(-c,a)$，溺水者位于点 $B(c,b)$($c>0$)。设救生员入水位置在点 $(x,0)$。则救生员用时为

$$T(x)=\frac{\sqrt{(x+c)^2+(0-a)^2}}{v_1}+\frac{\sqrt{(x-c)^2+(0-b)^2}}{v_2}.$$

易见T 连续，且 $\lim\limits_{x\to\infty}T(x)=+\infty$，所以 T 有最小值。

由

$$T'(x)=\frac{x+c}{v_1\sqrt{(x+c)^2+(0-a)^2}}+\frac{x-c}{v_2\sqrt{(x-c)^2+(0-b)^2}}$$

知当 $x<-c$ 时 $T'(x)<0$，当 $x>c$ 时 $T'(x)>0$。所以 T 的最小值点 $x_0\in[-c,c]$。且由 $T'(x_0)=0$ 知

$$\frac{x_0+c}{v_1\sqrt{(x_0+c)^2+a^2}}=\frac{c-x_0}{v_2\sqrt{(x_0-c)^2+b^2}}.$$

这表明

$$\frac{\sin\theta_1}{v_1}=\frac{\sin\theta_2}{v_2}.$$

在光学中，这是折射定律，其中 θ_1 为入射光线与法线的夹角，θ_2 为反射光线与法线的夹角。

因为

$$T^{(2)}(x)=\frac{a^2}{v_1\left((x+c)^2+a^2\right)^{3/2}}+\frac{b^2}{v_2\left((x-c)^2+b^2\right)^{3/2}}>0,$$

所以 T' 是严格增函数，从而 T' 有唯一零点(即 x_0)，并且它是 T 的最小值点。 □

习题5.2

1. 证明定理5.2.3的(5)和(6)。
2. 设 $0<a<b$，试比较 a^b 与 b^a 的大小。
3. 设 $\lambda\in\mathbb{R}$。讨论数列 $a_n=\left(1+\dfrac{1}{n}\right)^{n+\lambda}$ 的单调性。

4. 讨论函数 $f(x) = \frac{1-\cos x}{x-\sin x}$ 在区间 $(0, 2\pi]$ 上的单调性。

5. 设 $0 < a < b$，讨论函数 $f(x) = \begin{cases} \left(\frac{a^x+b^x}{2}\right)^{\frac{1}{x}}, & x \neq 0; \\ \sqrt{ab}, & x = 0 \end{cases}$ 的单调性。

6. 比较 $\frac{x}{\sin x}$ 和 $\frac{\tan x}{x}$ $(0 < x < \frac{\pi}{2})$的大小。

7. 如果例5.2.4中的海岸线不是直线，结论会怎样？

8. 直角坐标平面中，点 A, B 位于 $\mathscr{C}^1$ 曲线 $y = f(x)$ 的同一侧，求曲线 $y = f(x)$ 上的点 P 使得线段 AP 与 BP 的长度之和最小。

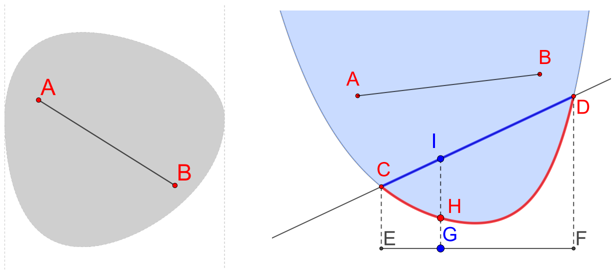
9. 如果甲乙两个人站在一个圆形的房间内，房间的内墙是反光镜，那么甲可以在墙面的镜子中看到乙的几个像？请用实验检查你的理论结果。

10. (快速检验) 据《长江日报》报道，2020年5月14日0时至6月1日24时，武汉对全市9,899,828人进行了新冠肺炎病毒核酸检测。在此之前，4月28日，世界知名医学期刊《柳叶刀》发布的一篇文章建议通过混合样本池的方法来提高检测速度，在保证灵敏度和准确性的前提下，最多可一次混合30份样本。这个方法是美国统计学家 Robert Dorfman 于1943年提出的混合抽样(sample pooling)方法[1]，目的是在数量众多的群体中快速查找患病个体，做法是：先把人群分组，每组样本混合后进行一次检验；如果检验结果呈阳性，那么再对该组成员进行逐一检验。试讨论
(1) 混合抽样是否可以提高检验效率？
(2) 为了最大可能提高检验效率，如何确定分组人数？

11. Young 不等式。设正数 p, q 满足 $\frac{1}{p} + \frac{1}{q} = 1$。证明对任意正数 x, y 都成立不等式
$$\frac{x^p}{p} + \frac{y^q}{q} \geq xy.$$
并讨论等号成立的条件。(提示：考虑函数 $f(x) = \frac{x^p}{p} + \frac{y^q}{q} - xy$。)

5.3 函数的凹凸性

定义 5.3.1. 称集合 $D \subset \mathbb{R}^2$是一个**凸集**，如果对任意 $A, B \in D$，线段 $AB \subseteq D$。



定义 5.3.2. 设函数 f 在区间 $I \subseteq \mathbb{R}$ 上是一个**凸函数**，如果 f 在 I 上有定义，且

$$D = \{(x, y) \in \mathbb{R}^2 | x \in I, y \geq f(x)\}$$

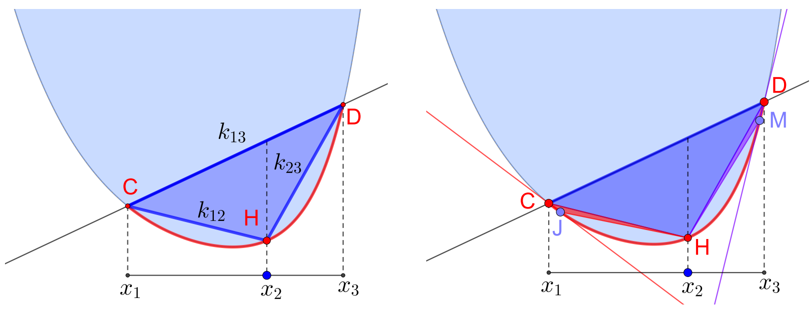
是凸集，即对任意 $x, y \in I$ 以及任意 $0 < t < 1$，

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

称 f 在区间 I 上是一个**严格凸函数**，如果 f 在 I 上是凸函数，并且上述不等式中等号成立当且仅当 $x = y$。

称函数 f 在区间 $I \subseteq \mathbb{R}$ 上是一个**凹函数(严格凹函数)**，如果 $-f$ 在 I 上是一个凸函数(严格凸函数)。

称 x_0 是 f 的一个**拐点**，如果 f 在 x_0 连续，且 f 在 x_0 两侧有相反的凹凸性。



上述定义的几何意义是：凸函数图像上任意两点之间的部分（弧）位于这两点所连直线（弦）的下方。

引理 5.3.3. 设 $A, B > 0$，$a, b \in \mathbb{R}$。若 $\frac{a}{A} < \frac{b}{B}$，则 $\frac{a}{A} < \frac{a+b}{A+B} < \frac{b}{B}$。 □

[1]Dorfman R. (1943) The detection of defective members of large populations. The Annals of Mathematical Statistics, 14(4): 436-440

以下引理把上述定义转化为弦的斜率关系。

引理 5.3.4. *(1) f 在区间 I 上是一个凸函数，当且仅当对任意 $x_1,x_2,x_3 \in I$，如果 $x_1<x_2<x_3$，则*

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_1)}{x_3-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

(2) f 在区间 I 上是一个严格凸函数，当且仅当对任意 $x_1,x_2,x_3 \in I$，如果 $x_1<x_2<x_3$，则

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} < \frac{f(x_3)-f(x_1)}{x_3-x_1} < \frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

证明. 取 $t=\dfrac{x_2-x_1}{x_3-x_1}$。对 x_1,x_3,t 用定义即可。 □

推论 5.3.5. *(1) 设 f 在区间 I 上可微，*

1. *f 在区间 I 上是一个凸函数当且仅当 f' 单调不减。*
2. *f 在区间 I 上严格凸当且仅当 f' 严格增。*

(2) 设 f 在区间 I 上二阶可微，则

1. *f 在区间 I 上是一个凸函数当且仅当 $f''\ge 0$。*
2. *若 $f''>0$，则 f 在区间 I 上是一个严格凸函数。*

证明. (1) 充分性：任取 $x_1,x_2,x_3\in I(x_1<x_2<x_3)$，则存在 $\xi\in(x_1,x_2),\eta\in(x_2,x_3)$ 使得

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}=f'(\xi),\quad f'(\eta)=\frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

若 f' 单调不减(严格增)，则 $f'(\xi)\le f'(\eta)$ (相应地，$f'(\xi)<f'(\eta)$)。从而

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} < \frac{f(x_3)-f(x_2)}{x_3-x_2}.$$

由上述两个引理知道 f 在区间 I 上是凸函数（严格凸函数）。

必要性：设 f 在区间 I 上是凸函数。则对任意 $x_1,x_2\in I$ ($x_1<x_2$)，记 $x_3=\frac{x_1+x_3}{2}$，任取充分小的 $h>0$，使 $x_1-h<x_1<x_1+h<x_3<x_2-h<x_2<x_2+h$，从而由已知条件得到

$$\begin{aligned}\frac{f(x_1-h)-f(x_1)}{-h} &\le \frac{f(x_1+h)-f(x_1)}{h} \le \frac{f(x_3)-f(x_1)}{x_3-x_1}\\ &\underset{(<,\text{当 } f \text{ 严格凸时})}{\le} \frac{f(x_2)-f(x_3)}{x_2-x_3} \le \frac{f(x_2-h)-f(x_2)}{-h} \le \frac{f(x_2+h)-f(x_2)}{h}.\end{aligned}$$

让 $h \to 0$，得到 $f'_-(x_1) \le f'_+(x_1) \le f'_-(x_2) \le f'_+(x_2)$（当 f 严格凸时，$f'_-(x_1) \le f'_+(x_1) < f'_-(x_2) \le f'_+(x_2)$）。因此 f' 单调不减（当 f 严格凸时，f' 严格增）。

(2) 对二阶可微函数 f，f' 单调不减当且仅当 $f'' \ge 0$。

若 $f'' > 0$，则 f' 严格增。 □

例 5.3.6. 若 $p > 1$ 或 $p < 0$，则 x^p 在 $(0, +\infty)$ 上严格凸；
若 $0 < p < 1$，则 x^p 在 $(0, +\infty)$ 上严格凹；
若 $a > 0$ 且 $a \ne 1$，则 a^x 在 $\mathbb{R}$ 上严格凸。
若 $a > 1$，则 $\log_a$ 在 $\mathbb{R}$ 上严格凹。
若 $0 < a < 1$，则 $\log_a$ 在 $\mathbb{R}$ 上严格凸。

定理 5.3.7. 设 f 在区间 I 上是凸函数，则对任意 $x_1, \ldots, x_n \in I$ 以及任意非负实数 $t_1, \ldots, t_n$，若 $t_1 + \cdots + t_n = 1$, 则

$$f(t_1x_1 + \cdots + t_nx_n) \le t_1f(x_1) + \cdots + t_nf(x_n).$$

如果 f 严格凸，则上述不等式中的等号成立当且仅当存在 $x_0 \in I$ 使得 $\{x_k | t_k > 0, 1 \le k \le n\} = \{x_0\}$。

例 5.3.8 (Young 不等式)**.** 设 $p_1, \ldots, p_n > 0$ 满足 $\dfrac{1}{p_1} + \cdots + \dfrac{1}{p_n} = 1$。则对任意 $x_1, \ldots, x_n \ge 0$，

$$\frac{x_1^{p_1}}{p_1} + \cdots + \frac{x_n^{p_n}}{p_n} \ge x_1 \cdots x_n.$$

其中等号成立当且仅当 $x_1^{p_1} = \cdots = x_n^{p_n}$。特别地，如果取 $p_1 = \cdots = p_n = n$，$x_k = a_k^{1/n}$，则得到算术-几何平均值不等式

$$\frac{a_1 + \cdots + a_n}{n} \ge \sqrt[n]{a_1 \cdots a_n}.$$

等号成立当且仅当 $a_1 = \cdots = a_n$。

证明. 不妨设 $x_1, \ldots, x_n > 0$，取 $u_k = p_k \ln x_k$。则上述不等式等价于

$$\frac{\mathrm{e}_1^{u_1}}{p_1} + \cdots + \frac{\mathrm{e}_n^{u_n}}{p_n} \ge \mathrm{e}^{\frac{u_1}{p_1}} \cdots \mathrm{e}^{\frac{u_n}{p_n}} = \mathrm{e}^{\frac{u_1}{p_1} + \cdots + \frac{u_n}{p_n}}.$$

因为 $f(u) = \mathrm{e}^u$ 是凸函数，所以对 f 和 $t_k = \frac{1}{p_k}$ 使用定理5.3.7，得到上述不等式。 □

关于凸函数的单调性和极值性质，我们有如下结论。

定理 5.3.9. 设 f 在区间 I 上是可微的凸函数，x_0 是 I 的内点满足 $f'(x_0) = 0$，则 x_0 是 f 在 I 上的最小值点。如果 f 是严格凸函数，则 f 在 I 上要么严格单调，要么有唯一的临界点，这个临界点是 f 的最小值点。

习题5.3

1. 证明定理5.3.7。

2. 证明定理**??**。

3. 设 f 在区间 I 上有定义，$x_0 \in I$，记

$$g_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0}, \quad x \in I \backslash \{x_0\}.$$

证明

(a) f 在区间 I 上是一个凸函数，当且仅当对任意 $x_0 \in I$， $g_{x_0}(x)$ 在 $I \backslash \{x_0\}$ 上单调不减；

(b) f 在区间 I 上是一个严格凸函数，当且仅当对任意 $x_0 \in I$， $g_{x_0}(x)$ 关于 x 在 $I \backslash \{x_0\}$ 上严格增。

(c) 如果 f 在区间 I 上是一个凸函数，则 f 在 I 的任何内点 x_0 处有有限的**左、右导数**

$$f'_-(x_0) := \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0}, \qquad f'_+(x_0) := \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}.$$

且 $f'_-(x_0) \leq f'_+(x_0)$。从而 f 在 I 的每个内点处连续。

(d) 如果 f 在区间 I 上是一个凸函数，x_1, x_2 是 I 的内点，$x_1 < x_2$，则 $f'_+(x_1) \leq f'_-(x_2)$。如果 f 是严格凸的，则 $f'_+(x_1) < f'_-(x_2)$。

(e) 设 f 在区间 I 上是一个凸函数。如果 f'_- 在 I 的内点 x_0 处连续，则 $f'_-(x_0) = f'_+(x_0)$，从而 f 在 x_0 可微。

(f) 设 f 在区间 I 上是一个凸函数。则 f 在 I 去掉一个最多可数的集合以外处处可微。

(g) 可导的凸函数都是 $\mathscr{C}^1$ 的。

(h) 若 f 是区间 I 上的凸函数，则要么 f 在 I 上单调，要么存在 $x_0 \in I$ 使得 f 在 $(-\infty, x_0) \cap I$ 上单调不增，在 $(x_0, +\infty,) \cap I$ 上单调不减。

4. （**Minkowsky 不等式**）设 $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$，$p \geq 1$。记 $\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$。证明 $\|x + y\|_p \leq \|x\|_p + \|y\|_p$。

5. （**Hölder 不等式**）设正数 p, q 满足 $\frac{1}{p} + \frac{1}{q} = 1$。证明对任意正数 x_1, x_2, y_1, y_2 都成立不等式

$$(x_1^p + x_2^p)^{\frac{1}{p}} (y_1^q + y_2^q)^{\frac{1}{q}} \geq x_1 y_1 + x_2 y_2.$$

（提示：考虑到不等式两边关于 x_1, x_2 都是一次齐次的，所以不妨设 $x_1 = 1$。于是可以考虑函数 $f(t) = (1 + t^p)^{\frac{1}{p}} (y_1^q + y_2^q)^{\frac{1}{q}} - y_1 - t y_2$。）

6. 设正数 $p_1,\dots,p_n,q_1,\dots,q_n$ 满足 $p_1+\cdots+p_n=q_1+\cdots+q_n=1$，证明

$$-\sum_{k=1}^{n}p_k\ln q_k\ge-\sum_{k=1}^{n}p_k\ln p_k,$$

并且其中等号成立当且仅当 $p_1=q_1,\dots,p_n=q_n$。从而

$$-\sum_{k=1}^{n}p_k\ln p_k\le\ln n.$$

最后这个不等式中等号成立当且仅当 $p_1=\cdots=p_n=\frac{1}{n}$。称 $-\sum\limits_{k=1}^{n}p_k\ln p_k$ 为概率分布 $p_1,\dots,p_n>0$ 的**熵**，这个不等式说明对有限多个状态的概率模型，等可能概率分布具有最大的熵。

7. 设 f 在区间 (a,b) 上是可微函数，证明：f 在区间 (a,b) 上是凸函数当且仅当对任意 $x_0\in I$，

$$f(x)\ge f(x_0)+f'(x_0)(x-x_0),\quad\forall x\in(a,b).$$

并给出一个几何解释。并给出拐点的一个几何解释。

8. 讨论 $(1+x)^\alpha$ 和 $1+\alpha x$ (其中 $x>-1$)的大小。

9. 设 f 在区间 $(a,+\infty)$ 上是凸函数，$y=kx+b$ 是 $y=f(x)$ 在 $x\to+\infty$ 时的渐近线，证明

$$f(x)\ge kx+b,\quad\forall x>a.$$

10. 设 $f:\mathbb{R}\to\mathbb{R}$ 是凸函数，且 f 有上界。证明 f 为常值函数。

11. 称 $f:I\to\mathbb{R}$ 是区间 I 上的**中点凸**函数，如果对任意 $x_1,x_2\in I$，

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

证明在以下条件之一成立时 f 在 I 上是凸函数。

(a) f 是连续函数；

(b) f 是单调函数；

(c) f 只有第一类间断点。

是否每个中点凸函数都是凸函数？

12. （**平均值不等式**）设 $\lambda_1,\lambda_2,\dots,\lambda_n$ 是正数，满足

$$\lambda_1+\lambda_2+\cdots+\lambda_n=1.$$

设 $a_1,a_2,\dots,a_n$ 是正数，不全相同。对非零实数 x，定义

$$f(x)=\left(\lambda_1a_1^x+\lambda_2a_2^x+\cdots+\lambda_na_n^x\right)^{\frac{1}{x}}.$$

求极限 $\lim\limits_{x\to0}f(x)$，并证明 f 是严格增函数。

13. 由平均值不等式知，对任意 $p>1$，

$$\lambda_1 a_1+\cdots+\lambda_n a_n \le (\lambda_1 a_1^p+\cdots+\lambda_n a_n^p)^{\frac{1}{p}}.$$

利用这个结论证明 **Hölder 不等式**：若正数 p,q 满足 $\frac{1}{p}+\frac{1}{q}=1$，则对任意非负实数 $x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_n$，都有

$$x_1y_1+x_2y_2+\cdots+x_ny_n \le \left(x_1^p+x_2^p+\cdots+x_n^p\right)^{\frac{1}{p}}\left(y_1^q+y_2^q+\cdots+y_n^q\right)^{\frac{1}{q}}.$$

5.4 应用： Newton 法的收敛性

定理 5.4.1 (Newton 法的局部收敛性与二阶收敛性). *设 $f:[a,b]\to\mathbb{R}$ 是 $\mathscr{C}^1$ 函数，$f(x^*)=0$，$f'(x^*)\neq 0$。则*

(1) 在 x^* 的一个邻域中， Newton 迭代

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$$

得到的数列 $\{x_n\}_{n\ge 0}$ 收敛于 x^*；

(2) 若 f 是 $\mathscr{C}^2$ 函数，则存在常数 C 使得

$$|x_{n+1}-x^*|\le C\,|x_n-x^*|^2.$$

证明. (1) 不妨设 $f'(x^*)>0$(否则可以考虑$-f$)。因为 f' 在 x^* 连续，所以存在 $\delta>0$ 使得 $\forall x\in(x^*-\delta,x^*+\delta)$，

$$\frac{10}{11}<\frac{f'(x)}{f'(x^*)}<\frac{11}{10}$$

从而对任意 $x,y\in(x^*-\delta,x^*+\delta)$，

$$\frac{4}{5}<\left(\frac{10}{11}\right)^2<\frac{f'(x)}{f'(y)}=\frac{f'(x)}{f'(x^*)}\frac{f'(x^*)}{f'(y)}<\left(\frac{11}{10}\right)^2<\frac{5}{4},$$

从而

$$\left|1-\frac{f'(x)}{f'(y)}\right|\le\frac{1}{4}.$$

因此当 $x_n\in(x^*-\delta,x^*+\delta)$ 时，存在介于 x_n 与 x^* 之间的 ξ_n 使得

$$\begin{aligned}|x_{n+1}-x^*|&=\left|x_n-x^*-\frac{f(x_n)-f(x^*)}{f'(x_n)}\right|\\&=\left|x_n-x^*-\frac{f'(\xi_n)(x_n-x^*)}{f'(x_n)}\right|\\&=|x_n-x^*|\left|1-\frac{f'(\xi_n)}{f'(x_n)}\right|<\frac{1}{4}\,|x_n-x^*|,\end{aligned}$$

$x_{n+1}\in(x^*-\delta,x^*+\delta)$。因此只要 $x_0\in(x^*-\delta,x^*+\delta)$，数列 x_n 就收敛于 x^*。

(2) 不妨设 $f'(x) \neq 0(\ \forall x \in [a,b]\ )$。存在介于 x_n 与 ξ_n 之间的 η_n 使得

$$|x_{n+1}-x^*| = |x_n - x^*|\left|1-\frac{f'(\xi_n)}{f'(x_n)}\right| = |x_n-x^*|\left|\frac{f''(\eta_n)}{f'(x_n)}\right||\xi_n - x^*|$$
$$\leq |x_n-x^*|^2\,\frac{\max\limits_{x\in[a,b]}|f''(x)|}{\min\limits_{x\in[a,b]}|f'(x)|}$$

□

从上述证明中的最后一个不等式可以看出，当 $y=f(x)$ 近似直线（二阶导数 f'' 绝对值小）且斜率较大（一阶导数 f' 绝对值大）时， Newton 法有很好的收敛性。

定理 5.4.2 (Newton 法的全局收敛性). 设 $f:[a,b]\to\mathbb{R}$ 是严格增的 $\mathscr{C}^1$ 凸函数，$f(a)<0<f(b)$。则对于任意 $x_0\in(a,b)$，只要$x_0-\frac{f(x_0)}{f'(x_0)}<b$，由 *Newton* 迭代

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$$

得到的数列 $\{x_n\}_{n\geq 0}$ 就收敛于 f 的（唯一）零点 x^*。

证明. f 是凸函数，所以对任意 $x,y\in(a,b)$，

$$f(x)\geq f(y)+f'(y)(x-y),$$

所以

$$f(x_{n+1})\geq f(x_n)+f'(x_n)(x_{n+1}-x_n)=0.$$

又 $f'(x_{n+1})>0$，所以

$$x_{n+2}=x_{n+1}-\frac{f(x_{n+1})}{f'(x_{n+1})}\leq x_{n+1}.$$

因此 $x^*\leq x_{n+2}\leq x_{n+1}$。从而只要 $x_1\in(a,b)$，数列 $\{x_n\}_{n\geq 0}$ 就必然收敛，记 $A=\lim\limits_{n\to+\infty}x_n$。对

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$$

取极限，因为 f' 连续且 $f'(A)>0$，所以得到

$$A=A-\frac{f(A)}{f'(A)},$$

所以 $f(A)=0$。又因为 x^* 是 f 的唯一零点，所以 $A=x^*$。 □

习题5.4

1. 讨论求解方程 $\arctan x=0$ 的 Newton 迭代数列的收敛性与初值 x_0 的关系。

2. 设 x^* 是 f 的 m 重零点，即 $f(x)=(x-x^*)^m g(x)$，其中 $g(x^*)\neq 0$。证明
   (a) 此时，Newton 法是线性收敛的，即 $|x_{n+1}-x^*|\le C|x_n-x^*|$；
   (b) 如果用以下改进的 Newton 迭代
   $$x_{n+1}=x_n-m\frac{f(x_n)}{f'(x_n)},$$
   则 $|x_{n+1}-x^*|\le C|x_n-x^*|^2$。即这个改进的迭代是二阶收敛的。
   (c) x^* 是 $\frac{f(x)}{f'(x)}$ 的单重零点，于是可以对方程 $\frac{f(x)}{f'(x)}=0$ 使用 Newton 法。试写出这个迭代的表达式，并讨论它的收敛阶数。

3. Newton 法虽然具有较快的收敛性，但每次迭代需要重新计算导数。而割线法迭代只需要计算函数值：考虑过点 $(x_n,f(x_n))$ 和 $(x_{n+1},f(x_{n+1}))$ 的直线与 x 轴的交点坐标
   $$x_{n+2}=x_{n+1}-\frac{f(x_{n+1})(x_{n+1}-x_n)}{f(x_{n+1})-f(x_n)}.$$
   讨论割线法迭代的收敛性和收敛的阶数。

5.5 L'Hôpital 法则

定理 5.5.1 ($\frac{0}{0}$ **型 L'Hôpital**)**.** 设 f 在 $a\in\mathbb{R}\cup\{-\infty,+\infty\}$ 的一个（单侧）去心邻域中可微，且
$$\lim_{x\to a}f(x)=\lim_{x\to a}g(x)=0,\quad \lim_{x\to a}\frac{f'(x)}{g'(x)}=A\in\mathbb{R}\cup\{-\infty,+\infty,\infty\}.$$
则
$$\lim_{x\to a}\frac{f(x)}{g(x)}=A.$$

证明. 不妨设 $A>-\infty$，否则可以考虑用 $-f(x)$ 代替 $f(x)$。不妨设 $a=x_0^+$ 或 $a=-\infty$，否则可以考虑 $f(-x)$ 和 $g(-x)$。

由 Darboux 定理，不妨设 $g'(x)>0$，否则可以考虑 $-g(x)$。于是 g 在 a 右侧严格增。因此在 a 右侧 $g(x)>0$。

对任意 $A_1\in(-\infty,A)$，存在 a 的去心邻域 U 使得对任意 $x\in U$，
$$\frac{f'(x)}{g'(x)}>A_1.$$
因此
$$[f(x)-A_1g(x)]'>0,$$
于是 $f(x)-A_1g(x)$ 在 a 右侧单调增。从而
$$f(x)-A_1g(x)>\lim_{x\to a}[f(x)-A_1g(x)]=0.$$

从而

$$\frac{f(x)}{g(x)} > A_1.$$

因此当 $A=+\infty$ 时，这证明了 $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}=A$。

类似证明当 $A<+\infty$ 时，对任意 $A_2\in(A,+\infty)$，存在 a 的去心邻域 U 使得对任意 $x\in U$，

$$\frac{f(x)}{g(x)} < A_2.$$

于是当 $A<+\infty$ 时，$\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}=A$。

同理可证当 $A=\infty$ 时，$\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}=A$。 □

定理 5.5.2 ($\frac{\infty}{\infty}$ 型 L'Hôpital). 设 f 在 $a\in\mathbb{R}\cup\{-\infty,+\infty\}$ 的一个（单侧）去心邻域中可微，且

$$\lim_{x\to a} g(x)=\infty,\quad \lim_{x\to a}\frac{f'(x)}{g'(x)}=A\in\mathbb{R}\cup\{-\infty,+\infty\}.$$

则 $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$ 存在，且 $\lim\limits_{x\to a}\frac{f(x)}{g(x)}=A$。

证明. 我们对 $A\in\mathbb{R}$ 给出证明，读者不难发现它依然适用于 $A=+\infty$ 或 $A=-\infty$。

不妨设 $a=x_0^-$ 或 $a=+\infty$，否则可以考虑 $f(-x)$ 和 $g(-x)$。由 Darboux 定理，不妨设 $g'(x)>0$，否则可以考虑 $-g(x)$。从而当 $x\to a$ 时，$g(x)$ 单调增，且 $\lim\limits_{x\to a}g(x)=+\infty$。因此不妨设 $g(x)>0$。

任取实数 A_1,A_2,A_3,A_4 使得 $A_1<A_3<A<A_4<A_2$，存在 a 的左侧去心邻域 U 使得对任意 $x\in U$，

$$A_3<\frac{f'(x)}{g'(x)}<A_4.$$

因此 $(f(x)-A_3g(x))'>0>(f(x)-A_4g(x))'$。任取 $x_0\in U$。于是，

$$f(x)-A_3g(x)>f(x_0)-A_3g(x_0),\quad f(x)-A_4g(x)<f(x_0)-A_4g(x_0),\quad \forall x\in(x_0,a).$$

从而

$$A_3+\frac{f(x_0)-A_3g(x_0)}{g(x)}<\frac{f(x)}{g(x)}<A_4+\frac{f(x_0)-A_4g(x_0)}{g(x)}.$$

因为

$$\lim_{x\to a}\left(A_3+\frac{f(x_0)-A_3g(x_0)}{g(x)}\right)=A_3>A_1,$$
$$\lim_{x\to a}\left(A_4+\frac{f(x_0)-A_4g(x_0)}{g(x)}\right)=A_4<A_2,$$

所以存在 a 的去心邻域 $V \subset U$，使得对任意 $x \in V$，

$$A_1 < \frac{f(x)}{g(x)} < A_2.$$

因此 $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = A$。 □

例 5.5.3. 求极限 $\lim\limits_{x\to+\infty} \left(\frac{\pi}{2} - \arctan x\right)^{\frac{1}{\ln x}}$。

解： 取对数

$$\frac{\ln\left(\frac{\pi}{2} - \arctan x\right)}{\ln x},$$

因为

$$\frac{\left(\ln\left(\frac{\pi}{2} - \arctan x\right)\right)'}{(\ln x)'} = \frac{\frac{1}{\frac{\pi}{2}-\arctan x}\left(-\frac{1}{1+x^2}\right)}{\frac{1}{x}} = \frac{-x^2}{(1+x^2)} \cdot \frac{\frac{1}{x}}{\frac{\pi}{2} - \arctan x},$$

对 $\frac{\frac{1}{x}}{\frac{\pi}{2}-\arctan x}$ 用 L'Hôpital 法则，得到

$$\lim_{x\to+\infty} \frac{\frac{1}{x}}{\frac{\pi}{2} - \arctan x} = \lim_{x\to+\infty} \frac{-\frac{1}{x^2}}{-\frac{1}{1+x^2}} = 1,$$

所以根据 L'Hôpital 法则

$$\lim_{x\to+\infty} \frac{\ln\left(\frac{\pi}{2} - \arctan x\right)}{\ln x} = \lim_{x\to+\infty} \frac{\left(\ln\left(\frac{\pi}{2} - \arctan x\right)\right)'}{(\ln x)'} = -1,$$

从而 $\lim\limits_{x\to+\infty} \left(\frac{\pi}{2} - \arctan x\right)^{\frac{1}{\ln x}} = \mathrm{e}^{-1}$。 □

习题5.5

1. 设 f 在 $(x_0 - \delta_0, x_0 + \delta_0)$ 上连续，在 $(x_0 - \delta_0, x_0 + \delta_0)\backslash\{x_0\}$ 可微，且 $\lim\limits_{x\to x_0} f'(x) = A \in \mathbb{R}$。证明 f 在 x_0 可微，且 $f'(x_0) = A$，即 f' 在 x_0 连续。

2. 设 $f(x) = \mathrm{e}^{-2x}(\cos x + 2\sin x)$，$g(x) = \mathrm{e}^{-x}(\cos x + \sin x)$。则 $\lim\limits_{x\to 0} f(x) = \lim\limits_{x\to 0} g(x) = 0$，

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{-5\mathrm{e}^{-2x}\sin x}{-2\mathrm{e}^{-x}\sin x} = 0,$$

   所以 $\lim\limits_{x\to 0} \frac{f(x)}{g(x)} = 0$。

   请问这个推理有问题吗？

3. 设 $y = f(x)$ 在 $x \to +\infty$ 时有渐近线，且 $\lim\limits_{x\to+\infty} f'(x) = k$。证明 $y = f(x)$ 的渐近线为 $y = kx + b$，其中 $b = \lim\limits_{x\to+\infty} (f(x) - kx)$。

4. 求极限 $\lim\limits_{x\to 0}\frac{1}{x^{n+1}}\left(1-\sum\limits_{k=0}^{n}\mathrm{e}^{-x}\frac{x^k}{k!}\right)$。

5. 设 $f\in\mathscr{C}^\infty(\mathbb{R})$，$f(0)=0$。记

$$g(x)=\begin{cases}\frac{f(x)}{x}, & x\neq 0;\\ f'(0), & x=0.\end{cases}$$

证明 $g\in\mathscr{C}^\infty(\mathbb{R})$。

6. **Stolz 定理**(离散形式的 L'Hôpital 法则)
设 $A\in\mathbb{R}\cup\{-\infty,+\infty\}$，数列 a_n 和 b_n 满足：a_n 单调且

$$\lim_{n\to+\infty}\frac{b_n-b_{n-1}}{a_n-a_{n-1}}=A\in\mathbb{R}\cup\{-\infty,+\infty\}.$$

若 $\lim\limits_{n\to+\infty}a_n=\infty$ 或 $\lim\limits_{n\to+\infty}a_n=\lim\limits_{n\to+\infty}b_n=0$，则 $\lim\limits_{n\to+\infty}\dfrac{b_n}{a_n}=A$。

5.6 Taylor 公式

多项式只涉及有限多次乘法和加法运算，是相对简单的一类函数。Taylor 公式的目的就是把一般的可微函数转换成多项式处理。

定义 5.6.1. 设 f 在 x_0 处 n 阶可微。称多项式

$$Tf_{n,x_0}(x)=f(x_0)+f'(x_0)(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n,$$

称为 f 在 x_0 处的 n 阶 **Taylor 多项式**。

注 5.6.2. 从定义不难发现，若 f,g 在 x_0 处 n 阶可微，则
(1) 对常数 λ,μ，$T(\lambda f+\mu g)_{n,x_0}(x)=\lambda Tf_{n,x_0}(x)+\mu Tg_{n,x_0}(x)$；
(2) 对 $h(t)=f(\lambda t+\mu)$ 及 $t_0=\frac{x_0-\mu}{\lambda}$，成立 $Th_{n,t_0}(t)=Tf_{n,x_0}(\lambda t+\mu)$。
(3) $(Tf_{n,x_0}(x))'=T(f')_{n-1,x_0}(x)$。

定理 5.6.3 (Peano 余项 Taylor 公式，Taylor 多项式的唯一性). 设 f 在 x_0 处 n 阶可微。则多项式

$$P_n(x)=a_0+a_1(x-x_0)+\frac{a_2}{2!}(x-x_0)^2+\cdots+\frac{a_n}{n!}(x-x_0)^n$$

满足 $f(x)=P_n(x)+o\big((x-x_0)^n\big),x\to x_0$ 当且仅当 $Tf_{n,x_0}(x)=P_n(x)$。

证明. (唯一性) 设 P_n, Q_n 都是不超过 n 次的多项式，且满足

$$P_n(x) - Q_n(x) = o((x-x_0)^n), \quad x \to x_0.$$

如果 $P_n \neq Q_n$，则存在非负整数 k 以及 $c_k \neq 0$ 使得 $0 \leq k \leq n$ 且

$$P_n(x) - Q_n(x) = c_k(x-x_0)^k + o((x-x_0)^k), \quad x \to x_0.$$

于是

$$c_k(x-x_0)^k + o((x-x_0)^k) = o((x-x_0)^n), \quad x \to x_0.$$

从而

$$c_k = -o(1) + o((x-x_0)^{n-k}) = o(1), \quad x \to x_0,$$

于是 $c_k = 0$ 矛盾。所以 $P_n = Q_n$。唯一性得证。于是定理的必要性是充分性和唯一性的推论。

(充分性) 当 $n = 1$ 时，因 f 可微，所以

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + o(x-x_0), \quad x \to x_0.$$

因此 $n = 1$ 时充分性成立。

假设 n 时充分性结论成立。设 f 在 x_0 处 $n+1$ 阶可微。则由 L'Hôpital 法则和归纳假设知

$$\lim_{x\to x_0} \frac{f(x) - Tf_{n+1,x_0}(x)}{\dfrac{(x-x_0)^{n+1}}{(n+1)!}} = \lim_{x\to x_0} \frac{f'(x) - T(f')_{n,x_0}(x)}{\dfrac{(x-x_0)^n}{n!}} = 0.$$

所以充分性对 $n+1$ 也成立。所以充分性对任意正整数 n 成立。 □

注 5.6.4. 上述定理给出了在一点附近逼近 n 阶可微函数的多项式的存在性和唯一性(在限定次数的情况下)。但是一个函数在一点附近可用多项式逼近，并不能保证这个函数有高阶可微性。例如

$$f(x) = \begin{cases} x^{100}, & x \in \mathbb{Q}, \\ 0, & x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$$

满足 $f(x) = o(x^{99}), x \to 0$，但 f 在 $x = 0$ 处不是二阶可微的。

推论 5.6.5. 设 f 在 x_0 处 n 阶可微，且

$$f(x) = a_0 + a_1(x-x_0) + \frac{a_2}{2!}(x-x_0)^2 + \cdots + \frac{a_n}{n!}(x-x_0)^n + o((x-x_0)^n), \quad x \to x_0.$$

则

$$f'(x) = a_1 + a_2(x-x_0) + \cdots + \frac{a_n}{(n-1)!}(x-x_0)^{n-1} + o((x-x_0)^{n-1}), \quad x \to x_0.$$

证明. 由上述定理知 $a_0+a_1(x-x_0)+\frac{a_2}{2!}(x-x_0)^2+\cdots+\frac{a_n}{n!}(x-x_0)^n$ 是 $f(x)$ 在 x_0 处的 Taylor 多项式，从而 $a_1+a_2(x-x_0)+\cdots+\frac{a_n}{(n-1)!}(x-x_0)^{n-1}$ 是 $f'(x)$ 在 x_0 处的 Taylor 多项式。再由上述定理知

$$f'(x)=a_1+a_2(x-x_0)+\cdots+\frac{a_n}{(n-1)!}(x-x_0)^{n-1}+o((x-x_0)^{n-1}),\quad x\to x_0.$$

□

例 5.6.6. 求基本初等函数的 Taylor 展开。

解： (1) 设 e^x 的 Taylor 展开式为

$$\mathrm{e}^x=\sum_{k=0}^{n}\frac{a_k}{k!}(x-x_0)^k+o\big((x-x_0)^n\big),\quad x\to x_0.$$

则根据上述推论，

$$(\mathrm{e}^x)'=\sum_{k=1}^{n}\frac{a_k}{(k-1)!}(x-x_0)^{k-1}+o\big((x-x_0)^{n-1}\big),\quad x\to x_0.$$

再由 $(\mathrm{e}^x)'=\mathrm{e}^x$ 以及 Taylor 展开的唯一性，比较多项式对应项系数，得到

$$\frac{a_{k+1}}{k!}=\frac{a_k}{k!},\quad k=0,\ldots,n-1,$$

因此 $a_n=a_{n-1}=\cdots=a_1=a_0=\mathrm{e}^{x_0}$，从而

$$\mathrm{e}^x=\sum_{k=0}^{n}\frac{\mathrm{e}^{x_0}}{k!}(x-x_0)^k+o\big((x-x_0)^n\big)=\mathrm{e}^{x_0}\sum_{k=0}^{n}\frac{(x-x_0)^k}{k!}+o\big((x-x_0)^n\big),\quad x\to x_0.$$

对任意 $a>0$，

$$\begin{aligned}a^x=\mathrm{e}^{x\ln a}&=\mathrm{e}^{x_0\ln a}\sum_{k=0}^{n}\frac{1}{k!}(x\ln a-x_0\ln a)^k+o\big((x\ln a-x_0\ln a)^n\big)\\&=a^{x_0}\sum_{k=0}^{n}\frac{(\ln a)^k}{k!}(x-x_0)^k+o\big((x-x_0)^n\big),\quad x\to x_0.\end{aligned}$$

特别是 $x_0=0$ 时，

$$\mathrm{e}^x=\sum_{k=0}^{n}\frac{x^k}{k!}+o\big(x^n\big),\quad x\to 0.$$

$$a^x=\sum_{k=0}^{n}\frac{(x\ln a)^k}{k!}+o\big(x^n\big),\quad x\to 0.$$

(2) 设 $\sin x$ 和 $\cos x$ 在 $x=0$ 处的 Taylor 展开分别为

$$\sin x=\sum_{k=0}^{n}\frac{s_k}{k!}x^k+o(x^k),\qquad \cos x=\sum_{k=0}^{n}\frac{c_k}{k!}x^k+o(x^k),\qquad x\to 0.$$

因为 $\sin' = \cos$，$\cos' = -\sin$，并且 $\sin 0 = 0$，$\cos 0 = 1$，所以由 Taylor 展开的唯一性通过比较系数得到

$$s_k = c_{k-1}, \quad c_k = -s_{k-1}, \quad s_0 = 0, \quad c_0 = 1.$$

因此 $s_1 = c_0 = 1$，$c_1 = -s_0 = 0$，$s_{2k} = c_{2k+1} = 0$，$s_{2k+1} = c_{2k} = (-1)^k$，

$$\sin x = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o\big(x^{2n+2}\big), \quad \cos x = \sum_{k=0}^{n} \frac{(-1)^k x^{2k}}{(2k)!} + o\big(x^{2n+1}\big), \quad x \to 0.$$

(3) 对 $\alpha \in \mathbb{R}$，设

$$(1+x)^\alpha = \sum_{k=0}^{n} \frac{a_k}{k!} x^k + o(x^n), \quad x \to 0.$$

注意到

$$(1+x)\big((1+x)^\alpha\big)' = \alpha(1+x)^\alpha,$$

代入 $(1+x)^\alpha$ 及其导数的 Taylor 展开式，比较系数得到

$$a_{k+1} + k a_k == \alpha a_k.$$

因此

$$a_{k+1} = (\alpha - k) a_k, k = 0, \ldots, n-1; \quad a_0 = 1.$$

从而

$$a_k = (\alpha - k + 1)(\alpha - k + 2) \cdots \alpha, k = 1, \ldots, n; \quad a_0 = 1.$$

所以 $(1+x)^\alpha$ 在 $x = 0$ 处的 Taylor 展开为

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

其中 $\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$。

当 α 是正整数 m 时，

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

这是经典的二项展开式。

特别对 $\alpha = -1$，

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

因为 $\big(\ln(1+x)\big)' = \dfrac{1}{1+x}$，$\ln(1+0) = 0$，所以利用注5.5.2

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + \frac{(-1)^n x^{n+1}}{n+1}$$

是 $\ln(1+x)$ 在 $x=0$ 处的 Taylor 多项式。因此

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots+\frac{(-1)^n x^{n+1}}{n+1}+o\big(x^{n+1}\big),\quad x\to 0.$$

因为 $\big(\arctan x\big)'=\dfrac{1}{1+x^2}$，$\arctan 0=0$ 及

$$\frac{1}{1+x^2}=1-x^2+x^4-\cdots+(-1)^n x^{2n}+o(x^{2n}),\quad x\to 0,$$

所以利用利用注5.5.2得到

$$\arctan x=x-\frac{x^3}{3}+\frac{x^5}{5}-\cdots+\frac{(-1)^n x^{2n+1}}{2n+1}+o(x^{2n+1}),\quad x\to 0.$$

利用 $\big(\arcsin x\big)'=(1-x^2)^{-\frac{1}{2}}$，$\arcsin 0=0$ 可以求得 $\arcsin x$ 在 $x=0$ 处的 Taylor 展开式。

再利用 $\arccos x=\dfrac{\pi}{2}-\arcsin x$ 得到 $\arccos x$ 在 $x=0$ 处的 Taylor 展开式。

□

例 5.6.7. 求 $\ln(1+\sin x)$ 在 $x=0$ 处的带有 Peano 余项的 4 阶 Taylor 展开式。

解：

$$\begin{aligned}
&\ln(1+\sin x)\\
=&\sin x-\frac{\sin^2 x}{2}+\frac{\sin^3 x}{3}-\frac{\sin^4 x}{4}+o\left(\sin^4 x\right)\\
=&\left(x-\frac{x^3}{3!}+o(x^4)\right)-\frac{x^2}{2}\left(1-\frac{x^2}{3!}+o(x^3)\right)^2+\frac{x^3}{3}\left(1+o(x)\right)^3-\frac{x^4}{4}+o(x^4)\\
=&x-\frac{x^2}{2}+\frac{x^3}{6}-\frac{x^4}{12}+o(x^4).
\end{aligned}$$

由定理5.5.3知这是 $\ln(1+\sin x)$ 在 $x=0$ 处的 4 阶 Taylor 展开。 □

利用 Peano 余项的 Taylor 公式，我们可以更方便地计算一些复杂函数的极限。

例 5.6.8. 求 $\displaystyle\lim_{x\to 0}\frac{\ln(1+\sin x)-\sin(\ln(1+x))}{\mathrm{e}^{\arcsin x}-1-\arcsin(\mathrm{e}^x-1)}$。

解： 由上例知

$$\ln(1+\sin x)=x-\frac{x^2}{2}+\frac{x^3}{6}-\frac{x^4}{12}+o(x^4).$$

同理可得

$$\begin{aligned}
\sin(\ln(1+x))&=x-\frac{x^2}{2}+\frac{x^3}{6}+0x^4+o(x^4),\\
\mathrm{e}^{\arcsin x}-1&=x+\frac{x^2}{2}+\frac{x^3}{3}+\frac{5x^4}{24}+o\left(x^4\right),\\
\arcsin(\mathrm{e}^x-1)&=x+\frac{x^2}{2}+\frac{x^3}{3}+\frac{7x^4}{24}+o\left(x^4\right),
\end{aligned}$$

所以

$$\lim_{x\to 0}\frac{\ln(1+\sin x)-\sin(\ln(1+x))}{\mathrm{e}^{\arcsin x}-1-\arcsin(\mathrm{e}^x-1)}=\lim_{x\to 0}\frac{-\dfrac{x^4}{12}+o\left(x^4\right)}{-\dfrac{x^4}{12}+o\left(x^4\right)}=1.$$

□

例 5.6.9. 求极限 $\lim\limits_{x\to+\infty}\left(\frac{\pi}{2}-\arctan x\right)^{\frac{1}{\ln x}}$。

解： 令 $y=\frac{\pi}{2}-\arctan x$，则 $x\to+\infty$ 时，$y\to 0^+$，$x=\frac{1}{\tan y}$。于是

$$\begin{aligned}\ln\left[\left(\frac{\pi}{2}-\arctan x\right)^{\frac{1}{\ln x}}\right]&=-\frac{\ln y}{\ln\tan y}=-\frac{\ln y}{\ln(y+o(y))}=-\frac{\ln y}{\ln y+\ln(1+o(1))}\\&=-\frac{1}{1+\frac{\ln(1+o(1))}{\ln y}}\to -1,\quad y\to 0^+,\end{aligned}$$

所以 $\lim\limits_{x\to+\infty}\left(\frac{\pi}{2}-\arctan x\right)^{\frac{1}{\ln x}}=\mathrm{e}^{-1}$。 □

例 5.6.10. 求 $y=f(x)$ 在 $P(x_0,f(x_0))$ 处的密切圆半径和曲率。

解： 适当平移坐标系，不妨设 $x_0=0$，$f(x_0)=0$。于是过 $P(x_0,y_0)$ 的圆的方程为

$$(x-a)^2+(y-b)^2=(x_0-a)^2+(y_0-b)^2=a^2+b^2=R^2$$

解得

$$\begin{aligned}y&=b\pm\sqrt{R^2-(x-a)^2}=b\pm\sqrt{b^2+2ax-x^2}\\&=b-b\sqrt{1+\frac{2ax-x^2}{b^2}}\\&=-\frac{1}{2}\frac{2ax-x^2}{b}-b\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)}{2}\left(\frac{2ax-x^2}{b^2}\right)^2+o(x^2)\\&=-\frac{a}{b}x+\frac{R^2}{2b^3}x^2+o(x^2).\end{aligned}$$

这个圆与 $y=f(x)$ 在 $P(x_0,y_0)$ 二阶相切，当且仅当它们在 $x=x_0$ 处有相同的二阶 Taylor 展开式，即

$$f'(x_0)=-\frac{a}{b},\qquad f''(x_0)=\frac{R^2}{b^3}.$$

因此

$$\left(1+f'(x_0)^2\right)^{3/2}=\frac{R^3}{|b|^3},$$

所以 $y=f(x)$ 在 $P(x_0,f(x_0))$ 处的曲率半径和曲率分别为

$$R=\frac{\left(1+f'(x_0)^2\right)^{3/2}}{|f''(x_0)|},\quad \kappa=\frac{1}{R}=\frac{|f''(x_0)|}{(1+f'(x_0)^2)^{3/2}}.$$

□

Peano 余项的 Taylor 公式只提供了 $f(x)$ 在 x 接近 x_0 时的局部信息。而下面这个 Lagrange 余项的 Taylor 公式则给出了 $f(x)$ 在一个区间上的整体性质，并且对余项有了更细致的刻画，当然这需要 f 满足更强的条件。

定理 5.6.11 (Lagrange 余项 Taylor 公式). 设 f 在区间 I 上n 阶可微，$x_0 \in I$。则对任意 $x \in I$，存在介于 x, x_0 之间的 ξ 使得

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x-x_0)^n.$$

证明. 记 $x = x_0 + th$，考虑

$$F(t) = f(x_0+th) - \sum_{k=0}^{n-1} t^k \frac{f^{(k)}(x_0)}{k!} h^k.$$

则 F 在区间 $[0,1]$ 上n 阶可微，且对 $m = 1, 2, \ldots, n-1$，

$$F^{(m)}(t) = f^{(m)}(x_0+th)h^m - \sum_{k=m}^{n-1} t^{k-m} \frac{f^{(k)}(x_0)}{(k-m)!} h^k.$$

于是 $F(0) = F'(0) = \cdots = F^{(n-1)}(0) = 0$。

$$F^{(n)}(t) = f^{(n)}(x_0+th)h^n.$$

因此只需证明存在 $0 < \theta < 1$ 使得 $F(1) = \dfrac{F^{(n)}(\theta)}{n!}$。

令 $G(t) = \dfrac{t^n}{n!}$，则 $G^{(m)}(t) = \dfrac{t^{n-m}}{(n-m)!)}$。于是由 Cauchy 微分中值定理，

$$\frac{F(1)}{G(1)} = \frac{F(1)-F(0)}{G(1)-G(0)} = \frac{F'(\theta_1)}{G'(\theta_1)} = \frac{F'(\theta_1)-F'(0)}{G'(\theta_1)-G'(0)} = \frac{F^{(2)}(\theta_2)}{G^{(2)}(\theta_2)} = \cdots = \frac{F^{(n)}(\theta_n)}{G^{(n)}(\theta_n)}.$$

所以

$$F(1) = \frac{F^{(n)}(\theta_n)}{n!} = \frac{f^{(n)}(\xi)}{n!}h^n.$$

□

推论 5.6.12. 设 f 在一个含x_0 在内的区间内有连续的 $n+1$ 阶导函数，则当 $x \to x_0$ 时

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + O\left((x-x_0)^{n+1}\right).$$

例 5.6.13. 估算 $\sqrt{10}$ 的值。

解： 因为 $2^{10} = 1024 = 100(10+0.24)$，所以 $2^5 = 10\sqrt{10(1+0.024)}$。

$$\begin{aligned}\sqrt{10} &= \frac{32}{10}(1+0.024)^{-1/2} \\ &= \frac{32}{10}\left(1 - \frac{1}{2}0.024 + \frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)}{2!}0.024^2 + \frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)\left(-\frac{1}{2}-2\right)}{3!(1+\xi)^{\frac{3}{2}}}0.024^3\right)\end{aligned}$$

其中

$$-0.000013824=-\frac{32}{10}\frac{15}{8}\frac{1}{3!}0.024^3<\frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)\left(-\frac{1}{2}-2\right)}{3!(1+\xi)^{\frac{3}{2}}}0.024^3<0,$$

所以 $3.1622737<\sqrt{10}<3.1622912$。 □

例 5.6.14. 证明对任意 $x\in\mathbb{R}$，$\mathrm{e}^x=\lim\limits_{n\to+\infty}\left(1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}\right)$。我们把后面这个极限记为 $\sum\limits_{n=0}^{\infty}\frac{x^n}{n!}$ 或者 $1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}+\cdots$。因此

$$\mathrm{e}^x=1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}+\cdots.$$

证明. 任取正整数 $N>0$ 和 $\varepsilon>0$，取正整数 $n>\max\left\{3N,\ln\left(\frac{\mathrm{e}^{4N}N^{3N}}{(3N)!\varepsilon}\right)\right\}$。对任意 $x\in[-N,N]$，存在介于 0 与 x 之间的 ξ 使得

$$\mathrm{e}^x=1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}+\frac{\mathrm{e}^{\xi}x^{n+1}}{(n+1)!}.$$

因而 $|\xi|\le N$，于是

$$\begin{aligned}\left|\mathrm{e}^x-\left(1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}\right)\right|&=\left|\frac{\mathrm{e}^{\xi}x^{n+1}}{(n+1)!}\right|\\&<\frac{\mathrm{e}^NN^{n+1}}{(n+1)!}<\frac{\mathrm{e}^NN^{3N}N^{n+1-3N}}{(3N)!(3N)^{n+1-3N}}\\&\le\frac{\mathrm{e}^NN^{3N}}{(3N)!}\left(\frac{1}{3}\right)^{n+1-3N}\\&<\frac{\mathrm{e}^NN^{3N}}{(3N)!}\left(\frac{1}{\mathrm{e}}\right)^{n+1-3N}=\frac{\mathrm{e}^{4N}N^{3N}}{(3N)!\mathrm{e}^{n+1}}<\varepsilon.\end{aligned}$$

□

事实上，上述证明所得的结果比我们预先想要的结果要强，我们实际上证明了在任何有界闭区间上 e^x 的 Taylor 多项式随次数 n 趋于无穷大会一致收敛于 e^x，即对给定的逼近精度 ε，我们对 n 的要求可以适用于这个有界闭区间中的所有 x。关于函数的一致逼近问题，我们以后会更详细的讨论。

习题5.6

1. 设 f 在 x_0 处 n 阶可微。证明对任何次数不大于 n 的多项式 $P(x)$，

$$\lim_{x\to x_0}\frac{f(x)-T_{f,n,x_0}(x)}{f(x)-P(x)}=\begin{cases}1,& P(x)=T_{f,n,x_0}(x),\\0,& P(x)\ne T_{f,n,x_0}(x).\end{cases}$$

这说明当用多项式作为函数 f 的近似时，相较于其他多项式，Taylor 多项式具有最小的相对误差。

2. 证明对每个 $\mathscr{C}^\infty$ 函数 f，都存在 $\mathscr{C}^\infty$ 函数 g 使得 $f \neq g$，但 f, g 的各阶 Taylor 多项式都相同。

3. 求以下函数在 $x = 0$ 处的 4 阶 Taylor 多项式：
(1) $\arcsin x$；(2) $\sin(1 + \ln x)$； (3) $\mathrm{e}^{\arcsin x}$；
(4) $\arcsin(\mathrm{e}^x - 1)$； (5) $f(x) = \begin{cases} \frac{\mathrm{e}^x - 1}{x}, & x \neq 0; \\ 1, & x = 0. \end{cases}$；
(6) $f(x) = \begin{cases} \frac{x}{\mathrm{e}^x - 1}, & x \neq 0; \\ 1, & x = 0. \end{cases}$

4. 我们知道 $\lim\limits_{n\to+\infty} n^{\frac{1}{n}} = 1$。请估计当 $n \to +\infty$ 时，无穷小量 $n^{\frac{1}{n}} - 1$ 的阶。

5. 利用 $\sqrt{10} = 3\sqrt{1 + \frac{1}{9}}$ 以及 $\sqrt{1 + x}$ 在 $x = 0$ 处的 Taylor 展开计算 $\sqrt{10}$ 的近似值，并与例5.5.14比较。

6. 设 A 是非负实数。证明：函数
$$f(x) = \begin{cases} x + Ax^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$
在 $x = 0$ 的某个邻域中严格增当且仅当 $A < 1$。

7. 求极限 $\lim\limits_{x\to 0} \dfrac{\tan(\sin x) - \sin(\tan x)}{\arctan(\arcsin x) - \arcsin(\arctan x)}$。

8. 设可微函数 f 有可微的反函数，f 在 $x = 0$ 处的 Taylor 展开为 $x + Ax^k + o(x^k)$，证明 f^{-1} 在 $x = 0$ 处的 Taylor 展开为 $x - Ax^k + o(x^k)$。

9. 设 f, g 是 $\mathscr{C}^\infty$ 的奇函数，满足 $f'(0) = g'(0) = 1$。证明 $f(g(x)) - g(f(x)) = O(x^7), x \to 0$。

10. 求参数 c 的取值范围使得对任意实数 x 都有
$$\frac{\mathrm{e}^x + \mathrm{e}^{-x}}{2} \leq \mathrm{e}^{cx^2}.$$

11. 证明对所有非负整数 n 和所有实数 x 都有
$$1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!} > 0.$$

12. 证明对任意 $x \in \mathbb{R}$，
$$\sin x = \lim_{n\to+\infty} \left(x - \frac{x^3}{3!} + \cdots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} \right),$$
$$\cos x = \lim_{n\to+\infty} \left(1 - \frac{x^2}{2!} + \cdots + \frac{(-1)^n x^{2n}}{(2n)!} \right).$$

13. 求参数 c 的范围使得对任意 $-\frac{\pi}{2} \le x \le \frac{\pi}{2}$ 都有

$$\left(\frac{\sin x}{x}\right)^c \ge \cos x.$$

14. **(复指数函数)** 对复数 $z \in \mathbb{C}$ 和正整数 n，记

$$E_n(z) = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^n}{n!}.$$

(1) 证明对任意 $z \in \mathbb{C}$，

$$|E_n(z)| \le E_n(|z|) \le E_{n+1}(|z|);$$

并且对任意 $x > 0$，$E_n(x)$ 关于 x 严格增。
(2) 证明对任意 $M \in \mathbb{N}^*$ 和任意 $n \in \mathbb{N}^*$，

$$|E_n(z)| \le 2E_{4M}(M), \quad \forall |z| \le M.$$

(3) 证明对任意 $M \in \mathbb{N}^*$，存在常数 $C > 0$，使得对任意 $n, m \in \mathbb{N}^*$ 以及任意 $|z| \le M$，

$$|E_{n+m}(z) - E_n(z)| \le \frac{C|z|^{n+1}}{(n+1)!}.$$

从而 $\{E_n(z)\}_{n\ge 1}$ 是 $\mathbb{C}$ 中的一个 Cauchy 列，从而存在极限 $E(z) = \lim_{n\to+\infty} E_n(z)$。

(4) 证明对任意 $M \in \mathbb{N}^*$，存在常数 $C > 0$，使得对任意 $n \in \mathbb{N}^*$ 以及对任意 $|z|, |w| \le M$，

$$|E_{2n}(z+w) - E_n(z)E_n(w)| \le \frac{CM^{n+1}}{(n+1)!}.$$

从而对任意 $z, w \in \mathbb{C}$，$E(z+w) = E(z)E(w)$。
(5) 证明对任意 $z \in \mathbb{C}$，$E'(z) = E(z)$。
(6) 证明对任意 $x \in \mathbb{R}$，$E(x) = \mathrm{e}^x$。
综合以上结论，对 $z \in \mathbb{C}$，可以在 $\mathbb{C}$ 上定义指数函数

$$\mathrm{e}^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots + \frac{z^n}{n!} + \cdots.$$

它满足

$$\mathrm{e}^{z+w} = \mathrm{e}^z \mathrm{e}^w, \quad \mathrm{e}^1 = \mathrm{e}.$$

15. **(三角函数的严格定义)** 对任意实数 x，定义

$$\cos x = \frac{\mathrm{e}^{\mathrm{i}x} + \mathrm{e}^{-\mathrm{i}x}}{2}, \qquad \sin x = \frac{\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}}.$$

(1) 证明

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

(2) 证明对任意 $k \in \mathbb{N}^*$，以及任意 $0 < x < 2$，

$$\frac{x^k}{k!} - \frac{x^{k+2}}{(k+2)!} > 0.$$

(3) 证明 cos 在区间 $[0,2]$ 上严格减，$1 = \cos 0 > 0 > \cos 2$。从而 cos 在区间 $[0,2]$ 上有唯一零点 $x_0 \in (0,2)$。记 $\pi = 2x_0$

(4) 证明对任意 $x, y \in \mathbb{R}$，

$$\cos(x+y) = \cos x \cos y - \sin x \sin y,$$
$$\sin(x+y) = \sin x \cos y + \cos x \sin y.$$

从而

$$\cos^2 x + \sin^2 x = 1.$$

5.7 应用：用 Taylor 公式修正圆周率近似值

记单位圆(即半径长为1的圆)内接正 n 边形的边长为 a_n ，周长为 $2\pi_n$。则利用勾股定理可得

$$a_{2n} = \sqrt{\left(1 - \sqrt{1 - \frac{a_n^2}{4}}\right)^2 + \frac{a_n^2}{4}} = \sqrt{2 - \sqrt{4 - a_n^2}} = \frac{a_n}{\sqrt{2 + \sqrt{4 - a_n^2}}}.$$

因为 $\pi_n = \frac{1}{2} n a_n$，所以

$$\pi_{2n} = n a_{2n} = \frac{n a_n}{\sqrt{2 + \sqrt{4 - a_n^2}}} = \frac{\pi_n}{\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{\pi_n^2}{n^2}}}}.$$

利用 $\pi_3 = 3$ (单位圆内接正六边形半周长)和上述递推关系，可以得到单位圆内接正 $2^n \times 3$ 边形半周长。

另一方面，因为 $\pi_n = n \sin \frac{\pi}{n}$。所以 $\lim\limits_{n \to +\infty} \pi_n = \pi$。 我们可以利用 Taylor 展开对上述结果进行改进。由

$$\pi_n = n \sin \frac{\pi}{n} = \pi - \frac{\pi^3}{6n^2} + \frac{\pi^5}{120 n^4} + O\left(\frac{1}{n^6}\right),$$
$$\pi_{2n} = 2n \sin \frac{\pi}{2n} = \pi - \frac{\pi^3}{24 n^2} + \frac{\pi^5}{120 \times 16 n^4} + O\left(\frac{1}{n^6}\right),$$

表 5.1: 利用正多边形周长逼近圆周率

边数 n	π_n	$\hat{\pi}_n$	π_n^*
6	3		
12	3.1058285412	3.1411047216	
24	3.1326286133	3.1415619706	3.1415924539
48	3.1393502030	3.1415907330	3.1415926505
96	3.1410319509	3.1415925335	3.1415926535
192	3.1414524723	3.1415926461	3.1415926536
384	3.1415576079	3.1415926531	3.1415926536
768	3.1415838921	3.1415926536	3.1415926536
1536	3.1415904632	3.1415926536	3.1415926536
$\vdots$	$\vdots$	$\vdots$	$\vdots$
196608	3.1415926535	3.1415926536	3.1415926536

利用线性组合消去其中的 $\frac{1}{n^2}$ 项，得到

$$\hat{\pi}_{2n} = \frac{4\pi_{2n} - \pi_n}{3} = \pi - \frac{\pi^5}{480n^4} + O\left(\frac{1}{n^6}\right).$$

进一步用类似办法得到

$$\pi_{4n}^* = \frac{16\hat{\pi}_{4n} - \hat{\pi}_{2n}}{15} = \pi + O\left(\frac{1}{n^6}\right).$$

从上表中我们可以看到，由正 $6, 12, 24, 48, 96$ 边形的周长进过上述修正得的圆周率的近似值 π_{96}^*，就已经达到了正 196608 边形周长给出的结果。这显示出这个修正算法的强大威力。

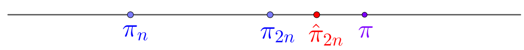


图 5.3: 外推法

$$\pi_{2n} = \frac{\pi_n}{\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{\pi_n^2}{n^2}}}} > \pi_n.$$

所以 $\pi_6, \pi_{12}, \pi_{24}, \pi_{48}, \ldots$ 是一个单调增数列。上述修正数列

$$\hat{\pi}_{2n} = \pi_{2n} + \frac{1}{3}(\pi_{2n} - \pi_n)$$

提供了一个**外推**(extrapolation)修正，它利用已有数据经简单操作就获得了更高精度的近似值。另外，因为 sin 在区间 $[0,\pi]$ 上是凹函数，所以

$$\hat{\pi}_{4n}-\hat{\pi}_{2n}=\frac{4\pi_{4n}-5\pi_{2n}+\pi_n}{3}=\frac{10n}{3}\left(\frac{8}{5}\sin\frac{\pi}{4n}+\frac{1}{10}\sin\frac{\pi}{n}-\sin\frac{\pi}{2n}\right)>0,$$

因此数列 $\hat{\pi}_{12},\hat{\pi}_{24},\hat{\pi}_{48},\ldots$ 也是单调增的，所以 $\hat{\pi}_{2n}$ 总是比 π_{2n} 更好的近似值，而

$$\pi^*_{2n}=\hat{\pi}_{2n}+\frac{1}{15}(\hat{\pi}_{2n}-\hat{\pi}_n)$$

是进一步的外推修正。

习题5.7

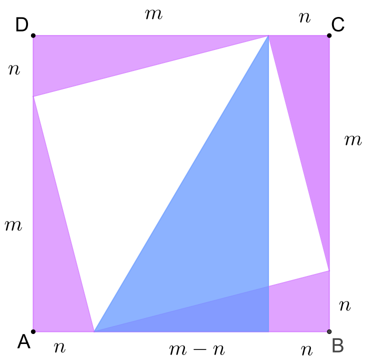
1. 仿照正文中的做法，利用圆外切正多边形周长构造圆周率的近似值。

2. 设 $x_{n+1}=\sin x_n$。请判断 x_n 是否有形如 $x_n=\frac{A}{n^\alpha}+o\left(\frac{1}{n^\alpha}\right)$ 的渐进展开？如果有，请给出 A,α 的值；如果没有，请说明理由。

3. 如图所示，$ABCD$是一个正方形。证明

$$\arctan\frac{n}{m}+\arctan\frac{m-n}{m+n}=\frac{\pi}{4}.$$

   并利用它计算 π 的十进制小数表示的近似值，要求精确到小数点后面前 20 位小数。



习题讨论课4

1. Darboux定理。设 f 在区间 (a,b) 内可导，且 $f'_+(a), f'_-(b)$ 存在，$f'_+(a) < f'_-(b)$。则对任意 $c \in (f'_+(a) < f'_-(b))$，存在 $\xi \in (a,b)$ 使得 $f'(\xi) = c$。

2. 设 $\alpha \in \mathbb{R}$ 为参数，讨论方程 $\mathrm{e}^x - x^\alpha = 0$ 的解的个数。

3. 求函数 $f(x) = \frac{x(2+\cos x)}{\sin x}$ 的值域。

4. 证明对任意 $0 < x < \frac{\pi}{2}$，$2\sin x + \tan x > 3x$。

5. 常系数线性微分方程与常数变易法。

   (a) $y'' - 3y' + 2y = 0$，

   (b) $y'' - 2y' + y = 0$，

   (c) $y'' + y = 0$。

6. 设 f 在 $[a,b]$ 上一阶可导，在 (a,b) 内二阶可导，$f(a) = f(b) = 0$，$f'(a)f'(b) > 0$。证明存在 $\zeta \in (a,b)$ 使得 $f''(\zeta) = f(\zeta)$。

7. 讨论 $f(x) = \left(1 + \frac{1}{x}\right)^{x+\alpha}$ 在区间 $(0, +\infty)$ 上的单调性，其中 $\alpha \in \mathbb{R}$ 是参数。

8. 讨论方程 $\ln\left(1 + \frac{1}{x}\right) = \frac{x+\alpha}{x(x+1)}$ 的解的存在性和唯一性，以及当方程有唯一解时，解关于参数 $\alpha \in \mathbb{R}$ 的可微性。

9. 设 $p \geq 1$。证明对任意实数 $x_1, \ldots, x_n, y_1, \ldots, y_n$，都成立
$$\left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p\right)^{\frac{1}{p}} \geq \left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{p}}.$$

10. 求当 $n \to +\infty$ 时，$\sqrt[n]{n}$ 的渐进展开式。

11. 考虑方程 $xy + \mathrm{e}^x - \mathrm{e}^y = 0$。

    (a) 证明对于任意 $x \in \mathbb{R}$，上述方程关于 y 在区间 $[x, +\infty)$ 中有唯一解 $y = f(x)$。

    (b) 证明 $f : \mathbb{R} \to \mathbb{R}$ 是 $\mathscr{C}^\infty$ 函数。

    (c) 求 $f(x)$ 在 $x = 0$ 处的四阶Taylor展开式。

    (d) 讨论 f 的单调性，估计 f 的值域。

    (e) 讨论当 $x \to -\infty$ 和 $x \to +\infty$ 时的渐近线。

12. 考察方程 $\mathrm{e}^{x+y} - \sin(x+y) = x^3$。

(a) 证明对于任意 $x \geq 1$，上述方程关于 y 在区间 $[-x, +\infty)$ 中有唯一解 $y = f(x)$。

(b) 证明 $f : (1, +\infty) \to \mathbb{R}$ 是 $\mathscr{C}^\infty$ 函数。

(c) 证明 $f(x) = -x + 3\ln x + o(\ln x), x \to +\infty$。

(d) 求 $f(x) = -x + \sqrt{3\ln x} + o(\ln x), x \to 1^+$。

13. 讨论函数 $f(x) = \frac{2^x - a}{x}$ 的图像随参数 a 的变化。

第 6 章　不定积分

6.1　原函数与不定积分

定义 6.1.1. 称函数 F 是函数 f（或者更确切地，微分 $f(x)\mathrm{d}x$）在 $I \subseteq \mathbb{R}$ 上的一个**原函数**，如果对任意 $x \in I$，$\mathrm{d}F_x = f(x)\mathrm{d}x$。在不特别声明的情况下，习惯上认为 I 是 f 的定义域。

记 $\int f(x)\mathrm{d}x$ 为 f 在 I 上的所有原函数组成的集合，称为 f 在 I 上的**不定积分**，称 f 为这个不定积分中的**被积函数**。

原函数与不定积分对应于最简单的微分方程

$$y'(x) = f(x).$$

f 的原函数 F 就是这个微分方程的一个解。这个微分方程的所有的解组成的集合就是 f 的不定积分。

这个微分方程在 (x, y) 坐标平面的每个点 (x, y) 处都指定了一个斜率值 $f(x)$，或者说指定了过点 (x, y) 且斜率值为 $f(x)$ 的直线，因此几何上这个微分方程对应于 (x, y) 坐标平面的一个**斜率场**或**方向场**。$y(x)$ 是这个微分方程的解，当且仅当由 $y = y(x)$ 所确定的曲线在其经过的每个点处都与该点处的方向场(即过该点且斜率为 $f(x)$ 的直线)相切。此时称曲线 $y = y(x)$ 为这个方向场的**积分曲线**。

沿平行于 y 轴的方向移动时，x 保持恒定，方向场 $f(x)$ 不变，这是这个方向场的一种对称性。由于这种对称性，该方向场的任何一条积分曲线沿平行于 y 轴的方向移动所得到的曲线始终与该方向场相切，因此也是这个方向场的积分曲线。这就是下述定理的第二个结论的几何意义。后面我们会进一步看到，微分方程或方向场的对称性对微分方程的解或方向场的积分曲线有着特别重要的意义。

定理 6.1.2. *(1) 如果 I 是区间，$f : I \to \mathbb{R}$ 连续，则 f 在 I 上有原函数。进一步，对任意 $x_0 \in I$ 及任意 $y_0 \in \mathbb{R}$，f 有唯一的原函数 F 满足 $F(x_0) = y_0$。*

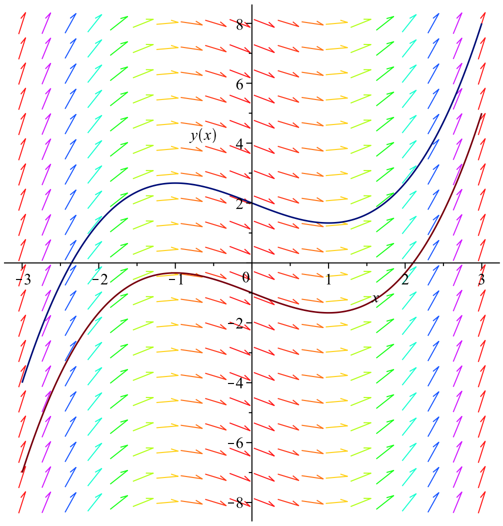


图 6.1: 方向场 x^2-1 和它的积分曲线

(2) 如果 I 是区间，F,G 都是 f 在 I 上的原函数，则 $F-G$ 是常值函数。从而

$$\int f(x)\mathrm{d}x=\left\{F+C|C\in\mathbb{R}\right\}.$$

习惯上写成

$$\int f(x)\mathrm{d}x=F(x)+C,\qquad C\text{是任意常数}.$$

例 6.1.3. 在 $\mathbb{R}$ 上，

$$\int \mathrm{e}^x\mathrm{d}x=\mathrm{e}^x+C,$$

$$\int \cos x\mathrm{d}x=\sin x+C,$$

$$\int \sin x\mathrm{d}x=-\cos x+C,$$

$$\int x^n\mathrm{d}x=\frac{x^{n+1}}{n+1}+C,\qquad n\in\mathbb{N},$$

$$\int \frac{1}{1+x^2}\mathrm{d}x=\arctan x+C.$$

在 $(0,+\infty)$ 上，

$$\int x^{\alpha}\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1} + C, \qquad \alpha \in \mathbb{R}\backslash\{-1\}.$$

在 $\mathbb{R}\backslash\{0\}$ 上，

$$\int \frac{1}{x}\mathrm{d}x = \begin{cases} \ln x + C_1, & x > 0; \\ \ln(-x) + C_2, & x < 0. \end{cases}$$

$$\int \frac{1}{x^n}\mathrm{d}x = \begin{cases} \frac{1}{(n-1)x^{n-1}} + C_1, & x > 0; \\ \frac{1}{(n-1)x^{n-1}} + C_2, & x < 0. \end{cases} \qquad n \in \mathbb{N}^*, n > 1.$$

在 $(-1,1)$ 上，

$$\int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \arcsin x + C.$$

在 $\left(k\pi - \dfrac{\pi}{2}, k\pi + \dfrac{\pi}{2}\right)$ 上，

$$\int \frac{1}{\cos^2 x}\mathrm{d}x = \tan x + C.$$

在 $(k\pi, k\pi+\pi)$ 上，

$$\int \frac{1}{\sin^2 x}\mathrm{d}x = -\cot x + C.$$

例 6.1.4.

$$F(x) = \begin{cases} |x|^{\frac{4}{3}} \sin \dfrac{1}{x}, & x \neq 0; \\ 0, & x = 0 \end{cases}$$

是

$$f(x) = \begin{cases} \frac{4}{3}x^{\frac{1}{3}} \sin \dfrac{1}{x} - |x|^{-\frac{2}{3}} \cos \dfrac{1}{x}, & x \neq 0; \\ 0, & x = 0 \end{cases}$$

的原函数，但 f 不仅在 $x=0$ 不连续，甚至在 $x=0$ 附近无界。

6.2 不定积分的运算性质

(1) **线性**

定理 6.2.1. 设 F, G 分别是 f, g 在 I 上的原函数，则 $\lambda F + \mu G$ 是 $\lambda f + \mu g$ 的原函数。

$$\int f(x) + g(x)\mathrm{d}x = \int f(x)\mathrm{d}x + \int g(x)\mathrm{d}x,$$

$$\int \lambda f(x)\mathrm{d}x = \lambda \int f(x)\mathrm{d}x.$$

注 6.2.2. 等式

$$\int f(x)+g(x)\mathrm{d}x=\int f(x)\mathrm{d}x+\int g(x)\mathrm{d}x$$

含义是：对 f 的任一原函数 F，g 的任一原函数 $G\in B$，以及 $f+g$ 的任一原函数 H，$H-(F+G)$ 是 0 的一个原函数，即 $(H-(F+G))'(x)=0$，也即 $H'=F'+G'$。如果 I 是区间，则这等价于：存在常数 C 使得 $H(x)=F(x)+G(x)+C$。

等式

$$\int \lambda f(x)\mathrm{d}x=\lambda\int f(x)\mathrm{d}x$$

的含义是：对 f 的任一原函数 F，λf 的任一原函数 H，$H-\lambda F$ 是 0 的一个原函数，即 $H'=\lambda F'$。如果 I 是区间，则这等价于：存在常数 C 使得 $H(x)=\lambda F(x)+C$。

例 6.2.3. 多项式积分 $\int a_0+a_1x+\cdots+a_nx^n\mathrm{d}x=C+a_0x+\frac{a_1x^2}{2}+\cdots+\frac{a_nx^{n+1}}{n+1}$。

例 6.2.4. 求 $\int\cos^2 x\mathrm{d}x$。

解： 用二倍角公式降次

$$\int\cos^2 x\mathrm{d}x=\int\frac{1+\cos 2x}{2}\mathrm{d}x=C+\frac{x}{2}+\frac{\sin 2x}{4}.$$

□

例 6.2.5. 求 $\int\frac{1}{x^2-1}\mathrm{d}x$。

解：

$$\begin{aligned}\int\frac{1}{x^2-1}\mathrm{d}x&=\int\frac{1}{2}\frac{1}{x-1}-\frac{1}{2}\frac{1}{x+1}\mathrm{d}x\\&=\frac{1}{2}\int\frac{1}{x-1}\mathrm{d}x-\frac{1}{2}\int\frac{1}{x+1}\mathrm{d}x\\&=\frac{\ln|x-1|}{2}-\frac{\ln|x+1|}{2}+C\end{aligned}$$

其中 C 是 $(-\infty,-1)\cup(-1,1)\cup(1,+\infty)$ 上的分段常值函数。 □

一般地，任何一个有理分式$\frac{P(x)}{Q(x)}$（其中 P,Q 是关于 x 的多项式）可以分解为一个多项式和有限多个以下形式的最简分式的线性组合

$$\frac{1}{(x-a)^k},\quad\frac{(x-p)}{[(x-p)^2+q^2]^m},\quad\frac{1}{[(x-p)^2+q^2]^m}.$$

所以有理分式的积分归结为这些最简分式的积分。前两类最简分式很容易找到相应的原函数，对第三类最简分式需要进行换元操作。

(2) **换元**

定理 6.2.6. *(1)* 设 F 是 f 在 I 上的原函数，$g:J\to I$ 是可微函数，则 $F(g(x))$ 是 $f(g(x))g'(x)$ 在 J 的原函数。

$$\int f(g(x))g'(x)\mathrm{d}x=\int f(u)\mathrm{d}u\Big|_{u=g(x)}=F(g(x))+C.$$

作为特例，

$$\int f(ax+b)\mathrm{d}x=\frac{1}{a}F(ax+b)+C.$$

(2) 设 h 是一一对应，有可微的反函数 h^{-1}，G 是 $f(h(u))h'(u)$ 的原函数，则

$$\int f(x)\mathrm{d}x=\int f(h(u))h'(u)\mathrm{d}u\Big|_{u=h^{-1}(x)}=G(h^{-1}(x))+C.$$

注 6.2.7. 上述定理(1)是把被积函数凑成一个函数的导数，形式上

$$f(g(x))g'(x)\mathrm{d}x=F'(u)\mathrm{d}u=\mathrm{d}F(u)=\mathrm{d}\big(F(g(x))\big),$$

所以这个换元方法也称为“凑微分法”。相当于引入中间变量 $u=g(x)$。

定理(2)是主动寻求换元，$x=h(u)$，变成以 u 为自变量的函数计算不定积分，但要把最终结果通过$u=h^{-1}(x)$ 变回以 x 为自变量的函数。

无论哪一种换元，目的都是为了简化不定积分中的函数形式，使不定积分更容易计算。换元法是一个很灵活的办法，需要积累一定的经验。

例 6.2.8. 求 $\int x\mathrm{e}^{-x^2}\mathrm{d}x$。

解： 凑微分

$$\begin{aligned}\int x\mathrm{e}^{-x^2}\mathrm{d}x&=\int \mathrm{e}^{-x^2}\mathrm{d}\frac{x^2}{2}\\&=\frac{1}{2}\int \mathrm{e}^{-u}\mathrm{d}u,\qquad u=x^2\\&=-\frac{1}{2}\mathrm{e}^{-u}+C\\&=-\frac{1}{2}\mathrm{e}^{-x^2}+C.\end{aligned}$$

□

例 6.2.9. 求 $\int\cos^3 x\mathrm{d}x$。

解： 凑微分

$$\int\cos^3 x\mathrm{d}x=\int(1-\sin^2 x)\mathrm{d}\sin x=C+\sin x-\frac{\sin^3 x}{3}.$$

一般地，

$$\int f(\sin x)\cos x\mathrm{d}x=F(\sin x)+C,$$

其中 F 是 f 的一个原函数。 □

例 6.2.10. 求 $\int\sqrt{1-x^2}\mathrm{d}x$。

解： 主动换元 $x=\sin\theta$($\theta\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$)，

$$\begin{aligned}\int\sqrt{1-x^2}\mathrm{d}x&=\int\cos\theta\mathrm{d}\sin\theta\\&=\int\cos^2\theta\mathrm{d}\theta=\frac{\sin 2\theta}{4}+\frac{\theta}{2}+C\\&=\frac{x\sqrt{1-x^2}}{2}+\frac{\arcsin x}{2}+C.\end{aligned}$$

这里三角换元 $x=\sin\theta$ 是为了处理被积函数 $\sqrt{1-x^2}$ 中的开方。 □

如果被积函数中出现 $\sqrt{1+x^2}$，可以考虑 $x=\tan\theta,\quad\theta=\arctan x$。

如果被积函数中出现 $\sqrt{x^2-1}$，可以考虑 $x=\frac{1}{\cos\theta},\quad\theta=\arccos(1/x)$。

(3) **分部积分**

定理 6.2.11. 设 f,g 是 $\mathscr{C}^1$ 函数。则

$$\int f(x)g'(x)\mathrm{d}x=f(x)g(x)-\int f'(x)g(x)\mathrm{d}x.$$

通常也把它写成更方便记忆的形式：

$$\int f(x)\mathrm{d}g(x)=f(x)g(x)-\int g(x)\mathrm{d}f(x).$$

这个性质源于乘积函数的导数公式。分部积分的目的是通过求导数简化被积函数中一部分因子$f(x)$ 的形式。

例 6.2.12. 求 $\int x^\alpha\ln x\mathrm{d}x$。

解： 当 $\alpha\neq-1$ 时，

$$\begin{aligned}\int x^\alpha\ln x\mathrm{d}x&=\int\ln x\mathrm{d}\frac{x^{\alpha+1}}{\alpha+1}\\&=\frac{x^{\alpha+1}\ln x}{\alpha+1}-\int\frac{x^{\alpha+1}}{\alpha+1}\mathrm{d}\ln x\\&=\frac{x^{\alpha+1}\ln x}{\alpha+1}-\int\frac{x^{\alpha}}{\alpha+1}\mathrm{d}x\\&=\frac{x^{\alpha+1}\ln x}{\alpha+1}-\frac{x^{\alpha+1}}{(\alpha+1)^2}+C.\end{aligned}$$

当 $\alpha\neq-1$ 时，

$$\int x^{-1}\ln x\mathrm{d}x=\int\ln x\mathrm{d}\ln x=\frac{(\ln x)^2}{2}+C.$$

□

例 6.2.13. 求 $\int\cos^n x\mathrm{d}x$。

解：

$$\begin{aligned}\int \cos^n x\mathrm{d}x &= \int \cos^{n-1} x\mathrm{d}\sin x\\ &= \cos^{n-1} x\sin x - \int \sin x\mathrm{d}\cos^{n-1} x\\ &= \cos^{n-1} x\sin x + (n-1)\int \sin^2 x\cos^{n-2} x\mathrm{d}x\\ &= \cos^{n-1} x\sin x + (n-1)\int (1-\cos^2 x)\cos^{n-2} x\mathrm{d}x\\ &= \cos^{n-1} x\sin x + (n-1)\int \cos^{n-2} x\mathrm{d}x - (n-1)\int \cos^n x\mathrm{d}x.\end{aligned}$$

所以

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

从而得到递推关系。而

$$\begin{aligned}&\int \cos x\mathrm{d}x = \sin x + C,\\ &\int \cos^2 x\mathrm{d}x = \int \frac{\cos 2x+1}{2}\mathrm{d}x = \frac{\sin 2x}{4} + \frac{x}{2} + C.\end{aligned}$$

□

另解：

$$\int \cos^{2n+1} x\mathrm{d}x = \int (1-\sin^2 x)^n\mathrm{d}\sin x = C + \sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^k\sin^{2k+1} x}{2k+1},$$

$$\int \cos^{2n} x\mathrm{d}x = \int \left(\frac{1+\cos 2x}{2}\right)^n\mathrm{d}x = \frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}\int \cos^k(2x)\mathrm{d}x.$$

其中后者进入递归过程。 □

换元、线性和分部积分，提供了把复杂函数不定积分计算朝更简单的方向转化的可能，但具体实现需要多加尝试以及积累经验。另外不是所有原函数都可以写成初等函数形式[1]，比如 $\int \mathrm{e}^{x^2}\mathrm{d}x, \int \frac{\mathrm{e}^x}{x}\mathrm{d}x, \int \frac{\sin x}{x}\mathrm{d}x, \int \frac{1}{\ln x}\mathrm{d}x$ 等。

习题6.2

1. $\int x^n\mathrm{e}^x\mathrm{d}x$，$\int x^n\sin x\mathrm{d}x$，$\int \mathrm{e}^{\lambda x}\sin x\mathrm{d}x$，$\int \sin\alpha x\cos\beta x\mathrm{d}x$，$\int x^n\arctan x\mathrm{d}x$，$\int x^{2n+1}\operatorname{arcsin} x\mathrm{d}x$

2. $\int \tan x\mathrm{d}x$。

[1]关于不定积分是否可以写成初等函数，属于微分代数理论研究的内容，Liouville 在这方面做出了最早的重要发现，一个简要的介绍可以参考 R.C.Churchill的讲义 Liouville’s Theorem on Integration in Terms of Elementary Functions，该讲义可在以下网址获取，http://math.hunter.cuny.edu/ksda/papers/rick_09_02.pdf

6.3 有理函数的不定积分以及可转化为有理函数的不定积分

有理函数的不定积分

所谓有理函数就是形如

$$\frac{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0}$$

的函数。我们称形如

$$\frac{1}{(x-a)^k}, \quad \frac{ax+b}{(x^2+px+q)^k}, \qquad \left(k \in \mathbb{N}^*, p^2 < 4q\right)$$

的有理函数为**最简分式**。

定理 6.3.1. 任何实数系数有理函数都可表示为一个多项式与有限多个最简分式的线性组合。

例 6.3.2.

$$\int \frac{1}{(x-a)^k}\mathrm{d}x = \begin{cases} \dfrac{1}{(1-k)(x-a)^{k-1}} + C, & k > 1, \\ \ln|x-a| + C, & k = 1. \end{cases}$$

例 6.3.3.

$$\int \frac{x}{(x^2+1)^k}\mathrm{d}x = \int \frac{1}{2(x^2+1)^k}\mathrm{d}(x^2+1) = \begin{cases} \dfrac{1}{2(1-k)(x^2+1)^{k-1}} + C, & k > 1, \\ \dfrac{1}{2}\ln(x^2+1) + C, & k = 1. \end{cases}$$

例 6.3.4.

$$\begin{aligned} I_k &= \int \frac{1}{(x^2+1)^k}\mathrm{d}x \\ &= \int \frac{1+x^2}{(x^2+1)^k}\mathrm{d}x - \int \frac{x^2}{(x^2+1)^k}\mathrm{d}x \\ &= I_{k-1} + \frac{1}{2(k-1)}\int x\mathrm{d}\frac{1}{(x^2+1)^{k-1}} \\ &= \frac{1}{2k-2}\frac{x}{(x^2+1)^{k-1}} + \frac{2k-3}{2k-2}I_{k-1}. \end{aligned}$$

最终是 $\arctan x, \frac{x}{x^2+1}, \dots, \frac{x}{(x^2+1)^{k-1}}$ 的线性组合。

对一般情形$\dfrac{ax+b}{(x^2+px+q)^k}$，可以通过换元

$$x = -\frac{p}{2} + \sqrt{q - \frac{p^2}{4}}y, \quad \tilde{a} = \frac{a}{\left(q - \frac{p}{4}\right)^{k-1}}, \quad \tilde{b} = \frac{b - \frac{pa}{2}}{\left(q - \frac{p}{4}\right)^{k}},$$

得到

$$\frac{ax+b}{(x^2+px+q)^k}=\frac{a\left(x+\dfrac{p}{2}\right)+b-\dfrac{pa}{2}}{\left(\left(x+\dfrac{p}{2}\right)^2+q-\dfrac{p^2}{4}\right)^k}=\frac{\tilde{a}y+\tilde{b}}{(y^2+1)^k}.$$

推论 6.3.5. *任何实数系数有理函数的不定积分都是初等函数。* □

因此计算有理函数的不定积分，关键是把有理函数分解为最简分式的线性组合。我们仅举一例说明实现这个分解的算法，定理6.3.1也是循着这个办法证明的。这个办法的第一步是要把有理函数的分母多项式写成一次或二次多项式的乘积，而这是由代数学基本定理**??**保证的。

例 6.3.6. 把 $f(x)=\dfrac{2x^2+2x+13}{(x-2)^2(x^2+1)^2}$ 分解为最简分式的线性组合，并求它的不定积分。

解： 设

$$\frac{2x^2+2x+13}{(x-2)^2(x^2+1)^2}=\frac{A}{(x-2)^2}+\frac{Bx+C}{(x^2+1)^2}+\frac{P_1(x)}{(x-2)(x^2+1)}.$$

两边乘以 $(x-2)^2(x^2+1)^2$ 得

$$2x^2+2x+13=A(x^2+1)^2+(Bx+C)(x-2)^2+P_1(x)(x^2+1)(x-2).$$

取 $x=2$，得 $A=1$；取 $x=\mathrm{i}$，得 $B=2,C=1$。

在第一个等式中代入 $A=1,B=2,C=1$，并化简得到

$$\frac{P_1(x)}{(x-2)(x^2+1)}=\frac{-x-4}{(x-2)(x^2+1)},$$

再用上述方法得到

$$\frac{-x-4}{(x-2)(x^2+1)}=\frac{-\frac{6}{5}}{x-2}+\frac{\frac{6}{5}x+\frac{7}{5}}{x^2+1}.$$

从而

$$f(x)=\frac{1}{(x-2)^2}+\frac{2x+1}{(x^2+1)^2}-\frac{6}{5}\cdot\frac{1}{x-2}+\frac{1}{5}\cdot\frac{6x+7}{x^2+1}.$$

设

$$\int f(x)\mathrm{d}x=\frac{E}{x-2}+F\ln|x-2|+\frac{Gx+H}{x^2+1}+K\ln(x^2+1)+L\arctan x+C.$$

两边求导得到

$$f(x)=-\frac{E}{(x-2)^2}+\frac{F}{x-2}+\frac{-2Hx+2G}{(x^2+1)^2}+\frac{2Kx+L-G}{x^2+1},$$

比较系数得到

$$E=-1,\quad F=-\frac{6}{5},\quad K=\frac{3}{5},\quad H=-1,\quad G=\frac{1}{2},\quad L=\frac{19}{10},$$

从而

$$\int f(x)\mathrm{d}x = \frac{-1}{x-2} - \frac{6}{5}\ln|x-2| + \frac{\frac{1}{2}x-1}{x^2+1} + \frac{3}{5}\ln(x^2+1) + \frac{19}{10}\arctan x + C.$$

□

三角有理分式

形如

$$\int R(\cos\theta,\sin\theta)\mathrm{d}\theta$$

其中 $R(x,y)$ 是关于变量 x,y 的有理分式，即 $R(x,y)=\dfrac{P(x,y)}{Q(x,y)}$，其中 $P(x,y)$ 和 $Q(x,y)$ 是关于变量 x,y 的多项式。

可以采用万能公式

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

得到

$$\int R\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right)\mathrm{d}2\arctan t=\int\frac{2}{1+t^2}R\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right)\mathrm{d}t,$$

对

$$\int R(\tan\theta)\mathrm{d}\theta$$

可以采用 $t=\tan\theta$，

$$\int R(\tan\theta)\mathrm{d}\theta=\int\frac{R(t)}{1+t^2}\mathrm{d}t.$$

某些根式

某些特殊的根式的不定积分可以通过适当的换元转化为有理函数或三角有理分式的不定积分，例如

$$\int R(x,\sqrt{1-x^2})\mathrm{d}x=\int R(\cos\theta,\sin\theta)(-\sin\theta)\mathrm{d}\theta,\qquad \theta=\arccos x.$$

$$\int R(x,\sqrt{1+x^2})\mathrm{d}x=\int R\left(\frac{t^2-1}{2t},\frac{t^2+1}{2t}\right)\frac{t^2+1}{2t^2}\mathrm{d}t,\quad x=\frac{t^2-1}{2t}(t>0).$$

$$\int R(x,\sqrt{x^2-1})\mathrm{d}x=\int R\left(\frac{t^2+1}{2t},\frac{t^2-1}{2t}\right)\frac{t^2-1}{2t^2}\mathrm{d}t,$$

$$x=\frac{t^2+1}{2t}(t>1\text{ 或 }-1<t<0).$$

$$\int R\left(x,\sqrt[n]{\frac{px+q}{rx+s}}\right)\mathrm{d}x=\int R\left(\frac{st^n-q}{p-rt^n},t\right)\mathrm{d}\left(\frac{st^n-q}{p-rt^n}\right),\quad t=\sqrt[n]{\frac{px+q}{rx+s}}.$$

以上情形可以统一为：对平面上的有理曲线 $\gamma:(x(t),y(t))$，其中 $x(t),y(t)$ 都是关于参数 t 的有理函数，则对有理函数 $R(x,y)$，沿着曲线 γ，

$$\int R(x,y)\mathrm{d}x=\int R(x(t),y(t))\mathrm{d}x(t)$$

是关于变量 t 的有理函数的积分。

圆锥曲线都是有理曲线：

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,\quad \begin{cases}\frac{x}{a}=\frac{1-t^2}{1+t^2},\\ \frac{y}{b}=\frac{2t}{1+t^2}.\end{cases}$$

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1,\quad \begin{cases}\frac{x}{a}=\frac{t^2+1}{2t},\\ \frac{y}{b}=\frac{t^2-1}{2t}.\end{cases}$$

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=-1,\quad \begin{cases}\frac{x}{a}=\frac{t^2-1}{2t},\\ \frac{y}{b}=\frac{t^2+1}{2t}.\end{cases}$$

$$y=ax^2,\quad \begin{cases}x=t,\\ y=at^2.\end{cases}$$

$(rx+s)y^n=px+q$是有理曲线，

$$\begin{cases}x=\frac{q-st^n}{rt^n-p},\\ y=t.\end{cases}$$

$x^3+y^3=3axy$ 是有理曲线：

$$\begin{cases}x=\frac{3at}{1+t^3},\\ y=\frac{3at^2}{1+t^3}.\end{cases}$$

习题6.3

1. 关于最简分式分解[2]。

   (1) 设 $P(x),Q_1(x)$ 是实系数多项式，$Q(x)=(x-\alpha)^kQ_1(x)$，$Q_1(\alpha)\neq 0$。证明：存在实数 A 以及实系数多项式 $P_1(x)$ 使得

$$\frac{P(x)}{(x-\alpha)^kQ_1(x)}=\frac{A}{(x-\alpha)^k}+\frac{P_1(x)}{(x-\alpha)^{k-1}Q_1(x)}.$$

   (2) 设 $P(x),Q_1(x)$ 是实系数多项式，$Q(x)=(x^2-px+q)^kQ_1(x)$ (其中 p,q 为实数满足$p^2<4q$)，$Q_1(\alpha)\neq 0$ (其中 α 是 x^2-px+q 的复数根)。证明：存在实数 A,B 以及实系数多项式 $P_1(x)$ 使得

$$\frac{P(x)}{(x^2-px+q)^kQ_1(x)}=\frac{Ax+B}{(x^2-px+q)^k}+\frac{P_1(x)}{(x^2-px+q)^{k-1}Q_1(x)}.$$

[2]这个方法属于Euler，见《无穷分析引论(上)》第二章

2. 关于代数曲线的有理参数化[3]。

 (1) 任何二次曲线 $Ax^2+Bxy+Cy^2+Dx+Ey+F=0$ 都具有有理函数形式的参数方程。具体方法是选取该曲线上一点 $P_0(x_0,y_0)$，向曲线上另一点 $P(x,y)$ 引射线，以射线 P_0P 的斜率 t 为参数，证明：坐标 x,y 是关于 t 的有理函数。

 (2) 按(1)中方法写出圆 $x^2+y^2=1$ 的参数方程，选 P_0 为 $(-1,0)$。

 (3) 按(1)中方法写出双曲线 $x^2-y^2=1$ 的参数方程，选 P_0 为 $(-1,0)$。

 (4) 按(1)中方法写出抛物线 $y=x^2$ 的参数方程，选 P_0 为 $(0,0)$。

 (5) 按(1)中方法写出三次曲线

$$Ax^3+Bx^2y+Cxy^2+Dy^3+Ex^2+Fxy+Gy^2=0$$

 的有理函数形式的参数方程。

6.4 线性微分方程和常数变易法

一阶线性微分方程

$$\frac{\mathrm{d}y}{\mathrm{d}x}=a(x)y+f(x)$$

是一阶线性微分方程，当 $f(x)$恒为零时，称为**齐次**线性方程，当 $f(x)$ 不恒为零时，称为**非齐次**线性方程。

设 $A(x)$ 为 $a(x)$ 的一个原函数。则

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{A(x)}=\mathrm{e}^{A(x)}A'(x)=\mathrm{e}^{A(x)}a(x),$$

即 $y_0(x)=\mathrm{e}^{A(x)}$ 是齐次方程

$$\frac{\mathrm{d}y}{\mathrm{d}x}=a(x)y$$

的一个非零解，对任意常数 C，$Cy_0(x)$ 也是齐次方程的解。

我们利用齐次方程的一个非零解 $y_0(x)$ 和下面的**常数变易法**求解非齐次方程。设 $y(x)$ 是非齐次方程的解。令 $C(x)=\mathrm{e}^{-A(x)}y(x)$。则 $y(x)=\mathrm{e}^{A(x)}C(x)$(形式上这是把齐次方程的解 $Cy_0(x)$ 中的常数 C换成函数 $C(x)$，这就是“常数变易”这个名字的含义)，代入非齐次方程得到

$$\mathrm{e}^{A(x)}C'(x)+\mathrm{e}^{A(x)}a(x)C(x)=a(x)\mathrm{e}^{A(x)}C(x)+f(x),$$

即

$$C'(x)=\mathrm{e}^{-A(x)}f(x),$$

因此

$$C(x)=\int \mathrm{e}^{-A(x)}f(x)\mathrm{d}x.$$

[3]这个方法属于Euler，见《无穷分析引论(上)》第三章

从而

$$y(x) = \mathrm{e}^{A(x)} \int \mathrm{e}^{-A(x)} f(x) \mathrm{d}x.$$

注意，上式等号右端积分前面的 $\mathrm{e}^{A(x)}$ 不能消去被积函数中的 $\mathrm{e}^{-A(x)}$。所以更准确的结果是，取 $\mathrm{e}^{-A(x)} f(x)$ 的一个原函数 $F(x)$，则非齐次方程的通解为

$$y(x) = \mathrm{e}^{A(x)} F(x) + \mathrm{e}^{A(x)} C,$$

其中 C 是任意常数。

高阶线性微分方程

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} + a_{n-1}(x) \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1(x) \frac{\mathrm{d}y}{\mathrm{d}x} + a_0(x) y = f(x)$$

是 n 阶线性微分方程，当 $f(x)$恒为零时，称为**齐次**线性方程，当 $f(x)$ 不恒为零时，称为**非齐次**线性方程。

设 $y_0(x)$ 是齐次方程的一个非零解，$y(x) = C(x) y_0(x)$ 是非齐次方程的解。则

$$C(x) \left[\frac{\mathrm{d}^n y_0}{\mathrm{d}x^n} + a_{n-1}(x) \frac{\mathrm{d}^{n-1} y_0}{\mathrm{d}x^{n-1}} + \cdots + a_1(x) \frac{\mathrm{d}y_0}{\mathrm{d}x} + a_0(x) y_0 \right] + \\ C^{(n)}(x) y_0(x) + C^{(n-1)}(x) b_{n-1}(x) + \cdots + C'(x) b_1(x) = f(x),$$

即

$$C^{(n)}(x) y_0(x) + C^{(n-1)}(x) b_{n-1}(x) + \cdots + C'(x) b_1(x) = f(x),$$

这是关于 C 的一个 n 阶线性方程，但也是关于 C' 的一个 $n-1$ 阶线性方程。

所以只要我们对每一个线性齐次方程都能找到一个非零解，我们就可以不断使用常数变易法把非齐次方程的阶数降低，直到降为一阶方程。最终求得原高阶方程的通解。

高阶常系数线性微分方程

$y_0(x) = \mathrm{e}^{\lambda x}$ 是

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} + a_{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = 0$$

的解，当且仅当 λ 是特征指数，即特征多项式

$$\lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0$$

的根。

根据代数学基本定理**??**，特征多项式有如下形式的因子分解

$$\lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_m)^{n_m},$$

则

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n} + a_{n-1} \frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1 \right)^{n_1} \cdots \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_m \right)^{n_m} y.$$

因此，只要

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}-\lambda_k\right)^{n_k} y=0,$$

y 就是原n 阶齐次线性微分方程的解。

不难验证对每个 $k=1,\ldots,m$，

$$\mathrm{e}^{\lambda_k x},\quad x\mathrm{e}^{\lambda_k x},\quad \ldots,\quad \frac{x^{n_k-1}}{(n_k-1)!}\mathrm{e}^{\lambda_k x}$$

是

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}-\lambda_k\right)^{n_k} y=0$$

的解。并且这样得到的 n 个解是线性无关的。齐次方程

$$\frac{\mathrm{d}^n y}{\mathrm{d}x^n}+a_{n-1}\frac{\mathrm{d}^{n-1} y}{\mathrm{d}x^{n-1}}+\cdots+a_1\frac{\mathrm{d}y}{\mathrm{d}x}+a_0 y=0$$

的通解是这 n 个解的任意线性组合。

对于实系数的特征多项式而言，复特征指数 $\lambda=\alpha+\mathrm{i}\beta$($\beta>0$)的复共轭$\overline{\lambda}=\alpha-\mathrm{i}\beta$ 也是特征指数。

$$\mathrm{e}^{\alpha x}\cos\beta x=\frac{\mathrm{e}^{\lambda x}+\mathrm{e}^{\overline{\lambda}x}}{2},\quad \mathrm{e}^{\alpha x}\sin\beta x=\frac{\mathrm{e}^{\lambda x}-\mathrm{e}^{\overline{\lambda}x}}{2},$$

从而

$$\begin{gathered}\mathrm{e}^{\alpha x}\cos\beta x,\quad \mathrm{e}^{\alpha x}\sin\beta x,\quad x\mathrm{e}^{\alpha x}\cos\beta x,\quad x\mathrm{e}^{\alpha x}\sin\beta x,\\ \ldots,\quad \frac{x^{n_k-1}}{(n_k-1)!}\mathrm{e}^{\alpha x}\cos\beta x,\quad \frac{x^{n_k-1}}{(n_k-1)!}\mathrm{e}^{\alpha x}\sin\beta x,\end{gathered}$$

代替了

$$\mathrm{e}^{\lambda x},\quad x\mathrm{e}^{\lambda x},\quad \ldots,\quad \frac{x^{n_k-1}}{(n_k-1)!}\mathrm{e}^{\lambda x},\qquad \mathrm{e}^{\overline{\lambda}x},\quad x\mathrm{e}^{\overline{\lambda}x},\quad \ldots,\quad \frac{x^{n_k-1}}{(n_k-1)!}\mathrm{e}^{\overline{\lambda}x}.$$

习题6.4

1. 设 $\lambda_1,\ldots,\lambda_m$ 是 m 个不同的复数，证明对任意自然数 n，

$$\begin{gathered}\mathrm{e}^{\lambda_1 x},\quad x\mathrm{e}^{\lambda_1 x},\quad \ldots,\quad \frac{x^n}{n!}\mathrm{e}^{\lambda_1 x};\\ \mathrm{e}^{\lambda_2 x},\quad x\mathrm{e}^{\lambda_2 x},\quad \ldots,\quad \frac{x^n}{n!}\mathrm{e}^{\lambda_2 x};\\ \vdots\\ \mathrm{e}^{\lambda_m x},\quad x\mathrm{e}^{\lambda_m x},\quad \ldots,\quad \frac{x^n}{n!}\mathrm{e}^{\lambda_m x}\end{gathered}$$

是线性无关的。(提示：用 $\frac{\mathrm{d}}{\mathrm{d}x}-\lambda_k$ 作用于上述函数的线性组合)

第 7 章　定积分与广义积分

7.1　定积分的概念

例 7.1.1. (由速度计算路程) 路程 $S(t)$ 是时间 t 的函数，速率 $v(t)$ 是 $S(t)$ 的导数。已知函数 $v(t)$，求在时间段 $[a,b]$ 上运动的路程。

解： 我们把时间段 $[a,b]$ 分成一些很短的时间段：

$$a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b.$$

于是对每个 $k : 1 \leq k \leq n$，存在 $\eta_k \in (t_{k-1}, t_k)$ 使得

$$S(t_k) - S(t_{k-1}) = S'(\eta_k)(t_k - t_{k-1}) = v(\eta_k)(t_k - t_{k-1}). \tag{7.1}$$

因此

$$S(b) - S(a) = \sum_{k=1}^{n} \big(S(t_k) - S(t_{k-1})\big) = \sum_{k=1}^{n} v(\eta_k)(t_k - t_{k-1}).$$

然而我们并不知道满足等式(7.1)的 η_k 的值，所以无法利用上式得到 $S(b)-S(a)$ 的值。

　　我们任取 $\xi_k \in [t_{k-1}, t_k]$，这样可以计算

$$\sum_{k=1}^{n} v(\xi_k)(t_k - t_{k-1})$$

的值。这相当于把在每个时间段 $[t_{k-1}, t_k]$ 内的运动近似看成是速率为 $v(\xi_k)$ 的直线匀速运动并把所对应的路程相加，这是实际运动路程的一个近似。

$$\begin{aligned}
&\left| \sum_{k=1}^{n} v(\xi_k)(t_k - t_{k-1}) - \sum_{k=1}^{n} v(\eta_k)(t_k - t_{k-1}) \right| \\
\leq & \sum_{k=1}^{n} |v(\xi_k) - v(\eta_k)| \, (t_k - t_{k-1}) \\
\leq & \max_{1 \leq k \leq n} |v(\xi_k) - v(\eta_k)| \sum_{k=1}^{n} (t_k - t_{k-1}) = \max_{1 \leq k \leq n} |v(\xi_k) - v(\eta_k)| (b - a).
\end{aligned}$$

如果 $v(t)$ 连续，则 $v(t)$ 在有界闭区间 $[a,b]$ 上一致连续，所以对任意 $\varepsilon>0$，存在 $\delta>0$ 使得只要 $t,t'\in[a,b]$ 满足 $|t-t'|<\delta$，就有 $|v(t)-v(t')|<\varepsilon$。从而当 $\max\limits_{1\le k\le n}(t_k-t_{k-1})<\delta$ 时，$|\xi_k-\eta_k|<\delta$，$|v(\xi_k)-v(\eta_k)|<\varepsilon$，因此

$$\left|S(b)-S(a)-\sum_{k=1}^{n}v(\xi_k)(t_k-t_{k-1})\right|\le(b-a)\varepsilon.$$

所以对 $[a,b]$ 的任意足够细(即 $\max\limits_{1\le k\le n}(t_k-t_{k-1})$ 足够小)的分割

$$a=t_0<t_1<t_2<\cdots<t_{n-1}<t_n=b$$

以及任意的 $\xi_k\in[t_{k-1},t_k]$，$\sum\limits_{k=1}^{n}v(\xi_k)(t_k-t_{k-1})$ 都是路程 $S(b)-S(a)$ 的足够好的近似值。 □

例 7.1.2. (由密度计算质量) 一根直的金属丝，质量 $m(l)$ 是长度参数 $l\in[0,L]$ 的函数，线密度为 $\rho(l)$。已知函数 $\rho(l)$，求金属丝的质量。

解： 我们把金属丝按长度区间 $[0,L]$ 分成一些很短的片段：

$$0=l_0<l_1<l_2<\cdots<l_{n-1}<l_n=L.$$

任取 $\xi_k\in[l_{k-1},l_k]$，这样可以计算

$$\sum_{k=1}^{n}\rho(\xi_k)(l_k-l_{k-1})$$

的值。这相当于把在每一段金属丝 $[l_{k-1},l_k]$ 近似看成质量按密度 $\rho(\xi_k)$ 均匀分布。于是

$$m\approx\sum_{k=1}^{n}\rho(\xi_k)(l_k-l_{k-1}).$$

□

例 7.1.3. (曲边梯形的面积) 设 D 是平面直角坐标系中由直线 $x=a,x=b,y=0$ 和 $y=f(x)$ (其中 $f(x)>0$)围成的平面区域。求区域 D 的面积。

解： 对一般的区域，我们尚未定义面积这个概念。所以这里我们不仅要解决区域 D 面积大小的计算，更重要的是我们要给这个面积一个合理的定义。

我们把区间 $[a,b]$ 分成一些很短的区间：

$$a=x_0<x_1<x_2<\cdots<x_{n-1}<x_n=b.$$

于是对每个 $k:1\le k\le n$，取 $\xi_k\in[x_{k-1},x_k]$，计算

$$\sum_{k=1}^{n}f(\xi_k)(x_k-x_{k-1})$$

的值，这相当于用一组矩形近似代替了区域 D，而这个和是这组矩形面积的和，它是 D 的面积(尚未定义)的一个近似值。

如果 f 连续，则按照前例中的办法可以证明，对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得当 $\max\limits_{1\le k\le n}(x_k - x_{k-1}) < \delta$ 时，对任意 $\xi_k, \eta_k \in [x_{k-1}, x_k]$，

$$\left|\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) - \sum_{k=1}^{n} f(\eta_k)(x_k - x_{k-1})\right| \le (b-a)\varepsilon.$$

这表明：如果 f 连续，则对 $[a,b]$ 的任意足够细的分割

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b$$

任何两个近似值 $\sum\limits_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$，$\sum\limits_{k=1}^{n} f(\eta_k)(x_k - x_{k-1})$ 都近乎相等。所以取其中任何一个作为 D 的面积的近似值都是合理的。而 D 的面积应该是一个常数，它可以由越来越精细的分割所对应的近似值 $\sum\limits_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$ 所逼近。

□

上述各例可以统一为以下形式：对函数 $f : [a,b] \to \mathbb{R}$，以及区间 $[a,b]$ 的划分

$$P : a = x_0 < x_1 < \cdots < x_n = b$$

和标志点组 $\xi = \{\xi_k | 1 \le k \le n\}(\xi_k \in [x_{k-1}, x_k])$，记

$$S(f, P, \xi) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}),$$

称它为 f 关于 P 和 ξ 的 **Riemann 和**。

我们希望这些 Riemann 和以某种方式存在极限。这需要解决因标志点组选取的随意性导致的问题，并克服由划分造成的障碍。

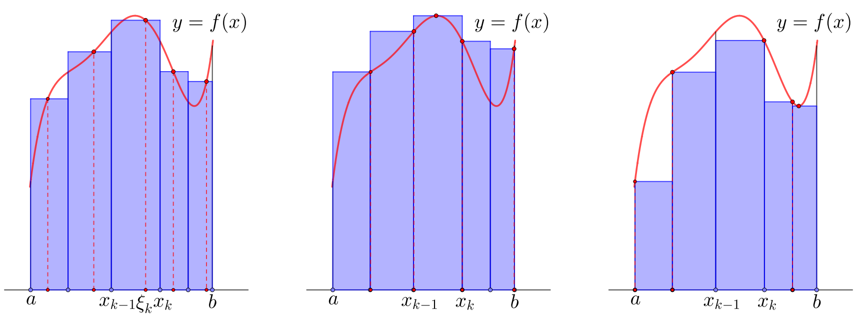


图 7.1: Riemann和，Darboux上和，Darboux下和

对有界函数数 $f:[a,b]\to\mathbb{R}$ 以及区间 $[a,b]$ 的划分 P，记

$$\overline{S}(f,P)=\sum_{k=1}^{n}\sup_{x\in[x_{k-1},x_k]}f(x)(x_k-x_{k-1}),$$
$$\underline{S}(f,P)=\sum_{k=1}^{n}\inf_{x\in[x_{k-1},x_k]}f(x)(x_k-x_{k-1}),$$

分别称为 f 关于 P 的 **Darboux 上和**与 **Darboux 下和**。易见，

$$\underline{S}(f,P)=\inf_{\xi\in\Xi(P)}S(f,P,\xi),\quad \overline{S}(f,P)=\sup_{\xi\in\Xi(P)}S(f,P,\xi),$$

其中 $\Xi(P)$ 为对应于划分 P 的所有标志点组。从而对任意标志点组 ξ 和 η，都有

$$\underline{S}(f,P)\le S(f,P,\xi)\le\overline{S}(f,P),$$
$$|S(f,P,\xi)-S(f,P,\eta)|\le\overline{S}(f,P)-\underline{S}(f,P).$$

这样，通过 Darboux 上下和我们控制了 Riemann 和，使其不受标志点组的随意性的影响。

接下来，我们希望 Darboux 上下和 $\overline{S}(f,P)$ 与 $\underline{S}(f,P)$ 能够相互靠近。我们称划分$\tilde{P}$是划分P的一个**加细**，如果 P 中每个分点 x_k 都是 $\tilde{P}$ 中的分点。 易见

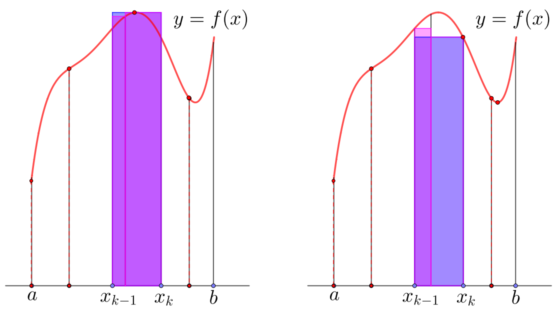


图 7.2: 划分加细后的Darboux上和与Darboux下和

对任何划分 P 和它的加细划分 $\tilde{P}$，

$$\underline{S}(f,P)\le\underline{S}(f,\tilde{P})\le\overline{S}(f,\tilde{P})\le\overline{S}(f,P).$$

从而对任何划分 P 和 Q，以 P,Q 中的所有分点形成的划分 $P\vee Q$ 是 P 和 Q 的加细划分，满足

$$\underline{S}(f,P)\le\underline{S}(f,P\vee Q)\le\overline{S}(f,P\vee Q)\le\overline{S}(f,Q).$$

因此，存在

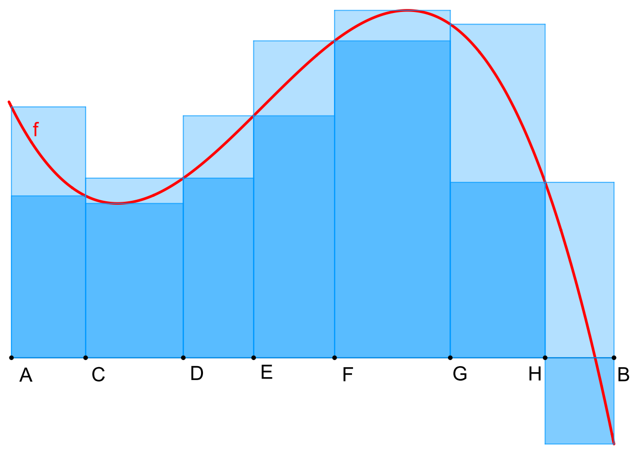
$$\overline{\int_a^b}f(x)\mathrm{d}x=\inf_{P\in\mathcal{P}}\overline{S}(f,P),\qquad \underline{\int_a^b}f(x)\mathrm{d}x=\sup_{P\in\mathcal{P}}\underline{S}(f,P),$$

分别称为 f 在区间 $[a,b]$ 上的 **Darboux 上积分**和 **Darboux 下积分**，其中 $\mathcal{P}$ 为区间 $[a,b]$ 的所有划分。

注 7.1.4.

$$\overline{\int_a^b} f(x)\mathrm{d}x = \inf_{P\in\mathcal{P}} \sup_{\xi\in\Xi(P)} S(f,P,\xi), \qquad \underline{\int_a^b} f(x)\mathrm{d}x = \sup_{P\in\mathcal{P}} \inf_{\xi\in\Xi(P)} S(f,P,\xi),$$

这类似于数列上下极限的想法。



定义 7.1.5. 称有界函数 f 在区间 $[a,b]$ 上 **Riemann 可积** (记 $f\in R[a,b]$)，如果

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

此时记上述共同的值为 $\int_a^b f(x)\mathrm{d}x$，称为 f 在区间 $[a,b]$ 上的 **Riemann 积分**(也称为**定积分**)。

定理 7.1.6. *设函数 $f:[a,b]\to\mathbb{R}$ 有界，则以下陈述等价：*

1. *f 在区间 $[a,b]$ 上可积；*
2. *存在实数 A 使得对任意 $\varepsilon>0$，存在 $[a,b]$ 的划分 P 使得对任意标志点组 ξ，$|S(f,P,\xi)-A|<\varepsilon$；*

3. 对任意 $\varepsilon>0$，存在 $[a,b]$ 的划分 P 使得 $\overline{S}(f,P)-\underline{S}(f,P)<\varepsilon$。

证明. 留给读者自己完成。 □

注 7.1.7. 上述定理表明，从几何上看，函数 $f:[a,b]\to\mathbb{R}$ 在区间 $[a,b]$ 上 Riemann 可积当且仅当存在着面积总和足够小的有限多个长方形，它们覆盖 $y=f(x)$ 在区间 $[a,b]$ 上的图像。

例 7.1.8. Dirichlet 函数 $D(x)=\begin{cases}1, & x\in\mathbb{Q};\\ 0, & x\in\mathbb{R}\backslash\mathbb{Q}\end{cases}$ 在任何区间 $[a,b]$ 上不可积。对区间 $[a,b]$ 的任意分割 P，取 $\xi_k\in\mathbb{Q}$，$\eta_k\in\mathbb{R}\backslash\mathbb{Q}$，则

$$S(D,P,\xi)=1,\quad S(D,P,\eta)=0.$$

所以 D 在 $[a,b]$ 上不是 Riemann 可积的。

定理 7.1.9 (Newton 1687)**.** *若 $f:[a,b]\to\mathbb{R}$ 单调，则对于 $[a,b]$ 的任意划分 P，都成立*

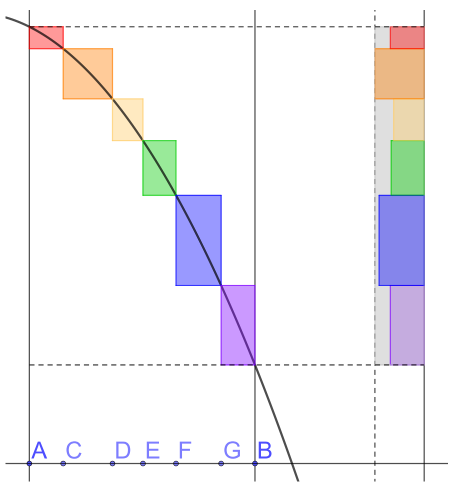
$$\overline{S}(f,P)-\underline{S}(f,P)\leq|f(b)-f(a)|\max(x_k-x_{k-1}).$$

从而 $f\in R[a,b]$，且对任意 $\varepsilon>0$ 以及任意划分 P，只要

$$\max(x_k-x_{k-1})<\frac{\varepsilon}{|f(b)-f(a)|+1},$$

就有对任意标志点组 ξ 都成立

$$\left|S(f,P,\xi)-\int_a^b f(x)\mathrm{d}x\right|<\varepsilon.$$



证明. [1] 不妨设 f 单调不增。则对任意划分 P，

$$\begin{aligned}\overline{S}(f,P)-\underline{S}(f,P)&=\sum_{k=1}^{n}(f(x_{k-1})-f(x_k))(x_k-x_{k-1})\\&\le\max_{1\le k\le n}(x_k-x_{k-1})\sum_{k=1}^{n}(f(x_{k-1})-f(x_k))\\&=\max_{1\le k\le n}(x_k-x_{k-1})(f(a)-f(b)).\end{aligned}$$

于是对任意 $\varepsilon>0$，当 $\max\limits_{1\le k\le n}(x_k-x_{k-1})|f(b)-f(a)|<\varepsilon$ 时（例如 P 是 n 等分满足 $\frac{(b-a)|f(b)-f(a)|}{n}<\varepsilon$），$|\overline{S}(f,P)-\underline{S}(f,P)|<\varepsilon$。因此 $f\in R[a,b]$。而且对任意标志点组 ξ，

$$\left|S(f,P,\xi)-\int_a^b f(x)\mathrm{d}x\right|\le\overline{S}(f,P)-\underline{S}(f,P)<\varepsilon.$$

□

例 7.1.10. 设 $\mu>0$，求 $\int_0^b x^\mu\mathrm{d}x$。
解: 因为函数 $f(x)=x^\mu$ 在区间 $[0,b]$ 上单调增，所以可积。我们把区间 $[0,b]$ 进行 n 等分

$$P_n:0=x_0<x_1<x_2<\cdots<x_{n-1}<x_n=b,\qquad x_k=\frac{kb}{n}.$$

取 $\xi_k=x_k$，则

$$\begin{aligned}\lim_{n\to\infty}S(f,P_n,\xi)&=\lim_{n\to\infty}\sum_{k=1}^{n}\left(\frac{kb}{n}\right)^\mu\frac{b}{n}=b^{\alpha+1}\lim_{n\to\infty}\frac{\sum\limits_{k=1}^{n}k^\mu}{n^{\mu+1}}\\&=b^{\mu+1}\lim_{n\to\infty}\frac{n^\mu}{n^{\mu+1}-(n-1)^{\mu+1}}\qquad(\text{Stolz 定理})\\&=b^{\mu+1}\lim_{n\to\infty}\frac{1}{n\left(1-(1-\frac{1}{n})^{\mu+1}\right)}=\frac{b^{\mu+1}}{\mu+1}.\end{aligned}$$

所以由前面的定理推论，

$$\int_0^b x^\mu\mathrm{d}x=\frac{b^{\mu+1}}{\mu+1}.$$

□

定理 7.1.11. *若 $f:[a,b]\to\mathbb{R}$ 有界，f 在 (a,b) 上连续，则 $f\in R[a,b]$，并且对任意 $\varepsilon>0$，存在 $\delta>0$ 使得对 $[a,b]$ 的任意划分 P 只要 $\max\limits_{1\le k\le n}(x_k-x_{k-1})<\delta$，就有*

$$\overline{S}(f,P)-\underline{S}(f,P)<\varepsilon,$$

[1]这个结论和证明见于 Newton 的《自然哲学的数学原理》中引理2和引理3。

从而对任意标志点组 ξ，

$$\left|S(f,P,\xi)-\int_a^b f(x)\mathrm{d}x\right|<\varepsilon.$$

证明. 设 $|f(x)|\le M(\ \forall x\in[a,b]\ )$。对任意 $\varepsilon>0$，取 $\delta_1>0$ 使得

$$\delta_1<\frac{b-a}{10},\quad 4M\delta_1<\frac{\varepsilon}{3}.$$

取 $u=a+\delta_1, v=b-\delta_1$。因为 f 在 $[u,v]$ 上连续，所以 f 在 $[u,v]$ 上一致连续，因此存在 $\delta_2>0$ 使得对任意 $x,y\in[u,v]$，只要 $|x-y|<\delta_2$，就有 $|f(x)-f(y)|<\frac{\varepsilon}{3(b-a)}$。

任取 $[a,b]$ 的分割 P 满足 $\max\limits_{1\le j}(x_j-x_{j-1})<\min\{\delta_1,\delta_2\}$，

$$a=x_0<\cdots<x_k\le u<x_{k+1}<\cdots<x_l<v\le x_{l+1}<\cdots<x_n=b.$$

则

$$\overline{S}(f,P)-\underline{S}(f,P)<2M(x_{k+1}-a)+\frac{\varepsilon}{3(b-a)}\cdot(x_l-x_{k+1})+2M(b-x_l)<\varepsilon.$$

□

定理 7.1.12. 若 $a<b<c$，$f\in R[a,b]\bigcap R[b,c]$，则 $f\in R[a,c]$，且

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

证明. 对任意 $\varepsilon>0$，存在 $[a,b]$ 的划分 P 和 $[b,c]$ 的划分 Q 使得

$$\int_a^b f(x)\mathrm{d}x-\varepsilon\le\underline{S}(f,P)\le\overline{S}(f,P)\le\int_a^b f(x)\mathrm{d}x+\varepsilon,$$
$$\int_b^c f(x)\mathrm{d}x-\varepsilon\le\underline{S}(f,Q)\le\overline{S}(f,Q)\le\int_b^c f(x)\mathrm{d}x+\varepsilon.$$

于是 $P\cup Q$ 是 $[a,c]$ 的一个划分，满足

$$\begin{aligned}\int_a^b f(x)\mathrm{d}x+\int_b^c f(x)\mathrm{d}x-2\varepsilon&\le\underline{S}(f,P\cup Q)=\underline{S}(f,P)+\underline{S}(f,Q)\\&\le\overline{S}(f,P)+\overline{S}(f,Q)=\overline{S}(f,P\cup Q)\le\int_a^b f(x)\mathrm{d}x+\int_b^c f(x)\mathrm{d}x+2\varepsilon,\end{aligned}$$

从而 f 在 $[a,c]$ 上可积，且

$$\int_a^b f(x)\mathrm{d}x+\int_b^c f(x)\mathrm{d}x-2\varepsilon\le\int_a^c f(x)\mathrm{d}x\le\int_a^b f(x)\mathrm{d}x+\int_b^c f(x)\mathrm{d}x+2\varepsilon.$$

所以

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

□

推论 7.1.13. 如果 f 在区间 $[a,b]$ 上有界，且至多只有有限多个间断点，则 f 在区间 $[a,b]$ 上可积。 □

受以上结果启发，我们给出 Riemann 对积分的定义。

定义 7.1.14. 称函数 f 在区间 $[a,b]$ 上 **Riemann 可积**(记 $f\in R[a,b]$)，如果存在数 I 使得对任意 $\varepsilon>0$，存在 $\delta>0$ 使得：对区间 $[a,b]$ 的任意分割：

$$P: a=x_0<x_1<x_2<\cdots<x_{n-1}<x_n=b,$$

只要 $\|P\|:=\max\limits_{1\le k\le n}\Delta x_k<\delta$ (其中 $\Delta x_k=x_k-x_{k-1}$)，就有对任意 $\xi_k\in[x_{k-1},x_k]$，都有

$$\left|\sum_{k=1}^n f(\xi_k)\Delta x_k-I\right|<\varepsilon.$$

记 $I=\int_a^b f(x)\mathrm{d}x$，称为 f 在区间 $[a,b]$ 上的 **Riemann 积分**(也称为**定积分**)。记

$$S(f,P,\xi)=\sum_{k=1}^n f(\xi_k)\Delta x_k,$$

称为 f 关于分割P 和标志点组 $\xi=\{\xi_k\}_{k=1}^n$ 的 **Riemann 和**。

定理 7.1.15 (Lebesgue)**.** 以下结论等价：

1. *f 在区间 $[a,b]$ 上 Riemann 可积；*
2. *f 在区间 $[a,b]$ 上有界，且 Darboux 可积；*
3. *f 在区间 $[a,b]$ 上有界，且 f 在区间 $[a,b]$ 中的间断点组成一个长度为零的集合*

例 7.1.16. Riemann 函数

$$R(x)=\begin{cases}\frac{1}{p}, & x=\frac{q}{p}\text{ 是既约分数，}p\text{ 是正整数;}\\ 0, & x\in\mathbb{R}\backslash\mathbb{Q}\end{cases}$$

在任何区间 $[a,b]$ 上都是 Riemann 可积的：$0\le R(x)\le 1$，从而有界；$R(x)$ 的间断点集为 $\mathbb{Q}$ 是长度为零的集合。对任意划分 P ，$\underline{S}(R,P)=0$，所以 $\int_a^b R(x)\mathrm{d}x=0$。

习题7.1

1. 针对给定的区间 $[a,b]$，函数 f、分割 $P=\{x_k\}_{k=0}^n$ 和标志点 $\{\xi_k\}_{k=1}^n$，计算 Riemann 和 $S(f,P,\{\xi_k\})$ 以及极限 $\lim\limits_{n\to+\infty}S(f,P,\{\xi_k\})$。

(a) $[a,b]$，$f(x)=x^2$，$P: a=x_0<x_1<\cdots<x_n=b$，$\xi_k=\sqrt{\dfrac{x_{k-1}^2+x_{k-1}x_k+x_k^2}{3}}$。

(b) $[a,b]$($a>0$)，$f(x)=x^r$($r\in\mathbb{R}$)，$x_k=a\left(\dfrac{b}{a}\right)^{k/n}$，$\xi_k=x_k$。

(c) $[0,b]$($b>0$)，$f(x)=x^r$($r>0$)，$\xi_k=x_k=b\lambda^{n-k}(0<\lambda<1, 1\le k\le n)$，$x_0=0$。

(d) $[a,b]$，$f(x)=\mathrm{e}^x$，$x_k=a+\frac{k}{n}(b-a)$，$\xi_k=x_k$。

(e) $[a,b]$，$f(x)=\mathrm{e}^{\mathrm{i}x}$，$x_k=a+\frac{k}{n}(b-a)$，$\xi_k=x_k$。

(f) $[a,b]$，$f(x)=\sin x$，$x_k=a+\frac{k}{n}(b-a)$，$\xi_k=x_k$。

2. 证明 Dirichlet 函数
$$D(x)=\begin{cases}1, & x\in\mathbb{Q},\\ 0, & x\in\mathbb{R}\backslash\mathbb{Q}\end{cases}$$
在任何区间 $[a,b]$ 上都不是 Riemann 可积函数。

3. 证明：如果 f 在区间 $[a,b]$ 上 Riemann 可积，则 f 在 $[a,b]$ 上有界。(提示：用反证法。设 $y_k\in[a,b]$ 满足 $f(y_k)>k^2$。对任意正整数 n，考虑区间 $[a,b]$ 的 n 等分以及适当选取的标志点集使相应的 Riemann 和无界)

4. 证明：若 f 在区间 $[a,b]$ 上 Riemann 可积，则 f 在 $[a,b]$ 的任何闭子区间上也可积。

7.2 定积分的性质

(1) **线性**

定理 7.2.1. 对任何 $f,g\in R[a,b]$ 以及 $\lambda,\mu\in\mathbb{R}$，都有 $\lambda f+\mu g\in R[a,b]$，且
$$\int_a^b \lambda f(x)+\mu g(x)\mathrm{d}x=\lambda\int_a^b f(x)\mathrm{d}x+\mu\int_a^b g(x)\mathrm{d}x.$$

证明. 由 $f,g\in R[a,b]$ 和 Lebesgue 定理知，f,g 有界，它们在区间 $[a,b]$ 中的间断点集是长度为零的集合。

因为 $|\lambda f(x)+\mu g(x)|\le|\lambda||f(x)|+|\mu||g(x)|$，所以 $\lambda f+\mu g$ 有界。

另外，$\lambda f+\mu g$ 的间断点是 f 的间断点或者是 g 的间断点，从而 $\lambda f+\mu g$ 在区间 $[a,b]$ 中的间断点集是长度为零的集合。因此由 Lebesgue 定理知 $\lambda f+\mu g\in R[a,b]$。

对任意 $\varepsilon > 0$，存在 $\delta > 0$，使得对任意划分 P，只要 $\|P\| < \delta$，就有对任意标志点组 ξ，

$$\left|S(f,P,\xi) - \int_a^b f(x)\mathrm{d}x\right| < \varepsilon,$$

$$\left|S(g,P,\xi) - \int_a^b g(x)\mathrm{d}x\right| < \varepsilon,$$

$$\left|S(\lambda f+\mu g,P,\xi) - \int_a^b \lambda f(x)+\mu g(x)\mathrm{d}x\right| < \varepsilon.$$

而

$$S(\lambda f+\mu g,P,\xi) = \lambda S(f,P,\xi) + \mu S(g,P,\xi),$$

所以

$$\left|\lambda\int_a^b f(x)\mathrm{d}x + \mu\int_a^b g(x)\mathrm{d}x - \int_a^b \lambda f(x)+\mu g(x)\mathrm{d}x\right| < (|\lambda|+|\mu|+1)\varepsilon,$$

因此

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

□

(2) **保序性**

定理 7.2.2. 设 $f,g \in R[a,b]$，且对任意 $x \in [a,b]$，$f(x) \le g(x)$，则

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

进一步，如果 f,g 在某个 $x_0 \in [a,b]$ 处连续且 $f(x_0) < g(x_0)$，则

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

证明. 对任意 $\varepsilon > 0$，取 $[a,b]$ 的划分 P 使得

$$\overline{S}(g,P) \le \int_a^b g(x)\mathrm{d}x + \varepsilon.$$

对于每个子区间 $I_k = [x_{k-1},x_k]$ 以及任意 $x \in I_k$， $f(x) \le g(x) \le \sup\limits_{t\in I_k} g(t)$，从而 $\sup\limits_{t\in I_k} f(t) \le \sup\limits_{t\in I_k} g(t)$。于是

$$\int_a^b f(x)\mathrm{d}x \le \overline{S}(f,P) \le \overline{S}(g,P) \le \int_a^b g(x)\mathrm{d}x + \varepsilon.$$

所以

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

如果 f,g 在某个 $x_0\in[a,b]$ 处连续且 $f(x_0)<g(x_0)$，则存在常数 A,B 以及 x_0 的邻域 $U\subset[a,b]$，使得对任意 $x\in U$，

$$f(x)<A<B<g(x).$$

把 U 的端点加入到划分 P 中得到划分 Q，于是

$$\int_a^b g(x)\mathrm{d}x+\varepsilon-\int_a^b f(x)\mathrm{d}x\geq\overline{S}(g,Q)-\overline{S}(f,Q)\geq(B-A)|U|.$$

从而

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

□

推论 7.2.3. 设 $f\in R[a,b]$，则 $|f|\in R[a,b]$ 且 $\left|\int_a^b f(x)\mathrm{d}x\right|\leq\int_a^b|f(x)|\mathrm{d}x$。

证明. 易见 $|f|$ 有界，且 $|f|$ 的间断点集是 f 的间断点集的子集，从而是长度为零的集合，因此 $|f|\in R[a,b]$。

再由 $-|f(x)|\leq f(x)\leq|f(x)|$ 和积分保序性以及线性得到 $\left|\int_a^b f(x)\mathrm{d}x\right|\leq\int_a^b|f(x)|\mathrm{d}x$。 □

(3) **积分平均值定理和 Cauchy-Schwarz 不等式**

定理 7.2.4 (积分平均值定理，第一积分中值定理)**.** 设 $f\in C[a,b]$，$g\in R[a,b]$ 且对任意 $x\in[a,b]$，$g(x)\geq 0$，则存在 $\xi\in[a,b]$ 使得

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

若进一步 $\int_a^b g(x)\mathrm{d}x>0$，则

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

证明. 由 Lebesgue 准则知对 $f,g\in R[a,b]$，$fg\in R[a,b]$。

由于 $f\in C[a,b]$，所以 f 有最小值 m 和最大值 M。从而由 g 非负以及保序性知

$$m\int_a^b g(x)\mathrm{d}x\leq\int_a^b f(x)g(x)\mathrm{d}x\leq M\int_a^b g(x)\mathrm{d}x.$$

若 $\int_a^b g(x)\mathrm{d}x=0$，则 $\int_a^b f(x)g(x)\mathrm{d}x=0$，此时任取 $\xi\in[a,b]$，都有

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

若 $\int_a^b g(x)\mathrm{d}x>0$，则

$$m\leq\frac{\int_a^b f(x)g(x)\mathrm{d}x}{\int_a^b g(x)\mathrm{d}x}\leq M.$$

再由连续函数的介值性质知存在 $\xi \in [a,b]$ 使得

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

□

定理 7.2.5 (Cauchy-Schwarz不等式)**.** 设 $f, g \in R[a,b]$，则 $fg \in R[a,b]$ 且

$$\left|\int_a^b f(x)g(x)\mathrm{d}x\right| \leq \sqrt{\int_a^b |f(x)|^2\mathrm{d}x \int_a^b |g(x)|^2\mathrm{d}x},$$

其中等号成立当且仅当存在常数 t 使得在一个长度为零的集合以外，$f(x) = tg(x)$ 或 $g(x) = tf(x)$。

证明. 由线性和保序性，得到对任意 $t \in \mathbb{R}$，

$$\int_a^b |f(x) - tg(x)|^2\,\mathrm{d}x = t^2\int_a^b |g(x)|^2\mathrm{d}x - 2t\int_a^b f(x)g(x)\mathrm{d}x + \int_a^b |f(x)|^2\mathrm{d}x \geq 0.$$

若 $\int_a^b |g(x)|^2\mathrm{d}x = 0$，则 $\int_a^b f(x)g(x)\mathrm{d}x = 0$，定理中的不等式(此时为等式)成立。此时，在除去 g 的间断点外，$g(x) = 0$，$g(x) = 0 \cdot f(x)$。

若 $\int_a^b |g(x)|^2\mathrm{d}x \neq 0$，则取

$$t^* = \frac{\int_a^b f(x)g(x)\mathrm{d}x}{\int_a^b |g(x)|^2\mathrm{d}x},$$

则

$$\int_a^b |f(x)|^2\mathrm{d}x - \frac{\left(\int_a^b f(x)g(x)\mathrm{d}x\right)^2}{\int_a^b |g(x)|^2\mathrm{d}x} = \int_a^b |f(x) - t^*g(x)|^2\,\mathrm{d}x \geq 0,$$

从而定理中不等式成立，并且其中等式成立当且仅当在除去 f, g 的间断点外，$f(x) = t^*g(x)$。 □

注 7.2.6. $\int_a^b f(x)g(x)\mathrm{d}x$ 在 $R[a,b]$ 上给出一个内积。

7.3 微积分基本定理与 Newton-Leibniz 公式

(1) **积分的有向性**

规定

$$\int_a^a f(x)\mathrm{d}x = 0, \qquad \int_b^a f(x)\mathrm{d}x = -\int_a^b f(x)\mathrm{d}x.$$

这样 $\int_a^b f(x)\mathrm{d}x$ 成为一个有方向的积分。

我们把 $f(x)$ 看成数轴上位于 x 处的沿数轴的一个向量，并把它理解为 x 处的力。当一个质点从 a 运动到 b 时，这个力场对它所做的功为 $\int_a^b f(x)\mathrm{d}x$。而当质点从 b 运动到 a 时，这个力场做功为 $\int_b^a f(x)\mathrm{d}x = -\int_a^b f(x)\mathrm{d}x$。

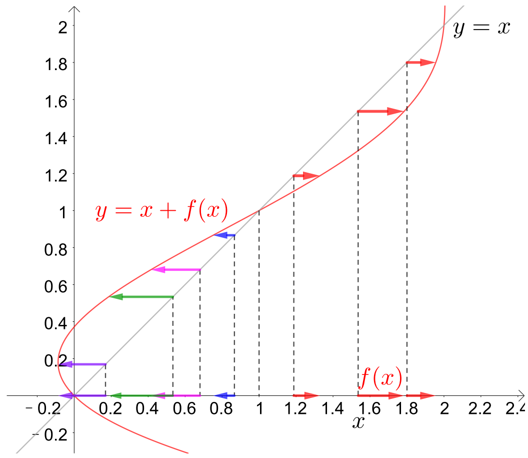


图 7.3:

另一方面，如果 f 在区间 $[a,b]$ 上可积，则任取单调减数列 $\{x_k\}_{k=0}^n$ 使得 $x_0=b, x_n=a$，并且对介于 x_{k-1} 和 x_k 之间的任意 ξ_k，

$$\sum_{k=1}^n f(\xi_k)(x_k-x_{k-1})=\sum_{k=1}^n -f(\xi_k)(x_{k-1}-x_k)$$

后者是 $-f$ 的一个 Riemann 和，因而形式上也有

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

当 $a=b$ 时，

$$\int_a^a f(x)\mathrm{d}x=-\int_a^a f(x)\mathrm{d}x=0.$$

定理 7.3.1. *设 f 在区间 I 上可积。则对于任意 $a,b,c\in I$，都有*

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

证明. 当 $a=b$ 或 $b=c$ 或 $c=a$ 时，结论成立。

若 $a<b$，则当 $c>b$ 时，此前的定理表明本定理结论中的等式成立。当 $c<b$ 时，则

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

从而

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

即结论成立。

若 $a > b$，则

$$\int_b^c f(x)\mathrm{d}x = \int_b^a f(x)\mathrm{d}x + \int_a^c f(x)\mathrm{d}x = -\int_a^b f(x)\mathrm{d}x + \int_a^c f(x),$$

因此

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

$\square$

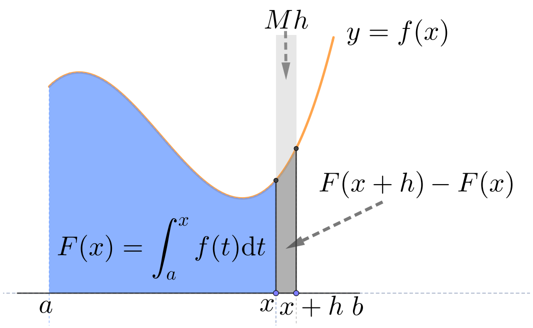


图 7.4:

定理 7.3.2. 设 f 在区间 I 的任何有界闭子区间上可积。对任意 $a \in I$，记

$$F(x) = \int_a^x f(t)\mathrm{d}t$$

则

1. $F : I \to \mathbb{R}$ 是连续函数；
2. 若 f 在 $x_0 \in I$ 连续，则 F 在 x_0 处可微，且 $F'(x_0) = f(x_0)$，即
$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\mathrm{d}t\bigg|_{x=x_0} = f(x_0);$$
3. 若 $f \in C(I)$，则 $F \in \mathscr{C}^1(I)$，且对任意 $x \in I$，$F'(x) = f(x)$。

证明. (1) 因为 f 可积，所以 f 有界，存在 $M > 0$ 使得 $|f(x)| \le M(\ \forall x \in I\ )$。从而对任意$u, v \in I$，

$$|F(v) - F(u)| = \left|\int_a^v f(x)\mathrm{d}x - \int_a^u f(x)\mathrm{d}x\right| = \left|\int_u^v f(x)\mathrm{d}x\right| \le M|v - u|.$$

所以 F 在 I 上(一致)连续。

(2) 设 $x_0, x_0+h \in I(h\neq 0)$。则对于任意 $\lambda\in[0,1]$，

$$F(x_0+\lambda h)-F(x_0)-f(x_0)\lambda h=\int_{x_0}^{x_0+\lambda h}\big(f(t)-f(x_0)\big)\mathrm{d}t.$$

因为 f 在 x_0 连续，所以对任意 $\varepsilon>0$，存在 $\delta>0$ 使得当 $x\in I$ 且 $|x-x_0|<\delta$ 时， $|f(x)-f(x_0)|<\varepsilon$，所以当 $0<\lambda<\frac{\varepsilon}{|h|}$ 时，

$$|F(x_0+\lambda h)-F(x_0)-f(x_0)\lambda h|=\left|\int_{x_0}^{x_0+\lambda h}\big(f(t)-f(x_0)\big)\mathrm{d}t\right|\leq\varepsilon|\lambda h|.$$

所以 F 在 x_0 处可微，且 $F'(x_0)=f(x_0)$。 □

定义 7.3.3. 称函数 F 是函数 f 在集合 I 上的一个**原函数**，如果对任何 $x\in I$，$F'(x)=f(x)$。

称函数 f 在集合 I 上的所有原函数组成的集合为 f 的**不定积分**，记为 $\int f(x)\mathrm{d}x$。

推论 7.3.4. *在一个区间上，所有连续函数都有原函数。* □

定理 7.3.5. *若 $f\in R[a,b]$ 且 F 是 f 在区间 $[a,b]$ 上的任何一个原函数，则*

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

证明. 对 $[a,b]$ 的任意分割

$$P: a=x_0<x_1<\cdots<x_n=b,$$

存在标志点集 $\xi(P)=\{\xi_k\}_{k=1}^n$ 使得对每个 k，$\xi_k\in[x_{k-1},x_k]$ 且

$$F(x_k)-F(x_{k-1})=F'(\xi_k)(x_k-x_{k-1})=f(\xi_k)(x_k-x_{k-1}).$$

于是

$$F(b)-F(a)=\sum_{k=1}^n\big(F(x_k)-F(x_{k-1})\big)=S(f,P,\xi(P)).$$

因为 $f\in R[a,b]$，所以对任意 $\varepsilon>0$，存在 $\delta>0$ 使得只要 $\|P\|<\delta$，就有

$$\left|S(f,P,\xi(P))-\int_a^b f(x)\mathrm{d}x\right|<\varepsilon.$$

因此

$$\left|F(b)-F(a)-\int_a^b f(x)\mathrm{d}x\right|<\varepsilon.$$

从而

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

□

从定理把积分计算的问题转化为找原函数的问题。因此上述定理称为微积分基本定理。

例 7.3.6. 设函数 f 在区间 I 上连续，$u, v : (\alpha, \beta) \to I$ 可导，则 $F(x) = \int_{u(x)}^{v(x)} f(t)\mathrm{d}t$ 关于 x 可导，且

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

证明. 任取 $a \in I$，令 $G(u) = \int_a^u f(t)\mathrm{d}t$。则 $F(x) = G(v(x)) - G(u(x))$。于是

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

□

例 7.3.7 (定积分的数值计算)**.** 虽然连续函数必然存在原函数，但是很多连续函数的原函数并不是初等函数，于是人们提出了一些数值算法计算积分的近似值。

设函数 f 具有我们希望的可微性。记

$$\begin{aligned} F(h) &= \int_{a-\frac{h}{2}}^{a+\frac{h}{2}} f(x)\mathrm{d}x, \\ F_{\text{矩}}(h) &= f(a)h, && \text{(矩形公式)} \\ F_{\text{梯}}(h) &= \frac{f(a+\frac{h}{2}) + f(a-\frac{h}{2})}{2}h. && \text{(梯形公式)} \end{aligned}$$

则 $F_{\text{矩}}(h), F_{\text{梯}}(h)$ 是 $F(h)$ 的近似值，我们估计它们的误差。

$$F'(h) = f(a+\frac{h}{2})\frac{1}{2} - f(a-\frac{h}{2})\left(-\frac{1}{2}\right) = \frac{f(a+\frac{h}{2}) + f(a-\frac{h}{2})}{2}$$

在 $h = 0$ 处进行 Taylor 展开，得到

$$\begin{aligned} F(h) &= f(a)h + \frac{f''(a)}{24}h^3 + O(h^5), \\ F_{\text{梯}}(h) &= f(a)h + \frac{f''(a)h^3}{8} + O(h^5). \end{aligned}$$

由此可见，对严格凸函数 f，矩形公式结果偏小，梯形公式结果偏大；对严格凹函数，结论相反。

由

$$(1-\lambda)F_{\text{矩}}(h) + \lambda F_{\text{梯}}(h) - F(h) = \left(\frac{\lambda}{8} - \frac{1}{24}\right) f''(a)h^3 + O(h^5)$$

知当且仅当 $\lambda=\frac{1}{3}$ 时，$(1-\lambda)F_{矩}(h)+\lambda F_{梯}(h)-F(h)=o(h^3)$。这样就得到如下辛普森公式，它是矩形公式和梯形公式的**内插修正**，

$$F_{辛}(h)=\frac{2F_{矩}(h)+F_{梯}(h)}{3}.$$

易见相比于矩形公式和梯形公式，辛普森公式具有更好的精度，

$$\begin{aligned}F_{矩}(h)-F(h)&=O(h^3),\\F_{梯}(h)-F(h)&=O(h^3),\\F_{辛}(h)-F(h)&=O(h^5).\end{aligned}$$

另外，辛普森公式对所有不超过三次的多项式都可以给出积分的精确值。

对于区间 $[a,b]$ 的 $2n$ 等分，

$$x_k=\frac{k(b-a)}{2n},\quad k=0,1,\ldots,2n,$$

辛普森公式为

$$\int_a^b f(x)\mathrm{d}x\approx\frac{f_0+4f_1+2f_2+4f_3+2f_4+\cdots+2f_{2n-2}+4f_{2n-1}+f_{2n}}{3}\frac{b-a}{2n}.$$

例如计算 $I=\int_0^1\frac{\sin x}{x}\mathrm{d}x$。我们把区间 $[0,1]$ 等分为10段，按矩形公式、梯形公式和辛普森公式分别得到 0.9462085788，0.9458320719，0.9460830765。我们用maple计算得到的结果是 0.9460830704。由此可见辛普森公式的优势。 □

习题7.3

1. 求 $\int_0^1\frac{1}{1+x^2}\mathrm{d}x$ 的值，并借助矩形公式、梯形公式、辛普森公式求圆周率 π 的近似值。

2. 求 $\int_1^2\frac{1}{x}\mathrm{d}x$ 的值，并借助矩形公式、梯形公式、辛普森公式求圆周率 $\ln 2$ 的近似值。

3. 设 $f:[a,b]\to\mathbb{R}$ 为凸函数，证明

$$f\left(\frac{a+b}{2}\right)(b-a)\le\int_a^b f(x)\mathrm{d}x\le\frac{f(a)+f(b)}{2}(b-a).$$

4. 函数

$$F(x)=\begin{cases}\sqrt[3]{x^4}\sin\frac{1}{x}, & x\neq 0;\\ 0, & x=0\end{cases}$$

处处可导，但 $f=F'$ 在区间 $[0,1]$ 上不可积。

7.4 积分计算

(1) **原函数与不定积分**

$$\int x^{\mu}\mathrm{d}x=\frac{x^{\mu+1}}{\mu+1}+C,\qquad \int_a^b x^{\mu}\mathrm{d}x=\frac{b^{\mu+1}-a^{\mu+1}}{\mu+1},\quad \mu\neq-1,$$

$$\int \mathrm{e}^{\lambda x}\mathrm{d}x=\frac{\mathrm{e}^{\lambda x}}{\lambda}+C,\qquad \int_a^b \mathrm{e}^{\lambda x}\mathrm{d}x=\frac{\mathrm{e}^{\lambda b}-\mathrm{e}^{\lambda a}}{\lambda},$$

$$\int \cos\lambda x\mathrm{d}x=\frac{\sin\lambda x}{\lambda}+C,\qquad \int_a^b \cos\lambda x\mathrm{d}x=\frac{\sin\lambda b-\sin\lambda a}{\lambda},$$

$$\int \sin\lambda x\mathrm{d}x=-\frac{\cos\lambda x}{\lambda}+C,\qquad \int_a^b \sin\lambda x\mathrm{d}x=-\frac{\cos\lambda b-\cos\lambda a}{\lambda},$$

$$\int \frac{1}{\cos^2 x}\mathrm{d}x=\tan x+C,\qquad \int_a^b \frac{1}{\cos^2 x}\mathrm{d}x=\tan b-\tan a,$$

$$\int \frac{1}{\sin^2 x}\mathrm{d}x=-\cot x+C,\qquad \int_a^b \frac{1}{\sin^2 x}\mathrm{d}x=\cot a-\cot b,$$

$$\int \frac{1}{1+x^2}\mathrm{d}x=\arctan x+C,\qquad \int_a^b \frac{1}{1+x^2}\mathrm{d}x=\arctan b-\arctan a,$$

$$\int \frac{1}{\sqrt{1-x^2}}\mathrm{d}x=\arcsin x+C,\qquad \int_a^b \frac{1}{\sqrt{1-x^2}}\mathrm{d}x=\arcsin b-\arcsin a,$$

$$\int \cosh x\mathrm{d}x=\sinh x+C,\qquad \cosh x=\frac{\mathrm{e}^x+\mathrm{e}^{-x}}{2},u$$

$$\int \sinh x\mathrm{d}x=\cosh x+C,\qquad \sinh x=\frac{\mathrm{e}^x-\mathrm{e}^{-x}}{2},$$

$$\int \frac{1}{\sqrt{x^2+1}}\mathrm{d}x=\sinh^{-1}x+C,\qquad \sinh^{-1}x=\ln(x+\sqrt{x^2+1}),$$

$$\int \frac{1}{\sqrt{x^2-1}}\mathrm{d}x=\cosh^{-1}x+C,\qquad \cosh^{-1}x=\ln(x+\sqrt{x^2-1}),\quad x>1.$$

(2) **线性**

定理 7.4.1. (线性) 设 $f,g\in R[a,b]$，则 $\lambda f+\mu g\in R[a,b]$，且

$$\int_a^b\big(\lambda f(x)+\mu g(x)\big)\mathrm{d}x=\lambda\int_a^b f(x)\mathrm{d}x+\mu\int_a^b g(x)\mathrm{d}x.$$

即 $R[a,b]$ 是一个线性空间，$\int_a^b: R[a,b]\to\mathbb{R}$ 是一个线性函数。

例 7.4.2. $\frac{x^{n+1}}{n+1}$ 是 x^n的原函数，所以任何多项式

$$P(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$$

有原函数，并且

$$\int_a^b P(x)\mathrm{d}x=a_0(b-a)+a_1\frac{b^2-a^2}{2}+\cdots+a_n\frac{b^{n+1}-a^{n+1}}{n+1},$$

$$\int P(x)\mathrm{d}x = C + a_0x + \frac{a_1x^2}{2} + \cdots + \frac{a_nx^{n+1}}{n+1}.$$

例 7.4.3. 设 $|\alpha| \neq |\beta|$。求 $\int_a^b \cos\alpha x\cos\beta x\mathrm{d}x$

解： 我们利用三角函数积化和差公式

$$\cos\alpha x\cos\beta x = \frac{\cos(\alpha+\beta)x + \cos(\alpha-\beta)x}{2}$$

把乘积转化为线性组合，进而利用线性计算积分。

$$\begin{aligned}
&\int_a^b \cos\alpha x\cos\beta x\mathrm{d}x \\
=&\frac{1}{2}\int_a^b \cos(\alpha+\beta)x\mathrm{d}x + \frac{1}{2}\int_a^b \cos(\alpha-\beta)x\mathrm{d}x \\
=&\frac{\sin(\alpha+\beta)b - \sin(\alpha+\beta)a}{\alpha+\beta} + \frac{\sin(\alpha-\beta)b - \sin(\alpha-\beta)a}{\alpha-\beta}.
\end{aligned}$$

□

例 7.4.4. $\int_a^b \cos^4 x\mathrm{d}x$

解： 我们用二倍角公式降低幂的次数并把被积函数写为三角函数的线性组合。

$$\begin{aligned}
\cos^4 x &= \left(\frac{1+\cos 2x}{2}\right)^2 = \frac{1}{4} + \frac{\cos 2x}{2} + \frac{\cos^2 2x}{4} \\
&= \frac{1}{4} + \frac{\cos 2x}{2} + \frac{1}{4}\left(\frac{1}{2} + \frac{\cos 4x}{2}\right) \\
&= \frac{3}{8} + \frac{\cos 2x}{2} + \frac{\cos 4x}{8}, \\
\int_a^b \cos^4 x\mathrm{d}x &= \left(\frac{3x}{8} + \frac{\sin 2x}{4} + \frac{\sin 4x}{32}\right)\Big|_a^b.
\end{aligned}$$

□

(3) **换元**

定理 7.4.5. (换元公式) 设 f 在区间 I 上 *Riemann* 可积，$\varphi \in \mathscr{C}^1[a,b]$ 满足 $\varphi[a,b] = [\alpha,\beta] \subseteq I$，则

$$\int_a^b f(\varphi(t))\varphi'(t)\mathrm{d}t = \int_{\varphi(a)}^{\varphi(b)} f(x)\mathrm{d}x,$$

等价地，

$$\int_{[a,b]} f(\varphi(t))|\varphi'(t)|\mathrm{d}t = \int_{[\alpha,\beta]} f(x)\mathrm{d}x,$$

证明. 我们在附加假设 $f \in C[a,b]$ 下证明定理结论。

考虑

$$F(u) = \int_a^u f(\varphi(t))\varphi'(t)\mathrm{d}t - \int_{\varphi(a)}^{\varphi(u)} f(x)\mathrm{d}x.$$

则 $F(a)=0$，并且对任意 $u\in[a,b]$，

$$F'(u)=f(\varphi(u))\varphi'(u)-f(\varphi(u))\varphi'(u)=0.$$

所以 $F(b)=F(a)=0$。 □

注 7.4.6. 上述换元公式不仅为定积分计算提供了一个转化手段，更重要的是它表明定积分不仅可以看成是一个函数在区间上的积分，也可以看成在一维直线上沿一条路径的积分。

我们把 $\varphi:[a,b]\to I$ 看成是以 $[a,b]$ 为参数空间、在 I 中的一条路径，它以 $\varphi(a)$ 为起点，以 $\varphi(b)$ 为终点，其中 $\varphi(a)<\varphi(b)$，$\varphi(b)<\varphi(a)$，$\varphi(a)=\varphi(b)$ 都有可能，且在 t 从 a 到 b 的过程中，$\varphi(t)$ 可以位于以 $\varphi(a)$ 和 $\varphi(b)$ 为端点的区间之外，也就是说以 $\varphi(a)$ 和 $\varphi(b)$ 为端点的区间有可能只是像集 $\varphi[a,b]$ 的一个真子集。

上式左端的积分就是 f 沿着路径 φ 的积分。

换元公式表明，在一维情形，沿路径的积分只于路径的起点和终点有关，与路径的选取无关：若 $\psi:[\alpha,\beta]\to I$ 是另一条 $\mathscr{C}^1$ 路径，满足 $\psi(\alpha)=\varphi(a)$，足 $\psi(\beta)=\varphi(b)$，则

$$\int_a^b f(\varphi(t))\varphi'(t)\mathrm{d}t=\int_\alpha^\beta f(\psi(s))(\psi'(s)\mathrm{d}s.$$

积分换元公式还有个物理解释。我们视 x 为空间位置，视 $f(x)$ 为一个一维力场，$x=\varphi(t)$ 为一个运动，其中 t 为时间，$\varphi'(t)$ 为速度。于是

$$\int_a^b f(\varphi(t))\varphi'(t)\mathrm{d}t$$

为在运动 $\varphi(t)$ 的过程中，力场 f 所做的功。如果 $f(x)$ 是一个连续力场，则 $f(x)$ 有原函数 $F(x)$，$F(x)$ 为势能函数。此时

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

因此换元公式表明一维连续力场做的功等于势能的变化。

例 7.4.7. $\int_a^b\cos^3 x\mathrm{d}x$

解：

$$\int_a^b\cos^3 x\mathrm{d}x=\int_a^b(1-\sin^2 x)\mathrm{d}\sin x=\left(\sin x-\frac{\sin^3 x}{3}\right)\Big|_a^b.$$

此方法可用于形如

$$\int_a^b f(\sin x)\cos x\mathrm{d}x$$

的积分。 □

例 7.4.8 (从积分定义对数函数、指数函数和幂函数)**.** 对正数 x，定义对数函数为

$$\ln(x)=\int_1^x \frac{1}{t}\mathrm{d}t.$$

证明

(1) 对任意正数 x,y，$\ln(xy)=\ln(x)+\ln(y)$。

(2) $\ln(1+x)=x+o(x),\quad x\to 0$。

(3) $\ln:(0,+\infty)\to\mathbb{R}$ 是 $\mathscr{C}^\infty$ 的严格增满射，$(\ln)'(x)=\frac{1}{x}$。

(4) $\ln:(0,+\infty)\to\mathbb{R}$ 有 $\mathscr{C}^\infty$ 反函数，记为 $\exp:\mathbb{R}\to(0,+\infty)$，称为指数函数。则 $\exp:\mathbb{R}\to(0,+\infty)$ 是一个 $\mathscr{C}^\infty$ 的严格增满射，满足 $\exp(x+y)=\exp(x)\exp(y)$，且 $\exp'(x)=\exp(x)$。

(5) 对任意实数 α 和任意正数 x 定义，$x^\alpha=\exp(\alpha\ln(x))$。则对任意实数 α,β 以及任意正数 x,y，

$$x^\alpha x^\beta=x^{\alpha+\beta},\quad (x^\alpha)^\beta=x^{\alpha\beta},\quad x^\alpha y^\alpha=(xy)^\alpha.$$

(6) 记 $\mathrm{e}=\exp(1)$，则 $\mathrm{e}^x=\exp(x)$，$\lim\limits_{n\to+\infty}\left(1+\frac{1}{n}\right)^n=\mathrm{e}$。

证明. (1) 利用积分换元公式[2]得

$$\ln(xy)-\ln(y)=\int_1^{xy}\frac{1}{t}\mathrm{d}t-\int_1^{y}\frac{1}{t}\mathrm{d}t=\int_y^{xy}\frac{1}{t}\mathrm{d}t\overset{t=ys}{=}\int_1^x\frac{1}{s}\mathrm{d}s=\ln(x).$$

(2) 利用后面介绍的积分保序性结论，对 $x>0$，

$$\frac{-x^2}{1+x}=\int_1^{1+x}\left(\frac{1}{1+x}-1\right)\mathrm{d}t<\ln(1+x)-x=\int_1^{1+x}\left(\frac{1}{t}-1\right)\mathrm{d}t<0.$$

对 $-1<x<0$，

$$\frac{x^2}{1+x}=\int_{1+x}^{1}\left(\frac{1}{1+x}-1\right)\mathrm{d}t>x-\ln(1+x)=\int_{1+x}^{1}\left(\frac{1}{t}-1\right)\mathrm{d}t>0.$$

从而 $\ln(1+x)=x+o(x),x\to 0$。

(3) 由定义知当 $x>1$ 时，$\ln(x)>0$。

对任意 $0<x<y$，由(1)知

$$\ln(y)=\ln(x)+\ln\left(\frac{y}{x}\right)>\ln(x).$$

所以 ln 是严格增函数。

由(1)知

$$\ln(2^n)=n\ln(2)>n\ln(1)=0.$$

所以 $\lim\limits_{n\to+\infty}\ln(2^n)=+\infty$。另一方面，$\lim\limits_{n\to+\infty}\ln(\frac{1}{2^n})=-\lim\limits_{n\to+\infty}\ln(2^n)=-\infty$。又根据微积分基本定理知 ln 连续，所以 $\ln:(0,+\infty)\to\mathbb{R}$ 是满射。

[2]事实上，也可以直接用 Riemann 积分定义验证

所以 $\ln : (0, +\infty) \to \mathbb{R}$ 有连续的反函数。

由微积分基本定理知 ln 可微，且

$$\ln'(x) = \frac{1}{x} > 0,$$

所以 ln 和它的反函数都是 $\mathscr{C}^\infty$ 的。

(4)是(1)和(3)的推论。

(5)

$$\begin{aligned}
x^\alpha x^\beta &= \exp(\alpha \ln(x)) \exp(\beta \ln(x)) = \exp((\alpha+\beta)\ln(x)) = x^{\alpha+\beta}, \\
(x^\alpha)^\beta &= \exp(\beta \ln(\exp(\alpha \ln(x)))) = \exp(\beta\alpha \ln(x)) = x^{\alpha\beta}, \\
x^\alpha y^\alpha &= \exp(\alpha \ln(x)) \exp(\alpha \ln(y)) = \exp(\alpha \ln(xy)) = (xy)^\alpha.
\end{aligned}$$

(6)

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

$$\left(1+\frac{1}{n}\right)^n = \exp\left(\ln\left(\left(1+\frac{1}{n}\right)^n\right)\right) \to \exp(1), \quad n \to +\infty.$$

$$\mathrm{e}^x = \exp(x \ln(\mathrm{e})) = \exp(x \cdot 1) = \exp(x).$$

□

例 7.4.9. $\int_a^b \frac{1}{(x-c)^n} \mathrm{d}x$

解：

$$\begin{aligned}
\int_a^b \frac{1}{(x-c)^n} \mathrm{d}x &= \int_{a-c}^{b-c} \frac{1}{y^n} \mathrm{d}y \qquad (y = x - c) \\
&= \begin{cases} \left.\frac{-1}{(n-1)y^{n-1}}\right|_{a-c}^{b-c}, & n > 1; \\ \ln|y|\big|_{a-c}^{b-c}, & n = 1. \end{cases} \\
&= \begin{cases} \frac{1}{n-1}\left(\frac{1}{(a-c)^{n-1}} - \frac{1}{(b-c)^{n-1}}\right), & n > 1; \\ \ln\left|\frac{b-c}{a-c}\right|, & n = 1. \end{cases}
\end{aligned}$$

□

例 7.4.10. $\int_a^b \frac{x}{(1+x^2)^n} \mathrm{d}x$

解：

$$\begin{aligned}
\int_a^b \frac{x}{(1+x^2)^n} \mathrm{d}x &= \frac{1}{2} \int_{1+a^2}^{1+b^2} \frac{1}{y^n} \mathrm{d}y \qquad (y = 1 + x^2) \\
&= \begin{cases} \left.\frac{-1}{2(n-1)y^{n-1}}\right|_{1+a^2}^{1+b^2}, & n > 1; \\ \frac{1}{2}\ln|y|\big|_{1+a^2}^{1+b^2}, & n = 1. \end{cases} \\
&= \begin{cases} \frac{1}{2(n-1)}\left(\frac{1}{(1+a^2)^{n-1}} - \frac{1}{(1+b^2)^{n-1}}\right), & n > 1; \\ \ln\sqrt{\frac{1+b^2}{1+a^2}}, & n = 1. \end{cases}
\end{aligned}$$

□

例 7.4.11. $\int_a^b \frac{1}{\cos x}\mathrm{d}x$

解：

$$\begin{aligned}\int_a^b \frac{1}{\cos x}\mathrm{d}x &= \int_a^b \frac{1}{\cos^2 x}\mathrm{d}\sin x\\ &= \int_a^b \frac{1}{1-\sin^2 x}\mathrm{d}\sin x\\ &= \int_{\sin a}^{\sin b} \frac{1}{1-y^2}\mathrm{d}y \qquad y=\sin x\\ &= \frac{1}{2}\int_{\sin a}^{\sin b} \frac{1}{1-y}+\frac{1}{1+y}\mathrm{d}y\\ &= \frac{1}{2}\ln\frac{1+\sin b}{1-\sin b}-\frac{1}{2}\ln\frac{1+\sin a}{1-\sin a}.\end{aligned}$$

□

例 7.4.12. $\int_0^{\frac{\pi}{2}} \frac{x}{\sin x+\cos x}\mathrm{d}x$

解：

$$\begin{aligned}I &= \int_{\frac{\pi}{2}}^{0} \frac{\frac{\pi}{2}-y}{\sin(\frac{\pi}{2}-y)+\cos(\frac{\pi}{2}-y)}\mathrm{d}(\frac{\pi}{2}-y) \qquad x=\frac{\pi}{2}-y\\ &= \int_0^{\frac{\pi}{2}} \frac{\frac{\pi}{2}-y}{\cos y+\sin y}\mathrm{d}y = \int_0^{\frac{\pi}{2}} \frac{\frac{\pi}{2}}{\cos x+\sin x}\mathrm{d}x - I,\end{aligned}$$

所以

$$\begin{aligned}I &= \frac{\pi}{4}\int_0^{\frac{\pi}{2}} \frac{1}{\sin x+\cos x}\mathrm{d}x = \frac{\pi}{4}\int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{2}\cos(x-\frac{\pi}{4})}\mathrm{d}x = \frac{\pi}{4\sqrt{2}}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \frac{1}{\cos u}\mathrm{d}u\\ &= \frac{\pi}{4\sqrt{2}}\ln\frac{1+\frac{\sqrt{2}}{2}}{1-\frac{\sqrt{2}}{2}} = \frac{\pi}{2\sqrt{2}}\ln(\sqrt{2}+1).\end{aligned}$$

□

注 7.4.13. $\frac{x}{\sin x+\cos x}$ 的原函数不是初等函数，所以不能直接利用 Newton-Leibniz 公式计算上述积分。

例 7.4.14. $\int_a^b \sqrt{1+x^2}\mathrm{d}x$

解:

$$\begin{aligned}
\int_a^b \sqrt{1+x^2}\mathrm{d}x &= \int_{\arctan a}^{\arctan b} \sqrt{1+\tan^2\theta}\mathrm{d}\tan\theta, \qquad x=\tan\theta \\
&= \int_{\arctan a}^{\arctan b} \frac{1}{\cos^3\theta}\mathrm{d}\theta \\
&= \int_{\arctan a}^{\arctan b} \frac{1}{(1-\sin^2\theta)^2}\mathrm{d}\sin\theta \\
&= \int_{\frac{a}{\sqrt{1+a^2}}}^{\frac{b}{\sqrt{1+b^2}}} \frac{1}{(1-y^2)^2}\mathrm{d}y \qquad y=\sin\theta, \quad \sin\arctan b = \frac{b}{\sqrt{1+b^2}} \\
&= \frac{1}{4}\int_{\frac{a}{\sqrt{1+a^2}}}^{\frac{b}{\sqrt{1+b^2}}} \frac{1}{(1+y)^2} + \frac{1}{(1-y)^2} + \frac{1}{1-y} + \frac{1}{1+y}\mathrm{d}y \\
&= \frac{1}{4}\left[\frac{1}{1-y} - \frac{1}{1+y} + \ln\frac{1+y}{1-y}\right]\Bigg|_{\frac{a}{\sqrt{1+a^2}}}^{\frac{b}{\sqrt{1+b^2}}}.
\end{aligned}$$

□

(4) **分部积分**

定理 7.4.15. (分部积分) 设 $f,g\in\mathscr{C}^1[a,b]$，则

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

证明. 令

$$F(t)=\int_a^t f(x)g'(x)\mathrm{d}x - \left[f(t)g(t)-f(a)g(a)-\int_a^t f'(x)g(x)\mathrm{d}x\right].$$

则 $F(a)=0$，

$$F'(t)=f(t)g'(t)-[f'(t)g(t)+f(t)g'(t)-f'(t)g(t)]=0.$$

所以 $F(b)=F(a)=0$。 □

例 7.4.16. 求 $\int_0^{\frac{\pi}{2}}\cos^n x\mathrm{d}x$。

解: 记 $I_n=\int_0^{\frac{\pi}{2}}\cos^n x\mathrm{d}x$。则

$$I_0=\int_0^{\frac{\pi}{2}}\mathrm{d}x=\frac{\pi}{2}, \quad I_1=\int_0^{\frac{\pi}{2}}\cos x\mathrm{d}x=1.$$

当 $n>1$ 时，

$$\begin{aligned}
I_n &= \int_0^{\frac{\pi}{2}}\cos^{n-1}x\mathrm{d}\sin x = -\int_0^{\frac{\pi}{2}}\sin x\mathrm{d}\cos^{n-1}x \\
&= (n-1)\int_0^{\frac{\pi}{2}}\sin^2 x\cos^{n-2}x\mathrm{d}x = (n-1)\int_0^{\frac{\pi}{2}}(1-\cos^2 x)\cos^{n-2}x\mathrm{d}x \\
&= (n-1)I_{n-2}-(n-1)I_n \\
&= \frac{n-1}{n}I_{n-2}
\end{aligned}$$

所以当 $n=2m$ 是偶数时，

$$I_n=\frac{(n-1)!!}{n!!}I_0=\frac{(n-1)!!}{n!!}\frac{\pi}{2}=\frac{n!}{(n!!)^2}\frac{\pi}{2}=\frac{(2m)!}{m!m!}\frac{\pi}{2^{2m+1}}.$$

当 $n=2m+1$ 是偶数时，

$$I_n=\frac{(n-1)!!}{n!!}I_1=\frac{(2m)!!}{(2m+1)!!}=\frac{2^{2m}m!m!}{(2m+1)!}.$$

□

例 7.4.17. $I_n=\int_a^b\frac{1}{(1+x^2)^n}\mathrm{d}x$

解：

$$I_1=\arctan b-\arctan a,$$

对 $n>1$，

$$\begin{aligned}I_{n-1}-I_n&=\int_a^b\frac{x^2}{(1+x^2)^n}\mathrm{d}x\\&=\frac{1}{2(1-n)}\int_a^b x\mathrm{d}\frac{1}{(1+x^2)^{n-1}}\\&=\frac{1}{2(1-n)}\left[x\frac{1}{(1+x^2)^{n-1}}\Big|_a^b-\int_a^b\frac{1}{(1+x^2)^{n-1}}\mathrm{d}x\right]\\&=\frac{1}{2(1-n)}\left[x\frac{1}{(1+x^2)^{n-1}}\Big|_a^b-I_{n-1}\right]\end{aligned}$$

于是

$$I_n=\frac{2n-3}{2n-2}I_{n-1}+\frac{x}{(2n-2)(1+x^2)^{n-1}}\Big|_a^b.$$

□

例 7.4.18. 设 f 在区间 I 上 $n+1$ 阶可微，且 $f^{(n+1)}\in R(I)$，则对 $a,x\in I$，

$$f(x)=f(a)+f'(a)(x-a)+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\int_a^x\frac{f^{(n+1)}(t)}{n!}(x-t)^n\mathrm{d}t.$$

证明. $n=1$ 时，

$$f(x)=f(a)+\int_a^x f'(t)\mathrm{d}t.$$

假设对正整数 n，

$$f(x)=f(a)+f'(a)(x-a)+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\int_a^x\frac{f^{(n+1)}(t)}{n!}(x-t)^n\mathrm{d}t.$$

则

$$\begin{aligned}f(x)&=f(a)+f'(a)(x-a)+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n-\int_a^x f^{(n+1)}(t)\mathrm{d}_t\frac{(x-t)^{n+1}}{(n+1)!}\\&=f(a)+\cdots+\frac{f^{(n)}(a)}{n+1!}(x-a)^{n+1}+\int_a^x f^{(n+2)}(t)\frac{(x-t)^{n+1}}{(n+1)!}\mathrm{d}t.\end{aligned}$$

所以定理结论成立。 □

有理函数积分

习题7.4

1. 求 $\int_0^1 \left(\int_x^1 \mathrm{e}^{-y^2}\mathrm{d}y\right)\mathrm{d}x$ 的值。(提示：分部积分)

2. 计算 $\int_0^1 \frac{x^4(1-x)^4}{1+x^2}\mathrm{d}x$ 并证明 $\pi < \frac{22}{7}$。对 π 的类似估计结果，见英文版维基百科词条"Proof of 22/7 exceeds π"。

7.5 定积分几何应用与物理应用

定义 7.5.1. 称 γ 为 $\mathbb{R}^n$ 中的一条 $\mathscr{C}^r$ **曲线**，如果存在 $\mathscr{C}^r$ 映射 $\mathbf{x}: I \to \mathbb{R}^n$(这里 I 是个区间)使得 $\gamma = \{\mathbf{x}(t)|t \in I\}$。此时称 $\mathbf{x}(t)$ 为 γ 的一个 $\mathscr{C}^r$ 参数表示。所谓 $\mathbf{x}: I \to \mathbb{R}^n$ 是 $\mathscr{C}^r$ 映射，即 $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$，其中每个函数 $x_k: [a,b] \to \mathbb{R}$ 都是 $\mathscr{C}^r$ 的。

称曲线 γ 的 $\mathscr{C}^r$ 参数表示 $\mathbf{x}(t)$ 是**正则的**，如果对任意 $t \in I$，$\mathbf{x}'(t) \neq 0$。

曲线的弧长

设 γ 是一条正则 $\mathscr{C}^1$ 曲线，$\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$($t \in [a,b]$)是 γ 的一个 $\mathscr{C}^1$ 正则参数表示。

对 $[a,b]$ 的任意划分 $P: a = t_0 < t_1 < \ldots < t_m = b$，记 $C_k = \mathbf{x}(t_k)$，则折线 $C_1C_2\ldots C_m$ 的长度为

$$\begin{aligned} L(\mathbf{x}, P) &= \sum_{k=1}^m |\mathbf{x}(t_k) - \mathbf{x}(t_{k-1})| \\ &= \sum_{k=1}^m \sqrt{(x_1(t_k) - x_1(t_{k-1}))^2 + \cdots + (x_m(t_k) - x_n(t_{k-1}))^2} \\ &= \sum_{k=1}^m \sqrt{x_1'(\xi_{1,k})^2 + \cdots + x_n'(\xi_{n,k})^2}(t_k - t_{k-1}). \end{aligned}$$

易见对区间 $[a,b]$ 的任意划分 P, Q，$L(\mathbf{x}, P) \leq L(\mathbf{x}, P \vee Q)$。

对任意 $\varepsilon > 0$，存在 $\delta > 0$ 使得对任意 $\xi, \eta \in [a,b]$ 满足 $|\xi - \eta| < \delta$，就有

$$|x_k'(\xi)^2 - x_k'(\eta)^2| < \varepsilon, \quad k = 1, 2, \ldots, n.$$

记 $\mathcal{P}_\delta$ 为区间 $[a,b]$ 满足 $\max\limits_{1\le k\le m}(t_k-t_{k-1})<\delta$ 的所有划分 P 组成的集合。因此对任意 $P\in\mathcal{P}_\delta$，

$$\begin{aligned}
&\left|L(\mathbf{x},P)-\sum_{k=1}^m\sqrt{x_1'(t_k)^2+\cdots+x_n'(t_k)^2}(t_k-t_{k-1})\right|\\
\le&\sum_{k=1}^m\left|\sqrt{x_1'(\xi_{1,k})^2+\cdots+x_n'(\xi_{n,k})^2}-\sqrt{x_1'(t_k)^2+\cdots+x_n'(t_k)^2}\right|(t_k-t_{k-1})\\
\le&\sum_{k=1}^m\frac{|x_1'(\xi_{1,k})^2-x_1'(t_k)^2|+\cdots+|x_n'(\xi_{n,k})^2-x_n'(t_k)^2|}{\sqrt{x_1'(t_k)^2+\cdots+x_n'(t_k)^2}}(t_k-t_{k-1})\\
\le&\frac{n\varepsilon}{\alpha}(b-a),
\end{aligned}$$

其中 $\alpha=\min\limits_{t\in[a,b]}\sqrt{x_1'(t)^2+\cdots+x_n'(t)^2}>0$。因此

$$s(\|\mathbf{x}'\|,P)-\frac{n\varepsilon}{\alpha}(b-a)\le L(x,P)\le S(\|\mathbf{x}'\|,P)+\frac{n\varepsilon}{\alpha}(b-a).$$

再取划分 P_ε 使得

$$\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t-\frac{n\varepsilon}{\alpha}(b-a)\le s(\|\mathbf{x}'\|,P_\varepsilon)\le S(\|\mathbf{x}'\|,P_\varepsilon)\le\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t+\frac{n\varepsilon}{\alpha}(b-a).$$

则对任意划分 Q，取划分 $P\in\mathcal{P}_\delta$，则 $Q\vee P\vee P_\varepsilon\in\mathcal{P}_\delta$，于是

$$\begin{aligned}
L(\mathbf{x},Q)&\le L(\mathbf{x},Q\vee P\vee P_\varepsilon)\\
&\le S(\|\mathbf{x}'\|,Q\vee P\vee P_\varepsilon)+\frac{n\varepsilon}{\alpha}(b-a)\\
&\le\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t+\frac{2n\varepsilon}{\alpha}(b-a),\\
L(\mathbf{x},Q\vee P\vee P_\varepsilon)&\ge s(\|\mathbf{x}'\|,Q\vee P\vee P_\varepsilon)-\frac{n\varepsilon}{\alpha}(b-a)\\
&\ge\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t-\frac{2n\varepsilon}{\alpha}(b-a),
\end{aligned}$$

所以

$$\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t-\frac{2n\varepsilon}{\alpha}(b-a)\le\sup_{P\in\mathcal{P}}L(\mathbf{x},P)\le\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t+\frac{2n\varepsilon}{\alpha}(b-a),$$

$$\left|\sup_{P\in\mathcal{P}}L(\mathbf{x},P)-\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t\right|\le\frac{2n\varepsilon}{\alpha}(b-a),$$

因此

$$\sup_{P\in\mathcal{P}}L(\mathbf{x},P)=\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t,$$

其中 $\mathcal{P}$ 是区间 $[a,b]$ 的所有划分组成的集合。

定义 7.5.2. $\mathscr{C}^1$ 曲线 $\gamma:\mathbf{x}(t)=(x_1(t),x_2(t),\ldots,x_n(t))$ $(a\le t\le b)$的弧长为

$$L(\gamma)=\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t=\int_a^b\sqrt{\langle\mathbf{x}'(t),\mathbf{x}'(t)\rangle}\mathrm{d}t=\int_a^b\sqrt{x_1'(t)^2+\cdots+x_n'(t)^2}\mathrm{d}t.$$

注 7.5.3. 1. 曲线的弧长与 $\mathbb{R}^m$ 中直角坐标系(即单位正交基)的选取无关。设 $\mathbf{y}(t)=A\mathbf{x}(t)+b$，其中 A 是正交矩阵，则 $\mathbf{y}'(t)=A\mathbf{x}'(t)$，于是

$$\int_a^b\|\mathbf{y}'(t)\|\mathrm{d}t=\int_a^b\|A\mathbf{x}'(t)\|\mathrm{d}t=\int_a^b\|\mathbf{x}'(t)\|\mathrm{d}t.$$

曲线的弧长与曲线的参数表示的选取无关。

2. 设 $t=t(s)$ 是区间 $[\alpha,\beta]$ 到 $[a,b]$ 的一一对应，且 $t(s)$ 和它的反函数 $s(t)$ 都是 $\mathscr{C}^1$ 的，$\mathbf{y}(s)=\mathbf{x}(t(s))$，则

$$\begin{aligned}\int_{[\alpha,\beta]}\|\mathbf{y}'(s)\|\mathrm{d}s&=\int_{[\alpha,\beta]}\|\mathbf{x}'(t(s))t'(s)\|\mathrm{d}s=\int_{[\alpha,\beta]}\|\mathbf{x}'(t(s))\||t'(s)|\mathrm{d}s\\&=\int_{[a,b]}\|\mathbf{x}'(t)\|\mathrm{d}t.\end{aligned}$$

上式最后一个等式利用了积分换元公式。

弧长参数：

$$l(t)=\int_a^t\|\mathbf{x}'(s)\|\mathrm{d}s,\quad \frac{\mathrm{d}l}{\mathrm{d}t}=\|\mathbf{x}'(t)\|>0,$$

所以 $l(t)$ 关于 t 严格增，称 l 为曲线 γ 的弧长参数。

$$\left\|\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l}\right\|=\left\|\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\cdot\frac{\mathrm{d}t}{\mathrm{d}l}\right\|=1.$$

于是

$$0=\frac{\mathrm{d}}{\mathrm{d}l}\left\langle\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l},\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l}\right\rangle=2\left\langle\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}l^2},\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l}\right\rangle,$$

如果视曲线的参数方程为运动方程，那么当时间为在弧长时，速率恒定为1，加速度向量总与速度向量垂直。

例 7.5.4 (圆周、弧长和三角函数)**.** 圆周 $x^2+y^2=R^2$ 可以写成 $x=\pm\sqrt{R^2-y^2}$ $(-R\le y\le R)$。对 $x>0$ 部分的半圆周 $x=\sqrt{R^2-y^2}$ $(-R\le y\le R)$，从点 $(R,0)$ 沿圆周到点 (x_0,y_0) 处的弧长为

$$l=\int_0^{y_0}\sqrt{1+\left(\frac{\mathrm{d}x}{\mathrm{d}y}\right)^2}\mathrm{d}y=\int_0^{y_0}\frac{R}{\sqrt{R^2-y^2}}\mathrm{d}y=R\int_0^{y_0/R}\frac{1}{\sqrt{1-t^2}}\mathrm{d}t,\quad y=Rt.$$

记

$$\arcsin u = \int_0^u \frac{1}{\sqrt{1-t^2}}\mathrm{d}t,$$

则

$$(\arcsin u)' = \frac{1}{1-u^2} > 0, \quad \forall u \in (-1,1),$$

所以 $\arcsin : (-1,1) \to \mathbb{R}$ 是严格增的可微函数，其反函数 $\sin$ 也是严格增的可微函数，并且

$$(\sin t)' = \frac{1}{(\arcsin u)'} = \sqrt{1-u^2} = \sqrt{1-\sin^2 t},$$
$$(\sin t)'' = \frac{-\sin t(\sin t)'}{\sqrt{1-\sin^2 t}} = -\sin t.$$

易见

$$l = R\arcsin\left(\frac{y_0}{R}\right),$$

从而

$$y_0 = R\sin\left(\frac{l}{R}\right), \quad x_0 = \sqrt{R^2-y_0^2} = R\sqrt{1-\sin^2\left(\frac{l}{R}\right)}.$$

这是半圆周 $x = \sqrt{R^2-y^2}$ $(-R \le y \le R)$由弧长表示的参数方程。 □

曲线的曲率

设 $\mathbf{x}(l)$ 为曲线 γ 在弧长参数下的参数表示。记 $\theta(l,\Delta l)$ 为曲线 γ 在 $\mathbf{x}(l)$ 处的切线与 $\mathbf{x}(l+\Delta l)$ 处的切线的夹角。则 $\theta(l,\Delta l)$ 是以 $\mathbf{x}'(l)$ 和 $\mathbf{x}'(l+\Delta l)$ 为腰的等腰三角形的顶角，于是

$$\sin\frac{\theta(l,\Delta l)}{2} = \frac{\|\mathbf{x}'(l+\Delta l)-\mathbf{x}'(l)\|}{2} = \frac{1}{2}\|\mathbf{x}''(l)\Delta l\| + o(\Delta l),$$

所以

$$\frac{\theta(l,\Delta l)}{2} + o(\theta(l,\Delta l)) = \frac{1}{2}\|\mathbf{x}''(l)\Delta l\| + o(\Delta l),$$

因此

$$\theta(l,\Delta l) = \|\mathbf{x}''(l)\|\Delta l + o(\Delta l).$$

于是在弧长参数下，沿曲线 γ 以单位速率前进时，曲线方向在 $\mathbf{x}(l)$ 处的瞬时变化率(称为 γ 在 $\mathbf{x}(l)$ 处的**曲率**)为

$$\kappa = \lim_{\Delta l\to 0}\frac{\theta(l,\Delta l)}{\Delta l} = \|\mathbf{x}''(l)\|.$$

对一般的参数表达 $\mathbf{x}(t)$，

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l}\cdot\frac{\mathrm{d}l}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l}\|\mathbf{x}'(t)\|,$$

$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}l^2}\|\mathbf{x}'(t)\|^2 + \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}l}\frac{\langle\mathbf{x}''(t),\mathbf{x}'(t)\rangle}{\|\mathbf{x}'(t)\|} = \frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}l^2}\|\mathbf{x}'(t)\|^2 + \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\frac{\langle\mathbf{x}''(t),\mathbf{x}'(t)\rangle}{\|\mathbf{x}'(t)\|},$$

所以
$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}l^2}=\frac{1}{\|\mathbf{x}'(t)\|^2}\left[\mathbf{x}''(t)-\left\langle\mathbf{x}''(t),\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\right\rangle\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\right].$$
因此曲率
$$\kappa=\frac{1}{\|\mathbf{x}'(t)\|^2}\left\|\mathbf{x}''(t)-\left\langle\mathbf{x}''(t),\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\right\rangle\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\right\|,$$
其中 $\mathbf{x}''(t)-\left\langle\mathbf{x}''(t),\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\right\rangle\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$ 是曲线 γ 位于 $\mathbf{x}(t)$ 处的加速度向量 $\mathbf{x}''(t)$ 在曲线法空间上的正交投影。

例 7.5.5. 直线 $\mathbf{x}(t)=\alpha t+\beta$ 的曲率 $\kappa=0$。

如果曲线 γ 的曲率为零，则考虑弧长参数下的参数方程 $\mathbf{x}(l)$，$\|\mathbf{x}''(l)\|=0$，从而 $\mathbf{x}''(l)=0$，因此 $\mathbf{x}(l)=\alpha l+\beta$。 □

例 7.5.6. 圆周 $x^2+y^2=R^2$ 在点$(x_0,y_0)(y_0>0)$附近等价于
$$y=\sqrt{R^2-x^2}.$$
则
$$\mathbf{x}'(x)=\left(1,\frac{-x}{\sqrt{R^2-x^2}}\right)=\left(1,\frac{-x}{y}\right),\quad\|\mathbf{x}'(x)\|=\frac{R}{y},$$
$$\mathbf{x}''(x)=\left(0,\frac{-R^2}{y^3}\right),$$
$$\mathbf{x}''(t)-\left\langle\mathbf{x}''(t),\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}\right\rangle\frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}=\left(\frac{-x}{y^2},\frac{-1}{y}\right),$$
于是曲率
$$\kappa=\frac{1}{R}.$$
类似可知在圆周的其他部分，曲率也是 $\frac{1}{R}$。 □

定理 7.5.7. *平面曲线 γ 具有非零的常数曲率当且仅当 γ 是圆周。*

证明. 充分性已经在上面例子中解决。下证必要性。设 γ 是一条平面曲线，曲率 κ 为非零常数。取 γ 的弧长参数表示 $z(l)$，则 $|z'(l)|=1$。于是设 $z'(l)=\mathrm{e}^{\mathrm{i}\varphi(l)}$。于是
$$\kappa=|z''(l)|=|\mathrm{e}^{\mathrm{i}\varphi(l)}\mathrm{i}\varphi'(l)|=|\varphi'(l)|.$$
因为 $\kappa\neq0$，$\varphi'(l)$具有介值性质，所以 $\varphi'(l)=\kappa(\ \forall l\ )$，或$\varphi'(l)=-\kappa(\ \forall l\ )$，因此 $\varphi(l)=\kappa l+\varphi_0(\ \forall l\ )$或$\varphi(l)=-\kappa l+\varphi_0(\ \forall l\ )$。

不妨设 $\varphi(l)=\kappa l+\varphi_0(\ \forall l\ )$。则 $z'(l)=\mathrm{e}^{\mathrm{i}\kappa l+\mathrm{i}\varphi_0}(\ \forall l\ )$，从而
$$z(l)=\frac{1}{\mathrm{i}\kappa}\mathrm{e}^{\mathrm{i}\kappa l+\mathrm{i}\varphi_0}+C_0$$
所以
$$|z(t)-C_0|=\frac{1}{\kappa}.$$
因此 γ 是圆周。 □

由封闭曲线围城的平面区域的面积

设 $f(x) \le g(x)(\ \forall x \in [a,b]\ )$。则曲线 $y=f(x)$ 与 $y=g(x)(\ a \le x \le b\ )$以及直线 $x=a$、$x=b$ 所围成的有界区域

$$D=\{(x,y)|a \le x \le b, f(x) \le y \le g(x)\}$$

的面积为

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

例 7.5.8. 设 $P(u,v)$ 是双曲线 $x^2-y^2=1$ 上一点，满足 $u,v>0$。求由线段 OP 、双曲线和 x 轴围成的平面有界区域的面积。

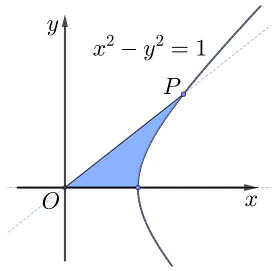


图 7.5:

解： 该区域为

$$\{(x,y)|0 \le y \le \frac{v}{\sqrt{1+v^2}}x, x^2-y^2 \le 1\},$$

即

$$0 \le \frac{\sqrt{1+v^2}}{v}y \le x \le \sqrt{1+y^2}, \quad 0 \le y \le v.$$

其面积为

$$\begin{aligned}
&\int_0^v \left(\sqrt{1+y^2}-\frac{\sqrt{1+v^2}}{v}y\right)\mathrm{d}y \\
=&\int_0^{\sinh^{-1} v}\left(\cosh t-\frac{\sqrt{1+v^2}}{v}\sinh t\right)\mathrm{d}\sinh t \qquad y=\sinh t \\
=&\int_0^{\sinh^{-1} v}\cosh^2 t\mathrm{d}t-\frac{\sqrt{1+v^2}}{2v}\sinh^2 t\Big|_0^{\sinh^{-1} v} \\
=&\int_0^{\sinh^{-1} v}\frac{\mathrm{e}^{2t}+2+\mathrm{e}^{-2t}}{4}\mathrm{d}t-\frac{v\sqrt{1+v^2}}{2} \\
=&\left(\frac{\sinh t\cosh t}{2}+\frac{t}{2}\right)\Big|_0^{\sinh^{-1} v}-\frac{v\sqrt{1+v^2}}{2}=\frac{\sinh^{-1} v}{2}.
\end{aligned}$$

□

注 7.5.9. 把上述结论与圆 $x^2+y^2=1$ 时的相应结论做对比，你应该理解 $\sinh, \cosh$ 被称为**双曲正弦函数**和**双曲余弦函数**的原因。

设 $\gamma:(x(t),y(t))(\,a\le t\le b\,)$是一条简单封闭曲线，即

$$\begin{cases} x(t)=x(s),\\ y(t)=y(s),\\ t<s \end{cases} \iff \begin{cases} t=a,\\ s=b. \end{cases}$$

Jordan定理指出：平面上每条简单封闭曲线都把平面分成三个互不相交的部分：曲线自身、一个有界区域以及一个无界区域。我们通常称其中有界的区域为曲线所围的内部区域，称无界的区域为曲线的外部区域。对 $\mathscr{C}^1$ 封闭曲线，其指向外部区域的法向量称为外法向量，指向内部区域的法向量称为内法向量。当你沿曲线移动时，如果内部区域总是在你的左侧，那么你前进的方向就是曲线的自然正向，反之，你前进的方向就是曲线自然正向的反向。对曲线 γ 的参数化表达 $(x(t),y(t))$，如果 t 增加时，点 $(x(t),y(t))$ 沿 γ 的自然正向移动，那么它就是一个符合自然正向的参数化表达，否则它是一个与自然正向相悖的参数化表达。

定理 7.5.10. *设 γ 是一条简单封闭的 $\mathscr{C}^1$ 平面曲线。则对于 γ 的符合自然正向的 $\mathscr{C}^1$ 参数化表达$(x(t),y(t))(\,a\le t\le b\,)$，积分*

$$-\int_a^b y(t)x'(t)\mathrm{d}t=\int_a^b x(t)y'(t)\mathrm{d}t=\frac{1}{2}\int_a^b x(t)y'(t)-y(t)x'(t)\mathrm{d}t,$$

且它们的值为 γ 所围成的有界区域的面积。我们依次记上述积分为

$$-\int_\gamma y\mathrm{d}x,\quad \int_\gamma x\mathrm{d}y,\quad \frac{1}{2}\int_\gamma x\mathrm{d}y-y\mathrm{d}x.$$

注 7.5.11. (a) $\int_\gamma y\mathrm{d}x$ 的值与曲线 γ 的同向参数化形式无关，但与 γ 的方向有关，它是一个有向积分。设 $t=t(s)$，$(\tilde{x}(s),\tilde{y}(s))=(x(t(s)),y(t(s)))$。则

$$\int_\alpha^\beta \tilde{y}(s)\tilde{x}'(s)\mathrm{d}s=\int_\alpha^\beta y(t(s))x'(t(s))t'(s)\mathrm{d}s=\int_{t(\alpha)}^{t(\beta)} y(t)x'(t)\mathrm{d}t.$$

(b) 对平面封闭曲线 γ，$-\int_\gamma y\mathrm{d}x=\int_\gamma x\mathrm{d}y$，

$$\int_a^b y(t)x'(t)\mathrm{d}t+\int_a^b x(t)y'(t)\mathrm{d}t=\int_a^b \big((x(t)y(t)\big)'\mathrm{d}t=x(b)y(b)-x(a)y(a)=0.$$

(c) 对平面封闭曲线 γ，$\int_\gamma x\mathrm{d}y-y\mathrm{d}x$ 的值与平面上直角坐标系的选取无关。设

$$(X(t),Y(t))=(x(t)\cos\theta+y(t)\sin\theta+A,-x(t)\sin\theta+y(t)\cos\theta+B).$$

则

$$\begin{aligned}
&X(t)Y'(t)-Y(t)X'(t)\\
=&(x(t)\cos\theta+y(t)\sin\theta+A)(-x'(t)\sin\theta+y'(t)\cos\theta)\\
&\quad-(-x(t)\sin\theta+y(t)\cos\theta+B)(x'(t)\cos\theta+y'(t)\sin\theta)\\
=&x(t)y'(t)-y(t)x'(t)-(A\sin\theta+B\cos\theta)x'(t)+(A\cos\theta-B\sin\theta)y'(t).
\end{aligned}$$

沿封闭曲线 γ，

$$\int_a^b(A\sin\theta+B\cos\theta)x'(t)\mathrm{d}t=(A\sin\theta+B\cos\theta)(x(b)-x(a))=0.$$

同理，

$$\int_a^b(A\cos\theta-B\sin\theta)y'(t)\mathrm{d}t=0.$$

所以

$$\int_a^b X(t)Y'(t)-Y(t)X'(t)\mathrm{d}t=\int_a^b x(t)y'(t)-y(t)x'(t)\mathrm{d}t.$$

(d) 设 $f(x)<g(x)(\ \forall x\in[a,b]\ )$。则平面区域

$$D=\{(x,y)|a\le x\le b,f(x)\le y\le g(x)\}$$

由四条曲线

$$\gamma_1:\begin{cases}x(t)=t,\\ y(t)=f(t),\end{cases}\quad(a\le t\le b),$$

$$\gamma_2:\begin{cases}x(t)=b,\\ y(t)=(1-t)f(b)+tg(b),\end{cases}\quad(0\le t\le 1),$$

$$\gamma_3:\begin{cases}x(t)=a+b-t,\\ y(t)=g(a+b-t),\end{cases}\quad(a\le t\le b),$$

$$\gamma_4:\begin{cases}x(t)=a,\\ y(t)=(1-t)g(a)+tf(a),\end{cases}\quad(0\le t\le 1),$$

顺次相连组成的简单闭曲线围成，于是

$$\begin{aligned}
&-\int_{\gamma_1}y\mathrm{d}x-\int_{\gamma_2}y\mathrm{d}x-\int_{\gamma_3}y\mathrm{d}x-\int_{\gamma_4}y\mathrm{d}x\\
=&-\int_a^b f(t)\mathrm{d}t-\int_a^b g(a+b-t)(-1)\mathrm{d}t\\
=&\int_a^b g(x)\mathrm{d}x-\int_a^b f(x)\mathrm{d}x.
\end{aligned}$$

(e) 如果曲线 γ 在极坐标系下的方程为 $r = r(\theta)(\ \alpha \le \theta \le \beta\ )$，则它在直角坐标系下的方程为

$$x(\theta) = r(\theta)\cos\theta, \quad y(\theta) = r(\theta)\sin\theta,$$

则

$$\mathrm{d}x(\theta) = \cos\theta \mathrm{d}r(\theta) - r(\theta)\sin\theta\mathrm{d}\theta, \quad \mathrm{d}y(\theta) = \sin\theta\mathrm{d}r(\theta) + r(\theta)\cos\theta\mathrm{d}\theta,$$

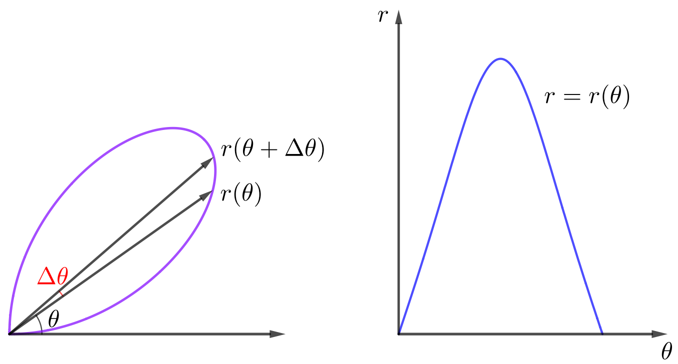
从而

$$x(\theta)\mathrm{d}y(\theta) - y(\theta)\mathrm{d}x(\theta) = r(\theta)^2\mathrm{d}\theta,$$

所以 γ 和射线 $\theta = \alpha$ 以及 $\theta = \beta$ 围成的有界区域的面积为

$$\frac{1}{2}\int_\alpha^\beta r(\theta)^2\mathrm{d}\theta.$$

例 7.5.12. 平面直角坐标系中的曲线 $x^3 + y^3 = 3xy$ 在第一象限围成了一个有界区域 D，求 D 的面积。



解： 取极坐标 (r, θ)，把 $x = r\cos\theta, y = r\sin\theta$ 代入曲线方程得到极坐标方程

$$r^3(\cos^3\theta + \sin^3\theta) = 3r^2\cos\theta\sin\theta.$$

在第一象限中 $0 \le \theta \le \frac{\pi}{2}$，$\cos^3\theta + \sin^3\theta > 0$，于是解得

$$r(\theta) = \frac{3\cos\theta\sin\theta}{\cos^3\theta + \sin^3\theta}, \quad 0 \le \theta \le \frac{\pi}{2}.$$

从而区域 D 的面积为

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} \frac{1}{2}r(\theta)^2\mathrm{d}\theta &= \int_0^{\frac{\pi}{2}} \frac{1}{2}\left(\frac{3\cos\theta\sin\theta}{\cos^3\theta + \sin^3\theta}\right)^2\mathrm{d}\theta \\
&= \int_0^{\frac{\pi}{4}} \left(\frac{3\cos\theta\sin\theta}{\cos^3\theta + \sin^3\theta}\right)^2\mathrm{d}\theta \quad (\text{曲线关于对角线 } y = x \text{ 对称}) \\
&= \int_0^{\frac{\pi}{4}} \left(\frac{3\tan\theta}{1 + \tan^3\theta}\right)^2\frac{1}{\cos^2\theta}\mathrm{d}\theta \quad (\text{换元 } t = 1 + \tan^3\theta) \\
&= \int_1^2 \frac{3}{t^2}\mathrm{d}t = \frac{3}{2}.
\end{aligned}$$

上述积分过程中使用了曲线的对称性，从而避免了后续积分中出现广义积分。

□

例 7.5.13. (**有心力场与 Kepler 第二定律**) $\mathbb{R}^n$ 中形如

$$F(x) = f(r)x, \quad r = \|x\|$$

的力场称为**有心力场**。即在 x 处的力$F(x)$ 的方向总是指向(或背离)一个固定点(选为坐标系原点)。

根据 Newton 第二定律，有心力场作用下质点的运动$x(t)$服从以下二阶微分方程

$$x'' = F(x) = f(\|x\|)x.$$

我们首先证明这个运动总是在一个固定的平面中。为此任取与 $x(t_0)$ 和 $x'(t_0)$ 正交的向量 $\mathbf{n}$。记 $y(t) = \langle x(t), \mathbf{n}\rangle$，则 $y(t)$ 是以下初值问题的解

$$\begin{cases} y'' = \langle x'', \mathbf{n}\rangle = f(\|x(t)\|)y. \\ y(t_0) = 0, \quad y'(t_0) = 0. \end{cases}$$

由线性微分方程初值问题解的存在唯一性，$y(t) = 0(\ \forall t\ )$。因此 $x(t)$ 永远位于由 $x(t_0)$ 和 $x'(t_0)$ 所确定的平面内。

在该平面内引入极坐标，并令 $z(t) = r(t)\mathrm{e}^{\mathrm{i}\theta(t)}$，则

$$z''(t) = r''(t)\mathrm{e}^{\mathrm{i}\theta(t)} + 2r'(t)\mathrm{e}^{\mathrm{i}\theta(t)}\mathrm{i}\theta'(t) + r(t)\mathrm{e}^{\mathrm{i}\theta(t)}\mathrm{i}^2\theta'(t)^2 + r(t)\mathrm{e}^{\mathrm{i}\theta(t)}\mathrm{i}\theta''(t),$$

代入

$$z''(t) = f(r(t))z(t)$$

得到

$$\begin{cases} r''(t) - r(t)\theta'(t)^2 = f(r(t))r(t), \\ 2r'(t)\theta'(t) + r(t)\theta''(t) = 0. \end{cases}$$

由第二个方程得到

$$\left[r(t)^2\theta'(t)\right]' = 0,$$

从而

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}r(t)^2\theta'(t) = 常数$$

其中 $A(t)$ 为在时段 $[t_0, t]$ 中连接原点到 $x(t)$ 的线段扫过的平面区域的面积，这就是 Kepler 第二定律。 □

旋转体的体积

设 $y = f(x)(\ a \le x \le b\ )$是直角坐标系中的一段连续曲线。它绕 x 轴旋转形成三维空间里一个旋转面 $\sqrt{y^2+z^2} = |f(x)|$，它围成一个有界区域

$$\Omega = \{(x, y, z) | \sqrt{y^2+z^2} \le |f(x)|, a \le x \le b\}.$$

我们计算 Ω 的体积。

任给 $[a,b]$ 一个划分 $P: x_0=a<x_1<\ldots<x_{n-1}<x_n=b$，在三维空间中用平面 $x=x_k$把 Ω 分成若干段，

$$\Omega_k=\{(x,y,z)|\sqrt{y^2+z^2}\le|f(x)|,x_{k-1}\le x\le x_k\}$$

是其中一段。取 M_k 和 m_k 分别是 $|f|$ 在 $[x_{k-1},x_k]$ 上的最大值和最小值，则

$$\begin{aligned}&\{(x,y,z)|\sqrt{y^2+z^2}\le m_k,x_{k-1}\le x\le x_k\}\subseteq\Omega_k\\&\qquad\qquad\subseteq\{(x,y,z)|\sqrt{y^2+z^2}\le M_k,x_{k-1}\le x\le x_k\},\end{aligned}$$

因此 Ω_k 的体积

$$\pi m_k^2\Delta x_k\le V(\Omega_k)\le\pi M_k^2\Delta x_k.$$

从而

$$\sum_{k=1}^n\pi m_k^2\Delta x_k\le V(\Omega)\le\sum_{k=1}^n\pi M_k^2\Delta x_k.$$

因此

$$V(\Omega)=\int_a^b\pi|f(x)|^2\mathrm{d}x.$$

这相当于在 x 处取 Ω 的一个正交于 x 轴的截面，再沿 x 轴取一段长度为 $\mathrm{d}x$ 的高，所成的柱体体积为 $\pi|f(x)|^2\mathrm{d}x$，上述积分相当于把这些无穷小柱体体积加起来得到 Ω 的体积。

设平面区域 D 由参数方程 $x=x(t),y=y(t)$ $(a\le t\le b)$刻画的简单封闭曲线围成，则 D 绕 x 轴旋转所得旋转体的体积为

$$\left|\int_a^b\pi y(t)^2\mathrm{d}x(t)\right|.$$

例如，如果曲线 γ 在极坐标系下的方程为 $r=r(\theta)$ $(\alpha\le\theta\le\beta)$，则 γ 和射线 $\theta=\alpha$ 以及 $\theta=\beta$ 围成的平面区域绕 x 轴旋转所得的旋转体体积为

$$\begin{aligned}&\left|\int_\alpha^\beta\pi y(\theta)^2\mathrm{d}x(\theta)-\int_0^{r(\beta)}\pi y(r)^2\mathrm{d}x(r)+\int_0^{r(\alpha)}\pi y(r)^2\mathrm{d}x(r)\right|\\=&\left|\int_\alpha^\beta\pi(r(\theta)\sin\theta)^2\mathrm{d}(r(\theta)\cos\theta)-\left.\frac{r(\theta)^3\sin^2\theta\cos\theta}{3}\right|_\alpha^\beta\right|\\=&\left|\int_\alpha^\beta\frac{2\pi}{3}r(\theta)^3\sin\theta\mathrm{d}\theta\right|.\end{aligned}$$

用上述方法我们可以类似求得 n 维欧氏空间中球的体积。

记 $V_n(R)$ 为 $\mathbb{R}^n$ 中半径为 R 的球 $x_1^2+\cdots+x_n^2\le R^2$ 的体积。则 $V_1(R)=2R$；对 $n>1$，在 $x_n=t$ 处，截面为

$$x_1^2+\cdots+x_{n-1}^2\le R^2-t^2,$$

n	1	2	3
$V_n(R)$	$2R$	πR^2	$\frac{4\pi}{3}R^3$
$A_n(R)$		$2\pi R$	$4\pi R^2$

它的 $n-1$ 维体积为 $V_{n-1}(\sqrt{R^2-t^2})$，于是

$$V_n(R)=\int_{-R}^{R}V_{n-1}(\sqrt{R^2-t^2})\mathrm{d}t=\int_{-1}^{1}V_{n-1}(R\sqrt{1-u^2})R\mathrm{d}u.$$

于是可以由数学归纳法证明 $V_n(R)=C_nR^n$。因此 $C_1=2$，

$$C_n=C_{n-1}\int_{-1}^{1}(1-u^2)^{\frac{n-1}{2}}\mathrm{d}u=2C_{n-1}\int_0^{\frac{\pi}{2}}\cos^n\theta\mathrm{d}\theta=C_{n-1}\frac{(n-1)!!}{n!!}b_n,$$

其中 $b_n=\begin{cases}\pi, & n\text{是偶数}\\ 2, & n\text{是奇数}.\end{cases}$ 所以

$$C_n=\frac{C_n}{C_{n-1}}\cdots\frac{C_2}{C_1}C_1=\frac{1}{n!!}b_n\cdots b_2C_1=\frac{1}{n!!}\pi^{[\frac{n}{2}]}2^{n-[\frac{n}{2}]},$$

所以

$$V_n(R)=\frac{2^n}{n!!}\left(\frac{\pi}{2}\right)^{[\frac{n}{2}]}R^n.$$

设 $A_n(R)$ 为半径为 R 的球$x_1^2+\cdots+x_n^2\le R^2$ 的表面积。则由薄球壳的体积

$$V_n(R+\Delta R)-V_n(R)=A_n(R)\Delta R+o(\Delta R)$$

得到

$$A_n(R)=\frac{\mathrm{d}V_n(R)}{\mathrm{d}R}=\frac{2^n}{(n-2)!!}\left(\frac{\pi}{2}\right)^{[\frac{n}{2}]}R^{n-1}.$$

旋转体的侧面积

平面直角坐标系中的 $\mathscr{C}^1$ 曲线 γ 绕 x 轴旋转得到一个旋转面，其中对应于从 x 到 $x+\mathrm{d}x$ 的部分近似为一个圆台，其面积为：

$$\frac{1}{2}2\pi(y+\mathrm{d}y)(l+\mathrm{d}l)-\frac{1}{2}2\pi yl=\pi(y\mathrm{d}l+l\mathrm{d}y+\mathrm{d}y\mathrm{d}l)=2\pi y\mathrm{d}l+o(\mathrm{d}l).$$

于是旋转面的面积为

$$\int_\gamma 2\pi y\mathrm{d}l=\int_a^b 2\pi y(t)\sqrt{x'(t)^2+y'(t)^2}\mathrm{d}t.$$

当曲线 γ 为函数图像 $y=f(x)$($a\le x\le b$)时，旋转面面积为

$$\int_a^b 2\pi|f(x)|\sqrt{1+f'(x)^2}\mathrm{d}x.$$

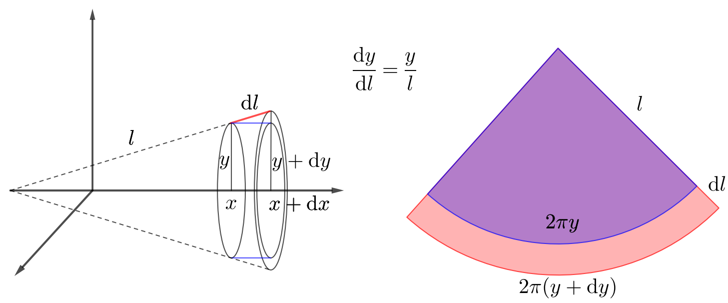


图 7.6:

当曲线 γ 极坐标方程为 $r=r(\theta)(\ a\le\theta\le b\ )$时，

$$\mathrm{d}l=\sqrt{[(r(\theta)\cos\theta)']^2+[(r(\theta)\sin\theta)']^2}\mathrm{d}\theta=\sqrt{r(\theta)^2+r'(\theta)^2}\mathrm{d}\theta,$$

从而旋转面面积为

$$\int_a^b 2\pi|r(\theta)\sin\theta|\sqrt{r(\theta)^2+r'(\theta)^2}\mathrm{d}\theta.$$

例 7.5.14. $\mathbb{R}^3$ 中半径为 R 的二维球面的表面积。对

$$x(\theta)=R\cos\theta,\quad y(\theta)=R\sin\theta,\quad (0\le\theta\le\pi),$$

成立

$$A=\int_0^\pi 2\pi R\sin\theta\sqrt{R^2\sin^2\theta+R^2\cos^2\theta}\mathrm{d}\theta=4\pi R^2.$$

□

例 7.5.15. 已知线密度，求质量、质心。

解： 质心是一个物体的各点按其所在位置的质量分布进行加权平均得到的平均位置。

曲线 $\gamma(t)=(x(t),y(t),z(t))$，线密度 $\rho=\rho(x,y,z)$。质量

$$m=\int_\gamma \rho\mathrm{d}l=\int_a^b\rho(\gamma(t))\|\gamma'(t)\|\mathrm{d}t=\int_a^b\rho(\gamma(t))\sqrt{x'(t)^2+y'(t)^2+z'(t)^2}\mathrm{d}t,$$

质心是位置向量按质量分布所做的加权平均，所以

$$\begin{pmatrix}\bar{x}\\ \bar{y}\\ \bar{z}\end{pmatrix}=\frac{\int_\gamma\gamma(t)\rho\mathrm{d}l}{\int_\gamma\rho\mathrm{d}l}=\frac{1}{\int_a^b\rho(\gamma(t))\|\gamma'(t)\|\mathrm{d}t}\begin{pmatrix}\int_a^b x(t)\rho(\gamma(t))\|\gamma'(t)\|\mathrm{d}t\\ \int_a^b y(t)\rho(\gamma(t))\|\gamma'(t)\|\mathrm{d}t\\ \int_a^b z(t)\rho(\gamma(t))\|\gamma'(t)\|\mathrm{d}t\end{pmatrix}$$

□

例 7.5.16. 求位于 $(a,0,0)$ 处的质点 m 受到密度为 1 的球壳 $x^2+y^2+z^2=1$ 的引力。

解： 设 $a>1$。由对称性，向量 F 应该与 x 轴平行。在 (x,y) 平面中半圆周 $x=\cos\theta, y=\sin\theta$($0\le\theta\le\pi$)绕 x 轴旋转得到三维空间中的单位球面，于是

$$\begin{aligned} F &= \int_0^\pi \frac{a-\cos\theta}{\sqrt{(a-\cos\theta)^2+\sin^2\theta}} \frac{Gm2\pi\sin\theta\mathrm{d}\theta}{(a-\cos\theta)^2+\sin^2\theta} \\ &= -2\pi mG\int_0^\pi \frac{(a-\cos\theta)\mathrm{d}\cos\theta}{(a^2+1-2a\cos\theta)^{3/2}} \\ &= \frac{G\pi m}{2a^2}\int_{(a-1)^2}^{(a+1)^2} \frac{(t+a^2-1)\mathrm{d}t}{t^{3/2}} \qquad (t=a^2+1-2a\cos\theta) \\ &= \frac{4\pi Gm}{a^2}. \end{aligned}$$

这相当于把球壳质量全部集中在球心时质点 m 受到的引力。

当 $0<a\le 1$ 时，引力的计算留作练习。 □

习题7.5

1. 设 f 是一元可微函数。求曲线 $y=f(x)$($a\le x\le b$)的弧长与曲率。

2. 设平面曲线 γ 的极坐标方程为 $r=r(\theta)$($\alpha\le\theta\le\beta$)，求 γ 的弧长与曲率。

3. 求三维欧氏空间中等距螺旋线 $x=R\cos t, y=R\sin t, z=ht$ ($0\le t\le 2\pi$)的弧长和曲率。

4. 证明旋转体体积 $\int_a^b \pi y(t)^2\mathrm{d}x(t)$ 和旋转面面积 $\int_a^b 2\pi y(t)\sqrt{x'(t)^2+y'(t)^2}\mathrm{d}t$ 的值与曲线 γ 的参数表示 $(x(t),y(t))$ 的选择无关。

5. 求圆周 $x^2+(y-a)^2=b^2$ ($a>b>0$)绕 x 轴旋转所得的曲面的面积以及该曲面所围成的有界区域的体积。

6. 求星形线 $x^{2/3}+y^{2/3}=a^{2/3}$ $(a>0)$的弧长、所围有界区域的面积、绕 x 轴旋转所得旋转体的体积和旋转面的面积。

7. 心脏线的极坐标方程为 $r=a(1+\cos\theta)(-\pi\le\theta\le\theta)$。求

   (a) 心脏线的弧长；

   (b) 心脏线在弧长参数下的方程，以及其各点处的曲率；

   (c) 心脏线的质心，假设线密度为1；

   (d) 心脏线所围成的平面有界区域的面积，该区域的质心(假设面密度为1)；

(e) 心脏线绕其对称轴所成的曲面的面积，曲面的质心(假设面密度为1)；

(f) 心脏线绕其对称轴所围成的三维区域的体积，该区域的质心(假设体密度为1)。

7.6 广义积分的概念

例 7.6.1. 求半径为 R 的圆所围区域的面积以及周长。

解： $x^2+y^2=R^2$，考虑 $y=\sqrt{R^2-x^2}(\,-R\le x\le R\,)$。由对称性

$$A=2\int_{-R}^{R}\sqrt{R^2-x^2}\mathrm{d}x=4R^2\int_0^1\sqrt{1-t^2}\mathrm{d}t.$$

记 $\pi=4\int_0^1\sqrt{1-t^2}\mathrm{d}t$。则 $A=\pi R^2$。

$$P=2\int_{-R}^{R}\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x=2\int_{-R}^{R}\sqrt{1+\frac{x^2}{R^2-x^2}}\mathrm{d}x=4R\int_0^1\sqrt{\frac{1}{1-t^2}}\mathrm{d}t.$$

因为

$$\begin{aligned}\int_0^1\sqrt{1-t^2}\mathrm{d}t&=t\sqrt{1-t^2}\Big|_0^1-\int_0^1 t\mathrm{d}\sqrt{1-t^2}\\&=\int_0^1\frac{t^2}{\sqrt{1-t^2}}\mathrm{d}t\\&=\int_0^1\frac{1}{\sqrt{1-t^2}}\mathrm{d}t-\int_0^1\sqrt{1-t^2}\mathrm{d}t\\&=\frac{1}{2}\int_0^1\frac{1}{\sqrt{1-t^2}}\mathrm{d}t\end{aligned}$$

所以

$$4\int_0^1\frac{1}{\sqrt{1-t^2}}\mathrm{d}t=8\int_0^1\sqrt{1-t^2}\mathrm{d}t=2\pi,$$

从而 $P=2\pi R$。 □

请找出上例论证中的问题。

例 7.6.2. 把地球表面质量为 m 的物体发送到距离地球无限远的太空，至少需要多大的初始速度？

解： 动能的改变量等于引力做的功

$$\frac{1}{2}mv_\infty^2-\frac{1}{2}mv_0^2=-\int_R^{+\infty}\frac{GMm}{x^2}\mathrm{d}x=\frac{GMm}{x}\Big|_R^{+\infty}=-\frac{GMm}{R},$$

所以

$$v_0=\sqrt{\frac{2GM}{R}+\frac{1}{2}mv_\infty^2}\ge\sqrt{\frac{2GM}{R}}.$$

□

上面两个例题中的积分都不是 Riemann 积分，因为 Riemann 积分中既要求积分区间是有界区间(第二个例子中的积分区间是 $[R,+\infty)$)，同时要求被积函数是有界函数(第一个例子中的被积函数 $\frac{1}{\sqrt{1-t^2}}$ 在区间 $(0,1]$ 上是无界的)。这说明 Riemann 积分是有缺陷的，它限制了被积函数以及积分区域，因此我们需要推广积分的概念。

定义 7.6.3. 设函数 $f:[a,+\infty)\to\mathbb{R}$ 满足：

1. 对任何 $A>a$，f 在区间 $[a,A]$ 上 Riemann 可积；并且
2. 极限 $\lim\limits_{A\to+\infty}\int_a^A f(x)\mathrm{d}x$ 收敛。

则记

$$\int_a^{+\infty} f(x)\mathrm{d}x=\lim_{A\to+\infty}\int_a^A f(x)\mathrm{d}x,$$

称它为 f 在区间 $[a,+\infty)$ 上的**广义积分**，称 f 在区间 $[a,+\infty)$ 上**广义可积**或**广义积分** $\int_a^{+\infty} f(x)\mathrm{d}x$ **收敛**。

类似定义广义积分 $\int_{-\infty}^a f(x)\mathrm{d}x$。

f 在区间 $(-\infty,+\infty)$ 上**广义可积**或**广义积分**$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x$**收敛**，如果两个广义积分 $\int_{-\infty}^a f(x)\mathrm{d}x$ 和 $\int_a^{+\infty} f(x)\mathrm{d}x$ 都收敛，并且记

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=\int_{-\infty}^a f(x)\mathrm{d}x+\int_a^{+\infty} f(x)\mathrm{d}x.$$

定义 7.6.4. 设函数 $f:[a,b]\to\mathbb{R}$ 满足：

1. f 在区间 $[a,b]$ 上无界；
2. 对任何 $c\in(a,b)$，f 在区间 $[a,c]$ 上 Riemann 可积；并且
3. 极限 $\lim\limits_{c\to b^-}\int_a^c f(x)\mathrm{d}x$ 收敛。

则记

$$\int_a^b f(x)\mathrm{d}x=\lim_{c\to b^-}\int_a^c f(x)\mathrm{d}x,$$

称它为 f 在区间 $[a,b]$ 上的**瑕积分**，b 为 f 的一个**瑕点**，称 f 在区间 $[a,b]$ 上**广义可积**或**瑕积分** $\int_a^b f(x)\mathrm{d}x$ **收敛**。

定义 7.6.5. 如果函数 $f: I \to \mathbb{R}$ 满足：

1. f 在区间 I 上只有有限多个瑕点；
2. 把区间 I 分解成有限多个不相交的区间 I_k 的并，每个 I_k 为有界区间或单侧有界区间，f 在每个有界区间 I_k 中至多只有一个瑕点，且该瑕点为区间 I_k 的端点
3. 每个广义积分 $\int_{I_k} f(x)\mathrm{d}x$ 都收敛。

则记

$$\int_I f(x)\mathrm{d}x = \sum_k \int_{I_k} f(x)\mathrm{d}x,$$

称它为 f 在区间 I 上的**广义积分**，称 f 在区间 I 上**广义可积**或**广义积分** $\int_I f(x)\mathrm{d}x$ **收敛**。否则称**广义积分** $\int_I f(x)\mathrm{d}x$ **发散**。

例 7.6.6. 讨论广义积分 $\int_0^{+\infty} \mathrm{e}^{-\lambda x}\mathrm{d}x$ 的收敛性。
解：

$$F(A) = \int_0^A \mathrm{e}^{-\lambda x}\mathrm{d}x = \begin{cases} \frac{1-\mathrm{e}^{-\lambda A}}{\lambda}, & \lambda \neq 0; \\ A, & \lambda = 0. \end{cases}$$

所以，当 $\lambda \leq 0$ 时，$F(A) \to +\infty(A \to +\infty)$，广义积分 $\int_0^{+\infty} \mathrm{e}^{-\lambda x}\mathrm{d}x$ 发散。

当且仅当 $\lambda > 0$ 时，广义积分 $\int_0^{+\infty} \mathrm{e}^{-\lambda x}\mathrm{d}x$ 收敛，且

$$\int_0^{+\infty} \mathrm{e}^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda}.$$

□

例 7.6.7. 讨论 $\int_1^{+\infty} \frac{1}{x^p}\mathrm{d}x$，$\int_2^{+\infty} \frac{1}{x(\ln x)^p}\mathrm{d}x$，$\int_3^{+\infty} \frac{1}{x\ln x(\ln\ln x)^p}\mathrm{d}x$ 的收敛性。
解： 因为

$$\int_1^A \frac{1}{x^p}\mathrm{d}x = \int_0^{\ln A} \frac{1}{\mathrm{e}^{pt}}\mathrm{d}\mathrm{e}^t = \int_0^{\ln A} \mathrm{e}^{-(p-1)t}\mathrm{d}t, \quad x = \mathrm{e}^t,$$

所以，当且仅当 $p > 1$ 时，广义积分 $\int_1^{+\infty} \frac{1}{x^p}\mathrm{d}x$ 收敛，且

$$\int_1^{+\infty} \frac{1}{x^p}\mathrm{d}x = \frac{1}{p-1}.$$

因为

$$\int_2^A \frac{1}{x(\ln x)^p}\mathrm{d}x = \int_{\ln 2}^{\ln A} \frac{1}{t^p}\mathrm{d}t, \quad t = \ln x,$$

所以，当且仅当 $p > 1$ 时，广义积分 $\int_2^{+\infty} \frac{1}{x(\ln x)^p}\mathrm{d}x$ 收敛。同理，当且仅当 $p > 1$ 时，广义积分 $\int_3^{+\infty} \frac{1}{x\ln x(\ln\ln x)^p}\mathrm{d}x$ 收敛。 □

例 7.6.8. 讨论 $\int_0^1 \frac{1}{x^p}\mathrm{d}x$ 的收敛性。

解： 当 $p \le 0$ 时，这是 Riemann 积分。当 $p > 0$ 时，$\frac{1}{x^p}$ 在区间 $(0,1)$ 上无界，但在任何区间 $[\varepsilon, 1]$ 上连续，所以 $x = 0$ 是 $\frac{1}{x^p}$ 的唯一瑕点。

$$\int_\varepsilon^1 \frac{1}{x^p}\mathrm{d}x = \int_0^{-\ln\varepsilon} \mathrm{e}^{pt}\mathrm{d}\mathrm{e}^{-t} = \int_0^{\ln A} \mathrm{e}^{-(1-p)t}\mathrm{d}t, \quad x = \mathrm{e}^{-t},$$

所以，当且仅当 $0 < p < 1$ 时，瑕积分 $\int_0^1 \frac{1}{x^p}\mathrm{d}x$ 收敛，且

$$\int_0^1 \frac{1}{x^p}\mathrm{d}x = \frac{1}{1-p}.$$

□

利用定义可以证明广义积分的线性、 Newton-Leibniz 公式、换元公式、分部积分等。

对瑕积分

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

其中 a 为唯一瑕点，考虑 $x = a + \mathrm{e}^{-t}$，则

$$\int_{a+\varepsilon}^b f(x)\mathrm{d}x = \int_{-\ln\varepsilon}^{-\ln(b-a)} f(a+\mathrm{e}^{-t})\mathrm{d}\mathrm{e}^{-t} = \int_{-\ln(b-a)}^{-\ln\varepsilon} \mathrm{e}^{-t} f(a+\mathrm{e}^{-t})\mathrm{d}t.$$

所以瑕积分 $\int_a^b f(x)\mathrm{d}x$ 收敛当且仅当广义积分 $\int_{-\ln(b-a)}^{+\infty} \mathrm{e}^{-t} f(a+\mathrm{e}^{-t})\mathrm{d}t$ 收敛，且

$$\int_a^b f(x)\mathrm{d}x = \int_{-\ln(b-a)}^{+\infty} \mathrm{e}^{-t} f(a+\mathrm{e}^{-t})\mathrm{d}t.$$

习题7.6

1. 平面直角坐标系中的曲线 $x^3 + y^3 = 3xy$ 在第一象限围成了一个有界区域 D，求 D 的面积。

2. 求平面直角坐标系中的曲线 $x^3 + y^3 = 3xy$ 和它的渐近线之间的无界区域的面积。

7.7 广义积分的收敛性

在实际应用中经常会遇到广义积分，判断积分的收敛性是首要的任务，但其中大部分积分因为被积函数没有初等原函数，所以无法通过广义积分定义中那样的方式进行检查。

Cauchy 准则

定理 7.7.1. 设 $f:[a,+\infty)\to\mathbb{R}$ 在任何有界区间 $I\subset[a,+\infty)$ 上 *Riemann* 可积。则广义积分 $\int_a^{+\infty}f(x)\mathrm{d}x$ 收敛当且仅当对任意 $\varepsilon>0$，存在 $N(\varepsilon)>a$ 使得对任意 $A_2>A_1\geq N(\varepsilon)$，

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

证明. 令

$$F(A)=\int_a^A f(x)\mathrm{d}x.$$

则要证明的结论就是函数极限 $\lim\limits_{A\to+\infty}F(A)$ 收敛的 Cauchy 准则。 □

虽然上述定理就是函数 $F(A)=\int_a^A f(x)\mathrm{d}x$ 在 $A\to+\infty$ 时的存在极限的 Cauchy 判别法， Cauchy 收敛准则使我们在不知道极限值的情况下也可以判断是否存在极限，但对广义积分而言，我们甚至无法求得函数 $F(A)$ 的值。所以我们通常需要对积分 $\int_{A_1}^{A_2}f(x)\mathrm{d}x$ 进行适当的不等式放缩。

定理 7.7.2. $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x$ 收敛当且仅当存在 $I\in\mathbb{R}$ 使得对任意 $\varepsilon>0$，存在 $N(\varepsilon)>0$ 使得对任意 $A<-N(\varepsilon)$ 以及任意 $B>N(\varepsilon)$，

$$\left|\int_A^B f(x)\mathrm{d}x-I\right|<\varepsilon.$$

证明. (必要性)设 $\int_{-\infty}^{+\infty}f(x)\mathrm{d}x$ 收敛。则 $I_1=\int_0^{+\infty}f(x)\mathrm{d}x$ 和 $I_2\int_{-\infty}^0 f(x)\mathrm{d}x$ 都收敛。于是对任意 $\varepsilon>0$，存在 $N(\varepsilon)>0$ 使得对任意 $A<-N(\varepsilon)$ 以及任意 $B>N(\varepsilon)$，

$$\left|\int_0^B f(x)\mathrm{d}x-I_1\right|<\frac{\varepsilon}{2},\quad \left|\int_A^0 f(x)\mathrm{d}x-I_2\right|<\frac{\varepsilon}{2}.$$

从而

$$\left|\int_A^B f(x)\mathrm{d}x-(I_1+I_2)\right|\leq\left|\int_0^B f(x)\mathrm{d}x-I_1\right|+\left|\int_A^0 f(x)\mathrm{d}x-I_2\right|<\varepsilon.$$

(充分性)设存在 $I\in\mathbb{R}$ 使得对任意 $\varepsilon>0$，存在 $N(\varepsilon)>0$ 使得对任意 $A<-N(\varepsilon)$ 以及任意 $B>N(\varepsilon)$，

$$\left|\int_A^B f(x)\mathrm{d}x-I\right|<\varepsilon.$$

则对任意 $B_2>B_1>N(\varepsilon)$，

$$\begin{aligned}\left|\int_{B_1}^{B_2}f(x)\mathrm{d}x\right|&=\left|\int_{-2N(\varepsilon)}^{B_2}f(x)\mathrm{d}x-\int_{-2N(\varepsilon)}^{B_1}f(x)\mathrm{d}x\right|\\&\leq\left|\int_{-2N(\varepsilon)}^{B_2}f(x)\mathrm{d}x-I\right|+\left|\int_{-2N(\varepsilon)}^{B_1}f(x)\mathrm{d}x-I\right|\\&<2\varepsilon.\end{aligned}$$

因此由 Cauchy 准则，$\int_0^{+\infty} f(x)\mathrm{d}x$ 收敛。同理可证 $\int_{-\infty}^{0} f(x)\mathrm{d}x$ 收敛。因此 $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x$ 收敛。 □

例 7.7.3. $\frac{x}{1+x^2}$ 是连续的奇函数，从而对任意 $A>0$，

$$\int_{-A}^{A} \frac{x\mathrm{d}x}{1+x^2}=0.$$

但是对任意正整数 N，

$$\int_{N}^{4N} \frac{x\mathrm{d}x}{1+x^2}=\frac{1}{2}\ln\left(\frac{1+16N^2}{1+N^2}\right)>\frac{\ln 9}{2}>1,$$

所以广义积分 $\int_0^{+\infty} \frac{x\mathrm{d}x}{1+x^2}$ 不收敛，从而广义积分 $\int_{-\infty}^{+\infty} \frac{x\mathrm{d}x}{1+x^2}$ 不收敛。 □

绝对收敛与比较法

定义 7.7.4. 称广义积分 $\int_a^{+\infty} f(x)\mathrm{d}x$ **绝对收敛**或f 在区间 $[a,+\infty)$ 上**绝对可积**，如果广义积分 $\int_a^{+\infty} |f(x)|\mathrm{d}x$ 收敛。

定理 7.7.5. *若广义积分 $\int_a^{+\infty} f(x)\mathrm{d}x$ 绝对收敛，则 $\int_a^{+\infty} f(x)\mathrm{d}x$ 收敛，并且*

$$\left|\int_a^{+\infty} f(x)\mathrm{d}x\right|\leq\int_a^{+\infty} |f(x)|\mathrm{d}x.$$

证明. 由 Cauchy 收敛准则和以下不等式得到。

$$\left|\int_{A_1}^{A_2} f(x)\mathrm{d}x\right|\leq\int_{A_1}^{A_2} |f(x)|\mathrm{d}x.$$

□

定理 7.7.6 (比较法). *设对任意 $N>a$，f,g 在有界区间$[a,N]$上Riemann可积，当 $x\to+\infty$ 时 $f(x)=O(g(x))$，即存在 $N_0>0$ 和 $M>0$ 使得对任意 $x>N_0$，$|f(x)|\leq M|g(x)|$。则*

1. *若广义积分 $\int_a^{+\infty} g(x)\mathrm{d}x$ 绝对收敛，则 $\int_a^{+\infty} f(x)\mathrm{d}x$ 绝对收敛。*
2. *若 $\int_a^{+\infty} f(x)\mathrm{d}x$ 不是绝对收敛的，则 $\int_a^{+\infty} g(x)\mathrm{d}x$ 不是绝对收敛的。*

证明. 若广义积分 $\int_a^{+\infty} g(x)\mathrm{d}x$ 绝对收敛，则对于任意 $\varepsilon>0$，存在 $N_\varepsilon>N_0$ 使得对任意 $A_2>A_1>N$，$\int_{A_1}^{A_2}|g(x)|\mathrm{d}x<M^{-1}\varepsilon$。于是

$$\int_{A_1}^{A_2}|f(x)|\mathrm{d}x\leq M\int_{A_1}^{A_2}|g(x)|\mathrm{d}x<\varepsilon.$$

因此 $\int_a^{+\infty} f(x)\mathrm{d}x$ 绝对收敛。

第二个结论是第一个结论的逆否命题。 □

推论 7.7.7. 设

$$\lim_{x\to+\infty}\left|\frac{f(x)}{g(x)}\right|<+\infty.$$

则

1. 若广义积分 $\int_a^{+\infty}g(x)\mathrm{d}x$ 绝对收敛，则 $\int_a^{+\infty}f(x)\mathrm{d}x$ 绝对收敛。
2. 若 $\int_a^{+\infty}f(x)\mathrm{d}x$ 不是绝对收敛的，则 $\int_a^{+\infty}g(x)\mathrm{d}x$ 不是绝对收敛的。

□

例 7.7.8 (Gamma函数). 记

$$\Gamma(\alpha)=\int_0^{+\infty}x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x.$$

证明：

1. 上述广义积分收敛当且仅当 $\alpha>0$；
2. 对任意 $\alpha>0$，$\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$；
3. 对任意非负整数 n，$\Gamma(n+1)=n!$。

证明. (1) 因为对任意实数 α，

$$x^{\alpha-1}\mathrm{e}^{-x}=x^{\alpha-1}\mathrm{e}^{-x/2}\mathrm{e}^{-x/2}=O(\mathrm{e}^{-x/2}),\quad x\to+\infty,$$

$\int_1^{+\infty}\mathrm{e}^{-\frac{x}{2}}\mathrm{d}x$ 收敛，所以 $\int_1^{+\infty}x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x$ 收敛。

因为 当$x\to0^+$时 $x^{\alpha-1}\mathrm{e}^{-x}$ 与 $x^{\alpha-1}$ 同阶，$\int_0^1x^{\alpha-1}\mathrm{d}x$ 收敛当且仅当 $1-\alpha<1$ 即 $\alpha>0$，所以 $\int_0^1x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x$ 收敛当且仅当 $\alpha>0$。

因此 $\int_0^{+\infty}x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x$ 收敛当且仅当 $\alpha>0$。

(2)

$$\begin{aligned}\Gamma(\alpha+1)&=\int_0^{+\infty}x^{\alpha}\mathrm{e}^{-x}\mathrm{d}x=-\int_0^{+\infty}x^{\alpha}\mathrm{d}\mathrm{e}^{-x}\\&=-\,x^{\alpha}\mathrm{e}^{-x}\Big|_0^{+\infty}-\alpha\int_0^{+\infty}x^{\alpha-1}\mathrm{e}^{-x}\mathrm{d}x=\alpha\Gamma(\alpha).\end{aligned}$$

(3)

$$\Gamma(1)=\int_0^{+\infty}\mathrm{e}^{-x}\mathrm{d}x=1.$$

利用 $\Gamma(n+1)=n\Gamma(n)$ 和数学归纳法可以证明 $\Gamma(n+1)=n!$。 □

例 7.7.9 (Beta函数).

$$B(\alpha,\beta)=\int_0^1x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x.$$

证明：
(1) 上述积分收敛当且仅当 α, β 都是正数；
(2) 对任意$\alpha > 0$，$B(\alpha, 1) = \frac{1}{\alpha}$；
(3) 对于任意正整数 m, n，$B(m, n) = \frac{(n-1)!(m-1)!}{(m+n-1)!} = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$。

证明. (1) 当 $x \to 0^+$ 时，

$$x^{\alpha-1}(1-x)^{\beta-1} = x^{\alpha-1}(1+o(1)),$$

所以 $\int_0^{1/2} x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x$ 收敛当且仅当 $\int_0^{1/2} x^{\alpha-1}\mathrm{d}x$ 收敛，后者当且仅当 $\alpha > 0$。

当 $x \to 1^-$ 时，

$$x^{\alpha-1}(1-x)^{\beta-1} = (1-x)^{\beta-1}(1+o(1)),$$

所以 $\int_0^{1/2} x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x$ 收敛当且仅当 $\int_{1/2}^{1}(1-x)^{\beta-1}\mathrm{d}x$ 收敛，后者当且仅当 $\beta > 0$。

所以 $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x$ 收敛当且仅当 α, β 都是正数。
(2) 直接计算即得。
(3) 对 $\alpha > 0$ 以及 $\beta > 1$，

$$\begin{aligned}
B(\alpha, \beta) &= \frac{1}{\alpha}\int_0^1 (1-x)^{\beta-1}\mathrm{d}x^{\alpha} \\
&= \frac{1}{\alpha}(1-x)^{\beta-1}x^{\alpha}\Big|_0^1 + \frac{\beta-1}{\alpha}\int_0^1 x^{\alpha}(1-x)^{\beta-2}\mathrm{d}x \\
&= \frac{\beta-1}{\alpha}B(\alpha+1, \beta-1).
\end{aligned}$$

因此对于任意正整数m, n，

$$\begin{aligned}
B(m,n) &= \frac{n-1}{m}B(m+1, n-1) = \frac{(n-1)(n-2)}{m(m+1)}B(m+2, n-2) = \cdots \\
&= \frac{(n-1)!}{m(m+1)(m+n-2)}B(m+n-1, 1) = \frac{(n-1)!(m-1)!}{(m+n-1)!} = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.
\end{aligned}$$

□

条件收敛与Dirichlet-Abel判别法

定义 7.7.10. 称广义积分 $\int_a^{+\infty} f(x)\mathrm{d}x$ **条件收敛**，如果它收敛但不是绝对收敛。

例 7.7.11. 讨论 $\int_0^{+\infty} \frac{\sin x}{x^{\alpha}}\mathrm{d}x$ 收敛性。
解： 分成瑕积分 $\int_0^1 \frac{\sin x}{x^{\alpha}}\mathrm{d}x$ 和无穷积分 $\int_0^{+\infty} \frac{\sin x}{x^{\alpha}}\mathrm{d}x$，分别讨论它们的收敛性。

当 $x \to 0^+$时，$\frac{\sin x}{x^\alpha} = \frac{1}{x^{\alpha-1}}(1+o(1))$，所以 $\int_0^1 \frac{\sin x}{x^\alpha}\mathrm{d}x$ 收敛当且仅当 $\int_0^1 \frac{1}{x^{\alpha-1}}\mathrm{d}x$ 收敛，后者当且仅当 $\alpha-1<1$ 即 $\alpha<2$。

所以，当 $\alpha \geq 2$ 时，$\int_0^1 \frac{\sin x}{x^\alpha}\mathrm{d}x$ 发散。

下面讨论 $\int_1^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 的收敛性。

由

$$\begin{aligned}\left|\int_A^B \frac{\sin x}{x^\alpha}\mathrm{d}x\right| &= \left|-\left.\frac{\cos x}{x^\alpha}\right|_A^B - \alpha\int_A^B \frac{\cos x}{x^{1+\alpha}}\mathrm{d}x\right| \\ &\leq \frac{1}{B^\alpha}+\frac{1}{A^\alpha}+|\alpha|\int_A^B \frac{1}{x^{1+\alpha}}\mathrm{d}x\end{aligned}$$

知：当 $\alpha>0$ 时，对任意 $\varepsilon>0$，存在 $N(\varepsilon)>1$ 使得对任意 $B>A>N(\varepsilon)$，

$$\frac{1}{B^\alpha}<\frac{\varepsilon}{3},\quad \frac{1}{A^\alpha}<\frac{\varepsilon}{3},\quad \alpha\int_A^B \frac{1}{x^{1+\alpha}}\mathrm{d}x<\frac{\varepsilon}{3}.$$

因此 $\int_1^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 收敛。

当 $\alpha>1$ 时，

$$\left|\frac{\sin x}{x^\alpha}\right| \leq \frac{1}{x^\alpha},$$

而 $\int_1^{+\infty} \frac{1}{x^\alpha}\mathrm{d}x$ 收敛，所以 $\int_1^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 收敛绝对收敛。

当 $0<\alpha\leq 1$ 时，

$$\begin{aligned}\int_{n\pi}^{2n\pi}\left|\frac{\sin x}{x^\alpha}\right|\mathrm{d}x &\geq \sum_{k=1}^{n}\int_{(n+k-1)\pi+\frac{\pi}{6}}^{(n+k)\pi-\frac{\pi}{6}}\frac{1}{2x}|\mathrm{d}x \\ &= \sum_{k=1}^{n}\ln\left(\frac{(n+k)\pi-\frac{\pi}{6}}{(n+k-1)\pi+\frac{\pi}{6}}\right) > \sum_{k=1}^{n}\ln\left(1+\frac{\frac{2}{3}}{n+k}\right) \\ &> \sum_{k=1}^{n}\frac{\frac{1}{3}}{n+k} > \sum_{k=1}^{n}\frac{\frac{1}{3}}{2n} = \frac{1}{6}\end{aligned}$$

所以 $\int_1^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 不是绝对收敛的。因此 $0<\alpha\leq 1$ 时，$\int_1^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 是条件收敛的。

当 $\alpha\leq 0$ 时，

$$\left|\int_{n\pi}^{(n+1)\pi}\frac{\sin x}{x^\alpha}\mathrm{d}x\right| = \int_{n\pi}^{(n+1)\pi}\left|\frac{\sin x}{x^\alpha}\right|\mathrm{d}x \geq \int_{n\pi}^{(n+1)\pi}|\sin x|\,\mathrm{d}x = 2.$$

所以当 $\alpha\leq 0$ 时，$\int_1^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 发散。

总结：

- 当 $\alpha\geq 2$ 或 $\alpha\geq 0$ 时，积分 $\int_0^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 发散；
- 当 $1<\alpha<2$ 时，积分 $\int_0^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 绝对收敛；
- 当 $\alpha<1$ 时，积分 $\int_0^{+\infty} \frac{\sin x}{x^\alpha}\mathrm{d}x$ 条件收敛。

□

定理 7.7.12. 设 g 单调，记 $F(x)=\int_a^x f(t)\mathrm{d}t$。则当下述两组条件之一成立时，$\int_a^{+\infty} f(x)g(x)\mathrm{d}x$ 收敛。

1. (**Dirichlet**) $F(x)$ 有界，$\lim\limits_{x\to+\infty} g(x)=0$；
2. (**Abel**) $\lim\limits_{x\to+\infty} F(x)$ 收敛，g 有界。

证明. (1) 设 f 连续，g 连续可微。则对任意 $B>A>0$，

$$\begin{aligned}\left|\int_A^B f(x)g(x)\mathrm{d}x\right| &= \left|\int_A^B g(x)\mathrm{d}F(x)\right| \\ &= \left|F(B)g(B)-F(A)g(A)-\int_A^B F(x)g'(x)\mathrm{d}x\right| \\ &\le M|g(B)|+M|g(A)|+M\int_A^B |g'(x)|\,\mathrm{d}x \\ &= M|g(B)|+M|g(A)|+M\left|\int_A^B g'(x)\mathrm{d}x\right| \quad (g\text{ 单调}) \\ &\le 2M\left(|g(B)|+|g(A)|\right)\le 4M|g(A)| \quad (g\text{ 单调}) \\ &\to 0, \quad (\text{当 } A\to+\infty).\end{aligned}$$

所有由Cauchy准则，$\int_a^{+\infty} f(x)g(x)\mathrm{d}x$ 收敛。

(2) 令 $\tilde{g}(x)=g(x)-\beta$，其中 $\beta=\lim\limits_{x\to+\infty} g(x)$。则由 Dirichlet，广义积分 $\int_a^{+\infty} f(x)\tilde{g}(x)\mathrm{d}x$ 收敛。所以

$$\int_a^{+\infty} f(x)g(x)\mathrm{d}x=\int_a^{+\infty} f(x)\tilde{g}(x)\mathrm{d}x+\beta\int_a^{+\infty} f(x)\mathrm{d}x$$

收敛。 □

习题7.7

1. 讨论积分 $\int_0^{\pi/2}(\sin x)^{\alpha}(\cos x)^{\beta}\mathrm{d}x$ 的收敛性，并用 Beta 函数表示它的值。
2. 讨论积分 $\int_0^{+\infty}\frac{x^{\alpha}}{1+x^{\beta}|\sin x|}\mathrm{d}x$ 的收敛性。

第 8 章　常微分方程

8.1　基本概念

微分方程就是一个或几个未知函数以及它们的导数满足的等式，微分方程的**阶**就是方程中出现的未知函数的最高阶导数的阶。如果一个微分方程中的未知函数都是关于同一个变量的一元函数，那么这个方程称为一个**常微分方程**。

$$F(x, y, y', \ldots, y^{(n)}) = 0$$

是关于未知函数 y 的 n 阶常微分方程，它通常被称为**隐式**常微分方程。在**显式**常微分方程中，未知函数 y 的最高阶导数是 y 和它的低阶导数的函数，即

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}).$$

微分方程组由关于一组未知函数的多个微分方程所组成。一阶（显式）常微分方程组的一般形式为

$$\mathbf{y}' = \mathbf{F}(x, \mathbf{y}), \qquad \begin{cases} y_1' = F_1(x, y_1, \ldots, y_n), \\ y_2' = F_2(x, y_1, \ldots, y_n), \\ \vdots \\ y_n' = F_n(x, y_1, \ldots, y_n). \end{cases}$$

高阶微分方程可以改写为一阶微分方程组。令

$$\mathbf{y} = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix},$$

则

$$\mathbf{y}' = \begin{pmatrix} y' \\ y'' \\ \vdots \\ y^{(n)} \end{pmatrix} = \begin{pmatrix} y' \\ y'' \\ \vdots \\ f(x, y, y', \cdots, y^{(n-1)}) \end{pmatrix} = \mathbf{F}(x, \mathbf{y}).$$

因此，高阶方程的性质可以通过研究一阶微分方程组而得到。

线性微分方程组形如

$$\mathbf{y}' = \mathbf{A}(x)\mathbf{y} + \mathbf{f}(x),$$

其中 $\mathbf{A}(x)$ 是 n 行 n 列的矩阵，$\mathbf{f}(x)$ 是 n 维列向量。$\mathbf{f}(x)$ 是这个线性方程组的**非齐次项**。若 $\mathbf{f}(x) = 0$，则该方程组称为**线性齐次微分方程组**；否则称为**线性非齐次微分方程组**。若矩阵 $\mathbf{A}$ 是常值矩阵，则该方程组称为**常系数线性微分方程组**；否则称为**变系数线性方程组**。

对 n 阶线性微分方程

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

可以类似定义**齐次**或**非齐次**、**常系数**或**变系数**线性微分方程，f 为**非齐次项**。

对一阶常微分方程组，

$$\begin{cases} \mathbf{y}' = \mathbf{F}(x, \mathbf{y}), \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

称为**初值问题**，其中 $\mathbf{y}(x_0) = \mathbf{y}_0$ 称为**初始条件**。相应地，对高阶微分方程，

$$\begin{cases} y^{(n)} = f(x, y, y', \dots, y^{(n-1)}), \\ y(x_0) = y_0, \\ y'(x_0) = y_1, \\ \vdots \\ y^{(n-1)}(x_0) = y_{n-1} \end{cases}$$

称为**初值问题**，$\begin{cases} y(x_0) = y_0, \\ y'(x_0) = y_1, \\ \vdots \\ y^{(n-1)}(x_0) = y_{n-1} \end{cases}$ 称为**初始条件**。一般而言，一个微分方程的解由初始条件确定，所以通常一个 n 阶微分方程的解会含有 n 个任意常数，这样的解称为微分方程的**通解**(general solution)。不含任何任意常数的解称为微分方程的**特解**。需要指出的是，对一些微分方程，并非每个特解都包含在通解之中，这时通解不是微分方程的所有的解。

8.2 降阶与降维

降低微分方程的阶数（降阶）和减少方程组中未知函数的个数（降维）是求解微分方程的常用手段。

降阶

(I) **不显含未知函数 y 的微分方程**，即形如

$$F(x, y^{(k)}, y^{(k+1)}, \dots, y^{(n)}) = 0$$

的微分方程，它等价于方程组

$$\begin{cases} F(x,, z, z', \dots, z^{(n-k)}) = 0, \\ y^{(k)} = z. \end{cases}$$

由这个方程组中的第一个方程解得 z，再利用第二个方程通过 k 次积分得到原方程的解 y。而第一个方程为 $n-k$ 阶微分方程。

例 8.2.1 (**线性方程利用常数变易法降阶**). 如果 $y_0(x)$ 是线性齐次方程

$$Ly = y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = 0$$

的一个恒不为零的解，则 $y(x) = C(x)y_0(x)$ 是 $Ly = f$ 的解当且仅当 $C(x)$ 是一个线性 n 阶方程 $L(Cy_0) = f$ 的解，这个方程是一个可降阶方程。

例如，e^x 是

$$xy'' - (x+1)y' + y = 0$$

的解。于是，$\mathrm{e}^x C(x)$ 是

$$xy'' - (x+1)y' + y = f(x)$$

的解当且仅当

$$x\mathrm{e}^x \left[C'' + 2C' + C\right] - (x+1)\mathrm{e}^x \left[C' + +C\right] + \mathrm{e}^x C = f(x),$$

即

$$xC'' + (x-1)C' = \mathrm{e}^{-x} f(x),$$

这是关于 C 的一个可降阶的二阶线性方程。 □

(II) **不显含自变量 x 的微分方程**，即形如

$$F(y, y', y'', \dots, y^{(n)}) = 0$$

的微分方程。令 $z = y'$ ，则

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x} = z\frac{\mathrm{d}u}{\mathrm{d}y},$$

于是

$$\frac{\mathrm{d}^k y}{\mathrm{d}x^k} = \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k y = \left(z\frac{\mathrm{d}}{\mathrm{d}y}\right)^k y = \left(z\frac{\mathrm{d}}{\mathrm{d}y}\right)^{k-1} \left(z\frac{\mathrm{d}y}{\mathrm{d}y}\right) = \left(z\frac{\mathrm{d}}{\mathrm{d}y}\right)^{k-1} z,$$

因此原方程等价于以下方程组

$$\begin{cases} F\left(y, z, \left(z\frac{\mathrm{d}}{\mathrm{d}y}\right)z, \ldots, \left(z\frac{\mathrm{d}}{\mathrm{d}y}\right)^{n-1} z\right) = 0, \\ \frac{\mathrm{d}y}{\mathrm{d}x} = z. \end{cases}$$

其中第一个方程是 $n-1$ 阶微分方程，由它解得 $z = z(y)$，再由第二个方程 $\frac{\mathrm{d}y}{\mathrm{d}x} = z(y)$ 解得 $y = y(x)$。

(III) **利用算子因子分解降阶**。

例 8.2.2. 考虑微分方程

$$Ly = xy'' - (x+1)y' + y = f(x).$$

其中 L 是微分算子

$$x\frac{\mathrm{d}^2}{\mathrm{d}x^2} - (x+1)\frac{\mathrm{d}}{\mathrm{d}x} + 1 = x\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^2 - (x+1)\frac{\mathrm{d}}{\mathrm{d}x} + 1 = \left(x\frac{\mathrm{d}}{\mathrm{d}x} - 1\right)\left(\frac{\mathrm{d}}{\mathrm{d}x} - 1\right).$$

所以原方程等价于以下一阶微分方程组

$$\begin{cases} \left(x\frac{\mathrm{d}}{\mathrm{d}x} - 1\right) z = f(x), \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - 1\right) y = z. \end{cases}$$

可以从第一个方程解出 $z(x)$，再把它代入第二个方程解得 $y(x)$。 □

例 8.2.3. 对常系数线性微分方程

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

微分算子 L 总是可以因式分解成如下形式的

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right)\cdots\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_n\right),$$

其中 $\lambda_1, \ldots, \lambda_n$ 是特征多项式

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

的所有根。于是原方程等价于以下一阶微分方程组

$$\begin{cases} \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right) z_1 = f(x), \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_2\right) z_2 = z_1, \\ \vdots \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_n\right) y = z_{n-1}. \end{cases}$$

从这些方程依次解出 $z_1, z_2, \ldots, z_{n-1}$ 和 y。 □

例 8.2.4. Euler方程是一种特殊类型的变系数线性微分方程，具有如下形式

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

它可以因式分解成如下形式

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right)\cdots\left(x\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_n\right) y = f.$$

于是原方程等价于以下一阶微分方程组

$$\begin{cases} \left(x\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right) z_1 = f(x), \\ \left(x\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_2\right) z_2 = z_1, \\ \vdots \\ \left(x\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_n\right) y = z_{n-1}. \end{cases}$$

从这些方程依次解出 $z_1, z_2, \ldots, z_{n-1}$ 和 y。 □

降维

减少微分方程组中未知函数之间的耦合，可以降低方程的复杂度，从而使方程更易于求解。如果一阶微分方程组具有如下形式

$$\begin{cases} y_1' = f_1(x, y_1), \\ y_2' = f_2(x, y_1, y_2), \\ \vdots \\ y_n' = f_n(x, y_1, y_2, \ldots, y_n), \end{cases}$$

那么我们可以从第一个方程解出 $y_1(x)$，将其代入第二个方程，得到

$$y_2' = f(x, y_1(x), y_2),$$

进而解出 $y_2(x)$。如此可依次解出 $y_1, y_2, \ldots, y_n$。本节中提到的那些例子，最终都转化为了这种形式的方程组。

例 8.2.5. 对矩阵

$$A = \begin{pmatrix} \lambda & 1 & \cdots & 0 & 0 \\ 0 & \lambda & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

常微分方程组

$$\mathbf{y}' = A\mathbf{y} + \mathbf{f}(x)$$

即

$$\begin{cases} y_1' - \lambda y_1 = y_2 + f_1(x), \\ y_2' - \lambda y_2 = y_3 + f_2(x), \\ \vdots \\ y_n' - \lambda y_n = f_n(x). \end{cases}$$

于是可以借由这些一维的一阶微分方程依次解出 $y_n, y_{n-1}, \ldots, y_2, y_1$。 □

例 8.2.6. 根据矩阵的特征值、特征向量以及Jordan标准型的相关结论，对任何复数常数矩阵 A，都存在可逆矩阵 P 使得 $P^{-1}AP = \Lambda$ 是对角块矩阵

$$\Lambda = \begin{pmatrix} \Lambda_1 & & & \\ & \Lambda_2 & & \\ & & \ddots & \\ & & & \Lambda_k \end{pmatrix}$$

于是经坐标变换

$$\begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_k \end{pmatrix} = P^{-1} \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_k \end{pmatrix}$$

常微分方程组

$$\mathbf{y}' = A\mathbf{y} + \mathbf{f}(x)$$

变为

$$\mathbf{z}' = P^{-1}\mathbf{y}' = P^{-1}\left(A\mathbf{y} + \mathbf{f}(x)\right) = \Lambda\mathbf{z} + P^{-1}\mathbf{f}(x),$$

后者由一些独立的低维微分方程组组成。也就是利用系数矩阵的特征值所对应的不变子空间分解，实现微分方程的解耦。此时，要么 Λ_j 是一维的，要么 Λ_j 是形如上例中的Jordan块矩阵。 □

8.3 一阶微分方程

一阶微分方程及其背后的几何

显示的一阶微分方程形如

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y).$$

此时二元函数 f 在 xy 坐标平面中的每个点 (x, y) 处指定了一个斜率值 $f(x, y)$，此时函数 $y = y(x)$ 是上述微分方程的解当且仅当该函数的图像，即曲线 $(x, y(x))$，在其经过的每个点处的切线斜率恰好为 $f(x, y(x))$。所以，我们把 f

视为 xy 坐标平面上的一个**斜率场**，称曲线 $(x, y(x))$ 为这个斜率场的一条**积分曲线**。

上述微分方程也可以改写为

$$f(x,y)\mathrm{d}x - \mathrm{d}y = 0.$$

从而可以考虑如下形式的微分方程

$$P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y = 0.$$

此时，称曲线 $\gamma : (x(t), y(t))$ 是这个微分方程的一条**积分曲线**，如果

$$P(x(t),y(t))x'(t) + Q(x(t),y(t))y'(t) = 0,$$

即凡在曲线 γ 所经之处，曲线 γ 的切线总是与向量 $(P(x,y), Q(x,y))^T$ 正交。因此在 xy 坐标平面上，向量场 $(P(x,y), Q(x,y))^T$ 的正交方向给出一个**方向场**，上述微分方程的积分曲线也就是这个方向场的积分曲线。而之前所说的斜率场其实也刚好确定了一个方向场。

如果 $\gamma : (x(t), y(t))$ 是

$$P(x,y)\mathrm{d}x + Q(x,y)\mathrm{d}y = 0$$

的一条积分曲线，且 $x'(t) \neq 0$，则通过反函数 $t = t(x)$ 得到曲线 γ 的另一个表达 $(x, \tilde{y}(x)) = (x, y(t(x)))$。

最简单的一阶微分方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x),$$

它的解就是 f 的所有原函数。如果 $y = F(x)$ 是该微分方程(即斜率场 f)的一条积分曲线，则 $y = F(x) + C$ 也是。从几何上看，这个方程的斜率场或方向场只依赖于 x 坐标，因而沿 y 轴平移保持不变，所以它的任何一条积分曲线沿 y 轴平移后仍是积分曲线。

分离变量的微分方程

所谓分离变量的微分方程方程就是形如

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{f(x)}{g(y)}$$

的微分方程，它等价于

$$g(y)\mathrm{d}y = f(x)\mathrm{d}x.$$

上式两边积分得到

$$\int g(y)\mathrm{d}y = \int f(x)\mathrm{d}x.$$

所以

$$G(y) = F(x) + C,$$

其中 F, G 分别是 f, g 的原函数，C 是任意常数。这给出了平面上的一条曲线。当 $g(y) \neq 0$ 时，G 有反函数，并解得

$$y = G^{-1}(F(x) + C).$$

几何上看，如果考虑坐标变换

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

则

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}x} = g(y) \cdot \frac{f(x)}{g(y)} = f(x),$$

即在xz坐标系中，原微分方程变为最简单的一阶微分方程。

例 8.3.1 (最速降线). 1696年 Johann Bernoulli 向全世界提出了如下这个挑战问题：给定竖直平面中不在同一竖线上的两点 A 和 B，A 高于 B，求连接 A 和 B 的一条曲线 γ，使得一个初始速度为零的珠子在重力作用下沿 γ 从 A 到 B 所用时间最短。这就是著名的“最速降线问题”。

解： 以 A 为原点建立坐标系，x 轴为水平方向，y 轴正向竖直向下，B点位于第一象限。

根据机械能守恒

$$\frac{1}{2}mv^2 = mgy.$$

所以

$$v = \sqrt{2gy}.$$

因此珠子的运动是变速运动，在同一高度速度相同。

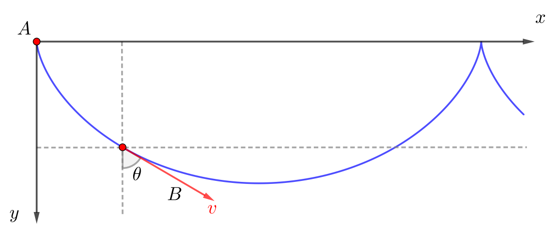


图 8.1:

Johann Bernoulli 根据 Fermat 原理和折射定理指出

$$\frac{v}{\sin\theta} = 常数C.$$

其中 θ 是曲线与 y 轴正向的夹角。

于是

$$\cot\theta=\tan\left(\frac{\pi}{2}-\theta\right)=\frac{\mathrm{d}y}{\mathrm{d}x}.$$

因此

$$C^2=\frac{v^2}{\sin^2\theta}=v^2\left(1+\cot^2\theta\right)=2gy\left[1+(y')^2\right],$$

即

$$y\left[1+(y')^2\right]=2C_1.$$

于是

$$y'=\sqrt{\frac{2C_1}{y}-1}.$$

这是一个分离变量方程，

$$\frac{\mathrm{d}y}{\sqrt{\frac{2C_1}{y}-1}}=\mathrm{d}x,$$

令 $y=2C_1\sin^2\psi$ ($0<\psi<\pi$)，得到

$$4C_1\sin^2\psi\mathrm{d}\psi=\mathrm{d}x,$$

解得

$$\begin{cases}x=2C_1\psi-C_1\sin2\psi+C_2,\\ y=C_1(1-\cos2\psi).\end{cases}$$

这是摆线的参数方程。

在 A 点，$x=y=0$，所以 $\psi=0$，$C_2=0$，于是

$$\begin{cases}x=C_1(2\psi-\sin2\psi),\\ y=C_1(1-\cos2\psi).\end{cases}$$

在 B点，$x=x_b$，$y=y_b$，

$$\frac{y_b}{x_b}=\frac{1-\cos2\psi}{2\psi-\sin2\psi}.$$

右端函数关于 ψ 在区间 $(0,\pi)$ 上是一个严格减的连续函数，值域为 $(0,+\infty)$，所以上述等式有唯一解 $\psi_0\in(0,\pi)$。于是 $C_1=\frac{y_b}{1-\cos2\psi_0}$。从而曲线 γ 为

$$\begin{cases}x=\dfrac{x_b(2\psi-\sin2\psi)}{2\psi_0-\sin2\psi_0},\\ y=\dfrac{y_b(1-\cos2\psi)}{1-\cos2\psi_0}.\end{cases}$$

□

一些可以转化为分离变量形式的微分方程

例 8.3.2.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(ax+by+c).$$

即 $p = ax+by+c$。则

$$1 = a\frac{\mathrm{d}x}{\mathrm{d}p} + b\frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}p} = (a+bf(p))\,\frac{\mathrm{d}x}{\mathrm{d}p},$$

即

$$\frac{\mathrm{d}x}{\mathrm{d}p} = \frac{1}{a+bf(p)},$$

这是个最简单的微分方程，由它解得

$$x = \int \frac{\mathrm{d}p}{a+bf(p)} = G(ax+by+c)+C,$$

其中 G 是 $\frac{1}{a+bf(p)}$ 的一个原函数，C 是任意常数。 □

例 8.3.3 (**齐次方程**)**.**

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f\left(\frac{y}{x}\right).$$

记 $t = \frac{y}{x}$，$u = \ln|x|$。则 $y = xt$，两边对 t 求导得到

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

因此

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{x}\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{x}\cdot\frac{x}{f(t)-t} = \frac{1}{f(t)-t},$$

这是最简单的微分方程，积分得到

$$u = \int \frac{\mathrm{d}t}{f(t)-t},$$

从而 $x = C\mathrm{e}^{G(y/x)}$，其中 G 是 $\frac{1}{f(p)-p}$ 的一个原函数，C 是任意常数。 □

例 8.3.4. xy 直角坐标平面中位于点 $(a,0)$ 处的一只狼看见一只兔子沿 y 轴正向以速度 v 匀速前进恰好经过原点。狼开始追逐兔子，狼始终朝向兔子，并保持以速度 1 匀速前进。问狼是否能追上兔子以及狼追多久能追到兔子？

解： 设时刻 t 时狼位于点 $(x(t),y(t))$。于是

$$\frac{vt-y(t)}{x(t)} = \tan(\pi-\theta(t)) = -\tan\theta(t) = -\frac{\mathrm{d}y}{\mathrm{d}x},$$

因此

$$vt-y(t) = -x(t)\frac{\mathrm{d}y}{\mathrm{d}x},$$

于是两边对 x 求导，得到

$$v\frac{\mathrm{d}t}{\mathrm{d}x} - \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\mathrm{d}y}{\mathrm{d}x} - x\frac{\mathrm{d}^2y}{\mathrm{d}x^2};$$

另一方面，

$$-\frac{\mathrm{d}x}{\mathrm{d}t}\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}=1,$$

所以

$$\boxed{v\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}=x\frac{\mathrm{d}^2y}{\mathrm{d}x^2}.}$$

这是可降阶方程，令 $z=\frac{\mathrm{d}y}{\mathrm{d}x}$，则

$$v\sqrt{1+z^2}=x\frac{\mathrm{d}z}{\mathrm{d}x}.$$

这是一个分离变量方程，解得

$$z=\frac{Ax^v}{2}-\frac{1}{2Ax^v}.$$

因此

$$y=\begin{cases}\dfrac{Ax^{v+1}}{2(1+v)}-\dfrac{x^{1-v}}{2A(1-v)}+B, & v\neq 1;\\ \dfrac{Ax^{v+1}}{2(1+v)}-\dfrac{\ln x}{2A}+B, & v=1.\end{cases}$$

若 $v\geq 1$，则 $\lim\limits_{x\to 0^+}y(x)=+\infty$，所以 $y=y(x)$ 与 y 轴无交点，于是狼追不上兔子。

若 $v<1$，则

$$y=\frac{Ax^{v+1}}{2(1+v)}-\frac{x^{1-v}}{2A(1-v)}+B.$$

又因为 $y(a)=0$，$y'(a)=0$，而

$$y'=\frac{Ax^v}{2}-\frac{x^{-v}}{2A},$$

所以 $A=a^{-v}$，所以

$$y=\frac{x^{v+1}}{2a^v(1+v)}-\frac{a^vx^{1-v}}{2(1-v)}+\frac{av}{1-v^2}.$$

于是

$$y(0)=\frac{av}{1-v^2}.$$

此时狼跑的路程

$$L=\int_0^a\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x=\int_0^a\frac{x}{v}\frac{\mathrm{d}^2y}{\mathrm{d}x^2}\mathrm{d}x=\frac{x}{v}\frac{\mathrm{d}y}{\mathrm{d}x}\bigg|_0^a-\int_0^a\frac{1}{v}\frac{\mathrm{d}y}{\mathrm{d}x}\mathrm{d}x=\frac{a}{1-v^2},$$

这也正是狼跑的时间，此时兔子恰好也位于 $\frac{av}{1-v^2}$ 处。所以狼追到兔子。 □

习题8.3

1. 设 $P(x,y), Q(x,y)$ 都是 k 次齐次函数，即对任意正数 λ 以及任意 $(x,y)\neq(0,0)$，都有

   $$P(\lambda x,\lambda y)=\lambda^k P(x,y),\quad Q(\lambda x,\lambda y)=\lambda^k Q(x,y).$$

   因此在xy坐标平面中沿从原点出发的每条射线平移，微分方程

   $$P(x,y)\mathrm{d}x+Q(x,y)\mathrm{d}y=0$$

   对应的方向场保持不变。证明：在平面极坐标系下，上述微分方程是分离变量的。

2. 证明形如

   $$\frac{\mathrm{d}y}{\mathrm{d}x}=f\left(\frac{ax+bx+c}{\alpha x+\beta y+\gamma}\right)$$

   的方程可以在适当的坐标系下改写为分离变量形式。

3. 解方程

   $$z'=-q-pz-z^2,$$

   其中 p,q 是实数。

8.4 线性微分方程

我们考察

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

其中系数函数 $a_0,a_1,\ldots,a_{n-1}$ 和非齐次项 f 都是区间 I 上的连续函数。

因为微分算子 L 是线性的，所以

定理 8.4.1 (线性叠加原理). 设 $Ly=f$，$Lz=g$。则对任意复数 λ,μ，

$$L(\lambda y+\mu z)=\lambda f+\mu g.$$

□

由此得到

定理 8.4.2 (齐次方程和非齐次方程的解空间). 齐次方程的所有的解组成的集合 $V_0=\{y|Ly=0\}$ 是一个线性空间；非齐次方程所有的解组成一个仿射空间 $V_f=\{y|Ly=f\}$，且对任意$y_1\in V_f$，$V_f=\{y_1+y|Ly=0\}$，即非齐次方程的通解是齐次方程的通解与非齐次方程的某个特解的和。 □

所以，求解线性微分方程就归结为如何求解齐次线性方程以及如何找到非齐次方程的一个特解。

一阶线性方程

对一阶线性微分方程

$$y' + a(x)y = f(x),$$

它对应的齐次方程

$$y' + a(x)y = 0$$

是分离变量方程，

$$\frac{\mathrm{d}y}{y} = -a(x)\mathrm{d}x.$$

积分得到

$$y(x) = C\exp\left(-\int_{x_0}^{x} a(t)\mathrm{d}t\right).$$

为了验证这是齐次方程的所有的解以及求解非齐次方程，我们可以使用**常数变易法**。

设 $Y_0(x)$ 是齐次方程 $y' + a(x)y = 0$ 的一个解，满足 $Y_0(x) \neq 0(\ \forall x\ )$。则对任意 $y(x)$，令 $C(x) = \frac{y(x)}{Y_0(x)}$，于是 $y(x) = Y_0(x)C(x)$。 $y(x) = Y_0(x)C(x)$ 是方程

$$y' + a(x)y = f(x)$$

的解，当且仅当

$$Y_0(x)C'(x) + Y_0'(x)C(x) + a(x)Y_0(x)C(x) = f(x)$$

即

$$Y_0(x)C'(x) = f(x),$$

也即

$$C'(x) = \frac{f(x)}{Y_0(x)}.$$

于是

$$C(x) = C_0 + \int_{x_0}^{x} \frac{f(t)}{Y_0(t)}\mathrm{d}t.$$

因此

$$y' + a(x)y = f(x)$$

的解为

$$y(x) = Y_0(x)C_0 + Y_0(x)\int_{x_0}^{x} \frac{f(t)}{Y_0(t)}\mathrm{d}t.$$

其中 $Y_0(x)C_0$ 是齐次方程的所有的解，$Y_0(x)\int_{x_0}^{x} \frac{f(t)}{Y_0(t)}\mathrm{d}t$ 是非齐次方程的一个特解。因此

定理 8.4.3. *1.* 齐次方程的通解为 $y(x)=Y_0(x)C$，齐次方程解空间是一个一维的线性空间，非齐次方程的解空间是一个一维的仿射空间；

2. 初值问题

$$y'+a(x)y=f(x),\quad y(x_0)=y_0$$

有唯一解

$$y(x)=Y_0(x)y_0+Y_0(x)\int_{x_0}^{x}\frac{f(t)}{Y_0(t)}\mathrm{d}t,$$

其中 $Y_0(x)=e^{-\int_{x_0}^{x}a(t)\mathrm{d}t}$。

从上述计算，我们发现利用常数变易法求非齐次线性方程特解时需要计算积分。在有些特殊情况下，可以利用一些特殊方法求得特解而无需积分。

例 8.4.4. 对微分方程

$$y'+ay=\mathrm{e}^{\mu x}P(x),\quad (\mu\in\mathbb{C})$$

设 $y=\mathrm{e}^{\mu x}z(x)$，则得到

$$z'+(a+\mu)\,z(x)=P(x).$$

这样就可以把非齐次项中的指数函数和三角函数(看成是复数指数的指数函数)因子剔除。

如果 $P(x)$ 是多项式，则上述方程有多项式解。

若 $a+\mu=0$，则

$$z(x)=\int P(x)\mathrm{d}x.$$

若 $a+\mu\neq 0$，则

$$z(x)=\frac{P(x)}{a+\mu}-\frac{P'(x)}{(a+\mu)^2}+\frac{P''(x)}{(a+\mu)^3}+\cdots+(-1)^k\frac{P^{(k)}(x)}{(a+\mu)^{k+1}},$$

其中 k 是多项式 $P(x)$ 的次数。 □

例 8.4.5. 求以下微分方程的通解。

$$x\frac{\mathrm{d}z}{\mathrm{d}x}-z=x^2\mathrm{e}^x.$$

解： 齐次方程

$$x\frac{\mathrm{d}z}{\mathrm{d}x}-z=0,$$

是分离变量方程，其通解为 $z=Cx$。

设 $u(x)\mathrm{e}^x$ 是非齐次方程的一个特解，则

$$xu'+xu-u=x^2.$$

如果 $u(x)$ 是多项式，则上述方程左侧中以 xu 次数最高，经观察知 $u(x)=x$ 是特解。所以非齐次方程通解为 $z=Cx+x\mathrm{e}^x$。 □

例 8.4.6 (Bernoulli方程). 形如

$$y' + a(x)y = b(x)y^\beta, \quad \beta \neq 0, 1$$

的一阶微分方程称为**Bernoulli方程**。它不是线性方程，但是很容易变形为线性方程：两边除以 y^β，得到

$$\left(\frac{1}{y^{\beta-1}}\right)' + (1-\beta)a(x)\left(\frac{1}{y^{\beta-1}}\right) = (1-\beta)b(x),$$

这是关于 $\frac{1}{y^{\beta-1}}$ 的一阶线性方程。由此可以解出 $y(x)$。 □

高阶线性方程

高阶线性微分方程可以看做关于微分算子 $\frac{\mathrm{d}}{\mathrm{d}x}$ 的多项式，因此可以尝试通过“因子分解”降阶。

例 8.4.7. 求以下方程的通解

$$xy'' - (x+1)y' + y = x^2\mathrm{e}^x.$$

解： 由待定系数法

$$x\frac{\mathrm{d}^2}{\mathrm{d}x^2} - (x+1)\frac{\mathrm{d}}{\mathrm{d}x} + 1 = \left(x\frac{\mathrm{d}}{\mathrm{d}x} - \alpha\right)\left(\frac{\mathrm{d}}{\mathrm{d}x} - \beta\right) = x\frac{\mathrm{d}^2}{\mathrm{d}x^2} - (\beta x + \alpha)\frac{\mathrm{d}}{\mathrm{d}x} + \alpha\beta.$$

可得 $\alpha = \beta = 1$，因此原方程即

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x} - 1\right)\left(\frac{\mathrm{d}}{\mathrm{d}x} - 1\right)y = x^2\mathrm{e}^x.$$

令 $z = \frac{\mathrm{d}y}{\mathrm{d}x} - y$。则原方程等价于由两个一阶线性方程组成的方程组

$$\begin{cases} x\frac{\mathrm{d}z}{\mathrm{d}x} - z = x^2\mathrm{e}^x, \\ \frac{\mathrm{d}y}{\mathrm{d}x} - y = z. \end{cases}$$

由第一个方程解得其通解 $z = C_1x + x\mathrm{e}^x$。再由第二个方程利用叠加原理解得

$$y = C_2\mathrm{e}^x - C_1(x+1) + \frac{x^2}{2}\mathrm{e}^x.$$

□

例 8.4.8. 求 $x^2y'' - 2y = x^4$ 的通解。

解： 微分方程可以分解为

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x} + 1\right)\left(x\frac{\mathrm{d}}{\mathrm{d}x} - 2\right)y = x^4,$$

从而可以仿照上例的办法求解(过程略)。

然而，我们通过直接观察 $y_1(x) = x^2$ 是齐次方程 $x^2y'' - 2y = 0$ 的一个解，$y_0(x) = \frac{x^4}{10}$ 是非齐次方程 $x^2y'' - 2y = x^4$ 的一个解。

用常数变易法，令 $y(x) = x^2C(x)$ 是齐次方程的解。则

$$x^2(x^2C'' + 4xC') = 0,$$

即

$$xC'' + 4C' = 0.$$

解得

$$C(x) = \frac{C_2}{x^3} + C_1,$$

从而原方程通解为

$$y(x) = \frac{x^4}{10} - \frac{C_2}{x} + C_1x^2.$$

□

例 8.4.9. 求常系数线性微分方程的通解

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = f(x).$$

解： 特征多项式的因式分解

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = \prod_{j=1}^{k}(\lambda - \lambda_j)^{m_j},$$

意味着微分方程的分解

$$\prod_{j=1}^{k}\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_j\right)^{m_j} y = f(x),$$

所以原方程等价于

$$\begin{cases} \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_1\right)^{m_1} y = z_1, \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_2\right)^{m_2} z_1 = z_2, \\ \vdots \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_k\right)^{m_1} z_{k-1} = f(x). \end{cases}$$

而每个形如

$$\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda\right)^{m} u = g(x)$$

的方程又等价于

$$\begin{cases} \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda\right) u = u_1, \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda\right) u_1 = u_2, \\ \vdots \\ \left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda\right) u_{m-1} = g(x). \end{cases}$$

该方程组是由一系列可以依次求解的一阶线性微分方程组成的。 □

高阶线性方程降阶的一般方法是常数变易法。

例 8.4.10 (高阶线性微分方程和常数变易法). 如果 $Y_0(x)$ 是齐次线性方程

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = 0$$

的一个解，满足 $Y_0(x) \neq 0(\,\forall x\,)$，则 $y(x) = Y_0(x)C(x)$ 是

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

的解当且仅当

$$\begin{aligned}\sum_{k=1}^{n}\binom{n}{k}Y_0^{(n-k)}(x)C^{(k)}(x) + a_{n-1}(x)\sum_{k=1}^{n-1}\binom{n-1}{k}Y_0^{(n-1-k)}(x)C^{(k)}(x)&\\ + \cdots + a_1(x)Y_0(x)C'(x)&\\ + (Y_0^{(n)} + a_{n-1}(x)Y_0^{(n-1)} + \cdots + a_1(x)Y_0' + a_0(x)Y_0)C(x) = f(x)&\end{aligned}$$

即

$$\begin{aligned}\sum_{k=1}^{n}\binom{n}{k}Y_0^{(n-k)}(x)C^{(k)} + a_{n-1}(x)\sum_{k=1}^{n-1}\binom{n-1}{k}Y_0^{(n-1-k)}(x)C^{(k)}&\\ + \cdots + a_1(x)Y_0(x)C' = f(x),&\end{aligned}$$

这是关于 C' 的一个 $n-1$ 阶线性微分方程，它的非齐次项 $f(x)$ 就是原方程的非齐次项(所以原方程为齐次方程时，这个方程也是齐次的)。当原方程是常系数齐次微分方程时，取 $Y_0(x) = \mathrm{e}^{\lambda x}$，其中 λ 是原方程的一个特征指数，于是这个关于 C' 的方程仍然是常系数线性齐次方程。 □

8.5 一阶线性微分方程组

一阶线性微分方程组是指形如

$$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$$

的方程，其中 $A(x)$ 是 n行n列的矩阵，

$$\mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}, \quad \mathbf{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}.$$

它的分量形式为

$$\begin{cases} y_1' = a_{11}(x)y_1 + \cdots + a_{1n}(x)y_n + f_1(x), \\ y_2' = a_{21}(x)y_1 + \cdots + a_{2n}(x)y_n + f_2(x), \\ \vdots \\ y_n' = a_{n1}(x)y_1 + \cdots + a_{nn}(x)y_n + f_n(x). \end{cases}$$

例 8.5.1. 记

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0(x) & -a_1(x) & -a_2(x) & \cdots & -a_{n-1}(x) \end{pmatrix},$$

$$\mathbf{f}(x) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ f(x) \end{pmatrix}.$$

则 $\mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}$ 是一阶线性方程组$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$ 的解，当且仅当 $\mathbf{y}(x) = \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix}$，其中 $y(x)$ 是以下高阶方程的解

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = f(x).$$

□

定理 8.5.2 (叠加原理). *若 $\mathbf{y}(x)$ 和 $\mathbf{z}(x)$ 分别是 $\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$ 和 $\mathbf{z}' = A(x)\mathbf{z} + \mathbf{g}(x)$ 的解，则 $\alpha\mathbf{y}(x) + \beta\mathbf{z}(x)$ 是 $\mathbf{u}' = A(x)\mathbf{u} + \alpha\mathbf{f}(x) + \beta\mathbf{g}(x)$ 的解。* □

推论 8.5.3. *1. 齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的所有解组成一个线性空间 V_0。*

2. 非齐次方程 $\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$ 的所有解组成一个基于 V_0 的仿射空间 V：对任意 $\mathbf{y}, \mathbf{z} \in V$，$\mathbf{y} - \mathbf{z} \in V_0$。

3. 非齐次方程 $\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$ 的通解是该方程的一个特解与齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的通解之和。

□

定理 8.5.4. 设 $A(x)$ 连续，则对任何 x_0 及 $\mathbf{y}_0 \in \mathbb{R}^n$，初值问题

$$\begin{cases} \mathbf{y}' = A(x)\mathbf{y}, \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

有解。

设 $\mathbf{y}_1(x), \mathbf{y}_2(x), \ldots, \mathbf{y}_n(x)$ 是齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的 n 个解。记

$$U(x) = (\mathbf{y}_1(x), \mathbf{y}_2(x), \ldots, \mathbf{y}_n(x)),$$

即 $U(x)$ 是以 $\mathbf{y}_1(x), \mathbf{y}_2(x), \ldots, \mathbf{y}_n(x)$ 为列向量的 n 行n 列的矩阵。则

$$\begin{aligned} U' &= (\mathbf{y}_1'(x), \mathbf{y}_2'(x), \ldots, \mathbf{y}_n'(x)) \\ &= (A(x)\mathbf{y}_1(x), A(x)\mathbf{y}_2(x), \ldots, A(x)\mathbf{y}_n(x)) \\ &= A(x)U(x). \end{aligned}$$

因此我们称 $U(x)$ 为齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的矩阵解。记

$$W(x) = \det U(x) = \det(\mathbf{y}_1(x), \mathbf{y}_2(x), \ldots, \mathbf{y}_n(x)),$$

称为 n 个解 $\mathbf{y}_1(x), \mathbf{y}_2(x), \ldots, \mathbf{y}_n(x)$ 的 Wronsky 行列式。

定理 8.5.5 (Liouville). 设 $W(x)$ 为齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的 n 个解的 *Wronsky* 行列式。则

$$W'(x) = \mathrm{tr}A(x)W(x).$$

从而 $W(x) = W(x_0)e^{\int_{x_0}^{x} \mathrm{tr}A(s)\mathrm{d}s}$。

证明. 由行列式的定义和求导的Leibniz法则知

$$W'(x) = \sum_{k=1}^{n} W_k(x),$$

其中 $W_k(x)$ 是对行列式 $W(x)$ 的第 k 行求导而其他行不变所得到的行列式。

由微分方程组知

$$W(x) \text{ 的第 } k \text{ 行的导数} = a_{kk}(x) \times W(x) \text{ 的第} k \text{行} + \text{其他行的线性组合}.$$

所以由行列式的性质知 $W_k(x) = a_{kk}(x)W(x)$，因此

$$W'(x) = \mathrm{tr}A(x)W(x).$$

从而 $W(x) = W(x_0)\mathrm{e}^{\int_{x_0}^{x} \mathrm{tr}A(s)\mathrm{d}s}$。 □

推论 8.5.6. *1.* 设 $W(x)$ 为齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的 n 个解的 *Wronsky* 行列式。则对任意 x，$W(x) = 0$ 当且仅当存在 x_0 使得 $W(x_0) = 0$。

2. 设 $\mathbf{y}_1(x),\ldots,\mathbf{y}_n(x)$ 是齐次方程 $\mathbf{y}'=A(x)\mathbf{y}$ 的 n 个解。则对任意 x，$\mathbf{y}_1(x),\ldots,\mathbf{y}_n(x)$线性无关当且仅当存在 x_0 使得 $\mathbf{y}_1(x_0),\ldots,\mathbf{y}_n(x_0)$线性无关。

□

定理 8.5.7. 设 $A(x)$ 和 $\mathbf{f}(x)$ 连续，则对任何 x_0 及 $\mathbf{y}_0\in\mathbb{R}^n$，初值问题

$$\begin{cases}\mathbf{y}'=A(x)\mathbf{y}+\mathbf{f}(x),\\ \mathbf{y}(x_0)=\mathbf{y}_0\end{cases}$$

有唯一解。

证明. 由解的存在性定理知初值问题

$$\begin{cases}\mathbf{y}'=A(x)\mathbf{y},\\ \mathbf{y}(x_0)=\mathbf{e}_k\end{cases}$$

有解 $\mathbf{y}_k(x)$。再由 Liouville 定理推论知对任意 x，$\mathbf{y}_1(x),\ldots,\mathbf{y}_n(x)$ 线性无关，从而对任意 x，矩阵 $U(x)=(\mathbf{y}_1(x),\mathbf{y}_2(x),\ldots,\mathbf{y}_n(x))$ 可逆。

[**常数变易法**] $\mathbf{y}(x)=U(x)\mathbf{C}(x)$ 是 $\mathbf{y}'=A(x)\mathbf{y}+\mathbf{f}(x)$ 的解当且仅当

$$U(x)\mathbf{C}'(x)=\mathbf{f}(x),$$

即

$$\mathbf{C}'(x)=U(x)^{-1}\mathbf{f}(x).$$

因此

$$\mathbf{C}(x)=\mathbf{C}_0+\int_{x_0}^{x}U(t)^{-1}\mathbf{f}(t)\mathrm{d}t.$$

于是

$$\mathbf{y}(x)=U(x)\mathbf{C}_0+U(x)\int_{x_0}^{x}U(t)^{-1}\mathbf{f}(t)\mathrm{d}t$$

是 $\mathbf{y}'=A(x)\mathbf{y}+\mathbf{f}(x)$ 的通解。

当 $x=x_0$ 时，$U(x_0)=I_n$，从而 $\mathbf{y}(x_0)=\mathbf{C}_0$。所以

$$\mathbf{y}(x)=U(x)\mathbf{y}_0+U(x)\int_{x_0}^{x}U(t)^{-1}\mathbf{f}(t)\mathrm{d}t$$

是初值问题

$$\begin{cases}\mathbf{y}'=A(x)\mathbf{y}+\mathbf{f}(x),\\ \mathbf{y}(x_0)=\mathbf{y}_0\end{cases}$$

的解。

当 $\mathbf{f} \equiv 0$ 时，$\mathbf{y}(x) = U(x)\mathbf{C}_0$ 是齐次方程 $\mathbf{y}' = A(x)\mathbf{y}$ 的通解。若 $\mathbf{y}(x_0) = \mathbf{0}$，则 $\mathbf{C}_0 = \mathbf{0}$，从而 $\mathbf{y}(x) \equiv \mathbf{0}$。即初值问题

$$\begin{cases} \mathbf{y}' = A(x)\mathbf{y}, \\ \mathbf{y}(x_0) = \mathbf{0} \end{cases}$$

只有零解。

若 $\mathbf{y}_1(x)$ 和 $\mathbf{y}_2(x)$ 都是初值问题

$$\begin{cases} \mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x), \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

的解，则 $\mathbf{y}(x) = \mathbf{y}_1(x) - \mathbf{y}_2(x)$ 是初值问题

$$\begin{cases} \mathbf{y}' = A(x)\mathbf{y}, \\ \mathbf{y}(x_0) = \mathbf{0} \end{cases}$$

的解，从而 $\mathbf{y}(x) = \mathbf{y}_1(x) - \mathbf{y}_2(x) \equiv \mathbf{0}$，即 $\mathbf{y}_1(x) \equiv \mathbf{y}_2(x)$。因此初值问题

$$\begin{cases} \mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x), \\ \mathbf{y}(x_0) = \mathbf{y}_0 \end{cases}$$

有唯一解。 □

上述证明中的办法是解一阶线性方程组的常数变易法。

推论 8.5.8. $\mathbf{y}' = A(x)\mathbf{y}$ 的解空间 V_0 是 n 维线性空间；$\mathbf{y}' = A(x)\mathbf{y} + \mathbf{f}(x)$ 的解空间是基于 V_0 的一个 n 维仿射空间。 □

例 8.5.9 (简谐振动与共振). 求 $y'' + y = \sin(\omega x)$ 的通解。

解： $y'' + y = \sin(\omega x)$对应于方程组

$$\mathbf{y}' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{y} + \begin{pmatrix} 0 \\ \sin(\omega x) \end{pmatrix}.$$

因为 $\sin x$ 和 $\cos x$ 是齐次方程 $y'' + y = 0$ 的解，所以

$$\begin{pmatrix} \cos x \\ \cos' x \end{pmatrix} = \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix}, \quad \begin{pmatrix} \sin x \\ \sin' x \end{pmatrix} = \begin{pmatrix} \sin x \\ \cos x \end{pmatrix}$$

是上述方程组的齐次部分的解，

$$U(x) = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}$$

是相应的矩阵解，满足 $\det U(x) = 1$。设 $\mathbf{y}(x) = U(x)\mathbf{C}(x)$ 是上述非齐次线性方程组的解。则

$$U(x)\mathbf{C}'(x) = \begin{pmatrix} 0 \\ \sin(\omega x) \end{pmatrix}.$$

解得

$$\begin{aligned}\mathbf{C}'(x) = U(x)^{-1}\begin{pmatrix} 0 \\ \sin(\omega x) \end{pmatrix} &= \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}\begin{pmatrix} 0 \\ \sin(\omega x) \end{pmatrix} \\ &= \begin{pmatrix} -\sin x \sin(\omega x) \\ \cos x \sin(\omega x) \end{pmatrix} = \begin{pmatrix} \frac{\cos(1+\omega)x - \cos(1-\omega)x}{2} \\ \frac{\sin(1+\omega)x - \sin(1-\omega)x}{2} \end{pmatrix}.\end{aligned}$$

积分得到

$$\mathbf{C}(x) = \begin{pmatrix} \frac{\sin(1+\omega)x}{2(1+\omega)} - \frac{\sin(1-\omega)x}{2(1-\omega)} + A \\ \frac{\cos(1-\omega)x}{2(1-\omega)} - \frac{\cos(1+\omega)x}{2(1+\omega)} + B \end{pmatrix}, \quad \text{若 } \omega \neq 1;$$

或

$$\mathbf{C}(x) = \begin{pmatrix} \frac{\sin 2x}{4} - \frac{x}{2} + A \\ -\frac{\cos 2x}{4} + B \end{pmatrix}, \quad \text{若}\omega = 1;$$

因此，当 $\omega \neq 1$ 时，

$$\begin{aligned} y(x) &= \left(\frac{\sin(1+\omega)x}{2(1+\omega)} - \frac{\sin(1-\omega)x}{2(1-\omega)} + A\right)\cos x \\ &\quad + \left(\frac{\cos(1-\omega)x}{2(1-\omega)} - \frac{\cos(1+\omega)x}{2(1+\omega)} + B\right)\sin x \\ &= \frac{\sin \omega x}{1-\omega^2} + A\cos x + B\sin x \end{aligned}$$

是原二阶方程通解，它们都是有界的。

当 $\omega = 1$ 时，

$$\begin{aligned} y(x) &= \left(\frac{\sin 2x}{4} - \frac{x}{2} + A\right)\cos x \\ &\quad + \left(-\frac{\cos 2x}{4} + B\right)\sin x \\ &= \left(B + \frac{1}{4}\right)\sin x + A\cos x - \frac{x}{2}\cos x \end{aligned}$$

是原二阶方程通解，它们都不是有界的。

$y'' + y = 0$刻画了一个简谐振动，其频率为 1。$y'' + y = \sin(\omega x)$ 的非齐次项 $\sin \omega x$ 是一个周期外力，其频率为 ω。当外力频率与系统固有频率不同时，所有解都是有界解。当外力频率与系统固有频率相同时，所有解都是无界的，这就是共振现象。 $\square$

定义 8.5.10. 设 $y_1(x),\dots,y_n(x)$ 是 n 阶线性齐次方程

$$y^{(n)}+a_{n-1}(x)y^{(n-1)}+\cdots+a_1(x)y'+a_0(x)y=0$$

的 n 个解。称 $y_1(x),\dots,y_n(x)$**线性无关**，如果如下矩阵可逆

$$\begin{pmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1'(x) & y_2'(x) & \cdots & y_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{pmatrix}.$$

设 $y_1(x),\dots,y_n(x)$ 是 n 阶线性齐次方程

$$y^{(n)}+a_{n-1}(x)y^{(n-1)}+\cdots+a_1(x)y'+a_0(x)y=0$$

的 n 个线性无关解。则对非齐次方程

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

的任何解 $y(x)$，可设

$$\begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix} = \begin{pmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1'(x) & y_2'(x) & \cdots & y_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{pmatrix} \begin{pmatrix} C_1(x) \\ C_2(x) \\ \vdots \\ C_n(x) \end{pmatrix}.$$

于是

$$\begin{pmatrix} y_1(x) & y_2(x) & \cdots & y_n(x) \\ y_1'(x) & y_2'(x) & \cdots & y_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_n^{(n-1)}(x) \end{pmatrix} \begin{pmatrix} C_1'(x) \\ C_2'(x) \\ \vdots \\ C_n'(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ f(x) \end{pmatrix}.$$

这就是求解高阶线性微分方程的常数变易法。从这个(代数)线性方程组解得 $C_k'(x)$ ($k=1,\dots,n$)，然后积分得到 $C_1(x),\dots,C_n(x)$，从而得到微分方程

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

的通解

$$y(x)=C_1(x)y_1(x)+\cdots+C_n(x)y_n(x).$$

8.6 常系数线性微分方程(组)

常系数线性高阶微分方程

对常系数线性方程

$$y^{(n)}+a_{n-1}y^{(n-1)}+\cdots+a_1y'+a_0y=f(x),$$

我们要解决两个问题：构造齐次方程的 n 个线性无关解，找到非齐次方程的一个特解。

求非齐次方程的特解的一般方法是常数变易法(见上节)，但这涉及到解代数方程组以及积分。对特殊形式的非齐次项，我们可以有更简便的办法。

我们称 $f(x)$ 是一个**拟多项式**，如果 $f(x)=P(x)\mathrm{e}^{\lambda x}$，其中 $P(x)$ 是多项式，λ 可以是实数，也可以是复数，也就是包括 $P(x)\mathrm{e}^{\alpha x}\sin\beta x$ 或 $P(x)\mathrm{e}^{\alpha x}\cos\beta x$。

这时我们设 $y(x)=z(x)\mathrm{e}^{\lambda x}$。于是

$$z^{(n)}\mathrm{e}^{\lambda x}+b_{n-1}z^{(n-1)}\mathrm{e}^{\lambda x}+\cdots+b_1z'\mathrm{e}^{\lambda x}+b_0z\mathrm{e}^{\lambda x}=P(x)\mathrm{e}^{\lambda x},$$

从而

$$z^{(n)}+b_{n-1}z^{(n-1)}+\cdots+b_1z'+b_0z=P(x).$$

这是关于 z 的一个常系数 n 阶线性微分方程，非齐次项是关于 x 的多项式。然后可以用待定系数法求得它的一个多项式解 $Q(x)$。从而 $Q(x)\mathrm{e}^{\lambda x}$ 是原方程的一个特解。

例 8.6.1. 求方程 $y''+y=\sin(\omega x)$ $(\omega>0)$的解。

解： 微分方程 $y''+y=\mathrm{e}^{\mathrm{i}\omega x}$ 的虚部就是原方程。下面求解这个复方程。设 $y(x)=Q(x)\mathrm{e}^{\mathrm{i}\omega x}$。代入上述复方程，得到

$$Q''+2\mathrm{i}\omega Q'+(1-\omega^2)Q=1.$$

当 $\omega\neq 1$ 时，考虑 Q 是多项式，由于上式等号右端是常数，所以 Q 为常数，易见

$$Q=\frac{1}{1-\omega^2},$$

于是

$$y=\frac{1}{1-\omega^2}\mathrm{e}^{\mathrm{i}\omega x}$$

是复方程的一个特解。易见 $\mathrm{e}^{\mathrm{i}x},\mathrm{e}^{-\mathrm{i}x}$ 是齐次方程的解，由常数变易法知齐次方程通解为

$$y=C_1\mathrm{e}^{\mathrm{i}x}+C_2\mathrm{e}^{-\mathrm{i}x},$$

从而非齐次方程通解为

$$y=C_1\mathrm{e}^{\mathrm{i}x}+C_2\mathrm{e}^{-\mathrm{i}x}+\frac{1}{1-\omega^2}\mathrm{e}^{\mathrm{i}\omega x}.$$

取它的虚部，得到原方程通解为

$$y = A\cos x + B\sin x + \frac{1}{1-\omega^2}\sin(\omega x).$$

此时，所有解都是有界的振荡。

当 $\omega = 1$ 时，

$$Q'' + 2\mathrm{i}Q' = 1.$$

易见

$$Q = \frac{x}{2\mathrm{i}}$$

是该方程特解，于是

$$y = \frac{x}{2\mathrm{i}}\mathrm{e}^{\mathrm{i}x}$$

是复方程的一个特解。类似前一情形的讨论，得到原方程通解

$$y = A\cos x + B\sin x - \frac{x}{2}\cos x.$$

此时，虽然非齐次项（外力）是有界的，但方程的每个解都是无界的振荡。 □

对常系数齐次线性方程

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0,$$

由特征多项式的因式分解

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = \prod_{j=1}^{k}(\lambda - \lambda_j)^{m_j},$$

我们知道

$$\mathrm{e}^{\lambda_j x}, x\mathrm{e}^{\lambda_j x}, \ldots, \frac{x^{m_j-1}}{(m_j-1)!}\mathrm{e}^{\lambda_j x}, \quad j = 1, \ldots, k$$

是齐次方程的 n 个解。我们证明它们是线性无关的。

事实上，我们证明对任意 k 个彼此不同的复数 $\lambda_1, \ldots, \lambda_k$，以及任意正整数 N，我们证明

$$\mathrm{e}^{\lambda_j x}, x\mathrm{e}^{\lambda_j x}, \ldots, \frac{x^N}{N!}\mathrm{e}^{\lambda_j x}, \quad j = 1, \ldots, k$$

是线性无关的。这 $k(N+1)$ 个函数是 $k(N+1)$ 阶齐次线性微分方程

$$\prod_{j=1}^{k}\left(\frac{\mathrm{d}}{\mathrm{d}x} - \lambda_j\right)^{N+1} y = 0$$

的解。

记 $M=k(N+1)$，上述函数为 $y_1(x),y_2(x),\dots,y_M(x)$。假设它们线性相关，则存在 x_0 及不全为零的常数 $C_1,\dots,C_M$ 使得

$$\begin{pmatrix} y_1(x_0) & y_2(x_0) & \cdots & y_M(x_0) \\ y_1'(x_0) & y_2'(x_0) & \cdots & y_M'(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(M-1)}(x_0) & y_2^{(M-1)}(x_0) & \cdots & y_M^{(M-1)}(x_0) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ C_M \end{pmatrix} = 0.$$

于是 $y(x)=C_1y_1(x)+C_2y_2(x)+\cdots+C_My_M(x)$ 是该齐次线性微分方程满足初始条件

$$y(x_0)=y'(x_0)=\cdots=y^{(M-1)}(x_0)=0$$

的解。由线性方程解的唯一性知 $y(x)\equiv 0$，即

$$C_1y_1(x)+C_2y_2(x)+\cdots+C_My_M(x)=0,\quad \forall x.$$

注意到

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}-\mu\right)\frac{x^m}{m!}\mathrm{e}^{\lambda x}=\frac{x^{m-1}}{(m-1)!}\mathrm{e}^{\lambda x}+(\lambda-\mu)\frac{x^m}{m!}\mathrm{e}^{\lambda x},$$

所以由

$$\left(\frac{\mathrm{d}}{\mathrm{d}x}-\lambda_1\right)^N\prod_{j=2}^{k}\left(\frac{\mathrm{d}}{\mathrm{d}x}-\lambda_j\right)^{N+1}(C_1y_1(x)+C_2y_2(x)+\cdots+C_My_M(x))=0$$

得到 $C_1y_1(x)+C_2y_2(x)+\cdots+C_My_M(x)$ 中 $\frac{x^N}{N!}\mathrm{e}^{\lambda_1 x}$ 的系数为零。同理可证每个 $y_j(x)$ 的系数都是零。这与 $C_1,\dots,C_M$ 不全为零矛盾。因此

$$\mathrm{e}^{\lambda_j x},x\mathrm{e}^{\lambda_j x},\dots,\frac{x^N}{N!}\mathrm{e}^{\lambda_j x},\quad j=1,\dots,k$$

线性无关。

Euler方程

形如

$$x^ny^{(n)}+a_{n-1}x^{n-1}y^{(n-1)}+\cdots+a_1xy'+a_0y=f(x)$$

的线性方程称为**Euler方程**。

记 $t=\ln x$，则 $x=\mathrm{e}^t$。对任何函数 $u=u(x)=u(\mathrm{e}^t)$，

$$\frac{\mathrm{d}u}{\mathrm{d}t}=\frac{\mathrm{d}u}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}=x\frac{\mathrm{d}u}{\mathrm{d}x}.$$

所以

$$\frac{\mathrm{d}}{\mathrm{d}t}=x\frac{\mathrm{d}}{\mathrm{d}x}.$$

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

所以

$$x^2\frac{\mathrm{d}^2}{\mathrm{d}x^2}=\frac{\mathrm{d}^2}{\mathrm{d}t^2}-\frac{\mathrm{d}}{\mathrm{d}t}=\left(\frac{\mathrm{d}}{\mathrm{d}t}-1\right)\frac{\mathrm{d}}{\mathrm{d}t}.$$

猜测

$$x^k\frac{\mathrm{d}^k}{\mathrm{d}x^k}=\left(\frac{\mathrm{d}}{\mathrm{d}t}-(k-1)\right)\cdots\left(\frac{\mathrm{d}}{\mathrm{d}t}-1\right)\frac{\mathrm{d}}{\mathrm{d}t}.$$

这个猜测可以用数学归纳法证明其成立。因为

$$x\frac{\mathrm{d}}{\mathrm{d}x}\left(x^k\frac{\mathrm{d}^k}{\mathrm{d}x^k}\right)=x^{k+1}\frac{\mathrm{d}^{k+1}}{\mathrm{d}x^{k+1}}+kx^k\frac{\mathrm{d}^k}{\mathrm{d}x^k},$$

所以

$$x^{k+1}\frac{\mathrm{d}^{k+1}}{\mathrm{d}x^{k+1}}=\left(x\frac{\mathrm{d}}{\mathrm{d}x}-k\right)\left(x^k\frac{\mathrm{d}^k}{\mathrm{d}x^k}\right)=\left(\frac{\mathrm{d}}{\mathrm{d}t}-k\right)\left(\frac{\mathrm{d}}{\mathrm{d}t}-(k-1)\right)\cdots\frac{\mathrm{d}}{\mathrm{d}t}.$$

后者是关于变量 t 的 $k+1$ 阶常系数线性微分算子。因此Euler方程可以写成

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n}+b_{n-1}\frac{\mathrm{d}^{n-1}y}{\mathrm{d}t^{n-1}}+\cdots+b_1\frac{\mathrm{d}y}{\mathrm{d}t}+b_0y=f(\mathrm{e}^t)$$

的形式，这是常系数线性微分方程。

常系数一阶线性微分方程组

考虑常系数线性齐次方程组

$$\mathbf{y}'=A\mathbf{y}.$$

如果 $\mathbf{v}_1,\ldots,\mathbf{v}_k$ 是 A 的一组线性无关的特征向量，分别对应于 A 的特征值 $\lambda_1,\ldots,\lambda_k$，即 $A\mathbf{v}_j=\lambda_j\mathbf{v}_j$。不难验证对任意常数 $C_1,\ldots,C_k$，

$$\mathbf{y}(x)=C_1\mathrm{e}^{\lambda_1 x}\mathbf{v_1}+\cdots+C_k\mathrm{e}^{\lambda_k x}\mathbf{v_k}$$

是上述齐次方程的解。如果 $k=n$，则这给出齐次方程的通解。

然而很多矩阵不具备构成空间的一组特征向量，用线性代数的语言，有的特征值其几何重数小于代数重数。这时需要考虑广义特征向量。

对任何复数 λ，考虑线性子空间序列

$$\{\mathbf{0}\}\subseteq\mathrm{Ker}(A-\lambda I)\subseteq\mathrm{Ker}(A-\lambda I)^2\subseteq\cdots$$

如果 $\mathrm{Ker}(A-\lambda I)^k=\mathrm{Ker}(A-\lambda I)^{k+1}$，则

$$\begin{aligned}(A-\lambda I)^{k+2}\mathbf{v}=0&\Rightarrow(A-\lambda I)\mathbf{v}\in\mathrm{Ker}(A-\lambda I)^{k+1}=\mathrm{Ker}(A-\lambda I)^k\\&\Rightarrow(A-\lambda I)^{k+1}\mathbf{v}=0,\end{aligned}$$

从而 $\mathrm{Ker}(A-\lambda I)^{k+1}=\mathrm{Ker}(A-\lambda I)^{k+2}$。

如果 λ 是 A 的特征值，则 $\mathrm{Ker}(A-\lambda I)\neq\{\mathbf{0}\}$，$\dim\mathrm{Ker}(A-\lambda I)$ 是 λ 的几何重数。由于维数的限制，所以上述子空间序列中存在 k 使得，$\mathrm{Ker}(A-$

$\lambda I)^k = \operatorname{Ker}(A-\lambda I)^{k+1}$，并且对任意 $j<k$，$\operatorname{Ker}(A-\lambda I)^j \subset \operatorname{Ker}(A-\lambda I)^{j+1}$。$\dim \operatorname{Ker}(A-\lambda I)^k$ 就是 λ 的代数重数。

设 λ 是 A 的一个特征值，

$$\mathbf{v}_0 = \mathbf{0} \leftarrow \mathbf{v}_1 \leftarrow \mathbf{v}_2 \leftarrow \cdots \leftarrow \mathbf{v}_k$$

是相应于 λ 的一个特征向量-广义特征向量链，即 $(A-\lambda I)\mathbf{v}_j = \mathbf{v}_{j-1}$。设

$$\mathbf{y}(x) = c_1(x)\mathbf{v}_1 + c_2(x)\mathbf{v}_2 + \cdots + c_k(x)\mathbf{v}_k$$

是 $\mathbf{y}' = A\mathbf{y}$ 的解，则

$$\begin{aligned}
&c_1'(x)\mathbf{v}_1 + c_2'(x)\mathbf{v}_2 + \cdots + c_k'(x)\mathbf{v}_k \\
=&A\left(c_1(x)\mathbf{v}_1 + c_2(x)\mathbf{v}_2 + \cdots + c_k(x)\mathbf{v}_k\right) \\
=&c_1(x)\lambda\mathbf{v}_1 + c_2(x)(\mathbf{v}_1 + \lambda\mathbf{v}_2) + \cdots + c_k(x)(\mathbf{v}_{k-1} + \lambda\mathbf{v}_k) \\
=&(\lambda c_1(x) + c_2(x))\mathbf{v}_1 + \cdots + (\lambda c_{k-1}(x) + c_k(x))\mathbf{v}_{k-1} + \lambda c_k(x)\mathbf{v}_k.
\end{aligned}$$

所以

$$\begin{cases}
c_1' - \lambda c_1 = c_2, \\
c_2' - \lambda c_2 = c_3, \\
\vdots \\
c_{k-1}' - \lambda c_{k-1} = c_k, \\
c_k' - \lambda c_k = 0
\end{cases}$$

所以

$$\begin{cases}
c_1(x) = C_1\mathrm{e}^{\lambda x} + C_2 x\mathrm{e}^{\lambda x} + C_3\frac{x^2}{2}\mathrm{e}^{\lambda x} + \cdots + C_k\frac{x^{k-1}}{(k-1)!}\mathrm{e}^{\lambda x} \\
c_2(x) = C_2\mathrm{e}^{\lambda x} + C_3 x\mathrm{e}^{\lambda x} + \cdots + C_k\frac{x^{k-2}}{(k-2)!}\mathrm{e}^{\lambda x} \\
\vdots \\
c_{k-1}(x) = C_{k-1}\mathrm{e}^{\lambda x} + C_k x\mathrm{e}^{\lambda x}, \\
c_k(x) = C_k\mathrm{e}^{\lambda x}
\end{cases}$$

因此

$$\begin{aligned}
\mathbf{y}(x) &= c_1(x)\mathbf{v}_1 + c_2(x)\mathbf{v}_2 + \cdots + c_k(x)\mathbf{v}_k \\
&= C_1\mathrm{e}^{\lambda x}\mathbf{v}_1 + C_2\mathrm{e}^{\lambda x}(x\mathbf{v}_1 + \mathbf{v}_2) + \cdots \\
&\quad + C_k\mathrm{e}^{\lambda x}\left(\frac{x^{k-1}}{(k-1)!}\mathbf{v}_1 + \frac{x^{k-2}}{(k-2)!}\mathbf{v}_2 + \cdots + x\mathbf{v}_{k-1} + \mathbf{v}_k\right)
\end{aligned}$$

8.7 从线性振动到单摆

前面我们介绍了求解微分方程的一些办法，然而对绝大多数微分方程没有初等函数形式的解。下面我们以二阶微分方程为例，说明如何利用微分方程本身来研究它的解的性质。

胡克定律与简谐振动

根据胡克定律和牛顿第二定律，弹簧振子的运动由以下微分方程刻画

$$y'' = -qy,$$

其中 $y(t)$ 为时刻 t 时系在弹簧自由端的物体偏离平衡位置的位移。胡克定律说，弹簧的回复力的方向与位移方向相反，大小与位移大小成比例。

记 $E = \frac{(y')^2}{2} + \frac{qy^2}{2}$，则 E 是机械能，它是动能$\frac{(y')^2}{2}$ 与弹性势能$\frac{qy^2}{2}$ 之和。

$$\frac{\mathrm{d}E}{\mathrm{d}t} = y'(y'' + qy) = 0.$$

所以机械能守恒。

对给定的能量 $E > 0$，由 $\frac{qy^2}{2} \le E$解得 $-\sqrt{\frac{2E}{q}} \le y \le \sqrt{\frac{2E}{q}}$，

$$T = 4\int_0^{\frac{T}{4}} \mathrm{d}t = 4\int_0^{\sqrt{\frac{2E}{q}}} \frac{\mathrm{d}y}{\sqrt{2E - qy^2}} = \frac{4}{\sqrt{q}}\int_0^1 \frac{\mathrm{d}u}{\sqrt{1-u^2}} = \frac{2\pi}{\sqrt{q}}, \quad y = u\sqrt{\frac{2E}{q}}.$$

所以除零解外，所有解都是周期运动，周期与 E 无关，只与 q 有关。

阻尼振动

简谐振动方程是如下二阶线性微分方程的特殊情况：

$$y'' + py' + qy = 0,$$

它的解由初值 $y(t_0)$ 和 $y'(t_0)$唯一决定。我们称以 (y, y') 为坐标的平面为上述微分方程的**相平面**，相平面中的点为**相点**。上述微分方程确定了相平面中的一个向量场

$$\mathbf{v}(y, y') = \begin{pmatrix} y' \\ -qy - py' \end{pmatrix}$$

它就是微分方程组

$$\begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix}\begin{pmatrix} y \\ y' \end{pmatrix}$$

的右端。

上述微分方程的特征多项式为

$$\lambda^2 + p\lambda + q,$$

其零点 λ_1, λ_2 为特征根，也是上述方程组系数矩阵的特征值。

$$\lambda_1 + \lambda_2 = -p, \quad \lambda_1\lambda_2 = q.$$

记

$$\Delta = p^2 - 4q.$$

当 $\Delta > 0$ 时，两个特征值是相异实数，且对应于每个特征值 λ，$\begin{pmatrix} 1 \\ \lambda \end{pmatrix}$ 是相应的特征向量。在特征向量所在的直线(以下称为特征直线)上，

$$\mathbf{v}(k, k\lambda) = \begin{pmatrix} k\lambda \\ -qk - pk\lambda \end{pmatrix} = \lambda \begin{pmatrix} k \\ k\lambda \end{pmatrix} = k\lambda$$

即那里的向量场与特征直线平行，从而从一条特征直线上出发的相曲线永远位于这条特征直线上。特征直线的存在限制了相曲线在相平面中的旋转。如果特征值 $\lambda < 0$，则对应的特征直线上的相曲线随 $t \to +\infty$ 而趋于原点；如果特征值 $\lambda > 0$，则对应的特征直线上的相曲线随 $t \to -\infty$ 而趋于原点；如果特征值 $\lambda = 0$，则对应的特征直线上的每个点都保持静止。

(I) 当 $q < 0$ 时，两个特征值满足 $\lambda_1 < 0 < \lambda_2$。此时原点称为**鞍点**。相曲线在 $t \to +\infty$ 时趋于零点，当且仅当它位于 λ_1 对应的特征直线上；相曲线在 $t \to -\infty$ 时趋于零点，当且仅当它位于 λ_2 对应的特征直线上。其他相曲线在 $t \to \pm\infty$ 时都会远离原点。你可以想象马鞍上的水珠在重力作用下流淌的样子。

(II) 当 $q > 0$ 时，方程

$$y'' = -qy - py',$$

表明除了线性回复力，还有另外一个外力$-py'$，它的大小与速度 y'成正比。

机械能

$$E = \frac{qy^2}{2} + \frac{(y')^2}{2}$$

的水平集(等能量线)是以原点为中心的椭圆。

$$\frac{\mathrm{d}E}{\mathrm{d}t} = qyy' + y'y'' = y'(y'' + qy) = -p(y')^2.$$

(II.1) 当 $p = 0$ 时，即简谐振动，能量守恒，E 为常值，相曲线位于一条椭圆上。此时所有解都是围绕原点(平衡位置)的周期运动，原点是稳定的——初值在原点附近的相曲线永远位于原点附近，称原点为**中心**。

(II.2) 当 $p < 0$ 时，令 $s = -t$，则

$$\frac{\mathrm{d}y}{\mathrm{d}s} = -\frac{\mathrm{d}y}{\mathrm{d}t}, \qquad \frac{\mathrm{d}^2y}{\mathrm{d}s^2} = \frac{\mathrm{d}^2y}{\mathrm{d}t^2},$$

因此 $y(t)$ 是微分方程

$$\frac{\mathrm{d}^2y}{\mathrm{d}t^2} + p\frac{\mathrm{d}y}{\mathrm{d}t} + qy = 0$$

的解，当且仅当 $z(s) = y(-s)$ 是微分方程

$$\frac{\mathrm{d}^2z}{\mathrm{d}s^2} - p\frac{\mathrm{d}z}{\mathrm{d}s} + qz = 0$$

的解。所以通过相平面和时间的坐标变换

$$\begin{cases} z = y, \\ z' = -y', \\ s = -t, \end{cases}$$

情形(II.2)就变为情形(II.3)。

(II.3) 当 $p > 0$ 时，$\frac{\mathrm{d}E}{\mathrm{d}t} < 0$，$E$ 单调减，除原点外所有相曲线当 t 增加时会从每一条当前所在的椭圆形能量水平集进入该椭圆的内部。因此所有相曲线都最终不会远离原点。原点是**稳定**的。

(II.3.1) 当 $0 < p < 2\sqrt{q}$ 时，$\Delta < 0$，此时特征值为一对具有负实部的共轭的复特征值，此时除原点外的所有相曲线都是绕原点旋转且当 $t \to +\infty$ 时趋于原点的螺旋线，这意味着运动是围绕平衡位置不停摆动但振幅越来越小的振动。这是原点是渐近稳定的**焦点**。这是小阻尼情形。

(II.3.2) 当 $p > 2\sqrt{q}$ 时，$\Delta > 0$，此时特征值为两个不同的负数，向量场有两条相交的特征直线，它们限制了相曲线的旋转。除原点外，所有相曲线当 $t \to +\infty$ 时趋于原点，当 $t \to -\infty$ 时趋于无穷远。此时原点是渐近稳定的**结点**。这对应大阻尼情形。

(II.3.3) 当 $p = 2\sqrt{q} > 0$ 时，$\Delta > 0$，此时特征值为两个相等的负数，向量场只有一条特征直线，它限制了相曲线的旋转。除原点外，所有相曲线当 $t \to +\infty$ 时趋于原点，当 $t \to -\infty$ 时趋于无穷远。此时原点是渐近稳定的**结点**。这对应临界阻尼情形。

单摆

$$ml\theta'' = -mg\sin\theta.$$

即

$$\theta'' + \frac{g}{l}\sin\theta = 0.$$

记

$$E = \frac{(\theta')^2}{2} + \frac{g}{l}(1 - \cos\theta),$$

即 E 为单摆的机械能。

于是

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \theta'\left(\theta'' + \frac{g}{l}\sin\theta\right) = 0,$$

这意味着机械能守恒，相曲线位于等能量线上。

当 $E < \frac{2g}{l}$ 时，由

$$\theta' = \sqrt{2E - \frac{2g}{l}(1 - \cos\theta)}$$

得

$$T = 4\int_0^{T/4} \mathrm{d}t = 4\int_0^{\Theta} \frac{\mathrm{d}\theta}{\sqrt{2E - \frac{2g}{l}(1 - \cos\theta)}} = 2\sqrt{\frac{2l}{g}}\int_0^{\Theta} \frac{\mathrm{d}\theta}{\sqrt{\cos\theta - \cos\Theta}}$$

其中 $E=\frac{g}{l}(1-\cos\Theta)$ 即 $\Theta=\arccos(1-\frac{lE}{g})$。

因为

$$\int_0^{\Theta}\frac{\mathrm{d}\theta}{\sqrt{\cos\theta-\cos\Theta}}=\int_0^{\Theta}\frac{\mathrm{d}s}{\sqrt{\cos(\Theta-s)-\cos\Theta}}$$
$$=\int_0^{\Theta}\frac{\mathrm{d}s}{\sqrt{\cos\Theta(\cos s-1)+\sin\Theta\sin s}},$$

$$\frac{1}{\sqrt{\cos\Theta(\cos s-1)+\sin\Theta\sin s}}=\frac{1}{\sqrt{\sin\Theta}}\frac{1}{\sqrt{s}}(1+o(s)),\quad s\to 0,$$

所以当 $0<\Theta<\pi$ 即 $E<\frac{2g}{l}$ 时，积分 $\int_0^{\Theta}\frac{\mathrm{d}\theta}{\sqrt{\cos\theta-\cos\Theta}}$ 收敛，从而 $T<+\infty$，即对应的解都是周期解。E 的最小值 $E=0$ 对应运动的平衡态 $(\theta,\theta')=(0,0)$，它是中心，是稳定的。

当 $\Theta=\pi$ 即 $E=\frac{2g}{l}$ 时，

$$\frac{1}{\sqrt{\cos\Theta(\cos s-1)+\sin\Theta\sin s}}=\frac{1}{s}(1+o(s)),\quad s\to 0,$$

此时积分 $\int_0^{\Theta}\frac{\mathrm{d}\theta}{\sqrt{\cos\theta-\cos\Theta}}$ 发散，$T=+\infty$。

当 $E>\frac{2g}{l}$ 时，$|\theta'|=\sqrt{2E-\frac{2g}{l}(1-\cos\theta)}\geq\sqrt{2E-\frac{4g}{l}}>0$，所以 θ 关于 t 单调且无界。

$(\theta,\theta')=(\pi,0)$ 是不动点。在它附近，记 $\theta=\pi+u$

$$0=\theta''+\frac{g}{l}\sin\theta=u''-\frac{g}{l}u+o(u),\quad u\to 0.$$

所以 $(\theta,\theta')=(\pi,0)$ 是鞍点，是不稳定的。

习题8.7

1. 证明：$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}=f(x)$ 下的运动都是“能量”守恒的。

$$E=\frac{{x'}^2}{2}+F(x),\quad F(x)\text{是 }-f(x)\text{ 的一个原函数。}$$

并且对 F 的任何极小值点 x_0，平衡解 $x(t)=x_0$ 是稳定的。

2. 对二阶线性微分方程

$$y''+py'+qy=0,$$

令 $z=y'/y$，证明

$$z'=-q-pz-z^2.$$

讨论这个方程的解的极限，说明相曲线随 $t\to+\infty$ 时的渐近线，并说明原方程解的渐近行为。

3. 对单摆，运动的周期 T 依赖于能量 E：

$$T(E) = 2\sqrt{\frac{2l}{g}}\int_0^{\Theta(E)} \frac{\mathrm{d}\theta}{\sqrt{\cos\theta - \cos\Theta(E)}},$$

其中 $\Theta(E) = \arccos\left(1 - \frac{lE}{g}\right)$。

(a) 求 $T(0) = \lim\limits_{E\to 0} T(0)$ 和 $T'(0)$。

(b) 证明 $T(E)$ 关于能量 E 是增函数。

索引