微积分学笔记

zhcosin

东汉建武

序

微积分是近代建立起来的数学理论，历经数代数学家不断添砖加瓦，构成了今天我们所看到的一部宏伟而严丝合缝的理论，它是高等数学的开端，使人类的数学思维和手段经历了一次质的飞跃。

微积分理论体系严谨，整个理论体系大致分为极限论、微分学、积分学和级数理论四部分，其中极限论是其余各部分的基础，也是初等数学和高等数学 (局限于分析学) 的分界点，包含数列极限和函数极限两部分内容，微分学包含导数和微分、微分中值定理，还有延伸学科微分方程 (又细分为常微分方程和偏微分方程)，积分学包含不定积分、定积分、无穷积分与瑕积分、重积分、曲线积分与曲面积分，级数理论则包括正项级数理论和一般项级数理论、傅里叶级数以及无穷乘积等内容。

我写这份笔记的主要参考书是前苏联数学家菲赫金哥尔茨的《微积分学教程》(高等教育出版社)，这部书篇幅巨大（三卷本），内容丰富，除了微积分的理论体系之外，还包含大量的例子和材料，那些例子并不仅仅是单纯的例题，而是微积分理论中由众多数学家所推导出的大量重要的结果，而应用材料则囊括了数学上的、物理上的、工程学上的、天文学上的，几乎遍及微积分的每一个角落，所以此书取材的广度和深度，都让人叹为观止。该书另一个突出特点是，在严格性和易读性之间，把握了一个极好的平衡，在不失严格性的同时，又保持了极好的可读性，所以该书无论是用于自学，还是用于教学，都具有极大的参考价值。

之所以写作这份笔记，是因为上大学期间贪玩，微积分学的非常粗糙，然而像这样的理论体系，如果不仔细的弄清楚它的每一个细节，将永远只是一个模糊的印象，这份笔记的意义便在于此，远期也可以考虑把它写成一份可以供他人自学的材料。

致谢也是不能缺少的，首先要感谢的是本书参考文献的作者们，是他们让我接触到了这么多精彩的数学内容，当然还有一些书没有在参考文献之列（其实是我也想不起来是从哪儿看到的了)，也一并感谢。还要感谢的是《计算机程序设计艺术》一书的作者 Donald E. Knuth，他开发的 TEX 排版系统以及由之发展而来的 LATEX 系统，使得我排版自己的书籍成为可能，还有编辑器 Emacs 的作者 Richard Stallman，这个编辑器所带来的强大的功能和编辑体验对我完成这份笔记功不可没。最后还要感谢我的夫人和女儿，女儿的降生给了我们这个家庭前所未有的欢乐，我对她的信心是在她的学生生涯，数学学科不至于成为她的拦路虎。夫人在照顾女儿上的付出才让我得以有精力来完成这份笔记。还有我的父母和哥哥。

zhcosin<zhcosin@163.com>
2017-05-25 于成都华阳

目录

第一章　极限

1.1　实数理论

分析学的基础建立在实数的公理化体系之上，在讨论极限理论之前，先来讨论一下实数的理论。

1.1.1　实数的十进制表示与大小关系

在人类历史上，为了计数而引进了自然数，最初以算筹的数量代表对应的数字，但这对于较大的数比较困难，为了表示数 100 就需要 100 根算筹，于是发明了十进制，这样所需的算筹数量就大大减少，之所以是十进制很可能是因为人正好有十根手指头，便于比划数字。后来为了解决多人平分食物等生活资料的问题又引进了整数之比即有理数的概念，再往后毕达哥拉斯学派根据勾股定理，发现了边长为 1 的正方形的对角线的长度不是有理数，引发了第一次数学危机，这次危机随着无理数的引入而得以解决。有理数与无理数一起，构成了全体实数。但在实数范围内，像 $x^2+1=0$ 这样的代数方程没有解，为了从理论上解决这个问题而引入了虚数的概念，实数与虚数一起构成了复数，代数方程的理论在复数范围内得到彻底的解决。

本节只讨论实数。在十进制下，一个实数 x 具有如下表示:

$$x = a_n a_{n-1} \cdots a_1 a_0 . a_{-1} a_{-2} \cdots \tag{1.1.1}$$

其中 $a_i \in \{0,1,2,3,4,5,6,7,8,9\}$，并且最左边的数位 a_n 非零（否则省略这一位不写），十进制就是说，这个式子表示的数值其实是

$$x = 10^n a_n + 10^{n-1} a_{n-1} + \cdots + 10 a_1 + a_0 + \frac{a_{-1}}{10} + \frac{a_{-2}}{10^2} + \cdots$$

即式 1.1.1实际表示的数值是它每一位数字与该数位上的权值之积的和，这一点是十分重要的，因为这样我们就只需要 $0-9$ 这十个数符就可以表示出任意实数，而不必为每一个数都去发明一个对应的数符，那样既是不可能的，也是很难使用的。

在这种表示下，数位 a_0 称为个位，a_1 称为十位，a_2 称为百位，依次类推，在 a_0 以后的部分称为小数部分，a_0 以及 a_0 左边的部分称为整数部分，两部分之间用小数点来分隔出明确的界限。

需要说明的是，实数十进制表示的小数部分是可以无限延伸的，但整数部分只能是有限位，并且规定，如果小数部分从某一位起全部都是零，则可以省写这些零，这样的小数称为有限小数，否则便称为无限小数。

如果无限小数的小数部分有连续重复出现的片段，例如 $0.12345678678678678\cdots$，这以后的数位全是重复的片段 678，就称这小数为循环小数，并简写为 $0.12345\dot{6}7\dot{8}$，即在循环片段的首尾两个数字上加点。如果没有这样的连续重复出现片段，则称为无限不循环小数。

关于整数的一个极为深刻的结论是

定理 1.1.1 (带余除法). *对任意两个整数 a 和 b，其中 b 为正整数，则存在唯一一对整数 q 与 $r(0 \leqslant r < b)$，使得 $a = qb + r$ 成立. 这整数 q 及 r 分别称为 a 除以 b 所得的商和余数.*

证明 以 b 的倍数为界点将全体实数划分为区间序列 $\ldots, [-2b, -b), [-b, 0), [0, b), [b, 2b), \ldots$，这些左闭右开区间两两无交集，且它们的并集就是全体实数，那么整数 a 必定从属于其中某一个区间，假定是 $[mb, (m+1)b)$，则取 $q = m, r = a - mb$ 即满足定理条件，反过来，如果还有另一组 q_1 及 r_1 满足定理中条件，那么有 $q_1 b \leqslant a < (q_1+1)b$，这即表明 $q_1 = m$，从而 $r_1 = a_{mb}$，这就证得了商及余数的唯一性。 □

利用带余除法，可以证明

定理 1.1.2. *有理数都是有限小数或者无限循环小数.*

证明 设有理数 $\frac{a}{b}$，其中 a 与 b 是整数，由于这结论与数的符号无关，所以假定这分子分母还是正的。这个证明过程其实就是两个正整数做除法的过程，思路就是在这个除法过程中，每一步所得的余数，或者是零从而被除尽，或者便要重复出现.

先用 a 除以 b，记商与余数分别为 q 及 r，即 $a = qb + r(0 \leqslant r < b)$，如果 $r > 0$，再用 $10r$ 除以 b，所得的商与余数分别记为 q_1 与 r_1，如果仍然有 $r_1 > 0$，则再将 $10r_1$ 除以 b 得到商 q_2 与余数 r_2，依次类推，得到序列 q_i 与 r_i，这时有 $q_i(i \leqslant 1)$ 只能取 0 到 9 中的数字，这是因为 $10r_{i-1} = q_i b + r_i$，而 $0 \leqslant r_{i-1} < b$，所以 q_i 不能超过 9，而由于 $0 \leqslant r_i < b$，所以 r_i 也只能在集合 $\{0, 1, 2, \ldots, b-1\}$ 这个有限集中取值，如果某一次取到了零 $r_m = 0$，则这个除法过程就结束了，而最终有

$$\frac{a}{b} = q + \sum_{i=0}^{m-1} \frac{q_i}{10^i} = q.q_1 q_2 \cdots q_{m-1}$$

即为有限小数。如果 r_i 始终不能取到零，那么必然存在某个 i 及 $j(> i)$ 使得 $r_i = r_j$，既然出现了相同的余数，那么在分别用 $10r_i$ 和 $10r_j$ 去除以 b 时也会得出相同的商 q_{i+1} 和 q_{j+1}，于是进一步出现相同的 r_{i+1} 与 r_{j+1}，这个过程将无限重复下去，这时就有

$$\frac{a}{b} = q.q_1 q_2 \cdots q_i q_{i+1} \cdots q_j q_{j+1} \cdots$$

这里从 q_i 到 q_{j-1} 便是一个重复片段，为小数的循环部分（不一定是最小循环片段），即为无限循环小数。 □

反之，有限小数与无限循环小数也都是有理数，但无限循环小数是有理数的证明涉及无限个数相加的和，这里暂不讨论。

我们规定，循环部分为一个单 9 的实数，等于将后面的 9 全部收上来的有限小数，即 $1 = 0.\dot{9}$，等等。

接下来讨论实数的大小问题，我们首先规定 0-9 这 10 个数符的大小关系，即 0 小于 1，1 小于 2，如此递推，直到 8 小于 9，在此基础上，我们定义实数的大小关系:

定义 1.1.1. *对于两个实数 $a = a_n a_{n-1} \cdots a_1 a_0.a_{-1}a_{-2}\cdots$ 与 $b = b_n b_{n-1} \cdots b_1 b_0.b_{-1}b_{-2}\cdots$ (如果它俩最高位不是同一数位，可以将最高位权重较低的那个前面补零)，如果存在某个整数 $m(\leqslant n)$，使得 $a_m < b_m$，并且对于所有大于 m 的整数 i 都有 $a_i = b_i$，则称实数 a 小于实数 b，记作 $a < b$，这时也称实数 b 大于实数 a，记作 $b > a$.*

这定义就是说，实数 a 小于实数 b 的充分必要条件是，从左边开始，第一个数符不同的数位上，a 在该数位上的数符小于 b 在该数位上的数符。

1.1.2 最小自然数原理

原理 1.1.1 (最小自然数原理). *任意非空的自然数集合中，必定存在一个最小的自然数。*

这个原理虽然看起来显而易见，但它是实数公理化体系的一部分。

1.1.3 确界定理

对于一个实数集，如果存在实数 M，使得集合中的全部数 x 都满足 $x \leqslant M$，则称实数 M 是这数集的一个上界，如果不等式是反向的，则称这实数是这数集的一个下界，显然，如果 M 是某个数集的上界，则比 M 大的所有实数也都是这数集的上界，对下界亦有类似结论。

如果数集既有上界又有下界，则称数集有界，有界数集的所有项的数值能够被某个区间所全部包含。

有界的另一种表述是，存在正实数 $M > 0$，使得数集的全部数 x 都满足 $|x| \leqslant M$，这与前述说法是等价的。

定义 1.1.2. *对于一个有上界的实数集，如果某个实数 M 满足: (1) 它是这数集的上界. (2) 对于无论多么小的正实数 ϵ，总存在数集中的数 x 使得 $x > M - \epsilon$，则称实数 M 是这数集的上确界，类似的有下确界的定义.*

显然，上确界是最小的上界，下确界是最大的下界。

定理 1.1.3 (确界定理). *若实数集合 (无论有限集无限集) 有上界，则有上确界，下界亦有相应结论。*

证明 设实数集合 A 有上界 M，我们先构造出一个数 K，再证明构造出的这个数正是这集合的上确界。

根据最小数原理，集合 A 中元素的整数部分有最大值，令 K 的整数部分与之相同，这整数部分记作 K_0。

再将集合 A 中所有元素乘以 10 后舍去小数部分，这些新数组成的新集合记作 A_1，这集合有上界 $10M$，因此按最小数原理，它也有最大值，而且这最大值除个位以外的

部分正是 K_0(按 K_0 的定义)，取这最大值的个位作为 K 的十分位。K 的其余数位依次类推，K 在 10^{-n} 上的数位是将集合 A 中全体元素乘以 10^n 后舍去小数部分所得新集合中最大数的个位数。

现在证明，数 K 是集合 A 的上确界，先证明它是上界，反证法，若它不是上界，则 A 中存在比它更大的数 x_0，那么按实数大小关系定义，在比较 x_0 与 K 时，从左边开始往右比较，第一个不相同的数位上，x_0 在该数位上的数大于 K 在该数位上的数，但这与 K 在这一数位上的数值的确定方法相矛盾，所以 K 是上界。其次需要证明，K 是最小的上界，设 L 是一个小于 K 的实数，那么它与 K 相比，从左边开始第一个不相同的数位上，它对应的数较小，假定这数位就是 10^{-n}，并设 K 和 L 在舍去这一数位以后的全部数位后所得的数分别是 K_n 和 L_n，那么 $K_n > L_n$，但根据 K 的确定过程可知，对于任何正整数 n，A 中都存在不小于 K_n 的数，自然这数也就大于 L_n，因此 K 是最小的上界，即为上确界。 □

1.1.4 实数指数幂

在中学数学里，我们已经接触过指数运算了，给出了指数为有理数的幂的定义，还引入了定义在 $\mathbb{R}$ 上的指数函数，但实际上，那时并没有给出无理指数幂的定义，这是因为，无理指数幂的定义，有赖于对实数理论的进一步讨论。在这一小节，我们利用确界定理来给出无理指数幂的定义，从而形成完整的指数运算。

先回顾一下有理指数幂的定义，设 a 是一个不等于 1 的正实数，首先它的正整数幂是被定义为连乘积:

$$a^n = a \cdot a \cdots a (n \in \mathbb{N}+)$$

这里是 n 个 a 相乘，然后把这定义推广到整数集上，规定 $a^0 = 1$，并且有如下的负整数指数幂定义

$$a^{-n} = \frac{1}{a^n}$$

于是指数运算被推广到指数是全体整数的情形了，再继续推广到有理数，设 $r = \dfrac{m}{n}$ 是一个有理数，这里 m、n 互素且 $n > 0$，则定义

$$a^{\frac{m}{n}} = \sqrt[n]{a^m}$$

这便是中学数学中对指数运算的定义。

不难得出，在此种定义下，有如下定理

定理 1.1.4. *设 $a > 0$ 且 $a \neq 1$，定义在 $\mathbb{Q}$ 上的指数函数 a^r 有如下基本性质*

1. $a^r > 0$
2. $a^{r_1+r_2} = a^{r_1} \cdot a^{r_2}$，$a^{r_1-r_2} = \frac{a^{r_1}}{a^{r_2}}$
3. *若 $a > 1$，则 a^r 随 r 严格递增，若 $0 < a < 1$，则是严格递减.*
4. *(连续性) 固定 $r_0 \in \mathbb{Q}$，对任意 $\varepsilon > 0$，都存在 $\delta > 0$，使得只要 $|r - r_0| < \delta$，就有 $|a^r - a^{r_0}| < \varepsilon$.*

前三条都容易，这里详细证明一下第四条，即连续性，在此只证明 $a>1$ 的情况，因为 $0<a<1$ 的情况也是完全类似的。证明 $r_0=0$ 的情形，即先证明定义在 $\mathbb{Q}$ 上的指数函数在 0 处是连续的，为此我们先建立一个不等式

引理 1.1.1. 设 $a>1$ 且 n 是一个正整数，则有不等式

$$0<\sqrt[n]{a}-1<\frac{a-1}{n}$$

证明一 由公式

$$s^n-t^n=(s-t)(s^{n-1}+s^{n-2}t+\cdots+st^{n-2}+t^{n-1})$$

令 $s=\sqrt[n]{a},t=1$，可得

$$a-1=(\sqrt[n]{a}-1)(s^{n-1}+s^{n-2}+\cdots+s+1)$$

显然 $s>1$，因此有

$$a-1>(\sqrt[n]{a}-1)\cdot n$$

由此即得证. □

证明二 设 $b=\sqrt[n]{a}-1>0$，则

$$a=(1+b)^n>1+nb$$

即得

$$b<\frac{a-1}{n}$$

□

借助这个引理，我们来证明定义在 $\mathbb{Q}$ 上的指数函数在 0 处是连续的，即要证明，对任意 $\varepsilon>0$，存在 $\delta>0$，使得对一切满足 $|r|<\delta$ 的有理数 r 都成立 $|a^r-1|<\varepsilon$。

为了寻找这个 δ，先考虑形式为 $\dfrac{1}{n}$ 的有理数，首先由刚证明的引理，要使 $|a^{\frac{1}{n}}-1|<\varepsilon$，只要使 $\dfrac{a-1}{n}<\varepsilon$ 就行了，这只需 $n>\frac{a-1}{\varepsilon}$，我们取一个 $n_0=\left[\dfrac{a-1}{\varepsilon}\right]+1$，再令 $\delta=\dfrac{1}{n_0}$，再根据第二条的单调性，只要有理数 r 满足 $0<r<\dfrac{1}{n_0}$，就有 $0<a^r-1<\varepsilon$，但还需要寻找左半区间的 δ，这时同样考虑形式为 $-\dfrac{1}{n}$ 的有理数，要使 $|a^{-\frac{1}{n}}-1|<\varepsilon$，因为

$$|a^{-\frac{1}{n}}-1|=\frac{|a^{\frac{1}{n}}-1|}{a^{\frac{1}{n}}}<|a^{\frac{1}{n}}-1|<\frac{a-1}{n}$$

所以仍然取 $\delta=\dfrac{1}{n_0}$，就能保证对任意满足 $-\delta<r<0$ 的有理数 r，有 $-\varepsilon<a^r-1<0$，这样就证明了，定义在 $\mathbb{Q}$ 上的指数函数在 0 处是连续的。

再考虑在任意有理数 r_0 处的连续性，对任意 $\varepsilon>0$，要寻找一个 $\delta>0$，使得对一切满足 $|r-r_0|<\delta$ 的有理数 r 都有 $|a^r-a^{r_0}|<\varepsilon$. 因为

$$|a^r-a^{r_0}|=a^{r_0}|a^{r-r_0}-1|$$

根据函数在 0 处的连续性，对于 $\varepsilon_1 = \frac{\varepsilon}{a^{r_0}} > 0$，存在某个 $delta > 0$，使得对一切满足 $|r - r_0| < \delta$ 的有理数 r 都成立 $|a^{r-r_0} - 1| < \varepsilon_1$，即 $|a^r - a^{r_0}| < \varepsilon$，即得证。

在此之上，还可以得出一些其它有用的性质，例如:

推论 1.1.1. *若 $a > 1$，则当 $r > 0$ 时有 $a^r > 1$，在 $r < 0$ 时 $0 < a^r < 1$，同样，若 $0 < a < 1$，则当 $r > 0$ 时有 $0 < a^r < 1$，在 $r < 0$ 时 $a^r > 1$.*

现在开始引入无理指数幂的概念，设实数 $a > 1(0 < a < 1$ 的情况也是类似的)，x 是一个无理数，而 r 和 s 是两个有理数，且 $r < x < s$，则显然有

$$a^r < a^s$$

分别作集合

$$A = \{a^r | r < x, r \in \mathbb{Q}\},\ B = \{a^s | s > x, s \in \mathbb{Q}\}$$

则 $\forall t_a \in A, \forall t_b \in B$，有 $t_a < t_b$，于是 A 有上界，B 有下界，于是 A 有上确界 L，B 有下确界 U，显然必有 $L \leqslant U$，实际上必须 $L = U$，不然的话，设 $L < U$，由确界定义，分别存在有理数 r_1 和 s_1，使得 $r_1 < x < s_1$，且 $\frac{L}{U}L < a^{r_1} < L < U < a^{s_1} < \frac{U}{L}U$，从而

$$L < \sqrt{a^{r_1} a^{s_1}} < U$$

即

$$L < a^{\frac{r_1 + s_1}{2}} < U$$

但显然 $\dfrac{r_1 + s_1}{2}$ 也是有理数，如果它小于 x，则它的函数值应在集合 A 中，从而不应当超过 L，同理，如果它大于 x，则它的函数值不应当小于 U，而它也更不可能等于 x，所以 $L < U$ 是不可能的，从而只能 $L = U$，即 A 的上确界即为 B 的下确界，记此共同的确界为 K，那么，K 是否在集合 A、B 之中呢？答案是不可能，这是因为，假如 $K \in A$，则等于就是说，对于一个无理数 x，存在一个满足 $r_0 < x$ 的有理数 r_0，使得所有满足 $r < x$ 的有理数 r 都有 $a^r \leqslant a^{r_0}$，这显然是不可能的，因为在 r_0 与 x 之间显然还存在着无穷多有理数，这些有理数的函数值都比 a^{r_0} 要大。同样，K 也不可能属于 B，也就是说，K 不可能是 a 的某个有理数次幂，很自然的，我们就把这个 K 定义为 a 的 x 次幂，即 $a^x = K$，这就是无理指数幂的概念。

定义 1.1.3. *设 $a > 0$ 且 $a \neq 1$，x 是一个无理数，定义集合 $A = \{a^r | r < x, r \in \mathbb{Q}\}$ 与 $B = \{a^s | s > x, s \in \mathbb{Q}\}$，则这两个集合中，一个有上确界，另一个有下确界，且二者相等，这个共同的确界，就称为 a 的 x 次幂，即为 a^x.*

有了无理指数幂，指数函数的定义域就可以拓广到全体实数上了，而且它还保持着在有理数集上所具有的良好性质，即:

定理 1.1.5. *设 $a > 0$ 且 $a \neq 1$，定义在 $\mathbb{R}$ 上的指数函数 a^x 仍然有如下性质*

1. $a^x > 0$

2. $a^{x_1 + x_2} = a^{x_1} \cdot a^{x_2}$，$a^{x_1 - x_2} = \frac{a^{x_1}}{a^{x_2}}$

3. 若 $a>1$，则 a^x 随 x 严格递增，若 $0<a<1$，则是严格递减.

4. （连续性）固定 $x_0\in\mathbb{R}$，对任意 $\varepsilon>0$，都存在 $\delta>0$，使得只要 $|x-x_0|<\delta$，就有 $|a^x-a^{x_0}|<\varepsilon$.

这就是说，定义在 $\mathbb{R}$ 上的指数函数，仍然保持着在有理数集上的运算法则、单调性和连续性，这些结论的证明并不难，只要充分利用定义就可以了，此处略去。

1.1.5 伯努利不等式

定理 1.1.6 (伯努利 (Bernoulli) 不等式)**.** 设实数 $x\geqslant -1$，则对任意正整数 n 成立不等式

$$(1+x)^n\geqslant 1+nx \tag{1.1.2}$$

等号成立的充分必要条件是 $x=0$ 或者 $x=-1,n=1$.

证明 在 $x\geqslant 0$ 时，左边按二项式定理展开可以看到它是成立的，所以关键是如何证明 $x<0$ 的情形。

在等式 (它可以经由数学归纳法得出)

$$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+\cdots+ab^{n-2}+b^{n-1})$$

在式中令 $a=1+x$，$b=1$ 便得

$$(1+x)^n-1=x\left[(1+x)^{n-1}+(1+x)^{n-2}+\cdots+(1+x)+1\right]$$

如果 $x>0$，则右边中括号内的部分显然大于 n，于是要证的不等式成立，而如果 $-1<x<0$，则中括号内的部分小于 n，再与负数 x 相乘，不等式反向即证出伯努利不等式。 □

1.1.6 复数

数系的扩充与解方程密切相关，为了解一元一次方程引入了有理数，数系被扩充为有理数集，为了解二次方程又发现了无理数，由此我们的数系被扩充到了实数集。但在实数范围内，方程 $x^2+1=0$ 仍然无解，于是引入了虚数，虚数与实数一起构成了更加广阔的复数集。

复数集的引入，理由虽然比较牵强，但后来的事实表明，很多数学理论在复数范围内都能得到完美的解决，例如代数学基本定理就表明，任何一个关于某未知数的 n 次方程，在复数范围内都有且仅有 n 个根 (重根按重数计数)，又例如，有些数列，它的每一项都是整数，然而它的通项，却只能借助复数来表达，又比如，在复数范围内，指数函数将与三角函数产生密切关系，欧拉公式 $e^{ix}=\cos x+i\sin x$ 便揭示了这一点，这在实数范围内是很难看清二者的关系的。诸如此类，许多理论都表明复数具有重大的理论意义，这一点随着我们对高等数学的更多了解，将能有更多体会。

为了解方程 $x^2+1=0$，引入一个虚数 i，规定 $i^2=-1$，因为实数范围内是不可能有某个数的平方是负的，所以这引入的 i 便称为一个虚数，并且称它是虚数单位。规定虚数可以与实数相加与相乘，并且符合实数运算所满足的交换律、结合律、分配律。

那么将实数与虚数进行混合加法与乘法运算，会有什么结果呢，将 i 与实数 b 相乘得出 bi，再将 bi 与实数 a 相加得到 $a+bi$，这个形式无法继续化简了，它就是复数的一般形式，也就是任意一个复数都具有这种形式，我们就来证明它。

定理 1.1.7. *将实数与虚数单位 i 进行有限次加法与乘法的混合运算，得出的结果都具有形式 $a+bi(a,b\in R)$.*

证明 只利用加法与乘法，运算对象为实数与虚数单位 i，那么最终结果是关于 i 的实系数多项式

$$a_ni^n+a_{n-1}i^{n-1}+\cdots+a_2i^2+a_1i+a_0$$

根据虚数单位 i 的定义，可知其乘幂 i^n 依次循环取值 $i,-1,-i,1$，由此便知定理成立。 □

由此，复数都具有形式 $z=a+bi(a,b\in R)$，a 称为它的实部，记作 $Re(z)$，b 称为它的虚部，记作 $Im(z)$，显然复数 $a+bi$ 与数对 (a,b) 一一对应，于是便与坐标平面上的点一一对应，于是坐标平面上的每一个点都对应着一个复数，于是这平面便被称为*复平面*.

规定，两个复数相等当且仅当它们的实部和虚部分别相等. 此外，复数集上没有大小关系。

称复数 $a-bi$ 是复数 $z=a+bi$ 的*共轭复数*，记作 $\overline{z}$，即共轭的两个复数实部相等而虚部互为相反数，显然

$$z+\overline{z}=2Re(z),\ z-\overline{z}=2iIm(z),\ z\overline{z}=|z|^2$$

再定义减法为加法的逆运算，那么按照实数的运算定律，有

$$(a_1+b_1i)\pm(a_2+b_2i)=(a_1\pm a_2)+(b_1\pm b_2)i$$

这显然也可以推广到任意有限个复数相加的情形，显然，复数的加减法对应着数对的加减法，也就对应着向量的加减法。

不难验证，关于共轭复数有

$$\overline{z_1+z_2}=\overline{z_1}+\overline{z_2}$$

关于乘法，有

$$\begin{aligned}(a_1+b_1i)(a_2+b_2i) &= a_1a_2+(a_1b_2+a_2b_1)i+b_1b_2i^2\\ &= (a_1a_2-b_1b_2)+(a_1b_2+a_2b_1)i\end{aligned}$$

同样规定除法为乘法的逆运算，有

$$\begin{aligned}\frac{a_1+b_1i}{a_2+b_2i} &= \frac{(a_1+b_1i)(a_2-b_2i)}{(a_2+b_2i)(a_2-b_2i)}\\ &= \frac{(a_1a_2+b_1b_2)+(b_1a_2-a_1b_2)i}{a_2^2+b_2^2}\end{aligned}$$

同样可以验证

$$\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2},\ \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\overline{z_1}}{\overline{z_2}}$$

例 1.1.1 从数对引入复数 我们从另一角度引入复数，实数对应着数轴上的点，属于一维数，我们认为经过推广后的复数为二维数，它与坐标平面上的点 (a,b) 一一对应，即复数 z 就是一个数对 (a,b)，当 $b=0$ 时，它就是实数，即复数 $(a,0)$ 就是实数 a，现在定义加法如下: 若 $z_1=(a_1,b_1)$，$z_2=(a_2,b_2)$，则

$$z_1+z_2=(a_1+a_2,b_1+b_2)$$

再定义复数的乘法是

$$z_1z_2=(a_1a_2-b_1b_2,a_1b_2+a_2b_1)$$

在这定义下，显然加法和乘法都满足交换律，且乘法对于加法的分配律也是容易验证的，此外，容易验证

$$(a,b)(1,0)=(a,b)$$

即复数 $(1,0)$(即实数 1) 在复数范围内仍然是乘法的单位元。还可以验证

$$(a,b)=(a,0)(1,0)+(b,0)(0,1)$$

即任意实数都经由两个复数单位 $(1,0)$ 与 $(0,1)$ 的实系数线性组合来表出，于是记纵轴上的单位复数 $(0,1)$ 为 i，则上式可以简写为 $(a,b)=a+bi$，这样，我们从数对出发，通过引入加法和乘法的定义也引出了复数的概念。 ■

利用变换 $x=r\cos\theta$，$y=r\sin\theta$，复数 $z=a+bi$ 可以改写为

$$z=r(\cos\theta+i\sin\theta)$$

这称为复数的三角形式，其中 $r=\sqrt{x^2+y^2}$ 称为复数的模，记作 $|z|$，角 θ 称为这复数的辐角，记作 $Arg(z)$，由于三角函数的周期性，将满足 $0\leqslant\theta<2\pi$ 的那个辐角称为辐角主值，记作 $arg(z)$.

显然，如果两个复数相等，它们的模相等，并且它们辐角集合相等。

我们看一下在这种形式下复数的乘除法运算:

$$\begin{aligned}
& r_1(\cos\theta_1+i\sin\theta_1)\cdot r_2(\cos\theta_2+i\sin\theta_2) \\
= & r_1r_2[(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)+(\cos\theta_1\sin\theta_2+\cos\theta_2\sin\theta_1)i] \\
= & r_1r_2(\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2))
\end{aligned}$$

而

$$\begin{aligned}
& \frac{r_1(\cos\theta_1+i\sin\theta_1)}{r_2(\cos\theta_2+i\sin\theta_2)} \\
= & \frac{r_1(\cos\theta_1+i\sin\theta_1)\cdot r_2(\cos\theta_2-i\sin\theta_2)}{r_2(\cos\theta_2+i\sin\theta_2)\cdot r_2(\cos\theta_2-i\sin\theta_2)} \\
= & \frac{r_1r_2[(\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2)+i(\sin\theta_1\cos\theta_2-\cos\theta_1\sin\theta_2)]]}{r_2^2(\cos^2\theta_2+\sin^2\theta_2)} \\
= & \frac{r_1}{r_2}[\cos(\theta_1-\theta_2)+i\sin(\theta_1-\theta_2)]
\end{aligned}$$

可见两个复数相乘除，就是将两个复数的模相乘除得到乘积或商的模，两个复数的辐角相加减得到乘积或商的辐角，这与向量的乘法 (无论内积还是外积) 不再一致，复数乘法在三角形式下变得相当简单，而且这显然可以推广到任意有限个复数相乘的情形。

例 1.1.2 利用共轭复数证明余弦定理 取复数 $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$，$z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$，则有

$$z_1\overline{z_2} = r_1 r_2(\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2))$$

因此，$z_1\overline{z_2}$ 的实数部分就是 z_1 与 z_2 两个复数对应的两个向量的内积，同理 $\overline{z_1}z_2$ 的实数部分也是这内积，所以这内积等于

$$\frac{1}{2}(z_1\overline{z_2} + \overline{z_1}z_2)$$

由恒等式

$$(z_1 - z_2)(\overline{z_1} - \overline{z_2}) = z_1\overline{z_1} + z_2\overline{z_2} - (z_1\overline{z_2} + \overline{z_1}z_2)$$

左边就是 $|z_1-z_2|^2$，右边前两项分别是 $|z_1|^2$ 和 $|z_2|^2$，最后的两项就是 z_1 与 z_2 对应两个向量的内积的 2 倍，于是便得出余弦定理

$$|z_1-z_2|^2 = |z_1|^2 + |z_2|^2 - 2|z_1||z_2|\cos(\theta_1-\theta_2)$$

■

更特别的是复数的乘幂，容易知道

$$[r(\cos\theta + i\sin\theta)]^n = r^n(\cos n\theta + i\sin n\theta) \tag{1.1.3}$$

这就是复数乘幂的*棣莫弗公式*.

例 1.1.3 正余弦的 n 倍角公式与切比雪夫多项式 利用棣莫弗公式，我们可以得到正余弦的 n 倍角公式，在棣莫弗公式中令 $r=1$，得

$$\cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n$$

将右边利用二项式定理展开，得

$$\cos n\theta + i\sin n\theta = \sum_{k=0}^{n} C_n^k i^k \cos^{n-k}\theta\sin^k\theta$$

当 k 为偶数时，求和中的通项变为实数，当 k 为奇数时，则它为虚数，据此可以将上式右端的实数和虚部分开

$$\begin{aligned}&\cos n\theta + i\sin n\theta\\ =&\sum_{0\leqslant 2r\leqslant n}^{n}(-1)^r C_n^{2r}\cos^{n-2r}\theta\sin^{2r}\theta + i\sum_{0\leqslant 2r+1\leqslant n}(-1)^r C_n^{2r+1}\cos^{n-2r-1}\theta\sin^{2r+1}\theta\end{aligned}$$

于是得到

$$\cos n\theta = \sum_{0\leqslant 2r\leqslant n}^{n}(-1)^r C_n^{2r}\cos^{n-2r}\theta\sin^{2r}\theta \tag{1.1.4}$$

$$\sin n\theta = \sum_{0\leqslant 2r+1\leqslant n}(-1)^r C_n^{2r+1}\cos^{n-2r-1}\theta\sin^{2r+1}\theta \tag{1.1.5}$$

这就是余弦和正弦的 n 倍角公式.

根据这公式，由于 $\cos n\theta$ 的每一项中的正弦的指数都是偶数，所以都可以化为余弦，于是 $\cos n\theta$ 可以展开为 $\cos\theta$ 的 n 次多项式，这就是第一类切比雪夫多项式，即

$$T_n(x)=\sum_{0\leqslant 2r\leqslant n}(-1)^r C_n^{2r}x^{n-2r}(1-x^2)^r$$

在 $\sin n\theta$ 的展式中，$\sin\theta$ 的次数都是奇数，所以 $\dfrac{\sin(n+1)\theta}{\sin\theta}$ 也可以展开为 $\cos\theta$ 的 n 次多项式，这就是第二类切比雪夫多项式，即

$$U_n(x)=\sum_{0\leqslant 2r+1\leqslant n+1}(-1)^r C_{n+1}^{2r+1}x^{n-2r}(1-x^2)^r$$

关于切比雪夫多项式的更多讨论参见 [6]. ■

例 1.1.4 现在来求和下面两个表达式

$$\begin{aligned}A_n &= 1+r\cos\theta+r^2\cos 2\theta+\cdots+r^n\cos n\theta\\ B_n &= r\sin\theta+r^2\sin 2\theta+\cdots+r^n\sin n\theta\end{aligned}$$

令 $z=r(\cos\theta+i\sin\theta)$，则

$$\begin{aligned}&A_n+iB_n\\ =\;& 1+r(\cos\theta+i\sin\theta)+r^2(\cos 2\theta+i\sin 2\theta)+\cdots+r^n(\cos n\theta+i\sin n\theta)\\ =\;& 1+z+z^2+\cdots+z^n\\ =\;& \frac{1-z^{n+1}}{1-z}\\ =\;& \frac{1-r^{n+1}(\cos(n+1)\theta+i\sin(n+1)\theta)}{1-r(\cos\theta+i\sin\theta)}\\ =\;& \frac{1-r^{n+1}(\cos(n+1)\theta+i\sin(n+1)\theta)}{1-r(\cos\theta+i\sin\theta)}\cdot\frac{1-r\cos\theta+ir\sin\theta}{1-r\cos\theta+ir\sin\theta}\\ =\;& \frac{1-r\cos\theta+r^{n+2}\cos n\theta-r^{n+1}\cos(n+1)\theta}{1-2r\cos\theta+r^2}+\\ & i\frac{r\sin\theta+r^{n+2}\sin n\theta-r^{n+1}\sin(n+1)\theta}{1-2r\cos\theta+r^2}\end{aligned}$$

于是比较实部和虚部可得

$$\begin{aligned}A_n &= \frac{1-r\cos\theta+r^{n+2}\cos n\theta-r^{n+1}\cos(n+1)\theta}{1-2r\cos\theta+r^2}\\ B_n &= \frac{r\sin\theta+r^{n+2}\sin n\theta-r^{n+1}\sin(n+1)\theta}{1-2r\cos\theta+r^2}\end{aligned}$$

■

有了棣莫弗公式，我们来讨论一下复数的开方。

记 $z=r(\cos\theta+i\sin\theta)$，现在来求它的 n 次方根，设 $z'=r'(cos\theta'+i\sin\theta')$ 是它的一个 n 次方根，按棣莫弗公式，应有

$$z'^n=r'^n(\cos n\theta'+i\sin n\theta')$$

因为 $z'^n = z$，所以有 $r'^n = r$，以及 $n\theta' = \theta + 2m\pi(m \in Z)$，即

$$r' = \sqrt[n]{r},\ \theta' = \frac{\theta + 2m\pi}{n}(m = 0, 1, \ldots, n-1)$$

或者写成

$$\sqrt[n]{z} = \sqrt[n]{r}\left(\cos\frac{\theta + 2m\pi}{n} + i\sin\frac{\theta + 2m\pi}{n}\right),\ (m = 0, 1, \ldots, n-1)$$

根据周期性，可知 z 的 n 次方根正好有 n 个，它们均匀分布在复平面上以原点为圆心，以 $|z|$ 为半径的圆上，于是在复数范围内，任何数都可以开 n 次方.

特别的是当 $z = 1$ 时，对 1 进行开 n 次方，因为 $1 = \cos 0 + i\sin 0$，于是得它的根

$$\varepsilon_i = \cos\frac{2i\pi}{n} + i\sin\frac{2i\pi}{n},\ i = 0, 1, \ldots, n-1$$

这些根 $\varepsilon_i(i = 0, 1, \ldots, n-1)$ 称为*n 次单位根*，根据复数乘法可以知道

$$\varepsilon_{i+1} = \varepsilon_1\varepsilon_i$$

实际上更一般的结论是

$$\varepsilon_i\varepsilon_j = \varepsilon_{i+j}$$

1.2 数列极限

1.2.1 数列极限的概念

在中学数学中，我们熟知反比例函数 $f(x) = 1/x$ 的图象在向无穷远处延伸时，它会无限的向横轴靠近，当 x 取正值并无限的增大时，其第一象限的一支会无限的向 x 轴正半轴靠近，但无论 x 取多大，因为 $1/x > 0$，所以它始终不会与 x 轴相交，这给了我们一种“无限接近但是不会相等”的直观感受。

同样的情况还有许多，我们就准备来详细的讨论下这种“无限接近又不相等”的现象。

上面反比例函数的例子是针对函数而言的，我们先从较为简单的数列开始，同样可以得到数列 $x_n = 1/n$ 这个数列，在 n 取正整数并无限增大时，数列的值无限的接近零，但却总是大于零，我们来从这种现象中提取 *极限*的概念。

首先要指出的是这里“无限接近但不等于”中的“不等于”其实是无关紧要的，例如在数列 $1/n$ 中把下标为偶数的项全部换成零，那么这个”无限”接近并没有被破坏，而且它仍然给我们以极限的印象，只是它在下标增大的过程中，可以无限次的取极限值，而且从这个例子中还可得知，数列的单调性也不是必要的。

我们先给出一个初步的定义: 如果数列 x_n 在随着 n 的无限增大过程中可以无限的接近一个常数 A，则称 A 是这数列当 n 趋于无穷大时的极限。

这个定义不会令人满意，因为作为数学上的一个定义，它需要具备精确性，而这个定义中出现了“无限接近”这样含义模糊不清的描述，利用这个定义，我们很难说明一个给定常数是否是一个数列的极限，我们需要将它严格化。

所谓数列 x_n“无限接近”于常数 A，自然指的是差值 $|x_n - A|$ 可以任意的小，所以我们进行第一步严格化：把数列 x_n 无限接近常数 A 严格化成差值 $|x_n - A|$ 可以任意小，于是极限的定义可以重新叙述为: 对于数列 x_n 和常数 A，如果数列 x_n 当 n 无限增大时差值 $|x_n - A|$ 可以任意的小，则称常数 A 是数列 x_n 当 n 趋向于无穷大时的极限。

然后我们考虑如何刻画“可以任意的小”，那就是说，差值 $|x_n - A|$ 可以小于任意的正实数 ε，而不管这正实数 ε 有多小。初看起来，“可以小于任意的正实数”，似乎只要存在正整数 n，使得 $|x_n - A| < \varepsilon$ 就可以了，也就是如下的极限定义: 对于数列 x_n 和常数 A，如果对于无论多么小的正实数 ε，总存在正整数 N，使得 $|x_N - A| < \varepsilon$ 成立，则称常数 A 是数列 x_n 在下标趋于无穷大时的极限。

这个定义看上去似乎非常符合 $1/n$ 这个数列的特征，不管多么小的正实数 ε，总能找到使 $1/n < \varepsilon$ 成立的 n，只要 n 取的足够大。然而这个定义却有一个严重的问题，我们把数列 $1/n$ 中下标为偶数的项全部换成 1，所得到的新数列显然不应该有极限，因为它的奇数下标项趋于零而偶数下标项恒为 1，按我们的直观感受，它的值并不无限靠近零，也不无限靠近 1，0 和 1 都不应该是它的极限，但是按照上面的定义，它却是符合条件的！

问题出在哪呢，仍然以这个把偶数下标项都替换为 1 的新数列为例，显然，存在正整数 N 使得 $|x_n - A| < \varepsilon$ 这个条件，是无法保证数列的全部项都向常数 A 靠近的，它只能保证数列中有一部分项会向常数 A 靠近，刚才这个例子也说明了这一点，所以我们需要一个更强的能保证数列的所有项都要向常数 A 靠近，我们把存在正整数 N 使得 $|x_N - A| < \varepsilon$ 成立，改为存在正整数 N，使得 $n > N$ 时 $|x_n - A| < \varepsilon$ 恒成立，这样一来这个新数列就不满足这条件了，而原来的数列 $1/n$ 却满足这条件。

这个新的条件，利用 $n > N$ 时 $|x_n - A| < \varepsilon$ 恒成立，来保证了数列向 A 靠近的总体趋势。这就是我们最终的极限定义，这个定义，从模糊到精确，别看在这几段话就给出了，实际上在历史上经过了几代数学家的努力，最后才由德国数学家魏尔斯特拉斯 (Weierstrass, 1815.10.31-1897.2.19) 在总结前人成果的基础上给出，这个精确定义，是人类智慧的结晶。

定义 1.2.1 (数列极限). *对于实数数列 a_n 和实数 a，如果对于任意小的正实数 ε，都存在某一下标 N，使得该数列在这之后的所有项 (即 $n > N$) 都满足*

$$|a_n - a| < \varepsilon \tag{1.2.1}$$

则称该数列存在极限，实数 a 称为该数列的极限。也称该数列为收敛数列，并且收敛到实数 a，记为

$$\lim_{n\to\infty} x_n = a \tag{1.2.2}$$

极限为零的数列称为无穷小数列，简称*无穷小*。如果数列不存在有限的极限，称为数列为*发散数列*。

如果数列无论对于多大的实数 $M > 0$，总能从某项开始，后续的全部项都有 $a_n > M$，则称数列为*正无穷大*。类似的也有负无穷大和 (绝对值) 无穷大的概念。

我们引入邻域的概念:

定义 1.2.2 (邻域). 以实数 x 为中心半径为 ε 的开区间 $(x-\varepsilon, x+\varepsilon)$ 称为实数 x 的 $\varepsilon-$ 邻域，记作 $U(x,\varepsilon)$，若从中去掉实数 x，即 $(x-\varepsilon, x)\cup(x, x+\varepsilon)$ 称为 x 的 $\varepsilon-$ 空心邻域，记作 $\mathring{U}(x,\varepsilon)$，有时并不关心半径是多少，直接称为邻域和空心邻域，简记为 $U(x)$ 和 $\mathring{U}(x)$.

借助邻域概念，我们可以这样刻画数列的极限：数列收敛到 A，就是说，对于 A 的任意小的邻域 $U(A)$，下标超过某一值以后，都将全部落在该邻域中，从而只有有限项落在该邻域外部。

1.2.2 一些极限的例

例 1.2.1 前面在提炼极限定义时用了 $\dfrac{1}{n}$ 在 n 无限增大时的情形，现在利用极限定义来证明这个极限:

$$\lim_{n\to\infty}\frac{1}{n}=0$$

对于任意小的正实数 ε，要使 $\dfrac{1}{n}<\varepsilon$，只要 $n>\dfrac{1}{\varepsilon}$ 就可以了，于是可以取 $N=1+\left[\dfrac{1}{\varepsilon}\right]$ 就可以了，这就证明了它的极限是零，实际上这里的 N，只要比 $\dfrac{1}{\varepsilon}$ 大就都可以了，以后我们就不专门作取整处理了。

仿此还可以得到，当 p 是正有理数时，有

$$\lim_{n\to\infty}\frac{1}{n^p}=0$$

■

例 1.2.2 设实数 q 满足 $|q|<1$，则

$$\lim_{n\to\infty}q^n=0$$

对于任意小的正实数 ε，要使不等式 $|q^n-0|<\varepsilon$，只要 $n>\dfrac{\ln\varepsilon}{\ln|q|}$ 就行了，这就证明得了结论。 ■

例 1.2.3 设实数 $a>1$，则

$$\lim_{n\to\infty}\frac{1}{a^n}=0$$

这实际上是例 1.2.2的特例，但这里换一种方法来证明它.

对于无论多么小的正实数 ε，为了找到极限定义中所要求的 N，考虑不等式

$$\frac{1}{a^n}<\varepsilon$$

也就是 $a^n>1/\varepsilon$，设 $a=1+\lambda$，则 $\lambda>0$，按二项式定理有[1]

$$a^n=(1+\lambda)^n=1+n\lambda+\frac{n(n-1)}{2!}\lambda^2+\cdots+\lambda^n>1+n\lambda$$

[1]我们这里并没有从 $a^n>1/\varepsilon$ 中直接使用对数来得出 $n>\log_a(1/\varepsilon)$，这是因为尽管中学数学中已经学过对数概念，但那时还没有给出无理指数幂的定义，所以指数的定义是不完整的，因此我们无法确认，对于底数 a，正实数 $1/\varepsilon$ 的对数是否存在，以后我们将在小节 1.4.9中专门讨论指数的定义和值域问题。

所以只要 $1+n\lambda > 1/\varepsilon$，便能保证 $a^n > 1/\varepsilon$ 成立，也就是只需要 $n > (1/\varepsilon - 1)/(a-1)$ 就行了，所以只要选择 $N > (1/\varepsilon - 1)/(a-1)$ 就行了，这就证得了此极限。 ■

例 1.2.4 我们来建立一个基本的极限，对于大于 1 的正实数 a，有

$$\lim_{n\to\infty} \frac{n}{a^n} = 0$$

事实上，仍同例 1.2.3一样，记 $a = 1+\lambda(\lambda > 0)$，则

$$a^n = (1+\lambda)^n = \sum_{i=0}^{n} C_n^i \lambda^i > C_n^2 \lambda^2$$

因此有

$$\frac{n}{a^n} < \frac{n}{C_n^2\lambda^2} = \frac{2}{(n-1)\lambda^2}$$

在 $n > 2$ 时，又有 $n-1 > \dfrac{n}{2}$，所以此时更有

$$\frac{n}{a^n} < \frac{4}{n\lambda^2}$$

所以对于任意小的正实数 ε，只要 $n > \max\{2, \dfrac{4}{\lambda^2\varepsilon}\}$，便有 $\dfrac{n}{a^n} < \varepsilon$，这就证得所要的极限。 ■

例 1.2.5 在例 1.2.4的基础上，还可以建立下面的极限

$$\lim_{n\to\infty} \frac{\ln n}{n} = 0$$

对于任意小的正实数 ε，为了使得 n 充分大时 $\dfrac{\ln n}{n} < \varepsilon$，只要使 $n < (e^\varepsilon)^n$ 就可以了，而根据例 1.2.4中所建立起的极限，这是可以在 n 充分大时恒成立的，故此就建立了此处的极限。 ■

例 1.2.6 设实数 $a > 1$，则

$$\lim_{n\to\infty} \frac{a^n}{n!} = 0$$

将原式写成

$$a \cdot \frac{a}{2} \cdots \frac{a}{n}$$

任取正整数 M 满足 $M > a$，根据上式因式中分母与 M 的大小分成两部分 (限定 $n > M$)

$$\left(a \cdot \frac{a}{2} \cdots \frac{a}{M}\right)\left(\frac{a}{M+1} \cdot \frac{a}{M+2} \cdots \frac{a}{n}\right)$$

将后一个括号中的因式全部放大为 $\dfrac{a}{M}$，成为

$$\left(a \cdot \frac{a}{2} \cdots \frac{a}{M}\right)\left(\frac{a}{M}\right)^{n-M}$$

它具有形式

$$cq^n$$

其中 $q = \dfrac{a}{M}$ 满足 $|q| < 1$，而 c 是一个正常数，所以此式极限是零，而原式也就以零为极限。 ■

例 1.2.7 设实数 $a > 1$ 且 $a \neq 1$，则 $\lim_{n\to\infty} \sqrt[n]{a} = 1$

证明一 利用乘法公式 $x^n-1=(x-1)(x^{n-1}+x^{n-2}+\cdots+1)$ 可得

$$\sqrt[n]{a}-1=\frac{a-1}{(\sqrt[n]{a})^{n-1}+(\sqrt[n]{a})^{n-2}+\cdots+1}<\frac{1}{n}(a-1)$$

于是对于任意正实数 ε，只要取 $N>\frac{a-1}{\varepsilon}$ 便能保证 $n>N$ 时有 $0<\sqrt[n]{a}-1<\varepsilon$, 所以这极限得证。 □

证明二 设 $z_n=\sqrt[n]{a}-1$，则

$$a=(1+z_n)^n=1+nz_n+\frac{n(n-1)}{2!}z_n^2+\cdots+z_n^n>1+nz_n$$

所以得到

$$0<z_n<\frac{1}{n}(a-1)$$

下同证明一. □

■

例 1.2.8 我们来证明下面这个极限

$$\lim_{n\to\infty}\sqrt[n]{n}=1$$

显然在 $n>1$ 时有 $\sqrt[n]{n}>1$，只要证明当 n 大于某一正整数时恒有

$$\sqrt[n]{n}<1+\varepsilon$$

就可以了，这里 ε 是一个任意小的正实数，而这只需要

$$n<(1+\varepsilon)^n$$

就行了，根据例 1.2.4中所得到的极限，这显然是可以办到的。 ■

例 1.2.9 现在证明

$$\lim_{n\to\infty}\frac{1}{\sqrt[n]{n!}}=0$$

任意固定一个正整数 M，在 $n>M$ 时显然有

$$n!>M!M^{n-M}$$

即

$$\frac{1}{\sqrt[n]{n!}}<\frac{1}{M\sqrt[n]{\frac{M!}{M^M}}}$$

而 $\dfrac{M!}{M^M}<1$，根据例 1.2.7的结论，当 n 充分大时，将有

$$\sqrt[n]{\frac{M!}{M^M}}>\frac{1}{2}$$

从而此时有

$$\frac{1}{\sqrt[n]{n!}}<\frac{2}{M}$$

即对于任何确定的 M，当 n 充分大时都有上式成立，所以对于任意小的正实数 ε，只要让 $\dfrac{2}{M}<\varepsilon$ 就可以了，这就证明了结论。 ■

例 1.2.10 假定数列 a_n 以 A 为极限，我们来考察一下它的前 n 项的算术平均序列

$$M_n=\frac{x_1+x_2+\cdots+x_n}{n}$$

作为一个数列的收敛情况。

先可以直观的想象一下，数列以 A 为极限，那么数列的项随着下标的无限增大将与 A 无限接近，除前面的项外，越往后的连续片段的算术平均也应该与 A 无限接近，而是越着 n 的无限增大，数列最前面的与 A 相差的项对算术平均的影响也将越来越小，因此猜想 M_n 也收敛到 A.

事实上也的确如此，因为对于任意正实数 ε，存在正整数 N，使得 $n>N$ 时恒有 $|a_n-A|<\varepsilon$，即 $A-\varepsilon<a_n<A+\varepsilon$，于是有

$$(n-N)(A-\varepsilon)<a_{N+1}+a_{N+2}+\cdots+a_n<(n-N)(A+\varepsilon)$$

所以

$$\sum_{i=1}^{N}a_i+(n-N)(A-\varepsilon)<a_1+a_2+\cdots+a_n<\sum_{i=1}^{N}a_i+(n-N)(A+\varepsilon)$$

从而

$$A-\varepsilon+\frac{1}{n}\left(\sum_{i=1}^{N}a_i-N(A-\varepsilon)\right)<M_n<A+\varepsilon+\frac{1}{n}\left(\sum_{i=1}^{N}a_i-N(A+\varepsilon)\right)$$

因为 $\lim\limits_{n\to\infty}\dfrac{1}{n}=0$，所以对于前述正实数 ε，令

$$\varepsilon_1=\frac{\varepsilon}{\left|\sum_{i=1}^{N}a_i-N(A-\varepsilon)\right|}$$

和正实数

$$\varepsilon_2=\frac{\varepsilon}{\left|\sum_{i=1}^{N}a_i-N(A+\varepsilon)\right|}$$

再令 $\varepsilon_0=\max\{\varepsilon_1,\varepsilon_2\}$，则存在正整数 N_0，使得当 $n>N_0$ 时，$\dfrac{1}{n}<\varepsilon_0$，于是在 $n>N_1=\max\{N,N_0\}$ 时有

$$A-2\varepsilon<M_n<A+2\varepsilon$$

这便说明 $\lim\limits_{n\to\infty}M_n=A$，需要提醒的是这里出现的 2ε 似乎与极限定义不一致，但实际上，只要在这整个过程将 ε 替换为 $\dfrac{1}{2}\varepsilon$，就可以与极限定义完全一致，而且从另一个角度来说，极限定义中只是说 ε 是任意正实数，那么如果 ε 是任意正实数，2ε 照样可以取遍任意正实数，所以这并没有本质上的不同，关于这一点以后将不再说明了。 ■

例 1.2.11 无穷级数的和 对于数列 a_n，我们可以作出它的前 n 项的和

$$S_n=\sum_{i=1}^{n}a_i=a_1+a_2+\cdots+a_n$$

依赖于 n 的取值，这些和将构成一个新的数列 $S_1, S_2, \ldots$，现在我们把加法推广到无穷个数相加的情形，称无穷个数相加的表达式

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots$$

为无穷级数，而 S_n 称为它的第 n 个部分和，如果 S_n 有 (有限的) 极限，即

$$\lim_{n\to\infty} S_n = S$$

则称无穷级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛，这个极限 S 称为级数的和，记作

$$\sum_{n=1}^{\infty} a_n = S$$

如果部分和数列 S_n 没有极限，则称无穷级数是发散的。 ■

1.2.3 极限的性质与运算

定理 1.2.1 (极限唯一性). *如果数列 x_n 收敛，则极限唯一。*

证明 反证法，假若有两个数 $a1$ 和 $a2$ 都是数列的极限，假定 $a_1 < a_2$，则对于任意正数 $\varepsilon > 0$，数列都能从某项起同时成立着 $|x_n - a_1| < \varepsilon$ 和 $|x_n - a_2| < \varepsilon$，于是取 $\varepsilon < (a_2 - a_1)/2$，则前述两个不等式因为无交集而产生矛盾。 □

定理 1.2.2. *如果数列 x_n 收敛到实数 a，则对于任意一个小于 a 的实数 x，数列都能从某项起恒大于 x，同样对于任意一个大于 a 的实数 y，数列也能从某项起恒大于 y。*

证明 只要在极限的定义中取 $\varepsilon < a - x$ 即可得前半部分结论，同样再取 $\varepsilon < y - a$ 即得后半部分结论。 □

推论 1.2.1 (保号性). *如果数列收敛到一个正的实数，则数列必从某项起恒保持正号，同样，若收敛到一个负的实数，则必从某项起恒保持负号。*

定理 1.2.3 (收敛数列的有界性). *收敛数列必有界。*

证明 这其实从定义就可以得出了，随便取一个 $\varepsilon > 0$，即知数列从某项起全部落在区间 $(a - \varepsilon, a + \varepsilon)$ 内，这里 a 是数列极限，再扩大此区间把前面的那些项 (有限个) 包含进来，于是数列便有界。 □

定理 1.2.4 (保不等式性). *设数列 a_n 与数列 b_n 分别收敛到 A 与 B，且当 n 充分大时恒有 $a_n < b_n$，则必有 $A \leqslant B$.*

证明 反证法，若 $A > B$，则取 $\varepsilon = \dfrac{1}{2}(A - B)$，在 n 充分大时必同时有 $a_n > A - \varepsilon = \dfrac{1}{2}(A + B)$ 以及 $b_n < B + \varepsilon = \dfrac{1}{2}(A + B)$ 成立，这时显然有 $a_n > b_n$，与定理条件矛盾，所以 $A \leqslant B$. □

要注意的是由定理中的条件并不能得出 $A < B$ 的结论来，例如 $a_n = \dfrac{1}{n^2}$ 与 $b_n = \dfrac{1}{n}$.

定理 1.2.5 (夹逼准则，迫敛性). *若三个数列 a_n、b_n、c_n 在 n 充分大时恒有 $a_n \leqslant b_n \leqslant c_n$，并且 a_n 与 c_n 都收敛到同一极限 M，则 b_n 亦必收敛到此极限值.*

证明 对任意小的正实数 ε，显然当 n 充分大时有

$$M - \varepsilon < a_n < b_n < c_n < M + \varepsilon$$

由此即得定理. □

定理 1.2.6. *如果数列 x_n 和 y_n 分别收敛到 x 和 y，则数列 $x_n + y_n$、$x_n - y_n$、$x_n y_n$、x_n/y_n 都收敛，而且极限分别是 $x + y$、$x - y$、xy、x/y，在商的情况要求 $y \neq 0$。*

这定理可以推广到任意有限个数列的情形。

证明 和差的情况是容易证明的，只证明积和商的情况。

先证明乘积的情形，由

$$|x_n y_n - xy| = |(x_n y_n - x y_n) + (x y_n - xy)| \leqslant |y_n||x_n - x| + |x||y_n - y|$$

任取 $\varepsilon > 0$，则存在 $N > 0$，使得 $n > N$ 时同时恒有 $|x_n - x| < \varepsilon$ 和 $|y_n - y| < \varepsilon$[2]，另外再由收敛数列的有界性，存在 $M > 0$，使得 $|y_n| < M$，于是就有 $|x_n y_n - xy| < (M + |x|)\varepsilon$，所以 $x_n y_n$ 收敛到 xy.

再来证明商的情况，先证明一个结论，如果数列 y_n 收敛到一个非零实数 y，那么数列 $1/y_n$ 必收敛，且收敛到 $1/y$，这是因为

$$\left|\frac{1}{y_n} - \frac{1}{y}\right| = \left|\frac{y_n - y}{y y_n}\right|$$

对于任意正实数 $\varepsilon > 0$，上式的分子能从某一个下标 N 开始恒小于 ε，同时再取另外一个正实数 $|y|/2$，数列能从某项起恒有 $|y_n| > |y|/2$，于是从某个下标开始，上式就能恒小于 $2\varepsilon/y^2$，所以数列 $1/y_n$ 收敛到 $1/y$，再将 x_n/y_n 视为 x_n 乘以 $1/y_n$ 并利用乘积的结果，便得商的情形。 □

例 1.2.12 给定数列 x_n 的前两项 x_1 与 x_2，以后的项由公式

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

确定，我们来讨论一下它的收敛情况。

递推公式可以改写成

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

因此，递推下去有

$$x_{n+1} - x_n = \left(-\frac{1}{2}\right)^{n-1}(x_2 - x_1)$$

[2]本来对同一个 ε，两个数列的 N 是不同的，但是可以取比这两个 N 都大的 N，这时就同时有那两个不等式。

于是

$$\begin{aligned} x_n &= x_1 + \sum_{k=2}^{n}(x_k - x_{k-1}) \\ &= x_1 + (x_2 - x_1)\sum_{k=2}^{n}\left(-\frac{1}{2}\right)^{k-2} \\ &= x_1 + (x_2 - x_1)\cdot\frac{2}{3}\left(1 - \left(-\frac{1}{2}\right)^{n-1}\right) \end{aligned}$$

于是有

$$\lim_{n\to\infty} x_n = x_1 + \frac{2}{3}(x_2 - x_1) = \frac{x_1 + 2x_2}{3}$$

■

例 1.2.13 几何级数 设实数 q 满足 $|q| < 1$，称级数

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + \cdots$$

为几何级数，它的部分和

$$S_n = \sum_{i=0}^{n} q^i = \frac{1 - q^{n+1}}{1 - q}$$

因为 $|q| < 1$，所以可以得出

$$\lim_{n\to\infty} S_n = \frac{1}{1 - q}$$

也就是说，几何级数的和是

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

■

推论 1.2.2. *若数列 a_n 收敛到实数 A，则对于固定的整数 m，数列 a_n^m 收敛到实数 A^m.*

例 1.2.14 设实数 $a > 0$ 且 $a \neq 1$，证明极限 $\lim_{n\to\infty} \sqrt[n]{a} = 1$.

我们在 例 1.2.7中已经证明了 $a > 1$ 时的情形，现在假设 $0 < a < 1$，则有

$$\lim_{n\to\infty} \sqrt[n]{a} = \lim_{n\to\infty} \frac{1}{\sqrt[n]{\frac{1}{a}}} = 1$$

■

例 1.2.15 在例 1.2.4中，我们已经证明了下面的极限

$$\lim_{n\to\infty} \frac{n}{a^n} = 0$$

其中 $a > 1$，现在我们将它推广，设 m 是一个正整数，则

$$\lim_{n\to\infty} \frac{n^m}{a^n} = 0$$

这是因为，由于 $\sqrt[m]{a} > 1$，根据例 1.2.4中的结论，有

$$\lim_{n\to\infty} \frac{n}{(\sqrt[m]{a})^n} = 0$$

再由推论 1.2.2，便得出

$$\lim_{n\to\infty}\frac{n^m}{a^n}=0$$

这表明在 n 趋于无限大时，多项式与指数相比，它将变得微不足道，也就是说，指数是比多项式更高阶的无穷大。 ■

1.2.4 无穷小与无穷大

定义 1.2.3. 如果数列收敛到零，则称它在 n 无限增大时是一个无穷小.

定义 1.2.4. 如果数列 a_n 满足: 对任意大的正实数 M, 总存在正整数 N, 使得当 $n>N$ 时恒有 $|a_n|\geqslant M$，则称这数列在 n 无限增大时是一个无穷大，如果 a_n 在 n 充分大时还保持着一定的符号，则称之为正无穷大或者负无穷大，此时也称数列收敛到正无穷或者收敛到负无穷，记作 $\lim\limits_{n\to\infty}a_n=+\infty$ 或者 $\lim\limits_{n\to\infty}a_n=-\infty$.

显然，当 a_n 是非零的无穷小时，$\dfrac{1}{a_n}$ 是无穷大，反之也对。

例如 $\dfrac{1}{n}$ 便是无穷小的例子，而 n^2 则是一个无穷大的例子。

为讨论的方便，我们引入无穷大的邻域概念.

定义 1.2.5. 正无穷大的邻域是开区间 $(a,+\infty)$，负无穷大的邻域是开区间 $(-\infty,a)$.

例 1.2.16 今来考虑关于 n 的多项式

$$a_mn^m+a_{m-1}n^{m-1}+\cdots a_1n+a_0$$

在 $n\to\infty$ 时的极限情况。

将最高次幂提出来，多项式写成下面的形状

$$n^m\left(a_m+\frac{a_{m-1}}{n}+\frac{a_{m-2}}{n^2}+\cdots+\frac{a_0}{n^m}\right)$$

在 $n\to\infty$ 时，显然括号中的部分将趋于极限值 a_m，但外面有一个无穷大因子 n^m，所以我们得到结论，如果最高次项系数为正，则多项式将趋于正无穷，如果这系数为负，则它将趋于负无穷。 ■

例 1.2.17 再来考察关于 n 的有理式子的极限，即关于 n 的两个多项式之比的极限:

$$\frac{a_mn^m+a_{m-1}n^{m-1}+\cdots a_1n+a_0}{b_ln^l+b_{l-1}n^{l-1}+\cdots b_1n+b_0}$$

这里 $a_m\neq 0$，$b_l\neq 0$.

如果 $m=l$，则分子分母同时除以最高次幂变形为

$$\frac{a_m+\frac{a_{m-1}}{n}+\cdots+\frac{a_0}{n^m}}{b_m+\frac{b_{m-1}}{n}+\cdots+\frac{b_0}{n^m}}$$

显然它的极限是 $\dfrac{a_m}{b_m}$，即最高次项系数之比。

如果 $m<l$，则分子分母同时除以 n^l 后，原式成为

$$\frac{\frac{a_m}{n^{l-m}}+\cdots+\frac{a_0}{n^l}}{b_m+\frac{b_{m-1}}{n}+\cdots+\frac{b_0}{n^l}}$$

分子趋于 0，而分母趋于 b_m，因此此时极限为零。

同理，当 $m>l$ 时，分子分母同时除以 n^l 后，分母趋于有限的极限 b_m，但分母却是无穷大，所以此时它的极限是无穷大，其符号由分子分母的最高次项系数共同决定，相同时为正无穷，相反时为负无穷。 ■

定义 1.2.6. *设数列 a_n 与 b_n 都是无穷小，即在 n 无限增大时都收敛到零，*
(1). 若 $\lim\limits_{n\to\infty}\dfrac{a_n}{b_n}=0$，则称 a_n 是 b_n 的高阶无穷小，记作 $a_n=o(b_n)$.
(2). 若存在正实数 U 和 V，使得当 n 充分大时有 $U\leqslant\left|\dfrac{a_n}{b_n}\right|\leqslant V$，则称 a_n 与 b_n 是同阶无穷小，特别的，如果 $\lim\limits_{n\to\infty}\dfrac{a_n}{b_n}=1$，则称它俩是等价无穷小，记作 $a_n\sim b_n$.

例如，$\dfrac{1}{n^2}$ 是 $\dfrac{1}{n}$ 的高阶无穷小，而 $\dfrac{1}{n}$ 与 $\dfrac{1+\sin n}{n}$ 则是同阶无穷小，这个概念体现了无穷小之间的比较，需要说明的是，两个无穷小之间并不必然有阶的高低之分，因为它俩之比可以不收敛。

关于无穷小，还有如此结论

定理 1.2.7. *如果数列 a_n 是无穷小，数列 b_n 有界，则数列 a_nb_n 是无穷小。*

证明 由条件，存在正实数 M，使得 b_n 中的所有项都满足 $|b_n|\leqslant M$，所以 $|a_nb_n|\leqslant M|a_n|$ 总是成立的，而 a_n 为无穷小，所以对于无论多么小的正实数 ε，数列 a_n 总能从某项起恒满足 $|a_n|<\varepsilon/M$，从而 $|a_nb_n|<\varepsilon$，于是结论成立。 □

由此知道，数列 $\dfrac{\sin n}{n}$ 是无穷小.

定理 1.2.8. *有限个无穷小之和仍是无穷小.*

利用极限定义即可简单证明，略去。

关于同阶无穷小，则有如下结论

定理 1.2.9. *若 a_n 与 b_n 是同阶无穷小，c_n 是任一数列，则数列 a_nc_n 与 b_nc_n 的收敛性情况相同，即要么都收敛要么都发散，并且在收敛的情况下，如果又是等价无穷小，则还具有相同的极限值.*

这定理关于等价无穷小的结论是非常有用的，它表明在计算极限的过程中，可以将一个无穷因子替换为与它等价的无穷小。

与无穷小类似，我们可以定义高阶无穷大的概念。

定义 1.2.7. *设数列 x_n 和 y_n 在 $n\to\infty$ 时都是无穷大，如果 $\lim\limits_{n\to\infty}\left|\dfrac{x_n}{y_n}\right|=+\infty$，则称 x_n 是 y_n 的高阶无穷大.*

x_n 是 y_n 的高阶无穷大，就意味着当 $n \to \infty$ 时，y_n 的大小与 x_n 相比完全可以忽略不计，尽管 y_n 自身也是无穷大。

例 1.2.18　根据例 1.2.5、例 1.2.6、例 1.2.15的结果，我们可以知道，在对数、多项式、指数、阶乘这些无穷大中，后面的无穷大，都是前面的无穷大的高阶无穷大，如果我们用 $x_n \prec y_n$ 来表示 y_n 是 x_n 的高阶无穷大，那么有

$$\ln n \prec n \prec n^2 \prec \cdots \prec n^m \prec a^n \prec n!$$

式中 $m > 2$ 是正整数，$a > 1$ 为一个实数。 ■

1.2.5　Stolz 定理

定理 1.2.10 (Stolz 定理). *对于两个数列 a_n 和 b_n，其中 b_n 是一个 (至少从某一项开始) 严格增加到正无穷或者严格减小到负无穷的数列，如果 $\lim\limits_{n\to\infty} \dfrac{a_{n+1} - a_n}{b_{n+1} - b_n} = M$，那么有 $\lim \dfrac{a_n}{b_n} = M$，这里的 M 可以是一个实数，也可以是正无穷或者负无穷。*

证明　先证明 b_n 严格增加到正无穷以及 l 是一个有限数的情况，此时由两个数列增量之比收敛到 l，所以对于任意小的正数 ε，当 n 充分大时 $(n \geqslant N)$ 恒有

$$(l-\varepsilon)(b_{n+1} - b_n) < a_{n+1} - a_n < (l+\varepsilon)(b_{n+1} - b_n)$$

累加可得

$$(l-\varepsilon)(b_n - b_N) < a_n - a_N < (l+\varepsilon)(b_n - b_N)$$

三边同时加上 a_N 再除以 b_n 可得 (b_n 必定从某一项开始恒保持正号)

$$l - \varepsilon + \frac{a_N - (l-\varepsilon)b_N}{b_n} < \frac{a_n}{b_n} < l + \varepsilon + \frac{a_N - (l+\varepsilon)b_N}{b_n}$$

因为 b_n 收敛到正无穷，所以当 n 充分大时，上式左右两边以 b_n 为分母的两项的绝对值可以任意小，从而当 n 充分大时就有

$$l - 2\varepsilon < \frac{a_n}{b_n} < l + 2\varepsilon$$

这就表明两个数列之比也收敛到 l.

要说明的是，上述证明过程中，N 的取值并不是恒定不变的，每出现一次“当 n 充分大时”之类的字眼，就意味着 N 可能需要取更大的值，以使得前面的不等式与新引入的不等关系能够同时成立。

再证明 l 是正无穷大的情形，这时可以利用已经证明了的有限数的结果，因为这时显然有

$$\lim_{n\to\infty} \frac{b_{n+1} - b_n}{a_{n+1} - a_n} = 0$$

只需要说明 a_n 能够从某项起单调增加到正无穷大就可以了。对于任意大的正实数 M，当 n 充分大时显然有

$$a_{n+1} - a_n > M(b_{n+1} - b_n)$$

而显然从某一个正整数开始恒有 $b_{n+1} - b_n > 1$，所以此时有 $a_{n+1} - a_n > M$，这便说明 a_n 能够从某一项开始单调增加到正无穷了，于是便可以得到我们要的结果。 □

推论 1.2.3. 如果数列 x_n 收敛到 A，则 $(x_1+x_2+\cdots+x_n)/n$ 也收敛到 A，反之亦然。

这正是我们在例 1.2.10中所得到的结论，那里是用极限定义证明了这个事实，现在有了 Stolz 定理，它就是一个显而易见的结论了。

证明 只要在 Stolz 定理中取 $a_n=x_1+x_2+\cdots+x_n$，$b_n=n$ 便得此结论。 □

例 1.2.19 我们证明，对任意确定的正整数 m，有

$$\lim_{n\to\infty}\frac{1^m+2^m+\cdots+n^m}{n^{m+1}}=\frac{1}{m+1}$$

只要取 $a_n=1^m+2^m+\cdots+n^m$ 与 $b_n=n^{m+1}$，有

$$\lim_{n\to\infty}\frac{a_{n+1}-a_n}{b_{n+1}-b_n}=\lim_{n\to\infty}\frac{(1+n)^m}{(1+n)^{m+1}-n^{m+1}}=\lim_{n\to\infty}\frac{n^m+\cdots}{(m+1)n^m+\cdots}=\frac{1}{m+1}$$

根据 Stolz 定理，便得原式亦有相同极限。

在初等数学中，我们已经知道自然数的幂和 $S(n,m)=\sum\limits_{i=1}^{n}i^m$ 是一个关于 n 的 $m+1$ 次多项式，并且它的最高次项系数就是 $\dfrac{1}{m+1}$，所以这里的极限是毫不意外的。

■

例 1.2.20 现在把上一例的结果进行加强，上例中求出了

$$\lim_{n\to\infty}\frac{1^m+2^m+\cdots+n^m}{n^{m+1}}=\frac{1}{m+1}$$

那么左端的数列减去极限值便是一个无穷小，这个无穷小的阶是如何的呢，现在我们拿它与 $\dfrac{1}{n}$ 进行比较，即要求极限

$$\lim_{n\to\infty}n\left(\frac{1^m+2^m+\cdots+n^m}{n^{m+1}}-\frac{1}{m+1}\right)$$

将原式通分变形为

$$\frac{(m+1)(1^m+2^m+\cdots+n^m)-n^{m+1}}{(m+1)n^m}$$

将分子作为数列 x_n，分母作为数列 y_n，显然 y_n 是单调增加到正无穷的，而

$$\begin{aligned}\frac{x_{n+1}-x_n}{y_{n+1}-y_n}&=\frac{(m+1)(n+1)^m-((n+1)^{m+1}-n^{m+1})}{(m+1)((n+1)^m-n^m)}\\&=\frac{(m+1)\sum_{i=0}^m C_m^i n^i-\sum_{i=0}^m C_{m+1}^i n^i}{(m+1)\sum_{i=0}^m C_m^i n^i}\end{aligned}$$

可以发现，分子分母都是关于 n 的 $m-1$ 次多项式，因此极限应对应的最高次项系数之比，即

$$\frac{(m+1)C_m^{m-1}-C_{m+1}^{m-1}}{(m+1)C_m^{m-1}}=\frac{1}{2}$$

于是根据 Stolz 定理，就有

$$\lim_{n\to\infty}n\left(\frac{1^m+2^m+\cdots+n^m}{n^{m+1}}-\frac{1}{m+1}\right)=\frac{1}{2}$$

这就是说，上例中的数列与极限作差的无穷小，与 $\dfrac{2}{n}$ 是等价无穷小. ■

例 1.2.21 我们证明一个关于调和级数与对数之间关系的一个重要极限:

$$\lim_{n\to\infty}\frac{1+\frac{1}{2}+\cdots+\frac{1}{n}}{\ln n}=1$$

取 $a_n=1+\dfrac{1}{2}+\cdots+\dfrac{1}{n}$，$b_n=\ln n$，则

$$\lim_{n\to\infty}\frac{a_{n+1}-a_n}{b_{n+1}-b_n}=\lim_{n\to\infty}\frac{1}{(n+1)\ln\left(1+\frac{1}{n}\right)}=1$$

显然 b_n 是单调增加到正无穷的，因此由 Stolz 定理即得所要证的极限。

这个例子的意义在于，它揭示了调和级数的增长速度，与对数是同一级别的，以后我们还将看到，它俩不但增长速度接近，事实上它俩之差会趋于一个固定值 (欧拉常数 $c=0.577\cdots$). ■

1.2.6 单调有界定理

定理 1.2.11. *单调递增有上界的数列必定收敛，而且收敛到它的上确界。单调递减有下界的数列也类似。*

证明 只证明单调递增有上界的情况，假如数列 x_n 就是这样的数列，它的上确界是 M，则数列中的全部项都满足 $x_n\leqslant M$，另外，对于任意小的正实数 ε，由上确界定义，总存在某个 x_N 满足 $x_N>M-\varepsilon$，再由单调性即知对于 $n>N$ 恒有 $M-\varepsilon<x_n\leqslant M<M+\varepsilon$，所以 M 就是这数列的极限。 □

例 1.2.22 考虑下面的具有 n 重根号的式子

$$\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}$$

为求其无限个根号时的极限，用数列来归纳的定义它,$x_1=\sqrt{2}$，$x_{n+1}=\sqrt{2+x_n}$.

显然这是一个正项数列，由 $x_n<2$ 可得到 $x_{n+1}=\sqrt{2+x_n}<\sqrt{2+2}=2$，又 $x_1<2$，因此数列有上界，又

$$x_{n+1}^2-x_n^2=2+x_n-x_n^2=(2-x_n)(1+x_n)>0$$

所以数列单调增加，于是这数列有极限，设其极限为 x，则在递推式两边取极限得 $\sqrt{2+x}=x$，解之得 $x=2$，即它收敛到 2.

可以看到，如果 x_1 的取值大于 2，则这数列将单调减少并恒大于 2，但会收敛到 2.

类似的，对于递推数列 $x_{n+1}=\sqrt{c+x_n}$，这里 $c>0$ 为常数，如果数列收敛，设极限为 x，显然 $x=\sqrt{c+x}$，即 $x=\dfrac{1+\sqrt{1+4c}}{2}$，又如果 $x_n<x$，则显然 $x_{n+1}<x$，如果 $x_n>x$，则 $x_{n+1}>x$，也就是说，只要首项 x_1 在 x 的某一边，则数列的全部项都在同一侧，而

$$x_{n+1}^2-x_n^2=c+x_n-x_n^2=\left(x_n-\frac{1+\sqrt{1+4c}}{2}\right)\left(\frac{1-\sqrt{1+4c}}{2}-x_n\right)$$

所以如果 $x_1 < x = \dfrac{1+\sqrt{1+4c}}{2}$，则 x_n 单调增加有上界，如果 $x_1 > x$，则数列单调减少有下界，无论哪种情况，都会收敛到 $\dfrac{1+\sqrt{1+4c}}{2}$. ■

例 1.2.23 序列

$$1+\frac{1}{2^2}+\cdots+\frac{1}{n^2}$$

显然是单调增加的，又

$$\begin{aligned} & 1+\frac{1}{2^2}+\cdots+\frac{1}{n^2} \\ < \ & 1+\frac{1}{1\times 2}+\cdots+\frac{1}{(n-1)n} \\ = \ & 1+\left(1-\frac{1}{2}\right)+\cdots+\left(\frac{1}{n-1}-\frac{1}{n}\right) \\ = \ & 2-\frac{1}{m}<2 \end{aligned}$$

因此它有上界，于是收敛，也就是说，级数 $\lim\limits_{n=1}^{\infty}\dfrac{1}{n^2}$ 收敛，以后会看到，这个和是 $\dfrac{\pi^2}{6}$. ■

例 1.2.24 序列

$$1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}$$

它显然是单调增加的，而关于它上有界，则有两种放缩方式可以证明，其一是利用 $\dfrac{1}{n!}<\dfrac{1}{2^{n-1}}(n\geqslant 2)$，其二是利用 $\dfrac{1}{n!}<\dfrac{1}{(n-1)n}=\dfrac{1}{n-1}-\dfrac{1}{n}$，因此它也有上界，从而有极限，这就是说: 级数 $\sum\limits_{n=0}^{\infty}\dfrac{1}{n!}$ 收敛，以后将看到，它会收敛到 e，自然对数的底数。 ■

例 1.2.25 序列

$$\left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)\cdots\left(1-\frac{1}{2^n}\right)$$

显然是单调减少有下界的，因此它收敛，不过它的极限，却不能从递推式取极限而得出来了。 ■

例 1.2.26 等差-等比中项 考虑两个数列 a_n 和 b_n，任取 $a_1=a>0$，$b_1=b>0$，以后的项由下式确定

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

现考察一下这两个数列的极限情况。

显然，两个数列的所有项都是正的，而且无论两个数列的第一项如何取值，从第二项起便恒有 $a_n\leqslant b_n$，因此

$$a_{n+1}=\sqrt{a_nb_n}\geqslant a_n$$

即 a_n 从第二项开始单调增加 (不减)，同时有

$$b_{n+1}=\frac{a_n+b_n}{2}\leqslant b_n$$

即 b_n 从第二项开始单调减少 (不增)，另一方面，容易发现两个数列都是有界的:

$$a_2\leqslant a_3\leqslant\cdots\leqslant a_n\leqslant\cdots\leqslant b_n\leqslant\cdots\leqslant b_3\leqslant b_2$$

即 a_n 是单调增加有上界，而 b_n 是单调减少有下界，因此它俩都有极限，再在 a_n 的递推式两端取极限便知它们有着共同的极限值，毫无疑问，这个极限值是由两个数列的首项 a 和 b 所唯一确定的，它称为 a 和 b 的*等差-等比中项*，但要想在这里求出它的具体表达式，却是不可能的，它不可能用目前我们已知的任何表达式写出来，以后将会看到，它可以利用椭圆积分表出。 ■

例 1.2.27 调和-算术中项 与例 1.2.26相仿，不过现在我们将两个平均数分别替换为调和平均与算术平均:

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

类似的可得出

$$a_2\leqslant a_3\leqslant\cdots\leqslant a_n\leqslant\cdots\leqslant b_n\leqslant\cdots\leqslant b_3\leqslant b_2$$

因此这两个数列仍有极限，这极限称为数 a 与 b 的*调和-算术平均*，与算术-几何平均不同的是，调和-算术平均有简单表达式，因为恒有关系

$$a_{n+1}b_{n+1}=\frac{2a_nb_n}{a_n+b_n}\cdot\frac{a_n+b_n}{2}=a_nb_n$$

因此恒有 $a_nb_n=a_1b_1=ab$，所以这个极限是 $\sqrt{ab}$，即两个数的调和-算术平均，就是两者的几何平均。 ■

例 1.2.28 调和级数的发散性 在这个例子中，我们来考虑*调和级数*

$$\sum_{n=1}^{\infty}\frac{1}{n}=1+\frac{1}{2}+\frac{1}{3}+\cdots$$

的收敛性，它的部分和

$$H_n=1+\frac{1}{2}+\cdots+\frac{1}{n}$$

虽然 H_n 随着 n 的增大，它的增量逐渐减小，但可以证明它可以大于任意正实数，因为我们有

$$\begin{aligned}
\frac{1}{3}+\frac{1}{4} &> 2\times\frac{1}{4}=\frac{1}{2}\\
\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8} &> 4\times\frac{1}{8}=\frac{1}{2}\\
&\cdots\cdots\\
\frac{1}{2^m+1}+\frac{1}{2^m+2}+\cdots\frac{1}{2^{m+1}} &> 2^m\times\frac{1}{2^{m+1}}=\frac{1}{2}
\end{aligned}$$

因此

$$\begin{aligned}
H_{2^m} &= 1+\frac{1}{2}+\sum_{i=1}^{m}\left(\frac{1}{2^i+1}+\frac{1}{2^i+2}+\cdots\frac{1}{2^{i+1}}\right)\\
&> 1+\frac{1}{2}+\frac{1}{2}m
\end{aligned}$$

这就表明 H_{2^m} 是一个无穷大，再由 H_n 的单调性，可知它是一个正无穷大，也就是说调和级数是发散的。 ■

例 1.2.29 数列

$$\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n}$$

它共有 n 个加数，且依次减小，把所有加数都放大为第一个，则得出它小于 $\frac{n}{n+1}<1$，这表明它有上界，再看单调性，当 n 变为 $n+1$ 时，末尾会增加两项 $\frac{1}{2n+1}$ 和 $\frac{1}{2n+2}$，但前面会少一项 $\frac{1}{n+1}$，因此其增量为

$$\frac{1}{2n+1}+\frac{1}{2n+2}-\frac{1}{n+1}>2\cdot\frac{1}{2n+2}-\frac{1}{n+1}=0$$

故它又是递增的，因此有极限，且极限不超过 1，以后将看到 (例 1.2.31)，它的极限是 $\ln 2$. ■

1.2.7 数 e

这一小节我们来证明下面这个数列有极限:

$$x_n=\left(1+\frac{1}{n}\right)^n$$

证明一 [3] 由多元均值不等式，把 x_n 看成 n 个 $(1+1/n)$ 的乘积，再添加上一个因数 1 构成 $n+1$ 个数的乘积，有

$$\left(1+\frac{1}{n}\right)^n=1\cdot\left(1+\frac{1}{n}\right)^n<\left(\frac{1+n\left(1+\frac{1}{n}\right)}{n+1}\right)^{n+1}=\left(1+\frac{1}{n+1}\right)^{n+1}$$

这便表明它是递增的。

下证它是有上界的，把 $n+1$ 拆分成 $\frac{5}{6}$ 和 n 个 $1+\frac{1}{6n}$，由均值不等式得

$$n+1=\frac{5}{6}+n\left(1+\frac{1}{6n}\right)>(n+1)\sqrt[n+1]{\frac{5}{6}\cdot\left(1+\frac{1}{6n}\right)^n}$$

整理即得

$$\left(1+\frac{1}{6n}\right)^n<\frac{6}{5}$$

所以

$$\left(1+\frac{1}{6n}\right)^{6n}<\left(\frac{6}{5}\right)^6<3$$

于是由单调性便知

$$\left(1+\frac{1}{n}\right)^n<3$$

所以数列 x_n 单调增加且有上界，故此存在极限。 □

证明二 [4] 把 x_n 按二项式定理展开得

$$\begin{aligned}x_n &= 1+\sum_{i=1}^{n}C_n^i\frac{1}{n^i}\\ &= 1+\sum_{i=1}^{n}\frac{1}{i!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{i-1}{n}\right)\end{aligned}$$

[3]这个证明来自参考文献 [?].

[4]这个证明来自于参考文献 [3].

易见对于 x_{n+1} 而言，在上式的基础上会多出 $i=n+1$ 的一个正项，并且其它项是把上式中每一个项中的每一个因子 $1-\frac{i}{n}$ 更换为更大的因子 $1-\frac{i}{n+1}$，所以 $x_{n+1}>x_n$，这是一个递增的数列.

将每一个项中的所有 $(1-i/n)$ 因子全部放大为 1，则有

$$x_n<1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}$$

接下来有两种放缩方式都可以证明它有上限:

$$\frac{1}{k!}<\frac{1}{k(k-1)}=\frac{1}{k-1}-\frac{1}{k}$$

和

$$\frac{1}{k!}<\frac{1}{2^{k-1}}$$

于是

$$x_n<2+\sum_{i=2}^{n}\left(\frac{1}{i-1}-\frac{1}{i}\right)=3-\frac{1}{n}<3$$

或者

$$x_n<2+\sum_{i=2}^{n}\frac{1}{2^{i-1}}=3-\frac{1}{2^{n-1}}<3$$

所以数列单调递增有上界，故此有极限. □

证明三 [5] 设实数 $a>b>0$，有

$$a^{n+1}-b^{n+1}=(a-b)(a^n+a^{n-1}b+\cdots+ab^{n-1}+b^n)<(n+1)a^n(a-b)$$

即

$$b^{n+1}>a^n(a-(n+1)(a-b))=a^n((n+1)b-na) \tag{1.2.3}$$

在此式中取 $a=1+\dfrac{1}{n}, b=1+\dfrac{1}{n+1}$，由于 $(n+1)b-na=1$，故得

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

即数列 x_n 单调增加，再在前式中取 $a=1+\dfrac{1}{2n}$,$b=1$，则 $(n+1)b-na=\dfrac{1}{2}$，则得

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

即

$$\left(1+\frac{1}{2n}\right)^{2n}<4$$

因为 x_n 是递增的，又对于任意正整数 n，都存在大于它的偶数，所以得到

$$\left(1+\frac{1}{n}\right)^n<4$$

故数列有界，从而存在极限。 □

[5]这个证明来自参考文献 [3].

我们用字母 e 来表示这个数列的极限，它的值是

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=\mathrm{e}=2.718281828459045\cdots$$

例 1.2.30 这个极限值 e 是一个非常重要的常数，对于数列 $x_n=\left(1+\frac{1}{n}\right)^n$，我们有如下一个解释，假如银行一年期定期存款的年利率是 r（通常是 3% 左右），那么如果将本金 1 万元存入银行，一年后所得收入是 $1+r$ 万元，而如果我们每个月按月利率 $\frac{r}{12}$ 计算一次月利息，并按复利方式（即把上一期的本金和利息一起计入下一期的本金）计算，那么一年后收入将是 $\left(1+\frac{r}{12}\right)^{12}$ 万元，由伯努利不等式可知，这个值比一年算一次利息的收入大了些，于是我们猜想，如果我们不嫌麻烦，按日利率 $\frac{r}{365}$ 每天计算一次利息，那么一年后的收入则是 $\left(1+\frac{r}{365}\right)^{365}$，这个值又比之前按月计息的收入高了些 (数列 x_n 的单调性)，进一步想象，如果我们继续减小计息周期，每分钟一次、每秒一次，甚至每毫秒计息一次，我们的收入会不会变得无穷大呢？这不是一本万利的事情吗，然而

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

这表明，在无限缩小计息周期的时候，总收入是有一个上确界的，会无限的逼近 e^r 但永远不会达到它，这一本万利的好事只能是个梦罢了。在这个意义上，也可以认为，常数 e 是个自然界中客观存在的常量，它还是自然对数的底数，之所以选择它作为对数的底，是因为涉及对数的很多公式，在选择它作底的时候，往往能够使公式具有最简单的形式。 ■

在前述证明二中我们已经证得

$$x_n<1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}=y_n$$

并且证明二就是利用了 y_n 有上界而得出 x_n 有上界的，显然 y_n 也是单调增加的，所以 y_n 也有极限，其实它也收敛到 e，在此证明这一点。

在证明二的过程中已经得到

$$x_n=1+\sum_{i=1}^{n}\frac{1}{i!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{i-1}{n}\right)$$

任选整数 k 满足 $0<k<n$，并在上式右端的求和中舍弃 $i>k$ 的加项，得 (注意求和的上限)

$$x_n>1+\sum_{i=1}^{k}\frac{1}{i!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{i-1}{n}\right)\tag{1.2.4}$$

固定整数 k 不变，则右端是有限个数之和，让 n 无限增大两端取极限，得

$$\mathrm{e}\geqslant 1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{k!}=y_k$$

于是得出

$$x_n<y_n\leqslant\mathrm{e}$$

由夹逼准则，就得出

$$\lim_{n\to\infty}\left(1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}\right)=\mathrm{e}$$

或者说，有下面的级数

$$\sum_{n=0}^{\infty}\frac{1}{n!}=\mathrm{e} \tag{1.2.5}$$

现在考虑一个问题，我们是由 $x_n<y_n\leqslant \mathrm{e}$ 以及 $\lim\limits_{n\to\infty}x_n=\mathrm{e}$ 得出 $\lim\limits_{n\to\infty}y_n=\mathrm{e}$ 的，但如果我们根据 y_n 单调增加并有上界从而定义 e 为它的极限，那又如何说明 x_n 也收敛到 e 呢，此时由 $x_n<y_n\leqslant \mathrm{e}$ 以及 y_n 收敛到 e 显然是无法得出 x_n 的极限的。

这点也不难，此时在式 1.2.4中固定 k 并令 $n\to\infty$ 有 $\lim\limits_{n\to\infty}x_n>y_k$，又 $x_n<y_n$，故得

$$x_n<y_n<\lim_{n\to\infty}x_n$$

由于 x_n 与 y_n 都收敛，对上式取极限就得

$$\lim_{n\to\infty}x_n\leqslant\lim_{n\to\infty}y_n\leqslant\lim_{n\to\infty}x_n$$

于是就有

$$\lim_{n\to\infty}x_n=\lim_{n\to\infty}y_n=\mathrm{e}$$

现在再考虑另一个数列

$$z_n=\left(1+\frac{1}{n}\right)^{n+1}$$

显然它有下界，我们再证明它是单调递减的，由伯努利不等式 (定理 1.1.6) 有

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

整理即得

$$\left(1+\frac{1}{n-1}\right)^n>\left(1+\frac{1}{n}\right)^{n+1}$$

即 $z_n>z_{n+1}$，于是它单调减少，从而收敛，又根据

$$z_n=\left(1+\frac{1}{n}\right)x_n$$

知它与 x_n 有共同的极限

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

事实上，x_n 与 z_n 构成一个闭区间套序列，而由这区间套确定的实数就是 e。

关于 x_n 趋于极限的速度，我们有

$$0<e-\left(1+\frac{1}{n}\right)^n<\frac{3}{n}$$

证明一 根据前面的结论，数列

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

单调减少并收敛到 e，因此

$$0<\mathrm{e}-\left(1+\frac{1}{n}\right)^n<\left(1+\frac{1}{n}\right)^{n+1}-\left(1+\frac{1}{n}\right)^n=\frac{1}{n}\left(1+\frac{1}{n}\right)^n<\frac{3}{n}$$

即得结论. □

证明二　我们只要证明数列

$$\left(1+\frac{1}{n}\right)^n+\frac{3}{n}$$

单调减少并收敛到 e 就可以了。

为了证明单调减少，只要证明

$$\left(1+\frac{1}{n}\right)^n+\frac{3}{n}>\left(1+\frac{1}{n+1}\right)^{n+1}+\frac{3}{n+1}$$

即要证

$$\left(1+\frac{1}{n+1}\right)^{n+1}-\left(1+\frac{1}{n}\right)^n<\frac{3}{n(n+1)}$$

设 $a>b>0$，则

$$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+\cdots+ab^{n-2}+b^{n-1})>nb^{n-1}(a-b)$$

取 $a=1+\dfrac{1}{n}$，$b=1+\dfrac{1}{n+1}$，有

$$\left(1+\frac{1}{n}\right)^n-\left(1+\frac{1}{n+1}\right)^n>\frac{1}{n+1}\left(1+\frac{1}{n+1}\right)^{n-1}$$

所以

$$\begin{aligned}
&\left(1+\frac{1}{n+1}\right)^{n+1}-\left(1+\frac{1}{n}\right)^n\\
=\ &\left(1+\frac{1}{n+1}\right)\left(1+\frac{1}{n+1}\right)^n-\left(1+\frac{1}{n}\right)^n\\
=\ &\left(1+\frac{1}{n+1}\right)^n-\left(1+\frac{1}{n}\right)^n+\frac{1}{n+1}\left(1+\frac{1}{n+1}\right)^n\\
<\ &-\frac{1}{n+1}\left(1+\frac{1}{n+1}\right)^{n-1}+\frac{1}{n+1}\left(1+\frac{1}{n+1}\right)^n\\
=\ &\frac{1}{(n+1)^2}\left(1+\frac{1}{n+1}\right)^{n-1}\\
<\ &\frac{3}{(n+1)^2}<\frac{3}{n(n+1)}
\end{aligned}$$

于是就证明了它是单调减少的，显然它有下界（比如零是它的一个下界），所以它有极限，而

$$\lim_{n\to\infty}\left(\left(1+\frac{1}{n}\right)^n+\frac{3}{n}\right)=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n+\lim_{n\to\infty}\frac{3}{n}=\mathrm{e}+0=\mathrm{e}$$

于是便得结论

$$0<e-\left(1+\frac{1}{n}\right)^n<\frac{3}{n}$$

□

利用式 1.2.5，我们可以用不多的几项就得出 e 精确度较高的近似值，这是因为阶乘的增长速度相当快，比指数还快，因此这个级数收敛到它的和的速度也是相当快。为此考虑用它的第 n 个部分和去近似 e 时的误差限，我们有

$$\begin{aligned}
y_{n+m}-y_n &= \frac{1}{(n+1)!}+\frac{1}{(n+2)!}+\cdots+\frac{1}{(n+m)!}\\
&= \frac{1}{n!}\left(\frac{1}{n+1}+\frac{1}{(n+1)(n+2)}+\cdots+\frac{1}{(n+1)(n+2)\cdots(n+m)}\right)\\
&< \frac{1}{n!}\left(\frac{1}{n+1}+\frac{1}{(n+1)^2}+\cdots+\frac{1}{(n+1)^m}\right)\\
&= \frac{1}{n!}\cdot\frac{1}{n}\left(1-\frac{1}{(n+1)^m}\right)
\end{aligned}$$

固定 n，令 $m\to\infty$，就得到 (下式本来应该是小于等于，但只要在上面取更强的放缩，比如是提 $\dfrac{1}{(n+1)!}$ 而不是 $\dfrac{1}{n!}$，取极限之后就可以得到一个小于等于 $\dfrac{1}{nn!}$ 的量)

$$\mathrm{e}-y_n<\frac{1}{nn!}$$

于是对于任意正整数 n，存在实数 $\theta_n\in(0,1)$，使得

$$\mathrm{e}=1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}+\frac{\theta_n}{nn!} \tag{1.2.6}$$

利用式 1.2.6，我们可以证明 e 是无理数，因为如果它是有理数，则有

$$\frac{m}{n}=1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}+\frac{\theta_n}{nn!}$$

注意这里是先用有理数 $\dfrac{p}{q}$ 表示 e，再将右端展开到对应的 $n=q$ 的，然后只要在此式的两端同时乘以 $n!$，左端得出一个整数，但是右端却是一个整数加上一个真分数 $\dfrac{\theta_n}{n}$，矛盾，所以 e 是无理数，如此也可以说明式 1.2.6中的 θ_n 不可能取 1.

现在利用式 1.2.6估算一下，只要取 $n=20$，就有

$$\frac{1}{20\times 20!}=\frac{1}{48658040163532800000}=2.0551\ldots\times10^{-20}<3\times10^{-20}$$

再考虑前面前面 20 项的舍入误差，每一项都保留 20 位小数，此时的舍入误差 $<20\times 10^{-20}$，因此用前 20 项的和去逼近 e 的误差不超过 23×10^{-20}，因此至少具有 17 位精确小数。

1.2.8 欧拉常数 C

我们已经证明了

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

对上式取以 e 为底数的对数得

$$n\ln\left(1+\frac{1}{n}\right)<1<(n+1)\ln\left(1+\frac{1}{n}\right)$$

即

$$\frac{1}{n+1} < \ln\left(1+\frac{1}{n}\right) < \frac{1}{n}$$

而

$$\ln\left(1+\frac{1}{n}\right) = \ln(n+1) - \ln n$$

所以

$$\ln(n+1) - \ln n < \frac{1}{n} < \ln n - \ln(n-1) \tag{1.2.7}$$

这是关于对数与自然数倒数的一个重要不等式，以后还会用到它，在式中取 $n = 1, 2, \ldots$ 进行累加 (右端要从 2 开始)，可得

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

于是得

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

这表明序列

$$x_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \ln n$$

是有上界的，而作差便知它是单调增加的，因此它收敛，这个结论属于欧拉，所以它的极限被称为*欧拉常数*，其值是

$$C = 0.577216\ldots$$

利用欧拉常数，调和级数可以写成

$$H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \ln n + C + \gamma_n$$

这里 γ_n 是一个无穷小。

我们早已证明过，在 $n \to \infty$ 时,

$$1 + \frac{1}{2} + \cdots + \frac{1}{n}$$

与 $\ln n$ 是等价无穷大，而这里的结果显示，不仅如此，这两者的差值将趋于常数，这表明用对数来逼近调和级数时的误差始终是一个常数，而相对误差则渐趋于零。

例 1.2.31 利用欧拉常数，可以求下式的极限

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}$$

因为

$$H_n = \ln n + C + \gamma_n,\ H_{2n} = \ln 2n + C + \gamma_{2n}$$

所以

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} = \ln 2n - \ln n + \gamma_{2n} - \gamma_n$$

两边取极限便得

$$\lim_{n\to\infty}\left(\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n}\right) = \ln 2$$

利用式 1.2.7也可以证明这个极限，因为可以得到

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

即

$$\ln\frac{2n+1}{n+1}<\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n}<\ln 2$$

对不等式取极限便得结论。 ■

1.2.9 圆周率 π

1.2.10 闭区间套定理

定理 1.2.12 (闭区间套定理)**.** *设闭区间序列 $[a_n, b_n](n=1,2,\ldots)$ 满足:*

1. $[a_1,b_1]\supseteq[a_2,b_2]\supseteq\cdots[a_n,b_n]\supseteq[a_{n+1},b_{n+1}]\supseteq\cdots$
2. $\lim\limits_{n\to\infty}(b_n-a_n)=0$

则存在唯一实数 x，使得 $x\in[a_n,b_n](n=1,2,\ldots)$.

证明 显然 a_n 是一个单调增加有上界的序列，所有的 b_m 都是它的上界，同样，b_n 是单调减少有下界的数列，所有的 a_m 都是它的下界，因此两个数列均有极限，分别记为 A 和 B，由 $a_n<b_n$ 恒成立知 $A\leqslant B$，又如果 $A<B$，那么有 $a_n\leqslant A<B\leqslant b_n$，这表明 $b_n-a_n>B-A$，这与区间长度趋于零相矛盾，所以只能 $A=B$，并且它同时位于所有的闭区间上，如果还有另一个与之不同的点也能同处于所有的闭区间上，则显然与区间长度趋于零相矛盾，所以这个共同极限是唯一同时位于所有区间上的数。 □

1.2.11 柯西收敛准则

定理 1.2.13 (柯西收敛准则)**.** *数列 x_n 收敛的充分必要条件是，对于任意正实数 ε，总存在正整数 $N>0$，使得任意 $n_1>N$ 和任意 $n_2>N$ 及任意恒有 $|x_{n_1}-x_{n_2}|<\varepsilon$。*

证明 先证必要性. 如果数列 x_n 收敛到 x，那么对于任意正实数 ε，都有正整数 N，使得 $n>N$ 时恒有 $|x_n-x|<\varepsilon/2$，于是对于任意 $n_1>N$ 及 $n_2>N$，便有 $|x_{n_1}-x_{n_2}|=|(x_{n_1}-x)-(x_{n_2}-x)|\leqslant|x_{n_1}-x|+|x_{n_2}-x|<\varepsilon/2+\varepsilon/2=\varepsilon$。必要性得证。

再证充分性，先任取一个正实数 ε_1，数列中的项能够从某一项开始全部落在一个宽度为 $2\varepsilon_1$ 的闭区间 I_1，再取 $\varepsilon_2=\dfrac{1}{2}\varepsilon_1$，则数列又能够从另一项（一般来说需要向后推移若干个下标）开始尽数落在一个宽度为 $2\varepsilon_2$ 的闭区间 I_2 上，且 $I_1\supseteq I_2$，依次类推，得出一个闭区间序列 $I_1,I_2,\ldots$，其中每一个闭区间都能包含数列从某一项开始之后的全部项，且这些闭区间的长度趋于零，因此按闭区间套定理，存在唯一实数 M 同时处于所有的闭区间上，它将是这数列的极限，这是因为，对于任意正实数 ε，显然存在某个闭区间 I_m，使得 $I_m\subset(M-\varepsilon,M+\varepsilon)$，而数列能够从某一项开始全部落在闭区间 I_m 上，也就从这一项开始全部落在 M 的 ε 邻域内，所以 M 就是这数列的极限。 □

我们还可以得出柯西收敛准则的否定叙述：数列 a_n 发散的充分必要条件是，存在正实数 r，使得无论对于多大的正整数 N，总是存在两个正整数 n_1 和 n_2，使得 $|a_{n_1}-a_{n_2}|\geqslant r$.

例 1.2.32 前 n 个正整数的平方倒数和

$$S_n=1+\frac{1}{2^2}+\cdots+\frac{1}{n^2}$$

对它的片段有

$$\begin{aligned}S_{m+p}-S_m &= \frac{1}{(m+1)^2}+\frac{1}{(m+2)^2}+\cdots+\frac{1}{(m+p)^2}\\ &< \frac{1}{m(m+1)}+\frac{1}{(m+1)(m+2)}+\cdots+\frac{1}{(m+p-1)(m+p)}\\ &= \left(\frac{1}{m}-\frac{1}{m+1}\right)+\left(\frac{1}{m+1}-\frac{1}{m+2}\right)+\cdots+\left(\frac{1}{m+p-1}-\frac{1}{m+p}\right)\\ &= \frac{1}{m}-\frac{1}{m+p}<\frac{1}{m}\end{aligned}$$

所以对于任意正实数 ε，只要取 $N>1/\varepsilon$，就能保证柯西条件成立，于是数列 S_n 有极限，不过这极限值在此处是求不出来的，在以后我们将会利用无穷级数理论，得到它的极限值，这极限值与圆周率有关:

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{i^2}=\frac{\pi^2}{6}$$

■

例 1.2.33 前 n 个正整数的阶乘的倒数和

$$T_n=1+\frac{1}{2!}+\cdots+\frac{1}{n!}$$

它的片段和

$$\begin{aligned}T_{m+p}-T_m &= \frac{1}{(m+1)!}+\frac{1}{(m+2)!}+\cdots+\frac{1}{(m+p)!}\\ &< \frac{1}{2^{m+1}}+\frac{1}{2^{m+2}}+\cdots+\frac{1}{2^{m+p}}\\ &= \frac{1}{2^m}\left(1-\frac{1}{2^p}\right)<\frac{1}{2^m}\end{aligned}$$

可见它也满足柯西收敛条件，所以这个数列也有极限，它的极限便是自然对数的底数 e:

$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{i!}=e$$

■

例 1.2.34 级数收敛的柯西准则 因为级数收敛等同于它的部分和数列收敛，据此可以得出级数收敛的柯西准则：级数 $\sum_{n=1}^{\infty}a_n$ 收敛的充分必要条件是，对于任意小的正实数 ε，总存在正整数 n 和正整数 p，使得 $|a_{n+1}+a_{n+2}+\cdots+a_{n+p}|<\varepsilon$。

同样可以得出它的反面叙述：级数 $\sum_{n=1}^{\infty} a_n$ 发散的充分必要条件是，存在一个正实数 r，使得无论对于多大的 N，总存在该下标之后的某一片段和满足 $|a_{n+1}+a_{n+2}+\cdots+a_{n+p}|\geqslant r$. ■

例 1.2.35 调和级数的发散性 在例 1.2.28中，我们已经证明了调和级数

$$\sum_{n=1}^{\infty}\frac{1}{n}=1+\frac{1}{2}+\frac{1}{3}+\cdots$$

的发散性，现在我们可以用柯西收敛准则来说明，但是所用的放缩仍然与那里是一致的，根据柯西准则，为了说明它的发散性，只要证明对于存在某个正实数 r，无论对于多大的正整数 N，总存在两个正整数 n_1 与 n_2，使得 $|H_{n_1}-H_{n_2}|\geqslant r$ 就可以了。

仍然由

$$\frac{1}{2^m+1}+\frac{1}{2^m+2}+\cdots+\frac{1}{2^{m+1}}>2^m\cdots\frac{1}{2^{m+1}}=\frac{1}{2}$$

可知，对于任意正整数 N，只要取一个 $n_1=2^{m_1}>N$，再取 $n_2=2^{m_1+1}$，就有 $|H_{n_1}-H_{n_2}|>\frac{1}{2}$，于是部分和数列 H_n 发散，即调和级数发散。

要说明的是，这里的放缩方式不是唯一的，比如说从级数中的任一项开始，作片段和

$$\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+m}>\frac{m}{n+m}$$

为了让右端能够大于某个正整数，取 $m=n$，即片段的长度要依赖于片段中首项的下标，下标越大，片段越长，这里有

$$\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n}>\frac{1}{2}$$

于是我们将调和级数按下标划分片段，1 和 2 分别为单独的片段，从 3 到 4 为一段，从 5 到 8 为一段，9 到 16 为一段，这样划分之后，每一段上的和都大于 $\frac{1}{2}$，但是这样的划分与前面的放缩其实是一致的，但是我们可以改为这里的 m，以得出不同的划分，例如取 $m=2n$，就得

$$\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{3n}>\frac{2}{3}$$

于是片段的划分方式是，1 是单独的片段，从 2 到 3 为一段，从 4 到 9 为一段，从 10 到 27 为一段，依次类推，每个片段上的倒数和都大于 $\frac{2}{3}$，于是级数发散。 ■

1.2.12 聚点定理

定义 1.2.8. *对于无穷数集 A 和某个实数 x，若 x 的任意小的邻域 $(x-\varepsilon,x_\varepsilon)$ 内都包含了 A 中的无穷多个数，则称数 x 是数集 A 的一个聚点.*

例如，零是数集 $\left\{\frac{1}{n}\right\}$ 的一个聚点。显然，定义中的邻域改成空心邻域也没有什么影响。

聚点可以在数集中，也可以不在集合中，事实上，如果聚点在集合中，那就从集合中除去此数，它仍然是新数集的聚点，并不受影响。

定理 1.2.14. *任意有界的无限数集，至少有一个聚点。*

证明 因为数集有界，设它的全部元素被限于某个闭区间 I_1 中，它含有无限个元素，则将此闭区间对半切割为两个子闭区间，则至少其中一个子闭区间上含有原数集中的无穷多个元素，记此子闭区间为 I_2，同样再对 I_2 进行对半切割，则可以得出一个含有原数集中无穷多个元素的子闭区间 I_2，依次类推，得到一个闭区间序列，该序列中每个闭区间都包含了原数集中的无穷多个元素，并且这闭区间序列符合闭区间套定理的条件，因此由这闭区间序列确定出一个数 K，它就是原数集的一个聚点，这是因为，对于任意小的正实数 ε，显然前述闭区间中存在某个闭区间 $I_m \in (K-\varepsilon, K+\varepsilon)$，而 I_m 中包含了原数集的无穷个元素，因此 K 的邻域也就包含了原数集中的无穷个元素，即为它的一个聚点。 □

从数列中剔除掉一些项，剩下的项按照原来的顺序排列而成的新数列，称为原数列的一个子数列.

推论 1.2.4 (博雷尔-魏尔斯特拉斯引理). *任何有界数列，必存在收敛的子数列。*

证明 将数列的所有项作成一个有界数集，如果该数集是有限集，则表明数列中有无穷多个项取相同的数值，那么取这些项作成的子列就是收敛的，如果此数集是无限集，由聚点定理知它至少有一个聚点 x，那么任意取定一个正实数 ε_1，则数集中有无穷多个数落在 $(x-\varepsilon_1, x+\varepsilon_1)$ 内并且不等于 x，任取其中一个记为 a_{n_1}，即它在原数列中的下标是 n_1，则再取 $\varepsilon_2 = \min\{\frac{1}{2}\varepsilon_1, |x-a_{n_1}|\}$，又可以选落在区间 $(x-\varepsilon_2, x+\varepsilon)$ 中并且不等于 x 的项 a_{n_2}，这里 n_2 的选法要保证 $n_2 > n_1$，因为是有无穷多个数落在区间内，所以这总是可以办到的。依次类推，得出一个子列 $a_{n_1}, a_{n_2}, \ldots$，显然这个子列是收敛到 x 的。 □

现在讨论下无穷聚点，在讨论之前，我们先用邻域重新描述一下聚点：数 x 是无穷数集 A 的一个聚点的充分必要条件是，x 的任意邻域都包含了数集 A 中的无穷多个数。我们已经引进过，正无穷的邻域是开区间 $(a, +\infty)$，负无穷的邻域是开区间 $(-\infty, a)$，因此我们把聚点的这个用邻域的定义照搬到正无穷和负无穷上，就得出无穷聚点的定义。

定义 1.2.9. *如果对于任意实数 a，正无穷的邻域 $(a, +\infty)$ 中总是包含着无穷数集 A 中的无穷多个数，则称正无穷是数集 A 的一个无穷聚点，同理可以定义负的无穷聚点.*

在引入了无穷聚点之后，聚点定理中的数集有界的要求就可以去掉了，因为如果数集无界，显然它必然有无穷聚点。同样，博雷尔-魏尔斯特拉斯引理中数列有界的限制也就可以去掉了，以后我们所说的聚点，均可以是有限的也可以是无限的，在这种意义上，我们把前面的相关定理修正为:

定理 1.2.15. *任意无穷数集必有聚点，且必有最大聚点和最小聚点.*

必有聚点这是显然的，所有聚点组成一个集合，这集合有上下确界（可以无穷），这上下确界也必然是其聚点，关于这一点我们需要详细的证明，即要证明下面的定理

定理 1.2.16. *设 A 是一个无穷数集，而 E 是它的聚点集合，如果 E 也有聚点，则这聚点也必定是 A 的聚点. 换句话说，集合 E 具有特点：它的聚点必定也在自身中.*

证明 只证明有限的情况，设 x 是聚点 E 的一个有限聚点，则它的任意小空心邻域内必然包含了无穷多个 E 中的点，而根据聚点定义，显然这个空心邻域也就必然包含了 A 中的无穷多个点，因此它也是 A 的聚点. □

聚点集的这种特性称为是闭的，我们定义

定义 1.2.10. 如果一个无穷数集的所有聚点也在这集合中，则称这集合是闭集.

与闭集相对，我们还有开集的概念，为此先引入内点和界点的定义.

定义 1.2.11. 设 A 是一个数集，如果 $x \in A$ 并且存在 x 的某个邻域 $(x-\delta, x+\delta)$ 整个都被 A 所包含，则称为 x 是 A 的一个内点.

定义 1.2.12. 设 A 是一个数集，x 是一个实数，如果 x 的任意小的空心邻域内都同时包含有 A 中的点和不在 A 中的点，则称 x 是 A 的一个界点.

下面定义开集:

定义 1.2.13. 如果一个数集中的任一个点都是这数集的内点，则称这数集是开集.

注意闭集与开集并不互斥，比如实数集 $\mathbb{R}$，它既是开集又是闭集.

1.2.13 上极限与下极限

这一节利用聚点的概念来定义上极限与下极限，并在此基础上讨论极限存在的条件。

首先我们证明一个结论，前面已经知道，有界数集存在聚点，在此基础上还有如下结论

定理 1.2.17. 有界数集存在最大的聚点和最小的聚点。

证明 记有界数集 A 的聚点集合是 B，设 A 的上下确界分别是 M 和 L，显然任何大于 M 或者小于 L 的数都不可能是 A 的聚点，也就是说，B 也以 M 和 L 为上界和下界（但不能保证是确界），记 B 的上确界为 P，我们来证明，P 也是 A 的聚点，也就是 $P \in B$.

对于任意正实数 ε，存在 $b_1 \in B$，使得 $P-\varepsilon < b_1 < P$，但 b_1 是 A 的一个聚点，因此它的任意邻域内都含有 A 中的无限个元素，取适当的邻域，使其被包含于开区间 $(P-\varepsilon, P)$ 中，于是开区间 $(P-\varepsilon, P)$ 也就包含了数集 A 中的无限多个元素，从而 P 是数集 A 的聚点，同理，B 的下确界 Q 也是数集 A 的聚点，它们分别是数集 A 的最大聚点和最小聚点。 □

如果考虑到无穷聚点，则上述定理仍然成立。

以后我们会称数列的聚点，它是指由数列所有项的数值所组成的数集 (若有数值相同的项，则该值按重复次数计算) 的聚点，以后不再特别说明，显然有如下定理。

定理 1.2.18. 数 A 是数列 $\{a_n\}$ 的一个聚点的充分必要条件是，这数列存在收敛到 A 的子数列.

定义 1.2.14. *数列 a_n 的最大聚点 (可以无穷) 称为数列的上极限，记作 $\varlimsup\limits_{n\to\infty} a_n$，而最小的聚点称为它的下极限，记作 $\varliminf\limits_{n\to\infty} a_n$.*

显然，数列必同时存在上极限与下极限 (可以无穷)。

定理 1.2.19. *数列存在极限的充分必要条件是，它的上极限与下极限相等。*

证明 必要性，如果数列 a_n 收敛到 A，显然 A 是它的唯一聚点，上极限与下极限都是它。

充分性，设上极限与下极限都是 A，显然 A 的任意邻域内含有数列中的无限多项，而在此邻域外只能含有数列中的有限多项，否则在此邻域外还有别的聚点存在，所以数列只能收敛到此值。 □

1.2.14 无穷级数

1.2.15 复数数列的极限

1.2.16 题选

题 1.2.1 已知非负实数数列 $\{a_n\}$ 对任意两个正整数 n 和 m 都满足 $a_{n+m} \leqslant a_n + a_m$，求证: 数列 $\{\frac{a_n}{n}\}$ 收敛。 ■

题目出处 http://kuing.orzweb.net/viewthread.php?tid=4633

解答日期 2017-05-24

这道题目跟 1997 年的一道 CMO 试题的条件一模一样，只是结论不同，题目如下: 已知非负实数数列 $\{a_n\}$ 对任意两个正整数 n 和 m 都满足 $a_{n+m} \leqslant a_n + a_m$，求证: 对任意 $n \geqslant m$ 都有 $a_n \leqslant ma_1 + \left(\frac{n}{m} - 1\right) a_m$。

由于这个关联性，所以原题目的解答过程参考了后者的思路[6].

证明 设 $n > m$，有

$$\begin{aligned}\frac{a_n}{n} - \frac{a_m}{m} <& \frac{a_{n-m} + a_m}{n} - \frac{a_m}{m} \\ =& \frac{n-m}{n}\left(\frac{a_{n-m}}{n-m} - \frac{a_m}{m}\right)\end{aligned}$$

这有点类似于辗转相除法，如果还有 $n - m > m$，则继续上述步骤，经过有限步之后，必然得出

$$\frac{a_n}{n} - \frac{a_m}{m} \leqslant \frac{s}{n}\left(\frac{a_s}{s} - \frac{a_m}{m}\right)$$

其中正整数 s 满足 $1 \leqslant s \leqslant m$，但是这个 s，一般的说是依赖于 n 的，但是它的取值集合却是有限个，所以右端除去因子 $\frac{1}{n}$ 以外的部分是有界的, 所以我们在上式中令 $n \to \infty$，便能得出右端在 $n \to \infty$ 时趋于零，这就说明: 对于每一个固定的 m 和任意小的正实数 δ，当 n 充分大时恒有

$$\frac{a_n}{n} \leqslant \frac{a_m}{m} + \delta$$

[6] 见参考文献 [?].

这便是我们解答问题的关键所在,因为显然还能得出 $a_n \leqslant a_{n-1}+a_1 \leqslant (a_{n-2}+a_1)+a_1 \leqslant \cdots \leqslant na_1$，从而 $0 \leqslant \frac{a_n}{n} \leqslant a_1$ 说明数列 $\{\frac{a_n}{n}\}$ 是有上下界的，因此我们只要证明如下这个引理就可以了：

引理: 如果数列 $\{x_n\}$ 有界，并且对于数列中的每一项 x_m 和任意小的正实数 δ，当 n 充分大时恒有 $x_n \leqslant x_m+\delta$，那么数列收敛。

证明是很容易的，既然数列有下界，便有下确界，我们将证明这下确界便是其极限，设其下确界为 M，那么对于无论多么小的正实数 ε，存在数列中的某一项 x_m 满足 $M \leqslant x_m < M+\varepsilon$，然而由条件，当 n 充分大时恒有 $x_n \leqslant x_m+\delta$，我们让这个 $\delta < M+\varepsilon-x_m$，从而这时就有 $M \leqslant x_n < M+\varepsilon$，按极限定义，这下确界 M 便是其极限。 □

题 1.2.2 已知 $m+1$ 个实数 $a_i(i=0,1,\ldots,m)$ 满足

$$a_0+a_1+\cdots+a_m=0$$

求证

$$\lim_{n\to\infty}(a_0\sqrt{n}+a_1\sqrt{n+1}+\cdots a_m\sqrt{n+m})=0$$

■

题目出处 http://kuing.orzweb.net/viewthread.php?tid=4640

解答日期 2017-05-27

证明 由条件有

$$a_m=-(a_0+a_1+\cdots+a_{m-1})$$

所以

$$\begin{aligned}
&\sum_{i=0}^{m}a_i\sqrt{n+i}\\
=&\sum_{i=0}^{m-1}a_i(\sqrt{n+i}-\sqrt{n+m})\\
=&-\sum_{i=0}^{m-1}\frac{(m-i)a_i}{\sqrt{n+i}+\sqrt{n+m}}
\end{aligned}$$

上式每一个项的极限都是零，且项数 m 不随 n 变化，故原式极限为零。 □

1.3 函数的极限

1.3.1 趋无穷的极限

与数列的极限有所区别，函数的极限过程有两大类，一类是自变量趋于正无穷或者负无穷的极限，一类是自变量趋于某个固定点的极限。

定义 1.3.1. 设函数 $f(x)$ 在无穷区间 $(a,+\infty)$ 上有定义，A 是一个实数，如果对于任意小的正实数 ε，总存在实数 $X(>a)$，使得 $x>X$ 时恒有 $|f(x)-A|<\varepsilon$ 成立，则称数 A 是函数 $f(x)$ 在自变量 x 趋于正无穷大时的 极限，记作:

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

类似的可以得函数当自变量趋于负无穷大的极限定义，并且如果函数当自变量趋于正无穷大和负无穷大时都有极限而且极限相同，则称函数当自变量趋于无穷大时有极限，这也可以从绝对值来定义而不考虑自变量的符号。

1.3.2 趋点极限与单侧极限

当自变量趋于某点的极限定义如下:

定义 1.3.2. 设函数 $f(x)$ 在 x_0 的某空心邻域内有定义，A 是一个实数，如果对于任意小的正实数 $\varepsilon>0$，总存在另一正实数 $\delta>0$，使得定义域中满足 $0<|x-x_0|<\delta$ 的数 x 恒有 $|f(x)-A|<\varepsilon$，则称 A 是函数 $f(x)$ 当自变量趋于 x_0 时的极限，记作

$$\lim_{x\to x_0} f(x)=A$$

要指出的是，1. 定义中限定了 $x\neq x_0$ 是因为函数在该点处的情况与该点处是否存在极限无任何关系. 2. 函数 $f(x)$ 在自变量趋于 x_0 时即使收敛，其极限值也并不一定等于 $f(x_0)$，实际上函数在 x_0 也并不一定有定义。

考虑到 x 趋于 x_0 的方式，它可以从小于 x_0 的一侧去靠近它，也可以从大于 x_0 的一侧去靠近它，也可以时而在大于 x_0 的一侧，时而位于小于 x_0 的一侧的方式去接近它，所以在这里我们给出 单侧极限的概念。

定义 1.3.3. 如果函数 $f(x)$ 在 x_0 的某个右空心邻域内有定义，A 是一个实数，如果对于任意小的正实数 $\varepsilon>0$，都存在另一个正实数 $\delta>0$，使得当 $x_0<x<x_0+\delta$ 时恒有 $|f(x)-A|<\varepsilon$ 成立，则称 A 是函数 $f(x)$ 在 x_0 处的 右极限，记作

$$\lim_{x\to x_0^+} f(x)=A$$

类似的，把右空心邻域改为左空心邻域，把不等式 $x_0<x<x_0+\delta$ 换成 $x_0-\delta<x<x_0$，就可以得到 左极限的定义，左极限记作:

$$\lim_{x\to x_0^-} f(x)=A$$

显然，$\lim_{x\to x_0} f(x)=A$ 的充分必要条件是 $\lim_{x\to x_0^+} f(x)=\lim_{x\to x_0^-} f(x)=A$.

例 1.3.1 设 $n\in\mathbb{Z}$，我们来证明

$$\lim_{x\to x_0} x^n=x_0^n$$

由

$$x^n-x_0^n=(x-x_0)(x^{n-1}+x^{n-2}x_0+\cdots+xx_0^{n-2}+x_0^{n-1})$$

限定 $x \in (x_0-\delta, x_0+\delta)$，令 $M=\max\{|x_0-\delta|, |x_0+\delta|\}$，则

$$|x^n - x_0^n| < nM^{n-1}|x-x_0|$$

注意这里的 nM^{n-1} 是常数，以 P 表之，则对于任意小的正实数 ε，只要取 $\delta = \dfrac{\varepsilon}{P}$，就能保证 $|f(x)-f(x_0)| < \varepsilon$ 对任意 $x \in (x_0-\delta, x_0) \cup (x_0, x_0+\delta)$ 恒成立，即得结论. ■

例 1.3.2 考虑函数 $f(x) = \sin\dfrac{1}{x}$ 在 $x \to 0$ 时的极限情况，此时函数在-1 与 1 之间完成无穷次振动，它的函数值不可能无限靠近任意一个实数，所以它在 $x=0$ 处不存在极限. 但若将它稍加改造，考虑函数 $f(x) = x\sin\dfrac{1}{x}$，则此时它的图象将被限定在 $y=x$ 和 $y=-x$ 之间并且在靠近零时，无限的在两个函数之间振动，此时显然有 $|x\sin\dfrac{1}{x}| \leqslant |x|$，所以显然就有

$$\lim_{x\to 0} x\sin\frac{1}{x} = 0$$

■

例 1.3.3 狄利克雷 (Dirichlet，德国数学家) 函数 $D(x)$ 是一个指示一个实数是有理数还是无理数的标志函数，它定义在全体实数上，当 x 是有理数时，$D(x)=1$，当 x 是无理数时，$D(x)=0$，即

$$D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

现在来证明它在任意 $x_0 \in \mathbb{R}$ 处的极限情况，当 x 取有理数并趋于 x_0 时，函数值趋于 1，但当 x 取无理数趋于 x_0 时，函数值趋于零，因而函数在 x_0 处没有极限，即在任意点处都不存在极限。 ■

例 1.3.4 高斯函数是一种取整函数，它通常用专门的记号 $[x]$ 来表示，在数论中有着广泛的应用，它的定义是，对任意实数 x，函数值是不超过 x 的最大整数，例如，$[2.3]=2$,$[\pi]=3$,$[-4.5]=-5$，现在考虑一下高斯函数在自变量取整数处的极限情况。

设 $x_0 \in \mathbb{Z}$，那么对于它的邻域 (x_0-1, x_0+1) 来说，右邻域 (x_0, x_0+1) 上的函数值都是 x_0，但左邻域 (x_0-1, x_0) 上的函数值都是 x_0-1，所以有

$$\lim_{x\to x_0^+}[x] = x_0,\quad \lim_{x\to x_0^-}[x] = x_0-1$$

但是 $\lim\limits_{x\to x_0}[x]$ 不存在. ■

例 1.3.5 黎曼 (Riemann，德国数学家) 函数 $R(x)$ 是一个有趣的例子，它定义在区间 $[0,1]$ 上，具体定义如下，

$$R(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \in [0,1], p, q \in \mathbb{Z}, q > 0, (p,q) = 1 \\ 0 & x \in [0,1] - \mathbb{Q} \end{cases}$$

简单来说，当 x 是有理数 $\dfrac{p}{q}$ 时，$R(x) = \dfrac{1}{q}$，注意这里 p、q 是一对既约整数（即最大公因数为 1），且 q 是正的。而当 x 是无理数时，$R(x)=0$.

黎曼函数可以看成是狄利克雷函数的一个改进，我们现在来证明，黎曼函数在任意点处都有极限为零。

任取数 $x_0 \in [0,1]$，只需要证明 $\lim\limits_{x \to x_0} R(x) = 0$ 就可以了，对于任意小的正实数 $\varepsilon > 0$，我们要找一个 $\delta > 0$，使得 $\forall x \in (x_0-\delta, x_0) \cup (x_0, x_0+\delta)$ 都有 $|R(x)| < \varepsilon$ 成立，显然无理数是无关紧要的，这个 δ 只需要由有理数来决定，当 x 是有理数 $\dfrac{p}{q}(q > 0)$ 时，我们就是要使

$$\frac{1}{q} < \varepsilon$$

成立，这就需要让 $(x_0-\delta, x_0) \cup (x_0, x_0+\delta)$ 上的一切有理数 $\dfrac{p}{q}$ 都满足 $q > \dfrac{1}{\varepsilon}$，这能否做到呢？是可以的，因为不满足这个条件的有理数是有限的（小于 $\dfrac{1}{\varepsilon}$ 的正整数 q 是有限的，而小于这些 q 的非负整数 p 也是有限的），从而把这些有理数记为 $x_1, x_2, \ldots, x_m$(注意这里的 m 跟 ε 有关)，则只要让 δ 同时满足 $\delta < |x_i - x_0| (i = 1, 2, \ldots, m)$，则这些有理数都不会落在范围 $(x_0-\delta, x_0) \cup (x_0, x_0+\delta)$ 内，从而对于该范围内的一切有理数，前面所需要的条件就成立了，于是函数在 x_0 处有极限为零。 ■

1.3.3 函数极限的性质

与数列极限的性质相仿，函数极限具有类似的性质，以下定理都以 $x \to x_0$ 为例，但它们对于自变量趋于无穷大时的极限也是成立的，由于与数列极限的相应结论对应，这里就省去了证明。

性质 1.3.1 (唯一性). *函数极限 $\lim_{x \to x_0} f(x)$ 若存在必唯一.*

定理 1.3.1 (局部有界性). *设函数 $f(x)$ 在 x_0 的某空心邻域内有定义，若 $\lim_{x \to x_0} f(x)$ 存在 (非无穷的有限值)，则 $f(x)$ 在 x_0 的某个空心邻域内有界。*

定理 1.3.2 (局部保号性). *若函数 $f(x)$ 在 $x \to x_0$ 处的极限存在为 A，则对于任意 $r < A$ 都存在 x_0 的某个空心邻域内，在这邻域内恒有 $f(x) > r$，同样，对于任意 $r > A$，都存在 x_0 的某空心邻域，在这邻域内恒有 $f(x) < r$.*

定理 1.3.3 (保不等式性). *如果函数 $f(x)$ 和 $g(x)$ 都在 x_0 的某个空心邻内有定义，且在这邻域内上恒满足 $f(x) \geqslant g(x)$，那么如果当 $x \to x_0$ 时两个函数分别有极限 A 和 B，则必有 $A \geqslant B$.*

定理 1.3.4 (夹逼定理). *在 x_0 的某空心邻域内有定义的三个函数 $f(x)$、$g(x)$ 和 $h(x)$，如果在这邻域内恒满足 $f(x) \leqslant h(x) \leqslant g(x)$，那么如果当 $x \to x_0$ 时 $f(x)$ 和 $g(x)$ 都收敛到 A，那么这时 $h(x)$ 也必收敛，且也收敛到 A.*

定理 1.3.5 (四则运算法则). *如果在 x_0 的某空心邻域内有定义的二函数 $f(x)$ 和 $g(x)$ 在 $x \to x_0$ 时分别收敛到 A 和 B，那么这由两个函数的和、差、积、商作成的新函数也收敛，并分别收敛到原先两个极限值和和、差、积、商，在商的情况下，要求分母不为零.*

写成公式就是，在 $\lim_{x\to x_0} f(x)$ 和 $\lim_{x\to x_0} g(x)$ 都存在的前提下，有

$$\begin{aligned}\lim_{x\to x_0}(f(x)\pm g(x)) &= \lim_{x\to x_0} f(x)\pm\lim_{x\to x_0} g(x)\\ \lim_{x\to x_0} f(x)g(x) &= \lim_{x\to x_0} f(x)\cdot\lim_{x\to x_0} g(x)\\ \lim_{x\to x_0}\frac{f(x)}{g(x)} &= \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}\end{aligned}$$

例 1.3.6 设多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0(a_n\neq 0)$$

现在考虑当 $x\to\infty$、$x\to+\infty$、$x\to-\infty$ 时的极限，无论是哪一种趋限，都有 $|x|\to\infty$，为此将多项式改写为

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

显然，当 $|x|\to\infty$ 时，括号中带分式的项全部趋于零，因而此时的极限全部取决于 a_nx^n，因此结论是，当 $|x|\to\infty$ 时，$f(x)$ 为无穷大，如果指定 x 趋于正无穷或者负无穷，则相应的极限也是带符号的无穷大，其符号由 a_n 的符号决定。 ■

例 1.3.7 再考虑一般的有理函数

$$f(x)=\frac{a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0}{b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0}$$

考虑它在无穷远处的极限，通过与上例类似的方法，可以得到，如果 $m=n$，则极限为 $\dfrac{a_n}{b_n}$，若 $m>n$，则极限为零，若 $m<n$，则极限为无穷大. ■

例 1.3.8 我们来建立一个有用的极限，设 r 为一个有理数，则有

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

$r=0$ 时结论显然成立，先来证明 r 是正整数的情况，由二项式定理有

$$\frac{(1+x)^n-1}{x}=\frac{C_n^1x+C_n^2x^2+\cdots+C_n^nx^n}{x}=C_n^1+C_n^2x+\cdots+C_n^nx^{n-1}$$

所以得到

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

而对于 r 为负整数的情形，有

$$\frac{(1+x)^{-n}-1}{x}=\frac{1-(1+x)^n}{x}\cdot\frac{1}{(1+x)^n}$$

令 $x\to 0$，上式极限即为 $-n$，所以结论对于 r 为一切整数的情形都成立.

再设 r 是有理数 $\dfrac{n}{m}$，令 $t=(1+x)^{n/m}$，由公式

$$a^m-b^m=(a-b)(a^{m-1}+a^{m-2}b+\cdots+ab^{m-2}+b^{m-1})$$

得

$$\frac{(1+x)^{n/m}-1}{x}=\frac{t-1}{x}=\frac{t^m-1}{x(1+t+\cdots+t^{m-1})}=\frac{(1+x)^n-1}{x}\cdot\frac{1}{1+t+\cdots+t^{m-1}}$$

当 $x\to 0$ 时，第一个因子趋于 n，第二个因子由于 $t\to 1$ 而趋于 m，因此整体趋于极限 $\dfrac{n}{m}$，即 r，得证。

这就是说，在 $x\to 0$ 时，有 $(1+x)^r=1+rx+o(x)$.

在以后，我们将把这结论推广到 r 为无理数的情形，从而对于指数为一切实数，结论都是成立的. ■

1.3.4 复合函数的极限

关于复合函数的极限，有如下结论

定理 1.3.6. *设有如下函数极限*

$$\lim_{x\to x_0} g(x)=u_0,\ \lim_{u\to u_0} f(u)=y_0$$

且 $g(x)$ 在 x_0 的某空心邻域上恒有 $g(x)\neq u_0$，则有

$$\lim_{x\to x_0} f(g(x))=y_0$$

证明 因为当 $u\to u_0$ 时，$f(u)\to y_0$，所以对于无论多么小的正实数 ε，总存在另一正实数 δ，使得当 $0<|u-u_0|<\delta$ 时恒有 $|f(u)-y_0|<\varepsilon$，而又由于当 $x\to x_0$ 时 $g(x)\to u_0$，所以对于前面提到的 δ，存在另一正实数 r，使得当 $0<|x-x_0|<r$ 时恒有 $|g(x)-u_0|<\delta$，而 $g(x)$ 又在 x_0 的某个半径为 r' 的空心邻域上函数值恒不为 u_0，于是 $\forall x, 0<|x-x_0|<\min\{r,r'\}$，有 $0<|g(x)-u_0|<\delta$，从而 $|f(g(x))-y_0|<\varepsilon$，所以最终得到的结论就是：对于无论多么小的正实数 ε，总存在另一正实数 $r_1=\min\{r,r'\}$，使得当 $0<|x-x_0|<r_1$ 时，恒有 $|f(g(x))-y_0|<\varepsilon$，这就证得结论。 □

要说明的是，这个定理中的条件是充分条件，但不是必要条件，考察下面这个例子就知道了:

$$f(x)=\begin{cases} x+1 & x\neq 0,1\\ 2 & x=0\\ 1 & x=1\end{cases}$$

与

$$g(x)=\begin{cases} x-1 & x\neq 1,2\\ 1 & x=1\\ 0 & x=2\end{cases}$$

显然有 $f(g(x))=x$ 在 $x=1,2$ 处都连续，但显然不满足定理条件。

另外一点是，定理中 $g(x)$ 在 x_0 的某个空心邻域内的函数值恒不取 u_0 这个条件是不能缺少的，考察下面这个例子

$$g(x)=\begin{cases} \frac{1}{q} & x=\frac{p}{q}, p,q\in\mathbb{Z}, q>0\\ 0 & x\in\mathbb{R}-\mathbb{Q}\end{cases}$$

及

$$f(x)=\begin{cases}0 & x\neq 0\\ 1 & x=0\end{cases}$$

在 $x=0$ 处的情况，$g(x)$ 即黎曼函数，我们已经知道它在任意无理点处的极限为零，在任意有理点处不存在极限。

1.3.5 与数列极限的关系

以趋点极限为例，函数极限与数列极限之间有如下结论:

定理 1.3.7 (函数极限与数列极限的关系)**.** *在 x_0 的某空心邻域内有定义的函数 $f(x)$，当 $x\to x_0$ 时存在极限 (记为 A) 的充分必要条件是，对于任意一个在这空心邻域内取值并以 x_0 为极限的数列 x_n，$\lim_{n\to\infty} f(x_n)$ 都存在而且都等于 A。*

证明 先证必要性，如果 $\lim_{x\to x_0} f(x)=A$，那么对于任意小的正实数 $\varepsilon>0$，都存在另一个正实数 $\delta>0$，使得对这邻域内满足 $|x-x_0|<\delta$ 的实数 x 都成立不等式 $|f(x)-A|<\varepsilon$，那么对于任意一个也在这空心邻域内取值并以 x_0 为极限的数列 x_n，因为它以 x_0 为极限，所以对于这个 $\delta>0$，就必然能够从某一项 x_N 开始，后面的所有项都满足 $|x_n-x_0|<\delta$，于是就有 $|f(x_n)-A|<\varepsilon$，这就表明 $f(x_n)$ 当 $n\to\infty$ 时以 A 为极限，必要性得证。

再证充分性，如果对于任意一个在这空心邻域内取值并收敛到 x_0 的数列 x_n，对应的函数值数列 $f(x_n)$ 都收敛到同一实数 A，我们将证明，函数 $f(x)$ 在 $x\to x_0$ 时也必将收敛到 A. 采用反证法，假使函数 $f(x)$ 当 $x\to x_0$ 时不以数 A 为极限，那么必然存在某个 $\varepsilon_0>0$，使得无论把另一个正实数 $\delta>0$ 限制得多么小，总有满足 $|x-x_0|<\delta$ 的实数 x 能够使得 $|f(x)-A|\geqslant\varepsilon_0$ 成立，于是先取 $\delta=1$，得出一个符合这条件的实数 x_1，然而取 $\delta=\min\{\frac{1}{2},|x_1-x_0|\}>0$，又可以选出 x_2，依次这样下去，逐个令 $\delta_n=\min\{\frac{1}{n},|x_{n-1}-x_0|\}$，就可以挑选出 x_{n+1}，这样就作出一个数列 x_n，由 $|x_n-x_0|<\delta_n<\frac{1}{n}$ 可知 x_n 收敛到 x_0，但是由于 $|f(x_n)-A|\geqslant\varepsilon$ 恒成立，可知数列 $f(x_n)$ 并不收敛到 A，这样，我们就证明了如果函数 $f(x)$ 当 $x\to x_0$ 时不以 A 为极限，那么就可以构造出一个以 x_0 为极限的数列 x_n，使得 $f(x_n)$ 也不以 A 为极限，这与我们的条件是矛盾的，所以充分性得证。 □

事实上，如果任意以 x_0 为极限的数列 x_n，函数值数列 $f(x_n)$ 都收敛的话，这些极限值也必然相同，这是因为，如若不然，假如两个数列 x_n 和 r_n 分别以 A 和 B 为极限，那么在这两个数列中交错的取项构成另一数列 s_n，显然 s_n 也以 x_0 为极限，而函数值数列 $f(s_n)$ 中的奇数下标子列和偶数下标子列分别以 A 和 B 为极限，由条件知 $f(s_n)$ 应有极限，所以 $A=B$. 有了这结论，上述定理中的条件可以适当减弱。

1.3.6 单调有界定理

与数列的单调有界定理相仿，我们有以下定理

定理 1.3.8. *如果函数 $f(x)$ 在 x_0 的某左空心邻域内单调递增且有上界，则函数 $f(x)$ 在 x_0 处的左极限存在，右极限也有类似的结论。*

证明 证明很简单，只要在这左邻域内任取一单调增加并以 x_0 为极限的数列 x_n，则函数值数列 $f(x_n)$ 亦必是单调增加的数列，而它又有上界，所以它有极限，设这极限为 A，则易证 A 便是函数在这左空心邻域内的上确界，那么对于无论多么小的正实数 ε，都存在正整数 N，使得当 $n > N$ 时恒有 $|f(x_n) - A| < \varepsilon$ 成立，于是取 $\delta = x_0 - x_{N+1} > 0$，则对于任意满足 $x_0 - \delta < x < x_0$ 的实数 x，有 $x_N + 1 < x < x_0$，因而 $A - \varepsilon < f(x_{N+1}) < f(x) \leqslant A$，这表明 A 就是 $f(x)$ 在 x_0 处的左极限。 □

1.3.7 柯西收敛准则

仿照数列的柯西收敛准则，有

定理 1.3.9. *函数 $f(x)$ 在 x_0 的某空心邻域内有定义，则它在该存在极限的充分必要条件是: 任给无论多么小的正实数 ε，恒存在另一正实数 δ，使得对任意满足 $|x - x_0| < \delta$ 的两个实数 x_1 和 x_2 都成立 $|f(x_1) - f(x_2)| < \varepsilon$.*

证明 必要性是容易证明的，略去，下证充分性，取正实数数列 $\varepsilon_n = 1/n(n = 1, 2, \ldots)$，存在另一单调递减的正实数数列 δ_n，使得对于任意满足 $|x_1 - x_0| < \delta_n$ 和 $|x_2 - x_0| < \delta_n$ 的 x_1, x_2 都成立 $|f(x_1) - f(x_2)| < \varepsilon_n$，于是函数 $f(x)$ 在 x_0 的以 δ_n 为半径的空心邻域内的函数值都必将限于一个长度为 ε 的闭区间 $[m_n, M_n]$ 上，显然这个闭区间序列符合闭区间套定理的条件，所以有唯一实数 A 从属于所有的闭区间，显然这实数 A 也是函数 $f(x)$ 在 x_0 处的极限。 □

1.3.8 两个重要的函数极限

这一节的主要任务是建立两个重要极限.

定理 1.3.10. *对于正弦函数，有如下极限*

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

证明 如图 1.1，在单位圆中，点 A 和 C 是圆周上两点, 并且 $\angle AOC$ 的弧度值为 x，x 是锐角，OC 与点 A 处的切线相交于点 B, 由图可得

$$S_{\triangle AOC} < S_{\text{扇形}AOC} < S_{\triangle AOB}$$

于是便得

$$\sin x < x < \tan x$$

从而

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

而当 $x \to 0$ 时, $\cos x \to 1$，所以最终得极限

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

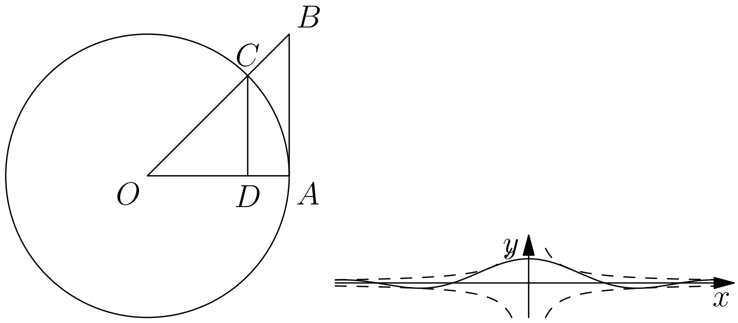

图 1.1

□

函数 $f(x) = \sin x/x$ 在 $x = 0$ 附近的图象如图 1.1所示，它的图象反得在函数 $y = 1/x$ 和 $y = -1/x$ 之间振荡，在 $x = 0$ 附近，函数值趋于 1，但它在 $x = 0$ 处并无定义。

定理表明，在 $x \to 0$ 时，正弦函数 $\sin x$ 与 x 是等价无穷小，即 $\sin x = x + o(x)$.

例 1.3.9 关于余弦函数 $\cos x$ 在 $x \to 0$，也有如下结果

$$\lim_{x\to 0}\frac{1-\cos x}{x^2} = \frac{1}{2}$$

这是由于

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

取极限并利用平方函数的连续性即得结论，于是就有，在 $x \to 0$ 时，有 $\cos x = 1 - \frac{1}{2}x^2 + o(x^2)$. ■

例 1.3.10 对于正切函数，我们有

$$\lim_{x\to 0}\frac{\tan x}{x} = 1$$

这从 $\tan x = \dfrac{\sin x}{\cos x}$ 便可以得出，于是在 $x \to 0$ 时，有 $\tan x = x + o(x)$. ■

例 1.3.11 我们来证明正弦函数 $f(x) = \sin x$ 在任意 $x_0 \in \mathbb{R}$ 处的极限是 $\sin x_0$，即

$$\lim_{x\to x_0}\sin x = \sin x_0$$

因为

$$|\sin x - \sin x_0| = 2\left|\cos\frac{x+x_0}{2}\sin\frac{x-x_0}{2}\right| \leqslant 2\left|\sin\frac{x-x_0}{2}\right|$$

根据我们已经知道的不等式,$\forall x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 有 $|\sin x| \leqslant |x|$,所以只要限定 $|x - x_0| < \frac{\pi}{2}$，就有

$$|\sin x - \sin x_0| \leqslant |x - x_0|$$

于是对任意正实数 $\varepsilon > 0$，只要取 $0 < \delta < \min\{\varepsilon, \frac{\pi}{2}\}$，就能保证 $\forall x \in (x_0 + \delta, x_0 + \delta)$ 都有

$$|\sin x - \sin x_0| < \varepsilon$$

成立，即证明了结论。

类似的，对于余弦函数，也可以证明

$$\lim_{x \to x_0} \cos x = \cos x_0$$

■

例 1.3.12 我们来证明下面的两个极限，这个极限在后面学习导数的时候将会用到

$$\lim_{x \to x_0} \frac{\sin x - \sin x_0}{x - x_0} = \cos x_0$$

以及

$$\lim_{x \to x_0} \frac{\cos x - \cos x_0}{x - x_0} = -\sin x_0$$

只证明第一个，第二个也是类似的，固定 x_0，有

$$\frac{\sin x - \sin x_0}{x - x_0} = 2\cos\frac{x + x_0}{2} \cdot \frac{\sin\frac{x - x_0}{2}}{x - x_0}$$

显然当 $x \to x_0$ 时，后一个因式趋于 $\dfrac{1}{2}$，而第一个因式我们已经证明了它趋于 $\cos x_0$，因而便得结论。 ■

我们在数列极限的部分曾经证明过下面这个数列存在极限

$$x_n = \left(1 + \frac{1}{n}\right)^n$$

并把它的极限记作 e，今来把它推广成函数的极限，这就是以下的定理:

定理 1.3.11.

$$\lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = e$$

这里 e 是自然对数的底数.

证明 设实数 x 的整数部分为 n，即 $n \leqslant x < n + 1$，有

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

显见左右两边作为两个数列，都以 e 为极限，我们把这左右两边转化为两个函数，即设 x 的整数部分为 n，定义

$$h_1(x) = \left(1 + \frac{1}{n+1}\right)^n,\ h_2(x) = \left(1 + \frac{1}{n}\right)^{n+1}$$

于是就有

$$\lim_{x \to +\infty} h_1(x) = \lim_{x \to +\infty} h_2(x) = e$$

所以由夹逼准则即得

$$\lim_{x \to +\infty} \left(1 + \frac{1}{x}\right)^x = e$$

这是函数在正无穷远处的极限，再来看负无穷处的极限，当 x 以负值趋于负无穷时，令 $x=-t$，则 x 趋于负无穷等价于 t 趋于正无穷，而

$$\left(1+\frac{1}{x}\right)^x = \frac{1}{\left(1-\frac{1}{t}\right)^t} = \left(1+\frac{1}{t-1}\right)^t = \left(1+\frac{1}{t-1}\right)^{(t-1)\cdot\frac{t}{t-1}}$$

利用复合函数的极限结果 (定理 1.3.6)，便知当式右端当 $t\to+\infty$ 时的极限是 e，所以当 $x\to-\infty$ 时左端的极限也就是 e，所以最终当 $x\to\infty$ 时，函数 $(1+1/x)^x$ 的极限都是 e. □

在此基础上，我们有更为一般性的结论:

定理 1.3.12. *设函数 $f(x)$ 在某一极限过程中为无穷小，那么在同样的极限过程中有*

$$\lim(1+f(x))^{1/f(x)} = \mathrm{e}$$

证明 以 $x\to\infty$ 为例，因为 $\lim\limits_{x\to\infty}(1+\dfrac{1}{x})^x=\mathrm{e}$，所以对于无论多么小的正实数 ε，都存在正实数 M，使得满足 $|x|>M$ 的一切实数 x 都符合 $|(1+\dfrac{1}{x})^x-\mathrm{e}|<\varepsilon$，又因为 $f(x)$ 为无穷小，所以对于这个正实数 M，存在另一正实数 K，使得当 $|x|>K$ 时，有 $|f(x)|<\dfrac{1}{M}$ 成立，于是 $\left|\frac{1}{f(x)}\right|>M$，从而就有

$$\left|(1+f(x))^{1/f(x)}-\mathrm{e}\right|<\varepsilon$$

这就证得结论。 □

1.3.9 无穷小与无穷大

与数列类似，我们可以在函数极限中引入无穷小与无穷大的定义，如果函数在某一极限过程中以零为极限，则称其在此极限过程中是无穷小，同样，如果它在这一极限过程中其绝对值能够大于任意的正数，则称其是一个无穷大，类似的可以得到正无穷大和负无穷大的概念。函数极限过程中的无穷小和无穷大有着与无穷小数列和无穷大数列类似的性质。

与数列一样，函数极限过程中的无穷小与无穷大也可以引入阶的比较概念。

1.3.10 曲线的渐近线

1.4 函数的连续性

在初等数学中，我们接触到了很多函数，例如幂函数 $y=x^n(n\in\mathbb{Z})$，三角函数与反三角函数、指数函数与对数函数。它们的图象都是一段连续的曲线，或者是由若干段连续的曲线所构成，在这一节，我们将对“连续”这个概念作一个精确的定义，并进而讨论连续函数的性质。

1.4.1 连续与单侧连续

定义 1.4.1. 如果函数 $f(x)$ 在 x_0 处存在极限，且极限值正好是该点处的函数值 $f(x_0)$，则称函数在 x_0 处 连续，即

$$\lim_{x\to x_0} f(x) = f(x_0)$$

如果是该点处的左极限等于该点处函数值，则称函数在该点处 左连续，类似的有右连续的概念。

连续用极限的精确语言描述就是，对于无论多么小的正实数 ε，恒存在另一正实数 δ，使得对区间 $(x_0-\delta, x_0+\delta)$ 上的一切实数成立着 $|f(x)-f(x_0)|<\varepsilon$ 成立。

如果函数在某点处不连续，则称该点是函数的一个 间断点。

定义 1.4.2. 如果函数在某个区间上处处连续，则称函数在这区间上连续，或者说它是这区间上的连续函数。

例 1.4.1 在 例 1.3.11中，我们证明了

$$\lim_{x\to x_0} \sin x = \sin x_0, \ \lim_{x\to x_0} \cos x = \cos x_0$$

因此正弦函数和余弦函数都是 $\mathbb{R}$ 上的连续函数. ■

例 1.4.2 从 例 1.3.4的讨论得知，对于高斯函数，设 $x_0 \in \mathbb{Z}$，有

$$\lim_{x\to x_0^+} [x] = x_0, \ \lim_{x\to x_0^-} [x] = x_0 - 1$$

因为 $[x_0] = x_0$，所以高斯函数在每个整数处右连续，但不左连续，所以每个整点都是间断点，易知除整点外的其余点处都连续。 ■

例 1.4.3 在 例 1.3.2中，我们知道了下面的极限

$$\lim_{x\to 0} x \sin\frac{1}{x} = 0$$

但由于函数在 $x=0$ 处没有定义，故而不连续，但我们可以补充定义它在该点处的函数值，即

$$f(x) = \begin{cases} x\sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

那么 $f(x)$ 就在 $x=0$ 处连续了。 ■

例 1.4.4 我们举一个定义在 $\mathbb{R}$ 上但处处不连续的函数的例子，狄利克雷 (Dirichlet，德国数学家) 函数 $D(x)$，这是一个指示一个实数是有理数还是无理数的标志函数，它定义在全体实数上，当 x 是有理数时，$D(x)=1$，当 x 是无理数时，$D(x)=0$，即

$$D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

现在来证明它在 R 上处处不连续，即在任意 $x_0 \in R$ 处都不连续，这是因为，当 x 取有理数并趋于 x_0 时，函数值趋于 1，但当 x 取无理数趋于 x_0 时，函数值趋于零，因而函数在 x_0 处没有极限，从而也不连续。 ■

例 1.4.5 在狄利克雷函数的基础上，我们可以构造出一个定义在 $\mathbb{R}$ 上，但仅在一个点处连续的函数，定义函数 $f(x) = xD(x)$，我们来证明，它仅在 $x = 0$ 处连续。

先证明它在 $x = 0$ 处连续，因为当 $x \to 0$ 时，$x \to 0$ 为无穷小，而 $D(x)$ 是有界量，无穷小与有界量的乘积仍是无穷小，所以 $\lim\limits_{x\to 0} f(x) = 0 = f(0)$，所以函数在 $x = 0$ 处连续。

再证明它在任意 $x_0 \neq 0$ 处都不连续，与狄利克雷函数相仿，当 x 取有理数值并趋于 x_0 时，函数值将趋于 x_0，但当 x 取无理数值并趋于 x_0 时，函数值则趋于 0，但 $x_0 \neq 0$，所以函数在 x_0 处不存在极限，更不连续。 ■

例 1.4.6 黎曼 (Riemann，德国数学家) 函数 $R(x)$ 是一个有趣的例子，它定义在区间 $[0,1]$ 上，在任意无理数处都连续，但在任意有理数处不连续，它的定义如下，

$$R(x) = \begin{cases} \frac{1}{q} & x = \frac{p}{q} \in [0,1], p, q \in \mathbb{Z}, q > 0, (p,q) = 1 \\ 0 & x \in [0,1] - \mathbb{Q} \end{cases}$$

简单来说，当 x 是有理数 $\dfrac{p}{q}$ 时，$R(x) = \dfrac{1}{q}$，注意这里 p、q 是一对既约整数（即最大公因数为 1），且 q 是正的。而当 x 是无理数时，$R(x) = 0.$

黎曼函数可以看成是狄利克雷函数的一个改进，我们现在来证明，黎曼函数在所有无理数处连续。

任取无理数 x_0，只需要证明 $\lim\limits_{x\to x_0} R(x) = 0$ 就可以了，对于任意小的正实数 $\varepsilon > 0$，我们要找一个 $\delta > 0$，使得 $\forall x \in (x_0 - \delta, x_0 + \delta)$ 都有 $|R(x)| < \varepsilon$ 成立，显然无理数是无关紧要的，这个 δ 只需要由有理数来决定，当 x 是有理数 $\dfrac{p}{q}(q > 0)$ 时，我们就是要使

$$\frac{1}{q} < \varepsilon$$

成立，这就需要让区间 $(x_0 - \delta, x_0 + \delta)$ 上的一切有理数 $\dfrac{p}{q}$ 都满足 $q > \dfrac{1}{\varepsilon}$，这能否做到呢？是可以的，因为不满足这个条件的有理数是有限的（小于 $\dfrac{1}{\varepsilon}$ 的正整数 q 是有限的，而小于这些 q 的非负整数 p 也是有限的），从而把这些有理数记为 $x_1, x_2, \ldots, x_m$(注意这里的 m 跟 ε 有关)，则只要让 δ 同时满足 $\delta < |x_i - x_0|(i = 1, 2, \ldots, m)$，则这些有理数都不会落在区间 $(x_0 - \delta, x_0 + \delta)$ 上，从而对于该区间上的一切有理数，前面所需要的条件就成立了，于是函数在 x_0 处连续，即在任意无理数处连续。

但对于 x_0 为有理数的情形，只要让 x 取无理数并趋于 x_0，则函数值将趋于 0 而非 $R(x_0)$，所以 $R(x)$ 在 x_0 处不存在极限，也就不连续。 ■

1.4.2 间断点及其分类

如果函数在某点处不连续但是存在极限，只是这极限与函数值不相等或者该点根本就没有定义函数值，那么称这点是 *可去间断点*，可以通过改变或者定义该点的函数值

为该点的极限值的方式来将函数进行 *连续开拓*。如果函数在某点处分别存在左极限和右极限，但是两个极限不相等，则称该点是函数的 *跳跃间断点*，跳跃间断点和可去间断点统称 *第一类间断点*，第一类间断点的特征是函数在该点存在两个方向的单侧极限。除第一类间断点之外的其它间断点统称 *第二类间断点*，显然，第二类间断点处至少有一个单侧极限不存在。

1.4.3 连续函数的性质

由连续的定义可见，函数在某点处连续，等价于它在该点处存在极限并且正好等于该点处的函数值，因此把函数极限的性质照搬过程，便得函数在某点处连续时所具有的性质:

定理 1.4.1 (局部有界性)**.** *若函数在某点处连续，则必在该点的某邻域上有界。*

定理 1.4.2 (局部保号性)**.** *若函数在 x_0 处连续，则对于任意小于 $f(x_0)$ 的实数 r，存在 x_0 的某邻域 $(x_0-\delta, x_0+\delta)$，使得函数在该邻域内恒有 $f(x)>r$，类似的，对于任意大于 $f(x_0)$ 的实数 r，也存在 x_0 的某邻域 $(x_0-\delta, x_0+\delta)$，使得函数在该区间上恒有 $f(x)<r$.*

推论 1.4.1. *如果函数在某点处连续，且该点处函数值为正，则存在该点的某邻域内，函数在这邻域内恒为正号，同理，如果该点函数值为负，则函数必在该点的某邻域内恒保持负号。*

定理 1.4.3. *如果函数 $f(x)$ 和 $g(x)$ 都在 x_0 处连续，则它们的和、差、积、商所作成的函数在该点也连续，在商的情形，要求 $g(x_0)\neq 0$。*

例 1.4.7 在**??** 中已经证明了指数函数 $f(x)=a^x(a>1)$ 在 $\mathbb{R}$ 上连续，这里来证明 $0<a<1$ 的情况下，指数函数同样是 $\mathbb{R}$ 上的连续函数，这是因为，在 $0<a<1$ 时，有

$$a^x=\frac{1}{\left(\frac{1}{a}\right)^x}$$

显然，分子分母均是 $\mathbb{R}$ 上的连续函数，从这一节的结果就证明了结论。 ■

1.4.4 复合函数的连续性

由复合函数极限存在的条件，可得

定理 1.4.4 (复合函数的连续性)**.** *设函数 $g(x)$ 在 x_0 处连续，记 $u_0=g(x_0)$，若另一函数 $f(u)$ 在 u_0 处连续，则复合函数 $f(g(x))$ 在 x_0 处连续。*

到现在，我们知道把某点处连续的函数进行四则混合运算，以及进行复合运算，所得到的新函数仍然在该点连续，由此我们可以迅速知道大批的连续函数。

1.4.5 有限覆盖定理

为了下一小节作准备，我们来证明一个非常深刻的关于实数连续性的定理，这个结果属于魏尔斯特拉斯。

定理 1.4.5 (魏尔斯特拉斯有限覆盖定理). *如果无穷多个开区间的并集覆盖了某个闭区间，则存在其中的有限个开区间，它们的并集就能够覆盖前述闭区间。*

证明 设开区间系 $U_i(i=1,2,\ldots)$ 的并集覆盖了闭区间 $[a,b]$，我们定义一个集合 A，如果实数 $x(a\leqslant x\leqslant b)$ 使得 $[a,x]$ 有有限覆盖，则 $x\in A$，显然 A 为有界集，因此有上确界，我们证明这个上确界必然是 b，设此上确界是 r，如果 $r<b$，则闭区间 $[a,r]$ 有有限覆盖，即能够选出有限个开区间覆盖它，把这些开区间再并上覆盖实数 r 的任一个开区间 $(r-\delta_1,r+\delta_2)$，就构成一个 $[a,r+\frac{\delta_2}{2}]$ 的一个有限覆盖，这表明 $r+\frac{\delta_2}{2}\in A$，这与 r 是 A 的上确界矛盾，所以 $r=b$，这就证明了定理. □

1.4.6 闭区间上的连续函数

定理 1.4.6. *闭区间上的连续函数在此区间上必定有界.*

证明 因为函数在闭区间上连续，所以对于此区间上任一点，函数都能在它的某个邻域上有界，显然这些邻域覆盖了整个闭区间，根据有限覆盖定理，能在其中选出有限个开区间，它们的并就能覆盖住这个闭区间，而函数在这有限个开区间中每一个上都是有界的，因此只要取最大的上界和最小的下界，便是它在这闭区间上的界。 □

定理 1.4.7. *闭区间上的连续函数存在最大值和最小值。*

证明 根据定理 1.4.6，函数在闭区间上连续则有界，从而便有上下确界，只要证明这两个确界能够取到就行了，只证明上确界可以取到，下确界也是同理的。设函数 $f(x)$ 在这闭区间上的上确界为 M，则存在收敛到 M 的函数值序列 $f(x_1),f(x_2),\cdots$，而自变量序列 $x_1,x_2,\cdots$ 是有界的，所以它存在收敛的子列，仍以 $x_1,x_2,\cdots$ 来表示它，设这子列的极限为 x_M，显然 x_M 仍然在这闭区间上，而且 $\lim\limits_{n\to\infty}f(x_n)=M$，则由连续性得有 $f(x_M)=f(\lim\limits_{n\to\infty}x_n)=\lim\limits_{n\to\infty}f(x_n)=M$. □

定理 1.4.8 (介值定理). *闭区间上的连续函数，其函数值能够遍历最大值与最小值之间的任意值。*

证明 根据前面的结果，这最大最小值都能够取到，不妨设 $f(a_1)=m$ 为最小值，$f(b_1)=M$ 为最大值，并且假设 $a_1\leqslant b_1$(反过来也是类似的)，实数 $m<\mu<M$，取闭区间 $[a_1,b_1]$ 的中点 $\frac{a_1+b_1}{2}$，如果 $f(\frac{a_1+b_1}{2})=\mu$ 则结论就成立了，否则如果 $\frac{a_1+b_1}{2}<\mu$ 就取后一半子区间 $[\frac{a_1+b_1}{2},b_1]$ 作为 $[a_2,b_2]$，如果 $f(\frac{a_1+b_1}{2})>\mu$ 则取前一半子区间 $[a_1,\frac{a_1+b_1}{2}]$ 作为 $[a_2,b_2]$，如此反复下去，最后的结果就是，或者能够在某个闭区间 $[a_n,b_n]$ 的中点处刚好有 $f(\frac{a_n+b_n}{2})=\mu$，或者能够得到一个闭区间序列 $[a_1,b_1],[a_2,b_2],\ldots,[a_n,b_n],\ldots$ 使得每一个闭区间都是前一个闭区间的一半并且两个端点的函数值分别大于和小于 μ，在后一种情形，显然这闭区间序列确定出一个数 x_0，使得 x_0 同时位于这所有的闭区间

上，那么必然有 $f(x_0)=\mu$，用反证法，如果 $f(x_0)>\mu$，则因为函数在 x 处连续，存在 x_0 的某个邻域 $(x_0-\delta,x_0+\delta)$，使得在此区间上恒有 $f(x)>\mu$，但这与前述闭区间序列显然矛盾，因为那闭区间序列充分靠后的闭区间必然落在此处的邻域内部，自然不可能在两个端点处的函数值分别大于和小于 μ，同理可以证明 $f(x_0)$ 也不可能小于 μ，所以 $f(x_0)=\mu$. □

推论 1.4.2 (零点定理). *若闭区间上连续函数在两个端点处函数值异号，则它在这区间上必存在至少一个零点。*

1.4.7 反函数的连续性

关于反函数的连续性，有如下结论:

定理 1.4.9. *如果函数 $f(x)$ 在闭区间 $[a,b]$ 上严格单调并连续，则其函数 $x=\varphi(y)$ 在其定义域 $[f(a),f(b)]$ 或者 $[f(b),f(a)]$ 上也连续。*

证明 假定原函数 $f(x)$ 是严格增加的，只要证明反函数 $x=\varphi(y)$ 在任意一点 $y_0\in[f(a),f(b)]$ 处连续就行，只证明 y_0 属于开区间的情形，端点处的情形也是类似的，设 $y_0=f(x_0)\in(f(a),f(b))$，需要证明对任意 $\delta>0$，都存在正实数 $\varepsilon>0$，以使得 $|y-y_0|<\varepsilon$ 时就有 $|x-x_0|<\delta$，这是很容易的，对于确定的 δ，则两个函数值 $U=y_0-\delta$ 和 $V=y_0+\delta$，如果 U 或 V 超出了闭区间 $[f(a),f(b)]$ 的范围，就适当向 $f(x_0)$ 的方向作一定程度的调整，以使得闭区间 $[U,V]\subseteq[f(a),f(b)]$，且 $U<y_0<V$，如此一来，由介值性，存在 $x_1,x_2\in[a,b]$，使得 $U=f(x_1)$ 和 $V=f(x_2)$，又由单调性，显然 $x_1<x_0<x_2$，显然当 $x\in(x_1,x_2)$ 时，因此可以取 $\varepsilon=\min\{x_0-x_1,x_2-x_0\}$ 就符合条件. □

从证明过程可以看出，定理中的闭区间是可以适当放宽的，为了证明反函数在某点处连续，只要原来的函数在包含对应点的某个闭区间内连续就可以了，下面这个例子便应用了这一点。

例 1.4.8 由指数函数 $y=a^x(a>0,a\neq 1)$ 在 $\mathbb{R}$ 上的连续性，便可以得出对数函数 $x=\log_a y$ 在 $(0,+\infty)$ 上也是连续函数。 ■

例 1.4.9 前面已经证明了，正弦函数 $y=\sin x$ 和余弦函数 $y=\cos x$ 都是 $\mathbb{R}$ 上的连续函数，因此由这里的定理便可知道，反正弦函数 $x=\arcsin y(y\in[-1,1])$ 以及反余弦函数 $x=\arccos y(y\in[-1,1])$ 便也在定义域上连续。 ■

1.4.8 一致连续

定义 1.4.3. *设函数 $f(x)$ 在某个区间上连续，如果对于任意小的正实数 $\varepsilon>0$，都存在某个 $\delta>0$，使得区间上任意满足 $|x_1-x_2|<\delta$ 的 x_1,x_2，都有 $|f(x_1)-f(x_2)|<\varepsilon$ 成立，则称函数在此区间上一致连续。*

注意一致连续与在某点处连续是有区别的，函数在某点 x_0 处连续，那么对于一个 ε，定义中的 δ，一般而言是与 x_0 有关的，即 $\delta=\delta(x_0)$，例如反比例函数 $f(x)=\frac{1}{x}$，如

果 x_0 越靠近 0，则 δ 就要越小，而一致连续则表明，存在对区间上所有点都适用的 δ，反比例函数显然无法满足这一点，所以反比例函数在定义域的两个区间上虽然连续，但不是一致连续。

例 1.4.10 利普希茨连续 设 $f(x)$ 是定义在某个区间上的函数，利普希茨条件是指，存在一个常数 $L>0$，使得对区间上任意两个实数 x_1, x_2，都有 $|f(x_1)-f(x_2)| \leqslant |x_1-x_2|$. 显然满足利普希茨条件的函数一定连续，并且一致连续。 ■

定理 1.4.10. *如果函数在某个有限区间上一致连续，则在此区间上必定有界。*

证明 固定一个正实数 $\varepsilon>0$，则按一致连续能够确定出一个正实数 $\delta>0$，使得当 $|x_1-x_2|<\delta$ 时有 $|f(x_1)-f(x_2)|<\varepsilon$，于是将区间划分成若干个闭区间的并集，其中每个闭区间的长度都小于 δ 但大于 $\frac{\delta}{2}$，则显然在每个闭区间上都有界，而原来区间的长度是有限的，所以这样的闭区间只能是有限个，因而函数在整个区间上也就有界。 □

定理 1.4.11. *闭区间上的连续函数在此闭区间上必定一致连续。*

证明 因为函数在闭区间上连续，所以对于任意一个确定的正实数 $\varepsilon>0$，在每个 x_0 处都存在一个相应的 $\delta=\delta(x_0)$，使得函数在区间 $(x_0-\delta, x_0+\delta)$ 上的函数值都落在区间 $(f(x_0)-\frac{\varepsilon}{2}, f(x_0)+\frac{\varepsilon}{2})$ 中，所有这些开区间 $(x_0-\delta, x_0+\delta)$ 显然覆盖了闭区间，依有限覆盖定理，存在其中有限个开区间 $(x_i-\delta_i, x_i+\delta_i)(i=1,2,\ldots,n)$，它们的并集就覆盖了此闭区间，取这有限个开区间的诸 δ_i 中的最小者的一半记为 δ_m，再取这有限个开区间中的诸重叠部分的最小长度为 δ_w，记 $\delta'=\min\{\delta_m, \delta_w\}$，则在闭区间上任意取两个实数 r_1, r_2 并满足 $|r_1-r_2|<\delta'$，则两数必然能够位于前述有限个开区间中的同一个 $(x_k-\delta_k, x_k+\delta_k)$ 上，则显然有 $|r_1-r_2|<\delta'\leqslant\frac{\delta_k}{2}$，从而

$$|f(r_1)-f(r_2)| \leqslant |f(r_1)-f(x_k)|+|f(r_2)-f(x_k)| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon$$

这便表明函数在闭区间上是一致连续的。 □

1.4.9 无理指数幂

在中学数学里，我们已经有了指数的概念，但那时的指数，受限于有理数的情形，虽然给出了定义在全体实数上的指数函数，却没有说明当指数是无理数时，这个幂是何种意义，本小节就来解决这个问题，我们将通过极限来定义无理指数幂。

先回顾一下有理指数幂的定义，设实数 $a>0$ 且 $a\neq 1$，它的正整数 n 次幂定义为

$$a^n = aa\cdots a(n\text{个}a)$$

在这定义下，显然有 $a^n>0$(正值性)，而且对于两个正整数 n 和 m 有

$$a^{n+m}=a^n a^m \tag{1.4.1}$$

这称为指数运算的乘法公式。

如果正整数 $n<m$，则在 $a>1$ 时成立

$$a^n<a^m$$

在 $0<a<1$ 时不等式反向，这就是说，定义在正整数集上的指数函数是单调函数，$a>$ 时是单调增加的，$0<a<1$ 时是单调减少的。

现在来把指数推广到任意整数，我们希望上述乘法公式在推广后对任意整数都成立，所以在其中令 $m=0$ 得 $a^n=a^n\cdot a^0$，这要对任意正整数 n 都成立则必须 $a^0=1$，于是我们就把这作为零次幂的定义，于是指数的正值性仍然成立，而定义在非负整数集上的指数函数仍然保持着它在正整数集上的单调性，接着在上述乘法公式中取 $m=-n$，则得到

$$a^{-n}=\frac{1}{a^n}$$

于是我们把它作为负整数指数幂的定义，这样，我们就把指数概念推广到了任意整数的情形，而且上述乘法公式对于任意两个整数都成立，而且还可以得到下面的公式

$$a^{n-m}=\frac{a^n}{a^m}$$

容易证明，在把指数函数的定义域从正整数集扩充到全体整数后，正值性和单调性仍然成立，这里就单调性作一证明。

定理 1.4.12. *设两个整数 $n<m$，实数 $a>1$ 且 $a\neq 1$，则在 $a>1$ 时有 $a^n<a^m$，在 $0<a<1$ 时有 $a^n>a^m$.*

证明 在 $a>1$ 的情况下，如果 n 是负整数而 m 是非负整数，则利用定义在正整数集上的指数函数的单调性得

$$a^n=\frac{1}{a^{-n}}<\frac{1}{a^0}=1=a^0\leqslant a^m$$

而在 n 和 m 都是负整数的情形，$-n$ 和 $-m$ 是两个正整数并且 $-n>-m$，所以利用定义在正整数集上的指数函数的单调性，有

$$a^n=\frac{1}{a^{-n}}<\frac{1}{a^{-m}}=a^m$$

这就证得 $a>1$ 时，定义在整数集上的指数是增函数，而在 $0<a<1$ 时，由

$$a^n=\frac{1}{a^{-n}}$$

知它是减函数。 □

由这定理即有如下推论

推论 1.4.3. *对于 $a>1$ 和正整数 n，有 $a^n>1$，而对于负整数 n，则 $0<a^n<1$，而对于 $0<a<1$，正整数 n 则是 $0<a^n<1$，对负整数 n 则是 $a^n>1$.*

我们再继续把指数向有理数范围内推广，我们先证下面的结论

定理 1.4.13. *设 n 和 m 是任意两个整数,实数 $a>0$ 且 $a\neq 1$,则 $(a^m)^n=a^{mn}=(a^n)^m$.*

证明 先证 n 和 m 都是正整数的情形，$(a^m)^n$ 代表 n 个 a^m 相乘，而 a^m 代表 m 个 a 相乘，所以最终便是 mn 个 a 相乘，所以 $(a^m)^n=a^{mn}$，同理 $(a^n)^m=a^{nm}=a^{mn}$

如果 m 和 m 中至少有一个是零，则结论显然是成立的，然后按负正整数指数幂的定义也容易得出结论对于 n 和 m 中至少有一个是负整数时也是成立的。 □

有了这个定理，我们来考虑有理整数幂，设 $x = n/m$ 为有理数，其中 n 和 m 是一对互素的整数并且 m 是正整数，我们推广的依据是使得刚才定理中的结论对有理数也成立，这就是说，有下式成立

$$(a^{\frac{n}{m}})^m = a^{\frac{n}{m}m} = a^n$$

于是得到

$$a^{\frac{n}{m}} = \sqrt[m]{a^n}$$

我们就把它作为有理指数幂和定义，乘法公式和负指数幂的公式仍然是成立的，并且指数函数在有理数集上的正值性和单调性仍然是成立的，这里我们证明一下单调性。

根据有理指数幂的定义，不难证明推论 1.4.3在有理数上也是成立的，假定 x 和 y 是两个有理数并且 $x < y$，则由正值性知 $a^x > 0$,$a^y > 0$，在 $a > 1$ 时

$$\frac{a^x}{a^y} = a^{x-y}$$

如果 $a > 1$,那么由 $x-y$ 是一个负有理数,利用推论 1.4.3 在有理数的情形即知 $a^{x-y} < 1$，所以 $a^x < a^y$，同理可证 $0 < a < 1$ 的情形，这就证明了定义在有理数集上的指数函数的单调性。

现在，我们再作一次推广，把指数推广到全体实数，这只要定义无理指数幂就可以了，但是这时指数的乘法公式似乎已经对我们没什么帮助了，我们转而寻求用有理数逼近无理数时有理指数幂的极限来定义无理指数幂，这就是以下的定义

定义 1.4.4. *设实数 $a > 0$ 且 $a \neq 1$，各项均为有理数的数列 r_n 收敛到一个无理数 r，则实数 a 的 r 次幂定义为*

$$a^r = \lim_{n\to\infty} a^{r_n}$$

这里有几个疑问：这个极限存在吗？对于收敛到无理数 r 的所有有理数列，这个极限都相等吗？下面这个定理就肯定了这一点。

定理 1.4.14. *设实数 $a > 0$ 且 $a \neq 1$，r 是一个无理数，则对于任意收敛到 r 的有理数数列 $\{r_n\}$，数列 $\{a^{r_n}\}$ 都收敛，而且极限值都相同 (即极限值只与 r 有关)。*

先证明下面的引理

引理 1.4.1. *设有理数数列 r_n 收敛到零,实数 $a > 0$ 且 $a \neq 1$,则有极限 $\lim_{n\to\infty} a^{r_n} = 1$.*

证明 我们在例 1.2.14中已经证得 $\lim_{n\to\infty} \sqrt[n]{a} = 1$，因此对于任意小的正实数 ε，都存在正整数 N，使得 $n > N$ 时恒有 $|\sqrt[n]{a} - 1| < \varepsilon$，在 $a > 1$ 时，就是 $1 < \sqrt[n]{a} < 1 + \epsilon$，现在就任意取定一个 $n_0 > N$，从而有 $1 < \sqrt[n_0]{a} < 1 + \varepsilon$ 而由引理条件，r_n 以零为极限，所以对于正实数 $1/n_0$，存在正整数 N_1，使得当 $n > N_1$ 时有 $|r_n| < 1/n_0$，这时按照定义在有理数集上的指数函数的单调性 (下式中做了限定 $\varepsilon < 1$)

$$1 - \varepsilon < \frac{1}{1+\varepsilon} < \frac{1}{\sqrt[n_0]{a}} < a^{r_n} < \sqrt[n_0]{a} < 1 + \varepsilon$$

由此即知 $|a^{r_n} - 1| < \varepsilon$，所以只要 $n > \max\{N, N_1\}$ 时便能保证 $|a^{r_n} - 1| < \varepsilon$，这即表明引理中的极限成立，而类似的可以证明 $0 < a < 1$ 的情况。 □

现在回过头来证明前面的定理

证明 我们先通过两个特殊的有理数数列来确定出这个极限值来，再证明所有收敛到无理数 r 的有理数数列都必以它为极限。

取无理数 r 的 n 位不足近似值 x_n 和 n 位过剩近似值 y_n，即 x_n 是把无理数 r 的第 n 位小数以后的小数全部舍去而得的有理数，y_n 是把它第 n 位小数以后的小数收上来而得到的有理数，显然 $x_n < r < y_n$ 并且 $y_n - x_n = 10^{-n}$, 同时，x_n 单调不减，而 y_n 单调不增，考虑由它们构成的两个有理指数幂的数列 a^{x_n} 和 a^{y_n}，由定义在有理数集上的指数函数的单调性，在 $a > 1$ 的假定下有

$$a^{x_n} \leqslant a^{x_{n+1}} \leqslant \cdots \leqslant a^{y_{n+1}} \leqslant a^{y_n}$$

于是作闭区间序列 $U_n = [a^{x_n}, a^{y_n}]$，则显见 $U_{n+1} \subset U_n$，而区间的长度 $a^{y_n} - a^{x_n} = a^{x_n}(a^{y_n - x_n} - 1)$，因为 $y_n - x_n = 10^{-n} \to 1$，由刚才所证的引理 1.4.1，$a^{y_n - x_n} - 1$ 是一个无穷小，而前面的因子 $a^{x_n} < a^{y_n} \leqslant a^{y_1}$ 是有界量，所以这闭区间的长度序列趋于零，于是由闭区间套定理，存在唯一实数 K，使得 $a^{x_n} < K < a^{y_n}$ 对一切正整数 n 成立，显然 $\lim_{n\to\infty} a^{x_n} = \lim_{n\to\infty} a^{y_n} = K$。

接下来我们需要证明，对于其它任何收敛到无理数 r 的有理数数列 z_n，也必将有 $\lim_{n\to\infty} a^{z_n} = K$，这是因为

$$\lim_{n\to\infty} \frac{a^{z_n}}{a^{x_n}} = \lim_{n\to\infty} a^{z_n - x_n} = 1$$

因此

$$\lim_{n\to\infty} a^{z_n} = \lim_{n\to\infty} a^{x_n} \frac{a^{z_n}}{a^{x_n}} = \lim_{n\to\infty} a^{x_n} \cdot \lim_{n\to\infty} \frac{a^{z_n}}{a^{x_n}} = K \cdot 1 = K$$

这就在 $a > 1$ 的情况下证明了定理，而 $0 < a < 1$ 时的情况完全类似。 □

这样无理指数幂定义的存在性和唯一性问题就解决了，我们把指数幂和概念推广到了指数是任意实数的场合，这时有一个问题就冒出来了，那就是我们是用的极限而不是指数乘法公式来推导无理指数幂的，那么推广后的实数指数幂是否仍满足前面的指数乘法公式呢? 进一步，在有理数上成立的那些运算性质，是否都仍然成立呢? 这由以下定理回答

定理 1.4.15. *设实数 $a > 0$ 且 $a \neq 1$，x 和 y 是任意两个实数，则 (1).*

$$a^{x+y} = a^x a^y$$

(2).

$$a^{-x} = \frac{1}{a^x}$$

为了证明它，先证一个引理

引理 1.4.2. *设实数 $a > 0$ 且 $a \neq 1$，r_n 是一个有理数数列，并且收敛到一个有理数 r，则 $\lim_{n\to\infty} a^{r_n} = a^r$.*

证明 由有理数数列 r_n 收敛到有理数 r 即知有理数数列 $r_n - r$ 收敛到零，由引理 1.4.1即知 $\lim_{n\to\infty} a^{r_n - r} = 1$，从而

$$\lim_{n\to\infty} a^{r_n} = \lim_{n\to\infty} a^{r+(r_n-r)} = \lim_{n\to\infty} a^r a^{r_n-r} = a^r \lim_{n\to\infty} a^{r_n-r} = a^r$$

□

现在来证明定理 1.4.15

证明 (1). 只需证明 x 和 y 中至少有一个无理数的情形，假定 x 是无理数，设 x_n 是一个以 x 为极限的有理数数列，则有

$$a^{x_n+y} = a^{x_n} \cdot a^y$$

显然 $x_n + y$ 是一个以 $x + y$ 为极限的有理数数列，而 $x + y$ 为无理数，所以上式左边的极限是 a^{x+y}，显然右端的极限是 $a^x a^y$，由极限的唯一性即得

$$\lim_{n\to\infty} a^{x_n+y} = \lim_{n\to\infty} a^{x_n} \cdot a^y$$

这就表明

$$a^{x+y} = a^x a^y$$

当 x 和 y 都是无理数时，设 x_n 和 y_n 是两个分别收敛到 x 和 y 的有理数数列，有

$$a^{x_n+y_n} = a^{x_n} a^{y_n}$$

显然右端以 $a^x a^y$ 为极限，对于左边，如果 $x + y$ 是无理数，则 $x_n + y_n$ 是收敛到无理数 $x + y$ 的有理数数列，所以它的极限是 a^{x+y}，如果 $x + y$ 是有理数，则 $x_n + y_n$ 是收敛到有理数 $x + y$ 的有理数数列，由引理 1.4.2知左边极限也是 a^{x+y}，所以无论 $x + y$ 是有理数还是无理数，左边都以 a^{x+y} 为极限，由极限的唯一性，得 $a^{x+y} = a^x a^y$.

(2). 同样只需要证明 x 为无理数的情形，设有理数数列 x_n 以 x 为极限，则显然有

$$a^{-x_n} = \frac{1}{a^{x_n}}$$

显然 $-x_n$ 是以无理数 $-x$ 为极限的有理数数列，所以上式左边以 a^{-x} 为极限，右边显然以 $1/a^x$ 为极限，由极限的唯一性，结论成立。 □

有了定理中的这两条，显然对于实数 x 和 y，也有

$$a^{x-y} = \frac{a^x}{a^y}$$

所以有理指数幂的运算性质，在实数范围内仍然是成立的。

既然指数扩展到了全体实数，那么我们也可以将指数函数的定义域扩充到全体实数上了，我们先来证明指数函数在 R 都是单调的

定理 1.4.16. *设实数 $a > 1$ 且 $a > 0$，指数函数 $f(x) = a^x$ 在 $a > 1$ 时是 R 上的严格递增函数，在 $0 < a < 1$ 时是严格递减函数。*

证明 先证 $a>1$ 的情况，这时任取两个实数 $x<y$，只需证明 $a^x<a^y$，此处只需要证明 x 和 y 中至少有一个无理数的情形，先假定 x 是无理数而 y 是有理数，则可以在 x 和 y 之间取一个有理数 z，即 $x<z<y$(这总是可以办到的)，然后设 x_n 是一个收敛到 x 的有理数数列，则由定义在有理数集上的指数函数的单调性和数列极限的保号性，在 x_n 中必定从某项起恒成立下面不等式

$$a^{x_n}<a^z<a^y$$

对上式取极限即得

$$a^x\leqslant a^z<a^y$$

这就证明了 $a^x<a^y$，这里引入 z 就是为了去掉 $a^x\leqslant a^y$ 中的等号。

同理可证 x 为有理数而 y 为无理数的情形，现在来看 x 和 y 都是无理数的情形，这时通过有理数列逼近的方法只能得出 $a^x\leqslant a^y$，为了去掉等号，我们在 x 和 y 之间插入两个有理数 r 和 s，即 $x<r<s<y$，这时令 x_n 和 y_n 是两个分别收敛到 x 和 y 的有理数数列，就有

$$a^{x_n}<a^r<a^s<a^{y_n}$$

取极限即得

$$a^x\leqslant a^r<a^s\leqslant a^y$$

所以 $a^x<a^y$，这就表明 $a>1$ 时，指数函数 $0<a<1$ 是 R 上的递增函数，而对于 $0<a<1$ 的情况，容易知道 $a^{-x}=1/a^x$ 对无理数 x 也是成立的，所以由这关系即可知道 $0<a<1$ 时指数函数是单调递减的。 □

接着考虑指数函数在 R 上的连续性，结论是，指数函数是实数集 R 上的连续函数。为着证明这一点，我们需要先证一个极限

引理 1.4.3. 设实数 $a>0$ 且 $a\neq 1$，则有极限 $\lim_{x\to 0}a^x=1$

证明 只证明 $a>1$ 的情形，$0<a<1$ 是类似的。

由引理引理 1.4.1，当 x 取有理数并趋于零时，有 a^x 趋于 1，所以对于无论多么小的正实数 ε，总存在另一正实数 δ，使得满足 $|x|<\delta$ 的一切有理数 x 都成立 $1-\varepsilon<a^x<1+\varepsilon$，而对于满足 $|x|<\delta$ 的任一无理数 x，必然可以在区间 $(-\delta,\delta)$ 上找到一个有理数 x' 和，使得 $0<|x|<|x'|<\delta$，这时在 $a>1$ 的情况下，利用前面已经证过的指数函数在实数集上的单调性，就有

$$1-\varepsilon<a^{-x'}<a^x<a^{x'}<1+\varepsilon$$

这就表明区间 $(-\delta,\delta)$ 上的全体实数 x 无论有理数无理数都满足 $|a^x-1|<\varepsilon$，所以此极限得证。 □

定理 1.4.17. 定义在 R 上的指数函数 $f(x)=a^x(a>0$ 且 $a\neq 1)$，是 R 上的连续函数。

证明 只要证明它在 R 上任何一点 x_0 处都连续即可，因为

$$a^{x_0+h} - a^{x_0} = a^{x_0}(a^h - 1)$$

由引理 1.4.3即知 $\lim_{h\to 0} a^{x_0+h} = a^{x_0}$，这就表明指数函数在 x_0 处是连续的，由 x_0 的任意性即得知它在整个 R 上都是连续的。 □

现在，我们有了完整的指数定义，就可以考虑它的逆运算了，也就是对数，设实数 $a > 0$ 且 $a \neq 1$，y 是一个正实数，如果实数 x 满足方程 $a^x = y$，则称 x 是 y 的以 a 为底的 对数，显然，指数和对数互为逆运算。

对数的定义有一个问题，满足方程 $a^x = y$ 的实数 x 是否一定存在呢，实数 y 必须是正实数吗？为解决这个问题，我们还需要证明指数函数的另一个性质

定理 1.4.18. *设实数 $a > 0$ 且 $a \neq 1$，则指数函数 $f(x) = a^x$ 的函数值可以取遍一切正实数，换句话说，它的值域是 $(0, +\infty)$.*

这利用连续函数在闭区间上的介值性便可以证明，我们在本节的后文给出。

1.4.10 初等函数的连续性

在中学里，我们接触过几类 *基本初等函数*: 幂函数、指数函数、对数函数、三角函数. 我们在这一小节里来证明这些函数在它们的定义域的各个区间上都是连续函数，在有了这个结论之后，根据连续的性质，所有的初等函数就都是连续函数了。

1. 幂函数

定理 1.4.19. *幂函数 $f(x) = x^p$ 在定义域的各个区间上连续。*

证明 因为如果 $p < 0$，有 $f(x) = 1/x^{-p}$，如果分母是连续的，则 $f(x)$ 就是连续的，所以只要证明 $p > 0$ 的情况就可以了。

先证明 p 是正整数的情况，这时由

$$(x_0 + h)^p - x_0^p = \sum_{i=1}^{n} C_n^i x_0^{p-i} h^i$$

显然当 $h \to 0$ 时，右边的各项 (有限项) 都趋于 0，因此 $(x_0 + h)^p \to x_0^p$，所以函数在 x_0 处连续，由 x_0 的任意性，p 为正整数的情形得证。 □

2. 指数函数与对数函数

定理 1.4.20. *指数函数 $f(x) = a^x (a > 0, a \neq 1)$ 是 R 上的连续函数。*

这在上一小节我们已经证明过了。

3. 三角函数.

定理 1.4.21. *正弦函数 $f(x) = \sin x$ 在 R 上连续，余弦函数 $g(x) = \cos x$ 在 R 上连续，正切函数 $h(x) = \tan x$ 在定义域的每一个区间上都是连续函数。*

证明 先证明正弦函数，任取 $x_0 \in R$，有

$$\sin(x_0+r)-\sin x_0 = 2\cos\left(x_0+\frac{r}{2}\right)\sin\frac{h}{2}$$

我们在定理 1.3.10中就已经知道，不等式 $|\sin x| \leqslant |x|$ 对一切实数 x 恒成立，所以当 $r \to 0$ 时，上式右端是一个有界量和一个无穷小的乘积，也收敛到零，所以

$$\lim_{r\to 0}\sin(x_0+r) = \sin x_0$$

从而正弦函数在 x_0 处连续，由 x_0 的任意性，它在 R 上都是连续的。

对于余弦函数，把它写成

$$\cos x = \sin\left(x+\frac{\pi}{2}\right)$$

由正弦函数的连续性和关于复合函数连续性的定理 1.4.4即知余弦函数也是连续的。

正切函数，把它表为正弦函数和余弦函数的商，由商函数的连续性即知它在定义域的各区间上也都是连续的。 □

1.4.11 利用函数连续性求极限

利用函数的连续性，我们可以求得一些有用的极限。

例 1.4.11 关于对数函数有极限

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

因为

$$\frac{\log_a(1+x)}{x} = \log_a(1+x)^{1/x}$$

由 $\lim_{x\to 0}(1+x)^{1/x} = \mathrm{e}$ 及对数函数的连续性即得结论，特别的有

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = 1$$

这就是说，在 $x \to 0$ 时，有 $\ln(1+x) = x + o(x)$. ■

例 1.4.12 对于指数函数，有

$$\lim_{x\to 0}\frac{a^x-1}{x} = \ln a$$

只要令 $t = a^x - 1$，就有

$$\frac{a^x-1}{x} = \frac{t}{\log_a(1+t)}$$

而 $x \to 0$ 时，$t \to 0$，借用上例的结论，便知上式有极限 $\ln a$，特别情况是

$$\lim_{x\to 0}\frac{e^x-1}{x} = 1$$

这就是说，在 $x \to 0$ 时，有 $e^x = 1 + x + o(x)$ ■

例 1.4.13 在前两个例子的基础上，我们证明下面这个重要的极限，设 μ 是任意实数，则

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

■

1.4.12 题选

题 1.4.1 设定义在 $\mathbb{R}$ 上的连续函数 $f(x)$，对于任意 $x \in \mathbb{R}$ 都成立 $f(2x) = f(x)$，问该函数是否是常数函数? ■

解答 首先易知 $f(0) = 0$，并且对于任意实数 x 和任意正整数 n，有

$$f(x) = f\left(\frac{x}{2^n}\right)$$

由函数 $f(x)$ 在 $x = 0$ 处的连续性，$\forall \varepsilon > 0$，$\exists \delta > 0$，使得当 $|x| < \delta$ 时，恒有 $|f(x)| < \varepsilon$.

任取两个不相等的实数 x 和 y，对于上述 δ，必然当正整数 n 充分大时有

$$\left|\frac{x}{2^n}\right| < \delta,\ \left|\frac{y}{2^n}\right| < \delta$$

因此有

$$\left|f\left(\frac{x}{2^n}\right)\right| < \varepsilon,\ \left|f\left(\frac{y}{2^n}\right)\right| < \varepsilon$$

从而

$$\left|f\left(\frac{x}{2^n}\right) - f\left(\frac{y}{2^n}\right)\right| < 2\varepsilon$$

即

$$|f(x) - f(y)| < 2\varepsilon$$

由 ε 的任意性，有 $f(x) = f(y)$，因此函数是常数函数，并且由过程可知，函数其实保证在 $x = 0$ 处连续就能保证是常数函数了。 □

题 1.4.2 设函数 $f(x) \in C^1(0,1)$，

$$\alpha = \frac{f(x_1) - f(x_2)}{x_1 - x_2} < \beta = \frac{f(x_3) - f(x_4)}{x_3 - x_4}$$

其中 $x_1, x_2, x_3, x_4 \in (0,1)$，求证: 对每个 $\lambda \in (\alpha, \beta)$，存在 $x_5, x_6 \in (0,1)$，使得

$$\lambda = \frac{f(x_5) - f(x_6)}{x_5 - x_6}$$

■

题目出处 北京大学 2018 年硕士入学考试试题

解答日期 2018-02-28

证明 这题目有点意思，导函数具备介值性大家都知道，其证明需要用到微分中值定理，然而这题却来了个割线斜率的介值性，而且函数只是连续，并没有可导性，但是其实从直观上一想，又是显然的，思路也是超级简单。

α, β 是两段割线的斜率，我们只要让其中一段割线连续的变动，直到跟另一段割线重合，这样就得到一个连续函数，从而根据连续函数的介值性解决问题。

以下为方便，将割线 $(x_1, f(x_1)) \to (x_2, f(x_2))$ 简单的记为 $L(x_1, x_2)$.

现在就来让割线 $L(x_1,x_2)$ 连续的变动到 $L(x_3,x_4)$，变动过程中，每次变动我们只移动割线的一个端点而保持另一个端点不动。下面以 $x_1<x_2<x_3<x_4$ 为例，将割线 $L(x_1,x_2)$ 的右端点在曲线上向右移动直到与 x_4 重合，即

$$L(x_1,x_2)\to L(x_1,x_4)$$

然后再移动左端点直到与 x_3 重合，即

$$L(x_1,x_4)\to L(x_3,x_4)$$

这样就将割线 $L(x_1,x_2)$ 连续的变动到了 $L(x_3,x_4)$，我们可以把这过程写出一个分段函数来，根据 $f(x)$ 的连续性即知此分段函数的连续性，于是由其介值性就得结论。

要说明的是，这个端点移动方式有原则的，在一个端点移动过程中，不能跨过当前割线的另一个端点，否则就会出现需要对割线求极限即用导数值来定义该点函数值的问题，而原来的函数并没有可导的条件，比如说，在前面的情况中，你的割线变动方式就不能是

$$L(x_1,x_2)\to L(x_3,x_2)\to L(x_3,x_4)$$

在前一段移动中，左端点越过 x_2 时，以及后一段移动中，右端点越过 x_3 时都面临上述情况。

可以验证，无论 x_1,x_2,x_3,x_4 大小情况如何，都存在相应的割线变动方式，因此结论得证。

最后，为了证明的完整性，我们把前面的分段函数明确写出来，第一段是由把割线 $L(x_1,x_2)$ 的右端点从 x_2 移动到 x_4，而从 x_2 到 x_4 的线性变化可由函数

$$u(t)=(1-t)x_2+tx_4$$

给出，于是这一段的移动过程就可以由函数

$$h(t)=\frac{f(x_1)-f(u(t))}{x_1-u(t)}$$

给出，其中 $0\leqslant t\leqslant 1$，第二段是将割线 $L(x_1,x_4)$ 的左端点从 x_1 移动到 x_3，这个线性变化由函数

$$u(t)=(1-t)x_1+tx_3$$

给出，但为了与第一段相衔接，这一段的参数变化由 $(0,1)$ 平移为 $(1,2)$，于是这一段的变化端点是

$$u(t)=(2-t)x_1+(t-1)x_3$$

而这段的斜率是

$$h(t)=\frac{f(u(t))-f(x_4)}{u(t)-x_4}$$

其中 $1\leqslant t\leqslant 2$，这样就得出了分段函数 $h(t)$，变量 t 在 $[0,2]$ 内变化，并且 $h(0)=\alpha,h(2)=\beta$，在 $t=1$ 处两段正好衔接上，又因为 $f(x)$ 是连续的，由复合函数的连续性即知 $h(t)$ 在定义域上连续，从而具备介值性。

如果有强迫症，函数 $h(t)$ 可以写为如下:

$$h(t)=\begin{cases}\frac{f(x_1)-f((1-t)x_2+tx_4)}{x_1-((1-t)x_2+tx_4)} & 0\leqslant t\leqslant 1\\ \frac{f((2-t)x_1+(t-1)x_3)-f(x_4)}{((2-t)x_1+(t-1)x_3)-x_4} & 1<t\leqslant 2\end{cases}$$

□

第二章　导数与微分

2.1　导数与微分

2.1.1　概念与几何意义

我们先来看产生导数概念的物理学和几何学背景。

在运动物理学中，假如一个质点的位移随时间变化的方程是 $s = s(t)$，现在考虑质点在 $t = t_0$ 时刻的瞬间速度，我们用它在 t_0 附近很小的一段时间内的平均速度来逼近它，给时间一个增量 Δt，则在时间段 $[t_0, t_0 + \Delta t]$ 内的平均速度是

$$\frac{s(t_0 + \Delta t) - s(t_0)}{\Delta t}$$

如果令 Δt 无限缩小，即让 $\Delta t \to 0$，那么这个平均速度将无限逼近质点在 $t = t_0$ 处的瞬时速度。

为了找到一条曲线在其上某点处的切线，假如恰当的选择坐标系后，曲线可以用一个连续函数 $y = f(x)$ 来表示，点 $P(x_0, f(x_0))$ 是这函数图象上一点，为了寻找函数图象在该点处的切线，可以从这图象上靠近点 P 的另一点 $Q(x, f(x))$，并出割线 PQ，其斜率是

$$\frac{f(x) - f(x_0)}{x - x_0}$$

在 $x \to x_0$ 时，割线将无限的与函数图象在 x_0 处的切线靠近，其极限位置就是该点处的切线，因此，上式的极限便是函数在 x_0 处的切线的斜率，设其为 K，则曲线在 x_0 处的切线方程就是

$$y - y_0 = K(x - x_0)$$

这两个例子都提出了一个问题，即对于给定函数 $f(x)$，要求它在某点 $x = x_0$ 处的下面这个表达式的极限值:

$$\frac{f(x) - f(x_0)}{x - x_0}$$

或者写成另一种形式

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

这个比值实际上就是函数值对自变量的变化率，这就引入了导数的概念。

定义 2.1.1. 设函数 $f(x)$ 在 x_0 的某邻域内有定义，如果极限

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

存在，则称函数 $f(x)$ 在 x_0 处可导，也称为可微，此极限值称为 $f(x)$ 在 x_0 处的导数，并记为 $f'(x_0)$，或者称为微商 [1]，记为 $\frac{\mathrm{d}y}{\mathrm{d}x}\big|_{x=x_0}$，即

$$f'(x_0)=\frac{\mathrm{d}y}{\mathrm{d}x}\Big|_{x=x_0}=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$$

如果前述极限不存在，则称函数在该点处不可导. 类似的，把极限过程分别替换为左极限和右极限，就得出左导数$f'_-(x_0)$ 和右导数$f'_+(x_0)$ 的概念，即

$$f'_-(x_0)=\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0},\ f'_+(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$$

显然，函数在某点处可导的充分必要条件是，该点既左可导又右可导，且左右导数相等。

容易看出，如果在某点可导，则必在该点处连续。

如果函数 $f(x)$ 在 x_0 处可导，则在此处有

$$\frac{\Delta y}{\Delta x}=f'(x_0)+\alpha$$

其中 α 是一个无穷小，于是有

$$\Delta y=f'(x_0)\Delta x+o(\Delta x)$$

即由一个 Δx 的一个线性式子加上一个 Δx 的高阶无穷小. 这公式可以用于对函数进行估值，只要给定了 x_0 处的函数值以及该点处的导数，就可以近似的估计该点附近的函数值。

下面引入导函数的概念，考虑前面的瞬时速度的例子，给定位移随时间变动的函数 $s=s(t)$，我们已经知道它在指定时刻 $t=t_0$ 的瞬时速度就是位移函数在 $t=t_0$ 处的导数值，显然，瞬时速度也是随着时间变化的物理量，即速度函数 $v=v(t)$，它的特点是，它在任意时刻 $t=t_0$ 的函数值，正好是位移函数 $s=s(t)$ 在该处的导数值，即它是位移函数的导数函数，由此给出导函数的定义:

定义 2.1.2. *如果函数 $f(x)$ 在某个区间上有定义，并且在该区间上处处可导（闭区间上的端点处只要单侧可导），则称函数在该区间上可导，而由自变量映射导数值构成一个新的函数 $f'(x)$，称为函数 $f(x)$ 的导函数，与之相应的，函数 $f(x)$ 则称为函数 $f'(x)$ 的 (一个) 原函数.*

例 2.1.1　常数函数 $f(x)=C$，由于函数的增量始终为零，所以它在任意点处都可导，且导数为零。 ■

例 2.1.2　设 $n\in\mathbb{Z}$，我们来求幂函数 $f(x)=x^n(n\in N_+)$ 在 x_0 处的导数，因为

$$\frac{x^n-x_0^n}{x-x_0}=x^{n-1}+x^{n-2}x_0+\cdots+xx_0^{n-2}+x_0^{n-1}$$

上式右端在 $x\to x_0$ 时显然以 nx_0^{n-1} 为极限，于是该函数的导函数是

$$(x^n)'=nx^{n-1}$$

[1]这个称呼及记号的意义在介绍了微分的概念后就清楚了。

■

例 2.1.3 函数 $f(x)=\frac{1}{x}$ 在 $x_0\neq 0$ 处的导数，有

$$\frac{\frac{1}{x}-\frac{1}{x_0}}{x-x_0}=-\frac{1}{xx_0}$$

令 $x\to x_0$，便知函数在 x_0 处的导数值是 $-\frac{1}{x_0^2}$，于是有

$$\left(\frac{1}{x}\right)'=-\frac{1}{x^2}$$

■

例 2.1.4 函数 $f(x)=\sqrt{x}$ 在 x_0 处，有

$$\frac{\sqrt{x}-\sqrt{x_0}}{x-x_0}=\frac{1}{\sqrt{x}+\sqrt{x_0}}$$

因此，函数在 x_0 处的导数值是 $\frac{1}{2\sqrt{x_0}}$，即它的导函数是

$$(\sqrt{x})'=\frac{1}{2\sqrt{x}}$$

■

例 2.1.5 我们在 例 1.3.12 中已经证明了下面两个极限:

$$\lim_{x\to x_0}\frac{\sin x-\sin x_0}{x-x_0}=\cos x_0,\ \lim_{x\to x_0}\frac{\cos x-\cos x_0}{x-x_0}=-\sin x_0$$

于是有

$$(\sin x)'=\cos x,\ (\cos x)'=-\sin x$$

■

例 2.1.6 现在来求指数函数 $f(x)=a^x(a>0,a\neq 1)$ 在任意点 x_0 处的导数。利用在 例 1.4.12 中求得的极限，有

$$\lim_{x\to x_0}\frac{a^x-a^{x_0}}{x-x_0}=\lim_{x\to x_0}a^{x_0}\frac{a^{x-x_0}-1}{x-x_0}=\lim_{h\to 0}a^{x_0}\frac{a^h-1}{h}=a^{x_0}\ln a$$

于是有

$$(a^x)'=a^x\ln a$$

特别的有 $(e^x)'=e^x$，这就是说，以 e 为底的指数函数的导函数就是它自己，这非常重要且有趣. ■

例 2.1.7 我们在例 1.4.5中利用狄利克雷函数构造了仅在一个点处连续的函数例子:$f(x)=xD(x)$，我们指出，这个函数虽然在 $x=0$ 处连续，但在此处却不可导，因为当 x 分别取有理数值和无理数值并趋于零时，变化率

$$\frac{f(x)-f(0)}{x-0}=\frac{f(x)}{x}=D(x)$$

分别趋于两个不同的极限值，所以不可导，但是如果我们把函数中的无穷小因子升次，变成

$$f(x)=x^2D(x)$$

此时有

$$\lim_{x\to 0}\frac{f(x)-f(0)}{x-0}=\lim_{x\to 0}xD(x)=0$$

这表明这个新构造的函数在 $x=0$ 处可导，且导数为零，然而它在除 $x=0$ 之外的其它点处均不连续，所以也就不可导，于是它成为了仅在一个点处连续且可导的函数例子。

■

2.1.2 求导法则

定理 2.1.1. 设函数 $f(x)$ 和 $g(x)$ 在 x_0 的某个邻域内有定义且在 x_0 处可导，则函数 $u(x)=f(x)+g(x)$ 和 $v(x)=f(x)-g(x)$ 在 x_0 处也可导，并且

$$u'(x_0)=f'(x_0)+g'(x_0),\ v'(x)=f'(x_0)-g'(x_0)$$

只要注意到 $\Delta u=\Delta f+\Delta g$ 以及 $\Delta v=\Delta f+\Delta g$ 即可得知结论成立。

定理 2.1.2. 设函数 $f(x)$ 和 $g(x)$ 在 x_0 的某个邻域内有定义并且在 x_0 处可导，则函数 $u(x)=f(x)g(x)$ 在 x_0 处也可导，并且

$$u'(x_0)=f'(x_0)g(x_0)+f(x_0)g'(x_0)$$

证明 因为

$$\begin{aligned}u(x)-u(x_0) &= f(x)g(x)-f(x_0)g(x_0)\\ &= [f(x)-f(x_0)]g(x)+f(x_0)[g(x)-g(x_0)]\end{aligned}$$

于是

$$\frac{u(x)-u(x_0)}{x-x_0}=\frac{f(x)-f(x_0)}{x-x_0}g(x)+f(x_0)\frac{g(x)-g(x_0)}{x-x_0}$$

两边令 $x\to x_0$ 便得结论. □

利用数学归纳法，可以把这结论推广到多个函数相乘的情况:

定理 2.1.3. 设有若干个函数 $f_i(x)(i=1,2,\ldots,n)$ 都在 x_0 的某个邻域内有定义，又若它们都在 x_0 处可导，则函数 $P(x)=\prod\limits_{i=1}^{n}f_i(x)$ 在 x_0 处也可导，并且

$$P'(x_0)=\prod_{i=1}^{n}f_i(x_0)\cdot\sum_{i=1}^{n}\frac{f_i'(x_0)}{f_i(x_0)}$$

定理 2.1.4. 设函数 $f(x)$ 在 x_0 的某邻域内有定义且 $f(x)\neq 0$，并且在 x_0 处可导，则函数 $u(x)=\dfrac{1}{f(x)}$ 在 x_0 处也可导，且有

$$u'(x)=-\frac{f'(x_0)}{f^2(x_0)}$$

证明 因为

$$\begin{aligned}\frac{u(x)-u(x_0)}{x-x_0} &= \frac{\frac{1}{f(x)}-\frac{1}{f(x_0)}}{x-x_0} \\ &= -\frac{1}{f(x)f(x_0)}\cdot\frac{f(x)-f(x_0)}{x-x_0}\end{aligned}$$

令 $x\to x_0$ 取极限即得结论. □

由这个定理可得下面推论，当然这推论直接使用导数定义也是可以验证的。

推论 2.1.1. 函数 $y=Cf(x)$ 的导函数是 $y'=Cf'(x)$，如果 $f(x)$ 存在导函数，这里 C 是任意实常数。

由以上两个定理，只要把两个函数的商函数 $\dfrac{f(x)}{g(x)}$ 写成 $f(x)\cdot\dfrac{1}{g(x)}$，就得到

定理 2.1.5. 设函数 $f(x)$ 和 $g(x)$ 都在 x_0 的某个邻域内有定义且 $g(x)\neq 0$，如果两者都在 x_0 处可导，则 $u(x)=\dfrac{f(x)}{g(x)}$ 在 x_0 处可导，并且有

$$u'(x_0)=\frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g^2(x_0)}$$

例 2.1.8 考察多项式函数

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

现在容易得知它的导函数是

$$f'(x)=na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+2a_2x+a_1$$

显然，n 次多项式函数的导函数是 $n-1$ 次多项式函数。 ■

例 2.1.9 由商的求导法则，可得出正切函数的导函数

$$\begin{aligned}\tan' x &= \left(\frac{\sin x}{\cos x}\right)' \\ &= \frac{(\sin x)'\cos x-\sin x(\cos x)'}{\cos^2 x} \\ &= \frac{\cos^2 x+\sin^2 x}{\cos^2 x} \\ &= \frac{1}{\cos^2 x}=\sec^2 x\end{aligned}$$

以及

$$(\sec x)'=\frac{\sin x}{\cos^2 x},\ (\csc x)'=-\frac{\cos x}{\sin^2 x}$$

■

2.1.3 反函数的导数

定理 2.1.6. *如果函数 $f(x)$ 在 x_0 的某个邻域上严格单调，且在 x_0 处存在非零的导数，则其反函数 $x=\varphi(y)$ 在 $y_0=f(x_0)$ 处亦可导，且导数值*

$$\varphi'(y_0)=\frac{1}{f'(x_0)}$$

证明 由于严格单调且连续，所以 $x\to x_0$ 等价于 $y\to y_0$，因此有

$$\lim_{y\to y_0}\frac{\varphi(y)-\varphi(y_0)}{y-y_0}=\lim_{x\to x_0}\frac{x-x_0}{f(x)-f(x_0)}=\frac{1}{f'(x_0)}$$

□

例 2.1.10 由指数函数的导数，可以得到对数函数的导数，函数 $y=a^x(a>0,a\neq 1)$ 的导函数是 $y'=a^x\ln a$，于是它的反函数 $x=\log_a y$ 的导函数就是 $x'=\frac{1}{a^x\ln a}=\frac{\log_a \mathrm{e}}{y}$，特别的是 $(\ln x)'=\frac{1}{x}$. ■

例 2.1.11 来求一下反三角函数的导数，因为 $(\sin x)'=\cos x$，所以

$$(\arcsin y)'=\frac{1}{(\sin x)'}=\frac{1}{\cos x}=\frac{1}{\sqrt{1-y^2}}$$

注意反正弦的取值范围是 $(-\frac{\pi}{2},\frac{\pi}{2})$，所以这里余弦是正值，这就是反正弦函数的导数，同样可求得反余弦导函数

$$(\arccos y)'=-\frac{1}{\sqrt{1-y^2}}$$

对于反正切函数，则有

$$(\arctan y)'=\frac{1}{(\tan x)'}=\cos^2 x=\frac{1}{1+\tan^2 x}=\frac{1}{1+y^2}$$

■

2.1.4 复合函数的导数

定理 2.1.7. *设函数 $u=\varphi(x)$ 在 x_0 处可导，而函数 $y=f(u)$ 在 $u_0=\varphi(x)$ 处亦可导，且 $y=f(u)$ 在 u_0 的某空心邻域内函数值恒不与 $f(u_0)$ 相同，则复合函数 $y=f(\varphi(x))$ 在 x_0 处可导，且导数值为 $f'(u_0)\varphi'(x_0)$.*

证明 由 $x=\varphi(x)$ 在 x_0 处的连续性，当 $x\to x_0$ 时亦必有 $u\to u_0$，于是在等式

$$\frac{f(\varphi(x))-f(\varphi(x_0))}{x-x_0}=\frac{f(\varphi(x))-f(\varphi(x_0))}{\varphi(x)-\varphi(x_0)}\cdot\frac{\varphi(x)-\varphi(x_0)}{x-x_0}=\frac{f(u)-f(u_0)}{u-u_0}\cdot\frac{\varphi(x)-\varphi(x_0)}{x-x_0}$$

两边令 $x\to x_0$ 便得结论. □

根据定理，函数 $y=f(\varphi(x))$ 的导函数是 $y'=f'(\varphi(x))\varphi'(x)$.

例 2.1.12 对于双曲函数的导函数，将双曲余弦 $y=\cos x=\frac{e^x+e^{-x}}{2}$ 看成 $\varphi(x)=e^x$ 以及 $f(x)=\frac{u+\frac{1}{u}}{2}$ 的复合，就有

$$(\cosh x)'=\left(\frac{\mathrm{e}^x+\mathrm{e}^{-x}}{2}\right)'=\frac{\mathrm{e}^x-\mathrm{e}^{-x}}{2}=\sinh x$$

同样有

$$(\sinh x)' = \cosh x$$

这跟三角函数的求导函数非常类似，以后我们还会看到，双曲函数在很多方面都跟三角函数有类似的性质。 ■

例 2.1.13 我们也可以把余弦函数 $y = \cos x$ 写成 $y = \sin\left(x + \frac{\pi}{2}\right)$，从而将它看成是由 $y = \sin u$ 以及 $u = x + \frac{\pi}{2}$ 复合而成，于是便得

$$(\cos x)' = (\sin u)' \cdot u'(x) = \cos u = \cos\left(x + \frac{\pi}{2}\right) = -\sin x$$

■

例 2.1.14 我们已经知道幂函数 $x^n (x > 0, n \in \mathbb{N}_+)$ 的导函数是 nx^{n-1}，现在我们将这一结果推广为，任意幂函数 $x^p (x > 0, p \in \mathbb{R})$ 的导函数都是 px^{p-1}.

显然 $p = 0$ 时结论成立，如果 p 为有理数 $\frac{m}{n}$，那么对于下式

$$\frac{(x + \Delta x)^{m/n} - x^{m/n}}{\Delta x}$$

令 $a = (x + \Delta x)^{m/n}$ 及 $b = x^{m/n}$ 并借用公式

$$a - b = \frac{a^n - b^n}{a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}}$$

可得

$$\frac{(x + \Delta x)^{m/n} - x^{m/n}}{\Delta x} = \frac{(x + \Delta x)^m - x^m}{\Delta x(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})}$$

令 $\Delta x \to 0$，显然有

$$\frac{(x + \Delta x)^m - x^m}{\Delta x} \to mx^{m-1}$$

而因子 $a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}$ 中的每一项都趋于 $x^{m(n-1)/n}$，因此整个式子的极限便是

$$\frac{mx^{m-1}}{nx^{m(n-1)/n}} = \frac{m}{n}x^{\frac{m}{n}-1}$$

于是结论对于 $p \in \mathbb{Q}$ 都成立。

但当考虑到 p 为无理数的情形时，上面的方法就失效了，我们不得不寻求别的方法，在这里，我们把 x^p 写成复合函数的形式 $y = \mathrm{e}^{p\ln x}$，于是可以求得

$$(x^p)' = (\mathrm{e}^{p\ln x})' = \mathrm{e}^{p\ln x} \cdot (p\ln x)' = \mathrm{e}^{p\ln x} \cdot \frac{p}{x} = px^{p-1}$$

于是就证明了结论。 ■

例 2.1.15 仿上例的方法，设函数 $u(x)$ 和 $v(x)$ 均可导，我们来求 $y = u(x)^{v(x)}$ 的导函数，同样把它写成 $y = \mathrm{e}^{v(x)\ln u(x)}$ 的形式，可求得

$$y' = e^{v(x)\ln u(x)}(v(x)\ln u(x))' = u(x)^{v(x)}(v'(x)\ln u(x) + \frac{v(x)u'(x)}{u(x)})$$

■

例 2.1.16 假如函数 $f(x)$ 的定义域关于 $x = 0$ 对称，如果它是奇函数，即对任意 x 成立 $f(-x) = -f(x)$，那么可以证明，它的导函数 $f'(x)$ 一定是偶函数，反之，如果 $f(x)$ 是偶函数，那么它的导函数 $f'(x)$ 一定是奇函数，这在正弦函数和余弦函数的身上已经得到验证。 ■

2.1.5 求导公式表

我们把到现在为止，我们已经求得的导数公式列在这里:

1. 常函数 $(C)' = 0$

2. 幂函数 $(x^p)' = px^{p-1}, (p \in \mathbb{R}, x > 0)$.

3. 指数函数 $(a^x)' = a^x \ln a$，对数函数 $(\log_a x)' = \frac{\log_a \mathrm{e}}{x}$，特殊情况是 $(\mathrm{e}^x)' = \mathrm{e}^x$ 与 $(\ln x)' = \frac{1}{x}$.

4. 三角函数与反三角函数 $(\sin x)' = \cos x$，$(\cos x)' = -\sin x$，$(\tan x)' = \sec^2 x$，$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$，$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$，$(\arctan x)' = \frac{1}{1+x^2}$.

2.1.6 参变量函数的导数

前面已经提过，引入导数概念的一个几何背景就是求曲线上某点处的切线，这个问题在曲线方程是用函数方程 $y = f(x)$ 来表达的情形下已经解决了，但是通常一般的曲线并不能表达为一个函数，而是由参数方程表达的，比如圆的参数方程是 $x = \cos\theta, y = \sin\theta$，假定一段曲线的参数方程是

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

假定参数 t 在 t_0 的某个邻域内变化，且两个分函数都在这个邻域上可导，那么有

$$\frac{\Delta y}{\Delta x} = \frac{y(t+\Delta t) - y(t)}{x(t+\Delta t) - x(t)} = \frac{y(t+\Delta t) - y(t)}{\Delta t} \cdot \frac{\Delta t}{x(t+\Delta t) - x(t)}$$

令 $\Delta t \to 0$ 便得

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y'(t)}{x'(t)}$$

当然这里需要 $x'(t) \neq 0$，这就是参变量函数的导数公式.

例 2.1.17 对于圆来说，它的参数方程是

$$\begin{cases} x = \cos\theta \\ y = \sin\theta \end{cases}$$

于是它在 $\theta = \theta_0$ 处的导数是

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\cot\theta_0$$

∎

对于函数 $y = f(x)$ 来说，它也可以看成是由以下的参数方程决定的曲线

$$\begin{cases} x = t \\ y = f(t) \end{cases}$$

2.1.7 高阶导数

如果一个函数连续，则它的图象是一条不连续不间段的曲线，而如果它又可导，则它的曲线呈现出光滑无转折点的特征，那么由于导函数的存在，自然会问到导函数是否也可导这样一个问题，也就是导数的导数，即高阶导数。

定义 2.1.3. 如果函数 $f(x)$ 在某个区间上存在导函数，且导函数也可导，则称导函数的导函数是 $f(x)$ 的二阶导数，记作 $f''(x)$ 或者 $\frac{\mathrm{d}^2y}{\mathrm{d}x^2}$，类似的可以得到三阶导数$\frac{\mathrm{d}^3y}{\mathrm{d}x^3}$ 或者 $y^{(3)}$ 或者 $f^{(3)}(x)$、四阶导数$\frac{\mathrm{d}^4y}{\mathrm{d}x^4}$ 或者 $f^{(4)}(x)...n$ 阶导数 $\frac{\mathrm{d}^ny}{\mathrm{d}x^n}$ 或者 $f^{(n)}(x)$，这些统称高阶导数.

下面考虑基本函数的高阶导数。

1. 幂函数 $y = x^p(x > 0, p \in \mathbb{R})$.

显然有

$$
\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= px^{p-1} \\
\frac{\mathrm{d}^2t}{\mathrm{d}x^2} &= p(p-1)x^{p-2} \\
&\cdots \\
\frac{\mathrm{d}^ny}{\mathrm{d}x^n} &= p(p-1)\cdots(p-n+1)x^{p-n}
\end{aligned}
$$

对于指数是正整数 m 的情形来说，函数 $x^m(m \in \mathbb{N}_+)$ 的一阶导数是 $m-1$ 次多项式，二阶导数是 $m-2$ 阶多项式，m 阶导数则成为常数，阶数更高阶的导数将恒为零函数。

2. 指数函数 $y = a^x(a > 0, a \neq 1)$.

这时有

$$
\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= a^x \ln a \\
\frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= a^x(\ln a)^2 \\
\frac{\mathrm{d}^3y}{\mathrm{d}x^3} &= a^x(\ln a)^3 \\
&\cdots \\
\frac{\mathrm{d}^ny}{\mathrm{d}x^n} &= a^x(\ln a)^n
\end{aligned}
$$

对于 $y = \mathrm{e}^x$ 来说，则因为 $\frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^x = y$，所以它的任意阶的导函数仍然是它自己。

3. 对数函数 $y = \log_a x(a > 0, a \neq 1)$.

因为求一阶导数后就成为幂函数，所以可以应用幂函数的高阶导数结果.

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{1}{x\ln a} \\
\frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= -\frac{1}{x^2\ln a} \\
&\cdots \\
\frac{\mathrm{d}^ny}{\mathrm{d}x^n} &= \frac{(-1)^{n-1}(n-1)!}{x^n\ln a}
\end{aligned}$$

特别情况是

$$\frac{\mathrm{d}^n(\ln x)}{\mathrm{d}x^n} = \frac{(-1)^{n-1}(n-1)!}{x^n}$$

4. 三角函数

正弦函数 $y=\sin x(x\in\mathbb{R})$

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \cos x = \sin\left(x+\frac{\pi}{2}\right) \\
\frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= -\sin x = \sin(x+\pi) \\
&\cdots \\
\frac{\mathrm{d}^ny}{\mathrm{d}x^n} &= \sin\left(x+\frac{n\pi}{2}\right)
\end{aligned}$$

同理，对于余弦也有

$$\frac{\mathrm{d}^n(\cos x)}{\mathrm{d}x^n} = \cos\left(x+\frac{n\pi}{2}\right)$$

对于正切函数 $y=\tan x$，先求出低数较高的几个导数如下

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= 1+\tan^2 x \\
\frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= 2\tan x(1+\tan^2 x) \\
\frac{\mathrm{d}^3y}{\mathrm{d}x^3} &= (2+6\tan^2 x)(1+\tan^2 x)
\end{aligned}$$

由复合函数求导法则可知，其 n 阶导数中必然包含因式 $1+\tan^2 x$，而剩下的部分是一个关于 $\tan x$ 的 $n-1$ 次多项式，即

$$\frac{\mathrm{d}^n(\tan x)}{\mathrm{d}x^n} = (1+\tan^2 x)g_n(\tan x)$$

求导便得

$$\frac{\mathrm{d}^{n+1}y}{\mathrm{d}x^{n+1}} = (1+\tan^2 x)[2\tan x g_n(\tan x) + (1+\tan^2 x)g_n'(\tan x)]$$

即 $g_n(t)$ 具有如下的递推公式

$$g_{n+1}(t) = 2tg_n(t) + (1+t^2)g_n'(t)$$

且 $g_1(t)=1$.

5. 对于反三角函数.

2.1.8 莱布尼茨公式

设两个函数 $f(x)$ 与 $g(x)$ 都存在高阶导数，来讨论一下它俩之积的高阶导数，设 $h(x) = f(x)g(x)$，根据乘积函数的求导法则，可以求得

$$\begin{aligned} h'(x) &= f'(x)g(x) + f(x)g'(x) \\ h''(x) &= f''(x)g(x) + 2f'(x)g'(x) + f(x)g''(x) \\ h'''(x) &= f'''(x)g(x) + 3f''(x)g'(x) + 3f'(x)g''(x) + f(x)g'''(x) \end{aligned}$$

易见这与二项式定理非常相似，利用数学归纳法可以证明下面一般性的结果:

定理 2.1.8. 设函数 $f(x)$ 与 $g(x)$ 都在某个共同的区间上存在直到 n 阶的导函数，则乘积 $h(x) = f(x)g(x)$ 的 n 阶导数是

$$h^{(n)}(x) = \sum_{i=0}^{n} C_n^i f^{(n-i)}(x)g^{(i)}(x)$$

2.1.9 微分与高阶微分

现在来讨论一个重要的概念：微分。

假定函数 $f(x)$ 在 x_0 的某邻域内有定义且在 x_0 处可导，那么有

$$\Delta y = f'(x_0)\Delta x + o(\Delta x)$$

即函数值的增量 Δy 可以写成一个自变量增量的倍数 $f'(x_0)\Delta x$ 与一个自变量增量的高阶无穷小之和，也就是说，函数值的增量 Δy 作为一个无穷小，我们从中分离出了主要成分 $f'(x_0)\Delta x$，而剩下的部分与主要成分相比可以忽略不计（高阶无穷小），这个主要成分正好是自变量增量的倍数，称为线性主部，于是提出微分的概念如下:

定义 2.1.4. 设函数 $f(x)$ 在 x_0 的某个邻域内有定义，如果存在常数 A，使得当自变量增量 Δx 趋于零时，函数值增量 Δy 与 $A\Delta x$ 相差一个 Δx 的高阶无穷小，即

$$\Delta y = A\Delta x + o(\Delta x)$$

则称函数 $f(x)$ 在 x_0 处可微，而表达式 $A\Delta x$ 就称为函数 $f(x)$ 在 x_0 处的微分，记作 $\mathrm{d}y|_{x=x_0}$ 或者 $\mathrm{d}f(x_0)$.

微分概念的核心就是从无穷小中分离出最主要的线性成分，这部分就是函数的微分。

如果函数 $f(x)$ 在 x_0 处可导，显然就有

$$\Delta y = f'(x_0)\Delta x + o(\Delta x)$$

这表明，$f'(x_0)\Delta x$ 就是函数在该处的微分，也就是 $A = f'(x_0)$，微分定义中显然 A 若存在也不可能有多个值，所以得到结论：函数在某点处可微的充分必要条件是在该点处可导。而微分就是导数与自变量的增量之积。

对于自变量而言，如果把它看成它自己的函数，显然有 $\mathrm{d}x = \Delta x$，所以微分表达式通常写成

$$\mathrm{d}y = f'(x_0)\,\mathrm{d}x$$

而不是写成 $\mathrm{d}y = f'(x_0)\Delta x$，把上式改写成

$$f'(x_0) = \frac{\mathrm{d}y}{\mathrm{d}x}$$

这就表明，函数在某点处的导数值，正是该点处函数的微分与自变量比的微分之比，这就是导数被称为微商以及导数符号 $\frac{\mathrm{d}y}{\mathrm{d}x}$ 的由来。

微分可以用于一些近似计算.

例 2.1.18 ，例如，要计算 $\sqrt{4.01}$，由开方函数的微分 $\mathrm{d}(\sqrt{x}) = \frac{\mathrm{d}x}{2\sqrt{x}}$ 得

$$\sqrt{4.01} \approx \sqrt{4} + \frac{0.01}{2\sqrt{4}} = 2.0025$$

这与精确值 $\sqrt{4.01} = 2.0024984\ldots$ 接近的很好。 ∎

有了导数与微分的关系，根据基本函数的导法公式，可以得出对应的微分公式

1. $y = c$, $\mathrm{d}y = 0$.
2. $y = x^p (p \in \mathbb{R}, x > 0)$, $\mathrm{d}y = px^{p-1}\,\mathrm{d}x$.
3. $y = a^x (a > 0, a \neq 1)$,$\mathrm{d}y = a^x \ln a\,\mathrm{d}x$, 特殊情况:$\mathrm{d}(\mathrm{e}^x) = \mathrm{e}^x\,\mathrm{d}x$.
4. $y = \log_a x (a > 0, a \neq 1, x > 0)$,$\mathrm{d}y = \frac{\mathrm{d}x}{x \ln a}$，特殊情况:$\mathrm{d}(\ln x) = \frac{\mathrm{d}x}{x}$.
5. 三角函数

$$\mathrm{d}(\sin x) = \cos x\,\mathrm{d}x,\ \mathrm{d}(\cos x) = -\sin x\,\mathrm{d}x,\ \mathrm{d}(\tan x) = \sec^2 x\,\mathrm{d}x$$

6. 反三角函数

$$\mathrm{d}(\arcsin x) = \frac{\mathrm{d}x}{\sqrt{1-x^2}},\ \mathrm{d}(\arccos x) = -\frac{\mathrm{d}x}{\sqrt{1-x^2}},\ \mathrm{d}(\arctan x) = \frac{\mathrm{d}x}{1+x^2}$$

以及微分法则

1. $\mathrm{d}(cf(x)) = c\,\mathrm{d}(f(x))$，这里 C 是常数.
2. $\mathrm{d}(f(x) \pm g(x)) = \mathrm{d}(f(x)) \pm \mathrm{d}(g(x))$.
3. $\mathrm{d}(f(x)g(x)) = \mathrm{d}(f(x))g(x) + f(x)\,\mathrm{d}(g(x))$.
4. $\mathrm{d}\left(\frac{f(x)}{g(x)}\right) = \frac{\mathrm{d}(f(x))g(x) - f(x)\,\mathrm{d}(g(x))}{f^2(x)}$.

由反函数的求导法则，我们可以建立反函数的微分法则，设 $y = f(x)$ 的反函数为 $x = \varphi(y)$，那么有

$$\mathrm{d}y = f'(x)\,\mathrm{d}x,\ \mathrm{d}x = \varphi'(y)\,\mathrm{d}y$$

而 $f'(x)\varphi'(y) = 1$，所以

$$\mathrm{d}x = \frac{\mathrm{d}y}{f'(x)}$$

导数有高阶导数，微分也可以建立高阶微分的概念，对函数 $f(x)$ 来说，它的微分

$$\mathrm{d}y = f'(x)\,\mathrm{d}x$$

既是 x 的函数，同时也是自变量增量 $\mathrm{d}x$ 的函数，现在把 x 看成主元，即保持 $\mathrm{d}x$ 为常量，再次进行微分，得

$$\mathrm{d}^2y = \mathrm{d}(f'(x)\,\mathrm{d}x) = \mathrm{d}(f'(x))\,\mathrm{d}x = f''(x)\,\mathrm{d}x^2$$

这里 $\mathrm{d}x^2 = (\mathrm{d}x)^2$ 是一个特别记号，以与函数 $y = x$ 的二阶微分符号 d^2x 相区别。上式就是函数 $f(x)$ 的二阶微分，同样，还可以求三阶微分

$$\mathrm{d}^3y = \mathrm{d}(\mathrm{d}^2f(x)) = \mathrm{d}(f''(x)\,\mathrm{d}x^2) = f^{(3)}(x)\,\mathrm{d}x^3$$

容易得知，对任意正整数 n，有

$$\mathrm{d}^ny = f^{(n)}(x)\,\mathrm{d}x^n$$

这就是 n 阶微分，而这符号也正是高阶导数符号的来源。

2.1.10 微分的形式不变性

前面已经讲过，设复合函数 $h(x) = f(g(x))$ 是由 $y = f(u)$ 及 $u = g(x)$ 复合而成，那么有求导公式

$$h'(x) = f'(g(x))g'(x)$$

因此它的微分为

$$\mathrm{d}(h(x)) = f'(g(x))g'(x)\,\mathrm{d}x$$

但是我们知道，上式右端中的因子 $g'(x)\,\mathrm{d}x$ 正是函数 $g(x)$ 的微分，即

$$\mathrm{d}(g(x)) = g'(x)\,\mathrm{d}x$$

因此前一式子可以写成

$$\mathrm{d}(f(g(x))) = f'(g(x))\,\mathrm{d}(g(x))$$

这与函数 $f(x)$ 的微分表达式 $\mathrm{d}(f(u)) = f'(u)\,\mathrm{d}u$ 一模一样，只是在其中把 u 用 $u = g(x)$ 替换掉了而已，这表明一个事实，在微分表达式中，如果一个变量又是别的变量的函数，可以在替换变量的同时把该变量的微分替换为相应函数的微分就行，微分表达式仍然成立，这就是*微分的形式不变性*.

复合函数的导数规则看起来不如微分的形式不变性来得简单。

但是对于高阶微分，这个性质就不再了，以二阶微分为例，函数 $y = f(u)$ 的二阶微分是

$$\mathrm{d}^2y = f''(u)\,\mathrm{d}u^2$$

但假若 $u = g(x)$，那么现在的二阶微分是

$$\mathrm{d}^2y = \mathrm{d}(\mathrm{d}(f(g(x)))) = \mathrm{d}(f'(g(x))g'(x)\,\mathrm{d}x) = (f''(g(x))g'(x) + f'(g(x))g''(x))\,\mathrm{d}x^2$$

但若直接在 $y = f(u)$ 的二阶微分中以 $u = g(x)$ 替换，则得出

$$f''(g(x))(\mathrm{d}g(x))^2 = f''(g(x))g'^2(x)\,\mathrm{d}x^2$$

可知两者并不相等，因此，高阶微分不再具有形式不变性。

2.1.11 微分用于近似计算

微分提供了一个近似计算的思路，设函数 $y = f(x)$ 在 $x = x_0$ 处可微，则在 x_0 处给自变量一个微小的增量 Δx，函数值也相应有一个增量 Δy，有

$$\Delta y = f(x_0 + \Delta x) - f(x_0) = f'(x_0)\Delta x + o(\Delta x)$$

于是

$$f(x_0 + \Delta x) = f(x_0) + f'(x_0)\Delta x + o(\Delta x)$$

于是有如下的近似公式

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x$$

这就是说，可以用曲线在某点处的切线来逼近它，并且由于误差是 Δx 的高阶无穷小，所以相对误差可以任意小，只要 Δx 足够小。

例 2.1.19 可以建立如下的近似公式，当 $|x|$ 非常小时，有

$$\begin{aligned}
(1+x)^\alpha &\approx 1 + \alpha x \\
\mathrm{e}^x &\approx 1 + x \\
\ln(1+x) &\approx x \\
\sin x &\approx x \\
\tan x &\approx x
\end{aligned}$$

其中第一个式子中，取 $\alpha = \frac{1}{2}$，可得

$$\sqrt{1+x} \approx 1 + \frac{1}{2}x$$

■

2.2 微分中值定理

2.2.1 费马极值定理

定理 2.2.1. *设函数 $f(x)$ 在 x_0 的某邻域内有定义，如果它在 x_0 处取得极值 (极大或者极小均可) 且在 x_0 处可导，则必有 $f'(x_0) = 0$.*

证明 假如函数在 x_0 处取的是极大值，用反证法证明该点处如果可导，则导数只能是零，这是因为，假设 $f'(x_0)>0$，则必定存在 x_0 的某个充分小的邻域，在此邻域上恒有

$$\frac{f(x)-f(x_0)}{x-x_0}>0$$

于是在这个充分小的邻域的右侧部分，就有 $f(x)>f(x_0)$，这与函数在 x_0 处达到极大值相矛盾. 同理，如果 $f'(x_0)<0$，则在 x_0 的某个充分小的左邻域上，亦必有 $f(x)>f(x_0)$，同样导致矛盾，因此，如果 $f'(x_0)$ 存在，则只能是零。 □

2.2.2 罗尔定理

定理 2.2.2 (罗尔 (Rolle) 定理). *设函数 $f(x)$ 在闭区间 $[a,b]$ 内连续，在开区间 (a,b) 内可导，且有 $f(a)=f(b)$，则开区间 (a,b) 上至少存在一个点 x_0，使得 $f'(x_0)=0$.*

证明 因为闭区间上的连续函数必定同时存在着最大值和最小值，如果两个最值相等，则函数值恒为常数，此时结论自然是成立的，如果最大值和最小值不相等，则两个最值必然有一个与两个端点处的函数值不同，从而只能在开区间内取得，于是按费马极值定理，该点处的导数便是零。 □

2.2.3 拉格朗日中值定理

定理 2.2.3 (拉格朗日 (Lagrange) 中值定理). *设函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，在开区间 (a,b) 内可导，则存在 $x_0\in(a,b)$，使得*

$$f'(x_0)=\frac{f(b)-f(a)}{b-a}$$

这定理的几何意义是，闭区间上的连续函数的图象上，存在某个点处的切线平行于两个端点的连线。

证明 两个端点相连直线对应的线性函数是

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

构造函数

$$L(x)=f(x)-g(x)$$

容易验证，函数 $L(x)$ 满足罗尔定理的条件，由罗尔定理即得结论。 □

在证明了拉格朗日定理之后，我们注意到一个的现象，罗尔定理是拉格朗日定理的特殊情形，但由特殊情形的罗尔定理可以得到一般情形的拉格朗日定理，这与我们通常的认识相违背，但这揭示了一种特殊与一般之间的等价关系，这在数学上很多地方都可以看到这种关系，有时候为了证明一个定理，我们往往先证明定理的特殊情形，再由特殊情形通过某种变换，可以得出一般情形的结论。

2.2.4 柯西中值定理

定理 2.2.4 (柯西 (Cauchy) 中值定理). 设函数 $f(x)$ 和 $g(x)$ 都在闭区间 $[a,b]$ 上连续且在开区间 (a,b) 内可导，并且两个函数的导数不同时为零以及 $g(a)\neq g(b)$，则存在 $x_0\in(a,b)$，使得

$$\frac{f'(x_0)}{g'(x_0)}=\frac{f(b)-f(a)}{g(b)-g(a)}$$

证明 同拉格朗日定理相仿，作函数

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

再作

$$C(x)=f(x)-h(x)$$

则可以验证，$C(x)$ 符合罗尔定理的条件，由罗尔定理即得结论。 □

柯西中值定理可看作是参变量曲线的拉格朗日定理，假定一条曲线由参数方程

$$\begin{cases}x=g(t)\\ y=f(t)\end{cases}$$

确定，参数范围是 $a\leqslant t\leqslant b$，那么它在 $t=t_0$ 处的切线斜率便是 $\frac{f'(t_0)}{g'(t_0)}$，这定理表明曲线上存在某点，该点处的切线平行于曲线两个端点的连线。

2.2.5 导函数的进一步性质

定理 2.2.5 (导函数介值定理). 设函数 $f(x)$ 在闭区间 $[a,b]$ 内可导，且 $f'_+(a)\neq f'_-(b)$，则对于介于 $f'_+(a)$ 与 $f'_-(b)$ 之间的任意实数 k，都存在 $x_0\in(a,b)$，使得 $f'(x_0)=k$.

证明 作函数 $g(x)=f(x)-kx$，则 $g'(x)=f'(x)-k$，显然 $g'(a)g'(b)<0$，不妨假设是 $g'(a)>0$ 而 $g'(b)<0$，则函数 $g(x)$ 在 a 的某个右邻域上单调增加，同时在 b 的某个左邻域上单调减小，因此，连续函数 $g(x)$ 必然在开区间 (a,b) 内的某点处达到它的最大值，记此点为 x_0，由费马极值定理可知 $g'(x_0)=0$，即 $f'(x_0)=k$. □

定理 2.2.6. 设函数 $f(x)$ 在 x_0 的某邻域内连续，且在其空心邻域内可导，如果导数函数 $f'(x)$ 在 x_0 处存在左右极限，则函数在 x_0 处存在左导数和右导数，并且有

$$\begin{aligned}f'_-(x_0)&=\lim_{x\to x_0^-}f'(x)\\ f'_+(x_0)&=\lim_{x\to x_0^+}f'(x)\end{aligned}$$

证明 只证明右侧导数的部分，左导数也是完全类似的。设导函数 $f'(x)$ 在 x_0 处的右极限是 A，只要证明 $f(x)$ 在 x_0 处右可导且导数值也是 A 即可，设 $x>x_0$，则函数在 $[x_0,x]$ 上显然满足拉格朗日中值定理的条件，于是存在 $x_1\in(x_0,x)$，使得

$$\frac{f(x)-f(x_0)}{x-x_0}=f'(x_1)$$

当 $x \to x_0^+$ 时，显然亦必有 $x_1 \to x_0^+$，而导函数 $f'(x)$ 在 x_0 处有右极限为 A，所以

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x_1 \to x_0^+} f'(x_1) = A$$

这就表明函数 $f(x)$ 在 x_0 处右可导，且导数值为 A. □

这两个定理表明，导函数不可能有第一类间断点，在一定程度上具有连续函数的某些性质，也正是因此，不是任何一个函数都可以成为某个函数的导函数的。

2.2.6 洛必达法则

洛必达法则是一系列函数极限公式的统称。

第一个定理是关于 $\frac{0}{0}$ 型的不定式的。

定理 2.2.7. *设 $f(x)$ 与 $g(x)$ 都在 x_0 的某个空心邻域内可导，且*

1. $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0$.
2. *在点 x_0 的某空心邻域内两者都可导，且 $g'(x) \neq 0$.*
3. $\lim_{x \to x_0} \frac{f'(x)}{g'(x)} = A$，*这里 A 可以是有限实数，也可以是无穷大或者带符号的无穷大.*

则有

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = A$$

证明 如果 $f(x_0)$ 与 $g(x_0)$ 不为零，则对它们进行连续开拓，命 $f(x_0) = g(x_0) = 0$，于是利用柯西中值定理，对任意 $x \neq x_0$ 都存在介于 x_0 与 x 之间的 x_1 使得

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(x_1)}{g'(x_1)}$$

再令 $x \to x_0$，亦必有 $x_1 \to x_0$，于是便得出结论。 □

例 2.2.1 今来证明

$$\lim_{x \to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$$

可以验证分子分母两个函数符合上述定理的条件，并且导函数在 $x = 0$ 处的极限是

$$\lim_{x \to 0} \frac{\sin x}{2x} = \frac{1}{2}$$

因此便得结论，这个极限的思路是这样来的，因为余弦函数在 $x = 0$ 处是连续的，所以我们知道 $1 - \cos x$ 是一个无穷小，即

$$\cos x = 1 + o(1)$$

我们将这个无穷小与 x 这个简单的无穷小进行比较，也就是考虑

$$\lim_{x \to 0} \frac{1 - \cos x}{x}$$

根据前面证明，显然这个极限是零，也就是说 $1-\cos x$ 是 x 的高阶无穷小，即

$$\cos x = 1 - o(x)$$

于是我们再将它与 x^2 进行比较，前述结果表明它俩是同阶无穷小，或者说，$1-\cos x$ 与 $\frac{1}{2}x^2$ 是等价无穷小，于是得

$$\cos x = 1 - \frac{1}{2}x^2 + o(x^2)$$

这个过程还可以继续进行下去,只要得到一个无穷小,就将它与最简单的无穷小 $x, x^2, \ldots, x^n, \ldots$ 依次进行比较，直到得出同阶无穷小为止，然后再反复进行此步骤，依次剥离出无穷小的主要成分，剩下的是越来越高阶的无穷小，下面就是将 $1-\cos x$ 与 $\frac{1}{2}x^2$ 作差之后与 x^3 比较，仍然由洛必达法则可得

$$\lim_{x\to x_0} \frac{(1-\cos x) - \frac{1}{2}x^2}{x} = 0$$

说明这个差是 x 的高阶无穷小，于是将它与 x^2 进行比较，得

$$\lim_{x\to x_0} \frac{(1-\cos x) - \frac{1}{2}x^2}{x^2} = \lim_{x\to 0} \frac{\sin x - x}{2x} = \lim_{x\to 0} \frac{\sin x}{2x} - \frac{1}{2} = 0$$

说明 x^2 的阶数仍然低了，同样的办法可以知道 x^3 仍然低，这个差与 x^4 是同阶无穷小，比值的极限是 $-\frac{1}{24}$，这就是说

$$1 - \cos x - \frac{1}{2}x^2 \sim -\frac{1}{24}x^4$$

或者写成

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + o(x^4)$$

因为余弦函数存在任意阶的导数，所以这个过程可以无限进行下去，接下来进行几步之后看起来是这样的

$$\cos x = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \frac{1}{40320}x^8 + o(x^8)$$

这意味着我们可以用多项式函数逐步逼近余弦函数，同样，对于正弦函数，我们也可以用同样的办法，可以得出

$$\sin x = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + o(x^7)$$

这提供了一个非常重要的思路，就是可以用多项式函数来无限逼近一个函数，这就是后面我们推导函数的多项式展开即泰勒公式的思想。 ■

例 2.2.2 依照上例，从我们已知的极限

$$\lim_{x\to 0} \frac{\mathrm{e}^x - 1}{x} = 1$$

出发，可以得出

$$\mathrm{e}^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5 + o(x^5)$$

同样，由

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=1$$

出发，可以得出

$$\ln(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3+\frac{1}{4}x^4-\frac{1}{5}x^5+o(x^5)$$

■

第二个是关于 $\frac{\infty}{\infty}$ 型的不定式.

定理 2.2.8. *设函数 $f(x)$ 与 $g(x)$ 都在 x_0 右侧的某个空心邻域内可导，且*

1. $\lim_{x\to x_0^+} f(x)=\lim_{x\to x_0^+} g(x)=\infty$.
2. *两个函数都在 x_0 的该空心邻域内可导，且 $g'(x)\neq 0$.*
3. $\lim_{x\to x_0^+}\frac{f'(x)}{g'(x)}=A$，*这里 A 可以是实数，也可以是无穷大或者带符号的无穷大.*

证明 在 x_0 的右侧任意取定一个数 x_1，设 a_n 是一个数列，它的所有项都被包含在开区间 (x_0,x_1) 上并且严格减少并收敛到 x_0，那么两个函数在这个数列上对应着两个值的数列:$f(a_n)$ 和 $g(a_n)$，显然这两个数列都是无穷大，并且根据柯西中值定理，存在 $x_n\in(a_{n+1},a_n)$，使得

$$\frac{f(a_{n+1})-f(a_n)}{g(a_{n+1})-g(a_n)}=\frac{f'(x_n)}{g'(x_n)}$$

显然当 $n\to\infty$ 时 $x_n\to x_0$，因而有

$$\lim_{n\to\infty}\frac{f(a_{n+1})-f(a_n)}{g(a_{n+1})-g(a_n)}=A$$

于是由 Stolz 定理，可知

$$\lim_{n\to\infty}\frac{f(a_n)}{g(a_n)}=A$$

然后由于数列 a_n 的任意性以及数列极限与函数极限的关系，可得

$$\lim_{x\to x_0^+}\frac{f(x)}{g(x)}=A$$

□

从证明过程可以看出，这个定理与 Stolz 定理非常相似，Stolz 定理可以看作是这个定理的离散版本，而这个定理可以看作是 Stolz 定理的连续版本。

2.2.7 泰勒公式与麦克劳林公式

在洛必达法则的例子中已经看到，反复应用洛必达法则可以得出一个函数的多项式逼近，这一节就来讨论这个问题。

假定函数 $f(x)$ 在 x_0 处具有各阶导数（即直到接下来出现到的阶的导数），那么由连续性，$f(x)-f(x_0)$ 必然是一个无穷小 $(x\to x_0)$，为了寻求这个无穷小的阶，将其与 $x-x_0$ 这个基本无穷小进行比较，即求极限

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$$

由洛必达法则，假定函数的一阶导数连续 (尤其是二阶可导)，上式的极限就是 $f'(x_0)$，于是

$$\frac{f(x)-f(x_0)}{x-x_0}=f'(x_0)+o(1)$$

或者改写为

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+o(x-x_0)$$

记

$$T_1(x)=f(x_0)+f'(x_0)(x-x_0)$$

显然 $T_1(x)$ 是一个一次多项式，并且与 $f(x)$ 在 x_0 处具有相同的函数值以及一阶导数值，即

$$f(x_0)=T_1(x_0),\ f'(x_0)=T_1'(x_0)$$

再考虑 $f(x)-T_1(x)$，它是 $x-x_0$ 的高阶无穷小，因此将它与 $(x-x_0)^2$ 比较，即求极限

$$\lim_{x\to x_0}\frac{f(x)-T_1(x)}{(x-x_0)^2}$$

假定 $f(x)$ 二阶可导并且二阶导数连续，连续两次应用洛必达法则，得

$$\lim_{x\to x_0}\frac{f(x)-T_1(x)}{(x-x_0)^2}=\lim_{x\to x_0}\frac{f'(x)-T_1'(x)}{2(x-x_0)}=\lim_{x\to x_0}\frac{f''(x)}{2}=\frac{1}{2}f''(x_0)$$

于是

$$\frac{f(x)-T_1(x)}{(x-x_0)^2}=\frac{1}{2}f''(x_0)+o(1)$$

或者写成

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2+o((x-x_0)^2)$$

并记

$$T_2(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2$$

则 $T_2(x)$ 与 $f(x)$ 在 x_0 处具有相同的函数值以及一阶导数和二阶导数。

同样，假定函数存在连续的三阶导数，那么还有

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\frac{f'''(x_0)}{3!}(x-x_0)^3+o((x-x_0)^3)$$

如果还具有连续的四阶导数，则还有

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\frac{f'''(x_0)}{3!}(x-x_0)^3+\frac{f^{(4)}(x_0)}{4!}(x-x_0)^4+o((x-x_0)^4)$$

照此下去，如果函数 $f(x)$ 在 x_0 的邻域上存在直到 n 阶的连续导函数，那么

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}x^n+o((x-x_0)^n)$$

记

$$T_n(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}x^n$$

这个多项式就称为函数 $f(x)$ 在 x_0 处的泰勒多项式，它的主要特征是，它与 $f(x)$ 在 x_0 处具有相同的函数值以及一阶导数、二阶导数直到 n 阶导数 (函数值也可以被视为零阶导数)，即

$$f(x_0)=T_n(x_0),\ f'(x_0)=T_n'(x_0),\ f''(x_0)=T_n''(x_0),\ \ldots,\ f^{(n)}(x_0)=T_n^{(n)}(x_0)$$

而函数的展开式则可以写成

$$f(x)=T_n(x)+o((x-x_0)^n)$$

称为函数 $f(x)$ 在 x_0 处的泰勒展开公式.

于是我们得到如下的著名定理

定理 2.2.9 (泰勒 (Taylor) 定理)**.** *如果函数 $f(x)$ 在 x_0 的邻域内有直到 n 阶的导函数，作泰勒多项式*

$$T_n(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}x^n$$

则有

$$f(x_0)=T_n(x_0),\ f'(x_0)=T_n'(x_0),\ f''(x_0)=T_n''(x_0),\ \ldots,\ f^{(n)}(x_0)=T_n^{(n)}(x_0)$$

并且有下式成立

$$f(x)=T_n(x)+o((x-x_0)^n)$$

这就是 (n 阶) 泰勒 (Taylor) *公式，其中的余项 $o((x-x_0)^n)$ 称为佩亚诺余项.*

我们用数学归纳法来证明.

证明 我们记

$$R(x)=f(x)-T_n(x)$$

显然有

$$R(x_0)=R'(x_0)=\cdots=R^{(n)}(x_0)=0$$

因此只要证明在这个条件下有

$$R(x)=o((x-x_0)^n)$$

就可以了。

当 $n=1$ 时，由导数意义，有

$$\frac{R(x)-R(x_0)}{x-x_0}=R'(x_0)+o(1)$$

注意到 $R(x_0)=R'(x_0)=0$，所以

$$R(x)=o(x-x_0)$$

结论成立，假定定理对于 n 也是成立的，则看 $n+1$ 的情形，此时设

$$S(x) = R'(x)$$

就有

$$S(x_0) = S'(x_0) = \cdots = S^{(n)}(x_0) = 0$$

因此由归纳假设，有

$$S(x) = o((x-x_0)^n)$$

再由洛必达法则，有

$$\lim_{x\to x_0} \frac{R(x)}{(x-x_0)^{n+1}} = \lim_{x\to x_0} \frac{S(x)}{(n+1)(x-x_0)^n} = 0$$

即

$$R(x) = o((x-x_0)^{n+1})$$

由归纳法，定理成立。 □

注意到，在前面的推导过程中，我们要求函数具有直到 n 阶的连续导函数，但在这个定理的证明过程中，第 n 阶导函数是否连续并没有用到，这是为什么呢，用前面推导到二阶展开式那里的情况说明，那时有

$$\lim_{x\to x_0} \frac{f(x)-T_1(x)}{(x-x_0)^2} = \lim_{x\to x_0} \frac{f'(x)-T_1'(x)}{2(x-x_0)} = \lim_{x\to x_0} \frac{f''(x)}{2} = \frac{1}{2}f''(x_0)$$

最后一步，求极限 $\lim_{x\to x_0} f''(x_0)$，那里我们是假定二阶导函数连续，得出这个极限为 $f''(x_0)$ 的，但实际上由导函数连续性定理可知，如果这个极限存在，设为 K，那么函数在 x_0 处必然存在二阶导数，且 $f''(x_0) = K$，所以我们不要求二阶导数在 x_0 处连续，仅要求 $f''(x_0)$ 存在就行，只要存在就必然连续，所以定理中的条件就减弱为在 x_0 的邻域上存在直到 n 阶的导函数。

在泰勒公式中，取 $x_0 = 0$，我们就得到如下的*麦克劳克 (Maclaurin) 公式*:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + o(x^n)$$

例 2.2.3 既然公式是函数的多项式展开，那么多项式函数自己的展开式会是什么样呢？设多项式函数

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

先写出它的 n 阶泰勒多项式

$$T_n(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}x^n$$

可以发现，它也是一个 n 次多项式，按泰勒定理，这两个多项式相差一个 $(x-x_0)^n$ 的高阶无穷小，我们指出，对于 n 次多项式函数 $f(x)$ 而言，它与它的 n 阶泰勒展开式的差值是零，也就是

$$f(x) = T_n(x)$$

我们利用归纳法证明下面的结论，有了下面这个结论，这里的结论也就成立了。

命题 2.2.1. 如果两个次数不超过 n 次的多项式 $f(x)$ 与 $g(x)$ 在 x_0 处直到 n 阶的导数值都相同，即

$$f'(x_0)=g'(x_0),\ f''(x_0)=g''(x_0),\ \ldots,\ f^{(n)}(x_0)=g^{(n)}(x_0)$$

则这两个多项式仅相差一个常数，即 $f(x)-g(x)=C$，这里 C 是一个固定的常数.

显然结论对于常数多项式和一次多项式均成立，假设结论对于 $\leqslant n$ 的情况都成立，现在假定 $f(x)$ 和 $g(x)$ 都是 $n+1$ 次多项式，并且在 x_0 处两者具有相同的直到 $n+1$ 阶的导数，那么它们的导数 $f'(x)$ 和 $g'(x)$ 在 x_0 处就具有相同的直到 n 阶导数，根据归纳假设，存在常数 C_1，使得 $f'(x)-g'(x)=C_1$，但由于 $f'(x_0)=g'(x_0)$，所以 $C_1=0$，即 $f'(x)=g'(x)$，于是对于多项式 $h(x)=f(x)-g(x)$ 有 $h'(x)=0$，于是 $h(x)$ 只能是常数多项式，即 $h(x)=f(x)-g(x)=C$. ■

有了这结论之后，由于 n 次多项式 $f(x)$ 与它的 n 阶泰勒多项式在 x_0 具有相同的直到 n 阶的导数，因此它俩仅相差一个常数，但又由于 $f(x_0)=T_n(x_0)$，所以有 $f(x)=T_n(x)$.

2.2.8 基本函数的泰勒展式

在这一小节，我们来求一些基本函数在 $x=0$ 处的泰勒展开式，即麦克劳林展开式。

1. 幂函数 $f(x)=x^p(x>0)$.

由于它已经是在 $x=0$ 处的展开式了，所以我们考虑它在 $x=1$ 处的泰勒展开，在作代换 $t=x-1$ 之后，实际就是要求函数 $f(t)=(1+t)^p$ 在 $t=0$ 处的展开式，我们还是用 x 来替换 t，那么根据前面已经求得的高阶导数值，有

$$(1+x)^p=1+px+\frac{p(p-1)}{2!}x^2+\cdots+\frac{p(p-1)\cdots(p-n+1)}{n!}x^n+o(x^n)$$

显然，如果 p 是正整数，由于 n 次多项式的 n 阶导数成为常数，更高阶的导数则恒零，因此这个展开式就成为二项式定理了

$$(1+x)^n=1+nx+\frac{n(n-1)}{2!}x^2+\frac{n(n-1)(n-2)}{3!}x^3+\cdots+x^n$$

因此这式子可以看作是二项式定理的推广，它被称为牛顿二项式定理.

在式中令 $p=-1$，并将 x 替换为 $-x$，得

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+o(x^n)$$

同样，取 $p=\frac{1}{2}$，得

$$\sqrt{1+x}=1+\frac{1}{2}x-\frac{1}{8}x^2+\frac{3}{48}x^3+o(x^3)$$

取 $p=-\frac{1}{2}$，得

$$\frac{1}{\sqrt{1+x}}=1-\frac{1}{2}x+\frac{3}{8}x^2-\frac{15}{48}x^3+o(x^3)$$

1. 指数函数 $f(x)=\mathrm{e}^x$.

根据在高阶导数那一小节我们求得的指数函数的高阶导数公式，我们得出

$$\mathrm{e}^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + o(x^n)$$

2. 对数函数 $f(x)=\ln x$.

由于对数函数在 $x=0$ 处无定义，所以我们考虑它在 $x=1$ 处的展开式，实际上就是求 $\ln(1+x)$ 在 $x=0$ 处的展开式，这时有

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

3. 三角函数

对正弦函数，有

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

对于余弦函数，则是

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

2.2.9 余项的其它形式

2.2.10 插值公式

2.2.11 微分方程

我们知道指数函数 $y=f(x)=\mathrm{e}^x$ 的导函数就是它自己，用微分的形式写出来就是

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = f(x)$$

或者简单写成

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y$$

或者更加简洁的写成

$$y' = y$$

像这样把函数自身与其微分 (或者导数) 联系起来的方程，就称为*微分方程*。

上面的例子就表明指数函数 $y=\mathrm{e}^x$ 就满足微分方程

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y$$

我们还可以写出一些微分方程

$$\begin{aligned} y'' + y' + y &= 0 \\ y'^2 + x^2 &= 0 \\ y'' + 2y'y + y^2 &= 0 \end{aligned}$$

如果一个函数与它的导函数满足一个给定的微分方程，就称这函数是这微分方程的一个解. 例如，函数 $y=\mathrm{e}^x$ 就是微分方程 $y'=y$ 的一个解，但并不是唯一解，因为零函数也满足这方程。

例 2.2.4 通过求导可以验证，函数

$$y=C_1\cos x+C_2\sin x$$

满足微分方程

$$y''+y=0$$

这里 C_1 和 C_2 是常数。 ■

例 2.2.5 函数

$$y=C_1\cosh x+C_2\sinh x$$

满足微分方程

$$y''-y=0$$

■

例 2.2.6 可以验证，函数

$$y=C_1\mathrm{e}^{\lambda_1 x}+C_2\mathrm{e}^{\lambda_2 x}$$

满足微分方程

$$y''-(\lambda_1+\lambda_2)y'+\lambda_1\lambda_2 y=0$$

像这样由 y、y'、y'' 等各阶导数的线性组合 (系数可以是常数或者自变量 x 的表达式) 而成的微分方程，称为*线性微分方程*，如果各项的系数又是常数，则称为*常系数线性微分方程*。 ■

例 2.2.7 在这个例子中，我们来解下面这个微分方程

$$y'-\lambda y=0$$

作函数 $u=y\mathrm{e}^{-\lambda x}$，那么有

$$u'=\mathrm{e}^{-\lambda x}(y'-\lambda y)=0$$

根据前面的结果，函数 u 只能是常数函数，即

$$y\mathrm{e}^{-\lambda x}=C$$

从而

$$y=C\mathrm{e}^{\lambda x}$$

这样就解出了这个微分方程. ■

2.3 利用导数研究函数的性质

2.3.1 函数为常数的条件

关于可导函数为常数函数的条件，有如下定理

定理 2.3.1. 某区间上的可导函数的函数值恒保持为某个常数的充分必要条件是，它的导函数为零函数.

证明 先证充分性，如果函数的导函数为常数函数，那么对于定义域上的任何两个不同的实数 x_1, x_2，存在介于它俩之间的另一实数 ξ，使得

$$\frac{f(x_1)-f(x_2)}{x_1-x_2}=f'(\xi)=0$$

于是 $f(x_1)=f(x_2)$，由 x_1, x_2 的任意性，知函数为常数函数，另外，常数函数的导函数为零函数，定理得证。 □

进一步有如下结论

推论 2.3.1. 如果某区间上的两个可导函数的导函数恒相等，则这两个函数仅相差一个常数。

这只要考察两个函数的差函数就清楚了。

2.3.2 单调性与极值

导数可以用来研究函数的单调性和极值，我们先建立如下关于单调性的定理

定理 2.3.2. 如果函数 $f(x)$ 的一阶导函数 $f'(x)$ 在某个区间上恒满足 $f'(x)\geqslant 0$，那么函数在此区间上单调不减，如果不等式反向，即 $f'(x)\leqslant 0$ 恒成立，那么函数在此区间上单调不增。如果不等式中的等号总不成立或者至多仅在有限个点处成立，那么函数相应的是严格单调的。

证明 如果函数在某区间上恒成立 $f'(x)\geqslant 0$，那么对于区间上任意两个不同的数 x_1 与 x_2，设 $x_1<x_2$，按照拉格朗日中值定理，存在 $\xi\in(x_1,x_2)$，使得

$$f'(\xi)=\frac{f(x_1)-f(x_2)}{x_1-x_2}$$

由 $f'(\xi)\geqslant 0$ 即得 $f(x_1)\leqslant f(x_2)$，于是函数单调不减。函数单调不增的证明也是完全类似的。

显然，如果上面的证明中不等式 $f'(x)\geqslant 0$ 中的等号永远不成立，即永远只能取大于号，那么相应的得出的结论是 $f(x_1)<f(x_2)$，即是严格递增的。

现在来证明，如果导函数 $f'(x)\geqslant 0$ 恒成立，但等号仅在有限个点处取得，那么函数仍然是严格增加的。

首先可以肯定函数是单调不减的了，把区间划分成多个小区间，使得每个小区间上至多只有一个导函数零点，如果函数在每个小区间上是严格增加的，而连续性又保证了

端点处的接续，那么就能得出函数在整个区间上都是严格增加的，所以只需要证明导函数只在一个点处取零点的情形就可以了。

假定导函数 $f'(x)$ 仅在区间 $[a,b]$ 内的某一点 c 处取零值，采用反证明法来证明函数是严格增加的，若不然，假如存在 $x_1, x_2 \in [a,b]$ 使得 $x_1 < x_2$ 且 $f(x_1) = f(x_2)$，那么 $f'(x_1)$ 与 $f'(x_2)$ 中至多只有一个为零，而另一个必为正值，不妨设就是 $f'(x_1) > 0$，于是存在 x_1 的某个右邻域 $(x_1, x_1+\delta)$，在此右邻域上恒有 $f(x) > f(x_1) = f(x_2)$，只要把 δ 限制的充分小，就能保证 $x_1 + \delta < x_2$，于是函数在 $(x_1, x_1+\delta)$ 上的函数值就都大于 $f(x_2)$，这与函数的单调不减是矛盾的，从而得证。 □

例 2.3.1 关于定理中导函数恒非负，仅在有限个点处取零值也能保证函数的严格增加这一点，函数 $y = x^3$ 提供了一个例子，它的导函数 $y' = 3x^2 \geqslant 0$，但仅在 $x = 0$ 处取零值，并不影响原来函数是严格增加的这一事实。 ■

例 2.3.2 需要说明的是，对于函数单调不减与单调不增来讲，条件 $f'(x) \geqslant 0$ 或者 $f'(x) \leqslant 0$ 是充分必要条件（在函数可导的前提下），但对于严格单调来说，不等式成立且至多仅在有限个点处取等号，则只是充分条件而非必要条件，我们会举一个反例加以说明。

在正方形区域 $[-1,1]\times[-1,1]$ 中，用 $x = y = \pm\frac{1}{n}(n = 1, 2, \ldots)$ 划分网格，并在每一个小正方形区域 $[\frac{1}{n+1}, \frac{1}{n}] \times [\frac{1}{n+1}, \frac{1}{n}]$ 内，将左下角的角点 $\left(\frac{1}{n+1}, \frac{1}{n+1}\right)$ 与右上角的角点 $\left(\frac{1}{n}, \frac{1}{n}\right)$ 用一段正弦曲线连接起来，这段正弦曲线是将 $y = \sin x$ 在 $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ 中的部分进行平移和沿坐标轴方向的拉伸变换得来，使得正好被小正方形区域框住，这样一段一段的正弦曲线相接，构成一个全新的函数，在每一个区间 $[\frac{1}{n+1}, \frac{1}{n}]$ 上，函数的表达式是:

$$f(x) = \frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right) + \frac{1}{2n(n+1)} \sin \pi n(n+1) \left[x - \frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right)\right]$$

在 $x = 0$ 处令 $f(0) = 0$，然后将曲线中心对称到第三象限，得到另一半图象，从而得出完整的函数图象，显然这个函数是严格增加的，但导函数在每一个 $x = \pm\frac{1}{n}$ 处都取零值，我们再证明 $f'(0) = 1$ 就可以了。

设

$$x = \frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right) + t,\ |t| \leqslant \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+1}\right) = \frac{1}{2n(n+1)}$$

于是

$$\begin{aligned}
\frac{f(x) - f(0)}{x - 0} &= \frac{f(x)}{x} \\
&= \frac{\frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right) + \frac{1}{2n(n+1)} \sin \pi n(n+1)t}{\frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right) + t} \\
&= \frac{2n + 1 + \sin \pi n(n+1)t}{2n + 1 + 2n(n+1)t}
\end{aligned}$$

注意到 $|2n(n+1)t| \leqslant 1$，得出

$$\lim_{x\to 0} \frac{f(x)}{x} = 1$$

即 $f'(0) = 1$.

这个例子表明，即便导函数有无穷多个零点，函数也是有可能严格单调的，那么这个条件到底放宽到何种程度，方能得出充分必要条件呢，这个问题在实变函数的测度理论中方能给出答案。 ■

关于极值，费马极值定理已经表明，极值点处如果可导，则导数只能是零，但是如何判断导函数的零点是否是原来函数的极值点呢，有如下定理

定理 2.3.3. *设函数 $f(x)$ 在 x_0 的邻域内可导，如果它满足以下两条，那么函数在 x_0 处取极值.*

1. $f'(x_0)=0$.
2. *存在 x_0 的某个足够小的邻域，使得函数在两侧空心邻域内各自保持恒定的符号，且两侧的符号正好相反，具体的说，如果左正右负，则函数在 x_0 处取极大值，反之，若左负右正，则函数在 x_0 处取极小值。*

证明 只证明极大值的情况，如果函数 $f(x)$ 满足 $f'(x_0)=0$，且在它的某个邻域的左侧 $(x_0-\delta, x_0)$ 上有 $f'(x)>0$，而在右侧 $(x_0, x_0+\delta)$ 上有 $f'(x)<0$，我们来证明 $f(x_0)$ 是一个极大值，因为在左侧导函数恒正，函数严格增加，同理函数在右侧严格减少，所以 $f(x_0)$ 是一个极大值。极小值也同理可证。 □

例 2.3.3 注意这个定理中的条件是充分条件，但不是必要条件 (假定函数总是可导的)。实际上，极值的邻域内都并不一定有确定的单调性，可以仿照前面的例子构造相应的反例，此处从略。 ■

例 2.3.4 光的折射定律 假定光在甲、乙两种介质中的传播速度分别是 v_1 和 v_2，如图所示，图中直线是两种介质的分界面，现在光从介质甲中的 A 点发出，经分界面折射后，经过介质乙中的 B 点，光在同一种介质中是必定沿直线传播的，我们知道，光总是沿着用时最短的路径前进，那么问题就来了，光线应该在分界面上何处折射，才能使得传播用时最短？

设 A、B 两点与分界面的距离分别记为 a、b，并且沿分界面的距离是 l，假定光线到达分界面上的点 P 处，点 P 沿分界面与 A、B 的距离分别是 x 和 $l-x$，那么光线传播所用的时间是

$$f(x)=\frac{\sqrt{x^2+a^2}}{v_1}+\frac{\sqrt{(l-x)^2+b^2}}{v_2}$$

为了求得最小值，求导并令其为零，得

$$f'(x)=\frac{x}{v_1\sqrt{x^2+a^2}}-\frac{l-x}{v_2\sqrt{(l-x)^2+b^2}}$$

根据物理意义，这最小值一定存在，设当 $x=x_0$ 时用时最短，则必有 $f'(x_0)=0$，于是有

$$\frac{x_0}{v_1\sqrt{x_0^2+a^2}}=\frac{l-x_0}{v_2\sqrt{(l-x_0)^2+b^2}}$$

通常用入射角 (入射光线与法线的夹角) 和折射角（折射光线与法线的夹角）来标记点 P 的位置，而不是用 x_0，设入射角为 α，折射角为 β，上式即为

$$\frac{\sin\alpha}{v_1}=\frac{\sin\beta}{v_2}$$

光线在介质中的传播速度与在真空中的传播速度之比，就是该介质的折射率，设这两种介质的折射率分别为 n_1 和 n_2，上式就成为

$$\frac{\sin\alpha}{n_1}=\frac{\sin\beta}{n_2}$$

或者写成

$$\frac{\sin\alpha}{\sin\beta}=\frac{n_1}{n_2}$$

上式就是光线传播最短路径所应满足的条件，其被称为光的折射定律。 ■

2.3.3 证明不等式

2.3.4 函数的凸性与拐点

函数的凸性是关于函数图象的拱形特征的刻画，如果函数的图象在区间上向上拱起，则它图象上任意两点间的部分，都位于这两点连线段的上方，从而引出如下定义

定义 2.3.1. 如果定义在区间 (a,b) 上的函数 $f(x)$ 对区间上的任意两个数 x_1,x_2 以及任意满足 $\alpha+\beta=1(\alpha\geqslant 0,\beta\geqslant 0)$ 的一对实数 α、β 都成立

$$f(\alpha x_1+\beta x_2)\geqslant \alpha f(x_1)+\beta f(x_2)$$

则称函数在这区间上是上凸函数，如果式中的等号总是取不到，则称为严格上凸函数，把不等式反向，则得到下凸函数和严格下凸函数的定义。

显然，把函数乘以一个正常数，不改变凸性，若乘以一个负常数，则凸性与原来相反，若干个具有相同凸性的函数相加，结果仍具备相同的凸性。

例 2.3.5 二次函数 $y=x^2$ 在 $\mathbb{R}$ 上是下凸函数，因为

$$(\alpha x_1+\beta x_2)^2-(\alpha x_1^2+\beta x_2^2)=-\alpha\beta(x_1-x_2)^2\leqslant 0$$

并且等号仅在 $x_1=x_2$ 时成立，所以它不但是下凸，还是严格下凸，一般的二次函数通过平移，都可以化为 $y=ax^2$ 的形式，而平移不改变凸性，但相应的区间也会相应的平移，所以结论就是，如果二次函数的二次系数为正，则为下凸，否则为上凸。 ■

把定义中居于中间的 $\alpha x_1+\beta x_2$ 看成第三个变量，就得下来这个定理

定理 2.3.4. 函数 $f(x)$ 在某区间上为上凸函数的充分必要条件是，对于区间上任意三个实数 $x_1<x_2<x_3$，都成立

$$\frac{f(x_1)-f(x_2)}{x_1-x_2}\geqslant\frac{f(x_2)-f(x_3)}{x_2-x_3}$$

这只要注意到中间的 x_2 可以用 x_1 与 x_3 的加权平均表达出来就明白了:

$$x_2=\frac{x_2-x_3}{x_1-x_3}x_1+\frac{x_1-x_2}{x_1-x_3}x_3$$

于是只需要取

$$\alpha=\frac{x_2-x_3}{x_1-x_3},\ \beta=\frac{x_1-x_2}{x_1-x_3}$$

再代回到定义中的不等式，便得知定理成立。

对于连续函数而言，上面的定义与下面这个定义等价

定义 2.3.2. 如果区间上的连续函数 $f(x)$ 对区间上的任意 x_1, x_2 都成立

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

则称它是区间上的上凸函数，类似的可得到严格上凸、下凸、严格下凸的定义。

即对于连续函数来说，只需要前面定义中 $\alpha=\beta=\frac{1}{2}$ 就够了，这一点在我关于初等数学的笔记中已经证明过，这里从略。

由定义立即可以得到著名的 琴生不等式.

定理 2.3.5. 对于某区间上的上凸函数而言,在区间上任意取定 n 个实数 $x_i(i=1,2,\ldots,n)$，有

$$f\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \geqslant \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

这定理的证明同样见于我的初等数学笔记，仍然从略。

例 2.3.6 琴生不等式是一系列重要不等式的来源，例如，在后面我们将证明，对数函数 $y=\ln x$ 在 $(0,+\infty)$ 上是上凸的，于是套用琴生不等式，对任意 n 个正实数 $x_i(i=1,2,\ldots,n)$，就有

$$\ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \geqslant \frac{1}{n}\sum_{i=1}^{n} \ln x_i$$

即

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geqslant \sqrt[n]{\prod_{i=1}^{n} x_i}$$

这就是著名的均值不等式. ■

利用数学归纳法，可以得到

定理 2.3.6. 如果函数 $f(x)$ 在某区间上上凸，那么对于该区间上任意 n 个实数 $x_i(i=1,2,\ldots,n)$ 以及满足

$$\sum_{i=1}^{n} \alpha_i = 1(\alpha_i \geqslant 0, i=1,2,\ldots,n)$$

的一组权值 $\alpha_i(i=1,2,\ldots,n)$，成立不等式

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \geqslant \sum_{i=1}^{n} \alpha_i f(x_i)$$

现在我们利用导数工具来研究函数的凸性，自然的，仅限于可导函数的凸性。

定理 2.3.7. 如果函数的导函数在某区间上是单调不增的，那么函数在该区间上是上凸的。

由拉格朗日定理结合前面的定理 2.3.4就清楚了，当然，这只是充分条件，不是必要条件。

自然的也有如下结论

推论 2.3.2. 如果函数在某区间上二阶可导，且 $f''(x) \leqslant 0$，那么函数在该区间上是上凸的，如果不等式反向，则是下凸的。

例 2.3.7 由此可以知道，指数函数在 $\mathbb{R}$ 上是下凸的，而对数函数在 $(0, +\infty)$ 上是上凸的。正弦函数在 $(0, \pi)$ 上是上凸的，而在 $(-\pi, 0)$ 上是下凸的，余弦函数在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 上是上凸的，而在 $(\frac{\pi}{2}\ \frac{3\pi}{2})$ 上是下凸的。 ■

例 2.3.8 幂函数的凸性与幂平均值不等式 对于幂函数 $y = x^p$，二阶导函数 $y'' = p(p-1)x^{p-2}$，如果 $p > 1$，则它在 $(0, +\infty)$ 上是下凸函数，否则是上凸函数。我们对一般的幂函数应用琴生不等式，即对任意 n 个正实数 $x_i (i = 1, 2, \ldots, n)$，如果 $p > 1$，则有

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^p \leqslant \frac{1}{n}\sum_{i=1}^{n} x_i^p \tag{2.3.1}$$

如果 $p < 1$，不等式反向。

现在设实数 $p > q > 0$，那么由上述不等式有

$$\left(\frac{1}{n}\sum_{i=1}^{n} t_i\right)^{\frac{p}{q}} \leqslant \frac{1}{n}\sum_{i=1}^{n} t_i^{\frac{p}{q}}$$

再令 $t_i = x_i^q$，即得

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^q\right)^{\frac{1}{q}} \leqslant \left(\frac{1}{n}\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \tag{2.3.2}$$

这表明，函数

$$f(\alpha) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\right)^{\frac{1}{\alpha}}$$

关于指数 α 严格增加。

不等式式 2.3.1以及式 2.3.2称为幂平均值不等式. ■

2.3.5 方程的近似解

第三章　不定积分与定积分

3.1　定积分与不定积分的原理及两者之间的关系

在初等范围内，我们经常见到一个量对另一个量的累积 (乘积)，例如物体的面积是纵横两个方向上的累积，路程是速度对时间的累积，功是力对位移的累积，等等。但是以前我们通常只会处理最简单的情况，我们只会求规则图形的面积，不会计算不规则图形的面积，只会处理匀速运动的路，不会计算任意变速运动的路程，只会计算恒定力做功，不会计算变力做功，定积分就是为了处理这类问题而被发现的。

后来牛顿与莱布尼茨又发现了定积分与函数的原函数之间存在着直接又简单的联系，于是积分学与微积分学之间的深刻关系也被揭示在世人面前，其影响之大，使得这个结论直接被冠之以微积分学基本定理，成为微积分学的基石。

3.1.1　黎曼和与定积分的概念

我们先来看几个例子。

例 3.1.1 曲边梯形的面积

如 图 3.1 所示，定义在区间 $[a,b]$ 上的正值函数 $y = f(x)$ 的图象与直线 $x = a$、$x = b$ 以及 x 轴围成了一个曲边梯形，我们考虑它的面积，也就是函数图象下方的面积。

通过在区间 $[a,b]$ 内插入一些点把区间 $[a,b]$ 划分成 n 个小区间 (不必是等分):

$$a = x_0 < x_1 < \cdots < x_{i-1} < x_i < \cdots < x_n = b$$

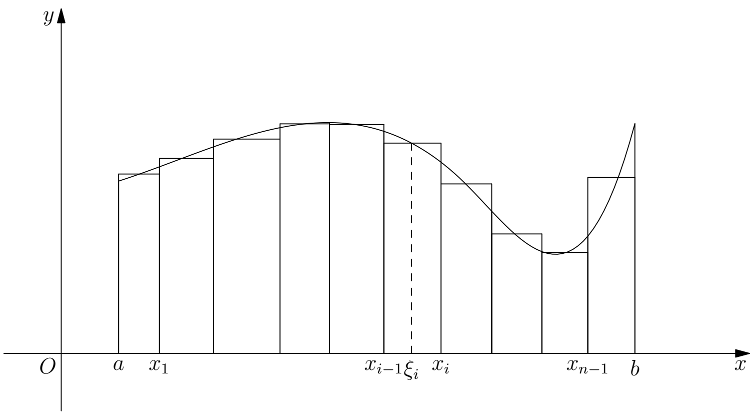

图 3.1 曲边梯形的面积

这样曲边梯形就被划分成了 n 个小的曲边梯形，设曲边梯形面积是 S，第 i 个区间上的曲边梯形面积是 S_i，那么

$$S = \sum_{i=1}^{n} S_i$$

对于每个小曲边梯形，所以我们用矩形的面积来近似其面积，在每个小区间上任意各取定一个点:$\xi_i \in [x_{i-1}, x_i](i = 1, 2, \ldots, n)$，用 $y_i = f(\xi_i)$ 来作矩形的上边界，得到一个小矩形，其面积是 $f(\xi_i)\Delta x_i$，这里 $\Delta x_i = x_i - x_{i-1}$ 是第 i 个小区间的长度，于是

$$S_i \approx f(\xi_i)\Delta x_i$$

那么曲边梯形的总面积，就可以用这 n 个小矩形的面积之和来近似代替:

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

显然，小区间越多，各个小区间长度越小，这个误差就越小，考虑划分的极限情况，即划分的粒度 (各小区间的最大长度) 趋于零时，上式右端将以左边为极限。 ■

例 3.1.2 变速直线运动的位移 假定某质点做变速直线运动，其速度是时间的函数 $v = v(t)$，现在我们要计算它从 $t = a$ 到 $t = b$ 这段时间段内发生的位移. 如果速度恒定不变，那么只要将速度与时间段长度相乘即可得出结果，但是现在速度是个时时刻刻都在改变的量，这个办法行不通了，但是我们可以将该时间段划分成很多很小的时间内，在每一个很小的时间内，速度的改变量很小，因此可以近似的看成是匀速运动，从而质点在这很小的时间内所发生的位移是可以近似求出的，于是总的位移也就可以近似求出。

在时间段 $t \in [a, b]$ 内插入 $n-1$ 个时间点将其划分为 n 个小的时间区间 (不必等分):

$$a = t_0 < \cdots < t_{i-1} < t_i < \cdots < t_n = b$$

在每个小区间段内随机取一个点 $t = \xi_i$，那么质点在该小的时间段内可以视为以大小为 $v(\xi_i)$ 的速度匀速运动，于是在这个小区间上发生的位移 S_i 便有近似

$$S_i \approx v(\xi_i)\Delta t_i$$

这里 $\Delta t_i = t_i - t_{i-1}$ 是该小时间段的长度. 于是质点在区间 $[a, b]$ 内所发生的总位移有如近似:

$$S = \sum_{i=1}^{n} S_i \approx \sum_{i=1}^{n} v(\xi_i)\Delta t_i$$

如果时间段划分得越小，那么上式的误差就越小，在各小时间段的最大长度趋于零时，误差即趋于零。 ■

例 3.1.3 变力做功 考虑一个变化的力 F 推动某物体移动了一段直线路程 L，我们把这一段直线路程对应为区间 $[a, b]$，这里 $L = b - a$，而变力 F 在该路程每个位置处的大小是 $F = F(l), l \in [a, b]$，同样的，我们把这段路程划分成很多小区间:

$$a = l_0 < \cdots l_{i-1} < l_i < \cdots < b = l_n$$

在每个小区间上可以近似是恒力做功 (只要区间划分得很小，那么在小区间上力的改变量也可以很小)，在每个小区间上任意取一点 $\xi \in [x_{i-1}, x_i]$，在这一段小区间上可以视为以恒力 $F_i = F(\xi_i)$ 进行的恒力做功，因此这一段小区间上做的功有近似:

$$W_i \approx F(\xi_i)\Delta l_i$$

其中 $\Delta l_i = l_i - l_{i-1}$ 是小区间长度. 变力 F 在区间 $[a, b]$ 内总共所做的功

$$W = \sum_{i=1}^{n} W_i \approx \sum_{i=1}^{n} F(\xi_i)\Delta l_i$$

■

总结上面三个例子，有很多情况下我们需要求出一个量 A 对另一个量 B 的累积 P(即乘积)，如果量 A 不依赖于量 B 的变化而变化，那么我们可以很容易的得出这个乘积:

$$P = A \cdot \Delta B = A(B' - B_0)$$

但是通常量 A 它不是恒定不变的，而是依赖于量 B 的，或者说，它是量 B 的函数 $A = A(B)$，这时为了求出这个累积量，可以通过将量 B 的变化区间 $[B_0, B']$ 划分成许多小区间，然后在每个小区间上将量 A 视为恒定不变的，从而该小区间上的累积量 P_i 可以近似得出，于是在整个区间上的累积量 P 也就可以用它们之和来近似代替，只要把区间划分得足够细，就可以使得误差足够小，在小区间的最大长度趋于零的极限情况，就得出了累积量的精确值了，从这个意义上说，定积分其实就是一种广义的乘法运算。

这里需要说明的一点是，这个近似的误差，是在小区间的最大长度趋于零的情况下才会足够小，而不是小区间数目趋于无穷的情况下，因为即便小区间数目再多，也不能保证有个别区间长度比较大，从而该小区间上的误差无法解决，例如用 $x_i = \frac{1}{i}$ 来划分区间 $[0, 1]$，这一点务必注意。

现在，是时候给出定积分的定义了:

定义 3.1.1 (积分和与定积分)**.** *设函数 $f(x)$ 在闭区间 $[a, b]$ 上有定义，在该闭区间上插入一系列分点:*

$$a = x_0 < \cdots < x_{i-1} < x_i < \cdots < x_n = b$$

将之划分为 n 个小区间，并在每一个小区间 $[x_{i-1}, x_i]$ 上任意取一个点 $\xi_i \in [x_{i-1}, x_i]$，作和式 (称为积分和或黎曼和)

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

这里 $\Delta x_i = x_i - x_{i-1}$ 是第 i 个小区间的长度。如果存在一个数 P，它满足: 对于任意小的正实数 $\varepsilon > 0$，都存在另一个正实数 $\delta > 0$，使得所有划分中，只要它的最大小区间的长度小于 δ，那么无论怎么选取各个小区间上的 ξ_i 的值，都有:

$$\left| \sum_{i=1}^{n} f(\xi_i)\Delta x_i - P \right| < \varepsilon$$

那么就称函数 $f(x)$ 在闭区间 $[a,b]$ 上是可积的，而数 P，就称为是函数 $f(x)$ 在闭区间 $[a,b]$ 上的定积分值，用下面符号表示:

$$P = \int_a^b f(x)dx$$

其中的 a、b 称为积分上限和积分下限.

关于这个定义，作几点说明.

定义的后半部分其实就是极限的 $\varepsilon - \delta$ 语言，为什么不直接使用极限符号呢，因为这里有一个问题是，极限符号只能体现出小区间的最大长度趋于零，但却体现不出各小区间上的点 ξ_i 取法的任意性。

定积分的这个符号，其实它就是黎曼和的极限形式，黎曼和的 Σ 直接变成了被拉长的字母 S(sum 代表求和)，黎曼和上下标变成了区间的上下限，而黎曼和中的区间长度 Δx_i 变成了自变量 x 的微分，可以理解为无限小的区间长度，这么一看，定积分符号就是黎曼和的极限形式就容易理解了。

所谓定积分，就是指定了积分上限和积分下限的积分，后面我们还见到不定积分，即没有指定上下限的积分，积分这个概念，指的就是函数 $f(x)$ 对自变量 x 的累积。

前面举了曲边梯形面积的例子，需要注意的是，如果函数 $f(x)$ 在闭区间上不能保证恒为正值，那么它在区间上的定积分 (如果存在的话) 等于位于 x 轴上方的面积减去 x 轴下方的面积，即代数面积。

3.1.2 可积条件

现在我们研究函数在闭区间上可积的条件。

先提出达布和的概念:

定义 3.1.2. 设函数 $f(x)$ 在闭区间 $[a,b]$ 上有定义，对于该闭区间的某个划分:

$$a = x_0 < \cdots < x_{i-1} < x_i < \cdots < x_n = b$$

设 $f(x)$ 在每个小区间 $[x_{i-1}, x_i]$ 上的上确界为 M_i，下确界为 m_i，作和式:

$$s = \sum_{i=1}^{n} m_i \Delta x_i$$

称为下积分和，再作和式

$$S = \sum_{i=1}^{n} M_i \Delta x_i$$

称为上积分和，上积分和与下积分和都称为达布和.

由积分布和的概念，在每个小区间 $[x_{i-1}, x_i]$ 上任意取定点 ξ_i，有

$$m_i \leqslant f(\xi_i) \leqslant M_i$$

因此有

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

即是说，对于同一个划分而言，积分和介于上积分和与下积分和之间，而上下积分和分别是积分和的上确界和下确界。

达布和具有两个性质:

性质 3.1.1. *如果在分划中增加一些新的分点，那么上积分和不会增加，下积分和不会减小。*

证明 假如第 i 个区间 $[x_{i-1}, x_i]$ 中新插入一个分点 x',设函数 $f(x)$ 在区间 $[x_{i-1}, x']$ 上的上下确界分别是 M_{i1} 和 m_{i1}，而在区间 $[x', x_i]$ 上的上下确界分别是 M_{i2} 和 m_{i2}，那么必定有

$$m_i \leqslant m_{i1}, m_i \leqslant m_{i2}$$

以及

$$M_i \geqslant M_{i1}, M_i \geqslant M_{i2}$$

而下积分和原来对应区间 $[x_{i-1}, x_i]$ 上的那一项

$$m_i(x_i - x_{i-1})$$

将由下面的两项代替:

$$m_{i1}(x' - x_{i-1}) + m_{i2}(x_i - x')$$

显然它的值不会减小，只会增加或者保持不变. 同样上积分和中原来的那一项

$$M_i(x_i - x_{i-1})$$

将被下面两项代替:

$$M_{i1}(x' - x_{i-1}) + M_{i2}(x_i - x')$$

同样，它的值不会增加，只会减小或者保持不变。这是只添加一个分点的情况，而多个分点可以逐个添加，每添加一个分点，上积分和都保持不增加，下积分和不减小，所以结论得证。 □

性质 3.1.2. *上积分和永远不小于下积分和，无论是对于同一个划分还是不同的划分均是如此.*

证明 对于同一个划分来说，上积分和不小于下积分和是显然的，所以需要证明的是两个不同的划分 A 和 B，A 的上积分和不会小于 B 的下积分和。

我们可以将两个划分的分点合并在一起组成一个新的更细的划分 C，它是在 A 的分点中加入了一些新的分点 (B 的分点)，因此按刚证明过的上一个性质，知 C 划分的上积分和不会大于 A 的上积分，同样，C 也是在 B 的分点中加入了一些新的分点 (A 的分点)，所以 C 划分的下积分和不会小于 B 划分的下积分和，于是我们有

$$S_C \leqslant S_A, s_C \geqslant s_B$$

而又由于

$$s_C \leqslant S_C$$

因此

$$s_B \leqslant S_A$$

即得证. □

达布和与黎曼和相比，它不需要在每个小区间上取一个点，从而对它的讨论将与这个点的选择无关。回忆定积分的定义中，我们本应该直接使用极限语言，直接说当各小区间的最大长度趋于零时，黎曼和的极限便是定积分，但是无奈还有各个小区间上的点怎么选择的干扰，所以不得不使用了 $\varepsilon-\delta$ 语言来描述，并要求是无论怎么选择那些点，都不影响误差限的成立，而现在达布和没有这个问题，我们可以直接使用极限语言了。

利用达布和，我们有如下的可积条件:

定理 3.1.1. *函数 $f(x)$ 在闭区间上可积的充分必要条件是:*

$$\lim_{\lambda\to 0}(S-s)=0$$

这里 λ 表示划分下的各小区间的最大长度.

证明 先证必要性，假定已经有函数 $f(x)$ 在闭区间 $[a,b]$ 上可积，且定积分值为 P，那么就是说，对于任意小的正实数 $\varepsilon>0$，存在另一个充分小的正实数 $\delta>0$，使得只要划分的小区间最大长度小于 δ，则无论怎么选定各个小区间上的 ξ_i，都有

$$\left|\sum_{i=1}^{n} f(\xi_i)\Delta x_i\right| < \varepsilon$$

那么作为黎曼和上下确界的达布和，自然也有

$$|s-P|\leqslant\varepsilon, |S-P|\leqslant\varepsilon$$

于是

$$|S-s|=|(S-p)-(s-p)|\leqslant|S-p|+|s-p|\leqslant 2\varepsilon$$

这就是说，$|S-s|$ 也是可以任意小的，因此必要性成立.

再来证明充分性，如果已经有上下积分和之差在小区间最大长度趋于零时也趋于零，那么由于任意一个上积分和不小于任意一个下积分和，也就是上积分的集合有下界，而下积分的集合有上界，于是各有确界，并且上积分和的下确界必定等于下积分和的上确界 (否则两者之差不可能趋于零)，设此共同的确界是 I，那么我们来证明，这个 I 便是函数在这闭区间上的定积分。

对于任意小的正实数 $\varepsilon>0$，存在另一个充分小的正实数 $\delta>0$，使得只要划分的最大小区间长度小于 δ，就有 $|S-s|<\varepsilon$，但是又由于

$$s\leqslant I\leqslant S$$

因此对于这个 s 和 S 有

$$I-\varepsilon\leqslant s\leqslant I$$

以及

$$I \leqslant S \leqslant I + \varepsilon$$

而在此区间上的黎曼和 σ 满足 (无论怎么选择 ξ_i)

$$s \leqslant \sigma \leqslant S$$

所以

$$I - \varepsilon \leqslant \sigma \leqslant I + \varepsilon$$

这便表示函数 $f(x)$ 在闭区间 $[a, b]$ 上可积，并且积分值为 I. □

在某个划分中，对于每一个小区间 $[x_{i-1}, x_i]$ 而言，$M_i - m_i$ 可以视作函数 $f(x)$ 在该小区间上的振幅，记作 ω_i，于是上述定理中的等式也可以写为

$$\lim_{\lambda \to 0} \sum_{i=1}^{n} \omega_i \Delta x_i = 0$$

在接下来我们将会用到这种形式，为了方便引用定理编号，将其写成如下形式:

定理 3.1.2. *函数 $f(x)$ 在闭区间上可积的充分必要条件是:*

$$\lim_{\lambda \to 0} \omega_i \Delta x_i = 0$$

这里 λ 表示划分下的各小区间的最大长度.

例 3.1.4 对于狄利克雷函数

$$D(x) = \begin{cases} 1 & x \in Q \\ 0 & x \notin Q \end{cases}$$

它在任意闭区间 $[a, b]$ 上均不可积，因为无论将区间划分得多么小，它在每个小区间上的振幅差都是 1. ■

例 3.1.5 定义在区间 $[0, 1]$ 上的黎曼函数 $R(x)$ 如下，

$$R(x) = \begin{cases} 0 & x = 0 \\ \frac{1}{q} & x = \frac{p}{q} \in [0, 1], p, q \in \mathbb{Z}, q > 0, (p, q) = 1 \\ 0 & x \in [0, 1] - \mathbb{Q} \end{cases}$$

当 x 是有理数 $\dfrac{p}{q}$ 时，$R(x) = \dfrac{1}{q}$，注意这里 p、q 是一对既约整数（即最大公因数为 1），且 q 是正的。而当 x 是无理数时，$R(x) = 0$.

我们来证明，我们在 例 1.4.6中证明过，黎曼函数在 $[0, 1]$ 上的无理点处均连续，$(0, 1)$ 上的有理点都是它的间断点。现在我们证明，它在 $[0, 1]$ 上是可积的，这就是一个拥有大量稠密间断点的非连续函数但却可积的例子.

对于任意小的正实数 $\varepsilon > 0$，我们需要找到另一个充分小的正实数 $\delta > 0$，使得对于满足小区间长度全部都小于 δ 的任意一个分划，都有

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

针对 ε，我们取定一个充分大的正整数 N(具体需要大到何种程度待定)，注意到，满足函数值 $R(x) \geqslant \frac{1}{N}$ 的自变量是有限的: $\frac{1}{2}$、$\frac{1}{3}$、$\frac{2}{3}$、$\frac{1}{4}$、$\frac{3}{4}$、…, $\frac{1}{N}$、…、$\frac{N-1}{N}$，一共不超过

$$1 + 2 + \cdots + (N-1) = \frac{1}{2}(N-1)(N-2) < N^2$$

个，记这些自变量的集合为 X，即

$$X = \{x | R(x) \geqslant \frac{1}{N}\}$$

上式即是说，X 中的元素个数不会超过 N^2. 对于一个充分小的实数 δ，作闭区间 $[0,1]$ 的任意一个满足所有小区间长度都小于 δ 的划分，这些小区间可以分为两类，一类包含 X 中的点，一类不包含 X 中的点，由于 X 是有限集，甚至可以将 δ 取得足够小使得每个小区间上至多只包含一个 X 中的点，不过这并不是必要的. 对于包含 X 中的点的那些小区间，有 $\omega_i < 1$，并且小区间的个数不超过 X 中元素个数，自然也就不会超过 N^2, 从而

$$\sum_1 \omega_i \Delta x_i < \sum_1 \Delta x_i < N^2 \delta$$

而对于那些不包含 X 中元素的小区间，有 $\omega_i < \frac{1}{N}$，并且区间总长度不超过 1，从而

$$\sum_2 \omega_i \Delta x_i < \frac{1}{N} \sum_2 \Delta x_i < \frac{1}{N}$$

因而有

$$\sum_{i=1}^{n} \omega_i \Delta x_i = \sum_1 \omega_i \Delta x_i + \sum_2 \omega_i \Delta x_i < N^2 \delta + \frac{1}{N}$$

现在只要取

$$N > \frac{2}{\varepsilon}$$

及

$$\delta < \frac{\varepsilon}{2N^2}$$

就有

$$\sum_{i=1}^{n} \omega_i \Delta x_i < N^2 \delta + \frac{1}{N} < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

因而可积. ∎

推论 3.1.1. *闭区间上的无界函数不可积.*

因为对于闭区间上的无界函数，无论把小区间划分得有多细，总存在一个振幅无穷大的小区间，按前述定理，这个定积分不存在。

前面已经说明了，上积分和虽然是与划分有关的，但是它有下界，任何一个下积分和都是它的下界，那么它的下确界，显然是与划分无关的，同样，下积分和也是有上确界的，于是我们类似于积分和与定积分的关系，定义上积分与下积分的概念:

定义 3.1.3 (上积分与下积分). *上积分和的下确界称为上积分，记作 I^*，下积分和的上确界称为下积分，记作 I_*.*

显然，上积分与下积分只与被积函数和被积区间有关，与区间的划分无关，并且，只要函数在闭区间上是有界的，那么上积分与下积分就总是存在的。

虽然上积分与下积分是用确界定义的，但很自然的，它们也可以看作极限:

定理 3.1.3. *上积分是上积分和的极限，下积分是下积分和的极限，即:*

$$I^* = \lim_{\lambda\to 0} S, I_* = \lim_{\lambda\to 0} s$$

这里 λ 表划分中子区间的最大长度.

证明 只证明上积分的结论，只需要证明：对于任意的正实数 $\varepsilon > 0$，都存在一个非常小的正实数 $\delta > 0$，使得对于最大子区间长度小于 δ 的任意划分，都能保证其上积分和 S 与上积分 I 相差不超过 ε，即

$$S < I^* + \varepsilon$$

首先根据确界概念，对于任意的正实数 $\varepsilon > 0$，都存在一个划分 (记作 A)，能够使得其上积分和满足:

$$S_A < I^* + \frac{1}{2}\varepsilon$$

虽然我们现在还未确定 δ 是多少，但是我们考虑最大子区间长度小于 δ 的任意一个划分 X，如果将划分 A 的分点全部加进 X 的分点中去，得到一个新划分 Y，那么由上积分和的性质有 $S_Y < S_X$，因此有

$$S_Y < I^* + \frac{1}{2}\varepsilon$$

接着我们考虑: 在 X 划分中加入若干分点到底能引起上积分和产生多大的变化，也就是估计 S_X 与 S_Y 之差，因为这个划分 A 我们是针对 ε 取定的，我们假设它有 l 个分点，那么在划分 X 中插入 l 个分点，最多只能影响 X 中原来 l 个小区间上的积分和，且每个小区间上的积分和改变量不会超过 $(M-m)\delta$，于是有

$$|S_Y - S_X| < l(M-m)\delta$$

这里 M 和 m 分别是函数在整个要积分的闭区间上的上确界和下确界. 因此如果取 δ 使其满足

$$\delta < \frac{\varepsilon}{2l(M-m)}$$

那么便有

$$|S_X - S_Y| < \frac{1}{2}\varepsilon$$

从而

$$S_X < I^* + \varepsilon$$

□

有了这个定理，那么我们现在可以利用上积分与下积分，把前面的可积条件写成下面的形式.

定理 3.1.4. 函数在闭区间上可积的充分必要条件是: 上积分与下积分相等，即

$$I^* = I_*$$

并且在上下积分相等的情况下，这个共同值便是定积分的值.

这个定理的理论意义是如此之重要，因为前面已经说明首先闭区间上的无界函数是不可积的，而对于有界函数来说，它在闭区间上的上积分与下积分是一定存在的，所以定积分是否存在的问题，便归结于这个上积分与下积分是否相等的问题了。

3.1.3 可积函数类

虽然我们已经得出了函数在闭区间上可积的充分必要条件，但那个条件更多的是理论上的意义，对于一个给定的函数，要快速的判断可积性是比较困难的，所以这一节我们总结一些常用的可积函数和不可积函数。

定理 3.1.5. 闭区间上的连续函数总是可积的.

证明 由于闭区间上的连续函数必然是一致连续，即任给正实数 $\varepsilon > 0$，都存在另一个正实数 $\delta > 0$，使得对于该闭区间上任意两个自变量 x_1, x_2，只要 $|x_1 - x_2| < \delta$，就能保证 $|f(x_1) - f(x_2)| < \varepsilon$. 也就是说，只要划分的最大子小区间长度小于 δ，那么所有子区间上的振幅都会小于 ε，于是它对应的上下积分和之差

$$\sum_{i=1}^{n} \omega_i \Delta x_i < \varepsilon \sum_{i=1}^{n} \Delta x_i < \varepsilon(b-a)$$

这里 a 和 b 分别是闭区间的两个端点. 上式便表明了函数的可积性。 □

定理 3.1.6. 如果函数在闭区间上只有有限多个间断点，那么它是可积的.

为了方便证明这个，我们把 定理 3.1.2 改述为以下的形式:

定理 3.1.7. 函数 $f(x)$ 在闭区间上可积的充分必要条件是: 对于任意小的两个正实数 $\varepsilon > 0$ 和 $\sigma > 0$，都存在一个很小的正实数 $\delta > 0$，使得任何一个最大小区间长度小于 δ 的划分都满足：那些振幅大于 ε 的小区间长度总和不超过 σ.

证明 充分性，在定理所述条件下，振幅和被分为两部分：振幅不超过 ε 的小区间上的振幅和，与振幅超过 ε 的小区间上的振幅和，对于前者，这部分振幅和不超过 $\varepsilon(b-a)$，对于后者，这部分振幅和不超过 $\sigma(M-m)$，于是整个振幅和将不会超过

$$\varepsilon(b-a) + \sigma(M-m)$$

由 ε 和 σ 的任意性，即知该振幅和可以任意小，即可积。

必要性是显然的。 □

有了这个描述，现在来证明在闭区间上有有限个间断点的函数是可积的:

证明 对于某个划分而言，我们将这些小区间分成两类，一类是包含间断点的，一类是不包含间断点的，只要划分足够细，包含间断点的那些小区间的总长度可以任意小，而不包含间断点的那些小区间上的振幅可以任意小，因此，函数在区间上可积。 □

定理 3.1.8. *如果函数在闭区间上单调有界，则它是可积的.*

3.1.4 定积分的性质

性质 3.1.3. *如果函数 $f(x)$ 在 $[a,b]$ 上可积，则它在反向区间 $[b,a]$ 上也可积，且有*

$$\int_b^a f(x)dx = -\int_a^b f(x)dx$$

证明 这只要在划分区间时进行反向划分就可以得出，对于区间 $[a,b]$ 的任一分划 $a = x_0 < x_1 < \cdots < x_n = b$，它也能成为反向区间 $[b,a]$ 的一个分划，但在计算积分和时，它们的 Δx_i 分别为 $x_i - x_{i-1}$ 和 $x_{i-1} - x_i$，符号正好相反，因此对应于该分划的积分和也互为相反数，因此便有此结论. □

性质 3.1.4. *设函数 $f(x)$ 在 $[a,b]$ 上可积，r、s、t 是该区间上任意三个实数 (大小关系任意)，恒有*

$$\int_r^s f(x)dx = \int_r^t f(x)dx + \int_t^s f(x)dx$$

性质 3.1.5. *如果函数 $f(x)$ 在 $[a,b]$ 上可积，则 $kf(x)$ 也在该区间上可积，且*

$$\int_a^b f(x)dx = k\int_a^b f(x)dx$$

性质 3.1.6. *如果函数 $f(x)$、$g(x)$ 均在 $[a,b]$ 上可积，则 $f(x) \pm g(x)$ 在该区间上也可积，且*

$$\int_a^b [f(x) \pm g(x)]dx = \int_a^b f(x)dx \pm \int_a^b f(x)dx$$

性质 3.1.7. *如果函数 $f(x)$ 在 $[a,b]$ 上可积，且恒有 $f(x) \geqslant 0$，则*

$$\int_a^b f(x)dx \geqslant 0$$

更进一步，在上述条件下，只要恒有 $f(x) > 0$，就有

$$\int_a^b f(x)dx > 0$$

性质 3.1.8. *如果函数 $f(x)$、$g(x)$ 均在 $[a,b]$ 上可积，且恒有 $f(x) \geqslant g(x)$ ，则有*

$$\int_a^b f(x)dx \geqslant \int_a^b g(x)dx$$

如果前面不等式恒取大于号，则积分不等式也恒取大于号.

性质 3.1.9. *如果函数 $f(x)$ 在区间 $[a,b]$ 上可积，且在该区间上函数有上界 M 和下界 m，则有*

$$m(b-a) \leqslant \int_a^b f(x)dx \leqslant M$$

3.1.5 积分中值定理

假定函数 $f(x)$ 在闭区间 $[a,b]$ 上可积，依 推论 3.1.1，它必定是有界的，记它在该区间上的上下确界分别为 M 和 m，则对于它的任意一个分划上的积分和，都有

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i \geqslant \sum_{i=1}^{n} m\Delta x_i = m(b-a)$$

及

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i \leqslant \sum_{i=1}^{n} M\Delta x_i = M(b-a)$$

最终

$$m(b-a) \leqslant \sum_{i=1}^{n} f(\xi_i)\Delta x_i \leqslant M(b-a)$$

作为极限的积分，自然也有

$$m(b-a) \leqslant \int_a^b f(x)dx \leqslant M(b-a)$$

因此便得到

定理 3.1.9. *设函数 $f(x)$ 在闭区间 $[a,b]$ 上可积，同时它在该区间上的上下确界分别为 M 和 m，则存在实数 μ 符合 $m \leqslant \mu \leqslant M$，满足*

$$\int_a^b f(x)dx = \mu(b-a)$$

如果函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，则依介值定理，还有

定理 3.1.10. *若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，且上下确界分别为 M 和 m，则存在 $\xi \in [a,b]$，使得*

$$\int_a^b f(x)dx = f(\xi)(b-a)$$

3.1.6 变动上限的积分函数

假定函数 $f(x)$ 在闭区间 $[a,b]$ 上可积，则依定积分的性质，它在 $[a,b]$ 的任意子区间 $[a,x]$ 上也是可积的, 这里 $a \leqslant x \leqslant b$，据此我们可以作出一个函数

$$\Phi(x) = \int_a^x f(t)dt$$

函数 $\Phi(x)$ 以变动的积分上限作为自变量，并定义在区间 $[a,b]$ 上，为与自变量 x 区分，积分变量用 t 表示. 它的几何意义就是变动右边界的曲边梯形面积.

我们研究下它的连续性和可导性，由

$$\Phi(x+\Delta x) - \Phi(x) = \int_a^{x+\Delta x} f(t)dt - \int_a^x f(t)dt = \int_x^{x+\Delta x} f(t)dt = \mu\Delta x$$

其中最后一步是运用积分中值定理得出，且 μ 介于函数 $f(x)$ 在该闭区间上的上下确界之间，上式最右边在 $\Delta x \to 0$ 时显然是一个无穷小，因此 $\Phi(x)$ 在闭区间 $[a,b]$ 上是连续的. 由此得到

定理 3.1.11. *如果函数 $f(x)$ 在闭区间 $[a,b]$ 上可积，则变动上限的积分函数*

$$\Phi(x)=\int_a^x f(t)dt$$

在闭区间 $[a,b]$ 上处处连续.

再讨论可导性，如果再假定函数 $f(x)$ 在闭区间 $[a,b]$ 上还是连续的，就有

$$\begin{aligned}\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x}&=\frac{1}{\Delta x}\int_x^{x+\Delta x} f(t)dt\\&=f(\xi)\end{aligned}$$

其中 ξ 介于 x 与 $x+\Delta x$ 之间，当 $\Delta x\to 0$ 时，有 $\xi\to x$，再由连续性有 $f(\xi)\to f(x)$，因此有

$$\begin{aligned}\Phi'(x)&=\lim_{\Delta x\to 0}\frac{\Phi(x+\Delta x)-\Phi(x)}{\Delta x}\\&=\lim_{\Delta x\to 0} f(\xi)\\&=f(x)\end{aligned}$$

即函数 $\Phi(x)$ 可导，且导函数满足

$$\Phi'(x)=f(x)$$

于是有

定理 3.1.12. *如果函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，则变动上限的积分函数 $\Phi(x)=\int_a^x f(t)dt$ 在区间 $[a,b]$ 上处处可导，且导函数就是 $f(x)$，即*

$$\Phi'(x)=f(x)$$

这表明：变动右边界的曲边梯形的面积，以函数值自身为其变化率，反过来说，对于闭区间上的连续函数，变动上限的定积分也正好是它的一个原函数.

于是得到一个重要结论:

推论 3.1.2. *闭区间上的连续函数必定存在原函数，变动上限的定积分便是其一.*

由于初等函数在其有定义的各个区间上都是连续函数，因此它们在各个区间上就都存在原函数，不过，初等函数的原函数不一定还是初等函数，后面将会见到，还有可能是非常高深的函数.

3.1.7 牛顿-莱布尼茨公式

继续上一小节的讨论，我们知道，函数的原函数如果存在，则会是一个函数族，它们都相差一个常数，假定 $F(x)$ 是 $f(x)$ 的任一原函数，则必有

$$F(x)=\Phi(x)+C$$

这里 C 是一个常数，注意到 $\Phi(a)=0$，代入上式即得 $C=F(a)$，即

$$\Phi(x)=F(x)-F(a)$$

那么作为定积分的曲边梯形面积就有

$$\int_a^b f(t)dt=\Phi(b)=F(b)-F(a)$$

这就是说，闭区间上的连续函数的定积分值，等于它的任一原函数在这个区间两端点处的函数值之差! 这个结论既为我们指出了定积分与原函数之间的关系，也给出了一种计算定积分值的途径. 这个公式就叫做 *牛顿-莱布尼茨公式*，而鉴于原函数与下文即将介绍的不定积分的关系，这个公式直接沟通了定积分与不定积分的联系，因此其理论价值奠定了它在微积分学中的基石地位，因而被称为*微积分学基本定理*，鉴于它的重要性，我们专门列出

定理 3.1.13 (微积分学基本定理/牛顿-莱布尼茨公式). *设函数 $f(x)$ 在闭区间 $[a,b]$ 上连续，$F(x)$ 是它的任一原函数，则关于 $f(x)$ 在该区间上的定积分有*

$$\int_a^b f(x)dx=F(b)-F(a)$$

在此给出另外一种证明:

证明 对于闭区间 $[a,b]$ 的任一分划 $a=x_0<x_1<\cdots<x_n=b$, 利用原函数 $F(x)$ 的可导性以及拉格朗日中值定理有

$$\begin{aligned}F(b)-F(a)&=\sum_{i=1}^n(F(x_i)-F(x_{i-1}))\\&=\sum_{i=1}^n f(\xi_i)(x_i-x_{i-1})\end{aligned}$$

其中 $\xi_i\in(x_{i-1},x_i)$ 为拉格朗日中值定理中存在的中值，非任意取定.

上式对于任一分划都存在相对应的一组 $\xi_i(i=1,2,\ldots,n)$，而由于 $f(x)$ 可积，因而对于任意小的正实数 $\varepsilon>0$，均存在另一个很小的正实数 $\delta>0$，使得对于满足所有小区间长度都小于 δ 的任意一个划分，在各个小区间上任意选定一组 $\eta_i(i=1,2,\ldots,n)$，都有

$$\left|\sum_{i=1}^n f(\eta_i)\Delta x_i-\int_a^b f(x)dx\right|<\varepsilon$$

既然任意选定 $\eta_i(i=1,2,\ldots,n)$ 都成立这不等式，那么对于同一个划分及其上确定的中值组 $\xi_i(i=1,2,\ldots,n)$，自然也有

$$\left|\sum_{i=1}^n f(\xi_i)\Delta x_i-\int_a^b f(x)dx\right|<\varepsilon$$

也就是说

$$\left|[F(b)-F(a)]-\int_a^b f(x)dx\right|<\varepsilon$$

因而有

$$F(b) - F(a) = \int_a^b f(x)dx$$

$\square$

3.1.8 积分第二中值定理

积分第二中值定理是关于两个函数乘积的积分的:

定理 3.1.14 (积分第二中值定理)**.** 设函数 $f(x)$ 和 $g(x)$ 均在 $[a,b]$ 上有定义，且 $f(x)$ 单调递减并且函数值非负，同时 $g(x)$ 在该区间上可积，则存在 $\xi \in [a,b]$，使得

$$\int_a^b f(x)g(x)dx = f(a)\int_a^\xi g(x)dx$$

类似的，如果 $f(x)$ 是单调递增而其它条件不变 (仍要求函数值非负)，则存在 $\xi \in [a,b]$，使得

$$\int_a^b f(x)g(x)dx = f(b)\int_\xi^b g(x)dx$$

推论 3.1.3. 设函数 $f(x)$ 和 $g(x)$ 均在 $[a,b]$ 上有定义，且 $f(x)$ 单调 (不要求函数值非负)，且 $g(x)$ 在该区间上可积，则存在 $\xi \in [a,b]$，使得

$$\int_a^b f(x)g(x)dx = f(a)\int_a^\xi g(x)dx + f(b)\int_\xi^b g(x)dx$$

3.1.9 不定积分概念与性质, 基本积分表

不定积分概念

3.2 定积分与不定积分的计算

3.2.1 不定积分的换元积分法

3.2.2 不定积分的分部积分法

3.2.3 有理式的不定积分

3.2.4 根式函数的不定积分

3.2.5 含指对函数的不定积分

3.2.6 椭圆积分

3.2.7 直接使用积分和计算定积分

3.2.8 定积分的换元积分法和分部积分法

3.2.9 积分的近似计算

3.3 积分学在几何与物理中的应用

3.4 反常积分

3.5 含参积分

第四章　常数项无穷级数

4.1　正项级数的收敛性判别

对于级数 $\sum_{n=1}^{\infty} a_n$，如果它的每一项都是正数（或者当 n 充分大时恒保持正号），我们有一系列的判别方法可以判断它的收敛性。

由于正项级数的部分和是单调增加的，所以回想起数列极限的单调有界定理，我们就有

定理 4.1.1. *正项级数收敛的充分必要条件是它的部分和有界。*

例 4.1.1　级数 $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ 的部分和 $S_n = 1 - \frac{1}{n+1} < 1$，所以级数收敛。 ■

例 4.1.2　级数 $\sum_{n=1}^{\infty} \frac{1}{n!}$ 的部分和 $S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}$，我们在小节 1.2.7便已经得到过 $S_n < 3$，所以级数收敛。 ■

例 4.1.3　级数 $\sum_{n=1}^{\infty} \ln\left(1 + \frac{1}{n}\right)$，注意到 $\ln\left(1 + \frac{1}{n}\right) = \ln(n+1) - \ln n$，所以部分和 $S_n = \ln(n+1)$，显然无界，所以级数发散。 ■

4.1.1　比较判别法

如下的比较判别法是相当重要的:

定理 4.1.2 (比较判别法)**.** *对于两个正项级数 $\sum_{n=1}^{\infty} a_n$ 和 $\sum_{n=1}^{\infty} b_n$，如果从某一项起恒有 $a_n \leqslant b_n$，那么由 $\sum_{n=1}^{\infty} b_n$ 的收敛便能推得 $\sum_{n=1}^{\infty} a_n$ 也收敛，同理，由 $\sum_{n=1}^{\infty} a_n$ 发散便能推得 $\sum_{n=1}^{\infty} b_n$ 也发散。*

证明　由条件，存在下标 N，使得当 $n > N$ 时恒有 $a_n \leqslant b_n$，分别用 A_n 和 B_n 表示两个级数的部分和，则

$$A_n = A_N + (a_{N+1} + \cdots + a_n),\ B_n = B_N + (b_{N+1} + \cdots + b_n)$$

显然 $A_n - A_N \leqslant B_n - B_N$，所以如果 $\sum_{n=1}^{\infty} b_n$ 收敛，则 B_n 有上界，从而 A_n 也有上界，所以 $\sum_{n=1}^{\infty} a_n$ 也收敛。而如果 $\sum_{n=1}^{\infty}$ 是发散的，那么 A_n 必定没有上界，从而 B_n 也不可能有上界，因而 $\sum_{n=1}^{\infty} b_n$ 也必然发散。 □

实际上，条件 $a_n \leqslant b_n$ 可以改成 $a_n \leqslant \lambda b_n$，其中 λ 是一个正常数，这是因为正项级数 $\sum_{n=1}^{\infty} b_n$ 跟正项级数 $\sum_{n=1}^{\infty} \lambda b_n$ 的收敛性是相同的。

这个判别法还有以下的极限形式

定理 4.1.3. 如果正项级数 $\sum_{n=1}^{\infty} a_n$ 和正项级数 $\sum_{n=1}^{\infty} b_n$ 的通项之比有极限 *(有限的或无穷的均可)*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = K$$

则如果 K 是正常数，那么两个级数同时收敛同时发散。如果 $K=0$，则由 $\sum_{n=1}^{\infty} b_n$ 收敛可推得 $\sum_{n=1}^{\infty} a_n$ 也收敛，如果 $K=+\infty$，由由 $\sum_{n=1}^{\infty} b_n$ 发散可推得 $\sum_{n=1}^{\infty} a_n$ 也发散。

例 4.1.4 设 $a_n=\frac{1}{n}$，$b_n=\ln\left(1+\frac{1}{n}\right)$，由极限

$$\lim_{n\to\infty} \frac{\frac{1}{n}}{\ln\left(1+\frac{1}{n}\right)} = 1$$

知道它们同时收敛同时发散，而在例 4.1.3中已经知道 $\sum_{n=1}^{\infty} b_n$ 是发散的，所以 $\sum_{n=1}^{\infty} a_n$ 也是发散的。 ■

例 4.1.5 我们来研究一个重要的级数，即级数 $\sum_{n=1}^{\infty} \frac{1}{n^s}$ 的收敛性。

如果 $s=1$，则级数成为

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1+\frac{1}{2}+\frac{1}{3}+\cdots$$

有如下的片段和估计

$$\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{2n} > n\cdots\frac{1}{2n} = \frac{1}{2}$$

片断和不能任意小，违反柯西收敛准则，所以 $s=1$ 时级数发散，于是 $s<1$ 时级数也都发散，因为部分和会更大。

而对于 $s>1$ 的情况，我们将证明它是收敛的，这是因为我们有

$$\frac{1}{(n+1)^s}+\frac{1}{(n+2)^s}+\cdots+\frac{1}{(2n)^s} < n\cdot\frac{1}{n^s} = \frac{1}{n^{s-1}}$$

所以我们把正整数分段，每一段以 2^k+1 作为开始，以 2^{k+1} 作为结尾，就有

$$\sum_{i=1}^{2^n} \frac{1}{i^s} = 1+\sum_{k=0}^{n}\sum_{i=1}^{2^k} \frac{1}{(2^k+i)^s} < 1+\sum_{k=0}^{n} \frac{1}{(2^k)^{s-1}} = 1+\sum_{k=0}^{n}\left(\frac{1}{2^{s-1}}\right)^k$$

注意到 $s>1$，所以上式最右边是一个公比小于 1 的等比级数，显然它是收敛的，于是左边的部分和有上界，从而级数收敛。

这个级数的和作为 s 的函数，便是著名的 *黎曼函数* $\zeta(s)$，即对于 $s>1$，

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

这函数在数论中有非常重要的地位。 ■

我们还有另一种形式的比较判别法

定理 4.1.4. 对于两个正项级数 $\sum_{n=1}^{\infty} a_n$ 和 $\sum_{n=1}^{\infty} b_n$，如果从某一项起恒有

$$\frac{a_{n+1}}{a_n} \leqslant \frac{b_{n+1}}{b_n}$$

那么由 $\sum_{n=1}^{\infty} b_n$ 收敛可推出 $\sum_{n=1}^{\infty} a_n$ 也收敛，同样，由 $\sum_{n=1}^{\infty} a_n$ 发散也可以推出 $\sum_{n=1}^{\infty} b_n$ 发散。

证明 设从 $n > N$ 时就有条件中的不等式恒成立，则可得

$$\frac{a_n}{a_N} \leqslant \frac{b_n}{b_N}$$

由定理 4.1.2即得结论。 □

利用比较判别法，我们把给定级数与一些已知为收敛的级数相比较，可以开发出一系列更具体的判别法，下文的判别法，基本都是如此。

4.1.2 柯西判别法与达朗贝尔判别法

因为几何级数 $\sum_{n=1}^{\infty} q^n$ 在 $0 < q < 1$ 时收敛，我们以它为比较标准，就可以得到柯西判别法和达朗贝尔判别法。

定理 4.1.5 (柯西判别法)**.** 对于正项级数 $\sum_{n=1}^{\infty} a_n$，作柯西序列

$$\mathcal{C}_n = \sqrt[n]{a_n}$$

如果存在正实数 $0 < q < 1$，使得当 n 充分大时恒有 $\mathcal{C}_n \leqslant q$，那么级数 $\sum_{n=1}^{\infty} a_n$ 收敛，如果当 n 充分大时恒有 $\mathcal{C}_n \geqslant 1$，那么级数发散。

证明 设当 $n > N$ 时恒有 $\mathcal{C}_n \leqslant q$，那么此时有 $a_n \leqslant q^n$，而级数 $\sum_{n=1}^{\infty} q^n$ 在 $0 < q < 1$ 时是收敛的，由比较判别法可得 $\sum_{n=1}^{\infty} a_n$ 也收敛。而如果当 $n > N$ 时恒有 $\mathcal{C}_n \geqslant 1$，那么此时恒有 $a_n \geqslant 1$，通项不能趋于零，故级数发散。 □

柯西判别法也有极限形式

推论 4.1.1. 对于正项级数 $\sum_{n=1}^{\infty} a_n$，如果柯西序列有极限 $\lim_{n\to\infty} \mathcal{C}_n = q$，那么当 $0 < q < 1$ 时级数收敛，当 $q > 1$ 时级数发散，当 $q = 1$ 时可能收敛也可能发散。

定理 4.1.6 (达朗贝尔判别法)**.** 对于正项级数 $\sum_{n=1}^{\infty} a_n$，作达朗贝尔序列

$$\mathcal{D}_n = \frac{a_{n+1}}{a_n}$$

如果当 n 充分大时有 $\mathcal{D}_n \leqslant q$，其中 $0 < q < 1$ 为常数，那么级数收敛，如果当 n 充分大时恒有 $\mathcal{D}_n \geqslant 1$，那么级数发散。

证明 设当 $n > N$ 时，$\mathcal{D}_n \leqslant q$，其中 $0 < q < 1$，那么自然就有 $a_n \leqslant a_N q^{n-N}$，由比较判别法即知原级数收敛。而如果当 $n > N$ 时 $\mathcal{D}_n \geqslant 1$，那么自然有 $a_n \geqslant a_N$，通项不趋于零，级数发散。 □

达朗贝尔判别法的极限形式

推论 4.1.2. 对于正项级数 $\sum_{n=1}^{\infty} a_n$，如果达朗贝尔序列有极限 $\lim_{n\to\infty} \mathcal{D}_n = q$，在 $0 < q < q$ 时级数收敛，在 $q > 1$ 时级数发散，$q = 1$ 时级数可能收敛也可能发散。

由达朗贝尔的条件出发，可以得它也满足柯西判别法的条件，因为由 $a_n \leqslant a_N q^{n-N}$ 可得 $\sqrt[n]{a_n} \leqslant q\sqrt[n]{a_N/q^N}$，后一根式极限为 1，所以当 n 充分大时它可以保证 $q\sqrt[n]{a_N/q^N} < q' < 1$，这里 q' 是比 q 稍大些但仍然小于 1 的常数，这样就得出了柯西判别法的条件，所以能用达朗判别法判断为收敛的级数，也能用柯西判别法来判别，但不一定能有达朗贝尔判别法来得方便。

4.1.3 拉阿伯判别法

我们已知级数 $\sum_{n=1}^{\infty} \frac{1}{n^s}$ 在 $s > 1$ 时收敛，在 $s \leqslant 1$ 时发散，把给定的级数与这个级数相比较，就得出拉阿伯判别法。

定理 4.1.7 (拉阿伯判别法)**.** 对于正项级数 $\sum_{n=1}^{\infty} a_n$，记拉阿伯序列

$$\mathcal{R}_n = n\left(\frac{a_n}{a_{n+1}} - 1\right)$$

如果存在常数 $r > 1$，使得当 n 充分大时恒有 $\mathcal{R}_n \geqslant r$，则级数收敛，如果当 n 充分大时恒有 $\mathcal{R}_n \leqslant 1$，则级数发散。

证明 假定当 $n > N$ 时恒有 $\mathcal{R}_n \geqslant r$，则可得

$$\frac{a_{n+1}}{a_n} \leqslant \frac{n}{n+r}$$

对于右边的分式，我们将说明存在 $s > 1$，使得当 n 充分大时恒有

$$\frac{n}{n+r} \leqslant \left(\frac{n}{n+1}\right)^s$$

这 s 只要满足下式

$$s \leqslant \frac{\ln\left(1+\frac{r}{n}\right)}{\ln\left(1+\frac{1}{n}\right)}$$

借由当 $x \to 0$ 时的等价无穷小 $\ln(1+x) \sim x$，我们可得出上式右边以 r 为极限，而 $r > 1$，所以取 $1 < s < r$，则上式右边充分大时便能恒大于 s，于是

$$\frac{a_{n+1}}{a_n} \leqslant \left(\frac{n}{n+1}\right)^s$$

便能在 n 充分大时恒成立，由级数 $\sum_{n=1}^{\infty} \frac{1}{n^s}$ 在 $s > 1$ 时收敛知原级数收敛。

如果当 n 充分大时 $\mathcal{R}_n \leqslant 1$，则此时

$$\frac{a_{n+1}}{a_n} \geqslant \frac{n}{n+1}$$

由级数 $\sum_{n=1}^{\infty} \frac{1}{n}$ 的发散知原级数发散。 □

拉阿伯判别法的极限形式是

推论 4.1.3. 对于正项级数 $\sum_{n=1}^{\infty} a_n$，如果拉阿伯序列存在极限 $\lim_{n\to\infty} \mathcal{R}_n = R$，则 $R > 1$ 时级数收敛，$R < 1$ 时级数发散，$R = 1$ 时级数可能收敛也可能发散。

4.1.4 库默尔判别法

库默尔判别法是一个泛型化的判别法，利用它可以构造一系列具体的判别法，包括达朗贝尔判别法和拉阿伯判别法。

定理 4.1.8 (库默尔 (Kummer) 判别法)**.** 对于正项级数 $\sum_{n=1}^{\infty} a_n$，我们选取一个序列 c_n，使得级数 $\sum_{n=1}^{\infty} \frac{1}{c_n}$ 是发散的，作库默尔序列

$$\mathcal{K}_n = c_n \cdot \frac{a_n}{a_{n+1}} - c_{n+1}$$

如果存在正常数 r，使得当 n 充分大时恒有 $\mathcal{K}_n \geqslant r$，那么级数 $\sum_{n=1}^{\infty} a_n$ 收敛，如果当 n 充分大时恒有 $\mathcal{K}_n \leqslant 0$，则级数 $\sum_{n=1}^{\infty}$ 发散。

证明 如果当 $n \geqslant N$ 时有 $\mathcal{K}_n \geqslant r > 0$，那么此时有 $c_n a_n - c_{n+1}a_{n+1} \geqslant r a_{n+1} > 0$，这表明数列 $c_n a_n$ 是一个单调递减的正项数列，因而它存在极限，进而有

$$c_N a_N - c_{n+1}a_{n+1} = \sum_{i=N}^{n}(c_i a_i - c_{i+1}a_{i+1}) \geqslant r\sum_{i=N}^{n} a_{i+1}$$

由于 $c_n a_n$ 存在极限，因而正项 (至少从 N 开始为正) 级数 $\sum_{n=1}^{\infty}(c_n a_n - c_{n+1}a_{n+1})$ 收敛，于是由比较判别法，级数 $\sum_{n=1}^{\infty} a_n$ 收敛，而且从这证明过程中可以看到，条件 $\sum_{n=1}^{\infty} \frac{1}{c_n}$ 发散并没有用到，但这条件在证明定理的后一半结论时会用到。

如果当 n 充分大时恒有 $\mathcal{K}_n \leqslant 0$，便有

$$\frac{a_{n+1}}{a_n} \leqslant \frac{\frac{1}{c_n}}{\frac{1}{c_{n+1}}}$$

由级数 $\sum_{n=1}^{\infty} \frac{1}{c_n}$ 的发散，知级数 $\sum_{n=1}^{\infty} a_n$ 发散。 □

在库默尔判别法中，取 $c_n = 1$，就得出达朗贝尔判别法，取 $c_n = n$，就得出拉阿伯判别法。

库默尔判别法的极限形式是

定理 4.1.9. 对于正项级数 $\sum_{n=1}^{\infty} a_n$ 及选定的序列 c_n,如果库默尔序列有极限 $\lim_{n\to\infty} \mathcal{K}_n = K$，则如果 $K > 0$，则级数收敛，如果 $K < 0$，则级数发散，而对于 $K = 0$，则级数可能收敛也可能发散。

4.1.5 积分判别法

可以利用反常积分来判别一类级数的敛散性.

定理 4.1.10 (积分判别法)**.** 若可积函数 $f(x)$ 在 $[0, +\infty)$ 上恒取非负值且单调递减，则反常积分 $\int_0^{+\infty} f(x)dx$ 与级数 $\sum_{n=1}^{\infty} f(n)$ 有相同的敛散性。

证明 令 $S_n = \sum_{i=1}^{n} f(i)$，$I_n = \int_0^n f(x)dx$，则 S_n 和 I_n 都是单调增加的数列，又

$$I_n = \sum_{i=1}^{n}\int_{i-1}^{i} f(x)dx$$

由单调性，有

$$\sum_{i=1}^{n} f(i) \leqslant I_n \leqslant \sum_{i=1}^{n} f(i-1)$$

即

$$S_n \leqslant I_n \leqslant f(0) + S_{n-1}$$

可见，如果级数 $\sum_{n=1}^{\infty} f(n)$ 收敛，则 S_n 有极限 S，则 I_n 单调增加并有上界 $f(0)+S$，因而 I_n 有极限 (不一定是 S)，即反常积分 $\int_{i=1}^{+\infty} f(x)dx$ 收敛。反之，如果反常积分收敛，则 I_n 有极限 I，这时 S_n 单调增加并有上界 I，因此 S_n 收敛，即级数 $\sum_{i=1}^{\infty} f(n)$ 收敛。 □

需要说明的是，上述定理中的区间可以是任意的左闭右开区间 $[a,+\infty)$，这时级数只要从大于 a 的任一正整数开始即可，从定理证明过程可以看出这并没有什么影响。

例 4.1.6　对于正实数 p, 考虑如下的级数

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

对应的函数 $f(x)=1/x^p$ 在 $[1/2,+\infty)$ 上单调递减，由于 $p>1$ 时反常积分 $\int_{1}^{+\infty} f(x)dx$ 收敛，所以此时级数也收敛，在 $p \leqslant 1$ 时反常积分是发散的，所以此时级数也是发散的。

■

4.2　一般项级数

4.2.1　交错级数

定义 4.2.1. 如果序列 $\{a_n\}$ 的任意相邻两项的符号都相反，即整个序列交错的取正值和负值，则称其为交错序列，而对应级数 $\sum_{i=1}^{\infty} a_n$ 为交错级数.

对于交错级数的收敛性有如下结论

定理 4.2.1. 如果数列 $\{a_n\}(n=0,1,\ldots)$ 单调递减并趋于零，则级数 $\sum_{i=0}^{\infty}(-1)^n a_n$ 收敛。

证明　级数 $\sum_{i=0}^{\infty}(-1)^{n-1}a_n$ 与级数 $\sum_{i=0}^{n}(-1)^n a_n$ 的收敛性是相同的，这里只是为了方便而让首项 a_0 的符号是正的。

作级数的部分和 $S_n = a_0 - a_1 + a_2 - \cdots + (-1)^n a_n$, 显然

$$S_{2n+1} = (a_0 - a_1) + (a_2 - a_3) + \cdots + (a_{2n} - a_{2n+1})$$

由 $\{a_n\}$ 单调递减可知 S_{2n+1} 是单调增加的正项数列，但是

$$S_{2n+1} = a_0 - (a_1 - a_2) - (a_3 - a_4) - \cdots - (a_{2n-1} - a_{2n}) - a_{2n+1}$$

显然就有 $S_{2n+1} < a_0$，即 S_{2n+1} 又有上界，所以 S_{2n+1} 收敛，同样的方法还可得出 S_{2n} 是单调递减有下界，因而也收敛，而由 $S_{2n+1} = S_{2n} + a_{2n+1}$ 及 $a_n \to 0(n \to \infty)$ 可知这两个子列只能收敛到相同的极限，即级数 $\sum_{i=0}^{\infty}(-1)^n a_n$ 收敛。 □

4.2.2 绝对收敛级数及其性质

定义 4.2.2. 对于级数 $\sum_{n=1}^{\infty} a_n$，如果它的各项的绝对值相加而得的新级数 $\sum_{n=1}^{\infty} |a_n|$ 收敛，则称原来的级数是 绝对收敛 的.

定理 4.2.2. 如果一个级数绝对收敛，则它一定收敛。

证明 由柯西准则，如果级数 $\sum_{n=1}^{\infty} a_n$ 绝对收敛，则对于任意 $\varepsilon > 0$，存在正整数 N，使得对于任意 $n > N$ 和任意正整数 m 成立

$$|a_{n+1}| + |a_{n+2}| + \cdots + |a_{n+m}| < \varepsilon$$

于是也有

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+m}| < \varepsilon$$

所以原来的级数也收敛。 □

注意收敛级数并一定都是绝对收敛，如交错级数 $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$。

下面的定理提示了绝对收敛级数的一个非常重要的性质

定理 4.2.3. 如果一个级数绝对收敛，则将它的项任意重新排列后所得新级数仍然绝对收敛，且其和不变。

证明 对于绝对收敛的级数 $\sum_{n=1}^{\infty} a_n$，将它的项任意重新排列后所得新级数记为 $\sum_{n=1}^{\infty} b_n$，记 $A_n = \sum_{i=1}^{n} |a_i|$，$B_n = \sum_{i=1}^{n} |b_i|$，则对于任一 B_n，组成它的各个 b_i 在原来的级数中的下标最大值记为 m，则显然 $B_n \leqslant A_m$，于是 B_n 单调增加有上界，故级数 $\sum_{n=1}^{\infty} b_n$ 绝对收敛。 □

以下定理深刻提示了绝对收敛级数与非绝对收敛级数之间的区别。

定理 4.2.4 (黎曼定理). 如果一个级数收敛但非绝对收敛，则将它的项进行适当的重新排列后，可使新级数收敛到任意预先指定的实数。

参考文献

[1] [前苏联] 菲赫金哥尔茨. 微积分学教程. 高等教育出版社, 2006.

[2] 华罗庚. 高等数学引论. 高等教育出版社,2009.

[3] 华东师范大学数学系, 数学分析. 高等教育出版社, 2004.

[4] [美]F. 约翰,R. 柯朗. 微积分与数学分析引论. 科学出版社, 2001.

[5] [前苏联] 吉米诺维奇. 数学分析习题集. 高等教育出版社, 2004.

[6] zhcosin, 初等数学笔记, 网络电子书.