我的数学分析习题课讲义（上）

陈宇轩

chen_yuxuan@pku.edu.cn

2026年1月23日，版本3.5

目录

序

我写给你的如此贫乏。而我不能写的

像老式飞艇不断膨胀

最终穿过夜空消失。

——托马斯·特朗斯特罗姆《致防线后面的朋友》

这是我第二次担任北京大学数学分析I课程助教。上一次是在两年前，疫情三年中间的一段平淡时光。那时我刚大四，才经历了一个学期的跌宕：四月保研面试失败，五月丘赛只拿到铜奖，暑假全用来念书，九月终于赶上末班车保研北大。爱伦堡大概说过这样的话："生活并非按照努力的方向前进，而是往往像在抽彩票"。想要写下一些东西留下记录，于是有了这份讲义。自然要问为何不是写在两年前，理由是简单的：当时我在写另一份本科水平的分析学小册子，在[redacted]的建议下取名《分毫析厘》。[redacted]，[redacted]。

回到这份讲义上来。首先需要注明的是，标题里"我的"并非暗示原创性，而是说其结构和内容的编排如实记录了我在这两学期的教学工作、并反映了我对数学分析I应当教什么的思考和实践。内容方面，除了大量题目之外，还涉及一些相对较深的补充内容。尽管这些补充内容对考试难有益处，但都是未来学习数学不可避免的话题，因此尽早接触并无坏处。一次习题课只有两个小时，不可能覆盖讲义的全部，因此必然根据实际情况进行取舍。

归根到底，助教工作对学生学习水平的影响是有限的，更大程度上在于学生本人是否愿意花费精力去读书。从某些角度看，这份讲义是有欠缺的，比如基础难度的题目较少。但如果通过习题课能让学生看到更多有意思、有意义的数学问题，从而激发其思考和探索，那么一定是比纯粹讲解习题更有意义的。从这个角度上讲，我对这份讲义是满意的。

陈宇轩

完稿于2024年1月10日

过去一年里，对若干笔误进行了修订。只能侥幸期望没有对上一个版本的读者造成认知偏差。此外，增加了几道新的题目，据我所知尚未在其他关于数学分析的资料中出现。

v2.0更新于2025年2月7日

本学期起北大改用新的"101计划"系列教材，因此习题课内容也适当调整。另给每讲开头加了一段"彩蛋"性质的引用，无疑是受Ravi Vakil和陶哲轩启发。

v3.1更新于2025年10月15日

致谢

感谢范辉军教授、刘培东教授和童嘉骏教授，他们指导了我的教学工作，并且提供了许多题目。感谢教我数学分析的李伟固教授以及助教吴志昂，部分难题是从他们的课程上了解到的。2023–24学年秋季学期习题课的许多学生帮我指出了初稿中存在的许多笔误，一些题目是在与他们的交流中得知或者产生的，未能记录下他们的名字属实愧疚。许多衍生问题是经过与浮生群众多群友的讨论方才解决，群友对趣味数学的热情以及奇思妙想的头脑令我倾佩。最后，照例应该感谢家人的支持和女友的陪伴，他们使得我在数学之外的生活成为可能。

记号约定

$x := y$或$y =: x$表示将x定义为y。

$|A|$或$\#A$表示有限集A的元素个数。$A \hookrightarrow B$表示集合A到集合B的单射，$A \twoheadrightarrow B$表示A到B的满射。$\bigsqcup_{\alpha\in I} A_\alpha$表示一族集合$A_\alpha\ (\alpha \in I)$的无交并，即这些$A_\alpha$两两不交。在讨论集合的运算（尤其是取交集）时，总是默认在某个大集合X中进行。

$\mathbb{N} = \{1, 2, \dots\}$表示正整数集，$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$表示非负整数集。$\lfloor x\rfloor$表示不超过$x$的最大整数，$\lceil x\rceil$表示不小于$x$的最小整数，$\{x\} := x - \lfloor x\rfloor$表示$x$的小数部分。对于$a, b \in \mathbb{Z}$和$n \in \mathbb{N}$，同余式$a \equiv b \pmod n$表示$n \mid a - b$，即$a - b$是$n$的整数倍。$\mathbb{R}[x]$、$\mathbb{Z}[x]$分别表示实系数、整系数多项式环。规定$\infty$表示正无穷，而非不带符号的无穷大。

$\log x$表示以自然常数e为底的对数函数，而非像高中课本中以10为底。

$f(x) = o(g(x))$或$f(x) \ll g(x)$表示$\frac{f(x)}{g(x)} \to 0$；$f(x) = O(g(x))$表示$|f(x)| \le C|g(x)|$，其中常数C不依赖x。$f(x_0+) := \lim_{x\to x_0+} f(x)$表示$f$在$x_0$处的右极限。按照这一约定，$f'(x_0+)$表示$f$的导数在0处的右极限，而非$f$在$x_0$处的右导数。右导数的记号是$f'_+(x_0)$。

$C^\infty(I)$表示区间I上的无穷次可导函数（端点处考虑单侧导数），$C^n(I)$表示I上具有直到n阶连续导数的函数，$C(I)$或$C^0(I)$表示I上的连续函数。规定光滑函数指C^∞函数，而非C^1函数。

带有*号的题目意味着难度较大或/且与课程内容联系较弱。部分带有*号的定理没有提供证明，读者容易从常见资料中找到，亦不妨忽略其证明。

第1讲：集合与函数

清醒的人会从简单的等式入手，不会上来就去接触巴拿赫(Banach)空间拓扑学。

——乔治·斯坦纳《巴别塔之后》

这一讲对应课程所需的一些常识，大致包括简单的形式逻辑、朴素集合论[1]（尤其是可数集）、实数的定义、函数的基本性质（奇偶性、周期性）等。由于没有太多非平凡的题目可供讲解，花费了大量篇幅谈论集合的势，由此引进偏序集、选择公理与Zorn引理等。新教材涉及了多项式和代数数的内容，却缺乏对多项式基本理论的介绍，这里一并给出。

例题

例 1.1. 设非空集合$D \subset \mathbb{R}$，$f, g\colon D \to \mathbb{R}$是两个非负有界函数，证明：

$$(\inf_D f)(\inf_D g) \le \inf_D (fg) \le (\inf_D f)(\sup_D g) \le \sup_D (fg) \le (\sup_D f)(\sup_D g).$$

证明. 对任意$x \in D$，有$f(x) \ge \inf_D f \ge 0$和$g(x) \ge \inf_D g \ge 0$，因此

$$f(x)g(x) \ge (\inf_D f)(\inf_D g).$$

由$x \in D$的任意性得到第一个不等号。第二个不等号是由于

$$\inf_D (fg) \le f(x)g(x) \le f(x)(\sup_D g),$$

然后对$x \in D$取下确界。后两个不等号与前两个类似，留作练习。 □

例 1.2. 设$f\colon (0, \infty) \to \mathbb{R}$满足$\frac{f(x)}{x}$关于$x$单调递减，证明：对任意$x_1, x_2 > 0$，有

$$f(x_1 + x_2) \le f(x_1) + f(x_2).$$

证明. 因为$x_1 + x_2 \ge x_1$，所以$\frac{f(x_1+x_2)}{x_1+x_2} \le \frac{f(x_1)}{x_1}$，即

$$\frac{x_1}{x_1 + x_2} f(x_1 + x_2) \le f(x_1).$$

交换x_1, x_2，又有

$$\frac{x_2}{x_1 + x_2} f(x_1 + x_2) \le f(x_2).$$

上面两个式子相加就得到欲证的结论。 □

[1]早期版本错误使用了专业术语“描述集合论”，感谢浮生群群友迷亭、小乌龟指出。

注. 一般的，如果$b,d>0$且$\frac{a}{b}\le\frac{c}{d}$，则有

$$\frac{a}{b}\le\frac{a+c}{b+d}\le\frac{c}{d}.$$

这一简单观察将在*Stolz*定理的证明中发挥作用。除了直接通分证明之外，另一种看法是

$$\frac{a+c}{b+d}=\frac{b}{b+d}\cdot\frac{a}{b}+\frac{d}{b+d}\cdot\frac{c}{d},$$

即将$\frac{a+c}{b+d}$表示为$\frac{a}{b}$和$\frac{c}{d}$的加权平均。后一种观点可以称为“糖水不等式”：将两杯一定浓度的糖水混合之后，浓度一定介于原来两杯的浓度之间。

例 1.3. 设$f\colon\mathbb{R}\to\mathbb{R}$是最小正周期为$T>0$的周期函数，并且满足$f(x)\ne f(0)\,(\forall x\in(0,T))$，证明：$x\mapsto f(x^2)$不是周期函数。

证明. 用反证法，假设$S>0$是$f(x^2)$的周期。存在$n\in\mathbb{N}_0$使得$nT\le S^2<(n+1)T$，于是

$$f(0)=f(S^2)=f(S^2-nT).$$

因为$S^2-nT\in[0,T)$，由条件知只能$S^2=nT$，从而$n>0$。对任意$m\in\mathbb{N}$，有

$$f(0)=f(mT)=f((\sqrt{mT}+S)^2)=f((\sqrt{m}+\sqrt{n})^2T)=f(2\sqrt{mn}T).$$

根据与上面相同的论证可知$2\sqrt{mn}\in\mathbb{N}$。但这件事不可能对任意m都成立，矛盾！ □

注. 如果去掉$f(x)\ne f(0)\,(\forall x\in(0,T))$的条件，则结论不成立。反例的构造主要涉及代数和数论知识，因此这里不作讨论[2]。如果再去掉最小正周期的限制，则有简单的反例如下。称$\alpha\in\mathbb{R}$是代数数*(algebraic number)*，如果存在$P\in\mathbb{Q}[x]\setminus\{0\}$使得$P(\alpha)=0$。不是代数数的实数称为超越数*(transcendental number)*。可以证明，两个代数数进行加减乘除，得到的还是代数数；以及如果$Q\in\mathbb{Q}[x]\setminus\{0\}$且$Q(\alpha)$是代数数，则$\alpha$也是代数数。定义

$$f(x)=\begin{cases}1, & x\text{是代数数},\\ 0, & x\text{是超越数}.\end{cases}$$

事实上，f的周期恰好由所有代数数组成，并且$f(x^2)=f(x)$。

练习题

设A,B是两个集合，定义对称差(symmetric difference)为

$$A\Delta B:=(A\setminus B)\cup(B\setminus A)=(A\cup B)\setminus(A\cap B).$$

[2]此事在mathoverflow上早有讨论，亦可参考知乎用户“cyb酱”的文章《周期函数的一件琐事（短文）》。

用自然语言描述，$A\Delta B$由那些恰好只在A,B之一中出现的元素组成。

在下一题中，用示性函数(characteristic function)刻画集合会带来很多方便，即定义

$$\chi_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

反过来，任给一个取值$\{0,1\}$的函数f，都可以唯一确定一个集合$f^{-1}(1)$。因此示性函数与集合是一一对应的，要证明两个集合相等只需说明它们的示性函数相同。

容易看出集合运算与示性函数有如下关系：

$$\chi_{A^C} = 1 - \chi_A, \quad \chi_{A\cap B} = \chi_A\chi_B,$$

$$\chi_{A\cup B} = \chi_A + \chi_B - \chi_A\chi_B, \quad \chi_{A\Delta B} \equiv \chi_A + \chi_B \pmod 2.$$

前两个等式是直截了当的，而后两个可以在此基础上结合De Morgan律以及$\chi_A^2 = \chi_A$推出：

$$\chi_{A\cup B} = \chi_{(A^C\cap B^C)^C} = 1 - (1-\chi_A)(1-\chi_B) = \chi_A + \chi_B - \chi_A\chi_B,$$

$$\chi_{A\Delta B} = \chi_{A\cup B}(1-\chi_{A\cap B}) = (\chi_A + \chi_B - \chi_A\chi_B)(1-\chi_A\chi_B) = \chi_A + \chi_B - 2\chi_A\chi_B.$$

值得注意的是，第四个等式是模2意义下的，但示性函数取值$\{0,1\}$，因此也就唯一确定了。

题 1.4. 设A,B,C是三个集合，证明对称差具有如下性质：

(a) $A\Delta(B\Delta C) = (A\Delta B)\Delta C$。

(b) $A\cap(B\Delta C) = (A\cap B)\Delta(A\cap C)$。

证明. (a) $\chi_{A\Delta(B\Delta C)} \equiv \chi_A + \chi_B + \chi_C \equiv \chi_{(A\Delta B)\Delta C} \pmod 2$。

(b) $\chi_{A\cap(B\Delta C)} \equiv \chi_A(\chi_B + \chi_C) \equiv \chi_{A\cap B} + \chi_{A\cap C} \equiv \chi_{(A\cap B)\Delta(A\cap C)} \pmod 2$。 □

称一个集合$X\subset\mathbb{R}$是稠密的(dense)，如果对任意非空开区间I，都有$X\cap I\neq\emptyset$。非平凡的例子是$\mathbb{Q}$和$\mathbb{R}\setminus\mathbb{Q}$都在$\mathbb{R}$中稠密。

题 1.5. 设$f\colon\mathbb{R}\to\mathbb{R}$是一个函数，定义$f$的全体“周期”组成的集合

$$P(f) := \{T\in\mathbb{R} \mid f(x+T) = f(x),\ \forall x\in\mathbb{R}\}.$$

证明：要么$P(f)$稠密，要么存在唯一的$\alpha\geq 0$使得$P(f) = \{n\alpha \mid n\in\mathbb{Z}\}$。

证明. 如果f不是周期函数，则$P(f) = \{0\}$，结论对$\alpha = 0$成立。以下假设f有正周期。显然结论中两种情况互斥，并且满足要求的α唯一。欲证二者必居其一，按照非空集合$P_+(f) := \{T\in P(f) \mid T>0\}$是否有最小元素讨论。注意到$P(f)$关于加法和整数乘法封闭。

情况*1*：如果$\alpha := \inf P_+(f) > 0$且$\alpha \in P_+(f)$。对任意$T \in P(f) \setminus \{0\}$，显然$|T| \in P_+(f)$。设$n\alpha \leq |T| < (n+1)\alpha (n \in \mathbb{N})$，则$|T| - n\alpha \in P(f)$且$0 \leq |T| - n\alpha < \alpha$。由$\alpha$的最小性可知$|T| = n\alpha$，进而易见$P(f) = \{n\alpha \mid n \in \mathbb{Z}\}$。

情况*2*：如果$\alpha := \inf P_+(f) > 0$且$\alpha \notin P_+(f)$，则存在$T_1, T_2 \in P_+(f)$使得$\alpha < T_1 < T_2 < 2\alpha$，于是$T_2 - T_1 \in P_+(f)$且$T_2 - T_1 < \alpha$，与$\alpha$是下确界矛盾！这种情况不存在。

情况*3*：如果$\inf P_+(f) = 0$，则对任意非空开区间(a, b)，存在$T \in P_+(f)$满足$T < b - a$。设$nT \leq a < (n+1)T \, (n \in \mathbb{Z})$，则$(n+1)T \in P(f) \cap (a, b)$。故$P(f)$稠密。 □

注. 非常值周期函数未必有最小正周期，比如*Dirichlet*函数$D(x) = \chi_Q(x)$以任何有理数为周期。在引进连续性之后，利用本题的结果容易证明，非常值连续函数一定有最小正周期。

用$\mathbb{R}[x]$表示全体实系数多项式组成的集合。下一题需要用到关于多项式差分的简单事实：设$P(x)$是$n \, (n \geq 1)$次多项式，则对任意$T > 0$，有$P(x+T) - P(x)$是$n-1$次非零多项式。这是因为按照二项式展开，除了$(x+T)^n - x^n$中出现的nTx^{n-1}之外不会再产生$n-1$次项。一般教材上的习题是证明$\sin x^2$不是周期函数，事实上x^2可以换成任何至少二次的多项式。

题 1.6 *. 设$P \in \mathbb{R}[x]$且$\deg P \geq 2$，证明：

(a) $\sin P(x)$不是周期函数。

(b) 设$f \colon \mathbb{R} \to \mathbb{R}$是非常值连续周期函数，则$f \circ P$不是周期函数。

证明. 先给出(a)的一个不本质的做法（依赖三角函数的性质），显然(b)蕴含更一般的情况。

(a) 反证法，假设$T > 0$是$\sin P(x)$的周期，则

$$0 = \sin P(x+T) - \sin P(x) = 2 \cos \frac{P(x+T) + P(x)}{2} \sin \frac{P(x-T) - P(x)}{2}.$$

断言右边两项只会有可数个零点。事实上，如果$Q(x)$是非常值多项式，则

$$(\sin Q(x))^{-1}(0) = \bigcup_{n \in \mathbb{Z}} Q^{-1}(n\pi).$$

而$Q(x) = n\pi$最多只有$\deg Q$个解，有限集的可数并是可数集，所以$\sin Q(x)$只有可数个零点。同理可知$\cos Q(x)$也只有可数个零点。因为$\deg P \geq 2$，所以$P(x+T) \pm P(x)$都不是常数，这就说明了上式右边只有可数个零点，不可能恒为0，矛盾！

(b) 反证法，假设$f \circ P$是周期函数，目标是证明f只能是常值函数，从而导致矛盾。不妨设P的首项系数为正（否则用$f(-x)$代替$f(x)$），则当x充分大时$P(x)$和$P'(x)$都单调递增

趋于∞（这里用到了$\deg P \geq 2$）。假设f有周期$T > 0$，则存在$N > 0$使得当$n > N$时，$P(x) = nT$在$[0, \infty)$上有唯一解x_n，其中x_n关于n单调递增且$x_{n+1} - x_n \to 0$。注意

$$f(P(x_n)) = f(nT) = f(0).$$

利用$x_{n+1} - x_n \to 0$和$f \circ P$的周期性可知$\{x \mid f(P(x)) = f(0)\}$稠密（思考理由），结合连续性可知$f \circ P$是常数。最后，$P$的值域包含$[P(0), \infty)$，再由$f$的周期性得$f$是常数。 □

题 1.7 *. 设$a, b, c, d \in \mathbb{R}$，且$ad - bc = 1$，证明：对任意$z \in \mathbb{C}$，有

$$|(az + b)(cz + d)| \geq |\operatorname{Im}(z)|,$$

并讨论取等的充要条件。

证明*1.* 设$z = x + iy (x, y \in \mathbb{R})$，则

$$\begin{aligned} &|(az + b)(cz + d)| \geq |\operatorname{Im}(z)|, \\ \Longleftrightarrow &((ax + b)^2 + a^2y^2)((cx + d)^2 + c^2y^2) \geq y^2, \\ \Longleftrightarrow &(ax + b)^2(cx + d)^2 + (c^2(ax + b)^2 + a^2(cx + d)^2 - 1)y^2 + a^2c^2y^4 \geq 0. \end{aligned}$$

因为$1 = ad - bc = a(cx + d) - c(ax + b)$，所以$y^2$前面的系数又可以写为

$$c^2(ax + b)^2 + a^2(cx + d)^2 - (a(cx + d) - c(ax + b))^2 = 2ac(ax + b)(cx + d).$$

这就证明了

$$|(az + b)(cz + d)| \geq |\operatorname{Im}(z)| \Longleftrightarrow ((ax + b)(cx + d) + acy^2)^2 \geq 0.$$

所以该不等式成立，并且取等的充要条件是$(ax + b)(cx + d) + acy^2 = 0$。 □

证明*2.* 只需考虑$\operatorname{Im}(z) \neq 0$的情况。因为$a, c$不全为0，并且条件和结论都关于$a, c$对称，因此不妨设$c \neq 0$。考虑辅助函数（注意分母非零）

$$f(z) = \frac{az + b}{cz + d},$$

则直接计算可知

$$\operatorname{Im}(f(z)) = \frac{\operatorname{Im}((az + b)(c\overline{z} + d))}{|cz + d|^2} = \frac{(ad - bc)\operatorname{Im}(z)}{|cz + d|^2} = \frac{\operatorname{Im}(z)}{|cz + d|^2}.$$

代入$|\operatorname{Im}(f(z))| \leq |f(z)|$即证，取等条件是$f(z)$是纯虚数（或$z \in \{-b/a, -d/c\}$）。 □

注. 证明*2*中的辅助函数$f(z)$被称为分式线性变换*(fractional linear transformation)*。在双曲几何中，上半平面$\{z \in \mathbb{C} \mid \operatorname{Im}(z) > 0\}$连同双曲度量$\frac{dz}{\operatorname{Im}(z)}$构成双曲平面的模型，而全体分式线性变换构成其自同构群。特别的，f保持双曲度量，即$\frac{|f'(z)|}{\operatorname{Im} f(z)} = \frac{1}{\operatorname{Im}(z)}$，这是证明的关键。

补充：集合的势

正如可数集表述为存在到$\mathbb{N}$有双射一样，可以用存在双射作为两个集合大小相同的标准。更一般的，可以用单射或满射去比较两个集合的大小。在这一小节中总是假设集合非空。

定义 1.8 (集合的势)**.** 设A, B是两个非空集合，如果存在单射$A \hookrightarrow B$，则称A的势不大于B，记为$A \prec B$；如果存在满射$B \twoheadrightarrow A$，则称B的势不小于A，记为$B \succ A$；如果存在A到B的双射，则称A与B等势，记为$A \sim B$。

势的大小关系没有公认的记号。这里采用$A \prec B$并不排除二者的势相等，正如集合的包含关系$\subset$通常不否认相等。势也被称为基数(cardinality)，势的大小即基数的大小。有了定义立即产生的两个问题：(1)$A \prec B$是否等价于$B \succ A$？(2)势的大小关系是否可以类比于数的大小关系，换言之势的大小是否满足合理的性质？

前一个问题是相对容易回答的，但需要用到选择公理(axiom of choice)。用自然语言描述，选择公理是说："任给一些非空集合，可以同时从每个集合里挑选一个元素。" 这看起来很显然，但非平凡之处在于元素是同时被取出的。尤其在集合个数不可数时，没有办法逐集合选取元素。数学中很多重要的存在性结果和反例都离不开选择公理，例如无限维线性空间中基的存在性、交换幺环中极大理想的存在性、Lebesgue不可测集的存在性等等。

选择公理可以用Cartesian积刻画。若A_1, A_2非空，则$A_1 \times A_2$的元素唯一确定了一个映射$f\colon \{1,2\} \to A_1 \cup A_2$，满足$f(i) \in A_i \, (i = 1, 2)$；反过来，如果$f$是这样一个映射，则唯一确定了$A_1 \times A_2$中的元素。一般的，可以定义一族集合$A_\alpha \, (\alpha \in I)$的Cartesian积为：

$$\prod_{\alpha \in I} A_\alpha := \{f\colon I \to \bigsqcup_{\alpha \in I} A_\alpha \mid f(\alpha) \in A_\alpha, \ \forall \alpha \in I\}.$$

选择公理等价于说，如果A_α都非空，则$\prod_{\alpha \in I} A_\alpha$也非空。

题 1.9. 设A, B是两个非空集合，证明：$A \prec B$当且仅当$B \succ A$。

证明. 先证明"仅当"。如果$f\colon A \hookrightarrow B$是单射，则对任意$b \in B$，原像$f^{-1}(b)$至多有一个点。任取$a \in A$，定义$g\colon B \to A$为（这里对一元集和其元素不加区分，仅仅在记号上是不严谨的）

$$g(b) := \begin{cases} f^{-1}(b), & b \in f(A), \\ a, & b \notin f(A). \end{cases}$$

则g是良定义的满射。接下来证明"当"。如果$g\colon B \twoheadrightarrow A$是满射，则$g^{-1}(a) \, (a \in A)$非空。根据选择公理，存在$f\colon A \to B$使得$f(a) \in g^{-1}(a)$。因为原像集是两两不交的，所以$f$是单射。 □

下面来看后一个问题。期待的性质被归入所谓偏序集的概念中，需要首先引进二元关系。

定义 1.10 (二元关系)**.** 设X, Y是非空集合，X与Y上的二元关系是指$X \times Y$的子集R。用xRy表示$(x, y) \in R$，意为二元对(x, y)满足R。当$X = Y$时，简称为X上的二元关系。

比如说，映射$f\colon X \to Y$给出一个二元关系，即xRy当且仅当$y = f(x)$。由映射给出的二元关系R满足对每个$x \in X$，都存在唯一的y使得xRy成立。反过来，任何一个这样的二元关系都给出一个映射。集合X上一类重要的关系被称为等价关系。

定义 1.11 (等价关系)**.** 设X是非空集合，X上的二元关系$\sim$称为等价关系，如果

(a) (自反性)对任意$x \in X$，有$x \sim x$；

(b) (对称性)对任意$x, y \in X$，如果$x \sim y$，则$y \sim x$；

(c) (传递性)对任意$x, y, z \in X$，如果$x \sim y$且$y \sim z$，则$x \sim z$。

从精神上说，任何对象之间相等或者等价的合理定义都应该是一种等价关系。

题 1.12. 集合的等势是一个等价关系。

证明. 恒同映射给出自反性，对称性是因为取逆保持双射，传递性是因为复合保持双射。 □

注. *Russell*悖论说，所有集合不构成集合，而是一个比集合更大的概念：类*(class)*。后文会给出*Russell*悖论的原理。因此严格来讲，应该先定义类的概念，再将二元关系的定义推广到类上。然而，最简单的处理方法就是直接忽略这个问题。

在定义商(quotient)这样的代数结构时，等价关系的概念必不可少，尤其是需要用到如下的等价类划分。有时还需要用选择公理，从每个等价类中取出一个代表元组成集合。

题 1.13. 设$\sim$是集合X上的等价关系，对任意$x \in X$，定义$[x]$是"x所在的等价类"，即

$$[x] := \{y \in X \mid y \sim x\}.$$

证明：对任意$x, y \in X$，要么$[x] = [y]$，要么$[x] \cap [y] = \emptyset$。

证明. 假设$z \in [x]$，则对任意$w \in X$，根据$z \sim x$以及传递性可知$w \sim x$当且仅当$w \sim z$，这就说明$[z] = [x]$。对于任何两个等价类$[x], [y]$，如果$z \in [x] \cap [y]$，则$[x] = [z] = [y]$。 □

另一类重要的关系被称为偏序关系，由此给出偏序集的概念。

定义 1.14 (偏序集)**.** 设X是非空集合，X上的二元关系$\prec$称为偏序关系，如果

(a) (自反性)对任意$x \in X$，有$x \prec x$；

(b) (反对称性)对任意$x, y \in X$，如果$x \prec y$且$y \prec x$，则$x = y$；

(c) (传递性)对任意$x, y, z \in X$，如果$x \prec y$且$y \prec z$，则$x \prec z$。

二元对$(X, \prec)$称为偏序集*(partially ordered set)*。

比如说，$(\mathbb{R}, \leq)$就是一个偏序集，其中$\leq$就是通常的实数大小关系。另一个例子是集合的包含关系：设X是非空集合，$2^X := \{A \mid A \subset X\}$称为$X$的幂集，则$(2^X, \subset)$也构成偏序集。事实上，集合连同势的大小关系，在商掉等势之后构成偏序集。所谓“商掉等势”，就是等势的集合被认为是相同的。此时自反性和传递性是显然的，下面给出反对称性的证明。

题 1.15 *. 设A, B是非空集合，如果$A \prec B$且$B \prec A$，则$A \sim B$。

证明. 存在单射$f \colon A \hookrightarrow B$和$g \colon B \hookrightarrow A$。记$A_0 = A$和$B_0 = B$，递归定义$A_n = g(B_{n-1})$和$B_n = f(A_{n-1})$ $(n \in \mathbb{N})$。再记$A_\infty = \bigcap_{n \in \mathbb{N}_0} A_n$和$B_\infty = \bigcap_{n \in \mathbb{N}_0} B_n$。有如下观察：

观察*1*：归纳易知，$A_0 \supset A_1 \supset \cdots$，$B_0 \supset B_1 \supset \cdots$。

观察*2*：f限制在$A_n \setminus A_{n+1}$上给出到$B_{n+1} \setminus B_{n+2}$的双射，$g$限制在$B_n \setminus B_{n+1}$上给出到$A_{n+1} \setminus A_{n+2}$的双射。只需验证$f(A_n \setminus A_{n+1}) \subset B_n \setminus B_{n+1}$以及单性、满性，留作练习。

观察*3*：f限制在A_∞上给出到B_∞的双射，g限制在B_∞上给出到A_∞的双射。对此只证f的情况。首先说明f将A_∞映到B_∞。事实上，如果$a \in A_n$ $(\forall n \in \mathbb{N}_0)$，则$f(a) \in B_{n+1}$ $(\forall n \in \mathbb{N}_0)$；又因为$B_n$单调递减，所以$f(a) \in B_\infty$。其次，作为限制映射自动有单性，下证满性。设$b \in B_\infty$，因为$b \in B_{n+1} (\forall n \in \mathbb{N}_0)$，所以存在$a_n \in A_n$使得$f(a_n) = b$；又因为$f$是单射，所以这些$a_n$彼此相等、是同一个$a \in A_\infty$，故这是满射。

有了这些观察，剩下的证明就是将A, B分解成

$$A = \left(\bigsqcup_{n=0}^{\infty} A_{2n} \setminus A_{2n+1} \right) \sqcup \left(\bigsqcup_{n=0}^{\infty} A_{2n+1} \setminus A_{2n+2} \right) \sqcup A_\infty,$$

$$B = \left(\bigsqcup_{n=0}^{\infty} B_{2n+1} \setminus B_{2n+2} \right) \sqcup \left(\bigsqcup_{n=0}^{\infty} B_{2n} \setminus B_{2n+1} \right) \sqcup B_\infty.$$

根据上面的观察可知f给出$A_{2n} \setminus A_{2n+1}$到$B_{2n+1} \setminus B_{2n+2}$的双射，$g^{-1}$给出$A_{2n+1} \setminus A_{2n+2}$到$B_{2n} \setminus B_{2n+1}$的双射，$f$（或$g^{-1}$）给出$A_\infty$到$B_\infty$的双射。具体写出来，定义$h \colon A \to B$为

$$h(a) = \begin{cases} f(a), & a \in (\bigsqcup_{n=0}^{\infty} A_{2n} \setminus A_{2n+1}) \sqcup A_\infty, \\ g^{-1}(a), & a \in \bigsqcup_{n=0}^{\infty} A_{2n+1} \setminus A_{2n+2}. \end{cases}$$

则h是A到B的双射，故$A \sim B$。 □

以上说明了全体集合连同势的大小关系构成偏序集。还可以问这是否是“全序”，即任意两个集合是否总可以比大小。如果偏序集$(X, \prec)$中的两个元素x, y满足$x \prec y$或$y \prec x$，则

称x和y是可比较的。任何两个元素都可比较的偏序集称为全序集。应当注意的是，一般的偏序集未必是全序的，比如$(2^X, \subset)$在X包含至少两个元素时就不是全序集。

下面需要用到Zorn引理，首先介绍若干概念。设$(X, \prec)$是一个偏序集，$Y \subset X$是任意子集，$x \in X$是任意元素。如果对任意$y \in Y$都有$y \prec x$，则称x是Y的上界(upper bound)。如果对任意$z \in X$，$x \prec z$蕴含着$z = x$，则称x是X中的极大元(maximal element)。如果Y中任何两个元素都是可比较的（即Y是全序的），则称Y是一条链(chain)。

定理 1.16 * (Zorn引理)**.** *如果非空偏序集$(X, \prec)$中任何链都有上界，则X有极大元。*

可以证明Zorn引理与选择公理是等价的，但由选择公理推出Zorn引理是非平凡的，这里不作讨论；由Zorn引理推出选择公理的方法在后续的证明中也会用到，因此在这里给出。

用Zorn引理证明选择公理. 假设$A_\alpha(\alpha \in I)$是一族非空集合，需要证明$\prod A_\alpha$非空。定义集合X由所有符合如下要求的映射$f\colon D(f) \to \cup A_\alpha$组成：$\emptyset \neq D(f) \subset I$，并且对任意$\alpha \in D(f)$都有$f(\alpha) \in A_\alpha$。（严格来讲应该写成二元对$(f, D(f))$，因为定义域也是需要指明的。）显然$X$是非空集合。在$X$上定义偏序关系如下：$f \prec g$当且仅当$D(f) \subset D(g)$且$g|_{D(f)} = f$。换言之，$f \prec g$意味着$g$是$f$的延拓，或者说$f$是$g$的限制。

容易验证$(X, \prec)$构成偏序集，下面验证Zorn引理的条件成立。假设$Y \subset X$是一条链，其中的元素记为$f_\beta(\beta \in Y)$。定义$F\colon D(F) \to \cup A_\alpha$如下：定义域$D(F) = \cup_\beta D(f_\beta)$，对于$x \in D(f_\beta)$，定义$F(x) := f_\beta(x)$。首先需要检查$F$是否良定义，这里可能出现的问题是，如果$x$同时属于$D(f_{\beta_1})$和$D(f_{\beta_2})$，是否有$f_{\beta_1}(x) = f_{\beta_2}(x)$？答案是肯定的：因为$Y$是一条链，可以不妨设$f_{\beta_1} \prec f_{\beta_2}$，又因为$x \in D(f_{\beta_1})$，所以$f_{\beta_1}(x) = f_{\beta_2}|_{D(\beta_1)}(x) = f_{\beta_2}(x)$。根据定义，显然$F \in X$，且对任意$f_\beta \in Y$都有$f_\beta \prec F$，所以$F$是$Y$的上界。

根据Zorn引理，X有极大元，记为G。断言$D(G) = I$。事实上，如果存在$\gamma \in I \setminus D(G)$，则任取$a_\gamma \in A_\gamma$，定义$D(H) := D(G) \cup \{\gamma\}$和映射$H\colon D(H) \to \cup A_\alpha$为

$$H(\alpha) := \begin{cases} G(\alpha), & \alpha \in D(G), \\ a_\gamma, & \alpha = \gamma, \end{cases}$$

则$H \in X$。易见$G \prec H$且$G \neq H$，与G是极大元矛盾！所以$D(G) = I$，即$G \in \prod A_\alpha$。 □

题 1.17 *. *设A, B是非空集合，证明：$A \prec B$和$B \prec A$中至少有一个成立。*

证明. 定义集合X由所有满足如下要求的映射$f\colon D(f) \to B$组成：$\emptyset \neq D(f) \subset A$，并且$f$是单射。显然$X$是非空的，同上面的论证在$X$上定义偏序关系为$f \prec g$当且仅当$g$是$f$的延拓。为了验证Zorn引理的条件，方法和上面一样，但还需验证F是单射。这是因为f_β都是单射，并且构成链，具体细节留作练习。根据Zorn引理，X中有极大元F，分两种情况：

情况1：如果$D(F) = A$，则F是A到B的单射，从而$A \prec B$。

情况2：如果$D(F) \neq A$，断言$F(A) = B$（即F是满射），从而$B \prec A$。事实上，如果$F(A) \neq B$，则可以取$c \in A \setminus D(F)$和$d \in B \setminus F(A)$，定义$D(G) := D(F) \cup \{c\}$和映射

$$G(a) = \begin{cases} F(a), & a \in D(F), \\ d, & a = c, \end{cases}$$

则$G\colon D(G) \to B$也是单射，从而与F的极大性矛盾！ □

最后一题说明不存在势最大的集合。

题 1.18 *. 设X是非空集合，则2^X的势严格大于X的势。

证明. 显然$X \prec 2^X$。用反证法证明不能取等，假设存在满射$f\colon X \to 2^X$，考虑集合

$$Y := \{y \in X \mid y \notin f(y)\} \subset X.$$

因为f是满射，所以存在$z \in X$使得$f(z) = Y$。那么，无论z是否属于Y都会导致矛盾：如果$z \in Y$，则根据Y的定义有$z \notin f(z) = Y$；如果$z \notin Y = f(z)$，则根据定义有$z \in Y$。 □

注. 这个证明可以与*Russell*悖论对比，即如果存在一个集合X由全体集合构成，考虑集合

$$Y := \{A \in X \mid A \notin A\} \subset X,$$

并考虑$Y \in Y$是否成立，便可得到悖论。若仔细思考，会发现$A \in A$是一个很奇怪的条件，被所谓“正则公理”所排除。由此发展出的*ZFC*公理体系通过对集合运算加以限制，用逻辑语言代替自然语言定义集合，从而保证了这一悖论仅是佯谬。

补充：多项式

多项式的理论应属代数学，然而新教材第一章习题却屡次涉及。为了避免循环论证，建立基础的理论并理清逻辑关系是必要的。本节考虑一般的域K上的多项式，不了解域的定义的读者可以始终认为$K = \mathbb{Q}, \mathbb{R}$或$\mathbb{C}$，但诸如有限域$\mathbb{F}_p$、$p$进数域$\mathbb{Q}_p$、数域（即$\mathbb{Q}$的有限扩张）、以及有理函数域$\mathbb{C}(x)$等例子在代数、数论中格外重要。

熟知多项式是指形式的表达式

$$P(x) = c_n x^n + \cdots + c_0, \quad c_i \in F.$$

通常为了明确起见假设$a_n \neq 0$。将P的次数记为$\deg(P) = n$，并规定0多项式的次数为$-\infty$。全体多项式组成的集合记为$K[x]$。从一般的定义来讲，x并不仅仅代表F中的元素，而是抽象的变量；P也不仅仅视为F到自身的映射，而是表达式本身。定义两个多项式相等当且仅当次

数相同且对应项的系数分别相等。当F是有限域时，P,Q作为多项式相等比$P(a)=Q(a)$对任意$a\in K$成立更强；但在无限域上两种定义是等价的（稍后会给出证明）。经典的例子是$\mathbb{F}_p:=\mathbb{Z}/p\mathbb{Z}$上的多项式$x^p$和$x$，二者并不相等，但在每个整数处的取值都模$p$相同（即Fermat小定理）。此外，只有在这样的定义下才有消去律：如果$P\neq 0$且$PQ_1=PQ_2$，则$Q_1=Q_2$。

多项式的乘法按照自然的方式定义。容易看出，如果P,Q都是非零多项式，则

$$\deg(PQ)=\deg(P)+\deg(Q).$$

如果认为$-\infty$与任意有限数或$-\infty$相加仍是$-\infty$，那么上式对任何$P,Q\in K[x]$都成立。

定义 1.19 (多项式的整除)**.** 对于$P,Q\in K[x]$，称$Q\mid P$，如果存在$R\in K[x]$使得$P=QR$。

当$Q\mid P$时，商P/Q也是多项式。与整数类似，引进带余除法的概念十分重要。事实上，以下大部分论证都可以推广到一般的Euclid整环，整数以及多项式是其中两个最重要的模型。

题 1.20. 设$P,Q\in K[x]$且$Q\neq 0$，证明：存在唯一的$A,B\in K[x]$满足$\deg(B)<\deg(Q)$且

$$P(x)=A(x)Q(x)+B(x).$$

证明. $\deg(Q)=0$的情况是显然的，不妨设$\deg(Q)>0$。先证明唯一性，假设

$$P(x)=A_1(x)Q(x)+B_1(x)=A_2(x)Q(x)+B_2(x),$$

并且$\deg(B_1),\deg(B_2)<\deg(Q)$，则

$$(A_1(x)-A_2(x))Q(x)=B_2(x)-B_1(x).$$

如果$A_1\neq A_2$，则左边的次数至少是$\deg(Q)$；而右边的次数严格小于$\deg(Q)$，矛盾！因此只能$A_1=A_2$，从而易见$B_1=B_2$。

下证存在性，对$\deg(P)$归纳。设$\deg(Q)=n$且Q的首项系数为$b\neq 0$。如果$\deg(P)<n$，取$A=0$和$B=P$即可，这就完成了奠基。如果结论对$\deg(P)=k-1$的情况成立，考虑$\deg(P)=k\geq n$的情况。设$P(x)$的首项为ax^k，其中$a\neq 0$。于是

$$P(x)-\frac{ax^{k-n}Q(x)}{b}$$

是次数至多为$k-1$次的多项式（注意这里用到了$k\geq n$）。根据归纳假设可知

$$P(x)-\frac{ax^{k-n}Q(x)}{b}=A_0(x)P(x)+B(x),$$

其中$\deg B<n$。取$A(x)=A_0(x)+ax^{k-n}/b$，则有$P(x)=A(x)Q(x)+B(x)$。 $\square$

注. 作为唯一性的推论，$Q\mid P$当且仅当带余除法中$B=0$。

下一题是多项式的根（即零点）的重要等价刻画。

题 1.21. 设$P \in K[x]$，对于$a \in K$，证明：$P(a) = 0$当且仅当$x - a \mid P(x)$。

证明. 考虑$P(x)$关于$x - a$的带余除法，存在$Q \in K[x]$使得

$$P(x) = (x - a)Q(x) + c,$$

其中c是次数严格小于1的多项式，从而只能是常数。令$x = a$可知$c = P(a)$，即

$$P(x) = (x - a)Q(x) + P(a).$$

由上一题的注可知结论成立。 □

建立根与整除的关系之后，可以引进重根的概念。

定义 1.22 (重根)**.** 设$P \in K[x]\backslash\{0\}$，称$a \in K$是$P$的$n$重根，如果$(x-a)^n \mid P$且$(x-a)^{n+1} \nmid P$。

这里0重根即不是根。根据定义，a是$P(x)$的n重根且$n \geq 1$，则a是$P(x)/(x - a)$的$n - 1$重根。反复利用上一题可知，a是$P(x)$的n重根当且仅当存在$Q \in K[x]$满足$Q(a) \neq 0$且

$$P(x) = (x - a)^n Q(x).$$

直观上看对任意$b \neq a$，应该有b在P中的重数等于b在Q中的重数。然而这件事的证明依赖仅仅看似平凡的事实：如果$(x - b)^k \mid (x - a)^n Q(x)$，是否一定有$(x - b)^k \mid Q(x)$？最简单的证明是对$k$归纳，但缺乏一般性。为此，以下引进公因式的概念，并类比整数互素的性质。

题 1.23. 设$P, Q \in K[x]$，记P, Q的公因式集合为

$$\Gamma(P, Q) := \{R \in K[x] \mid R \mid P \text{且} R \mid Q\}.$$

(a) 证明：对任意$A \in K[x]$，有$\Gamma(P - AQ, Q) = \Gamma(P, Q)$。

(b) 证明：存在$A, B \in K[x]$，使得$R \in \Gamma(P, Q)$当且仅当$R \mid AP + BQ$。

证明. (a) 根据定义验证即可。

(b) 这个方法叫辗转相除法，来自Euclid的《几何原本》。如果$\deg(P) \geq \deg(Q)$且$Q \neq 0$，考虑带余除法$P = AQ + B$，则$\Gamma(P, Q) = \Gamma(B, Q)$，此时$\deg(B)$严格比$\deg(P)$更小。如果$B$不是0，那么又可以对$Q$关于$B$进行带余除法。如此下去，每一步都会让$\deg(P) + \deg(Q)$严格减小，直到某一项变为0，为止。于是最终得到一个多项式S使得$\Gamma(P, Q) = \Gamma(S, 0)$，而显然$\Gamma(S, 0)$由全体$S$的因式构成。于是$R \in \Gamma(P, Q)$当且仅当$R \mid S$。

最后，注意到带余除法的每一步得到的式子都形如$AP + BQ$，从而结论得证。 □

注. 注意到*(b)*中给出的S也是P,Q的公因子，并且是全体公因子中次数最大的，因此这被称为P,Q的最大公因子。乘以一个非零常数不会改变最大公因子的性质，因此通常只说首一多项式是最大公因子，记为$\gcd(P,Q)$。如果P,Q不全为0，则$\gcd(P,Q)\neq 0$（至少1是一个公因式）。能够将$\gcd(P,Q)$表示为$AP+BQ$的形式，这一事实被称为*Bézout*定理。

特别的，如果$Q(x)=x-a$是一次多项式，则$\gcd(P,x-a)$只能是1或者$x-a$。由此可知，a不是$P(x)$的根当且仅当$\gcd(x-a,P(x))=1$。这是互素的特例。

定义 1.24 (多项式的互素)**.** 称$P,Q\in K[x]$互素，如果$\gcd(P,Q)=1$。

互素的多项式具有与互素整数类似的性质。

题 1.25. 设$P,Q,R\in K[x]$且$\gcd(P,Q)=1$，证明：$\gcd(P,QR)=\gcd(P,R)$。

证明. 只需证明公因式集合相同。显然P,R的公因式都是P,QR的公因式，只需证明反过来的包含关系。根据Bézout定理，存在$A,B\in K[x]$使得$AP+BQ=1$，两边乘以R可知

$$R=AR\cdot P+B\cdot QR$$

如果$S\in K[x]$是P和QR的公因式，则上式表明$S\mid R$，证毕。 □

类比整数的素因子分解，多项式也可以唯一分解为不可约因式的乘积。

定义 1.26 (不可约多项式)**.** 称$P\in K[x]$是不可约多项式，如果$\deg(P)>1$，并且P不能写成两个非常值多项式乘积。

一次多项式必然是不可约的，而$\mathbb{Q}[x]$中存在任意次数的不可约多项式（Eisenstein判别法的经典应用）。如果P是不可约多项式，则P的因式（在相差常数倍的意义下）只有1和P。从而对任意$Q\in K[x]$，必有$\gcd(P,Q)=1$或P。作为上一题的推论，如果P不可约且$P\mid QR$，则$P\mid Q$或$P\mid R$。因此不可约多项式类比于素数，下面给出唯一分解定理。

题 1.27. *(a)* 设$P\in K[x]\setminus\{0\}$，则存在有限多个不可约多项式$Q_1,\dots,Q_n$使得

$$P=Q_1\cdots Q_n.$$

进一步，n由P唯一确定，并且$Q_1,\dots,Q_n$在至多相差排列顺序的意义下也是唯一的。

(b) 设$P\in K[x]\setminus\{0\}$，则P至多有$\deg(P)$个根（计重数）。进一步，设P的所有互异根为$a_1,\dots,a_k$，重数分别为$e_1,\dots,e_k$，则

$$P(x)=Q(x)\prod_{i=1}^{k}(x-a_i)^{e_i}.$$

其中$Q\in K[x]$是没有根的多项式。

证明. (a) 先证明存在性。如果P不可约，取$n=1$和$Q_1=P$即可；特别的，$\deg(P)=1$时这样的分解存在。对$\deg(P)$归纳，如果可约，那么P可以写成两个多项式的乘积$P=QR$，并且Q,R的次数都严格比P更小，利用归纳假设即可。

再证明唯一性。如果另有一种分解$P=R_1\cdots R_m$，其中$R_1\cdots R_m$是不可约多项式，则$Q_1\mid P$蕴含了存在$1\le k\le m$使得$Q_1\mid R_k$。因为$\deg(Q_1)\ge 1$且R_k不可约，所以只能$Q_1=R_k$。现在可以约去Q_1与R_k。剩下的容易归纳完成。

(b) 只需证明如果P有互异根$a_1,\dots,a_k$，重数分别至少为$e_1,\dots,e_k$，则存在$Q\in K[x]$使得

$$P(x)=Q(x)\prod_{i=1}^{k}(x-a_i)^{e_i}.$$

事实上，如果这一断言得证，比较两边次数可知$\sum_{i=1}^{k}e_i\le\deg(P)$；而如果进一步$P$没有其他根，那么$Q$也就没有根。为此，只需注意到$x-a_i$都是不可约多项式，并且$a_i$的重数至少是$e_i$保证了$P$的不可约多项式分解中$x-a_i$至少出现$e_i$次（见注）。利用(a)即证。 □

注. 利用唯一分解可以重新理解整除关系，$P\mid Q$当且仅当任何不可约多项式在P中的次数都不高与在Q中的次数。这一描述类似于整数的整除只需比较素因子次数。特别的，a是P的n重根当且仅当不可约多项式分解中$x-a$出现了n次。

如果不借助不可约分解，那么一些看似显然的事情也并不直接。读者可以尝试按定义证明$(x-1)^n$的因式只有$(x-1)^k$ $(k=0,1,\dots,n)$。

注. 如果不超过n次的多项式P至少有$n+1$个根，由*(b)*可知$P=0$。进一步，如果P,Q是两个不超过n次的多项式，且$P(x)=Q(x)$至少有$n+1$个根，则$P=Q$。这说明在无限域K上，多项式由其在K上的取值唯一确定。

如果n次多项式P有n个根$a_1,\dots,a_n$（计重数），则

$$P(x)=c\prod_{i=1}^{n}(x-a_i).$$

其中$c\in K$是P的首项系数。进一步，如果$P(x)=c_nx^n+c_{n-1}x^{n-1}+\cdots+c_0$，比较系数可得Vieta定理：对任意$1\le k\le n$都有

$$\sum_{i_1<\cdots<i_k}a_{i_1}\cdots a_{i_k}=(-1)^kc_{n-k}/c_n.$$

自然的问题是多项式是否有根？在$\mathbb{R}$中答案是未必，比如x^2+1。然而可以证明任何$\mathbb{C}[x]$中的多项式在$\mathbb{C}$中一定有根，即如下的代数基本定理。进一步，归纳可知任何n次复系数多项式都恰好有n个复根（计重数），从而可以按照根进行因式分解。

定理 1.28 (代数基本定理). *设$P \in \mathbb{C}[x]$且$\deg(P) \geq 1$，则存在$a \in \mathbb{C}$使得$P(a) = 0$。*

证明. 证明分为两步，第一步是利用连续性说明$|P(x)|$有最小值点，第二步是说明$|P(x)|$不可能有除了零点之外的最小值点。在学过连续函数的基本性质之后之后，第一步是非常容易的（结合基本的事实$\lim_{|x|\to\infty} |P(x)| = \infty$）。因此这里只关注第二步，通过平移自变量并用$P(x)/P(1)$代替$P(x)$，不妨设0是$|P|$的最小值点且$P(0) = 1$。于是

$$P(x) = a_n x^n + \cdots + a_k x^k + 1,$$

其中$a_n, a_k \neq 0$，且x^k是除了零次项之外最小的非零项。记全体高于k次的项为Q，即$P(x) = Q(x) + a_k x^k + 1$。直观的想法是可以选取适当的辐角让$a_k x^k$是负数，这样$a_k x^k + 1$就严格比1小；只要$x$还充分小，那么$Q(x)$会比$x^k$衰减更快；形如$1 - \varepsilon + o(\varepsilon)$的量在$\varepsilon$充分小时比1严格更小。

具体来说，因为$Q(x)$中每一项都至少是$k+1$次，所以存在$r > 0$，使得当$|x| < r$时$|Q(x)| < |a_k x^k|/2$。设$a_k = e^{i\theta}|a_k|\,(\theta \in \mathbb{R})$，令$t = e^{i(\pi-\theta)/k}\varepsilon$，其中$0 < \varepsilon < r$。注意到辐角的选取保证了

$$a_k t^k = -|a_k t^k|$$

再要求ε充分小，使得上式严格小于1。于是得到

$$|P(t)| \leq |a_k t^k + 1| + |Q(t)| = 1 - |a_k t^k| + |Q(t)| \leq 1 - |a_k t^k|/2.$$

这与$|P(0)| = 1$是最小值矛盾！ □

全体复数中，有理系数多项式的根称为代数数。这一定义可以抽象地推广到一般的域上：如果K是L的子域，即K, L都是域且$K \subset L$（代数数的情形无非是$K = \mathbb{Q}, L = \mathbb{C}$），则可以定义$L$中的$K$-代数元为$K[x]$中多项式的根。注意$K$-代数元与域$L$的选取无关。

下面用初等方法证明代数元构成域，即两个代数元经过加减乘除得到的仍然是代数元。特别的，一个复数是代数数当且仅当实部、虚部都是代数数。例如，对任意$n \in \mathbb{N}$，对n次单位根取实部可知$\cos(2\pi/n)$都是代数元；而若要找到$P \in \mathbb{Q}[x]$使得$P(\cos(2\pi/n)) = 0$则并不平凡，这与今后会介绍的Chebyshev多项式有关。证明需要借助对称多项式：以$x_1, \ldots, x_n$为变元的全体n元多项式的集合记为$K[x_1, \ldots, x_n]$，称$P \in K[x_1, \ldots, x_n]$是对称多项式，如果对任意双射（称为置换）$\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$都有

$$P(x_1, \ldots, x_n) = P(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

题 1.29. *(a)* *对任意$n \in \mathbb{N}$，定义n个初等对称多项式为*

$$\sigma_k = \sigma_k(x_1, \ldots, x_n) := \sum_{i_1 < \cdots < i_k} x_{i_1} \cdots x_{i_k}, \quad 1 \leq k \leq n.$$

证明：对任意对称多项式$P \in K[x_1, \ldots, x_n]$，都存在$Q \in K[\sigma_1, \ldots, \sigma_n]$使得

$$P(x_1, \ldots, x_n) = Q(\sigma_1, \ldots, \sigma_n).$$

(b) 设K是L的子域，如果n次多项式$P \in K[x]$在L中恰有n个根$x_1, \ldots, x_n \in L$（计重数），证明：对任意n元对称多项式R，都有$R(x_1, \ldots, x_n) \in K$。

(c) 设K是L的子域，证明：L中全体K-代数元构成域。

证明. (a) 每个n元单项式$x_1^{e_1} \cdots x_n^{e_n}$由次数$(e_1, \ldots, e_n)$（$n$元有序组）唯一确定，对全体次数引进字典序：$(e_1, \ldots, e_n) \geq (f_1, \ldots, f_n)$当且仅当二者相等或者存在$1 \leq k \leq n$满足$e_i = f_i (1 \leq i < k)$且$e_k > f_k$。换言之，字典序就是从前往后依次比较每个分量。这样就可以谈论次数的高低，并定义n元多项式的次数。

证明的想法是每次构造一个$\sigma_1, \ldots, \sigma_n$的多项式，消去$P$中最高次项，从而有限步之后结束构造。先给出一个例子，考虑$x_1^3 + x_2^3 + x_3^3$，最高次项是x_1^3，因此考虑σ_1^3（注意x_1^3也是σ_1^3中的最高次项），计算得

$$x_1^3 + x_2^3 + x_3^3 = \sigma_1^3 - 3(x_1^2(x_2 + x_3) - x_2^2(x_1 + x_3) + x_3^2(x_1 + x_2)) - 6x_1x_2x_3.$$

剩下的最高次项是$-3x_1^2x_2$，因此考虑$-3\sigma_1^2\sigma_2$，计算得

$$x_1^3 + x_2^3 + x_3^3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3x_1x_2x_3.$$

剩下的项正好是$3\sigma_3$。

一般的证明是对P的次数归纳。根据对称性，最高次项$x_1^{e_1} \cdots x_n^{e_n}$满足$e_1 \geq \cdots \geq e_n$。补充定义$e_{n+1} = 0$，考虑

$$\prod_{k=1}^{n} \sigma_k^{e_k - e_{k+1}},$$

容易证明其中的最高次项也是$x_1^{e_1} \cdots x_n^{e_n}$。由此消去$P$中的最高次项，细节留作练习。

(b) Vieta定理保证了$\sigma_k \in K\ (k = 1, \ldots, n)$，结合(a)即可。

(c) 显然取逆保持代数数，因此只需证明如果α, β都是代数数，则$\alpha + \beta, \alpha\beta$也是代数数。首先证明$\alpha + \beta$是代数数，为此需要承认所谓“分裂域”的存在性，即存在K的扩域M使得K中任何多项式都在M中有根。比如，如果$K = \mathbb{Q}$，则可以取$M = \mathbb{C}$；这也是主要关心的情形，因此不再探讨一般情形下分裂域的存在性。

设$P,Q\in K[x]$使得$P(\alpha)=Q(\beta)=0$（不妨设都是首一多项式），则存在$x_1=\alpha,x_2,\ldots,x_m\in M$和$y_1=\beta,y_2,\ldots,y_n\in M$使得

$$P(x)=\prod_{i=1}^{m}(x-x_i),\quad Q(x)=\prod_{j=1}^{n}(x-y_j).$$

考虑$M[x]$中的多项式

$$R(x):=\prod_{i=1}^{m}\prod_{j=1}^{n}(x-x_i-y_j).$$

可以将R改写为

$$R(x)=\prod_{i=1}^{m}Q(x-x_i).$$

容易看出展开之后每一项x^k的系数都是关于$x_1,\ldots,x_n$的对称多项式，结合(b)可知$R\in K[x]$。注意到$\alpha+\beta=x_1+y_1$是R的根，这就证明了$\alpha+\beta$是K-代数数。

类似可以证明$\alpha\beta$也是K-代数数，具体细节留作练习。 □

注. 借助域的理论可以给出更简单的证明：$\alpha\in L$是K上的代数元当且仅当$K(\alpha)$是K的有限扩张（其中$K(\alpha)$是指域K添加α生成的扩域），又等价于存在K的有限扩张M包含α；如果α,β都是代数数，则$K(\alpha)$是K的有限扩张，$K(\alpha,\beta)$是$K(\alpha)$的有限扩张；于是$K(\alpha,\beta)$是K的有限扩张，其中包含了$\alpha+\beta$以及$\alpha\beta$。类似的方法还可以证明所谓“代数整数”构成环。

提高观点的意义在于得到更一般的结果。利用域扩张的方法容易证明，如果$P(\alpha)=0$且$P\in\mathbb{C}[x]$的系数都是代数数，则α也是代数数。这件事并不具有轻松的初等证明。

下一题的(a)给出Diophantus逼近的经典结果，即d次无理代数数的无理测度至少是d（无理测度是指使得结论成立的指数下确界）；其逆否命题可以用于证明Liouville数$\sum_{n=1}^{\infty}\frac{1}{10^{n!}}$是超越数。更精细的结果是Roth定理，即无理代数数的无理测度都是2（应当指出2确实是下确界而不是最小值），这是获得Fields奖工作。而超越数的无理测度未必巨大，比如π的无理测度是有限的，由此容易证明

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

这是十分经典的钓鱼题，今后也会偶尔给出几道看似能做但实则深刻的题目，以此告诉读者不要一味追求只用初等方法做题而忽视对更深入知识的学习（所谓“思而不学则殆”）。

题 1.30 *. *(a)* 设$P\in\mathbb{Q}[x]$且$\deg(P)=d\,(d\in\mathbb{N})$，无理数$\alpha$是$P(x)$的根。证明：存在$c>0$，使得对任意$m\in\mathbb{Z}$和$n\in\mathbb{N}$都有

$$\left|\alpha-\frac{m}{n}\right|\geq\frac{c}{n^d}.$$

(b) (2017 CMO) 已知正整数q不是立方数，证明：存在正实数c，使得

$$\{\sqrt[3]{q}n\}+\{\sqrt[3]{q^2}n\}\geq n^{-\frac{1}{2}}c$$

对任意正整数n成立。（其中$\{x\}$表示x的小数部分）

证明. (a) 通过乘上一个常数，不妨设$P\in\mathbb{Z}[x]$。关键的观察在于：对任意$m\in\mathbb{Z}$和$n\in\mathbb{N}$，按定义$P(m/n)$是$(m/n)^k$ $(k=0,1,\ldots,d)$的整系数线性组合，在通分之后分子是整数、分母至多是n^d；因此，只要$P(m/n)\neq 0$，就有$|P(m/n)|\geq 1/n^d$。

设P的互异复根分别为$\alpha=\alpha_1,\ldots,\alpha_k$，则

$$P(x)=a\prod_{i=1}^{k}(x-\alpha_i)^{e_i},$$

其中$a\neq 0$是首项系数。记$r:=\min\{|\alpha_i-\alpha_j|\mid \alpha_i\neq\alpha_j\}$，容易证明$P$至少有两个不同的根，所以$r>0$。如果$|m/n-\alpha|<r/2$，则对任意$\alpha_j\neq\alpha$都有$|m/n-\alpha_j|\geq r/2$；此外，$\alpha$不是无理数保证了$m/n\neq\alpha$，因此$m/n$不是$P$的根。另一方面，记$R:=\max\{|\alpha_i-\alpha_j|\mid \alpha_i\neq\alpha_j\}$，则对$j\neq 1$有$|m/n-\alpha_j|<R+r/2$。结合先前的观察可知，

$$\frac{1}{n^d}\leq|P(m/n)|\leq|a|\left|\frac{m}{n}-\alpha\right|(R+r/2)^{d-1}.$$

由于a,r,R,d都是常数，因此存在$c_0>0$使得当$|m/n-\alpha|<r/2$时，

$$\left|\frac{m}{n}-\alpha\right|\geq\frac{c_0}{n^d}.$$

最后，取$c=\min\{c_0,r/2\}$，则可以兼顾$|m/n-\alpha|\geq r/2$的情况。

(b) 关键是注意到

$$a^3+b^3+c^3-3abc=\frac{1}{2}(a+b+c)((a-b)^2+(b-c)^2+(c-a)^2).$$

根据均值不等式，对任意$a,b,c\geq 0$，左边都是非负的；并且可以想象当a,b,c很接近时左边很小；而与(a)类似，如果还能说明左边是整数，那么1就是很好的下界估计。

为此，取$a=q^{2/3}\lfloor nq^{1/3}\rfloor$，$b=q^{1/3}\lfloor nq^{2/3}\rfloor$，$c=nq$，则$q$不是立方数保证了$a,b,c$不全相等，并且显然$a^3,b^3,c^3,abc$都是整数，于是

$$a^3+b^3+c^3-3abc\geq 1.$$

另一方面，$a+b+c\leq 3nq$，且

$$|a-b|=q^{2/3}\{nq^{1/3}\},\quad |a-c|=q^{1/3}\{nq^{1/3}\}.$$

结合$(a-b)^2=((a-c)-(b-c))^2\le 2((a-c)^2+(b-c)^2)$可知

$$(a-b)^2+(b-c)^2+(c-a)^2\le 3((a-b)^2+(a-c)^2)\le C(\{nq^{1/3}\}^2+\{nq^{2/3}\}^2).$$

结合两方面的估计可得，

$$1\le a^3+b^3+c^3-3abc=\frac{1}{2}(a+b+c)((a-b)^2+(b-c)^2+(c-a)^2)\le Cn(\{nq^{1/3}\}^2+\{nq^{2/3}\}^2).$$

因此存在常数$c_0>0$使得$\{nq^{1/3}\}^2+\{nq^{2/3}\}^2\ge n^{-1}c_0$。取$c:=\sqrt{c_0}$，则有

$$\{nq^{1/3}\}+\{nq^{2/3}\}\ge(\{nq^{1/3}\}^2+\{nq^{2/3}\}^2)^{1/2}\ge n^{-1/2}c. \qquad \square$$

下一题讨论$P(x)=\alpha x$小数部分的稠密性；这是Weyl一致分布定理的简单情形。

题 1.31 *. *(a)* 设$\alpha\in\mathbb{R}\setminus\mathbb{Q}$，证明：$\{\{n\alpha\}\mid n\in\mathbb{N}\}$在$[0,1)$中稠密。

(b) 证明：$\inf\{|\sin n|\mid n\in\mathbb{N}\}=0$。

证明. (a) 这是抽屉原理的经典应用。记这个集合为X，容易证明稠密性等价于如下性质：对任意$m\in\mathbb{N}$，将$[0,1)$等分为m个小区间$[(i-1)/m,i/m)\,(i=1,\dots,m)$，则每个小区间中都包含$X$的元素。下面验证这个性质。

事实上，根据抽屉原理，至少有一个小区间包含无穷多个$\{n\alpha\}\,(n\in\mathbb{N})$。任取其中两个相减可知，存在$k\in\mathbb{N}$和$b\in\mathbb{Z}$使得$|k\alpha-b|<1/m$。注意$\alpha\notin\mathbb{Q}$保证了$k\alpha$不是整数。

如果$k\alpha-b>0$，则可设$k\alpha=b+r$，其中$0<r<1/m$。只要$j\in\mathbb{N}$满足$jr<1$，则有$\{jk\alpha\}=jr$。由此可知，X至少包含$r,2r,\dots,Lr$，其中L是使得$Lr<1$的最大整数。因为$r<1/m$，这就保证了每个区间$[(i-1)/m,i/m)$中都存在X的元素。

而如果$k\alpha-b<0$，则可设$k\alpha=b-r$，其中$0<r<1/m$。此时可以证明对任意满足$jr<1$的$j\in\mathbb{N}$都有$\{jk\alpha\}=1-jr$，余下的证明是类似的。

(b) 因为$|\sin x|$以π为周期，所以

$$|\sin n|=|\sin\pi\{n/\pi\}|.$$

由(a)可知（当然，这需要预先知道π是无理数）$\{n/\pi\}$可以任意小，从而结论成立。 $\square$

注. 今后会使用*Fourier*级数证明*Weyl*一致分布定理：对任意$0<a<b<1$，都有

$$\lim_{N\to\infty}\frac{\#\{n\in\{1,\dots,N\}\mid\{n\gamma\}\in(a,b)\}}{N}=b-a.$$

从遍历论的观点来看，这一事实本质上意味着圆周上的无理旋转是唯一遍历的。通过研究高维环面上的唯一遍历映射，*Furstenberg*将这一命题推广到无理多项式的情形，证明了只要$P(x)=a_kx^k+\cdots+a_0$的系数$a_k,\dots,a_0$都是实数并且$a_k,\dots,a_1$中至少有一个是无理数，则$P(n)$的小数部分在$[0,1)$中一致分布。而即使只是说明其稠密性，看似也没有初等方法。

第2讲：数列的极限（上）

……他说，这样的自由组合，故事的变化是无尽的，其实，严格计算起来，依照史特连(Stirling)的方程式应该只有十的二百六十三次方组合而已。

——西西《像我这样的一个读者》

这一讲主要包括数列极限的基本性质、自然常数e、单调有界收敛原理、Cauchy收敛原理等。习题的重点在于用不等式估计的方法求一些趋于0或∞的极限，以及Stolz定理的应用（新教材给出了Stolz定理的证明，而不再是仅仅作为一道习题，实在是一件善事）。无穷级数与数列极限本是一回事，然而不知何故教材将其放入数学分析II，这里简单进行考察。

未加说明时，这一讲考虑的极限都是关于离散变量$n \to \infty$而言；但有时，借助关于连续变量$x \to \infty$或$x \to 0$的极限会让证明更方便。正如数列极限与无穷级数本质相同一样，离散极限与连续极限也没有太大区别，对后一项避而不谈只会徒增烦恼。

例题

设$f(n), g(n)$是非负函数，如果$\frac{f(n)}{g(n)} \to 0$，则称f的阶比g更低，记为$f \ll g$或$f = o(g)$。

例 2.1. *(a)* 证明：$\lim_{n\to\infty} \sqrt[n]{n} = 1$。

(b) 证明：$\lim_{x\to\infty} x^{\frac{1}{x}} = 1$。

(c) 设$p > 0$，$a > 1$，证明：当$n \to \infty$时，

$$\log n \ll n^p \ll a^n \ll n!.$$

证明. (a) 对任意$\varepsilon > 0$，利用二项式展开可知$(1+\varepsilon)^n \geq 1 + n\varepsilon \geq n\varepsilon$，所以

$$1 \leq \sqrt[n]{n} \leq \varepsilon^{-\frac{1}{n}}(1+\varepsilon).$$

当$n \to \infty$时，右边的极限是$1+\varepsilon$，因此存在$N > 0$，使得当$n > N$时

$$1 \leq \sqrt[n]{n} \leq 1 + 2\varepsilon.$$

由ε任意性可知$\lim \sqrt[n]{n} = 1$。

(b) 设$n \leq x < n+1$，则当$x \to \infty$时也有$n \to \infty$。因为当$x > 1$时

$$n^{\frac{1}{2n}} \leq n^{\frac{1}{n+1}} \leq x^{\frac{1}{x}} \leq (n+1)^{\frac{1}{n}} \leq (n+1)^{\frac{2}{n+1}},$$

并且由(a)的结果可知左右两边的极限都是1，根据夹逼原理即证。

(c) 先证$\log n \ll n^p$，取指数只需证明$n^{\frac{1}{n^p}} \to 1$，再p次方即$(n^p)^{\frac{1}{n^p}} \to 1$，而这已经由(b)保证。再证$n^p \ll a^n$，开$p$次方并用$a^{\frac{1}{p}}$代替$a$，可以不妨设$p=1$，再结合二项式定理

$$a^n \geq 1 + n(a-1) + \frac{n(n-1)}{2}(a-1)^2 \geq \frac{n(n-1)}{2}(a-1)^2,$$

以及$a > 1$即证。最后证明$a^n \ll n!$，当$n > N := \lceil 2a \rceil$时，

$$\frac{a^n}{n!} \leq \frac{a^n}{(2a)^{n-N}} = \frac{a^N}{2^{n-N}} \to 0. \qquad \square$$

注. *(a)的一种经典错误论证在于，直接对第一步的不等式两边取极限得到*

$$1 \leq \lim_{n\to\infty} \sqrt[n]{n} \leq 1 + \varepsilon.$$

再结合ε的任意性得到$\lim_{n\to\infty} \sqrt[n]{n} = 1$。问题出在并不事先知道极限存在。除了上面给出的正确论证之外，亦可使用上、下极限的语言：

$$1 \leq \liminf_{n\to\infty} \sqrt[n]{n} \leq \limsup_{n\to\infty} \sqrt[n]{n} \leq 1 + \varepsilon.$$

此时利用ε的任意性可知上、下极限都等于1，从而极限存在且等于1。

设$n \in \mathbb{N}$，定义双阶乘为

$$n!! = \begin{cases} n(n-2)\cdots 4\cdot 2, & 2 \mid n, \\ n(n-2)\cdots 3\cdot 1, & 2 \nmid n. \end{cases}$$

例 2.2. 证明：$\lim_{n\to\infty} \frac{(2n-1)!!}{(2n)!!} = 0$.

证明*1.* 对任意$n \in \mathbb{N}$有$(1-\frac{1}{n})^n = (1+\frac{1}{n-1})^{-n} \leq e^{-1}$，即$1 - \frac{1}{n} \leq e^{-\frac{1}{n}}$，所以

$$\frac{(2n-1)!!}{(2n)!!} = \left(1-\frac{1}{2}\right)\left(1-\frac{1}{4}\right)\cdots\left(1-\frac{1}{2n}\right) \leq e^{-\frac{1}{2}(1+\frac{1}{2}+\cdots+\frac{1}{n})}.$$

熟知$\lim_{n\to\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}\right) = \infty$，所以上式极限是0。 $\square$

证明*2.* 将式子平方，注意到$(2n-1)(2n+1) = (2n)^2 - 1 < (2n)^2$，所以

$$\left(\frac{(2n-1)!!}{(2n)!!}\right)^2 = \frac{1\cdot 3}{2^2}\cdot\frac{3\cdot 5}{4^2}\cdots\frac{(2n-1)(2n+1)}{(2n)^2}\cdot\frac{1}{2n+1} < \frac{1}{2n+1}.$$

即$\frac{(2n-1)!!}{(2n)!!} \leq \frac{1}{\sqrt{2n+1}}$，从而极限是0。 $\square$

证明*3.* 用归纳法可以证明$\frac{(2n-1)!!}{(2n)!!} \leq \frac{1}{\sqrt{3n+1}}$，具体细节留作练习。 $\square$

注. 有如下的*Wallis*公式：

$$\lim_{n\to\infty}\frac{1}{2n+1}\left(\frac{(2n)!!}{(2n-1)!!}\right)^2=\frac{\pi}{2}.$$

常见的证明需要用到定积分：记

$$I_n:=\int_0^{\pi/2}\sin^n x\,dx.$$

分部积分可知$I_{n+2}=\frac{n+1}{n+2}I_n$，于是

$$I_n=\begin{cases}\frac{(n-1)!!}{n!!}, & n\text{是奇数},\\ \frac{(n-1)!!}{n!!}\cdot\frac{\pi}{2}, & n\text{是偶数}.\end{cases}$$

注意到$\frac{n+1}{n+2}I_n=I_{n+2}\le I_{n+1}\le I_n$，所以$I_{n+1}/I_n\to 1$，由此可得*Wallis*公式。

另有基于面积法的初等证明可见*J. Wästlund, An elementary proof of the Wallis product formula for pi, Amer. Math. Monthly* **114** *(2007), no. 10, 914–917*。

熟知Euler常数$\gamma:=\lim\left(\sum_{k=1}^n\frac{1}{k}-\log n\right)\approx 0.57721\cdots$，其收敛性来自于单调有界。

例 2.3. *(a)* 求$\lim\left(\frac{1}{n+1}+\cdots+\frac{1}{2n}\right)$和$\lim\left(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots+\frac{1}{2n-1}-\frac{1}{2n}\right)$。

(b) 将无穷级数$\sum_{n=1}^\infty\frac{(-1)^{n-1}}{n}$重新排序，按顺序交替取$p$个正项、$q$个负项，即

$$1+\frac{1}{3}+\cdots+\frac{1}{2p-1}-\frac{1}{2}-\frac{1}{4}-\cdots-\frac{1}{2q}+\frac{1}{2p+1}+\cdots.$$

证明：新的级数收敛于$\log 2+\frac{1}{2}\log\frac{p}{q}$。

证明. (a) 记$\gamma_n=\sum_{k=1}^n\frac{1}{k}-\log n$，则

$$\frac{1}{n+1}+\cdots+\frac{1}{2n}=\gamma_{2n}+\log(2n)-(\gamma_n+\log(n))=\log 2+\gamma_{2n}-\gamma_n\to\log 2.$$

至于后一个极限，实际上与前一个是相等的：

$$1-\frac{1}{2}+\cdots+\frac{1}{2n-1}-\frac{1}{2n}=(1+\frac{1}{2}+\cdots+\frac{1}{2n})-2\left(\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2n}\right)=\frac{1}{n+1}+\cdots+\frac{1}{2n}.$$

(b) 考虑按照正负号分段之后的级数$a_1-b_1+a_2-b_2+\cdots$，因为

$$a_1+\cdots+a_n=1+\frac{1}{3}+\cdots+\frac{1}{n(2p-1)}=\log(2np)+\gamma+o(1)-\frac{1}{2}(\log(np)-\gamma-o(1)),$$

$$b_1+\cdots+b_n=\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2nq}=\frac{1}{2}(\log(nq)+\gamma+o(1)).$$

两式相减，结合$a_n,b_n\to 0$可知该级数收敛到$a_1-b_1+a_2-b_2+\cdots=\log 2+\frac{1}{2}\log\frac{p}{q}$。 □

注. 与本题形式类似有若干关于交错级数的结论，比如：

$$1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots=\frac{\pi}{4},$$

$$1-\frac{1}{3^3}+\frac{1}{5^3}-\frac{1}{7^3}+\cdots=\frac{\pi^3}{2},$$

$$1-\frac{1}{2}+\frac{1}{4}-\frac{1}{5}+\cdots=\frac{\pi}{3\sqrt{3}}.$$

这些结果的证明往往只能具体处理，而没有放之四海而皆准的方法。事实上，它们与数论中的*Dirichlet L*函数有关，反映了代数整数环的一些性质。

注. *(b)*说明无穷级数在重排之后可能极限改变（当然，也有可能导致极限不存在），稍后将证明更一般的结果：条件收敛的无穷级数通过重排可以收敛到任何数。

练习题

熟知平方差公式$x^2-y^2=(x+y)(x-y)$，由此产生类似于分母有理化的技巧，即

$$x-y=\frac{x^2-y^2}{x+y}.$$

一般的，有n次方差公式

$$x^n-y^n=(x-y)(x^{n-1}+x^{n-2}y+\cdots+y^{n-1}).$$

题 2.4. 求下列极限：

(a) $\lim(\sqrt{n^2+2n}-\sqrt{n^2+n})$。

(b) $\lim(\sqrt[4]{n^4+2n^3}-\sqrt[4]{n^4+n^3})$。

证明. (a) 利用平方差公式得，

$$\sqrt{n^2+2n}-\sqrt{n^2+n}=\frac{n}{\sqrt{n^2+2n}+\sqrt{n^2+n}}=\frac{1}{\sqrt{1+\frac{2}{n}}+\sqrt{1+\frac{1}{n}}}\to\frac{1}{2}.$$

(b) 现在需要用四次方差公式。记$a_n=\sqrt[4]{n^4+2n^3}$和$b_n=\sqrt[4]{n^4+n^3}$，则$\frac{a_n}{n},\frac{b_n}{n}\to 1$，所以

$$a_n-b_n=\frac{n^3}{a_n^3+a_n^2b_n+a_nb_n^2+b_n^3}\to\frac{1}{4}.$$

□

注. 在学了*Taylor*公式之后，这样的题目不需要再进行变形，而是写成

$$\sqrt{n^2+2n}-\sqrt{n^2+n}=n\left(\sqrt{1+\frac{2}{n}}-\sqrt{1+\frac{1}{n}}\right)$$
$$=n\left(1+\frac{1}{n}+o\left(\frac{1}{n}\right)-\left(1+\frac{1}{2n}+o\left(\frac{1}{n}\right)\right)\right)=\frac{1}{2}+o(1)\to\frac{1}{2}.$$

这种方式可以处理一般的次数，而不再需要是有理数。

题 2.5. *(a)* (Cauchy根式判别法)设数列$a_n\geq 0$，证明：若$\lim\sqrt[n]{a_n}<1$，则$\lim a_n=0$。

(b) 求极限$\lim\frac{n(1+\frac{1}{n})^{n^2}}{n^2+n2^n+3^n}$。

证明. (a) 因为$\lim\sqrt[n]{a_n}<1$，所以存在$c\in(0,1)$和$N>0$，使得当$n>N$时$\sqrt[n]{a_n}<c$，即$a_n\leq c^n$。又因为$c^n\to 0$，所以$a_n\to 0$。

(b) 记$a_n=\frac{n(1+\frac{1}{n})^{n^2}}{n^2+n2^n+3^n}$，则$\lim\sqrt[n]{a_n}=\frac{e}{3}<1$（具体细节留作练习），所以$a_n\to 0$。 □

注. 在学了上极限之后，*(a)*中的条件可以改进为$\limsup\sqrt[n]{a_n}<1$。此外，如果$\liminf\sqrt[n]{a_n}>1$，则容易证明$a_n\to\infty$。但如果二者都不成立，就没法判断a_n的极限。

题 2.6. *(a)* (d'Alembert比值判别法)设数列$a_n\geq 0$，证明：若$\lim\frac{a_{n+1}}{a_n}<1$，则$\lim a_n=0$。

(b) 求极限$\lim\frac{\left(\frac{n}{2}\right)^n}{(2n+1)!}$。

证明. (a) 存在$c\in(0,1)$和$N>0$，使得当$n>N$时$\frac{a_{n+1}}{a_n}\leq c$，所以$a_n\leq c^{n-N}a_N\to 0$。

(b) 记$a_n=\frac{\left(\frac{n}{2}\right)^n}{(2n+1)!}$，则$\frac{a_{n+1}}{a_n}=\frac{(n+1)(1+\frac{1}{n})^n}{2(2n+3)}\to\frac{e}{4}<1$，所以$a_n\to 0$。 □

注. 同上一题的情况，*(a)*中的条件可以改进为$\limsup\frac{a_{n+1}}{a_n}<1$。此外，用*(b)*的方法可以得到一个有意思的结论，即如果$0<c<e$，则$n!/(n/c)^n\to 0$；而如果$c>e$，则极限是∞。由此猜想$\left(\frac{n}{e}\right)^n$和$n!$应当相近，但实际上这种临界情形仍有$n!/(n/e)^n\to\infty$。

更精确的结果是如下的*Stirling*公式：

$$\lim_{n\to\infty}\frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{n!}=1.$$

证明思路先验证$n!/(n^{n+1/2}e^n)$单调递减，再用积分估计得到正下界（尚不知这一步是否有足够初等的方法），于是可设其极限为$c>0$；代入*Wallis*公式简单计算可知

$$\frac{\pi}{2}=\lim_{n\to\infty}\frac{1}{2n+1}\left(\frac{(2n)!!}{(2n-1)!!}\right)^2=\lim_{n\to\infty}\frac{2^{4n}(n!)^4}{(2n)!(2n+1)!}=\frac{c^2}{4}.$$

这就说明了$c=\sqrt{2\pi}$，从而*Stirling*公式得证。

下一题需要用到均值不等式：若$a,b\ge 0$，则$a+b\ge 2\sqrt{ab}$。

题 2.7. *(a)* 设$0<a_1<b_1$，递归定义

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

证明：a_n,b_n收敛到公共极限。

(b) 设$0<a_1<b_1$，递归定义

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

证明：$\lim a_n=\lim b_n=\sqrt{a_1b_1}$。

*(c)** 令$a_1=\frac{1}{2\sqrt{3}}$和$b_1=\frac{1}{3}$，递归定义

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

证明：$\lim a_n=\lim b_n=\frac{1}{\pi}$。

证明. (a) 用均值不等式和归纳法容易证明$a_n<a_{n+1}<b_{n+1}<b_n$，根据单调有界收敛原理可知a_n,b_n收敛。此外，因为

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

所以$b_n-a_n\to 0$，即$\lim a_n=\lim b_n$。[另一种办法是对$b_{n+1}=\frac{a_n+b_n}{2}$两边取极限。]

(b) 同(a)的方法可证a_n,b_n收敛到公共极限，再注意到$a_nb_n=a_1b_1$即可，具体留作练习。

(c) 可以归纳证明

$$a_n=\frac{1}{3\cdot 2^n}\cot\left(\frac{\pi}{3\cdot 2^n}\right)\quad 且\quad b_n=\frac{1}{3\cdot 2^n}\csc\left(\frac{\pi}{3\cdot 2^n}\right).$$

然后利用当$x\to 0$时$\tan x,\sin x\sim x$即可。 □

注. *(a)*中的极限称为算术几何平均*(AGM, 德语arithemetrisch-geometrisches Mittel)*，最早由*Gauss*研究并发现其深刻性质（顺带一提，*Gauss*使用的记号是*agM*，可能是因为德语里名词首字母要大写而形容词不需要）。他证明了*AGM*与椭圆积分的联系：

$$\frac{1}{\mathrm{AGM}(a,b)}=\frac{2}{\pi}\int_0^{\frac{\pi}{2}}\frac{d\theta}{\sqrt{a^2\sin^2\theta+b^2\cos^2\theta}}.$$

与*(b)*的平凡结果作比较，可见*Gauss*的眼光独到。*AGM*以指数速率收敛，可以作为一种有效的椭圆积分算法。*Gauss*还研究了复数情形的*AGM*，新的现象来自于$\sqrt{a_n b_n}$有两种取法。*Gauss*发现这一问题与同余子群、进而与模形式的理论有密切联系。相关内容可见*D. A. Cox. The Arithmetic-Geometric Mean of Gauss. In: Pi: A Source Book. Springer, 2004*。

题 2.8. *(a)* 设$f\colon \mathbb{R}\to\mathbb{R}$满足$|f(x)-f(y)|\le|x-y|^2\ (\forall x,y\in\mathbb{R})$，证明：$f$是常值函数。

*(b)** 设$f\colon[0,1]\to\mathbb{R}$满足$|f(x)-f(y)|\le|x-y|^\alpha\ (\forall x,y\in[0,1])$，其中常数$\alpha\in(\frac{1}{2},1]$。证明：

$$\lim_{n\to\infty}\sum_{k=1}^{n}\left(f\left(\frac{k}{n}\right)-f\left(\frac{k-1}{n}\right)\right)f\left(\frac{k}{n}\right)=\frac{1}{2}(f(1)^2-f(0)^2).$$

证明. (a) 对任意$a<b$，取n等分点$a=x_0<x_1<\cdots<x_n=b$，即$x_i:=a+\frac{i}{n}(b-a)$，则

$$|f(b)-f(a)|\le\sum_{i=1}^{n}|f(x_i)-f(x_{i-1})|\le\sum_{i=1}^{n}|x_i-x_{i-1}|^2=n\cdot\frac{(b-a)^2}{n^2}\to 0,$$

即$f(b)=f(a)$。由a,b任意性可知f是常数。

(b) 注意到$\alpha>\frac{1}{2}$，于是

$$\left|\sum_{k=1}^{n}\left(f\left(\frac{k}{n}\right)-f\left(\frac{k-1}{n}\right)\right)^2\right|\le n\cdot\frac{1}{n^{2\alpha}}\to 0.$$

从而上式左边的极限是0，故

$$\begin{aligned}
&\sum_{k=1}^{n}\left(f\left(\frac{k}{n}\right)-f\left(\frac{k-1}{n}\right)\right)f\left(\frac{k}{n}\right)\\
=&\frac{1}{2}\sum_{k=1}^{n}\left(f\left(\frac{k}{n}\right)+f\left(\frac{k-1}{n}\right)\right)^2+\frac{1}{2}\sum_{k=1}^{n}\left(f\left(\frac{k}{n}\right)^2-f\left(\frac{k-1}{n}\right)^2\right)\\
\to&0+\frac{1}{2}(f(1)^2-f(0)^2)=\frac{1}{2}(f(1)^2-f(0)^2).
\end{aligned}$$

□

注. *(a)*更常见的做法是由条件得出f可导且导数为0，所以是常数。但这里的方法说明直接的估计往往也是有效的。条件右边$|x-y|^2$可以改进为$|x-y|^{1+\varepsilon}$，其中ε是任意正实数。所谓*Hölder*连续函数是指存在$\alpha\in(0,1]$和常数C使得$|f(x)-f(y)|\le C|x-y|^\alpha$，特别的，当$\alpha=1$时称为*Lipschitz*连续。本题说明了为何不考虑$\alpha>1$阶的*Hölder*连续函数。

注. 在*(b)*中如果假设f光滑，则直观上这个式子近似为*Riemann*和

$$\sum_{k=1}^{n}f(x_k)f'(x_k)\Delta x_k,$$

其中$x_k = \frac{k}{n}\,(k=0,1,\ldots,n)$。于是应当收敛到

$$\int_0^1 f(x)f'(x)\,dx = \frac{1}{2}(f(1)^2 - f(0)^2).$$

本题说明为了让*Riemann–Stieljis*积分$\int_0^1 f\,df$收敛，只需要f有$> \frac{1}{2}$阶的*Hölder*连续性。一般的，要让$\int_0^1 f\,dg$有定义，只需要f和g的*Hölder*指标相加> 1。

题 2.9 *. 证明：$\lim_{n\to\infty} \frac{n^{1/1}+n^{1/2}+\cdots+n^{1/n}}{n} = 2$。

证明*1.* 分子上的n显然贡献了1，只需证明$\frac{1}{n}\sum_{k=2}^n n^{1/k} \to 1$。为了估计$n^{1/k}$的大小，最合理（尽管看似复杂）的方法是用Taylor公式，将$n^{1/k}$写成$e^{(\log n)/k} = \sum_{m=0}^\infty \frac{(\log n/k)^m}{m!}$可知

$$\frac{1}{n}\sum_{k=2}^n n^{1/k} = \frac{n-1}{n} + \frac{1}{n}\sum_{k=2}^n\sum_{m=1}^\infty \frac{((\log n)/k)^m}{m!}.$$

因此只需证明后一项求和的贡献是0。注意到对每个固定的m都有

$$\frac{1}{n}\sum_{k=2}^n \frac{((\log n)/k)^m}{m!} \le \frac{1}{n}\sum_{k=2}^n \frac{(\log n)^m}{k} \le \frac{(\log n)^m(1+\log n)}{n} \to 0.$$

对任意$\varepsilon > 0$，存在$M \in \mathbb{N}$使得$\sum_{k=2}^\infty 1/k^M < \varepsilon$，所以（显然有限求和与无穷级数可交换）

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n}\sum_{k=2}^n\sum_{m=1}^\infty \frac{((\log n)/k)^m}{m!} &= \limsup_{n\to\infty} \frac{1}{n}\sum_{m=M}^\infty\sum_{k=2}^n \frac{((\log n)/k)^m}{m!}\\ &\le \frac{\varepsilon}{n}\sum_{m=M}^\infty \frac{(\log n)^m}{m!} \le \frac{\varepsilon}{n}\cdot e^{\log n} = \varepsilon.\end{aligned}$$

由ε任意性可知上极限是0，从而结论成立。 □

证明*2.* 同样只需证明$\frac{1}{n}\sum_{k=2}^n n^{1/k} \to 1$。根据均值不等式可知（注意这里$k \ge 2$的意义）

$$n^{1/k} = \sqrt[k]{1\times\cdots\times 1\times\sqrt{n}\times\sqrt{n}} \le \frac{k-2}{k} + \frac{2\sqrt{n}}{k} \le 1 + \frac{2\sqrt{n}}{k}.$$

求和并利用$\sum_{k=2}^n 1/k \le \log n$便得到

$$\frac{n-1}{n} \le \frac{1}{n}\sum_{k=2}^n n^{1/k} \le \frac{n-1}{n} + \frac{2\log n}{\sqrt{n}}.$$

结合夹逼定理即证。 □

题 2.10 *. 将给定的n个实数$a_1,\ldots,a_n$排成一圈，每一步操作将所有位置上的数都替换成它和它下一个位置的数的算术平均，即将a_k换成$(a_k+a_{k+1})/2$（约定$a_{n+1} = a_1$）。证明：随着操作

无限进行，每个位置上的数都会收敛到$(a_1+\cdots+a_n)/n$。

证明*1.* 记第m步操作之后的n个数为$a_1^{(m)},\ldots,a_n^{(m)}$，定义极差$V_m$以及平方和$D_m$为

$$V_m=\max a_k^{(m)}-\min a_k^{(m)},\quad D_m=(a_1^{(m)})^2+\cdots+(a_n^{(m)})^2.$$

根据定义有

$$D_{m+1}=\sum_{k=1}^{n}\left(\frac{a_k^{(m)}+a_{k+1}^{(m)}}{2}\right)^2=D_m-\sum_{k=1}^{n}\left(\frac{a_k^{(m)}-a_{k+1}^{(m)}}{2}\right)^2.$$

因此D_m非负单减，从而极限存在。注意到存在$1\le k\le n$使得$|a_k^{(m)}-a_{k+1}^{(m)}|\ge V_m/n$，所以

$$D_{m+1}\le D_m-\left(\frac{V_m}{2n}\right)^2.$$

由D_m收敛可知$\lim_{m\to\infty}V_m=0$。特别的，对任意$i\neq j$，都有$\lim_{m\to\infty}(a_i^{(m)}-a_j^{(m)})=0$。结合$a_1^{(m)}+\cdots+a_m^{(m)}$始终等于定值可知，每个$a_k^{(m)}$都会收敛到平均值$(a_1+\cdots+a_n)/n$。 □

证明*2.* 本质上是线性代数问题，用向量$A_m=(a_1^{(m)},\ldots,a_n^{(m)})^T$表示第$m$步操作之后的状态，则有$A_{m+1}=BA_m$，其中$B$是$n$阶方阵（空位表示0）

$$B=\frac{1}{2}\begin{pmatrix}1&1&&\\&\ddots&\ddots&\\&&1&1\\1&&&1\end{pmatrix}$$

核心是研究B^m的极限行为。容易计算B的特征多项式为$(x-1/2)^n-1/2^n$，因此B有n个互异特征值$\lambda_k=(1+\omega^k)/2\,(1\le k\le n)$，其中$\omega=e^{2\pi i/n}$是$n$次单位根。记相应的特征向量为$v_1,\ldots,v_n$，则这些向量给出$\mathbb{R}^n$的一组基。注意到$\lambda_1=1$且$|\lambda_k|<1\,(2\le k\le n)$，所以

$$\lim_{m\to\infty}B^m v_k=\begin{cases}v_1, & k=1,\\ 0, & 2\le k\le n.\end{cases}$$

设$A_0=(a_1,\ldots,a_n)$在这组基下的分解为$A_0=\sum_{k=1}^n e_k v_k$，其中$e_k\in\mathbb{R}$，则有

$$A_m=B^mA_0\to e_1v_1.$$

容易计算出$\lambda_1=1$对应的特征向量，可以取$v_1=(1,\ldots,1)$，于是每个$a_k^{(m)}$都收敛到e_1。

最后，为了求出e_1，定义线性泛函T将向量$v=(x_1,\ldots,x_n)^T$映为各个分量之和

$$Tv=x_1+\cdots+x_n.$$

容易验证对任意$v \in \mathbb{R}^n$都有$TBv = Tv$。于是对$2 \le k \le n$都有

$$\lambda_k T v_k = T(\lambda_k v_k) = TBv_k = Tv_k.$$

结合$\lambda_k \neq 1\,(2 \le k \le n)$可知，$Tv_2 = \cdots = Tv_n = 0$。于是

$$a_1 + \cdots a_n = TA_0 = \sum_{k=1}^{n} e_k T v_k = e_1 T v_1 = n e_1.$$

故$\lim_{m\to\infty} a_k^{(m)} = e_1 = (a_1 + \cdots + a_n)/n$。 □

证明3. 证明2蕴含了B^m的每个位置都收敛到$1/n$，下面给出代数证明。设$B = (I+D)/2$，其中I是n阶单位阵，D的i行j列元素是1当且仅当$i + 1 \equiv j \pmod n$、其他位置是0。直接计算可知D^m的i行j列元素是1当且仅当$i + m = j \pmod n$、其他位置是0，特别的有$D^n = I$。于是$I = D^0, D^1, \ldots, D^{n-1}$分别对应了$n$个主对角线方向，要说明的是对每个$0 \le k \le n-1$，对$B^m = (I+D)^m/2^m$进行二项式展开之后，等于$D^k$的项的系数之和收敛到$1/n$，即

$$\lim_{m\to\infty} \frac{1}{2^m} \sum_{0 \le l \le m,\ l \equiv k \pmod n} C_m^l = \frac{1}{n}.$$

为此，考虑n次单位根$\omega = e^{2\pi i/n}$，熟知对任意$l \in \mathbb{Z}$都有

$$\frac{1}{n}\sum_{t=0}^{n-1} \omega^{lt} = \begin{cases} 1, & n \mid l, \\ 0, & n \nmid l. \end{cases}$$

取定$0 \le k \le n$，令$f_m(x) = x^{-k}(1+x)^m$，则有

$$\begin{aligned} \frac{1}{n}\sum_{t=0}^{n-1} f_m(\omega^t) &= \frac{1}{n}\sum_{t=0}^{n-1}\sum_{l=0}^{m} C_m^l \omega^{t(l-k)} \\ &= \frac{1}{n}\sum_{l=0}^{m} C_m^l \sum_{t=0}^{n-1} \omega^{t(l-k)} = \frac{1}{2^m} \sum_{0 \le l \le m,\ l \equiv k \pmod n} C_m^l. \end{aligned}$$

最后，注意到

$$\lim_{m\to\infty} \frac{1}{2^m} f_m(\omega^l) = \lim_{m\to\infty} \big((1+\omega^l)/2\big)^m = \begin{cases} 1, & l = 0, \\ 0, & 1 \le l \le n-1. \end{cases}$$

所以

$$\lim_{m\to\infty} \frac{1}{2^m} \sum_{0 \le l \le m,\ l \equiv k \pmod n} C_m^l = \frac{1}{n} \lim_{m\to\infty} \frac{1}{2^m} \sum_{l=0}^{n-1} f_m(\omega^l) = \frac{1}{n}.$$ □

Stolz定理

Stolz定理可能对求数列极限帮助最大的定理，这一小节给出其证明以及若干习题。应当注意的是，有时并不直接使用这一定理本身，而是用证明的思想和方法。为了写起来简单，回忆上一讲提到过的简单性质：如果$b, d > 0$且$\frac{a}{b} \leq \frac{c}{d}$，则$\frac{a}{b} \leq \frac{a+c}{b+d} \leq \frac{c}{d}$。

定理 2.11 (Stolz定理). *(a) 设数列c_n严格单增且$c_n \to \infty$，如果$\frac{b_n - b_{n-1}}{c_n - c_{n-1}} \to a$，其中$a \in [-\infty, \infty]$（即$a$可以是实数、$-\infty$或者$\infty$），证明：$\frac{b_n}{c_n} \to a$。*

(b) 设数列c_n严格单减且$b_n, c_n \to 0$，如果$\frac{b_n - b_{n-1}}{c_n - c_{n-1}} \to a$，其中$a \in [-\infty, \infty]$，证明：$\frac{b_n}{c_n} \to a$。

证明. (a) 先处理$a \in \mathbb{R}$的情况。根据定义，对任意$\varepsilon > 0$，存在$N > 0$，使得当$n \geq N$时

$$a - \varepsilon \leq \frac{b_{n+1} - b_n}{c_{n+1} - c_n} \leq a + \varepsilon.$$

因为$c_{n+1} - c_n > 0$，所以对$n > N$有

$$a - \varepsilon \leq \frac{\sum_{k=N}^{n-1}(b_{k+1} - b_k)}{\sum_{k=N}^{n-1}(c_{k+1} - c_k)} = \frac{b_n - b_N}{c_n - c_N} \leq a + \varepsilon.$$

于是

$$\begin{aligned}\left|\frac{b_n}{c_n} - a\right| &= \left|\frac{c_n - c_N}{c_n}\left(\frac{b_n - b_N}{c_n - c_N} - a\right) + \frac{b_N}{c_n} + \frac{c_N}{c_n}a\right| \\ &\leq \left|\frac{c_n - c_N}{c_n}\right| \cdot \left|\frac{b_n - b_N}{c_n - c_N} - a\right| + \left|\frac{b_N}{c_n}\right| + \left|\frac{c_N}{c_n}a\right| \leq \left|\frac{c_n - c_N}{c_n}\right|\varepsilon + \left|\frac{b_N}{c_n}\right| + \left|\frac{c_N}{c_n}a\right|.\end{aligned}$$

因为$c_n \to \infty$，所以当$n \to \infty$时右边的极限是ε。由ε任意性可知$\frac{b_n}{c_n} \to a$。

最后，如果$a = \infty$，则同上面的方法可得，对任意$A > 0$，存在$N > 0$使得当$n > N$时

$$\frac{b_n - b_N}{c_n - c_N} > A.$$

随后用类似的估计可知$\frac{b_n}{c_n} \to \infty$。具体细节以及$a = -\infty$的情况留作练习。

(b) 先处理$a \in \mathbb{R}$的情况。同(a)可知，对任意$\varepsilon > 0$，存在$N > 0$，使得当$n > m > N$时，

$$a - \varepsilon \leq \frac{b_n - b_m}{c_n - c_m} \leq a + \varepsilon.$$

令$n \to \infty$，利用$b_n, c_n \to 0$可知

$$a - \varepsilon \leq \frac{b_m}{c_m} \leq a + \varepsilon.$$

上式对任意$m > N$成立，结合ε任意性可知，当$m \to \infty$时$\frac{b_m}{c_m} \to a$。

最后，$a = \infty$或$-\infty$的情况是类似的，具体细节留作练习。 □

注. *(a)*中没有对b_n作进一步要求，而*(b)*中假设$b_n \to 0$是必要的，比如b_n是常数就是反例。此外，*Stolz*定理的逆命题通常是不成立的，即$\frac{b_n}{c_n} \to a$并不能推出$\frac{b_{n+1}-b_n}{c_{n+1}-c_n} \to a$。问题出在$b_n$正负交错的情况下，比如在*(a)*的条件下有反例$b_n = (-1)^n$，$c_n = n$，在*(b)*的条件下也可以构造类似的反例。最后，要求c_n单调也是必要的，反例的构造留作练习。

题 2.12. *(a)* 设$\alpha > -1$，求极限$\lim_{n\to\infty} \frac{1+2^\alpha+\cdots+n^\alpha}{n^{\alpha+1}}$。

(b) 设$\alpha > 0$，求极限$\lim_{n\to\infty} \left(n \left(\frac{1^\alpha+2^\alpha+\cdots+n^\alpha}{n^{\alpha+1}} - \frac{1}{\alpha+1} \right) \right)$。

*(c)** 设$0 < s < 1$，证明：当$n \to \infty$时，$\sum_{k=1}^n \frac{1}{k^s} - \frac{n^{1-s}}{1-s}$收敛。

证明. (a) 直接用Stolz定理（初学时记得验证条件），得

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

这里最后一步用到了$\lim_{x\to 0} \frac{(1+x)^\beta-1}{x} = \beta\,(\forall \beta \in \mathbb{R})$。

(b) 先通分，分母是$(\alpha+1)n^\alpha$，当$\alpha > 0$时可以用Stolz定理，得

$$\begin{aligned}
&\lim_{n\to\infty} n \left(\frac{1^\alpha+\cdots+n^\alpha}{n^{\alpha+1}} - \frac{1}{\alpha+1} \right) \\
&= \lim_{n\to\infty} \frac{(\alpha+1)(n+1)^\alpha - (n+1)^{\alpha+1} + n^{\alpha+1}}{(\alpha+1)((n+1)^\alpha - n^\alpha)} \\
&= \lim_{n\to\infty} \frac{(\alpha+1)\left(1+\frac{1}{n}\right)^\alpha - n\left(\left(1+\frac{1}{n}\right)^{\alpha+1}-1\right)}{(\alpha+1)\left(\left(1+\frac{1}{n}\right)^\alpha - 1\right)} \\
&= \lim_{n\to\infty} \frac{(\alpha+1)\left(1+\frac{\alpha}{n}+o\left(\frac{1}{n}\right)\right) - n\left(\frac{\alpha+1}{n}+\frac{\alpha(\alpha+1)}{2n^2}+o\left(\frac{1}{n^2}\right)\right)}{(\alpha+1)\left(\frac{\alpha}{n}+o\left(\frac{1}{n}\right)\right)} \\
&= \lim_{n\to\infty} \frac{\frac{\alpha(\alpha+1)}{2n}+o\left(\frac{1}{n}\right)}{\frac{\alpha(\alpha+1)}{n}+o\left(\frac{1}{n}\right)} = \frac{1}{2}.
\end{aligned}$$

这里用到了$x \to 0$时的Taylor展开式$(1+x)^\beta = 1+\beta x+\frac{\beta(\beta-1)}{2}x^2+o(x^2)\,(\forall \beta \in \mathbb{R})$。

(c) 证明的关键在于，当$n \geq 1$时

$$\frac{(n+1)^{1-s}-n^{1-s}}{1-s} \leq n^{-s} \leq \frac{n^{1-s}-(n-1)^{1-s}}{1-s}.$$

这一不等式的标准证明需要借助定积分，即$(n+1)^{-s} \leq \int_n^{n+1} x^{-s} \leq n^{-s}$。有了这一结果，结合单调有界收敛原理容易证明极限存在。 □

注. *(a)*的条件$\alpha > -1$是最优的，因为当$\alpha \le -1$时分子发散到∞而分母收敛。于是$1^\alpha + \cdots + n^\alpha$的渐近展开首阶项是$n^{\alpha+1}/(\alpha+1)$。直观上，下一阶项应该形如$cn^\alpha$，于是*(b)*表明当$\alpha > 0$是确实如此，并且常数$c = 1/2$不依赖$\alpha$，即

$$\sum_{k=1}^{n} k^\alpha = \frac{n^{\alpha+1}}{\alpha+1} + \frac{n^\alpha}{2} + o(n^\alpha).$$

而如果$\alpha \in (-1, 0)$，则*(c)*表明在$n^{\alpha+1}$与n^α之间还会出现常数项。事实上，*(c)*中的极限等于$\zeta(s)$，其中ζ是$Riemann\ \zeta$函数。于是当$\alpha \in (-1, 0)$时，正确的渐进展开为

$$\sum_{k=1}^{n} k^\alpha = \frac{n^{\alpha+1}}{\alpha+1} + \zeta(-\alpha) + o(1).$$

当$s > 1$时，$Riemann\ \zeta$函数定义为（稍后会证明收敛性）

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

但在$s \to 1+$时$\zeta(s)$会遇到发散到∞。注意到

$$\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^\infty \frac{1}{x^s}\,dx = \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx.$$

后一项积分大约是n^{-s-1}，当$s > 0$时求和绝对收敛。因此$\zeta(s) - 1/(s-1)$对任意$s > 0$良定义，由此可以将ζ延拓到$(0,1)$上。这样的延拓是$\zeta(s)$在$(0,1)$上唯一自然的定义，复变函数的理论称之为解析延拓。实际上，ζ可以延拓为整个复平面上的亚纯函数，并且只有1这一个极点。

回到*(c)*所求的极限，根据$s \in (0,1)$时$\zeta(s)$的定义可知

$$\zeta(s) = \frac{1}{s-1} + \lim_{n\to\infty} \sum_{k=1}^{n} \int_k^{k+1} \left(\frac{1}{k^s} - \frac{1}{x^s}\right) dx = \lim_{n\to\infty} \sum_{k=1}^{n} \left(\frac{1}{k^s} - \frac{(n+1)^{1-s}}{1-s}\right).$$

由于$(n+1)^{1-s} - n^{1-s} \to 0$，所以*(c)*的结果确实等于$\zeta(s)$。

下一题与Stolz定理无关，仅仅是为了讨论当$k \in \mathbb{N}$时$1^k + \cdots + n^k$的表达式。特别的，不借助Stolz定理就可以知道，对任意正整数k都有$\frac{1^k+\cdots+n^k}{n^{k+1}} \to \frac{1}{k+1}$。

题 2.13. *(a)* 证明：对任意$n \in \mathbb{N}$，有

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2},$$

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6},$$

$$1^3+2^3+\cdots+n^3=(1+2+\cdots+n)^2=\frac{n^2(n+1)^2}{4}.$$

(b) 证明：对于固定的$k\in\mathbb{N}_0$和任意$n\in\mathbb{N}$，$S_k(n):=1^k+2^k+\cdots+n^k$是关于$n$的$k+1$次多项式，且首项系数为$\frac{1}{k+1}$。

证明. (a) 对n归纳即可。至于这些等式是如何被发现的，可以参考(b)的方法。

(b) 对k归纳，$k=0$时显然有$S_0(n)=n$。当$k\geq 1$时，注意到

$$(m+1)^{k+1}-m^{k+1}=\sum_{i=0}^{k}C_{k+1}^{i}m^i.$$

对m从1到n累加得

$$(n+1)^{k+1}-1=\sum_{i=0}^{k}C_{k+1}^{i}S_i(n).$$

由此可知

$$S_k(n)=\frac{1}{k+1}\left((n+1)^{k+1}-1-\sum_{i=0}^{k-1}C_{k+1}^{i}S_i(n)\right).$$

由归纳假设可知$S_k(n)$是$k+1$次多项式，且首项仅由$(n+1)^{k+1}$贡献，故系数是$\frac{1}{k+1}$。 □

注. 根据上一题的*(b)*可知，$k\geq 1$时$S_k(n)$的第二项（即x^n项）系数为$\frac{1}{2}$。现在可以不使用*Stolz*定理而直接看出：$S_k(n)$中的x^n项有两处来源，即$(n+1)^{k+1}$的第二项和$S_{k-1}(n)$的最高次项，简单计算即可。今后会介绍*Bernoulli*数并给出$S_k(n)$的一般项系数的表达式。

题 2.14. *(a)* 设$a_n\to a$且$b_n\to b(a,b\in\mathbb{R})$，证明：

$$\lim_{n\to\infty}\frac{a_1b_n+a_2b_{n-1}+\cdots+a_nb_1}{n}=ab.$$

(b) 设数列p_n满足$p_n>0$，且$\frac{p_n}{p_1+\cdots+p_n}\to 0$。证明：如果数列$a_n\to a\,(a\in\mathbb{R})$，则

$$\lim_{n\to\infty}\frac{p_1a_n+\cdots+p_na_1}{p_1+\cdots+p_n}=a.$$

(c) 设$|q|<1$，如果数列$x_n\to x\,(x\in\mathbb{R})$，证明：

$$\lim_{n\to\infty}(x_n+qx_{n-1}+\cdots+q^nx_0)=\frac{x}{1-q}.$$

证明. (a) 用a_n-a替换a_n，因为Stolz定理保证了$\frac{b_1+\cdots+b_n}{n}\to b$，所以可以不妨设$a=0$；同理，再用$b_n-b$替换$b_n$，可以不妨设$b=0$。由收敛知$|a_n|,|b_n|$有界，记其公共上界为$M$。

对任意$\varepsilon>0$，存在$N>0$，使得当$n>N$时$|a_n|,|b_n|<\varepsilon$。当$n>2N$时，因为a_kb_{n+1-k}中至少有一个的下标大于N，所以$|a_kb_{n+1-k}|\le M\varepsilon$，故

$$\left|\frac{a_1b_n+a_2b_{n-1}+\cdots+a_nb_1}{n}\right|\le M\varepsilon.$$

由ε的任意性即证。

(b) 用a_n-a替换a_n，可以不妨设$a=0$。收敛数列一定有界，记$M=\sup|a_n|<\infty$。对任意$\varepsilon>0$，存在$N>0$使得当$n>N$时$|a_n|<\varepsilon$，所以

$$\begin{aligned}\left|\frac{p_1a_n+\cdots+p_na_1}{p_1+\cdots+p_n}\right|&\le\varepsilon\cdot\frac{p_1+\cdots+p_{n-N}}{p_1+\cdots+p_n}+M\cdot\frac{p_{n-N+1}+\cdots+p_n}{p_1+\cdots+p_n}\\&\le\varepsilon+M\sum_{k=0}^{N-1}\frac{p_{n-k}}{p_1+\cdots+p_n}.\end{aligned}$$

根据条件，对任意固定的$k\in\mathbb{N}_0$，当$n\to\infty$时都有

$$\frac{p_{n-k}}{p_1+\cdots+p_n}\le\frac{p_{n-k}}{p_1+\cdots+p_{n-k}}\to0.$$

所以当$n\to\infty$时上式右边的第二项趋于0，结合ε的任意性即证。

(c) 仿照(b)的证明或者直接运用(b)的结论即可（当$q>0$时可以取$p_n=q^{n-1}$，而当$q<0$时则需要按奇偶性分开处理），具体细节留作练习。 □

下面给出Cesàro求和以及反过来的Tauber型定理（即(b)(c)两个命题）。一般的，Cesàro可求和比真正的可求和严格更弱。事实上，如果取$a_n=(-1)^n$，则$(s_n)_{n=0}^\infty=(1,0,1,0,\dots)$发散；而$\sigma_n=\frac{\lceil\frac{n}{2}\rceil}{n}\to\frac12$，因此在Cesàro求和的意义下

$$1-1+1-1+\cdots=\frac12.$$

题 2.15 (Cesàro求和). *设$a_n\in\mathbb{R}\,(n\in\mathbb{N}_0)$，记$s_n:=a_0+\cdots+a_n$以及$\sigma_n:=\frac{s_0+\cdots+s_n}{n+1}$，证明：*

(a) 若$s_n\to s\in[-\infty,\infty]$，则$\sigma_n\to s$。

(b) 若$\sigma_n\to\sigma\in[-\infty,\infty]$，且$na_n\to0$，则$s_n\to s$。

(c) 若$\sigma_n\to\sigma\in[-\infty,\infty]$，且存在常数$M>0$使得$|na_n|\le M$，则$s_n\to s$。*

证明. (a) 这是Stolz定理的推论。

(b) 虽然这是(c)的推论，但因为证明会简单很多所以单独列出。注意到

$$s_n-\sigma_n=\frac{a_1+2a_2+\cdots+na_n}{n+1}.$$

Stolz定理保证了右边的极限是0，从而结论成立。

(c) [3] 注意到对任意$m > n \geq 0$，有

$$(m+1)\sigma_m - (n+1)\sigma_n - (m-n)s_n = \sum_{i=n+1}^{m} (m+1-i)a_i.$$

于是

$$s_n = \frac{m+1}{m-n}\sigma_m - \frac{n+1}{m-n}\sigma_n - \frac{1}{m-n}\sum_{i=n+1}^{m}(m+1-i)a_i.$$

由此得到

$$s_n - \sigma_n = \frac{m+1}{m-n}(\sigma_m - \sigma_n) - \frac{m+1}{m-n}\sum_{i=n+1}^{m}\left(\frac{1}{i} - \frac{1}{m+1}\right) i a_i.$$

利用$\frac{1}{i} \leq \log\frac{i}{i-1}$，得

$$0 \leq \sum_{i=n+1}^{m}\left(\frac{1}{i} - \frac{1}{m+1}\right) \leq \log\frac{m}{n} - \frac{m-n}{m+1}.$$

所以

$$|s_n - \sigma_n| \leq \frac{m+1}{m-n}|\sigma_m - \sigma_n| + M\left(\frac{m+1}{m-n}\log\frac{m}{n} - 1\right).$$

对任意$\varepsilon > 0$，取$m = m_n := \lfloor (1+\varepsilon)n \rfloor$，则当$n \to \infty$时$m_n > n$，且

$$\frac{m_n}{n} \to 1 + \varepsilon,$$

$$\frac{m_n + 1}{m_n - n} \to \frac{1+\varepsilon+0}{1+\varepsilon-1} = 1 + \frac{1}{\varepsilon},$$

$$|\sigma_{m_n} - \sigma_n| \to 0.$$

因此，取上极限得到

$$\limsup_{n\to\infty} |s_n - \sigma_n| \leq M\left(\left(1 + \frac{1}{\varepsilon}\right)\log(1+\varepsilon) - 1\right).$$

注意到当$\varepsilon \to 0$时右边的极限是0，所以$s_n - \sigma_n \to 0$，从而结论成立。 □

注. *从(b)的证明可以看出，条件可以减弱为$\frac{a_1+2a_2+\cdots+na_n}{n} \to 0$，但(c)的结论依然严格更强。*

通过指对数映射转换，可以将Cesàro求和转换为乘积版本：如果$a_n > 0$，且$\frac{a_{n+1}}{a_n} \to a$，则$\sqrt[n]{a_n} \to a$。在此举一个经典例子，并给出一种直接证明，其好处是能够给出误差估计。

[3]这个复杂而精妙的证明来自W. Rudin, *Principles of mathematical analysis*, third edition, McGraw-Hill, 1976。包括本题在内的若干Tauber型定理最早都是由Hardy–Littlewood给出。

题 2.16. *(a)* 证明：$\lim \frac{n}{\sqrt[n]{n!}} = e$。

(b) 考虑等差数列$a_n = a + nd$，其中$a, d \in \mathbb{R}$且$d \neq 0$。记

$$A_n = (a_1 + \cdots + a_n)/n, \quad G_n = \sqrt[n]{a_1 \cdots a_n}.$$

分别是前n项的算术平均和几何平均，求$\lim_{n\to\infty} A_n/G_n$。

证明. (a) 记$a_n = \frac{n^n}{n!}$，则$\frac{a_{n+1}}{a_n} = \left(1 + \frac{1}{n}\right)^n \to e$，所以$\sqrt[n]{a_n} \to e$。

另一种方法常见于教材，却不容易想到。因为$\left(1+\frac{1}{n}\right)^n \le e \le \left(1+\frac{1}{n}\right)^{n+1}$，所以

$$\frac{(n+1)^n}{n!} = \prod_{k=1}^{n}\left(1+\frac{1}{k}\right)^k \le e^n \le \prod_{k=1}^{n}\left(1+\frac{1}{k}\right)^{k+1} = \frac{(n+1)^{n+1}}{n!}.$$

由此可知

$$\frac{e}{(n+1)^{\frac{1}{n}}} - \frac{1}{n!} \le \frac{n}{\sqrt[n]{n!}} \le e - \frac{1}{n!}.$$

左右两边的极限都是e，所以结论得证。

(b) 同(a)的第一种方法，利用Stolz定理可知

$$\lim_{n\to\infty} \frac{A_n}{n} = \lim_{n\to\infty} \frac{na + (1+2+\cdots+n)d}{n^2} = \frac{d}{2},$$

$$\lim_{n\to\infty} \frac{G_n}{n} = \lim_{n\to\infty} \frac{(a+(n+1)d)/(n+1)^{n+1}}{1/n^n} = \lim_{n\to\infty} \frac{(a+(n+1)d)/(n+1)}{(1+1/n)^n} = \frac{d}{e}.$$

所以$\lim_{n\to\infty} A_n/G_n = e/2$。 □

题 2.17. *(a)* 给定$\alpha \in (0,\infty)$，设$x_1 \in (0,1)$，递归定义$x_{n+1} = x_n(1 - x_n^\alpha)$，证明：

$$\lim_{n\to\infty} n^{1/\alpha} x_n = \alpha^{-1/\alpha}.$$

(b) 给定$\delta \in (0,1)$，设$x_1 \in (0,\infty)$，递归定义$x_{n+1} = x_n + x_n^\delta$，证明：

$$\lim_{n\to\infty} n^{1/(1-\delta)} x_n = \delta^{1/(1-\delta)}.$$

(c) 设实数列$a_n\,(n\in\mathbb{N})$满足$a_n\left(\sum_{k=1}^n a_k^2\right) = 1$，证明：$\lim_{n\to\infty} \sqrt[3]{3n}a_n = 1$。

证明. (a) 归纳容易证明$x_n \in (0,1)$且单调递减趋于0。根据Stolz定理（注意验证条件），得

$$\lim_{n\to\infty} n x_n^\alpha = \lim_{n\to\infty} \frac{n}{\frac{1}{x_n^\alpha}} = \lim_{n\to\infty} \frac{1}{\frac{1}{x_{n+1}^\alpha} - \frac{1}{x_n^\alpha}} = \lim_{t\to 0} \frac{1}{\frac{1}{t^\alpha(1-t^\alpha)^\alpha} - \frac{1}{t^\alpha}} = \lim_{s\to 0} \frac{s}{(1-s)^{-\alpha} - 1} = \frac{1}{\alpha}.$$

(b) 由Stolz定理可知（注意$\delta\in(0,1)$）

$$\begin{aligned}\lim_{n\to\infty}\frac{x_n^{1-\delta}}{n}&=\lim_{n\to\infty}(x_{n+1}^{1-\delta}-x_n^{1-\delta})=\lim_{t\to\infty}\left((t+t^\delta)^{1-\delta}-t^{1-\delta}\right)\\&=\lim_{t\to\infty}t^{1-\delta}((1+t^{\delta-1})^\delta-1)=\lim_{t\to\infty}t^{1-\delta}((1+\delta t^{\delta-1}+o(t^{\delta-1}))-1)=\delta.\end{aligned}$$

(c) 由定义容易看出$a_n>0$且单调递减趋于0。记$S_n=a_1^2+\cdots+a_n^2$，则易见$S_n\to\infty$。条件表明$(S_n-S_{n-1})S_n^2=1(n>1)$，由此可知$\frac{S_{n+1}}{S_n}\to 1$，所以

$$\begin{aligned}\lim_{n\to\infty}\frac{n}{S_n^3}&=\lim_{n\to\infty}\frac{1}{S_{n+1}^3-S_n^3}=\lim_{n\to\infty}\frac{1}{(S_{n+1}-S_n)(S_{n+1}^2+S_{n+1}S_n+S_n^2)}\\&=\lim_{n\to\infty}\frac{S_{n+1}^2}{(S_{n+1}-S_n)(S_{n+1}^2+S_{n+1}S_n+S_n^2)}=\frac{1}{3}.\end{aligned}$$

即$3n/S_n^3\to 1$，故

$$\lim_{n\to\infty}\sqrt[3]{3n}a_n=\lim_{n\to\infty}\sqrt[3]{\frac{3n}{S_n^3}}=1.\qquad\square$$

注. *今后会解释(a)(b)中的首阶项次数$n^{1/\alpha}$和$n^{1/(1-\delta)}$是如何找到的。(b)是Bourgain证明二维环面上非线性Schrödinger方程解的高阶Sobolev范数具有至多多项式增长的关键，而在此之前人们只能得到指数甚至双指数增长。提起这件事只是为了说明部分题目在科研中意义重大。*

无穷级数

无穷级数的基本结论大多与数列极限平行。新的内容在于求和次序对无穷级数的影响。

定义 2.18 (无穷级数)**.** *设$a_n\in\mathbb{R}\,(n\in\mathbb{N})$，记$s_n=a_1+\cdots+a_n$，如果$s_n$收敛，则称无穷级数$\sum_{n=1}^\infty a_n$收敛，并记*

$$\sum_{n=1}^\infty a_n:=\lim_{n\to\infty}s_n.$$

对于正项级数，即$a_n\geq 0$的情况，单调收敛原理保证了s_n总是广义收敛（即允许极限是∞），但只有当$\lim s_n<\infty$时才称$\sum a_n$收敛。但在书写时，为了方便，对于正项级数总是允许写$\sum_{n=1}^\infty a_n$，即使它等于∞。下一题给出无穷级数的基本性质。

题 2.19. *(a) 设$a_n\geq 0$，则$\sum_{n=1}^\infty a_n$收敛当且仅当存在$M<\infty$使得$s_n<M\ (\forall n\in\mathbb{N})$。*

(b) 如果$\sum_{n=1}^\infty a_n$收敛，则$a_n\to 0$。

(c) (Cauchy收敛原理)*无穷级数$\sum_{n=1}^\infty a_n$收敛的充要条件是，对任意$\varepsilon>0$，存在$N>0$，使得当$m>n>N$时，$|\sum_{k=n}^m a_k|<\varepsilon$。*

(d) 如果$\sum_{n=1}^{\infty}|a_n|$收敛（称为绝对收敛），则$\sum_{n=1}^{\infty}a_n$收敛。

证明. (a)实际上已经包含在上面的一段说明里，(b)(c)由数列极限的性质推出，具体细节留作练习。最后，(d)是由$|\sum_{k=n}^{m}a_k|<\sum_{k=n}^{m}|a_k|$以及(c)推出。 □

为了证明级数收敛，通常可以证明其绝对收敛，而这又可以通过与常见的收敛正项级数比较。下面给出三个常用的级数，其中(c)为了方便只考虑$x>0$，稍后会处理任意$x\in\mathbb{R}$。

题 2.20. *(a)* 设$|q|<1$，证明：$\sum_{n=0}^{\infty}q^n=\frac{1}{1-q}$。

(b) 证明：当$0<s\le 1$时，$\sum_{n=1}^{\infty}\frac{1}{n^s}$发散；当$s>1$时，$\sum_{n=1}^{\infty}\frac{1}{n^s}$收敛。

(c) 证明：当$x>0$时，$\sum_{n=0}^{\infty}\frac{x^n}{n!}=e^x$。

证明. (a) 根据等比数列求和公式（或者n次方差公式），

$$1+q+\cdots+q^n=\frac{1-q^{n+1}}{1-q}.$$

因为$|q|<1$，所以右边的极限是$\frac{1}{1-q}$。

(b) 先证明$0<s\le 1$时发散，又只需考虑$s=1$的情况。这是因为对任意$N\in\mathbb{N}$，有

$$\sum_{n=1}^{2^N}\frac{1}{n}=\sum_{k=1}^{N}\sum_{n=2^{k-1}}^{2^k}\frac{1}{n}\ge\sum_{k=1}^{N}\sum_{n=2^{k-1}}^{2^k}\frac{1}{2^k}=\sum_{k=1}^{N}\frac{1}{2}=\frac{N}{2}.$$

所以部分和没有上界。下面考虑$s>1$的情况，用类似的估计，

$$\sum_{n=1}^{2^N}\frac{1}{n^s}=\sum_{k=1}^{N}\sum_{n=2^{k-1}}^{2^k}\frac{1}{n^s}\le\sum_{k=1}^{N}\sum_{n=2^{k-1}}^{2^k}\frac{1}{2^{(k-1)s}}=\sum_{k=1}^{N}\frac{1}{2^{(k-1)(s-1)}}.$$

因为$s>1$，所以$q:=2^{-(s-1)}\in(0,1)$，由(a)可知上式不超过$\frac{1}{1-q}$。

(c) 根据二项式定理，对任意$n\in\mathbb{N}$，有

$$\left(1+\frac{x}{n}\right)^n=\sum_{k=0}^{n}C_n^k\frac{x^k}{n^k}=\sum_{k=0}^{n}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\frac{x^k}{k!}.$$

括号里的每一项都不超过1，结合当$n\to\infty$时左边趋于e^x，可知

$$e^x\le\sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

反过来，对任意给定的$m\in\mathbb{N}$，当$n>m$时，

$$\left(1+\frac{x}{n}\right)^n\ge\sum_{k=0}^{m}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\frac{x^k}{k!},$$

取极限得（注意这里求和的项数是固定的，因此可以分别求极限再相加）

$$e^x \geq \sum_{k=0}^{m} \frac{x^k}{k!}.$$

由m的任意性即证。 □

注. 如果$\mathrm{Re}(s) > 1$，由*(b)*可知*Riemann* ζ函数$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$良定义。著名的*Basel*问题是说

$$\zeta(2) = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = ?$$

*Euler*首先给出了答案$\zeta(2) = \frac{\pi^2}{6}$。受限于当时复变函数的理论尚不完全，*Euler*并没有给出严格的证明，但他通过观察零点敏锐地看出

$$\frac{\sin(\pi z)}{\pi z} = \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right).$$

比较两边展开式中z^2的系数就得到了这一结果。今后会给出一个初等证明，并借助*Bernoulli*数求出$\zeta(2n)$ $(n \in \mathbb{N})$的表达式。而ζ函数在奇数处的值充满神秘，比如目前只知道$\zeta(3)$是无理数，至于是否是超越数仍是难题。

下一题的(b)说明，对任意$x \in \mathbb{C}$都可以定义

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

用幂级数定义指数函数的好处在于，幂级数总是绝对收敛，而$\left(1 + \frac{x}{n}\right)^n$并不显然具有极限。今后还会看到，幂级数具有诸多良好性质，比如可以逐项求导等等。

题 2.21. *(a)* (Bernoulli不等式)设$x_1, \ldots, x_n \in [0, 1]$，证明：

$$(1 - x_1) \cdots (1 - x_n) \geq 1 - x_1 - \cdots - x_n.$$

(b) 证明：对任意$x \in \mathbb{C}$和$n \in \mathbb{N}$，有

$$\left| \left(1 + \frac{x}{n}\right)^n - \sum_{k=0}^{n} \frac{x^k}{k!} \right| \leq \frac{|x|^2 e^{|x|}}{2n}.$$

*(c)** 证明：

$$\lim_{n \to \infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

证明. (a) 对n归纳，$n=1$是平凡的。假设结论对$n-1$成立，由$x_i\in[0,1]$可知

$$\begin{aligned}(1-x_1)\cdots(1-x_n)\geq&(1-x_1-\cdots-x_{n-1})(1-x_n)\\=&1-x_1-\cdots-x_n+x_n(x_1+\cdots+x_n)\\\geq&1-x_1-\cdots-x_n.\end{aligned}$$

(b) 根据二项式展开及Bernoulli不等式，

$$\begin{aligned}\left|\left(1+\frac{x}{n}\right)^n-\sum_{k=0}^{n}\frac{x^k}{k!}\right|=&\left|\sum_{k=0}^{n}\left(1-\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\right)\frac{x^k}{k!}\right|\\\leq&\sum_{k=0}^{n}\left(\frac{1}{n}+\cdots+\frac{k-1}{n}\right)\frac{|x|^k}{k!}\\=&\frac{1}{2n}\sum_{k=2}^{n}\frac{|x|^k}{(k-2)!}\leq\frac{|x|^2}{2n}\sum_{k=0}^{\infty}\frac{|x|^k}{k!}=\frac{|x|^2e^{|x|}}{2n}.\end{aligned}$$

(c) 这可以由概率论的中心极限定理得到。假设$X_1,X_2,\ldots$都是服从参数为1的Poisson分布的独立随机变量，则根据中心极限定理，

$$e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=P(X_1+\cdots+X_n\leq n)=P\left(\frac{X_1+\cdots+X_n-n}{\sqrt{n}}\leq 0\right)\to\Phi(0)=\frac{1}{2},$$

其中Φ是标准正态分布的分布函数。尚不知本题是否具有合理的分析方法。 □

题 2.22 *. 设$x>1$，证明：

$$\sum_{n=1}^{\infty}\frac{n!}{(x+1)(x+2)\cdots(x+n)}=\frac{1}{x-1}.$$

证明. 裂项，对$n\geq 2$有

$$\begin{aligned}&\frac{n!}{(x+1)(x+2)\cdots(x+n)}=\frac{1}{x-1}\frac{n!((x+n)-(n+1))}{(x+1)(x+2)\ldots(x+n)}\\=&\frac{1}{x-1}\left(\frac{n!}{(x+1)(x+2)\cdots(x+n-1)}-\frac{(n+1)!}{(x+1)(x+2)\cdots(x+n)}\right).\end{aligned}$$

所以

$$
\begin{aligned}
&\sum_{n=1}^{\infty}\frac{n!}{(x+1)(x+2)\cdots(x+n)}\\
=&\frac{1}{x+1}+\frac{1}{x-1}\sum_{n=2}^{\infty}\left(\frac{n!}{(x+1)(x+2)\cdots(x+n-1)}-\frac{(n+1)!}{(x+1)(x+2)\cdots(x+n)}\right)\\
=&\frac{1}{x+1}+\frac{1}{x-1}\lim_{n\to\infty}\left(\frac{2}{x+1}-\frac{(n+1)!}{(x+1)(x+2)\cdots(x+n)}\right)\\
=&\frac{1}{x+1}+\frac{1}{x-1}\cdot\frac{2}{x+1}=\frac{1}{x-1}.
\end{aligned}
$$

注意这里用到了当$x>1$时$\frac{(n+1)!}{(x+1)\cdots(x+n)}\to 0$，具体细节留作练习。 □

题 2.23 * (2016年丘成桐大学生数学竞赛–分析与方程决赛). 任取$x_1\in(0,\pi/2)$，递归定义$x_{n+1}=\sin x_n$，证明：$\sum_{n=1}^{\infty}x_n^2=\infty$。

证明. 注意到对任意$n\in\mathbb{N}$有

$$\frac{1}{\sqrt{n}}-\frac{1}{\sqrt{n+1}}=\frac{1}{\sqrt{n(n+1)}(\sqrt{n}+\sqrt{n+1})}\ge\frac{1}{2}\cdot\frac{1}{(n+1)\sqrt{n}}\ge\frac{1}{4}\cdot\frac{1}{n\sqrt{n}}.$$

可以归纳证明$x_n\ge\frac{C}{\sqrt{n}}$，其中C是正常数。事实上，利用$\sin x\ge x-\frac{x^3}{6}$以及单调性，只需

$$\frac{1}{\sqrt{n}}-\frac{1}{\sqrt{n+1}}\ge\frac{C^2}{n\sqrt{n}}.$$

因此只要取$C=\min\left\{x_1,\frac{1}{2}\right\}$即可。故$\sum_{n=1}^{\infty}x_n^2\ge\sum_{n=1}^{\infty}\frac{C^2}{n}=\infty$。 □

注. 易见x_n单调递减趋于0，利用*Stolz*定理和*Taylor*公式可知

$$\lim_{n\to\infty}nx_n^2=\lim_{n\to\infty}\frac{1}{\frac{1}{x_{n+1}^2}-\frac{1}{x_n^2}}=\lim_{t\to 0}\left(\frac{1}{\sin^2 t}-\frac{1}{t^2}\right)^{-1}=3.$$

即$x_n\sim\frac{\sqrt{3}}{\sqrt{n}}$。至此已经见过几例无穷小量的首阶项结果，本题给出的证明很好地解释了首阶项是如何找到的：$\sin x$的*Taylor*展开式前两项是$x-\frac{x^3}{6}$，如果$x_n\approx n^{\alpha}(\alpha<0)$，则应该期待

$$n^{\alpha}-(n+1)^{\alpha}\approx n^{3\alpha}.$$

根据*Taylor*公式（或者说导数的定义），

$$n^{\alpha}-(n+1)^{\alpha}=n^{\alpha}\left(1-\left(1+\frac{1}{n}\right)^{\alpha}\right)\approx-\alpha n^{\alpha-1}.$$

因此应当选择$\alpha-1=3\alpha$，即$\alpha=-\frac{1}{2}$。类似可以解释先前的题目中，为何递推关系$x_{n+1}=x_n(1-x_n)\,(x_1\in(0,1))$会导致$x_n$的首阶项形如$c/n$。

对于无穷级数$\sum_{n=1}^{\infty} a_n$，所谓重新排序是指对于一个双射$\tau\colon \mathbb{N} \to \mathbb{N}$，考虑$\sum_{n=1}^{\infty} a_{\tau(n)}$的收敛性以及是否与原级数相等，即求和是否与次序无关。下面将证明对于正项级数和绝对收敛的级数，求和是与次序无关的；但对于条件收敛的级数，求和次序至关重要。

题 2.24. *(a)* 设$a_n \geq 0$，证明：

$$\sum_{n=1}^{\infty} a_n = \sup\left\{\sum_{i\in I} a_i \mid I \subset \mathbb{N}\text{是有限集}\right\}.$$

由此可知正项级数求和与次序无关。

*(b)** 设$a_\alpha \geq 0\,(\alpha \in I)$是一族非负实数，其中指标集$I$是任意的，定义

$$\sum_{\alpha\in I} a_\alpha := \sup\left\{\sum_{\alpha\in J} a_\alpha \mid J \subset I\text{是有限集}\right\}.$$

证明：如果$\sum_{\alpha\in I} a_\alpha < \infty$，则$\{\alpha \in I \mid a_\alpha \neq 0\}$是可数集。

(c) 设$\sum_{n=1}^{\infty} |a_n| < \infty$，证明：对任意双射$\tau\colon \mathbb{N} \to \mathbb{N}$，都有

$$\sum_{n=1}^{\infty} a_{\tau(n)} = \sum_{n=1}^{\infty} a_n.$$

(d) 设$\sum_{n=1}^{\infty} a_n$收敛但$\sum_{n=1}^{\infty} |a_n| = \infty$（称为条件收敛），证明：对任意$A \in [-\infty, \infty]$，存在双射$\tau\colon \mathbb{N} \to \mathbb{N}$，使得

$$\sum_{n=1}^{\infty} a_{\tau(n)} = A$$

证明. (a) 记右边为A，则显然有$\sum_{n=1}^{\infty} a_n \geq A$。另一方面，对任意$M < \sum_{n=1}^{\infty} a_n$，存在$N > 0$使得$M < \sum_{n=1}^{N} a_n < A$，这就说明$\sum_{n=1}^{\infty} a_n \leq A$，从而结论得证。[应该注意的是，这样的写法对$\sum_{n=1}^{\infty} a_n = \infty$的情况同样适用。]

(b) 对任意$n \in \mathbb{N}$，显然有

$$\sum_{\alpha\in I} a_\alpha \geq \frac{1}{n} \cdot \left|\left\{\alpha \in I \mid a_\alpha > \frac{1}{n}\right\}\right|,$$

所以$\{\alpha \in I \mid a_\alpha > \frac{1}{n}\}$是有限集。最后，因为

$$\{\alpha \in I \mid a_\alpha \neq 0\} = \bigcup_{n=1}^{\infty} \left\{\alpha \in I \mid a_\alpha > \frac{1}{n}\right\},$$

所以左边是可数集。

(c) 设$\sum_{n=1}^{\infty} a_n = A$，则对任意$\varepsilon > 0$，存在$M > 0$使得

$$\sum_{n=M+1}^{\infty} |a_n| < \varepsilon \quad 且 \quad \left| A - \sum_{n=1}^{N} a_n \right| < \varepsilon.$$

存在$N > 0$使得$\{1, \ldots, M\} \subset \{\tau(1), \ldots, \tau(N)\}$，因此当$n > N$时，

$$\left| \sum_{k=1}^{n} a_{\tau(k)} - A \right| \le \left| \sum_{k=1}^{n} a_{\tau(k)} - \sum_{k=1}^{M} a_k \right| + \left| \sum_{k=1}^{M} a_k - A \right| \le \sum_{k=M+1}^{\infty} |a_k| + \varepsilon \le 2\varepsilon.$$

这就证明了$\sum_{n=1}^{\infty} a_{\tau(n)} = A$。

(d) 设$x_1, x_2, \ldots$是这些a_n中全体大于等于0的数的排列，$y_1, y_2, \ldots$是其中全体小于0的数的排列，则容易证明$\sum |x_n| = \sum |y_n| = \infty$且$x_n, y_n \to 0$。

先证明$A \in \mathbb{R}$的情况，递归定义$p(n), q(n)$如下：$p(1)$是使得$x_1 + \cdots + x_m \ge A$的最小正整数m，$q(1)$是使得$x_1 + \cdots + x_{p(1)} + y_1 + \cdots + y_m < A$的最小正整数$m$；一般的，假设$p(n-1), q(n-1)$都已经定义，则

$$p(n) := \min\{m > p(n) \mid x_1 + \cdots + x_m + y_1 + \cdots + y_{q(n-1)} \ge A\},$$

$$q(n) = \min\{m > q(n) \mid x_1 + \cdots + x_{p(n)} + y_1 + \cdots + y_{q(n)} < A\}.$$

考虑如下无穷级数：

$$(x_1 + \cdots + x_{p(1)}) + (y_1 + \cdots + y_{q(1)}) + (x_{p(1)+1} + \cdots + x_{p(2)}) + \cdots.$$

则这是$\sum_{n=1}^{\infty} a_n$的重新排序，下面证明这个无穷级数等于A。记S_n是前n个括号内的项之和，则根据定义有$|S_{2n} - A| < \max\{x_{p(n)}, y_{q(n)}\}$和$|S_{2n+1} - A| \le \max\{x_{p(n+1)}, y_{q(n)}\}$，因为$p(n), q(n) \ge n$且$x_n, y_n \to 0$，所以$S_n \to A$。最后，根据$x_n \ge 0$和$y_n < 0$可知，第$n$个括号到第$n+1$个括号之间的部分和落在$S_n$和$S_{n+1}$之间，这就说明了这个无穷级数收敛到$A$。[一般的，假设对一个无穷级数加了若干括号，满足每个括号内的项符号都相同，如果得到的新级数收敛，则原来的级数求和也收敛，并且二者相等。]

最后，$A = \infty$和$-\infty$的情况是类似的，具体细节留作练习。 □

注. *(c)也可以借助(a)简单证明：注意到$a_n = a_n^+ - a_n^-$，其中$x^+ := \max\{x, 0\}$，$x^- := -\min\{x, 0\}$；于是$\sum a_n^+$和$\sum a_n^-$都是收敛的正项级数，求和不依赖次序，从而二者作差也不依赖求和次序。*

第3讲：数列的极限（下）

"不，被告不可能逃走。法庭从毁坏的门延伸到任何地方。"

——安部公房《S·卡尔玛先生的罪行》

这一讲的难点是上、下极限。上极限通常有两种刻画：一种是$\lim_{k\to\infty}\sup_{n\geq k}a_n$，另一种是对所有子列极限取上确界。在题目中需要灵活使用两种定义。此外，也涉及到关于实数的拓扑性质的内容，比如聚点的概念和有限覆盖定理等，在将来研究连续函数的时候会展现威力。补充内容有两部分：一是介绍无穷级数中名为Abel变换（或者说分部求和）的重要技巧，二是引进度量空间的语言，作为谈论实数拓扑性质的推广、以及今后进入拓扑学的敲门砖。

例题

例 3.1. 是否存在$[0,1]$中的数列聚点是$[0,1]$？又能否是$[0,1]\setminus\left\{\frac{1}{2}\right\}$？

证明. 取x_n是$\mathbb{Q}\cap[0,1]$的一个排列，则容易证明x_n的聚点集是整个$[0,1]$。但要让聚点集是$[0,1]\setminus\left\{\frac{1}{2}\right\}$是不可能的，因为"聚点的聚点还是聚点"，具体细节留作练习。 □

注. 自然的问题是，$\mathbb{R}$的子集X何时能称为某个数列的聚点集。从上面的证明可以看出，一个必要条件是X的聚点仍然属于X。这样的集合被称为闭集*(closed set)*。今后会证明这一条件也是充分的，即$\mathbb{R}$中的闭集都是聚点集。

除了单调收敛以及Cauchy列，下一题展示了（在不求出极限的情况下）证明收敛性的另一种方法，即证明上、下极限相等。但要注意通常还需要说明有限性，才能得出真正意义下的收敛，否则只能得到在$[-\infty,\infty]$中广义收敛。

例 3.2. 设$x_n\in\mathbb{R}\,(n\in\mathbb{N})$满足$x_{m+n}\leq x_m+x_n\,(\forall m,n\in\mathbb{N})$，证明：$\frac{x_n}{n}$在$[-\infty,\infty)$中收敛（即允许极限为$-\infty$，但不可以是$\infty$），且

$$\lim_{n\to\infty}\frac{x_n}{n}=\inf_{n\in\mathbb{N}}\frac{x_n}{n}.$$

证明. 任意取定$m\in\mathbb{N}$，则每个$n\in\mathbb{N}$可以唯一表示为$n=pm+q$，其中$p,q\in\mathbb{N}_0$且$1\leq q\leq m$。反复利用条件，并结合$|p-n/m|\leq 1$可知

$$x_n\leq px_m+x_q\leq n\cdot\frac{x_m}{m}+|x_m|+\sup\{|x_1|,\ldots,|x_m|\}.$$

两边除以n并取上极限，得

$$\limsup_{n\to\infty}\frac{x_n}{n}\leq\frac{x_m}{m}.$$

由m的任意性可知
$$\limsup_{n\to\infty}\frac{x_n}{n}\le\inf_{n\in\mathbb{N}}\frac{x_n}{n}\le\liminf_{n\to\infty}\frac{x_n}{n}.$$
显然$\inf_{n\in\mathbb{N}}\frac{x_n}{n}\in[-\infty,\infty)$，所以结论成立。 □

注. 满足$x_{m+n}\le x_m+x_n$的数列称为次可加的。这一命题在分析学中具有实际应用，比如动力系统中定义的*Lyapunov*指数、测度熵和拓扑熵等。

练习题

题 3.3. 证明：$\mathbb{R}$的不可数子集一定有聚点。

证明. 如果$X\subset\mathbb{R}$不可数，则因为$X=\bigcup_{n\in\mathbb{N}}(X\cap[-n,n])$可数，所以存在$n\in\mathbb{N}$使得$X\cap[-n,n]$是无限集（否则可数个有限集之并至多可数，矛盾），从而X在$[-n,n]$中有聚点。 □

题 3.4. *(a)* 设数列$a_n\in\mathbb{R}$，证明：
$$\liminf_{n\to\infty}a_n\le\liminf_{n\to\infty}\frac{a_1+\cdots+a_n}{n}\le\limsup_{n\to\infty}\frac{a_1+\cdots+a_n}{n}\le\limsup_{n\to\infty}a_n.$$

(b) 设数列$a_n>0$，证明：
$$\liminf_{n\to\infty}\frac{a_{n+1}}{a_n}\le\liminf_{n\to\infty}\sqrt[n]{a_n}\le\limsup_{n\to\infty}\sqrt[n]{a_n}\le\limsup_{n\to\infty}\frac{a_{n+1}}{a_n}.$$

证明. (a) 只证最右边的不等式，左边是类似的。对任意$k\in\mathbb{N}$，有
$$\limsup_{n\to\infty}\frac{a_1+\cdots+a_n}{n}\le\limsup_{n\to\infty}\frac{a_1+\cdots+a_k+(n-k)\sup_{n>k}a_n}{n}=\sup_{n>k}a_n.$$
对k取极限即证。

(b) 本质上是(a)的乘法版本，具体细节留作练习。 □

注. *(b)*说明*Cauchy*根式判别法比*d'Alembert*比值判别法更强。今后在定义幂级数的收敛半径时，用的也是*Cauchy*根式判别法，主要是为了兼顾某些$|a_n|=0$的情况。

题 3.5. 设数列$a_n\in\mathbb{R}$，记$S_n=a_1+\cdots+a_n$。若$\limsup_{n\to\infty}\sqrt[n]{|a_n|}\ge 1$，证明：
$$\limsup_{n\to\infty}\sqrt[n]{|a_n|}=\limsup_{n\to\infty}\sqrt[n]{|S_n|}.$$

证明. 先证$\le$，如果$\limsup_{n\to\infty}\sqrt[n]{|S_n|}=\infty$，则自动成立，因此不妨设其有限。对任意正实数$A>\limsup_{n\to\infty}\sqrt[n]{|S_n|}$，存在$N\in\mathbb{N}$使得当$n\ge N$时$S_n\le A^n$。于是对$n\ge N+1$有
$$|a_n|=|S_n-S_{n-1}|\le|S_n|+|S_{n-1}|\le A^n+A^{n-1}=(1+A^{-1})A^n.$$

由此可知$\limsup_{n\to\infty}\sqrt[n]{|a_n|}\le A$，结合$A$的任意性即可。[这一步没有用到任何条件。]

再证$\ge$，同样不妨设$\limsup_{n\to\infty}\sqrt[n]{|a_n|}<\infty$。对任意正实数$B>\limsup_{n\to\infty}\sqrt[n]{|a_n|}$，存在$M\in\mathbb{N}$使得当$n\ge M$时$|a_n|\le B^n$。注意到条件保证了$B>1$，所以对$n>N$有

$$|S_n|\le|S_N|+|a_{N+1}|+\cdots+|a_n|\le|S_N|+B^{N+1}+\cdots+B^n\le|S_N|+nB^n.$$

由此可知$\limsup_{n\to\infty}\sqrt[n]{|S_n|}\le B$，结合$B$的任意性即可。 □

注. 条件$\limsup_{n\to\infty}\sqrt[n]{|a_n|}\ge 1$是必要的，否则取$a_n=2^{-n}$可得反例。特别的，这一条件可以由$\sum_{n=1}^{\infty}a_n$发散保证。本题也可以借助幂级数说明：形式上

$$(1-x)\sum_{n=1}^{\infty}S_nx^n=\sum_{n=1}^{\infty}a_nx^n,$$

在已知右边的收敛半径不超过1的前提下，可以说明左边的收敛半径等于右边。

题 3.6 *. 设数列$x_n>0$，证明：$\limsup\left(\frac{1+x_{n+1}}{x_n}\right)^n\ge e$，并且右边的常数$e$是不可改进的。

证明*1.* 先给出一种巧妙的反证法。若结论不成立，则存在$\varepsilon>0$和$N>0$，使得当$n\ge N$时

$$\left(\frac{1+x_{n+1}}{x_n}\right)^n<e-\varepsilon.$$

将N适当放大，可以让$\left(1+\frac{1}{n}\right)^n>e-\varepsilon$对一切$n\ge N$成立，于是

$$\frac{1+x_{n+1}}{x_n}\le 1+\frac{1}{n}.$$

这表明

$$\frac{x_n}{n}\ge\frac{1}{n+1}+\frac{x_{n+1}}{n+1}.$$

递推可知对任意$k\in\mathbb{N}$都有

$$\frac{x_n}{n}\ge\frac{1}{n+1}+\frac{1}{n+2}+\cdots+\frac{1}{n+k}+\frac{x_{n+k}}{n+k}.$$

但$\sum_{k>n}\frac{1}{k}=\infty$，矛盾！所以本题的结论成立。

最后，取$x_n=kn$，其中$k>0$待定，则

$$\left(\frac{1+x_{n+1}}{x_n}\right)^n=\left(1+\frac{1+k}{kn}\right)^n\to e^{\frac{1+k}{k}}.$$

当$k\to\infty$时右边的极限是e，所以本题的结论中e不能换成任何更大的数。 □

证明2. 下面提供一种更直观的思路。如果当$n\geq N$时$\left(\frac{1+x_{n+1}}{x_n}\right)^n\leq e-\varepsilon$，则

$$x_{n+1}\leq(e-\varepsilon)^{\frac{1}{n}}x_n-1.$$

归纳可知

$$x_n\leq(e-\varepsilon)^{\frac{1}{N}+\frac{1}{N+1}+\cdots+\frac{1}{n-1}}x_N-(n-N)=e^{(\frac{1}{N}+\cdots+\frac{1}{n-1})\log(e-\varepsilon)}-(n-N).$$

由$\frac{1}{N}+\frac{1}{N+1}+\cdots+\frac{1}{n-1}\sim\log n$可知前一项约为$Cn^{\log(e-\varepsilon)}$，又$n^{\log(e-\varepsilon)}\ll n$，可知$x_n\geq 0$不可能对充分大的$n$成立，严格论证留作练习。说明$e$不能改进的例子同上。 □

题 3.7 *. 证明下列无穷乘积公式：

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^3}\right)=\frac{\cosh\left(\frac{\sqrt{3}}{2}\pi\right)}{3\pi},$$

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^4}\right)=\frac{\sinh\pi}{4\pi}.$$

证明. 需要用到Euler Γ函数，已经远超本课程的范围。有无穷乘积展开

$$\Gamma(x)=\frac{1}{x}e^{-\gamma x}\prod_{n=1}^{\infty}\left(1+\frac{x}{n}\right)^{-1}e^{x/n}.$$

记$\omega=e^{2\pi i/x}$，考察$\Gamma(-\omega)\Gamma(-\omega^2)$。一方面用乘积展开计算，结合$(1-\omega x)(1-\omega^2x)=1+x+x^2$容易看出与本题关系；另一方面，利用余元公式

$$\Gamma(\alpha)\Gamma(1-\alpha)=\frac{\pi}{\sin(\pi\alpha)},$$

可以求出最终的结果。[一般的，即使将减号换成加号，或者把指数换成任意大于1的整数，则结果都可以用Γ函数在单位根处的取值表示，只是通常无法继续化简。] □

注. 之所以将这个问题放在这里，是因为容易看出

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right)=\lim_{N\to\infty}\prod_{n=2}^{N}\frac{(n+1)(n-1)}{n^2}=\lim_{N\to\infty}\frac{N+1}{2N}=\frac{1}{2}.$$

然而若将二次改为三次，则基本的极限理论难以对付。还可以问若是将减号变成加号又该如何，答案亦与Γ函数有关，比如二次的情况是（也可由$\sin(x)$的无穷乘积展开得到）

$$\prod_{n=1}^{\infty}\left(1+\frac{1}{n^2}\right)=\frac{\sinh x}{\pi}.$$

这里的教训在于：不要做来路不明的题。

补充：Abel变换

在考察形如$\sum a_nb_n$的求和时，如果有关于a_n的部分和$S_n := a_1 + \cdots + a_n$、以及$b_n$的相邻两项之差$b_{n+1} - b_n$的信息，则一个常用的手法是Abel变换，即

$$\sum_{k=1}^{n} a_kb_k = S_1b_1 + \sum_{k=2}^{n}(S_k - S_{k-1})b_k = \sum_{k=1}^{n-1} S_k(b_k - b_{k+1}) + S_nb_n.$$

题 3.8 (Abel求和). 设$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \le 1$，则当$|r| < 1$时$\sum_{n=0}^{\infty} a_nr^n$绝对收敛。证明：

(a) 如果$\sum_{n=0}^{\infty} a_n$收敛，则

$$\lim_{r\to 1-}\sum_{n=0}^{\infty} a_nr^n = \sum_{n=0}^{\infty} a_n.$$

*(b)** 如果a_n是*Cesàro*可求和的，则也是*Abel*可求和的，且两种求和的结果相同。

*(c)** 存在数列a_n是*Abel*可求和的但不是*Cesàro*可求和的。

*(d)** 如果a_n是*Abel*可求和的，且$a_n = o\left(\frac{1}{n}\right)$，则$\sum_{n=0}^{\infty} a_n$收敛。

*(e)** 如果a_n是*Abel*可求和的，且$\lim_{n\to\infty} \frac{a_1+2a_2+\cdots+na_n}{n} = 0$，则$\sum_{n=0}^{\infty} a_n$收敛。

*(f)** 如果a_n是*Abel*可求和的，且$a_n = O\left(\frac{1}{n}\right)$，则$\sum_{n=0}^{\infty} a_n$收敛。

证明. (a) 用$a_0 - \sum_{n=0}^{\infty} a_n$代替$a_0$，可以不妨设$\sum_{n=0}^{\infty} a_n = 0$。记$S_n = a_0 + \cdots + a_n$，则

$$\sum_{n=0}^{N} a_nr^n = S_Nr^N + \sum_{n=0}^{N-1} S_n(r^n - r^{n+1}).$$

令$N \to \infty$，结合$|S_Nr^N| \to 0$可知

$$\sum_{n=0}^{\infty} a_nr^n = (1-r)\sum_{n=0}^{\infty} S_nr^n.$$

对任意$\varepsilon > 0$，存在$N > 0$使得当$n > N$时$|S_n| < \varepsilon$，所以

$$\left|\sum_{n=0}^{\infty} a_nr^n\right| \le \left|(1-r)\sum_{n=0}^{N} S_nr^n\right| + (1-r)\sum_{n=N+1}^{\infty} \varepsilon r^n \le \left|(1-r)\sum_{n=0}^{N} S_nr^n\right| + \varepsilon.$$

由此可知

$$\limsup_{r\to 1-}\left|\sum_{n=0}^{\infty} a_nr^n\right| \le \varepsilon.$$

从而结论成立。

(b) 回忆在Cesàro求和的定义中，$\sigma_n = \frac{S_0+\cdots+S_n}{n+1}$。同(a)的方法可知

$$\sum_{n=0}^{\infty} a_n r^n = (1-r)^2 \sum_{n=0}^{\infty} (n+1)\sigma_n r^n.$$

需要证明当$r \to 1-$时，右边的极限是$\lim_{n\to\infty} \sigma_n$。这和(a)的办法是一样的，需要用到

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

为了证明这件事，一种方法是对$\frac{1}{1-r} = \sum_{n=0}^{\infty} r^n$两边求导，但需要事先证明右边可以逐项求导（今后将证明幂级数都可以逐项求导）；另一种更初等的办法是，首先说明左边是绝对收敛的，这是因为任取$s \in (r,1)$，则有$(n+1)r^n \ll s^n$；记求和为R，则

$$rR = \sum_{n=0}^{\infty} (n+2) r^{n+1} - \sum_{n=0}^{\infty} r^{n+1} = R - 1 - \left(\frac{1}{1-r} - 1\right) = R - \frac{1}{1-r}.$$

由此解出$R = \frac{1}{(1-r)^2}$。

(c) 取$a_n = (-1)^{n+1} n$。先考虑Cesàro求和，则$(S_n)_{n=0}^{\infty} = (0,1,-1,2,-2,3,-3,\dots)$，从而

$$S_1 + \cdots + S_n = \begin{cases} 0, & 2 \mid n, \\ \frac{n+1}{2}, & 2 \nmid n. \end{cases}$$

特别的，$\sigma_{2n+1} \to 0$而$\sigma_{2n} \to \frac{1}{2}$，因此$a_n$不是Cesàro可求和的。

再考虑Abel求和，则

$$\sum_{n=0}^{\infty} a_n r^n = -\sum_{n=0}^{\infty} (n+1)(-r)^n + \sum_{n=0}^{\infty} (-r)^n = -\frac{1}{(1+r)^2} + \frac{1}{1+r} = \frac{r}{(1+r)^2}.$$

所以a_n是Abel可求和的，并且结果是$\frac{1}{4}$。

(d) 显然(f)⇒(e)⇒(d)，但证明的难度不同，因此分别列出。对任意$n \in \mathbb{N}$，取$r_n = 1 - \frac{1}{n}$，利用Bernoulli不等式可知对$0 \le k \le n$有$|1 - r_n^k| \le \frac{k}{n}$，所以

$$\left| \sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_k r_n^k \right| \le \sum_{k=0}^{n} |a_k| \cdot \frac{k}{n} = \frac{1}{n} \sum_{k=0}^{n} k|a_k|.$$

因为$na_n \to 0$，所以当$n \to \infty$时上式极限是0。记a_n的Abel求和是A，则对任意$\varepsilon > 0$，存在$N > 0$使得当$n > N$时以下三个式子成立：

$$n|a_n| < \varepsilon, \quad \left| \sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_k r_n^k \right| < \varepsilon, \quad \left| A - \sum_{k=0}^{\infty} a_k r_n^k \right| < \varepsilon.$$

由此可知

$$
\begin{aligned}
\left|\sum_{k=0}^{n} a_k - A\right| &\le \left|\sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_k r_n^k\right| + \left|\sum_{k=n+1}^{\infty} a_k r_n^k\right| + \left|A - \sum_{k=0}^{\infty} a_k r_n^k\right| \\
&\le \varepsilon + \sum_{k=n+1}^{\infty} \frac{\varepsilon}{k} \cdot r_n^k + \varepsilon \le 2\varepsilon + \frac{\varepsilon}{n} \sum_{k=0}^{\infty} r_n^k \\
&= 2\varepsilon + \frac{\varepsilon}{n} \cdot \frac{1}{1-r_n} = 3\varepsilon.
\end{aligned}
$$

这就证明了$\sum_{k=0}^{\infty} a_n = A$。

(e) 不妨设$a_0 = 0$。记$b_n = a_1 + 2a_2 + \cdots + na_n$，则$\frac{b_n}{n} \to 0$。利用Abel变换得

$$
\begin{aligned}
\sum_{n=1}^{\infty} a_n r^n = \lim_{N\to\infty}\left(\sum_{n=1}^{N} a_n r^n\right) &= \lim_{N\to\infty}\left(\sum_{n=1}^{N-1} b_n\left(\frac{r^n}{n} - \frac{r^{n+1}}{n+1}\right) + \frac{b_N}{N} r^n\right) \\
&= \sum_{n=1}^{\infty} b_n\left(\frac{r^n}{n} - \frac{r^{n+1}}{n+1}\right) = (1-r)\sum_{n=1}^{\infty} \frac{b_n}{n+1} r^n + \sum_{n=1}^{\infty} \frac{b_n}{n(n+1)} r^n.
\end{aligned}
$$

注意这里运算的合法性：第二行的第一个等号用到了$\frac{b_N}{N} \to 0$，第二个等号是分解成两个收敛的无穷级数之和。仿照(a)的证明，由$\frac{b_n}{n} \to 0$可知

$$
\lim_{r\to 1-} (1-r)\sum_{n=1}^{\infty} \frac{b_n}{n+1} r^n = 0.
$$

因此在上式两边令$r \to 1-$可知$\frac{b_n}{n(n+1)}$是Abel可求和的，并且结果等于a_n的Abel求和，记为A。又因为$\frac{b_n}{n(n+1)} = o\left(\frac{1}{n}\right)$，由(d)可知$\sum_{n=1}^{\infty} \frac{b_n}{n(n+1)} = A$。故

$$
\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty}\left(\sum_{n=1}^{N} a_n\right) = \lim_{N\to\infty}\left(\sum_{n=1}^{N-1} \frac{b_n}{n(n+1)} + \frac{b_N}{N}\right) = \sum_{n=1}^{\infty} \frac{b_n}{n(n+1)} = A.
$$

(f) [4] 定义截断函数$f\colon [0,1] \to \mathbb{R}$为（这里的$\frac{1}{2}$并不重要，换成任何$(0,1)$中的数都行）

$$
f(x) := \begin{cases} 0, & 0 \le x \le \frac{1}{2}, \\ 1, & \frac{1}{2} < x \le 1. \end{cases}
$$

断言： $\lim_{x\to 1-} \sum_{n=0}^{\infty} a_n f(x^n) = A$，其中$A$是$a_n$的Abel求和。

易见这一断言能够推出原题。其证明被分成若干引理陈述。

[4]这个复杂而精妙的证明来自J. Korevaar, *Tauberian theory: A Century of Developments*, Springer, 2004。和之前关于Cesàro求和的Tauber定理一样，最早也是由Hardy–Littlewood给出。

引理*1*：设$P(x)$是多项式且$P(0)=0$，则$\lim_{x\to 1-}\sum_{n=0}^{\infty}a_nP(x^n)=P(1)A$。

引理*1*的证明：设$P(x)=c_mx^m+\cdots+c_1x$，则

$$\lim_{x\to 1-}\sum_{n=0}^{\infty}a_nP(x^n)=\lim_{x\to 1-}\sum_{k=1}^{m}c_k\sum_{n=0}^{\infty}a_n(x^k)^n=\sum_{k=1}^{m}c_kA=P(1)A.$$

注意这里用到了有限和与无穷级数、极限都可交换，以及当$x\to 1-$时$x^k\to 1-$。

引理*2*：设$g(x)$在$[0,1]\setminus\left\{\frac{1}{2}\right\}$上连续，且在$\frac{1}{2}$处是跳跃间断点，则对任意$\varepsilon>0$，存在多项式$P(x)$使得$P(x)\geq g(x)$，且$\int_0^1|P-g|<\varepsilon$。

引理*2*的证明：容易找到连续函数$\varphi\colon[0,1]\to\mathbb{R}$使得$\varphi(x)\geq g(x)-\frac{\varepsilon}{3}$且$\int_0^1|\varphi-g|\leq\frac{\varepsilon}{3}$。根据Weierstrass逼近定理，存在多项式$Q$使得$\sup_{[0,1]}|Q-\varphi|<\frac{\varepsilon}{3}$，取$P:=Q+\frac{2\varepsilon}{3}$即可。

引理*3*：设$g(x)$在$[0,1]\setminus\left\{\frac{1}{2}\right\}$上连续，且在$\frac{1}{2}$处是跳跃间断点，则

$$\lim_{x\to 1-}(1-x)\sum_{n=1}^{\infty}x^ng(x^n)=\int_0^1 g(t)\,dt.$$

引理*3*的证明：两边作差，有

$$\int_0^1 g(t)\,dt-(1-x)\sum_{n=1}^{\infty}x^ng(x^n)=\int_x^1 g(t)\,dt+\sum_{n=1}^{\infty}\int_{x^{n+1}}^{x^n}(g(t)-g(x^n))\,dt.$$

因为$\frac{1}{2}$是跳跃间断点，对任意$\varepsilon>0$，存在$\delta>0$使得当s,t位于$\frac{1}{2}$同侧且$|s-t|<\delta$时$|g(s)-g(t)|<\varepsilon$。当$x\to 1$时，区间$[x^{n+1},x^n]$的长度不超过$1-x$，从而也就小于δ。此外，至多只有一个区间包含$\frac{1}{2}$，并且对应的积分不超过$2M(1-x)$，其中$M:=\sup_{[0,1]}|g|$。因此

$$\begin{aligned}\left|\int_0^1 g(t)\,dt-(1-x)\sum_{n=1}^{\infty}x^ng(x^n)\right|&\leq(1-x)M+\sum_{n=1}^{\infty}\varepsilon(x^n-x^{n+1})+2M(1-x)\\&=3M(1-x)+\varepsilon x.\end{aligned}$$

当$x\to 1-$时，右边的极限是ε，于是结论得证。

引理*4*：$\limsup_{x\to 1-}\sum_{n=0}^{\infty}a_nf(x^n)\leq A$。

引理*4*的证明：定义$g(x)=\frac{f(x)-x}{x(1-x)}$，则$g$是$[0,1]\setminus\{\frac{1}{2}\}$上的连续函数且$\frac{1}{2}$为跳跃间断点。根据引理2，对任意$\varepsilon>0$，存在多项式$P$使得$P(x)\geq g(x)$且$\int_0^1|P-g|<\varepsilon$。令$Q(x)=x+x(1-x)P(x)$，则$Q$也是多项式，满足$Q(0)=0$，$Q(1)=1$以及$Q(x)\geq f(x)$。

因为na_n有界，可设$na_n\geq -C\,(C>0)$。对于$x\in(0,1)$，注意到$|1-x|^n=|1-x|\cdot|1+$

$x + \cdots + x^{n-1}| \le n|1-x|$，所以

$$\sum_{n=0}^{\infty} a_n f(x^n) - \sum_{n=0}^{\infty} a_n Q(x^n) = -\sum_{n=1}^{\infty} a_n (Q(x^n) - f(x^n))$$
$$\le \sum_{n=0}^{\infty} \frac{C}{n}(Q(x^n) - f(x^n)) \le C\sum_{n=1}^{\infty} \frac{1-x}{1-x^n}(Q(x^n) - f(x^n)) =: C(1-x)\sum_{n=1}^{\infty} \phi(x^n).$$

其中$\phi(t) := \frac{Q(t)-f(t)}{1-t} = t(P(t) - g(t))$。由引理3，当$x \to 1-$时右边趋于

$$C\int_0^1 (P-g) \in (0, C\varepsilon).$$

由ε的任意性可知左边的上极限不超过0。结合引理1得

$$\limsup_{x\to 1-} \sum_{n=0}^{\infty} a_n f(x^n) \le \lim_{x\to 1-} \sum_{n=0}^{\infty} a_n Q(x^n) = Q(1)A = A.$$

引理*5*: $\liminf_{x\to 1-} \sum_{n=0}^{\infty} a_n f(x^n) \ge A$。

引理*5*的证明：用$-a_n$代替a_n，由引理4立即得到。

由引理4、5就完成了断言的证明，从而结论得证。 □

注. 还有很多广义求和的办法，比如*Borel*求和

$$\lim_{x\to\infty} e^{-x} \sum_{n=0}^{\infty} \frac{S_n}{n!} x^n.$$

最著名的一个广义求和的例子是（注意这是*Abel*不可求和的）

$$1 + 2 + 3 + 4 + \cdots = -\frac{1}{12}.$$

这来自于*Riemann* ζ函数在-1处的解析延拓，利用*(c)*的例子容易给出形式的解释。

数学分析II中将会介绍两种无穷级数的收敛性判别法，这里提前给出。

题 3.9. *(a)* (Abel判别法)如果$\sum_{n=1}^{\infty} a_n$收敛，b_n单调有界，证明：$\sum_{n=1}^{\infty} a_n b_n$收敛。

(b) (Dirichlet判别法)如果$\sum_{n=1}^{N} a_n$关于N有界，b_n单调趋于0，证明：$\sum_{n=1}^{\infty} a_n b_n$收敛。

证明. (a) 设$M = \sup |b_n| < \infty$。对任意$\varepsilon > 0$，存在$N > 0$使得当$m > n > N$时$|a_n + \cdots + a_m| < \varepsilon$。对于$k \ge n$，记$S_k = a_n + a_{n+1} + \cdots + a_k$，结合$b_n$的单调性得

$$\left|\sum_{k=n}^{m} a_k b_k\right| = \left|S_m b_m + \sum_{k=n}^{m-1} S_k (b_k - b_{k+1})\right| \le \varepsilon(|b_m| + \sum_{k=n}^{m-1} |b_k - b_{k+1}|) \le 3M\varepsilon.$$

由Cauchy收敛准则可知$\sum_{n=1}^{\infty} a_n b_n$收敛。

(b) 和(a)类似，利用$|S_n|<M$和$|b_m|+\sum_{k=n}^{m-1}|b_k-b_{k+1}|\to 0$，具体细节留作练习。 □

上面使用Abel变换都是利用b_n的单调性，下面一题用的是$|b_{n+1}-b_n|$的小性。

题 3.10. 设$f\colon[0,1]\to\mathbb{R}$是连续函数，如果$a_n\in\mathbb{R}$满足$\sum_{n=0}^{N}a_n$关于$N\in\mathbb{N}_0$有界，证明：

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n}a_kf\left(\frac{k}{n}\right)=0.$$

证明. 记$S_n=a_1+\cdots+a_n$，设$M=\sup|S_n|$。连续函数在有限闭区间上一致连续，对任意$\varepsilon>0$，存在$N>0$使得当$n>N$时$|f\left(\frac{k}{n}\right)-f\left(\frac{k+1}{n}\right)|<\varepsilon$，所以

$$\left|\sum_{k=0}^{n}a_kf\left(\frac{k}{n}\right)\right|=\left|S_nf(1)+\sum_{k=1}^{n-1}S_k\left(f\left(\frac{k}{n}\right)-f\left(\frac{k+1}{n}\right)\right)\right|\le M(|f(1)|+\varepsilon n).$$

两边除以n并取极限即可。 □

注. 本题很好地解释了*Abel*变换的作用。直观上看，让结论成立的有两种效应：*(1)*如果$a_n\ge 0$且$\sum_{n=0}^{\infty}a_n<\infty$，则$|\sum_{k=0}^{n}a_kf\left(\frac{k}{n}\right)|$不会超过$\sup|f|\cdot\sum_{n=0}^{\infty}a_n$，除以$n$之后趋于0；*(2)*如果$a_n$具有某种抵消作用，比如$a_n=(-1)^{n+1}$，则可以利用$f\left(\frac{k+1}{n}\right)-f\left(\frac{k}{n}\right)$的小性。条件正是这两种效应的叠加，但一般无法各自拆分出来，而*Abel*变换能够将其集中体现。

补充：度量空间

度量空间(metric space)是$\mathbb{R}$连同欧氏距离的一种推广。

定义 3.11 (度量空间)**.** 设X是非空集合，$d\colon X\times X\to\mathbb{R}$称为$X$上的度量，如果

(a) (正定性)对任意$x,y\in X$，有$d(x,y)\ge 0$，且$d(x,y)=0$当且仅当$x=y$；

(b) (对称性)对任意$x,y\in X$，有$d(x,y)=d(y,x)$；

(c) (三角形不等式)对任意$x,y,z\in X$，有$d(x,y)\le d(x,z)+d(y,z)$。

集合X连同其上的一个度量d构成的二元对(X,d)称为度量空间。

$\mathbb{R}$连同$d(x,y):=|x-y|$构成一个度量空间。另一个平凡的例子是

$$d(x,y):=\begin{cases}1, & x\ne y,\\ 0, & x=y.\end{cases}$$

这称为集合X上的离散度量。如果(X,d)是一个度量空间，则对任意非空子集Y，d在Y上的限制给出Y上的一个度量，从而$(Y,d|_{Y\times Y})$也是一个度量空间，称为X的子空间。对于度量空间的子集，如果没有特殊说明，都是取继承的度量。

定义 3.12 (收敛)**.** *设(X,d)是度量空间，若$x_n,x\in X$且$d(x_n,x)\to 0$，则称x_n收敛到x。*

度量空间中的收敛与数的收敛具有类似的性质。

题 3.13. *设(X,d)是度量空间，证明：*

(a) 如果$x_n\to x$且$x_n\to y$，则$x=y$。

(b) 如果$x_n\to x$，则对任意$\varepsilon>0$，存在$N>0$使得当$m,n>N$时$d(x_m,x_n)<\varepsilon$。

证明. (a) 对任意$\varepsilon>0$，存在$N>0$使得当$n>N$时$d(x_n,x)<\varepsilon$且$d(x_n,y)<\varepsilon$，所以

$$d(x,y)\le d(x_n,x)+d(x_n,y)\le 2\varepsilon.$$

由ε任意性得$d(x,y)=0$，从而$x=y$。

(b) 存在$N>0$使得当$n>N$时$d(x_n,x)<\frac{\varepsilon}{2}$，从而对任意$m,n>N$有

$$d(x_m,x_n)\le d(x_m,x)+d(x_n,x)<\varepsilon.$$ □

满足上面的(b)的点列称为Cauchy列，于是收敛列一定是Cauchy列。但一般的，Cauchy列未必收敛。比如取$X=(0,1)$，则$\frac{1}{n}$是Cauchy列但不收敛，直观上看是因为其“极限”0并不在X中。由此引进完备性的概念。一般的，任何度量空间都可以（以某种意义下唯一的方式）完备化；常见的方法是Cauchy列的等价类，类似于通过十进制小数定义实数。

定义 3.14 (完备性)**.** *设(X,d)是度量空间，如果X中的Cauchy列都收敛，则称X是完备度量空间(complete metric space)。*

回忆$\mathbb{R}$的子集称为闭集，当且仅当X的聚点都包含在X中。按照定义容易说明，如果X是$\mathbb{R}$的子集，带有继承自$\mathbb{R}$的度量，则X完备当且仅当X是闭集。

引进完备度量空间的一个好处是可以证明压缩映射原理。今后会看到除了$\mathbb{R}$之外，在分析学中还有很多自然的完备度量空间，于是这些结果可以应用在对应的模型中，这就揭开了泛函分析的一角。例如连续函数空间$C([a,b])$连同距离

$$d(f,g):=\max_{x\in[a,b]}|f(x)-g(x)|$$

构成完备度量空间，由此可以证明一阶常微分方程初值问题解的存在唯一性。

题 3.15. *(a)* (Banach不动点定理/压缩映射原理)设(X,d)是完备度量空间，若$f\colon X\to X$满足对任意$x,y\in X$都有$d(f(x),f(y))<cd(x,y)$，其中$c\in(0,1)$是常数，证明：f有唯一不动点。事实上，任取$x\in X$，则$x,f(x),f(f(x)),\ldots$收敛到f的唯一不动点。

(b) 如果将*(a)*的条件改为“当$x\neq y$时$d(f(x),f(y))<d(x,y)$”，是否f一定有不动点？

*(c)** 若(X,d)是紧致度量空间，$f\colon X\to X$满足当$x\neq y$时$d(f(x),f(y))<d(x,y)$，证明：f有唯一不动点。进一步，任取$x\in X$，则$x,f(x),f(f(x)),\ldots$收敛到f的唯一不动点。

证明. (a) 唯一性几乎是显然的：如果x,y都是不动点，则$d(x,y)=d(f(x),f(y))<cd(x,y)$，从而$d(x,y)=0$，所以$x=y$。下面验证存在性，任取$x_0\in X$，递归定义$x_n=f(x_{n-1})$，则$d(x_{n+1},x_n)\leq c^n d(x_1,x_0)$。于是对任意$N>0$，当$m>n>N$时

$$d(x_m,x_n)\leq\sum_{k=n}^{m-1}d(x_{k+1},x_k)\leq\sum_{k=n}^{m-1}c^k d(x_1,x_0)\leq\frac{c^N}{1-c}d(x_1,x_0).$$

这就说明x_n是Cauchy列，根据完备性可知$x_n\to x$。因为$d(x_n,f(x))<cd(x_{n-1},x)\to 0$，所以又有$x_n\to f(x)$，这就说明$x=f(x)$。

(b) 取$X=[1,\infty)$和$f(x)=x+\frac{1}{x}$即可，具体验证留作练习。

(c) 唯一性同上，下面先证明不动点的存在性。可以证明$x\mapsto d(x,f(x))$是连续函数，由X的紧性可知存在最小值点x_0。如果$x_0\neq f(x_0)$，则$d(f(x_0),f(f(x_0)))<d(x_0,f(x_0))$，与$x_0$是最小值点矛盾！故$d(x_0,f(x_0))=0$，即$f(x_0)=x_0$。

最后，任取$x_1\in X$，递归定义$x_{n+1}=f(x_n)$，需要证明$f(x_n)\to x_0$。事实上，由于

$$d(x_{n+1},x_0)=d(f(x_n),f(x_0))<d(x_n,x_0).$$

可知$d(x_n,x_0)$单减，从而存在极限$c\geq 0$。如果$c>0$，因为X是紧集，存在子列x_{n_k}收敛到某个极限$y\in X$，从而$d(y,x_0)=c$；结合条件可知$x_{n_k+1}=f(x_{n_k})$收敛到$f(y)$，从而$d(f(y),x_0)=c$，但这与

$$d(f(y),x_0)=d(f(y),f(x_0))<d(y,x_0)$$

矛盾！故$c=0$，即$x_n\to x_0$。 □

下面给出一般的度量空间上开集、闭集的概念，今后讨论连续性时开集会起到重要作用。

定义 3.16 (开集)**.** 设(X,d)是度量空间，集合$U\subset X$称为开集，如果对任意$x\in U$，存在$r>0$使得$B(x,r)\subset U$，其中$B(x,r):=\{y\in X\mid d(x,y)<r\}$是球心为$x$、半径为$r$的开球。

根据定义，$\mathbb{R}$中的开区间都是开集；但闭区间不是开集，因为端点不满足开集条件。值得说明的是任何开球$B(x,r)$都是开集，这是因为如果$y \in B(x,r)$，则$r - d(x,y) > 0$，由三角形不等式可知$B(y, r-d(x,y)) \subset B(x,r)$。

题 3.17. 证明开集具有如下性质：

(a) $\emptyset$和X是开集。

(b) 如果$U_\alpha\,(\alpha \in I)$是任意一族开集，则$\bigcup_{\alpha \in I} U_\alpha$也是开集。

(c) 如果$U_1, \ldots, U_n$是有限个开集，则$U_1 \cap \cdots \cap U_n$也是开集。

证明. (a) 按定义是显然的，唯一值得说明的是"空集的元素具有任何性质"。

(b) 若$x \in \bigcup_{\alpha \in I} U_\alpha$，则$x$属于某个$U_\beta\,(\beta \in I)$，从而存在$r > 0$使得$B(x,r) \subset U_\beta \subset \bigcup_{\alpha \in I} U_\alpha$。

(c) 如果$x \in U_1 \cap \cdots \cap U_n$，则存在$r_i > 0$使得$B(x, r_i) \subset U_i\,(i = 1, \ldots, n)$，取$r = \min\{r_1, \ldots, r_n\}$，则$B(x,r) \subset U_1 \cap \cdots \cap U_n$。 □

注. 由于$\bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$不是开集，所以*(c)*不能改成无限个开集。

一般的，非空集合X上的拓扑结构就是满足上面三条性质的集族$\mathcal{T}$，其中的集合称为开集。很多性质只与拓扑结构有关，而不依赖度量，但为了简便这份讲义中只考虑度量空间。比如收敛性可以用开集刻画[5]：$x_n \to x$当且仅当对任意包含x的开集U，都存在$N > 0$使得当$n > N$时$x_n \in U$。今后还会说明连续性也可以用开集刻画。即使在分析学中，研究一般的拓扑空间也是有意义的，比如泛函分析中会遇到不可度量化的局部凸拓扑线性空间，所谓分布（即物理学家所说的广义函数）的理论就是最经典的例子。

定义 3.18 (闭集)**.** 设(X,d)是度量空间，集合$F \subset X$称为闭集，如果F^C是开集。利用*De Morgan*律以及开集的性质，可知闭集满足如下性质：

(a) $\emptyset$和X是闭集。

(b) 如果$F_\alpha(\alpha \in I)$是任意一族闭集，则$\bigcap_{\alpha \in I} F_\alpha$也是开集。

(c) 如果$F_1, \ldots, F_n$是有限个闭集，则$F_1 \cup \cdots \cup F_n$也是闭集。

回忆之前对$\mathbb{R}$中的闭集给出了另一个定义，即包含所有聚点。容易证明两种定义等价。事实上，在任何度量空间中，都可以仿照$\mathbb{R}$的情形定义聚点，并且闭集等价于包含所有聚点。

[5]在不可度量化的拓扑空间中，点列的收敛性可能非常不自然，比如一列点可以同时收敛到多个不同的极限。因此开集语言对于研究一般的空间至关重要。尽管点列极限在一般的拓扑空间中用途有限，但至少在度量空间中点列极限与开集的效力几乎是相同的。点列收敛在一般的拓扑空间中应当推广为网(net)的收敛。

定义 3.19 (闭包). 设(X,d)是度量空间，对任意集合A，定义A的闭包为

$$\overline{A} := \bigcap_{\text{闭集}F\supset A} F.$$

则$\overline{A}$是包含A的最小闭集。

这里对“包含A的最小闭集”需要进行说明。因为闭集的任意交还是闭集，所以$\overline{A}$是闭集。任何包含A的闭集都在右边的交里。因此这个称呼是准确的。注意闭包依赖于大空间X，即如果$A\subset Y\subset X$，则A在Y中的闭包（将Y视为子度量空间）通常不等于A在X中的闭包。事实上，容易证明A在Y中的闭包是$\overline{A}\cap Y$。特别的，如果Y是闭集，则A在Y中的闭包就是$\overline{A}$。

根据先前对聚点的说明，闭包其实就是A添上所有聚点形成的集合。特别的，稠密集的定义可以更新为闭包等于全集。更进一步，如果$A\subset Y\subset X$，则可以将A在Y中稠密定义为$Y\subset\overline{A}$，或者说A在Y中的闭包是Y。当然，可以用最朴素（也最繁琐）的语言去说：对任意开集U，如果$U\cap Y\neq\emptyset$，则$U\cap A\neq\emptyset$。

定义 3.20 (可分性). 设(X,d)是度量空间，如果X存在可数稠密子集，则称X是可分度量空间*(separable metric space)*。

$\mathbb{R}$就是可分度量空间，因为全体有理数构成可数稠密子集。有趣的是，虽然无理数集不可数，但同样可以找到可数个无理数在$\mathbb{R}$中稠密（构造类似于下一题的(a)）。

题 3.21. *(a)** 证明：可分度量空间的子空间也是可分的。

(b) 设$A\subset\mathbb{R}$，证明：A是某个数列的聚点集当且仅当A是闭集。

证明. (a) 设(X,d)有可数稠密集$\{x_m\mid m\in\mathbb{N}\}$。对任意$m,n\in\mathbb{N}$，考虑开球$B\left(x_m,\frac{1}{n}\right)$是否与$Y$相交，如果是则取出一个交点$y_{m,n}\in Y$。断言这些$y_{n,m}$在$Y$中稠密。

事实上，对任意与Y交非空的开集U，由U是开集知存在$y\in Y$和$r>0$使得$B(y,r)\subset U$。取$n>\frac{1}{2r}$，由$\{x_n\}$的稠密性可知存在$m\in\mathbb{N}$使得$d(x_m,y)<\frac{1}{n}$，从而

$$y\in B\left(x_m,\frac{1}{n}\right)\subset B(y,\frac{1}{r})\subset U.$$

特别的，$B\left(x_m,\frac{1}{n}\right)$与$Y$交非空，根据定义取出过一个点$y_{m,n}$，从而$y_{m,n}\in Y\cap U$。

(b) 之前已经说明过聚点集一定是闭集。反过来，如果A是闭集，则根据(a)可知存在可数个点$x_1,x_2,\ldots$使得$\{x_n\mid n\in\mathbb{N}\}$的闭包是$A$。因为$\mathbb{N}\times\mathbb{N}$是可数集，可以构造数列$a_1,a_2,\ldots$，使得每个$x_n$都出现无穷多次，于是每个$x_n$都在数列的聚点集中，从而聚点集包含$A$。反过来，$A$中的数列的聚点一定属于$A$，所以$a_n$的聚点集等于$A$。 $\square$

第4讲：连续函数（上）

因此，解构实践也称为“文本骚扰”，或者“对抗性阅读”，其目的是呈现文本内部的矛盾或者不连贯，呈现隐藏于表面的连续之下的不连续性。

——彼得·巴里《理论入门：文学与文化理论导论》

这一讲主要关于连续极限以及连续函数的定义。连续极限与离散极限在方法上并没有太多不同，只要熟练运用诸如$\lim_{x\to 0}\frac{\sin x}{x}=1$、$\lim_{x\to 0}\frac{e^x-1}{x}=1$等结论。所谓运用，包括对这些公式进行换元得到的常见结果，比如说对于$a>0$有

$$\lim_{x\to 0}\frac{a^x-1}{x}=\lim_{x\to 0}\frac{e^{x\log a}-1}{x}=\log a.$$

此外，给出了常见不等式的初等证明，可以与今后利用导数或凸函数性质的证明进行对比。补充内容中引进了一致收敛的概念，并证明了连续函数的一致收敛极限仍然连续。

例题

例 4.1. 求下列极限：

(a) $\lim_{n\to\infty}\sin\left(\sqrt{n^2+1}\pi\right)$。

(b) $\lim_{n\to\infty}n\left(\frac{\pi}{2}-\arccos\frac{1}{n}\right)$。

(c) $\lim_{x\to\infty}\frac{\sqrt{x+\sqrt{x}}}{x+\sqrt{x}+1}$。

(d) $\lim_{x\to\infty}\left(\sqrt{x+\sqrt{x+\sqrt{x}}}-\sqrt{x}\right)$。

(e) $\lim_{x\to 0}\frac{a^x-b^x}{c^x-d^x}$，其中$a,b,c,d>0$且$c\neq d$。

解. (a) 利用$\sin n\pi=0$可知

$$|\sin\left(\sqrt{n^2+1}\pi\right)|=\left|\sin\left((\sqrt{n^2+1}-n)\pi\right)\right|=\left|\sin\frac{\pi}{\sqrt{n^2+1}+n}\right|\to 0.$$

所以$\lim_{n\to\infty}\sin\left(\sqrt{n^2+1}\pi\right)=0$。

(b) 当$x\in(0,1)$时$\frac{\pi}{2}-\arccos x=\arcsin x$，再换元$y=\arcsin\frac{1}{n}$，得

$$\lim_{n\to\infty}n\left(\frac{\pi}{2}-\arccos\frac{1}{n}\right)=\lim_{t\to 0}\frac{\sin t}{t}=1.$$

(c) 容易看出分子$\sim\sqrt{x}$，而分母$\sim x$，所以极限是0。

(d) 利用平方差公式，有

$$\lim_{x\to\infty}\left(\sqrt{x+\sqrt{x+\sqrt{x}}}-\sqrt{x}\right)=\lim_{x\to\infty}\frac{\sqrt{x+\sqrt{x}}}{\sqrt{x+\sqrt{x+\sqrt{x}}}+\sqrt{x}}=\frac{1}{2}.$$

(e) 将b^x和d^x提出来，有

$$\lim_{x\to 0}\frac{a^x-b^x}{c^x-d^x}=\lim_{x\to 0}\frac{\left(\frac{a}{b}\right)^x-1}{x}\cdot\frac{x}{\left(\frac{c}{d}\right)^x-1}=\frac{\log a-\log b}{\log c-\log d}.$$

□

例 4.2. *(a)* 设$f\colon(0,\infty)\to\mathbb{R}$满足$\lim_{x\to\infty}xf(x)=a(a\in\mathbb{R})$，证明：

$$\lim_{x\to\infty}(1+f(x))^x=e^a.$$

换一种写法即

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

(b) 求$\lim_{n\to\infty}\left(\cos\frac{\pi}{\sqrt{n}}\right)^n$。

(c) 求$\lim_{x\to\infty}\left(\sin\frac{1}{x}+\cos\frac{1}{x}\right)^x$。

证明. (a) 因为$xf(x)\to a$，所以$f(x)\to 0$。于是当x充分大时可以取对数，只需证

$$\lim_{x\to\infty}x\log(1+f(x))=a.$$

事实上，因为$\lim_{t\to 0}\frac{\log(1+t)}{t}=1$，所以上式由$xf(x)\to a$保证。

(b) $\lim_{n\to\infty}\left(\cos\frac{\pi}{\sqrt{n}}\right)^n=\lim_{n\to\infty}\left(1-\frac{\pi^2}{2n}+o\left(\frac{1}{n}\right)\right)^n=e^{-\frac{\pi^2}{2}}$。

(c) $\lim_{x\to\infty}\left(\sin\frac{1}{x}+\cos\frac{1}{x}\right)^x=\lim_{x\to\infty}\left(1+\frac{1}{x}+o\left(\frac{1}{x}\right)\right)^x=e$。

□

例 4.3. *(a) Riemann*函数$R(x)\colon(0,1)\to\mathbb{R}$定义为

$$R(x):=\begin{cases}0, & x\notin\mathbb{Q},\\ \frac{1}{n}, & x=\frac{m}{n},\ m,n\in\mathbb{N},\ \gcd(m,n)=1.\end{cases}$$

证明：R在$(0,1)\setminus\mathbb{Q}$上连续、$(0,1)\cap\mathbb{Q}$上间断。

(b) 设$N\subset\mathbb{R}$是可数集，证明：存在$f\colon\mathbb{R}\to\mathbb{R}$使得间断点集恰好等于$N$。

*(c)** 设$F_n\subset\mathbb{R}$是一列闭集，证明：存在$f\colon\mathbb{R}\to\mathbb{R}$使得间断点集恰好等于$\bigcup_{n\in\mathbb{N}}F_n$。

证明. (a) 对任意$n \in \mathbb{N}$，只有有限个点处$R(x) \geq \frac{1}{n}$。由此可知$R(x)$在每一点处的极限都是0，从而只在无理点处连续。

(b) 记$N = \{x_n \mid n \in \mathbb{N}\}$。仿照Riemann函数的情况，容易证明

$$f(x) = \begin{cases} 0, & x \notin \{x_n \mid n \in \mathbb{N}\}, \\ \frac{1}{n}, & x = x_n, \end{cases}$$

满足要求。[应该注意的是，如果定义$f(x_n) = 1$，则在x_n的聚点处f也不连续。]

(c) 不妨设F_n单调递增，否则用$F_1 \cup \cdots \cup F_n$代替F_n即可。先前证明过可以取一列点$x_{1k} \in F_1\ (k \in \mathbb{N})$在$F_1$中稠密。接下来，可以添加$F_2$中的一列点$x_{2k}$使得$\{x_{1k}, x_{2k} \mid k \in \mathbb{N}\}$在$F_2$中稠密；进一步，可以不妨设这些$x_{1k}, x_{2k}$互异。一般的，可以递归构造可数个互异的点$x_{nk} \in F_n\ (n, k \in \mathbb{N})$使得对任意$n \in \mathbb{N}$都有$\{x_{mk} \mid 1 \leq m \leq n,\ k \in \mathbb{N}\}$是$F_n$中的稠密集。

现在定义

$$f(x) = \begin{cases} \frac{1}{n}, & x = x_{nk}, \\ 0, & x \notin \{x_{nk} \mid n, k \in \mathbb{N}\}. \end{cases}$$

仿照(a)容易证明（唯一的区别是此时满足$f(x) \geq 1/n$的集合不再是有限个点，而是包含于闭集F_n），如果$x_0 \notin \bigcup_{n \in \mathbb{N}} F_n$，则$\lim_{x \to x_0} f(x) = 0 = f(x_0)$，所以$f$在$(\bigcup_{n \in \mathbb{N}} F_n)^C$上连续。下面只需证明$f$在任意$x_0 \in F_n$处间断。事实上，因为$f$只在可数个点处非0，所以$\liminf_{x \to x_0} f(x) = 0$；又因为$x_0 \in F_n$，存在一列$x_{m_j k_j}\ (j \in \mathbb{N})$满足$1 \leq m_j \leq n$且$x_{m_j k_j} \to 0$，结合$f(x_{m_j k_j}) = 1/m_j \geq 1/n$可知$\limsup_{x \to x_0} f(x) \geq 1/m$；故$f$在$x_0$处间断。 □

注. *(b)还可以要求f严格单调递增。事实上，记Heaviside函数$H := \chi_{[0,\infty)}$，则可以取*

$$f(x) = x + \sum_{n=1}^{\infty} 2^{-n} H(x - x_n).$$

注. *(c)说明任何可数个闭集的并（通常记为F_σ集）都可以是某个函数的间断点集，今后会借助振幅的概念证明反过来的事实：任何函数的间断点集都是F_σ集。此外，在证明Baire纲定理之后容易说明$\mathbb{R} \setminus \mathbb{Q}$不是$F_\sigma$集。因此不可能存在一个函数在有理点连续、而在无理点间断（注意这与Riemann函数恰恰相反）。*

练习题

题 4.4. 设$x_1,\dots,x_n>0$，证明：

$$\lim_{p\to 0}\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}}=\sqrt[n]{x_1\cdots x_n},$$

$$\lim_{p\to \infty}\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}}=\max\{x_1,\dots,x_n\},$$

$$\lim_{p\to -\infty}\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}}=\min\{x_1,\dots,x_n\}.$$

证明. 为了求第一个极限，两边取对数，利用$\lim_{x\to 1}\frac{\log x}{x-1}=1$，得

$$\begin{aligned}\lim_{p\to 0}\frac{1}{p}\log\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)&=\lim_{p\to 0}\frac{1}{p}\left(\frac{x_1^p+\cdots+x_n^p}{n}-1\right)\\&=\lim_{p\to 0}\frac{1}{n}\sum_{i=1}^{n}\frac{x_i^p-1}{p}=\frac{1}{n}\sum_{i=1}^{n}\log x_i.\end{aligned}$$

下面验证第二个极限，记$M=\max\{x_1,\dots,x_n\}$，则对$p>0$有

$$\left(\frac{M^p}{n}\right)^{\frac{1}{p}}\le\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}}\le M.$$

利用夹逼定理即证。第三个极限是类似的，具体细节留作练习。 □

题 4.5. *(a)* 设$f\colon(0,\infty)\to\mathbb{R}$满足$f(0+)=0$和$\lim_{x\to 0+}\frac{f(2x)-f(x)}{x}=0$，证明：

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

(b) 设$f\colon\mathbb{R}\to\mathbb{R}$在0处连续，且$\lim_{x\to 0}\frac{f(2x)-f(x)}{x}$收敛，证明：$f$在0处可导，且

$$f'(0)=\lim_{x\to 0}\frac{f(2x)-f(x)}{x}.$$

证明. (a) 对任意$\varepsilon>0$，存在$\delta>0$使得当$0<x<\delta$时$|f(2x)-f(x)|<\varepsilon x$。于是

$$\begin{aligned}|f(x)|&=\left|\sum_{n=0}^{N-1}(f(2^{-n}x)-f(2^{-n-1}x))+f(2^{-N}x)\right|\\&\le\sum_{n=0}^{N-1}\varepsilon\cdot 2^{-n-1}x+|f(2^{-N}x)|\le\varepsilon x+|f(2^{-N}x)|,\end{aligned}$$

其中N是任意正整数。令$N\to\infty$，利用$f(0+)=0$得$|f(x)|\le\varepsilon x$。结合ε任意性即证。

(b) 记$A=\lim_{x\to 0}\frac{f(2x)-f(x)}{x}$，考虑$g(x):=f(x)-f(0)-Ax$，则$g(x)$和$g(-x)$都满足(a)的条件，于是$g$在0处的左、右导数都是0，也就是说$f$在0处可导且$f'(0)=A$。 □

下一题的最后一问需要用到线性代数的知识，即任何域F上的线性空间V一定有一组基，也就是说存在$\{v_\alpha \mid \alpha\in I\}$使得任何向量$v\in V$可以唯一表示成$v=\sum_{\alpha\in I}x_\alpha v_\alpha$，且只有有限个$x_\alpha\neq 0$。标准的证明用到Zorn引理：极大线性无关组一定构成V的一组基。

题 4.6. *设$f\colon\mathbb{R}\to\mathbb{R}$满足Cauchy方程，即*

$$f(x+y)=f(x)+f(y),\quad \forall x,y\in\mathbb{R}.$$

(a) 证明：对任意$q\in\mathbb{Q}$，有$f(qx)=f(x)\,(\forall x\in\mathbb{R})$。

(b) 若f在某一点处连续，证明：$f(x)=xf(1)$。

(c) 若f在某一点的邻域上有界，证明：$f(x)=xf(1)$。

(d) 若f的图像在$\mathbb{R}^2$中不稠密，证明：$f(x)=xf(1)$。*

(e) 证明：若不添加额外条件，则未必有$f(x)=xf(1)$成立。*

证明. (a) 令$x=y=0$，可知$f(0)=0$；再令$y=-x$，可知$f(-x)=-f(x)$。归纳可知$f(nx)=nf(x)$，再用$\frac{x}{n}\,(n\neq 0)$代替x可知$f(x)=nf\left(\frac{x}{n}\right)$。由此容易证明结论。

(b) 显然(d)⇒(c)⇒(b)，但证明难度不同，因此分别给出。若f在x_0的邻域上连续，则对任意$\varepsilon>0$，存在$\delta>0$，使得当$|x_0-y|<\delta$时$|f(x_0-y)|=|f(x_0)-f(y)|<\varepsilon$。也就是说，对任意$|x|<\delta$有$|f(x)|<\varepsilon$。对任意$x\in\mathbb{R}$，取一列有理数$q_n\in\mathbb{Q}$收敛到$x$，则有

$$|f(x)-xf(1)|=\lim_{n\to\infty}|f(x)-f(q_n)|=\lim_{n\to\infty}|f(x-q_n)|<\varepsilon.$$

由ε任意性得$f(x)=xf(1)$。

(c) 由(b)的方法可知，存在$M>0$使得当$|x|<\delta$时$|f(x)|<2M$。对任意$\varepsilon>0$，取$n>\frac{2M}{\varepsilon}$，则当$|x|<\frac{\delta}{n}$时

$$|f(x)|=\frac{|f(nx)|}{n}<\frac{2M}{n}<\varepsilon.$$

这就说明f在0处连续，从而由(b)的结论得证。

(d) 若否，则存在$x\in\mathbb{R}\,(x\neq 0)$，使得$v:=(x,f(x))$与$w:=(1,f(1))$不平行，从而$\mathbb{R}v\oplus\mathbb{R}w=\mathbb{R}^2$。另一方面，对任意$p,q\in\mathbb{Q}$有$f(px+q)=f(px)+f(q)=pf(x)+qf(1)$，所以$f$的图像包含$\mathbb{Q}v\oplus\mathbb{Q}w$，而后者在$\mathbb{R}v\oplus\mathbb{R}w$中稠密。

(e) 由(a)可知f满足Cauchy方程等价于f是$\mathbb{Q}$-线性映射。取$\mathbb{R}$作为$\mathbb{Q}$-线性空间的一组基，并取出一个元素α，定义$f(\alpha)=1$并在这组基的其他元素处取0，然后线性延拓到整个$\mathbb{R}$上，则f是Cauchy方程的非平凡解。 □

注. *(e)中的构造是抽象的，如果取$\alpha=1$，则$f(x)=x$对$x\in\mathbb{Q}$成立，但并不是说在无理数处的取值都是0。这是因为如果β是这组基中的另一个元素，则对任意$x\in\mathbb{Q}$也有$f(\beta+x)=x$。*

注. *还可以构造Cauchy方程的解满足f在任何非空开区间上的值域都是$\mathbb{R}$（特别的，这样的f具有介值性质）。[6]具体来说，取$\{\alpha_i\mid i\in I\}$是$\mathbb{R}$在$\mathbb{Q}$上的一组基，任取$i_0\in I$，则存在双射$\tau\colon I\setminus\{i_0\}\to I$（这里用到$I$是无限集）；定义$f(\alpha_{i_0})=0$以及$f(\alpha_i)=\alpha_{\tau(i)}\,(i\neq i_0)$，并线性延拓为Cauchy方程的解；注意到任何数都可以通过加上α_{i_0}的有理数倍平移到给定开区间中，于是容易验证f满足所需的性质。*

注. *还可以考虑乘法版本的Cauchy方程，即$f\colon(0,\infty)\to\mathbb{R}$满足$f(xy)=f(x)f(y)\,(\forall x,y>0)$，其与算子半群有天然的联系。尽管取对数可以化归到加法版本的Cauchy方程（容易证明$f(x)$如果有零点则恒为0，而无零点则恒大于0），但如果推广为取值在一般的Banach代数的情况（比如$f\colon(0,\infty)\to GL(n,\mathbb{R})$），则需要另外的证明，今后将利用定积分给出。*

题 4.7*. *设P是n次多项式，证明：P可以写成$n+1$个周期函数之和。*

证明. 任取$n+1$个在$\mathbb{Q}$上线性无关的实数$\xi_1,\cdots,\xi_{n+1}$，并将其扩充成$\mathbb{R}$在$\mathbb{Q}$上的一组基。则$x\in\mathbb{R}$可以唯一表示成$r+q_1\xi_1+\cdots+q_{n+1}\xi_{n+1}$，其中$q_i\in\mathbb{Q}$，$r$由这组基中其他元素生成。考虑$f(x)=f(r+q_1\xi_1+\cdots+q_{n+1}\xi_{n+1})$，将其视为关于$r,\xi_1,\ldots,\xi_{n+1}$的$n$次多项式，展开之后一定有某一个$\xi_i$没有出现。将这些项分成$P=f_1+\cdots+f_{n+1}$，使得$f_i$不含$\xi_i$，则$f_i$以$\xi_i$为周期（这是因为$x+\xi_i$的表达式只是让$q_i$变成$q_i+1$而其他不变）。 □

注. *考虑n阶差分可知n次多项式不能写成n个周期函数之和。*

间断点通常按照左右极限是否都存在分为第一、二类，第一类间断点又按照左右极限是否相等细分为跳跃、可去。[7]下一题说明第一类间断点至多可数，并且可去间断点恰如其名。

题 4.8. *(a) 证明：$\mathbb{R}$中两两不交的开集最多只有可数个。*

(b) 设$f\colon\mathbb{R}\to\mathbb{R}$单调，证明：$f$只有至多可数个间断点。

(c) 证明：任何函数$f\colon\mathbb{R}\to\mathbb{R}$的第一类间断点只有至多可数个。*

[6]感谢浮生群群主笑姐姐指出这一点。

[7]新教材将可去间断点称为第三类间断点，而第一类间断点仅仅包含跳跃间断点。这份讲义中采用更常用的分类，否则要说“第一类和第三类间断点满足某某性质”实在奇怪。

(d) 设$f\colon \mathbb{R}\to\mathbb{R}$的间断点集为$A$，其中可去间断点集为$B\subset A$，定义函数

$$g(x)=\begin{cases} f(x), & x\notin B,\\ \lim_{y\to x} f(y), & x\in B.\end{cases}$$

证明：g的间断点集包含于$A\setminus B$。

证明. (a) 这是因为非空开集都包含有理数，而$\mathbb{Q}$是可数集。[严格来讲用到了选择公理。]

(b) 不妨设f单增。根据单调有界收敛原理，f在x处间断当且仅当$f(x-)\neq f(x+)$。对于每个间断点，考虑开区间$(f(x-),f(x+))$，易见它们两两不交，从而只能有可数个。

(c) 将第一类间断点集分解成

$$\left(\bigcup_{n=1}^{\infty}\left\{f(x)>f(x+)+\frac{1}{n}\right\}\right)\cup\left(\bigcup_{n=1}^{\infty}\left\{f(x)<f(x+)-\frac{1}{n}\right\}\right)$$
$$\cup\left(\bigcup_{n=1}^{\infty}\left\{f(x)>f(x-)+\frac{1}{n}\right\}\right)\cup\left(\bigcup_{n=1}^{\infty}\left\{f(x)<f(x-)-\frac{1}{n}\right\}\right).$$

这是可数个集合的并，只需证明每个都是可数集。又只需证明第一类集合

$$\{x\in\mathbb{R}\mid f(x)>f(x+)+\varepsilon\},$$

对任意$\varepsilon>0$是可数集，用$-f(x)$，$f(-x)$，$-f(-x)$替换$f(x)$可得后三类的情况。

若$f(x)>f(x+)+\varepsilon$，则存在$\delta>0$，使得在$I_x:=(x,x+\delta)$上恒有$f(x)>f(t)+\varepsilon$以及$|f(s)-f(t)|<\varepsilon\,(\forall s,t\in I_x)$。断言这些$I_x$是不交的，从而只能有可数个。事实上，如果$I_x$与$I_y$交非空，则$x\in I_y$或$y\in I_x$。不妨设是$x\in I_y$，取$t\in I_x\cap I_y$，则$t\in I_x$保证了$f(x)>f(t)+\varepsilon$，而$x,t\in I_y$又导致了$|f(x)-f(t)|<\varepsilon$，矛盾！这就证明了结论。

(d) 首先断言对任意$x\in\mathbb{R}$都有

$$\limsup_{y\to x} g(x)\le\limsup_{y\to x} f(x)\quad 和\quad \liminf_{y\to x} g(x)\ge\liminf_{y\to x} f(x).$$

根据对称性只需证明前者。事实上，对任意子列$y_n\to x\,(y_n\neq x)$，如果$y_n\notin B$，定义$z_n=y_n$；而如果$y_n\in B$，根据构造方式可知存在$z_n\neq x$使得$|y_n-z_n|<1/n$且$f(z_n)\ge g(y_n)-1/n$。于是$z_n\to x\,(z_n\neq x)$且

$$\limsup_{n\to\infty} f(y_n)\le\limsup_{n\to\infty}(f(z_n)+1/n)=\limsup_{n\to\infty} f(z_n)\le\limsup_{z\to x} f(z).$$

结合y_n的任意性即证。最后，对任意$x\notin A\setminus B$，根据断言容易推出

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

而这根据定义恰好等于$g(x)$。故g在$A \setminus B$上连续。 □

注. *Riemann*函数给出了具有可数个第一类间断点的例子。一般的，任给可数个实数x_n，例题中构造了恰好以这些点为第一类间断点的函数。

注. *(d)*中函数g间断点集可能严格包含于$A \setminus B$。比如考虑f在全体$1/n\,(n \in \mathbb{N})$处等于1、其他位置为0，则f的第二类间断点0不再是g的间断点。

题 4.9 *. 记$R\colon (0,1) \to \mathbb{R}$是*Riemann*函数。证明：存在一列连续函数$f_n\colon (0,1) \to \mathbb{R}$，使得对任意$x \in (0,1)$都有$\lim_{n\to\infty} f_n(x) = R(x)$。

证明. 对任意$n \in \mathbb{R}$，设$x_1, \ldots, x_k$是所有满足$R(x) \ge \frac{1}{n}$的点，取$\delta < \frac{1}{n}$充分小，使得$(x_i - \delta, x_i + \delta) \subset (0,1)$两两不交，定义$f$在这些$(x_i - \delta, x_i + \delta)$之外是0、在$[x_i - \delta, x_i + \delta]$上的图像是连接$(x_i - \delta, 0)$、$(x_i, R(x_i))$和$(x_i + \delta, 0)$的两段折线。易见$f_n$是连续函数，下证$f_n$逐点收敛到$R$。

如果$y \in (0,1) \cap \mathbb{Q}$，则存在$N > 0$使得当$n > N$时$R(y) \ge \frac{1}{n}$，从而根据定义有$f_n(y) = R(y)$。而如果$y \notin \mathbb{Q}$，则对任意$m \in \mathbb{N}$，满足$R(x) \ge \frac{1}{m}$的只有有限个异于$y$的点，这些点在每个$f_n\,(n \ge m)$中对应的小区间的长度$\le \frac{1}{n}$，因此当$n$充分大时不包含$y$；此外，除了这些点对应的小区间之外，总有$f_n(x) \le \frac{1}{m}$，所以$\limsup_{n\to\infty} f_n(y) \le \frac{1}{m}$，由$m$的任意性可知$\lim_{n\to\infty} f_n(y) = 0 = R(y)$。 □

注. 连续函数列的逐点极限被称为*Baire*第一类函数。一般的，*Baire*第n类函数的逐点极限称为*Baire*第$n+1$类函数。今后会用*Baire*纲定理证明，*Baire*第一类函数的连续点集稠密。特别的，*Dirichlet*函数$D(x) = \chi_{\mathbb{Q}}(x)$不是*Baire*第一类函数；事实上，$D(x)$是*Baire*第二类函数：

$$D(x) = \lim_{m\to\infty} \lim_{n\to\infty} |\cos(m!x\pi)|^n.$$

下面将用到“连续性方法”，尽管对于这个名称并没有一个明确的定义，但凡使用连续性（尤其是某种延拓性）的证明技巧都可称为连续性方法。这方面的最简单例子可能是有限覆盖定理的一种证明：若$U_\alpha(\alpha \in I)$是$[a,b]$的开覆盖，考虑集合

$$X := \{x \in [a,b] \mid [a,x]\text{可以被有限个}U_\alpha\text{覆盖}\}.$$

则可以证明：(1)$X \ne \emptyset$（因为$a \in x$），于是$\sup X \in [a,b]$；(2)$\sup X \in X$（因为存在某个U_α包含x，从而由开集定义可知覆盖了$(x-\delta, x]$，而$[a, x-\delta]$已被有限覆盖）；(3)如果$x \in X$且$x < b$，则存在$\delta > 0$使得$x + \delta \in X$（证明类似于(2)）。这三点推出$b = \sup X \in X$，从而$[a,b]$具有有限覆盖。在具体问题中，集合X的构造往往具有技巧性，更多例子今后还会遇到。

题 4.10. 设$f\colon \mathbb{R} \to \mathbb{R}$连续。如果对任意$x \in \mathbb{R}$，以及任意$\delta > 0$，都存在$y \in (x - \delta, x)$使得$f(y) \le f(x)$，证明：$f(x)$单调递增。

证明. 只需证明对任意$a<b$，都有$f(a)\le f(b)$。记

$$X := \{x \in [a,b] \mid f(x) \le f(b)\}.$$

由连续性可知X是闭集。又因为$b \in X$，所以X非空。记$c := \inf X$，则$c \in [a,b] \cap X$。如果$c > a$，则根据条件存在$d \in (c,a)$使得$f(d) \le f(c) \le f(b)$，从而$d \in X$，与$c = \inf X$矛盾！故$c=a$，即$f(a) \le f(b)$。 □

常见不等式

这一节给出若干常见不等式的初等证明。今后还可以用导数的方法给出另外的证明。为了节省篇幅，没有对取等条件进行讨论，但容易从证明中观察得出。

题 4.11. *(a)* (均值不等式)设$x_1,\ldots,x_n \ge 0$，证明：

$$\frac{x_1+\cdots+x_n}{n} \ge \sqrt[n]{x_1\cdots x_n}.$$

*(b)** 设$f\colon \mathbb{R}\to\mathbb{R}$满足

$$f\left(\frac{x+y}{2}\right) \le \frac{f(x)+f(y)}{2}, \quad \forall x,y \in \mathbb{R}.$$

证明：对任意$p \in (0,1)\cap\mathbb{Q}$，都有

$$f(px+(1-p)y) \le pf(x)+(1-p)f(y), \quad \forall x,y\in\mathbb{R}.$$

(c) (Young不等式)设$x,y\ge 0$，$p,q>0$且$\frac{1}{p}+\frac{1}{q}=1$，证明：

$$\frac{x^p}{p}+\frac{y^q}{q} \ge xy.$$

(d) (Bernoulli不等式)设$\alpha \ge 1$且$x \ge -1$，证明：

$$(1+x)^\alpha \ge 1+\alpha x.$$

证明. (a) 这个方法称为反向归纳，首先对$k\in\mathbb{N}$归纳，容易看出结论对$n=2^k$成立。一般的，设$2^{k-1}<n<2^k$，取$x_{n+1}=\cdots=x_{2^k}=\frac{x_1+\cdots+x_n}{n}$，代入$2^k$的不等式可得

$$\frac{x_1+\cdots+x_n}{n} \ge \sqrt[2^k]{x_1\cdots x_n\left(\frac{x_1+\cdots+x_n}{n}\right)^{2^k-n}}.$$

化简立即得到n的情况。

(b) 还是利用反向归纳的技巧，可以证明对任意$x_1,\dots,x_n\in\mathbb{R}$，都有（具体细节留作练习）

$$f\left(\frac{x_1+\cdots+x_n}{n}\right)\le\frac{f(x_1)+\cdots+f(x_n)}{n}.$$

设$p=\frac{m}{n}\,(m<n)$，则在上式中取$x_1=\cdots=x_m=x$和$x_{m+1}=\cdots=x_n=y$即可。

(c) 根据连续性，只需证$p\in\mathbb{Q}$的情况。设$p=\frac{m}{n}(m<n)$，则等价于证明

$$mx^{\frac{m}{n}}+(n-m)y^{\frac{n-m}{m}}\ge nxy.$$

将左边视为m个$x^{\frac{m}{n}}$再加上$n-m$个$y^{\frac{n-m}{n}}$，由均值不等式即证。

(d) 记$y=x+1>0$，则等价于证明

$$y^{\alpha}+(\alpha-1)\ge\alpha y.$$

取$p=\alpha$和$q=\frac{\alpha}{\alpha-1}$并用Young不等式即可。 □

注. *(b)*的条件可以称为“中点凸”，而所谓凸函数是指$f(px+(1-p)y)\le pf(x)+(1-p)f(y)$ $(\forall p\in[0,1])$。从*(b)*的结论可以看出，中点凸离真正的凸函数还差了连续性，比如*Cauchy*方程的非连续解是中点凸的但不是凸函数。教材今后会证明凸函数（除了区间端点处）必然连续。

注. 在*Bernoulli*不等式中，如果$x\ge 0$且$\alpha\ge 1$是有理数，则可以通过二项式定理和比较系数得出。至少在$\alpha\in\mathbb{N}$时，这是*Bernoulli*不等式的动机。而α是无理数的情况可以用有理数逼近。

下面需要用到齐次不等式中的常见技巧，即可以不妨设某个齐次量为一常数。具体来讲，假设需要证明$F(x_1,\dots,x_n)\ge 0$，其中$x_1,\cdots x_n>0$。如果F是k次齐次函数，即对任意$\lambda>0$有$F(\lambda x_1,\dots,\lambda x_n)=\lambda^kF(x_1,\dots,x_n)$，则对自变量同时乘上非负常数不改变$F$的符号，因此只需对$x_1+\cdots+x_n=1$的情况证明。当然，也只需对关于$x_1,\cdots,x_n$的任意齐次量是常数的情况证明。至于取哪个齐次量以及设它等于何值，需要对F的具体形式进行观察。

题 4.12. *(a)* (幂平均不等式)设$x_1,\dots,x_n>0$，对于$p\in\mathbb{R}$，定义

$$f(p):=\begin{cases}\left(\frac{x_1^p+\cdots+x_n^p}{n}\right)^{\frac{1}{p}}, & p\ne 0,\\ \sqrt[n]{x_1\cdots x_n}, & p=0.\end{cases}$$

证明：$f(p)$是单调递增的连续函数。

(b) 设$x_1,\dots,x_n>0$，$p\ge q>0$，证明：

$$(x_1^p+\cdots+x_n^p)^{\frac{1}{p}}\le(x_1^q+\cdots+x_n^q)^{\frac{1}{q}}.$$

(c) (Hölder不等式)设$x_1,\dots,x_n,y_1,\dots,y_n\ge 0$，$p,q>0$且$\frac{1}{p}+\frac{1}{q}=1$，证明：

$$(x_1^p+\cdots+x_n^p)^{\frac{1}{p}}(y_1^q+\cdots+y_n^q)^{\frac{1}{q}}\ge x_1y_1+\cdots+x_ny_n.$$

特别的，当$p=q=2$时，上式被称为*Cauchy–Schwarz*不等式。

(d) (Minkowski不等式)设$x_1,\dots,x_n,y_1,\dots,y_n\ge 0$，$p\ge 1$，证明：

$$((x_1+y_1)^p+\cdots+(x_n+y_n)^p)^{\frac{1}{p}}\le (x_1^p+\cdots+x_n^p)^{\frac{1}{p}}+(y_1^p+\cdots+y_n^p)^{\frac{1}{p}}.$$

证明. (a) $p\ne 0$时$f(p)$显然连续，而$p=0$处的连续性已经由之前的练习题给出。当$p<0$时，$f(p)$相当于$\frac{1}{x_1},\cdots\frac{1}{x_n}$的$|p|$次幂平均取倒数，因此只需证明如果$p>q>0$，则$f(p)>f(q)$。因为是齐次不等式，不妨设$x_1^p+\cdots+x_n^p=n$，则只需证明

$$x_1^q+\cdots+x_n^q\le n.$$

根据Young不等式以及$q<p$，可知$x_1^q\le\frac{q}{p}\cdot x_1^p+\frac{p-q}{p}$，故

$$\sum_{i=1}^n x_i^q\le\sum_{i=1}^n\left(\frac{q}{p}\cdot x_1^p+\frac{p-q}{p}\right)=\frac{q}{p}\sum_{i=1}^n x_i^p+\frac{p-q}{p}\cdot n=\frac{q}{p}\cdot n+\frac{p-q}{p}\cdot n=n.$$

(b) 由齐次性，不妨设$x_1^q+\cdots+x_n^q=1$，则只需证明

$$x_1^p+\cdots+x_n^p\le 1.$$

因为$x_i\in(0,1)$，所以$x_i^p\le x_i^q$，对i求和即可。

(c) 由齐次性，不妨设$x_1^p+\cdots+x_n^p=y_1^p+\cdots+y_n^q=1$。Young不等式表明

$$\sum_{i=1}^n x_iy_i\le\sum_{i=1}^n\left(\frac{x_i^p}{p}+\frac{y_i^q}{q}\right)=\frac{1}{p}+\frac{1}{q}=1.$$

(d) 不妨设$p>1$，记$q=\frac{p}{p-1}$。令$z_i:=(x_i+y_i)^{\frac{p}{q}}$，由Hölder不等式可知

$$(x_1^p+\cdots+x_n^p)^{\frac{1}{p}}(z_1^q+\cdots+z_n^q)^{\frac{1}{q}}\ge x_1z_1+\cdots+x_nz_n,$$
$$(y_1^p+\cdots+y_n^p)^{\frac{1}{p}}(z_1^q+\cdots+z_n^q)^{\frac{1}{q}}\ge y_1z_1+\cdots+y_nz_n.$$

两式相加得

$$\begin{aligned}&\left((x_1^p+\cdots+x_n^p)^{\frac{1}{p}}+(y_1^p+\cdots+y_n^p)^{\frac{1}{p}}\right)(z_1^q+\cdots+z_n^q)^{\frac{1}{q}}\\ \ge&(x_1+y_1)z_1+\cdots+(x_n+y_n)z_n\\ =&z_1^{1+\frac{q}{p}}+\cdots+z_n^{1+\frac{q}{p}}=z_1^q+\cdots+z_n^q.\end{aligned}$$

两边约去$(z_1^q+\cdots+z_n^q)^{\frac{1}{q}}$即可。 □

注. *Hölder*不等式的条件实际上保证了$p, q \geq 1$。满足$\frac{1}{p}+\frac{1}{q}=1$的指标称为共轭的，通常认为$(p,q)=(1,\infty)$也是一对共轭指标。后一种情况下的不等式是显然的，只要把∞次幂平均按照极限解释为最大值。类似的，*Minkowski*不等式也对$p=\infty$成立。

注. *Minkowski*不等式表明$\mathbb{R}^n$上可以定义范数$\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$，其中$1 \leq p \leq \infty$。所谓范数*(norm)*，是指从$\mathbb{R}$（或$\mathbb{C}$）上的线性空间$X$到$[0,\infty)$的映射$\|\cdot\|$，满足：

(a) (正定性)对任意$x \in X$都有$\|x\| \geq 0$，并且$\|x\|=0$等价于$x=0$；

(b) 对任意$\lambda \in \mathbb{R}$（或$\mathbb{C}$）和任意$x \in X$，有$\|\lambda x\| = |\lambda|\|x\|$；

(c) (三角形不等式)对任意$x, y \in X$，有$\|x+y\| \leq \|x\| + \|y\|$。

二元组$(X, \|\cdot\|)$称为赋范空间*(normed space)*。容易看出，如果$\|\cdot\|$是X上的范数，则$d(x,y) := \|x-y\|$给出X上的度量。当然，并非所有度量都由范数给出。如此一来，$\mathbb{R}^n$上给出了不可数种度量，但好在这些$\|\cdot\|_p$是两两等价的，即

$$\|x\|_\infty \leq \|x\|_p \leq n^{\frac{1}{p}} \|x\|_\infty, \quad \forall x \in \mathbb{R}^n.$$

于是点列的收敛性在不同范数下是相同的，从而并不需要建立不可数种$\mathbb{R}^n$上的分析学。一般的，不难证明X上的两种范数$\|\cdot\|$与$\|\cdot\|'$诱导的拓扑相同当且仅当存在$C_1, C_2 > 0$使得

$$C_1\|x\| \leq \|x\|' \leq C_2\|x\|, \quad \forall x \in X.$$

这样的两个范数称为等价的（容易看出这确实是等价关系）。

题 4.13 * (Carleman不等式). 设$a_n \geq 0 (n \in \mathbb{N})$且$\sum_{n=1}^\infty a_n < \infty$，证明：

$$\sum_{n=1}^\infty \sqrt[n]{a_1 \cdots a_n} \leq e \sum_{n=1}^\infty a_n.$$

并且右边的常数e不可改进。

证明. 取$c_n = \frac{(n+1)^n}{n^{n-1}} (n \in \mathbb{N})$，则$c_1 \cdots c_n = (n+1)^n$。根据均值不等式可知

$$\sqrt[n]{a_1 \cdots a_n} \leq \frac{c_1 a_1 + \cdots + c_n a_n}{n \cdot \sqrt[n]{c_1 \cdots c_n}} = \frac{c_1 a_1 + \cdots + c_n a_n}{n(n+1)}.$$

所以

$$\sum_{n=1}^\infty \sqrt[n]{a_1 \cdots a_n} \leq \sum_{n=1}^\infty c_n a_n \sum_{k=n}^\infty \frac{1}{k(k+1)} = \sum_{n=1}^\infty \frac{c_n}{n} \cdot a_n.$$

注意到$\frac{c_n}{n} = \left(1+\frac{1}{n}\right)^n < e$，因此不等式得证。

下面证明右边的常数e不可改进。因为$\frac{c_n}{n}\to e$，所以对任意$\varepsilon>0$，存在$M>0$使得当$n>M$时$\frac{1}{c_n}\leq\frac{1}{n(e-\varepsilon)}$。对任意$N>M$，定义

$$a_n=\begin{cases}\frac{1}{c_n}, & 1\leq n\leq N,\\ 0, & n>N.\end{cases}$$

记$S_M:=\sum_{n=1}^M a_n$，则

$$\sum_{n=1}^{\infty}a_n\leq S_M+\frac{1}{e-\varepsilon}\sum_{n=M+1}^{N}\frac{1}{n}.$$

又$\sum_{n=1}^{\infty}\sqrt[n]{a_1\cdots a_n}=\sum_{n=1}^{N}\frac{1}{n+1}$，于是

$$\frac{\sum_{n=1}^{\infty}\sqrt[n]{a_1\cdots a_n}}{\sum_{n=1}^{\infty}a_n}\geq\frac{\sum_{n=1}^{N}\frac{1}{n+1}}{S_M+\frac{1}{e-\varepsilon}\sum_{n=M+1}^{N}\frac{1}{n}}=\frac{\log N+O(1)}{\frac{1}{e-\varepsilon}\log N+O(1)}.$$

令$N\to\infty$可知不等式右边的常数不能小于$e-\varepsilon$，再结合ε任意性即证。 □

补充：一致收敛

这一小节讨论函数列的一致收敛(uniform convergence)。为了说话方便通常假设定义域是$\mathbb{R}$或$[0,1]$，但实际上这里的证明并不依赖定义域。之前已经给出过例子，说明连续函数的逐点极限未必连续。因此为了保持连续性需要引进更强的概念：

定义 4.14 (一致收敛). 设$E\subset\mathbb{R}$，函数$f_n,f\colon E\to\mathbb{R}\,(n\in\mathbb{N})$。如果对任意$\varepsilon>0$，存在$N>0$使得当$n>N$时，对任意$x\in E$都有

$$|f_n(x)-f(x)|<\varepsilon,$$

则称f_n（在E上）一致收敛到f，记为$f_n\rightrightarrows f$。

一致收敛最重要的好处是保证了某些二重极限的顺序可以交换，比如要问连续函数列的极限是否连续，本质上就是说是否有

$$\lim_{x\to x_0}\lim_{n\to\infty}f_n(x)=\lim_{n\to\infty}\lim_{x\to x_0}f_n(x).$$

题 4.15. 设$f_n\colon\mathbb{R}\to\mathbb{R}$是一列连续函数，如果$f_n$一致收敛到$f$，证明：$f$是连续函数。

证明. 任意取定$x_0\in\mathbb{R}$。对任意$\varepsilon>0$，存在$n\in\mathbb{N}$，使得对任意$x\in\mathbb{R}$都有$|f(x)-f_n(x)|<\varepsilon$。因为$f_n$连续，所以存在$\delta>0$使得当$|x-x_0|<\delta$时$|f_n(x)-f_n(x_0)|<\varepsilon$，从而

$$|f(x)-f(x_0)|\leq|f(x)-f_n(x)|+|f_n(x)-f_n(x_0)|+|f(x_0)-f_n(x_0)|\leq 3\varepsilon.$$

由ε任意性可知f在x_0处连续。 □

注. 从证明可以看出，一致收敛的条件可以大大减弱。事实上，对于任意函数$f\colon \mathbb{R}\to\mathbb{R}$和$x\in\mathbb{R}$，$f$在$x$处连续的充要条件是对任意$\varepsilon>0$，存在函数$g$在$x$处连续且$|f(y)-g(y)|<\varepsilon$对某个$\delta>0$和任意$y\in(x-\delta,x+\delta)$成立。但往往这样弱的条件没法验证，相比之下一致收敛有简单的判别法，尤其是下面将会提到的大M判别法。

记$C([0,1])$（更准确的写法是$C([0,1];\mathbb{R})$）是$[0,1]$上全体实值连续函数组成的集合。对于$f,g\in C([0,1])$，定义L^∞范数（或最大模范数）

$$\|f\|_{L^\infty}=\sup_{x\in[0,1]}|f(x)|,$$

并定义$d(f,g)=\|f-g\|_{L^\infty}$，则容易看出$d$是$C([0,1])$上的度量。

题 4.16. 证明：$(C([0,1]),d)$是完备度量空间。

证明. 只需验证完备性。设f_n是Cauchy列，则对每个固定的$x\in[0,1]$，$f_n(x)$都是$\mathbb{R}$中的Cauchy列，所以$f_n(x)$逐点收敛到某个函数$f(x)$。根据Cauchy列的定义，对任意$\varepsilon>0$，存在$N>0$使得当$m,n>N$时对任意$x\in[0,1]$都有$|f_m(x)-f_n(x)|<\varepsilon$。令$m\to\infty$，就得到$|f(x)-f_n(x)|<\varepsilon\,(\forall x\in[0,1])$，所以$f_n$一致收敛到$f$。最后，上面的命题保证了$f\in C([0,1])$，所以$d(f_n,f)\to 0$，即$f_n$在$C([0,1],d)$中收敛到$f$。 □

从证明可以看出，$C([0,1])$中度量意义下的收敛就是连续函数的一致收敛。

定理 4.17 (Weierstrass优势级数判别法/大M判别法)**.** 设$a_n\geq 0$且$\sum_{n=1}^\infty a_n<\infty$。如果函数列$f_n\colon\mathbb{R}\to\mathbb{R}$满足$\sup_{x\in\mathbb{R}}|f_n(x)|\leq a_n$，证明：$\sum_{n=1}^\infty f_n$一致收敛。特别的，如果这些$f_n$都是连续函数，则$\sum_{n=1}^\infty f_n$也是连续函数。

证明. 对任意$\varepsilon>0$，存在$N>0$使得$\sum_{n=N}^\infty a_n<\varepsilon$。于是当$m>n>N$时，

$$\left|\sum_{k=n}^m f_k(x)\right|\leq\sum_{k=n}^m a_n<\varepsilon,\quad\forall x\in\mathbb{R}.$$

从而$\sum_{k=1}^n f_k$关于L^∞范数是Cauchy列，仿照上一题的证明可知一致收敛。 □

在数学分析II中还会详细研究一致收敛的性质，包括其与可微性以及Riemann积分的联系。因此这一小节仅对在不久的将来就会用到事情进行讨论，其余按下不表。

第5讲：连续函数（下）

随后，他像是自言自语地说："如果空间是无限的，我们就处在空间的任何一点。如果时间是无限的，我们就处在时间的任何一点。"

——豪尔赫·路易斯·博尔赫斯《沙之书》

这一讲针对连续函数的性质，主要是介值原理和一致连续性；补充介绍了上/下半连续函数和连续性的拓扑刻画，以及Baire纲定理的证明和应用。最后一点难度较大，本应是实变函数或泛函分析内容，但相关的问题在今天已经成为经典。

例题

例 5.1. 设$f\colon \mathbb{R}\to\mathbb{R}$是连续单射，证明：$f$严格单调。

证明. 只需证明单调，严格是自动的。若不然，则存在$a<b<c$使得$f(b)>f(a)$且$f(b)>f(c)$，或者$f(b)<f(a)$且$f(b)<f(c)$。不妨设是前一种情况，则根据介值原理可知，对充分小的δ，$(f(b)-\delta, f(b))$中的数在(a,b)和(b,c)上都能被取到，从而与单射矛盾！ □

例 5.2. 设函数$f\colon (a,b)\to\mathbb{R}\,(-\infty<a<b<\infty)$，证明：$f$一致连续当且仅当可以连续延拓到$[a,b]$上，即存在连续函数$F\colon [a,b]\to\mathbb{R}$使得$F|_{(a,b)}=f$。

证明. 充分性是显然的，下证必要性。对任意$\varepsilon>0$，存在$\delta>0$使得当$|x-y|<\delta$时$|f(x)-f(y)|<\varepsilon$。特别的，对任意$x,y\in(a,a+\delta)$都有$|f(x)-f(y)|<\varepsilon$，由Cauchy收敛准则可知$f(a+)$存在。同理$f(b-)$存在，容易验证如此补充定义之后的f在$[a,b]$上连续。 □

注. 如果条件换成无限区间(a,∞)，则一致连续函数可以连续延拓到$[a,\infty)$，但反过来$[a,\infty)$上的连续函数当然未必一致连续。一般的，度量空间中定义在一个子集上的一致连续函数可以唯一地连续延拓到闭包上，而紧集上的连续函数一定一致连续。

例 5.3. *(a)* 给定$\alpha>0$，证明：对任意$x,y>0$都有

$$|x^\alpha-y^\alpha|\le\begin{cases}|x-y|^\alpha, & 0<\alpha\le 1,\\ \alpha\max\{x,y\}^{\alpha-1}|x-y|, & \alpha>1.\end{cases}$$

(b) 给定$\alpha>0$，证明：$x^\alpha\sin\frac{1}{x}$在$(0,1)$上的*Hölder*连续。

证明. (a) 不妨设$x>y>0$。当$\alpha\le 1$时，先前证明过

$$(|x-y|^\alpha+y^\alpha)^{1/\alpha}\ge|x-y|+y=x.$$

两边α次方再移项即可。而如果$\alpha > 1$，则利用Bernoulli不等式得（注意$|(x-y)/x| < 1$）

$$\frac{y^\alpha}{x^\alpha} = \left(1 - \frac{x-y}{x}\right)^\alpha \geq 1 - \frac{\alpha(x-y)}{x}.$$

简单整理即可。[另一种方法是使用Lagrange中值定理。]

(b) 记$\beta = \min\{\alpha/2, 1\}$。任取$0 < y < x < 1$，则

$$\left|x^\alpha \sin\frac{1}{x} - y^\alpha \sin\frac{1}{y}\right| \leq |x^\alpha - y^\alpha| + y^\alpha \left|\sin\frac{1}{x} - \sin\frac{1}{y}\right|.$$

由(a)可知

$$|x^\alpha - y^\alpha| \leq C|x-y|^{\min\{\alpha,1\}} \leq C|x-y|^\beta.$$

此外，注意到

$$\left|\sin\frac{1}{x} - \sin\frac{1}{y}\right| \leq \min\left\{2, \left|\frac{1}{x} - \frac{1}{y}\right|\right\} = \min\left\{2, \frac{|x-y|}{xy}\right\}.$$

根据定义有$\beta \leq 1$，所以

$$\left|\sin\frac{1}{x} - \sin\frac{1}{y}\right| \leq 2^{1-\beta}\left(\frac{|x-y|}{xy}\right)^\beta = C\frac{|x-y|^\beta}{x^\beta y^\beta}.$$

故对于$0 < y < x < 1$有

$$\left|x^\alpha \sin\frac{1}{x} - y^\alpha \sin\frac{1}{y}\right| \leq C|x-y|^\beta + C\cdot\frac{y^{\alpha-\beta}}{x^\beta}\cdot|x-y|^\beta \leq C|x-y|^\beta.$$

最后一步用到了$\alpha - \beta \geq \beta$以及$y < x$。这就证明了$\beta$阶Hölder连续性。 □

注. *(b)中考虑$x, y = 1/((2n \pm 1/2)\pi)$ $(n \in \mathbb{N})$可知$\min\{\alpha/2, 1\}$阶Hölder连续性是最优的。同样的构造可以说明$\frac{\sin(1/x)}{\log x}$可以延拓为$[0, 1/2]$上的连续函数，但并不是Hölder连续函数。*

练习题

题 5.4. 是否存在连续函数$f\colon \mathbb{R} \to \mathbb{R}$将有理数映成无理数、且将无理数映成有理数？

解. 不存在。反证法，假设f满足要求，则f不是常数，再结合连续性可知，存在$a < b$ $(a, b \in \mathbb{Q})$使得$f(a) \neq f(b)$，不妨设$f(a) < f(b)$，注意到$[f(a), f(b)]$中有不可数个无理数，根据介值定理它们都必须是$[a, b]$中某个有理数的像。但有理数只有可数个，矛盾！ □

题 5.5. 设$f\colon \mathbb{R} \to \mathbb{R}$是非常值周期函数，且至少有一个连续点，证明：$f$有最小正周期。

证明. 之前证明过任何函数的周期组成的集合要么稠密，要么有某个最小正周期生成。只需证明如果f的周期稠密，则对任意$x \in \mathbb{R}$都有$f(x) = f(z)$，其中z是f的一个连续点。取一列周期$T_n \to z - x$，则

$$f(x) = f(x + T_n) \to f(z).$$

最后一步用到了$x + T_n \to x + z - x = z$以及$f$在$z$处的连续性。 □

题 5.6. 设f, g是非常值连续周期函数，且最小正周期之比是无理数，证明：$f + g$不是周期函数。

证明*1.* 用反证法。记f, g的最小正周期为T_f, T_g。因为$\frac{T_f}{T_g} \notin \mathbb{Q}$，所以$h := f + g$不是常数，从而有最小正周期$T$。注意到

$$h(x + T_f) - h(x) = g(x + T_f) - g(x),$$

左边以T为周期，右边以T_g为周期，并且两边定义了同一个非常值连续周期函数，所以$T/T_g \in \mathbb{Q}$。同理可知$T/T_f \in \mathbb{Q}$，从而与$T_f/T_g \notin \mathbb{Q}$矛盾。 □

证明*2.* 反证假设同上，于是对任意$x \in \mathbb{R}$都有$f(x+T) + g(x+T) = f(x) + g(x)$。令

$$h(x) := f(x + T) - f(x) = g(x) - g(x + T).$$

则h是连续函数且T_f, T_g都是周期，这就导致h没有最小正周期，从而只能是常数。易见连续周期函数一定有界，如果$h(x) =: h \neq 0$，则

$$f(x + nT) - f(x) = h(x) + h(x + T) + \cdots + h(x + (n-1)T) = nh,$$

在$n \to \infty$时无界，故只能$h = 0$。于是T同时是f和g的周期，与$\frac{T_f}{T_g} \notin \mathbb{Q}$矛盾！ □

后一种证明看似麻烦，但其想法可以用于讨论离散调和函数的Liouville定理。

题 5.7 *. 设$n \in \mathbb{N}$，称$f \colon \mathbb{Z}^n \to [0,1]$是离散调和函数，如果

$$f(x_1, \dots, x_n) = \frac{1}{2n} \sum_{|x_1 - y_1| + \cdots + |x_n - y_n| = 1} f(y_1, \dots, y_n), \quad \forall (x_1, \dots, x_n) \in \mathbb{Z}^n.$$

(a) 证明：如果f是有界离散调和函数，则f是常数。

(b) 证明：如果f是下界（或者有上界）的离散调和函数，则f是常数。

证明. (a) 不妨设$f(x) \in [0,1]$。记$g(x,y) = f(x+1,y) - f(x,y)$，则$g(x,y)$有界并且满足相同的函数方程。此外，根据定义，对任意$n \in \mathbb{N}$和任意$x, y \in \mathbb{Z}$，都有

$$|g(x,y) + g(x+1,y) + \cdots + g(x+n,y)| \leq 2.$$

下证g只能恒等于0。反证法，假设$\sup|g|>0$，不妨设是$M:=\sup g>0$，则对任意$\varepsilon>0$，存在$(x,y)\in\mathbb{Z}^2$使得$g(x,y)\geq M-\varepsilon$。利用函数方程可以看出$g(x+1,y)\geq M-4\varepsilon$，进一步归纳可知$g(x+k,y)\geq M-4^k\varepsilon$。于是对于任意给定的$n\in\mathbb{N}$，总可以取充分小的$\varepsilon$使得$g(x,y),g(x+1,y),\cdots g(x+n,y)>\frac{M}{2}$，从而它们的和至少是$\frac{M}{2}\cdot n$，取$n$充分大可得矛盾。因此$f(x+1,y)=f(x,y)$，即$f$与$x$无关。同理可知$f$与$y$无关，从而是一个常数。

(b) 通过平移、伸缩（有上界时还需用$-f$代替f），只需证明：如果严格正的离散调和函数满足$f(0)=1$（注意区分n维向量0与实数0），则f恒等于1。记这样的函数组成的集合为S。

为了说话方便，将定义中的求和范围称为与x相邻的点；如果y与x相邻，则x也与y相邻。如果f是非负调和函数且x,y是相邻的，则$f(x)\geq f(y)/2n$，交换x,y可得$f(x)\leq 2nf(y)$。易见任何$x\in\mathbb{Z}^n$都可以通过有限多个相邻的点到达0，因此存在常数$C_x>1$使得

$$C_x^{-1}f(0)\leq f(x)\leq C_xf(0).$$

特别的，对于$f\in S$，有$f(x)\in[C_x^{-1},C_x]$。考虑乘积空间$\Gamma=\prod_{x\in\mathbb{Z}^n}[C_x^{-1},C_x]$。将$S$按此自然的单射$f\mapsto(f(x))_{x\in\mathbb{Z}^n}$等同于$\Gamma$的子集，并赋予子拓扑（对应于函数的逐点收敛），则按定义容易证明S是闭凸集。根据Tychonoff定理可知Γ是紧集（由于是可数乘积，也可以先说明Γ可度量化，再用对角线方法证明列紧性），所以S是紧凸集。

核心的技巧（不本质，见注）在于构造连续的严格凸函数$T\colon\Gamma\to\mathbb{R}$，使得$T$的唯一最小值点是常值函数1。一旦这样的$T$存在，根据紧性可知$T$在$S$上存在最大值点$f\in S$；定义

$$f_i(x)=f(x+e_i)/f(e_i),\quad f_{n+i}=f(x-e_i)/f(-e_i),\quad i=1,\ldots,n,$$

其中e_i是第i个单位向量，以及$p_i=f(e_i)/(2n)$和$p_{n+i}=f(-e_i)/(2n)$，则调和函数满足

$$f=\sum_{k=1}^{2n}p_kf_k.$$

显然有$f_k\in S$，并且$p_k>0$满足$\sum_{k=1}^n p_k=f(0)=1$。所以

$$T(f)=T\left(\sum_{k=1}^n p_kf_k\right)\leq\sum_{k=1}^n p_kT(f_k)\leq T(f).$$

由于T严格凸，等号成立必须每个f_k都等于f。换言之，对$i=1,\ldots,n$都有

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

特别的，$f(-e_i)=f(e_i)^{-1}f(0)=f(e_i)^{-1}$。根据均值不等式可知

$$1=\sum_{k=1}^{2n}p_k=\frac{1}{2n}\sum_{i=1}^n(f(e_i)+f(e_i)^{-1})\geq 1.$$

等号成立必须每个$f(e_i)$都等于1，从而f关于n个方向平移不变，必然恒等于1。因为$f = 1$是T在S上的最大值点，同时也是T的唯一最小值点，所以S只包含$f = 1$一个函数。

最后，构造满足要求的T是容易的，比如类比二次函数定义

$$T((\gamma_x)_{x\in\mathbb{Z}^n}) := \sum_{x\in\mathbb{Z}^n} \frac{1}{(1+|x|)^{n+1}} \frac{(\gamma_x - 1)^2}{C_x^2}.$$

注意$|\gamma_x - 1| \le C_x$以及分母的$(1+|x|)^{n+1}$保证了绝对一致收敛，并且每一项都连续，从而极限函数T连续；而严格凸是因为每一项关于γ_x都是严格凸的；显然T非负，且唯一的零点是是所有γ_x恒为1。这就完成了证明。 □

注. 本质上*(b)*的方法是通过研究凸集的极端点*(extremal point)*来得到一般点的性质。所谓凸集的极端点是指不能写成凸集中任何互异两点的严格凸组合的点；例如对于$\mathbb{R}^n$中的凸多面体来讲，极端点就是所有顶点。泛函分析中的*Krein–Milman*定理表明，在局部凸拓扑线性空间（比如*Banach*空间）中，如果K是紧凸集，则K等于全体极端点的闭凸包；特别的，极端点一定存在。这一命题的经典应用是便理论中遍历测度的存在性。

具体来说，全体$\mathbb{Z}^n$上的函数关于逐点收敛构成局部凸拓扑线性空间（每个点处的绝对值构成一族半范数），*(b)*中定义的集合S构成紧凸集；如果能说明极端点都是常值函数，则根据*Krein–Milman*定理即证。极端点的定义代替了上面构造的严格凸函数T，由此直接得到f_k都等于f，余下是相同的。类似的方法可以更简洁地得到*(a)*。

注. 当$n = 2$时，*(b)*具有更简单的概率论证明。设X_n是$\mathbb{Z}^2$上的简单随机游走，则离散调和函数的定义保证了$f(X_n)$是鞅*(martingale)*。因为f非负，鞅收敛定理说明$f(X_n)$几乎必然收敛。而根据二维随机游走的常返性，每个点x都几乎必然会被X_n无穷多次走到，所以极限等于$f(x)$，结合x的任意性可知f是常数。这个论证在$n \ge 3$时失效，因为高维随机游走不是常返的。

题 5.8. 设$f\colon \mathbb{R} \to \mathbb{R}$是连续函数，满足$x = 0$是$f$的唯一零点。对于有界数列$x_n$，证明：如果$f(x_n) \to 0$，则$x_n \to 0$。

证明*1.* 容易证明有界数列收敛当且仅当任何收敛子列的极限都相同（注意这里子列的收敛性是凭空产生的，类似于泛函分析中的闭图像定理）。于是只需证明：如果子列$x_{n_k} \to A$，则$A = 0$。事实上，根据f连续以及$f(x_{n_k}) \to 0$可知$f(A) = 0$，再结合$f^{-1}(0) = \{0\}$即得。 □

证明*2.* 设$|x_n| \le M$。对任意$\varepsilon > 0$，函数$|f|$在$\{\varepsilon \le |x| \le M\}$上连续、有最小值、没有零点，所以有正下界$\delta > 0$。存在$N \in \mathbb{N}$使得当$n \ge N$时$|f(x_n)| < \delta$，因此只能$|x_n| < \varepsilon$。 □

注. 如果去掉x_n的有界性，则结论可能不成立，因为f可以在无穷远处衰减到0（类似于$f(x) = 1/x$且$x_n = n$）。如果再加上$\liminf_{|x|\to\infty} |f(x)| > 0$，则条件中的有界性可以去掉。

题 5.9. 设$f\colon \mathbb{R}\to\mathbb{R}$是连续函数，问能否每个点恰有2个原像？又能否恰有3个原像？

证明. 每个点恰有2个原像是不行的。反证法，假设$f\colon \mathbb{R}\to\mathbb{R}$满足每个点恰有2个原像，设两个零点是$a<b$，于是在$(a,b)$上$f$定号。用$-f$替换$f$，可以不妨设$f(x)>0(\forall x\in(a,b))$。记$M=\sup_{[a,b]} f>0$，则根据介值定理可知$(0,M)$中的值都在$(a,b)$中被取到至少2次，于是在$(-\infty,a)\cup(b,\infty)$上只能有$f(x)<0$。但这就导致$(M,\infty)$不在$f$的值域中，矛盾！

下面证明每个点恰有3个原像是可以的。定义$\varphi\colon[0,3)\to\mathbb{R}$的图像由连接$(0,0)$、$(1,1)$、$(2,0)$以及$(3,1)$的三段折线组成，再将$\varphi$以3为周期延拓到$\mathbb{R}$上。现在定义$f\colon\mathbb{R}\to\mathbb{R}$如下：当$x\in[3n,3n+3)(n\in\mathbb{Z})$时，令$f(x)=n+\varphi(x)$。容易看出$f(x)$是良定义的连续函数，并且每个点恰好有3个原像。事实上，根据“周期性”只需对$[0,1]$验证，而这是显然的。 $\square$

注. 仿照这里的证明和构造可知，如果n是偶数，不可能有$\mathbb{R}\to\mathbb{R}$的连续函数将每个点恰好映到n次；而如果n是奇数，则这样的函数是存在的。此外，如果只要求每个点的原像个数都是偶数，那么这样的连续函数可以存在；比如令$g(x)$在$(-\infty,0]$上是$-x$，在$(0,\infty]$上等于上面构造的$f(x)$，则$|g^{-1}(y)|$当$y<0$时为0，当$y=0$时为2，当$y>0$时为4。

题 5.10. *(a)* 设有界实数列x_n满足$x_n-x_{n-1}\to 0$，证明：$[\liminf x_n,\limsup x_n]$都是$x_n$的聚点。[实际上结论对无界实数列也成立，只是需要对∞是聚点给出定义。]

*(b)** 设$f\colon[0,1]\to[0,1]$连续，任取$x_0\in[0,1]$，递归定义$x_n=f(x_{n-1})\,(n\in\mathbb{N})$，证明：$x_n$收敛当且仅当$x_n-x_{n-1}\to 0$。

证明. (a) 记$l=\liminf x_n$以及$L=\limsup x_n$。反设某个$c\in(l,L)$不是聚点，则存在$\delta>0$使得$(c-\delta,c+\delta)\subset(l,L)$且只包含有限个$x_n$。当$n$充分大时，$|x_{n+1}-x_n|<\delta$，于是从某一项起$x_n$全都落在$c-\delta$的左边或者$c+\delta$的右边（思考理由），与$l,L$都是聚点矛盾！

(b) 必要性是显然的，下证用反证法证明充分性。记$l=\liminf x_n$以及$L=\limsup x_n$，反设$l<L$。下面证明$[l,L]$都是不动点，根据连续性，只需证明任意$c\in(l,L)$都是不动点。若不然，不妨设$f(c)>c$，则存在$\delta>0$使得$f(x)>x\,(\forall x\in(c-\delta,c+\delta))$。当$n$充分大时数列$x_n$中相邻两项之差小于$\delta$，于是$x_n$要从$c+\delta$右边走到$c-\delta$左边，必然要先落在$(c-\delta,c+\delta)$中，而这又导致$f(x_n)>x_n$，所以$x_n$永远无法从右边走到左边去（请思考如何严格书写），与l,L都是聚点矛盾！最后，由$x_{n+1}-x_n\to 0$可知一定存在某个n使得$x_n=(l,L)$，从而$x_n=x_{n+1}=\cdots$，与l,L是聚点矛盾！这就证明了结论。 $\square$

题 5.11. *(a)* 设$f,g\colon[0,1]\to[0,1]$连续，且满足$f\circ g=g\circ f$，证明：存在$x\in[0,1]$使得$f(x)=g(x)$。

(b) 设$f\colon [a,b]\to\mathbb{R}$连续，且存在一列$x_n\in[a,b]$使得

$$f(x_n)=g(x_{n+1}).$$

证明：存在$x\in[a,b]$使得$f(x)=g(x)$。

证明. (a) 反证法，若结论不成立，根据介值原理可以不妨设$f(x)>g(x)\,(\forall x\in[a,b])$。熟知$f$有不动点，记为$x_1\in[0,1]$，则

$$f(g(x_1))=g(f(x_1))=g(x_1).$$

所以$g(x_1)$也是f的不动点，于是可以递归定义$x_{n+1}=g(x_n)$都是f的不动点。注意到

$$x_{n+1}=g(x_n)<f(x_n)=x_n,$$

所以x_n单减有极限，记为z。对$f(x_n)=x_n$和$x_{n+1}=g(x_n)$分别取极限，由连续性可知

$$f(z)=z=g(z),$$

与反证假设矛盾。[注意这并没有说明f,g一定有公共不动点，这件事具有反例。]

(b) 仍然用反证法，不妨设$f>g$，则条件表明

$$g(x_{n+1})=f(x_n)>g(x_n),$$

结合g有界可知$g(x_n)$单增有极限，特别的任何子列也收敛到同一个极限。因为x_n有界，存在收敛子列$x_{n_k}\to z\in[a,b]$，结合连续性可知

$$g(z)=\lim_{k\to\infty}g(x_{n_k})=\lim_{n\to\infty}g(x_n).$$

因为$g(x_n)=f(x_{n-1})$，所以

$$g(z)=\lim_{n\to\infty}f(x_n)=\lim_{k\to\infty}f(x_{n_k})=f(z).$$

[这里的逻辑十分微妙：$g(x_n)$有极限是单调性，极限等于$g(z)$用到连续性以及z是x_n的聚点；极限等于$f(x_n)$的极限源于条件的等式，特别的也等于子列$f(x_n)$的极限$f(z)$。] $\square$

题 5.12 * (2018 IMO中国集训队选拔考试)**.** 设$p,q>0$且$p+q=1$，证明：对任意2017元实数组$(y_1,\ldots,y_{2017})$，存在唯一的2017元实数组$(x_1,\ldots,x_{2017})$，使得（约定$x_{2018}=x_1$）

$$p\max\{x_i,x_{i+1}\}+q\min\{x_i,x_{i+1}\}=y_i,\quad i=1,\ldots,2017.$$

证明. 对任意给定的$x_1 \in \mathbb{R}$，容易看出

$$x_2 \mapsto p\max\{x_1, x_2\} + q\min\{x_1, x_2\}$$

严格单增的折线函数，因此存在唯一的x_2使得结论对$i = 1$成立。容易看出这样定义出的x_2是x_1的单减连续函数，并且满足

$$\lim_{x_1 \to -\infty} x_2 = \infty \quad 和 \quad \lim_{x_1 \to \infty} x_2 = -\infty.$$

称这样的函数为好函数。一旦得到了x_2，就可以唯一确定x_3使得结论对$i = 2$成立，并且$x_2 \mapsto x_3$也是好函数；递推可以得到$x_2, \ldots, x_{2018}$使得结论对$i = 1, \ldots, 2017$成立（在不要求$x_{2018} = x_1$的前提下）。这里x_{2018}是关于x_1的函数，作为2017个好函数的复合显然也是好函数。由介值定理可知好函数一定有不动点，所以存在$x_1 \in \mathbb{R}$使得$x_{2018} = x_1$。 □

补充：上/下半连续函数

这一小节讨论上半连续函数，下半连续函数的理论是完全类似的。主要的结果是有界闭区间上的上半连续函数函数一定有最大值点，以及一族上半连续函数的下确界是上半连续的。先前说过连续函数的极限未必连续，但如果考虑一列递减的上半连续函数，则极限也是上半连续的。于是半连续性在某些情况下更容易保持，比如拓扑学中的Urysohn引理的证明用上/下半连续函数书写会更加自然[8]。

定义 5.13. 函数$f\colon \mathbb{R} \to [-\infty, \infty)$称为上半连续的，如果对任意$x \in \mathbb{R}$都有

$$f(x) \geq \limsup_{y \to x} f(y).$$

类似可以定义下半连续函数是指$f(x) \leq \liminf_{y \to x} f(y)$。容易发现，闭集的特征函数上半连续，而开集的特征函数下半连续。注意上半连续的定义中允许f取值$-\infty$，这是为了今后说话方便。此外，虽然定义中f的定义域是$\mathbb{R}$，但容易类似得到定义域是一般区间的情况。

题 5.14. 设$[a, b]$是有界闭区间，$f\colon [a, b] \to [-\infty, \infty)$上半连续，则$f$有最大值点（允许$-\infty$）。

证明. 由$f(x) \geq \limsup_{y \to x} f(y)$可知存在$\delta > 0$使得当$y \in [a, b]$满足$|y - x| < \delta$时$f(y) < f(x) + 1$，因此$f$在$(x - \delta, x + \delta)$上有界。根据有限覆盖定理可知$f$的上界$M$有限。取一列$x_n \in [a, b]$使得$f(x_n) \to M$，则存在$x_n$的收敛子列$x_{n_k} \to x$，根据上半连续性有$f(x) \geq \limsup_{k \to \infty} f(x_{n_k}) = M$，所以$f(x) = M$，即$x$是最大值点。 □

[8]相关论述可在W. Rudin, *Real and complex analysis*, third edition, McGraw-Hill, New York, 1987中找到。

下一题给出连续性的拓扑刻画。因为开集总是开区间的并，所以f连续又等价于开区间的原像是开集。事实上，任何两个拓扑空间之间的映射都可以用“开集的原像是开集”来定义连续性，而通常的连续函数与上半连续函数无非是在$\mathbb{R}$取了不同的拓扑结构。

题 5.15. 设函数$f\colon \mathbb{R} \to \mathbb{R}$，证明：

(a) f连续的充要条件是开集的原像是开集。

(b) f上半连续的充要条件是对任意$a \in \mathbb{R}$，$\{x \in \mathbb{R} \mid f(x) < a\}$是开集。

证明. 只要把话说明白就可以了，并没有本质困难。这里仅以(b)的必要性为例。对任意$a > f(x)$，因为$\{x \in \mathbb{R} \mid f(x) < a\}$是开集，所以存在$\delta > 0$使得当$|x-y| < \delta$时$f(y) < a$，进而$\limsup_{y\to x} f(y) \le a$。由$a$的任意性可知$\limsup_{y\to x} f(y) \le f(x)$，即$f$在$x$处上半连续。 □

下一题体现了拓扑语言的优势。若不借助开集，说明起来虽不太困难，但相当丑陋。

题 5.16. 设$f_\alpha\ (\alpha \in I)$是一族上半连续函数，则$f(x) := \inf_{\alpha\in I} f_\alpha(x)$也是上半连续函数。

证明. 只需注意到对任意$a \in \mathbb{R}$有

$$\{x \in \mathbb{R} \mid f(x) < a\} = \bigcup_{\alpha\in I}\{x \in \mathbb{R} \mid f_\alpha(x) < a\},$$

并利用任意一族开集的并还是开集即可。 □

注. 反过来，可以证明任何上半连续函数都是一列连续函数的单调递减逐点极限。构造方式与先前构造连续函数逐点收敛到*Riemann*函数有相似之处，这里为了简洁起见不展开说明。

补充：Baire纲定理

Baire纲定理应当属于拓扑学，但主要用在实变函数和泛函分析中。

定理 5.17* (Baire纲定理)**.** 设(X,d)是完备度量空间，证明：

(a) 如果$G_1, G_2, \ldots$是可数个稠密开集，则$\bigcap_{n=1}^\infty G_n$仍然稠密。

(b) 如果$F_1, F_2, \ldots$是可数个闭集，且$X = \bigcup_{n=1}^\infty F_n$，则一定某个$F_n$有内点。

证明. 容易看出在取补集之后，(b)是(a)的推论，具体细节留作练习，下面给出(a)的证明。只需证明对任意开球$B_0 = B(x_0, r_0)$，都有$B_0 \cap \bigcap_{n=1}^\infty G_n \ne \emptyset$。因为$G_1$是稠密开集，所以存在开球$B_1 = B(x_1, r_1)$，使得$0 < r_1 < 1$且$2B_1 := B(x_1, 2r_1) \subset B_0 \cap G_1$。一般的，可以递归找出一列开球$B_n = B(x_n, r_n)$，满足$0 < r_n < \frac{1}{n}$且

$$2B_n := B(x_n, 2r_n) \subset B_{n-1} \cap G_n.$$

对任意$m > n > N$，因为$x_m, x_n \in B_N$，所以$d(x_m, x_n) \le 2r_N \le \frac{2}{N}$，所以$\{x_n\}$是Cauchy列，根据完备性可知$x_n$收敛到某个点$x$。进一步，因为当$m > n$时$x_m \in B_{n+1}$，所以$x \in 2B_{n+1} \subset B_n$[因为取极限之后可能导致$d(x, x_{n+1}) = r_{n+1}$，所以在构造中采用了二倍的球]。这就证明了

$$x \in \bigcap_{n=0}^{\infty} B_n \subset B_0 \cap \bigcap_{n=1}^{\infty} G_n. \qquad \square$$

注. *满足Baire纲定理的空间也被称为Baire空间，于是完备度量空间都是Baire空间。另一类分析学中常见的Baire空间是局部紧Hausdorff空间，相应的定理证明是类似的。*

注. *从上述证明可以凝练出区间套定理的推广：在完备度量空间中，如果$B_1 \supset B_2 \supset \cdots$是一列递减的闭球（所谓闭球是指形如$\{y \in X \mid d(x,y) \le r\}$ $(x \in X,\ r > 0)$的集合），且B_n的半径趋于0，则$\bigcap_{n\in\mathbb{N}} B_n$恰有一个点。注意这里的非平凡之处在于闭球未必是紧集，但完备性足以保证交集非空。由此容易直接证明(b)。*

Baire纲定理中的“纲”代表集合的分类。经常要把集合被分为两类，一类是小的，另一类是大的：比如可数集与不可数集，离散点集与稠密集。设X是完备度量空间，如果一个集合的闭包没有内点，则称它是无处稠密的(nowhere dense)。无处稠密集的可数并称为第一纲集，不是第一纲集的集合称为第二纲集。Baire纲定理表明，可数个稠密开集的交是第二纲集。

为了说话简洁，今后将可数个开集的交称为G_δ集。以下是Baire纲定理的若干应用。

题 5.18. *设函数$f\colon \mathbb{R} \to \mathbb{R}$，本题希望证明$f$的连续点集不能等于$\mathbb{Q}$，具体分为以下步骤：*

(a) *对任意$x \in \mathbb{R}$，定义f在x处的振幅为*

$$\omega(x) := \lim_{\delta \to 0+} \sup_{y,z \in (x-\delta, x+\delta)} |f(y) - f(z)|.$$

证明：$\omega(x) \in [0, \infty]$是良定义的，并且$f$在$x$处连续当且仅当$\omega(f) = 0$。

(b) *证明：对任意$\varepsilon > 0$，$\{x \in \mathbb{R} \mid \omega(x) < \varepsilon\}$是开集。*

(c) *证明：f的连续点集是G_δ集。*

*(d)** *用Baire纲定理证明：$\mathbb{Q}$不是G_δ集。从而f的连续点集不能等于$\mathbb{Q}$。*

证明. (a) 对$y, z \in (x-\delta, x+\delta)$的sup是关于$\delta$单增的，根据单调有界收敛原理可知$\omega(x) \in [0, \infty]$。而连续等价于$\omega(x) = 0$无非是定义的重述，留作练习。

(b) 若$\omega(x) < \varepsilon$，则存在$\delta > 0$使得对任意$y, z \in (x-\delta, x+\delta)$，都有

$$|f(y) - f(z)| < \varepsilon.$$

特别的，对任意$w \in (x-\delta, x+\delta)$，记$\delta' = \delta - |w-x|$，结合单调性可知

$$\omega(w) \le \sup_{y,z \in (w-\delta', w+\delta')} |f(y)-f(z)| \le \sup_{y,z \in (x-\delta, x+\delta)} |f(y)-f(z)| < \varepsilon.$$

所以$\{x \in \mathbb{R} \mid \omega(x)\} < \varepsilon$是开集。

(c) 由(a)可知$\{f\text{的连续点}\} = \omega^{-1}(0) = \bigcap_{n=1}^{\infty} \left\{x \in \mathbb{R} \mid \omega(x) < \frac{1}{n}\right\}$，结合(b)即证。

(d) 反证法，假设$\mathbb{Q}$是可数个开集$G_1, G_2, \dots$的交。因为$\mathbb{Q}$可数，记$\mathbb{Q} = \{q_n \mid n \in \mathbb{N}\}$，则

$$\emptyset = (\mathbb{R} \setminus \mathbb{Q}) \cap \bigcap_{n=1}^{\infty} G_n = \left(\bigcap_{n=1}^{\infty} \{q_n\}^C\right) \cap \left(\bigcap_{n=1}^{\infty} G_n\right),$$

但右边是可数个稠密开集的交，矛盾！结合(c)(d)可知f的连续点集不能是$\mathbb{Q}$。 □

题 5.19 *. *(a)* 证明：没有孤立点的完备度量空间包含不可数个点。

(b) 证明：$(0,1)$不能写成可数个闭区间的无交并。

证明. (a) 反证法，假设(X,d)是完备度量空间且$X = \{x_n \mid n \in \mathbb{N}\}$。因为$x_n$不是孤立点，所以$\{x_n\}^C$是稠密开集。然而$\bigcap_{n=1}^{\infty}\{x_n\}^C = \emptyset$，矛盾！

(b) 反证法，假设$(0,1)$可以写成$I_n := [a_n, b_n]\,(n \in \mathbb{N})$的无交并。记这可数个端点组成的集合为$E := \{a_n, b_n \mid n \in \mathbb{N}\}$，则对任意$x \in (0,1) \setminus E$，存在$I_n$使得$x \in (a_n, b_n)$，结合$(a_n, b_n) \cap E = \emptyset$可知$E$是闭集。对任意$\delta > 0$，$(a_n - \delta, a_n) \cap (0,1) \ne \emptyset$，因此一定包含某个$I_k$的右端点，所以$a_n$是$E$的聚点，同理$b_n$也是$E$的聚点。因此$E$是没有孤立点的完备度量空间，由(a)知$E$是不可数集，矛盾！ □

注. 没有孤立点的闭集称为完全集*(perfect set)*。在*(b)*中使用的是*(a)*的推论：完备度量空间中的完全集一定不可数。这方面的典型例子是$Cantor$集。

题 5.20. 设$f\colon [0,\infty)$是一个函数，考虑如下问题：

(a) 如果f连续且对任意$x > 0$都有$\lim_{n\to\infty} f(x+n) = 0$，是否$\lim_{x\to\infty} f(x) = 0$?

(b) 如果f一致连续且对任意$x > 0$都有$\lim_{n\to\infty} f(x+n) = 0$，是否$\lim_{x\to\infty} f(x) = 0$?

(c) 如果对任意$x > 0$都有$\lim_{n\to\infty} f(nx) = 0$，是否$\lim_{x\to\infty} f(x) = 0$?

*(d)** 如果f连续且对任意$x > 0$都有$\lim_{n\to\infty} f(nx) = 0$，是否$\lim_{x\to\infty} f(x) = 0$?

解. (a) 不是。对于$k \in \mathbb{N}$，定义连续函数f_k在区间$I_k = (a_k, b_k) := (k + \frac{1}{2^{k+1}}, k + \frac{1}{2^k})$之外是0、在中点$\frac{a_k+b_k}{2}$处等于1、其余位置用两条线段连接。令$f(x) = \sum_{k=1}^{\infty} f_k(x)$，注意对任意$x \in$

$\mathbb{R}$，存在x的邻域$(x-\delta,x+\delta)$至多与一个I_k相交，所以f是连续函数。此外，至多存在一个$n\in\mathbb{N}$使得$x+n\in\bigcup_{k=1}^{\infty}I_k$，所以$\lim_{n\to\infty}f(x+n)=0$。但显然$\limsup_{x\to\infty}f(x)=1$。

(b) 是的。对任意$\varepsilon>0$，存在$M\in\mathbb{N}$，使得当$|x-y|\le\frac{1}{M}$时$|f(x)-f(y)|<\varepsilon$，记$0=x_0<x_1<\cdots<x_M=1$是$[0,1]$的M等分，即$x_i=\frac{i}{M}$，则对$0\le i\le M$，存在$N_i\in\mathbb{N}$，使得当$n\ge N_i$时$|f(x_i+n)|<\varepsilon$。记$N=\max\{N_1,\dots,N_M\}$，则对任意$x>N(x\in\mathbb{R})$，存在$x_i(0\le i\le M)$使得$|x-(x_i+\lfloor x\rfloor)|<\frac{1}{M}$，结合$\lfloor x\rfloor\ge N$可知

$$|f(x)|\le|f(x)-f(x_i+\lfloor x\rfloor)|+|f(x_i+\lfloor x\rfloor)|\le\varepsilon+\varepsilon=2\varepsilon.$$

由ε任意性可知$\lim_{x\to\infty}f(x)=0$。

(c) 不是。比如取一列互异的素数$p_n\to\infty$，定义f在$\sqrt{p_n}$处为1、其他位置为0。因为$\sqrt{p_n}$两两比值不会是有理数，所以条件成立，但$\limsup_{x\to\infty}f(x)=1$。

(d) 是的。任意取定$\varepsilon>0$，对于$n\in\mathbb{N}$，定义

$$Z_n:=\{x\in[0,\infty)\mid|f(kx)|\le\varepsilon,\ \forall k\ge n\}=\bigcap_{k=n}^{\infty}\{f\in[0,\infty)\mid|f(kx)|\le\varepsilon\}.$$

由f连续可知Z_n是闭集，并且条件表明$[0,\infty)=\bigcup_{n=1}^{\infty}Z_n$。根据Baire纲定理，存在$N\in\mathbb{N}$和非空开区间$(a,b)$使得$(a,b)\subset Z_N$，于是只要$n\ge N$，则对任意$x\in(na,nb)$都有$|f(x)|<\varepsilon$。再注意到当$n\ge M:=\max\{N,\lceil\frac{a}{b-a}\rceil\}$时，有$(n+1)a<nb$，从而$(Ma,\infty)=\cup_{n=M}^{\infty}(na,nb)$，所以当$x>Ma$时$|f(x)|<\varepsilon$。结合$\varepsilon$的任意性即证。 □

注. *考虑开集$\{x>0\mid|f(x)|>\varepsilon\}$可以将(d)翻译为初等问题：任给开集$U\subset(0,\infty)$，记*

$$U/n:=\{x/n\mid x\in U\},\quad n\in\mathbb{N},$$

则U无界当且仅当

$$\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}U/k\ne\emptyset.$$

左边被称为集合列U/n的上极限，由所有在无穷多个U/n中出现的元素组成。但这个初等问题处理起来并不比原问题更容易，最直接的方式也是借助Baire纲定理：当U无界时，注意到$\bigcup_{k=n}^{\infty}U/k$是稠密开集（稠密性本质上来自于上面证明的最后一步），所以可数交一定稠密。

题 5.21 *. 设$f\colon\mathbb{R}\to\mathbb{R}$连续，且每个点$x\in\mathbb{R}$都是极值点，证明：$f$是常数。

证明. 称$x\in\mathbb{R}$是f的局部常值点，如果存在$\delta>0$，使得f在$(x-\delta,x+\delta)$上是常数。记全体局部常值点的集合为C，则C显然是开集。对于$x\in C$，存在最大的包含x的开区间使得f在其上

是常数，称为x所在的最大常值区间。[实际上，区间端点的函数值也等于$f(x)$，但是端点一定不在C中。当然，最大常值区间可以是一端或两端无限的，但这不影响论述。]

若能证明$C=\mathbb{R}$，则易知f是常数（思考理由！）。下面反设$C\neq\mathbb{R}$，对任意$n\in\mathbb{R}$，定义

$$U_n:=\left\{x\in\mathbb{R}\setminus C \mid f(x)\geq f(y),\ \forall y\in\left(x-\frac{1}{n},x+\frac{1}{n}\right)\right\},$$

$$D_n:=\left\{x\in\mathbb{R}\setminus C \mid f(x)\leq f(y),\ \forall y\in\left(x-\frac{1}{n},x+\frac{1}{n}\right)\right\}.$$

由连续性可知U_n,D_n都是闭集，以U_n为例：如果$x_k\in U_n\to x$，则因为$x_k\notin C$，容易看出$x\notin C$；其次，若$|x-y|<\frac{1}{n}$，则存在$N>0$使得当$k>N$时$|x_k-y|<\frac{1}{n}$，从而$f(x_k)\geq f(y)$，令$k\to\infty$并利用连续性得$f(x)\geq f(y)$，故$x\in U_n$。因为每个点都是f的极值点，所以

$$\mathbb{R}\setminus C=\left(\bigcup_{n=1}^{\infty}U_n\right)\cup\left(\bigcup_{n=1}^{\infty}D_n\right).$$

根据Baire纲定理，存在某个U_n或者D_n有内点，不妨设是U_n（从证明过程可以看出D_n的情况是一样的）。也就是说，存在非空开区间(a,b)满足$(a,b)\subset U_n\cup C$且$(a,b)\cap U_n\neq\emptyset$。不妨设$|b-a|<\frac{1}{n}$，则对任意两点$x,y\in(a,b)\cap U_n$有$f(x)\geq f(y)\geq f(x)$，所以$f|_{(a,b)\cap U_n}$是一个常数，记为$c$。[应当注意的是，因为$(a,b)\cap U_n\neq\emptyset$，所以$c$是良定义的。这一观察是本题的关键。]假设$x\in(a,b)$是任意一个点，断言$f(x)=c$。如果断言成立，则与$(a,b)\cap U_n\neq\emptyset$矛盾！事实上，如果$x\in U_n$，则自动有$f(x)=c$；而如果$x\in C$，考虑$x$所在的最大常值区间$(l,L)$，则$l$或$L\in(a,b)$，否则导致$(a,b)\subset(l,L)$，矛盾！不妨设$l\in(a,b)$（$L$的情况是一样的），则$l\in U_n$，所以$c=f(l)=f(x)$。于是反证假设不成立，原题得证。 □

题 5.22 *. 设$f\in C^\infty(\mathbb{R})$满足对任意$x\in\mathbb{R}$，都存在$n\in\mathbb{N}$使得$f^{(n)}(x)=0$。证明：$f$是多项式。[注意条件中的$n$是依赖于$x$的。]

证明. 类似于上一题，可以定义局部多项式点和最大多项式区间。记全体局部多项式点为P，则P也是开集，只需证明$P=\mathbb{R}$（思考理由！）。

对于$n\in\mathbb{N}$，定义

$$Z_n=\{x\in\mathbb{R}\setminus P \mid f^{(n)}(x)=0\}.$$

显然Z_n是闭集，并且$\mathbb{R}\setminus P=\bigcup_{n=1}^{\infty}Z_n$。根据Baire纲定理，存在$n\in\mathbb{N}$和非空开区间$(a,b)$，使得$(a,b)\subset Z_n\cup P$且$(a,b)\cap Z_n\neq\emptyset$。现在，关键在于$(a,b)\cap Z_n$不会有孤立点！否则存在$x\in(a,b)\cap Z_n$和$\delta>0$，使得$f$在$(x-\delta,x)$和$(x+\delta)$上都是多项式，结合$f\in C^\infty(\mathbb{R})$易知$f$在$(x-\delta,x+\delta)$上是多项式（思考理由！）。特别的，对任意$x\in(a,b)\cap Z_n$，存在一列$x_k\in(a,b)\cap Z_n(x_k\neq x)$收敛到$x$，导致$f^{(n+1)}(x)=\lim_{k\to\infty}\frac{f^{(n)}(x_k)-f^{(n)}(x)}{x_k-x}=0$，所以$(a,b)\cap Z_n\subset Z_{n+1}$。归纳可知对于$x\in(a,b)\cap Z_n$和任意$m\geq n$，都有$f^{(m)}(x)=0$。

由此可知对任意$x \in (a,b)\cap P$，f在x的邻域上一定是一个小于n次的多项式。这是因为假设x所在的最大多项式区间是(l,L)，则可以不妨设$l \in (a,b)$，从而$l \in Z_n$；因为k次多项式的k阶导数是非零常数，所以f在(l,L)上的次数一定小于n。由此可知，在(a,b)上恒有$f^{(n)} = 0$，导致$f|_{(a,b)}$是小于n次多项式，矛盾！ □

下面的最后一题说明，连续函数的逐点极限至少在一个稠密集上是连续的。似乎很难不借助Baire纲定理说明f哪怕有一个连续点。作为推论，Dirichlet函数不能写成连续函数的逐点极限。这一事实可以对比之前证明过的，Riemann函数是连续函数的逐点极限。

题 5.23 *. *设$f_n\colon \mathbb{R}\to\mathbb{R}$是一列连续函数，满足$\lim_{n\to\infty} f_n(x) =: f(x)$对任意$x\in\mathbb{R}$都收敛。证明：$f$的连续点集是稠密$G_\delta$集。*

证明. 任意取定$\varepsilon > 0$，记

$$G_\varepsilon := \bigcup_{n=1}^{\infty}\{x\in\mathbb{R} \mid \exists\delta>0,\ \mathrm{s.t.}\,\forall k,m\geq n,\ \forall y\in(x-\delta,x+\delta), |f_m(y)-f_k(y)|<\varepsilon\}.$$

仿照“连续函数的一致收敛极限是连续函数”的证明，可以说明$\bigcap_{n=1}^\infty G_{\frac{1}{n}}$中的点都是$f$的连续点（具体细节留作练习）。显然$G_\varepsilon$总是开集，下面只需证明$G_\varepsilon$稠密。

事实上，对任意内部非空的闭区间$[a,b]$，记

$$Z_n := \{x\in[a,b] \mid |f_m(x)-f_k(x)|<\varepsilon,\ \forall k,m\geq n\}.$$

则Z_n是闭集，且由逐点收敛知$[a,b]=\bigcup_{n=1}^\infty Z_n$。根据Baire纲定理，存在开集$(c,d)\subset[a,b]$和$n\in\mathbb{N}$使得$(c,d)\subset Z_n$，由此可知$(c,d)\subset G_\varepsilon$，证毕。 □

以上题目大致属于实变函数的范围。稍后在研究解析函数时，还会用Baire纲定理说明“光滑函数的Taylor级数收敛半径有一致的正下界就能保证解析性”。而Baire纲定理在泛函分析中的主要用途是证明Banach–Steinhaus定理，后者在分析学中有着诸多重要应用，比如“连续函数的Fourier级数未必逐点收敛”。最后给出在泛函分析中的一个微小应用：

题 5.24 *. *设$(X,\|\cdot\|)$是（$\mathbb{R}$或$\mathbb{C}$上的）完备赋范线性空间。如果$\{e_\alpha \mid \alpha\in A\}$是$X$的基，即$X$中的任意元素都可以唯一地表示为有限多个$e_\alpha$的线性组合。证明：$A$是不可数集。*

证明. 反设A可数，不妨设$A=\mathbb{N}$。记F_n是$e_1,\dots,e_n$张成的线性子空间，则可以证明F_n是没有内点的闭集（要说明F_n是闭集也许比看起来的更困难，但实际上拓扑线性空间中的任何有限维子空间一定是闭集），但$X=\bigcup_{n\in\mathbb{N}}F_n$，与Baire纲定理矛盾。 □

注. *上述线性代数意义下的基被称为Hamel基，先前已经用Zorn引理说明过Hamel基的存在性。泛函分析中通常允许无穷个基向量线性组合，或者说只要求张成的线性子空间稠密。在后一种意义下，很多常见的Hilbert空间都具有可数基。*

第6讲：导数的定义

她从你的睫毛上梳理出盐，并与你分享

她从你的时间里听出沙子，然后端在你面前

——保罗·策兰《睡眠和进餐》

这一讲主要关于导数的基本性质，并讨论了处处连续不可导的函数，其构造与分形密切相关。导数的计算已在作业中充分练习过，因此没有列举相关题目。另有部分题目或多或少应该放在学过中值定理之后，不必着急强行绕开中值定理而介绍过于技巧性的办法。

微分的定义现阶段还难以解释，而教材又不愿放过。今后在学习多元微积分时，会介绍光滑流形以及相关概念，届时将可以完整回答“什么是微分”。概括起来，光滑流形上可以定义切空间，任何光滑函数$f\colon M\to N$诱导出切映射（即微分）df，将x处的切向量v映为$f(x)$处的切向量$df(v)$，而导数（高维情形是Jacobi矩阵）就是切映射df在局部坐标下的矩阵表示。这里切映射的定义可以朴素理解如下：考虑M,N都是一维曲线，则M在x处的切向量v就是割线$y-x$的极限，被f映为N在$f(x)$处的割线$f(y)-f(x)$；极限情况下将v变成了$\lim_{y\to x}\frac{f(y)-f(x)}{y-x}$倍，这就与导数的定义相符。事实上，一维空间到自身的线性映射与实数一一对应，因此在一维情形df和$f'(x)$容易被混淆；高维情形下，线性映射与其关于一组基的矩阵表示不是一回事。

特别的，如果f是流形M到$\mathbb{R}$的映射，因为$\mathbb{R}$的切空间可以等同于$\mathbb{R}$，所以$df\colon TM\to\mathbb{R}$可以视为切空间TM的对偶空间中的元素。后者称为余切空间，记作T^*M。如此一来就可以理解一些题目中出现的微分等式的含义，比如“在$x^2+y^2=1$上$xdx+ydy=0$”，指的是$x^2+y^2=1$定义出一个光滑流形M（即单位圆），x,y都是M上的光滑函数，所以$dx,dy\in T^*M$有定义；T^*是一个线性空间，于是可以谈论$xdx+ydy$等于T^*M中的零元。一般的，给定方程组$f_i(x_1,\dots,x_n)=0\,(i=1,\dots,m)$，可以借助隐函数定理判断是否定义出$\mathbb{R}^n$中的光滑流形。

例题

例 6.1. 给定函数$f\colon\mathbb{R}\to\mathbb{R}$，判断以下命题是否成立：

(a) 如果f处处可导，则导函数连续。

(b) 如果当$x\to 0$时$f(x)=f(0)+o(|x|)$，则f在0处可导且$f'(0)=0$。

(c) 如果当$x\to 0$时$f(x)=f(0)+o(|x|^2)$，则f在0处二阶可导且$f''(0)=0$。

(d) 如果f具有连续的一阶导数且*(c)*的条件成立，则f在0处二阶可导且$f''(0)=0$。

解. (a) 错误，比如

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

则f处处可导，且

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

由此可见f'在0处不连续。

(b) 正确，这是导数的等价刻画。

(c) 错误，比如$f(x)=x^3D(x)$，其中D是Dirichlet函数，则f只在0处连续且可导，在其邻域上根本没有导数，当然不可能二阶可导。

(d) 错误，比如

$$f(x)=\begin{cases}0, & x=0,\\ x^3\sin\frac{1}{x}, & x\neq 0.\end{cases}$$

则f处处可导，且

$$f'(x)=\begin{cases}0, & x=0,\\ 3x^2\sin\frac{1}{x}-x\cos\frac{1}{x}, & x\neq 0.\end{cases}$$

由此可见f'连续且f在0处不二阶可导。 □

注. *(c)(d)采用直接构造分段函数的方式更简单。事实上，任给满足$g(0)=0$的严格单增连续函数，容易构造分段线性函数$f\colon(0,1)\to[0,\infty)$满足$f(x)=O(g(x))$且在0处不是二阶可导的：任取一列$x_n$严格单减趋于0，存在$y_n\in(x_{n+1},x_n)$使得$g(y_n)/(x_n-y_n)$关于$n$趋于$\infty$，定义$f(x)$是将这些点$(x_n,g(y_n)),(y_n,0)$连接形成的折线函数，则$0\le f\le g$且对于$t\in(y_n,x_n)$上$f'(t)=g(y_n)/(x_n-y_n)$在$n\to\infty$时趋于$\infty$，所以在0处不可能二阶可导。这样构造的$f$并不在0的邻域上可导，但只要将折线连接改进为光滑连接（见下一讲的讨论）就能解决。*

注. *之前用Baire纲定理证明过，连续函数的逐点极限一定在稠密G_δ集上连续。而导函数可以表示为一列连续函数的逐点极限，所以总有很多连续点。*

例 6.2. 判断$(0,1)$上的*Riemann*函数$R(x)$的可导性。

解. 因为$R(x)$只在无理点处连续，所以只需考虑无理点处是否可导。对任意$x\in(0,1)\setminus\mathbb{Q}$。一方面，可以取一列无理数$x_n\to x$，从而$x$处如果可导则导数只能是0；另一方面，可以取形

如$q_n = \frac{m}{n}$的有理数使得$|x - q_n| \le \frac{1}{n}$且根据定义有$R(q_n) \ge \frac{1}{n}$，因此

$$\left|\frac{R(x) - R(q_n)}{x - q_n}\right| = \frac{R(q_n)}{|x - q_n|} \ge \frac{\frac{1}{n}}{\frac{1}{n}} = 1,$$

令$n \to \infty$可知x处的导数不可能为0。综上，$R(x)$处处不可导。 □

注. 当x是无理数时，更精细的估计可以证明

$$\limsup_{y \to x} \left|\frac{R(x) - R(y)}{x - y}\right| = \infty.$$

事实上，借助抽屉原理可以证明，存在无穷多对互素整数p, q使得$|x - p/q| < 1/q^2$（这与先前关于无理测度的讨论并不矛盾，因为其定义只要求右边的分子为常数$c > 0$）；显然对于每个q至多有限个p满足要求，所以可以取一列这样的$y = p/q$满足$q \to \infty$，即$y \to x$。

这样的估计还说明$R^2(x)$也是处处不可导的函数。更为神秘的是$R^3(x)$存在可导的点。例如，如果无理数x是二次代数数，先前证明过存在常数$c > 0$使得对任意互素整数p, q都有[9]

$$\left|x - \frac{p}{q}\right| \ge \frac{c}{q^2}.$$

于是当$p/q \to x$时，必然有$q \to \infty$，从而

$$\left|\frac{R^3(p/q) - R^3(x)}{p/q - x}\right| \le \frac{1/q^3}{c/q^2} = \frac{1}{cq} \to 0.$$

这就说明R^3在二次代数无理数处可导，且导数为0。

例 6.3. *(a)* 设$f\colon \mathbb{R} \to \mathbb{R}$在$z$处可导，如果$x_n < z < y_n$且$x_n - y_n \to 0$，证明：

$$\lim_{n \to \infty} \frac{f(x_n) - f(y_n)}{x_n - y_n} = f'(z).$$

(b) 设$f\colon \mathbb{R} \to \mathbb{R}$在0处连续，给定实数$A \in \mathbb{R}$，证明：$f$在0处可导且$f'(0) = A$的充要条件是

$$\lim_{x \to 0} \frac{f(x) - f(x^2)}{x - x^2} = A.$$

证明. (a) 因为f在z处可导，并注意到$|x_n - z|, |y_n - z| < |x_n - y_n| \to 0$，可知

$$f(x_n) = f(z) + f'(z)(x_n - z) + o(|x_n - z|) = f(z) + f'(z)(x_n - z) + o(|x_n - y_n|),$$

$$f(y_n) = f(z) + f'(z)(y_n - z) + o(|y_n - z|) = f(z) + f'(z)(y_n - z) + o(|x_n - y_n|).$$

两式相减得$f(x_n) - f(y_n) = f'(z)(x_n - y_n) + o(|x_n - y_n|)$，这正是要证的结论。

[9]可以证明无理数x满足这一性质当且仅当连分数展开是有界的。另一方面，x是二次代数数当且仅当连分数展开从某一项起是周期的。相关结论可见M. L. Einsiedler and T. B. Ward, Ergodic theory with a view towards number theory, Springer, 2011第三章。

(b) 必要性与(a)方法相同，此时$x, x^2 \in O(|x-x^2|)$是因为x^2是x的高阶无穷小量。充分性同关于“如果f在0处连续，且$(f(x)-f(x/2))/x \to 0$，则$f'(0)=0$”的题目（但要注意当$x<0$时$|x-x^2| = |x|+|x|^2$，这一点不影响证明）。 $\square$

注. 如果*(a)*的条件改为x_n, y_n都在z的同侧，则结论未必成立。比如说，令f在0的邻域上可导，但导函数在0处不连续，则可以取一列$x_n \to 0$使得$f'(x_n)$不趋于$f'(0)$，再取y_n充分接近x_n，则可以让$\frac{f(x_n)-f(y_n)}{x_n-y_n}$与$f'(x_n)$相差无几，从而不会收敛到$f'(0)$。如果再加上$f$的导函数连续，则根据*Lagrange*中值定理可知结论依然成立。从证明可以看出，问题出在$|x_n-z|$和$|y_n-z|$未必属于$O(|x_n-y_n|)$；在*(b)*的情况下，由于二者不是同阶无穷小量，这个障碍不会出现。

练习题

题 6.4. *(a)* 设I是包含0的开区间，$f\colon I \to \mathbb{R}$在0处可导，求极限

$$\lim_{n\to\infty}\left(f\left(\frac{1}{n^2}\right)+f\left(\frac{2}{n^2}\right)+\cdots+f\left(\frac{n}{n^2}\right)-nf(0)\right).$$

(b) 计算$\lim_{n\to\infty}\left(1+\frac{1}{n^2}\right)\left(1+\frac{2}{n^2}\right)\cdots\left(1+\frac{n}{n^2}\right)$。

解. (a) 因为$\frac{1}{n^2},\dots,\frac{n}{n^2} \le \frac{1}{n} \to 0$，所以对$k=1,\dots,n$有

$$f\left(\frac{k}{n^2}\right) = f(0)+\frac{k}{n^2}f'(0)+o\left(\frac{1}{n}\right).$$

因此（注意上面的$o(1/n)$关于不同的k可以取成同一个量，乘以n是$o(1)$，见注）

$$\sum_{k=1}^{n} f\left(\frac{k}{n^2}\right)-nf(0) = \frac{n+1}{2n}f'(0)+o(1) \to \frac{f'(0)}{2}.$$

(b) 令$f(x)=\ln(1+x)$，利用(a)再取指数可知答案是$\sqrt{e}$。 $\square$

注. 现在对*(a)*中“n个$o(1/n)$相加是$o(1)$”进行严格的讨论。首先，如果仅仅要求下三角矩阵$a_{n,k}\,(1 \le k \le n)$满足对任意k都有$\lim_{n\to\infty} na_{n,k}=0$，则未必有$\lim_{n\to\infty}\sum_{k=1}^{n} a_{n,k}=0$。反例只需让$a_{n,n}=1$而其他项都为0。因此通常来讲未必$n$个$o(1/n)$相加是$o(1)$，问题本质上来源于不一致收敛：如果条件改为

$$\lim_{n\to\infty} n \max_{1\le k\le n}|a_{n,k}|=0,$$

则显然有$\sum_{k=1}^{n} a_{n,k} \to 0$。

原题证明的正确性在于这些$o(1/n)$是同一个量：如果$a_n=o(1/n)$，当然$na_n=o(1)$。根据条件f在0处可导，存在单增函数φ满足$\varphi(x)=o(|x|)$且$|f(x)-f(0)-xf'(0)| \le \varphi(|x|)$；事

实上，只要定义$\varphi(t)$是左边关于$|x| \le t$的上确界即可。于是证明中的每个$o(1/n)$都可以实现为与k无关的$\varphi(1/n)$，从而证明是严谨的。

题 6.5. 设$f\colon \mathbb{R} \to \mathbb{R}$是连续函数，证明：

(a) 如果对任意$x \in \mathbb{R}$有（即右上导数非负）

$$\limsup_{y \to x+} \frac{f(y) - f(x)}{y - x} \ge 0,$$

则f单调递增。

(b) 如果对任意$x \in \mathbb{R}$有（即右导数为0）

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

则f是常值函数。

*(c)** 如果对任意$x \in \mathbb{R}$有（即对称导数为0）

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

则f是常值函数。

*(d)** 如果存在至多可数集$I \subset \mathbb{R}$，使得当$x \notin I$时f可导且$f'(x) = 0$，则f是常值函数。

证明. (a) 因为对自变量和函数值进行平移都不影响条件和结论，只需在$f(0) = 0$的前提下证明当$x > 0$时$f(x) \ge 0$。取定$x > 0$，对任意$\varepsilon > 0$，考虑集合

$$X := \{t \in [0, x] \mid f(t) \ge -\varepsilon t\}.$$

则X是闭集且$0 \in X$，因此$\sup X \in X$。断言$\sup X = x$，若不然，则存在$s \in (\sup X, x)$使得$\frac{f(s) - f(\sup X)}{s - \sup X} \ge -\varepsilon$，即

$$f(s) \ge f(\sup X) - \varepsilon(s - \sup X) \ge -\varepsilon \cdot \sup X - \varepsilon(s - \sup X) = -\varepsilon s.$$

于是$s \in X$，与$\sup X$的定义矛盾！因此$f(x) \ge -\varepsilon x$，再结合ε任意性得$f(x) \ge 0$。

(b) 由(a)可知f和$-f$都是单增函数，从而f是常数。

(c) 同(a)的开头以及(b)中的论述可知，只需在$f(0) = 0$的条件下证明当$x > 0$时$f(x) \ge 0$。方法与(a)类似，只是现在要考虑的集合是

$$X := \{t \in [0, x] \mid \forall y \in [0, t],\ f(y) \ge -\varepsilon - \varepsilon y\}.$$

这里额外的$-\varepsilon$保证了X包含充分小的$t>0$（从而$\sup X>0$）。具体细节留作练习。

(d) 同(c)，只需在$f(0)=0$的条件下证明当$x>0$时$f(x)\geq 0$。设$I=\{x_n \mid n\in\mathbb{N}\}$，定义

$$X:=\{t\in[0,x] \mid f(t)\geq -\varepsilon t-\sum_{x_n<t}2^{-n}\varepsilon\}.$$

容易证明$\sup X\in X$（虽然X通常不是闭集），这里的新意在于反设$\sup X<x$之后，如果$\sup X=x_n\in I$，则需要利用当$t>x_n$时后一项求和中多出了一项$2^{-n}\varepsilon$，结合连续性可知当$t-x_n$充分小时仍有$t\in X$，才能得到矛盾。具体细节留作练习。 □

上一题的(a)也可以由之前的一道习题得到：如果对任意x和任意$\delta>0$，都存在$y\in(x,x+\delta)$使得$f(x)\leq f(y)$，则f单调递增[10]；易见这一条件对$f(x)+\varepsilon x$成立，再由ε的任意性可知f单增。可以看见辅助函数$f(x)+\varepsilon x$的作用是将不严格的不等式转化为严格不等式。这种方法在偏微分方程领域称为构造闸函数(barrier function)，常被用于建立强极值原理。

下一题的方法与之类似。简单来说，记条件中的算子为$D^2f(x)$，则从定义可以看出，如果$D^2f(x)>0$，那么f不能在区间内部取到最大值点；如果$D^2f(x)<0$，那么f不能在区间内部取到最小值点；扰动之后的函数$f(x)\pm\varepsilon x^2$满足这样的严格不等式，所以最值点只能在端点处取到，从而只能是一次函数。今后还会给出一个构造非平凡闸函数的例子。

题 6.6 *. 设$f\colon\mathbb{R}\to\mathbb{R}$连续，且对任意$x\in\mathbb{R}$有

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2}=0.$$

证明：f是一次函数。

证明. 只需证明在任何闭区间$[a,b]$上f都是一次函数（思考理由）。对任意$\varepsilon>0$，考虑

$$g(x):=f(x)-f(a)-\frac{f(b)-f(a)}{b-a}(x-a)+\varepsilon(x-a)(x-b),$$

则$g(a)=g(b)=0$。下面说明$g(x)\leq 0$，若不然，g在$[a,b]$上的最大值点x_0落在(a,b)中，于是当h充分小时

$$\frac{g(x+h)+g(x-h)-2g(x)}{h^2}\leq 0.$$

另一方面，根据条件可知

$$\lim_{h\to 0}\frac{g(x+h)+g(x-h)-2g(x)}{h^2}=\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2}+\varepsilon=\varepsilon>0,$$

矛盾！所以$g(x)\leq 0$，再结合ε的任意性可知

$$f(x)\leq f(a)+\frac{f(b)-f(a)}{b-a}(x-a),\quad \forall x\in[a,b].$$

[10]用(c)的方法还可以推广：如果f连续且对任意x和$\delta>0$，存在$x-\delta<y<x\leq z<x+\delta$使得$f(y)\leq f(z)$，则$f$单增。

类似的，考虑

$$h(x) := f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a) - \varepsilon(x-a)(x-b),$$

可知$h(x) \geq 0$，即$f(x) \geq f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$，从而等号成立。 □

注. 定义一阶对称差分算子$\Delta_h f(x) = f(x+h) - f(x-h)/h$，则复合之后得到$n$阶差分算子$D_h^n$。本题以及上一题的*(c)*说明如果用$D^n f := \lim_{h\to 0} \Delta^n h$表示$n$阶对称导数，则对于连续函数$f$而言，$Df = 0$等价于$f$是常数（零次多项式），$D^2 f = 0$等价于$f$是一次多项式。自然猜想对任意$n$都有$D^n f = 0$蕴含$f$是$n-1$次多项式。然而结论对$n = 3$已经不成立，反例只需让$f$是偶函数且在$[0,\infty)$上是二次函数，比如令$f(x) = (|x|-1)^2$。

在学过定积分之后会讨论分布意义下导数的定义，届时将说明n阶分布导数是0等价于（至多）$n-1$次多项式。分布导数的另一个优势在于任何分布都可以求导。

处处连续处处不可导的函数

这一小节首先给出一种具体构造，用无穷级数定义比经典的三分法定义Weierstrass函数在技术上略微容易；然后用Baire纲定理说明$[0,1]$上处处连续处处不可导的函数实际上有很多，构成$(C([0,1]), \|\cdot\|_{L^\infty})$中的稠密$G_\delta$集。

题 6.7. 定义$\varphi_0(x)$以1为周期且当$|x| \leq \frac{1}{2}$时$\varphi_0(x) = |x|$，以及$\varphi_n = 2^{-n}\varphi_0(2^n x)\,(n \in \mathbb{N})$。易见$\varphi_n$是最小正周期为$2^{-n}$的连续周期函数，且取值范围是$[0, 2^{-n-1}]$。定义

$$\Phi(x) := \sum_{n=0}^{\infty} \varphi_n(x).$$

根据大M判别法可知$\Phi\colon \mathbb{R} \to \mathbb{R}$连续。本题的目标是说明*Hölder*连续函数可能处处不可导。

(a) 给定$x \in \mathbb{R}$，对任意$n \in \mathbb{N}$，定义

$$a_n = 2^{-n-1}\lfloor 2^{n+1}x \rfloor, \quad b_n = a_n + 2^{-n-1}.$$

证明：$\frac{\Phi(b_n)-\Phi(a_n)}{b_n-a_n}$是整数，且与$n$的奇偶性相反。由此得出结论：$\Phi$处处不可导。

*(b)** 证明：Φ不是*Lipschitz*连续函数，但对任意$\alpha \in (0,1)$都是α阶*Hölder*连续函数。

证明. (a) 一方面，对任意$0 \leq k \leq n$，都有φ_n在$[a_n, b_n]$上是斜率为± 1的一次函数；另一方面，对$k > n$都有$b_n - a_n = 2^{-n-1}$是φ_k的周期。因此

$$\frac{\Phi(b_n)-\Phi(a_n)}{b_n-a_n} = \sum_{k=0}^{n} \frac{\varphi_k(b_n)-\varphi_k(a_n)}{b_n-a_n}.$$

右边的每一项都是± 1，从而求和还是整数，且奇偶性与$n+1$相同。

最后，如果Φ在x处可导，则根据先前的例题可知$\frac{\Phi(b_n)-\Phi(a_n)}{b_n-a_n} \to \Phi'(x)$；但一列奇偶性交错的整数列不可能收敛，矛盾！

(b) 考虑$x_n = 2^{-n}$，则对$0 \le k \le n-1$都有$\varphi_k(x_n) = x_n$，所以

$$\Phi(x_n) \ge \sum_{k=0}^{n-1} \varphi_k(x_n) - \sum_{k=n}^{\infty} |\varphi_k(x_n)| \ge nx_n - \sum_{k=n}^{\infty} 2^{-k-1} = nx_n - 2^{-n} = (n-1)x_n.$$

因此$|\Phi(x_n) - \Phi(0)|/|x_n| \ge n-1$，这说明$\Phi$不是Lipschitz连续函数。

另一方面，任意取定$\alpha \in (0,1)$，要证明Hölder连续性，因为Φ是有界函数，只需考虑$|x-y| \le 1/2$的情况。取$n \in \mathbb{N}$待定，因为每个φ_k都是1-Lipschitz连续函数，所以

$$|\Phi(x) - \Phi(y)| \le \sum_{k=0}^{n-1} |\varphi_k(x) - \varphi_k(y)| + \sum_{k=n}^{\infty} |\varphi_k(x)| + \sum_{k=n}^{\infty} |\varphi_k(y)| \le n|x-y| + 2^{-n+1}.$$

现在令$n = \lfloor -\log_2 |x-y| \rfloor$，则$2^{-n} \le C|x-y|$。注意到由于$\alpha \in (0,1)$，存在$C > 0$使得对$t \in (0,1/2]$有$-t\log_2 t \le Ct^\alpha$（等价于$\lim_{s\to 0+} s\log s = 0$）。故当$|x-y| \le 1/2$时有

$$|\Phi(x) - \Phi(y)| \le -|x-y|\log_2|x-y| + C|x-y| \le C|x-y|^\alpha. \qquad \square$$

注. 今后会证明“单调函数、进而*Lipschitz*连续函数几乎处处可导”，本题表明任意$\alpha \in (0,1)$阶的*Hölder*连续函数都可以处处不可导。此外，这一结论还说明处处连续处处不可导的函数不可能有单调区间，即在任何（内部非空的）区间上都不单调。

处处连续但处处不可导的函数在连续函数中是很多的，相关的例子包括随机过程中的Brown运动以及动力系统中的分形，后者的经典例子是Koch雪花曲线。

记$I := [0,1]$，回忆在$C(I)$上可以引进度量

$$d(f,g) = \|f-g\|_{L^\infty} := \sup_{[0,1]} |f-g|,$$

使得$(C(I),d)$成为完备度量空间。I上的分段线性函数（或者说折线函数）是指存在划分$0 = x_0 < \cdots < x_n = 1$使得其在$[x_{i-1},x_i]\,(i=1,\ldots,n)$上都是一次函数。

题 6.8*. 本题的目标是说明处处不可导的函数至少包含$C(I)$中的稠密G_δ集。证明：

(a) 任给$\varepsilon > 0$和$f \in C(I)$，存在分段线性函数s，使得$\|f-s\|_{L^\infty} \le \varepsilon$。进一步，还可以要求$s$每一段斜率的绝对值大于等于任意给定的正数。

(b) 对任意$n\in\mathbb{N}$，定义$X_n\subset C(I)$如下：$f\in X_n$当且仅当$f\in C(I)$且存在$\varepsilon>0$，使得

$$\forall x\in I,\ \exists y\in I,\ \text{s.t.}\,|x-y|<\frac{1}{n}\ \text{且}\ |f(x)-f(y)|\geq n|x-y|+\varepsilon.$$

则X_n是稠密开集。

(c) 稠密G_δ集$Y:=\bigcap_{n=1}^\infty X_n$中的函数都处处不可导。事实上，对任意$f\in Y$有

$$\limsup_{y\to x}\left|\frac{f(x)-f(y)}{x-y}\right|=\infty,\quad \forall x\in I.$$

证明. (a) 根据一致连续性，存在$n\in\mathbb{N}$使得当$|x-y|\leq 1/n$时$|f(x)-f(y)|\leq\varepsilon/2$。定义$s$是将$(i/n,f(i/n))\,(i=0,1,\ldots,n)$用折线连接形成的分段线性函数，则$s(i/n)=f(i/n)$，且对任意$(i-1)/n\leq x\leq i/n$都有

$$\begin{aligned}|f(x)-s(x)|&\leq|f(x)-f(i/n)|+|f(i/n)-s(i/n)|+|s(i/n)-s(x)|\\&\leq\varepsilon/2+0+|s(i/n)-s((i-1)/n)|=\varepsilon/2+|f(i/n)-f((i-1)/n)|\leq\varepsilon.\end{aligned}$$

这里用到了s在$[(i-1)/n,i/n]$上是一次函数，从而任何两点的差不超过端点处的差。

最后，若要求s的每一段斜率绝对值都至少是给定的$M>0$，则只需将连接$((i-1)/n,f((i-1)/n))$与$(i/n,f(i/n))$的线段换成（有限多段）斜率至少为M的折线，并保证s在这个区间上的最大、最小值之差不超过$\varepsilon/2$即可。具体细节留作练习。

[在早期版本的讲义中，本题错误地加上了“s的每一段斜率都等于$\pm M$”的限制。容易看出，能够这样被逼近的函数当且仅当是M-Lipschitz连续的。]

(b) 先证明X_n是开集。设$f\in X_n$，对应的常数是ε。假设$g\in C(I)$满足$\|f-g\|_{L^\infty}<\frac{\varepsilon}{3}$，则对任意$x\in I$，存在$y\in I$使得$|f(x)-f(y)|\geq n|x-y|+\varepsilon$，从而

$$|g(x)-g(y)|\geq|f(x)-f(y)|-|g(x)-f(x)|-|g(y)-f(y)|\geq n|x-y|+\varepsilon-\frac{2}{3}\varepsilon=n|x-y|-\frac{\varepsilon}{3}.$$

这就说明$g\in X_n$，从而X_n是开集。

下证X_n稠密。这是因为每段斜率绝对值至少为$n+1$的分段线性函数一定包含在X_n中（具体细节留作练习，尤其需要说清楚ε只依赖函数本身而与x,y无关），再结合(a)即证。

(c) Baire纲定理保证了Y是稠密G_δ集。对于$f\in Y$和任意$x\in I$，条件表明存在一列$y_n\to x$和$\varepsilon_n>0$使得$|f(x)-f(y_n)|\geq n|x-y_n|+\varepsilon_n\geq n|x-y_n|$，这就说明

$$\limsup_{y\to x}\left|\frac{f(x)-f(y)}{x-y}\right|=\infty.\qquad\square$$

第7讲：微分中值定理（上）

> 然后他又怀疑自己的演算能力，打算开一个平方根来证明一下。但是他偏巧选择了2来开平方，结果发现开起来无穷无尽，……王仙客在被撵出宣阳坊之前，正手持一根竹竿，竹竿头上栓了一支毛笔，在天花板上写算式哪。据我所知，他是用麦克劳林(Maclaurin)级数开平方，已经算到了第五千项。
>
> ——王小波《寻找无双》

这一讲是微分中值定理的初步应用，主要涉及Lagrange中值定理以及Taylor公式；此外，在应对一些较弱的条件时，Darboux中值定理以及反证法十分有效。两道例题简要说明了为何对Cauchy中值定理以及L'Hôpital法则的讨论较少：取巧的辅助函数构造并不本质，强调相关技巧反而不明所以。然而，读者也不应该误以为Cauchy中值定理是不重要的；恰恰相反，Cauchy中值定理能够证明L'Hôpital法则以及建立Taylor公式余项表达式，后者在计算数学中用于误差估计。最后补充了关于解析函数的内容，所谓“解析函数”即每一点处的Taylor级数都在一个小邻域上收敛到自身的光滑函数。

例题

例 7.1. 给定$x \in \mathbb{R}$，$\delta > 0$和$A \in \mathbb{R}$，如果f在$[x_0, x_0+\delta)$上连续、在$(x_0, x_0+\delta)$上可导，且

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

证明：f在x_0处右可导，且右导数等于A。

证明. 根据Lagrange中值定理，对任意$x > x_0$，存在$\xi \in (x_0, x)$使得

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi).$$

当$x \to x_0+$时，有$\xi \to x_0+$，因此上式极限为A。 □

注. 如果已知f在x_0处右可导，且导数在x_0处的右极限存在，则*Darboux*中值定理保证了

$$f'_+(x_0) = f'(x_0+).$$

（注意区分记号。）本题说明假设x_0处的可导性是多余的。这实际上是*L'Hôpital*法则的特例。

记号$C^\infty(\mathbb{R})$表示$\mathbb{R}$上的无穷次可导函数，$C^n(\mathbb{R})(n \in \mathbb{N})$表示直到$n$阶可导且$n$阶导数连续的函数，$C(\mathbb{R})$或$C^0(\mathbb{R})$表示全体连续函数。

例 7.2. *(a)* 定义$f\colon \mathbb{R}\to\mathbb{R}$为

$$f(x) := \begin{cases} 0, & x \le 0, \\ e^{-\frac{1}{x}}, & x > 0. \end{cases}$$

证明：$f \in C^\infty(\mathbb{R})$。

(b) 构造$g \in C^\infty(\mathbb{R})$满足$g(x) \in [0,1]$且

$$g(x) = \begin{cases} 0, & x \le 0, \\ 1, & x \ge 1. \end{cases}$$

(c) 构造$h \in C^\infty(\mathbb{R})$满足$h(x) \in [0,1]$且

$$h(x) = \begin{cases} 0, & |x| \ge 2, \\ 1, & |x| \le 1. \end{cases}$$

证明. (a) 当$x \ne 0$时显然f无穷次可导，并且当$x > 0$时导函数形如

$$f^{(n)}(x) = P_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}},$$

其中P_n是多项式。这就说明当$x \to 0+$时$f^{(n)}(x) \to 0$，因此0处的各阶右导数都是0，又显然左导数也是0，所以$f \in C^\infty(\mathbb{R})$。[严格来讲这里应该对阶数归纳，需要用到$n-1$阶导数连续，才能由$\lim_{x\to 0} f^{(n)}(x) = 0$说明0处$n$阶可导且$f^{(n)}(0) = 0$。]

(b) 注意到$f(x) + f(1-x) > 0$，所以$g(x) := \frac{f(x)}{f(x)+f(1-x)} \in C^\infty(\mathbb{R})$并且满足要求。

(c) 将(b)中的函数g拼接起来即可。具体来说，

$$h(x) = \begin{cases} g(x+2), & x \le 0, \\ g(2-x), & x > 0. \end{cases}$$

注意在0的邻域上h恒等于1，所以在0附近也是C^∞函数。 □

注. *(a)*中的函数f在0处各阶导数都是0，但当$x > 0$时0处的*Taylor*级数不收敛到$f(x)$。这给出光滑函数不解析的例子，说明*Taylor*公式中的余项$o(|x|^n)$只是随$|x|$趋于0而并非随n趋于0。

注. *(c)*中的函数h称为截断函数*(cutoff function)*，在处理逼近、延拓等问题时非常重要。例如：*(1)*任给紧集$F \subset \mathbb{R}$和开集$U \supset F$，存在$\varphi \in C^\infty(\mathbb{R})$使得$\varphi|_F = 1$且$\varphi|_{U^C} = 0$（注意$C$无非是$F = [-1,1]$且$U = (-2,2)$的特例）。但证明最好留到定积分与卷积之后，或者借助多元微分

学中的单位分解。*(2)*任给数列a_n，构造$f \in C^\infty(\mathbb{R})$使得$f^{(n)} = a_n$。这一命题同样不适合在现在处理，而应该留到学完一致收敛与逐项求导的关系之后。

例 7.3. 设$a \in \mathbb{R}$，求$\lim_{x\to\infty} e^{-ax}\left(1+\frac{a}{x}\right)^{x^2}$。

解. 利用$\log\left(1+\frac{a}{x}\right) = \frac{a}{x} - \frac{a^2}{2x^2} + o\left(\frac{1}{x^2}\right)$，有

$$\lim_{x\to\infty} e^{-ax}\left(1+\frac{a}{x}\right)^{x^2} = \lim_{x\to\infty} e^{-ax+x^2\left(\frac{a}{x}-\frac{a^2}{2x^2}+o\left(\frac{1}{x^2}\right)\right)} = \lim_{x\to\infty} e^{-\frac{a^2}{2}+o(1)} = e^{-\frac{a^2}{2}}. \qquad \square$$

下面两例说明某些辅助函数构造本质上来自于定积分，因此这些问题并不适合在现阶段深入讨论。各种习题集乐于强调这样的技巧而不讲解背后的道理，现在新教材也将这些内容放入了例题。而对于真正的科研来说，这些所谓的技巧几乎没有用处。

例 7.4. 设$f\colon (0,\infty) \to \mathbb{R}$可导，且$\sqrt{x}f'(x)$有界，证明：$f$一致连续。

证明. 根据Cauchy中值定理，对任意$0 < x < y$，存在$\xi \in (x,y)$使得

$$\frac{f(x)-f(y)}{\sqrt{x}-\sqrt{y}} = \frac{f'(\xi)}{\frac{1}{2\sqrt{\xi}}} = 2\sqrt{\xi}f'(\xi).$$

记$M := \sup_{x\in(0,\infty)} |\sqrt{x}f'(x)| < \infty$，则

$$|f(x)-f(y)| \le M|\sqrt{x}-\sqrt{y}| \le M\sqrt{|x-y|}.$$

由此可知f一致连续。 $\square$

注. 一般的，如果g一致连续且g'非零，满足$\frac{f'}{g'}$有界，则同样的证明可以推出f也一致连续。若进一步假设f'和g'连续，并且由g'非零可以不妨设$g' > 0$，则对任意$0 < x < y$有

$$|f(y)-f(x)| = \left|\int_x^y f'(t)\,dt\right| \le C\int_x^y g'(t) = C(g(y)-g(x)).$$

有了这一观察，反过来考虑用*Cauchy*中值定理就是自然的。

例 7.5. *(a)* 设$f\colon (0,\infty) \to \mathbb{R}$可导，若$\lim_{x\to\infty}(f(x)+f'(x)) = 0$，证明：$\lim_{x\to\infty} f(x) = 0$。

*(b)** 设$f\colon (0,\infty) \to \mathbb{C}$可导，若$\lim_{x\to\infty}(f'(x)-zf(x)) = 0$，其中$z \in \mathbb{C}$是满足$\mathrm{Re}(z) < 0$的常数，证明：$\lim_{x\to\infty} f(x) = 0$。

*(c)** 设$n\,(n \ge 1)$次复系数多项式$P(z) = c_nz^n + \cdots + c_1z + c_0\,(c_n \ne 0)$的$n$个复根的实部都严格小于0，证明：如果$n$阶可导函数$f\colon (0,\infty) \to \mathbb{C}$满足

$$\lim_{x\to\infty}(c_nf^{(n)}(x) + \cdots + c_1f'(x) + c_0f(x)) = 0,$$

证明：对$k = 0,1,\ldots,n$都有$\lim_{x\to\infty} f^{(k)}(x) = 0$。

证明. (a) 巧妙而经典的方法是补上e^x再用L'Hôpital法则，可知

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{e^x f(x)}{e^x} = \lim_{x\to\infty} \frac{e^x(f(x)+f'(x))}{e^x} = 0.$$

(b) 直观上看可以仿照(a)的证明，只需将e^x替换为e^{-zx}，此时$\mathrm{Re}(z) < 0$保证了$|e^{-zx}| \to \infty$；然而问题在于中值定理未必对复值函数仍然成立，读者可以自行思考能否建立相应的理论。下面给出的方法本质上来自于微分方程中的能量法。

记$c = -\mathrm{Re}(z) > 0$，注意到

$$\begin{aligned}\frac{d}{dx}|f|^2 &= 2\,\mathrm{Re}(\overline{f}f') = 2\,\mathrm{Re}(\overline{f}(zf + o(1))) \\ &= 2\,\mathrm{Re}(z|f|^2) + 2\,\mathrm{Re}(\overline{f}\cdot o(1)) \le -2c|f|^2 + 2|\overline{f}\cdot o(1)|.\end{aligned}$$

利用均值不等式，有$2|f\cdot o(1)| \le c|f|^2 + o(1)/c$，这就证明了

$$\frac{d}{dx}|f|^2 + c|f|^2 \le o(1).$$

这结合非负性已经足够推出$|f|^2 \to 0$，本质上是复述L'Hôpital的证明。具体来说，根据$c > 0$以及Cauchy中值定理可知，对任意$r > 0$，都有

$$\begin{aligned}\limsup_{x\to\infty} |f|^2(x) &= \limsup_{x\to\infty} \frac{e^{cx}|f|^2(x) - e^{cr}|f|^2(r)}{e^{cx} - e^{cr}} \\ &\le \sup_{\xi > r} \frac{e^{c\xi}\left(\frac{d}{dx}|f|^2(\xi) + c|f|^2(\xi)\right)}{e^{c\xi}} \\ &= \sup_{\xi > r}\left(\frac{d}{dx}|f|^2(\xi) + c|f|^2(\xi)\right).\end{aligned}$$

右边关于r的上极限不超过0，这就证明了$|f|^2 \to 0$，从而$f \to 0$。

(c) 不妨设$P(z) = \prod_{k=1}^n (z - z_k)$，其中$\mathrm{Re}(z_k) < 0$。对$n$归纳，$n = 1$已经在(b)中解决。假设结论对$n-1$成立，考虑$n \ge 2$的情况，关键在于考虑$Q(x) = \prod_{k=1}^{n-1}(z - z_k)$。条件可以写为$P(D)f \to 0$，其中$D$表示求导算子，于是

$$\frac{d}{dx}(Q(D))f - z_n Q(D)f \to 0.$$

利用归纳假设可知$Q(D)f \to 0$，进而$f, f', \dots, f^{(n-1)} \to 0$，再代回条件得$f^{(n)} \to 0$。 □

注. *如果已知f'连续，则(a)更直接的方法是解常微分方程。记$f(x) + f'(x) = g(x)$，解得*

$$f(x) = e^{-x}\left(f(0) + \int_0^x e^t g(t)\,dt\right).$$

利用*L'Hôpital*法则可知

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{f(0) + \int_0^x e^t g(t)\,dt}{e^x} = \lim_{x\to\infty} \frac{e^x g(x)}{e^x} = 0.$$

有了这一观察，反过来考虑乘以e^x再用*L'Hôpital*就是自然的。*(b)*也可以类似处理。

注. *(a)*的条件并不能由$f + f' \to 0$换成$f - f' \to 0$，比如有反例$f(x) = e^x$；*(b)*的条件也不能由$\mathrm{Re}(z) < 0$减弱为$\mathrm{Re}(z) \le 0$，比如有反例$f(x) = e^{ix}$。本质上这来自于微分方程的稳定性理论，*(c)*对应了常系数情形的一般结果：如果$P(x)$有复根z_0，则$f(x) = e^{iz_0 x}$满足$c_0 f + \cdots + c_n f^{(n)} = 0$，从而要保证$f \to 0$必须$\mathrm{Re}(z_0) < 0$。

练习题

下一题需要用到广义Rolle中值定理：如果f在$[a,\infty)$上连续、(a,∞)上可导，满足$f(a) = 0$和$\lim_{x\to\infty} f(x) = 0$，则存在$\xi \in (a,\infty)$使得$f'(\xi) = 0$。证明采用反证法，如果结论不成立，根据Darboux中值定理可知f'在(a,∞)上恒大于0或者恒小于0，从而f在$[a,\infty)$上严格单增或者严格单减，都与$f(a) = 0$且$\lim_{x\to\infty} f(x) = 0$矛盾。

题 7.6. 对于$n \in \mathbb{N}_0$，定义第n个*Hermite*多项式为

$$H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n}(e^{-x^2}).$$

(a) 计算H_0, H_1, H_2, H_3, H_4。

(b) 证明：$H_n(x)$是n次多项式，且有n个互异实根。

(c) 证明：对$n \in \mathbb{N}$有$H_n'(x) = 2nH_{n-1}(x)$。

(d) 证明：对$n \in \mathbb{N}$有$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$。

(e) 证明：对$n \in \mathbb{N}_0$有$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0$。

证明. (a) 按定义计算即可，

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2,$$

$$H_3(x) = 8x^3 - 12x, \quad H_4(x) = 16x^4 - 48x^2 + 12.$$

(b) 对$n \in \mathbb{N}$有

$$H_n(x) = -e^{x^2} \frac{d}{dx}(H_{n-1}(x)e^{-x^2}) = 2xH_{n-1}(x) - H_{n-1}'(x).$$

归纳可知H_n是n次多项式，从而至多n个实根。要证明恰好有n个实根，因为$e^{x^2} > 0$，只需证明$f(x) := e^{-x}$的n阶导函数有n个互异零点。这也来自于对n归纳：奠基是显然的，假设已经证明了$f^{(n)}(x)$有n个零点$x_1 < \cdots < x_n$，则由$f^{(n)}(x) = (-1)^n H_n(x)e^{-x^2}$可知$\lim_{x\to\pm\infty} f^{(n)}(x) = 0$，再根据广义Rolle中值定理可知$f^{(n+1)}$在$n+1$个区间

$$(-\infty, x_1), (x_1, x_2), \ldots, (x_{n-1}, x_n), (x_n, \infty)$$

上各至少有一个零点，这就证明了$n+1$的情况。

(c) 对H_n求导并利用Leibniz公式，得

$$\begin{aligned} H_n'(x) &= 2xH_n(x) + (-1)^n e^{x^2} \frac{d^{n+1}}{dx^{n+1}}(e^{-x^2}) \\ &= 2xH_n(x) + (-1)^n e^{x^2} \frac{d^n}{dx^n}(-2xe^{-x^2}) \\ &= 2xH_n(x) + (-1)^{n+1} 2e^{x^2} \left(x\frac{d^n}{dx^n}(e^{-x^2}) + C_n^1 \frac{d^{n-1}}{dx^{n-1}}(e^{-x^2}) \right) \\ &= 2xH_n(x) - 2xH_n(x) + 2nH_{n-1}(x) = 2nH_{n-1}(x). \end{aligned}$$

(d) 结合(b)的证明中给出的递推式以及(c)即可。

(e) $n = 0, 1$是显然的，$n \geq 2$的情况是(c)和(d)的推论：

$$\begin{aligned} H_n''(x) - 2xH_n'(x) + 2nH_n(x) &= 4n(n-1)H_{n-2}(x) - 4nxH_{n-1}(x) + 2nH_n(x) \\ &= 2n(H_n(x) - 2xH_{n-1}(x) + 2(n-1)H_{n-2}(x)) = 0. \quad \square \end{aligned}$$

注. 许多特殊多项式在数学和物理中都十分重要，主要意义是作为某些算子的特征函数或者某些加权L^2空间的标准正交基。*Hermite*多项式给出*Fourier*变换的特征函数，并且是加权空间$L^2(\mathbb{R}, e^{-x^2}dx)$的一组正交基，其应用之一是随机分析中*Wiener*混沌的构造。

下一题是将Lagrange中值定理的证明进行衍生。

题 7.7. *(a)* 设f在$[a,b]$上具有介值性质且在(a,b)上可导，证明：存在$\xi \in (a,b)$使得

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

(b) 设f在$[a,b]$上连续、在(a,b)上右可导，证明：存在$\xi, \eta \in (a,b)$使得

$$f'_+(\xi) \geq \frac{f(b) - f(a)}{b - a} \geq f'_-(\eta).$$

(c) 设f在$[a,b]$上连续、在(a,b)上可导，证明：如果f不是一次函数，则存在$c \in (a,b)$使得

$$|f'(c)| > \left| \frac{f(b)-f(a)}{b-a} \right|.$$

证明. (a) 考虑辅助函数

$$F(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a) - f(a).$$

则F在(a,b)上可导，只需证明F'有零点。若不然，根据Darboux中值定理，F在(a,b)上严格单调，因此极限$f(a+)$和$f(b-)$广义收敛（可能是$\pm\infty$）。因为f具有介值性质，只能$f(a+) = f(a)$且$f(b-) = f(b)$。也就是说$F(a+) = F(b-) = 0$，与F严格单调矛盾！

(b) 仍考虑(a)中的辅助函数，则F在$[a,b]$上连续、在(a,b)上右可导，只需证明由$F(a) = F(b) = 0$能够推出存在$\xi, \eta \in (a,b)$使得$F'_+(\xi) \geq 0 \geq F'_+(\eta)$。此时Fermat引理不成立（最大值点处可能右导数严格小于0），但可以借助先前的习题：如果右导数严格恒大于0，则F严格单增。由此可知，为了让$F(a) = F(b)$，右导数必然存在非负的点、也存在非正的点。

(c) 不妨设$f(b) \geq f(a)$。仍考虑(a)中的辅助函数F，此时F在$[a,b]$上连续、在(a,b)上可导。如果结论不成立，则对任意$c \in (a,b)$都有$F'(c) \leq 0$，与$F(a) = F(b) = 0$矛盾！ □

注. *(a)*的条件并不比*Lagrange*中值定理强太多，因为最坏的情况也只是f在端点a, b处可能不连续（比如考虑$f(x) = \sin 1/x \, (x \in (0,1))$并补充定义$f(0) = 0$）。特别的，根据*Darboux*中值定理，如果$f$是某个函数的导函数，则介值性质自动成立。因此对于一些关于n阶导数的中值定理题目，往往无需假设$n-1$阶导数在端点处也连续。

注. *(b)*中的不等式可能是严格的，比如$[-1,1]$上的函数$f(x) = |x|$。

题 7.8. *(a)* 设f在$[0,\infty)$上连续、在$(0,\infty)$上可导，满足$|f'(x)| \leq A|f(x)| \, (\forall x > 0)$，其中$A > 0$是常数。证明：如果$f(0) = 0$，则$f$恒等于0。

(b) 去掉*(a)*中的条件“$f(0) = 0$”，证明：$|f(x)| \leq e^{Ax}|f(0)|$。

证明. (a) 首先证明对$x \in \left[0, \frac{1}{2A}\right]$有$f(x) = 0$。事实上，假设$t$是$|f|$在$\left[0, \frac{1}{2A}\right]$上的最大值点，则根据Lagrange中值定理，存在$\xi \in (0,t)$使得

$$|f(t)| = |f(t) - f(0)| = |f'(\xi)| \cdot |t| \leq A|f(\xi)| \cdot \frac{1}{2A} \leq \frac{|f(t)|}{2}.$$

所以$|f(t)| = 0$。接下来，利用$f\left(\frac{1}{2A}\right) = 0$以及条件，可以证明$f$在$\left[\frac{1}{2A}, \frac{2}{2A}\right]$上恒等于0。归纳可知$f$在任何$\left[\frac{n-1}{2A}, \frac{n}{2A}\right] \; (n \in \mathbb{N})$上都恒等于0。

(b) $f(0)=0$的情况已经在(a)中处理。用$f(x)/|f(0)|$代替f，可以不妨设$f(0)=1$。取定$x>0$，则需要证明$|f(x)|\le e^{Ax}$。如果f在$[0,x]$上有零点t，则$x\mapsto f(t+x)$满足(a)的条件，从而$f(x)=0$。下面考虑f在$[0,x]$上没有零点的情况，此时f在$[0,x]$上恒大于0。考虑辅助函数$g(x)=e^{-Ax}f(x)$，则$g'(x)=e^{-Ax}(f'(x)-Af(x))$，由条件可知$g'(x)\le 0$，故

$$f(x)=e^{Ax}g(x)\le e^{Ax}g(0)=e^{Ax}. \qquad \square$$

注. 此即*Gronwall*不等式的微分形式，在微分方程领域十分常用，今后会用积分的方法重新理解。仿照*(a)*的证明可知，如果$f\in C^n([0,\infty))$满足

$$|f^{(n)}(x)|\le C(|f(x)|+|f'(x)|+\cdots+|f^{(n-1)}(x)|),$$

并且$f(0)=f'(0)=\cdots=f^{(n-1)}(0)=0$，则$f$恒等于0。这件事可以用于证明线性常微分方程初值问题解的唯一性，而无需借助经典的*Wronski*行列式方法。

下一题的技巧是用Taylor公式的线性组合消去无关项，这种方法在下一讲还会多次出现。

题 7.9. 设f在$[a,b]$上可导、在(a,b)上二阶可导，且满足$f'(a)=f'(b)=0$，证明：存在$c\in(a,b)$使得

$$|f''(c)|\ge\frac{4}{(b-a)^2}|f(b)-f(a)|.$$

证明. 考虑$\frac{a+b}{2}$在a,b处的Taylor展式，结合$f'(a)=f'(b)$可知

$$f\left(\frac{a+b}{2}\right)=f(a)+\frac{(b-a)^2}{8}f''(\xi)=f(b)+\frac{(b-a)^2}{8}f''(\eta),$$

其中$\xi\in\left(a,\frac{a+b}{2}\right)$且$\eta\in\left(\frac{a+b}{2},b\right)$。所以

$$\frac{4}{(b-a)^2}|f(b)-f(a)|\le\frac{|f''(\xi)-f''(\eta)|}{2}\le\sup\{|f''(\xi)|,|f''(\eta)|\}. \qquad \square$$

题 7.10. 设$f\colon\mathbb{R}\to\mathbb{R}$可导，且对任意$x,y\in\mathbb{R}$都有

$$f(x)-f(y)=f'\left(\frac{x+y}{2}\right)(x-y).$$

证明：f是二次函数。

证明. 将y视为固定的常数，则$f'\left(\frac{x+y}{2}\right)=\frac{f(x)-f(y)}{x-y}$的右边在$x\ne y$处可导，从而$f'$在$x\ne y$处可导；进一步，这又导致右边在$x\ne y$处二阶可导，从而$f'$在$x\ne y$处二阶可导……归纳可知$f$在$x\ne y$处无穷次可导，再结合$y$的任意性可知$f\in C^\infty(\mathbb{R})$。

在原式两边对x求导，得

$$f'(x)=f''\left(\frac{x+y}{2}\right)\frac{x-y}{2}-f'\left(\frac{x+y}{2}\right).$$

将$\left(\frac{x+y}{2},\frac{x-y}{2}\right)$用$(x,y)$代替，得

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

再对y求导，得$f''(x+y)=f''(x)$，即f''是常数，所以f是二次函数。 □

注. 容易看出本题的条件对二次函数确实是成立的，即抛物线上中点处的切线平行于两点连线。利用这一观察可以得到抛物线的面积公式。设$f(t)=at^2+bt+c$是二次函数，对于$x\neq y$，考虑抛物线上两点$X=(x,f(x))$与$Y=(y,f(y))$，希望求出f的图像与线段XY围出的有界区域D的面积。记l_X,l_Y分别是X,Y处的切线，交点为P。再记$z=\frac{x+y}{2}$和$Z=(z,f(z))$，则Z处的切线l_Z平行于XY，进一步简单计算可知l_Z是$\triangle PXY$的中位线。所以$\triangle XYZ$的面积是$\triangle PXY$的$\frac{1}{2}$。记PX,PY的中点为X_1,Y_1，再对$\triangle XX_1Z$和$\triangle YY_1Z$进行相同的讨论，最终可知D的面积等于$\triangle PXY$的面积的$\frac{1}{2}\left(1+\frac{1}{4}+\frac{1}{4^2}+\cdots\right)=\frac{2}{3}$倍，即$\triangle XYZ$的面积的$\frac{4}{3}$倍。最后，用行列式可以算出$\triangle XYZ$的面积为$\frac{a}{8}|x-y|^3$，故区域$D$的面积是$\frac{a}{6}(y-x)^3$。诚然这一结果可以用定积分简单证明，但这里采用的方法早在古希腊时期就已经由*Apollonius*发现。

下一题的非平凡之处在于，高阶Taylor公式成立并不保证高阶导数存在。从Taylor公式看，如果$f\in C^n(\mathbb{R})$且$f(0)=0$，则当$x\to 0$时

$$\frac{f(x)}{x}=f'(0)+\frac{f''(0)}{2}x+\cdots+\frac{f^{(n)}(0)}{n!}x^{n-1}+o(|x|^{n-1}).$$

一旦知道$\frac{f(x)}{x}$在0处确实$n-1$次可导（这个函数在0处自然应该连续延拓为$f'(0)$），则Taylor公式的唯一性就保证了0处的$n-1$阶导数只能是$\frac{f^{(n)}(0)}{n}$。然而$\frac{f(x)}{x}$是否$n-1$次可导并不事先知道。

题 7.11. *(a)* 设f在$\mathbb{R}$上$n-1$次可导、在0处n次可导且$f(0)=0$。定义$g\colon\mathbb{R}\to\mathbb{R}$为

$$g(x):=\begin{cases}f'(0), & x=0,\\ \frac{f(x)}{x}, & x\neq 0.\end{cases}$$

证明：g在$\mathbb{R}$上$n-1$次可导、$n-1$次导数在0处连续、且$g^{(n-1)}(0)=\frac{f^{(n)}(0)}{n}$。

(b) 设$f\in C^\infty(\mathbb{R})$且$f$是偶函数，证明：$f(\sqrt{x})\in C^\infty([0,\infty))$。

证明*1.* (a) 只需证明g在0处$n-1$次可导，且$g^{(n-1)}$在0处连续。一旦这件事得证，g在$\mathbb{R}\backslash\{0\}$上$n-1$次可导是显然的，并且$g^{(n-1)}(0)$可以由Taylor级数的唯一性求出（见上文）。

为此，对m归纳证明加强的命题：对任意$m\in\mathbb{N}_0$以及$k\in\mathbb{N}$，如果f在0处$m+k$阶可导，且$f(0)=\cdots=f^{(k-1)}(0)=0$，则函数

$$g(x)=\begin{cases}\frac{f^{(k)}(0)}{k!}, & x=0,\\ \frac{f(x)}{x^k}, & x\neq 0,\end{cases}$$

在0附近m阶可导，且$g^{(m)}$在0处连续。特别的，取$(m,k)=(n-1,1)$就得到原题。

下面对m归纳证明（也就是说，对每个m证明结论对任意k成立）。$m=0$时结论显然，$m=1$时根据f在0处的$k+1=k+m$阶Taylor公式可知

$$\begin{aligned}
g'(0)&=\lim_{x\to 0}\frac{g(x)-g(0)}{x}=\lim_{x\to 0}\frac{f(x)-\frac{f^{(k)}(0)}{k!}x^k}{x^{k+1}}\\
&=\frac{\frac{f^{(k+1)}(0)}{(k+1)!}x^{k+1}+o(|x|^{k+1})}{x^{k+1}}=\frac{f^{(k+1)}(0)}{(k+1)!}.
\end{aligned}$$

因为f在0附近至少$m+k-1\geq m=1$阶可导，所以g在0的去心邻域上可导（导数的四则运算），再利用Taylor公式计算可知

$$\begin{aligned}
&\lim_{x\to 0}g'(x)=\lim_{x\to 0}\frac{xf'(x)-kf(x)}{x^{k+1}}\\
=&\frac{x\left(\frac{f^{(k)}(0)}{(k-1)!}x^{k-1}+\frac{f^{(k+1)}(0)}{k!}x^k+o(|x|^k)\right)-k\left(\frac{f^{(k)}(0)}{k!}x^k+\frac{f^{(k+1)}(0)}{(k+1)!}x^{k+1}+o(|x|^{k+1})\right)}{x^{k+1}}\\
=&\frac{f^{(k+1)}(0)}{(k+1)!}=g'(0),
\end{aligned}$$

所以g'在0处连续。这就证明了$m=1$的情形。

假设$m\,(m\geq 1)$的情形已证，考虑$m+1$的情形。用$f(x)-f^{(k)}(0)x^k/k!$代替$f(x)$，可以不妨设除了$f(0)=\cdots=f^{(k-1)}(0)$之外还有$f^{(k)}(0)=0$。特别的，由于$k\in\mathbb{N}$，所以$f'(0)=0$。根据$m=1$时的归纳奠基可知g可导且

$$g'(x)=\begin{cases}\frac{f^{(k+1)}(0)}{(k+1)!}, & x=0,\\ \frac{xf'(x)-kf(x)}{x^{k+1}}, & x\neq 0.\end{cases}$$

将g'分拆为$g'(x)=h_1(x)-kh_2(x)$，其中

$$h_1(x)=\begin{cases}\frac{f^{(k+1)}(0)}{k!}, & x=0,\\ \frac{f'(x)}{x^k}, & x\neq 0,\end{cases}\qquad h_2(x)=\begin{cases}\frac{f^{(k+1)}(0)}{(k+1)!}, & x=0,\\ \frac{f(x)}{x^{k+1}}, & x\neq 0.\end{cases}$$

因为$f'(x)$在0处$m+1+k-1=m+k$阶可导，并且0处的直到$k-1$阶导数都是0（注意这里用到了“不妨设”的内容），根据归纳假设可知h_1在0附近m阶可导，且$h_1^{(m)}$在0处连续。类似的，因为$f(x)$在0附近$m+1+k>m+k$阶可导，并且0处直到$k>k-1$阶导数都是0，根据归纳假设可知h_2也在0处m阶可导，且$h_2^{(m)}$在0处连续。总结起来，g'在0处m阶可导，即g在0处$m+1$阶可导，并且$g^{(m+1)}=(g')^{(m)}$在0处连续。这就完成了归纳。

(b) 只需归纳证明$g(x) := f(\sqrt{x})$的各阶导数在0处连续。事实上，对$x > 0$有

$$g'(x) = \frac{f'(\sqrt{x})}{2\sqrt{x}}.$$

因为f'是奇函数，由(a)可知$h(x) := \frac{f'(x)}{2x}$也是C^∞偶函数。只要结论对n成立，则由$g'(x) = h(\sqrt{x})$可知$g \in C^{n+1}([0,\infty))$，于是结论对$n+1$也成立。而奠基$n = 0$是显然的。 □

证明2. 以(a)为例说明直接计算也是可行的，即本质上只需要证明$\lim_{x\to 0} g^{(n-1)}(x) = \frac{f^{(n)}(x)}{n}$。而(b)涉及到复合函数的高阶导数更加复杂，因此不详细讨论。与例题类似，严格来讲需要由归纳假设保证$g^{(n-2)}$在0处连续，才能由$g^{(n-1)}$极限存在得到0处的可导性。

当$x \neq 0$时，由Leibniz公式可知

$$g^{(n-1)}(x) = \sum_{k=0}^{n-1} C_{n-1}^k f^{(n-k-1)}(x) \cdot (-1)^k \frac{k!}{x^{k+1}}.$$

按照Taylor公式，有

$$f^{(n-k-1)}(x) = \sum_{l=0}^{k+1} \frac{f^{(n-k-1+l)}(0)}{l!} x^l + o(|x|^{k+1}).$$

所以

$$g^{(n-1)}(x) = \sum_{k=0}^{n-1} \sum_{l=0}^{k+1} (-1)^k \frac{(n-1)!}{(n-k-1)!} \frac{f^{(n-k-1+l)}(0)}{l!} \frac{1}{x^{k+1-l}} + o(1).$$

注意到$k+1-l$的范围是$\{0,1,\ldots,n\}$，下面分别考虑每个$\frac{1}{x^t}$前面的系数。

$t = 0$的情况：即对$l = k+1$求和。因为$\sum_{k=0}^n (-1)^k C_n^k = 0$，所以常数项是

$$\sum_{k=0}^{n-1} (-1)^k \frac{(n-1)!}{(n-k-1)!(k+1)!} f^{(n)}(0) = \frac{f^{(n)}(0)}{n} \sum_{k=0}^{n-1} (-1)^k C_n^{k+1} = \frac{f^{(n)}(0)}{n}.$$

$1 \le t \le n-1$的情况：根据相同的原因（注意$n-t \ge 1$），$\frac{1}{x^t}$的系数是

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

$t = n$的情况：此时只有一项$k = n-1$且$l = 0$，对应的系数里包含了$f(0) = 0$，所以是0。[这种情况与第二种的区别在于，$\sum_{k=0}^n (-1)^k C_n^k$在$n = 0$时等于1而非0。]

综上所述，当$x \to 0$时

$$g^{(n-1)}(x) = \frac{f^{(n)}(0)}{n} + o(1).$$

由归纳假设可知$g^{(n-2)} \in C(\mathbb{R})$，这就说明$g^{(n-1)}$在0处连续且等于$\frac{f^{(n)}(0)}{n}$。 □

补充：无穷级数的乘积与二重级数

这一小节讨论无穷级数的乘积和二重级数，将在接下来研究解析函数时用到。

对于两个收敛的无穷级数$\sum a_m$和$\sum b_n$的乘积，如果把第n项想象成次数为n的部分，乘积里面先对n次项合并，则猜想$\sum a_m \sum b_n = \sum_n \left(\sum_{i+j=n} a_i b_j\right)$。这被称为Cauchy乘积公式，但通常是不成立的：比如考虑$a_n = b_n = \frac{(-1)^n}{n^{\frac{1}{4}}}$（显然此时$\sum a_n, \sum b_n$收敛），则

$$\left|\sum_{i=1}^{n} a_i b_{n+1-i}\right| = \sum_{i=1}^{n} \frac{1}{|i(n+1-i)|^{\frac{1}{4}}} \geq \sum_{i=1}^{n} \frac{1}{\left(\frac{n^2}{2}\right)^{\frac{1}{4}}} = 2^{\frac{1}{4}}\sqrt{n} \to \infty,$$

从而不可能求和收敛。但如果级数是绝对收敛的，则结论成立。

题 7.12. *(a)* 设$\sum_{n=0}^{\infty} a_n$和$\sum_{n=0}^{\infty} b_n$都绝对收敛，则$\sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i}\right)$也绝对收敛，且

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty}\left(\sum_{i=0}^{n} a_i b_{n-i}\right).$$

*(b)** 设$\sum_{n=0}^{\infty} a_n$和$\sum_{n=0}^{\infty} b_n$都收敛且至少有一个绝对收敛，则仍然有

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty}\left(\sum_{i=0}^{n} a_i b_{n-i}\right).$$

证明. (a) 绝对收敛是显然的：对任意$N \in \mathbb{N}$，易见

$$\sum_{n=0}^{N}\left|\sum_{i=0}^{n} a_i b_{n-i}\right| \leq \sum_{n=0}^{N}\sum_{i=0}^{n} |a_i||b_{n-i}| \leq \left(\sum_{n=0}^{N} |a_n|\right)\left(\sum_{n=1}^{N} |b_n|\right) \leq \left(\sum_{n=0}^{\infty} |a_n|\right)\left(\sum_{n=1}^{\infty} |b_n|\right).$$

余下的结论可以由(b)得出，但直接证明更容易：通过将a_n, b_n都分拆为正部、负部（参见之前关于无穷级数求和次序的注），只需处理$a_n, b_n \geq 0$的情况。此时观察求和项可知

$$\left(\sum_{n=0}^{\lfloor N/2 \rfloor} a_n\right)\left(\sum_{n=0}^{\lfloor N/2 \rfloor} b_n\right) \leq \sum_{n=0}^{N}\sum_{i=0}^{n} a_i b_{n-i} \leq \left(\sum_{n=0}^{N} a_n\right)\left(\sum_{n=0}^{N} b_n\right).$$

两边取极限结合夹逼即可。

(b) 不妨设$\sum |a_n| < \infty$。为了说话方便，记$A_n := a_0 + \cdots + a_n$，$B_n := b_0 + \cdots + b_n$，$c_n := a_0 b_1 + \cdots + a_n b_0$以及$S_n := c_0 + \cdots + c_n$，则

$$S_0 + \cdots + S_n = \sum_{k=0}^{n} a_k \left(\sum_{l=0}^{n-k} b_l\right) = \sum_{k=0}^{n} a_k B_{n-k}.$$

再记$A := \sum_{n=0}^{\infty} a_n$，$B := \sum_{n=0}^{\infty} b_n$，则

$$\left|AB - \sum_{k=0}^{n} S_k\right| = \left|\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} a_k(B - B_{n-k}) + \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} a_k(B - B_{n-k}) + \sum_{k=n+1}^{\infty} a_k B\right|$$

$$\le \sup_{0\le k\le \frac{n}{2}} |B - B_{n-k}| \cdot \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} |a_n| + \sup_{\frac{n}{2}<k\le n} |B - B_{n-k}| \cdot \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} |a_k| + |B| \sum_{k=n+1}^{\infty} |a_k|$$

$$\le \sup_{k>\frac{n}{2}} \left|\sum_{m=k}^{\infty} b_m\right| \cdot \sum_{k=0}^{\infty} |a_k| + \left(\sup_{k\in\mathbb{N}_0} \left|\sum_{n=k}^{\infty} b_k\right| + |B|\right) \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{\infty} |a_n|.$$

容易证明右边在$n \to \infty$时的极限是0，具体细节留作练习。 □

注. *(b)中等号右边的收敛通常不是绝对的，比如取*$(a_n) = (1, -1, 0, 0, \dots)$*和*$b_n = \frac{(-1)^n}{n}$。

下面考虑二重级数$a_{mn}\,(m, n \in \mathbb{N})$进行累次求和是否与次序有关，即$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}$是否等于$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn}$？一般来讲这是不对的，比如取$a_{n,n} = 1$，$a_{n,n+1} = -1$，而其他位置都是0。若干例子表明绝对收敛才是研究无穷级数的正确语境。在学习实变函数之后，（绝对收敛的）求和可以看成关于计数测度的积分，于是下面将证明的二重求和与累次求和的关系就是Fubini–Tonelli定理的推论。对于正项级数，总是认为求和允许广义收敛到∞，并约定∞与$a \in [0, \infty]$相加等于∞。于是容易看出正项二重级数的累次求和总是广义收敛的。

题 7.13. *(a) 如果*$a_{mn} \ge 0$，*则*$\sum_m \sum_n a_{mn} = \sum_n \sum_m a_{mn}$*（允许都等于*$\infty$*）。*

(b) 如果$\sum_m \sum_n |a_{mn}| < \infty$，*则*$\sum_m \sum_n a_{mn} = \sum_m \sum_n a_{mn}$。

证明. (a) 只需证明两种求和方式都等于

$$\sup\{\sum_{(i,j)\in I} a_{ij} \mid I \subset \mathbb{N}^2\text{是有限集}\}.$$

具体细节留作练习。[这个技巧曾被用于证明正项级数求和不依赖次序。]

(b) 简单的方法是将a_n拆成正部和负部之差并借助(a)（参见之前关于无穷级数求和次序的注），下面给出虽然略微复杂但更直接的方法。

由绝对收敛可知，对任意$\varepsilon > 0$，存在N使得

$$\sum_{(m,n)\in\mathbb{N}^2,\ m\text{或}n>N} |a_{mn}| < \varepsilon.$$

记$A_k := \sum_{m=1}^{k}\sum_{n=1}^{k} a_{mn}$，则由上式可知$A_k$构成Cauchy列，从而极限$A := \lim_{k\to\infty} A_k$存在。下面证明$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} = A$，同理可知$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn} = A$，从而二者相等。

由二重级数绝对收敛可知，对固定的m，$\sum_{n=1}^{\infty} a_{mn}$绝对收敛。当$k > N$时，注意到

$$\left|\sum_{m=1}^{k}\sum_{n=1}^{\infty} a_{mn} - A_k\right| = \left|\sum_{m=1}^{k}\sum_{n=k+1}^{\infty} a_{mn}\right| \leq \sum_{(m,n)\in\mathbb{N}^2,\ m\text{或}n>N} |a_{mn}| < \varepsilon.$$

由此可知$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} = A$。 □

题 7.14 * (怀新一题20230316)**.** 设$a_1, \dots, a_n \in (-1, 1)$，证明：

$$\prod_{1\leq i,j\leq n} \frac{1+a_ia_j}{1-a_ia_j} \geq 1.$$

证明. 取对数，只需证明

$$\sum_{1\leq i,j\leq n} (\log(1+a_ia_j) - \log(1-a_ia_j)) \geq 0.$$

当$|x| < 1$时$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k$，由此可知

$$\begin{aligned}\sum_{1\leq i,j\leq n} (\log(1+a_ia_j) - \log(1-a_ia_j)) &= \sum_{1\leq i,j\leq n}\sum_{k=1}^{\infty} \frac{(-1)^{k-1}+1}{k}(a_ia_j)^k \\ &= \sum_{l=1}^{\infty} \frac{2}{2l-1} \sum_{1\leq i,j\leq n} (a_ia_j)^{2l-1} \\ &= \sum_{l=1}^{\infty} \frac{2}{2l-1} \cdot (a_1^{2l-1} + \cdots + a_n^{2l-1})^2 \geq 0.\end{aligned}$$

故原不等式成立。（注意验证这里交换求和次序合法来自于$|a_i| < 1$以及绝对收敛。） □

补充：解析函数

截断函数的例子说明光滑函数并非总能在一个点的邻域上展成收敛的Taylor级数。能够展成Taylor级数的函数具有比光滑更强的性质，即这一小节将会讨论的解析函数(analytic function)。这里给出的证明都是使用数学分析的方法（或者说实方法），但讨论解析函数的正确语境是复平面。今后学习复变函数时会看到这些基本性质可以有更简洁的证明。

定义 7.15 (幂级数)**.** *以x_0为中心的幂级数是指$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$，其中$a_n \in \mathbb{R}$（或$\mathbb{C}$）。定义收敛半径$r := \frac{1}{\limsup \sqrt[n]{|a_n|}}$，其中$\frac{1}{0}$和$\frac{1}{\infty}$分别约定为$\infty$和0。*

在相差平移的意义下，只考虑$x_0 = 0$的情况不会带来损失。下面说明收敛半径恰如其名。

题 7.16. 设$f(x)$是以0为中心的幂级数，$r>0$是收敛半径，证明：

(a) 当$|x|<r$时，$f(x)$绝对收敛；当$|x|>r$时，$f(x)$发散。

(b) 对任意$0<\delta<r$，$f(x)$在$[-\delta,\delta]$上一致收敛，从而$f(x)$在$(-r,r)$上是连续函数。

(c) $f(x)$在$(-r,r)$上是C^∞函数。事实上，f可以逐项求导，即

$$f'(x)=\sum_{n=1}^{\infty} n a_n x^{n-1}.$$

由此可知，$f'(x)$也是收敛半径为r的幂级数，并且$a_n=\frac{f^{(n)}(0)}{n!}$。

证明. (a) 这是Cauchy根式判别法的推论。

(b) 任意取定$\eta\in(\delta,r)$，则存在$N>0$使得当$n>N$时$|a_n|\le\eta^{-n}$，于是对$|x|\le\delta$有$|a_nx^n|\le\left(\frac{\delta}{\eta}\right)^n$。因为$\delta<\eta$，所以右边构成一个收敛的正项级数，用大M判别法即证。

(c) 只需证明一阶导的表达式，然后归纳即可。注意到结论中f'的表达式也是收敛半径为r的幂级数，记为$g(x)$。取定$x\in(-r,r)$，当$|\delta|$充分小时$x+\delta\in(-r,r)$。根据Lagrange中值定理，存在$\xi\in(x,x+\delta)$（或$(x+\delta,x)$）使得

$$|(x+\delta)^n-x^n|=n\delta|\xi^{n-1}|\le n\delta(|x|+|\delta|)^{n-1}.$$

任意取定$N\in\mathbb{N}$，有

$$\begin{aligned}
&\left|\frac{f(x+\delta)-f(x)}{\delta}-g(x)\right|\\
\le&\left|\sum_{n=0}^{N}a_n\left(\frac{(x+\delta)^n-x^n}{\delta}-nx^{n-1}\right)\right|+\sum_{n=N+1}^{\infty}n|a_n|(((|x|+|\delta|)^{n-1}+|x|^{n-1})\\
\le&\sum_{n=0}^{N}|a_n|\left|\frac{(x+\delta)^n-x^n}{\delta}-nx^{n-1}\right|+2\sum_{n=N+1}^{\infty}n|a_n|(((|x|+|\delta|)^{n-1}.
\end{aligned}$$

取定$\eta\in(|x|,r)$，则当δ充分小时$|x|+|\delta|<\eta$。由(a)可知$\sum_{n=1}^{\infty}n|a_n|\eta^{n-1}$收敛，因此对任意$\varepsilon>0$，可以事先取定充分大的$N$使得$\sum_{n=N+1}^{\infty}n|a_n|\eta^{n-1}<\varepsilon$。当$\delta\to0$时，前一项求和都是趋于0的，因此

$$\limsup_{\delta\to0}\left|\frac{f(x+\delta)-f(x)}{\delta}-g(x)\right|\le\varepsilon.$$

由ε任意性可知$f'(x)=g(x)$对任意$x\in(-r,r)$成立。 □

注. 易见*(a)*的证明对任何满足模长小于r的复数都成立，*(b)(c)*也可以相应改进到复平面上（但涉及多元微分学）。这就解释了为何复平面才是研究幂级数和解析函数的正确语境。

定义 7.17 (解析函数)**.** 称函数f在x_0处解析，如果存在$r > 0$使得f在$(x_0-\delta, x_0+\delta)$上等于一个以x_0为中心的收敛幂级数。如果f在定义域上每一点都解析，则称f是解析函数。

下面说明解析函数关于加减乘除封闭，并且f总是在一个开集上解析。这些证明写得很繁琐，但实际上几乎是形式的：只需要验证二重求和换序合法。

题 7.18. *(a)* 证明：若f, g都在x_0处解析，则$f \pm g$和fg也都在x_0处解析。

(b) 证明：若以0为中心的幂级数f在$(-\delta, \delta)$上收敛，则f在$(-\delta, \delta)$上解析。进一步，对任意$x_0 \in (-\delta, \delta)$，$f$在$x_0$处的$Taylor$级数收敛半径至少是$\delta - |x_0|$。

(c) 证明：若f在x_0处解析，g在$f(x_0)$处解析，则$g \circ f$在x_0处解析。

(d) 证明：若f在x_0处解析且$f(x_0) \neq 0$，则$\frac{1}{f(x)}$在x_0处解析。

(e) 设f在0处解析且$f(0) = 0$，则$\frac{f(x)}{x}$也在0处解析。

证明. (a) 加减是显然的，下面证明fg在x_0处解析。不妨设$x_0 = 0$，且在$(-\delta, \delta)$上有

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad 和 \quad g(x) = \sum_{n=0}^{\infty} b_n x^n.$$

则对$x \in (-\delta, \delta)$，二者都是绝对收敛的无穷级数。因此

$$f(x)g(x) = \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} a_i b_{n-i} \right) x^n,$$

也是绝对收敛的幂级数。

(b) 假设在$(-\delta, \delta)$上$f(x) = \sum_{n=0}^{\infty} a_n x^n$。对于$x_0 \in (-\delta, \delta)$，当$|x| < \delta - |x_0|$时

$$f(x + x_0) = \sum_{n=0}^{\infty} a_n (x + x_0)^n = \sum_{n=0}^{\infty} a_n \sum_{k=0}^{n} C_n^k x_0^{n-k} x^k.$$

为了交换求和次序，需要验证绝对收敛，而这是因为$|x| + |x_0| < \delta$保证了

$$\sum_{n=0}^{\infty} |a_n| \sum_{k=0}^{n} |C_n^k x_0^{n-k} x^k| \leq \sum_{n=0}^{\infty} |a_n| \sum_{k=0}^{n} C_n^k |x_0|^{n-k} |x|^k = \sum_{n=0}^{\infty} |a_n| (|x| + |x_0|)^n < \infty.$$

因此

$$f(x + x_0) = \sum_{k=0}^{\infty} \left(\sum_{n=k}^{\infty} a_n C_n^k x_0^{n-k} \right) x^k.$$

(c) 假设在$(-\delta,\delta)$上$f(x+x_0)=f(x_0)+\sum_{n=1}^{\infty}a_nx^n$，并且在$(-2\eta,2\eta)$上$g(y+f(x_0))=\sum_{n=0}^{\infty}b_ny^n$。取$\delta$充分小，不妨设$\sum_{n=1}^{\infty}|a_n|\delta^n<\eta$。于是当$x\in(-\delta,\delta)$时

$$g(f(x+x_0))=\sum_{n=0}^{\infty}b_n\left(\sum_{k=1}^{\infty}a_kx^k\right)^n.$$

为了证明这是收敛的幂级数，由(b)可知$\left(\sum_{n=1}^{\infty}a_kx^k\right)^n=\sum_{k=0}^{\infty}c_{nk}x^k$且$\sum_{k=0}^{\infty}|c_{nk}||x|^k\leq\left(\sum_{n=1}^{\infty}|a_k|\delta^k\right)^n<\eta^n$。于是$\sum_n|b_n|\sum_k|c_{nk}||x|^k\leq\sum_n|b_n|\eta^n<\infty$，再交换次序即可。

(d) 熟知$\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n$对$|x|<1$成立。对任意$x_0\neq0$，当$|x|<x_0$时

$$\frac{1}{x_0-x}=\frac{1}{x_0}\cdot\frac{1}{1-\frac{x}{x_0}}=\frac{1}{x_0}\sum_{n=0}^{\infty}\frac{x^n}{x_0^n}.$$

因此$\frac{1}{x}$是$\mathbb{R}\setminus\{0\}$上的解析函数，再结合(c)即证。

(e) 这几乎是显然的。[而如果f仅仅是光滑函数，先前证明$\frac{f(x)}{x}$仍然光滑相当麻烦。] □

需要指出解析函数f的Taylor级数未必在f的整个定义域上都收敛，例如$1/x$在$x>0$处的Taylor级数收敛半径就是$|x|$（因为在0处有奇点）。即使f在$\mathbb{R}$上都解析，也可能会因为在复平面上有奇点导致障碍，例如$1/(x^2+1)$在x处的Taylor级数收敛半径是$\sqrt{1+|x|^2}$（因为$\pm i$是奇点）。因此，幂级数展开只是在局部上给出了f的表达式，但下一题说明解析函数具有很强的刚性：一点处的信息就能决定整个函数。

题 7.19. *设f是开区间(a,b)上的解析函数。*

(a) 如果存在$x_0\in(a,b)$使得$f^{(n)}(x_0)=0\,(\forall n\in\mathbb{N}_0)$，证明：$f$恒等于0。

(b) 如果$f^{-1}(0)$在(a,b)中有聚点，证明：f恒等于0。

证明. (a) 考虑$X:=\{x\in[x_0,b)\mid f^{(n)}(y)=0,\ \forall n\in\mathbb{N}_0,\ \forall y\in[x_0,x]\}$，则$x_0\in X$。下面证明$\sup X=b$，用反证法，如果$c:=\sup X<b$，则根据连续性可知$c\in X$。再结合$f$在$c$处解析性，存在$\delta>0$使得在$(c-\delta,c+\delta)$上$f$恒等于0，与$c=\sup X$矛盾！这就说明$f|_{[x_0,b)}=0$，同理可证$f|_{(a,x_0]}=0$。

(b) 设$x_0\in(a,b)$是$f^{-1}(0)$的聚点，则利用Lagrange中值定理容易证明$f^{(n)}(x_0)=0\,(\forall n\in\mathbb{N}_0)$，再结合(a)即证。另一种证明是如果$f$在$x_0$的各阶导数不全为0，则可设

$$f(x+x_0)=x^n\left(a_n+\sum_{k=1}^{\infty}a_{n+k}x^k\right),$$

其中$a_n\neq0$。容易证明当$|x|$充分小时$f(x+x_0)$不会有0之外的零点。 □

注. 拓扑空间X称为连通的，如果X中既开又闭的集合只有空集和全集。称X的子集Y是连通的，如果Y连同子拓扑构成的拓扑空间是连通的。容易证明，$\mathbb{R}$中的连通集就是所有（有限或无限、端点开或闭的）区间。上面的证明本质上用到的是区间的连通性：根据解析性可知满足各阶导数都为0的点构成开集，而连续性又保证了这样的点构成闭集，因此一旦非空就必然是全集。在处理高维情形时，连通性有更显著的作用。

下面将定义Bernoulli数，并利用生成函数(generating function)方法计算$1^k + \cdots + n^k$的表达式。在此之前先给出所谓指数型生成函数的乘积公式；先前关于两个解析函数乘积的计算无非是普通生成函数的乘积公式，而数论中更为重要的是所谓Dirichlet生成函数。

题 7.20. 假设下式左边的两个幂级数都具有正的收敛半径，证明：当$|x|$充分小时，

$$\left(\sum_{n=0}^{\infty} \frac{a_n}{n!} x^n\right) \left(\sum_{n=0}^{\infty} \frac{b_n}{n!} x^n\right) = \sum_{n=0}^{\infty} \frac{c_n}{n!} x^n,$$

其中$c_n = \sum_{k=0}^{n} C_n^k a_k b_{n-k}$。

证明. 按照Cauchy乘积展开：

$$c_n = n! \sum_{k=0}^{n} \frac{a_k}{k!} \cdot \frac{b_{n-k}}{(n-k)!} = \sum_{k=0}^{n} C_n^k a_k b_{n-k}.$$ □

题 7.21. 定义*Bernoulli*数$B_n\ (n \in \mathbb{N}_0)$为（易见左边是0处的解析函数）

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \frac{B_n}{n!} x^n.$$

(a) 求B_0, B_1, B_2, B_3, B_4的值。

(b) 证明：对于$n \geq 1$，有$B_{2n+1} = 0$。

(c) 证明：对任意$n \geq 2$有

$$C_n^0 B_0 + C_n^1 B_1 + \cdots + C_n^{n-1} B_{n-1} = 0.$$

证明. (a) 按定义可以求出$B_0 = 1$，$B_1 = -\frac{1}{2}$，$B_2 = \frac{1}{6}$，$B_3 = 0$，$B_4 = -\frac{1}{30}$。

(b) 注意到

$$\frac{x}{e^x - 1} + \frac{x}{2} = \frac{x}{2} \cdot \frac{e^x + 1}{e^x - 1},$$

是偶函数，所以当$n \geq 1$时Taylor展开式中x^{2n+1}项系数为0，即$B_{2n+1} = 0$。

(c) 比较$x = (e^x - 1) \sum_{k=0}^{\infty} \frac{B_n}{n!} x^n = \left(\sum_{k=1}^{\infty} \frac{x^k}{k!}\right) \left(\sum_{k=0}^{\infty} \frac{B_n}{n!} x^n\right)$两边$x^n$项系数即可。 □

题 7.22. 设$k \in \mathbb{N}_0$，本题的目标是给出$1^k + \cdots + n^k\ (n \in \mathbb{N})$的表达式。

(a) 利用指数函数的*Taylor*级数，证明：

$$e^x + \cdots + e^{nx} = \sum_{k=0}^{n} (1^k + \cdots + n^k) \frac{x^k}{k!}.$$

(b) 利用等比数列求和（或者说n次方差公式），证明：

$$e^x + \cdots + e^{nx} = \left(\sum_{k=0}^{\infty} \frac{B_k}{k!} x^k \right) \left(\sum_{k=0}^{\infty} \frac{(n+1)^{k+1} - 1}{k+1} \frac{x^k}{k!} \right).$$

(c) 比较系数，证明：

$$1^k + \cdots + n^k = \frac{1}{k+1} \sum_{i=0}^{k} C_{k+1}^{i+1} B_{k-i} ((n+1)^{i+1} - 1).$$

证明. (a)是显然的，而(b)是因为

$$\begin{aligned} e^x + \cdots + e^{nx} &= \frac{e^{(n+1)x} - e^x}{e^x - 1} = \frac{x}{e^x - 1} \cdot \frac{e^{(n+1)x} - e^x}{x} \\ &= \left(\sum_{k=0}^{\infty} \frac{B_k}{k!} x^k \right) \left(\sum_{k=0}^{\infty} \frac{(n+1)^{k+1} - 1}{k+1} \frac{x^k}{k!} \right). \end{aligned}$$

最后，比较(a)(b)右边x^k项系数，结合$\frac{1}{i+1} C_k^i = \frac{1}{k+1} C_{k+1}^{i+1}$就得到(c)：

$$1^k + \cdots + n^k = \sum_{i=0}^{k} C_k^i B_{k-i} \frac{(n+1)^{i+1} - 1}{i+1} = \frac{1}{k+1} \sum_{i=0}^{k} C_{k+1}^{i+1} B_{k-i} ((n+1)^{i+1} - 1). \qquad \square$$

注. 结合上一题的*(c)*可知，当$k \geq 1$时，每一个括号中额外的-1实际上没有贡献：

$$\sum_{i=0}^{k} C_{k+1}^{i+1} B_{k-i} = 0.$$

这里保留-1是为了让结论对$k = 0$依然成立。

题 7.23 *. 回忆对于$n \in \mathbb{N}_0$，第n个*Hermite*多项式定义为

$$H_n(x) := (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}).$$

(a) 设$x, t \in \mathbb{R}$，证明：

$$e^{2tx - t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} t^n.$$

(b) 记D是求导算子，即$Df := f'$，证明：

$$H_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{2^{2k} k!} D^{2k} (2x)^n.$$

(c) 设$a, b \in \mathbb{R}$且$a^2 + b^2 = 1$，证明：对$n \in \mathbb{N}$有

$$H_n(ax + by) = \sum_{k=0}^{n} C_n^k a^k b^{n-k} H_k(x) H_{n-k}(y).$$

证明. (a) 将x视为参数，则$f(t) := e^{2tx - t^2}$是关于t的解析函数。换元$s = x - t$可知

$$f^{(n)}(0) = e^{x^2} \frac{d^n}{dt^n} (e^{-(t-x)^2})|_{t=0} = (-1)^n e^{x^2} \frac{d^n}{ds^n} (e^{-s^2})|_{s=x} = H_n(x).$$

因此右边就是$f(t)$在0处的Taylor级数，从而在收敛半径内都收敛到$f(t)$自身。为了说明收敛半径是∞，可以直接估计$|H_n(x)|$的大小，但更简单的方式是借助复变函数的理论：$f(t)$在整个$\mathbb{C}$上解析，所以Taylor级数的收敛半径是∞。

(b) 由(a)可知

$$\sum_{n=0}^{\infty} \frac{H_n(x)}{n!} t^n = e^{2tx - t^2} = \left(\sum_{n=0}^{\infty} \frac{(2x)^n}{n!} t^n \right) \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} t^{2n} \right).$$

比较两边t^n项系数得

$$H_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{k!} \frac{n!}{(n-2k)!} (2x)^{n-2k} = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{2^{2k} k!} D^{2k} (2x)^n.$$

(c) 技巧在于将t^2表示为$a^2 t^2 + b^2 t^2$：

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{H_n(ax+by)}{n!} t^n &= e^{2t(ax+by) - t^2} = e^{2t(ax+by) + (a^2+b^2)t^2} = e^{2(at)x - (at)^2} \cdot e^{2(bt)t - (bt)^2} \\ &= \left(\sum_{n=0}^{\infty} \frac{H_n(x)}{n!} (at)^n \right) \left(\sum_{n=1}^{\infty} \frac{H_n(y)}{n!} (bt)^n \right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} C_n^k a^k H_k(x) b^{n-k} H_{n-k}(y) \right) \frac{t^n}{n!} \end{aligned}$$

比较两边t^n项系数即可。 □

注. 由于当$k > n/2$时$(-D)^k (x^n) = 0$，常常将*(b)*的表达式写为

$$H_n(x) = \sum_{k=0}^{\infty} \frac{1}{k!} (-D^2/4)^k (2x)^n =: e^{-D^2/4} (2x)^n.$$

最后讨论解析性由导数的增长行为给出的等价刻画。下面的(a)中用K^{n+1}而不是K^n，仅仅是为了兼顾$n=0$的情形，并没有本质影响。

题 7.24 *. *(a)* 设$f\in C^\infty(\mathbb{R})$，证明：$f$在$x_0$处解析当且仅当存在$\delta>0$和$K\geq 1$，使得

$$|f^{(n)}(x)|\leq K^{n+1}n!,\quad \forall x\in(x_0-\delta,x_0+\delta),\ \forall n\in\mathbb{N}_0.$$

(b) 设$f\in C^\infty(\mathbb{R})$且$f^{(n)}(x)\geq 0$对任意$n\in\mathbb{N}$和$x\in\mathbb{R}$成立，证明：$f$是解析函数，且每一点处*Taylor*级数的收敛半径都是∞。

(c) 设$f\in C^\infty(\mathbb{R})$，证明：$f$在$x_0$处解析当且仅当存在$\delta>0$，使得$f$在任意$x\in(x_0-\delta,x_0+\delta)$处的*Taylor*级数收敛半径有（不依赖$x$的）正下界。

证明. (a) 充分性：当$0<|x-x_0|<\delta$时，存在ξ满足$|\xi-x_0|<|x-x_0|$且

$$\left|f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k\right|=\left|\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}\right|\leq K^{n+2}|x-x_0|^{n+1}.$$

于是只要$|x-x_0|<\min\{\delta,\frac{1}{K}\}$，就知道$x_0$处的Taylor级数在$x$处收敛到$f(x)$。

必要性：不妨设$x_0=0$，相应的Taylor级数为$\sum_{k=0}^\infty a_k|x|^k$。因为收敛半径$r>0$，所以存在常数$C>0$使得$a_k\leq C(2/r)^k\ (\forall k\in\mathbb{N}_0)$。当$|x|<r$时，

$$f^{(n)}(x)=\sum_{k=0}^{\infty}(k+1)\cdots(k+n)a_{k+n}x^k.$$

所以

$$\left|\frac{f^{(n)}(x)}{n!}\right|\leq C(2/r)^n\sum_{k=0}^{\infty}C_{n+k}^{n}(2|x|/r)^k.$$

在$\frac{1}{1-t}=\sum_{k=0}^\infty t^k$两边求$n$阶导，可知当$|t|<1$时

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

因此只要$|x|<r/4$，就有

$$\left|\frac{f^{(n)}(x)}{n!}\right|\leq C(2/r)^n\sum_{k=0}^{\infty}C_{n+k}^{n}\frac{1}{2^k}=C(2/r)^n\cdot 2^{n+1}=2C(4/r)^n.$$

取$K=\max\{4/r,2C\}$，则有$f^{(n)}(x)\leq K^{n+1}n!$对任意$|x|<r/4$成立，结论得证。

(b) 任给$x\in\mathbb{R}$，对任意$y\in[x,x+1]$和任意$n\in\mathbb{N}_0$，存在$\xi\in(y,x+2)$使得

$$f(x+2)=\sum_{k=0}^{n}\frac{f^{(k)}(y)}{k!}(x+2-y)^k+\frac{f^{(n+1)}(\xi)}{(n+1)!}(x+2-\xi)^{k+1}\geq\sum_{k=0}^{n}\frac{f^{(k)}(y)}{k!}.$$

这就说明$f^{(k)}(y) \le f(x+2)\cdot k!$。根据(a)可知$f$在$(x,x+1)$上是解析函数，再结合$x$的任意性可知$f$在$\mathbb{R}$上解析。最后，从上式也可看出对任意$z>y$和任意$n\in\mathbb{N}$都有$f(z)\ge \sum_{k=0}^{n}\frac{f^{(k)}(y)}{k!}(z-y)^k$，于是$y$处的Taylor级数在$z$处收敛，进而收敛半径是$\infty$。

(c) 必要性由先前证明过的幂级数性质是显然的：如果$f(x)$在$(x-2r,x+2r)$上可以展成收敛的幂级数，则在$(x-r,x+r)$上的Taylor级数收敛半径都至少是r。

下证充分性。记x处的Taylor级数是

$$T_x(t)=\sum_{n=0}^{\infty}\frac{f^{(n)}(x)}{n!}t^n.$$

适当缩小δ，可设对任意$x\in I:=[x_0-\delta,x_0+\delta]$，都有$T_x$的收敛半径至少是$r$，其中$r>8\delta>0$。于是$T_x(\cdot-x)$在$(x_0-6\delta,x_0+6\delta)$上收敛，进而在$I$上解析。[注意如果已知$f$也在$I$上解析，则$f(\cdot)=T_x(\cdot-x)$；特别的，$T_x(\cdot-x)=T_y(\cdot-y)$，这是后续的关键。]

先前证明过f的解析点集U是开集，于是$X:=I\setminus U$是闭集。假设X非空，记

$$F_n:=\{x\in X \mid f^{(m)}(x)\le (2/r)^m m!,\ \forall m\ge n\}.$$

易见F_n是闭集，并且T_x的收敛半径至少是r保证了$X=\bigcup_{n\in\mathbb{N}}F_n$。根据Baire纲定理，存在$N\in\mathbb{N}$以及开区间$(a,b)\subset I$使得

$$(a,b)\cap F_N=(a,b)\cap X\ne\emptyset.$$

断言：对任意$x\in(a,b)$，都有

$$f^{(n)}(x)\le 2\cdot(4/r)^n n!,\quad \forall n\ge N.$$

如果断言成立，由(a)中充分性的证明可知f在(a,b)上解析，与$(a,b)\cap X\ne\emptyset$矛盾！故$X=\emptyset$；特别的，这就证明了f在x_0处解析。

最后，要证明断言，结合F_N的定义，只需考虑$x\in(a,b)\cap U$的情况。此时f在x处解析，设f在x处的"最大解析区间"为(c,d)（即U中包含x的最大开区间）。因为$(a,b)\cap X\ne\emptyset$，所以(a,b)不被(c,d)包含，不妨设$c\in(a,b)$（$d\in(a,b)$的情况类似）。关键的观察是：解析性保证了对任意$t\in(c,x)$都有$T_x(t-x)=f(t)$，从而$T_x^{(n)}(t-x)=f^{(n)}(t)$；又$f\in C^\infty$保证了各阶导数连续，从而

$$T_x^{(n)}(c-x)=f^{(n)}(c)=T_x^{(n)}(0),$$

这就说明$T_x^{(n)}(\cdot-x)$与$T_c^{(n)}(\cdot-c)$在c处的各阶导数都相同，且都在I上解析；根据先前讨论的解析函数刚性可知二者在I上相等，于是$f^{(n)}(x)$可以被c处的各阶导数控制。具体来

讲，当$n \geq N$时，利用$|x-c| \leq 2\delta < r/4$以及$c \in F_N$，仿照(a)的证明可知

$$\begin{aligned}\left|\frac{f^{(n)}(x)}{n!}\right| &= \left|\sum_{k=0}^{\infty} C_{n+k}^{n} \frac{f^{(n+k)}(c)}{(n+k)!}(x-c)^k\right| \\ &\leq \sum_{k=0}^{\infty} C_{n+k}^{n} \left(\frac{2}{r}\right)^{n+k} \left(\frac{r}{4}\right)^k = (2/r)^n 2^{n+1} = 2 \cdot (4/r)^n.\end{aligned}$$

这就证明了断言，从而原题得证。 □

注. 对于一列$M_n > 0$，定义$C^M(-1,1)$是$C^\infty(-1,1)$中满足

$$|f^{(n)}(x)| \leq K^{n+1} M_n, \quad \forall n \in \mathbb{N},\ x \in (-1,1),$$

的函数组成的集合，其中K是允许依赖于f的常数。如果$M_n = n!$，则*(a)*表明$C^M(-1,1)$就是解析函数类。可以问对于怎样的M_n，如果$f \in C^M((-1,1))$满足$f^{(n)}(0) = 0\,(\forall n \in \mathbb{N})$，则$f$在$(-1,1)$上恒等于0。这样的函数空间$C^M(-1,1)$称为拟解析函数类*(quasi-analytic class)*，其等价刻画由*Denjoy–Carleman*定理给出：记

$$a_n := \inf_{k \geq n} M_k^{1/k},$$

则$C^M(-1,1)$是拟解析函数类当且仅当$\sum_n \frac{1}{a_n} = \infty$。[11]

注. 按定义，*(b)*仅仅说明*Taylor*级数收敛是不够的，因为其未必收敛到函数自身。这里利用*(a)*的充分性说明解析，即使没有预设*(a)*的结论，这在得到$f^{(k)}(y) \leq Ck!$的估计之后也是相对容易的。也可以借助*(c)*的结论（但这是不平凡的）：既然*(b)*中的*Taylor*级数收敛半径都是∞，那么函数本身必然解析，从而*Taylor*级数收敛到函数自身。

[11] Denjoy–Carleman定理的复分析证明可见W. Rudin, *Real and complex analysis*, third edition, McGraw-Hill, New York, 1987，基于卷积的实方法可见L. V. Hörmander, *The analysis of linear partial differential operators. I*, second edition, Springer, Berlin, 1990。

第8讲：微分中值定理（下）

> 说话间，医生像变戏法一样从白大褂里掏出一本书。你接过来看了看，书名是《数学物理中的微分形式》，你猜想这不是一本文学书。“如果你想让你的目标尽快苏醒，那就给他读这本书，好吗？”医生说完就拉开门走了。室内一片寂静，外面的暴风雨反而加强了这份寂静。你拉过一把椅子，坐在摩德万的病床旁，尽量沉下心来，开始读这本《数学物理中的微分形式》，这里面的内容你一点儿也不懂，有些符号你不会念，只好跳过去，你甚至感到困窘。但你渐渐习惯了，不停歇地小声朗读着，你想这一定是些咒语，医学是如此神秘。
>
> ——朱岳《原路追踪》

这一讲讨论若干插值问题，构造辅助函数的常用工具是Lagrange插值公式；另一个主题是凸函数，但凸性真正发挥威力需待今后学习泛函分析。补充内容给出Bassel问题的两种证明：一是结合三角函数的Markov多项式与Vieta定理，二是利用$\cot(x)$的一个非平凡展开式。

例题

例 8.1. *(a)* 证明：e是无理数。

(b) 求$\lim_{n\to\infty} n\sin(2\pi n!e)$。

证明. (a) 对任意$n\in\mathbb{N}$，根据带Lagrange余项的Taylor公式，存在$\theta\in(0,1)$使得

$$e=\sum_{k=0}^{n}\frac{1}{k!}+\frac{e^\theta}{(n+1)!}.$$

在上式两边同乘$n!$，右边的求和中每一项都是整数，但余项$\frac{e^\theta}{n+1}\in\left(0,\frac{e}{n+1}\right)$在$n>1$时不是整数，所以$n!e\notin\mathbb{Z}$。由$n$的任意性可知$e\notin\mathbb{Q}$。

(b) 在(a)中用直到$n+1$阶导的Taylor公式代替n阶Taylor公式，可知

$$\{n!e\}=\frac{1}{n+1}+O\left(\frac{1}{n^2}\right).$$

所以

$$\lim_{n\to\infty} n\sin(2\pi n!e)=\lim_{n\to\infty} n\sin(2\pi\{n!e\})=\lim_{n\to\infty} n\left(\frac{2\pi}{n+1}+o\left(\frac{1}{n}\right)\right)=2\pi.\qquad\square$$

例 8.2. 设$f\colon\mathbb{R}\to\mathbb{R}$在$x_0$处$n+1$次可导，且$f^{(n+1)}(x_0)\neq 0$。根据带*Lagrange*余项的*Taylor*公

式，当$x \neq x_0$充分接近x_0时，存在$\theta = \theta(x) \in (0,1)$（未必唯一）使得

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1} + \frac{f^{(n)}(x_0+\theta(x-x_0))}{n!}(x-x_0)^n.$$

证明：$\lim_{x\to x_0}\theta(x) = \frac{1}{n+1}$。

证明. 当$x \to x_0$时$f^{(n)}(x_0+\theta(x-x_0)) = f^{(n)}(x_0) + \theta f^{(n+1)}(x-x_0) + o(|x-x_0|)$，于是

$$f(x) = \sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{n!} + \frac{\theta f^{(n+1)}(x_0)}{n!}(x-x_0)^{n+1} + o(|x-x_0|^{n+1}).$$

与Taylor公式对比可知$\left(\theta - \frac{1}{n+1}\right) f^{(n+1)}(x_0) = o(1)$，结合$f^{(n+1)}(x_0) \neq 0$可知$\theta \to \frac{1}{n+1}$。 □

注. 当$f^{(n+1)}(x_0) = 0$时结论通常不成立，比如考虑

$$\arcsin x = \frac{x}{\sqrt{1-(\theta x)^2}},$$

则当$x \to 0$时容易证明$\theta \to 1/\sqrt{3}$。结论本质上依赖于第一个使得$f^{(k)}(x_0) \neq 0$的$k > n$。

例 8.3. 设$f\colon I \to \mathbb{R}$二阶可导，其中I是区间。记$M_k = \sup_{x\in I}|f^{(k)}(x)|\,(k=0,1,2)$。

(a) 若$I = (0,\infty)$且$M_0, M_2 < \infty$，证明：$M_1 \leq 2\sqrt{M_0M_2}$。

(b) 若$I = \mathbb{R}$且$M_0, M_2 < \infty$，证明：$M_1 \leq \sqrt{2M_0M_2}$。

证明. (a) 任意取定$x \in (0,\infty)$，对任意$h > 0$，存在$\xi \in (x, x+h)$使得

$$f(x+h) = f(x) + f'(x)h + \frac{f''(\xi)}{2}h^2.$$

所以

$$|f'(x)| \leq \frac{|f(x+h)| + |f(x)|}{h} + \frac{|f''(\xi)|}{2}h \leq \frac{2M_0}{h} + \frac{M_2}{2}h.$$

取$h = 2\sqrt{\frac{M_0}{M_2}}$即可。[严格来讲这里需要讨论$M_2 = 0$的情况。]

(b) 除了(a)中等式外，还存在$\eta \in (x-h, x)$使得

$$f(x-h) = f(x) - f'(x)h + \frac{f''(\eta)}{2}h^2.$$

两式相减得

$$f(x+h) - f(x-h) = 2f'(x)h + \frac{f''(\xi) - f''(\eta)}{2}h^2.$$

由此可知

$$|f'(x)| \leq \frac{M_0}{h} + \frac{M_2}{2}h.$$

取$h = \sqrt{\frac{2M_0}{M_2}}$即可。 □

下一题的(a)(b)来自偏微分方程中的极值原理，其中(b)构造闸函数的想法不在这门课的范围内；而(c)关心初值问题采用的是完全不同的方法。

例 8.4. 设f在$[0,1]$上二阶可导，$b(x), c(x)\colon [0,1] \to \mathbb{R}$，满足

$$f''(x) + b(x)f'(x) + c(x)f(x) = 0.$$

(a) 如果$f(0) = f(1) = 0$，且$c(x) < 0$。证明：f恒等于0。

*(b)** 如果$f(0) = f(1) = 0$，$b(x), c(x)$有下界，且$c(x) \le 0$。证明：f恒等于0。

*(c)** 如果$f(0) = f'(0) = 0$，且$b(x), c(x)$有界。证明：f恒等于0。

证明. (a) 首先说明f在$[0,1]$上不能有非负值。否则一定有最大值点$x \in (0,1)$，从而$f''(x) \le 0$，$f'(x) = 0$，$f(x) > 0$，结合$c(x) < 0$可知（注意$c(x)f(x)$提供严格不等号）

$$f''(x) + b(x)f'(x) + c(x)f(x) < 0,$$

矛盾！因此在$[0,1]$上有$f \le 0$，用$-f$代替f可知$f \ge 0$，故$f = 0$。

(b) 对于待定的常数$\varepsilon, \alpha > 0$，考虑$g(x) := f(x) + \varepsilon e^{\alpha x}$，则

$$g''(x) + b(x)g'(x) + c(x)g(x) = 0 + \varepsilon(\alpha^2 + b(x)\alpha + c(x))e^{\alpha x}.$$

因为$b(x), c(x)$有下界，固定充分大的α，则上式右边严格大于0。仿照(a)可知$g(x)$不能在(a,b)内部取到非负最大值（注意这里用到了上式严格大于0，而导数条件只能给出不严格的不等号），所以

$$f(x) \le g(x) \le \sup\{0, g(0), g(1)\} = \sup\{0, \varepsilon, \varepsilon e^{\alpha}\}.$$

令$\varepsilon \to 0$可知$f(x) \le 0$。同理$f(x) \ge 0$，故$f = 0$。

(c) 由条件可知$|f''(x)| \le C(|f(x)| + |f'(x)|)$，结合$f(0) = f'(0) = 0$可知$f$恒等于0。[先前讨论过一阶的情况，当时的注指出了一般的情形是完全一样的。] □

练习题

下一题需要用到Vandermonde行列式：

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_0^{n-1} & x_1^{n-1} & \cdots & x_n^{n-1} \\ x_0^n & x_1^n & \cdots & x_n^n \end{vmatrix} = \prod_{i>j}(x_i - x_j).$$

在构造辅助函数时，想法是如果有两列相等则行列式为0。

题 8.5. 设f在$[a,b]$上连续且在(a,b)上n次可导，若$x_0,\dots,x_n$是$[a,b]$中$n+1$个互异的数，证明：存在$\xi\in(a,b)$，使得

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_0 & x_1 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_0^{n-1} & x_1^{n-1} & \cdots & x_n^{n-1} \\ f(x_0) & f(x_1) & \cdots & f(x_n) \end{vmatrix} = \frac{f^{(n)}(\xi)}{n!}\prod_{i>j}(x_i-x_j).$$

证明. 考虑函数$g\colon [a,b]\to\mathbb{R}$，定义如下：

$$g(x) := \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_0 & x_1 & \cdots & x_n & x \\ \vdots & \vdots & & \vdots & \vdots \\ x_0^{n-1} & x_1^{n-1} & \cdots & x_n^{n-1} & x^{n-1} \\ x_0^n & x_1^n & \cdots & x_n^n & x^n \\ f(x_0) & f(x_1) & \cdots & f(x_n) & f(x) \end{vmatrix}.$$

则$g(x)$有$n+1$个零点$x_0,\dots,x_n$。反复使用Rolle中值定理可知，存在$\xi\in(a,b)$使得$g^{(n)}(\xi)=0$。注意到$g(x)$的表达式中只有最后一列与x有关，求g的n阶导就是将最后一列变成n阶导，即$(0,\dots,0,n!,f^{(n)}(x))^T$，按最后一列展开即得结论。 □

下一题暂时偏离微分中值定理的讨论，借助Vandermonde行列式给出三次方程判别式的表达式。设n次多项式f有n个复根$z_1,\dots,z_n$，则$\Delta(f):=\prod_{i<j}(z_i-z_j)^2$称为$f$的判别式。判别式等于0当且仅当$f$有重根。在代数和数论中，判别式具有重要地位，例如可以判断素数在数域中是否分裂。注意到$\Delta(f)$是$z_1,\dots,z_n$的对称多项式，先前证明过对称多项式都可以用n个初等对称多项式表示，结合Vieta定理就可以用多项式本身的系数表示，类比本题的构造以及Newton多项式的递推关系可以给出了n次方程判别式的一般算法。

题 8.6 *. 已知三次多项式$f(x)=x^3+px+q$的三个根为$z_1,z_2,z_3\in\mathbb{C}$，求（用p,q表示）

$$\Delta(f) := (z_1-z_2)^2(z_1-z_3)^2(z_2-z_3)^2.$$

证明. 记$S_n=z_1^n+z_2^n+z_3^n$（即所谓Newton多项式），根据Vandermonde行列式以及矩阵乘法

可知

$$\Delta(f) = \begin{vmatrix} 1 & 1 & 1 \\ z_1 & z_2 & z_3 \\ z_1^2 & z_2^2 & z_3^2 \end{vmatrix} \cdot \begin{vmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \end{vmatrix} = \begin{vmatrix} S_0 & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{vmatrix}.$$

下面计算S_n。由Vieta定理可知$S_0 = 3$，$S_1 = 0$，$S_2 = S_1^2 - 2(z_1z_2 + z_1z_3 + z_2z_3) = -2p$。因为$a_n := z_i^n$满足递推公式$a_{n+3} + pa_{n+1} + qa_n = 0$，所以$S_n$也满足这一关系，由此可知

$$S_3 = -pS_1 - qS_0 = -3q \quad 和 \quad S_4 = -pS_2 - qS_1 = 2p^2.$$

最后，直接计算行列式得

$$\Delta(f) = \begin{vmatrix} 3 & 0 & -2p \\ 0 & -2p & -3q \\ -2p & -3q & 2p^2 \end{vmatrix} = -12p^3 + 8p^3 - 27q^2 = -4p^3 - 27q^2. \qquad \square$$

下一题的构造看似不讲道理，但实际上通过研究$n = 1, 2$就不难猜出。

题 8.7. *设f在$[a, b]$上$n-1$次连续可导、在(a, b)上n次可导，如果$f^{(i)}(a) = f^{(i)}(b) = 0\,(i = 0, 1, \ldots, n-1)$，证明：存在$\xi \in (a, b)$使得$f^{(n)}(\xi) = f(\xi)$。*

证明. 考虑辅助函数

$$F(x) := e^{-x}(f(x) + f'(x) + \cdots + f^{(n-1)}(x)).$$

则F在$[a, b]$上连续、在(a, b)上可导，且$F(a) = F(b) = 0$。根据Rolle中值定理，存在$\xi \in (a, b)$使得$F'(\xi) = e^{-x}(f(\xi) - f^{(n)}(\xi)) = 0$，即$f^{(n)}(\xi) = f(\xi)$。 $\square$

下一题拓展了例题，核心方法是对Taylor公式线性组合消去中间阶导数。

题 8.8 *. *设$f \in C^\infty(\mathbb{R})$，对于$k \in \mathbb{N}_0$，记$M_k := \sup_{\mathbb{R}} |f^{(k)}|$。设$n > m \geq 0$。*

(a) 若$M_m, M_n < \infty$，证明：对任意$m \leq k \leq n$都有$M_k < \infty$。

(b) 若$\lim_{x\to\infty} f^{(m)}(x)$收敛且$M_n$有界，证明：对任意$m < k < n$都有$\lim_{x\to\infty} f^{(k)}(x) = 0$。

(c) 若M_n有界，证明：f一致连续当且仅当$M_1 < \infty$。

证明. (a) 用$f^{(m)}$代替f，不妨设$m = 0$。任意取定$h > 0$[这里取$h = 1$就行了，但(c)会用到其他的h]，则对任意$x \in \mathbb{R}$，存在$\xi_i\,(i = 1, \ldots, n-1)$使得

$$f(x + ih) = f(x) + \frac{f'(x)}{1!}ih + \cdots + \frac{f^{(n-1)}(x)}{(n-1)!}(ih)^{n-1} + \frac{f^{(n)}(\xi_i)}{n!}(ih)^n, \quad i = 1, \ldots, n-1.$$

将其视为关于$f'(x),\dots,f^{(n-1)}(x)$的线性方程组，即

$$\begin{cases}\sum_{k=0}^{n-1}\frac{h^k}{k!}f^{(k)}(x)=f(x+h)-f(x)-\frac{f^{(n)}(\xi_1)}{n!}h^n,\\ \sum_{k=0}^{n-1}\frac{(2h)^k}{k!}f^{(k)}(x)=f(x+2h)-f(x)-\frac{f^{(n)}(\xi_2)}{n!}(2h)^n,\\ \cdots\\ \sum_{k=0}^{n-1}\frac{((n-1)h)^k}{k!}f^{(k)}(x)=f(x+(n-1)h)-f(x)-\frac{f^{(n)}(\xi_{n-1})}{n!}((n-1)h)^n.\end{cases}$$

利用Vandermonde行列式可知系数行列式不等于0，所以存在常数$a_{kj}\ (1\le k,j\le n-1)$（由$n,h$决定），使得

$$f^{(k)}(x)=\sum_{j=1}^{n-1}a_{kj}\left(f(x+jh)-f(x)-\frac{f^{(n)}(\xi_j)}{n!}(jh)^n\right).$$

特别的，a_{kj}与x无关，所以对$1\le k\le n-1$有

$$M_k\le\sum_{j=1}^{n-1}|a_{kj}|\left(2M_0+\frac{M_n}{n!}(jh)^n\right)<\infty.$$

(b) 不妨设$m=0$且$\lim_{x\to\infty}f(x)=0$，并重新定义$M_k:=\sup_{x\in(0,\infty)}|f^{(k)}(x)|$。根据(a)的证明可知$M_1,\dots,M_{n-1}<\infty$，因此只需证明$\lim_{x\to\infty}f'(x)=0$，然后递归即可。于是不妨设$(m,k,n)=(0,1,2)$，此时最快速的方法是借助例题：

$$\sup_{x\ge R}|f'(x)|\le C\sqrt{\sup_{x\ge R}|f(x)|\cdot\sup_{x\ge R}|f''(x)|}\le C\cdot o(1)\cdot M_2=o(1).$$

直接证明也并不复杂：对任意$\varepsilon>0$，因为$\lim_{x\to\infty}f(x)$收敛，存在$N>0$使得当$x>N$时$|f(x+\varepsilon)-f(x)|<\varepsilon^2$。根据Lagrange中值定理可知，存在$\xi\in(x,x+\varepsilon)$使得$f(x+\varepsilon)-f(x)=\varepsilon f'(\xi)$，从而$|f'(\xi)|<\varepsilon$。进而又有

$$|f'(x)|\le|f'(\xi)|+|f'(x)-f'(\xi)|\le\varepsilon+M_2|x-\xi|=(1+M_2)\varepsilon.$$

上式对任意$x>N$成立，由ε的任意性可知$f'(x)\to 0$。

(c) “当”是显然的，下证“仅当”。由一致连续可知，存在$\delta>0$使得当$|x-y|\le\delta$时$|f(x)-f(y)|<1$。在(a)的证明中取$h=\frac{\delta}{n-1}$，则对$1\le j\le n-1$都有$|f(x+jh)-f(x)|<1$，故

$$M_1\le\sum_{j=1}^{n-1}|a_{1j}|\left(1+\frac{M_n}{n!}(jh)^n\right)<\infty.\qquad\square$$

注. *(a)*可以做得更精细一些：将$Taylor$展开式视为关于$h^k f^{(k)}(x)$的方程组，则可以看出$a_{kj} = c_{kj}h^{-k}$，其中c_{kj}是不依赖h的常数，于是存在与h无关的常数C使得

$$M_k \le C(M_0 h^{-k} + M_n h^{n-k}).$$

要让右边尽可能小，应该让$M_0 h^{-k} \approx M_n h^{n-k}$，即$h \approx (M_0/M_n)^{1/n}$，即

$$M_k \le C M_0^{\frac{n-k}{n}} M_n^{\frac{k}{n}}.$$

一般的，对任意$m < k < n$都有

$$M_k \le C M_m^{\frac{n-k}{n-m}} M_n^{\frac{k-m}{n-m}}.$$

此外，*(a)(b)*的结论对在题述范围外的k通常不成立，比如$k < m$时考虑k次多项式，$k > n$时考虑$\frac{1}{x^n}\sin(x^2)$（这个函数不是C^∞的，但结论本质上只关心无穷远处，考虑光滑截断即可）。

以下三题至今仍然难以看清构造有何道理，只能说是结合经验与观察。

题 8.9 *. 设$f\colon [a,b] \to \mathbb{R}$可导，且$f'(a) = f'(b)$，证明：存在$\xi \in (a,b)$，使得

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

证明. 考虑辅助函数

$$F(x) = \begin{cases} f'(a), & x = a, \\ \frac{f(x)-f(a)}{x-a}, & x \in (a,b]. \end{cases}$$

则F连续且在$(a,b]$上可导。事实上，F的导函数是

$$F'(x) = \frac{f'(x)(x-a) - (f(x) - f(a))}{(x-a)^2} = \frac{1}{x-a}\left(f'(x) - \frac{f(x)-f(a)}{x-a}\right), \quad \forall x \in (a,b].$$

因此只需证明存在$\xi \in (a,b)$使得$F'(\xi) = 0$。由$f'(a) = f'(b)$得

$$F'(b) = \frac{1}{b-a}\left(f'(b) - \frac{f(b)-f(a)}{b-a}\right) = \frac{1}{b-a}\left(F(a) - F(b)\right).$$

这说明$F'(b)$与$F(a) - F(b)$同号。如果不存在$\xi \in (a,b)$使得$F'(\xi) = 0$，不妨设F'在(a,b)上严格大于0，则F严格单增，从而$F'(b) \ge 0$（注意这里并没有使用F'连续，而是由单调性推出$F'(b) \ge 0$）且$F(a) - F(b) < 0$，矛盾！从而结论成立。 □

题 8.10 *. 设f在$[0,1]$上连续可导、在$(0,1)$上二次可导，并且$f(0) = 2$，$f'(0) = -2$，$f(1) = 1$。证明：存在$\xi \in (0,1)$使得$f(\xi)f'(\xi) + f''(\xi) = 0$。

证明. 记$g(x)=f(x)^2+2f'(x)$，则结论等价于$g'(\xi)=0$。注意到$g(0)=0$，因此只需再证明g还有其他零点。形式上$g(x)/f(x)^2=1-2(1/f(x))'$，因此按照f是否有零点，分两种情况讨论。

如果恒有$f(x)>0$，则$\frac{1}{f(x)}$可导，存在$\eta\in(0,1)$使得

$$-\frac{f'(\eta)}{f(\eta)^2}=\frac{d}{dx}\left(\frac{1}{f}\right)(\eta)=\frac{1}{f(1)}-\frac{1}{f(0)}=\frac{1}{2}.$$

这正好是$g(\eta)=0$。

而如果$f(x)$在$[0,1]$上有零点，记最小和最大的零点分别为$x_1,x_2\in(0,1)$。因为$f(0)>0$，所以在$[0,x_1)$上$f(x)>0$，从而$f'(x_1)\leq 0$。同理，因为$f(1)>0$，所以$f'(x_2)\geq 0$。于是$g(x_1)=2f'(x_1)\leq 0$且$g(x_2)=2f'(x_2)\geq 0$，从而存在$\eta\in[x_1,x_2]$使得$g(\eta)=0$。

综上所述，g在$(0,1)$上一定还有零点，从而结论成立。 □

题 8.11*. 设$f\colon[0,1]\to\mathbb{R}$可导，且$f(0)=0$，$f(1)=1$。证明：对任意$k_1,\dots,k_n>0$，存在$0<x_1<\cdots<x_n<1$，使得$f'(x_i)\neq 0$且

$$\sum_{i=1}^{n}\frac{k_i}{f'(x_i)}=\sum_{i=1}^{n}k_i.$$

证明. 根据齐次性，不妨设$\sum_{i=1}^{n}k_i=1$。由连续性以及$f(0)=0$、$f(1)=1$可知，存在$0=y_0<y_1<\cdots<y_n=1$使得$f(y_0)=0$且$f(y_i)=k_1+\cdots+k_i\,(i=1,\dots,n)$。根据Lagrange中值定理，存在$x_i\in(y_{i-1},y_i)$使得

$$f'(x_i)(y_i-y_{i-1})=f(y_i)-f(y_{i-1})=k_i.$$

特别的，$f'(x_i)\neq 0$，并且

$$\sum_{i=1}^{n}\frac{k_i}{f'(x_i)}=\sum_{i=1}^{n}(y_i-y_{i-1})=y_n-y_0=1.$$ □

Lagrange插值公式的应用

经常需要直接或间接使用Lagrange插值公式，即任意给定$n+1$个互异实数$x_0,\dots x_n$和任意$a_0,\dots,a_n$，存在唯一的不超过n次的多项式P满足$P(x_i)=a_i\,(i=0,1,\dots n)$，即

$$P(x)=\sum_{i=0}^{n}a_i\prod_{j\neq i}\frac{x-x_j}{x_i-x_j}.$$

本质上这来自于线性方程组解的存在唯一性（以及Vandermonde行列式保证了系数矩阵可逆）。有时还希望指定P的导数值，则需要手动进行求解（当然，解未必存在唯一）。

第一个应用与微分中值定理无关，但结论本身有趣且在代数和数论中有用。

题 8.12 *. 设$x_0,\dots,x_n$是$n+1$个互异实数，证明：

$$\sum_{i=0}^{n}\frac{x_i^k}{\prod_{j\neq i}(x_i-x_j)}=\begin{cases}0, & k=0,1,\dots,n-1,\\ 1, & k=n.\end{cases}$$

证明. 由Lagrange插值公式可知，对$k=0,1,\dots,n$有

$$x^k=\sum_{i=0}^{n}x_i^k\prod_{j\neq i}\frac{x-x_j}{x_i-x_j}.$$

比较两边x^n项的系数即可。 □

题 8.13. *(a)* 设f在$[-1,1]$上二阶连续可导、且在$(-1,1)$上三阶可导，且$f(0)=f'(0)=0$，$f(1)=1$，$f(-1)=0$。证明：存在$\xi\in(-1,1)$使得$f'''(\xi)=3$。

(b) 设f在$[a,b]$上连续且在(a,b)上二阶可导，对任意$c\in(a,b)$，证明：存在$\xi\in(a,b)$，使得

$$\frac{f''(\xi)}{2}=\frac{f(a)}{(a-b)(a-c)}+\frac{f(b)}{(b-a)(b-c)}+\frac{f(c)}{(c-a)(c-b)}.$$

证明. (a) 记$P(x)=\frac{1}{2}x^2(x+1)$，则P与f在$0,-1,1$处的函数值相等、且在0处的导数值相等。反复利用Rolle中值定理可知，存在$\xi\in(-1,1)$使得$f'''(\xi)=P'''(\xi)=3$。

(b) 考虑辅助多项式（即Lagrange插值公式）

$$P(x)=f(a)\frac{(x-b)(x-c)}{(a-b)(a-c)}+f(b)\frac{(x-a)(x-c)}{(b-a)(b-c)}+f(c)\frac{(x-a)(x-b)}{(c-a)(c-b)}.$$

则P与f在a,b,c处相等，存在$\xi\in(a,b)$使得$f''(\xi)=P''(\xi)$，此即欲证的结论。 □

下一题的背景是Lagrange插值公式的误差估计（用$f-P$代替题中的f，其中多项式P满足$P(x_i)=f(x_i)\,(i=1,\dots,n)$且次数不超过$n-1$）。从另一种观点看，这是一种广义的Taylor展开式：Taylor展开式用满足$P^{(k)}(x)=f^{(k)}(x_0)\,(k=0,1,\dots,n-1)$的至多$n-1$次多项式$P$逼近$f$，并用$f^{(n)}(\xi)(x-x_0)^n/n!$作为误差估计；而如果换成用Lagrange多项式来逼近，即指定n个不同的点处的函数值，则本题给出相应的（带Lagrange余项的）展开式。例如，对于$f\in C^2[0,1]$有

$$f(x)=f(a)+(x-a)f(b)+\frac{f''(\xi)}{2}(x-a)(x-b),\ \xi\in(a,b).$$

题 8.14. 设f在$[a,b]$上$n-1$阶连续可导、且在(a,b)上n阶可导，若f在$[a,b]$上有n个互异零点$x_1,\dots,x_n$，证明：对任意$x\in[a,b]$，存在$\xi\in[a,b]$，使得

$$f(x)=\frac{f^{(n)}(\xi)}{n!}(x-x_1)\cdots(x-x_n).$$

证明. 若$x \in \{x_1, \ldots, x_n\}$，则结论显然成立。不妨设$x \notin \{x_1, \ldots, x_n\}$，考虑辅助多项式

$$P(t) := \frac{(t - x_1) \cdots (t - x_n)}{(x - x_1) \cdots (x - x_n)} f(x).$$

则$P(t) - f(t)$有$n+1$个不同的零点$x, x_1, \ldots, x_n$，因此存在$\xi \in (a, b)$使得$f^{(n)}(\xi) = P^{(n)}(\xi)$，即

$$f^{(n)}(\xi) = \frac{n!}{(x - x_1) \cdots (x - x_n)} f(x). \qquad \square$$

注. 如果将零点条件推广为存在k个互异的点$x_1, \ldots, x_k \in [a, b]$满足对$i = 1, \ldots, k$都有

$$f(x_i) = f'(x_i) = \cdots = f^{(e_i - 1)}(x_i) = 0,$$

其中$e_i \in \mathbb{N}$满足$e_1 + \cdots + e_k = n$，则同样的方法可以说明

$$f(x) = \frac{f^{(n)}(\xi)}{n!} \prod_{i=1}^{k} (x - x_i)^{e_i}, \quad \xi \in (a, b).$$

再次岔开话题。如果将上一题理解为Lagrange插值公式的误差估计，自然要问如何选取插值点$x_1, \ldots, x_n$才能使得$(x - x_1) \cdots (x - x_n)$尽可能小（并且不依赖$x \in [a, b]$）。那么问题化归为当$Q$取遍所有$n$次首一多项式时，$\max_{x \in [-1,1]} |Q(x)|$的最小值是多少？（原始的问题还应要求$Q$在$[-1, 1]$上有$n$个零点，但下面的构造自动满足这一条件。）

熟知$[-1, 1]$上n次Chebyshev多项式的定义是

$$Q_n(x) := \cos(n \arccos(x)).$$

为了说明Q_n是n次多项式，只需证明$\cos n\theta$是关于$\cos \theta$的n次多项式。事实上，

$$\begin{aligned}
\cos n\theta &= \mathrm{Re}(\cos(n\theta) + i \sin(n\theta)) = \mathrm{Re}(\cos\theta + i \sin\theta)^n \\
&= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k C_n^{2k} \cos^{n-2k}\theta \sin^{2k}\theta \\
&= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k C_n^{2k} \cos^{n-2k}\theta (1 - \cos^2\theta)^k.
\end{aligned}$$

由此还可看出Q_n的首项系数是$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} C_n^{2k} = 2^{n-1}$。又因为$|Q|$的最大值是1，所以$Q/2^{n-1}$是$n$次首一多项式，并且绝对值的最大值是$2^{1-n}$。下一题说明这一常数是最佳的。

题 8.15 *. 设$n \in \mathbb{N}$，若$P(x)$是n次首一多项式，证明：$\max_{x \in [-1,1]} |P(x)| \geq 2^{1-n}$。

证明. 注意到根据Q_n的定义，存在$n+1$个点$-1 \leq x_0 < \cdots < x_n \leq 1$使得$Q(x_i)$交错取值$\pm 1$。用反证法，假设$|P(x)| < 2^{1-n}$对任意$x \in [-1, 1]$成立，则$Q(x) - 2^{n-1} P(x)$是至多$n-1$次多项式，且在$x_i$处的符号是$(-1)^i$。根据介值原理可知$Q(x) - 2^{n-1} P(x)$至少有$n$个零点，所以只能恒为0，即$P(x) = Q(x)/2^{n-1}$，与$|P(x)| < 2^{1-n}$矛盾！ $\square$

下一题的背景是数值方法求定积分的矩形法、梯形法以及抛物线法的误差估计，今后还会用定积分的分部积分公式重新理解。按照Newton–Leibniz公式有$f(b)-f(a)=\int_a^b f'(x)\,dx$，于是(a)(b)(c)除去误差项分别是将$f'$的图像当作矩形（即$f'$为常数）、梯形（即$f'$为一次函数）、抛物线（即$f'$为二次函数）计算得到的近似结果。

题 8.16 *. *(a)* 设f在$[a,b]$上二次连续可导、在(a,b)上三次可导，证明：存在$\xi\in(a,b)$，使得

$$f(b)=f(a)+(b-a)f'\left(\frac{a+b}{2}\right)+\frac{(b-a)^3}{24}f'''(\xi).$$

(b) 设f在$[a,b]$上二次连续可导、在(a,b)上三次可导，证明：存在$\xi\in(a,b)$，使得

$$f(b)=f(a)+\frac{b-a}{2}(f'(a)+f'(b))-\frac{(b-a)^3}{12}f'''(\xi).$$

(c) 设f在$[a,b]$上四次连续可导、在(a,b)上五次可导，证明：存在$\xi\in(a,b)$，使得

$$f(b)=f(a)+\frac{b-a}{6}\left(f'(a)+4f'\left(\frac{a+b}{2}\right)+f'(b)\right)-\frac{(b-a)^5}{2880}f^{(5)}(\xi).$$

证明. 由于结论关于$\frac{a+b}{2}$对称，自然应该考虑关于中点对称化的奇函数：记$r=\frac{b-a}{2}$，定义

$$g(x)=f\left(\frac{a+b}{2}+x\right)-f\left(\frac{a+b}{2}-x\right),\quad x\in[-r,r].$$

(a) 因为g是奇函数，在0处的Taylor展式为

$$g(r)=rg'(0)+\frac{r^3}{6}g'''(\eta),\quad \eta\in(0,r).$$

上式即

$$f(b)-f(a)=(b-a)f'\left(\frac{a+b}{2}\right)+\frac{(b-a)^3}{48}\left(f'''\left(\frac{a+b}{2}+\eta\right)+f'''\left(\frac{a+b}{2}-\eta\right)\right).$$

根据Darboux中值定理，存在$\xi\in(a,b)$使得

$$f'''\left(\frac{a+b}{2}+\eta\right)+f'''\left(\frac{a+b}{2}-\eta\right)=2f'''(\xi).$$

带入上式可知结论成立。

(b) 同(a)，结合Darboux定理，只需证明存在$\eta\in(-r,r)$使得

$$g(r)=rg'(r)-\frac{r^3}{3}g'''(\eta).$$

为此，希望构造合适的三次多项式$P(x)$。由于上式含有$g(r)$和$g'(r)$，并且g是奇函数，自然希望$P(\pm r)=g(\pm r)$和$P'(\pm r)=g'(\pm r)$。这样的多项式一定形如Ax^3+Bx，容易解出

$$P(x)=\frac{rg'(r)-g(r)}{2r^3}x^3+\frac{3g(r)-cg'(r)}{2r}x.$$

（容易验证$P(x)$满足要求。）反复运用Lagrange中值定理可知，存在$\eta\in(-r,r)$使得

$$g'''(\eta)=P'''(\eta)=\frac{3(rg'(r)-g(r))}{r^3}.$$

整理即得要证的等式。

(c) 同(a)，只需证明存在$\eta\in(-r,r)$使得

$$g(r)=\frac{r}{3}(g'(r)+2g'(0))-\frac{r^5}{180}g^{(5)}(\eta).$$

考虑构造五次多项式$P(x)$，现在自然是在$0,\pm r$处对零阶导和一阶导提条件。一共有六个方程，由此确定出五次多项式为

$$P(x):=\frac{rg'(r)-3g(r)+2rg'(0)}{2r^5}x^5-\frac{rg'(r)-5g(r)+4rg'(0)}{2r^3}x^3+g'(0)x.$$

（容易验证确实有$P(0)=g(0)$，$P'(0)=g'(0)$，$P(\pm r)=g(\pm r)$，$P'(\pm c)=g'(\pm c)$。）于是存在$\eta\in(-r,r)$使得$g^{(5)}(\eta)=P^{(5)}(\eta)$，整理即可。 □

注. *(a)如果不借助$g(x)$，其实就是将$f(b)$和$f(a)$在$\frac{a+b}{2}$处的Taylor展开式相减。(b)还可以通过反复运用Cauchy中值定理得到，这是因为要证的结论即*

$$f'''(\xi)=\frac{\varphi(b)-\varphi(a)}{\psi(b)-\psi(a)},$$

其中

$$\varphi(x):=f(x)-f(a)-\frac{x-a}{2}(f'(a)+f'(x)),\quad \psi(x)=-\frac{(x-a)^3}{12}.$$

这种方法看似很有道理，但实际上对(a)难以实现（正负号会出问题）。

凸函数

开区间上的凸函数通常有三种理解方式：(1)代数观点即定义，或者等价于差商$\frac{f(x)-f(y)}{x-y}$ $(x\neq y)$关于x,y分别单增；(2)解析观点，左/右导数处处存在且单增；(3)几何观点，每一点处的支撑线存在，即对任何一点x，存在常数k使得

$$f(t)\geq f(x)+k(t-x),\quad \forall t.$$

稍后会证明后两种观点同样是凸性的等价刻画。这些视角具有不同的局部/整体特征：比如左/右导数就是局部性质，而支撑线是整体性质（稍后会证明局部支撑线也是足够的）。闭区间上的情况并没有本质不同，只是端点处可能有跳跃。以下为了避免麻烦都假设定义域是$\mathbb{R}$。

题 8.17. *(a)* 设$f\colon \mathbb{R} \to \mathbb{R}$连续，且满足对任意$x, y \in \mathbb{R}$有

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

证明：f是凸函数。

(b) 将*(a)*中的连续性减弱为“f在任何有限区间上都有上界”，证明结论依然成立。

(c) 将*(b)*进一步减弱为“f在某个非空开区间(a, b)上有上界”，证明结论依然成立。

证明. (a) 容易归纳证明对任意$n \in \mathbb{N}$和$m \in \{0, 1, \dots, 2^n\}$有

$$f\left(\frac{m}{2^n} \cdot x + \left(1 - \frac{m}{2^n}\right) \cdot y\right) \le \frac{m}{2^n} \cdot f(x) + \left(1 - \frac{m}{2^n}\right) \cdot f(y).$$

[事实上，之前指出过上式对于将$m/2^n$换成任意$q \in [0, 1] \cap \mathbb{Q}$都成立，这里不需要那么强的结论。]再结合形如$\frac{m}{2^n}$的有理数在$[0, 1]$中稠密，以及$f$的连续性即可。

(b) 不妨设$x = 0$且$y = 1$。用$f(x) - (f(1) - f(0))x - f(0)$代替$f(x)$，可设$f(0) = f(1) = 0$。只需证明当$x \in (0, 1)$时$f(x) \le 0$。递归定义$x_0 = x$，如果$x_{n-1}$已经定义，则

$$x_n = \begin{cases} 2x_{n-1}, & x_n \in [0, 1/2], \\ 2x_{n-1} - 1, & x \in (1/2, 1]. \end{cases}$$

则x_{n-1}是x_n与0或1的中点，根据条件以及归纳可知$f(x_n) \ge 2f(x_{n-1}) \ge 2^n f(x_0)$。因为$f$在$[0, 1]$上有上界，所以必然$f(x_0) \le 0$。

(c) 不妨设$(a, b) = (-1, 1)$。根据(b)，只需证明f在任何有限区间上都有上界；任取$R > 0$，根据对称性又只需证明f在$(0, R)$上有上界。取充分接近x/R且形如$m/2^n$的参数$q \in [0, 1] \in \mathbb{Q}$，则使得$x = qy + (1-q)R$的参数$y$满足$y \in (-1, 1)$。于是

$$f(x) \le qf(y) + (1-q)f(R) \le \sup_{t \in (-1,1)} f(t) + f(R),$$

右边不依赖x而仅与R有关，故f在$(0, R)$上有上界。 □

题 8.18 * (2018 USAMO). 求所有函数$f\colon (0, \infty) \to (0, \infty)$，满足当$x, y, z > 0$且$xyz = 1$时，

$$f\left(x + \frac{1}{y}\right) + f\left(y + \frac{1}{z}\right) + f\left(z + \frac{1}{x}\right) = 1.$$

证明. 通过标准的换元$x=a/b$，$y=b/c$，$z=c/a$，可知

$$f\left(\frac{b+c}{a}\right)+f\left(\frac{a+c}{b}\right)+f\left(\frac{a+b}{c}\right)=1,\quad \forall a,b,c>0.$$

考虑换元$g(x):=f(1/x-1)$，其中$x\in(0,1)$。上式改写为

$$g(x)+g(y)+g(z)=1,\quad \forall x,y,z>0,\ x+y+z=1.$$

注意到$0<g<1$，并且对任意$x,y\in(0,1/2)$都有

$$g(x)+g(y)=1-g(1-x-y)=2g\left(\frac{x+y}{2}\right).$$

由上一题(b)可知g在$(0,1/2)$上是一次函数，设为$g(x)=kx+m$。代入方程得$k+3m=1$，即

$$g(x)=kx+\frac{1-k}{3},\quad \forall x\in(0,1/2).$$

而当$x\in[1/2,1)$时，因为$(1-x)/2<1/2$，所以

$$g(x)=1-2g\left(\frac{1-x}{2}\right)=1-2\left(k\cdot\frac{1-x}{2}-\frac{1-k}{3}\right)=kx+\frac{1-k}{3}.$$

这就说明这一表达式对一切$x\in(0,1)$都成立。为了保证当$x\in(0,1)$时$g(x)>0$，k的范围是$[-1/2,1]$。最后，代回f的表达式可知

$$f(x)=\frac{k}{x+1}+\frac{1-k}{3},\quad k\in[-1/2,1].$$

容易验证这样的f满足要求。 □

题 8.19. 设$f\colon\mathbb{R}\to\mathbb{R}$是凸函数，证明：

(a) $\lim_{x\to\infty}\frac{f(x)}{x}$存在且属于$(-\infty,\infty]$，$\lim_{x\to-\infty}\frac{f(x)}{x}$存在且属于$[-\infty,\infty)$。

(b) 如果$\lim_{|x|\to\infty}\frac{f(x)}{x}=0$，则$f$是常数。

证明. (a) 由f是凸函数可知$\frac{f(x)-f(0)}{x-0}$关于$x>0$单调递增，从而$\frac{f(x)}{x}$在$x\to\infty$时极限存在（可能等于∞但不会是$-\infty$）。同理可得$x\to-\infty$的情况。

(b) 对任意$a<b$，由凸性可知对$x>b$和$y<a$有

$$\frac{f(b)-f(y)}{b-y}\le\frac{f(b)-f(a)}{b-a}\le\frac{f(x)-f(a)}{x-a}.$$

令$x\to\infty$和$y\to-\infty$可知$\frac{f(b)-f(a)}{b-a}=0$，即$f(b)=f(a)$。 □

最后几道题关心凸函数与导数、支撑线的关系，常用的性质是凸函数f的左、右导数$f'_\pm$（在定义域内部）存在、单调递增，并且$f'_-(x)\le f'_+(x)$。

题 8.20. 设$f\colon \mathbb{R}\to\mathbb{R}$是凸函数，证明：

(a) f局部*Lipschitz*连续（即f在任何有限闭区间$[a,b]$上*Lipschitz*连续）。

(b) f在除了至多可数个点之外可导。

证明. (a) 这是因为熟知对任意$a\le x<y\le b$都有

$$f'_+(a)\le \frac{f(x)-f(y)}{x-y}\le f'_-(b).$$

(b) 由于f'_+单增、不连续点至多可数，只需证明f在f'_+的连续点处可导。事实上，当$h>0$时

$$f'_+(x-2h)\le f'_-(x-h)\le f'_+(x)\le f'_-(x+h)\le f'_+(x+h).$$

由此只要f'_+在x处连续，则f'_-在x处的极限等于$f'_+(x)$，从而$f'_-(x)=f'_+(x)$。 □

题 8.21. *(a)* 证明：$f\colon \mathbb{R}\to\mathbb{R}$是凸函数当且仅当对任意$x\in\mathbb{R}$，存在$k\in\mathbb{R}$使得

$$f(y)\ge k(y-x)+f(x),\quad \forall y\in\mathbb{R}.$$

*(b)** 设$f\colon \mathbb{R}\to\mathbb{R}$连续，且对任意$x\in\mathbb{R}$，存在$\delta>0$和$k\in\mathbb{R}$使得

$$f(y)\ge k(y-x)+f(x),\quad \forall y\in(x-\delta,x+\delta),$$

证明：f是凸函数。

证明. (a) “仅当”是因为k可以取成$f'_+(x)$或$f'_-(x)$，下面证明“当”。对任意$x,y\in\mathbb{R}$和$\lambda\in[0,1]$，记$z:=\lambda x+(1-\lambda)y$。根据条件，存在由$z$决定的常数$k$使得

$$f(x)\ge k(x-z)+f(z)\quad 且\quad f(y)\ge k(y-z)+f(z).$$

注意到$\lambda(x-z)+(1-\lambda)(y-z)=0$，因此由上面两个不等式可知

$$\lambda f(x)+(1-\lambda)f(y)\ge f(z)=f(\lambda x+(1-\lambda)y).$$

(b) 为了说话方便，称题中的条件为f在x处是“好的”；进一步，如果条件加强为

$$f(y)>k(y-x)+f(x),\quad \forall y\in(x-\delta,x+\delta)\setminus\{x\},$$

则称f在x处是“严格好的”。在每一点处都好的函数称为好函数，在每一点处都严格好的函数称为严格好函数。因为凸函数在每一点处都是好的，所以f加上任何凸函数也是好的；进一步，f加上任何严格凸函数都是严格好函数。

任意取定$\varepsilon > 0$，令$g(x) = f(x) + \varepsilon x^2$，希望证明$g$是凸函数。如果这一结论得证，则令$\varepsilon \to 0$可知$f$是凸函数。为了证明$g$是凸函数，根据(a)可知，只需证明对任意$x \in \mathbb{R}$，都存在$k \in \mathbb{R}$使得对任意$y \neq x$都有

$$g(y) > k(y-x) + g(x).$$

因为g是严格好函数，可以取k使得上式对任意$y \in (x-\delta, x+\delta) \setminus \{x\}$成立，其中$\delta > 0$。用$h(y) := g(y) - k(y-x) - g(x)$代替$g(y)$，则$h$也是严格好函数，并且满足

$$h(y) > 0, \quad \forall y \in (x-\delta, x+\delta) \setminus \{x\}.$$

于是只需证明上式对任意$y \neq x$成立。

用反证法，如果结论不成立，不妨设存在$y > x$使得$h(y) \le 0$（$y < x$的情况是类似的）。记$z := \min\{y > x \mid h(y) \le 0\}$，因为这个集合非空，再结合条件可知$z > x$。于是在$(x, z)$上$h$严格大于0，可以取$w \in (x, z)$是$h$在$[x, z]$上的最大值点。因为$h$在$w$处是严格好的，存在$l \in \mathbb{R}$和$\eta > 0$，使得当$t \in (-\eta, \eta)$时有

$$h(w+t) > lt + h(w).$$

又因为$h(w+t) \le h(w)$，得到$lt < 0$对$t \in (-\eta, \eta)$恒成立，但这是不可能的！ □

注. *(b)中的连续性是不可或缺的，比如$f = -\chi_{\{0\}}$。从几何上看，证明是说如果g图像并不始终位于局部的支撑线l_x上方，则存在一个点w离支撑线具有最大的正距离，而这个点处不可能再有局部的支撑线l_w：因为l_w左边或者右边的点离l_x的距离更大（取决于l_x与l_w斜率的大小关系）。这里为了让等号变严格，再次用到构造闸函数εx^2来保证严格凸的技巧。*

题 8.22. 设$f\colon \mathbb{R} \to \mathbb{R}$连续，满足右导数$f'_+$存在且单增，证明：$f$是凸函数。

证明*1.* 先前证明过右导数的Lagrange中值定理：任给$x < y < z$，则

$$\frac{f(y)-f(x)}{y-x} \le f'_+(\xi), \quad \xi \in (x, y),$$

$$\frac{f(z)-f(y)}{z-y} \ge f'_+(\eta), \quad \eta \in (y, z).$$

单调性表明$f_+(\xi) \le f'_+(\eta)$，所以$\frac{f(y)-f(x)}{y-x} \le \frac{f(z)-f(y)}{z-y}$，熟知这是凸性的等价刻画。 □

证明*2.* 根据上一题的(a)，只需证明任给$x \in \mathbb{R}$，都有

$$f(y) \ge f(x) + f'_+(x)(y-x), \quad \forall y \in \mathbb{R}.$$

事实上，记$g(y)=f(y)-f'_+(x)(y-x)$，则g连续且$g'_+(y)=f'_+(y)-f'_+(x)$。当$y>x$时$g'_+(y)\geq 0$，先前证明过右导数非负的连续函数必然单增，因此当$y>x$时$g(y)$单增，特别的有$g(y)\geq g(x)$；同理，当$y<x$时$g'_+(y)\leq 0$，因此g单减，从而$g(y)\geq g(x)$。 □

注. 利用本题的结论容易证明：如果$a<b<c<d$，函数f在(a,c)以及(b,d)上分别是凸函数，则f在(a,d)上是凸函数。当然，直接按定义证也不困难，只是写起来略微复杂。

补充：Basel问题与$\zeta(2n)$

这一小节先给出Basel问题的初等做法，然后借助所谓Herglotz技巧证明cot的一个非平凡展开式，并利用其将$\zeta(2n)$用Bernoulli数表示。今后学习Fourier级数时还会给出另外的证明。

题 8.23. 本题的目标是证明$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$。

(a) 对任意$n\in\mathbb{N}$，证明：

$$\sin((2n+1)x)=\sum_{k=0}^{n}(-1)^k C_{2n+1}^{2k+1}\cos^{2n-2k}x\cdot\sin^{2k+1}x.$$

(b) 对任意$n\in\mathbb{N}$，证明：

$$\sum_{k=1}^{n}\cot^2\frac{k\pi}{2n+1}=\frac{n(2n-1)}{3}.$$

(c) 证明：对$0<x<\frac{\pi}{2}$有$\frac{1}{x^2}-1<\cot^2 x<\frac{1}{x^2}$。

(d) 证明：$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$。

证明. (a) 利用$\cos((2n+1)x)+i\sin((2n+1)x)=(\cos x+i\sin x)^{2n+1}$，比较两边虚部即可。

(b) 考虑n次多项式

$$P(x)=\sum_{k=0}^{n}(-1)^k C_{2n+1}^{2k+1}x^{n-k}.$$

则由(a)可知$\cot^2\frac{k\pi}{2n+1}(k=1,\ldots,n)$是$P$的$n$个互异实根。根据Vieta定理得

$$\sum_{k=1}^{n}\cot^2\frac{k\pi}{2n+1}=-\frac{(-1)^0C_{2n+1}^3}{(-1)^1C_{2n+1}^1}=\frac{n(2n-1)}{3}.$$

(c) 这不过是$\sin x<x<\tan x$的另一种写法。

(d) 结合(b)(c)可知

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

即
$$\frac{\pi^2}{3}\cdot\frac{n(2n-1)}{(2n+1)^2}<\sum_{k=1}^{n}\frac{1}{k^2}<\frac{\pi^2}{3}\cdot\frac{n(2n+2)}{(2n+1)^2}.$$
显然左、右两边的极限都是$\frac{\pi^2}{6}$，从而结论成立。 □

题 8.24 *. 考虑由如下递推关系定义的数列：
$$a_{n+1}=a_n+\frac{a_n^2}{n^2},\quad n\in\mathbb{N}.$$

(a) (2023阿里巴巴全球数学竞赛)设$a_1=\frac{2}{5}$，证明：对任意$n\in\mathbb{N}$有$a_n<1$。

(b) (2024阿里巴巴全球数学竞赛)设$0\le a_1<1$，证明：极限$\lim_{n\to\infty}a_n$存在且有限。

证明. (a) 显然a_n非负且单增。将递推式改写为
$$\frac{1}{a_{n+1}}=\frac{1}{a_n}-\frac{1}{a_n+n^2}.$$
于是当$n\ge 3$时，
$$\frac{1}{a_n}=\frac{1}{a_1}-\frac{1}{a_1+1}-\sum_{k=2}^{n-1}\frac{1}{a_k+k^2}\ge\frac{5}{2}-\frac{5}{7}-\sum_{k=2}^{\infty}\frac{1}{k^2}=\frac{25}{14}-\frac{\pi^2}{6}+1>1,$$
最后一步估算$\frac{25}{14}>\frac{\pi^2}{6}$是容易的。而上式即$a_n<1$。

(b) $a_1=0$的情况是显然的，因此不妨设$a_1\in(0,1)$。归纳可知$a_n\le na_1$（这里仅仅用到了$a_1\le 1$）。若结论不成立，则$1/a_n$单减趋于0，于是Stolz定理保证了
$$\lim_{n\to\infty}\frac{a_n}{n}=\lim_{n\to\infty}\frac{\frac{1}{n}}{\frac{1}{a_n}}=\lim_{n\to\infty}\frac{\frac{1}{n}-\frac{1}{n+1}}{\frac{1}{a_n}-\frac{1}{a_{n+1}}}=\lim_{n\to\infty}\frac{\frac{1}{n(n+1)}}{\frac{1}{a_n+n^2}}=\lim_{n\to\infty}\frac{n^2+a_n}{n^2}=1.$$
其中最后一步用到了$a_n/n^2\le a_1/n\to 0$。但这与$a_n/n\le a_1<1$矛盾！ □

最后，借助cot的展开式给出所有$\zeta(2n)=1^{-2n}+2^{-2n}+3^{-2n}+\cdots$的表达式。

题 8.25 *. 本题的目标是证明cot的如下展开式：
$$\pi\cot(\pi x)=\frac{1}{x}+\sum_{n=1}^{\infty}\left(\frac{1}{x-n}+\frac{1}{x+n}\right),\quad\forall x\notin\mathbb{Z}.$$
对于$x\notin\mathbb{Z}$，定义$f(x)=\pi\cot(\pi x)-\frac{1}{x}-\sum_{n=1}^{\infty}\frac{2x}{x^2-n^2}$，证明：

(a) f在$\mathbb{R}\setminus\mathbb{Z}$上是连续函数。

(b) $f(x+1)=f(x)$对任意$x\notin\mathbb{Z}$成立。

(c) f在$x \in \mathbb{Z}$处的极限都是0，从而f可以延拓为$\mathbb{R}$上周期为1的连续函数。

(d) 对任意$x \in \mathbb{R}$，有

$$f\left(\frac{x}{2}\right) + f\left(\frac{x+1}{2}\right) = 2f(x).$$

(e) f恒等于0，即$\pi\cot(\pi x) = \frac{1}{x} + \sum_{n=1}^{\infty}\left(\frac{1}{x-n} + \frac{1}{x+n}\right)$。

证明. (a) 当只需证明对每个$N \in \mathbb{N}$，f在$|x| \in (N-1, N)$上连续。将原式改写为

$$f(x) = \pi\cot(\pi x) - \frac{1}{x} - \sum_{n=1}^{N}\frac{2x}{x^2-n^2} - \sum_{n=N+1}^{\infty}\frac{2x}{x^2-n^2}.$$

除了最后一个无穷级数外，都是连续函数。注意到

$$\sum_{n=N+1}^{\infty}\left|\frac{2x}{x^2-n^2}\right| \leq \sum_{n=N+1}^{\infty}\frac{2N}{n^2-N^2} < \infty.$$

根据大M判别法，这个无穷级数一致收敛，故f在$|x| \in (N-1, N)$上连续。

(b) 将$f(x)$写成

$$f(x) = \pi\cot(\pi x) - \lim_{N\to\infty}\sum_{n=-N}^{N}\frac{1}{x+n}.$$

则

$$\begin{aligned}
f(x+1) - f(x) &= -\lim_{N\to\infty}\left(\sum_{n=-N}^{N}\frac{1}{x+n+1} - \sum_{n=-N}^{N}\frac{1}{x+n}\right) \\
&= -\lim_{N\to\infty}\left(\sum_{n=-N+1}^{N+1}\frac{1}{x+n} - \sum_{n=-N}^{N}\frac{1}{x+n}\right) \\
&= -\lim_{N\to\infty}\left(\frac{1}{x+N+1} - \frac{1}{x-N}\right) = 0.
\end{aligned}$$

(c) 根据周期性，只需证明f在0处的极限是0。同(a)可知当$|x| < \frac{1}{2}$时$\sum_{n=1}^{\infty}\frac{2x}{x^2-n^2}$一致收敛，并且显然在0处等于0。根据Taylor展式可知$\pi\cot(\pi x) - \frac{1}{x}$在0处极限也是0。

(d) 根据连续性，只需处理$x \notin \mathbb{Z}$的情况。记$g(x) = \pi\cot(\pi x)$和$h(x) = f(x) - g(x)$，则有

$$\begin{aligned}
g\left(\frac{x}{2}\right) + g\left(\frac{x+1}{2}\right) &= \pi\frac{2\cos^2\frac{\pi x}{2}}{\sin(\pi x)} + \pi\frac{2\cos^2\frac{\pi(x+1)}{2}}{\sin(\pi(x+1))} \\
&= 2\pi\frac{\cos^2\frac{\pi x}{2} - \sin^2\frac{\pi x}{2}}{\sin(\pi x)} = 2g(x),
\end{aligned}$$

以及

$$h\left(\frac{x}{2}\right)+h\left(\frac{x+1}{2}\right)=\lim_{N\to\infty}\sum_{n=-N}^{N}\left(\frac{2}{x+2n}+\frac{2}{x+2n+1}\right)$$
$$=2\lim_{N\to\infty}\sum_{n=-2N}^{2N+1}\frac{1}{x+n}=2h(x).$$

两式相加即证。

(e) 由周期性以及连续性可知，f有最大值点$x_0\in[0,1]$。由(d)可知

$$2\max f=2f(x_0)=f\left(\frac{x_0}{2}\right)+f\left(\frac{x_0+1}{2}\right)\le 2\max f.$$

因此$f\left(\frac{x_0}{2}\right)=f(x_0)$。同理可证对任意$n\in\mathbb{N}$都有$f\left(\frac{x_0}{2^n}\right)=f(x_0)$。又因为左边趋于$f(0)=0$，所以$f(x_0)=0$，即$f\le 0$。同理可证$f\ge 0$，故$f=0$。 □

题 8.26 *. 本题的目标是证明对任意$n\in\mathbb{N}$都有

$$\zeta(2n)=\frac{(-1)^{n-1}2^{2n-1}B_{2n}}{(2n)!}\pi^{2n}.$$

(a) 利用上一题的结果，证明：当$|x|$充分小时，

$$x\cot x=1-\sum_{n=1}^{\infty}\frac{2\zeta(2n)}{\pi^{2n}}x^{2n}.$$

(b) 证明：对于$z\in\mathbb{C}$，当$|z|$充分小时，

$$\frac{z}{e^z-1}=\sum_{k=0}^{\infty}\frac{B_k}{k!}z^k.$$

(c) 利用三角函数的指数形式，证明：当$|x|$充分小时，

$$x\cot x=\sum_{n=0}^{\infty}\frac{(-1)^n2^{2n}B_{2n}}{(2n)!}x^{2n}.$$

与*(a)*比较系数可得$\zeta(2n)$的表达式。

证明. (a) 简单展开即可，其中第二行交换顺序合法性容易利用$\sum 1/n^2<\infty$说明：

$$
\begin{aligned}
x\cot x &= 1-\sum_{n=1}^{\infty}\frac{2x^2}{n^2\pi^2-x^2} = 1-\sum_{n=1}^{\infty}\frac{2x^2}{n^2\pi^2\left(1-\frac{x^2}{n^2\pi^2}\right)} \\
&= 1-\sum_{n=1}^{\infty}\sum_{k=0}^{\infty}\frac{2x^2}{n^2\pi^2}\cdot\left(\frac{x^2}{n^2\pi^2}\right)^k = 1-\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{n^{2k}}\frac{2x^{2k}}{\pi^{2k}} = 1-\sum_{k=1}^{\infty}\frac{2\zeta(2k)}{\pi^{2k}}x^{2k}.
\end{aligned}
$$

(b) 如果事先建立了$\mathbb{C}$上的解析函数理论，那么结论就是显然的。先前所以没有那样做，主要是为了避开复可微的概念。这里简单就所需的事实进行说明。

设$f(x)=\sum_{n=0}^{\infty}a_nx^n$具有收敛半径$r>0$，则将$x$替换为任意满足$|z|<r$的复数$z$，右边都收敛，从而可以定义$f(z)$。如果$a_0=f(0)\neq 0$，则$1/f$在0附近可以展为收敛的幂级数

$$\frac{1}{f(x)}=\sum_{n=0}^{\infty}b_nx^n.$$

断言上式对x换成（模长充分小的）z也是成立的，这是因为只需验证

$$\left(\sum_{n=0}^{\infty}a_nz^n\right)\left(\sum_{n=0}^{\infty}b_nz^n\right)=1,$$

而这件事（根据Cauchy乘积公式）完全取决于a_n,b_n的代数运算，而与自变量是x还是z无关。回到原题，结论对应于$f(z)=(e^z-1)/z=\sum_{n=0}^{\infty}z^n/(n+1)!$的特例。

(c) 由(b)可知（注意$B_1=-\frac{1}{2}$且$B_{2n+1}=0$）

$$
\begin{aligned}
x\cot x &= \frac{ix(e^{2ix}+1)}{e^{2ix}-1} = ix+\frac{2ix}{e^{2ix}-1} \\
&= ix+\sum_{k=0}^{\infty}\frac{B_k}{k!}(2ix)^k = \sum_{n=0}^{\infty}\frac{(-1)^n2^{2n}B_{2n}}{(2n)!}x^{2n}.
\end{aligned}
$$

□

还可以考虑复数版本的推广。Gauss证明了

$$\sum_{(a,b)\in\mathbb{Z}^2\setminus\{(0,0)\}}\frac{1}{(a+bi)^4}=\frac{1}{15}\left(2\int_0^1\frac{dx}{\sqrt{1-x^4}}\right)^4.$$

Gauss通过类比$\pi=2\int_0^1\frac{dx}{\sqrt{1-x^2}}$，猜想椭圆积分$2\int_0^1\frac{dx}{\sqrt{1-x^4}}$也有重要地位。上面的结果揭示了冰山一角，其证明需要用到复变函数中关于双周期函数（即椭圆函数）的理论。这一结论实则给出模形式理论中Eisenstein级数的特殊值$E_2(i)$，具有深厚的代数和数论背景，与格的复乘(complex multiplication)有关。另一个相关的结果是，若记$\omega:=e^{2\pi i/3}$是三次单位根，则

$$\sum_{(a,b)\in\mathbb{Z}^2\setminus\{(0,0)\}}\frac{1}{(a+b\omega)^6}=\frac{27}{2240}\left(2\int_0^1\frac{dx}{\sqrt{1-x^3}}\right)^6.$$

第9讲：不定积分

“你确定那不是用来撬门的撬棍吗？”

——阿瑟·柯南·道尔《海军协定》

这一讲主要是不定积分的计算题，掌握换元法、分部积分法以及常见的原函数足矣。双曲三角换元往往容易被忽略，题目中给出了一例。实际上，常用的结果

$$\int \frac{dx}{\sqrt{x^2 \pm a^2}} = \log |x + \sqrt{x^2 \pm a^2}| + C$$

无非是反双曲三角函数。使用换元法时应当注意变量的范围，但这里为了简洁没有特别注明。忽视变量范围以及混淆不同变量容易导致错误，因此求导检验必不可少。

应当避免不加限制地使用复数，比如以下计算是错误的：

$$\int \frac{dx}{1+x^2} = \int \frac{1}{2i} \left(\frac{1}{x-i} - \frac{1}{x+i} \right) dx = \frac{1}{2i} \log \frac{x-i}{x+i} + C.$$

问题在于log在复平面上并不是良定义的单值函数。

新教材还讨论了连续函数原函数的存在性，为了避开定积分以及Newton–Leibniz公式，选择在不引进一致收敛等概念的前提下手动构造分段线性逼近，实属野蛮至极。今后学习定积分自会明白，如果$f \in C[a,b]$，则变上限积分$F(x) := \int_a^x f(t)\,dt$就是$f$的原函数。更复杂的问题可能涉及更深入的积分理论，最后一题就是这样的例子。

练习题

下一题的(f)(g)两问显然可以推广到$\int \frac{dx}{\sqrt{ax^2+bx+c}}$，但公式并不简洁，写下来意义不大。

题 9.1. 计算下列带有根式的不定积分：

(a) $\int x^3 \sqrt[3]{1+x^2}\, dx$。

(b) $\int \sqrt{1+x^2}\, dx$。

(c) $\int x^2 \sqrt{x^2+1}\, dx$。

(d) $\int \frac{dx}{x^2\sqrt{x^2-1}}$。

(e) $\int \frac{dx}{1+\sqrt{x}+\sqrt{1+x}}$。

(f) $\int \frac{dx}{\sqrt{(x-a)(b-x)}}$，其中$a < b$。

(g) $\int \frac{dx}{\sqrt{x^2+x}}$。

(h) $\int \frac{x^3}{\sqrt{x^4-2x^2-1}}\,dx$。

解. (a) 换元$t=\sqrt[3]{1+x^2}$，则$dt=\frac{2x}{3t^2}\,dx$，所以

$$\int x^3\sqrt[3]{1+x^2}=\int \frac{3t^3(t^3-1)}{2}\,dt=\frac{3}{14}t^7-\frac{3}{8}t^4+C=\frac{3}{14}(1+x^2)^{\frac{7}{3}}-\frac{3}{8}(1+x^2)^{\frac{4}{3}}+C.$$

(b) 分部积分，得

$$\int\sqrt{1+x^2}\,dx=x\sqrt{1+x^2}-\int\frac{x^2}{\sqrt{1+x^2}}\,dx=x\sqrt{1+x^2}-\int\sqrt{1+x^2}\,dx+\log(x+\sqrt{1+x^2})+C.$$

因此

$$\int\sqrt{1+x^2}\,dx=\frac{1}{2}\left(x\sqrt{1+x^2}+\log(x+\sqrt{1+x^2})\right)+C.$$

(c) 分部积分可知

$$\begin{aligned}\int x^2\sqrt{x^2+1}\,dx&=\frac{1}{3}\int xd(x^2+1)^{\frac{3}{2}}=\frac{1}{3}x(x^2+1)^{\frac{3}{2}}-\frac{1}{3}\int(x^2+1)^{\frac{3}{2}}\,dx\\&=\frac{1}{3}x(x^2+1)^{\frac{3}{2}}-\frac{1}{3}\int x^2\sqrt{x^2+1}\,dx-\frac{1}{3}\int\sqrt{1+x^2}.\end{aligned}$$

结合(b)得

$$\int x^2\sqrt{x^2+1}\,dx=\frac{1}{4}x(x^2+1)^{\frac{3}{2}}-\frac{1}{8}\left(x\sqrt{x^2+1}+\log(x+\sqrt{x^2+1})\right)+C.$$

(d) 换元$x=\cosh t\,(t\geq 0)$，则$dx=\sinh t\,dt$，所以

$$\int\frac{dx}{x^2\sqrt{x^2-1}}=\int\frac{dt}{\cosh^2 t}=\tanh t=\frac{\sqrt{x^2-1}}{x}+C.$$

(e) 注意到$\frac{1}{1+\sqrt{x}+\sqrt{1+x}}=\frac{1+\sqrt{x}-\sqrt{1+x}}{2\sqrt{x}}$，换元$x=t^2$并结合(b)得

$$\begin{aligned}\int\frac{dx}{1+\sqrt{x}+\sqrt{1+x}}&=\int(1+t-\sqrt{1+t^2})\,dt\\&=\sqrt{x}+\frac{1}{2}x-\frac{1}{2}\sqrt{x(1+x)}-\frac{1}{2}\log(\sqrt{x}+\sqrt{1+x})+C.\end{aligned}$$

(f) 熟知$\int\frac{dx}{\sqrt{c^2-x^2}}=\arcsin\frac{x}{c}+C\,(c>0)$，所以

$$\int\frac{dx}{\sqrt{(x-a)(b-x)}}=\int\frac{dx}{\sqrt{\left(\frac{b-a}{2}\right)^2-\left(x-\frac{a+b}{2}\right)^2}}=\arcsin\frac{2x-(a+b)}{b-a}+C.$$

另一种常见做法是考虑换元$t=\sqrt{(x-a)(b-x)}$，则$dx=\sqrt{(x-a)(b-x)}\,dt$，所以

$$\int \frac{dx}{\sqrt{(x-a)(b-x)}} = 2\arcsin\sqrt{\frac{x-a}{b-a}} + C.$$

[这两种表达式当然是相同的，因为$\arcsin(2u^2-1) = 2\arcsin u - \pi/2\,(u\in(0,\pi/2))$。]

(g) 熟知$\int \frac{dx}{\sqrt{x^2\pm a^2}} = \log|x+\sqrt{x^2\pm a^2}|+C$，所以

$$\begin{aligned}\int \frac{dx}{\sqrt{x^2+x}} &= \int \frac{dx}{\sqrt{\left(x+\frac{1}{2}\right)^2-\frac{1}{4}}} = \log\left|\left(x+\frac{1}{2}\right)+\sqrt{\left(x+\frac{1}{2}\right)^2-\frac{1}{4}}\right|+C \\ &= \log\left|x+\frac{1}{2}+\sqrt{x^2+x}\right|+C.\end{aligned}$$

(h) 换元$t=x^2$，则$dt=2x\,dx$，所以

$$\begin{aligned}\int \frac{x^3}{\sqrt{x^4-2x^2-1}}\,dx &= \int \frac{t}{2\sqrt{t^2-2t-1}}\,dt = \int \frac{d(t^2-2t-1)}{4\sqrt{t^2-2t-1}} + \int \frac{dt}{2\sqrt{(t-1)^2-2}}, \\ &= \frac{1}{2}\sqrt{t^2-2t-1} + \frac{1}{2}\log\left|t-1+\sqrt{t^2-2t-1}\right|+C \\ &= \frac{1}{2}\sqrt{x^4-2x^2-1} + \frac{1}{2}\log\left|x^2-1+\sqrt{x^4-2x^2-1}\right|+C.\end{aligned}$$

□

题 9.2. 计算下列带有三角函数的不定积分：

(a) $\int \frac{dx}{1-\cos x}$。

(b) $\int \frac{\sin x+\cos x}{\sqrt[3]{\sin x-\cos x}}\,dx$。

(c) $\int \tan^3 x\,dx$。

(d) $\int \frac{dx}{\sin x}$。

(e) $\int \sin^4 x\,dx$。

(f) $\int \csc^4 x\,dx$。

(g) $\int x\tan x\sec^2 x\,dx$。

(h) $\int \frac{x\arctan x}{(1+x^2)^2}\,dx$。

解. (a) 用二倍角公式，得

$$\int \frac{dx}{1-\cos x} = \int \frac{dx}{2\sin^2\frac{x}{2}} = -\cot\frac{x}{2}+C.$$

(b) 换元$t=\sin x-\cos x$，则$dt=(\cos x+\sin x)\,dx$，所以

$$\int \frac{\sin x+\cos x}{\sqrt[3]{\sin x-\cos x}}\,dx=\int \frac{dt}{\sqrt[3]{t}}=\frac{3}{2}t^{\frac{2}{3}}+C=\frac{3}{2}(\sin x-\cos x)^{\frac{2}{3}}+C.$$

(c) 换元$t=\cos x$，则

$$\int \tan^3 x\,dx=\int \frac{-(1-t^2)}{t^3}\,dt=\frac{1}{2\cos^2 x}+\log|\cos x|+C.$$

(d) 利用二倍角公式$\sin x=2\sin\frac{x}{2}\cos\frac{x}{2}$，得

$$\begin{aligned}\int \frac{dx}{\sin x}&=\int \frac{d\left(\sin\frac{x}{2}\right)}{\sin\frac{x}{2}\left(1-\sin^2\frac{x}{2}\right)}=\int \frac{d\left(\sin^2\frac{x}{2}\right)}{\sin^2\frac{x}{2}\left(1-\sin^2\frac{x}{2}\right)}\\&=\frac{1}{2}\left(\log\left|\sin^2\frac{x}{2}\right|-\log\left|1-\sin^2\frac{x}{2}\right|\right)+C=\log\left|\tan\frac{x}{2}\right|+C.\end{aligned}$$

(e) 反复运用二倍角公式，有$\sin^4 x=\left(\frac{1-\cos(2x)}{2}\right)^2=\frac{1}{4}-\frac{1}{2}\cos(2x)+\frac{1+\cos(4x)}{8}$，所以

$$\int \sin^4 x\,dx=\frac{3}{8}x-\frac{1}{4}\sin(2x)+\frac{1}{32}\sin(4x)+C.$$

(f) 换元$t=\cot x$，则$dt=-\csc^2 x\,dx$，所以

$$\int \csc^4 x\,dx=-\int(1+t^2)\,dt=-\cot x-\frac{1}{3}\cot^3 x+C.$$

(g) 注意到$\tan x\sec^2 x\,dx=d\left(\frac{1}{2}\sec^2 x\right)$，分部积分得

$$\int x\tan x\sec^2 x\,dx=\frac{1}{2}x\sec^2 x-\frac{1}{2}\int\sec^2 x=\frac{1}{2}x\sec^2 x-\frac{1}{2}\tan x+C.$$

(h) 注意到$\frac{x\,dx}{(1+x^2)^2}=-\frac{1}{2}d\left(\frac{1}{1+x^2}\right)$，所以

$$\begin{aligned}\int \frac{x\arctan x}{(1+x^2)^2}\,dx&=-\frac{1}{2}\left(\frac{\arctan x}{1+x^2}-\int\frac{dx}{(1+x^2)^2}\right)\\&=-\frac{1}{2}\cdot\frac{\arctan x}{1+x^2}-\frac{1}{4}\left(\frac{1}{x(1+x^2)}+\int\frac{dx}{x^2(1+x^2)}\right)\\&=-\frac{1}{2}\cdot\frac{\arctan x}{1+x^2}-\frac{1}{4}\left(\frac{1}{x(1+x^2)}-\frac{1}{x}-\arctan x\right)+C\\&=\frac{1}{4}\cdot\frac{x^2-1}{x^2+1}\cdot\arctan x+\frac{1}{4}\cdot\frac{x}{1+x^2}+C.\end{aligned}$$

□

题 9.3. 计算下列带有指数/对数函数的不定积分：

(a) $\int \frac{dx}{1+e^x}$。

(b) $\int \frac{dx}{\sqrt{e^x-1}}$。

(c) $\int \frac{x\log x}{(1+x^2)^2}\,dx$。

*(d)** $\int \left(1-\frac{2}{x}\right)^2 e^x\,dx$。

解. (a) 换元$t=1+e^x$，则

$$\int \frac{dx}{1+e^x} = \int \frac{dt}{t(t-1)} = \log\frac{t-1}{t} + C = \log\frac{e^x}{1+e^x} + C.$$

(b) 换元$t=\sqrt{e^x-1}$，则$dt=\frac{t^2+1}{2t}dx$，所以

$$\int \frac{dx}{\sqrt{e^x-1}} = \int \frac{2\,dt}{t^2+1} = 2\arctan\sqrt{e^x-1} + C.$$

(c) 注意到$\frac{x\,dx}{(1+x^2)^2} = -\frac{1}{2}d\left(\frac{1}{1+x^2}\right)$，所以

$$\begin{aligned}\int \frac{x\log x}{(1+x^2)^2}\,dx &= -\frac{1}{2}\left(\frac{\log x}{1+x^2} - \int\frac{dx}{x(1+x^2)}\right)\\ &= -\frac{1}{2}\left(\frac{\log x}{1+x^2} - \log|x| + \frac{1}{2}\log(1+x^2)\right) + C.\end{aligned}$$

(d) 分部积分可知

$$\int \frac{e^x}{x}\,dx = \frac{e^x}{x} + \int\frac{e^x}{x^2}\,dx.$$

所以

$$\int\left(1-\frac{2}{x}\right)^2 e^x\,dx = e^x - 4\int\frac{e^x}{x}\,dx + 4\int\frac{e^x}{x^2}\,dx = e^x - 4\cdot\frac{e^x}{x} + C.$$ □

注. $\int\frac{e^x}{x}\,dx$不是初等函数，因此*(d)*中利用分部积分进行配对是必要的。

题 9.4. 计算下列有理函数的不定积分：

(a) $\int\left(\frac{x}{x^2-3x+2}\right)^2 dx$。

(b) $\int\frac{x}{(x-1)^2(x^2+2x+2)}\,dx$。

(c) $\int\frac{dx}{(x+a)^2(x+b)^2}$，其中$a,b\in\mathbb{R}$且$a\neq b$。

解. (a) 注意到$\left(\frac{x}{x^2-3x+2}\right)^2 = \frac{1}{(x-1)^2} + \frac{4}{(x-2)^2} - \frac{4}{(x-1)(x-2)}$，所以

$$\int \left(\frac{x}{x^2-3x+2}\right)^2 dx = -\frac{1}{x-1} - \frac{4}{x-2} + 4\log\left|\frac{x-1}{x-2}\right| + C.$$

(b) 注意到$\frac{x}{(x-1)^2(x^2+2x+2)} = \frac{1}{5(x-1)^2} - \frac{1}{25(x-1)} - \frac{x+8}{25(x^2+2x+2)}$，所以

$$\begin{aligned}\int \frac{x}{(x-1)^2(x^2+2x+2)}\,dx &= -\frac{1}{5(x-1)} + \frac{1}{25}\log|x-1| - \frac{1}{25}\int \frac{(x+1)+7}{x^2+2x+2}\,dx \\ &= -\frac{1}{5(x-1)} + \frac{1}{25}\log|x-1| - \frac{1}{50}\log(x^2+2x+2) - \frac{7}{25}\arctan(x+1) + C.\end{aligned}$$

(c) 注意到$\frac{1}{(x+a)^2(x+b)^2} = \frac{1}{(b-a)^2}\left(\frac{1}{(x+a)^2} + \frac{1}{(x+b)^2} - \frac{2}{(x+a)(x+b)}\right)$，所以

$$\int \frac{dx}{(x+a)^2(x+b)^2} = \frac{1}{(b-a)^2}\left(-\frac{1}{x+a} - \frac{1}{x+b} - \frac{2}{b-a}\log\left|\frac{x+a}{x+b}\right|\right) + C. \qquad \square$$

题 9.5. *(a)* 求$I_n := \int \frac{dx}{(1+x^2)^n}$ $(n \in \mathbb{N})$的递推公式。

(b) 求$\int \frac{\sin(nx)}{\sin x}\,dx$，其中$n \in \mathbb{N}_0$。

解. (a) 对于$n \in \mathbb{N}$，分部积分得

$$\begin{aligned}I_n &= \frac{x}{(1+x^2)^n} + 2n\int \frac{x^2}{(1+x^2)^{n+1}}\,dx \\ &= \frac{x}{(1+x^2)^n} + 2n\int \left(\frac{1}{(1+x^2)^n} - \frac{1}{(1+x^2)^{n+1}}\right) dx \\ &= \frac{x}{(1+x^2)^n} + 2n(I_n - I_{n+1}).\end{aligned}$$

即

$$I_{n+1} = \frac{1}{2n}\cdot\frac{x}{(1+x^2)^n} + \frac{2n-1}{2n}I_n.$$

(b) 记这个式子为I_n，和差化积可知

$$I_{n+2} - I_n = \int 2\cos((n+1)x)\,dx = \frac{2\sin((n+1)x)}{n+1} + C.$$

再结合$I_0 = C$以及$I_1 = x + C$可知，对$n \in \mathbb{N}_0$有

$$I_{2n} = 2\sum_{k=1}^{n}\frac{\sin((2k-1)x)}{2k-1} + C \quad 和 \quad I_{2n+1} = x + 2\sum_{k=1}^{n}\frac{\sin(2kx)}{2k} + C. \qquad \square$$

题 9.6. 用配对法计算下列不定积分：

(a) $\int \frac{\sin x}{2\sin x+3\cos x}\,dx$。

(b) $\int \frac{dx}{1+x^4}$。

*(c)** $\int \frac{dx}{x^8+x^4+1}$。

解. (a) 记$I := \int \frac{\sin x}{2\sin x+3\cos x}\,dx$和$J := \int \frac{\cos x}{2\sin x+3\cos x}\,dx$，则

$$2I + 3J = \int dx = x + C,$$

$$-3I + 2J = \int \frac{d(2\sin x + 3\cos x)}{2\sin x + 3\cos x} = \log|2\sin x + 3\cos x|.$$

解得

$$I = \frac{2}{13}x - \frac{3}{13}\log|2\sin x + 3\cos x| + C.$$

(b) 记$I := \int \frac{dx}{1+x^4}$和$J := \int \frac{x^2}{1+x^4}\,dx$，则

$$I + J = \int \frac{1+\frac{1}{x^2}}{x^2+\frac{1}{x^2}}\,dx = \int \frac{d\left(x-\frac{1}{x}\right)}{\left(x-\frac{1}{x}\right)^2+2} = \frac{1}{\sqrt{2}}\arctan\frac{x-\frac{1}{x}}{\sqrt{2}} + C,$$

$$I - J = -\int \frac{1-\frac{1}{x^2}}{x^2+\frac{1}{x^2}}\,dx = -\int \frac{d\left(x+\frac{1}{x}\right)}{\left(x+\frac{1}{2}\right)^2-2} = -\frac{1}{2\sqrt{2}}\log\frac{x+\frac{1}{x}-\sqrt{2}}{x+\frac{1}{x}+\sqrt{2}} + C.$$

解得

$$I = -\frac{1}{4\sqrt{2}}\log\frac{x^2-\sqrt{2}x+1}{x^2+\sqrt{2}x+1} + \frac{1}{2\sqrt{2}}\arctan\frac{x^2-1}{\sqrt{2}x} + C,$$
$$J = \frac{1}{4\sqrt{2}}\log\frac{x^2-\sqrt{2}x+1}{x^2+\sqrt{2}x+1} + \frac{1}{2\sqrt{2}}\arctan\frac{x^2-1}{\sqrt{2}x} + C.$$

(c) 注意到$\frac{1}{x^8+x^4+1} = \frac{1}{2}\left(\frac{1+x^2}{x^4+x^2+1} + \frac{1-x^2}{x^4-x^2+1}\right)$，利用(b)的技巧容易将两项都积出来，得

$$\int \frac{dx}{x^8+x^4+1} = \frac{1}{2\sqrt{3}}\arctan\frac{x^2-1}{\sqrt{3}x} - \frac{1}{4\sqrt{3}}\log\frac{x^2-\sqrt{3}x+1}{x^2+\sqrt{3}x+1} + C.$$ □

注. *(b)*也可以直接用部分分式分解进行计算：利用共轭复根配对可知

$$\frac{1}{x^4+1} = \frac{1}{2\sqrt{2}}\left(\frac{x+\sqrt{2}}{x^2+\sqrt{2}x+1} - \frac{x-\sqrt{2}}{x^2-\sqrt{2}x+1}\right).$$

最后一题是为了说明，现阶段最好的做法是直接忽略关于原函数存在性的问题。今后自会有更一般的理论（前三种方法）给出简单证明，初等证明（第四种方法）复杂而多余。

题 9.7 *. 设$f\colon [a,b]\to\mathbb{R}$有界且存在原函数，$g\colon [a,b]\to\mathbb{R}$连续，证明：$fg$的原函数存在。

证明1. 记f的原函数为F，因为f有界，所以F是有界变差函数：记V_a^b是全变差，则

$$V_a^b(F)\le (b-a)\sup|f|.$$

因为g连续，可以定义变上限Riemann–Stieltjes积分

$$\Phi(x):=\int_a^x g\,dF.$$

容易证明Φ的导数就是fg：以a处的导数为例（其他点处是类似的），将$\Phi(x)$分解为

$$\Phi(x)=g(a)\int_a^x dF+\int_a^x (g-g(a))\,dF=g(a)(F(x)-F(a))+\int_a^x (g-g(a))\,dF.$$

因为F是f的原函数，所以前一项在a处的导数就是$f(a)g(a)$；而后一项的绝对值不超过

$$\sup_{t\in[a,x]}|g(t)-g(a)|\cdot V_a^x(F)\le o(1)\cdot|x-a|\sup|f|=o(|x-a|),$$

因此导数是0；故Φ在a处可导且$\Phi'(a)=f(a)g(a)$。 □

证明2. 根据实变函数的理论，f作为其原函数F的导数，必然Lebesgue可测，结合有界性可知Lebesgue可积。逐点可导加上导函数Lebesgue可积足以保证$F(x)=\int_a^x f\,dm$（这是一个较强版本的微积分基本定理）。因为g连续，所以fg有界可测也是Lebesgue可积的。定义$\Phi(x)=\int_a^x fg\,dm$（变上限Lebegue积分），可以证明这就是fg的原函数：仍然只以a处为例，

$$\Phi(x)=g(a)\int_a^x f\,dm+\int_a^x f(g-g(a))\,dm=g(a)F(x)+\int_a^x f(g-g(a))\,dm,$$

后一项绝对值不超过$o(1)\cdot\int_a^x|f|\,dm\le o(|x-a|)$，所以$\Phi'(a)=F'(a)g(a)=f(a)g(a)$。 □

证明3. 需要用到下学期关于函数极限的理论。想法是用折线函数逼近g，然后说明对于折线函数结论成立，最后证明结论在逼近下保持。

第一步是容易的。定义g_n是将g在$[a,b]$的n等分点处的函数值用折线连接得到的分段线性函数，因为g连续，容易证明g_n在$[a,b]$上一致收敛到g。又因为f有界，所以fg_n一致收敛到fg。

第二步需要说明fg_n具有原函数。为此，只需在每一段上证明原函数存在然后接起来，根据线性结构又只需证明$xf(x)$具有原函数。本质上就是分部积分：记f的原函数是F，因为F连续，所以F又存在原函数$\int F$，容易验证$xF(x)-\int F(x)$就是$xf(x)$的原函数。

第三步，需要用到如下结论：如果$[a,b]$上的一列函数都有原函数且一致收敛，则极限函数也有原函数。这来自于如下更基本的事实：如果函数列F_n在至少一点处收敛，且导函数F_n'一致收敛，则F_n一致收敛到极限函数F，满足F可导且$F'=\lim F_n'$（注意不需要假设F_n'连续）。回到原题，前两步说明了fg_n都有原函数，且一致收敛到fg，所以结论成立。 □

证明4. 如果只接受已学的知识、避免使用一致收敛的性质，则需手动处理证明3的最后一步。

首先，取Φ_n是fg_n的原函数且$\Phi_n(a)=0$。根据Lagrange中值定理

$$\begin{aligned}&|\Phi_m(x)-\Phi_n(x)| = |(\Phi_m(x)-\Phi_n(x))-(\Phi_m(a)-\Phi_n(a))|\\ \le&|x-a|\sup_{[a,x]}|\Phi_m'(x)-\Phi_n'(x)| \le |x-a|\cdot\sup_{[a,b]}|f|\cdot\sup_{[a,b]}|g_m-g_n|.\end{aligned}$$

因为f有界且g_n一致收敛，上式表明Φ_n一致收敛。[这里实际上只需要Φ_n逐点收敛。]

最后，证明Φ是fg的原函数。任意固定$x\in[a,b]$，需要说明当$y\to x$时

$$\Phi(y)-\Phi(x) = f(x)g(x)(y-x)+o(|y-x|).$$

上式当然来自于对Φ_n取极限，但此时的难点是f没有连续性，使用Lagrange中值定理时无法估计$f(x+\theta(y-x))$与$f(x)$的误差。解决办法是将右边的fg理解为$F'g$，让F'对应的分点与Φ_n'相同。具体来说，对$\Phi_n(\cdot)-g(x)F(\cdot)$用Lagrange中值定理可知，存在$\theta_n\in(0,1)$使得

$$\begin{aligned}&|\Phi_n(y)-\Phi_n(x)-g(x)(F(y)-F(x))|\\ =&|y-x|\cdot|\Phi_n'(x+\theta_n(y-x))-g(x)F'(x+\theta_n(y-x))|\\ =&|y-x|\cdot|F'(x+\theta_n(y-x))|\cdot|g_n(x+\theta_n(y-x))-g(x)|\\ \le&|y-x|\cdot\sup|f|\cdot|g_n(x+\theta_n(y-x))-g(x)|\\ \le&|y-x|\cdot\sup|f|\cdot(\sup_{[a,b]}|g_n-g|+\sup_{\theta\in(0,1)}|g(x+\theta(y-x))-g(x)|).\end{aligned}$$

（注意最后一步将g_n替换为g的放缩。）对于固定的y，令$n\to\infty$可知，

$$|\Phi(y)-\Phi(x)-g(x)(F(y)-F(x))| \le |y-x|\cdot\sup|f|\cdot\sup_{\theta\in(0,1)}|g(x+\theta(y-x))-g(x)|.$$

根据g的连续性可知最后一项是在$y\to x$时是$o(1)$，所以

$$\Phi(y)-\Phi(x) = g(x)(F(y)-F(x))+o(|y-x|).$$

两边除以$y-x$并取极限，结合$F'=f$就证明了$\Phi'=fg$。 □

注. f的有界性是必要的，反例可以考虑$[0,1]$上的函数（注意$g\in C[0,1]$）

$$f(x)=\begin{cases}\frac{1}{\sqrt{x}}\sin\frac{1}{x}, & x\in(0,1],\\ 0, & x=0,\end{cases}\qquad g(x)=\begin{cases}\sqrt{x}\sin\frac{1}{x}, & x\in(0,1],\\ 0, & x=0.\end{cases}$$

可以证明f有原函数，但fg没有原函数；关键在于0处是否与变上限积分的导数匹配：

$$\lim_{x\to0+}\frac{1}{x}\int_0^x\frac{1}{\sqrt{t}}\sin\frac{1}{t}\,dt=0,\quad \lim_{x\to0+}\frac{1}{x}\int_0^x\sin^2\frac{1}{t}=\frac{1}{2}\ne0.$$

第10讲*：单调函数几乎处处可导

它单调地哭泣，
像水在哭泣，
像风在雪上
哭泣。

——费德里科·加西亚·洛尔迦《吉他》

这一讲是补充内容，目的是证明单调函数几乎处处可导。这里介绍的Vitali覆盖尽管麻烦但相当初等，并且在几何测度论中还有进一步的应用。在学习一般的测度论知识之后，可以用Radon–Nikodym定理定理给出简单直接的证明。

Lebesgue测度

区间长度、图形面积等概念已经深入人心，自然的问题是能否给$\mathbb{R}$中的任意集合定义“长度”。具体来讲，是否存在映射$l\colon 2^{\mathbb{R}} \to [0,\infty]$满足如下一些合理的性质：

(a) 如果I是区间，则$l(I)$等于I的长度。

(b) 对任意$X \subset \mathbb{R}$和$t \in \mathbb{R}$有$l(X+t) = l(X)$，其中$X + t := \{x + t \mid x \in X\}$表示平移。

(c) 如果有可数个两两不交的集合$X_n \subset \mathbb{R}\,(n \in \mathbb{N})$，即$X_m \cap X_n = \emptyset\,(m \neq n)$，则

$$l\left(\bigcup_{n\in\mathbb{N}} X_n\right) = \sum_{n=1}^{\infty} l(X_n).$$

前两条性质无需多言，而(c)称为“可数可加性(countably additive)”。由可数可加性可以推出有限可加性，即有限个两两不交的集合之并的长度等于长度之和；此外，还可以推出单调性，即如果$X \subset Y$则$l(X) \leq l(Y)$。之所以要求可数可加而不仅仅是有限可加，在于分析中的实际需要：取极限。意外的是这样的l并不存在，反例的构造用到选择公理。

在$\mathbb{R}$上定义等价关系$\sim$，使得$x \sim y$当且仅当$x - y \in \mathbb{Q}$。每个等价类都有代表元在$(0,1)$中，根据选择公理，存在$X \subset (0,1)$使得X恰好包含每个等价类中的一个元素。对于$q \in (-1,1)\cap\mathbb{Q}$，定义$X_q := X + q$，则$X + q$两两不交，并且

$$(0,1) \subset \bigcup_{q\in(-1,1)\cap\mathbb{Q}} X_q \subset (-1,2).$$

因为$l(X_q) = l(X)$，由可数可加性和单调性可知$1 \leq \sum_{q\in(-1,1)\cap\mathbb{Q}} l(X) \leq 3$，无论$l(X) = 0$还是$l(X) > 0$都会导致矛盾！所以不存在满足要求的映射$l$。

解决方案是限制l的定义域，只对合理的集合定义长度，即所谓的“可测集(measurable set)”。如此可以产生一种自然的构造，即Lebesgue外测度(outer measure)。

定义 10.1 (Lebesgue外测度). 设$X \subset \mathbb{R}$，定义

$$m^*(X) := \inf\left\{\sum_{n=1}^{\infty}|I_n| \mid X \subset \bigcup_{n\in\mathbb{N}} I_n\right\},$$

其中I_n是开区间（允许是空集），$|I_n|$表示I_n的长度。若$m^*(X)=0$，则称X是零测集。

先前证明过不可数个正数之和一定是∞，因此这里只考虑可数个开区间并不失一般性。在涉及∞的运算中，总是规定$0\cdot\infty=0$，$a\cdot\infty=\infty\,(\forall a\in(0,\infty])$以及$b+\infty=\infty\,(\forall b\in[0,\infty])$。

题 10.2. 验证$m^*: 2^{\mathbb{R}} \to [0,\infty]$满足如下性质：

(a) $m^*(\emptyset)=0$，$m^*(X+t)=m^*(X)$，$m^*(\lambda X)=|\lambda|m^*(X)\,(\lambda\in\mathbb{R})$。

(b) 若$A\subset B$，则$m^*(A)\le m^*(B)$。

(c) 若$X_n\subset\mathbb{R}\,(n\in\mathbb{N})$，则$m^*\left(\bigcup_{n\in\mathbb{N}}X_n\right)\le\sum_{n=1}^{\infty}m^*(X_n)$。

(d) 若N是至多可数集，则$m^*(N)=0$。[这结合*(e)*给出$\mathbb{R}$不可数的另一种证明。]

(e) 若I是区间，则$m^*(I)=|I|$。

证明. (a)(b)根据定义是显然的。

(c) 不妨设每个$m^*(X_n)$都有限，对任意$\varepsilon>0$，存在可数个开区间$I_{ni}\,(i\in\mathbb{N})$覆盖X_n且$\sum_{i=1}^{\infty}|I_{ni}|\le m^*(X)+2^{-n}\varepsilon$。于是全体$I_{ni}\,(n,i\in\mathbb{N})$覆盖了$\bigcup_{n\in\mathbb{N}}X_n$，并且

$$\sum_{n=1}^{\infty}\sum_{i=1}^{\infty}|I_{ni}|\le\sum_{n=1}^{\infty}(m^*(X_n)+2^{-n}\varepsilon)=\sum_{n=1}^{\infty}m^*(X_n)+\varepsilon.$$

再根据定义即证。

(d) 显然单点集都是零测集，再结合(c)即可。

(e) 根据(a)(b)(d)，只需证明$I=[0,1]$的情况。因为$[0,1]\subset(-\varepsilon,1+\varepsilon)$，所以$m^*([0,1])\le 1+2\varepsilon$，进而$m^*([0,1])\le 1$。反过来，因为$[0,1]$的开覆盖都有有限子覆盖，不妨设有限个开区间$I_i=(a_i,b_i)\,(i=1,\ldots,n)$构成$[0,1]$的覆盖，并且这些$I_i$互不包含。因为互不包含，不妨设$a_1<\cdots<a_n$，从而$b_1<\cdots<b_n$；又因为覆盖$[0,1]$，所以$a_1<0$，$b_n>1$，并且对$i=1,\ldots,n-1$一定有$b_i>a_{i+1}$。故

$$\sum_{i=1}^{n}|I_i|=\sum_{i=1}^{n}(b_i-a_i)=b_n-a_1+\sum_{i=1}^{n-1}(b_i-a_{i+1})>1. \qquad \square$$

上题中的性质(c)称为“可数次可加性(countably subadditive)”；一般的，非空集合X上满足可数次可加性以及$m^*(\emptyset) = 0$的映射$m^*\colon X \to [0, \infty]$称为外测度。先前已经解释过，可数可加性不可能对所有集合成立，为此引进可测集的概念。

定义 10.3 (可测集). 集合$E \subset \mathbb{R}$称为*Lebesgue*可测的，如果对任意$X \subset \mathbb{R}$，都有

$$m^*(X) = m^*(X \cap E) + m^*(X \setminus E).$$

全体*Lebesgue*可测集组成的集族记为$\mathcal{M}$，$m := m^*|_{\mathcal{M}}$称为*Lebesgue*测度。

注. 根据次可加性，为了验证E可测，只需证明当$m^*(X) < \infty$时

$$m^*(X) \geq m^*(X \cap E) + m^*(X \setminus E).$$

值得注意的是，可测集的定义（结合归纳法）已经表明m具有有限可加性。另一个推论是，如果集合$A \subset B$，其中A是可测集且$m^*(A) < \infty$，则

$$m^*(B \setminus A) = m^*(B) - m^*(A).$$

注意条件$m^*(A) < \infty$是必要的，否则右边出现$\infty - \infty$没有意义。

下一题以及稍后的定理将说明m确实具有可数可加性，并且可测集是很多的。

题 10.4. *(a)* 证明：零测集都是可测集。

(b) 设$E_n \, (n \in \mathbb{N})$是两两不交的可测集，证明：[稍后会证明左边的$m^*$可以换成$m$]

$$m^*\left(\bigcup_{n \in \mathbb{N}} E_n\right) = \sum_{n=1}^{\infty} m(E_n).$$

(c) 如果$A, B \subset \mathbb{R}$且$\operatorname{dist}(A, B) := \inf\{|a - b| \mid a \in A,\ b \in B\} > 0$，证明：

$$m^*(A \cup B) = m^*(A) + m^*(B).$$

*(d)** 证明：闭集都是可测集。

证明. (a) 设$m^*(E) = 0$，根据上面的注以及$m^*(X \cap E) = 0$可知E可测。

(b) “$\leq$”由可数次可加性得到，“$\geq$”是因为对任意$N \in \mathbb{N}$都有

$$m^*\left(\bigcup_{n=1}^{\infty} E_n\right) \geq m^*(E_1 \cup \cdots \cup E_N) = \sum_{n=1}^{N} m(E_n).$$

(c) 只需证明$m^*(A\cup B)\geq m^*(A)+m^*(B)$，可以不妨设$m^*(A\cup B)<\infty$。对任意$\varepsilon>0$，设开区间$I_n\,(n\in\mathbb{N})$构成$A\cup B$的覆盖，使得$\sum_{n=1}^{\infty}|I_n|<m^*(A\cup B)+\varepsilon$。记$d=\operatorname{dist}(A,B)>0$，易见可以找到可数个开区间$I_{ni}\,(i\in\mathbb{N})$覆盖$I_n$，且满足$|I_{ni}|<d$和$\sum_{i=1}^{\infty}|I_{ni}|<|I_n|+2^{-n}\varepsilon$。于是全体$I_{ni}\,(n,i\in\mathbb{N})$也构成$A\cup B$的开覆盖，且每个$I_{ni}$至多与$A,B$之一相交，可以将其分成$A$的覆盖与$B$的覆盖，故

$$m^*(A\cup B)+2\varepsilon\geq\sum_{n=1}^{\infty}\sum_{i=1}^{\infty}|I_{ni}|\geq m^*(A)+m^*(B).$$

结合ε的任意性即可。

(d) 设F是闭集，只需证明当$m^*(X)<\infty$时$m^*(X)\geq m^*(X\cap F)+m^*(X\setminus F)$。记

$$F_n:=\left\{x\in\mathbb{R}\mid\operatorname{dist}(x,F)\leq\frac{1}{n}\right\},\quad n\in\mathbb{N},$$

其中$\operatorname{dist}(x,F):=\inf\{|x-t|\mid t\in F\}$。因为$F\subset F_n$且$\operatorname{dist}(F,F_n^C)\geq\frac{1}{n}$，由(c)可知

$$m^*(X\cap F)+m^*(X\setminus F_n)=m^*((X\cap F)\cup(X\setminus F_n))\leq m^*(X).$$

只需证明$\lim_{n\to\infty}m^*(X\setminus F_n)=m^*(X\setminus F)$，根据次可加性又只需证明

$$\lim_{n\to\infty}m^*(X\cap(F_n\setminus F))=0.$$

对$n\in\mathbb{N}$，记

$$E_n:=X\cap\left\{x\in\mathbb{R}\mid\frac{1}{n+1}<\operatorname{dist}(x,F)\leq\frac{1}{n}\right\}.$$

则$X\cap(F_n\setminus F)=\bigcup_{k\geq n}E_n$，由可数次可加性得

$$m^*(X\cap(F_n\setminus F))\leq\sum_{k=n}^{\infty}m^*(E_n).$$

因此只需证明$\sum_{n=1}^{\infty}m^*(E_n)<\infty$。最后，注意到当$m-n>2$时$\operatorname{dist}(E_m,E_n)>0$，对任意$N\in\mathbb{N}$，反复利用(b)得

$$\begin{aligned}\sum_{n=1}^{2N}m^*(E_n)&=m^*(E_1\cup E_3\cup\cdots\cup E_{2N-1})+m^*(E_2\cup E_4\cup\cdots\cup E_{2N})\\&\leq m^*(X)+m^*(X)=2m^*(X).\end{aligned}$$

令$N\to\infty$，结合$m^*(X)<\infty$即证。 □

注. *这里给出的(d)的证明并非最简单的，关键在于空间的分离性质。其好处是可以推广到一般的度量空间上，比如处理$\mathbb{R}^n$上的Lebesgue测度以及度量空间上的Hausdorff测度。*

下面说明可测集经过合理的运算仍然是可测集，为此引进σ-代数的概念。

定义 10.5 (σ-代数)**.** 设$\mathcal{A} \subset 2^{\mathbb{R}}$是一族集合，称$\mathcal{A}$是$\sigma$-代数，如果满足：

(a) $\emptyset \in A$;

(b) 若$A \in \mathcal{A}$，则$A^C \in \mathcal{A}$;

(c) 若$A_n \in \mathcal{A}\,(n \in \mathbb{N})$，则$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$。

注. 由*(b)(c)*以及*De Morgan*律可知σ-代数对可数交也封闭。名字中的“σ”指的是“可数”，如果条件*(c)*改为$\mathcal{A}$只对有限交封闭，则这样的集族称为代数。

定理 10.6 * (Carathéodory)**.** 全体*Lebesgue*可测集构成σ-代数。

证明. σ-代数的定义中条件(a)(b)是显然的，只需证明$\mathcal{M}$对可数并封闭。以引理的形式给出。
引理*1*：设E_1, E_2可测，则$E_1 \cup E_2$也可测。
引理*1*的证明：对任意$X \subset \mathbb{R}$，因为E_1, E_2可测，所以

$$\begin{aligned}
&m^*(X) = m^*(X \cap E_1) + m^*(X \setminus E_1)\\
=&m^*((X \cap E_1) \cap E_2) + m^*((X \cap E_1) \setminus E_2)) + m^*((X \setminus E_1) \cap E_2) + m^*((X \setminus E_1) \setminus E_2)\\
=&m^*(X \cap (E_1 \cap E_2)) + m^*(X \cap (E_1 \setminus E_2)) + m^*(X \cap (E_2 \setminus E_1)) + m^*(X \setminus (E_1 \cup E_2)).
\end{aligned}$$

注意到$X \cap (E_1 \cup E_2) = (X \cap (E_1 \cap E_2)) \cup (X \cap (E_1 \setminus E_2)) \cup (X \cap (E_2 \setminus E_1))$，由次可加性得

$$m^*(X) \geq m^*(X \cap (E_1 \cup E_2)) + m^*(X \setminus (E_1 \cup E_2)).$$

这就证明了$E_1 \cup E_2$可测。[写起来很复杂，画Venn图会更清晰：实际上就是E_1, E_2把X分为四份，再把除了$X \setminus (E_1 \cup E_2)$之外的三份用次可加性合起来。]
引理*2*：设$E_1 \subset E_2 \subset \cdots$都可测，则$E := \bigcup_{n\in\mathbb{N}} E_n$也可测。
引理*2*的证明：记$F_1 = E_1$和$F_n = E_n \setminus E_{n-1}\,(n \geq 2)$，因为$E_n$可测，由引理1知$F_n$可测，并且$E_n = F_1 \sqcup \cdots \sqcup F_n$，反复运用可测性的定义可知

$$m^*(X \cap E_n) = \sum_{k=1}^{n} m^*(X \cap F_k).$$

又因为$E_n \subset E$，所以

$$m^*(X) = m^*(X \cap E_n) + m^*(X \setminus E_n) \geq m^*(X \cap E_n) + m^*(X \setminus E).$$

结合以上两式并令$n \to \infty$，利用可数次可加性得

$$m^*(X) \geq \sum_{k=1}^{\infty} m^*(X \cap F_k) + m^*(X \setminus E) \geq m^*(X \cap E) + m^*(X \setminus E).$$

这就证明了E可测。

引理*3*：设$E_n\,(n\in\mathbb{N})$都可测，则$\bigcup_{n\in\mathbb{N}}E_n$也可测。于是原命题得证。

引理*3*的证明：由引理1可知$E_1\cup\cdots\cup E_n$可测，再由引理2可知$\bigcup_{n\in\mathbb{N}}E_n$可测。 □

注. *Carathéodory*定理的证明对任何非空集合X上的外测度都适用，但并不事先保证可测集构成的σ-代数有多大，比如$\{\emptyset,X\}$也构成σ-代数。

因为任何一族σ-代数的交仍然是σ-代数，所以任给一族集合$\mathcal{S}$，可以定义$\mathcal{S}$生成的σ-代数$\mathcal{A}(\mathcal{S})$为所有包含$\mathcal{S}$的σ-代数之交，则$\mathcal{A}(\mathcal{S})$是包含$\mathcal{S}$的最小σ-代数。包含全体闭集的最小σ-代数称为Borel σ-代数，记为$\mathcal{B}$，其中的集合称为Borel集。$\mathcal{B}$也可以定义为包含所有开集的最小σ-代数。至此证明了Borel集都是Lebesgue可测集。[12]

可以给出一般集合上测度的定义。特别的，Lebesgue测度的确是一个测度。

定义 10.7 (测度). 设$\mathcal{A}$是非空集合X上的σ-代数，映射$\mu\colon\mathcal{A}\to[0,\infty]$称为测度，如果：

(a) $\mu(\emptyset)=0$；

(b) 对任意两两不交的$E_n\in\mathcal{A}\,(n\in\mathbb{N})$，有$\mu\left(\bigcup_{n\in\mathbb{N}}E_n\right)=\sum_{n=1}^{\infty}\mu(E_n)$。

下一题说明测度与单调集合列极限的关系。特别的，对$\mathbb{R}$上的Lebesgue测度适用。

题 10.8. 设$\mathcal{A}$是非空集合X上的σ-代数，$\mu\colon\mathcal{A}\to[0,\infty]$是测度。证明：

(a) 若$E_n\in\mathcal{A}\,(n\in\mathbb{N})$单增，即$E_1\subset E_2\subset\cdots$，则

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\lim_{n\to\infty}\mu(E_n).$$

(b) 若$E_n\in\mathcal{A}\,(n\in\mathbb{N})$单减，即$E_1\supset E_2\supset\cdots$，且满足$\mu(E_1)<\infty$，则

$$\mu\left(\bigcap_{n=1}^{\infty}E_n\right)=\lim_{n\to\infty}\mu(E_n).$$

证明. (a) $\geq$由包含关系是显然的，只需证明$\leq$。记$F_1=E_1$和$F_n=E_n\setminus E_{n-1}\,(n\geq 2)$，则有$F_n\in\mathcal{A}$且$\bigcup_{n=1}^{\infty}E_n=\bigsqcup_{n=1}^{\infty}F_n$。由可数可加性得

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\mu\left(\bigsqcup_{n=1}^{\infty}F_n\right)=\sum_{n=1}^{\infty}\mu(F_n)=\lim_{N\to\infty}\sum_{n=1}^{N}\mu(F_n).$$

再注意到$E_N=\bigsqcup_{n=1}^{N}F_n$，结合有限可加性即证。

[12]但Lebesgue可测集未必是Borel集，原因是零测集未必Borel：任何Lebesgue可测集都可以写成Borel集并上一个零测集。

(b) 注意到$E_1 \setminus E_n \in \mathcal{A}$是单增集合列，由(a)可知

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu(E_1) - \mu\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right) = \mu(E_1) - \mu\left(\bigcup_{n=1}^{\infty} (E_1 \setminus E_n)\right)$$
$$= \mu(E_1) - \lim_{n\to\infty} \mu(E_1 \setminus E_n) = \lim_{n\to\infty} \mu(E_n).$$

注意$\mu(E_1) < \infty$的作用：运算中的集合测度都是有限的，不会出现$\infty - \infty$的问题。 □

注. *(b)中要求$\mu(E_1) < \infty$（或者减弱为存在n使得$\mu(E_n) < \infty$）是必要的，比如考虑$\mathbb{R}$上的集合$E_n = [n, \infty)$，则$m(E_n) = \infty$，但是$\bigcap_{n=1}^{\infty} E_n = \emptyset$。*

已经有三种视角界定一个集合是大还是小：(1)从势的观点看，可数集是小的，不可数集是大的；(2)从拓扑的观点看，稠密G_δ集是大的，第一纲集是小的；(3)从测度观点看，零测集是小的，正测集是大的。如果不能说明某个性质对所有点都成立，则可以退而求其次研究是否对大多数点成立。例如，此前已经遇到过“单调函数在除了可数个点之外连续”；此外，代数几何中亦有Zariski稠密反映类似的思想。

自然要关心三类小集合之间的关系。之前已经说明过$\mathbb{R}$中的可数集一定是Lebesgue零测集、一定不是稠密G_δ集，下面用Cantor集（以及类Cantor集）的例子说明零测集未必可数、第二纲集未必零测，因此这三种概念互不包含。

记$C_0 = [0,1]$，令$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$是$C_0$去掉中间比例为$\frac{1}{3}$的部分，$C_2$是$C_1$分别去掉两段区间中间比例为$\frac{1}{3}$的部分，即

$$C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

一般的，假设已经构造好C_n，由2^n段长为3^{-n}的闭区间组成，则定义C_{n+1}是C_n去掉每段中间$\frac{1}{3}$得到。最后，$C := \bigcup_{n\in\mathbb{N}} C_n$称为Cantor集。

因为$m(C_n) = \left(\frac{2}{3}\right)^n$，所以$m(C) = 0$。另一方面，$C$恰好由三进制表示中不含1的数组成[13]，从而$C$和$\{0,1\}^{\mathbb{N}}$等势；另一方面，考虑二进制可知$\{0,1\}^{\mathbb{N}}$与区间$[0,1]$等势，所以$C$和$\mathbb{R}$等势。特别的，$C$是不可数的零测集。[14]

因为C_n的每段区间长度是3^{-n}，所以C不包含任何区间；又因为C_n都是闭集，所以C是无处稠密的闭集。如果修改C_n的定义，将每一步去掉$\frac{1}{3}$变成每一步去掉比例为$\lambda_n \in (0,1)$的部分，则得到的集合$\tilde{C}$称为类Cantor集（或胖Cantor集），其测度为

$$m(\tilde{C}) = \prod_{n=1}^{\infty} (1 - \lambda_n).$$

[13]容易处理无限循环小数带来的问题，比如三进制下$0.1 = 0.0222\cdots$，这里不作详细说明。

[14]另一种证明C不可数的方法是借助Baire纲定理：先前证明过每个点都是聚点的完备度量空间一定不可数。但这并不能证明C与$\mathbb{R}$等势。是否存在$\mathbb{R}$的不可数子集严格比$\mathbb{R}$势更小就是所谓“连续统假设”，在ZFC公理体系下是不可判别真假的。

如果$\sum_{n=1}^{\infty}\lambda_n < \infty$，则$m(\tilde{C}) > 0$。和Cantor集的情况一样，类Cantor集$\tilde{C}$也不含内点，从而是无处稠密的闭集。这就给出无处稠密集具有正测度的例子。

Vitali覆盖

为了使用Lebesgue测度的可数可加性，需要保证集合两两不交，而Vitali覆盖定理就是具有这一功效的技术手段之一。这一小节要求I_α是闭区间，仅仅是为了说话方便，本质上不影响命题的正确性。下一题是简单的热身，其证明可以推广到一般的度量空间。

题 10.9 *. *设$I_1,\dots,I_n$是有限个闭区间，记$\tilde{I}_i$是与I_i具有相同的中心、且长度为3倍的闭区间。证明：存在$\{i_1,\dots,i_k\}\subset\{1,\dots,n\}$满足$I_{i_1},\dots,I_{i_k}$两两不交，且*

$$I_1\cup\cdots\cup I_n\subset\tilde{I}_{i_1}\cup\cdots\cup\tilde{I}_{i_k}.$$

证明. 取I_{i_1}是$I_1,\dots,I_n$中长度最大者（若不只一个，则任取一个）。假设$I_{i_1},\dots,I_{i_k}$已经确定，如果$I_1,\dots,I_n$中还剩下与$I_{i_1},\dots,I_{i_k}$都不交的区间，则取其中长度最大者为I_{k+1}。因为只有有限个区间，所以有限步之后操作一定停止。

当操作停止时，$I_1,\dots,I_n$都与$I_{i_1},\dots,I_{i_k}$之一相交，断言此时对$p=1,\dots,n$都有$I_p\subset\tilde{I}_{i_1}\cup\cdots\cup\tilde{I}_{i_k}$。不妨设$I_p$不在$I_{i_1},\dots,I_{i_k}$中，考虑最小的$j$使得$I_{i_j}\cap I_p\neq\emptyset$，则根据操作的规则必然$|I_{i_j}|\geq|I_p|$，由此可知$I_p\subset\tilde{I}_{i_j}$，结论得证。 □

定义 10.10 (Vitali覆盖)**.** *设$X\subset\mathbb{R}$，称一族闭区间$I_\alpha\,(\alpha\in A)$构成X的Vitali覆盖，如果$X\subset\bigcup_\alpha I_\alpha$，并且对任意$x\in X$和$\varepsilon>0$，存在$\alpha\in A$使得$|I_\alpha|<\varepsilon$且$x\in I_\alpha$。*

定理 10.11 * (Vitali覆盖定理)**.** *设$X\subset\mathbb{R}$且$m^*(X)<\infty$，如果$I_\alpha\,(\alpha\in A)$是X的Vitali覆盖，则对任意$\varepsilon>0$，存在其中有限个两两不交的$I_1,\dots,I_n$，使得$m^*(X\setminus(I_1\cup\cdots\cup I_n))<\varepsilon$。*

证明. 因为$m^*(X)<\infty$，存在开集G满足$X\subset G$且$m^*(G)<\infty$。只考虑那些包含在G中的I_α，仍然构成X的Vitali覆盖，因此不妨设$I_\alpha\subset G$。任取其中一个闭区间I_1，假设已经取好了有限个两两不交的闭区间$I_1,\dots,I_k$，如果$X\subset I_1\cup\cdots\cup I_k$，则结论已经成立。因此不妨设$I_1,\dots,I_k$没有覆盖$X$，考虑所有与这些$I_1,\dots,I_k$都不交的$I_\alpha$。由Vitali覆盖的定义以及$I_1\cup\cdots\cup I_k$是闭集可知，这样的$I_\alpha$一定存在，记这样的$I_\alpha$的长度上确界为$\delta_k$。因为$I_\alpha\subset G$且$m^*(G)<\infty$，所以$0<\delta_k<\infty$。于是能够取出$I_{k+1}$与$I_1,\dots,I_k$都不交，且$|I_{k+1}|\geq\frac{\delta_k}{2}$。

对于这一列两两不交的$I_1,I_2,\dots$，有

$$\sum_{k=1}^{\infty}|I_k|\leq m^*(G)<\infty.$$

由收敛性可知$\delta_k\to 0$，且存在$n\in\mathbb{N}$使得$\sum_{k=n+1}^{\infty}|I_k|<\frac{\varepsilon}{5}$。记$\tilde{I}_k$是与$I_k$具有相同的中心、且长

度为5倍的闭区间，断言

$$X \setminus (I_1 \cup \cdots \cup I_n) \subset \bigcup_{k=n+1}^{\infty} \tilde{I}_k.$$

事实上，若$x \in X \setminus (I_1 \cup \cdots \cup I_n)$，则存在$I_\alpha\,(\alpha \in A)$包含$x$且与$I_1, \ldots, I_n$不交。因为$|I_\alpha|$是定值且$\delta_k \to 0$，所有存在最小的$k \geq n$使得$I_\alpha$与$I_1, \ldots, I_k$不交且与$I_{k+1}$相交。注意到按定义有$|I_\alpha| \leq \delta_k \leq 2|I_{k+1}|$，所以$I_\alpha \subset \tilde{I}_{k+1}$，从而$x \in \tilde{I}_{k+1}$。于是断言得证，故

$$m^*(X \setminus (I_1 \cup \cdots \cup I_n)) \leq \sum_{k=n+1}^{\infty} m^*(\tilde{I}_k) = 5 \sum_{k=n+1}^{\infty} m^*(I_k) < \varepsilon. \qquad \square$$

注. 因为$I_1, \ldots, I_n$是可测集，所以结论等价于$m^*(X \cap (I_1 \cup \cdots \cup I_n)) > m^*(X) - \varepsilon$。

单调函数几乎处处可导

几乎处处(almost everywhere)指的是某个性质对除了一个零测集之外的点成立。这一小节将证明单调函数几乎处处可导。由于可数个零测集的并还是零测集，只需考虑定义域是$(0,1)$的有界单增函数（将定义域和值域都限制在有界集具有技术上的好处）。

为此，定义左上、左下、右上、右下导数

$$D^- f(x) := \limsup_{h \to 0-} \frac{f(x+h) - f(x)}{h}, \quad D_- f(x) := \liminf_{h \to 0-} \frac{f(x+h) - f(x)}{h},$$

$$D^+ f(x) := \limsup_{h \to 0+} \frac{f(x+h) - f(x)}{h}, \quad D_+ f(x) := \liminf_{h \to 0+} \frac{f(x+h) - f(x)}{h}.$$

（没有通用记号，这里采用上、下标表示上、下极限，$\pm$表示从左、右侧逼近。）可导等价于$D^\pm f$和$D_\pm f$都相等且有限。根据定义自动有$D^\pm f, D_\pm f \in [0, \infty]$，以及

$$D_- f(x) \leq D^- f(x) \quad \text{和} \quad D_+ f(x) \leq D^+ f(x).$$

因此只需证明几乎处处成立$D^\pm f < \infty$，以及$D^+ f \leq D_- f$和$D^- f \leq D_+ f$。

直观上来说，证明的核心在于导数的几何意义，即函数的局部缩放倍率：$f'(x) = A$表明x附近的小区间倍放大为A倍。如果在一个正测集上$D^\pm f = \infty$，则f的像集测度应该是∞，因此一旦事先假设f有界，则几乎处处$D^\pm f < \infty$是合理的预期。

题 10.12. 设$f\colon (0,1) \to [-M, M]\,(M > 0)$单调递增。本题的目标是证明$f$几乎处处可导。

(a) 设$[a_1, b_1], \ldots, [a_k, b_k]$是$(0,1)$中两两不交的闭区间，证明：

$$\sum_{i=1}^{k} (f(b_i) - f(a_i)) \leq 2M.$$

*(b)** 证明：$D^{\pm}f(x)<0$对几乎所有$x\in(0,1)$成立。

*(c)** 设$0<p<q$，记

$$A=A_{p,q}:=\{x\in(0,1)\mid D_-f(x)<p<q<D^+f(x)\},$$

$$B=B_{p,q}:=\{x\in(0,1)\mid D_+f(x)<p<q<D^-f(x)\},$$

证明：$m^*(A)=m^*(B)=0$。

(d) 证明：f几乎处处可导。

证明. (a) 不妨设$a_1<\cdots<a_k$，则根据不交可知$b_i<a_{i+1}\,(1\le i<k)$，再结合f单增得

$$\sum_{i=1}^{k}(f(b_i)-f(a_i))=f(b_k)-f(a_1)+\sum_{i=1}^{k-1}(f(b_i)-f(a_{i+1}))\le f(b_k)-f(a_1)\le 2M.$$

(b) 记$N:=\{x\in(0,1)\mid D^+f(x)=\infty\}$，下面证明$m^*(N)=0$。任意取定$\varepsilon>0$和$C>0$。对任意$x\in N$，存在一列$\delta_{x,n}\in(0,1-x)$满足$\lim_{n\to\infty}\delta_{x,n}=0$且

$$f(x+\delta_{x,n})-f(x)\ge C\delta_{x,n}.$$

于是这些闭区间$I_{x,n}:=[x,x+\delta_{x,n}]$构成$N$的Vitali覆盖。又因为$m^*(N)\le 1$，根据Vitali覆盖定理（以及定理下方的注），存在有限个两两不交的$I_{x_1,n_1},\dots,I_{x_k,n_k}$满足

$$|I_{x_1,n_1}|+\cdots+|I_{x_k,n_k}|\ge m^*(N\cap(I_{x_1,n_1}\cup\cdots\cup I_{x_k,n_k}))\ge m^*(N)-\varepsilon.$$

因为I_{x_i,n_i}两两不交，结合(a)可知

$$m^*(N)-\varepsilon\le\sum_{i=1}^{k}\delta_{x_i,n_i}\le\frac{1}{C}\sum_{i=1}^{k}(f(x_k+\delta_{x_k,n_k})-f(x_k))\le\frac{2M}{C}.$$

先令$\varepsilon\to0$，再令$C\to\infty$，可知$m^*(N)=0$。这就说明$D^+f<\infty$几乎处处成立，同理可以证明$D^-f<\infty$几乎处处成立。

(c) 只需证$m^*(A)=0$，同理可得$m^*(B)=0$。任意取定$\varepsilon>0$，则存在开集$G\subset(0,1)$使得$A\subset G$且$m^*(G)\le m^*(A)+\varepsilon$。对任意$x\in A$，存在一列$\delta_{x,n}$满足$[x,x+\delta_{x,n}]\subset G$，$\lim_{n\to\infty}\delta_{x,n}=0$且

$$f(x)-f(x-\delta_{x,n})\le p\delta_{x,n}.$$

这些$I_{x,n}:=[x-\delta_{x,n},x]$构成$A$的Vitali覆盖，存在有限个$I_{x_1,n_1},\dots,I_{x_k,n_k}$满足

$$m^*\left(A\setminus(I_{x_1,n_1}\cup\cdots\cup I_{x_k,n_k})\right)<\varepsilon.$$

因为$I_{x,n} \subset G$，所以

$$m^*(A) + \varepsilon \geq m^*(G) \geq \sum_{i=1}^{k} \delta_{x_i,n_i} \geq \frac{1}{p} \sum_{i=1}^{k} (f(x_i) - f(x_i - \delta_{x_i,n_i})).$$

记$\tilde{A} := A \cap \bigcup_{i=1}^{k} (x_i - \delta_{x_i,n_i}, x_i)$，因为有限个点的测度是0，所以

$$m^*(\tilde{A}) = m^*(A \cap (I_{x_1,n_1} \cup \cdots \cup I_{x_k,n_k})) \geq m^*(A) - \varepsilon.$$

对任意$y \in \tilde{A}$，存在$1 \leq i \leq k$使得$y \in A \cap (x_i - \delta_{x_i,n_i}, x_i)$，于是存在一列$\eta_{y,m} \in (0, x_i - y)$满足$\lim_{m\to\infty} \eta_{y,m} = 0$且

$$f(y + \eta_{y,m}) - f(y) \geq q\eta_{y,m}.$$

这些$\tilde{I}_{y,m} := [y, y + \delta_{y,m}]$构成$\tilde{A}$的Vitali覆盖，所以存在有限个$\tilde{I}_{y_1,m_1}, \ldots, \tilde{I}_{y_l,m_l}$满足

$$m^*(\tilde{A} \setminus (\tilde{I}_{y_1,m_1} \cup \cdots \cup \tilde{I}_{y_l,m_l}) < \varepsilon.$$

于是

$$m^*(A) - 2\varepsilon \leq m^*(\tilde{A}) - \varepsilon \leq \sum_{j=1}^{l} \eta_{y_j,m_j} \leq \frac{1}{q} \sum_{j=1}^{l} (f(y_j + \eta_{y_j,m_j}) - f(y_j)).$$

最后，因为每个$\tilde{I}_{y_j,m_j}$都包含在某个I_{x_i,n_i}中，利用(a)的方法可知

$$\sum_{j=1}^{l} (f(y_j + \eta_{y_j,m_j}) - f(y_j)) \leq \sum_{i=1}^{k} (f(x_i) - f(x_i - \delta_{x_i,n_i})).$$

这就证明了

$$p(m^*(A) + \varepsilon) \geq q(m^*(A) - 2\varepsilon).$$

令$\varepsilon \to 0$，结合$p < q$可知$m^*(A) = 0$。

(d) 根据先前的说明，f不可导的点可以分解为$\{D^+f = \infty\} \cup \{D^-f = \infty\} \cup \{D^+f > D_-f\} \cup \{D^-f > D^+f\}$。由(b)可知前两个集合是零测集。再注意到

$$\{D^+f > D^-f\} = \bigcup_{p,q \in \mathbb{Q},\ p<q} \{x \in (0,1) \mid D_-f(x) < p < q < D^+f(x)\}.$$

由(c)可知右边的每一项都是零测集，而可数个零测集的并还是零测集，所以$m^*(\{D^+f > D^-f\}) = 0$。同理可知$m^*(\{D^-f > D_+f\}) = 0$。 □

最后，利用上题的结论容易证明：

定理 10.13. 设I是非空区间，$f\colon I \to \mathbb{R}$是单调函数，则f在I上几乎处处可导。

注. 结论不能加强为单调函数在除了一个可数集之外可导，经典的反例是所谓*Cantor*函数。今后还将利用*Cantor*函数给出关于定积分的若干反例，这里暂时不作讨论。

有界变差函数

这一小节讨论单调函数的推广，即有界变差函数(functions of bounded variance)。

定义 10.14 (有界变差函数)**.** 设$f\colon [a,b] \to \mathbb{R}$，定义$f$的全变差*(total variance)*为

$$V_a^b(f) := \sup_{a=x_0<\cdots<x_n=b,\ n\in\mathbb{N}} \sum_{i=1}^n |f(x_i) - f(x_{i-1})|.$$

如果$V_a^b(f) < \infty$，则称f是$[a,b]$上的有界变差函数。

显然$[a,b]$上的单调函数都是有界变差函数，并且有界变差函数构成线性空间。下一题说明有界变差函数就是两个单增函数之差。

题 10.15. *(a)* 证明：如果f是$[0,1]$上的有界变差函数，则f有界。

(b) 设$f\colon [0,1] \to \mathbb{R}$，证明：对任意$0 \le a < b < c \le 1$，都有$V_a^c(f) = V_a^b(f) + V_b^c(f)$。

(c) 证明：$f\colon [0,1] \to \mathbb{R}$有界变差当且仅当可以写成两个单调递增函数的差。

证明. (a) 对任意$x \in (0,1)$，有

$$V_0^1(f) \ge |f(x) - f(0)| + |f(1) - f(x)| \ge 2|f(x)| - |f(0)| - |f(1)|.$$

即$|f(x)| \le \frac{1}{2}(V_0^1(f) + |f(0)| + |f(1)|) < \infty$，所以$f$有界。

(b) 因为任何$[a,b]$的划分并上$[b,c]$的划分得到$[a,c]$的划分，所以$V_a^c(f) \ge V_a^b(f) + V_b^c(f)$。反过来，注意到添加任何分点会让求和变大，所以计算$V_a^c(f)$时只需考虑包含b为分点的划分，也就是$[a,b]$的划分并上$[b,c]$的划分，由此可得$V_a^c(f) \le V_a^b(F) + V_b^c(f)$。

(c) 充分性是因为单调函数都有界变差，且两个有界变差函数之和仍然有界变差。下面证明必要性，设f是$[0,1]$上的有界变差函数，记

$$F(x) := V_0^x(f) \quad 和 \quad G(x) := F(x) - f(x).$$

则显然有$f = F - G$以及F单调递增，只需证明G也单调递增。设$x < y$，则需要证明$G(x) \le G(y)$，即$f(y) - f(x) \le V_x^y(f)$，而这根据定义是显然的。 $\square$

注. *(c)*更美观的分解是（好处是两项都非负）

$$f(x) = \frac{V_0^x(f) + f(x)}{2} - \frac{V_0^x(f) - f(x)}{2}.$$

利用有界变差函数与单调函数的关系立即得到：

定理 10.16. 有界变差函数在除了至多可数个点之外连续，且几乎处处可导。

如果$f\colon [0,1] \to \mathbb{R}$是Lipschitz连续函数，即$|f(x) - f(y)| \le L|x-y|$，则$V_0^1(f) \le L$。因此Lipschitz连续函数都是有界变差函数，从而几乎处处可导。特别的，凸函数在定义域内部局部Lipschitz连续，因此凸函数几乎处处可导。尽管这一结论比先前证明过的凸函数在至多可数个点之外可导要弱，但今后会看到，对于$\mathbb{R}^n$中的凸函数来讲这一结论不能改进。

进一步，凸函数的导数是单增的，因此可以猜测凸函数几乎处处二阶可导。但仔细思考会发现，既然凸函数的导数未必在一个区间上有定义，自然就无法谈论二阶可导性。但下一题表明将二阶可导性减弱为二阶Taylor公式成立是可行的。

题 10.17 *. 设$f\colon \mathbb{R} \to \mathbb{R}$是凸函数，证明：对几乎所有$x_0 \in \mathbb{R}$，存在常数$a, b, c$（依赖$x_0$）使得当$x \to x_0$时

$$f(x) = a + b(x - x_0) + c(x - x_0)^2 + o(|x - x_0|^2).$$

证明. 记f'_+的不可微点集为N，则$m(N) = 0$。对任意$x_0 \in N$，当$x \to x_0$时

$$f'_+(x) = f'_+(x_0) + (f'_+)'(x_0)(x - x_0) + \varphi(x - x_0)(x - x_0).$$

其中$\varphi\colon [0,\infty) \to \mathbb{R}$满足$\lim_{t\to 0}\varphi(t) = 0$。今后会证明Newton–Lebniz公式对右导数成立，于是

$$\begin{aligned} f(x) - f(x_0) &= \int_{x_0}^x f'_+(t)\,dt = \int_{x_0}^x (f'_+(x_0) + (f'_+)'(x_0)(t - x_0) + \varphi(t - x_0)(t - x_0))\,dt \\ &= f'_+(x - x_0) + \frac{(f'_+)'}{2}(x - x_0)^2 + \int_{x_0}^x \varphi(t - x_0)(t - x_0)\,dt. \end{aligned}$$

最后，注意到$t - x_0$在积分区间上不变号，所以（本质上是积分第一中值定理）

$$\left|\int_{x_0}^x \varphi(t - x_0)(t - x_0)\,dt\right| \le \sup_{|s| \le |x - x_0|} |\varphi(s)| \cdot \left|\int_{x_0}^x (t - x_0)\,dt\right| = o(1)\cdot(x - x_0)^2 = o(|x - x_0|^2).$$

因此结论对任意$x \notin N$以及$a = f(x_0)$，$b = f'_+(x_0)$，$c = (f'_+)'(x_0)/2$成立。 □

Lipschitz连续函数、凸函数的概念容易直接推广到$\mathbb{R}^n$上的函数，而单调性则不然。今后将会证明，如果$f\colon \mathbb{R}^n \to \mathbb{R}$是Lipschitz连续函数，则$f$关于$\mathbb{R}^n$上的Lebsgue测度几乎处处可微；这就是所谓Rademacher定理。[15]特别的，还可以证明凸函数依然局部Lipschitz连续，所以$\mathbb{R}^n$上的凸函数也几乎处处可微。进一步，凸函数依然几乎处处有二阶Taylor展开式，这被称为Alexandrov定理。这些性质通向更深入的学科：凸分析以及几何测度论。

[15]关于Lipschitz连续函数的更多性质可见香蕉空间《Lipscthiz连续映射》。