微积分：数学家、计算机科学家和物理学家的微积分

《抽象数学导论》安德鲁·D·黄

内容

图目录

序言

微积分是古希腊传承给我们的重要知识传统的一部分。在接下来的几个世纪中，微积分得到了精炼，并被用于解决物理学中一系列真正令人难以置信的问题，并且在计算机科学、人口生物学和经济学等众多领域得到越来越广泛的应用。在学校中教授的“微积分”几乎完全指的是用于研究变化率的数学研究中的计算程序——一元函数的微分和积分。然而，微积分有一个美丽但鲜为人知的理论基础，它使我们能够一致且有意义地谈论无限大和无限小。这个基础是所有严肃的数学学生必须掌握的知识，它体现了数学中理论和应用的两个方面。

这本书的目标多种多样，但都源于作者希望尽可能直观、非正式地呈现美丽、有趣、生动的数学，同时不妥协逻辑严谨性。当然，你将巩固你的计算知识，因为在大多数应用中，这是最重要的技能。其次，你将获得对微积分理论基础的了解，基本上是从第一性原理开始的。如果你喜欢思考无限的概念，这本书的这一部分应该会吸引你。第三个主要目标是教你一些数学本身的本性和哲学。这一部分的展示旨在具有普遍吸引力，而不仅仅是针对那些打算在大学阶段主修数学的学生，而是针对对数学感兴趣的公众。微积分非常适合这个元课程，因为它的理论基础建立在无限的概念上，如果不小心发展，可能会导致明显的逻辑悖论。打个比方，仅仅通过直觉来理解微积分的逻辑性质，就像在多云的天气中降落飞机；如果你有人帮助，可能会成功。

可信的告诉你该做什么（或者如果一个作者陈述了没有证明的定理，甚至没有给出仔细的定义），但你获得的飞行技能很少可以在其他机场（或在其他数学情境中）应用。仔细的定义、逻辑和证明是雷达，让你能穿透直观的迷雾，解决（或避免）明显的矛盾，并理解什么是真实的以及为什么它是真实的。当你掌握了证明定理所需的思想组织时，定理就成为了你的一部分，任何人都不可能夺走或否认。相比之下，通过记忆获得的事实知识是浅薄而脆弱的：如果你被告知你学到的东西是错误的，你无法找到真相。继续使用驾驶比喻，如果你只知道在教练告诉你该做什么时如何飞行，你的技能通常不适用。雷达提供给你的信息（定理的陈述）很重要，但更重要的是你能够自己使用雷达，尤其是在新情境中。

第四和最终的大目标是展示重要的数学成果。尽管有普遍的看法，但数学是一门创造性的学科，其从业者常将其比作古典音乐。正如没有人会把音阶误认为是真正的音乐一样，也没有人会混淆拼写和语法与文学，任何数学家都不会把常规的书籍问题等同于真正的数学。当然，常规问题是重要的教学工具：它们让你有机会练习计算技巧，直到你熟练掌握。但除非你精通计算，否则你不太可能取得数学上的成功，你当然也不能仅仅通过大量做常规问题就能保证成功。如果你的数学经验仅限于学校课程，你可能在数学的本质上有很多东西需要重新学习。你最初遇到的材料可能看起来在质量上很陌生，但渐渐地，你的观点将从技术转向数学本质的基本概念和逻辑关系。在这个过程中，你还会遇到诸如

$$e^{i\pi} + 1 = 0 \qquad \text{and} \qquad \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6},$$

并且精确地陈述了包含你在学校学到的知识的断言，但用词经过仔细定义，逻辑结构清晰明了。

为了做一个最后的类比，把数学想象成地形。有些地方平坦且长满草地，适合愉快的散步；其他地方则是丘陵地带，但有磨损良好的小径。有些悬崖需要攀爬狭窄而困难的路径，并需要绳索，还有峡谷，在那里很容易看到如何穿越，但穿过灌木丛却很困难。这里和那里，山峰高耸于平原之上。从山顶的景色令人惊叹，让你可以看到广阔的视野及其联系，但只有通过漫长而艰难的攀登才能达到。尽管如此，小径已经被开辟，坚持不懈的工作将带你到达山顶。这本书是你的导游，向你展示“实分析”部分地图上的有趣小径和地理特征。你学到的技能会积累起来，让你能够应对更艰难的小径。沿途，有泉水、瀑布和野花作为稳步前进的奖励。行程中的所有地方都经过了详细的地图绘制，但有时我们会经过一个尚未被攀登的悬崖，而且几次你会瞥见广阔的、尚未绘制的后院，那里正在进行当前的研究。

第一章

数学的语言

> 数学家就像一种法国人；和他们说话时，他们会将其翻译成自己的语言，很快就会变得完全不同。
>
> 数学家就像法国人。无论你对你说什么，他们都会将其翻译成自己的语言，然后它就变成了完全不同的事情。——歌德

1.1 数学本质

数学是人类智力追求中独一无二的；它既不是艺术、哲学或科学，但与每一门学科都有某些共同点。作为一种语言，数学在以惊人的准确性表达物理世界特征方面是无与伦比的。同时，数学与现实世界没有内在的联系；数学的对象是概念，它们不以星星、分子或人们那样的方式存在于物理世界中。相反，星星、分子和人们不是数学对象，尽管它们具有（例如，高度、质量或温度）可以被数学概念所描述的属性。

数学语言能够用来描述和预测自然现象，这一点引人注目，但这一事实本身并非数学的一部分。数学概念似乎独立存在，在（比喻的）柏拉图式的宇宙中，独立于物理世界之外，等待被发现而不是被创造。数学在一定程度上受到美学的影响，尽管没有关于构成美丽数学的硕士课程，但数学家们一致同意

显著的广度在于判断数学的部分是否美丽、深刻（意味着“与主题广泛不同的部分相关联”）或正确（意味着“逻辑上一致”而不是“真实”）。

电子计算机为抽象和递归提供了良好的物理类比，这两者是本书的基本主题。简而言之，抽象是描述一个对象或结构所具有的属性的过程。递归是在更简单的对象的基础上定义结构的行为。抽象与实现相对立，实现是一种特定的构建或实现某物的特定方式。

例如，计算机将数据存储和操作为位或二进制数字的模式，通常称为0和1。位是数据的一种抽象，数据的抽象属性是那些与位的物理表示无关的属性。在许多理论考虑中，只有位模式、数据的抽象结构是相关的。

计算机科学的递归性质体现在这样一个事实：计算机程序不仅仅是一长串比特，这些比特首先被组织成“字节”（通常是8比特），然后是“字”（4或8字节），然后是汇编语言指令（对处理器有意义的词组），函数（执行程序员指定的某些操作的汇编指令组），以及库（相关函数的集合）。计算机程序——您的网页浏览器、文本编辑器或媒体播放器——是通过“链接”库中的函数来构建的。

数据存储可以以多种方式在物理上实现：

- 磁性区域——软盘和ZIP盘，PC硬盘。
- 亮暗斑或条纹——光盘和UPC符号。
- 孔和“无孔”的纸条——穿孔卡片和纸带。
- 充电和非充电电容器—RAM。

常见的抽象特征是一对对比状态。数学家或理论计算机科学家认为这些存储方案之间没有本质区别。根据上下文，我们可能会称这些对比状态为“黑白”、“0和1”、“真和假”或“开和关”，但无论名称如何，都有一个单一的底层结构。在数学中，你应该努力理解结构，而不是记住名称。

接下来，考虑以下由位表示构建的算术/逻辑运算：

- (二进制算术)将0视为表示任意偶数，将1视为表示任意奇数。也就是说，一个整数与其被2除的余数相对应。两个奇数的和总是偶数（“$1+1=0$”），偶数和奇数的乘积总是偶数（“$0 \cdot 1 = 0$”），等等。如果我们列出加法和乘法的结果，我们得到

$+$	0	1
0	0	1
1	1	0

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

- (布尔逻辑{v*}将F视为表示任意“错误”断言（例如“$2+2=5$”）和T视为表示任意“正确”句子（例如“$1+1=2$”）。由于“$2+2=5$ 或 $1+1=2$，但不是两者都”为真，我们写成“F xor T=T”。（“xor”代表“排他或”：一个陈述为真，但不是两者都。）由于“$2+2=5$ 和 $1+1=2$”为假，我们写成“F and T=F”。下表根据每个陈述是否为真或假，给出了由两个陈述连接而成的陈述的“真值”。

xor	F	T
F	F	T
T	T	F

and	F	T
F	F	F
T	F	T

每个表格对封装了关于数据位的一些结构。真正的数学观察是，表格的条目对应：在对应偶-假和奇-真的情况下，“加法（模2）”对应于“异或”，而“乘法（模2）”对应于“与”。上面的两个表格对是同一抽象结构的不同实现，甚至可能表示为

$\vee$	$\bullet$	$\circ$
$\bullet$	$\bullet$	$\circ$
$\circ$	$\circ$	$\bullet$

$\wedge$	$\bullet$	$\circ$
$\bullet$	$\bullet$	$\bullet$
$\circ$	$\bullet$	$\circ$

我们刚刚描述了抽象结构之间的一种抽象关系。最终，我们可能会在各种情境中发现这种关系，并给这个概念起一个名字，比如“同构”。制定这样的定义是一种递归行为：它将一类抽象结构之间的抽象关系组合在一起。随着时间的推移，我们可能会发现研究同构本身是有益的，从而进入更高层次的抽象。被称为范畴论的数学分支处理这类问题。

上述示例旨在不是让你因复杂性而气馁，而是为了说明数学关注概念系统的抽象属性，以及结构可以在去除其无关细节后进行组织和理解。

数学的对象

数学的基本对象是集合及其构成元素。集合是对象集合的非正式概念的抽象。1976年美国各州名称的集合有50个元素；“马萨诸塞州”是这个集合的一个元素，而“波多黎各”和“铀”则不是。作为一个更数学化的例子，素数集合，即大于1且除了1和自身没有其他因数的整数集合p和p，是一个集合。数字2、5和$2^{13466917}-1$是素数集合的元素，而4和$2^{13466917}$则不是。

在数学中，我们不讨论“集合究竟是什么”；这个问题是哲学性的，而不是数学性的。数学在很大程度上是建立在集合论之上的，就像计算机科学是建立在位串之上一样。在计算机科学中，研究对象是递归构建的；最终，所有东西都是用布尔运算和位词来定义的。此外，计算机程序除了物理存储和访问位的方式之外，还具有抽象的存在。同样，微积分的对象和结构——整数、实数、函数等等——都是用更简单的对象递归定义的，并且最终是由集合构建的。每个数学断言都可以解释为关于集合的断言，尽管即使是“$2+2=4$”用集合纯语言表达也异常困难。集合的本质是一个无关紧要的“实现细节”；相反，集合的性质至关重要。

数学和科学

数学曾被称为“最精确的科学”。在过去的两个世纪里，它已经相当清楚，数学在本质上根本不是一门科学。在数学中，一个想法被接受的标准是逻辑、演绎证明，关于这一点我们将在下面进行更多讨论。虽然数学家有时会进行“实验”，无论是用铅笔和纸，还是用电子计算机，但数学实验的结果从未被视为定论。相比之下，在物理学或化学中，实验是检验一个想法有效性的唯一标准。[1]

几分钟的思考应该会揭示这些截然不同的标准的原因。数学概念除了我们赋予它们的属性外没有其他相关属性，因此从原则上讲，数学家可以完全访问一个对象所拥有的所有属性。然而，在物理学中，研究的对象是现象，关于这些现象的信息只能通过实验获得。无论进行多少实验，科学家们永远不能确定他们的知识是完整的；一个更精细的实验可能与现有结果相冲突，这表明（至多）需要修改一个被接受的自然科学定律，或者（最坏的情况）有人收集数据时不够谨慎。实验结果永远不是数学上精确的，而是受“不确定性”或“实验误差”的影响。因此，在科学中，我们对我们研究对象的访问并不像在数学中那样。自然定律——现实某个方面的数学模型——几乎可以肯定只是近似的。

尽管目标和接受标准不同，数学和物理学在很大程度上相互丰富。最明显的影响方向是从数学到科学：自然现象的最佳描述是数学的，并且惊人地准确。例如，日全食可以提前数百年预测，甚至可以预测到全食发生的时间和地点。虽然不那么明显但同样重要的是物理学、生物学和经济学对数学的积极影响，尤其是在20世纪。无论什么原因，描述自然现象的数学与科学有着深刻的联系，充满了美丽而意外的结果。没有科学的指导影响，数学往往会变得僵化、专业化，仅仅局限于技术。

[1]This characterization of science is due to the physicist, R. P. Feynmann.

数学确定性

本章非正式，其目的是与您所了解的材料建立联系，而不是奠定数学的正式基础。尽管如此，确实存在形式基础，即集合论公理集的形式。通常的公理被称为ZFC，即泽梅洛-弗兰克尔和选择公理。

在数学中，没有绝对真理的概念，只有与公理系统的一致性，关于这一点我们将在下面进行更多说明。数学家们坚信ZFC是一致的。大多数数学家在ZFC框架内工作，并在ZFC中证明的结果被称为（口语上）是"真实的"。然而，重要的是要记住，数学实际上并不产生"客观真理"，而是确立陈述与ZFC的一致性。这种区别是基本的，并且经常被误解。说"$2+2=4$是一个普遍的、数学上的真理"是错误的；更准确的说法是（法律上的，但非平凡的）主张，"如果2、4、$+$和$=$的概念被适当地定义，以符合我们对计数的直觉，那么$2+2=4$。"数学真理，可以说是最确定的一种真理，总是相对于一个公理系统。可证性，而不是真理本身，是数学的中心关注点。

1.2 集合与运算

集合通常用大写字母表示，元素用小写字母表示。我们写 $x \in X$ 来表示 X 是一个集合，x 是 X 的一个元素。同样，我们写 $y \notin X$ 来表示 y 不是 X 的一个元素。

如果 Y 是一个具有性质 Y 的每个元素都是 X 的元素的集合，那么我们称 Y 是 X 的子集，并写作 $Y \subset X$。原始的十三个殖民地的名称是州名称集合的子集。素数集合是正整数集合的子集。偶数集合不是素数集合的子集。

为了避免逻辑矛盾（例如罗素悖论，见练习1.5），有必要固定一个宇宙，一个具有以下性质的集合$\mathcal{U}$：对于考虑的每个集合X，都有$X \subset \mathcal{U}$。在这本书中，宇宙通常被认为是$\mathbf{R}$，实数集，或者是定义域和值域为$\mathbf{R}$的"函数"集。（函数在第3章中定义。）

集合可以用维恩图来表示。宇宙被描绘为一个矩形，所考虑的集合是曲线的内侧，图1.1。

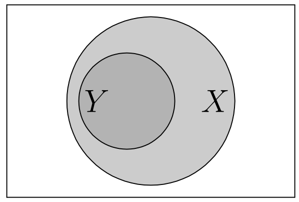


图1.1：关系$Y \subset X$的维恩图。

两个集合 X 和 Y 在且仅在 $X \subset Y$ 和 $Y \subset X$ 相等时相等，即每个集合具有相同的元素。具有有限个元素的集合可以表示为列表，如 $\{0, 1\}$。一个元素被列出多次无关紧要；集合 $\{0, 1\}$ 和 $\{0, 1, 1\}$ 是相等的。“集合构造表示法”通过指定宇宙以及描述元素的属性来描述一个集合。例如，$\mathbf{R}^+ := \{x \in \mathbf{R} \mid x > 0\}$ (读作“在 $\mathbf{R}$ 中大于 0 的 x 的集合”) 表示正实数集合。等号旁边的冒号表示等式的一侧正在被定义。

存在一个唯一的空集 $\varnothing$，它不包含任何元素。如果你对此感到难以置信，你会同情中世纪的哲学家们，他们对“零”的概念感到犹豫。然而，决定性因素是空集是有用的，并且它的存在与集合论公理的逻辑一致性相符合。

一个集合 x 的元素应与单元素集合 X（其唯一元素为 $\{x\}$）仔细区分。例如，关系 $x \in \{x\}$ 总是成立，而 $x \in x$ 很少成立。同样，元素和子集不应混淆，尽管它们在表面上似乎是相关概念：对于每个集合 X，我们有 $\varnothing \subset X$，但通常有 $\varnothing \notin X$。

一个集合可以用多种方式指定；例如，$\mathbf{R}^+$也可以表示为 $\{y \in \mathbf{R} \mid y$具有实数对数$\}$，或者表示为$\{x \in \mathbf{R} \mid x \neq 0$ 和 $x = \sqrt{x^2}\}$，而由两个元素组成的集合 $\{0, 1\}$ 可以写成 $\{x \in \mathbf{R} \mid x^2 = x\}$，例如。

空集的规格通常很有趣：

$$\varnothing = \{x \in \mathbf{R} \mid x \neq x\} = \{x \in \mathbf{R} \mid x^2 = -1\} = \{x \in \mathbf{Q} \mid x^2 = 2\}.$$

最后一种特征化将在下面讨论；“**Q**”是有理数集。空集必须与（非空）单元素为空集的集合$\{\varnothing\}$区分开来。尽管表面上看起来有些荒谬，但这一点并不愚蠢，请参阅第2章。

定义宇宙为$\mathcal{U}$。如果 $X \subset \mathcal{U}$ 是一个集合，那么 X 的补集，记为 $\sim X$ 或 X^c，是由 $X^c = \{y \in \mathcal{U} \mid y \notin X\}$ 定义的集合。非正式地说，X 的补集是“不在 X 中的对象的集合”。

存在从现有集合中形成新集合的操作。以下非正式描述了其中四个最重要的操作。如果 X 和 Y 是集合，那么我们可以形成它们的

- 并集 $X \cup Y$，由属于 X 或 Y 或两者都有的元素组成。（除非特别说明，否则数学家在非排他意义上使用“或”这个词。）

- 交集 $X \cap Y$，由同时存在于 X 和 Y 中的元素组成。如果集合 X 和 Y 的交集是空集，即它们没有共同元素，则这两个集合是互斥的。

- 笛卡尔积 {v*}，由所有有序对 (x, y) 组成，包含 $x \in X$ 和 $y \in Y$。如果 $X = \{A,\ B,\ \dots, H\}$ 和 $Y = \{1, 2,\ \dots, 8\}$，则 $X \times Y$ 是一个64元素集

  $$\{(A, 1),\ (A, 2),\ \dots,\ (A, 8),\ (B, 1),\ \dots,\ (H, 8)\},$$

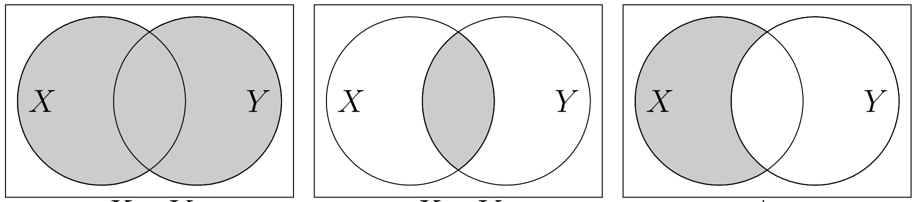
  当 X 和 Y 是实数区间时，它们的笛卡尔积是平面上的一个矩形。特别是，可以将平面视为两条直线的笛卡尔积。

- 差异 $X \setminus Y$，由所有不属于 Y 的 X 的元素组成。如果全集固定，那么 $X \setminus Y = X \cap (\sim Y)$。

并集和交集是布尔运算（分别是“或”和“与”），而笛卡尔积则从列表中创建表格。并集、交集和差的维恩图如下：

无限多个集合的并集和交集被定义为预期的：如果$\{X_\alpha\}_{\alpha \in I}$是某个宇宙$\mathcal{U}$ (中的一些集合的族，并且I是一个“指标集”)，那么

$$\bigcup_{\alpha \in I} X_\alpha = \{x \in \mathcal{U} \mid x \in X_\alpha \text{ for } \textit{some } \alpha \in I\},$$
$$\bigcap_{\alpha \in I} X_\alpha = \{x \in \mathcal{U} \mid x \in X_\alpha \text{ for } \textit{all } \alpha \in I\}.$$



$X \cup Y$ $X \cap Y$ $X \setminus Y$

图1.2：简单集合运算的文氏图。

在这本书中，无限指标集通常是正整数的集合，如{v*}

$$X = \bigcap_{n=1}^{\infty} (-\tfrac{1}{n}, \tfrac{1}{n}).$$

一个实数 x 是 X 的元素当且仅当对于每个正整数 n，都有 $-1/n < x < 1/n$。我们将看到，0 是唯一具有此性质的实数，因此 $X = \{0\}$。在练习 2.10 中给出了无限集操作的更多示例。

1.3 逻辑

如上所述，数学家并非真正找到“客观真理”，而是从被称为公设的假设开始，通过逻辑推理得出结论。在本节中，我们开始研究逻辑推理，强调与普通英语的语言差异。

数学家使用普遍认可的公理（基础假设）和逻辑推理规则。在这本书中，公理是集合论中的策梅洛-弗兰克尔-选择公理，推理规则是亚里士多德逻辑的规则。

在整个本节中，通过一系列的示例来展示所介绍的概念。数学追求普遍性，但人的思维往往偏爱特殊性，正是这些特殊性使数学直接有用。数学的目标是精确思考，而不是为了普遍性而普遍性。话虽如此，抽象（促进普遍性）有一个明确的目的：提取问题的本质特征并忽略无关细节。精确性很重要，因为直觉（尤其是关于无限的概念）往往具有误导性，有时甚至明显错误。逻辑推理是

数学的“卫生学”（继H. Weyl之后），检查直觉逻辑一致性的原则工具，以及避免错误思考。

语句和含义

一个陈述是一个具有真值的句子，即它对于某个公理系统来说是不含糊地要么为真要么为假。一个数学句子可能依赖于变量，因此可以总结一个陈述族，每个变量可能的赋值对应一个陈述。结果陈述的真值可能依赖于变量的值。重要的是，对于每个特定的变量选择，句子应该是真或假的。通过使用变量，定理通常封装了无限多个陈述。

这里有一些例子，其中变量 n 是一个整数。观察依赖于 n 的语句封装了无限多个语句。

- “-4 是一个偶数整数。” “π 的十进制展开是非重复的，并包含字符串 ‘999999’。” （正确）
- 对于每个整数 n，$n^2 - n$ 是一个偶数整数。（正确）
- “$2 + 2 = 5$.”（假）
- 对于某些整数 n，n 和 $n + 1$ 都是偶数整数。（错误）

句子不是陈述的包括“n是一个偶数整数”（其真值取决于n）和自指示例，“这个句子是真的”（其真值必须作为一个公理来指定）和“这个句子是假的”（无法一致地分配一个真值）。

语句通过逻辑蕴涵或if-then语句相互连接，例如，“如果 H，那么 C。”变量 H 是一个语句，称为蕴涵的假设，而变量 C 是一个语句，称为蕴涵的结论。我们认为 C 是从 H 推导或派生出来的。

亚里士多德逻辑的基本思想是，除非从一个真命题推导出假命题，否则一个蕴涵是有效的:

- 如果 $1 \neq 0$，那么 $1^2 \neq 0$。（有效）

- 如果 $1 \neq 0$，那么 $1^2 = 0$。(无效)
- 如果 $1 = 0$，那么 $1^2 \neq 0$。（有效）
- 如果 $1 = 0$，那么 $1^2 = 0$。(有效)

如果一个假设和结论通过有效的蕴涵相关联，那么就称这个假设蕴涵了结论。在这个观点下，从一个错误的假设中得出结论是有效的（不是逻辑上错误的）：如果我们从真理开始并做出有效的推理，我们只能得到真理，而不是谬误。具有错误假设的蕴涵被称为空洞的。为了强调，有效性具有可能反直觉的性质，即如果假设是错误的，那么每个结论都通过有效的蕴涵而成立。尽管这个惯例看起来很奇怪，但它并不允许我们从真理中推断出谬误。逻辑有效性是“证明”概念的核心，因此对于本书的其余部分（以及数学的一般而言）至关重要。

“imply”这个术语在逻辑学中的含义与在日常英语中非常不同。在英语中，“imply”意味着“暗示”或“建议”或“暗示”。在数学中，如果一个假设推导出结论，那么只要假设是真的，结论的真实性就是一个坚不可摧的确定性。术语“valid”也有一个精确的含义，这与日常英语中的含义并不完全相同。最后，请注意，每个陈述都有一个真值，但只有如果-那么陈述可以是有效的或无效的。

有趣的逻辑蕴涵通常取决于变量，因此蕴涵的真值取决于假设和结论的真值。当蕴涵取决于变量时，逻辑有效性的概念就变得尤为重要。以下例子说明了假设和结论中真值和假值的各种组合：

- 如果 n 是一个整数，并且如果 n 是偶数，那么 $n/$^2 是一个整数。（有效）
- 如果 n 是一个整数，那么 $n/2$ 是一个整数。（无效）
- 如果 n 是一个整数且 $n = n + 1$，那么 $2 + 2 = 5$。（有效）
- 如果 n 是一个整数且 $n = n + 1$，那么 $2 + 2 = 4$。(有效)

真理（适用于陈述）与有效性（适用于逻辑蕴涵）之间的区别可能一开始看起来有点繁琐。然而，重要的是要意识到这两个概念是不同的，尽管它们并非完全无关；当一个逻辑

蕴含具有明确的假设和结论，整个句子成为一个陈述，可能是真或假。蕴含的有效性表达了假设和结论的真值对结果陈述的真理。从这个意义上说，逻辑有效性是一种“元真理”。一个有效的蕴含通常会产生无限多个真命题。

逻辑一致性

如果在一个公理化系统中，可以证明某个命题 P 并且也可以证明其否定 $\neg P$，那么每个陈述 Q 都是可证明的，因为要么 $P \implies Q$ 要么 $\neg P \implies Q$ 是空洞地真实的。成对 $\{P, \neg P\}$ 被称为逻辑矛盾，如果一个公理系统中可以推导出矛盾，则称该系统是不一致的。虽然发现一个公理系统是不一致的可能是非常有趣的，但一个不一致的系统本身并不具有数学上的兴趣。

哥德尔在20世纪30年代的工作表明，除了使用某些其他（“更强大”）的公理系统之外，无法证明ZFC的一致性，而这些公理系统的相容性是未知的。为了安心，还知道如果ZFC中存在矛盾，那么普通算术中也存在矛盾。

定义和定理

数学定义建立术语，这是工作的共同基础。在制定“良好”定义的主要困难在于隔离或抽象出所需的概念属性。

数学定义被逐字解释。对于初学者来说，在解释定义时没有阅读比所陈述的更多内容，可能会成为一个严重的概念障碍。

> 一个物理学家、一个统计学家和一个数学家在苏格兰乡村驾车时，遇到了一群一百只羊，其中有一只是黑色的。物理学家说：“从这个情况我们可以推断出，有一百分之一的羊是黑色的。”统计学家说：“不，我们只知道在这100只羊中，有一只是黑色的。”数学家纠正道：“恐怕你们两个都错了。我们只知道在这100只羊中，有一只羊的一面是黑色的。”

当你被要求证明某事（在这本书或其他地方）时，首先要做的是确保你了解并理解所有涉及概念的定义。最终这将成为你的第二天性（或者你会在挫败中放弃数学），但在这个阶段经常提醒自己并没有坏处。

定义1.1 一个整数 n 是偶数，如果存在一个整数 m 使得 $n =^{2}m$。

对于每个整数 n，定义提供了一个判断 n 是否为偶数的标准。这个标准是一个及格/不及格测试，没有更多。定义还提供了一个条件，每个偶数都满足这个条件。如果一个整数 k 是偶数，你立即就会知道关于 k 的某些信息；例如，最后一位数字不能是“3”。

我们立即看出 $n = 6$ 是偶数，因为 $m = 3$ 满足定义的条件。相比之下，我们无法轻易地确定 5 是偶数还是不是；使用定义，我们必须以某种方式证明对于每个整数 m，$5 \neq 2m$，这是一项无限的任务。相反，我们必须找到非偶数的标准。例如，我们可能证明一个除以 2 余 1 的整数不是偶数。由于 5 满足这个标准，我们就会推断出 5 不是偶数。

一般来说，在解决问题时，定义扮演着测试的角色，例如“确定以下哪个整数是偶数...”，而在证明定理时，定义扮演着条件的角色，例如“如果n是一个偶数整数，那么n^2是一个偶数整数。”

在数学实践

上述困境（“5是偶数吗？”）是数学研究和学习的常态，可能会非常令人沮丧。数学标准通常比“偶数性”更为微妙，可能难以立即判断一个特定对象是否满足某个标准。不幸的是，如果你事先不知道什么是对的，你在试图证明某事时不知道如何进行！很快你就会遇到这个生存教训：在数学中，你必须不怕暂时工作，跟随死胡同，检查未知为真的假设的后果，并寻找可能不存在的例子。发现过程永远不会一帆风顺，数学也不例外。随着时间的推移，你将发展出关于可能富有成效的问题解决方法的直觉，以及可能陷入死胡同的想法。你将学会如何“玩弄”假设，如何观察特殊情况并制定

一般猜测，如何区分真实模式和幻觉模式，最后如何证明你找到的模式是真实的。

定理

一个定理是具有足够兴趣的有效推论，值得特别关注。引理是建立定理时主要具有技术兴趣的有效推论。如果你编程，将引理视为“逻辑子程序”可能有所帮助，这是一个反复使用的简短逻辑论证，为了清晰和简洁应该单独列出。命题是“一个具有独立兴趣的小定理”。在某种程度上，在特定情况下术语的选择是风格问题。

大多数数学断言都是以三种形式之一陈述的（按形式化的程度大致递减）：

- (有效的蕴含)：“如果n是一个偶数，那么n^2是4的倍数。”

- (对于每个偶数整数n，n^2是4的倍数。

- (直接陈述) “偶数的平方是4的倍数。”

每个都表达了这样一个事实：一个具有一个属性（一个偶数）的对象也具有另一个属性（它的平方是4的倍数）。一个更正式的说法，结合了蕴涵和量化，是“如果n是一个偶数，那么存在一个整数k，使得$n^2 =$是4的k倍。”

当一个有效蕴涵的每个假设都为真时，结论也必然为真。这是逻辑蕴涵所隐含或明确传达的唯一信息。特别是，如果某个假设为假，则不做出任何断言。为了强调：

> 一个定理除非每个假设都得到满足，否则绝对不传达任何信息。

一个常见的误解是记住定理的结论而忽略假设，这可能导致（至多）一个脱离语境的陈述或（最坏的情况）一个错误的解释。在20世纪90年代中期，一位流行的报纸专栏作家陷入了这个陷阱

关于A. Wiles对“费马大定理”的证明。Wiles使用了“双曲几何”中的技术，该专栏作家认为这是自相矛盾的，因为“在双曲几何中可以画圆的平方”，而该专栏作家那一代的每个学生都学过“画圆的平方是不可能的”。该专栏作家可能是在回忆19世纪数学的一个著名定理的结论：

定理1.2. 假设欧几里得几何的公理。如果给定一个单位长度的线段，那么仅使用直尺和圆规，在有限步骤内构造长度为 π 的线段是不可能的。因此，构造长度为 $\sqrt{\pi}$ 的线段，即“化圆为方”也是不可能的。

作为一个一般性的教训，定理1.2没有提及使用直尺和圆规以外的工具构造这样的线段的可能性，也没有提及使用直尺和圆规获得更好和更好的近似值的可能性，从而（在某种意义上）通过无限多步实现构造。这个故事中的相关缺陷是，除非假设欧几里得几何的公理，否则定理什么也没有说.[2]

如果从数学中汲取语言学的教训，那就是单词本身仅仅是概念的标签。虽然我们的思维对单词反应强烈，[3] 但逻辑、数学和现实的核心是潜在的概念。好的术语被选择来反映意义，但基于术语假设隐含意义是明显的，这是一个常见的、人类的错误。数学家们通过“红鲱鱼原则”提醒自己这一点：

在数学中，“红鲱鱼”可能既不是红色的，也不是鲱鱼。

定理1.2之所以引人注目，还因为另一个原因：它断言了一个在先验上可以想象的过程（即不是{v*}）的不可能性。

[2]The columnist's error was not this glaring; they argued that because theorems of hyperbolic geometry can be interpreted as statements in Euclidean geometry, a “hyperbolic” proof is self-contradictory. The resolution to this objection is that while “squaring the circle in hyperbolic geometry” can be interpreted as a statement about Euclidean geometry, the interpretation is markedly different from “squaring the circle in Euclidean geometry”, and does not contradict Theorem 1.2.

[3]To the extent that nonsensical rhetoric can be persuasive, or that it is illegal in the U.S. to broadcast certain words by radio or television.

明显矛盾）。[4] 这与说“人类知识目前没有‘化圆为方’的手段”完全是两回事。相反，这意味着欧几里得几何的公理与使用某些工具在有限步骤内构造某条线段在逻辑上是不兼容的。证明的逻辑将在后面简要介绍，在讨论了证明方法之后。

证明

为了更详细地阐述证明的思想，并以一些具有历史和数学重要性的例子来说明，考虑一个熟悉的事实：“$\sqrt{2}$ 是无理数。” 这个断言背后有一个定理，但目前的表述还有很多不足：例如，它忽略了诸如“什么是实数？”和“有理数和无理数（实数）之间的关系是什么？”等问题。一个勤奋的数学家可能更喜欢这样的断言：“2 没有有理数的平方根。” 然而，这也不是一个逻辑推论。精确地（并以一种便于证明的方式）表达 $\sqrt{2}$ 的无理性如下。

定理1.3. 如果 m 和 n 是正整数，那么 $(m/n)^2 \neq 2$.

定理1.3例证了本章开头的歌德名言，尽管经过一点练习，你将能够从非正式断言到精确陈述的逻辑蕴涵进行心算。你应该确信这个定理真的说的是“2没有有理数平方根。”另一种表述是量化句子，“对于每一个有理数x，$x^2 \neq 2$。”

定理1.3的证明通过以下观察得到加速：

引理1.4。如果k是一个偶数整数，并且存在一个整数m满足$m^2 = k$，那么k是4的倍数。

在引理1.4中，假设由两个陈述组成，“k是一个偶数”和“存在一个整数m，使得$m^2 = k$”（有时表述为“k是一个平方”）。结论是陈述“存在一个整数n，使得$k = 4n$。”正如所述，引理1.4在k不是偶数的情况下没有任何信息，也不会在k不是完全平方数的情况下提供信息。

[4]A popular cartoonist claimed that “It is impossible to prove the impossibility of something.” While this is arguably true of science, it is certainly not true of mathematics.

证明。引理1.4通过检查几个情况来建立。假设k是一个平方数，因此存在一个整数m满足$m^2 = k$。如果m是奇数，那么$k = m^2$也是奇数（为什么？），而如果m是偶数，那么$k = m^2$是4的倍数（为什么？）。因此，一个平方数要成为偶数，唯一的方式是它已经是4的倍数，正如所要证明的那样。 □

这个证明虽然从概念上来说是正确的，但确实有些不够正式；细节需要定义一个“奇数”整数，以及回答括号中的两个问题的步骤。定理1.3的直接证明可以基于引理1.4。稍后给出一个更标准的反证法证明。

证明。观察$(m/n)^2 = 2$ 与 $m^2 = 2n^2$ 表示相同的意思。通过将 m/n 写成最简形式，可以假设 m 和 n 没有公共因子；特别是，它们都不是偶数。

案例1：m是奇数。在这种情况下，m^2也是奇数。由于对于每个n，2n^2都是偶数，因此不存在一个整数n使得$m^2 =$等于$2n^2$。

案例2：n是奇数。那么$2n^2$是一个不是4的倍数的偶数整数。根据引理1.4，$2n^2$不是一个完全平方数。

这表明如果 m 和 n 是正整数，那么 $m^2 \neq 2n^2$，从而完成证明。 □

等价蕴含形式

可能最直观理解条件语句“如果 H，那么 C”的方式是通过集合和子集。让 $\mathcal{H}$ 表示满足假设 H 的所有对象的集合，让 $\mathcal{C}$ 表示满足结论 C 的所有对象的集合。逻辑蕴涵“如果 H，那么 C”的形式为 $\mathcal{H} \subset \mathcal{C}$，见下图 1.3。用文字来说，“每个满足假设 H 的对象也满足结论 C。”

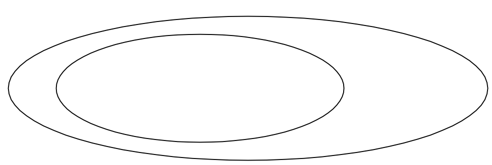


图1.3：“如果H，则C”的维恩图。

设 $\mathcal{H}^c$ 表示不满足假设 H 的对象集合，设 $\mathcal{C}^c$ 表示不满足结论 C 的对象集合。因此 $\mathcal{H}^c$ 是集合 $\mathcal{H}$ 的集合论补集。从图 1.3 和 1.4 可以清楚地看出 $\mathcal{H} \subset \mathcal{C}$ 与 $\mathcal{C}^c \subset \mathcal{H}^c$ 意义相同。相应的逻辑蕴涵，“如果不 C，那么不 H”，被称为命题“如果 H，那么 C”的逆否命题。每个蕴涵都与它的逆否命题在逻辑上等价。逆否命题有时表述为“只有当 H 时，才 C。”

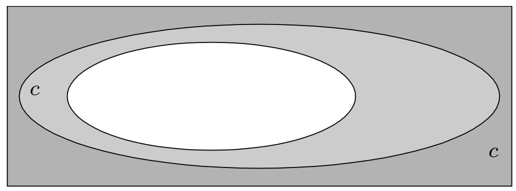


图1.4：逆否命题：“如果不 C，那么不 H”。

它通常使用符号 $H \implies C$，读作“H 蕴含 C”，而不是等价形式 $\mathcal{H} \subset \mathcal{C}$ 或“如果 H，那么 C。”逻辑学家用 $\neg C$ 代替“非 C。”在这种符号中，逆否命题写作 $\neg C \implies \neg H$。

然而，$\mathcal{H} \subset \mathcal{C}$的第三次重新表述是$\mathcal{H} \cap \mathcal{C}^c = \varnothing$。用话来说，没有对象既满足假设$H$又未能满足结论$C$。

Statement	Name	Set Interpretation
If H, then C.	Direct Implication	$\mathcal{H} \subset \mathcal{C}$
If not C, then not H.	Contrapositive	$\mathcal{C}^c \subset \mathcal{H}^c$

表1.1：直接蕴涵和逆否

以下句子（每个都与 $H \implies C$ 逻辑等价）可互换使用，使用主要受风格决定：如果 H，则 C；C 当且仅当 H；H 仅当 C；H 是 C 的充分条件；C 是 H 的必要条件。数学阅读需要高度关注精确措辞！

示例1.5（有效）的蕴涵“如果m和n是偶数整数，那么$m+n$是偶数”可以等价地写成

- $m+n$ 即使 m 和 n 是偶数。

- m 并且 n 仅当 $m+n$ 为偶数时才为偶数。
- 为了使 $m+n$ 为偶数，只需 m 和 n 为偶数即可。
- 为了使 m 和 n 为偶数，必须使 $m+n$ 为偶数。
- 如果 $m+n$ 不是偶数，那么 m 和 n 不都是偶数；换句话说，m 和 n 中至少有一个不是偶数。

仔细观察，最后一句中没有任何内容排除了m和n都不是偶数的可能性。假设是“$m+n$是奇数”，结论是“m和n不都是偶数”（或者说“至少有一个是奇数”）。虽然确实这两个数不能都是奇数（如果它们的和不是偶数），但这一事实并未被断言。你可能会说一个真实的数学断言不必说出全部真相。□

对数和逆函数

存在两个其他陈述在外观上与 $H \implies C$ 相似——但逻辑上不等价。第一个，称为逆命题，是 $C \implies H$；这声称如果结论得到满足，那么假设也得到满足。第二个，称为逆命题 $\neg H \implies \neg C$，声称如果假设没有得到满足，那么结论也没有得到满足。

Statement	Name	Set Interpretation
If C, then H.	Converse	$\mathcal{C} \subset \mathcal{H}$
If not H, then not C.	Inverse	$\mathcal{H}^c \subset \mathcal{C}^c$

表1.2：逆命题和反命题。

一个非常常见的错误是将一个陈述与其逆命题或逆否命题混淆。一般来说，一个陈述在逻辑上并不等同于其逆命题。在例1.5中，（非有效）的逆命题是“如果$m+n$是偶数，那么m和n都是偶数。”逆命题的错误性通过反例的存在得到体现：确实，两个奇数的和是偶数。

证明与反证方法

每个数学定理都（等价于）一个或多个逻辑蕴涵，尽管有时逻辑蕴涵被掩盖了

定理的措辞。一个“当且仅当”（有时写作“iff”或“$\Leftrightarrow$”）陈述是一对逻辑蕴涵，其中每个陈述都是另一个陈述的逆命题。粗略地说，一个真实的iff陈述就是全部真相。在定义中，将“如果”一词始终理解为“当且仅当”是一种风格传统。因此，上述定义1.1实际上意味着“一个整数n是偶数，当且仅当存在一个整数m，使得$n = 2m$。”

证明是用来确立逻辑蕴涵 $H \implies C$ 真理的论证。有三种证明方法，对应于蕴涵的三个集合论解释。

直接证明使用直接方法，通过展示如果 x 在 $\mathcal{H}$ 中，则 x 在 $\mathcal{C}$ 中，或者“对于每个 $x \in \mathcal{H}$，有 $x \in \mathcal{C}$。”用文字来说，选择一个满足假设 H 的“泛型”对象 x，并展示这个对象也必须满足结论 C。

逆否命题证明逆否命题依赖于 $\mathcal{H} \subset \mathcal{C}$ 和 $\mathcal{C}^c \subset \mathcal{H}^c$ 的等价性。这种方法可以被视为一种不同的证明方法，或者被视为一个不同（但逻辑上等价）命题的直接证明。在此方法中，选择一个“泛型”对象，该对象未能满足结论，并展示它也未能满足假设。

矛盾证明通过反证法依赖于 $\mathcal{H} \subset \mathcal{C}$ 和 $\mathcal{H} \cap \mathcal{C}^c = \varnothing$ 的等价性。方法是表明如果某个对象同时满足假设且不满足结论，那么数学在逻辑上是不一致的：存在一个陈述 P，使得 P 和 $\neg P$ 都为真。

示例1.6不存在2的有理平方根的标准证明是通过反证法，并且更多地依赖于引理1.4。假设m/n是最简形式，并且$(m/n)^2 = 2$，或者$m^2 = 2n^2$。这个方程意味着偶数$2n^2$是一个完全平方数，因此根据引理1.4，它是4的倍数。因此$m = 2\ell$，即m是偶数。将$2n^2 = m^2 = 4\ell^2$除以2得到$n^2 = 2\ell^2$。再次应用引理1.4，我们发现n^2是4的倍数，因此n是偶数。因此m/n不是最简形式，这与原始假设相矛盾。

总结来说，上述论点表明，如果 m/n 以最简形式表示且 $(m/n)^2 = 2$，那么 m/n 不是以最简形式表示。这表明

$(m/n)2 = 2$ 是不可能的，也就是说，2没有有理数的平方根。
□

证明反证法在逻辑上往往显得笨拙，通常不会在假设和结论之间建立逻辑联系。因此，反证法应被视为最后的手段。幸运的是，大多数反证法可以很容易地改写为逆否命题的证明，尽管在定理1.3中，可能需要适当地重新表述蕴含关系。

这是一个常见的错误，尤其是在考试压力下，从假设结论开始证明。这相当于假设要证明的内容，显然是错误的。为了强调：

> 未提供完整的源文本，无法进行翻译。请提供完整的源文本以便翻译。
> 假设。

让我们暂时回到定理1.2，它断言（松散地）在欧几里得几何中不可能将圆平方。以下是证明的基本思想：首先，如果可以用直尺和圆规在有限步骤内构造出长度为π的线段，那么实数π将满足一个有理系数的多项式关系——类似于$\pi^2 - 10 = 0$或$1 - \pi^2/6 + \pi^4/120 = 0$（这两个都不正确）。但是，出于纯粹的分析原因（遗憾的是，这些原因超出了本书的范围），这样的关系将意味着存在一个介于0和1之间的整数。由于不存在这样的整数，因此所谓的构造在意义上是不可能的，即与数字的基本性质不相容。

反例 前述项目涉及建立逻辑蕴涵的真实性。证明逻辑蕴涵的虚假性的对偶任务是通过对反例的运用来完成的。对断言$H \implies C$的反例是一个对象x，它既满足假设又未能满足结论。这样的x的存在证明了交集$\mathcal{H} \cap \mathcal{C}^c$非空，因此$\mathcal{H} \subset \mathcal{C}$是错误的，见图1.5。

虽然可以通过找到一个既满足假设又未能满足结论的对象来证明断言 $H \implies C$ 的错误，但无法通过找到一个同时满足假设和结论的例子 y 来证明陈述 $H \implies C$；在图 1.6 中，这样的例子存在，但陈述“$H \implies C$”是错误的。为了证明逻辑蕴涵，必须证明每个满足假设的对象都满足结论。

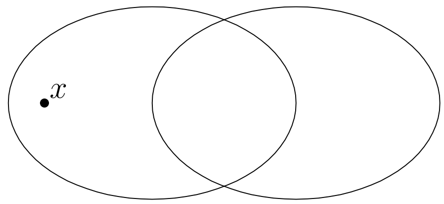


图1.5：一个反例。

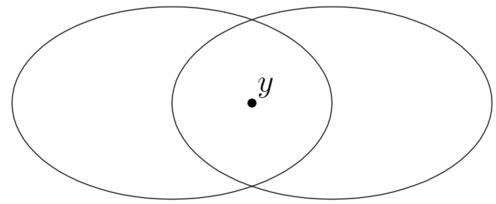


图1.6：一个误导性的例子。

摘要

三个主题贯穿于所有发达的数学知识体系：定义、定理和例子。定义可以自由地做出，但好的定义通常是来之不易的。当你第一次遇到一个定义时，你应该问自己：有哪些例子和非例子？哪些对象被定义所区分？通常，只指定可以用更简单的替代品来表征的对象的定义是糟糕的定义。定义也可能将大量微妙的信息压缩到几个看似简单的标准中。自然数的定义——本书中的第一个精确定义——是一个很好的例子。大致来说，定义的质量可以通过其简单性（与其易用性相关）和它产生的意外后果的数量来衡量。

例子使定义生动起来；最糟糕的定义是没有例子（或没有反例），因为定义要么逻辑上不一致，要么是同义反复。例子使定义能够用于模拟现实生活中的情况，而定义通常被选择来使某些现实世界的直觉变得精确。正如之前所提到的，验证一个特定对象是否是定义的例子可能非常困难！因此，简单的定义更可取：在处理特定例子时，它们更容易验证。然而，A. 爱因斯坦的建议是恰当的：尽可能使事物简单，但不要过于简单。

定理是定义的非显而易见的后果。它们对于分类例子（可能通过以等效且非显而易见的方式重新表述定义）组织概念之间的逻辑关系以及扩展关于对象类别的知识是有用的。通常，通过验证对象是某些定义的例子，然后使用保证所有此类对象都具有所需属性的定理，可以从中获得关于特定对象的知识。

数学是参与性的，因此这些评论在这个阶段可能意义不大（除非它们直观明显）。如果你在整本书中定期回顾它们，你应该会发现它们的意义变得越来越清晰。

1.4 微积分与“现实世界”

微积分通常被称为研究变化率的数学。它解决的问题包括在平面内找到曲线所围成的面积（以及定义曲线区域中的“面积”是什么意思），找到与曲线相切的直线（以及定义“相切”是什么意思），计算在非恒定重力场中移动物体所需的能量，预测化学反应随着反应物的消耗而进展的速度，找到从悬崖上掉落的石头的速度（以及定义“在某一时刻的速度”是什么意思），等等。作为科学工具，微积分包括三个主要方面：

- “实用”的一面，它将物理、化学、生物学、经济学——“现实世界”——等领域联系起来，并通过这种方式，微积分获得了其最大的消费者群体：科学家和工程师。

  应用形式为“自然法则”，这些是方程式，其解可以模拟物理世界的某些方面。在科学建模中，总是存在简单性和准确性之间的权衡。牛顿的万有引力定律简单，对于大多数实际用途来说是足够的，但它无法解释一些可观察的现象。爱因斯坦的广义相对论改进并扩展了牛顿定律，但它需要进行预测计算，需要困难和微妙的数学。

- “计算”方面，它允许将关于“无限小”的数量、空间中的“点”和时间中的“瞬间”的直观想法转化为符号表达式，并对其进行操作，以获得对上述类型问题的有用答案。

- “理论”基础，它以集合论（数学的“机器语言”）定义微分和积分，并证明这些操作在逻辑上是一致的，因此

> 即使技术细节和直觉不完全一致，所获得的答案在某种意义上仍然是合理的。

每个方面本身都很重要，并且每个方面都支持其他方面，尽管大多数用户不需要理解基础才能使用应用程序。（用机械比喻来说，你可以开车而不需要能够重新组装引擎。）然而，对某个主题有真正的理解比仅仅熟悉它更好；真正的知识是灵活的，可以在新情况下正确应用，而熟悉性仅限于之前遇到的情况。

用数学建模物理学

一个简单的物理例子将有助于说明这些方面以及由此产生的几个相关问题。假设一块石头从高 y_0 米的悬崖上掉下；石头落地需要多少秒，石头撞击地面时的速度有多快？在基础物理学中，这些答案被编码在经过 t 秒后石头的高度 $y(t)$ 的公式中:

$$y(t) = y_0 - 4.9t^2. \tag{1.1}$$

重要的是要记住，这个模型并不是对现实的精确描述。接下来，我们将详细讨论构建此模型所依据的假设，以说明数学与科学相互作用的实际方式。

为了使用数学来描述自然现象，必须采用一种近似和数学建模的往复程序。在像落石这样简单的情况下，这些心理活动往往是无意识的，但在复杂和新颖的问题中，有必要了解如何将数学翻译成“现实世界”以及如何反过来。

作为第一次近似，石头将被视为一个点粒子，除了位置和质量外没有其他属性，地球表面将通过一个平面来建模。不考虑风和科里奥利效应的影响，石头垂直下落，因此石头的运动可以通过知道它在释放后的每一秒（例如t ()）的高度（以米为单位，例如$y(t)$ ()）来决定，即地面以上。即使在非正式的英语中，石头的高度也被说成是“时间的函数。”函数的数学概念

这是基本的，在第3章中进行了详细讨论。因为所关注的数量——在时间 t 时石的高度——由一个单一的数字确定，因此这个模型被称为"一变量"问题。

要进一步说任何东西，有必要借用一些物理学结果；数学对石头下落的方式绝对什么也没说，更不用说一般性的除了数学之外的事情。在我们的理想化情况下，石头的运动由牛顿运动定律支配：有"力"在作用（重力和空气阻力是最重要的），这些力决定了石头的"加速度"。加速度是微积分的概念：石头的速度是它位置变化的速率（以每秒米为单位），而加速度是速度变化的速率（"每秒米"每秒）。在牛顿模型中，作用于石头的力决定了它的行为，正是这种预测能力回答了原始问题。

它方便进行一些理想化 s:

- 重力加速度在石头下落过程中是恒定的。
- 没有空气阻力。

第一个假设如下得到证明。根据牛顿的万有引力定律，作用于石头的力是石头质量 m 乘以地球质量 M 的一个常数 G 倍，除以石头到地球中心的距离 $R(t)$ 的平方，再除以时间 t。根据牛顿第三运动定律，作用于石头的合力等于石头质量乘以其加速度。用符号表示，

$$F = -G\frac{mM}{R(t)^2} = ma; \tag{1.2}$$

负号表示力指向地球中心。因为石头下落的距离与地球半径 R 相比非常小，所以比值 $R(t)/R$ 在石头下落过程中几乎等于1，因此方程（1.2）中的分母可以用 R^2 替换而不会造成太大的误差。没有空气阻力的假设并不现实，即使物体是完美的球形（另一个不现实的假设），在固体上模拟空气阻力也是非常复杂的。然而，要说明的是建模问题，忽略空气阻力可以说明这一点

有时你必须做出不切实际的简化，以便进行可处理的计算。

根据公式（1.2），石头的加速度（大约！）为 $a = -GM/R^2$。这个数值测量结果大约为地球表面每秒每秒 –9.8 米。最后我们准备好使用微积分。要使用的技术称为“积分”，从第7章开始详细研究。在这里我们给出直观描述；下文中的内容不应过于字面理解。

石头的加速度——速度的变化率——是恒定的，所以在每个时间瞬间，速度增加的量是相同的，经过 t 秒后等于 $v(t) = v_0 - at$ 米/秒；v_0 是初始速度，因为石头是掉落的，所以初始速度为零。

现在，在相同的时间瞬间，石头下落了$-at\,dt$米的距离。将这些无穷小距离相加（并省略细节）最终导致（1.1）。重复一遍，石头撞击前的海拔高度是

$$y(t) = y_0 - \frac{1}{2}at^2 = y_0 - 4.9\,t^2.$$

作为一个理智的检查，让我们验证这个公式的正确性：在时间点 t 和 $t + dt$ 之间，石头从高度 $y(t)$ 落到 $y(t + dt)$，距离为

$$dy := y(t + dt) - y(t) = -\frac{1}{2}a\Big((t + dt)^2 - t^2\Big) = -\frac{1}{2}a(2t\,dt + dt^2).$$

除以 dt 给石头一个瞬时的速度为

$$\frac{dy}{dt} = -at - \frac{a}{2}dt, \tag{1.3}$$

这本质上等同于 $v(t) = -at$，因为 dt 非常小。

再次强调，方程（1.1）是描述从y_0米高度释放后t秒的石块高度的模型。有了这个公式，我们可以回答“石头何时落地？”的问题，因为石头的撞击对应于模型中的条件$y(t) = 0$，并且这个方程可以很容易地解出t，以初始高度y_0为依据。石头的撞击速度甚至更容易找到: at_0米/秒（被丢弃的负号仅表示石头在向下移动时撞击地面）。就初始高度y_0 (米而言),

$$\text{Impact time} = \sqrt{\frac{y_0}{4.9}}, \qquad \text{Impact velocity} = 2\sqrt{4.9y_0}.$$

对于大多数微积分课程来说，这就是故事的结尾，确实，希望数学或物理对大家来说都不陌生。详细地通过这个简单例子，是为了指出模拟现实世界情况的特点，并提及在从方程（1.2）中获得石头的速度以及在验证方程（1.1）确实导致计算出的速度时发生的有趣且具有争议的步骤。

速度应表示“位置的变化率”，但这是什么意思呢？（如果这个问题无法回答，那么就没有理由相信方程 $v = -at$ 有意义或有用！）古希腊的芝诺发现了以下明显的悖论。想象一块石头在“某一瞬间”下落。石头有一个确定的位置，并且与在同一高度的静止石头无法区分。更具体地说，想象一个无限快的相机，可以捕捉到时间的真实瞬间。那么，就没有办法根据一张照片来区分运动物体和静止物体。但是，由于这个论点可以在每个瞬间进行，所以运动和静止之间没有区别！然而，下落的石头确实下落而不是悬浮，运动显然是可能的。论点在哪里走错了？为了直观地理解为什么没有悖论，想象一个可以捕捉任意小时间间隔的非常快的相机；千分之一秒的曝光，十亿分之一秒，或者十亿亿亿亿分之一秒的曝光都是可能的。对于这样的相机来说，运动的石头和静止的石头看起来并不相同；运动的石头因为其位置在快门开启时发生变化而形成略微模糊的图像。当然，随着快门速度的增加，下落石头的图像会变得更加清晰，但最终的照片永远不会与静止石头的照片相同。此外，随着快门速度的增加，石头行进的距离除以曝光时间“越来越接近一个极限值”；这个“极限值”的单位是米每秒，被解释为石头的“瞬时速度”。实际上，这个极限过程“将时间尺度放大了 ∞ 倍”。这个解释在很大程度上是不完整的，因为引号中的短语还没有被定义。直观地说，要点是:

- 静止和运动状态可以在任意小的时间间隔内区分开来，尽管它们在单个瞬间无法区分；

- 如果我们观察越来越小的时间间隔，运动看起来越来越像以恒定速度进行；从几何上看，随着时间的推移，位置作为时间的函数的图像在我们“放大”时间 t 时越来越像一条直线。

尽管这些评论不够精确，但它们包含了精确定义的萌芽，其中可以制定一个逻辑上一致的数学变化率理论。关于“瞬时速度”（距离除以时间）的朴素想法涉及到无意义的表达式 $0/0$。关于“极限”（第四章）的概念巧妙地绕过了这个困难，并允许对瞬时速度进行数学处理。

几个其他问题也隐含地出现了，尽管它们大多非数学性质。“时间瞬间”在物理上没有意义，空间点也是如此，尽管这两个概念作为有用的理想化模型，对于宏观事件来说相当准确。在构建“真实”情况的数学模型时，总是必须忽略某些效应，忘记所考虑对象的某些属性，甚至进行纯粹数学上的近似（例如，用数字计算机解决数值模型）。数学是一种精确而详细的语言，碰巧对描述许多观察到的现象很有用。许多自然法则可以方便地用微分方程的语言表达，这些方程描述了数量及其变化率。物理学和化学等科学试图将实验结果与数学描述联系起来，以便可以事先预测结果或以其他方式理解结果。这些数学模型——所谓的自然法则——可能非常准确，但迄今为止，没有一个模型是包罗万象的或完全准确到测量极限的。

有充分的理由断言数学在物理意义上并非“真实”；它是我们的大脑用来构建准确、预测现实模型的一种工具或语言。本书的目的是通过从集合论构建它，展示微积分背后的数学逻辑上是一致的，同时给出这些强大数学技术的有趣和实质性的应用。

练习

一些这些练习相当传统，并假设你熟悉标准数学符号。其他练习旨在让你思考语言、心智模型和语义。

唯一学习“用数学说话”的方法是通过实践；写作、重新表述和思考。可以通过阅读来获得熟悉感，但原创性只能通过参与获得。此外，数学使用英语单词和语法（至少在这本书中是这样），但它不是英语。一些练习说明了重要差异。数学定义和定理往往以“逻辑”形式陈述，而我们的头脑通常不擅长理解，参见练习1.10。对于大多数学生来说，认识到数学是一种具有相当严格语法的语言，日常用语通常相当不精确，以及通过联想提出的东西不是通过逻辑联系起来的，这是一个心理障碍，参见练习1.3、1.4和1.9。（几乎所有广告都依赖于联想与逻辑联系的混淆。）

练习1.1 证明如果整数 k 是一个完全平方数，那么 k 或 $k-1$ 可以被4整除。（比较引理1.4。）$\diamond$

练习1.2 设X和Y为集合。对称差$X \triangle Y$定义为$(X \setminus Y) \cup (Y \setminus X)$。用维恩图说明该定义；证明

$$X \triangle Y = (X \cup Y) \setminus (X \cap Y),$$

对称差对应于布尔运算异或（“排他或”）。

注意：为了证明两个集合 X 和 Y 相等，你必须证明两件事：$X \subset Y$，和 $Y \subset X$。$\diamond$

练习1.3 某人拿着一袋蓝色弹珠说：“这个袋子里的每颗弹珠要么是红色要么是蓝色。” 这是不是一个正确的陈述？如果不是，为什么不是？如果是，它“告诉了全部真相”吗？解释一下。$\diamond$

练习1.4 R. M. Waldo，有记录的最高人类，身高略低于9英尺。假设对于这个问题，他是曾经活着的最高的人类，并且他的身高正好是9英尺。

(a) 考虑以下说法，“人类最高不超过12英尺。”这个说法是真的吗？如果不是，为什么不是？如果是，你能将其写成“如果……，那么……”的形式吗？

(b) 考虑这个说法，“人类最多9英尺高。”这个说法“是否说了全部真相”？如果是，它是在什么确切的意义上做到的？

(c) 部分a中的断言是否意味着部分b中的断言，反之亦然?

数学上讲真话与在法庭上讲真话不同。（证明标准也不同，但这又是另一个问题。）⋄

练习1.5 虽然一开始看起来很奇怪，但一个集合可以是它自己的元素。集合论的一个早期问题是罗素悖论：设X是所有不是自身元素的集合的集合。证明X是它自己的元素当且仅当它不是它自己的元素。

这个问题通过更仔细地处理“集合”的概念得到了解决：一开始必须固定一个宇宙，而“所有集合的集合”不是一个“集合”，而是一个更大的对象，称为“类”。⋄

练习1.6 20世纪逻辑的一些最伟大成就由K.哥德尔完成。哥德尔提出了一个算术陈述，可以解释为“这个陈述不能被证明”。假设数学中没有矛盾，证明哥德尔陈述是真实的但不可证明的（在它所形成的公理框架中）。这摧毁了每个数学真理都可以在某个固定的公理系统中被证明的宝贵想法。⋄

许多分析中的陈述涉及量词“对于每个”（全称量词 $\forall$）和“存在”（存在量词 $\exists$）。尽管如此，我们不会使用这些符号，尽管提醒您，有些人喜欢写作，例如

$$(\forall\varepsilon > 0)\ (\exists\delta > 0)\ |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

下一个练习介绍了量词的一些基本性质。
练习1.7 考虑以下量化句子：

(a) 对于袋中的每一个弹珠 x，x 是蓝色的。 (b) 袋中存在一个红色弹珠 y。 (c) 对于每一个实数 $x > 0$，存在一个自然数 n 满足 $1/n < x$。

给出每个句子的否定。将（a）和（c）表达为条件语句，并给出它们的逆否命题。将（b）和（c）的否定表达为条件语句，并给出它们的逆否命题。

注意，“对于每一个”和“存在”在否定下被交换。$\diamond$

练习1.8 一个游戏节目主持人向参赛者展示等式“$a^2+b^2=c^2$”。参赛者回答，“什么是勾股定理？”

(a) 这真的是正确的问题吗？如果不是，你能添加一个条款来提出一个能满足数学家的问题吗？

(b) 以如果-那么形式陈述勾股定理。

一个定理不仅仅是它的结论。尽管如此，在A. Wiles宣布费马大定理的证明之后，伯克利数学科学研究所（MSRI）制作了印有信息“$a^n+b^n=c^n$：NOT！”的T恤。$\diamond$

练习1.9 总统，一个守法公民，总是说实话，在新闻发布会上还有时间回答一个更多是的/否的问题。为了羞辱总统，一名记者问道：“你是否已经停止向来访的国家元首提供非法毒品？”

(a) 哪个答案（“是”或“否”）在逻辑上是真实的？ (b) 假设总统回答“是”。公众能否得出总统向来访的国家元首提供了非法药物的结论？如果答案是“否”呢？ (c) 解释为什么两个答案都很尴尬。

如果总统是一位禅宗佛教徒，她可能会回答“无”，[5]意味着“你的问题在假设上存在太多缺陷，以至于无法有意义的回答。”$\diamond$

练习1.10人类大脑的一个显著特点是，它在以情感的方式看待某些情况方面“更擅长”于理解等效的逻辑公式。以下是一个例子。

(a)一副牌中的每张牌一面印有字母，另一面印有数字。确切地说，字母是“D”或“N”，数字是介于16和70之间的整数。牌的印刷组合没有限制。你的任务是评估牌是否满足

[5]Pronounced “moo”

单标准：“每张‘D’卡在背面都印有大于或等于21的数字。”您还必须将满足此标准的牌与不满足标准的牌分开。

将标准写成“如果……，那么……”的陈述，并确定以下哪些卡片满足该标准：

(i)	(ii)	(iii)	(iv)
D	D	N	N
20	46	16	25

(b) 你被展示了四张牌：

(i)	(ii)	(iii)	(iv)
18	35	D	N

哪些牌需要翻转以确定它们是否满足(a)部分的准则？

(c) 某个州的法定饮酒年龄为21岁。你在聚会上的任务是确保没有21岁以下的人饮酒，并报告那些正在饮酒的人。一个由四个人组成的小组包括一个20岁正在饮酒的人，一个46岁正在饮酒的人，一个16岁未饮酒的人，和一个25岁未饮酒的人。这些人中谁/哪些人违反了法律？

在报告这一事件后，你在酒吧里发现了四个人：一个18岁和一个35岁的人背对着你，还有两个年龄不明的人，其中一个人正在喝酒。你需要从哪个人那里获取更多信息，以判断他们是否违反了法律？

(d) 解释为什么卡片问题在逻辑上等同于饮酒问题。

Which did you find easier to answer correctly? ◇

第二章

数字

数学是一门研究数字的普遍但错误的观念；一些数学家甚至开玩笑说，公众认为数学研究就是将乘法表扩展到更高的因数。这种误解是由将常规计算视为数学主要目标的学校课程所培养的。事实上，计算是一种技能，它与数学的关系类似于拼写与文学创作的关系。

在学校，你学习了各种类型的数字：计数（自然）数、整数、分数（有理数）和小数（实数）。你可能甚至已经接触到了复数，或者至少是它们最著名的非实数成员，i，-1的平方根。本章的一个目标是让你（重新）熟悉这些数字集合，并介绍它们的抽象性质——算术的交换律、结合律和分配律、不等式的性质等等。同时，你将（概述地）看到这些数字集合是如何从集合论中构建的，并发现你迄今为止对数字的了解几乎完全是符号性的。关于整数不需要十进制符号，关于实数也不强制我们使用无限小数来表示它们。你将学会从公理的角度来理解数字，这些公理抽象了它们的性质。哲学问题“什么是实数？”将消失，留下的答案是“遵循几个公理的集合元素。”

一个误解应该立即消除：尽管 $\sqrt{2}$ 是一个“实”数，而 $\sqrt{-1}$ 是一个“虚”数，但每个符号都代表一个数学抽象，它们的存在并没有一个比另一个更“真实”或更不真实。没有任何物理量可以

明确表示 $\sqrt{\hat{}}2$，因为无法以任意精度进行测量。如果你在计算器上求2的平方根，你会得到一个有理数，其平方通常明显不等于2。甚至不能争论 $\sqrt{\hat{}}2$ 有几何意义（单位正方形的对角线）而 $\sqrt{-\hat{}1}$ 没有；虚数单位完全可以看作是平面的1/4旋转。执行这样的旋转两次（平方）会通过原点反射平面，这相当于乘以 $-\hat{}1$。复数乘法的几何图像"足够真实"，可以用来解释欧拉著名的等式 $e^{i\pi} +\hat{}1 =\hat{}0$，正如我们在第15章中将要看到的。

自然数足够简单，可以直接用集合来定义，但更复杂的数系——整数、有理数、实数和复数——则是依次定义的，基于前一种数系。有一个边缘好处：在从集合理论构建整数时自然出现的递归定义、数学归纳法和等价类技术，在数学中非常重要且有用。迄今为止，最复杂的步骤是用有理数来定义实数。如果递归定义被展开，一个实数就是一个令人难以置信的复杂集合。幸运的是，永远没有必要与"展开"的定义打交道；实数集合满足的抽象性质完全足够。

为了进行计算机类比，集合是位，自然数是字节，整数和有理数是单词，实数和复数是汇编语言，而微积分本身是应用程序。在每一个阶段，感兴趣的物体都是由下一级的物体构建而成。任何理智的人都不会用汇编语言编写电子表格程序，任何理智的人都不会尝试在自然数的层面上用集合来解释积分。微积分（或编程）的递归性质允许你忘记实现的细节，并专注于你构建块所具有的性质。

2.1 自然数

L. Kroeneker，19世纪的数学家，说（自由翻译）："只有自然数是上帝创造的；所有其他的都是人类[的]作品。"更现代（且世俗）的说法是，数学迫使我们研究计数数，但实数

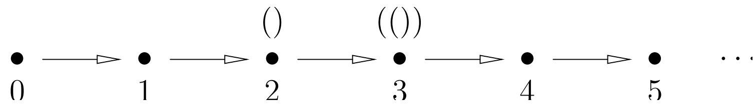


图2.1：自然数集。

是人类发明。

您已经遇到了自然数集合 **N**，它是计数数的集合，从0开始。抽象地讲，N 是一个具有三个特性的“无限列表”：

N.存在一个初始元素，暂时用 ϕ 表示。

N.2 存在继任的概念：每个自然数 n 都有一个唯一的“继任者” $\star n$ ，除了 ϕ 以外的每个自然数都是唯一自然数的继任者，即它的前驱。

N.对于每个非空子集 $A \subset$ **N**，存在一个最小元素，即一个元素 $n_0 \in A$，使得 A 中其他每个元素都出现在以 n_0 为起始的继承链中。

集合 **N** 在图2.1中展示：后继关系用箭头表示，**N** 的元素既用通常的方式（印度-阿拉伯数字）表示，也使用后继的概念表示。省略号表示“模式永远继续”。

属性 **N**.1 表示 **N** 非空，而 **N**.3，即良序性质，是数学归纳法的基础。可能地，**N**.1–**N**.3 在逻辑上是不一致的。为了证明它们不是，我们将构造一个集合 **N** 和一个后继概念，以满足 **N**.1–**N**.3。在大多数情况下，我们将 **N**.1–**N**.3 视为公理——其真实性不容置疑的陈述。换句话说，我们将“忘记定理 2.1 的实现细节”，并将 **N**.1–**N**.3 作为推导自然数性质的起点。

定理2.1. 存在一个集合**N**，具有继任关系概念，满足性质**N**.1–**N**.3。此集合在保持顺序同构的意义上是唯一的。

该术语“保持顺序同构”在第3章中有解释，但在当前情况下意味着“对元素的重命名，保留了后继关系的概念”。

证明。（草图）存在一个唯一的集合 {v*}，它不包含任何元素。定义 ϕ 为集合 $\varnothing$。现在定义后继关系：如果 $n \in \mathbf{N}$ 已经被定义（作为一个集合），那么 n 的后继被定义为集合

$$\star n := n \cup \{n\}. \tag{2.1}$$

展开定义得到，在印度-阿拉伯数字表示法中，

$$\begin{aligned}
1 &= \{\varnothing\} \\
2 &= 1 \cup \{1\} = \big\{\varnothing,\ \{\varnothing\}\big\} \\
3 &= 2 \cup \{2\} = \big\{\varnothing,\ \{\varnothing\},\ \big\{\varnothing,\ \{\varnothing\}\big\}\big\} \\
&\ \ \vdots
\end{aligned}$$

注意 $n \subset \star n$ 作为集合，以及（例如）集合“3”有三个元素：$\varnothing$、$\{\varnothing\}$ 和 $\big\{\varnothing,\ \{\varnothing\}\big\}$。

在一个仔细的证明中，有许多细节需要检查，例如发生构造的通用集的构造。还必须小心只使用集合论公理，而不要依赖（可能是愿望！）直觉。然而，以下断言应该是清晰的：

- 对于每一对不同的自然数 n 和 m（被视为集合），要么在以 n 为起点的继承链中出现 $n \subset m$ 和 m，反之亦然。自然数中“小于”的通常意义恰好是集合的严格包含。

- 集合 n 的“元素数量”正好是自然数 n 通常所表示的含义。这是因为集合 $\star n$ 比集合 n 多一个元素，即元素 n。（不要将嵌套在一组花括号内的元素与“$\varnothing$”符号的数量混淆。）

属性 **N**.1 通过构造是正确的，大多数 **N**.2 也是如此；唯一不明显的是自然数不能有两个不同的前驱。但承认刚才提出的这些事实，如果 $\star n_1 = \star n_2$，那么（如果需要的话重新标记）$n_1 \subset n_2$。如果 n_1 和 n_2 是不同的，那么我们就会得到 $\star n_1 \subseteq n_2$，这是不可能的，因为 $\star n_1 = \star n_2$。

为了证明每个非空集合 $A \subset \mathbf{N}$ 都有 s 最小元素，从 ϕ 开始并取后继。在每一阶段，你或者在 A，或者不在。如果你从 A 开始，则无需证明，而如果你从未到达 A 的元素，那么根据 $\mathbf{N}$ 的构造，集合 A 是

空。否则，存在第一次继任产生 A 的元素 n_0，而 A 的其他每个元素都必须在以 n_0 为起点的继任链中出现。

□

尽管(2.1)的递归结构简单，但展开的符号无法使用：写出整数100是不可能的，因为“$\varnothing$”符号的数量随着每次连续增加而翻倍。“井号”，例如／表示“5”，以更有效的方式（类似于罗马数字）表示自然数，但对于现代使用来说仍然不足。印度-阿拉伯数字极其紧凑；每个额外的数字描述的数字是前一个的十倍，使用位置上下文来赋予数字值（个位、十位等等）简化了计算，使得写出、加法和乘以巨大的自然数变得容易。

递归定义，归纳

考虑从头开始编写计算机程序的问题。通常，程序员会选择一种“计算机语言”，例如C，编写一个正式结构化但易于阅读的“源代码文件”，然后使用一种称为编译器的特殊程序将源代码转换为机器可以执行的字节序列。现在，现代C编译器是一个极其复杂、复杂的程序，这是太复杂而无法从头开始编写的。那么，没有编译器，人们该如何开始呢？答案是，他们首先用机器语言编写一个小的、功能不太强大的编译器，然后使用它来编译一个更强大的编译器。接下来，他们会用他们的“第二阶段”编译器编写一个完整的C编译器。这样装备后，他们会编写他们最初设定的程序。

通过类比，假设我们希望从头开始构建实数集。我们上面构建的 **N** 就像一台裸机，能够被编程但没有任何软件。自然数集并没有配备加法、乘法和指数运算；这些必须从后继的概念中构建，并且类似于我们手写的，“第一阶段”编译器。拥有自然数和算术运算后，我们继续构建整数和有理数，这类似于后续的编译器；只有到那时，我们才准备好构建实数集。

上述引入的思想体现了递归定义，其中

一个对象或结构的序列被定义，每个都是基于前一个定义的。甚至**N**上的算术运算的构造也具有强烈的递归性质。加法定义为迭代序列："$2+3$"表示"2的后继的后继的后继"（'将2的后继取3次'），见方程（2.2）。一旦加法可用，乘法定义为迭代加法，指数运算定义为迭代乘法。使用数学归纳法，一种建立无限多个适当"逻辑相关"断言的真实性的技术，将证明算术的熟悉性质。从这些结构中流淌出一条浩瀚而深邃的思想和定理之河，其范围远非完全映射。费马大定理仅在1994年得到证明；关于**N**的其他断言，如哥德巴赫猜想[1]或孪生素数猜想[2]，仍然悬而未决。

自然数之和

自然数的加法背后的直觉是"堆的聚集"，就像在{v*}

$\bullet\bullet + \bullet\bullet\bullet = \bullet\bullet\bullet\bullet\bullet$您应该记住，以下讨论只是这个想法的形式化，并记住歌德的引言。

假设我们希望编程计算机来加两个自然数，使用加法的定义作为迭代后继。虽然直观上表达式"$m+n$"表示"从m开始并取后继n次"，但这种规定并不立即有用，因为计算机有一个固定的处理器堆栈，而数字n可以是任意大的。所需的是一个只需要固定处理器内存的进程。

想象一下，此刻你不知道"$+$"的含义。符号的任何使用都必须用自然数和后继关系来解释，特别是没有理由假设像$m+n=n+m$这样的形式性质。设m为一个自然数，并定义$m+1=\star m$。现在，如果n是一个自然数，并且表达式"$m+n$"已经为所有$m\in\mathbf{N}$定义，那么我们定义

$$m+(n+1) := (m+n)+1 \qquad \text{for all } m\in\mathbf{N}. \tag{2.2}$$

有必要为 $m+1$ 制定一个单独的定义以启动过程。在不那么暗示的记法中，$m+(\star n) := \star(m+n)$ 对所有

[1] Every even number greater than 4 is a sum of two odd primes.

[2] There exist infinitely many pairs of primes of the form $\{p,\ p+2\}$.

$m \in \mathbf{N}$它不明显，对于所有 m 和 n，有 $m+n=n+m$；这是下面定理2.6的一部分。

方程（2.2）提供了一个“取n-次迭代后继”的算法：从两个“石头堆”开始，将石头从第二个堆移动到第一个堆，直到第二个堆中没有石头为止。在伪代码中，

```
让 m 和 n 为自然数；当 (n > 0) 时 {
用 m 的后继替换 m；用 n 的前驱替
换 n；} 返回 m；
```

数学归纳法

定理2.2，数学归纳原理，是自然数在分析中出现的最重要的地方之一。在许多情况下，可以通过证明其中之一（通常是但并不总是第一个）的真实性，并展示每个陈述都蕴含下一个陈述，来证明一个无限长的断言序列。归纳特别适合于已经递归定义了某物的情况。

定理2.2. 假设每个自然数$n \in \mathbf{N}$都与一个句子$P(n)$相关联，并且

I.1（基本情形）对于某些 $n_0 \in \mathbf{N}$，句子 $P(n_0)$ 是正确的；

I.2（归纳步骤）对于每个自然数 $k \geq n_0$，$P(k)$ 的真值蕴含 $P(k+$ 的真值 $1)$。

然后每个句子 $P(n)$ 与 $n \geq n_0$ 都是真实的。

证明。直观上，根据假设，$P(n_0)$ 是正确的，因此归纳步骤表明 $P(n_0+1)$ 也必须正确。但随后归纳步骤又表明 $P(n_0+2)$ 是正确的，以此类推，无穷无尽。因此，所有后续句子都是正确的。

形式上，令 $A \subset \mathbf{N}$ 为满足 $P(n)$ 为假的 $n \geq n_0$ 的集合。我们希望证明如果定理的假设成立，则 A 为空集。等价地，我们可以建立逆否命题：如果 A 不为空，则定理的假设不成立。

假设 A 不为空。根据性质 **N**.3，存在 A 的最小元素。令 n 为其前驱。根据 A 的定义，陈述 $P(n)$ 为真，而 $P(n+1)$ 为假；因此，归纳步骤对 $k=n$ 失败，所以定理的假设不成立。

□

示例2.3 假设$(a_k)=a_0$，a_1，a_2，... 是一个数列。我们使用递归定义来定义数列的部分和如下。首先我们设 $S_0 = a_0$；然后，对于 $n \geq 0$ 我们定义 $S_{n+1} = S_n + a_{n+1}$。这个定义依次展开到 $S_0 = a_0$，$S_1 = a_0 + a_1$，$S_2 = a_0 + a_1 + a_2$，等等。求和符号在此处非常有用；我们写

$$S_n = a_0 + a_1 + \cdots + a_n = \sum_{k=0}^{n} a_k \text{ for } n \geq 0.$$

右边的表达式读作“从 $k=0$ 到 n 的 a_k 的和”，称为序列 (a_k) 的 n 部分和。部分和的递归定义看起来像

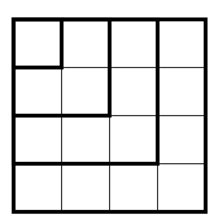
$$\sum_{k=0}^{n+1} a_k = a_{n+1} + \sum_{k=0}^{n} a_k \text{ for } n \geq 0. \tag{2.3}$$

符号 k，称为虚指数（求和的），是一个“局部变量”，在求和符号之外没有意义。它仅仅是一个占位符，用来提醒我们正在对有限序列的项进行求和，可以用任何尚未使用的字母替换，例如 i 或 j。相比之下，n 与项数相关，不是一个局部变量。

它很重要，要熟练掌握求和符号的表示法。我们将在适当的时候遇到许多例子。练习2.4提供了进一步的练习。□

示例2.4 假设我们希望找到一个公式来求前n个奇数的和。第一步是猜测答案！虽然数学是演绎的，但它通常是通过试错或有根据的猜测来发现的。证明的作用是验证猜测是否正确。奇数是1、3、5、7等等；第n个奇数是$2n-1$。前几个和是

$$\begin{aligned} 1 &= 1 \\ 1+3 &= 4 \\ 1+3+5 &= 9 \\ 1+3+5+7 &= 16 \end{aligned}$$



基于此证据，看起来前n个奇数的和是n^2。然而，尽管图像引人入胜，但无论进行多少逐个案例的检查，都无法证明这个论断对所有$n \in \mathbf{N}$都成立，因为提出了无限多个论断。为了尝试通过归纳法进行证明，考虑以下句子

$$P(n): \qquad \sum_{i=1}^{n}(2i-1) = 1+3+5+\cdots+(2n-1) = n^2.$$

在这个阶段，我们不知道句子 $P(n)$ 中的一些或全部是否为真；我们将通过展示条件 I.1 和 I.2 成立来尝试证明它们的真实性。

为了验证I.1，将n在$P(n)$中所有地方替换为1；这得到句子$1 = 1^2$，这是一个明显的真理。为了建立归纳步骤，假设第k个句子$P(k)$是正确的。在这个例子中，假设

$$P(k): \qquad \sum_{i=1}^{k}(2i-1) = k^2.$$

左边是前 k 个奇数的和；要得到前 $(k+1)$ 个奇数的和，我们在等式两边加上 $(k+1)$ 个奇数 $2(k+1)-1 = 2k+1$，并进行代数运算：

$$\begin{aligned}\sum_{i=1}^{k+1}(2i-1) &= \Big(\sum_{i=1}^{k}(2i-1)\Big) + (2k+1)\\ &= k^2 + (2k+1) = (k+1)^2.\end{aligned}$$

第二个等式使用了归纳假设 $P(k)$，而最后一步是一个代数恒等式。整个方程是断言 $P(k+1)$。因此，如果 $P(k)$ 是真的，那么 $P(k+1)$ 也是真的。根据归纳原理，$P(n)$ 对所有 $n \in \mathbf{N}$ 都是真的。

这个论点的最终结果是 一个有用的事实，并且据说它以“封闭形式”表达所给的求和。要找到前1000个奇数的和，不需要进行加法运算，只需平方1000即可。□

数学家以将问题简化为已解决的问题而闻名。在原则上，更难的问题也因此得到解决；在实践中，一个完整的解决方案可能极其复杂，

因为早期问题可能依赖于一个更简单问题的解决方案，以此类推。

示例2.5 汉诺塔谜题由7个按大小递减的盘子组成，最初堆叠在一起，可以放置在三个轴中的任何一个上，如图2.2所示。目标是根据以下两个规则将盘子堆从一处移动到另一处：

(i) 一次只能转移一个磁盘。

(ii) 磁盘只能放在更大的磁盘上或空位上。

问题有两方面：确定如何转移磁盘，以及找到最小的单个转移次数以移动整个堆栈。

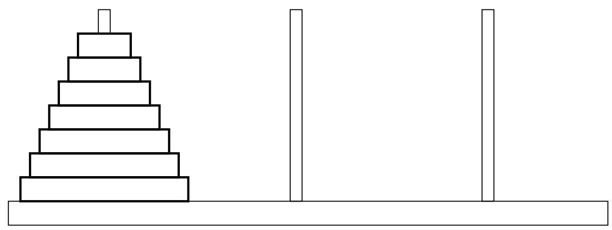


图2.2：汉诺塔。

更普遍地，游戏可以用 n 个盘子来玩。在继续阅读之前，你应该尝试解决 n 的较小值谜题；不同面额的硬币是很好的盘子。当 $n =$ 为 1 或 $n =$ 为 2 时，解决方案很明显，而当 $n =$ 为 3 或 $n =$ 为 4 时，解决方案应该很容易找到。根据传说，有一个婆罗门寺院，那里的僧侣们不停地劳作，将64个盘子从一个地方转移到另一个地方，遵循上述规则；当他们完成这项任务时，宇宙将在一声雷鸣中结束。我们不必担心这个传说的真实性，这一点很快就会变得明显。

汉诺塔具有美丽而简单的递归结构。让我们以管理者的方法来处理一般问题：假设我们知道如何将一叠 $(n-1)$ 个盘子在任意一对轴之间移动。然后我们可以通过将轴 1 上的最上面的 $(n-1)$ 个盘子移动到轴 2（图 2.3；这需要许多单独的转移，但可以视为一个单一操作），将最大的盘子从 1 移动到 3，最后将盘子堆从 2 移动到 3 来解决问题。这简化了解决

n 汉诺盘塔问题求解 $(n-1)$ 盘的汉诺盘塔。1 盘的塔是平凡的。如果你懂一种编程语言，你可能喜欢实现这个递归算法，并看看它用少量盘子运行需要多长时间。

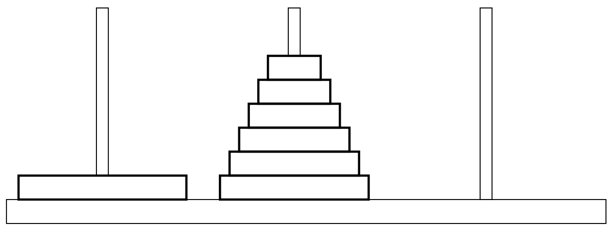


图2.3：递归解决汉诺塔问题。

要看到需要发生多少次单独的转移，请更仔细地检查解决方案的递归结构：想象有 n 个人，标记为 1 到 n，并且那个人 j 懂得解决 j 盘的汉诺塔问题。在解决方案中，所有 j 做的就是将两个任务委托给 $(j-1)$ 并移动一个磁盘。j 的权限下的总转移次数是一次加上 $(j-1)$ 的权限下的两倍。移动一个磁盘需要一次转移；因此移动两个磁盘的堆叠需要 $1+2{\cdot}1=3$ 次转移，移动三个磁盘的堆叠需要 $1+2\cdot3=7$ 次转移，移动四个磁盘的堆叠需要 $1+2\cdot7=15$ 次转移，以此类推。猜测移动 n 个磁盘堆叠所需的转移次数的公式，并通过数学归纳法证明这个猜测是正确的，留作练习。

应清楚为什么递归定义如此有用；大量复杂性可以编码在少量递归规则中。解决汉诺塔问题中的每个人只需要知道两件简单的事情，但通过协调的任务委派，他们解决了复杂的问题。然而，每增加一个磁盘，所需的转移次数实际上会翻倍。假设每秒可以移动一个磁盘。移动两个磁盘至少需要2秒，移动三个磁盘至少需要4秒，依此类推（这是一个下限，不是确切的数量）。移动7个磁盘的堆叠将需要超过一分钟。如果没有错误，13个磁盘的堆叠将需要大约一个小时，20个磁盘的堆叠将需要大约一周，35个磁盘的堆叠远远超过一个人的寿命，而60个磁盘的堆叠——如果每秒转移一次——将需要比宇宙存在的时间长得多。传说中的婆罗门祭司将不会

在地球被熟知的宇宙现象摧毁之前完成任务。□

加法性质

作为数学归纳法的最终应用，让我们看看加法的性质是如何从自然数的公理中得出的。这些论证相对复杂；它们除了使用归纳法之外，没有使用任何其他方法，但归纳陈述的选择有时很微妙。至少应该浏览一遍，尽管在第一次阅读时可以跳过定理2.6的证明。

定理2.6. 加法是结合律和交换律；也就是说，如果m，n和ℓ是自然数，那么$m + (n + \ell) = (m + n) + \ell$和$n + m = m + n$。

证明。方程（2.2）说明（用不同的字母）如下：

$$(*) \qquad p + (q + 1) = (p + q) + 1 \qquad \text{for all } p \text{ and } q \in \mathbf{N};$$

结合律对于 $\ell = 1$ 已内置于加法的定义中。为了证明一般情况下的结合律，考虑以下陈述

$$A(\ell): \qquad m + (n + \ell) = (m + n) + \ell \qquad \text{for all } m \text{ and } n \in \mathbf{N}.$$

基本情形由$(*)$给出。假设对于某些$k > 1$，$A(k)$成立；然后对于m和$n \in \mathbf{N}$的每个选择，

$$\begin{aligned} m + \big(n + (k + 1)\big) &= m + \big((n + k) + 1\big) && (*)\text{: } p = n,\ q = k \\ &= \big(m + (n + k)\big) + 1 && (*)\text{: } p = m,\ q = n + k \\ &= \big((m + n) + k\big) + 1 && \text{by } A(k) \\ &= (m + n) + (k + 1) && (*)\text{: } p = m + n,\ q = k \end{aligned}$$

因此，$A(k)$意味着$A(k+1)$；因为$A(1)$为真，定理2.2表明对于所有$\ell \in \mathbf{N}$，$A(\ell)$为真，即加法是结合的。

交换性通过两次归纳法证明；首先通过归纳在 n) 上证明 $n + 1 = 1 + n$ 对所有 $n \in \mathbf{N}$ (成立，然后证明 $n+m = m+n$ 对所有 m, $n \in \mathbf{N}$ (通过在 m) 上的归纳。结合性被多次使用。考虑以下陈述

$$P(n): \qquad n + 1 = 1 + n.$$

基本情形 $P(1)$ 表示 1 的后继是 1 的后继（或 $1+1=1+1$），这显然是正确的。现在假设 $P(k)$ 对于某些 $k \in \mathbf{N}$ 是正确的；我们希望证明 $P(k+1)$。但是

$$\begin{aligned}(k+1)+1 &= (1+k)+1 && \text{inductive hypothesis } P(k)\\ &= 1+(k+1) && \text{by associativity.}\end{aligned}$$

这证明了 $P(k+1)$，因此通过归纳 $P(n)$ 对所有 $n \in \mathbf{N}$ 都成立。现在考虑以下陈述

$$C(m): \qquad n+m=m+n \qquad \text{for all } n \in \mathbf{N}.$$

无限断言序列 $\{P(n) : n \in \mathbf{N}\}$ 正好是基本情形 $C(1)$。假设对于某个自然数 $C(k)$，k 为真，即对于所有 $n+k=k+n$，$n \in \mathbf{N}$， 对于所有 $n \in \mathbf{N}$,

$$\begin{aligned}n+(k+1) &= (n+k)+1 && \text{associativity}\\ &= (k+n)+1 && \text{inductive hypothesis, } C(k)\\ &= k+(n+1) && \text{associativity}\\ &= k+(1+n) && \text{by } P(n) \text{ above}\\ &= (k+1)+n, && \text{associativity}\end{aligned}$$

即，$C(k+1)$ 为真。通过归纳，对于所有 m，$C(m)$ 为真，因此加法是交换律的。 □

大量工作仅仅是为了将小学算术建立在集合论的基础上，尽管开发出的工具——递归定义和数学归纳法——是值得努力的。观察发现，“2”、“4”、“+”和“=”这些符号已经用集合来定义，并且定理“2+2 = 4”已经基本得到证明。

乘法和指数运算

正如自然数的加法被定义为迭代序列，乘法定义为迭代加法：$2 \times 3 = 2+2+2$（“将2加到自己身上”3次），例如。精确地，对于所有 $m \in \mathbf{N}$ 定义 $m \times 0 = 0$，然后对于 $n \geq 0$ 定义

$$m \times (n+1) = (m \times n) + m \qquad \text{for all } m \in \mathbf{N}. \tag{2.4}$$

注意，分配律已内置到乘法的递归规范中。一个很好的练习是模仿定理2.6的证明，展示乘法是结合的和交换的，并且对加法进行分配。

进一步，指数是迭代乘法：$2^3 = 2 \times 2 \times 2$ (‘将2自乘’3次)。精确地说，我们设所有 $m > 0$ 的 $m^0 = 1$，然后对于 $n \in \mathbf{N}$，定义，

$$m^{n+1} = (m^n) \times m \qquad \text{for } m > 0. \tag{2.5}$$

(我们做出特殊定义 $0^n = 0$ 对于所有 $n > 0$；表达式 0^0 未定义。) 你应该递归展开这些定义，并将算法作为练习写成伪代码。注意，当用后继表示法表达时，指数运算极其复杂。

加法和乘法是交换律和结合律的，怀疑这些操作通过连续迭代得到的操作也是真的并不无道理！然而，指数运算既不是交换律也不是结合律！事实上，除了平凡情况 $m = n$ 之外，只有一对自然数（2和4）的指数运算是交换的。思考为什么 $2^4 = 4^2$，以及为什么这些数字在这方面是异常的，是很有趣的。

关系

设 X 为一个集合。回忆 $X \times X$ 是 X 元素的有序对的集合。

定义2.7 在X上的关系是$R \subset X \times X$中从X中选取的有序对集合。如果x和y是X的元素，那么我们说“x通过R与y相关”或“xRy”，如果(x, y) $\in R$。

在th 以下示例中，实心圆表示元素 s of R。

示例2.8 设$X = \mathbf{N}$为自然数集，设$R \subset \mathbf{N} \times \mathbf{N}$为包含$m, n$的有序对集，见图2.4。在这种情况下，$m$与$n$相关当且仅当$m < n$。□

示例2.9 再次令 $X = \mathbf{N}$，令 R 为满足 $m + n$ 为偶数的对 (m, n) 的集合，图2.5。在这种情况下，当 m 和 n 具有相同的奇偶性，即都是偶数或都是奇数时，m 与 n 相关。□

[3]If generalization comes to mind, you are ready for the *Ackerman function*!

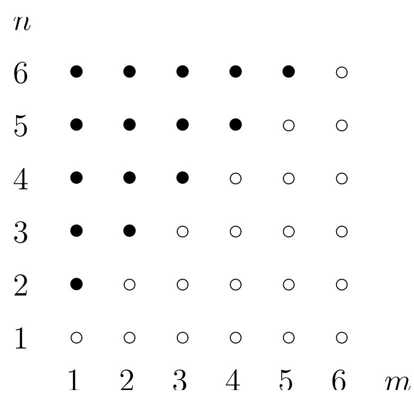


Figure 2.4: Less than.

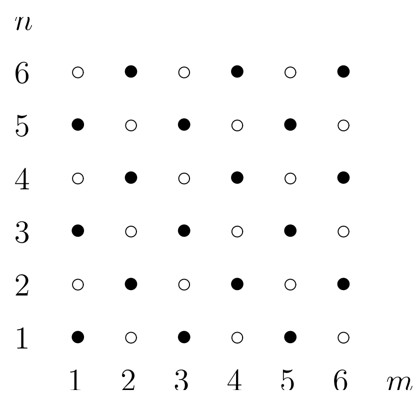


Figure 2.5: Parity.

示例2.10 设X为一个任意集合。等价关系由子集$\Delta = \{(x, x) : x \in X\}$定义，如图2.6。当且仅当它们相等时，$X$中的两个元素通过$\Delta$相关联。子集$\Delta$通常被称为$X \times X$的对角线。□

示例2.11 设X是一个包含多于一个元素的集合。不等关系由对角线的补集定义，如图2.7所示，即由$R = (X \times X) \setminus \Delta = \{(x, y) \in X \times X : x \neq y\}$定义。
□

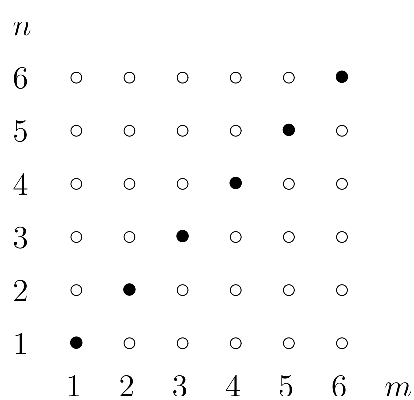


Figure 2.6: Equality.

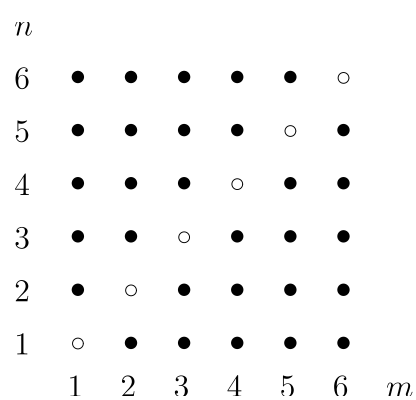


Figure 2.7: Inequality.

定义2.12 设X为一个集合。X上的一个等价关系是一个关系，通常表示为$\sim$，它满足

- (自反性) 对于所有 $x \in X$，$x \sim x$。用话来说，每个元素都与自身相关。
- (对称) 对于所有 x 和 $y \in X$，$x \sim y$ 当且仅当 $y \sim x$。粗略地说，这种关系只看到 x 和 y 是否相关，而不会区分元素对的其他方面。
- (传递性) 对于所有 x，y 和 $z \in X$，如果 $x \sim y$ 和 $y \sim z$，则 $x \sim z$。直观上，这种关系是包罗万象的；与 x 相关的一切都与 x 相关。

如果 $\sim$ 是集合 X 上的一个等价关系，那么 X 可以被“划分”为不相交的子集，称为等价类，每个等价类由通过 $\sim$ 互相相关联的元素组成。$x \in X$ 的等价类定义为

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

等价类集合表示为 $X/\sim$，读作 X 模等价。直观上，等价关系“对”某些区别“视而不见”；它不能区分单个等价类中的元素。**N** 上的“奇偶性”关系是一个等价关系，有两个等价类：[0] = {偶数} 和 [1] = {奇数}。这些类可以（无限地）以许多方式写出，例如，[1] = [3] = [329]。

备注2.13 “小于”关系是传递的，但既不是自反的也不是对称的。“不等式”关系是对称的，但既不是自反的也不是传递的（如果 $x \neq y$ 和 $y \neq z$，一般并不意味着 $x \neq z$）。 在非空集合上的“空”关系（$R = \varnothing$；没有任何东西与任何其他东西相关）是对称的和传递的（因为假设是空洞的），但不是自反的。

存在一个巧妙（但错误！）的论点，即对称且传递的关系必须是自反的：如果 x 通过一个对称关系与 y 相关，那么（按照这个论点）利用传递性属性取 $z = x$ 显示 x 与 x 相关。正如“空关系”所示，错误在于假设每个元素实际上都与某物相关。□

2.2 整数

自然数具有不对称性；虽然每个自然数都有一个后继数，但并非每个自然数都有一个前驱数。因此，像 $2 + x = 1$ 这样的方程在自然数集合中没有解。

自然数。为了克服这种不足，我们仅使用自然数和集合论运算构建了一个更大的“数字”集合，其中包含**N**的一个副本，并且方程 $n_1 + x = n_2$（当n_1和n_2是更一般类型的数字时）总有解。这个更大的数字集合是整数集合**Z**。[4]

在以下讨论中，我们谈论的对象是我们希望定义的，就好像它们已经存在一样。这并不逻辑上循环，因为我们只寻求激发最终的定义，而不是证明任何事情。

我们从方程 $n_1 + x = n_2$ 中获得启示：给定一个自然数有序对 (n_1, n_2)，其中 $n_1 \leq n_2$，存在一个唯一的自然数 x，满足 $n_1 + x = n_2$。假设我们“定义”一个整数为一个自然数的有序对，其想法是，有序对 $x = (n_1, n_2)$ 对应于 $n_1 + x = n_2$ 的解。这似乎很有希望，因为负数可以表示为有序对 $n_1 > n_2$；例如，有序对 (2, 1) 对应于 $2 + x = 1$ 的解，即整数 $x = -1$。

小问题是许多不同的对会代表相同的数字；（1, 4）、（6, 9）和（1965, 1968）都对应于数字3。事实上，当$n_2 - n_1 = m_2 - m_1$，即当$n_2 + m_1 = n_1 + m_2$时，两个对（n_1, n_2）和（m_1, m_2）恰好代表相同的数字。因此，我们定义了这种关系

$$(m_1, m_2) \sim (n_1, n_2) \text{ iff } n_2 + m_1 = n_1 + m_2. \tag{2.6}$$

这个方程定义了集合 $X = \mathbf{N} \times \mathbf{N}$ 上的一个等价关系，正如你可以检查的那样，并且只使用了自然数集合和加法运算。我们得到了关于集合的整数“实现”：

定义2.14 一个整数是关于关系（2.6）的等价类 $\mathbf{N} \times \mathbf{N}$。

例如，使用粗体数字以通常的方式表示整数，

$$\mathbf{2} = [(0, 2)] = [(5, 7)], \quad -\mathbf{2} = [(2, 0)] = [(7, 5)], \quad \mathbf{0} = [(0, 0)] = [(4, 4)].$$

在这个构造中，一个单独的整数由无限多对自然数组成！等价类 $\big[(0, n)\big] \in X$ 对应于自然数 n，因此我们成功在 **Z** 内部构建了 **N** 的副本。

[4]The abbreviation comes from German, probably from *Zahl* (for “number”) or *Zyklus* (for “cycle”, a reference to abstract algebra).

它还需要定义整数的加法和乘法。将每个整数（等价类）想象成一个包含无限多张纸条的帽子，每张纸条上写着有序的自然数对。我们知道，如果从单个帽子中抽取两张纸条（n_1, n_2）和（n'_1, n'_2），那么$n_2 + n'_1 = n'_2 + n_1$。要加（$\oplus$）或乘（$\odot$）两个整数，从相应的帽子中各取一张纸条；比如说，纸条是（m_1, m_2）和（n_1, n_2）。这些纸条

$$\begin{aligned}(m_1, m_2) \oplus (n_1, n_2) &:= (m_1 + n_1, m_2 + n_2) \\ (m_1, m_2) \odot (n_1, n_2) &:= (m_1 n_2 + n_1 m_2, m_1 n_1 + m_2 n_2)\end{aligned} \tag{†}$$

每个都确定一个等价类，这些将被称为两个原始纸条的和与积。（例如，将（$4, 2$）和（$1, 4$）代入；这些对对应哪些整数？根据这些公式，它们的“和”和“积”是什么？）这个“定义”是临时的，因为并不立即清楚“和”和“积”会因从同一顶帽子中抽取的纸条不同而相同。声称的是，虽然和与积纸条上的数字取决于原始纸条，但由(†)确定的等价类无论选择如何都不会改变。详细写出证明将确保你理解这些想法。

我们应该检查当一对整数对应自然数时，算术法则“以相同的方式”工作，就像自然数的加法和乘法规则一样。为此，将（$0, n$）和（$0, m$）代入（†）中:

$$\begin{aligned}(0, m) \oplus (0, n) &:= (0, m + n) \\ (0, m) \odot (0, n) &:= (0, mn)\end{aligned}$$

确保你理解为什么这些方程表明“将自然数视为整数时的算术运算与自然数的算术运算一样”。做出这个观察后，在加法或乘法运算整数时使用普通符号“$+$”和“$\cdot$”是安全的。

更令人启迪的是我们发现(†)的过程。方法是写下我们希望新运算具有的性质，然后以我们定义的整数来表示这些运算。因为我们不是在证明任何事情，所以不需要严谨。滑移(m_1, m_2)代表通常称为$m_2 - m_1$的整数，对于(n_1, n_2)也是如此。$m_2 - m_1$和$n_2 - n_1$的和与积，用普通符号表示，是

$$m_2 + n_2 - (m_1 + n_1) \quad \text{and} \quad m_1 n_1 + m_2 n_2 - (m_1 n_2 + n_1 m_2).$$

翻译成对的语言给出（†）。观察可知，规则（†）显然是交换的，并且+作为整数运算符是结合的；乘法也是结合的，但证明需要简短的计算。

负数在中世纪哲学家中被视为无意义，这种观点在现代笑话中被讽刺性地表达：

> 一个物理学家、一个生物学家和一个数学家看到两个人走进房子。后来三个人出来了。物理学家想，“最初的观察有误。”生物学家想，“他们已经繁殖了。”数学家想，“如果正好有一个人进入房子，它将再次空无一人。”

当然，问题不在于负数没有意义（或者更糟，与集合论相矛盾！），而在于它们不能被随意用来模拟现实世界的情况。运用常识，我们知道房子并不是一开始就是空的。

整数性质

我们已将整数“实现”为自然数对的等价类，但主要对其抽象的算术和顺序性质感兴趣，现在我们将转向这些性质。整数集 $\mathbf{Z}$ 与加法运算一起构成一个称为（交换）群的数学结构。为了将来使用，下面给出了一个抽象定义。定义2.15形式化了两个整数的和是整数，并且加法遵循某些性质的想法。

一个基本概念是在集合 G 上的二元运算。松散地说，G 上的二元运算 $\oplus$ 是一种“组合”集合 G 中的两个元素 x 和 y 以得到一个元素 $x \oplus y$ 的方法。两个整数的普通和是在 $\mathbf{Z}$ 上的二元运算，普通乘积也是如此。

定义2.15 设G为一个集合，设$\oplus$为G上的二元运算。如果以下条件成立，则对$(G, \oplus)$是一个交换群：

A.1（结合性）$(x \oplus y) \oplus z = x \oplus (y \oplus z)$ 对于 G 中的所有 x，y 和 z。

A.2（中性元素）存在一个$e \in G$，使得对于所有$x \in G$，有 $x \oplus e = e \oplus x = x$。

A.3（逆元的存在）对于每个 $x \in G$，存在一个元素 y 使得 $x \oplus y = e$ 。

A.4（交换律）对于所有在 G 中的 x，y，有 $x \oplus y = y \oplus x$。

我们的直接兴趣是整数集 $G = \mathbf{Z}$ 的情形，其中 $\oplus = +$ 是普通加法。然而，交换群的概念在其他情境中也会出现，而非常规的“$\oplus$”是为了提醒您，在一般情况下，群运算不必是字面上的加法。

要讨论整数的加法群，我们切换到更传统的“$+$”符号表示二元运算，并将“e”写作“0”。通常用“$-n$”表示n的加法逆元，但在此情况下，后者必须被视为一个单独的符号，而不是n和-1的乘积！减法运算仅仅是加法逆元的加法；然而，请注意，减法既不具有结合性也不具有交换性！

整数不由公理A1–A.4定义；也就是说，存在交换群在抽象上不与整数加法群等价。对这些问题的进一步研究属于抽象代数课程的内容。

公理 A.1–A.4 有一些简单但有用的结果。其中两个在此被证明，既为了说明，也为了强调它们并非定义的明确部分。

定理2.16. 在$\mathbf{Z}$中的中性元素是唯一的，且$\mathbf{Z}$中的每个元素都有一个唯一的加法逆元。

证明。假设0和$0'$是中性元素。因为0是中性元素，所以A.2公理与$a = 0'$结合意味着$0' = 0 + 0'$。交换0和$0'$的角色意味着$0 = 0 + 0'$；因此，$0 = 0'$如所声称。

现在假设 m 和 m' 是 n 的逆，即 $n+m = n+m' = 0$。那么

$$m = m + (n + m') = (m + n) + m' = m'$$

根据公理A.2，A.1和A.2分别。 □

公理系统，如A.1–A.4，扮演两个角色。它们可以被直接假设并作为抽象理论（在这种情况下为交换群）的起点。或者，公理可以被视为在某种构造的背景下必须建立的定理。后者观点用于证明公理系统的一致性。（未经证明且无法证明的假设是，集合论本身在逻辑上是一致的。）从现在起，我们将主要采用前者观点。幕后是理性、实数和复数系统的连续构造。

2.3 有理数

直观上，一个有理数 {v*} 是一个“当它自身加到自身 q 次时”得到的量。或者，p/q 由 p 个“货币”$1/q$ 的单位组成，其中 q 个单位加起来等于1。如果 k 是一个正整数，那么 p/q 和 kp/kq 表示相同的数，因为货币的量减少了 k 倍，但单位数量增加了 k 倍。逻辑一致性要求 $kp/kq = p/q$ 即使 k 是负数。

每个有理数都可以用整数的有穷比表示，有无限多种方式；例如，$1/2 = 5/10 = (-3)/(-6)$。如果一个比 p/q 的分子和分母都是0，并且它们没有大于1的公因数，那么这个比就是最简比。一个整数被视为分母为1的有理数。每个有理数都有唯一的最简比表示。

有理数对乘法的作用与整数对加法的作用相同：它们允许解任意方程 $qx = p$ 对于 $q \neq 0$，其中 p 和 q 是整数，并且作为额外的好处，当 p 和 q 是有理数（且 $q \neq 0$）时，这样的方程也能自动求解。如果 $q = 0$，那么表达式“$p/0$”应该表示方程 $0x = p$ 的解。然而，当 $p \neq 0$ 时，这个方程没有解，而当 $p = 0$ 时，每个有理数都是解。因此，我们不做尝试将商 $p/0$ 定义为有理数。非正式地说，除以零是未定义的。

如果要将两个有理数——比如说2/5和5/12——相加，它们必须“转换”成相同的货币。这通过交叉相乘来完成，在这种情况下，将给定的数字表示为24/60和25/60。现在可以按明显的方式相加（或相减）。这种推理是以下通用公式的背后：

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 q_2 + p_2 q_1}{q_1 q_2}, \tag{2.7}$$

这将用于激励从 $\mathbf{Z}$ 构建到 $\mathbf{Q}$ 的方程 (†) 的类似物。

如果将整数视为长度（数轴上某个单位长度的倍数），那么有理数也可以视为长度；可以使用直尺和圆规构造任意有理长度。了解有理数在数轴上的位置方式是很重要的。为此，固定一个正整数 q，并考虑整数倍 $\frac{1}{q}\mathbf{Z}$ 的集合，即形式为 p/q 的有理数集合，如图2.8所示。

集合 $\frac{1}{q}\mathbf{Z}$ 是 $\mathbf{Z}$ 的“缩放副本”，间距为 $1/q$，所有这些集合的并集是 $\mathbf{Q}$，参见图 2.9。

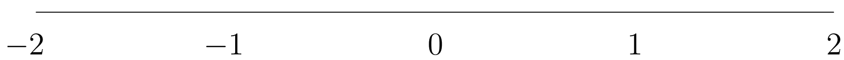


图2.8：分母为q的有理数集合。

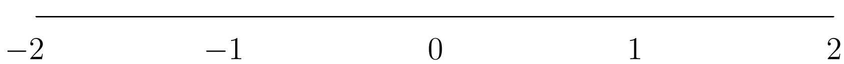


图2.9：分母不超过q的有理数集合。

不清楚数轴上的每个点是否都由有理数表示，但图2.8和2.9应该清楚地表明，线上的每个点都“任意接近”一个有理数。如果我们把古希腊关于数轴上的点作为距离的想法，那么可以合理地认为每个距离都是整数比。然而，尽管毕达哥拉斯学派将他们的哲学建立在整数比和谐性的理念上，但他们最终发现——据传说，他们相当沮丧——单位正方形的对角线不能表示为整数比，参见定理1.3。如果要将距离建模为数字，那么有理数集是不够的。

很遗憾，我们不清楚如何描述或量化“有理数之间的差距”，尤其是如果我们不能对外部参考数轴（毕竟，我们还没有在集合论术语中定义它）进行参考。R. Dedekind 和 G. Cantor 在 1872 年独立地、通过完全不同的方法找到了合适的描述。为了给出他们构造的指示，我们首先必须更详细地研究有理数集 $\mathbf{Q}$。

算术和顺序性质 $\mathbf{Q}$

有理数集与加法运算（方程（2.7））构成一个交换群（参见上述A.1–A.4）：加法是交换的和结合的，存在加法的单位元，每个元素都有一个加法逆元。然而，群$(\mathbf{Q},+)$比群$(\mathbf{Z},+)$更复杂。具体来说，$\mathbf{Z}$是

“由”元素1生成，在意义上，**Z**中的每个元素都是通过反复将1（或其加法逆元）加到自己身上得到的。相比之下，方程（2.7）表明，对于$(\mathbf{Q},+)$不存在有限个生成元：如果a_1，$\cdots$，a_n是有理数，且$a_i = p_i/q_i$为最简形式，那么通过求和和差不能生成分母大于$N = q_1 \cdots q_n$的数，$N = q_1 \cdots q_n$是各个分母的乘积。因此，由a_i生成的有理数集的分母是有界的，因此不能是所有的**Q**，见图2.9。

有理数集还有一个算术运算，乘法，定义为

$$\frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1 p_2}{q_1 q_2}. \tag{2.8}$$

设 $\mathbf{Q}^\times$ 表示非零有理数集。乘法将 $\mathbf{Q}^\times$ 中的两个元素结合成一个第三元素；因子的顺序不影响结果，而数字 1 起到中性元素的作用。最后，如果 $r = p/q$ 是一个非零有理数，那么 $r^{-1} = q/p$ 也是一个非零有理数，显然是 r 的乘法逆元。带有乘法运算的集合 $\mathbf{Q}^\times$ 满足 A.1–A.4：$(\mathbf{Q}^\times, \cdot)$ 是一个交换群。

如果 **Q** 只被赋予了两个独立的群运算，那么有理数算术将相对无趣。然而，这些运算通过分配律相互联系

$$a \cdot (b + c) = a \cdot b + a \cdot c \qquad \text{for all } a,\ b, \text{ and } c \in \mathbf{Q}.$$

类似地，通过乘法的交换律，右乘的类似情况成立。（注意加法不分配到乘法上的不对称性。）

许多初等算术可以由 **Q** 中的加法群公理、$\mathbf{Q}^\times$ 中的乘法群公理和分配律推导出来。这里给出几个例子来说明如何使用公理来证明更熟悉的性质。下面的第二个断言乍一看似乎是同义反复的，它断言 a 的加法逆元等于 a 和 -1 的乘积。（如果不是这样，那么“$-a$”这个符号就会极其误导！）

定理2.17。如果 $a \in \mathbf{Q}$，那么 $a \cdot 0 = 0$ 和 $-a = a \cdot (-1)$。

证明。第一个断言通过注意到$0+0 = 0$（将中立元素的定义应用于$a = 0$），因此将两边乘以a

并且使用分配律得到 $a \cdot 0 + a \cdot 0 = a \cdot 0$。加上 $a \cdot 0$ 的逆元意味着 $a \cdot 0 = 0$。

根据定理2.16，$-a$是满足$a + (-a) = 0$的唯一有理数。如果我们能证明$a \cdot (-1)$具有这个性质，那么定理就得到了证明。但是我们有

$$\begin{aligned} a + \big(a \cdot (-1)\big) &= \big(a \cdot 1\big) + \big(a \cdot (-1)\big) && \text{Definition of neutral element} \\ &= a \cdot \big(1 + (-1)\big) && \text{Distributive Law} \\ &= a \cdot 0 && \text{Definition of additive inverse,} \end{aligned}$$

并且这等于0，根据定理的第一部分。 □

抽象结构类似于 $\mathbf{Q}$ 的结构频繁出现，足以获得一个特殊的名称：

定义2.18 一个域($\mathbf{F}, +, \cdot$)由一个非空集合($\mathbf{F}$)和两个结合运算($+$)和($\cdot$)组成，使得

F.1 ($\mathbf{F}, +$) 是一个交换群，具有中性元素 $\mathbf{0}$。

F.2 集合 $\mathbf{F}^\times = \mathbf{F} \setminus \{\mathbf{0}\}$ 是集合 $\mathbf{F}$ 中非零元素的交换群，乘法下的单位元素为 $\mathbf{1} \neq \mathbf{0}$。

F.3 乘法 $\cdot$ 在加法 $+$ 上分配。

根据习惯的懒惰，数学家们在操作被理解时说“域$\mathbf{F}$”。观察($\mathbf{Z}, +, \cdot$)不是一个域；非零整数2没有乘法逆元。一般域的研究属于抽象代数，这里不详细探讨。给出公理是为了展示抽象的概念经济：像定理2.17这样的结果，只需用域的公理就能证明，对任意域都成立。

有限域

令人惊讶的是，存在有限元素的域。这些域出现在数学的各个部分，以及“公钥加密”中。当你查看以https://开头的“安全网站”时，你的网络浏览器使用有限域算术来安全地发送和接收数据。这里的“位数”（本文中通常为128）与域的元素数量相关。域越大，就越

加密难度越大，没有解码密钥就越难破解，但数据传输速度越慢。

示例2.19 一个域必须包含至少两个元素，因为我们要求 $\mathbf{0} \neq \mathbf{1}$。在有限域中，$\mathbf{1}$ 的连续求和不可能全部不同，因此必须存在一个最小的正整数 p，使得

$$\mathbf{1} + \cdots + \mathbf{1} = \mathbf{0} \qquad (p \text{ summands}).$$

字段公理意味着 p 是一个素数；例如，p 不能是 6，因为那样 $\mathbf{1}+\mathbf{1}$ 和 $\mathbf{1}+\mathbf{1}+\mathbf{1}$ 将是非零元素，它们的乘积是 $\mathbf{0}$。进一步的考虑表明，有限域中的元素数量——它的阶数——必须是 p 的幂。可以构造阶数为 p^k 的域，从而证明它们的存在。因此存在阶数为 2、3、$4 = 2^2$、5、7、$8 = 2^3$、$9 = 3^2$、11 等域；并且不存在阶数为 $6 = 2 \cdot 3$、$10 = 2 \cdot 5$、$12 = 2^2 \cdot 3$ 等域。

素域容易描述；对于{v*}5，例如，取{v*}0{v*}1{v*}2{v*}3{v*}4{v*}除以5的余数集合{v*}，并定义运算为通常的运算，除了和或积除以5并取余数。因此，在{v*}中，2 {v*} 3 {v*} 0（普通和为5，除以5余0）和2 {v*} 3 {v*} 1（积为6，除以5余1）。有限域中的算术看起来可能很奇怪。上面给出的两个例子可以合法地写成{v*}2 {v*} 3（或{v*}3 {v*} 2）和1{v*}2 {v*} 3（或1{v*}3 {v*} 2）。当然，符号2和3代表{v*}的元素，而不是有理数。写成{v*}[2] {v*} [3]和[2]{v*} [3]可能不那么挑衅。

一个包含$4 = 2^2$ 个元素的域（更一般地，具有 p^k 个元素，其中 p 是素数且 $k >$ 为1）在此未讨论。一个域可能具有无限多个元素，但仍能满足有限个加数时的 $\mathbf{1}+\cdots+\mathbf{1} = \mathbf{0}$。□

有限几何级数

本节介绍了一种在任意域中可以进行的实用计算，并提供了数学归纳法的一个良好示例。如果您愿意，可以将下面的域视为 $\mathbf{R}$。

示例2.20 设a和r是域中的元素，并设$n \in \mathbf{N}$。表达式

$$(2.9) \qquad \sum_{j=0}^{n} a\,r^j = a + a\,r + a\,r^2 + a\,r^3 + \cdots + a\,r^n$$

几何级数，其特征是相邻两项的比值相同（即 r）。如果 $r = 1$，则该级数由 $n + 1$ 项组成，每项等于 a，因此和为 $a(n + 1)$。（在一般域中，即使 $a \neq 0$，这也可能为 0！）当 $r \neq 1$ 时，可以通过将和乘以 $1 - r$ 并“缩并”项来猜测几何级数的封闭形式:

$$\begin{aligned}
&(1-r)(a + a\,r + a\,r^2 + a\,r^3 + \cdots + a\,r^n) \\
&= (a + a\,r + a\,r^2 + a\,r^3 + \cdots + a\,r^n) \\
&\qquad - \ (ar + a\,r^2 + a\,r^3 + \cdots + a\,r^n + a\,r^{n+1}) \\
&= a - ar^{n+1} = a\,(1 - r^{n+1}).
\end{aligned}$$

这个论点，就像许多包含省略号或“等等”一词的论点一样，可以通过数学归纳法来精确化。（就像奇数之和一样，非正式的论点需要猜测答案！）考虑以下陈述

$$P(n): \qquad (1-r)\sum_{j=0}^{n} a\,r^j = a\,(1 - r^{n+1}).$$

当 n = 为 0 时，这是恒等式 $a\,(1 - r) = a\,(1 - r)$。假设通过归纳，$P(k)$ 对某些 $k \in \mathbf{N}$ 是真的。通过在等式 $P(k)$ 的两边添加与 $j = k + 1$ 对应的项 $a\,r^{k+1}(1 - r)$ (并使用代数，推导出 $P(k + 1)$ 的真值。这留作练习。

当 r $= 1$ 时，$P(n)$ 仅断言 $0 = 0$，但上述非正式论证（可以通过归纳证明来补充）评估了总和。因此

$$\sum_{j=0}^{n} a\,r^j = \begin{cases} a\,\dfrac{1 - r^{n+1}}{1 - r} & \text{if } r \neq 1, \\[2ex] a\,(n+1) & \text{if } r = 1. \end{cases} \tag{2.10}$$

引人注目地，该方程的左侧由一个单一的和定义，而闭式形式需要两个情况；在 $r = 1$ 处的特殊定义恰好发生在 $r \neq 1$ 的表达式变为 $0/0$ 时。我们找到了一个可以赋予“0/0”表达式意义的上下文！ □

有序域和绝对值

有理数集在抽象意义上是"有序"的，参见定义2.21。**Q**上的顺序关系是构建实数集的关键。因为顺序性质是微积分的基本性质，所以我们在一个一般化的设置中引入它们。

定义2.21 有序域是一个域$(\mathbf{F},+,\cdot)$，它与满足以下条件的子集$P \subset \mathbf{F}$一起。

O.1（三分法）对于每个 $a \in \mathbf{F}$，以下三个命题中恰好有一个成立：$a \in P$，$-a \in P$，或 $a = 0$。O.2（在 $+$ 下的封闭性）如果 a，$b \in P$，那么 $a + b \in P$。O.3（在 $\cdot$ 下的封闭性）如果 a，$b \in P$，那么 $a \cdot b \in P$。

一个P的元素是正数。如果$-a \in P$，那么a是负数。

非正式地说，O.1 表示每个数要么是正数，要么是负数，要么是零。O.2 和 O.3 表示正数的和与积都是正数。与正数集 P 相关的顺序关系 $<$ 定义为

$$x < y \quad \text{iff} \quad y - x \in P.$$

一个域可能或可能不承认一个序关系。一个有限域从不承认序操作（为什么？），也不承认在-1有平方根的域。为了看到后者，首先观察在一个有序域中，如果$-a$和$-b$是负的（因此a和b是正的），那么定理2.17意味着

$$(-a)(-b) = (a)(-1)(b)(-1) = (-1)^2(a \cdot b) = a \cdot b,$$

这是正的，O.3。这个简单的代数事实在中世纪哲学家之间引起了激烈的争论。公理化与计算相对于直观论证具有一定的优势。特别是，在一个有序域中，每个非零元素的平方都是正的。这表明，在每一个有序域中，$\mathbf{1} = \mathbf{1}^2$是正的，$-\mathbf{1}$是负的。因此，正如所声称的，如果$-1$在$\mathbf{F}$中有平方根，那么$\mathbf{F}$不承认一个顺序关系。

命题2.22. 在有理域中，存在一个唯一的集合 $P \subset \mathbf{Q}$ 满足 O.1–O.3。

证明。假设 $P \subset \mathbf{Q}$ 满足 O.1–O.3。如上所示，$1 \in P$。因此，每个自然数 $q \neq 0$ 都在 P 中，根据 O.2。倒数 $1/q$ 也是正的，否则方程 $q \cdot (-1/q) = -1$ 将与 O.3 矛盾。最后，如果 p 和 q 是非零自然数，那么 $p/q = p \cdot (1/q)$ 是正的。我们声称

(2.11) $P = \{p/q \in \mathbf{Q} \mid p, q \text{ 互质正整数}\}$

我们已经证明，如果 $P \subset \mathbf{Q}$ 满足 O.1–O.3，那么 P 就是 (2.11) 中的集合，因此 P 的选择至多只有一个。但对于这个 P 的选择，O.1–O.3 都是明确的；参见练习 2.11。 $\square$

让 x 和 y 是有序域的元素。不等式 $x < y$ (“x 小于 y”) 也写作 $y > x$ (“y 大于 x”)。用“$x \leq y$”代替“$x < y$ 或 $x = y$”是方便的。根据三歧性，$x \leq y$ 是 $y < x$ 的否定。

传统上，较大的数字在数轴上越往右。图2.9清楚地表明，尽管有理数是有序的，有理数并不是像珠子一样沿着数轴排列；对于每个有理数x，都没有“下一个”有理数。特别是，没有最小的正有理数。

一个域 $\mathbf{F}$ 上的序关系 $<$ 确定了集合 $P = \{x \in \mathbf{F} \mid 0 < x\}$。因此，有序域可以表示为 $(\mathbf{F}, +, \cdot, P)$ 或 $(\mathbf{F}, +, \cdot, <)$。这种有些笨拙的记法强调了有序域的所有结构：一组元素，两个二元运算（受公理约束），以及一组“正”元素（受进一步公理约束）。总的来说，我们为有序域列出了十二条公理：四个加法群公理，四个乘法公理，一个分配律公理，以及三个序公理。

The order axioms O.1–O.3 are low-level instructions. In practice, we want to manipulate inequalities as fluently as equalities. Theorem 2.23 successively gives rules for adding and multiplying inequalities by fixed numbers, for taking reciprocals of an inequality, and for adding and multiplying two inequalities (of non-negative numbers). Each of these properties is deduced from the order axioms by one of a few standard arguments. This list is not all-encompassing, but the proof illustrates the most important ideas, and other similar properties should provide easy exercises.

定理2.23。设a，b，c和d为有序域中的元素。

(i) (v47 的传递性) 如果 $a < b$ 和 $b < c$，那么 $a < c$。

(ii) 如果 $a < b$，那么 $a + c < b + c$ 和 $-a > -b$。

(iii) 如果 $a < b$ 和 $0 < c$，那么 $ac < bc$。

(iv) 如果 $c < d < 0 < a < b$，则 $1/d < 1/c < 0 < 1/b < 1/a$。

(如果) $0 < a < b$ 和 $0 < c < d$，那么 $0 < a+c < b+d$ 和 $0 < ac < bd$。

如果将(i)、(ii)、(iii)和(v)中的“$<$”替换为“$\leq$”，则类似的断言成立。

证明。性质(i)是对O.2的重述；$a < b$表示$b - a \in P$，而$b < c$表示$c - b \in P$。根据O.2，$c - a = (c - b) + (b - a) \in P$，因此$a < c$。

(ii) 假设 $a < b$，即 $b - a \in P$。使用域公理，这可以重写为 $(b + c) - (a + c) \in P$，根据定义意味着 $a+c < b+c$。同样，$b-a = (-a)-(-b) \in P$，意味着 $-b < -a$。

要证明(iii)，观察假设 $b - a$ 和 c 在 P 中。根据 O.3，它们的乘积 $(b - a)c = bc - ac$ 在 P 中，这意味着 $ac < bc$。

(iv) 对于倒数，注意如果 $x > 0$，则 $1/x > 0$（因为否则根据 O.3，$-1 = x(-1/x)$ 将是正数）；同样，如果 $y < 0$，则 $1/y < 0$。假设 $0 < a < b$。那么根据 O.3，$0 < ab$，所以 $0 < 1/(ab)$ 如前所述，应用 (iii) 并用 $c = 1/(ab)$ 证明 (iv)。

属性（v）是（i）–（iii）的结果：在假设下，$a + c < b + c < b + d$ 和 $ac < bc < bd$。

最后，如果将 $<$ 替换为 $\leq$，则分别考虑两种情况 $a < b$ 和 $a = b$。在第一种情况下，已给出的证明意味着所期望的结果（例如，因为 $-a > -b$ 显然意味着 $-a \geq -b$）。如果 $a = b$，结论是明显的。 □

设 $(\mathbf{F}, +, \cdot, <)$ 为一个有序域。$a \in \mathbf{F}$ 的绝对值，记为 $|a|$，定义为

$$|a| = \begin{cases} a \text{ if } a \geq 0 \\ -a \text{ if } a < 0 \end{cases} \tag{2.12}$$

显然，对于所有 a，$0 \leq |a| = |-a|$。量 $|a - b| = |b - a|$ 可以解释为 a 和 b 在数轴上的距离，这解释了它在分析中的重要性。

大致上，取一个数的绝对值“如果有的话，会丢弃负号。”然而，请注意，“$|-a| = a$”通常是错误的。在符号计算中，$|a|$可以根据需要替换为a或$-a$，因此可以通过检查足够多的案例来建立关于绝对值的任何断言。这通常很繁琐，因为每个额外的绝对值符号都会使需要检查的案例数量翻倍。

极值和极小值

设$(\mathbf{F}, +, \cdot, <)$为一个有序域，并设$a$，$b \in \mathbf{F}$。$a$和$b$的最大值是这两个数中较大的一个，记作max$(a, b)$。最小值定义类似。绝对值的一个有趣且有用的应用是max和min的一对公式：

定理2.24。设a和b为有序域中的元素。那么

$$\max(a, b) = \frac{a + b + |a - b|}{2}, \qquad \min(a, b) = \frac{a + b - |a - b|}{2}.$$

证明。为了简单起见，写为 $M =$ max(a, b) 和 $m =$ min(a, b)。有两种可能性：$M = a$ 和 $m = b$，或 $M = b$ 和 $m = a$。（如果 $a = b$，则两者都为真。）在任何情况下，$M + m = a + b$：两个数的和就是它们的和！此外，$M - m$ 是非负的，等于 $a - b$ 或等于 $b - a$；因此，$M - m = |a - b|$。求解 M 和 m 证明了该定理。 □

一些绝对值和距离的一般性质收集在定理2.25中。它们将在定义和操作极限时被反复使用。

定理2.25。设a和b为有序域中的元素。那么

(i) $|a| \le b$ 当且仅当 $-b \le a \le b$。

(ii) $|ab| = |a| \cdot |b|$. (iii) $|a + b| \le |a| + |b|$ 和 $|a - b| \le |a| + |b|$. (iv) $|a - b| \ge \big||a| - |b|\big|$.

这些不等式之所以重要，是因为它们对所有 a 和 b 的选择都同时成立，因此代表了有序域的一般性质。

证明。第一个性质在单个绝对值不等式和一对普通不等式之间进行转换。为了证明它，注意 $a \le |a|$ 和 $-a \le |a|$；因此 $|a| \le b$ 当且仅当 $a \le |a| \le b$ 和 $-a \le |a| \le b$，而其中之一等价于 $-b \le a$。将这些结合起来给出 (i)。

属性（ii）和（iii）通过检查情况得到证明。因为如果将 a 和 b 都乘以 -1，每个命题都不变，所以只需假设其中一个数是非负的。如果将 a 和 b 交换，每个命题也不变，因此只需假设 $a \le b$。最后，（iii）的两个部分是等价的，因为用 $-b$ 替换 b 交换了两个断言。总之，只需在两种情况下验证（ii）和（iii）的第二部分：$0 \le a \le b$ 和 $a \le 0 \le b$。在定理2.24的记法中，$a = m$ 和 $b = M \ge 0$。

属性（ii）在每种情况下都很容易验证。 (iii) 的第二部分并不难。如果 $a \le 0 \le b$，那么 $m = a = -|a|$ 和 $M = b = |b|$，所以 $|a-b| = M-m = |a|+|b|$。如果相反 $a = m \ge 0$，那么 $|a-b| = M-m \le M+m = |a|+|b|$。这证明了 (iii)。

性质（iv）可以通过一些值得注意的体操从（iii）中推导出来：

$$|a| = |(a-b)+b| \le |a-b|+|b| \qquad \text{for all } a \text{ and } b,$$

同样 $|b| \le |b-a|+|a| = |a-b|+|a|$。从第一个不等式中减去 $|b|$，从第二个不等式中减去 $|a|$ 得到

$$|a|-|b| \le |a-b| \quad \text{and} \quad |b|-|a| = -\big(|a|-|b|\big) \le |a-b|,$$

并且这与(i)通过(iv)等价。 □

断言（iii）和（iv）通常被称为三角不等式和逆三角不等式，尤其是在写成{v*}

$$|x-z| \le |x-y|+|y-z| \tag{2.13}$$

$$|x-z| \ge \big||x-y|-|y-z|\big| \tag{2.14}$$

对于所有 x，y，z。

第一个正好是(iii)中的$a = x-y$和$b = y-z$，而第二个是(iv)中的 $a = x-y$和$b = z-y$。这些名称来源于将绝对值解释为距离；三角形的边长不大于其他两边长度之和，且至少与（它们的长度之差的绝对值）一样长。

最高

哪个州的最高点最低？ —Moxy Früvous

如果 y 是数轴上的一个点，那么存在“任意接近” y 的有理数，如图2.8所示。这种任意接近逼近的想法可能是微积分的基本思想，并且在消除 $\mathbf{Q}$ 中的“差距”方面是关键。

定义2.26 设$(\mathbf{F}, +, \cdot, <)$为一个有序域。若存在一个$M \in \mathbf{F}$使得对所有$x \in A$有$x \le M$，则子集$A \subset \mathbf{F}$是有上界的。若存在一个$a \in A$使得对所有$x \in A$有$x \le a$，则集合$A$有最大值。通过明显的修改，我们谈论有下界或具有最小值的集合。若集合既有上界又有下界，则称其为有界。

在一个有序域中，具有最大值的子集不仅是有上界的，而且有最大元素；然而，一个集合可以有上界但没有最大值。这有点反直觉，可能是因为有限个数的集合总是有最大元素。负有理数的集合被0有上界，但没有最大的负有理数。显然，如果M是A和$M \le M'$的上界，那么M'也是A的上界；上界不是唯一的。相比之下，一个集合最多只有一个最大值，因为如果a和a'都是A的最大值，那么$a' \le a$ (因为a是最大值)和$a \le a'$ (因为a'是最大值)，所以$a = a'$。

空集是有界的（条件是空洞的），尽管它没有最大值或最小值。每个非空有限集都有最大值和最小值，因此是有界的。有界性、最大值和最小值的概念对于无限集——那些具有无限多个元素的集合——最为重要。请注意，在日常用语中，“无限”通常用于表示“无界”（例如，“宇宙学家不知道宇宙是否无限”）。在数学上，不要混淆这些术语：一个无界集是无限的，[5] 但集合 $\{x \in \mathbf{Q} \mid 0 < x < 1\}$ 是无限的，但也是有界的。请注意，这个集合既没有最小值也没有最大值。

假设 $A \subset \mathbf{F}$ 是一个非空的上界集合。在理论和实践中，都希望找到 A 的“最佳可能上界”。“更好”意味着更小，因为较小的上界传达了更多信息；知道一个人身高不到6英尺比知道他们身高不到7英尺传达的信息更多。

[5] Contrapositively, a finite set is bounded.

考虑集合 A 的上界集。如果存在，A 的“最佳”上界将是上界集的最小值。如果存在，这个最小值（根据上述备注是唯一的）被称为集合 A 的最小上界或上确界，表示为 $\sup A$，见图 2.10。

如果 A 有最大值，那么 $\sup A = \max A$，但一个集合即使没有最大值也可能有上确界。负有理数集合 $\mathbf{Q}^-$ 没有最大值，但对于每个 $M \geq 0$，它被 M 上界限制。因此，0 是 $\mathbf{Q}^-$ 的上界，并且是最小的上界；$\sup \mathbf{Q}^- = 0$。

命题2.27。设$A \subset \mathbf{F}$是一个有上确界的集合，记为$a = \sup A$。那么对于每一个$\varepsilon > 0$，存在一个元素$x \in A$满足$a - \varepsilon < x$。

证明。令 $\varepsilon > 0$。因为 $a = \sup A$ 是 A 的最小上界，所以数字 $a - \varepsilon < a$ 不是 A 的上界。换句话说，存在一个元素 $x \in A$ 满足 $a - \varepsilon < x$。 □

大致上，一个集合的上确界与该集合的某个元素任意接近。这种直观的说法有点误导，因为两个不同的数永远不会“任意接近”。漏洞在于元素 x 通常依赖于 ε 的选择。更准确地说，$\sup A$ 与集合 A 任意接近。如果 $a \in A$ (即，A 包含其上确界)，那么就没有什么要证明的；对于每个 $\varepsilon > 0$，取 $x = a$。对于不包含其上确界的集合，命题 2.27 最有趣。

在数轴上，集合 A 的上界是一个位于 A 右侧的点。将上界向左移动（但保持在 A 右侧）“改善”了上界，而（如果存在）上确界是点不能移动而至少经过 A 中一个点的位置。

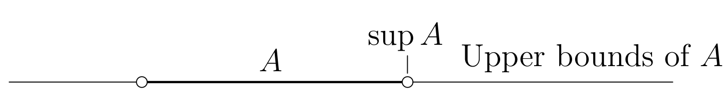


图2.10：上界和上确界。

如果集合 $A \subset \mathbf{F}$ 有上界，那么存在三种逻辑可能性：

- A 具有最大值。
- A 没有最大值，但有上确界。
- A 没有上确界。

给定边界和上确界的视觉描述，可能难以想象一个有上界的集合为何可能没有上确界。学校直觉强调数轴，确实，实数域所拥有的难以捉摸的性质是，如果 A 有上界，那么第三种可能性是不可能的。值得注意的是，一个有理数集合可以有上界但可能没有（有理）上确界：

命题2.28. 设$A = \{x \in \mathbf{Q} \mid x^2 \le 2\} \subset \mathbf{Q}$。集合$A$非空且有上界，但没有上确界。

证明。由于$0 \in A$，A 非空。因为2没有有理平方根，所以A定义中的"$\le$"可以用严格不等式替换。设x为一个正有理数，且$z = 2/x$。根据定理2.23，$x^2 > 2$ 当且仅当 $z^2 < 2$，即$2/x \in A$ 当且仅当 $x \notin A$。为了证明A没有上确界，我们首先证明集合$B = \{x \in \mathbf{Q}^+ \mid x^2 > 2\}$没有最小元素。

引理2.29。设$x = p/q$为一个最简正有理数，且$x^2 >$为2，并设

$$y = \frac{1}{2}\left(x + \frac{2}{x}\right) = \frac{1}{2}\left(\frac{p}{q} + \frac{2q}{p}\right).$$

然后 y 是一个正有理数，具有 $y^2 > 2$ 和 $y < x$。

显然，y 是一个正有理数。为了看出 $y^2 >$^2，写出 $x^2 = (p^2/q^2) =$ ^2 $+ \varepsilon$ (，用 $\varepsilon >$ = 0 通过假设) 并计算

$$\begin{aligned} y^2 &= \frac{1}{4}\left(\frac{p^2}{q^2} + 4 + \frac{4q^2}{p^2}\right) = \frac{1}{4}\left(2 + \varepsilon + 4 + \frac{4}{2+\varepsilon}\right) \\ &= \frac{1}{4}\left(\frac{4 + 4\varepsilon + \varepsilon^2 + 8 + 4\varepsilon + 4}{2+\varepsilon}\right) = 2 + \frac{1}{4} \cdot \frac{\varepsilon^2}{2+\varepsilon} \\ &> 2. \end{aligned}$$

最后，$0 < \varepsilon = (p^2/q^2) - 2$，所以

$$x - y = \frac{1}{2}\left(\frac{p}{q} - \frac{2q}{p}\right) = \frac{1}{2} \cdot \frac{q}{p}\left(\frac{p^2}{q^2} - 2\right) > 0,$$

证明 $y < x$ 如所述。这确立了引理。

设 $-B = \{x \in \mathbf{Q} \mid -x \in B\}$。引理2.29及前面的注释表明 $\mathbf{Q}$ 是三个不相交集的并集：$\mathbf{Q} = -B \cup A \cup B$。因此，有理数 x 是 A 的上界当且仅当 $x \in B$。但引理2.29表明 B 没有最小元素，因此 A 没有上确界。

□

注意，A没有最大值（因为最大值将是上确界）。引理2.29的证明也给出了一个直接证明。您可能会发现详细推导很有帮助。

2.4 实数

极限的概念是“最大值”对于无限集合的正确推广。相信微积分的大部分内容都建立在所有（非空）有上界的集合存在上确界的基础上并不困难。因此，下面的定理2.30是核心。完整的证明对本书的其余部分贡献不大，但基本思想很好地说明了在集合论中构造自然数和整数时讨论的几个想法。最好将定理2.30视为一种许可；通过向有序域的公理中添加一个公理，我们最终获得了一个适合无穷小量微积分的数系。定理2.30断言存在的有序域是实数域。

定理2.30。存在一个有序域$(\mathbf{R}, +, \cdot, <)$具有以下性质：

（完备性）如果 $A \subset \mathbf{R}$ 非空且有上界，那么 $\sup A \in \mathbf{R}$。

此字段在有序域同构下是唯一的，并包含有理数域的副本（$\mathbf{Q}, +, \cdot, <$）。

证明。（简要概述） Dedekind 的想法是将一个单独的实数定义为一定无限的理性数集，他称之为“割”。回顾这个想法是将实数与所有严格小于它的理性数集相关联。为了以一种不涉及任何其他东西的方式表达这一点，$\mathbf{Q}$ 的割是一个非空集合 $X \subset \mathbf{Q}$，使得

- X 是有上界且没有最大元素；
- 如果 $x \in X$ 和 $x' < x$，那么 $x' \in X$。

这个构造的美是多方面的：如果 X 和 Y 是割集，那么要么 $X \subseteq Y$ 要么 $Y \subseteq X$；$\mathbf{R}$ 上的顺序关系是由集合的包含关系诱导的。此外，有理数对应于已经具有上确界的割集；例如，负有理数集 $\mathbf{Q}^{-}$ 是一个割集，对应于数字 0。

完整性容易看出，因为如果 $A \subset \mathbf{R}$ 有上界（即，是一个有单个上界的割集），那么这些割集的并集本身就是一个割集，并且可以轻易地看出它是 A 的上确界。最后，加法很容易定义；两个割集的和是由将 X 的元素加到 Y 的元素上得到的和的集合。唯一令人烦恼的是，乘法定义稍微有些复杂；类似于加法的定义来取成对乘积的集合是不行的，因为割集包含绝对值很大的负数。一旦乘法被定义，就有许多细节需要检查，即检查割集的加法和乘法是否满足域公理，以及是否满足顺序公理。

康托尔对实数的构造——在第4章中简要概述——完全不同；他的定义比戴德金更复杂，但域和顺序公理更容易建立。

$\square$

存在一个域($\mathbf{R}, +, \cdot$)扩展了有理数系。可以在O.1-O.3顺序公理的意义下比较两个实数。完备性表明，任何从下往上任意接近地被实数逼近的所谓数量本身也是实数。正是在这里，有理数失败了，因为正如命题2.28中所述，单位正方形的对角线可以被有理数任意接近地逼近，但它本身不是有理数。这种不足是严重的，因为分析的基本操作——取“极限”——通常是通过从下往上逼近来完成的。

定理2.30中的唯一性断言意味着任何两个完备有序域公理的实现都是抽象等价的。对于有序域的公理则不能这么说：$\mathbf{Q}$和$\mathbf{R}$是有序域，但它们不是抽象等价的。例如，每个正实数都有一个实数平方根，但并非每个正有理数都有一个有理数平方根。

无穷小

所有关于上界的讨论都有一个版本，经过适当的修改，也适用于下界。假设A是一个非空集合（在某些有序域中），且是有下界的。如果下界集合有一个最大值，那么这个最大值（自然地）被称为A的最大下界或下确界。要看到上界和下界之间的关系，可以简单地考虑一个域中的非空集合A，以及集合$-A = \{-x\colon\ x \in A\}$。根据定理2.23，乘以$-1$会颠倒有序域中的顺序关系，因此$A$的上界的负值是$-A$的下界。命题2.27和2.28对于下确界有明显的重述，完备性可以用下确界来表述。下面的命题2.31列出了上确界和下确界之间的一些基本关系。这个陈述不应令人惊讶，证明留给读者作为练习。

有理数的密度

命题2.31. 设$A \subset$ **R**非空。那么sup$(-A) \leqslant = \leqslant - \leqslant$ inf A 且 inf $A \leq \leqslant$ sup A，当且仅当A仅包含一个元素时取等号。

从 **Q** 的构造中可以相当清楚地看出，**N** 在 **Q** 中没有上界。事实上，**N** 在 **R** 中也没有上界；这一事实，是 **R** 的阿基米德性质的特例，是完备性的结果，并在“极限”的研究中具有核心重要性。

定理2.32. 对于每一个 $a > 0$ 和每一个 $R \in$ **R**，存在一个 $n \in$ **N** 使得 $an > R$。

另一种说法是，如果 $a > 0$，那么集合 $a\mathbf{N} = \{an \mid n \in \mathbf{N}\}$ 在 **R** 中没有上界。在一种模糊的道家比喻中，“千里之行（R）始于足下（一次一步 a）。”

证明。固定一个实数 $a > 0$，并假设存在一个上界 a**N**。根据完备性，将存在一个最小上界，比如说 R。但这样 $R - a$ 就不再是 a**N** 的上界，因此将存在一个 $n \in$ **N** 满足 $an > R - a$。这反过来又意味着 $a(n + 1) > R$，由于 $n + 1 \in$ **N**，这又意味着 R 不是 a**N** 的上界，与 $R =$ sup a**N** 矛盾。

□

陈述这样一个“明显”的事实的原因是存在包含实数域的有序域；在这些域中，存在真正无限或无穷小的元素，并且在这样的域中，集合$\mathbf{N}$有上界。这样的“非阿基米德”域是不完备的；向$\mathbf{R}$中添加更多元素会引入新的间隙。

阿基米德性质可用于证明有理数和实数的根本近似性质：在两个不同的实数之间，存在一个有理数。特别是，每个实数都无限接近有理数集；我们说有理数集在$\mathbf{R}$中是稠密的。几何思想很简单：如果$x < y$，那么$y - x > 0$。选择一个正整数q，使得$\frac{1}{q} < y - x$。$\frac{1}{q}\mathbf{Z}$的连续元素比x和y更接近，所以至少有一个元素必须位于x和y之间。为了以后使用，这里给出了一个正式的陈述和证明。

定理2.33。设x和y是满足$x < y$的实数。存在一个有理数$r = \frac{p}{q}$使得$x < r < y$。

证明。首先考虑$0 < x < y$的情况。根据假设，$0 < y - x$，因此$0 < z := 1/(y-x)$ 由定理2.23得出。阿基米德性质意味着存在一个正整数$q > z$；从而$1/q < y - x$。现在考虑集合 $A = \{n \in \mathbf{N} \mid x < n/q\}$。根据阿基米德性质与 $R = x$ 和 $a = 1/q$，$A \neq \varnothing$。此外，$0 \notin A$ 因为 $0 < x$。性质 $\mathbf{N}.3$ 表明集合 A 有一个最小元素 $p > 0$。因为 $p - 1 \in \mathbf{N}$ 不在 A 中，所以 $(p-1)/q \leq x < p/q$。但是 $1/q < y - x$，所以

$$\frac{p}{q} = \frac{p-1}{q} + \frac{1}{q} \leq x + \frac{1}{q} < x + (y - x) = y,$$

证明 $p/q < y$。

如果 $x < y <$ 为 0，我们将上述论点应用于数字 $0 < -y < -x$，推导出存在一个有理数 r 满足 $-y < r < x$。定理 2.23 表明 $x < -r < y$，这证明了在这种情况下定理的正确性。在剩余的情况下，$x < 0 < y$，定理的结论是显然的。 $\square$

推论2.34。设$x \in \mathbf{R}$，设$\varepsilon > 0$为任意。存在一个有理数$r = \frac{p}{q}$，使得$|x - r| < \varepsilon$。

这是定理2.33的直接推论：例如，设$y = x + \varepsilon$。

完整性与几何

几何中有一个微妙之处，直到20世纪才被充分认识：欧几里得的《几何原本》并非完全基于他的公理，而是隐含地依赖于完备性。严格来说，他的许多定理在表述上是不正确的，当然，一旦几何的基础得到适当发展，欧几里得的证明是正确的。

实数集可以是公理化地或几何地看待，但尽管几何图像通常更直观地有说服力，我们总是需要将断言翻译成公理的语言以便验证它们。在这本书中，几何的作用是促进理解和发现，而域、顺序和完备性公理支持逻辑演绎证明。

公理体系：实数域

从现在起，我们将实数域视为由算术、顺序和完备性公理指定的；这些包含开发微分和积分微积分所需的所有关于 $\mathbf{R}$ 的信息，并为本书的其余部分提供了一个清晰的逻辑基础。为了记录，这些公理在此处收集。

定义2.35 实数域是一个集合 $\mathbf{R}$，以及二元运算 “$+$” 和 “$\cdot$”，以及满足以下条件的子集 P:

- 加法公理：$(\mathbf{R}, +)$ 是一个交换群

  – $+$ 是结合的：$(x+y)+z = x+(y+z)$ 对于所有 x，y，z 在 $\mathbf{R}$ – 在 $\mathbf{R}$ 中存在一个元素 0，使得 $x+0=x$ 对于所有 x 在 $\mathbf{R}$ – 对于 $\mathbf{R}$ 中的每个 x，存在一个 $\mathbf{R}$ 中的 y，使得 $x+y=0$
  – 运算 $+$ 是交换的：$x+y=y+x$ 对于所有 x，y 在 $\mathbf{R}$

- 乘法定理：$(\mathbf{R}^\times, \cdot)$ 是一个交换群 $(\mathbf{R}^\times := \mathbf{R} \setminus \{0\})$

  – 操作 $\cdot$ 是结合的：对于所有 x，y，z 在 $\mathbf{R}$ 中，(xy)
  $z = x(yz)$ – 在 $\mathbf{R}^\times$ 中存在一个元素 1，对于所有 x 在 $\mathbf{R}$ 中，有 $x \cdot 1 = x$

对于 x 在 $\mathbf{R}^\times$ 中，存在一个 y 在 $\mathbf{R}^\times$ 中，使得 $xy = 1$ – 操作 $\cdot$ 是交换的：对于所有 x，y 在 $\mathbf{R}$ 中

- 分配律：对所有 $x(y + z) = xy + xz$，x，y，z 在 $\mathbf{R}$ 中
- 订单公理：

  – 三分法：对于每个 x 在 $\mathbf{R}$ 中，$x \in P$、$-x \in P$ 或 $x =$ 中的 0 恰好成立一个。– 在 $+$ 下的闭包：如果 $x \in P$ 和 $y \in P$，则 $x + y \in P$ – 在 $\cdot$ 下的闭包：如果 $x \in P$ 和 $y \in P$，则 $xy \in P$

- 完整性：如果 $A \subset \mathbf{R}$ 非空且有上界，则 sup $A \in \mathbf{R}$。

表示实数

在“实际应用”中（以及许多数学应用中）需要一种具体的方式来表示实数，类似于将有理数表示为整数商的方式。最熟悉的方案无疑是十进制表示法，它紧凑地编码了一串有理数序列，这些数依次更好地逼近一个实数。例如，表达式 1.414213 . . . 代表的序列是

$$\frac{1}{1}, \quad \frac{14}{10}, \quad \frac{141}{100}, \quad \frac{1,414}{1,000}, \quad \frac{14,142}{10,000}, \quad \frac{141,421}{100,000}, \quad \frac{1,414,213}{1,000,000}, \ldots .$$

这个记法有一个小缺点，即某些表达式的配对对应于同一个实数：例如 $.9\bar{9} = 1$。

十进制表示法基于10的幂次除法，起源于人类有十个手指（在生物学中被称为数字）。对于每个正整数b，都有一个类似的“基b”表示法。可以说，唯一的“自然”表示法是二进制，或基2，尽管需要更长的表达式才能获得相同的精度（二进制表达式0.000001表示有理数1/64，这个数大于十进制0.01）。这些问题在练习2.17、2.18和2.19中进行了详细探讨。

辛普森一家粉丝可能记得，Homer 用八进制计数——基数 8。

一种更自然的表示形式是通过连分数。每个实数 x 可以唯一地表示为

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}},$$

与 a_0 为整数和 a_k 为正整数对于 $k \geq 1$。生成整数 a_k 的算法以及连分数的一些基本性质在练习 2.22 中给出。

在数学常数

随着电子计算器的出现，人们认为某些常数是由它们的十进制展开式定义的。这种信念是荒谬的，应该立即消除。除非知道某种规则来找到连续的数字，否则一个不重复的十进制展开式包含无限量的信息。符号“{v*}2”并不是定义为1{v*}4142135623 {v*}；定义必须指定被定义的东西，但省略号省略了无限多个数字。相反，数学家采取一种实用的观点：“{v*}2”是一个平方为2的正数。（这立即提出了更多问题：这样的数是否存在？是否可能有多个数具有这种属性？如何找到十进制表示？）像{v*}、{v*}或2{v*}这样的实数是由属性定义的，而不是作为十进制表示；有限的十进制表达式仅仅是理性的近似。因此，特定实数的定义通常隐藏着惊人的微妙定理。存在一个唯一的正实数{v*}，其平方为2，是示例4.47的特殊情况（一个更一般的结果由定理5.9给出），并且关于这个“明显的事实”，证明直到第4章才出现。（现在可以给出证明，但中间的材料引入了将使证明更容易理解并更通用的概念。）我们还将遇到其他实数，其中至少有一些是熟悉的：{v*}单位圆的面积，或基本三角函数周期的一半{v*}，{v*}自然对数的底数{v*}，{v*}欧拉常数{v*}，黄金比例{v*}等等。每个都由一个属性来表征，并且在每个情况下都有一个定理，该属性确实指定了一个唯一的实数。有时

可以证明看似无关的性质指定了相同的实数，参见第12、13和15章。

区间

在微积分中，最重要的数集之一是“区间”。区间存在于任意有序域中，但对于大多数有序域（如**Q**）来说，实数区间的特性特别简单，不成立。

定义2.36 设**F**为一个有序域。若对于所有满足$x < y$的x和$y \in I$，都有

$$x < z < y \implies z \in \mathbf{F}.$$

一个区间 I 被称为在 $\mathbf{F}$ 中有界，如果存在一个元素 $R \in \mathbf{F}$ 使得 $-R < x < R$ 对所有 $x \in I$ 成立，并且是开区间，如果它既不包含最小值也不包含最大值。

在一个有序域中，如果集合 I 中一对点之间的每个点也都是集合 I 和 I 中的点，那么这个集合 I 是一个区间。区间的最简单例子是由一对元素 a 和 $b \in \mathbf{F}$ 确定的以下区间:

$$(a, b) = \{x \in \mathbf{F} \mid a < x < b\}$$
$$[a, b] = \{x \in \mathbf{F} \mid a \le x \le b\}$$

您甚至可能见过开区间和闭区间被定义为形如(a, b)或$[a, b]$的集合，并且可能会 wonder 为什么定义2.36如此复杂。原因是，在一般域中——实际上，在除了实数域以外的每个有序域中——存在不是(a, b)形式的开区间！例如，集合

$$A = \{x \in \mathbf{Q} \mid x^2 \le 2\}$$

是一个开区间在 $\mathbf{Q}$ 中，但正如我们在命题2.28中看到的，A 不具有 (a, b) 的形式，无论 a 和 $b \in \mathbf{Q}$ 的选择如何。由于完备性性质，实数域在这方面是特殊的:

命题2.37。如果 $I \subset \mathbf{R}$ 是一个有界、非空的开放区间，那么存在实数 a 和 b 使得 $I = (a, b)$。

证明。因为 $I \subset \mathbf{R}$ 有界且非空，存在 $a := \inf I$ 和 $b := \sup I$。因为 I 是开集，a 和 b 不是 I 的元素，所以 $a < b$。我们试图建立两个包含关系：

$$I \subset (a,b) \quad \text{and} \quad (a,b) \subset I.$$

第一个断言是明确的；如果 $z \in I$，那么根据 inf 和 sup 的定义，我们有 $a \le z \le b$。由于 a 和 b 不是 I 的元素，我们必须有 $a < z < b$，因此 $z \in (a,b)$。

第二个断言同样简单。假设 $a < z < b$；我们必须证明存在 x 和 $y \in I$ 使得 $x < z < y$。现在，$\varepsilon := b - z > 0$，所以根据命题2.27，存在一个点 $y \in I$ 使得 $y > b - \varepsilon = b - (b - z) = z$。类似的论证表明存在一个 $x \in I$ 使得 $x < z$。

□

证明清楚地说明了为什么实数域是唯一一个开区间可以如此容易地被描述的有序域：在每一个其他有序域中，都存在一个有界非空集合，它没有上确界。

方括号用于表示端点的包含，例如闭区间 $[a,b]$
$= \{x \in \mathbf{R} \mid a \le x \le b\}$。半开区间 $[a,b)$ 和 $(a,b]$ 以明显的方式定义
。在几何上描绘区间时，惯例是使用实心点 “
●表示包含端点
，以及一个开放的点“◦”或没有点表示排除端点。

A- 符号表示

在科学中，实验结果从不精确，数据总是以误差区间呈现。例如，我们永远不能说金属棒上的两个标记在数学意义上正好相距一米，只能说它们（比如说）相距在0.9999和1.0001米之间。实际上，真实金属棒上的真实标记本身就有宽度，所以认为一个单一的数字可以准确地代表距离的物理概念是幼稚的。现在，你可能已经相信数学在数值问题上完全精确，但这种信念也不现实。π的确切数值是多少？你的计算器告诉你了吗？物理设备能给你提供π的所有小数位吗？实际上，π是如何定义的，更不用说计算了？这些问题值得回答（并且将在适当的时候得到回答），但此时我们希望研究数值数学的不精确方面。

在计算器普遍之前，大多数科学和工程专业的学生都知道 $\pi \simeq 3.1416$。然而，一个哲学上谨慎的学生通常很难试图弄清楚符号“$\simeq$”的确切含义，这个符号（众所周知！）代表约等于。

数学家对表达式“$\pi \simeq 3.1416$”的解释会让歌德感到自豪：$\pi = 3.1416 \pm 0.00005$。更正式地说，

$$3.14155 \le \pi \le 3.14165, \text{ or } |\pi - 3.1416| \le 0.00005.$$

大致上，不给出更多的小数位数，我们无法更准确地说明 π 等于多少。这种量化我们无知的方法非常有用，因此我们为它引入了特殊的符号：$\pi - 3.1416 = A(0.00005)$。右边的表达式读作“绝对值不超过 0.00005 的量”。

此“A表示法”使我们能够简洁地表达相对大小关系；当我们想表达“如果 $|x| \le 2$，那么 $|10^x| \le 100$”时，我们可以写成 $10^{A(2)} = A(100)$。当我们进行涉及多个不确定项的计算时，A表示法的便利性更加明显：

$$\bigl(2.5 + A(0.1)\bigr) + \bigl(1 + A(0.03)\bigr) = 3.5 + A(0.13)$$

和

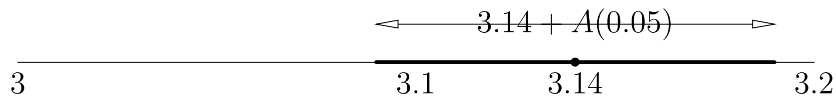
$$\begin{aligned}\bigl(2.5 + A(0.1)\bigr)\bigl(1 + A(0.03)\bigr) &= 2.5 + A(0.1) + A(0.075) + A(0.003)\\ &= 2.5 + A(0.178) = 2.5 + A(0.2).\end{aligned}$$

A 符号提供了一种“粗心大意”的“微积分”。

当然，你必须小心不要将 A 符号完全当作一个普通方程来处理；$1 = A(2)$ 和 $1.5 = A(2)$，但这并不意味着 $1 = 1.5$，也不意味着 $A(2) = 1$。在英语中，这是合理的：1 是一个绝对值最多为 2 的数，但绝对值最多为 2 的数不一定是（必然是）1。同样，$A(0.178) = A(0.2)$，但我们不能得出 $0.178 = 0.2$ 或 $A(0.2) = A(1.78)$ 的结论。然而，我们确实有 $A(0) = 0$。

存在一个关于 A 符号的几何解释。表达式 $A(0.05)$ 表示区间 $[-0.05, 0.05]$ 中的任意数，而表达式 $3.14 + A(0.05)$ 表示区间中的任意数

$$[3.14 - 0.05, 3.14 + 0.05] = [3.09, 3.19].$$



当我们写出 $A(0.1) = A(0.2)$ 时，它表示 $[-0.1, 0.1] \subset [-0.2, 0.2]$。“$A$” 符号也适用于“变量”表达式：

$$x = A(x), \quad 2x = A(1 + x^2), \qquad \text{and} \qquad \frac{x^2 - 1}{x^2 + 1} = A(1).$$

第一个很明显，而第二个是真的，因为

$$0 \le (1 \pm x)^2 = (1 + x^2) \pm 2x \qquad \text{for all real } x,$$

所以 $|2x| \le 1 + x^2$。最后一个留给你们，见练习2.7。

扩展实数系统

扩展实数系统通过在 $\mathbf{R}$ 后添加两个元素得到，分别表示为 $+\infty$ 和 $-\infty$。这些点不是实数，也不在数轴上。通过声明，

$$-\infty < x < +\infty \qquad \text{for all } x \in \mathbf{R}.$$

如果集合 $A \subset \mathbf{R}$ 上无界，则我们写 $\sup A = +\infty$。同样，当 A 下无界时，我们写 $\inf A = -\infty$。如果 $A = \varnothing$ 我们同意 $\sup A = -\infty$ 和 $\inf A = +\infty$。(你应该检查这些看似奇特的定义与空假设的逻辑是否一致！）因此，每个实数集合在扩展实数中都有一个上确界和下确界。注意，与 $\pm\infty$ 的算术运算未定义；某些表达式，如 $+\infty + (-\infty)$，不能以与域公理一致的方式定义。

如果 a 和 b 是扩展实数，那么相应的开区间是集合

$$(a, b) = \{x \in \mathbf{R} \mid a < x < b\}.$$

特别地，$\mathbf{R} = (-\infty, +\infty)$。

社区

区间[$\{a, b\}$]的中点（一个或两个端点可能包含）是$(a+b)/2$，半径是$|b-a|/2$，即长度的一半。（比较定理2.25。）指定以下内容通常很有用

一个开区间通过其中点和半径；关于 a 的 δ-区间是该区间

$$N_\delta(a) := (a-\delta, a+\delta) = \{x \in \mathbf{R} : |x-a| < \delta\}$$

由距离 a 小于 δ 的点组成。删除的 δ-区间不包含 a：

$$N_\delta^\times(a) := N_\delta(a) \setminus \{a\} = \{x \in \mathbf{R} : 0 < |x-a| < \delta\},$$

查看图2.11。有时将删除的 δ-区间视为一对区间，$(a-\delta, a)$
$\cup\,(a, a+\delta)$，是有用的。注意

$$x \in N_\delta(a) \quad \text{iff} \quad x = a + A(\delta') \text{ for some } \delta' < \delta.$$

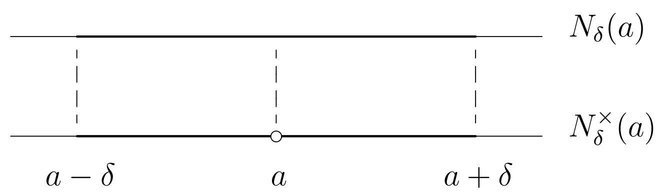


图2.11：关于a的开放和删除的δ-区间。

一个包含关于 a 的某些 δ 区间的 a 邻域是 $\mathbf{R}$ 的子集。一个开区间是其每个点的邻域；一个闭区间不是其端点的邻域。

所有关于 a 的 δ-区间的集合用于研究“函数”在 a 附近的行为。值得注意的是，尽管所有 δ-区间的交集是 $\{a\}$，但这样的区间集合捕捉到了其他情况下无法看到的信息；关于 a 的所有 δ-区间的集合是 a 的“无穷小邻域”。考虑已删除区间的理由是明确忽略点 a，专注于“无限接近”的点。（在实数域 $\mathbf{R}$ 中这种说法是矛盾的，但请记住歌德的引言！）包含某些关于 a 的已删除 δ-区间的集合 $A \subset \mathbf{R}$ 被说成包含足够接近 a 的所有点。

非标准分析

尽管微积分的传统设置是实数域，因为古希腊人设定的历史先例，但微积分可以

基于更大的数系——由A. Robinson发现的非标准实数，其中包含“无穷小”数，即比每个正实数都小的正数，以及比每个实数都大的“无穷大”。（这些无穷大比上面为排序目的引入的粗略符号$+\infty$要微妙得多。非标准实数域是有序的，“包含$\mathbf{R}$的一个副本”，但不是完备的；实际上，它不满足阿基米德性质。）由于复杂的原因，非标准分析在现代数学的主流之外占据一席之地。除了对它的历史偏见之外，还有两个技术原因导致偏见：首先，为了构建非标准实数，必须扩大集合论本身；其次，有一个定理（传递原理）表明，关于实数系统的每个具有“非标准”证明的定理也都有一个“真实”证明。简而言之，需要付出很多代价，但没有逻辑上的回报；没有不能用“标准”技术证明的新定理。因此，非标准分析在逻辑学家和集合论者中最为人所知。然而，新数学的发现过程更多的是试错，而不是严格的逻辑演绎链。I. Stewart认为，非标准分析作为发现实数定理的概念工具非常有用，这些定理否则可能无法被发现！非标准分析将对流体流动中湍流的发生或自发的对称破缺和相变等物理现象的研究产生重大影响的可能性并不小。

2.5 复数

在实数系统的不完善之处中，缺乏具有实系数的多项式方程的通解是一个；$x^2+1=0$没有实数解，例如。为了纠正这种情况的朴素尝试是假设存在-1的平方根，并看看会得出什么逻辑结论。由于历史原因，-1的平方根表示为i，称为“虚数单位”。如前所述，希腊人把“数字”看作“长度”，确实没有长度其平方是-1。从这个角度来看，i确实是虚数！然而，从代数角度来看，“数字”仅仅是域的元素，在一般域中不存在-1的平方根是不被阻止的。（正如我们所看到的，有序域不能包含这样的元素。）

通过一点想象力，人们会考虑具有 a 和 b 的表达式 $\alpha = a + bi$，并按照 $i^2 = -1$ 规定的算术运算以及希望域公理成立。复数就是这样被正式使用超过300年，直到19世纪初，C. F. Gauss 将复数定义为有序实数对，并定义了加法和乘法规则

$$(a,b) + (c,d) = (a+c, b+d),$$
$$(a,b)\cdot(c,d) = (ac-bd, ad+bc). \tag{2.15}$$

这些运算满足域公理，可以通过直接计算来验证。非零复数 $\alpha = a+bi$ 的倒数是

$$\frac{1}{\alpha} = \frac{a-bi}{a^2+b^2}.$$

实数 a 对应于复数 $(a, 0)$，操作 (2.15) 如预期所示：

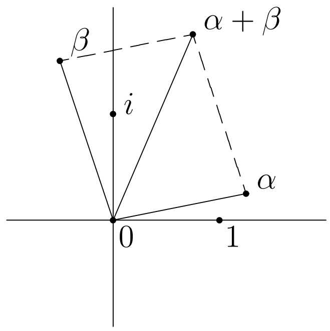
$$(a,0)+(c,0) = (a+c,0), \qquad (a,0)\cdot(c,0) = (ac,0).$$

在文字上，有一个$\mathbf{R}$的副本位于$\mathbf{C}$内部。立即可以看出，对$(0, 1)$来说，它满足$(0, 1)^2 = (-1, 0)$，因此复数系统包含-1的平方根。从我们现代的观点来看，高斯从实数域构造了复数域$\mathbf{C}$，这证明了$\sqrt{-1}$存在的逻辑一致性。

复数在几何上表示为数平面。复数的加法和乘法有美丽的几何描述，见图2.12。我们将在第15章证明这种描述是正确的。复数 α 的加法由平行四边形法则给出，该法则将原点转换为 α。乘以 $\alpha \in \mathbf{C}$ 是围绕原点旋转和缩放平面的操作，使得 1 被携带到 α。特别是，乘以 -1 对应于原点的反射（即，半圈旋转），而乘以 $i = (0, 1)$ 是平面的逆时针四分之一旋转。这幅图赋予了复数与实数一样“真实”的存在感。

实际上有两个虚数单位，i 和 $-i$；按照惯例，i 用点 $(0, 1)$ 表示。$\alpha = a+bi$ 的共轭复数是复数 $\bar{\alpha} = a - bi$。从几何上看，共轭数在水平轴上是对称的。简短的计算表明

$$\overline{(\alpha+\beta)} = \bar{\alpha}+\bar{\beta}, \qquad \overline{(\alpha\cdot\beta)} = \bar{\alpha}\cdot\bar{\beta}. \tag{2.16}$$



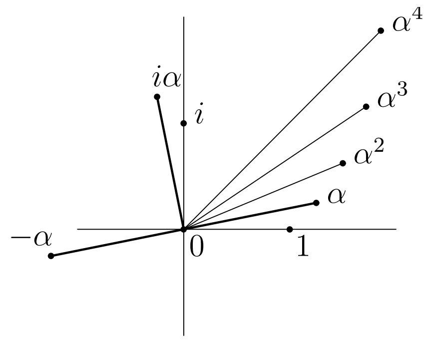


图2.12：复数加法和乘法的几何形状。

在文字上，场运算在共轭后工作方式相同：没有代数上的理由将配对（$0, 1$）与i而不是$-i$等同起来。

复数的范数（或绝对值）是到原点的距离，即
$|\alpha| = \sqrt{\alpha\bar{\alpha}} = \sqrt{a^2+b^2}$，$\alpha$ 和 β 之间的距离是 $|\alpha-\beta|$。这个定义与勾股定理一致，并满足三角形不等式和反三角形不等式：

T 定理2.38. 对于所有复数 α 和 β,

$$\big||\alpha|-|\beta|\big| \leq |\alpha-\beta| \leq |\alpha|+|\beta|.$$

证明——见第15章，定理15.1——如果组织得仔细，并不困难，但不是实数证明的明显推广。然而，几何解释是相同的，额外的好处是复平面上的“三角形”确实是三角形。

练习

练习2.1 将域$\mathbf{F}_2$视为集合$\{[0], [1]\}$，其中[0]是偶数集合，[1]是奇数集合。找出该域的加法和乘法表。–[1]在$\mathbf{F}_2$中是否有平方根？（提示：–[1]是什么？）使用对应关系[0] ↔ False，[1] ↔ True，找出对应加法和乘法的布尔运算。⋄

练习2.2 回想一下，方程（2.2），加法的定义，是结合律的基础情况。同样，乘法的定义，方程（2.4），是分配律的归纳基础情况。哪个已知的恒等式以方程（2.5），指数的定义，作为基础情况？ $\diamond$

练习2.3发现两个数学结构“抽象上是相同的”可以导致新的发现，因为从一个角度看难以看到的事物，从另一个角度看可能很容易看到。

这里是一个由马丁·加德纳提出的奇妙例子。九张卡片，标有1-9，按顺序放在桌子上：

1	2	3	4	5	6	7	8	9

两位玩家轮流抽牌。目标是抽三张牌，总和为15。例如，如果第一位玩家抽出的牌是3、8、5和4（按此顺序），那么牌$\{3, 4, 8\}$构成胜利（假设第二位玩家在此期间没有获胜）。如果两位玩家都没有成功，则游戏平局。

使用3×3 魔方

2	9	4
7	5	3
6	1	8

为了表明这个游戏与井字棋等价，即在两个游戏中存在获胜策略之间的对应关系。 $\diamond$

练习2.4 设(a_i)和(b_j)是序列，设c是一个数。使用数学归纳法证明以下命题。

(a) $\displaystyle\sum_{k=0}^{n}(a_k + b_k) = \sum_{i=0}^{n} a_i + \sum_{j=0}^{n} b_j$ for all $n \geq 0$.

(b) $\displaystyle\sum_{i=0}^{n}(ca_i) = c\sum_{i=0}^{n} a_i$ for all $n \geq 0$.

(c) $\displaystyle\left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{m} b_j\right) = \sum_{i=0}^{n}\left(\sum_{j=0}^{m} a_i b_j\right)$ for all m and $n \geq 0$.

◇

练习2.5（a）使用归纳法证明对于每个 $n \in \mathbf{N}$，有 $1 + n \leq 2^n$，当且仅当 $n = = 1$ 时取等号。

(b) 更一般地，对于所有 $r > \geqslant 0$ 和所有 $n \geq = 1$，证明 $1 + nr \leq (1+r)^n$，当且仅当 $n = = 1$ 时取等。这是下面二项式定理的平凡结果。这里你应该进行归纳。

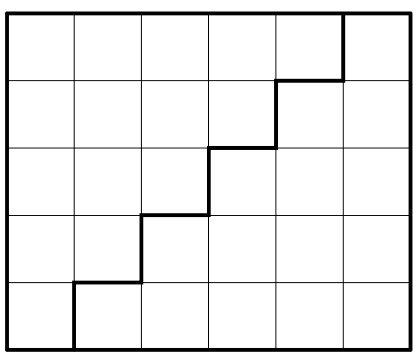
(c) 假设 $0 < r < 1$，并设 $\varepsilon > 0$ 已知。证明存在一个 $N \in \mathbf{N}$ 使得 $r^N < \varepsilon$。换句话说，“通过取足够大的指数，r 的幂可以变得任意小”。
提示：如果 $0 < r < 1$，则存在 $x > 0$ 使得 $r = 1/(1+x)$。◇

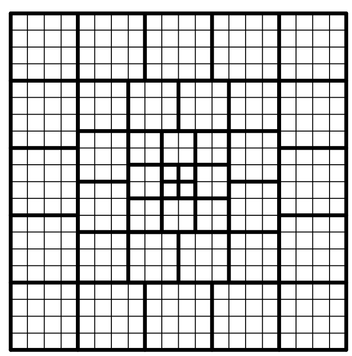
练习2.6 使用对 n 的归纳法来建立以下“幂和”恒等式：

(a) $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$　　(b) $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1}{6}$

(c) $\displaystyle\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4} = \left(\frac{n(n+1)}{2}\right)^2$

部分（a）和（c）中的关系可以被视为以下图表的结果：





在第二层，中心四个正方形的值为1；连续的层更大，第k层由4k个每个为$k \times k$的正方形组成。◇

练习2.7 设 x 为一个实数。建立以下断言。

(a) If $x = A(0.5)$, then $1 + x \geq 0.5$ and $1/(1+x) = A(2)$.

(b $\frac{2 \quad 1}{x}$ x^2+1
$=)$ $A(1)$.

(c) 如果 $x = A(1)$，则 $x^2 = A(x)$ 和 $\big(1 + A(x)\big)^2 = 1 + 3A(x)$。

$\diamond$

练习2.8 设$A \subset \mathbf{R}$非空。对于$c \in \mathbf{R}$，令$c + A = \{c + a \mid a \in A\}$和$cA = \{ca \mid a \in A\}$。

(a) 证明 inf $A \leq \leqslant$ sup A。如果这些数相等，你能说什么？(b) 如果 $A \subset B$，那么 inf $B \leq \leqslant$ inf $A \leq \leqslant$ sup $A \leq \leqslant$ sup B。(c) 证明 sup($c + A$) $\leqslant = c + \leqslant$ sup A。(d) 找出 sup(cA) 的表达式；答案将取决于 $c >$ 是否为 0，$c <$ 是否为 0，或 $c =$ 是否为 0。

以草图说明每个部分。 $\diamond$

练习2.9 设A和B是非空实数集。证明

$$\sup(A \cup B) = \max(\sup A, \sup B), \qquad \inf(A \cup B) = \min(\inf A, \inf B).$$

建议：要么证明右侧满足是上/下确界的条件，要么证明

$$\sup A,\ \sup B \leq \sup(A \cup B) \leq \max(\sup A, \sup B).$$

证明不存在这样的公式来表示集合交集的上确界和下确界。尽可能精确地阐述这一原则。$\diamond$

练习2.10 对于每个区间族，证明其并集或交集符合所述。

(a) $\displaystyle\bigcup_{n=1}^{\infty} [-n, n] = \mathbf{R}$.

(b) $\displaystyle\bigcap_{n=1}^{\infty} [0, 1 + 1/n) = [0, 1]$.

(c) $\bigcap_{n=1}^{\infty}(0,1/n]=\varnothing.$

$\diamond$

练习2.11 通过证明由（2.11）定义的集合P满足三个顺序公理来完成命题2.22的证明。$\diamond$

练习 2.12 证明命题 2.31。 $\diamond$

年金和摊销

练习2.13 早在巴比伦时代，商业就是数学发现的推动力。在这个练习中，你将从头开始找到一个有用的金融公式。

当金钱被贷款时，出借人通常向借款人收取一笔费用（称为利息），该费用与欠款金额成比例。通常，借款人以固定大小的分期付款在固定的时间间隔（如每月或每年）支付给出借人。每期付款的一部分用于支付累积的利息，另一部分用于减少借款金额（本金）。问题是确定每期付款的大小，给定借款金额、利率、偿还贷款所需的时间和付款次数。

(a) 令 r 为年利率（百分比，您可假设 $0 < r < 100$），[7] 并假设 $n \geq 1$ 每年支付一次。每期的利率为 r/n%，因此在一个特定期间产生的利息为 $\rho := r/(100n)$ 倍于该期初的欠款金额。在加上利息后，该期的付款被扣除，得到新的欠款金额。

设 A_{i-1} 为第 i 期初的欠款金额，设 P 为付款金额。以 A_{i-1}、ρ 和 P 为条件求 A_i。（使用上一段末尾的公式计算利息，然后减去付款金额。）如果 $P = \rho A_{i-1}$ 发生，财务上会发生什么？如果 $P < \rho A_{i-1}$ 发生?

(b) 令 A_0 为初始借款金额。确定在支付 i 次款项后应还金额 A_i。（索引与第一部分一致。）

提示：您的答案将取决于 A_0、ρ、P 和 i。您可能需要手动计算前几个 A_i，直到您发现一个模式。

[7]This ensures that $\rho < 1$ below. In fact, charging a rate over about 30% is a crime, called *usury*, though there is no mathematical reason to bound r.

然后使用部分(a)和关于i的归纳法证明你的模式是正确的。方程(2.10)在将答案转化为封闭形式时应有所帮助。请注意至少计算A_1、A_2和A_3；有些看起来可能正确的模式是错误的。

(c) 借款应在 $i = kn$ 时偿还；使用此方法来根据借款金额 A_0、每期“利率”乘数 ρ 和总支付次数 $N: = kn$ 计算支付 P。注意，支付与借款金额成比例，而对利率的依赖则更为复杂。

假设借入10,000元，年利率为4%，在五年内以等额月付款方式偿还。每期付款金额是多少？总共偿还了多少钱？ $\diamond$

二项式定理及其应用

设 $n \geq 0$ 为一个整数。n 的阶乘，记作 $n!$，通过递归定义如下

$$0! = 1, \qquad n! = n\,(n-1)! \quad \text{for } n \geq 1. \tag{2.17}$$

如果 $n \geq 1$，那么 $n!$ 是不超过 n 的正整数的乘积。0! 的约定由练习 2 .15 得到证明。双阶乘 $n!!$，不要与 (n!)! 混淆，定义为

$$0!! = 1!! = 1, \qquad n!! = n\,(n-2)!! \quad \text{for } n \geq 2. \tag{2.18}$$

例如，$6!! = 6 \cdot 4 \cdot 2$ 和 $7!! = 7 \cdot 5 \cdot 3 \cdot 1$。

练习2.14 证明对于所有 m，$n \in \mathbf{N}$，无论是非正式地还是通过数学归纳法，都有 $n! = n!!\,(n-1)!!$ 和 $(2m)!! = 2^m\,m!$。 $\diamond$

下一个练习介绍了“二项式系数”以及它们出现的几种自然方式 。如果 k 和 n 是非负整数，且 $0 \leq k \leq n$，那么二项式系数 $\binom{n}{k}$，读作 “n 选择 k”，被定义为

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}; \tag{2.19}$$

观察 $\binom{n}{n} = \binom{n}{0} = 1$ 对于 $n \geq 0$。根据定义，$\binom{n}{k} = 0$ 如果 $k < 0$ 或 $k > n$。尽管不是立即明显，二项式系数是整数；实际上，它们有一个有用的组合解释，参见部分（b）。

练习2.15（a）证明如果$0 \leq k < \ell \leq n/2$，那么 $\binom{n}{k} < \binom{n}{\ell}$。（这可以通过多种方式从定义中轻松完成。）

使用定义来证明

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} \qquad \text{for} \quad 1 \le k \le n, \tag{2.20}$$

并且使用这个观察结果制作一个二项式系数表，直到 $n = 5$。如果你将固定 n 的系数按行写出，那么下一行的条目是相邻条目的和，这个结果图被称为帕斯卡三角形。为了帮助你开始，前四行是：

$$\begin{array}{r|ccccccccccccc}
n=0 & \cdots & & 0 & & 0 & & 1 & & 0 & & 0 & & \cdots \\
1 & \cdots & 0 & & 0 & & 1 & & 1 & & 0 & & 0 & \cdots \\
2 & \cdots & & 0 & & 1 & & 2 & & 1 & & 0 & & \cdots \\
 & \cdots & 0 & & 1 & & 3 & & 3 & & 1 & & 0 & \cdots
\end{array}$$

方程（2.20）本质上描述了二项式系数；对于所有 $k \in \mathbf{Z}$（和某些 $n \ge 0$），知道 $\binom{n}{k}$ 唯一确定 $\binom{n+1}{k}$ 对于所有 $k \in \mathbf{Z}$。特别是，二项式系数是整数，因为对于每个 k，$\binom{0}{k}$ 是整数。

(b) 设$\underline{\mathbf{n}}$为一个恰好有$n \ge 0$个元素的有限集合；那么$\underline{\mathbf{0}} = \varnothing$，并且为了明确起见，假设$\underline{\mathbf{n}} = \{1, \dots, n\}$如果$n \ge 1$。定义$B(n, k)$为具有恰好$k$个元素的$\underline{\mathbf{n}}$的子集的数量。显然$B(0,0) = 1$，而当$B(n, k) = 0$时，如果$k < 0$或$k > n$。通过将$\{1, \dots, n+1\} = \{1, \dots, n\} \cup \{n+1\}$写成，证明

$$B(n+1, k) = B(n, k) + B(n, k-1) \qquad \text{for } n \ge 0,\ k \ge 1.$$

使用此来证明对于所有整数 k 和 n，有 $B(n, k) = \binom{n}{k}$。

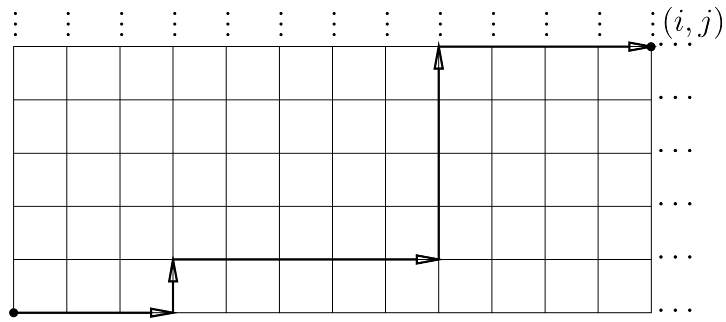
(©) 一个包含 $a + b$ 和 ^{n} 的表达式 a 和 b（实数或复数）和 $n \in \mathbf{N}$ 是一个二项式。如果这个表达式展开，结果将是形如 $a^k\, b^{n-k}$ 的项之和，因为乘积中的每个项的总次数为 n。证明二项式定理：

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k\, b^{n-k}. \tag{2.21}$$

给出尽可能多的论据；至少应该有一个是通过归纳的正式证明，但还有其他方式来记录，例如计算单项式 $a^k\, b^{n-k}$ 在乘积中出现的次数。使用二项式定理来证明

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n, \qquad \sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0.$$

(d) 考虑第一象限中的单位正方形网格。



设 $C(i, j)$ 为通过一系列向上和向右的网格步数将原点连接到点 (i, j) 的路径数。论证 $C(i, j) = C(i-1, j) + C(i, j-1)$ 对于 i，$j \in \mathbf{N}$ 成立，并且对于所有 $i \geq 0$，有 $C(i, 0) = 1$。

找到一个二项式系数，使其等于 $C(i, j)$。(建议：从 (0, 0) 到 (i, j) 需要多少步？其中有多少步是水平/垂直的？这应该可以帮助你猜测答案；然后你可以通过改变变量，使 $C(i, j)$ 的递归关系成为方程 (2.20)，或者通过步骤数量的归纳来验证你的猜测的正确性。)

(e) 使用一张坐标纸，绘制一个“模2的二项式系数帕斯卡三角形”，其中项为0或1，根据相应的二项式系数是偶数还是奇数。可以用实心或空心方块代替0和1。尽量包含至少32行。$\diamond$

表示实数的小数形式

练习2.16 设0 {v*} 1，并设 {v*}。方程（2.10）断言

$$\sum_{i=0}^{n} a\, r^i = a\, \frac{1 - r^{n+1}}{1 - r}.$$

使用练习2.5（c）来证明

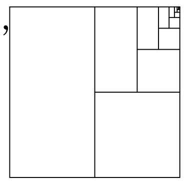
$$\sup_{n \in \mathbf{N}} \sum_{i=0}^{n} a\, r^i = \frac{a}{1 - r} \qquad \text{for } 0 < r < 1. \tag{2.22}$$

评估此公式对于 $a = 0.9$ 和 $r = 0.1$。 $\diamond$

直观上，方程（2.22）给出了一种找到“无限多个项之和”的方法，前提是这些项构成一个几何级数。实际上，并不存在“无限加法”；涉及的是上确界。尽管如此，标准表示法是

$$\sum_{i=0}^{\infty} ar^i = \frac{a}{1-r} \qquad \text{for } 0 < r < 1.$$

设置 $a = r = 1/2$ 在方程（2.22）中给出有启发性的结果

$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots = 1,$ 

练习2.18（c）给出了一个更实质性的应用。

练习2.17 本练习概述了将十进制算术纳入**R**的公理化描述中的步骤。尽管如此，你应该仅使用**R**和**N**的公理，尽管你总是可以自由地运用你的所有知识来提前猜测答案和策略。

在一个十进制表达式，如314.1592中，小数点左侧或右侧的数字位置确定一个相应的10的幂，该数字乘以这个幂。具体来说，如果$a_1, \dots, a_n$和$b_0, b_1, \dots, b_m$是数字（即集合$\{0, 1, 2, \dots, 9\}$的元素），那么表达式$b_m \cdots b_1 b_0.a_1 a_2 \cdots a_n$表示一个（正）有理数

$$\sum_{j=0}^{m} b_j\, 10^j + \sum_{i=1}^{n} a_i\, 10^{-i}.$$

如果存在无限多个非零 a_i，则故事会变得稍微复杂一些；这种情况在部分（c）中处理。

(a) 证明每个实数 {v*} 可以唯一地表示为 $N + d$ 的和，其中 N 是一个整数，且 $0 \le d < 1$。

注释：考虑情况 $x \ge 0$ 和 $x < 0$ 分别可能有所帮助。你可能需要 **R** 的阿基米德性质和属性 **N**.3。通常的表示法是 $N =: \lfloor x \rfloor$ 和 $d = x$ mod 1。当 x 为正时，N 和 d 被称为 x 的整数部分和小数部分；当 $x < 0$ 时，N 和 d 不是整数部分和小数部分。例如，如果 $x = -3.14159$，那么 $N = -4$ 和 $d = 0.85841$，而整数部分和小数部分是 -3 和 -0.14159。

(b) 证明每个自然数 N 可以唯一地表示为

$$(*) \qquad b_m \cdots b_1 b_0 := \sum_{j=0}^{m} b_j\, 10^j, \qquad b_j \in \{0,1,2,\ldots,9\}, \quad b_m \neq 0.$$

注释：这可能是一个根深蒂固的真理，以至于很难决定需要证明什么。记住，自然数是根据迭代后继性定义的；**N** 并没有本质上的“十进制”属性。这个问题更多的是为了让你思考，而不是要求你提供一个完整的证明，这个证明可能很长，取决于你的怀疑程度。以下是一些你应该考虑的问题：给定一个自然数 N，是否存在一个大于 N 的10的幂？一个小于 N 的最大的10的幂？如果是这样，这个幂可以减去多少次，直到差值变成负数？这能否将问题简化为一个更简单的情况？你如何确定两个表示法（$*$）中哪一个更大？

(c) 证明每个表达式

$$(\dagger) \qquad 0.a_1 a_2 \cdots a_n := \sum_{i=1}^{n} a_i\, 10^{-i}, \qquad a_i \in \{0,1,2,\ldots,9\}$$

是一个区间 [0, 1) 中的有理数。如果无穷多个 a_i 不为零，则定义

$$(\ddagger) \qquad \sum_{i=1}^{\infty} a_i\, 10^{-i} := \sup_{n\in\mathbf{N}} \sum_{i=0}^{n} a_i\, 10^{-i}.$$

证明每个表达式（‡）代表[0, 1]区间内的一个实数。反过来证明，每个$0 \le x \le 1$的实数x都可以表示成形式（‡）的表达式。

注释：(†)中的表达式小于1应该是对n，即数字个数的简单归纳。唯一其他微妙的部分是证明[0, 1]中的每个实数都有一个十进制展开；你可能不想写出所有细节，但至少应该让自己相信这是可以做到的。这个想法与部分(b)类似；想象单位区间被分成十分之一、百分之一等等可能会有所帮助。然后可以从x的十进制表示中“读出”x的位置。

(d) 在 (‡) 中的十进制表示不唯一；例如，1.00 = 0.99。证明这在以下意义上本质上只有这种歧义。两个十进制表达式

$$\sum_{i=1}^{\infty} a_i\, 10^{-i} \quad \text{and} \quad \sum_{i=1}^{\infty} a_i'\, 10^{-i}$$

表示在 $[0, 1)$ 中相同的实数当且仅当

- $a_i = a_i'$ 对于所有 $i \in \mathbf{N}$，或者
- 存在一个 $n \in \mathbf{N}$，使得 $a_i = a_i'$ 对于 $1 \le i < n$，$a_n = a_n' + 1 \le 9$，以及 $a_i = 0$，$a_i' = 9$ 对于 $i > n$。

例如，0.2499 = 0.25。在第二种情况下，可能需要颠倒 a_i 和 a_i' 的角色。◇

练习2.18 这继续了练习2.17，但由于那个练习的结果是熟悉的初级学校事实，因此可以独立完成。

(a) 证明每个有理数都有一个十进制表示，要么是终止的（有限的），要么最终是循环的（在有限个十进制位之后，有一段有限的数字重复无限次，如$1/12 = .0833$）。实际上，证明如果p/q是最简形式且具有最终循环的十进制展开，那么重复的数字串长度最多为$q-1$。

提示：有理数的十进制展开可以通过将 q 除以 p 进行长除法得到，并且在每一步过程中只有有限个可能的余数。

(b) 证明每个有限小数或最终循环小数都表示一个有理数。

提示：对于终止小数，这是明确的，参见练习2.17(c)。对于循环小数，只需证明

$$0.a_1a_2\cdots a_N\overline{a_1a_2\cdots a_N}$$

表示一个有理数（为什么？），这可以通过练习2.16（c）实现，使用 $r = 10^{-N}$ 和 $a = .a_1a_2\cdots a_N \in \mathbf{Q}$。

(c) 将 .1212 和 .87544544 写成分数形式，使其为最简形式。注释：当然，(c)部分是为了确保你理解(b)部分。总之，这个练习表明无理数有

非终止非循环小数表示，这些是唯一的。循环数字不全为“9”且不全为“0”的有理数也有唯一的十进制表示。终止的有理数恰好有两种表示。$\diamond$

练习2.19 如果 $b \geq$^2 是一个自然数，那么在练习2.17中完成的所有内容都可以使用 b 的幂而不是10的幂来表示。这种表示法被称为“基 b”，尽管特殊情况 $b =$^2、3、8 和 16 分别称为二进制、三进制、八进制和十六进制。为该练习的每个部分（尤其是，需要哪些符号，基-b 表达式代表什么有理数？）制定类似的声明，并确信它们的证明是十进制表示法论证的直接修改。

将十进制数64写成三进制、八进制和十六进制表示法。将分数1/3写成三进制和二进制。$\diamond$

连分数

为了印刷简便，一个表达式

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}} \tag{2.23}$$

对于所有 k，当 $a_k >$ 为 0 且 $k >$ 为 1 时，$a_n >$ 为 1 将表示为 $[a_1; a_2, \dots, a_n]$。在这些限制下，(2.23) 中的表达式被称为有限连分数，可以看作是有限小数的类似物。接下来的练习研究通过有限连分数逼近实数的可能性，这导致“无限”连分数，类似于无限小数。连分数相对于小数有一些理论优势：$x \in \mathbf{R}$ 的表示是唯一的，当 $x \in \mathbf{Q}$ 时是有限的，并且不依赖于基数 b 的选择。在某种意义上，连分数还提供了对无理数的“最优”逼近。接下来的两个练习涉及有理数和有限连分数；练习 2.22 处理无理数和无限连分数。

练习2.20 假设在整个过程中 $0 < q < p$，并且 p 和 q 没有公约数。

(a) 设置 $r_0 = p$ 和 $r_1 = q$，并递归定义 r_k 为 $k \geq 2$ 通过

$$r_{k-1} = a_k r_k + r_{k+1}, \qquad 0 \leq r_{k+1} < r_k. \tag{2.24}$$

(参见图（2.25)以下。）定义n为使得$a_k \neq 0$的最大索引。证明对于$1 \leq k \leq n$，有$a_k > 0$，并且

$$\frac{p}{q} = [a_1; a_2, a_3, \ldots, a_n].$$

每个有理数 x 都有一个唯一的有限连分数表示。（显然，每个有限连分数都代表一个有理数。）

(b) 使用(a)部分找到5/7, 8/5和355/113的连分数表示。

(c) 以 $[a_1; a_2, \ldots, a_n]$ 表示 q/p 的连分数，其中 $[a_1; a_2, \ldots, a_n]$ 是 p/q 的连分数。

(d) 增加 a_n 是否会使 $[a_1; a_2, a_3, \ldots, a_n]$ 变大或变小（或者这个问题比这更微妙，如果是这样，真正的答案是什么？）

如果你在最后一部分卡住了，做下一项练习。 $\diamond$

练习2.21固定一个有理数x =$[a_1; a_2, \ldots, a_n]$，对于1<$\leq k \leq n$定义$r_k = [a_k; a_{k+1}, \ldots, a_n]$。

(a) 证明 $r_{k+1} = 1/(r_k - a_k)$ 对于 $1 \leq k < n$ 成立。引用本章适当的结果，得出如果 r_{k+1} 变大，则 r_k 减少，反之亦然。

(b) 如果 $a_n = r_n$ 变大，$x = r_1$ 是增加还是减少?

这个问题只有初等代数，但结果很快就会非常有用。$\diamond$

练习2.22设$x \in \mathbf{R}$，并且如同练习2.17中所述，让$\lfloor x \rfloor$表示不大于x的最大整数。我们希望研究以下情况的可能性:

$$x = [a_1; a_2, a_3, \ldots] = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}$$

使用 a_k 个整数和 $a_k > $ 个 0 对于 $k \geq$ 个 1。暂时不考虑这些无限表达式意味着什么，为每个正整数 n 定义以下数字

$$x_n = [a_1; a_2, a_3, \ldots, a_n], \qquad r_n = [a_n; a_{n+1}, a_{n+2}, \ldots].$$

(a) 证明 $x_1 < x_3 < x_5 < \cdots < x_6 < x_4 < x_2$，即证明

$$x_{2k-1} < x_{2k+1} < x_{2k+2} < x_{2k} \qquad \text{for all } k \geq 1.$$

您应该能够从上一题的(b)部分读懂大部分内容。正式证明 $r_{k+1} = 1/(r_k - a_k)$ 对于 $k \geq 1$，参看练习2.21(a)。

(b) 给定 $x \in \mathbf{R}$，递归定义整数 a_k、p_k 和 q_k，以及实数 y_k，如下: 将 $y_1 = x$、$p_{-1} = q_0 =$ 设置为 0，将 $p_0 = q_{-1} =$ 设置为 1。然后定义，对于 $k \geq 1$,

$$\begin{aligned} a_k &= \lfloor y_k \rfloor \\ p_k &= a_k\, p_{k-1} + p_{k-2} \\ q_k &= a_k\, q_{k-1} + q_{k-2} \\ y_{k+1} &= \frac{1}{y_k - a_k} \end{aligned} \tag{2.25}$$

证明对所有的 $n \geq 1$，$a_n > 0$，$q_n < q_{n+1}$ 和 $p_n < p_{n+1}$，$x_n = p_n/q_n$，形式上展示 $y_n = r_n$。

(c) 递归证明 $|x - x_n| < 1/(q_n q_{n+1})$ 对于 $n \geq 1$。这给出了 x 由 x_n 近似的一个定量度量。

(d) 证明“连分数是 {v*} 的最佳有理近似”在以下意义上:

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2} \quad \text{iff} \quad \frac{p}{q} = \frac{p_n}{q_n} \quad \text{for some } n \in \mathbf{N}.$$

根据部分(a)，数 x_{2k+1} 的上确界存在，数 x_{2k} 的下确界存在。实际上，这些界限是相等的，它们的公共值被定义为无穷连分数 $[a_1; a_2, a_3, \ldots]$。我们将在第4章回到连分数，届时将更容易研究这些问题。◇

练习2.23 本练习探讨了“周期”连分数与二次多项式根之间众所周知的关系。

(a{v*} $2x=\sqrt{}$；计算 1$\leq n\leq$6 的 p_n 和 q_n。在列表 a_1、a_2、$a_3,\ldots$ 中找到一个规律，并证明你的猜测是正确的。计算 1$\leq n\leq$6 的 x_n^2 。在这个过程中，你不需要使用电子计算器；只需观察 1< $\sqrt{2}$< 1.5 即可。你可能发现制作以下表格（有足够的列）很有帮助：

	$n=-1$	$n=0$	$n=1$	$n=2$	$n=3$	$\cdots$
a_n	$-$	$-$				$\cdots$
p_n	0	1				$\cdots$
q_n	1	0				$\cdots$
y_n	$-$	$-$				$\cdots$
x_n	$-$	$-$				$\cdots$

(b) 对 $x=\sqrt{3}$ 重复。

(c) 设 a 和 b 为正整数。证明连分数 $[a;b,a,b,a,\ldots]$ 满足一个具有整数系数的二次多项式。

整数出现在 x 的连分数中，可以用作衡量“x 有多无理”的尺度。有理数有有限的连分数，而有理系数二次多项式的无理根有“周期为2”的“周期性”连分数。$\diamond$

第三章

函数

函数的概念在现代数学中绝对核心，特别是在微积分中是最重要的概念之一。这个词的日常含义——“你的考试成绩是你在学习上付出的努力的函数”或“你的生活水平是收入的函数”——与数学上的含义相似，表达了一种事物对另一事物的依赖关系。函数在用数学术语描述现实世界关系时非常有用。

数学上，一个函数可以被视为一种“规则”，它将一个“输出”值（位于某个特定集合中）分配给每个“输入”（位于另一个特定集合中）；这是函数作为“黑盒”的常见解释。分析通常涉及输出为实数且输入为实数有限列表的函数，这在经典上被称为“变量”。本书中避免使用“变量”一词，因为它鼓励用相同的符号表示两个完全不同的概念——数字和函数。在作者看来，“变量”是一个语言概念，而不是数学概念。尽管如此，有时使用这个术语是方便的，偶尔也会用到。如有混淆，定义总是最后的依据。

3.1 基本定义

一个函数 $f: X \to Y$（读作 “从 X 到 Y 的 f”）由三件事组成：

- 一个非空集合 X，称为 f 的定义域，其元素是 f 的“输入”；
- 一个非空集合 Y，称为 f 的范围，其元素是

"f 的潜在输出"；

- 每个元素 $x \in X$ 被分配一个唯一的元素 $y \in Y$ 的"规则"。元素 y 被称为 f 在 x 的值，通常写作 $y = f(x)$。

这三条信息在概念上与自然数的公理相对应：它们指定了实际中使用的函数的性质。相比之下，形式定义仅使用集合来实现这些性质。这是我们关于集合层面的最终定义；此后的一切都是使用公理层面的性质来定义的。

定义3.1 设X和Y非空。函数f：$X \to Y$是子集$\Gamma_f = \Gamma \subset X \times Y$，对于每个$x \in X$，存在一个唯一的$y \in Y$，使得$(x, y) \in \Gamma$。

许多本书中的函数是"实值"函数（意味着 $Y \subset \mathbf{R}$），是"实变量"的函数（意味着 $X \subset \mathbf{R}$）。对于这些函数，其图形可以看作是笛卡尔平面 $\mathbf{R} \times \mathbf{R}$ 的子集，图3.1。

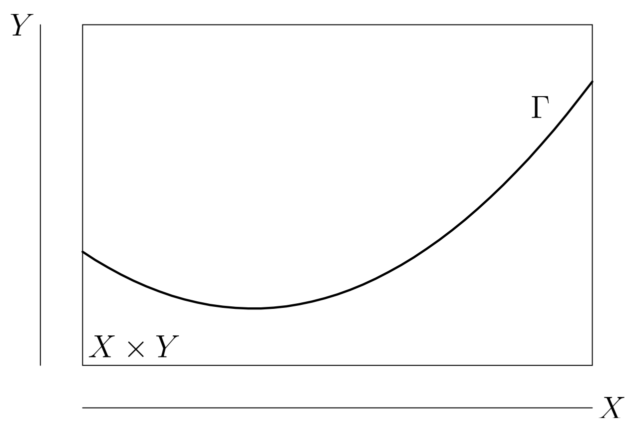


图3.1：一个函数作为图形。

这个定义与通常的微积分书定义之间最大的实际区别在于，定义域和值域是 f 的一个基本部分。改变定义域——即使是去掉一个点——也会导致不同的函数。一个函数不仅仅是一个规则或公式；像 $f(x) = x^2$ 这样的方程并不能定义一个函数；参见例3.2。

Γ 集合被称为 f 的图。根据我们的定义，一个函数就是它的图。我们通常说“一个函数 f”“具有图 Γ”，尽管根据定义我们完全可以这样说“函数 Γ”。

域的元素被称为点，并说它们被 f 映射到 Y 的元素；这有时表示为 $f\colon x \mapsto y$ 或 $x \mapsto y = f(x)$。程序上，从 x 开始，找到唯一的 $y \in Y$，使得 $(x, y) \in \Gamma$，图 3.2。

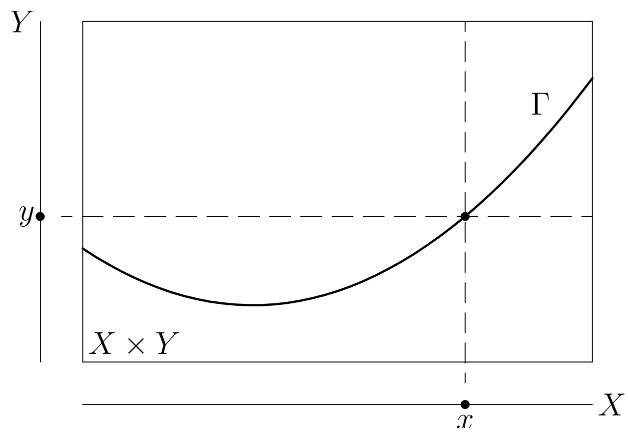


图3.2：一个映射的函数。

我们使用“函数”和“映射”这两个术语来表示相同的概念，尽管“函数”暗示了实数值，而“映射”则没有。

图和过程

函数可以被“静态地”或“动态地”看待。一个函数的图（静态地）在一个单独的图片中捕捉了函数的所有特征。有时，将函数视为一个黑盒是更好的选择，这样 $x \in X$ 就是一个“变量”，而输出 $y = f(x)$ 随着输入 x 的变化而变化（动态地）。在这个图中，每个 x 都是一个潜在的输入，但所有输入的集合并不是同时考虑的。

对于一个物理例子，考虑一个在垂直数线上运动的粒子，如图3.3所示。定义域是考虑运动的所有t的集合，而范围被取为整个实数线$\mathbf{R}$。在动态图中，取个别的时间值，并且当t增加时，粒子被视为向上或向下移动。在静态图中，粒子的“历史”是其世界线，它

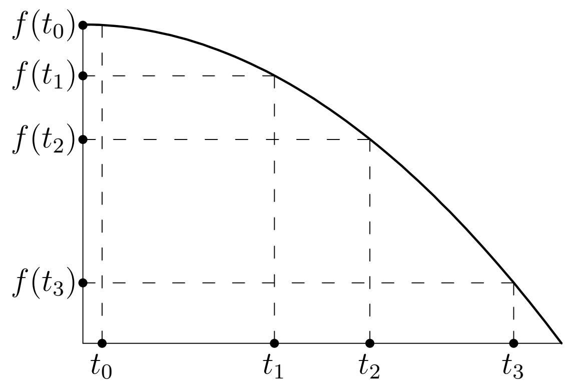


图3.3：函数的静态（粗体）和动态解释。

这正是 f 的图。重要的是要记住，尽管通常更方便以一种或另一种方式看待特定问题，但这些是同一数学情况下的两种不同视角。

射影性和注入性

一个函数在定义上对域中的每个点 x 都有一个唯一值 X。然而，定义中没有任何内容断言值域中的每个点 y 确实是函数值。函数的所有值的集合是其像，$f(X)=\{y \in Y\colon\ y=f(x)$ 对于某个 $x \in X\} \subset Y$:

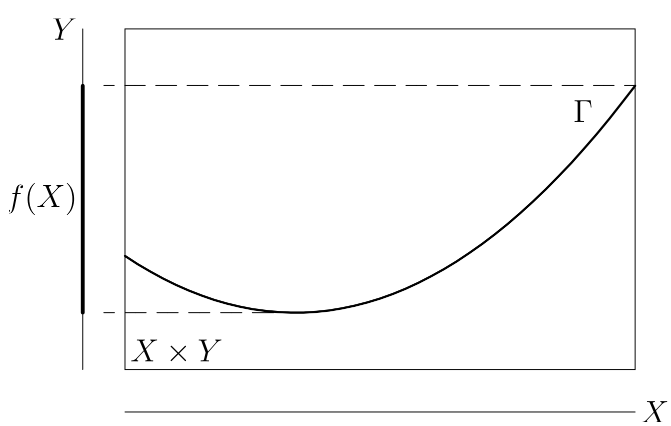


图3.4：函数的图像。

一个函数 f 将 X 映射到 Y，如果 $Y = f(X)$，即如果像是整个范围。“Onto” 作为形容词使用（“f 是 onto”），尽管术语“surjective”——由 N. Bourbaki[1] 创造——具有相同的意义。图 3.4 中描绘的函数不是 surjective。

大多数微积分书籍将函数的“值域”定义为我们所称的“像”；因此，满射性是一个多余的概

单射性

尽管每个 x 确定一个唯一的 y，但在定义中没有任何东西保证图像中的每个 y 都对应一个唯一的 x；域中的多个点可能映射到同一个 $y \in Y$，见图3.5。

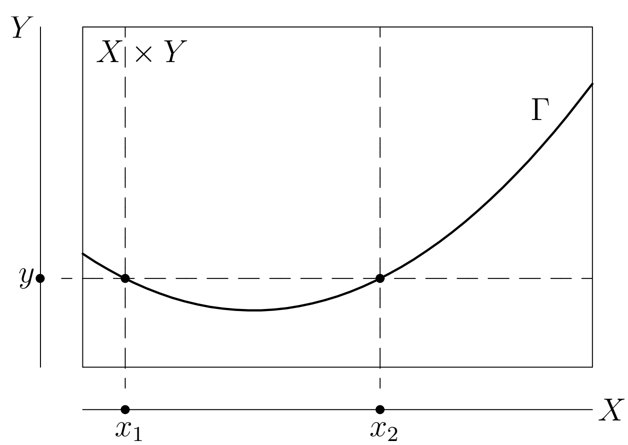


图3.5：不同的点可以映射到图像中的相同点。

如果图像中的每个 y 都是唯一 $x \in X$ 的 f 的值，那么函数 f 是一一对应的，或者说是一对一的。换句话说，一一对应性意味着“如果 $f(x_1) = f(x_2)$，那么 $x_1 = x_2$”。这也有助于逆否形式“如果 $x_1 \neq x_2$，那么 $f(x_1) \neq f(x_2)$”，或者以“不会发生 $x_1 \neq x_2$ 和 $f(x_1) = f(x_2)$”的形式。

[1] Boor bah KEY: The pen name of an influential group of French mathematicians.

垂直和水平线测试

假设 $f\colon[a,b]\to[c,d]$，并令 $R=[a,b]\times[c,d]$ 为由不等式 $a\le x\le b$ 和 $c\le y\le d$ 确定的 x-y 平面内的矩形。f 的图形是子集 $\Gamma\subset R$，它"通过垂直线测试"，即与每条垂直线 $x = x_0$ (相交恰好一次。确实，一条垂直线至少与图形相交一次，因为每个定义域的点都对应一个函数值，而该线最多与图形相交一次，因为函数是单值的。上述第三条规则保证了这些性质。

注入性和满射性的条件具有类似的空间几何解释，涉及水平线。假设 $\Gamma \subset R$ 是一个函数 f 的图。那么 f 是满射当且仅当每条 $y = y_0$ (线与 Γ 至少相交一次，而 f 是单射当且仅当每条水平线最多与图相交一次。

一个既是单射又是满射的函数是双射。上述注释表明，双射 $f\colon[a,b]\to[c,d]$ 的图像恰好与每条线 $y = y_0$ $(c \le y_0 \le d)$ 交一次；我们可以说图像"通过了水平线测试"。记住，一个函数是否是单射或满射不仅取决于定义函数的"规则"，还取决于定义域和值域。

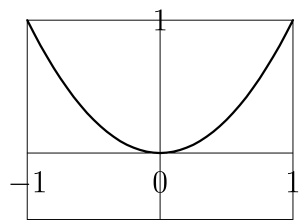


图3.6：与平方法则相关的函数。

示例3.2 这里是四个不同的函数，它们都通过平方规则 $f\colon x\mapsto x^2$ "定义"，但具有不同的定义域和/或值域。下面每个函数都是通过从图3.6中移除（可能为空）的一部分得到的。

- $X=[-1,1]$ 和 $Y=[-1,1]$。这个函数不是单射的，例如 $-1\ne 1$ 但 $f(-1)=f(1)$。该函数也不是满射的，因为例如 $y=-1$ 不是实数的平方，因此不在像集中。

- $X = [-1, 1]$ 和 $Y = [0, 1]$；图的下半部分被移除。此函数是满射的，因为每个实数 $y \in [0, 1]$ 是某个实数 $x \in [-1, 1]$ 的平方，根据定理5.8。与上面一样，此函数不是单射的。

- $X = [0, 1]$，$Y = [-1, 1]$；左半部分被移除。这个函数不是满射，如在第一个例子中。然而，这个函数是一对一的，因为正实数的两个平方根互为相反数，而其中只有一个在f的定义域内。形式上，如果$f(x_1) = f(x_2)$，那么

$$0 = f(x_1) - f(x_2) = x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2).$$

  现在，如果 x_1 和 x_2 是 f 的定义域中的点，那么 $x_1 + x_2 > 0$，因此前一个方程意味着 $x_1 = x_2$。因此 f 是单射的。

- $X = [0, 1]$，$Y = [0, 1]$；仅剩下右上象限。此函数是一个双射，如上所述的断言容易检查。

再次强调，改变定义域（和/或值域）会产生不同的函数。□

单调函数

设 X 是一组实数，例如一个区间。如果对于定义域中的所有 x_1 和 x_2，有 $x_1 < x_2$ 则称函数 $f : X \to \mathbf{R}$ 为增函数。从几何上看，图形“向右倾斜”。如果 n 是一个正整数，那么由 $f(x) = x^n$ 定义的第 n 次幂函数 $f : [0, +\infty) \to \mathbf{R}$ 根据定理 2.23 是增函数。类似地，如果 $x_1 < x_2$ 则称函数 $f : X \to \mathbf{R}$ 为减函数。既是增函数又是减函数的函数是严格单调的。

应用增函数保持不等式。例如，平方函数在正实数集上是增函数，且（1.7）$^2 = 2.89 < 3 < 3.24 = (1.8)^2$，因此如果存在一个正实数 $\sqrt{3}$，其平方是 3，那么 $1.7 < \sqrt{3} < 1.8$，见图 3.7。注意，严格单调函数是单射的。（如果这不明显，从定义中证明它！）同样，减函数的应用会反转不等式。定理 2.23 表示倒数函数在正实数集上是减函数：如果 $0 < x_1 < x_2$，那么 $1/x_1 > 1/x_2$。

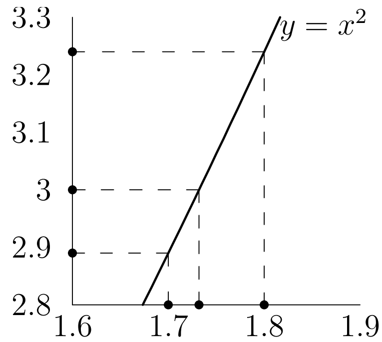


图3.7：增函数保持顺序关系。

一个函数 f 是非递减的，如果 $x_1 < x_2$ 推出 $f(x_1) \leq f(x_2)$。增函数一定是非递减的，但反之不一定。例如，一个常数函数是非递减的，但不是增函数。你应该没有困难地写下非增函数的定义。一个既是非递减又是非增的函数是单调的。

前像

设 $f\colon X \to Y$ 为一个函数。如果 $B \subset Y$，那么 f 在 B 下的前像是被 f 映射到 B 的 X 中所有点的集合：

$$f^{[-1]}(B) = \{x \in X \mid f(x) \in B\} \subset X, \tag{3.1}$$

查看图3.8。预像满足一些有用且易于验证的性质：

命题3.3。设$f\colon X \to Y$为一个函数。如果$A \subset X$和$B \subset Y$，那么

$$f\big(f^{[-1]}(B)\big) = B \quad \text{and} \quad A \subset f^{[-1]}\big(f(A)\big).$$

第二个包含通常是适当的。

证明。练习3.2和3.3。 □

预像 B 可能是空的，即使 B 不是空的（如果 f 不是满射），一个单点集的预像可能包含多个点（如果 f 不是单射）。因此，如果 f 不是双射，就无法将 $f^{[-1]}$ 视为从 Y 到 X 的映射。

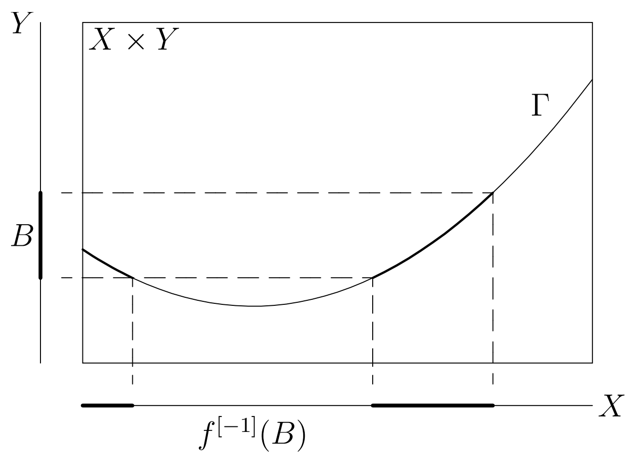


图3.8：集合 B 在 f 下的原像。

限制和扩展

一个可以对函数做的最简单的事情就是在不改变值域的情况下缩小其定义域；记住这会得到一个不同的函数。如果 $f: X \to Y$ 是一个函数且 $A \subset X$ 非空，那么 f 在 A 上的限制，记为 $f|_A$，形式上定义为 $\Gamma \cap (A \times Y)$，见图 3.9。粗略地说，限制“遵循与 f 相同的规则，但仅对 A 中的点定义。”换句话说，$f|_A : A \to Y$ 是由以下定义的函数

$$f|_A(x) = f(x) \quad \text{for } x \in A. \tag{3.2}$$

限制因此相当于忘记 f 的一部分，或者丢弃信息。参见练习3.6，了解注入性、满射性和限制之间的关系。

扩展

让 $f: A \to \mathbf{R}$ 为一个函数和 $B \supset A$。f 到 B 的扩展是一个函数 $F: B \to \mathbf{R}$，它在 A 上与 f 一致，即与 $F|_A = f$ 一致。扩展从不唯一，但在实际问题中通常有额外的约束，在这些约束下，扩展可能是唯一的，也可能根本不存在。

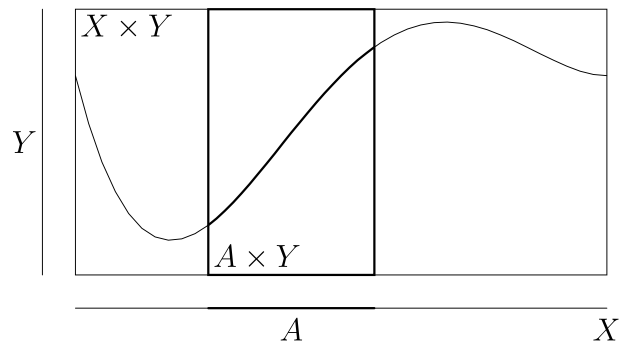


图3.9：函数的约束。

注意事项

当函数通过如下公式定义时

$$f(x) = \frac{x+1}{x-1}, \quad g(z) = \sqrt{1-z^2}, \quad \text{or } h(t) = \log t,$$

并且域和范围没有明确指定，通常同意范围是实数集 $\mathbf{R}$，而域是“自然域”，包括所有使表达式为实数的实数。当处理复变函数时，这种宽松性通常是不够的；必须仔细指定域。在任何情况下，稍微过于小心总比含糊不清要好。

很糟糕　不明确地谈论“函数 x^2”的两个　原因：

- 域和范围未指定（较轻）；
- 未提及“自变量”（这是一个严重的遗漏）。

未声明独立变量的名称是一个极其常见的错误，因为许多学生通过潜移默化地学习，认为x总是代表“独立变量”。这个观点在伯克利数学楼的一幅涂鸦中被幽默地推向了极致：

$\sqrt{3} > 2$ for sufficiently large 3.

我们如此习惯于符号“3”表示特定的实数，以至于给它赋予另一个意义是荒谬的。不幸的是，

约定用 {v*} 表示“自变量”的规则不能普遍适用，正如我们在后面的章节中将会看到的。规则 $x \mapsto x^2$ 与规则 $t \mapsto x^2$ 完全不同，然而在现实生活中，有时人们想要考虑一个关于 t 的常数函数，其值在每处都是 x^2。规则 $x \mapsto x^2$ 是一个程序，即取一个数并将其平方，而 x^2 仅仅是一个数，尽管它是当 x 作为输入时偶然出现的输出。用“平方规则”来表示可能更准确，即

$$\boxed{} \longmapsto \boxed{}^2,$$

在理解到表示一个数的任意表达式可以放入框中。这种表示法避免了将字母 x 锁定在特定角色上，但过于笨拙，不适合通用。在任何情况下，规则 $x \mapsto x^2$、$t \mapsto t^2$ 和 $\xi \mapsto \xi^2$ 在没有进一步了解字母 x、t 或 ξ 的情况下在数学上是相同的。尽管 x 常表示“通用”输入，y 表示相应的函数值，但也可以互换它们的作用，或者使用更复杂的表达式作为输入。函数指定的是输入值和输出值之间的关系，而不是表示法。符号和意义之间的区别是数学中最难吸收的心理要点之一。

常用的构造“$x = x(t)$”应极其谨慎地使用，或者（更好）完全避免。在左侧，x是一个数字；在右侧，x是一个函数。将x称为“变量”会导致语法模糊：x是函数还是数字？“$f(x)$”是函数值还是一个‘复合函数’？随着函数数量的增加，问题变得更加复杂，并且很容易导致两个不同的函数被赋予相同的名称。在最好的情况下，这会导致不必要的混淆，但如果在您使用符号操作程序时发生这种情况，您将陷入困境：计算机程序只能通过对象的字面名称来区分对象。

函数是“黑盒”

一个函数由其定义域、值域以及定义域中各点的取值完全确定。虽然这个陈述在抽象上听起来空洞，但在实践中可能会让人感到反直觉。例如，这三个公式中的每一个都在$\mathbf{R}$上定义了绝对值函数：

$|x| = \sqrt{x}2$;

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases} \qquad |x| = \begin{cases} x^2 & \text{if } x = -1,\, 0, \text{ or } 1 \\ \sqrt{x^2} & \text{otherwise} \end{cases}$$

它很容易找到无限多个定义相同函数的公式集合。虽然其中第二个恰好是定义，但它们之间没有一个是比另一个更正确的。这个例子有点愚蠢（尽管请注意第一个公式，这是一个经常被误记的恒等式），因为在这个阶段，我们除了将函数定义为代数公式的集合之外，几乎没有其他定义函数的方法，因此验证两个函数的相等性很可能是代数问题。在随后的章节中，研究具有复杂定义的函数（涉及极限和上确界，如导数、定积分和无穷级数）；所谓的“微积分基本定理”断言某些函数对是相等的。通常，一个函数有一个有趣的解释（但定义复杂），而另一个容易计算（但没有有趣的解释）。知道这些函数是相等的是有价值的信息。为了强调：“函数 f 和 g 的‘相等性’意味着 f 和 g 有相同的定义域和相同的值域，并且对于定义域中的所有 x，$f(x) = g(x)$ 都相同；$f(x)$ 和 $g(x)$ 可以通过完全不同的方式得到。

3.2 基本函数类

本节介绍了几个有趣的函数类别。其中一些（如多项式和向量）可能您已经熟悉；如果是这样，您应该尝试将您现有的知识与迄今为止解释的集合和函数框架相结合。

多项式与有理函数

一个非空集合 $A \subset \mathbf{R}$ 上的多项式函数是一个由公式定义的函数 $p : A \to \mathbf{R}$

$$p(x) = \sum_{k=0}^{n} a_k\, x^k = a_0 + a_1\, x + a_2\, x^2 + \cdots + a_n\, x^n \tag{3.3}$$

使用 $a_k \in \mathbf{R}$ 对 $k = 0, \dots, n$。多项式函数之所以重要，原因有很多。其中最不重要的是，多项式函数的评估仅使用加法和乘法。

右侧的表达式称为“多项式（在x中）”，应与多项式函数（具有指定的定义域A）区分开来。a_k被称为x^k的系数；系数a_0通常称为常数项。在表示（3.3）中，通常假设$a_n \neq 0$，在这种情况下，多项式p的次数为n，a_n被称为最高次或首项系数。如果首项系数为1，则多项式是单次的。

引理3.4. 设$p : \mathbf{R} \to \mathbf{R}$为(3.3)给出的多项式函数。如果对于所有$x$，$p(x) = 0$，那么对于所有$k$，$a_k = 0$。

证明。我们将对 p 的次数进行归纳，即 p 的次数。如果 $p(x) = a_0$ 是一个常数多项式，即如果 $p =$ 的次数为 0，那么结果是显然的。假设引理的结论对每个次数为 n 的多项式成立，并令 p 的次数为 ($n +$ + 1)。写

$$p(x) = \sum_{k=0}^{n+1} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots + a_{n+1} x^{n+1}.$$

根据假设，$0 = p(0) = a_0$，因此我们有

$$0 = p(x) = \sum_{k=1}^{n+1} a_k x^k = x(a_1 + a_2 x + \cdots + a_{n+1} x^n) =: xq(x)$$

对于所有 x；因此 $q(x) =$ 为 0，除非 $x =$ 为 0。如果我们证明 $q(0) =$ 为 0，那么我们可以应用归纳假设来得出结论，p 的其余系数都为 0。

我们证明逆命题：如果 $a_1 = q(0) \neq 0$，那么对于某个 x，$q(x) \neq 0$。思路是，对于“非常小”的 $|x|$，值 $q(x)$ “近似”于 a_1。精确地说，反向三角不等式断言

$$(*)\quad |q(x)| = |a_1 + a_2 x + \cdots + a_{n+1} x^n| \geq \big||a_1| - |x| \cdot |a_2 + \cdots + a_{n+1} x^{n-1}|\big|.$$

为了得到下界，我们寻求 $|x| \cdot |a_2 + \cdots + a_{n+1} x^{n-1}|$ 的上界。三角不等式意味着

$$(**)\qquad |x| \cdot |a_2 + \cdots + a_{n+1} x^{n-1}| \leq |x| \cdot \big(|a_2| + \cdots + |a_{n+1}|\,|x|^{n-1}\big).$$

如果我们选择 x 使得 $|x| < 1$，那么 $(**)$ 的右侧不大于 $|x| \cdot \big(|a_2| + \cdots + |a_{n+1}|\big)$。此外，如果我们取 $|x| < |a_1|/2 \big(|a_2| + \cdots + |a_{n+1}|\big)$，那么 $(**)$ 的右侧不大于 $\frac{1}{2}|a_1|$。这反过来又意味着 $(*)$ 的右侧至少为 $\frac{1}{2}|a_1|$。

总结来说，我们已经证明如果 $|a_1| > 0$，那么

$$0 < |x| < \min\Big(1, \frac{|a_1|}{2\big(|a_2| + \cdots + |a_{n+1}|\big)}\Big) \implies |q(x)| \geq \frac{|a_1|}{2} > 0.$$

这是所需的逆否命题。 □

对于每个固定的数 a，可以用 $x - a$ 的幂次来写出多项式。你可以将 x 和 $u := x - a$ 视为两个不同的坐标系；用 $x - a$ 的幂次来写多项式，描述的是在“u 世界”中观察者看到的相同多项式。例如，如果 $u = x - 2$，则 $1 + x^2 = 5 + 4u + u^2$。一般公式

$$\sum_{k=0}^{n} a_k\, x^k = \sum_{k=0}^{n} b_k\, (x-a)^k$$

确定 b_k 的值，通过展开并等于 x 的类似幂次。右侧的多项式表示法被称为 $x - a$ 的幂次。请注意，k 是一个虚指数，因此等式两边的单个项不相等；求和符号不能省略。最高次项系数总是相等：$a_n = b_n$。在第14章中，我们将获得一个快速计算过程，用于根据 a_k 求得 b_k 的值，参见示例14.12。

算术运算符作为函数

函数的定义域不必是实数集。加法和乘法可以看作是定义域为有序实数对集合 $\mathbf{R} \times \mathbf{R}$ 的实值函数；s 和 p：$\mathbf{R} \times \mathbf{R} \to \mathbf{R}$ (对于“和”和“积”) 的定义如下

$$s(x, y) = x + y, \qquad p(x, y) = x \cdot y.$$

系统研究“多个变量”的函数是在更高级的课程中进行的。定义“多个变量的函数”的概念很容易，但为它们发展微积分要困难得多，并且需要牢固掌握一元微积分。

Lagrange 插值多项式

一个常见的智力测试问题是给出一个有限序列的前几个项——例如1, 2, 3, 5，或$\otimes \oplus \oplus \otimes \otimes\otimes$——并要求给出下一个项，或者生成该序列的规则。具有讽刺意味的是，这样的问题是非数学的，因为无论给出什么模式，都有无限多种继续下去的方式。这些测试确实证明了人类大脑在辨别模式方面的非凡能力，即使没有逻辑上隐含的模式。

假设我们希望找到一个多项式 p，使其产生上述数值序列，即

$$(*) \qquad p(1) = 1, \quad p(2) = 2, \quad p(3) = 3, \quad p(4) = 5.$$

虽然存在无限多个多项式满足这四个方程，但存在一个唯一的多项式，其次数为3或更少。这并不完全明显，但可以如下看出。首先想象一下，我们手头有四个三次多项式 e_1、e_2、e_3 和 e_4，它们满足

$$\begin{array}{llll}
e_1(1) = 1 & e_1(2) = 0 & e_1(3) = 0 & e_1(4) = 0 \\
e_2(1) = 0 & e_2(2) = 1 & e_2(3) = 0 & e_2(4) = 0 \\
e_3(1) = 0 & e_3(2) = 0 & e_3(3) = 1 & e_3(4) = 0 \\
e_4(1) = 0 & e_4(2) = 0 & e_4(3) = 0 & e_4(4) = 1
\end{array}$$

然后 $p(x) = e_1(x)+2e_2(x)+3e_3(x)+5e_4(x)$ 将是一个满足 $(*)$ 的三次多项式，因为例如（从第三列向下阅读）

$$\begin{aligned}
p(3) &= e_1(3) + 2e_2(3) + 3e_3(3) + 5e_4(3) \\
&= 0 + (2 \cdot 0) + (3 \cdot 1) + (5 \cdot 0) = 3.
\end{aligned}$$

实际上，给定“魔法多项式”$\{e_i\}_{i=1}^4$，我们只需填写空白即可生成任意四个数字的序列：

$$(**) \qquad p(x) = \underline{\quad} e_1(x) + \underline{\quad} e_2(x) + \underline{\quad} e_3(x) + \underline{\quad} e_4(x).$$

现在，我们如何找到e_i呢？这比看起来要简单；多项式$\widetilde{e}_1(x) = (x-2)(x-3)(x-4)$在1处非零，在其他三个数处为零。除以$(1-2)(1-3)(1-4) = -6$调整1处的值，得到$e_1$：

$$e_1(x) = \frac{(x-2)(x-3)(x-4)}{(1-2)(1-3)(1-4)} = -\frac{x^3 - 9x^2 + 26x - 24}{6},$$

查看图3.10。类似推理告诉我们

$$e_2(x) = \frac{(x-1)(x-3)(x-4)}{(2-1)(2-3)(2-4)},$$

$$e_3(x) = \frac{(x-1)(x-2)(x-4)}{(3-1)(3-2)(3-4)},$$

$$e_4(x) = \frac{(x-1)(x-2)(x-3)}{(4-1)(4-2)(4-3)}.$$

当然，这些多项式可以进行展开，但给出的形式清楚地表明这些是所寻求的神奇多项式。

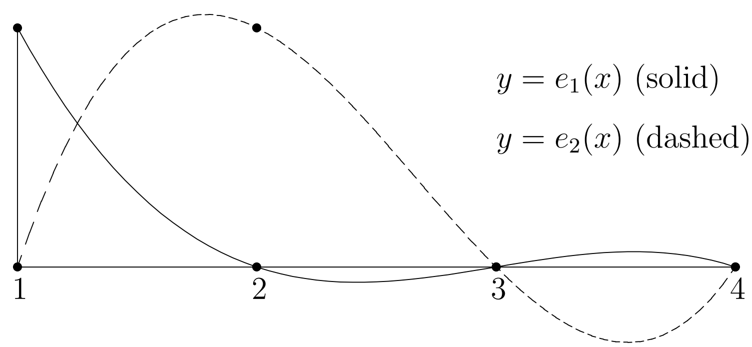


图3.10：用三次多项式插值四个点。

一旦你消化了这个解决方案，你就会意识到这个论证证明了更多:

定理3.5。设$B = \{b_i\}_{i=1}^n$是n个不同的实数或复数集合，设$C = \{c_i\}_{i=1}^n$是n个数的任意集合。存在一个唯一的次数至多为$n-1$的多项式p，使得

$$p(b_i) = c_i \qquad \textit{for all } i = 1, \ldots, n. \tag{3.4}$$

多项式 p 被称为给定数据 $\{b_i\}$ 和 $\{c_i\}$ 的拉格朗日插值多项式，其存在性由定理3.5断言。插值多项式的存在性证明了每个有限数列都可以由足够高次的多项式生成。

证明。你应该不难说服自己，关键是找到一个集合 n “魔法多项式” $\{e_i\}_{i=1}^n$，每个多项式的次数 $n-$ 为 1，在 b_i 处等于 1，在其他所有 b_j 处为零。这是对上述关于四个不同点的论证的直接推广。$(**)$ 的明显推广允许你插值任意一串 n 数字。

定理中唯一的新成分是唯一性断言。假设 q 是另一个次数最多为 $n-1$ 的多项式，满足 (3.4)。那么差分 $p-q$ 是一个次数最多为 $n-1$ 的多项式，它有 n 个不同的根，即 b_i。这表明 $p-q$ 是恒等于零的，因此 $p=q$。

□

分段多项式函数

一个在闭区间上的函数 f 是分段多项式，如果其定义域可以被划分为有限个区间，使得每个子区间上的 f 是多项式。一个例子是函数 f: $[-2,2] \to \mathbf{R}$，定义为

$$f(x) = \begin{cases} 0 & \text{if } -2 \le x < -1 \\ x^2 & \text{if } -1 \le x \le 1/4 \\ x - x^5 & \text{if } 1/4 < x \le 1 \\ 1/2 & \text{if } 1 < x \le 2 \end{cases}$$

图表显示在图3.11中，它还说明了使用圆圈和点来表示开区间或闭区间的用法。

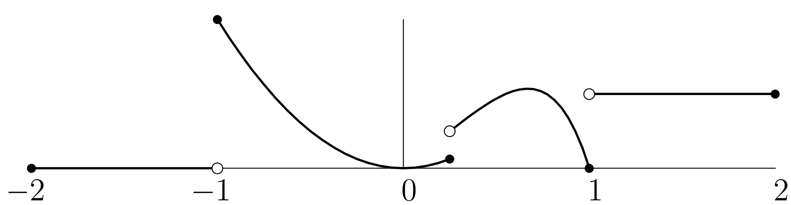


图3.11：分段多项式函数。

形式幂级数

一个多项式可以被视为一个表达式 $p(x) = \sum_{k=0}^{\infty} a_k x^k$，其中除了有限个系数 a_k 外，其余系数均为零。次数

这是对应系数非零的最大指标。用此符号，两个多项式的和与积为

$$\begin{aligned}\sum_{k=0}^{\infty} a_k\,x^k + \sum_{k=0}^{\infty} b_k\,x^k &= \sum_{k=0}^{\infty} (a_k + b_k)\,x^k,\\ \Big(\sum_{k=0}^{\infty} a_k\,x^k\Big) \cdot \Big(\sum_{k=0}^{\infty} b_k\,x^k\Big) &= \sum_{k=0}^{\infty} \Big(\sum_{i=0}^{k} a_i\,b_{k-i}\Big) x^k.\end{aligned} \tag{3.5}$$

第一个方程的意义应该是清晰的，而第二个方程说明要乘两个多项式，我们需要将第一个多项式的每一项乘以第二个多项式的每一项，然后合并相同幂次的 {v*}：{v^3} 的系数是

$$\sum_{i=0}^{k} a_i\,b_{k-i} = a_0 b_k + a_1 b_{k-1} + a_2 b_{k-2} + \cdots + a_{k-1} b_1 + a_k b_0.$$

特别地，多项式的和或积是一个多项式。这些方程对于以 a 为中心的多项式有明显的类似物。

如果我们放弃最多只有有限个系数非零的假设，那么方程（3.5）仍然有意义（每个系数的计算只涉及有限次的算术运算）。形式幂级数（称为{v*}）可以无歧义地相加和相乘。形式幂级数不能以明显的方式定义x的函数，因为将无限多个数相加是没有意义的。相反，形式幂级数应该被视为系数的无限列表（a_k）$_{k=0}^{\infty}$；方程（3.5）解释了如何相加和相乘两个这样的列表。形式幂级数在许多情况下都很有用。仅举一个例子，请注意，

$$(1 - x)(1 + x + x^2 + x^3 + \cdots) = 1.$$

是否这个方程在给 x 赋予特定数值时有任何意义完全是另一回事。有时可以“相加”一个形式幂级数的项（可能对 x 有所限制），从而创建出不是多项式的函数。这样的函数被称为“（实）解析的”。在十九世纪之前研究的函数——多项式、有理函数、指数函数、对数、三角函数以及在经典分析中遇到的各种更奇特的存在——都是解析的。当时的数学家们默认“函数”是解析的。最

著名的多产的18世纪数学家L. 欧拉[2]是一位在操作幂级数方面深刻的genius。我们在本书后面遇到的许多令人瞩目的结果都归功于欧拉。

多项式除法与因式分解

设 $\mathbf{F}$ 为一个域。若存在系数在 $\mathbf{F}$ 中的非常数多项式 p_1 和 p_2，使得它们满足 $p_1p_2 = p$，则称系数在 $\mathbf{F}$ 中的多项式 p 在 $\mathbf{F}$ 上可分解。例如，在 $\mathbf{R}$ 上我们有

$$x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2}), \qquad x^3 - 2x - 4 = (x - 2)(x^2 + 2x + 2),$$
$$x^4 - 1 = (x - 1)(x + 1)(x^2 + 1).$$

一个没有因子的多项式是不可约的（在 $\mathbf{F}$ 上）。第一个例子在 $\mathbf{Q}$ 上是不可约的，而二次多项式 $x^2 + 1 = (x + i)(x - i)$ 在 $\mathbf{R}$ 上是不可约的，但在 $\mathbf{C}$ 上有因子。扩大一个域使得分解多项式“更容易”。

存在一个带余数的多项式除法算法，类似于带余数的整数除法。以下特殊情况适用于第15章中的我们的目的。一般情况类似，参见练习3.12。

定理3.6. 设p是一个系数在域$\mathbf{F}$中的多项式，并设$a \in \mathbf{F}$。存在一个唯一的系数在$\mathbf{F}$中且次数小于deg p的多项式q，使得

$$p(x) = (x - a)q(x) + p(a).$$

证明。当 p 是单射时，只需证明该定理，因为我们可以在 q 中吸收一个乘积常数。我们通过 p 的次数进行归纳。如果 p 是常数，定理显然成立：取 $q =$ 为 0。

该陈述，“如果 p 是单射且次数为 k，则存在一个次数最多为 $k -$ 的多项式 q，使得 $p(x) = (x - a)(q(x) + p(a)$，”是我们第 k 级的归纳假设。假设 p 是次数为 ($k + 1$)) 的单射多项式；多项式 $p(x) - (x - a)x^k$ 的次数最多为 k（，因为我们已经减去了最高次项)。在提取首项系数后，我们可以应用归纳假设来找到一个次数最多为 ($k - 1$) 的 q，使得 $p(x) - (x - a)x^k = (x - a)q(x) + p(a)$，或者

$$p(x) = (x - a)\big(x^k + q(x)\big) + p(a).$$

[2]Pronounced “Oiler”.

这是第 $k+$ 级的归纳假设 1。 □

推论3.7.一个多项式p如果且仅如果$p(a)=$0能被$(x-a)$整除。

一个数 $a \in \mathbf{F}$ 是 p 的根当且仅当 $p(a)=0$。推论3.7表明根与线性因子之间存在对应关系。

有理函数

一个多项式的商确定一个“有理函数”。更确切地说，如果 $A \subset \mathbf{R}$ ，那么一个函数 $f\colon A \to \mathbf{R}$ 是有理函数，如果存在多项式 p 和 q，使得对所有 $x \in A$，有 $q(x) \neq 0$，对所有 $x \in A$，有 $f(x)=p(x)/q(x)$ 。通过限制（如果需要），多项式 q 可以在 A 上变得非零。通常假设 p 和 q 没有非平凡公因子，在这种情况下，分数 p/q 被称为已约简。表达式 $(1-x)/(1-x^2)$ 未约简，而 $1/(1+x)$ 和 $(1+x^2)/(1-x^2)$ 已约简。

自然定义域为 $f(x)=(1-x)/(1-x^2)$ 的实数集合，其中分母不为零，即 $\mathbf{R}\setminus\{\pm1\}$。在此例中，我们可以约去一个公因式，得到函数 $g(x)=1/(1+x)$，其自然定义域仅排除 $x=-1$。观察 $f(x)=g(x)$ 对于 $x \neq 1$，但形式上 $f(1)=0/0$，而 $g(1)=1/2$：约去公因式允许我们给 $f(1)$ 赋值。然而，不约去公因式允许隐式限制定义域，这在某些情况下是有用的。在寻找有理函数 $f=p/q$ 的自然定义域时，询问 $q(x)=0$ 的位置，而不约去公因式。

隐式和代数函数

一个函数不一定需要通过一个变量的公式“明确”给出，但可以通过两个变量的关系“隐式”指定。例如，方程 $x^2+y^2-1=0$ 定义了两个函数，$f(x)=\pm\sqrt{1-x^2}$ 对于 $|x|\leq 1$。如果我们设置 $y=f(x)$，那么隐式关系 $x^2+y^2-1=0$ 对于 f 的定义域中的所有 x 都成立。

设 $R=(a,b)\times(c,d)$ 是平面上的一个矩形，设 $F\colon R \to \mathbf{R}$ 是一个函数。我们说方程 $F(x,y)=0$ 定义了在 R 中的一个隐函数，如果存在一个唯一的函数 $f\colon(a,b)\to(c,d)$ 使得

$$F(x,y)=0 \quad \text{for } (x,y)\in R \iff y=f(x) \quad \text{for } x\in(a,b).$$

几个矩形在图3.12中描绘；每个“好”矩形确定一个隐函数，而每个“坏”矩形则不能。为了强调，一个方程是否定义隐函数不仅取决于方程本身，还取决于平面中的矩形R。取一个开矩形并不特别重要。

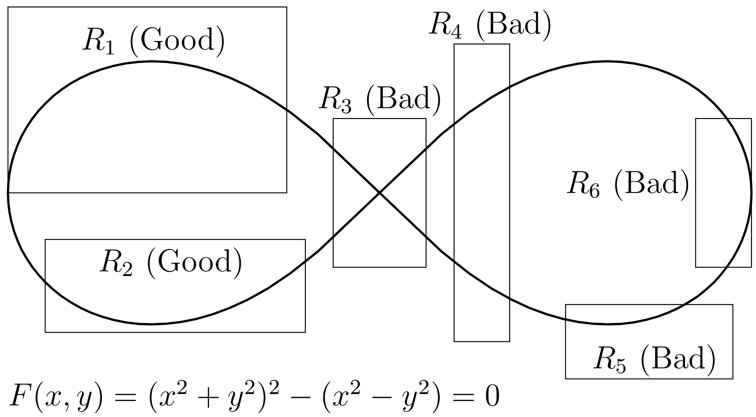


图3.12：代数方程的零点轨迹和隐函数。

示例3.8 方程 $x^2+y^2-1=0$ 定义了在矩形 $[-1,1]\times[0,1]$，$(0,1)\times(0,2)$，和 $[-0.1,0)\times(-1.1,0.8)$ 中的隐函数，但不在正方形 $[-1,1]\times[-1,1]$ 中，也不在矩形 $[1-\delta,1]\times[-1,1]$ 中，无论 $\delta>0$ 多小。你应该画一个草图并确信这些断言。 □

代数函数

让 $F\colon \mathbf{R}\times\mathbf{R}\to\mathbf{R}$ 为一个多项式函数；这意味着存在常数 a_{ij}，满足 i，$j\in\mathbf{N}$，并且只有有限个 a_{ij} 非零，使得

$$F(x,y)=\sum_{i,j=0}^{\infty} a_{ij}\,x^i y^j \quad \text{for all } (x,y)\in\mathbf{R}\times\mathbf{R}. \tag{3.6}$$

零点集为 F 是 $Z(F)=\{(x,y)\mid F(x,y)=0\}\subset\mathbf{R}\times\mathbf{R}$。

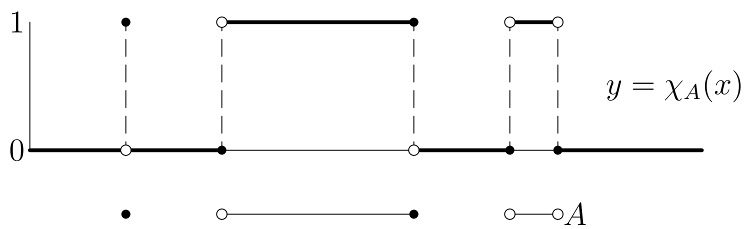
定义3.9 设$f:(a,b)\to(c,d)$为一个函数。如果f在矩形$(a,b)\times(c,d)$中通过一个多项式函数F隐式定义，那么我们称f为代数函数。

每个有理函数都是代数的（练习3.8），但反之则不然；如上所示，对于$-1 < x < 1, 0 \leq y < \infty$，函数$f(x) = \sqrt{1-x^2}$是代数的。一般来说，仅使用四种算术运算和根式提取来表达代数函数是不可能的，这不仅是一个实际问题，也是一个理论问题。

特性与步进函数

设X为一个非空集合，而$A \subseteq X$ (可能为空)。A在X (中的特征函数，有时称为“指示函数”)，是一个定义在$X \to \mathbf{R}$上的函数χ_A：

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A, \\ 0 \text{ if } x \notin A. \end{cases} \tag{3.7}$$



特征函数回答——对于每个 $x \in X$——问题是“你是 A 的一个元素吗？”并将响应转换为二进制（“是” $= 1$，“否” $= 0$）。布尔运算被转换为模 2 的算术，参见练习 2.1 和 3.4。计算机科学家可能会将范围视为有限域 $\mathbf{F}_2 = \{0, 1\}$ 而不是 $\mathbf{R}$，以利用模 2 算术的力量。

设 $I \subset \mathbf{R}$ 为一个区间。阶梯函数是一个函数 $f : I \to \mathbf{R}$，它只取有限个值，且其原像都是点和区间的有限并集。例如，对于 $k = 1, \dots, n$，设 I_k 为一个区间，χ_k 为 I_k 的指示函数，c_k 为一个实数，并假设区间 I_k 互不相交。该函数

$$f(x) = \sum_{k=1}^{n} c_k \chi_k(x) = \begin{cases} c_k & \text{if } x \in I_k \\ 0 & \text{if } x \notin I_k \text{ for all } k \end{cases} \tag{3.8}$$

这是一个阶梯函数。换句话说，阶梯函数不仅仅是分段多项式，而是“分段常数”。阶梯函数是积分理论（第7章）的基础，因为它们

可以使用仅加法和乘法“集成”，因为它们可以用来近似非常广泛的函数类。练习3.5以（3.8）的方式描述了沿线的阶梯函数。

向量与序列

最简单的函数是定义域为有限集的函数。具有 n 个元素的典型集合是初始段 $\underline{\mathbf{n}} := \{1, \ldots, n\}$。函数 $\mathbf{x} : \underline{\mathbf{n}} \to \mathbf{R}$ 是一组 n 有序对，$\{(k, x_k) \mid 1 \leq k \leq n\}$。相同的数据被编码为有序 n-元组 $\mathbf{x} = (x_1, x_2, \ldots, x_n)$，通常称为向量。拉格朗日插值多项式的存在表明每个向量都是多项式函数；这个观察结果主要是一个好奇。

所有在 $\underline{\mathbf{n}}$ 上的实值函数集合表示为 $\mathbf{R}^n$。定义域为一个点的函数本质上是一个实数，因此函数集合 $\mathbf{R}^1$ $\{1\} \to \mathbf{R}$ 可以看作是普通数轴。一个函数 $\mathbf{x} : \{1, 2\} \to \mathbf{R}$ 是一个有序对 (x_1, x_2)，所有此类函数的集合 $\mathbf{R}^2$ 可以看作是笛卡尔平面；例如，点 $(-2, 3)$ 是由 $\mathbf{x}(1) = -2$ 和 $\mathbf{x}(2) = 3$ 定义的函数。同样，$\underline{\mathbf{3}} = \{1, 2, 3\}$ 上的实值函数集合可以看作是笛卡尔空间。没有数学理由限制在三个点的域上，但函数空间变得难以可视化。

上述备注隐藏了一个微妙之处。如果 $A \subset \mathbf{R}$ 是一个无限集合，比如说闭单位区间 $[0, 1]$，那么在 A 上的实值函数集合是绝对庞大的；大致上，每个 A 的元素都有一个坐标轴，并且它们在某种意义上是相互垂直的！另一方面，这个集合的单个元素（即函数 $f : A \to \mathbf{R}$）可以想象成平面上的一个图形。换句话说，平面上的图形集合是庞大的。人们说，在 $\underline{\mathbf{n}}$ 上的实值函数依赖于“有限多个参数”，或者说实值函数空间 $\underline{\mathbf{n}}$ “具有 n 个自由度。”稍微拉伸一下语言，单个函数 f : $[0, 1] \to \mathbf{R}$ 依赖于无限多个参数，而 $[0, 1]$ 上的实值函数空间具有无限多个自由度。

序列

设X为一个非空集合。一个函数$a : \mathbf{N} \to X$被称为X中的序列，并且也记作$(a_k)_{k=0}^{\infty} \subset X$。从概念上讲，一个序列是

无限有序点列（可能非互异）在 X 中。正如符号所示，有序的 n 元组只是一个有限序列。

一个数列可能由一个公式定义，例如 $a_k = 1/(k+1)$ 或 $a_k = (-1)^k$，或者通过递归指定，如下所示

$$a_0 = 2, \qquad a_{k+1} = \frac{1}{2}\left(a_k + \frac{2}{a_k}\right) \quad \text{for } k \geq 0. \tag{3.9}$$

(比较引理2.29；这个序列依次给出对 $\sqrt{2}$.) 的更好近似。在实践中，序列通常递归出现，找到封闭公式是非常理想的（尽管有时困难或不可能）。更多例子在练习中给出。

序列是微积分中最重要的技术工具之一。康托对实数域的构造基于有理数序列。一般来说，如果一个人想要研究一个对象，比如像 π 这样的无理数，或者像 cos 这样的函数，或者平面上的曲线所围成的面积，一个自然的方法是考虑一个在某种意义上“近似”目标对象的序列。希望利用近似器的性质来推断“极限”对象的性质。

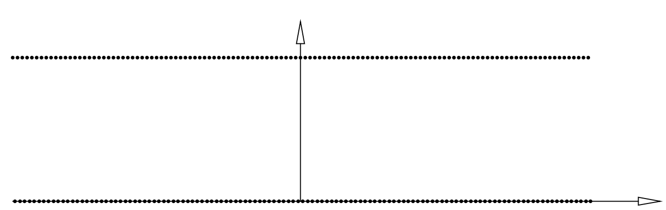
“病理学”示例

非常奇怪的函数可以通过一系列规则或公式来指定。重要的是，域中的每个点必须与一个确切值相关联。

示例 3.10 在 $\mathbf{R}$ 中，$\mathbf{Q}$ 的特征函数定义为

$$\chi_{\mathbf{Q}}(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

因为有理数在实数中是稠密的，参见定理2.33，图看起来像这样



描述性地说，该图看起来像两条水平线，理解是每条垂直线只与图相交于一点。当然，实际的图包含无限精细的细节，并且与图片不同，没有“连续”的点。□

示例3.11每个实数要么是有理数，要么是无理数，每个有理数都有唯一的最简形式表示为p/q。定义一个函数f：$\mathbf{R} \to \mathbf{R}$，通过

$$f(x) = \begin{cases} 1/q & \text{if } x = p/q \text{ in lowest terms,} \\ 0 & \text{if } x \text{ is irrational.} \end{cases} \tag{3.10}$$

图3.13展示了f的图形；由于打印的点无法分辨图中任意精细的细节，因此无法得到完全准确的图片。由于没有显示$q >$为40的点，因此在水平轴附近出现了白色带状区域。□

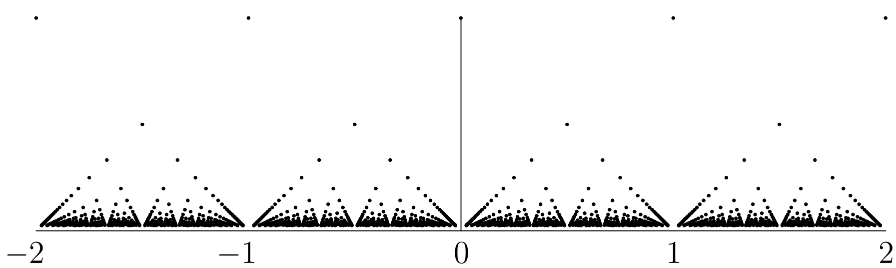


图3.13：分母函数。

示例3.12根据练习2.18，每个实数都有一个十进制展开，如果我们同意用一串9的无限序列来表示非零终止小数（例如0.25），那么这个展开是唯一的（如0.2499）。考虑到这个惯例（并写作“x”表示“x的十进制展开”），定义f : (0, 1) → $\mathbf{Z}$ 为

$$f(x) = \begin{cases} k & \text{if the digit 7 occurs exactly } k \text{ times in } x, \\ -1 & \text{if the digit 7 occurs infinitely many times in } x. \end{cases}$$

由于每个实数 x 都有一个唯一的十进制展开，数字 7 出现的次数是明确的。然而，对于特定的 x 选择，可能无法计算 $f(x)$；认为值 $f(\pi - 3)$ 是 -1，但这尚不清楚。

为了传达f的值是如何“混乱”地分布的，我们提出一个论点，即在每个开区间(a, b)——无论多小——函数f都能达到任意大的值。例如，在区间$(0, 10^{-10})$中，我们发现具有小数展开$x = 0.0\cdots07\cdots79\ldots$的数字，后面跟着一串$k$ 7，再后面跟着一个无限串的9；对于这个数，$f(x) = k$。应该清楚，本质上相同的推理适用于任意区间。这个函数的图像可以表示为一系列水平线，高度为-1，0，1，2等等，但有一个前提：这些线实际上并不是“实心”：每条垂直线恰好与图像相交一次。□

3.3 组合，迭代和逆元素

如果 $f: X \to Y$和 $g: Y \to Z$ 是函数（特别是，f 的像包含在 g 的定义域内），那么 g 和 f 的复合是函数 $g \circ f: X \to Z$，读作“g of f，”定义为对所有 $x \in X$，$(g \circ f)(x) = g\big(f(x)\big)$。换句话说，先对 f 应用 x，然后对输出应用 g。函数的复合是结合的，但通常不是交换的；一个复合可能被定义，而另一个则没有，即使两个复合都被定义，它们通常也是无关的。例如，如果 $f(x) = x + 1$（“加一”）和 $g(x) = x^2$ (“平方”)，那么

$$\begin{aligned}(g \circ f)(x) &= g(x + 1) = (x + 1)^2 = x^2 + 2x + 1,\\ (f \circ g)(x) &= f(x^2) = x^2 + 1.\end{aligned}$$

迭代

如果 $f: X \to X$，即 f 的定义域和值域是相同的集合，那么 f 可以反复与自己组合。(f 不一定是满射；为什么不是呢？) f 的第 n 次迭代是定义为“将 f 与自身组合 n 次”的函数。其形式递归定义是

$$f^{[0]} = I_X, \qquad f^{[n+1]} = f \circ f^{[n]} \quad \text{for } n \geq 0. \tag{3.11}$$

因此 $f^{[1]} = f$，$f^{[2]} = f \circ f$，$f^{[3]} = f \circ f \circ f$，以此类推。

序列　　由方程（3.9）定义的ence通过它获得　　　　评分

$$f(x) = \frac{1}{2}\left(x + \frac{2}{x}\right)$$

视为从（$0, +\infty$）到自身的映射。在此示例中，$x_0 = 2$，$x_{k+1} = f(x_k)$ 为 $k \geq 0$。一般来说，设 $f : X \to X$，并设一个“初始点” $x_0 \in X$ 已给出。由 $x_k = f^{[k]}(x_0)$ 定义的序列 $(x_n)_{n=1}^{\infty}$ 由映射 f (的重复应用下的 x_0 的后续位置组成，其“正向历史”)，集合 $\{x_k\}_{k=0}^{\infty} \subset X$ 是 x_0 在 f 迭代下的轨道。同时考虑 X 的所有点给出一个离散动力系统；集合 X 被视为“空间”，其点通过映射 f “混合”。特别感兴趣的是具有 $f(x) = x$ 的点 $x \in X$，即 f 的不动点。我们将在第4章中看到原因。

反转

如果一个函数 $f : X \to Y$ 被视为一个操作或过程，通常希望通过应用另一个函数来“撤销”f 的效果，即恢复每个输出值 $y = f(x)$ 的输入 x。并非每个函数都适合“求逆”。如果一个函数 $f : X \to Y$ 存在一个函数 $g : Y \to X$，使得

$$f \circ g = I_Y, \quad \text{that is, } (f \circ g)(y) = y \text{ for all } y \in Y, \tag{3.12}$$

和

$$g \circ f = I_X, \quad \text{that is, } (g \circ f)(x) = x \text{ for all } x \in X. \tag{3.13}$$

这两个方程在以下意义上是“对偶”的：同时交换 f 与 g 以及 X 与 Y 可以将每个方程转换为另一个方程。此外，它们在逻辑上是独立的，它们所指定的属性已经遇到过：

命题3.13。设$f : X \to Y$为一个函数。存在一个函数g满足方程（3.12），当且仅当f是满射，并且存在一个函数满足方程（3.13），当且仅当f是单射。

证明相当于重新表述定义；通过自己尝试证明这一点，你将比简单地阅读下面的证明学到更多。为了直观地理解命题的含义，考虑一个类比。假设你要搬到一套新公寓，并且从搬家公司购买了印有标签（“厨房”、“浴室”、“车库”等）的标签。你有几种类型的财产（餐具、玻璃杯、毛巾、衣服、书籍等），每个物品都被放入一个带有标签的盒子中。你的数学家朋友正在打包盒子，同时

你跑腿：每个盒子都根据其内容所属的房间进行标记。

待打包的物品类型是 X 的点，标签类型是 Y 的点，你朋友的标记方案是一个函数 f。你的目标是仅通过查看箱子上的标签来识别你的物品，即从其函数值中恢复 X 的点。在这种情况下，命题3.13的两个部分解决了以下问题:

- 对于每种房间标签，是否存在相应的盒子？如果标签函数是满射的，则存在一个函数 g，如（3.12）所示。
- 您能否仅通过查看标签就确定每个盒子的内容？如果标签函数是单射的，则存在一个函数 g，如（3.13）所示。

证明。如果存在一个满足方程（3.12）的函数，那么每个 $y \in Y$ 都在 f 的像中，因为元素 $x := g(y) \in X$ 由 f 映射到 y。反之，如果 f 是满射的，那么对于每个 $y \in Y$，集合 $f^{[-1]}(\{y\})$ 是非空的（根据定义）。因此，对于每个 $y \in Y$，可以选取一个元素 $x \in f^{[-1]}(\{y\})$，这样的选择族不过是一个函数 $g: Y \to X$。通过构造，方程（3.12）成立。

现在假设 f 是一一对应的。任意选取一个元素 $x_0 \in X$，并定义函数 g：$Y \to X$。

$$g(y) = \begin{cases} x & \text{where } f(x) = y, \text{ if } y \in f(X), \\ x_0 & \text{otherwise.} \end{cases}$$

这个处方定义了一个函数（即单值的），因为 f 是一一对应的，并且对于这个函数 g，显然方程（3.13）成立。反过来，假设方程（3.13）成立，并设 x_1 和 x_2 是 X 的元素，使得 $f(x_1) = f(x_2)$。我们想证明 $x_1 = x_2$。但这很清楚，因为

$$x_1 = g\big(f(x_1)\big) = g\big(f(x_2)\big) = x_2$$

通过方程（3.12）。 □

一个满足方程（3.12）的函数 g 被称为 f 的右逆，或 f^{-1} 的一个分支，而一个满足方程（3.13）的函数 g 是 f 的左逆。左逆很少出现，因为你可以用一个函数的像替换其值域，此时函数变为

全射。相比之下，f^{-1} 的分支在代数和三角学中自然出现。

一个函数可能有多个左逆或多个右逆；然而，如果一个函数 f 同时具有左逆 g 和右逆 h，那么 $g = h$:

$$g = g \circ (f \circ h) = (g \circ f) \circ h = h.$$

在这个事件中，命题3.13表明f是一个双射。因此，双射和可逆函数是同一件事。从概念上讲，双射不过是一种重新命名：X的元素是感兴趣的对象，而Y的元素是“标签”，双射f：$X \to Y$将一个唯一的标签与每个对象关联，并将一个唯一的对象与每个标签关联。无穷集合之间的双射一开始可能看起来很奇怪。以下例子给出了一些有趣的双射的小样本。

示例3.14恒等映射 $I_X : X \to X$ 对于每个非空集合 X 都是可逆的，并且是其自身的逆。如果 $a \in \mathbf{R}$，那么称为通过 a 平移的函数 $x \mapsto x + a$ 是一个双射，其逆是通过 $-a$ 平移。如果 $a \neq 0$，那么称为通过 a 缩放的函数 $x \mapsto ax$ 是一个双射，其逆是缩放$1/a$。类似函数在任意域中都是双射。每个单射函数都是到其像的双射。例如，由 $f(n) = 2n$ 定义的函数 $f : \mathbf{Z} \to \mathbf{Z}$ 是整数集和偶数集之间的双射。观察到一个无限集合可以与其自身的真子集一一对应。□

示例3.15 对数是一个双射 $L : (0, \infty) \to \mathbf{R}$，使得

$$L(xy) = L(x) + L(y) \qquad \text{for all } x, y \in \mathbf{R}.$$

第12章推导了对数的存在。在历史上，对数之所以重要，是因为它们将乘法转换为加法，前提是存在一种有效的从 $x \in (0, \infty)$ 到 $L(x) \in \mathbf{R}$ 的转换方法。在电子计算机时代之前，这种转换是通过对数表和计算尺完成的。对数在纯数学、科学和工程中都具有极其重要的意义；恒星亮度、响度（以分贝为单位）和酸度（pH值）都是使用对数尺度来测量的。□

示例3.16 存在计算方法来寻找由公式定义的函数的逆。在高中课程中，通常的做法是“在方程 $y = f(x)$ 中交换 x 和 y，然后

求解 y。”等价于求解 $y = f(x)$ 关于 x。这基本上是正确的，尽管必须注意定义域和值域，正如这个例子所示。

Let $f : [-1, 0] \to [0, 1]$ be defined by $f(x) = x^2$. This function is one -to-one and onto. Formal solution for x gives $x = \pm\sqrt{y}$. This “equation” (really a pair of equations) does not determine f^{-1}, though it narrows d own the possibilities enough that the inverse can be found by inspection. B ecause the domain of f is $[-1, 0]$, the range of f^{-1} must be this same inter val. Therefore, $f^{-1} = -\sqrt{\,}$, since by definition the square root function tak es non-negative values. □

示例3.17 集合 $\{a, \dots, z\}$ 和集合 $\{1, \dots, 26\} \subset \mathbf{N}$ 之间的“显然”双射可以用来将消息编码为数字。解码消息相当于反转最初编码消息的双射。更复杂的代码可以允许大写和小写字母、标点和数字。所谓的ASCII字符编码（在美国以外称为“ISO 8859-I”）就是这样一种对应关系，并且广泛用于文本存储。□

单调函数的逆

一个严格单调函数是单射的，因此是到其像上的双射。如果 f 是递增的，那么 f^{-1} 也是递增的：设 $y_1 < y_2$ 是像中的元素，并设 $x_i = f^{-1}(y_i)$。以下不等式 $x_1 < x_2$ 或 $x_1 > x_2$ 必须成立。因为 f 是递增的，所以第二种可能性不会发生。因此，如果 $y_1 < y_2$，那么 $f^{-1}(y_1) < f^{-1}(y_2)$。同样的论证可以证明，如果 f 是递减的，那么 f^{-1} 也是递减的。请注意，一个单射函数通常不是单调的。

排列与基数

回忆起对于 $n \in \mathbf{N}$，相应的初始段 $\underline{\mathbf{n}}$ 是集合 $\{1, \dots, n\}$。从 $\underline{\mathbf{n}}$ 到自身的双射称为 n 字母上的排列。在 n 个字母上有 $n!$ 个排列。直观上相当明显（可以通过数学归纳法证明），如果且仅当 $n \le m$ 存在，则存在一个注入 $i : \underline{\mathbf{n}} \to \underline{\mathbf{m}}$。

康托尔比较无限集“大小”的想法推广了这一点；如果存在一个双射 $f : X \to Y$，则两个集合 X 和 Y 具有相同的基数。更一般地，“X 的基数不大于

集合 Y 的基数当且仅当存在一个注入 $i: X \to Y$。(根据命题 3.13，这等价于存在一个满射 $p: Y \to X$。)如同例 3.14，一个无限集合的基数可以与一个真子集的基数相同。根据定义，一个集合 X 是可数的，如果存在一个双射 $f: \mathbf{N} \to X$，并且是至多可数的，如果它是有限的或可数的。Cantor 最初认为所有无限集合都是可数的。后来他证明了相反的结论，既通过一个一般性的论证（参见定理 3.18），又在一种引人注目的特殊情况中（定理 3.19）。Cantor 的工作在 19 世纪末的一些数学家那里遭到了激烈的反对，但现在被认为是根本正确的。

定理3.18。设X为一个集合，设$\mathcal{P}(X)$为其幂集，即X的所有子集的集合。那么不存在满射p：$X \to \mathcal{P}(X)$；每个集合的基数严格小于其幂集的基数。

证明。康托尔证明了如果 $p: X \to \mathcal{P}(X)$ 是一个任意函数，那么存在 $\mathcal{P}(X)$ 中的一个元素不在其像中。映射 p 将每个 $x \in X$ 与一个子集 $p(x) \subset X$ 相关联。对于每个 x，可以问 $x \in p(x)$ 是否成立，答案要么是“是”，要么是“否”（取决于 x 和 p）。考虑集合

$$A = \{x \in X \mid x \notin p(x)\} \subset X;$$

集合 $A \in \mathcal{P}(X)$ 依赖于 p，但定义明确。

对于每个 $x \in X$，要么 $x \in A$，要么 $x \notin A$。如果 $x \in A$，那么 $x \notin p(x)$，因此 $p(x) \neq A$ 作为集合（一个包含 x，另一个不包含）。另一方面，如果 $x \notin A$，那么 $x \in p(x)$，再次 $p(x) \neq A$。总之，如果 $x \in X$，那么 $p(x) \neq A$，即 p 不是满射的。 □

这个定理表明，对于每一个集合——可能是无限的——都存在另一个具有严格更大基数（{v*}）的集合！也许我们可以同情那些感到只有疯狂或语言迷雾（无限递增的更大无限）才存在于这个方向的数学家。以下定理，同样归功于康托尔，表明无理数比有理数“更多”。

定理3.19.有理数集是可数的；实数集是不可数的。具体来说，存在一个双射 $f: \mathbf{N} \to \mathbf{Q}$，但不存在一个满射 $p: \mathbf{N} \to \mathbf{R}$。

证明。（草图）从概念上讲，一个双射 $f: \mathbf{N} \to \mathbf{Q}$ 是列出 $\mathbf{Q}$ 中所有元素的方法。我们首先从 $\mathbf{N}$ 构造一个满射

到整数对（p,q）的集合中，其中 $q >$ 为 0，参见图 3.14，然后“删除”不是最简形式的对。这给出了所需的双射。注意，**Q** 按照升序排列 $<$ 并不给出双射，因为没有所谓的“连续”有理数对。

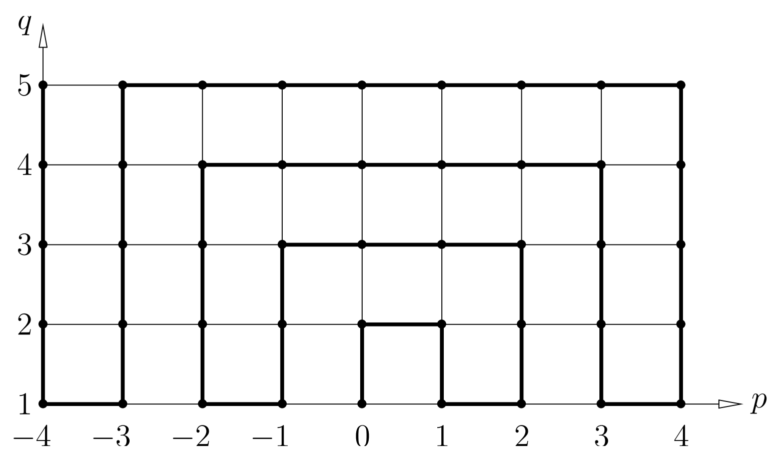


图3.14：从**N**到**Q**构造双射

根据大多数人的直觉，有理数比自然数多，因为每对自然数之间有无限多个有理数。图3.14中描绘的双射表明这种直觉是错误的；在计数无限集合时，元素的枚举顺序很重要，因为一个无限集合可以与一个适当的子集建立双射对应。

要证明 **R** 不可数似乎在上述论证之后变得不可能；如果我们找不到双射，也许我们只是不够聪明！正如您将注意到的，我们在这里使用的论证完全不同。只需证明 **R** 的某个子集不可数就足够了。考虑由只包含数字 0 和 1 的十进制表示的实数集合 X，并令 $f\colon \mathbf{N} \to X$ 为一个任意的映射。按照 (3.14) 中的图像列出像中的元素，该图像描述了 f 的“典型”选择：

$$\begin{aligned} f(1) &= 0.\mathbf{1}101110\ldots \\ f(2) &= 0.0\mathbf{1}00100\ldots \\ f(3) &= 0.11\mathbf{0}1100\ldots \\ f(4) &= 0.101\mathbf{0}010\ldots \end{aligned} \quad \rightsquigarrow \quad x = .0011\ldots \tag{3.14}$$

要证明 f 不是满射，只需构造一个不在 f 的像中的数 x。考虑第 k 位小数是第 k 个数 $f(k)$；如果这个位是 0，则将 k 位小数 x 设为 1，反之亦然。然后 $x \in X$，因为它的十进制展开只包含 0 和 1，但 x 不在 f 的像中，因为 x 和 $f(k)$ 在第 k 位小数上不同。我们已经证明了 f 不是满射；由于 f 是任意的，因此不存在满射 $f : \mathbf{N} \to X$，更不用说满射 $\mathbf{N} \to \mathbf{R}$。 □

大多数人第一次跟随这个证明时立即会问，“为什么不将新数字添加到列表中？”要理解这为什么是一个无关紧要的点，你必须正确解释这个主张：任意的映射 $f : \mathbf{N} \to X$ 不是满射。在构造数字 x 之前就指定了函数 f。你完全可以理解康托尔批评者的感受，但这个定理与定义完全一致。

同构

两个“抽象等价”的数学结构通常被认为是“相同的”。对于每个数学结构，都有一个同构的概念，它精确地定义了“抽象等价”的含义。之前遇到的数学结构包括集合、交换群、域和有序域。下一个例子将详细解释集合和交换群中的同构。

两个集合 X 和 Y 是同构的，当且仅当存在一个双射 $\phi : X \to Y$。直观上，一个集合除了“元素的数量”之外没有其他属性（这个数量可能是有限的或无限的）。映射 ϕ 是 X 和 Y 之间的同构，如上所述，“重命名” X 的元素。这些集合

$$X = \{0, 1\}, \qquad Y = \big\{\varnothing, \{\varnothing\}\big\}, \qquad \text{and} \quad Z = \{\text{True}, \text{False}\}$$

它们是相互同构的。如果 X 和 Y 是具有多个元素的同构集，那么它们之间有多个同构，通常没有理由选择一个而不是另一个。然而，如果 $X = Y$，那么恒等映射 I_X 在某种意义上是同构的“自然”选择。你可能会说[3]，具有相同元素数量的集合是同构的，但有些集合比其他集合更同构。

[3]With apologies to G. Orwell.

情况类似，但更有趣的是，当考虑具有附加结构的集合时。在这种情况下，“同构”应该“保留附加结构。”假设 $(\mathbf{G}, +)$ 和 $(\mathbf{H}, \oplus)$ 是交换群。这意味着 $\mathbf{G}$ 是一个非空集合，并且 $+$ 是一种“添加”$\mathbf{G}$ 的两个元素以得到第三个元素的方法，受第 51 页上的公理 A.1–A.4 的约束。对于 $\mathbf{H}$ 和 $\oplus$ 也适用类似的评论。$(\mathbf{G}, +)$ 和 $(\mathbf{H}, \oplus)$ 之间的同构是一个双射 $\phi : \mathbf{G} \to \mathbf{H}$，使得

$$\phi(x + y) = \phi(x) \oplus \phi(y) \qquad \text{for all } x, y \in \mathbf{G}. \tag{3.15}$$

方程（3.15）表明，在 ϕ 下，操作 $+$ 和 $\oplus$ 对应；在 $\mathbf{G}$ 中添加并重新标记（左侧）与重新标记然后添加 $\mathbf{H}$ (右侧) 是相同的。就交换群的性质而言，$(\mathbf{G}, +)$ 和 $(\mathbf{H}, \oplus)$ 是不可区分的。它们的元素集合和/或它们的“加法”操作可能看起来非常不同，但抽象上结构是相同的。对数（例3.15）是正实数乘法群 $(\mathbf{R}^+, \cdot)$ 和实数加法群 $(\mathbf{R}, +)$ 之间的同构 L。

同构的概念以直接的方式扩展到更复杂的数学结构。有序域($\mathbf{F}, +, \cdot, <$)由一个非空集合$\mathbf{F}$、两个运算$+$和$\cdot$以及一个在$\mathbf{F}$上的关系$<$组成，这些关系满足公理。有序域$(\mathcal{F}, \oplus, \odot, \prec)$与$(\mathbf{F}, +, \cdot, <)$（作为有序域）同构，如果存在一个双射$\phi$：$\mathbf{F} \to \mathcal{F}$，使得方程(3.15)的类似物对算术运算成立，并且顺序关系在以下意义上相对应：$x < y$当且仅当$\phi(x) \prec \phi(y)$。如上所述，同构的有序域在有序域的问题上是抽象上不可区分的。

一旦这些概念被理解，就可以精确地陈述定理2.30（实数的存在性和唯一性）。首先，存在一个完备有序域$(\mathbf{R}, +, \cdot, <)$。说“$\mathbf{R}$包含$\mathbf{Q}$”意味着存在一个注入$i : \mathbf{Q} \to \mathbf{R}$，它是一个同构（有序域的同构）到其像。唯一性意味着每个完备有序域$(\mathcal{R}, \oplus, \odot, \prec)$作为有序域同构于$(\mathbf{R}, +, \cdot, <)$。

3.4 线性算子

通过引入线性代数的一些术语，可以获得大量的概念经济学。微积分的基本运算——积分和微分——可以被视为“函数”，其定义域是函数空间。

向量空间

让 $X \subset \mathbf{R}$ 和让 $\mathscr{F}(X, \mathbf{R})$ 表示 X 上的实值函数集合。当 X 是一个有限集时，函数空间 $\mathscr{F}(X, \mathbf{R})$（本质上）是 $\mathbf{R}^n$，我们将一般元素视为“由 X 上的点索引的实数列表”。

我们感兴趣的运算包括函数的加法和“标量乘法”。如果 f 和 g 是 $\mathscr{F}(X, \mathbf{R})$ 的元素，并且 c 是一个实数，那么我们通过以下方式定义函数 $f+g$ 和 $cf \in \mathscr{F}(X, \mathbf{R})$:

$$\begin{aligned}(f+g)(x) &= f(x)+g(x) \\ (cf)(x) &= c \cdot f(x)\end{aligned} \qquad \text{for all } x \in X. \tag{3.16}$$

$\mathscr{F}(X, \mathbf{R})$的集合与这些代数运算一起是一个向量空间的例子。存在一个类似于域的定义的公理化定义，你将在线性代数课程中遇到。

一个非空子集 $V \subset \mathscr{F}(X, \mathbf{R})$ 是一个向量子空间，如果满足以下两个条件:

- (闭包于加法{v*} 如果 f 和 $g \in V$，那么 $f+g \in V$
- (闭包在标量乘法下{v*} 如果 $f \in V$，那么 $cf \in V$对于所有 $c \in \mathbf{R}$

例如，X上的多项式函数集是$\mathscr{F}(X, \mathbf{R})$的向量子空间，步骤函数集也是如此。$X$上的指示函数集不是向量子空间：两个指示函数之和通常不是指示函数。

线性映射

让 V和 W 是 $\mathscr{F}(X, \mathbf{R})$ 的向量子空间。映射 $L : V \to W$ 以一个函数 f 作为输入，并返回一个函数 Lf 作为输出。它是

习惯上写 Lf 而不是 $L(f)$ 以避免过多的括号。一个映射 $L: V \to V$ 被称为 V 上的运算符。

映射 $L: V \to W$ 是线性的，如果

$$\begin{aligned} L(f+g) &= Lf + Lg \\ L(cf) &= c \cdot Lf \end{aligned} \qquad \text{for all } f \text{ and } g \text{ in } V, \text{ all real } c. \tag{3.17}$$

您可以将线性映射视为“尊重向量空间结构”的一种映射。线性泛函[4] 是一个线性映射 $T: V \to \mathbf{R}$。

示例3.20 固定 $a \in \mathbf{R}$ 并定义一个算子 L_a: $\mathscr{F}(\mathbf{R},\mathbf{R}) \to \mathscr{F}(\mathbf{R},\mathbf{R})$ 为 $L_a f(x) = f(x-a)$。L_a 的作用是将域中的 f 向右“移动” a。线性性是直接的，你应该检查。

$\square$

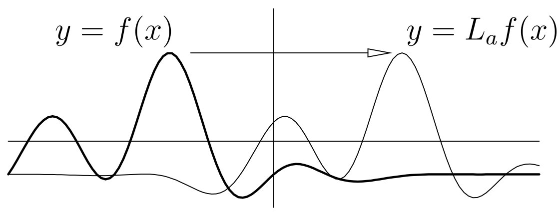


图3.15：翻译算子。

示例3.21一个与之密切相关的是反射算子 {v*}，由 $Rf(x) = f(-x)$ 定义。在几何上，R 将 f 的图形关于垂直轴进行反射。$\square$

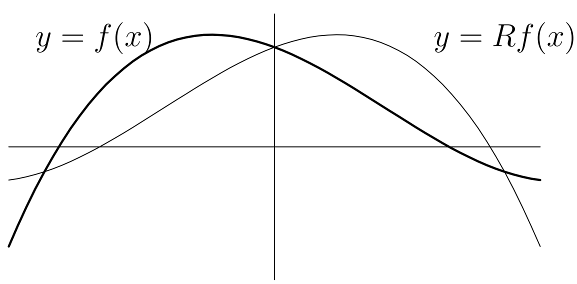


图3.16：域反射算子。

[4]In mathematics, “functional” is a noun!

示例3.22 修复 $x \in X$；评估泛函 $\mathrm{ev}_x : \mathscr{F}(X, \mathbf{R}) \to \mathbf{R}$ 定义为 ev $_x(f) = f(x)$，即“在 x 处对 f 的评估”。函数加法和数乘的定义说明 ev_x 是一个线性泛函。□

操作符 S 由 $Sf(x) = f(x)^2$ 定义的不是线性的；例如，将 f 乘以 2 并应用 S 将输出乘以 $2^2 = 4$。

函数的对称性

上述引入的反射和翻译算子使我们接触到一些有趣的函数类。

偶函数和奇函数

设 $A \subset \mathbf{R}$ 为形如 $[-a, a]$ 的区间，对于某个 $a > 0$。如果函数 f : $A \to \mathbf{R}$ 是偶函数，则

$$f(-x) = f(x) \qquad \text{for all } x \in A, \tag{3.18}$$

并且是奇数当且仅当

$$f(-x) = -f(x) \qquad \text{for all } x \in A, \tag{3.19}$$

查看图3.17。术语来源于单项式函数 $x \mapsto$

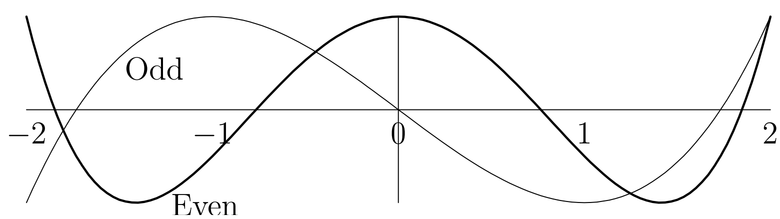


图3.17：偶函数和奇函数。

x^k 对于 k 是一个正整数；在上述意义上的单项式的“奇偶性”与指数 k 作为整数的奇偶性相同，因为

$$(-x)^k = (-1)^k x^k = \begin{cases} x^k & \text{if } k \text{ is even,} \\ -x^k & \text{if } k \text{ is odd.} \end{cases}$$

均匀性和奇偶性是“全局”属性：它们取决于f在整个域上的行为。

存在一个关于反射算子的美丽解释，即例3.21：一个函数是偶函数当且仅当$Rf = f$ (f在R)下不变，是奇函数当且仅当$Rf = -f$ (f在R)下反不变。对于每个函数f，算子R交换f和Rf，因为$R(Rf) = f$。

引理3.23.偶函数和奇函数在$A = [-a, a]$上的空间是$\mathscr{F}(A, \mathbf{R})$的向量子空间。

证明。这是对 R 线性性质的重新表述：如果 f 和 g 是偶数且 c 是实数，那么

$$R(f+g) = Rf + Rg = f + g, \qquad R(cf) = c \cdot Rf.$$

因此 $f + g$ 和 cf 是偶数，所以偶函数的集合是 $\mathscr{F}(A, \mathbf{R})$ 的向量子空间。奇函数的证明基本上是相同的。

□

由于偶函数之和是偶函数，一个多项式是偶函数当且仅当每个项的次数都是偶数。逆命题也成立，见下面的命题3.24。如果将“偶”一词替换为“奇”，这些论述仍然成立。

每个在 $\mathbf{R}$ 上的常函数都是偶函数。唯一的奇函数是零函数；实际上，零函数很容易被看出是唯一既是偶函数又是奇函数的函数。符号函数

$$\operatorname{sgn}(x) = \begin{cases} \dfrac{x}{|x|} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases} = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{3.20}$$

是奇数。

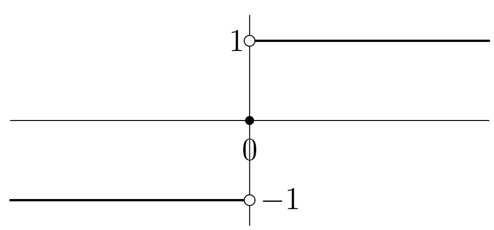


图3.18：符号函数。

大多数函数既不是偶函数也不是奇函数。然而，对称域上的每个函数 f 都可以表示（唯一地）为以下形式的和

偶函数 f_{even} 和奇函数 f_{odd}。事实上，由以下定义的函数

$$
\begin{aligned}
f_{\text{even}}(x) &= \tfrac{1}{2}\big(f(x) + f(-x)\big) \\
f_{\text{odd}}(x) &= \tfrac{1}{2}\big(f(x) - f(-x)\big)
\end{aligned}
\tag{3.21}
$$

这些公式是通过写出 $f = f_{\text{even}} + f_{\text{odd}}$ 并使用方程 (3.18) 和 (3.19) 得到的。观察可知，当 f 的奇数部分 f_{odd} 是零函数时，f 是偶数，并且当 f 的偶数部分恒为零时，f 是奇数。在 R 的术语中，方程 (3.21) 表示

$$f_{\text{even}} = \tfrac{1}{2}(f + Rf), \qquad f_{\text{odd}} = \tfrac{1}{2}(f - Rf).$$

为了从 f 获得一个偶函数，我们平均 f 和 Rf，而要获得一个奇函数，我们平均 f 和 $-Rf$：f 的偶部和奇部是通过“在 R 的作用下进行加权平均”获得的。

为了完成关于函数奇偶性的讨论，我们描述了偶数和奇数多项式。

命题3.24。设$p : \mathbf{R} \to \mathbf{R}$为一个多项式函数。那么$p$是偶函数当且仅当$p$的每一项都有偶数次幂，$p$是奇函数当且仅当每一项都有奇数次幂。

简洁地（如果不太透明），p 是偶数当且仅当存在一个多项式 q，使得对所有 $x \in \mathbf{R}$，有 $p(x) = q(x^2)$，并且 p 是奇数当且仅当存在一个多项式 q，使得对所有 $x \in \mathbf{R}$，有 $p(x) = x\,q(x^2)$。

证明。假设 p 是一个多项式，并令 p_e 和 p_o 分别表示偶次项之和和奇次项之和。如前所述，这些多项式函数分别是偶函数和奇函数，它们的和是 p。根据唯一性，它们必须是 p 的偶部和奇部。

□

周期函数

设 ℓ 为一个非零实数。函数 $f : \mathbf{R} \to \mathbf{R}$ 以 ℓ 为周期——或者说 ℓ-周期性，如果

$$f(x - \ell) = f(x) \qquad \text{for all } x \in \mathbf{R}. \tag{3.22}$$

关于示例3.20中的翻译算子 T_ℓ，f 当且仅当 ℓ 是 $T_\ell f = f$ 的周期。

通过归纳，对所有 $n \in \mathbf{Z}$ 有 $f(x+n\ell) = f(x)$；因此，一个 ℓ-周期函数的图像由长度为 ℓ 的“波形”组成，无限重复。将一个 ℓ-周期函数限制在长度为 ℓ 的区间上称为一个周期。显然，一个周期函数由其每个周期完全确定。反之，给定一个长度为 ℓ 的半开区间的函数，存在一个唯一的 ℓ-周期扩展。

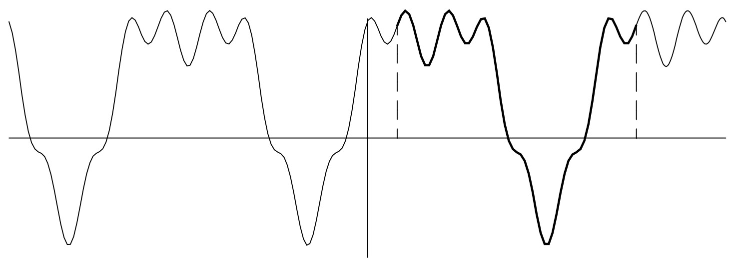


图3.19：一个周期函数及其一个完整周期。

示例3.25 Charlie Brown函数cb：$\mathbf{R} \to \mathbf{R}$是绝对值函数在$[-1, 1)$上的周期扩展，见图3.20。注意cb是分段多项式，实际上，是分段线性。□

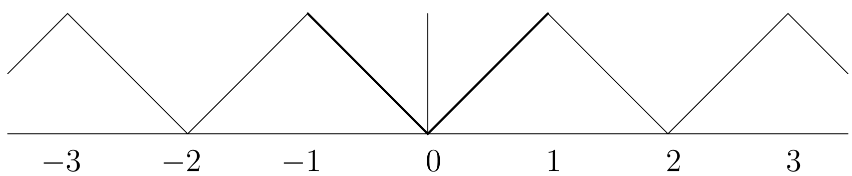


图3.20：查理·布朗函数。

正负部分

设 f 为定义在任意集合 X 上的实值函数。f 的正部分是函数 f_+: $X \to \mathbf{R}$，定义为

$$f_+(x) = \max\bigl(f(x), 0\bigr) = \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{if } f(x) < 0 \end{cases}$$

同样，f的负部分由$f_-(x) = \min\big(f(x), 0\big)$定义。你应该在图3.19中绘制函数的正负部分，以确保你理解定义。

注意，对于所有 x，有 $f(x) = f_+(x) + f_-(x)$；每个实值函数都是非负函数的差。这个观察结果在以后有有趣且重要的应用。例如，我们将能够证明某个大类的函数可以写成单调函数的差。

练习

练习3.1 打完一封长信后，你意识到你系统地交换了“there”和“their”这两个词。幸运的是，你的文本编辑器可以替换字符串的所有出现。你首先将“there”替换为“their”，然后将“their”替换为“there”。这会有预期的效果吗？将这些替换解释为从集合{there, their}到自身的函数。这些函数是一对一的吗？你该如何成功使用替换来交换“there”和“their”的所有出现？ $\diamond$

练习3.2 证明命题3.3。你必须使用函数和前像的定义来建立三个集合的包含关系。$\diamond$

练习3.3 给出一个函数 $f: X \to Y$和一个特定的 $A \subset X$，使得包含 $A \subset f^{[-1]}\big(f(A)\big)$ 是适当的。提示：你的函数不能是一对一的。图3.8可能会有帮助。$\diamond$

练习3.4 设A和B是$\mathbf{R}$的子集，设χ_A和χ_B是它们的指示函数，方程（3.7）。建立以下：

(a) $1 - \chi_A = \chi_{(\mathbf{R}\setminus A)}$，$A^c$的指标。

(b) $\min(\chi_A, \chi_B) = \chi_{A\cap B} = \chi_A \cdot \chi_B$. (c) $\max(\chi_A, \chi_B) = \chi_{A\cup B} = \chi_A + \chi_B - \min(\chi_A, \chi_B)$. (d) $\chi_A + \chi_B \pmod 2) = \chi_{A\triangle B}$. (参见练习1.2)

在词义上，集合上的布尔运算和特征函数密切相关。$\diamond$

练习3.5 本练习描述了阶梯函数。

(a) 对于 $k = 1, \ldots, n$，令 I_k 为一个区间，χ_k 为 I_k 的特征函数，c_k 为一个实数。使用练习3.4和关于 n 的归纳法来证明

$$f(x) = \sum_{k=1}^{n} c_k \chi_k(x) \tag{‡}$$

这是一个在 $\mathbf{R}$ 上的步函数。与方程（3.8）的区别在于，这里的区间不必两两不相交。

(b) 每个步骤函数都可以写成 (‡) 的形式吗？如果是，请证明；如果不是，需要修改集合 I_k 的哪些性质？(c) 部分 (b) 中的表示是唯一的吗？如果是，请证明；如果不是，需要修改集合 I_k 的哪些性质？

可能有助于绘制一些形式为 $c_1\chi_1 + c_2\chi_2$ 的函数，其中集合 I_1 和 I_2 是，或不是，不相交的。$\diamond$

练习3.6 在每一部分中，$f : X \to Y$ 是一个函数，$A \subset X$。确定以下每个蕴含是否有效（给出证明）或无效（找到反例）。

(a) 如果 f 是单射的，那么 $f|_A$ 也是单射的。

(b) 如果 $f|_A$ 是单射的，那么 f 也是单射的。

(a) 如果 f 是满射，那么 $f|_A$ 是满射。

(a) 如果 $f|_A$ 是满射，那么 f 是满射。

可能考虑逆否命题会有所帮助。 $\diamond$

有理函数与代数函数

练习3.7 令 $S^1 \subset \mathbf{R}^2$ 表示单位圆，令 $X \subset S^1$ 为点 (0, 1) 的补集。定义一个函数 $p : \mathbf{R} \to X$ 如图3.21所示。在几何上，通过一条直线将 (0, 1) $\in S^1$ 与 $(t, 0)$ 连接起来，并令 $p(t) = (x, y)$ 为与圆的交点。

(a) 使用相似三角形找到 (x, y) 关于 t 的公式。

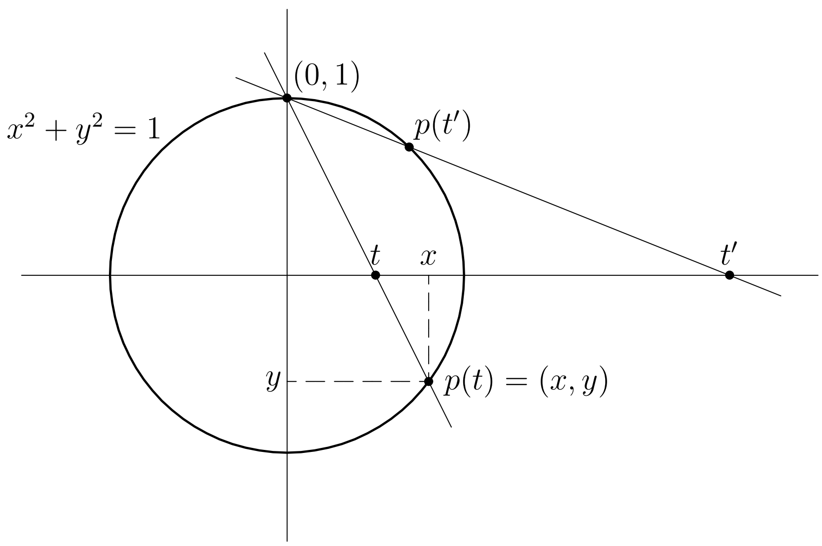


图3.21：正射投影。

(b) 证明 {v*} 在几何和代数上都是一一对应和满射的。找到一个公式来表示 p^{-1}（，即用 t 表示 x 和 y。）映射 p^{-1} 被称为球面投影。(c) 使用部分 (b) 来证明 t 是有理数当且仅当 x 和 y 是有理数。因此，球面投影表征了圆上的"有理点"。(d) 证明在球面投影下，映射 $f(t)=1/t$ 对应于通过水平轴的圆的反射。如果可能，给出代数和几何证明。(e) 证明有理映射 $f(t)=(t-1)/(t+1)$ 对应于圆逆时针旋转四分之一。提示：旋转映射 $(x,y)\mapsto(-y,x)$。

部分（d）表明，人们可能会说$1/0=\infty$和$1/\infty=0$。与第4章关于射无限极限的部分进行比较。$\diamond$

练习3.8 证明每个有理函数都是代数的。（形式上这是显而易见的，但请确保考虑到精确的定义，包括定义域和值域。）$\diamond$

练习3.9 设$F(x,y)=1+y+xy+xy^2$。求由F隐式定义的所有代数函数，并绘制零点集$Z(F)$。（建议：使用二次公式。）$\diamond$

练习3.10 设$F(x,y)=(x^2+y^2)^2-(x^2-y^2)$。找出由$F$定义的所有代数函数，并在图3.12中找到它们的图形。$\diamond$

练习3.11 绘制{v*}$1+x)(k+x)-y^2=0$的轨迹，对于$k=-1$，0和1。（先绘制$y=x(1+x)(k+x)$的图形可能会有帮助。）$\diamond$

练习3.12 设$\mathbf{F}$为一个域，设p和d是定义在$\mathbf{F}$上的非常数多项式，且$d<$的次数小于p的次数。证明存在唯一的定义在$\mathbf{F}$ (上的多项式q和r，使得商为)，余数为)，且满足

$$p(x)=d(x)q(x)+r(x).$$

提示：模仿定理3.6的证明。 $\diamond$

逆元

练习3.13 设$f:\mathbf{R}\to[0,\infty)$为平方函数。函数$\sqrt{}$在$x$处的值是$x$的非负平方根是$f^{-1}$的一个分支。证明$-\sqrt{}$是$f^{-1}$的另一个分支。找出此函数的所有分支。（存在许多“不连续”的f^{-1}分支。）$\diamond$

练习3.14 在这个练习中，你将建立一个有界区间与 $\mathbf{R}$ 之间的双射。定义 $f:(-1,1)\to\mathbf{R}$ 为 $f(x)=x/(1-x^2)$；参见图3.22以了解 f 的图形。

(a) 设置 $y=f(x)$ 并解出 x。建议：乘以 $(1-x^2)$ 并重新排列以得到二次方程 $yx^2+x-y=0$。如果 $y\neq 0$，则适用二次公式。

(b) 在部分(a)中，你找到了f的两个形式逆元，对应于二次公式中的两个根号符号。你知道f^{-1}必须在f的定义域中取值，即在区间$(-1,1$)内。哪种符号的选择是正确的？当$y=$等于0时会发生什么？将每个符号的选择与图3.22中的某一部分相对应。

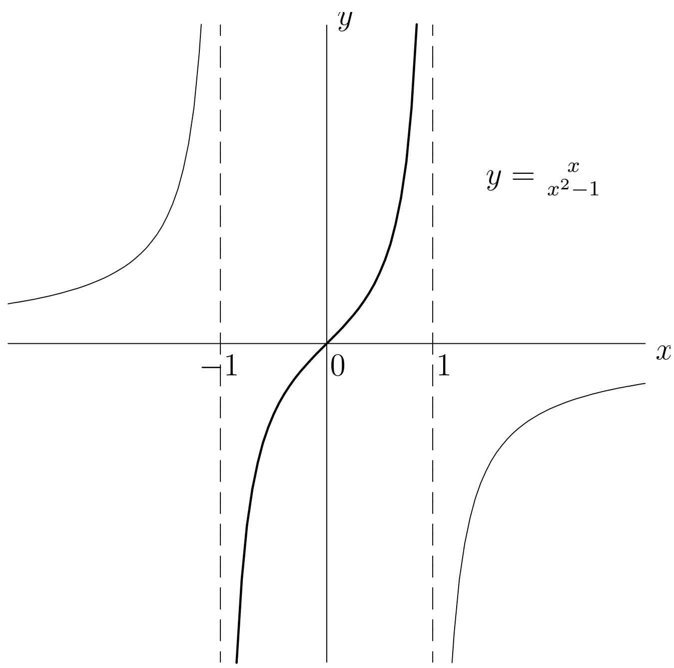


图3.22：一个从有界区间到**R**诱导双射的函数。

(c) 在这个阶段，已经找到了一个关于 f^{-1} 的假设公式。验证你找到的公式确实给出了 f 的双向逆。也就是说，直接验证方程（3.12）和（3.13），或者通过一般推理证明它们成立。

建议：如果 $y > 0$，则 $1 + 4y^2 < 1 + 4y + 4y^2 = (1 + 2y)^2$，所以

$$0 < \frac{-1 + \sqrt{1 + 4y^2}}{2y} < 1.$$

◇

函数的对称性

练习3.15 找出 $p(x) = x(x-1)^2$ 的偶部和奇部。找出 p 的正部和负部；将你的答案写成分段多项式函数。◇

练习3.16 找出 $p(x) = (1-x)^4$ 的偶数部分和奇数部分。提示：使用二项式定理展开 p。 ◇

练习3.17 假设 f 是偶数且 g 是奇数。关于它们的乘积 fg 你能说什么？如果它们都是奇数呢？证明你的所有主张。$\diamond$

练习3.18 完成引理3.23的证明，通过证明奇函数集在加法和数乘下是封闭的。$\diamond$

在每个以下练习中，T_ℓ 是由 $T_\ell f(x) = f(x-\ell)$ 定义的平移算子。

练习3.19 设$f = \chi_{\mathbf{Q}}$是$\mathbf{Q}$的特征函数，且ℓ是有理数。证明f是ℓ-周期的。$\diamond$

练习3.20 证明ℓ-周期函数的集合是$\mathscr{F}(\mathbf{R},\mathbf{R})$的向量子空间。$\diamond$

练习3.21 一个函数是"ℓ反周期"的，如果 $T_\ell f = -f$。证明这样的函数是2ℓ周期的。$\diamond$

练习3.22 假设f是ℓ-周期的。证明f的偶部和奇部是ℓ-周期的。$\diamond$

练习3.23 假设 f 是1周期性的，并且 g 是 ℓ-周期性的。

(a) 证明如果 ℓ 是有理数，那么 $f+g$ 是周期的。建议：将 $\ell = p/q$ 写成最简形式。(b) 假设 1 和 ℓ 是 f 和 g 的最小正周期。证明如果 ℓ 是无理数，那么 $f+g$ 不是周期的。

部分（b）需要认真使用 $\mathbf{Q}$ 的结构。 $\diamond$

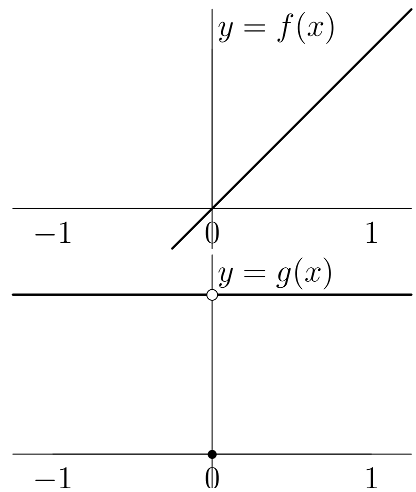
第四章

极限与连续性

“极限”的概念将分析（包括微分学的数学分支）与代数区分开来。极限的历史动机和主要实际用途是在各种情况下为“$0/0$”或“$0\cdot\infty$”等表达式赋予意义。正如我们在第一章中看到的，描述“某一时刻”的运动会导致形如（行驶距离）/（经过时间)$=$0/0的差商，而阿基米德的“穷竭法”（使他能够“分割”圆盘成矩形，见第13章）相当于将无限多个“无限薄的”三角形或矩形的面积相加，其总面积形式上为$0\cdot\infty$。

一个极限是在某些假设下，赋给函数 f 在某一点 a 的一个数。然而，与只考虑定义域中单个点的函数值 $f(a)$ 不同，如果存在)，f 在 a (的极限编码了 f “靠近” a 的行为，因此不能仅通过考虑有限多个点的值来确定！对于“连续”函数（包括多项式、三角函数和指数函数），“f 在 a 的极限”与 $f(a)$ 一致。一般来说，可能在函数值未定义的点处存在极限，或者函数值和极限在一点上可能都存在但不相等。在我们给出任何精确定义之前，让我们考虑两个简单但具有说明性的例子：

$$f(x)=x, \qquad g(x)=\begin{cases}1 & \text{if } x\neq 0\\ 0 & \text{if } x=0\end{cases}, \qquad x\in\mathbf{R}.$$



立即计算得出 $f(0) = g(0) = 0$；这些函数在原点处消失。如果我们尝试量化原点附近的性质，那么（在某种意义上，我们尚未明确）对于 $|x| \simeq 0$，我们有 $f(x) \simeq 0$ 和 $g(x) \simeq 1$。仔细思考这样一个断言的确切含义是一个非常有益的哲学练习。几分钟的思考应该会让你相信，仅考虑有限个函数值无法捕捉 f “靠近”（但不包括）a 的行为。相反，为了研究 f “靠近” a 的行为，我们将 f 限制在关于 a 的任意开区间内，并考虑限制的像。因此，发展适合研究函数值集合的符号对我们来说非常有利。第一个工具是 A 符号，我们在第二章中遇到过。下面介绍的两个辅助符号，“大 O” 和 “小 o”，也将扮演重要的角色。

在整个本章中，$f\colon X \to \mathbf{R}$ 是一个定义域为实数集（通常为一个区间）的实值函数。我们将使用 $\mathbf{R}$ 的顺序性质来定义极限的概念，并在证明极限定理时使用。可以在不进行域排序的情况下定义极限，但这样会有额外的技术开销，我们希望避免。

4.1 消失顺序

在分析中，我们允许自己稍微马虎一些；我们通常不在乎是否能精确解出方程（无论这意味着什么），我们只关心已知存在一个解，并且（比如说）它在3.14和3.1416之间。

存在一些演算（也称为计算过程），允许我们忽略不感兴趣的细节，并集中关注感兴趣的粗略细节。

A 符号综述

回忆一下，表达式 $f = A(\varepsilon)$，读作"f 的绝对值至多为 ε"，意味着对于 f 的定义域中的所有 x，都有 $|f(x)| \leq \varepsilon$。更一般地，如果 g 是一个其定义域包含 f 的定义域的函数，那么说 $f = A(g)$ 意味着对于所有 x，都有 $|f(x)| \leq |g(x)|$。例如，我们在 $(-1, 1)$ 上有 $x^2 = A(x)$，因为 $x^2 \leq |x|$ 对于 $-1 < x < 1$。

我们的这个术语的第一个扩展使我们能够通过一个未指定的数量来限制 f 的定义域。

定义4.1 如果 $\varepsilon > \neq 0$，那么我们称 $f = A(\varepsilon)$ 在 a 处 (或在 a) 附近局部 $N_\delta^\times(a)$，如果存在某个删除的开区间 $N_\delta^\times(a)$ 在其上 $f = A(\varepsilon)$。如果存在一个 $M > \neq 0$ 使得 $f = A(M)$ 在 a 附近，那么我们称 f 在 a 处局部有界。

请注意，此条件明确忽略在 a 发生的事情；我们可能有 $|f(a)| > \varepsilon$，或者 $f(a)$ 甚至可能未定义。

ε越小，条件$f = A(\varepsilon)$越严格。例如，如果f和g是上面引入的函数，那么对于每个$\varepsilon \geq 1$，我们在0处有$g = A(\varepsilon)$，而如果$\varepsilon < 1$，则$g = A(\varepsilon)$在0附近不成立。如果我们对f提出类似的问题，我们会发现一个可能令人惊讶的答案：如果给我们一个$\varepsilon > 0$，那么在开区间$N_\varepsilon(0) = (-\varepsilon, \varepsilon)$上我们有$f = A(\varepsilon)$。换句话说，对于每个$\varepsilon > 0$，我们在0处有$f = A(\varepsilon)$。仔细观察，这并不等同于说$f = A(0)$在0处！

存在一个可能令人困惑的点在最后一段：在询问是否在 a 处局部 $f = A(\varepsilon)$，我们首先给出 $\varepsilon > 0$，然后我们选择一个区间。分析中的许多概念同样依赖于做出一个或多个"选择"，并且按照商定的顺序做出选择至关重要。

在第二章中，我们非正式地看到了A符号在计算中的应用。现在我们更有理由证明这些操作。如果以下陈述看起来很显然，请记住$f = A(\varepsilon)$不是一个等式，而是"f的绝对值小于ε"的缩写。

命题 4.2. 如果 r_1 和 r_2 是正实数，那么

$$A(r_1) + A(r_2) = A(r_1 + r2)$$
$$A(r_1) \cdot A(r_2) = A(r_1 r2)$$

特别地，如果 $x > 0$ 和 $\varepsilon > 0$，那么 $x + A(\varepsilon) = A(x+\varepsilon)$ 和 $xA(\varepsilon) = A(x\varepsilon)$。

证明。第一个断言是三角不等式，但如果你不小心，不等式可能会看起来走向错误的方向。如果 $x = A(r_1)$ 和 $y = A(r_2)$，那么 $|x+y| \le |x| + |y| \le r_1 + r_2$，这意味着 $x + y = A(r_1 + r_2)$，正如所声称的。在相同的假设下，$|xy| = |x|\,|y| \le r_1 r_2$，这证明了 $xy = A(r_1 r_2)$。 □

请注意，$A(r_1+r_2) = A(r_1)+A(r_2)$是错误的。仅仅因为一个量不大于1，并不意味着它是两个量之和，每个量都不大于1/2。

表达式 $x = a + A(\delta)$ 表示 x 是形如 a 的数加上一个绝对值不超过 δ 的数。这相当于说 $x \in [a-\delta, a+\delta]$。同样，如果 f 是一个函数，那么 $f = b + A(\varepsilon)$ 表示 f 的像包含在区间 $[b-\varepsilon, b+\varepsilon]$ 内。注意，当数字 h 在关于 0 的区间内变化时，数字 $x = a + h$ 在关于 a 的区间内变化。因此，$x = a + A(\delta)$ 和 $x - a = h = A(\delta)$ 表示相同的意思。这些都是您应该掌握的 A 语法标准习语。

示例4.3 互反函数 $f(x) = 1/x$，$x \ne 0$，在 a 对于每个 $a \ne 0$ 是局部有界的，但在 0 处不是局部有界的。（0 不在定义域中无关紧要，因为局部有界性“忽略 a”。）为了看出 f 在 $a \ne 0$ 处是局部有界的，首先假设 $a > 0$，并令 $\delta = a/2$，图4.1。根据上一段，$x = a + A(\delta)$ 表示 $x \in [\frac{a}{2}, \frac{3a}{2}]$，或者 $0 < \frac{a}{2} \le x \le \frac{3a}{2}$。定理2.23（iv）意味着 $0 < \frac{2}{3a} \le x \le \frac{2}{a}$，这意味着 $f = A(2/a)$ 在以 $a/2$ 为半径、以 a 为中心的区间上。$a < 0$ 的情况类似：取 $\delta = -a/2 > 0$。□

在第3章中，我们看到了一些函数（例如计算其输入十进制展开中“7”的数量的函数）在a处不具有局部有界性，无论我们给出哪个a。考虑到这样一个函数在**R**的每个点上都有一个定义良好、有限的值，这种行为是引人注目的。你应该认识到局部性质与逐点性质在本质上不同。

O 符号

A 表示允许我们计算那些不完全知道但对其绝对值有界限的量。通常，

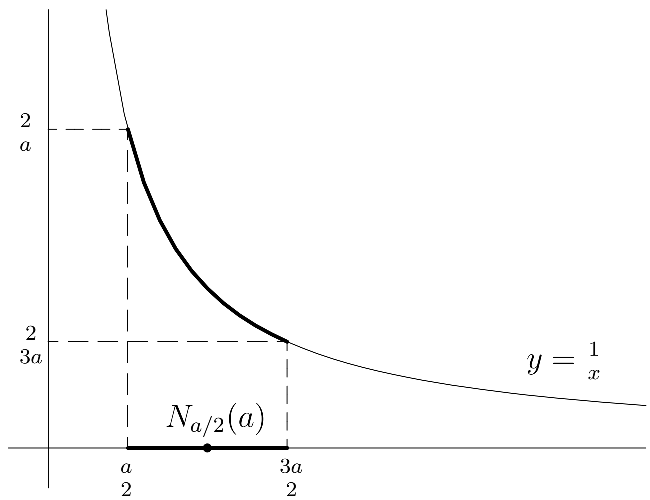


图4.1：界定倒数函数。

我们希望更加随意，忽略乘积常数。在这种情况下，我们会说在 $x=0$ 附近，x 和 $10x$ “大致相同”，而 1 明显更大，x^2 明显更小。如果您使用过计算机代数程序，您可能遇到过所谓的“O 符号”。

定义4.4 设f和g是定义域包含某些集合X的实值函数。我们说$f=O(g)$在X（上读作“f是g在X上的大O”），如果存在一个正实数C，使得对于所有$x\in X$有$|f(x)|\leq C|g(x)|$。

当使用 O 符号时，重要的是要提及集合 X，或者至少要记住不等式成立的限制。例如，我们在 $[0,10^{40}]$ 上有 $O(x^2)=O(x)$（取 $C=10^{40}$），但不在 $\mathbf{R}$ 上。

O 表示法比 A 表示法更对称。例如，$O(1)=O(10)$ 和 $O(10)=O(1)$ 都是真的。对于“$f=O(g)$ 在 a 处局部”有一个明显的定义，你应该给出。对于每个 $a\in\mathbf{R}$ (why?)，我们都有 $x=O(1)$，$10x=O(x)$ 和 $x^2=O(x)$ 在 a 处局部。然而，我们并没有在 0 附近有 $x=O(x^2)$。

与 A 符号表示法一样，我们可以使用 O 符号表示法来计算不等式：

$$\big(1+O(x)\big)^2=1+2O(x)+O(x)^2=1+O(x)+O(x^2)\qquad\text{on }\mathbf{R}.$$

特别地，$\big(1+O(x)\big)^2=1+O(x)$ 接近 0，并且是 $O(1)$ 接近 a 的

每个 $a \in \mathbf{R}$。这里总结了 O 符号的某些重要属性。

命题4.5。设$h > 0$，且设$k < \ell$为正整数。那么

$$\left.\begin{aligned} O(h^k) + O(h^\ell) &= O(h^k) \\ O(h^k) \cdot O(h^\ell) &= O(h^{k+\ell}) \end{aligned}\right\} \qquad near\ h = 0$$

如果 f 是一个有界函数，那么 $f \cdot O(h^k) = O(h^k)$。

您将没有困难证明这些断言。为了了解这些属性在实际中的工作方式，假设 f 是一个函数，使得

$$f(x+h) = f(x) + O(h) \text{ near } h = 0, \text{ for all } x \text{ in the domain.} \tag{4.1}$$

这样的函数在每个 x 上局部 $O(1)$，因为对于每个 x，我们有

$$\begin{aligned} f(x+h) &= f(x) + O(h) \\ &= f(x) + A(1) = A\big(|f(x)| + 1\big) \\ &= O(1) \qquad \text{near } h = 0, \end{aligned}$$

(记住 x 是固定的)。此外，如果 f 和 g 满足 (4.1)，那么 $f + g$ 和 fg 也同样满足：

$$\begin{aligned} (f+g)(x+h) &= f(x+h) + g(x+h) \\ &= f(x) + O(h) + g(x) + O(h) \\ &= f(x) + g(x) + O(h) = (f+g)(x) + O(h), \end{aligned}$$

和

$$\begin{aligned} (fg)(x+h) &= \big[f(x) + O(h)\big]\big[g(x) + O(h)\big] \\ &= f(x)g(x) + \big[f(x) + g(x)\big]O(h) + O(h^2) \\ &= (fg)(x) + O(h) \text{ near } 0. \end{aligned}$$

这里有一些在第8章中将很有用的例子。

示例4.6 练习2.15的二项式定理表明

$$(x+h)^n = x^n + nx^{n-1}h + O(h^2) \text{ near } h = 0, \text{ for all } n \in \mathbf{N}. \tag{4.2}$$

二项式定理精确地说明了 $O(h^2)$ 项等于什么，但对于许多目的，我们只需要 (4.2) 提供的信息。例如，我们得出以下结论：

$$\frac{(x+h)^n - x^n}{h} = nx^{n-1} + O(h) \text{ near } h = 0.$$

这很有用，因为我们不能在左边将 $h=0$ 设置为 0，但我们可以在右边设置，从而获得 0/0 的评估！□

示例4.7 假设 f 是一个满足以下条件的函数：存在一个实数 $f'(0)$ 使得

$$f(h) = f(0) + f'(0)h + O(h^2) \text{ near } h = 0.$$

直观上，“f 在0处直到二次项是线性的”。物理学家可能会将其写作 $f(h) \simeq f(0) + f'(0)h$ 对于 $h \simeq 0$，但我们的表达式有明确的精确解释。现在，假设 f 和 g 都满足这个条件。我们立即计算出

$$(f+g)(h) = f(0) + g(0) + \big[f'(0) + g'(0)\big]\, h + O(h^2)$$

和

$$\begin{aligned}(fg)(h) &= \big[f(0) + f'(0)\, h + O(h^2)\big]\big[g(0) + g'(0)\, h + O(h^2)\big] \\ &= f(0)g(0) + \big[f'(0)g(0) + f(0)g'(0)\big]\, h + O(h^2),\end{aligned}$$

这证明了$f+g$和fg也满足条件，并且（作为额外的好处）告诉我们($f+g$)′(0)和(fg)′(0)是什么。□

这些例子表明，O 符号形式化了科学家们经常进行的“信封背面”计算，以估计理论预测或实验结果。更多示例在练习中给出。

o 符号

我们的最终符号定义表面上看起来像 O 符号，但实际上封装了一个非常微妙且非平凡的属性。o 符号中的一个单独的表达式包含无限多个 A 表达式。

定义4.8 设f为一个实值函数。我们称$f = o(1)$在a处

(4.3) $$f = A(\varepsilon) \text{ locally at } a \text{ for every } \varepsilon > 0.$$。

(4.3) $$f = A(\varepsilon) \text{ locally at } a \text{ for every } \varepsilon > 0.$$

如果 g 是一个在某些关于 a 删除的区间上不消失的函数，那么我们称 $f = o(g)$ 在 a 处局部地等于 $f/g = o(1)$ 在 a 处局部地。

我们之前看到，$\mathbf{R}$ 的恒等函数在 0 处是 $o(1)$；非正式地说，$x = o(1)$ 在 $x = 0$ 处。请注意，x 可以被任何其他字母替换；

我们可以（并将）使用这个事实：在 $h=0$ 处有 $h=o(1)$。条件 $f=o(1)$ 在 a 处捕捉到直觉："通过将 x 足够接近 a，$f(x)$ 可以变得任意小"，而 $f=o(g)$ 在 a 处意味着"对于 x 接近 a，$f(x)$ 相比于 $g(x)$ 是极其微小的"。对于这些断言有意义，没有必要在 a 处定义 f。

o 和 O 符号表示消失阶（阶数）。在理论（无变量）和计算（有变量）环境中使用 O 和 o 符号都很方便。相应的符号略有不同，你应该努力在这两者之间保持流畅。例如，"$f=o(1)$ at x"（无变量）和"$f(x+h)=o(1)$ at $h=0$"（有变量）表示相同的意思。为了了解这些标准和它们之间的关系，你应该验证以下陈述（并且为了保险起见，根据需要将每个陈述翻译成"有变量"或"无变量"语言）：

- 如果 $f=o(1)$ 在 x，那么 $f=O(1)$ 近 x。
- 如果 $f(x+h)=O(h)$ 接近 $h=0$，那么 $f(x+h)=o(1)$ 在 $h=0$。
- $h2=o(h)$ 在 0；事实上，$O(h^2)=o(h)$ 在 0。
- 对于每个正整数 k，$o(h^k)=O(h^k)$ 接近 0。

原型消失行为在 a 上由 k 次幂函数 $g(x)=(x-a)^k$ 对于 $k\in\mathbf{N}$ 展示，我们将其视为在 a 处"消失到 k 阶"，更一般地说，我们说

$$\text{“}f\text{ vanishes to order }\geq k\text{ at }a\text{” if }f=O\big((x-a)^k\big)\text{ near }a$$
$$\text{“}f\text{ vanishes to order }> k\text{ at }a\text{” if }f=o\big((x-a)^k\big)\text{ near }a$$

我们将很快看到这个术语是自然的，并且符合我们对整数不等式的直觉。

对于在生活中遇到的（特别是，对于我们在第11章中遇到的"实分析"函数），条件

$$f=O\big((x-a)^{k+1}\big)\quad\text{and}\quad f=o\big((x-a)^k\big)$$

本质上等效。换句话说，"大多数"函数在整数阶上消失，除了可能在孤立点之外。然而，正如我们在第3章中看到的，有许多"古怪"的函数例子，并且一般来说，{v*}1{v*}比{v*}9要少得多。

代数计算中操作 o 表达式的规则与 O 符号表示的规则类似，但证明更为微妙。这是可以预料的，因为“$f = o(1)$”表示“对于每个 $\varepsilon > 0$ 的 $f = A(\varepsilon)$”，这是一个涉及无限多个标准的条件。

定理4.9。设f和g是在关于a的某个删除区间上定义的函数。如果在a处有$f = o(1)$和$g = o(1)$，那么在a处有$f + g = o(1)$。如果在a处有$f = o(1)$和$g = O(1)$，那么fg是$o(1)$在a处的值。

非正式地说，该定理表明

$$o(1) + o(1) = o(1), \qquad o(1) \cdot O(1) = o(1).$$

特别地，对于每个实数 c，有 $c \cdot o(1) = o(1)$。

证明。我们假设对于每个 $\varepsilon > 0$，都有 $f = A(\varepsilon)$，同样适用于 g。记住，ε 本身是一个“变量”，代表任意正数。如果方便，我们可以称它为其他名称，例如 $\varepsilon/2$。

我们希望证明 $f + g = o(1)$。设 $\varepsilon > 0$。因为 $f = o(1)$，存在一个关于 a 的开区间，在该区间上 $f = A(\varepsilon/2)$。同样，存在一个开区间，在该区间上 $g = A(\varepsilon/2)$。在这些区间中较小的一个（即它们的交集，见图 4.2），我们同时有 $f = A(\varepsilon/2)$ 和 $g = A(\varepsilon/2)$，因此

$$f + g = A(\varepsilon/2) + A(\varepsilon/2) = A(\varepsilon).$$

我们已经证明，如果 $\varepsilon > 0$，那么存在一个关于 a 的开区间，使得 $f + g = A(\varepsilon)$；这恰好意味着 $f + g = o(1)$。

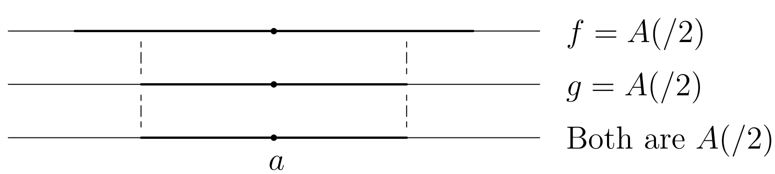


图4.2：确保单个开区间上的两个条件。

为了建立第二部分，首先假设 $g = O(1)$ 在 a。这意味着存在一个实数 $M > 0$ 和一个关于 a 的开区间，使得 $|g(x)| \le M$ 对于区间内的所有 x 成立。现在固定 $\varepsilon > 0$。因为 $f = o(1)$ 在 a，并且因为 ε/M 是一个正实数

数字，存在一个关于 a 的开区间，其中 $f = A(\varepsilon/M)$。与之前一样，考虑较小的开区间；在这个区间上，我们既有 $f = A(\varepsilon/M)$ 和 $g = A(M)$，所以

$$fg = A(\varepsilon/M) \cdot A(M) = A(\varepsilon).$$

我们已经证明，如果 $\varepsilon > 0$，那么 $fg = A(\varepsilon)$ 在关于 a 的某个开区间上。由于 $\varepsilon > 0$ 是任意的，我们已经证明 $fg = o(1)$ 在 a 上。

□

证明展示了分析中的一个标准技巧，如图4.2所示。如果给出了有限个条件，每个条件在关于a的一些区间上成立，那么通过取这些区间的交集，我们可以找到一个所有条件都成立的单个区间。在a处的有限个区间对应于有限个正实数（它们的半径），交集对应于集合中的最小数。相比之下，当有无限多个条件时，这个技巧通常不适用，因为关于a的无限多个开区间的交集不必是开区间！为了理解为什么，考虑半径为$1/n$关于a的区间$N_{1/n}(a)$。让我们确定当n在正整数集合上变化时，哪些实数x在所有这些区间中。显然$x = a$是，根据定义。然而，如果$x \neq a$，那么x不在交集中：存在一个n，使得$1/n < |x - a|$，根据$\mathbf{R}$的阿基米德性质，这个不等式意味着$x \notin N_{1/n}(a)$。因此，这些开区间的交集是单点$\{a\}$。为了给出一个更分析性的（更少几何性的）解释，回想一下，一个正实数的无限集合总是有一个下确界，但不一定有最小值，并且正实数集合的下确界可以是零。

4.2 限制

我们几乎准备好介绍“极限”的概念了。还有一个最后的技术问题需要提出，关于函数的定义域。到目前为止，我们对函数f的定义域X没有做出任何严肃的假设。然而，o的语言有一个小特殊性。假设f的定义域是单点$\{0\}$。我们仍然可以问，“$f = o(1)$接近1吗？”答案是“是的”，因为在以1为半径的开区间内，没有f的定义域中的点，因此空洞地我们有“对于每个$\varepsilon > 0$，$f = A(\varepsilon)$接近1”。我们希望消除这样的

空洞的断言来自我们的考虑。此外，我们希望在定义“极限”时忽略f在a的值，因为我们希望即使在f在a未定义的情况下也能使用极限的概念。这两个问题都得到了很好的解决。

定义4.10 设$X \subset \mathbf{R}$。若关于a的每个去心区间都包含X中的点，则点$a \in \mathbf{R}$是X的极限点。X中不是X的极限点的点是一个孤立点。

极限点的概念比乍看之下更为微妙。例如，a是否是X的极限点在逻辑上与$a \in X$是否是极限点无关。如果a是f的定义域的极限点，那么条件“$f - f(a) = o(1)$在a”非空。反之，如果a不是f的定义域的极限点，那么条件是空的。你应该没有困难验证一个稍微更强的条件，这为X中不是X的极限点的点提供了“孤立点”这一术语的合理性：

引理4.11。设$X \subset \mathbf{R}$，并设a是$\mathbf{R}$中的一个点。如果a不是X的极限点，那么存在一个关于a的删除区间，该区间不包含X中的任何点。

示例4.12 设$X = (0, +\infty)$为正实数集。我们断言X的极限点集是非负实数集。为了证明这一断言，我们将展示$[0, +\infty)$中的每个点都是X的极限点，并且不在$[0, +\infty)$中的每个点都不是极限点。“有趣”的点是0，它不是X的点，但却是X的极限点。

如果 $a \geq 0$，那么对于每个 $\varepsilon > 0$，点 $x = a + \varepsilon/2 > a$ 在 X 与 $N_\varepsilon^\times(a)$ 的交集中。这意味着 a 是 X 的一个极限点，并建立了第一个包含关系。为了证明另一个方向，假设 $a < 0$；我们希望证明 a 不是一个极限点 X。因为 a 到 X 的距离是 $|a|$，半径为 $\varepsilon = |a|/2 > 0$ 的删除区间与 X 不相交。因为存在这样的删除区间，所以点 a 不是一个极限点 X，正如我们希望证明的那样。将这个几何论证翻译成不等式的语言是一个很好的练习。

□

示例4.13 我们的第二个例子是 $X = \mathbf{Z}$，即整数集。这个集合没有任何极限点！（特别是，X 的一个点不一定是 X 的极限点，即使 X 有无限多个元素。）如果 $a \in X$，那么半径为1的删除区间不包含 X 的任何点，因此 a 不是 X 的极限点。如果 $a \notin X$，那么存在一个整数 n

如此，$n < a < n+1$（确保您只能使用 $\mathbf{N}$ 和 $\mathbf{R}$ 的公理来证明这个“显然”的断言！），所以取 ε 为 $a-n$ 和 $n+$ 中的较小者 $1-a$，我们看到半径为 $\varepsilon > 0$ 的删除区间不包含 X 的任何点，并且再次 a 不是 X 的极限点。□

示例4.14 我们的第三个例子是 $X = \mathbf{Q}$，有理数集。回忆（定理2.33）每个实数开区间都包含一个有理数。很容易得出每个实数都是 $\mathbf{Q}$ 的极限点：如果 $a \in \mathbf{R}$ 和 $\varepsilon >$ 不等于0，那么开区间 $(a, a+\varepsilon) \subset N_\varepsilon^\times(a)$ 包含 $\mathbf{Q}$ 的一个点。□

作为理解测试，确定（用证明！）区间（$0, 1$）的极限点集，以及集合 $X = \{1/n \mid n > 0\}$ 的极限点集：

$\ldots$

0 1

大多数X的点在此处未描绘；在左侧边缘，代表它们的刻度线重叠在一起。将问题分解为情况可能会有所帮助。首先处理情况$a < 0$和$a > 1$；然后考虑X本身的点，同时记住当您“放大”接近0时X看起来像什么；接下来考虑不是X元素的点$0 < a < 1$；最后考虑0。如果您感觉像在雾中降落飞机，请记住定义是你的雷达。

最后我们可以引入极限。我们想要表达一个实数 {v*} 是 {v*} 在 {v*} 处的“极限”，如果 {v*} 在 {v*}。细节是，我们希望避免一个空洞的陈述（因此我们要求 {v*} 是 {v*} 的定义域的极限点），我们希望忽略 {v*} 在 {v*} 的值，因此我们限制在删除的区间 {v*}。完全展开的定义是所有分析学生都熟知的“{v*}-{v*} 标准”。我们将其打包到 {v*} 符号中，以阐明概念内容。

定义4.15 设$f : X \to \mathbf{R}$为一个函数，且a是X的一个极限点。如果实数ℓ是f在a处的极限，则称其为f在a处的极限。在这种情况下，我们写作$\ell = \lim(f, a)$或$\ell = \lim\limits_{x\to a} f(x)$.

如果对于每个实数 ℓ，“$\lim(f, a) = \ell$” 是假的，那么我们说 f 在 a 处的极限不存在。按照我们设定的定义，除非 a 是 f 的定义域的极限点，否则询问 $\lim(f, a)$ 是否存在是没有意义的。

极限通常通过非正式的说法来解释，即“‘ℓ是f在a的极限’意味着‘当x足够接近a时，函数值$f(x)$可以任意接近ℓ’”。如果你在其他地方研究过极限，你很可能已经看到了以下定义：

> 实数 ℓ 是 f 在 a 处的极限，如果对于每个 $\varepsilon > 0$，存在一个 $\delta > 0$，使得 $0 < |x-a| < \delta$ 蕴含 $|f(x)-\ell| < \varepsilon$。

它很容易看出这个条件等价于定义4.15。短语“$0 < |x-a| < \delta$ 蕴含 $|f(x)-\ell| < \varepsilon$”意味着在关于 a 的删除 δ-区间上 $f = \ell + A(\varepsilon)$。子句“对于每个 $\varepsilon > 0$，存在 $\delta > 0$”在这个上下文中意味着“对于每个 $\varepsilon > 0$，存在一个删除的区间。. . ”。因此，上述定义意味着 $f = \ell + o(1)$ 在 a。

“lim({v*}) $= \ell$”这个表达式看起来像是一个等式，但你应该小心不要把它当作等式来处理；记住，“$f = o(g)$”这个表达式不是一个等式，而是一个缩写。具体来说，问题是可能存在多个数字可以作为f在a处的极限。首要任务是消除这种担忧：

引理4.16。如果ℓ和m在关于a的一些删除区间上是常数函数，并且如果$\ell - m = o(1)$在a处，那么$\ell = m$。换句话说，如果两个实数任意接近，那么它们是相等的。

证明。奇怪的是，假设由无限多个陈述组成，即对于每个 $\varepsilon > 0$，有 $\ell - m = A(\varepsilon)$ 在 a，但没有任何有限数量的这些假设能推出结论！相反，考虑其逆否命题：如果 $\ell \neq m$，那么存在 $\varepsilon > 0$ 使得 $\ell - m$ 不是 $A(\varepsilon)$。稍加思考，这很明显：如果 $\ell - m \neq 0$，那么取 $\varepsilon = |\ell - m|/2 > 0$。根据假设，我们在每个关于 a 的删除区间上都有 $|\ell - m| = 2\varepsilon$，并且有 $2\varepsilon > \varepsilon$ 因为 $\varepsilon > 0$。 $\square$

定理4.17。设f为一个函数。如果$\lim(f, a) = \ell$和$\lim(f, a) = m$，那么$\ell = m$。

非正式地，f 在 a 处最多有一个极限，如果我们同意极限可能不存在)，那么谈论 f 在 a (处的极限是有意义的。定理 4.17 也证明了当极限存在时，将表达式“$\lim(f, a) = \ell$”视为实数方程是合理的，并且从现在起我们将这样做。

证明。假设在 $f = \ell + o(1)$ 和 $f = m + o(1)$ 处。这意味着在 $f - \ell = o(1)$ 和 $m - f = o(1)$ 处。相加，我们得到 $\ell - m = o(1)$ 在 a 处，这由引理4.16得出 $\ell = m$。 $\square$

符号

$$\lim_{x\to a} f(x) = \ell \quad \text{and} \quad \lim(f,a) = \ell \tag{4.4}$$

“当 x 趋近于 a 时，$f(x)$ 的极限等于 ℓ” 和 “f 在 a 处的极限是 ℓ” 。在抽象情况下，后者更为简洁，因为它避免了引入虚假的“x。” 应该强调的是，虽然“$\lim_{x\to a} x^2$”是允许的（因为“$x \to a$”使“x^2 ”是 x 处的函数值这一点变得清晰），但表达式“$\lim(x^2, a)$”是模糊的，因为“x^2”没有定义一个函数。最有可能的两个含义是 $\lim_{x\to a} x^2 = a^2$ 和 $\lim_{t\to a} x^2 = x^2$。

为了使用极限，我们需要“常用工具”：给出极限一般性质的定理、具有和不具有极限的函数的例子，以及寻找和操作极限的简便计算方法。

示例4.18 两个极限是直接的：如果 c 是一个实数，那么 $\lim_{x\to a} c = c$ ，以及对于所有 a 的 $\lim_{x\to a} x = a$。在 o 符号中，$c = c+o(1)$ 和 $x = a+o(1)$ 在 $x = a$ 附近。第一个是显然的，因为 $0 = o(1)$，第二个是清晰的，因为 $h = o(1)$ 在 $h = 0$。□

定理4.19。设f和g是具有相同定义域X的函数。如果$\lim(f, a) = \ell$和$\lim(g, a) = m$，那么$\lim(f + g, a)$和$\lim(fg, a)$存在，并且分别等于$\ell+m$和ℓm。如果再加上$m \neq 0$，那么$\lim(f/g, a)$存在，并且等于ℓ/m。

证明。（求和）假设在 a 处有 $f = \ell + o(1)$ 和 $g = m + o(1)$；根据定理 4.9，我们有 $f+g = \ell+m+o(1)$ 在 a 处，这意味着 $\lim(f + g, a) = \ell + m = \lim(f, a) + \lim(g, a)$。

(产品) 在相同的假设下，我们有

$$\begin{aligned} fg &= \big(\ell + o(1)\big)\big(m + o(1)\big) && \text{根据假设} \\ &= \ell m + (\ell + m)o(1) + o(1)\cdot o(1) && \\ &= \ell m + o(1) && \text{定理4.9.} \end{aligned}$$

这表示 $\lim(fg, a) = \ell m = \lim(f, a) \cdot \lim(g, a)$.

(商) 这稍微复杂一些，但唯一的新成分是4.3例中的技术，见图4.1 。首先假设$m > 0$，并让$\varepsilon = m/2 > 0$。因为$g = m + o(1)$，所以在a处有一个关于$g = m + A(\varepsilon)$的删除区间，其中$g = m + A(\varepsilon)$，根据ε的选择意味着

$$\frac{m}{2} < g < \frac{3m}{2}, \quad \text{or} \quad \frac{2}{3m} < \frac{1}{g} < \frac{2}{m}.$$

换句话说，$1/g = O(1)$ 近 a。直接计算给出

$$\begin{aligned}\frac{1}{g} - \frac{1}{m} &= \frac{m-g}{gm} \\ &= -(g-m) \cdot \frac{1}{g} \cdot \frac{1}{m} \\ &= o(1) \cdot O(1) \cdot O(1) = o(1),\end{aligned}$$

或 $1/g = (1/m) + o(1)$ 在 a。乘以 $f = \ell + o(1)$，我们得到 $f/g = \ell/m + o(1)$，正如预期的那样。 □

推论4.20。设p和q是互质因子的多项式，设$f = p/q$是对应的有理函数。如果$q(a) \neq 0$，那么lim$(f, a) = p(a)/q(a)$。特别是，如果$p : \mathbf{R} \to \mathbf{R}$是多项式函数，那么对于所有$a \in \mathbf{R}$，lim$(p, a) = p(a)$。

用文字来说，有理函数的极限通过求值得到。例如，

$$\lim_{x \to a} \frac{1 + x + x^5}{4 - x^2} = \frac{1 + a + a^5}{4 - a^2}$$

如果 $a \neq \pm$^2。

证明。每个单项式函数 $x \mapsto a_n\, x^n$ 是一个常数与 n 个恒等函数的乘积，因此满足定理4.19的“乘积”部分所得到的结论。根据定理的“和”部分，多项式也是这样，而对于有理函数则是通过“商”部分。为了完全形式化这个论证，你需要通过数学归纳法进行几个论证。仔细证明单项式的断言是值得的，以了解涉及的内容。

□

推论4.21。假设lim$(f, a) = \ell$存在，但g在a处没有极限。那么$f + g$在a处没有极限，并且如果$\ell \neq 0$，那么fg在a处没有极限。

证明。设$h = f + g$；该定理断言，如果h在a处有极限，那么 $g = h + (-f)$也有极限。这是推论的反命题。关于fg的说法是练习4.5。 □

事实上限和逐点评估对于有理函数来说是同一件事，这可能会让人认为这两个概念始终相同。从哲学角度来看，几乎正好相反。定理4.22，

局部性原理，精确地阐述了不能通过观察有限个函数值来确定极限的断言。这种对有限数据的“盲视”似乎矛盾，但并没有逻辑上的不一致。真正的教训是数轴比我们第一眼看到的要复杂。

定理4.22。设f和g是定义域包含关于$a \in \mathbf{R}$的一些删除区间的实值函数，并假设$f(x) = g(x)$除可能在有限多个x处外。那么，f, a的极限=等于g, a的极限，在以下意义上：要么两个极限都存在且相等，要么两个极限都不存在。

证明。设$\{b_1, \dots, b_n\}$是$f(x) \neq g(x)$的互异点的列表。我们不妨假设b_i中没有一个是a，因为我们只对f和g在关于a的删除区间上的行为感兴趣。令$\delta > 0$为$|b_i - a|$的最小值，对于$1 \leq i \leq n$，即a到最近的b_i的距离。（这里使用了有限性；如果存在无限多个b_i，我们无法保证$\delta > 0$。）在删除区间$N_\delta^\times(a)$上，f和g是相同的函数，因此对于定义4.15的目的，$f = g$。

□

另一个有用、基本性质是“极限尊重 $\leq$”。你应该仔细注意，关于单一点的极限信息会告诉你关于开区间上函数的一些信息。

定理4.23。假设f和g具有相同的定义域，且f, a和g, a的极限存在，并且对于关于a的某个删除区间内的所有x，有$f(x) \leq g(x)$。那么f, a的极限等于g, a的极限。

证明。虽然定理中给出的陈述在应用中经常出现，但其逆否命题更易于证明：如果 $\lim(f, a) > \lim(g, a)$，那么存在一个关于 a 的删除区间，使得 $f > g$。

考虑函数 $h = f - g$，它有极限

$$\ell := \lim(f, a) - \lim(g, a) > 0$$

根据定理4.19。我们将证明存在一个关于a的删除邻域，在该邻域上$h > 0$。令$\varepsilon = \ell/2 > 0$。因为$h = \ell + A(\varepsilon)$在$a$处局部，存在一个关于$a$的删除区间，在该区间上$h > \ell - \varepsilon = \ell/2 > 0$。这已经证明了定理，但值得一提的是，不仅$h > 0$在$a$附近，而且$h$有一个正的下界，或者被限制在零之外。这个额外的信息通常很有用。

□

如果假设加强为“对于某些删除区间内的所有 x，有 $f(x) < g(x)$”，那么在一般情况下并不一定有 $\ell < m$（，尽管当然有 $\ell \le m$，根据定理)。你应该检查证明以了解为什么不是这样，并找到一个函数对 $f < g$，使得当 f, 趋于 0 时，极限 $=$ 等于当 g, 趋于 0 时的极限。

一个相关结果是所谓的“挤压定理”（或“夹逼定理”）。通过找到简单的上界和下界并应用挤压定理，评估了几个有趣的极限。

定理4.24. 假设在关于a的某个删除区间上有$f \le h \le g$，并且lim(f, a)和lim(g, a)都存在且等于ℓ。那么lim(h, a)存在且等于ℓ。

证明。固定 $\varepsilon > 0$，并选择 $\delta > 0$，使得 $f = \ell + A(\varepsilon)$ 和 $g = \ell + A(\varepsilon)$ 在关于 a 的已删除 δ-区间上。结合 $f \le h \le g$ 在 a 处局部成立的假设，我们得到 $-\varepsilon \le f - \ell \le h - \ell \le g - \ell < \varepsilon$ 在 $N_\delta^\times(a)$ 上，因此 $h = \ell + A(\varepsilon)$ 在 a 处局部成立。由于 $\varepsilon > 0$ 是任意的，所以 lim(h, a) $= \ell$。 □

示例4.25 设sgn : $\mathbf{R} \to \mathbf{R}$ 为符号函数。局部性原理意味着如果 $a \neq \neq$ 0，那么

$$\lim_{x \to a} \operatorname{sgn}(x) = \operatorname{sgn}(a) = \frac{a}{|a|} = \begin{cases} 1 & \text{if } a > 0, \\ -1 & \text{if } a < 0. \end{cases}$$

在0附近，然而，sgn取值为1和-1；这些值不在任何长度小于2的区间内，因此如果$\varepsilon \le 1$，则对于每个实数ℓ，条件sgn $= \ell + A(\varepsilon)$都是假的。这意味着lim(sgn, 0)不存在。□

示例4.26 这里有两个相对复杂的例子。第一个在定义域中的每个点 a 上在 a 处没有极限。第二个在每个点都有极限，但乍一看似乎在所有地方都没有极限。

考虑函数 $\chi_{\mathbf{Q}} : \mathbf{R} \to \mathbf{R}$，有理数集的特征函数，并固定一个实数 a。在关于 a 的每个删除区间中，都存在有理数和无理数，因此函数 $\chi_{\mathbf{Q}}$ 取 0 和 1 两个值。无论如何选择潜在的极限 ℓ，如果 $\varepsilon < 1/2$，那么我们在 a 处没有 $\chi_{\mathbf{Q}} = \ell + A(\varepsilon)$，因为“目标”区间 $(\ell - \varepsilon, \ell + \varepsilon)$ 的长度为 < 1。这意味着 $\chi_{\mathbf{Q}}$ 在 a 处没有极限，对于每一个实数 a。

第二个例子是“分母”函数 f，它是 Example 3.11 的。下面命题 4 .28 中描绘了图形的放大部分。这个函数与 $\mathbf{Q}$ 的特征函数有一定的相似性，尽管当分母较小时，函数值较小。为了研究 f 的极限行为，考虑集合是方便的

$$\mathbf{Q}(N) := \{p/q \in \mathbf{Q} : |q| \leq N\} = \bigcup_{q=1}^{N} \tfrac{1}{q}\mathbf{Z}$$

有理数的分母至多为 N 的集合，并将 $\mathbf{Q}$ 写作这些集合的并集，当 N 在正整数范围内变化时。

我们在例4.13中表明，整数集$\mathbf{Z}$没有极限点，对论据的明显修改证明，类似地，集合$\frac{1}{q}\mathbf{Z}$也没有极限点。以下注解表明，集合$\mathbf{Q}(N)$本身也没有极限点。

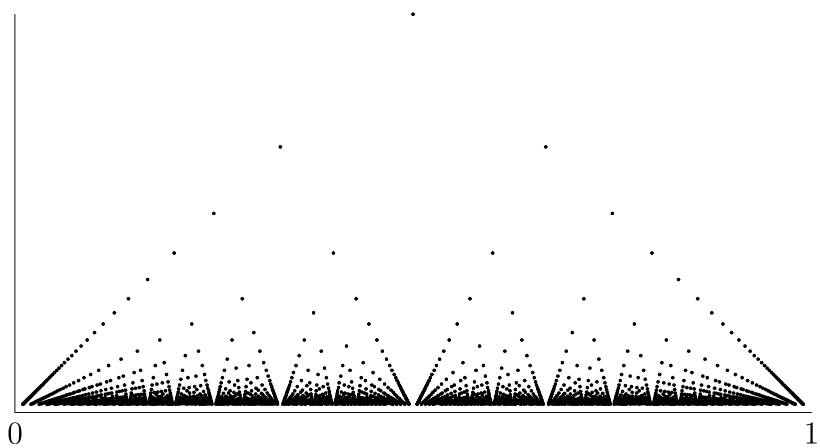
引理4.27. 设A_1, ..., A_N是$\mathbf{R}$的子集，且它们没有实数极限点。那么它们的并集没有极限点。

证明。设a是$\mathbf{R}$的一个点。因为A_1没有极限点，根据引理4.11，存在一个关于a的删除区间I_1，它不包含A_1的任何点。类似地，我们得到有限多个删除区间I_2，...，I_n，使得I_k不包含A_k的任何点。这些区间的交集是一个关于a的非空删除区间，它不包含$A_1 \cup \cdots \cup A_n$的任何点。特别是，a不是$A_1 \cup \cdots \cup A_n$的极限点。 $\square$

命题 4.28. 定义 $f : \mathbf{R} \to \mathbf{R}$ 为

$$f(x) = \begin{cases} 1/q & \textit{if } x = p/q \textit{ in lowest terms,} \\ 0 & \textit{if } x \textit{ is irrational.} \end{cases}$$

然后，对于每个 $a \in \mathbf{R}$，有 $\lim(f, a) = 0$。



证明。回忆 $\mathbf{Q}$ 是 $N \geq 1$ 的并集，其中包含集合 $\mathbf{Q}(N)$。根据 f 的定义，$\mathbf{Q}(N)$ 的元素恰好是满足 $|f(x)| \geq 1/N$ 的实数点 x。这一事实解释了图中靠近水平轴的窄带；只有 $N < 100$ 的点被显示出来。

设 a 为一个任意的实数。固定 $\varepsilon > 0$，并选择 $N \in \mathbf{N}$ 使得 1 $/N < \varepsilon$。根据引理 4.27，存在一个关于 a 的删除区间，该区间不包含 $\mathbf{Q}(N)$ 的任何点。由于此区间省略了所有满足 $f(x) \geq 1/N$ 的点 x，因此我们在 a 处有 $f = A(\varepsilon)$ 的局部存在。但是 ε 是任意的，因此我们已经证明了对于每个 a，在 a 处有 $f = o(1)$。 □

极限是关于函数的“新”信息，这些信息不能通过考虑单个函数值来看到。前一个例子使用了极限定义的全部力量，并展示了定义如何偏离直觉。尽管分母函数在无限多个点上不为零，但其极限存在，并且在每个实数 a 处等于零。命题4.28可以表述为一种近似原理：如果一个实数 a 被有理数 $x \neq a$ 近似，那么分母必须无限增长。即使 $a \in \mathbf{Q}$! □ 也是如此。

极限游戏

存在一种博弈论对极限的解释，许多人发现这很有帮助，可能是因为人脑更擅长理解竞争而不是存在量词。想象一个两人游戏，参与者为玩家 ε (“你的对手”) 和玩家 δ (“你”)。裁判事先指定一个函数 f ，一个点 a ，它是 f 的定义域的极限点，以及一个实数 ℓ (，假设的

极限)。玩家 ε 先走，选择一个正数 ε。这个数字是函数值的“容差”，并指定了以 ℓ 为中心的目标的半径。要在 f 的定义域中的点 x 处“满足”容差（或击中目标）意味着 $|f(x)-\ell|<\varepsilon$；使 ε 更小会使目标更难击中。

现在轮到你了：你想要选择一个“发射半径”($\delta>0$)，使得关于 a 删除的 δ-区间中起源的每个“射击”x都能击中目标。显然，选择较小的发射半径不会降低你的准确性。如果你成功了，那么你赢得了这一轮，我们称 $f=\ell+A(\varepsilon)$ 在 a 处局部；否则玩家 ε 赢。

等式 $\lim(f,a)=\ell$ 表示“你对完美玩家有制胜策略”：无论玩家 ε 如何缩小目标（记住，其大小为正），你都能赢得游戏。赢得一局和有对完美玩家的制胜策略之间的区别，是对于特定的 $\varepsilon>0$ 有 $f=\ell+A(\varepsilon)$ 和对于任意的 $\varepsilon>0$ 有 $f=\ell+A(\varepsilon)$ 之间的区别。

如前所述，ε 首先出牌至关重要。你是否拥有获胜策略取决于函数 f、位置 a 和假定的极限 ℓ；它不依赖于 ε 的选择。如果玩家 ε 出错，即使“$\lim(f,a)=\ell$”为假，你仍可能赢得一轮，比较示例 4.25。为了避免过于重复，赢得一轮不如拥有对抗完美玩家的获胜策略好。

如果您觉得博弈论解释有帮助，您可能还想考虑以下变体：游戏开始方式与之前完全相同，裁判选择 f、a 和 ℓ。您的对手选择 $\varepsilon>0$，而您选择 $\delta>0$，但现在您的对手在 f 的定义域内选择一个满足 $0<|x-a|<\delta$ 的 x。如果 $|f(x)-\ell|<\varepsilon$，那么您赢得这一轮，否则您输掉。同样，“$\lim(f,a)=\ell$” 等价于您对完美玩家拥有一个获胜策略。

存在其他由规则微小变化引起的数学上有趣的“极限游戏”。例如，裁判可能指定 $\ell=f(a)$；或者可能不指定数字 a，要求你同时以选择 $\ell=f(a)$ 赢得所有 $a\in A$；或者可能要求你以 $\ell=f(a)$ 赢得所有 $a\in A$，并且只能选择 δ。我们将在适当的时候遇到这些和其他极限游戏。

“极限游戏”以一种重要的方式理想化：预先指定了一个假设的极限 ℓ。在实际情况下，对 ℓ 了解甚少或一无所知。定义将只说明极限是否等于 ℓ，而不是极限是否实际存在。几乎一无所知。

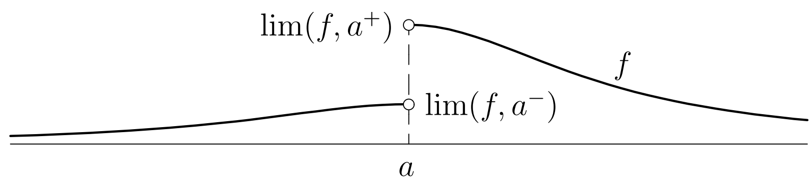


图4.3：f在a处的单侧极限。

学习如果尝试显示 $\lim(f, a) = \ell$ 失败；裁判提供了“错误”的数字 ℓ ，但可能没有“正确”的数字。在游戏类比中，如果你认为有一个以 ℓ 为中心的靶心，你射击准确但未命中，你不能推断出没有靶心，只能推断出你瞄准了错误的位置。有证明极限存在而不实际知道 ℓ 是什么的程序；这样的定理是一种“雷达”，它可以检测到目标但不定位它。这种知识本身可能有几个原因有用。首先，如果已知存在极限并且满足其他假设，则有定理可以找到极限；也许更重要的是，这意味着可以从给定的函数中提取极限来定义新的函数。导数、积分和幂级数——微积分的三个支柱——都属于这种类型。

单侧极限

极限 f 在 a 处考虑了 f 在关于 a 的删除区间上的行为。删除的区间 $N_r^\times(a)$ 是两个开区间的并集，即 $(a - r, a)$ 和 $(a, a + r)$，有时需要分别研究每个区间上的 f。

如果函数 f 在某个 $r > 0$ 的区间 $(a, a+r)$ 上定义，并且（在限制到这个开区间之后）$f = \ell + o(1)$ 在 a 附近，那么我们说 f 在 a (的右极限或“上方”极限存在，并写为

$$\lim(f, a^+) = \ell \qquad \text{or} \qquad \lim_{x \to a^+} f(x) = \ell \qquad \text{or} \qquad \lim_{x \searrow a} f(x) = \ell$$

您应将此条件翻译为 ε-δ 游戏，并应从左（即“从下方”）对极限的类似定义进行公式化。左极限的表示方法如预期所示：

$$\lim(f, a^-) = \ell \qquad \text{or} \qquad \lim_{x \to a^-} f(x) = \ell \qquad \text{or} \qquad \lim_{x \nearrow a} f(x) = \ell.$$

如果lim$(f, a) = \ell$，那么两个单侧极限都存在且等于ℓ。你应该能够证明，反之，如果f在a处的两个单侧极限都存在且等于ℓ，那么lim$(f, a) = \ell$。尽管这种紧密对应，单侧极限在重要情况下自然出现，尤其是在我们希望证明极限的存在性而没有关于f的具体信息时，参见定理4.30。

示例4.29 设sgn : $\mathbf{R} \to \mathbf{R}$ 为符号函数。因为sgn $x = 1$对所有 $x > 0$，我们有lim$(\text{sgn}, 0^+) = 1$；同样，lim$(\text{sgn}, 0^-) = -1$。结合前一段的观察，我们再次得出结论，lim$(\text{sgn}, 0)$不存在，比较示例4.25。□

单调函数的极限

单调函数具有简单的极限行为；这在 $\mathbf{R}$ 的完备性属性中得到了掩饰，并解释了完备性在分析中的关键作用。定理4.30是为非递减函数（特别是递增函数）而陈述的。对于非递增函数有一个明显的类似物。这一结果清楚地表明，“上确界”的概念是对于无限多个元素的有限集正确的“最大值”的推广。

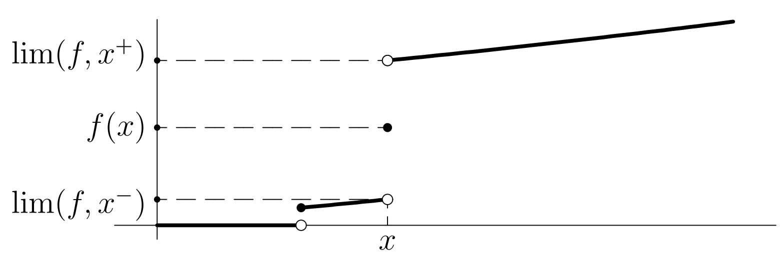


图4.4：单调函数的单侧极限。

定理4.30。设$A = (a, b)$为一个开区间，且设$f : A \to \mathbf{R}$为一个非递减函数。那么对于所有$x \in A$，单侧极限lim$(f, x^{\pm})$存在。

证明。要证明一个函数有极限，需要有一个候选极限。对于非递减函数，倾向于猜测在 x 处的下方极限是所有函数值 $f(y)$ 的“最大值”。

与 $y < x$。（我们不希望允许 $y = x$，因为“在 x 处的极限看不到函数值 x。”）不幸的是，集合 $\{f(y) \mid y < x\}$ 通常没有最大值；我们只知道集合是非空的，并且有上界（由 $f(x)$ 本身，因为 f 是非递减的）。然而，这些条件恰好足以暗示集合有一个上确界，并且我们推测如果 $f : (a, b) \to \mathbf{R}$ 是非递减的，那么

$$\begin{aligned} \ell^- &:= \lim(f, x^-) = \sup\{f(y) \mid a < y < x\}, \\ \ell^+ &:= \lim(f, x^+) = \inf\{f(y) \mid x < y < b\}. \end{aligned} \tag{4.5}$$

这个猜测不仅正确，而且可以从定义中轻松建立。固定 $\varepsilon >$ 为 0；根据上确界的定义，$\ell^- - \varepsilon$ 不是 $\{f(y) \mid a < y < x\}$ 的上界，因此存在一个 $z < x$ 满足 $\ell^- - \varepsilon < f(z) \leq \ell^-$。令 $\delta = x - z$；那么 $\delta >$ 为 0，并且因为 f 是非递减的，$z < y < x$ 意味着 $f(z) \leq f(y) \leq \ell^-$。但是 $z = x - \delta$，所以前面的蕴涵可以写成

$$x - \delta < y < x \implies \ell^- - \varepsilon < f(y) \leq \ell^-,$$

显然意味着 $|f(y) - \ell^-| < \varepsilon$。这意味着在 x 处的下极限是 ℓ^-，正如所声称的。另一个单侧极限通过完全相同的论证建立，参见练习4.19。 □

无限极限

回忆一下，通过添加两个非实数点 $+\infty$ 和 $-\infty$ 到 $\mathbf{R}$ 并扩展序关系，从而得到扩展实数。符号 $\pm\infty$ 不代表实数，目前对这些符号的算术运算尚未定义。直观上，$-\infty$ 和 $+\infty$ 是 $\mathbf{R}$ 的“左端点和右端点”。到目前为止，我们已考虑了函数在（有限）极限点的极限。对于许多实际和理论应用，我们希望在 $\pm\infty$ 处定义和使用极限。例如，物理系统的长期行为通常通过 $+\infty$ 处的极限来建模。

$+\infty$ 删除区间的类似物是某种形式的 $(R, +\infty)$ 区间，对于某些 $R \in \mathbf{R}$；当 R 变大时，该区间会缩小，即“更接近 $+\infty$。”同样，关于 $-\infty$ 的删除区间是 $(-\infty, R)$ 形式的区间。本节中的所有内容在 $-\infty$ 都有类似物，但为了简化说明，我们只将 $+\infty$ 明确提及。

定义和定理对于极限在这个新情况下保持不变。设 f 是一个定义域为 X 的实值函数。

一个集合 $\mathbf{R}$ 的子集。我们说 $+\infty$ 是 X 的一个极限点，如果 X 没有上界；直观上，X 包含接近 $+\infty$ 的点。在低级细节中，$+\infty$ 是 X 的极限点当且仅当对于每一个 $R \in \mathbf{R}$，存在一个点 $x \in X$ 满足 $x > R$。说 $f = A(\varepsilon)$ 在 $+\infty$ 附近是有意义的；这意味着如果我们把 f 限制在关于 $+\infty$ 的一个删除区间内，像在有限情况中一样，像在有限情况
中一样，像在有限情况中一样，像在有限情况中一样，像在有限情况
中一样，像在有限情况中一样，像在有限情况中一样，像在有限情况
中一样，像在有限情况中一样，像在有限情况中一样，像在有限情况
中一样，像在有限情况中一样，像在有限情况中一样，像在有限情况
中一样，像在有限情况中一样，像在有限情况中一样，像在有限情况

定义 4.31 设 $f : X \to \mathbf{R}$ 为一个定义域无上界的函数。若对于每一个 $\varepsilon > 0$，$f = \ell + o(1)$ 在 $+\infty$ 处，实数 ℓ 是 f 在 $+\infty$ 处的极限。

至于有限点的极限，$+\infty$ 处的极限是唯一的（如果存在），因此可以允许使用这些表达式

$$\lim_{x \to +\infty} f(x) = \ell \text{ and } \lim(f, +\infty) = \ell$$

作为实数的方程。我们将假设为有限点建立的定理可以推广到 $+\infty$ 的极限。

示例 4.32 每个常数函数在 $\pm\infty$ 处有明显的极限。恒等函数 $I(x) = x$ 在 $+\infty$ 处没有极限。（可以说极限是 $+\infty$，并且通过适当的定义这是一个定理。然而，一个函数有无限极限是没有实极限的特殊情况。）

在 $+\infty$ 处的极限出于不同原因不存在，当 $f : \mathbf{R} \to \mathbf{R}$ 是一个非常数周期函数时：该函数"无法决定接近哪个值"。正式地，设 x 和 y 是具有 $f(x) \neq f(y)$ 的点。在每个区间 $(R, +\infty)$ 上，f 取得 $f(x)$ 和 $f(y)$ 的值，但不存在一个数 ℓ，使得每个关于 ℓ 的区间都包含 $f(x)$ 和 $f(y)$。□

示例 4.33 回想一下，一个（实）序列是一个函数 {v*}：{v*} {v*}，通常表示为 ({v*})(v*)。如果一个序列 ({v*}) 收敛到 {v*}，则满足 lim {v*}。这个条件在 4.4 节中详细讨论。{v*}

倒数函数 $f(x) = 1/x$ 对于 $x \in (0, +\infty)$ 很容易看出当 $x \to +\infty$ 时趋近于 0：如果 $\varepsilon > 0$，那么 $f = A(\varepsilon)$ 在区间 $(1/\varepsilon, +\infty)$ 上。更一般地，如果 $k \geq 1$ 是一个整数，那么 $\lim_{x \to +\infty} 1/x^k = 0$:

$$\lim_{x \to +\infty} \frac{1}{x^k} = \left(\lim_{x \to +\infty} \frac{1}{x} \right)^k = 0^k = 0.$$

第一个等式是“一个乘积的极限是极限的乘积。”这个等式可以通过数学归纳法正式建立。或者，你可以使用夹逼定理。

使用这些初步结果，我们可以在 $+\infty$ 处找到有理函数的极限。在下一个例子中，将分子和分母除以 x^4 并使用定理 4.19 给出

$$\lim_{x\to+\infty}\frac{1-2x+3x^2}{x+x^4}=\lim_{x\to+\infty}\frac{(1/x^4)-(2/x^3)+(3/x^2)}{(1/x^3)+1}=\frac{0-2\cdot 0+3\cdot 0}{0+1},$$

所以极限是0。一般来说，如果分子次数不大于分母，则极限存在；如果分子和分母具有相同的次数，则极限非零:

定理4.34。设p和q是互质因子的多项式，即

$$p(x)=\sum_{k=0}^{n}a_k\,x^k,\qquad q(x)=\sum_{k=0}^{m}b_k\,x^k,\qquad a_n,\ b_m\neq 0.$$

如果 $n<m$，则 $\lim(p/q,+\infty)=0$，而如果 $n=m$，极限存在且等于 a_n/b_n。如果 $n>m$，极限不存在。

证明。直观上，最大度数的项是唯一重要的项，证明本质上利用了这个想法。首先，分母至多有有限个零，因此商在某个区间（$R,+\infty$）上定义。将分子和分母分别除以它们各自的最高次幂x；得到的结果是

$$\frac{p(x)}{q(x)}=x^{n-m}\cdot\left(\sum_{k=0}^{n}a_k\,x^{k-n}\right)\Big/\left(\sum_{k=0}^{m}b_k\,x^{k-m}\right).$$

括号内的各项单独有极限，因为它们是具有极限的单项式的和；分子趋近于a_n，分母趋近于b_m，所以它们的商根据定理4.19趋近于a_n/b_m
。

如果 $n\le m$，则该陈述是直接的，因为“剩余”项 x^{n-m} 当 $n<m$ 时趋近于 0，当 $n=m$ 时恒等于 1。如果 $n>m$，则该项在 $+\infty$ 处没有极限，因此根据推论 4.21，商 p/q 也没有极限。

□

如果f的定义域包含某些形如$(R,+\infty)$的区间，并且如果$\lim(f,+\infty)=\ell$，那么直线$y=\ell$被称为水平渐近线

图 f 的。定理 4.34 表示，有理函数 p/q 的图有水平渐近线，前提是分子的次数不超过分母的次数。

无限极限

有两个上下文中，“无限”可以被视为“极限”。在第一个上下文中，两个点（$+\infty$和$-\infty$）被添加到$\mathbf{R}$中，我们区分大的正函数值和大的负[1]函数值。在第二个上下文中，只有一个点，称为∞，被添加到$\mathbf{R}$中，我们不能再谈论涉及∞的顺序关系。每种技术在不同情况下都很有用，并且将依次进行研究。为了避免一定程度的重复，本节中所有函数都假定具有无界定义域。

扩展实数极限

有限极限的定义是有意义的，因为我们知道条件 $f = \ell + A(\varepsilon)$ 的含义。如果我们用 $+\infty$ 代替 ℓ，这就不再成立，因为我们不能对 $+\infty$ 进行代数运算。为了继续，我们必须以使“当 $\ell = +\infty$”有意义的方式重新表述有限极限的定义。说 $\lim(f, a) = \ell$ 意味着对于关于 ℓ 的每一个区间 I，都存在一个关于 a 的删除区间，使得 f 的限制的像包含在 I 中。这看起来不错，因为我们知道关于 $+\infty$ 的区间的含义，并且条件没有提到与 ℓ 的代数运算。

定义4.35 设f为一个函数。我们称f在a处的极限为$+\infty$ (或“当$f(x)$趋近于x时，a的极限是$+\infty$”)，如果对于每一个$R \in \mathbf{R}$，都存在一个关于a的删除区间，使得$f > R$。这个条件可以写成$\lim(f, a) = +\infty$或$\lim\limits_{x\to a} f(x) = +\infty$。

语言的一个特性是，如果 $\lim(f, a) = +\infty$，那么 $\lim(f, a)$ 不存在；极限的存在意味着极限是一个实数。当然，极限不存在并不意味着极限是 $+\infty$；想想0处的符号函数，或者 $\mathbf{Q}$ 在 $\mathbf{R}$ 中的特征函数。

[1]This is somewhat of an oxymoron, but is standard usage. “Large” refers to “large absolute value.”

如果通过限制关于 a 的某个删除区间，可以将 f 制成任意大的负数，那么我们称 $\lim(f,a)=-\infty$ 或 $\lim_{x\to a} f(x)=-\infty$。精确条件——除了一个小变化外——与定义 4.35 中的相同：结论是“$f<R$”而不是“$f>R$。”

某些涉及 $+\infty$ 的算术运算可以用极限来定义。像“$(+\infty)+(+\infty)=+\infty$”这样的等式是“如果 f 和 g 在 a 处有极限 $+\infty$，那么 $f+g$ 在 a 处有极限 $+\infty$”的缩写。在这种情况下，以下都是正确的（ℓ 表示任意正实数）：

$$
\begin{aligned}
+\infty \pm \ell &= (+\infty)+(+\infty)=(+\infty)\cdot\ell \\
&= (+\infty)\cdot(+\infty)=+\infty, \qquad \frac{\pm\ell}{(+\infty)}=0.
\end{aligned} \tag{4.6}
$$

证明几乎立即从定义中得出，留给你们（参见练习4.16）。以下表达式在以下意义上是不确定的[2]，即“答案”，如果它确实存在，取决于函数f和g，而不仅仅是它们的极限：

$$(+\infty)-(+\infty),\quad 0\cdot(+\infty),\quad \frac{\ell}{0},\quad \frac{0}{0},\quad \frac{(+\infty)}{(+\infty)}.$$

这里有一些典型的反例，其中 $a=$ 为 0：

- $(+\infty)-(+\infty)$ $f(x)=1/|x|$ 和 $g(x)=1/|x|-\ell$ 对于 ℓ 一个固定的实数；每个函数在 0 处有极限 $+\infty$，但它们的差有极限 ℓ。如果相反 $f(x)=1/x^2$，那么差有极限 $+\infty$。如果 $\lim(f,0)=+\infty$ 并且（比如说）$g=f-\chi_{\mathbf{Q}}$，那么 $\lim(g,0)=+\infty$ 但 $\lim(f-g,0)$ 不存在。

- $0\cdot(+\infty)$ $f(x)=\ell x^2$, $g(x)=1/x^2$；它们的乘积有极限 ℓ。如果 $f(x)=\pm|x|$，则乘积有极限 $\pm\infty$。如果 $f(x)=x$，那么乘积在 0 处无极限。

投影无穷极限

极限如$\lim_{x\to 0} 1/x$不存在，即使允许$\pm\infty$作为极限的可能值；在某种意义上，$1/x$从下方趋近于$-\infty$，但$1/x\ \to\ +\infty$从上方趋近于$x\ \to$。每当$f=p/q$是q有根的有理函数时，就会发生这种情况。

---

[2]Or “not consistently definable.”

奇数阶在 a (，通常假设 p 和 q 没有公因数)。解决这种烦恼的一种方法是将单个 ∞ 仅附加到 $\mathbf{R}$ 上。这相当于在扩展实数中将 $-\infty$ “粘合”到 $+\infty$ 上，并且当处理有理函数时更可取。关于 ∞ 的删除区间是 $\mathbf{R}$ 中闭、有界区间 $[\alpha, \beta]$ 的补集，并且集合 $\mathbf{R} \cup \{\infty\}$ 与这种“关于 ∞ 的区间”的概念构成了实射影线，表示为 $\widehat{\mathbf{R}}$ 或 $\mathbf{RP}^1$。在这种情况下，如果对于每个 $R > 0$，存在一个关于 a 的删除区间，使得 $|f| > R$，则称函数 f 在 a 处有极限 ∞。这与定义 4.35 的区别在于，这里我们不在乎 f 是正数还是负数，只在乎它有大的绝对值。

存在一些自然几何情况，其中单个“位于 ∞ 的点”比两个点更好。考虑通过原点且由其斜率描述的直线。具有大正斜率的直线几乎垂直，但具有大负斜率的直线也是如此。将斜率为 ∞ 的直线解释为垂直线，并且不区分 $+\infty$ 和 $-\infty$ 是有意义的。如果我们把实数 m 与通过 $(0, 0)$ 的斜率为 m 的直线相对应，那么每条非垂直线都对应一个唯一的实数，而 ∞ 对应于垂直轴，见图 4.5。

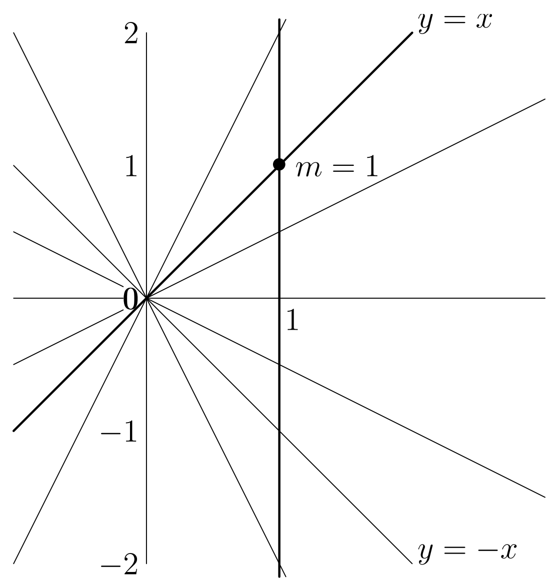


图4.5：平面通过原点的直线。

允许与 ∞ 进行的算术运算与允许与 $\pm\infty$ 进行的运算不同。例如，表达式 $\infty + \infty$

不确定，因为“$+\infty = -\infty$。”获得的优势是大多数除以0的情况都是明确的。精确地说，假设f和g是有理函数(g不为零)且$\lim(f,a) = \ell \neq 0$和$\lim(g,a) = 0$。那么$\lim(f/g,a) = \infty$；简而言之，$\ell/0 = \infty$对于$\ell \neq 0$。在类似假设下，$\ell/\infty = 0$和$\ell + \infty = \infty$对所有$\ell \in \mathbf{R}$成立。表达式$0/0$，$0 \cdot \infty$和∞/∞是不确定的；它们可以通过适当选择f和g来赋予任意值。

定理4.34的证明表明，如果f是一个有理函数，那么$\lim(f,-\infty)$和$\lim(f,+\infty)$存在且相等。将f的定义域视为$\widehat{\mathbf{R}}$而不是$\mathbf{R}$的子集，会得到一个非常令人满意的图像：当分母为零时，f的值是∞，而在∞处的值是极限值（这个极限值本身也可能是∞）。简而言之，有理函数可以被视为映射f: $\widehat{\mathbf{R}} \to \widehat{\mathbf{R}}$，参见练习4.18。

4.3 连续性

假设 {v*} 是在 {v*} 的一个区间上定义的函数。有两个数——函数值 {v*} 和极限 lim({v*})——分别描述 {v*} 在 {v*} 处和接近 {v*} 处的行为。极限不一定存在，因为 {v*} 在 {v*} 附近的行为可能“太复杂”而无法用单个数 {v*} 来描述。对于有理函数，我们已经看到这两个数是一致的。这种愉快的巧合值得一个名字：如果函数的定义域是一个开区间，那么当 lim({v*}) {v*} 时，该函数在 {v*} 处“连续”，如果它在定义域的每个点都连续，则简单地称为“连续”。要定义闭区间端点的连续性，需要单侧极限。而不是给出一系列特殊定义，我们以一种对所有函数都有意义的方式，对标准进行了一点不同的表述。

定义4.36 设$A \subset \mathbf{R}$非空。若函数f: $A \to \mathbf{R}$在$a \in A$处满足以下条件，则称其为在$a \in A$处连续：

> 对于每个 $\varepsilon > 0$，存在一个 $\delta > 0$，使得如果 x 是具有 $|x - a| < \delta$ 的 A 的一个点，那么 $|f(x) - f(a)| < \varepsilon$。

否则 f 在 a 处不连续。如果对于每个 $a \in A$，f 在 a 处连续，那么我们说 f 是连续的（在 A 上）。

不再需要规定 $x \neq a$，因为如果 $x = a$，结论 $|f(x) - f(a)| < \varepsilon$ 就是自动的。让我们比较一下

[3] This is even true when the value is ∞, and at the point ∞.

普通极限游戏；规则的主要变化以斜体表示。裁判指定函数 f 和点 a，该点必须是 f 的定义域中的点。假定的极限是 $\ell = f(a)$。玩家 ε 选择一个正的容差，玩家 δ 尝试控制射击半径，以确保每一枪都能击中目标。然而，玩家 δ 只需从 f 的定义域中的点射击，而 f 的定义域不必包含关于 a 的删除区间。

考虑一个序列 $f : \mathbf{N} \to \mathbf{R}$；如果 $n \in \mathbf{N}$，那么以 n 为中心的半径为 1 的删除区间不包含任何自然数。因此，玩家 δ 默认获胜，因为 f 的定义域包含 n 但不包含“邻近”的点。因此，对于每个 $n \in \mathbf{N}$，序列在 n 处自动连续。显然，如果 f 的定义域确实包含关于 a 的一个区间，那么当且仅当 $\lim(f, a) = f(a)$ 时，f 在 a 处是连续的。制定更一般的定义 4.36 的原因是，使用这个定义更容易陈述定理，并且连续性的“真正”性质不会被关于 f 的定义法问题所掩盖。

示例4.37 根据引理4.21，有理函数在其自然定义域上是连续的。符号函数在$\mathbf{R}^\times$的每个点处连续，但在0处不连续。$\mathbf{Q}$的特征函数在每处都不连续。示例3.11的分母函数在每个非有理点处连续，在每个有理点处不连续，根据命题4.28。□

一些连续函数的基本性质几乎立即从关于极限的类似结果中得出。对定理4.19的证明进行明显修改表明，连续函数的和或积是连续的，如果分母在a处非零，那么商也是连续的。连续函数的复合几乎显然是连续的:

命题4.38。设f和g是可复合的函数。假设f在a处连续，且g在$f(a)$处连续。那么复合函数$g \circ f$在a处连续。

证明。固定 $\varepsilon > 0$，并选择 $\eta > 0$，使得如果 y 是 g 的定义域内具有 $|y - f(a)| < \eta$ 的点，则 $\left|g(y) - g\big(f(a)\big)\right| < \varepsilon$。然后选择 $\delta > 0$，使得如果 x 是 f 的定义域内具有 $|x - a| < \delta$ 的点，则 $|f(x) - f(a)| < \eta$。这种选择 δ 的过程是 $g \circ f$ 在 a 处的连续性游戏中的获胜策略，因为如果 x 在 $g \circ f$ 的定义域内 (即 f) 的定义域内，那么

$$|x - a| < \delta \implies |f(x) - f(a)| < \eta \implies |(g \circ f)(x) - (g \circ f)(a)| < \varepsilon;$$

因此 $g \circ f$ 在 a 处连续。 $\square$

设 f 为一个定义域包含关于 a 的区间的函数。f 在 a 处的可能不连续性通常分为三种类型：点 a 是

- 可去间断点，如果存在$\lim(f, a)$但不等于$f(a)$。在这种情况下，"$f(a)$具有错误值。"通过在a处重新定义f为$\lim(f, a)$，可以消除间断点。说$\lim(f, a)$存在的那一点是可去间断点并不罕见，即使f在a处未定义。这通常是无害的，但并不完全正确。

- 如果一个单侧极限存在但不等于彼此（其中一个可能是 $f(a)$，也可能不是），则存在跳跃间断点。直观上，f 的图像在 a 处从 $\lim(f, a^-)$ 跳跃到 $\lim(f, a^+)$，见图 4.3。

- 如果至少有一个单侧极限不存在，则存在一个野[4]不连续性。

函数 $\chi_{\{0\}}$ (单点集 $\{0\}$) 的特征函数在 0 处有可去间断点，符号函数在 0 处有跳跃间断点，倒数函数 $x \mapsto 1/x$ 在 0 处有奇异的间断点。一个函数在有限区间内可能有无穷多个各种类型的间断点。分母函数在每个有理数处有可去间断点。根据定理 4.30，非递减函数只有跳跃间断点，并且不难安排有无限多个。最后，$\chi_{\mathbf{Q}}$ 在每个实数处都有奇异的间断点。间断点集有细微的限制，超出了本书的范围。例如，$f : \mathbf{R} \to \mathbf{R}$ 的间断点集不可能是无理数集。

连续性是一个局部性质，即它只依赖于函数在一点附近的性质。关于连续函数的许多最有趣的结果都是全局的；它们依赖于函数在其定义域内每处的性质。其中一些在第5章中介绍。

[4]This is not a standard term.

4.4 序列与级数

序列及其极限是分析中最重要的主题之一，无论是在理论上还是在应用中。可以说，序列的收敛性是函数具有极限的最简单方式。同时，序列以有趣的方式出现，例如映射的迭代、连分数和无穷级数。

设 $(a_n)_{n=0}^{\infty}$ 为一个序列。根据定义，如果对于每一个 $\varepsilon > 0$，都存在一个 $N \in \mathbf{N}$，使得 $|a_n - \ell| < \varepsilon$ 对于 $n > N$ 成立，那么这个序列有一个极限 $\ell \in \mathbf{R}$。这个条件有一个简单的表述，在区间上的函数中没有类似的形式：

> 一个序列（a_n）$_{n=0}^{\infty}$收敛到ℓ当且仅当关于ℓ的每个开区间都包含除了有限个之外的所有项a_n。

这些术语是根据它们在无限列表 a_0、a_1、a_2、... 中出现的次数来计算的，而不是根据图像中的点数来计算。例如，由 $a_n = (-1)^n$ 定义的序列的图像是 $\{-1, 1\}$，并且关于 1 的每个开区间都包含图像除了有限多个点之外的所有点。然而，半径小于 2 的开区间无法包含所有奇数项 a_{2k+1}、$k \in \mathbf{N}$，并且有无限多个这样的项。因此，这个序列不收敛到 1（也不收敛到任何其他 ℓ）。

如果 (a_n) 是 A 和 f 中的一个序列，并且 $A \to \mathbf{R}$ 是一个函数，那么 f 与 a 的复合是一个实序列 (b_n)，具有 $b_n = f(a_n)$。此类序列的收敛性可以用来确定 f 是否有极限；出于技术原因，人们使用不触及点 a 的序列。

定理4.39。设$f : A \to \mathbf{R}$是一个定义域包含关于a的删除区间的函数。那么f, a的极限存在且等于$\ell \in \mathbf{R}$，当且仅当对于每个收敛到a的序列(a_n)在$A \setminus \{a\}$中，$\lim\limits_{n\to\infty} f(a_n) = \ell$的极限存在。

证明。假设 $\lim_{f,a}= \ell$，并令 a_n 是 $A \setminus \{a\}$ 中收敛到 a 的一个序列。固定 $\varepsilon > = 0$ 并选择 $\delta > = 0$ 使得如果 $0 < < |x - a| < \delta$，则 $|f(x) - \ell| < \varepsilon$。然后选择 $N \in \mathbf{N}$ 使得如果 $n > N$，则 $0 < < |a_n - a| < \delta$；这是可能的，因为序列 a_n 收敛到 a 但永远不会等于 a。如果 $n > N$，则这些不等式意味着 $|f(a_n) - \ell| < \varepsilon$，因此 $f(a_n) \to \ell$ 随着 $n \to \infty$。(与命题4.38的证明进行比较。)

相反，假设“$\lim(f, a) = \ell$”是错误的，即极限存在但不等于 ℓ 或极限不存在。在博弈论

解释，玩家 ε 有一个获胜策略。让我们使用解释来追踪游戏的进程，其中玩家 ε 选择某个 $\varepsilon > 0$，玩家 δ 选择一个 $\delta > 0$，最后玩家 ε 选择一个点 x，其值为 $0 < |x - a| < \delta$。首先，玩家 ε 选择一个足够小的 $\varepsilon > 0$。因为玩家 δ 无法战胜 ε 的选择，因此对于每个自然数 n，都存在一个点 $a_n \in A \setminus \{a\}$，使得

$$|a_n - a| < \frac{1}{n} \quad \text{and} \quad |f(a_n) - \ell| \geq \varepsilon;$$

该点 a_n 是玩家 ε 对抗 $\delta = 1/n$ 的获胜选择。现在考虑序列 $(a_n)_{n=1}^{\infty}$。由于构造，这些项中没有一个是 a，但由于 $|a_n - a| < 1/n$，序列收敛到 a。最后，存在一个 $\varepsilon > 0$，使得对于所有 $n \in \mathbf{N}$，有 $|f(a_n) - \ell| \geq \varepsilon$，因此序列 (b_n) 在 $b_n = f(a_n)$ 下不收敛到 ℓ。 □

一个显然的证明修改适用于 $+\infty$ 或 $-\infty$ 的极限。定理 4.39 对于证明单侧极限不存在是有用的，否则这项任务可能会很混乱。例如，设 $f: \mathbf{R} \to \mathbf{R}$ 是一个非常数周期函数（例如，例如 3.25 节中的“查理·布朗”函数），并定义 $g: \mathbf{R}^{\times} \to \mathbf{R}$ 为 $g(x) = f(1/x)$。然后 lim(g, 0) 不存在。直观上，函数 g 在每个区间 $(0, \delta)$ 上振荡无限多次，因为函数 f 在区间 $(1/\delta, +\infty)$ 上振荡无限多次。为了给出一个正式的证明，构造两个正数序列，比如说 (a_n) 和 (b_n)，它们收敛到 0，但相应的函数值序列有不同的极限。为了明确起见，假设 f 的周期是 α。在 $(0, \alpha]$ 中选择点 x 和 y，使得 $f(x) \neq f(y)$；根据周期性，

$$f(x + n\alpha) = f(x) \neq f(y) = f(y + n\alpha) \qquad \text{for all } n \in \mathbf{N}.$$

设置 $a_n = 1/(x + n\alpha)$ 和 $b_n = 1/(y + n\alpha)$；这些序列具有所需的性质，您应进行验证。

定理4.39的证明可以很容易地修改以确立以下有用的结果。从概念上讲，评估一个连续函数与取一个序列的极限是可交换的，参见方程（4.7）。这个性质将在应用中反复使用。

推论4.40。设$f: A \to \mathbf{R}$为一个函数。那么f在$a \in A$处连续当且仅当

$$\lim_{n\to\infty} f(a_n) = f\Big(\lim_{n\to\infty} a_n\Big) \tag{4.7}$$

对于每个序列 $(a_n)_{n=1}^{\infty}$ 在 A 中收敛到 a。

定理4.41是序列定理4.30的类比；证明留作练习。这样的定理在无法事先猜测极限时，对于证明序列的收敛性是有用的。当然，对于非递增序列，也有类似的结果。

定理4.41。设(a_n)是一个非递减的实数序列。那么(a_n)收敛当且仅当它有上界。

柯西序列

如已暗示，序列的收敛可能是一个难以驾驭的理论条件，因为它除非已知极限，或者除非有其他信息（例如“序列是单调且有界的”）可用，否则无法验证特定序列。拥有一个不要求知道极限的一般收敛标准将是有用的。“柯西[5]准则”满足这一目的。

定义4.42 一个序列$(a_n)_{n=0}^{\infty}$是柯西序列，如果对于每一个$\varepsilon > 0$，存在一个$N \in \mathbf{N}$，使得m，$n \geq N$蕴含$|a_n - a_m| < \varepsilon$。

直观上，柯西序列的项可以通过在序列中足够远地前进而彼此任意接近。这个条件甚至不依赖于极限的存在，更不用说极限的确切值了。相比之下，序列的收敛意味着可以通过在序列中足够远地前进，使项任意接近一个固定的数（极限）。这种区别很微妙，但很重要。在继续阅读之前，简要考虑一下每个柯西序列是否收敛，或者每个收敛序列是否是柯西的，或者两者都是，或者都不是。作为一个提示，你应该没有困难地解决一个方向。

为了让你对柯西准则有一个感觉，这里有两个基本应用。

引理4.43。如果(a_n)是一个柯西序列，那么存在$R \in \mathbf{R}$使得$|a_n| \leq R$对所有$n \in \mathbf{N}$成立。简而言之，柯西序列是有界的。

证明。令$\varepsilon = 1$. 因为 (a_n) 是柯西的，存在 $N \in \mathbf{N}$ 使得 $|a_n - a_m| < 1$ 对于 n，$m \geq N$ 。设置 $m = N$ 并使用三角不等式

[5]KOH shee; in French, the *second* syllable is accented.

不等式，$|a_n| = |(a_n - a_N) + a_N| \leq 1 + |a_N|$ 对于所有 $n \geq N$ 。通过取.. . 得到所需的结论

$$R = \max(|a_1|, |a_2|, \ldots, |a_{N-1}|, 1 + |a_N|),$$

一个有限数字列表的最大值。 □

如果 (a_n) 收敛，那么 (a_n) 是柯西序列。

证明。令 ℓ 表示 (a_n) 的极限。固定 $\varepsilon >$ 为 0，并选择 $N \in \mathbf{N}$ 使得 $|a_n - \ell| < \varepsilon/$^2 对于 $n \geq N$ 成立。根据三角不等式，

$$|a_n - a_m| \leq |a_n - \ell| + |a_m - \ell| < \varepsilon$$

对于 m，$n \geq N$ 。 □

“逆命题”问题“每个柯西序列都收敛吗？”更为微妙；问题在于从一个柯西序列构造一个假设的极限 ℓ。完备性性质最终证明是至关重要的。

定理4.45。设$(a_k)_{k=0}^{\infty}$是实数的一个柯西序列。存在$\ell \in \mathbf{R}$使得$\lim\limits_{k\to\infty} a_k = \ell$。

证明。构造一个辅助序列 $(b_n)_{n=0}^{\infty}$ 如下：

$$b_n = \sup\{a_k \mid k \geq n\}. \tag{4.8}$$

在文字上，查看(a_k)的连续“尾部”，并取它们的“最大值”。序列(b_n)显然是非递增的：你不能通过移除元素来使集合的上确界变大！此外，引理4.43说(a_k)，因此(b_n)是有界的。定理4.41说明(b_n)有一个实数极限，我们称之为ℓ。

修复 $\varepsilon > 0$ 并使用柯西条件 (a_k) 来选择 $N_0 \in \mathbf{N}$ 使得 $|a_n - a_m| < \varepsilon$ 对于 n 和 $m \geq N_0$。接下来，选择 $N_1 \geq N_0$ 使得 $|b_n - \ell| < \varepsilon$ 对于 $n \geq N_1$。现在，$b_{N_1} = \sup\{a_k \mid k \geq N_1\}$，因此存在 $N \geq N_1$ 使得 $|b_{N_1} - a_N| < \varepsilon$。如果 $n \geq N$，那么

$$|a_n - \ell| \leq |a_n - a_N| + |a_N - b_{N_1}| + |b_{N_1} - \ell| < 3\varepsilon.$$

这意味着 $(a_k) \to \ell$。 □

康托尔对 $\mathbf{R}$ 的构造

作为一个简短的旁白（本书其他地方未使用），这里简要介绍康托尔构造实数的方法。首先，注意柯西序列的概念对于有理数序列是有意义的，即使一个人不知道实数；柯西准则纯粹是在序列本身的术语下定义的。

一个显著的缺陷是 $\mathbf{Q}$ 中存在柯西序列在 $\mathbf{Q}$ 中不收敛——没有有理极限。以下示例 4.47 中的 $x_0 = b = 2$（参见图 2.29）是有理数的柯西序列，但没有有理极限，因为不存在一个有理数的平方是 2。如果一个人生活在 $\mathbf{Q}$ 中，他只能同意这样的序列发散。

朴素地，人们可能会想定义一个实数为一个有理数的柯西序列；这在$\mathbf{Q}$的纯粹意义上是有意义的，并且考虑到一个实数的柯西序列是收敛的，我们会将一个序列与其（实）极限相对应。问题是许多柯西序列具有相同的极限，因此一个实数应该与具有相同实极限的柯西序列的集合相对应。不幸的是，这个“定义”不再仅指$\mathbf{Q}$，因此我们必须重新表述它。

康托尔声明，如果$\lim_{n\to\infty} |a_n - b_n| = 0$，则两个柯西序列$(a_n)$和$(b_n$
$)$是等价的。很容易看出，这在有理数柯西序列的集合上是一个等价关系（自反性和对称性是显然的，传递性由三角不等式得出）。正如你可能猜到的，康托尔将实数定义为在这种等价关系下的柯西序列的等价类。这个定义的美丽之处在于，可以使用我们已证明的定理（即极限保持算术和顺序关系，参见定理4.19和4.23）来检查域公理和顺序性质。

完整性是唯一剩余的性质，它通过“对角线论证”得到验证：如果对于每个$m \in \mathbf{N}$（，$A_m := (a_{n,m})_{n=0}^{\infty}$是一个柯西序列，即$(A_m)_{m=0}^{\infty})$是一个柯西序列的序列！并且如果$(A_m)$在顺序性质的意义上是单调递增且有上界的，那么（可以这样论证）存在一个递增的索引集$n_{k_{k=0}}^{\infty}$
，使得规则$b_k = a_{k,n_k}$定义了一个对应于集合$\{A_m\}$上确界的柯西序列
。在细节上，这证明了完备性。

如果您比较这是Cha中的程序构建 pt2

通过整数、有理数、实数和复数的依次构建，你会发现从有理数到实数的构建是最复杂的步骤，涉及到有理数的无限集合（割集）或柯西序列的等价类。你可能想知道是否存在一个更简单的构建，只涉及有理数的有限集合，或者可能是有限集合的等价类。（其他每个构建都使用前一种类型数对的等价类。）答案是可证明的没有。经过微小修改，定理3.19（由康托尔提出）表明:

> 设 $\mathbf{Q}^n$ 为有序的 n-元有理数序列（即长度为 n 的有理数序列），令 $\mathbf{Q}^\infty$ 表示 $\mathbf{Q}^n$ 在 $n \in \mathbf{N}$（上的并集，即所有有限有理数序列的集合）。不存在一个满射函数 $f : \mathbf{Q}^\infty \to \mathbf{R}$。

从有理数中仅使用有限个有理数集合构造实数是不可能的；有理数的有限集合不够用！

递归序列的极限

设 $X \subset \mathbf{R}$ 为一个区间（为了方便），设 $f : X \to X$ 为一个连续函数。f 的迭代在 X 上确定一个离散动力系统，一个点的轨道是定义为 $x_{n+1} = f(x_n)$ 的序列 $(x_n)_{n=0}^\infty$，对于 $n \geq 0$。有一种几何方法可以可视化一个点的轨道，见图 4.6。在正方形 $X \times X$ 中绘制 f 的图形，并绘制对角线。从水平轴上的 x_0 开始。上升到 f 的图形，然后向左或向右到对角线。重复，上升或下降到 f 的图形，然后向左或向右到对角线。生成的点的水平位置是 x_0 的迭代。

当前感兴趣的问题是：“如果(x_n)收敛，它的极限是什么？”一个很好的答案是，这样的序列必须收敛到f的固定点，即一个具有$f(\ell) = \ell$的点ℓ。极限和连续性的形式可以用来给出一个简单的证明：如果$\lim\limits_{n\to\infty} x_n = \ell$，那么

$$f(\ell) = f\Big(\lim_{n\to\infty} x_n\Big) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n = \ell.$$

第一个、第三个和第五个等式是定义。第二个等式是推论4.40，第四个等式从极限的定义中可以看出。

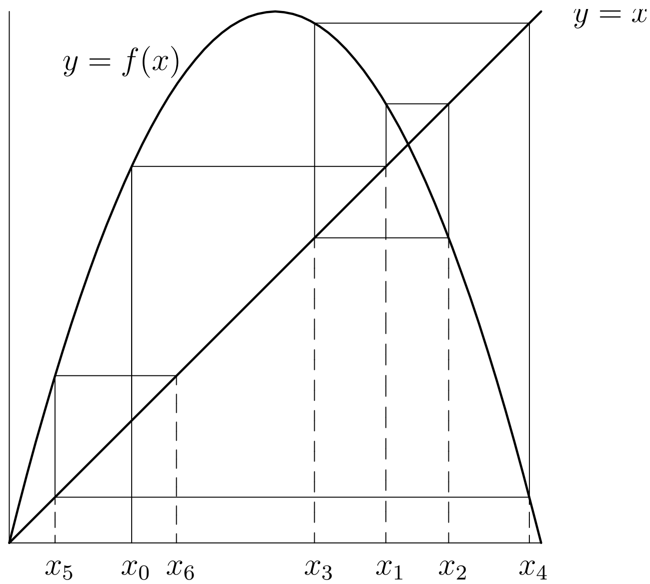


图4.6：点在 f 迭代下的轨道。

尽管这个观察并不总是定位递归序列的极限，但它将可能的选项减少到有限集，在许多有趣的情况下。

示例4.46 令 $f: [0, 1] \to [0, 1]$ 由 $f(x) = x^2$ 定义。f 的不动点是0和1，因此如果 x_0 在 f 下的轨道收敛，它必须收敛到这些点之一。如果 0 $\le x \le 1$，那么 $x^2 \le x$，因此对于每个 $x_0 \in [0, 1]$，x_0 的轨道是一个非递增序列，根据定理4.41，它收敛。对于这个函数，很容易看出1的轨道是对于所有 n 的常数序列 $x_n = 1$，而对于 $x_0 < 1$，x_0 的轨道收敛到0。

对于函数 g：由 $g(x) = -x$ 定义，$[-1, 1] \to [-1, 1]$，唯一的定点是 0。如果 $x_0 \neq 0$，那么 x_0 的轨道不收敛；在这个例子中，对定点的了解几乎毫无用处。□

示例4.47 设 $b >$ = 0；我们将证明 b 有一个实数平方根，即存在一个正实数 ℓ 满足 $\ell^2 = b$。详细信息和进一步信息留待练习4.23。Translated Text: 示例4.47 设 $b >$ = 0；我们将证明 b 有一个实数平方根，即存在一个正实数 ℓ 满足 $\ell^2 = b$。详细信息和进一步信息留待练习4.23。

修复 $x_0 > b$，并通过以下方式递归定义一个序列

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{b}{x_n}\right) \qquad \text{for } n \in \mathbf{N}. \tag{4.9}$$

代数和关于 n 的归纳法表明 (x_n) 在 0 以下有界且递减。根据定理 4.4 1，(x_n) 收敛于一个实数 $\ell \neq 0$。因为对于所有 n，$x_n > 0$，定理 4.23 表明 $\ell \geq 0$，因此 $\ell > 0$。另一方面，该序列是通过迭代连续函数 $f:(0,\infty) \to (0,\infty)$ 得到的，该函数定义为

$$f(x) = \frac{1}{2}\left(x + \frac{b}{x}\right),$$

所以 ℓ 是 f 的一个不动点。不动点方程 $f(\ell) = \ell$ 等价于 $\ell^2 = b$，这证明了 b 有一个平方根，因此记为 $\sqrt{b}$。实际上，这个序列非常迅速地收敛到 $\sqrt{b}$；练习 4.23 提供了收敛率的估计。□

这个例子典型地代表了分析问题“解”的情况。考虑某种无限对象（在这种情况下是一个序列），并希望得到一些信息（在这种情况下是极限）。通过一般结果证明了极限的存在；必须验证相关定理（定理）的假设以适用于特定情况。一个单独的结果（在这种情况下是固定点结果）将极限缩小到有限种可能性，并且一些额外的工作排除了所有选择除了一个。

如果您已经上过微积分课程，您至少熟悉本大纲的部分内容。为了找到“可微”函数的最大值，您将导数设为0，以排除除了（通常是）有限数量的可能位置之外的所有可能位置。一些额外的工作消除了所有错误的选择。剩下的部分是最大值的存在性；适当的存在性结果在第5章中得到证明。

无限级数

每个实数或复数序列 $(a_k)_{k=0}^{\infty}$（都与一个序列 $(s_n)_{n=0}^{\infty}$ 相关联，该序列为部分和（即“累计总和”），定义为

$$s_n = \sum_{k=0}^{n} a_k = a_0 + a_1 + \cdots + a_n. \tag{4.10}$$

原始序列可以通过以下方式从部分和中恢复：{v*}

$$a_0 = s_0, \qquad a_n = s_n - s_{n-1} \quad \text{for } n \geq 1,$$

所以你可能想知道我们为什么要费心计算总和。原因很简单，许多数学问题自然地表现为总和，而不是作为序列中的项。恒等式

$$s_n - s_m = \sum_{k=m+1}^{n} a_k, \qquad 0 \le m < n,$$

有用且将被反复使用。

定义4.48 设$(a_k)_{k=0}^{\infty}$是一个实数数列。如果部分和数列$(s_n)_{n=0}^{\infty}$收敛，则称(a_k)是可求和的。极限表示为

$$\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} s_n, \tag{4.11}$$

和称为级数的和。

另一种说法是，“级数”{v*}“收敛”。这种用法很方便，因为即使不知道序列（a_k）是否可求和，也常常写出表达式$\sum_k a_k$。然而，必须记住，序列（a_k）的收敛与级数（$\sum_k a_k$）的收敛是相当不同的；这两个概念之间唯一的一般关系在命题4.51中给出。

序列在逐项相加和数乘方面表现如预期。级数的乘积更为微妙，并且仅在附加假设下处理。你应该能够从定义中证明以下内容，而不会有任何困难。

定理4.49。如果(a_k)和(b_k)是可求和的，并且$c \in \mathbf{R}$，那么序列$(a_k + b_k)$和(ca_k)是可求和的，并且

$$\sum_{k=0}^{\infty}(a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k, \qquad \sum_{k=0}^{\infty}(ca_k) = c\sum_{k=0}^{\infty} a_k.$$

序列求和可以被视为“使用无限账本进行核算。”你有一个无限长的数字列表 a_0，a_1，a_2，...，依次将它们相加，并询问累积总和是否趋近于某个极限。如果是这样，该序列是可求和的，而这个极限（松散地）被视为列表中数字的“和”，按照呈现的顺序。（这个前提很重要。）序列的可求和性仅取决于序列的尾部。换句话说，丢弃有限个

序列的项不变不影响可和性（尽管当然它通常会改变实际的和）。显然，可以删除消失的项而不影响可和性或实际的和。此外，可以将任意多个零项随机排列到列表中，只要所有原始项保持相同的顺序。

可加性是一个微妙且略带反直觉的性质。在无限账本中，删除所有消失的项，并将项分为贷方（正项）和借方（负项）。如果原始序列是可加的，那么我们可以推断出以下两点之一：

- 信用和借方可以分别求和（有有限和），或者

- 信用和借方分别无法求和（具有无限和）。

在第一种情况下，总和不受重新排列的影响（这需要一些工作来证明），但在第二种情况下，账簿的总和类似于不确定表达式（$+\infty$）$-(+\infty)$，项的顺序很重要；"序列项的重新排列"可以改变可求和性，并可以改变总和的值！尽管有这种表示法，但一个级数实际上并不是无限多个项的总和，而是部分和的极限。

我们已经遇到了两个可求和序列的例子：几何级数（练习2.16）和实数的十进制表示。为了记录，以下是关于几何级数的完整故事。

命题4.50。设$a \neq 0, r \in \mathbf{R}$。序列$(ar^k)_{k=N}^{\infty}$可求和当且仅当$|r|<1$，并且在这种情况下

$$\sum_{k=N}^{\infty} ar^k = \frac{ar^N}{1-r}.$$

证明。对于每个 $n \in \mathbf{N}$,

$$\sum_{k=N}^{N+n} ar^k = ar^N \sum_{j=0}^{n} r^j.$$

从示例2.20中回忆起,

$$(*) \qquad \sum_{j=0}^{n} r^j = \begin{cases} \dfrac{1-r^{n+1}}{1-r} & \text{if } r \neq 1, \\ \\ (n+1) & \text{if } r = 1. \end{cases}$$

如果 $|r|\ >1$ 或 $r\ =1$，则部分和无限增长（绝对值），因此该级数不收敛。当 $r\ =\ -1$ 时，部分和交替为 1 和 0，因此在这种情况下该级数不收敛。

最后，如果 $|r|<1$，那么 $r^n\to 0$，根据练习2.5(c)的 $n\to\infty$，所以方程($*$)中的级数和为$1/(1-r)$。根据定理4.49，得出所需的结论。 □

几何级数是最基本的级数之一，因为它们可以明确地求和；也就是说，级数不仅已知收敛，而且和可以计算。十进制表示是特殊的，因为它们的项是非负的，所以部分和的序列是非递减的。一般来说，收敛级数并不那么行为良好（既不能明确求和，也没有单调的部分和），因此，拥有确定收敛性的通用理论工具是很有用的。推导级数收敛性测试的基本工具是柯西准则，应用于部分和的序列。收敛条件的非常简单的例子是消失准则：

命题4.51。如果$(a_k)_{k=0}^{\infty}$是一个可求和序列，那么$\lim_{k\to\infty} a_k=0$。

证明。(a_k) 的可和性意味着，根据定义，部分和序列 (s_n) 是收敛的，因此根据引理4.44是柯西序列。根据“柯西序列”的定义取 $m=n+1$，意味着对于所有 $\varepsilon>0$，存在一个 $N\in\mathbf{N}$ 使得 $n\geq N$ 蕴含 $|a_{n+1}|=|s_m-s_n|<\varepsilon$。这恰好意味着 $\lim_{n\to\infty} a_n=0$。 □

不能过分强调，没有适用于所有级数的“通用”收敛或发散测试，也没有评估可求和序列之和的程序。特别是，命题4.51的逆命题是错误的：一个序列可能收敛到0，但不可求和。对于正项序列，可求和度量以一种微妙的方式衡量项趋于0的“速率”。

比较测试

几乎所有关于级数收敛性的测试都依赖于与已知收敛的级数的比较。下一个定理是比较测试，而引理4.54是极限比较测试。

定理4.52. 设(b_k)是一个可求和的非负项序列。如果(a_k)是一个满足对所有$k \in \mathbf{N}$有$|a_k| \le b_k$的序列，那么(a_k)是可求和的。

因为可和性是 (a_k) 的尾部的一个性质，所以假设可以减弱为“存在 $N \in \mathbf{N}$ 使得 $|a_k| \le b_k$ 对于 $k \ge N$ 成立。”其逆否命题是一个发散性测试：如果 $(|a_k|)$ 不可和，那么 (b_k) 也不可和。

证明。设(s_n)为(a_k)的部分和序列，设(t_n)为(b_k)的部分和序列。根据三角不等式，如果$m < n$，那么

$$\begin{aligned}|s_n - s_m| = \left|\sum_{k=m+1}^{n} a_k\right| &\le \sum_{k=m+1}^{n} |a_k| \\ &\le \sum_{k=m+1}^{n} b_k = |t_n - t_m|.\end{aligned}$$

因为(b_k)是可求和的，(t_n)是柯西的，所以之前的估计表明(s_n)也是柯西的，即(a_k)是可求和的。 □

定理4.53。设(b_k)是一个可求和的正项序列。如果(a_k)是一个极限存在序列，那么(a_k)是可求和的。

证明。如果 $a_k/b_k \to \ell$ 作为 $n \to \infty$，那么 $|a_k|/b_k \to |\ell|$ 作为 $n \to \infty$。为了看到这一点，将引理4.40应用于绝对值函数。根据定理4.52，只需证明 $(|a_k|)$ 是可求和的。取一个实数 $M > |\ell|$ 并用 $\varepsilon > 0$ 写出 $M = |\ell| + \varepsilon$。选择 $N \in \mathbf{N}$ 使得

$$n \ge N \implies \left|\frac{|a_k|}{b_k} - |\ell|\right| < \varepsilon.$$

一点代数可以表明，如果 {v*}，那么 {v*}。序列 ({v*}) 可由定理 4.49 求和，因此 ({v*}) 可由定理 4.52 求和。

□

推论4.54。如果(a_k)和(b_k)是正项序列，并且如果

$$\lim_{k\to\infty} \frac{a_k}{b_k} = \ell \ne 0,$$

那么 (a_k) 是可求和的当且仅当 (b_k) 是可求和的。

绝对可和性

如前所述，可和性涉及两个技术问题：正项和负项分别是否有有限和，或者是否存在基于和项顺序的偶然抵消？在本节中，我们研究前一种情况。

定义4.55 设(a_k)为一个可求和序列。如果$(|a_k|)$也是可求和的，那么我们称(a_k)为绝对可求和的，并且称级数$\sum_k a_k$为绝对收敛的。如果$(|a_k|)$不可求和，那么(a_k)为条件可求和的，并且$\sum_k a_k$为条件收敛的。

每个序列 (a_k) 都恰好属于以下类别之一：不可求和、条件可求和或绝对可求和。显然，非负项序列要么是绝对可求和的，要么是不可求和的。

给定一个实数序列(a_k)，通过以下方式定义其正项和负项相关序列：

$$a_k^+ = \max(a_k, 0), \qquad a_k^- = \min(a_k, 0). \tag{4.12}$$

例如，如果 $a_k = (-1)^k k$，那么

k	0	1	2	3	4	5	6	$\cdots$
a_k	0	-1	2	-3	4	-5	6	$\cdots$
a_k^+	0	0	2	0	4	0	6	$\cdots$
a_k^-	0	-1	0	-3	0	-5	0	$\cdots$

命题4.56. 一个实数序列(a_k)是绝对可和的当且仅当两个序列$(a_k^\pm)$都是可和的。

证明。首先观察对于所有 k，有 $|a_k^+| = a_k^+$ 和 $|a_k^-| = -a_k^-$；因此，正项或负项序列可求和当且仅当它们是绝对可求和的。

现在，对于所有 k，$|a_k^\pm| \leq |a_k|$，因此根据定理4.52（比较测试），如果 (a_k) 绝对可求和，那么 $(a_k^\pm)$ 也是可求和的。反之，对于所有 k，$|a_k| = a_k^+ - a_k^-$，因此如果 $(a_k^\pm)$ 可求和，那么 (a_k) 是绝对可求和的。 □

$m : \mathbf{N} \to \mathbf{N}$ 是一个双射，(a_k) 是一个序列。由 $b_k = a_{m(k)}$ 定义的序列是一个重新排列。非正式地说，重新排列是同一张账簿，以不同的顺序阅读。注意，

然而，这种重新排列不能做像“读取所有偶数项，然后所有奇数项”这样的事情，因为偶数项的列表本身就是一个无限列表。我们的下一个目标是证明“重新排列一个绝对可求和的序列不会改变和”。

定理4.57。设(a_k)绝对可和，设(b_k)为一个重排。则(b_k)绝对可和，且$\sum_k b_k = \sum_k a_k$。

证明的想法很简单：取足够的项 a_k 来近似 $\sum_k a_k$，然后在重新排列中足够远，以包括所有选定的项。(b_k) 的对应部分和也接近 $\sum_k a_k$
。

证明。设A_n为(a_k)的第n个部分和，B_n为(b_k)的第n个部分和，$A =$为$\lim_n A_n$。我们希望证明$(B_n) \to A$。

修复 $\varepsilon > 0$，并使用绝对可加性选择 N_0 以使

$$\sum_{k=0}^{\infty} |a_k| - \sum_{k=0}^{N_0} |a_k| = \sum_{k=N_0+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

特别地，$|A - A_{N_0}| < \varepsilon/2$。现在选择$N \in \mathbf{N}$，使得加数$a_0$，...，$a_{N_0}$属于项$b_0$，$b_1$，...，$b_N$；这是可能的，因为$(b_k)$是$(a_k)$的重新排列。如果$n \geq N$，那么

$$|B_n - A_N| = \left| \left(\sum_{k=0}^{n} b_k \right) - (a_0 + a_1 + \cdots + a_N) \right| \leq \sum_{k=N+1}^{\infty} |a_k| < \frac{\varepsilon}{2},$$

所以 $|B_n - A| \leq |B_n - A_N| + |A_N - A| < \varepsilon$。 □

我们关于绝对可和性的最后一个一般结果涉及级数的乘积。令(a_k)和(b_ℓ)为可和序列。我们想要理解“无限双重和”$\sum_{k,\ell} a_k b_\ell$，并在可能的情况下，用(a_k)和(b_ℓ)的和来评估这个表达式。第一个问题是求和是在$\mathbf{N} \times \mathbf{N}$上进行的，虽然这个集合是可数的，但没有“自然”的枚举方法可以用来得到一个序列。第二个问题是，“和”可能取决于我们选择的枚举，就像级数的值可能在重新排列下改变，或者可能根本不存在。如果(a_k)和(b_ℓ)中没有一个绝对可和，这些潜在的障碍是真正的难题。如果任一序列是绝对可和的，那么双重和就是

个体和的乘积；然而，我们不需要这种一般性，而只处理两个序列都绝对可求和的情况。

双级数的“标准排列”由 (a_k) 和 (b_ℓ) 的柯西积提供，迭代和

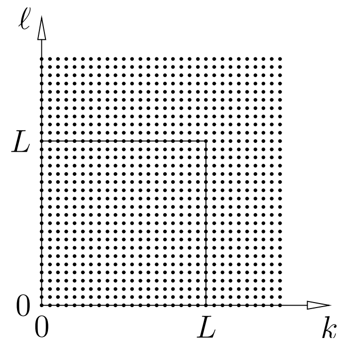
$$\sum_{n=0}^{\infty} \sum_{k+\ell=n} a_k b_\ell = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k}. \tag{4.13}$$

高斯乘积在之后极为有用，当我们用它来乘幂级数时。

定理4.58. 设(a_k)和(b_ℓ)是实数的绝对可和序列，且(c_m)是双无穷集合$\{a_k b_\ell \mid (k,\ell) \in \mathbf{N} \times \mathbf{N}\}$的任意枚举。那么$(c_m)$是绝对可和的，并且

$$\sum_{m=0}^{\infty} c_m = \Big(\sum_{k=0}^{\infty} a_k\Big)\Big(\sum_{\ell=0}^{\infty} b_\ell\Big).$$

空间的双序列索引可以如下可视化：



如同定理4.57中所述，思想是“大多数”对总和的贡献来自左下方的项；(c_n)的一些有限初始部分包括这些项，剩余项的总和很小。

证明。引入以下内容可以极大地简化符号：

$$A_n = \sum_{k=0}^{n} a_k, \qquad B_n = \sum_{\ell=0}^{n} b_\ell, \qquad C_n = \sum_{m=0}^{n} c_m,$$

$$\mathbf{A}_n = \sum_{k=0}^{n} |a_k|, \qquad \mathbf{B}_n = \sum_{\ell=0}^{n} |b_\ell|.$$

最后，让 $A = \lim_n A_n$，$B = \lim_n B_n$，$\mathbf{A} = \lim_n \mathbf{A}_n$，以及 $\mathbf{B} = \lim_n \mathbf{B}_n$。

序列 $P_n := A_n B_n$ 收敛于 AB，因为乘积的极限是极限的乘积。同样，$\mathbf{P}_n = \mathbf{A}_n \mathbf{B}_n$ 收敛于 $\mathbf{AB}$。

修复 $\varepsilon > 0$，并选择 $L \in \mathbf{N}$ 使得

$$(*) \qquad |AB - P_n| < \varepsilon \quad \text{and} \quad |\mathbf{AB} - \mathbf{P}_n| < \varepsilon \qquad \text{for } n \geq L.$$

现在选择 $N \geq L$，使得“方括号内”的每个项，即每个与 $a_k b_\ell$、k、$\ell < L$ 的乘积 c_0、c_1、...、c_N 都属于这些项。如果 $n \geq N$，那么

$$|C_n - P_n| = |C_n - A_n B_n| \leq \sum_{k \text{ or } \ell > L} |a_k| \cdot |b_\ell| = |\mathbf{AB} - \mathbf{P}_n| < \varepsilon,$$

根据 $(*)$ 的第二部分。因此，如果 $n \geq N$，那么

$$|AB - C_n| \leq |AB - P_n| + |P_n - C_n| < 2\varepsilon$$

通过$(*)$的第一部分。 □

比值测试和根测试

以上测试假定存在一个已知可求和的序列。接下来的两个测试，即比值测试和根值测试，可以通过将特定序列与几何级数进行比较来证明这些序列是可求和的。不幸的是，这两个测试都不太适用，尽管它们对“幂级数”都很有用，请参阅第11章。

定理4.59。设(a_k)是一个正项序列，并且假设

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \rho$$

存在。如果 $\rho < 1$，则 (a_k) 是绝对可求和的。如果 $\rho > 1$，则 (a_k) 不可求和。

证明。假设 $\rho < 1$. 选择 $r \in (\rho, 1)$，并设置 $\varepsilon = r - \rho > 0$。根据假设，存在一个 $N \in \mathbf{N}$，如下所示（见图4.7）

$$n \geq N \implies \left|\frac{a_{n+1}}{a_n}\right| < \rho + \varepsilon = r,$$

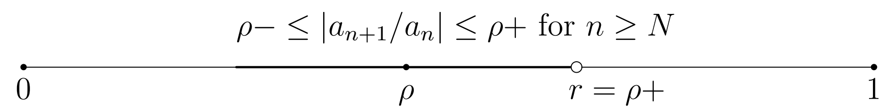


图4.7：比率测试中的边界比

或 $|a_{n+1}| \le |a_n| r$ 对于 $n \ge N$。通过 k 的归纳，

$$|a_{N+k}| \le |a_N| r^k \qquad \text{for } k \in \mathbf{N}.$$

因此，尾部$(a_{N+k})_{k=0}^{\infty}$的绝对值被$(|a_N| r^k)_{k=0}^{\infty}$所界定，这是一个收敛的几何级数。

如果 $\rho > 1$，那么选择 $r \in (1, \rho)$ 并如上所述进行论证表明 (a_k) 的尾部由一个发散的几何级数的项有下界。

□

定理4.60。设(a_k)是一个正项序列，并且假设

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = \rho$$

存在。如果 $\rho < 1$，那么 (a_k) 是绝对可求和的。如果 $\rho > 1$，那么 (a_k) 不可求和。

根值测试的证明类似于比值测试的证明，参见练习4.22。在这两个定理中，如果极限不存在，或者极限为1（在这种情况下，测试是“不确定的”），则无法得出任何结论。

示例4.61 假设我们希望测试以下级数

$$\sum_{k=0}^{\infty} \frac{k+1}{2^k} = 1 + \frac{2}{2} + \frac{3}{4} + \frac{4}{8} + \frac{5}{16} + \cdots$$

对于收敛；$a_k = (k+1)\,2^{-k}$。与几何级数 $\sum_k 2^{-k}$ 的比较没有帮助，因为所讨论的级数比这个几何级数大。因此我们尝试比值测试：

$$\frac{a_{n+1}}{a_n} = \frac{(n+2)\,2^{-(n+1)}}{(n+1)\,2^{-n}} = \frac{(n+2)}{2(n+1)}.$$

这个比率收敛到 $\rho = 1/2$，当 $n \to \infty$ 时，因此该级数通过比率测试绝对收敛。（求和该级数是另一回事。）□

示例4.62 令 $p > 0$，并考虑 p-系列

$$\sum_{k=1}^{\infty} \frac{1}{k^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \frac{1}{4^p} + \cdots .$$

目前我们开发的任何测试都无法解决这个级数的收敛性问题；项趋近于0，因此消失标准无法得出结论。项的减小速度比任意收敛几何级数的项慢，因此没有明显的比较可以做出。最后，比值测试和根值测试对于每个p-级数都返回$\rho = 1$，因此这些测试也无法得出结论。

□

下一定理，柯西测试，对于确定正项递减数列的收敛性很有用，并将使我们能够确定 p-级数的收敛性。

定理4.63。设$(a_k)_{k=1}^{\infty}$是一个正项序列，并假设对于所有$k \geq 1$。如果(b_n)是由$b_n = 2^n a_{2^n}$定义的序列，那么(a_k)和(b_n)同时可求和或不可求和。

证明。对于 $n \in \mathbf{N}$，定义序列中的“n 组”为那些其索引 k 在 $2^n + 1$ 和 2^{n+1} 之间的项：(a_k) 的 0 组是 $\{a_2\}$，第一组是 $\{a_3, a_4\}$，第二组是 $\{a_5, \dots, a_8\}$，第三组是 $\{a_9, \dots, a_{16}\}$，以此类推。第 n 组中有 2^n 项，除了前两项外，所有项都恰好属于一个组。

c_n 表示 (a_k) 的第 n 组项之和，即

$$c_n := \sum_{k=2^n+1}^{2^{n+1}} a_k.$$

序列 $(a_k)_{k=2}^{\infty}$ 和 $(c_n)_{n=0}^{\infty}$ 同时可求和或不可求和（因为它们由相同顺序的相同项组成）。

由于 (a_k) 是非递增的，

$$\frac{1}{2} b_{n+1} = 2^n a_{2^{n+1}} \leq c_n \leq 2^n a_{2^n+1} \leq 2^n a_{2^n} = b_n.$$

如果(b_n)是可求和的，那么(c_n)可以通过比较测试进行求和，而如果(c_n)是可求和的，那么$(b_n/)$²——因此(b_n)本身——是可求和的。 □

示例4.64（p-系列重访）固定$p >$为0并设置$a_k = k^{-p}$为$k \geq$为1。根据柯西测试，p-级数收敛当且仅当

$$\sum_{n=1}^{\infty} 2^n\, a_{2^n} = \sum_{n=1}^{\infty} 2^n\, (2^n)^{-p} = \sum_{n=1}^{\infty} (2^{1-p})^n$$

是收敛的。这是一个具有 $r = 2^{1-p}$ 的几何级数，因此当且仅当 $p > 1$ 时收敛。注意，这个论证证明了 p-级数在 $p > 1$ 时收敛；它并没有说明级数的和是多少，尽管它给出了上界和下界。截至本文写作时，3-级数（具有 $p = 3$ 的 p-级数）的确切和尚未知晓，尽管根据1978年Apéry的结果，其值已知是无理数。相比之下，2k-级数（k 为正整数）的值已知是某个有理数的倍数。本书的最终结果，依赖于即将到来的大部分材料，是对2-级数的评估。

如果您很挑剔，您完全可以合法地抱怨我们没有为非整数 p 定义 n^p。这一缺陷在第12章中得到纠正，之后您可以验证这里给出的估计可以推广到任意的实数 p。□

交错级数

设(a_k)为一个正项序列。该级数

$$\sum_{k=0}^{\infty} (-1)^k a_k = a_0 - a_1 + a_2 - a_3 + a_4 - a_5 + \cdots$$

称为交错级数。

交错级数在研究幂级数时经常出现，并且使用与我们迄今为止所用的不同技术进行研究。其思想是假设相邻项“相互抵消”，而不是假设绝对可和。基本充分条件归功于莱布尼茨，通常称为交错级数测试:

定理4.65。设(a_k)是一个趋于0的正项序列，设A_n为由此产生的交错级数的n项部分和。那么A：$=$ lim$_n$ A_n存在——该级数收敛——

$$A_{2n-1} < A < A_{2n}$$

对于所有 $n \geq 1$.

证明。直观上，证明很简单：0阶部分和是 a_0，后续部分和是通过交替向左和向右移动越来越小的量来获得的，步长趋于零。部分和能做什么，但会收敛?

一个形式化的证明是基于分别考虑偶数和奇数部分和，对上述论点进行解析性写作。

考虑偶数部分和 A_{2n}。下一个偶数部分和是

$$A_{2n+2} = A_{2n} - a_{2n+1} + a_{2n+2} = A_{2n} - (a_{2n+1} - a_{2n+2}) < A_{2n}.$$

证明奇数部分和构成一个递增序列是完全类似的。因为每个奇数部分和都小于下一个偶数和,

$$A_{2n-1} < A_{2n+1} < A_{2n+2} < A_{2n} \qquad \text{for all } n > 0. \tag{4.14}$$

归纳法证明，每个偶数和都大于每个奇数和。特别是，偶数和的集合有下界，奇数和的集合有上界。设L为奇数部分和的上确界，设U为偶数部分和的下确界。由于对所有k，夹逼定理表明$U = L$。

□

作为直接的结果，我们得到一个简单、明确的“误差界限”，该界限衡量通过部分和估计交错级数之和的准确性：误差不超过省略的第一个项的大小。

推论4.66. 在定理的记号中，$|A - A_n| < a_{n+1}$。

示例4.67 如果 p > = 0，则序列 (n^{-p})^$_{n=1}^{\infty}$ 是正的且递减的，因此莱布尼茨测试意味着该级数

$$\sum_{n=1} \frac{(-1)^{n-1}}{n^p} = 1 - \frac{1}{2^p} + \frac{1}{3^p} - \frac{1}{4^p} + \frac{1}{5^p} - \frac{1}{6^p} + \cdots \tag{4.15}$$

收敛。如上所述，当且仅当 $p > 1$ 时，此级数绝对可和。

该系列中 $p = 1$ 是条件收敛的交错调和级数,

$$\sum_{n=1} \frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots$$

这是一个级数，其和 S 可以被显式地找到，参见第14章。目前，我们从前三个部分和中可以看出

$$0.5 = 1 - \tfrac{1}{2} < S < 1 - \tfrac{1}{2} + \tfrac{1}{3} = 0.83\overline{3}.$$

交错调和级数也有相当引人注目的重新排列。不是交替取正负项，而是每次取一个正项和两个负项：

$$\begin{aligned}
1 - \tfrac{1}{2} - \tfrac{1}{4} + \tfrac{1}{3} - \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{5} - \tfrac{1}{10} - \tfrac{1}{12} + \cdots \\
&= \left(1 - \tfrac{1}{2}\right) - \tfrac{1}{4} + \left(\tfrac{1}{3} - \tfrac{1}{6}\right) - \tfrac{1}{8} + \left(\tfrac{1}{5} - \tfrac{1}{10}\right) - \tfrac{1}{12} + \cdots \\
&= \tfrac{1}{2} - \tfrac{1}{4} + \tfrac{1}{6} - \tfrac{1}{8} + \tfrac{1}{10} - \tfrac{1}{12} + \cdots \\
&= \tfrac{1}{2}\left(1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{10} + \tfrac{1}{5} - \tfrac{1}{6} + \cdots\right) = \tfrac{1}{2}S.
\end{aligned}$$

条件收敛级数的重排可以改变和！ □

关于交错调和级数重排的显著之处在于计算的明确性。考虑到下一个结果，这种行为本身并不令人惊讶。证明与莱布尼茨测试的证明有相似之处。

定理4.68。设$\sum_k a_k$为一个条件可求和级数。对于每一个实数A，存在一种排列，其收敛于A。还存在一些排列，其部分和发散到$+\infty$或$-\infty$。

一个形式化的证明很繁琐，但这个想法相当简单。正负项序列（a_k^+）和（a_k^-）都未能求和。假设$A >$为0。将（a_k^+）中的项相加，直到部分和大于A；这一定会发生，因为正项序列不可求和。现在开始从（a_k^-）中添加项，直到部分和小于A。再次利用（a_k^-）不可求和的事实。无限重复，按顺序选择正负项，使得部分和夹住目标A。

它相当明显，这个配方描述了原始序列的重新排列：每个项在新和中恰好出现一次。此外，和项趋近于零，因此括号内的和随着我们继续进行而越来越接近。括号内的和显然趋近于A，因此我们得到了所需的重新排列。

如果 $A < 0$，则从负项开始，否则使用相同的方法。为了得到一个发散到 $+\infty$ 的重排，将正项相加，直到部分和大于 1。然后减去一个负项，并加正项直到部分和大于 2。以此类推。你可以通过观察这个程序是否容易工作来衡量你对可数集的直觉的准确性。一个未经训练的人可能会反对说“正项比负项用得更快”；这不是问题，因为我们只为每个负项相加有限个正项，而且每个都有无限多个！

另一个潜在的问题是，一些负项的绝对值可能非常大。然而，根据消失准则（命题4.51），负项的序列收敛于零，因此经过有限个项之后，每个负项都不大于1/2（比如说），每个循环（添加一些正项并减去一个负项）至少将$1-\frac{1}{2}$添加到部分和中，因此部分和变得任意大。

练习

在所有要求你求极限的问题中，你应该给出完整的证明，无论是通过 ε-δ 论证，还是引用文本中的适当定理。对于有是/否答案的问题，根据适当的情况给出证明或反例。

练习4.1 以下陈述是对还是错?

(a) $\displaystyle\lim_{x\to 0}\frac{1}{x}=+\infty$　　(b) $\displaystyle\lim_{x\to 0}\frac{1}{x^2}=+\infty$

(c) $\displaystyle\lim_{x\to 0}\frac{1}{\sqrt{x}}=+\infty$　　(d) $\displaystyle\lim_{x\to 0}\frac{1}{\sqrt{|x|}}=+\infty$

在每一部分，你必须证明你的答案是正确的。 $\diamond$

练习4.2 设 $f:(a,b)\to\mathbf{R}$ 为一个开区间上的函数。

(a) 假设对于所有 x，$f(x+h)=f(x)+O(h^2)$ 接近 0。证明

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=0 \qquad \text{for all } x\in(a,b).$$

你能说更多吗?

(b) 假设存在一个在 (a,b) 上的函数 f' 具有如下性质

$$f(x+h)=f(x)+hf'(x)+o(h) \qquad \text{for all } x\in(a,b).$$

证明对于所有 x，有 $f'(x)=$ 的极限 $\lim\limits_{h\to 0}\dfrac{f(x+h)-f(x)}{h}$。

(c) 假设 f 和 g 满足部分 (b) 中的条件。证明 $f+g$ 和 fg 也满足该条件。作为这次计算的额外好处，你应该用 f、g、f' 和 g' 来表示 ($f+g)'$ 和 $(fg)'$。

$\diamond$

仔细阅读下一两个问题！

练习4.3 设$f:A\to\mathbf{R}$是一个定义域包含关于a的删除区间的函数。考虑以下条件：对于每一个$\delta>0$，存在一个$\varepsilon>0$使得

$$0<|x-a|<\varepsilon \implies |f(x)-\ell|<\delta.$$

这是否等价于“$\lim(f,a)=\ell$”？给出证明或反例。$\diamond$

练习4.4 设$f:A\to\mathbf{R}$是一个定义域包含关于a的删除区间的函数。考虑以下条件：对于每一个$\varepsilon>0$，存在一个$\delta>0$使得

$$0<|x-a|<\varepsilon \implies |f(x)-\ell|<\delta.$$

这是否等价于“$\lim(f,a)=\ell$”？给出证明或反例。$\diamond$

练习4.5 本练习与引理4.21相关。

(a) 证明如果 $\lim_{f, a}= ℓ ≠=0$ 且 $\lim_{g, a}$ 不存在，则 $\lim_{fg, a}$ 不存在。 (b) 找到一对函数 f 和 g，使得 $\lim_{f,=0}}==0$ 且 $\lim_{g,=0}}$ 不存在，但 $\lim_{fg,=0}}$ 存在。 (c) 找到一对函数 f 和 g，使得 $\lim_{f,=0}}$ 和 $\lim_{g,=0}}$ 不存在，但 $\lim_{fg,=0}}$ 存在。

◇

练习4.6 定义 $f:(-1,1)\to\mathbf{R}$ 通过 $f(x)=0$ 如果 $x\neq 0$, $f(0)=1$。求（带证明）lim(f, 0)，或者证明极限不存在。◇

练习4.7与文本中“上极限”的定义类比，仔细定义“f在a的下极限”。注意f的定义域，以及定义极限条件的量化句子。◇

练习4.8 假设 $p:\mathbf{R}\ \to\ \mathbf{R}$ 是一个非常数多项式函数。证明 lim($|p|,+\infty$) = $+\infty$。 ◇

练习4.9 精确定义“lim($f,+\infty$) = $+\infty$。” ◇

练习4.10 精确定义“lim(f,x_0) = $-\infty$。” ◇

练习4.11 设 $f:(0,+\infty)\to\mathbf{R}$ 为一个函数。证明

$$\lim_{x\to 0^+} f(\tfrac{1}{x}) = \lim_{x\to+\infty} f(x)$$

在两种情况下：要么两个极限都存在且相等，要么两个极限都不存在。◇

练习4.12 一个集合{v*}是稠密的，如果{v*}的每个区间都包含{v*}中的一个点。例如，{v*}是稠密的。假设{v*}和{v*}是在{v*}上的连续函数，并且对于某个稠密集合{v*}有{v*}。证明{v*}作为{v*}上的函数。

换句话说，在A上的连续函数最多只能有一个连续扩展到$\mathbf{R}$。◇

练习4.13 设 $f:\mathbf{R}\to\mathbf{R}$ 是一个连续的、非常数的、周期函数。证明存在一个最小的正周期。◇

练习4.14 设cb : $\mathbf{R}\to\mathbf{R}$ 为查理·布朗函数。

(a) $_{x\to+\infty}$ x cb(x) 作为扩展实数存在极限吗?

(b) 绘制函数 f 的图像：由 $f(x)$ = cb($1/x$) 定义，定义域为 (0, 1) $\to\mathbf{R}$。

(c) 极限 f, 0 存在吗?

◇

练习4.15 如果 f 和 g 是定义在 $\mathbf{R}$ 上的实值函数，那么 max(f, g) 是定义在 $\mathbf{R}$ 上的函数，其定义为

$$\max(f,g)(x) = \max\bigl(f(x), g(x)\bigr) = \begin{cases} g(x) \text{ if } f(x) \le g(x) \\ f(x) \text{ if } g(x) \le f(x) \end{cases}$$

并且 min(f, g) 定义方式类似。

(a) 证明如果 f 和 g 在 $\mathbf{R}$ 上是连续的，那么 max(f, g) 和 min(f, g) 在 $\mathbf{R}$ 上也是连续的。建议：使用定理2.24。

(b) 证明每个连续函数 $f\colon \mathbf{R} \to \mathbf{R}$ 可以写成连续、非负函数之差，即 $f = f_+ - f_-$。部分 (a) 和草图可能会有所帮助。

◇

练习4.16每个部分应按方程（4.6）的意义来解释。

(a) 证明对于所有 $x \in \mathbf{R}$，有 (+∞) $+ x = +\infty$。

(b) 证明 (+∞) $+ (+\infty) = +\infty$。

(证明：如果 $\ell > 0$，那么 $-\ell \cdot (+\infty) = -\infty$。)

◇

练习4.17每个部分都应在投影无穷极限的意义下解释。

(a) 证明对所有 $x \in \mathbf{R}$，有 $\infty + x = \infty$。

(b) 证明 $\infty + \infty$ 是不确定的。

(©) 证明对所有 $x \neq 0$，有 $x/0 = \infty$。

◇

练习4.18 设 $p\colon \mathbf{RP}^1 \to S^1$ 为球面投影，参见练习3.7。

(a) 证明在 {v*} 的意义上，$\lim\limits_{t\to\infty} p(t) = (0, 1)$，即每个分量函数都有所宣称的极限。

(b) 证明有理函数 $t \mapsto 1/t$ 在 p 下对应于单位圆在水平轴上的反射。

(©) 证明有理函数 $t \mapsto \frac{t-1}{t+1}$ 在 p 下对应于映射 $(x, y) \mapsto (-y, x)$，即绕原点逆时针旋转四分之一圈的单位圆。

$\diamond$

练习4.19 在定理4.30的假设下证明，{v*}

$$\ell^+ := \lim(f, x^+) = \inf\{f(y) \mid x < y < b\}.$$

给定一个非递减函数 f 的例子：$[0, 1] \to \mathbf{R}$，该函数有无限多个间断点。$\diamond$

练习4.20 证明推论4.40。 $\diamond$

练习4.21 证明定理4.41。 $\diamond$

练习4.22 证明定理4.60（根值测试）。 $\diamond$

练习4.23 此问题涉及示例4.47。写$a = \sqrt{b}$。

(a) 证明对所有 $n \in \mathbf{N}$，有 $x_n^2 > b$，并且序列 (x_n) 是递减的。得出结论：(x_n) 由 a 在下方有界。

(b) 对于 $n \in \mathbf{N}$，令 $e_n = x_n - a$ 为通过 x_n 估计 $\sqrt{b}$ 的误差。证明

$$e_{n+1} = \frac{e_n^2}{2x_n} < \frac{e_n^2}{2a} \qquad \text{for all } n \in \mathbf{N},$$

因此，这

$$e_{n+1} < 2a\left(\frac{e_1}{2a}\right)^{2^n} \qquad \text{for all } n \in \mathbf{N}. \tag{4.16}$$

这表示每次迭代中准确性的小数位数呈指数增长。

(c) 令 $b = 3$ 和 $x_0 = 2$。证明（当然不进行 $\sqrt{3}$ 的数值评估）$e_1/(2a) < 1/10$，并得出第六项近似 $\sqrt{3}$ 到 31 位小数，即 $|x_6 - \sqrt{3}| < 5 \cdot 10^{-32}$。建议：通过算术，$1.7 < \sqrt{3} < 1.8$。

$\diamond$

练习4.24根据示例4.47，存在一个函数 $\sqrt{\ }: [0,\infty) \to [0,\infty)$ 具有性质 $(\sqrt{x})^2 = x$ 对所有 $x \geq 0$ 成立。证明 $\sqrt{\ }$是连续的。建议：首先证明 $\sqrt{\ }$ 是递增的，然后使用恒等式 $(\sqrt{x}-\sqrt{a})(\sqrt{x}+\sqrt{a}) = x-a$。$\diamond$

练习4.25 求下列极限或证明其不存在。

(a) $\displaystyle\lim_{x\to-1} \frac{x^2-1}{x+1}$

(b) $\displaystyle\lim_{x\to 1} \frac{1-\sqrt{x}}{1-x}$

(c) $\displaystyle\lim_{x\to 0} \frac{1-\sqrt{1-x^2}}{x}$

(d) $\displaystyle\lim_{x\to 0} \frac{1-\sqrt{1-x^2}}{x^2}$

(e) $\displaystyle\lim_{x\to\infty} \sqrt{x^2+x}-x$

您可以使用第4.24题的结果。 $\diamond$

练习4.26 定义 $f: (0,\infty) \to (0,\infty)$ 为 $f(x) = \sqrt{2+x}$。

(a) 证明 f 是连续的，并找到 f 的不动点。

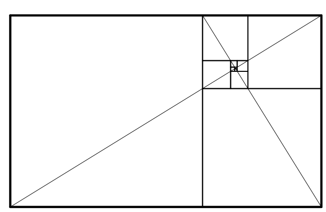
(b) 定义 $(x_n)_{n=0}^{\infty}$ 为 $x_0 = \sqrt{}$^2, $x_{n+1} = f(x_n)$，对于 $n \geq 1$；即，

$$x_1 = \sqrt{2+\sqrt{2}}, \qquad x_2 = \sqrt{2+\sqrt{2+\sqrt{2}}}, \qquad \dots$$

证明(x_n)收敛，并求出极限。

极限表示为 $\sqrt{2+\sqrt{2+\sqrt{2+\cdots}}}$。 $\diamond$

练习4.27 古希腊人认为最赏心悦目的矩形是黄金矩形，即当移除一个正方形时保持相同比例的矩形：



(a) 求黄金矩形宽高比 τ。计算 τ^2、$\tau - 1$ 和 $1/\tau$。

(b) 证明

$$\tau = [1; 1, 1, 1, \ldots] = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \ddots}}$$

(有几种进行的方法.)

(c) 在右手图的基础上解释部分(a)和(b)。

$\diamond$

练习4.28 评估无限连分数 $[2; 2, 2, 2, \ldots]$. $\diamond$

练习4.29 设(a_k)是一个实数或复数序列，而$b_k = a_{k+1} - a_k$是“差分”序列。

(a) 使用对 n 的归纳法证明

$$\sum_{k=0}^{n} b_k = a_{n+1} - a_0 \qquad \text{for all } n \in \mathbf{N}.$$

这种形式的和被称为望远镜和。

(b) 假设 $\lim_k a_k = \ell$ 存在；证明 (b_k) 是可求和的，并求出该级数的和。

(c) 评估无穷级数$\sum_{k=1}^{\infty} \frac{1}{k^2 + k} = \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+1} \right)$。

$\diamond$

练习4.30 评估级数

$$\sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{1}{4} + \frac{1}{16} + \frac{1}{36} + \frac{1}{64} + \cdots$$

$$\sum_{n=1}^{\infty} \frac{1}{(2n+1)^2} = \frac{1}{1} + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \cdots$$

关于 $S = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1} + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \frac{1}{36} + \frac{1}{49} + \cdots$。 $\diamond$

练习4.31 将前面的练习推广：设$p > 1$，使得p-级数$\sum_n n^{-p}$绝对可和，并设S为级数的和。求偶数项的和和奇数项的和。$\diamond$

练习4.32 假设 $f : (a, b) \to \mathbf{R}$ 是连续的，并且 $(x_n)_{n=0}^{\infty}$ 是 (a, b) 中的柯西序列。序列 $\big(f(x_n)\big)$ 必定是柯西序列吗？安心：决定答案是否为“是”或“否”可能是这个问题的最难部分。$\diamond$

第五章

区间上的连续性

在整个章节中，$f:[a,b]\to\mathbf{R}$ 是一个连续函数，即对于每个 $x\in[a,b]$，f 在 x 处是连续的。函数在一点的连续性是一个局部性质，甚至在区间上的函数连续性，表面上不过是一系列局部条件的集合。在本章中，我们推导出闭区间 $[a,b]$ 上连续函数的一些真正全局性质（例如有界性）。这些结果是微积分的技术基石，尽管其中一些陈述“直观上显然”，但它们都需要 $\mathbf{R}$ 的完备性公理作为基本要求。

一个主题贯穿本章，基于假设 f 的定义域是一个闭的、有界的区间。假设我们希望证明 f 满足某些性质 $\mathcal{P}$ (，例如在 $[a,b]$ 上的“有界性”)。根据 $\mathcal{P}$ 的性质，我们知道连续函数在局部（每个点的某个区间内）具有这个性质。我们从 a 开始，其中 f 具有这个性质。由于连续性，这个性质在某个包含 $t>a$ 的区间 $[a,t)$ 上成立。现在考虑满足在 $[a,t)$ 上 $\mathcal{P}$ 的 t 的上确界。如果 $t<b$，我们就会得到一个矛盾，因为 t 处的连续性意味着 $\mathcal{P}$ 在一个略微更大的区间 $[a,t')$ 上成立；因此 $t=b$ 和 $\mathcal{P}$ 在 $[a,b)$ 上成立。由于 f 在 b 处的连续性，这个性质在整个区间 $[a,b]$ 上成立。在这个非常粗略的草图上，三个条件“闭的、有界的区间”都得到了本质的应用。如果省略了这些假设中的任何一个，或者如果函数 f 在甚至一个点上不连续，那么 f 在一般情况下就不再具有性质 $\mathcal{P}$。

直到20世纪初，数学家们才找到了用公理标准来取代“闭区间、有界区间”的条件。这些标准是研究多个变量的函数所必需的，因为“区间”在这里没有意义。抽象条件是

称为“紧致性”（取代闭集和有界性）和“连通性”（取代区间）。这些条件中的每一个都可以用游戏来表示，就像我们在第4章中为连续性所做的那样，但解释游戏规则以及理解为什么一般标准是“正确”的，需要我们走得很远。对这些“拓扑”条件的深入研究属于更高级的课程，而且它们将不再被提及。

5.1 均匀连续性

我们注意到方程 $\ell = \lim(f, a)$ 可以被视为一个两人游戏。在本节中，我们遇到了一个“更强”的连续性版本，在 o 符号下无法自然表达。然而，存在一个自然的博弈论解释。记住 $f : [a, b] \to \mathbf{R}$ 在其定义域的每个点都是连续的。

假设的对手，玩家 ε 和 δ，在游戏中寻求变化，因为 ε 持续输掉。回想一下，他们已经得到了一个函数 f，并且他们同意在 $x \in$ 的连续性游戏中采取 $\ell = f(x)$ $[a, b]$。他们使用的函数 f 给玩家 δ 在每个 $x \in [a, b]$ 提供一个获胜策略。这意味着他们设定了一个 x，玩家 ε 选择一个容差 $\varepsilon > 0$，玩家 δ 成功“满足”这个容差。然而，根据规则，点 x 在选择 ε 或 δ 之前就已经指定。玩家 δ 在了解 ε 和 x 的情况下进行选择。为了使游戏对玩家 δ 更具挑战性，他们按照以下方式更改了规则：

函数 f 已给出。玩家 ε 选择一个公差。现在需要玩家 δ 使用一个适用于所有 x 的单个 δ 来满足公差。如果玩家 δ 有一个获胜策略（如之前，针对完美玩家），那么函数 f 是“一致连续的”。没有必要假设 f 的定义域是一个区间，更不用说是一个闭的、有界的区间了：

定义5.1 设$f : X \to \mathbf{R}$为一个函数。我们称f在X上是一致连续的，如果对于每一个$\varepsilon > 0$，存在一个$\delta > 0$，使得如果$x, y \in X$和$|x - y| < \delta$，那么$|f(x) - f(y)| < \varepsilon$。

均匀连续性是一个“全局”性质：它需要同时考虑 X 的所有点。这就是为什么 o 符号不适合的原因。即使 f 在 X 的每个点的某个邻域内“表现良好”（比如说常数），也不能推断出 f 在 X 上是均匀连续的，参见例5.3。

如果 f 在 X 上是一致连续的，那么 f 在非空子集上的限制也是一致连续的。显然，一致连续的函数在其定义域的每个点都是连续的，因为玩家 δ 可以懒惰并且使用相同的 δ 而不考虑 x。要理解 X 上的连续性和 X 上一致连续性的区别，最好的方法是考虑例子。引理 5.2 给出了一致连续性的一个必要（但非充分）条件。

引理5.2. 设$f : X \to \mathbf{R}$是在一个有界区间上的均匀连续函数。那么f是一个有界函数：存在一个$M > 0$，使得$f = A(M)$在X上。

证明。这个想法是通过一致连续性，存在一个 $\delta > 0$，使得 f 在每个长度为 δ 的区间内变化不超过 1。因为 X 是有界的，它可以被有限个长度为 δ 的区间所覆盖，所以 f 的总变化是有限的。

形式上，取 $\varepsilon = 1$；由一致连续性，存在一个 $\delta > 0$ 使得
$|x - y| < \delta$ 蕴含 $|f(x) - f(y)| < 1$（假设 x 和 y 在 X 中）。令 x_0 为 X 的中点，令 ℓ 表示 X 的半径。选择一个大于 N 的自然数 ℓ/δ。想法是每个点 $x \in X$ 都可以通过一个“链”与 x_0 “连接”，该链由长度小于 N 的重叠区间组成；在每个这样的区间上，函数值的变化不超过 1，因此根据三角不等式，函数值 $f(x)$ 与 $f(x_0)$ 的差不超过 N，无论 x 如何。

为记录起见，这里是一个完整的论证。设$x \in X$，并考虑将x_0与$x = x_N$连接的等距点序列$(x_i)_{i=0}^N$。甚至有一个简单的公式来表示第i个点：

$$x_i = x_0 + \tfrac{i}{N}(x - x_0) = \left(1 - \tfrac{i}{N}\right) x_0 + \tfrac{i}{N}\, x, \qquad 0 \le i \le N.$$

对于每个 i，$|x_i - x_{i-1}| = |x - x_0|/N < \ell/N < \delta$，因此
$|f(x_i) - f(x_{i-1})| < 1$. 根据三角不等式，

$$|f(x) - f(x_0)| = \left| \sum_{i=1}^{N} \Big(f(x_i) - f(x_{i-1}) \Big) \right| \le \sum_{i=1}^{N} \Big| f(x_i) - f(x_{i-1}) \Big| < N.$$

这个估计对所有 $x \in X$ 都成立，因此再次根据三角不等式，

$$|f(x)| \le |f(x) - f(x_0)| + |f(x_0)| < N + |f(x_0)| =: M$$

对于所有 $x \in X$。 □

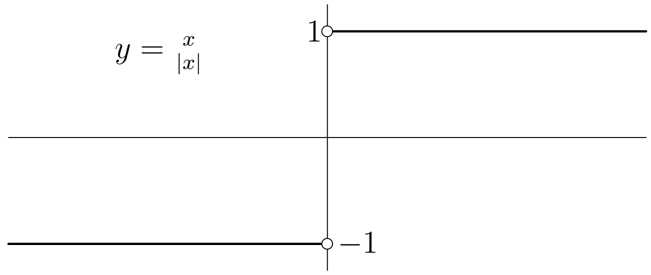


图5.1：一个局部常函数但不是一致连续的。

示例5.3恒等函数 $I:\mathbf{R}\to\mathbf{R}$ 是一致连续的；独立地取 $\delta=\varepsilon$ (给出一个获胜策略。尽管 I 在 $\mathbf{R}$ 上无界，但由于 $\mathbf{R}$ 不是一个有界区间，因此与引理5.2没有冲突。根据引理，I 在一个有界区间上的限制是有界的。

倒数函数 $f:(0,1)\to\mathbf{R}$ 在 $(0,1)$ 上连续，但无界，因此根据引理5.2，不是一致连续的，参见图5.2。

函数 $g(x)=x/|x|$ (符号函数在 $\mathbf{R}^\times$ 上的限制（图 5.1)）在局部是常数（对于每个 $a\neq 0$，都有一个关于 a 的开区间，使得 g 是常数），但不是一致连续的！取 $\varepsilon=1$；无论 $\delta>0$ 多小，点 $x=-\delta/3$ 和 $y=\delta/3$ 是 δ-接近的，但 $|g(x)-g(y)|=2>1=\varepsilon$。这个例子强调了均匀连续性的全局性质，并表明一个有界函数可能不是一致连续的。如果你倾向于抗议说 g 在 0 处不连续，请记住，“连续性”在不在 g 定义域内的点上是没有意义的。

□

均匀连续性具有几何解释：通过使区间足够短，可以使其像任意短。精确地说，对于每个 $\varepsilon>0$，存在一个 $\delta>0$，使得任意长度为 δ 的区间的像包含在某个长度为 ε 的区间内。根据这一观察重新考虑上述例子。长度为 δ 的区间具有形式 $(a,a+\delta)$；在倒数函数下，这样的区间可以具有任意长的像，见图5.2。在符号函数下，如果区间跨越原点，像就是两点集 $\{-1,1\}$，因此像不包含在长度小于2的区间内。相比之下，$(a,a+\delta)$ 在

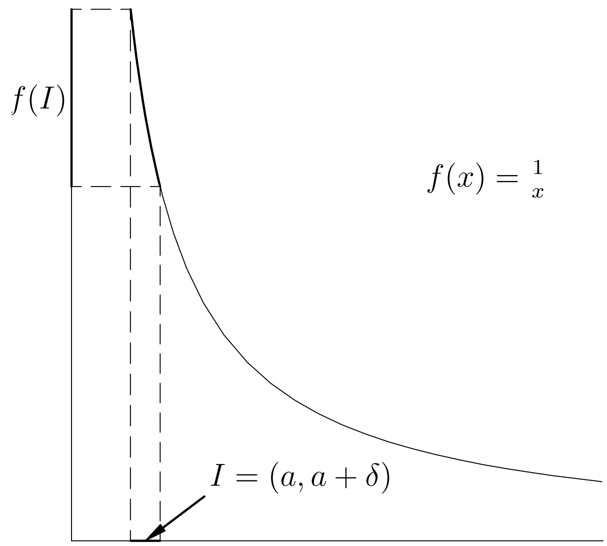


图5.2：将任意短区间映射到“长”区间。

恒等函数是区间$(a, a + \delta)$，通过取δ任意小，可以逻辑上任意缩小！

该节的主要结果，即下面的定理5.5，是一个简单的均匀连续性判据，在理论和实践上都很重要。为了将“均匀连续性”简化为更易管理的形式，我们引入一个辅助判据。对于固定的 $\varepsilon > 0$，我们说函数 f 在 X 上是“ε-驯服的”，如果存在一个 $\delta > 0$ 使得

$$x,\ y \in X \text{ and } |x - y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| \leq \varepsilon.$$

例如，如果 $f = A(M)$ 在 X 上，那么 f 是 $2M$ -tame 在 X 上，根据三角不等式。同样，每个特征函数都是 1-tame。在 X 上的均匀连续性与“对于每个 $\varepsilon > 0$，是 ε-tame 在 X 上”等价。

引理5.2的证明表明，在有限区间上的一个ε-驯服函数是有界的。因此，尽管倒数函数是连续的，但无论ε有多大，它都不是ε-驯服的。

ε-驯服的条件满足一个“修补”属性：

引理5.4. 假设 f 在闭区间 $[a, t]$ 和 $[t, c]$ 上是 ε-驯服的，并且 f 在 t 处连续。那么 f 在并集 $[a, c]$ 上是 ε-驯服的。

证明。利用 f 在 t 处的连续性来选择 $\delta_1 > 0$，使得 $|x-t| < \delta_1$ 蕴含 $|f(x)-f(t)| < \varepsilon/2$。三角不等式意味着 f 在 $(t-\delta_1, t+\delta_1)$ 上是 ε-驯服的。接下来选择 $\delta_2 > 0$ 和 $\delta_3 > 0$，分别对 $[a,t]$ 和 $[t,c]$ “起作用”，并设置 $\delta = \min(\delta_1, \delta_2, \delta_3) > 0$。

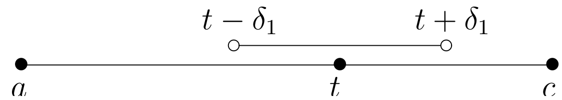


图5.3：f是ε-驯服的修补间隔。

一瞥图5.3可知，如果x和y是具有$|x-y| < \delta$的$[a,c]$中的点，那么这两个点都位于三个区间之一$[a,t]$、$[t,c]$或$(t-\delta_1, t+\delta_1)$中。因此，f在$[a,c]$上是ε-驯服的。 $\square$

定理5.5. 如果 $f : [a,b] \to \mathbf{R}$ 是一个在闭区间上的连续函数，那么 f 是一致连续的。

证明。我们将证明在假设下，对于每个 $\varepsilon > 0$，f 在 $[a,b]$ 上是 ε-驯服的。固定 $\varepsilon > 0$，并考虑集合

$$B = \{t \in (a,b] \mid f \text{ is } \varepsilon\text{-tame on } [a,t]\}.$$

我们的目标是证明 $b \in B$。这是通过本章引言中的草图，通过 f 的连续性、$\mathbf{R}$ 的完备性公理以及 $[a,b]$ 是一个闭区间和有界区间的实际情况来实现的。

因为 f 在 a 处连续，存在一个 $\delta > 0$ 使得 $f = A(\varepsilon/2)$ 在 $[a, a+2\delta)$ 上。根据三角不等式，f 在 $[a, a+\delta]$ 上是 ε-驯服的。因此 $a+\delta \in B$，所以集合 B 非空；显然，B 有上界（由 b 给出），所以根据完备性 B 有上确界，设为 β。

我们断言 $\beta = b$；为了看到这一点，假设 f 在 $[a,\alpha)$ 上关于 $\alpha < b$ 是 ε-驯服的。由于 f 在 α 处的连续性，存在一个 $\delta > 0$，使得 f 在 $[\alpha-\delta, \alpha+\delta]$ 上是 ε-驯服的，而根据假设，f 在 $[a, \alpha-\delta]$ 上是 ε-驯服的。引理5.4表明 f 在 $[a, \alpha+\delta]$ 上是 ε-驯服的，从而证明了 α 不是 B 的上界。这是“$\sup B = b$”的逆否命题。

我们仍然不知道 $b \in B$；可能是 $B = [a,b)$。然而，f 在 b 处连续，因此在某些区间 $[b-\delta, b]$ 上是 ε-tame 的，前一段说明 f 在每个区间 $[a,\alpha]$ 上都是 ε-tame 的，特别是 $[a, b-\delta]$。引理 5.4 的另一个应用证明了 f 在 $[a,b]$ 上是 ε-tame 的。由于 $\varepsilon > 0$ 是任意的，因此定理得证。 $\square$

定理5.5表明，例如，有理函数p/q在区间$[a, b]$上的限制是均匀连续的（假设q在$[a, b]$上非零）。因此，倒数函数$x \mapsto 1/x$在$[\delta, 1]$上对每个$\delta > 0$是均匀连续的。如上所述，限制到(0, 1]不是均匀连续的。f的均匀连续性既是定义域的性质，也是定义f的“规则”的性质。

5.2 连续函数的极值

定理5.5有一个重要的技术后果，它值得被称为“定理”而不是“推论”。定理5.6被称为极值定理。

定理5.6. 设$f : [a, b] \to \mathbf{R}$为一个连续函数。那么存在点$x_{\min}$和$x_{\max} \in [a, b]$，使得

$$f(x_{\min}) \le f(x) \le f(x_{\max}) \qquad \textit{for all } x \in [a, b].$$

这个定理并没有断言点 $x_{\min}$ 和 $x_{\max}$ 是唯一确定的；在“极端”情况下 f 是一个常数函数，区间内的每一个点既是 $x_{\min}$ 也是 $x_{\max}$。该定理也没有提供关于这些点可能在定义域中的位置的信息；必须使用其他工具来完成这个目的。最后，极值定理对具有甚至单个间断点的函数也没有任何说法，也没有关于定义域不是闭的、有界实数区间的函数的说法。该定理所断言的是，在某个无限数集（f 的值集合）中存在一个最大数和一个最小数。这比知道值集合有上界和下界更有用；这是函数最大值和最小值的“狩猎许可证”，在合适的假设下，保证猎物确实存在。

证明。根据定理5.5，f在$[a, b]$上是一致连续的，因此根据引理5.2，f在$[a, b]$上有界。考虑映射$f([a, b])$的像；它显然非空，并且如前所述，上下均有界。根据完备性，该像有上确界$y_{\sup}$和下确界$y_{\inf}$。我们想要证明这些数是函数值。这将通过证明如果$y_{\sup}$不是函数值，则定理5.5是错误的来完成。

如果 $y_{\sup}$ 不是一个函数值，那么对于每个 $x \in$，有 $f(x) < y_{\sup}$ [a, b]（注意严格不等式）。考虑函数 g：[a, b] 由以下定义→ $\mathbf{R}$

$$g(x) = \frac{1}{y_{\sup} - f(x)}.$$

通过假设，分母在 [a, b] 上非零，因此根据 f 的连续性和定理 4.19，函数 g 是连续的。然而，g 的分母可以任意小，因为根据上确界的定义，对于每个 $\varepsilon > 0$，存在一个 x_0 满足 $y_{\sup} - \varepsilon < f(x_0)$，即 0 $< y_{\sup} - f(x_0) < \varepsilon$。这表明 g 没有上界，这与所需的定理 5.5 相矛盾。使用函数 $h(x) = 1/(f(x) - y_{\inf})$ 的类似论证表明 f 实现了最小值。

□

在微积分课程中，极值定理通常没有证明地陈述，或者最多只是用一种可能性论证。必须警惕对可能性论证的信任，原因如下：极值定理的陈述对于函数 $f : \mathbf{Q} \to \mathbf{Q}$ 有意义，但尽管任何可能性论证都可能适用于此类函数，但定理的结论是错误的。以下给出了一些例子。

一个典型的可能性论证基于“定义”，即“连续函数”是指其图像可以在不拾起铅笔的情况下绘制出来.[1] 因为 f 的定义域是一个包含其端点的区间，所以 f 的图像必须在左侧“开始”并在右侧“结束”。因此，铅笔必须“因此”达到最高点和最低点；这些点的水平位置是 $x_{\max}$ 和 $x_{\min}$。

5.3 连续函数与中间值

闭区间上的连续函数 f 的另一个基本性质是“介值性质”。简单来说，这意味着如果 f 在某处为负，在另一处为正（否则），那么 f 必须在某处为零。一般来说，实数区间在连续函数下的像是一个区间。为了将此性质置于上下文中，我们简要回顾一下历史。

[1]For many common functions this is true enough, though it differs substantially from Definition 4.36. How many differences can you name?

分析中有理数的不足

在连续性的启发式定义中，介值性质是“显然的”：如果f的图像从x-轴下方开始并结束于x-轴上方，那么图像必须在某处穿过轴。但在仔细审视下，这个论点并不稳固。“当然，一个图像不可能不穿过轴就从轴下方到轴上方”听起来像是愿望式思考。事实上，这种事实显而易见的误解至少可以追溯到欧几里得，他假设通过圆心的每条直线都会在圆上相交于两个对径点。假设平面由所有具有有理坐标的点组成，即$\mathbf{Q}\times\mathbf{Q}$。（如果没有“平面是$\mathbf{R}\times\mathbf{R}$”的先见之明，就没有充分的理由假设“平面”不是$\mathbf{Q}\times\mathbf{Q}$。在我们看来，它们看起来是一样的。）以原点为中心、半径为1的圆的方程为$x^2+y^2=1$。这个方程在有理数中是有意义的，并且有无穷多个有理解。方程为$y=x$的直线也“存在于”有理平面中，并穿过圆心（参见图5.4的示意描绘）。同时解这些方程给出交点为$\pm(x,x)$，其中$x^2=1/2$。但是没有这样的有理数，所以这条直线根本不与圆相交！然而，有理平面$\mathbf{Q}\times\mathbf{Q}$满足欧几里得的所有公理，因此他的“不言而喻”的假设，即某些直线和圆相交，并不普遍成立。

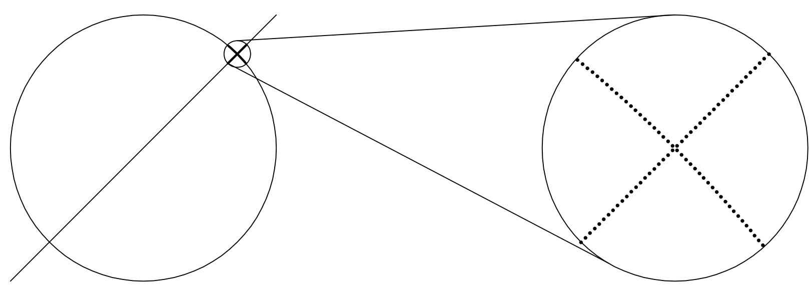


图5.4：圆是否与通过其中心的每条直线相交?

这个不足可以用函数来表述，表明中间值性质的启发式论证是不完整的。假设我们将数轴视为 $\mathbf{Q}$，并考虑由 $f(x)=x^2-2$ 定义的多项式函数 $f\colon\mathbf{Q}\to\mathbf{Q}$。在第 4 章中使用过的相同估计表明，该函数在 $\mathbf{Q}$ 上是连续的。在 $x=0$ 处，图形位于 x-轴下方，而在 $x=2$ 处，图形位于

以上。然而，在闭区间$[0, 2]$上，图形永远不会触及轴，因为对于所有有理数x，都有x^2 −^2 $\neq 0$。这不是一个孤立的例子；一个具有有理系数的一般多项式将在某处为正，在某处为负，但永远不会为零。

这些观察对于有理函数，例如 $g = 1/f : \mathbf{Q} \to \mathbf{Q}$ 由 $g(x) = 1/(x^2 - 2)$ 定义，可能更为显著。定义域和值域是正确的；右侧的表达式对每个有理 x 都有定义，并且本身也是有理的。这个函数在 $\mathbf{Q}$ 上是连续的，因为它是有理多项式的商，并且分母在所有地方都不为零！然而，g 在 $[0, 2]$ 上不是一致连续的，因为它在这个区间上是无界的。

您可能会争辩说："是的，但 $\mathbf{R}$ 没有间隙，所以这不可能在 $\mathbf{R}$ 中发生。"但什么是间隙，我们如何知道 $\mathbf{R}$ 没有间隙？为了强调，本章的定理以及我们对连续函数的直觉都是基于使用实数区间，而这种直觉是基于 $\mathbf{R}$ 的完备性属性。当粗心大意地应用可能性论证时，它们可能是错误的或严重误导的；只有通过精确的定义和逻辑、演绎的证明，才能确保避免错误。

这些评论应该使你质疑你的直觉。以免你觉得你所知道的一切都是没有根据的，让我们快速回顾相关事实。区间和绝对值在每一个有序域中都有意义，包括$\mathbf{Q}$和$\mathbf{R}$。因此，即使我们的数轴是$\mathbf{Q}$，我们也可以有意义地谈论函数的极限点和（均匀）连续性。引理5.2的证明，即一致连续的函数是有界的，只使用了有序域的性质，因此也在$\mathbf{Q}$中成立。然而，在证明定理5.5时，我们定义了一个特定的集合，然后取其上确界。同样，我们在证明极值定理时使用了上确界。因此，我们没有理由期望如果我们的数轴是$\mathbf{Q}$，这些定理的类似物仍然成立。上述讨论的反例证明了这种悲观是有根据的："$\mathbf{Q}$上的极值定理"根本不成立。此外，我们已经看到，在$\mathbf{Q}$上的连续函数可以取正值和负值，而不会在任何地方消失，这与我们的直觉相反。我们的下一个目标是证明我们对$\mathbf{R}$的直觉是正确的：在$\mathbf{R}$上的连续函数如果同时取正值和负值，那么它必须有一个零点。这个结论将有许多有用的后果；例如，这将意味着每个正实数都有一个立方根、四次根，以及一般地，每个整数的根。

中间值定理

下一个定义精确地表达了“如果 f 实现了两个值，那么它实现了介于这两个值之间的所有值”这一概念。注意这个标准如何简单地用区间来表述。

定义5.7 设f为一个函数。我们称f具有中间值性质，如果对于f的定义域中的每一个区间I，I的像$f(I)$也是一个区间。

要换一种说法，如果 $[a, b]$ 包含在 f 的定义域中，并且如果 $f(a) \neq f(b)$，那么对于 $f(a)$ 和 $f(b)$ 之间的每一个 c，都存在一个 $x \in (a, b)$ 满足 $f(x) = c$。你应该让自己相信这个条件与定义 5.7 在逻辑上是等价的。我们的几何直觉表明，连续函数（在区间上）确实具有介值性质。事实上，它们确实如此，根据介值定理：

定理5.8. 设f是一个定义在实数区间上的连续函数。那么f具有介值性质。

中间值定理在相同的意义上是一种“狩猎许可证”，就像定理5.6一样。我们不再寻找极值，而是现在寻求方程 $f(x) = c$ 的解，受限于 $a \leq x \leq b$。定理5.8表明，如果 f 是 $[a, b]$ 上的连续函数，并且如果数字 c 在 $f(a)$ 和 $f(b)$ 之间，那么方程 $f(x) = c$ 有一个解 $x \in (a, b)$。该定理并没有说只有一个解，也没有提供关于解的位置的任何信息。与极值定理一样，这些问题必须用其他工具来研究。

证明。因为连续函数的约束是连续的，所以只需证明如果 {v*} 在 [{v*}] 上连续，并且如果 {v*}，那么存在一个 {v*} 使得 {v*}。(通过考虑 {v*}，可以得出不等式反转的类似命题。) 事实上，只需假设 {v*} 0，因为 {v*} 可以被 {v*} 替换。直观的想法是“观察铅笔，看笔尖何时穿过 {v*}-轴”。当然，这样的推理将是循环的，因为我们试图证明铅笔穿过轴。相反，我们寻求“最大的” {v*} 使得 {v*} 在 [{v*}] 上是负的。现在，没有这样的 {v*}，但存在一个上确界；这就是 {v*} 的完备性如何进入画面。

回忆定理4.23的逆否命题：如果lim$(f, a) < 0$，那么在关于 a (的某个删除区间上 $f < 0$，即存在一个 $\delta > 0$ 使得

那 $f(x)$ <0 对于 $0<$ $|x-a|$ $<$ δ）。具有正极限的类似断言也成立。

假设 $f(a)<0<f(b)$，并考虑集合

$$B=\{x\in(a,b)\mid f(t)<0 \text{ for all } t\in[a,x]\}.$$

由于$\lim(f,a^+)=$ $f(a)$ <0，存在一个 δ >0 使得 f <0 在 $[a,a+\delta]$ 上；因此 $a+\delta\in B$，所以 B 非空。在另一个端点，$\lim(f,b^-)$ $=$ $f(b)$ >0，因此存在一个 δ' >0 使得 f >0 在区间 $[b-\delta',b]$ 上。这意味着 “$f(t)<0$ 对于所有 $t\in[a,b-\delta']$” 是错误的，或 $b-\delta'\notin B$。因此 B 被上界 $b-\delta'$ 限制。

让 $x_0=$ 上标 B；然后 $x_0\le b-\delta'<b$，和

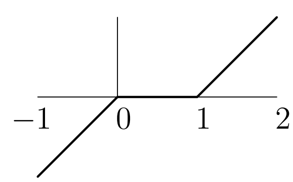
$$(*)\qquad f<0 \text{ on } [a,x] \text{ for every } x<x_0.$$

断言是 $f(x_0)=0$。为了证明这一点，我们表明 $f(x_0)\neq 0$ 蕴含 $x_0\neq$ sup B。

如果 $f(x_0)>0$，那么在某些关于 x_0 的开区间内，根据 f 的连续性，f >0，这与 $(*)$ 矛盾。另一方面，如果 $f(x_0)<0$，那么在某些关于 x_0 的开区间内，$f<0$，因此根据 $(*)$，$f<$ 在某些区间 $[a,x]$ 上为 0，与 $x>x_0$ 相关联，这意味着 x_0 不是 B 的上界。因此，剩余的可能性 $f(x_0)=0$ 必须成立，从而完成证明。 □

这个证明找到了区间 $[a,b]$ 中 $f(x)$ $=$ c 的最小解；特别是，存在一个最小解。也存在一个最大解，通过明显的修改论证得到。一般来说，不存在“第二小的”解。函数 $f\colon[-1,2]\to\mathbf{R}$ 定义为

$$f(x)=\begin{cases} x & \text{if } -1\le x<0\\ 0 & \text{if } 0\le x\le 1\\ x-1 & \text{if } 1<x\le 2\end{cases}$$



具有一个首位零和一个末位零，但没有“第二个”零。

5.4 应用程序

中间值定理可以用来证明关于数字和函数的几个有趣的事情。这里给出的采样包括

存在正实数的 n 次根，每个奇数次多项式存在实根，以及二分法，这是一个用于逼近连续函数 f 的 $f(x)\ =0$ 解的数值算法。

实数的根

设 $n\geq$^2 为一个整数。根据定义，一个数 c 的 n 次根是一个数 t (，通常与 $t^n=c$ 在同一个域) 中。我们现在已经拥有了足够的工具来证明每个正实数都有一个唯一的正 n 次根。对于负数来说，由于只考虑实数的偏见，情况更为混乱。当与复数一起工作时，情况更为满意：每个非零复数恰好有 n 个不同的（复数）n 次根。正如我们在第15章中将要看到的，这些根位于复平面上以0为中心的正 n 边形的顶点上
。

中间值定理几乎立即暗示了正实数的 n 次根的存在。技巧是构造一个其零点是所需根的函数。

定理5.9。设$c\ \in\ \mathbf{R}$为正数，设$n\ \geq 2$为一个整数。存在一个唯一的正实数t，使得$t^n=c$。

证明。唯一性是基本的：如果$0<\ a\ <\ b$，那么对于每个正整数 n，0
$<\ a^n\ <\ b^n$，所以 $a^n=c$ 对于至多一个 $a>0$ 成立。必须证明存在一个 n 次根。定义 $f:[0,\infty)\to\mathbf{R}$ 为 $f(x)=x^n-c$；我们希望证明 f 有一个正根。由于 f 是一个多项式函数，f 在 $[0,\infty)$ 上是连续的。通过直接求值，$f(0)=-c<0$，因此只需证明存在一个 $x>0$ 使得
$f(x)>0$。如果 $c>1$，则取 $x=c$：

$$f(c)=c^n-c=c\,(c^{n-1}-1)>0,$$

我们完成了。如果 $0<c<1$，那么 $c^n<1$，所以 $f(1)=1-c^n>0$，我们再次完成了。如果 $c=1$，结论是显而易见的。 □

令人欣慰的是，我们看到我们已经走了多远。除了 $\mathbf{R}$ 的构建细节外，这个推论中的所有内容都是建立在集合论和逻辑之上的。你还应该欣赏描述一个基本几何概念——单位正方形的对角线——需要多少抽象机械——用数字（更不用说集合）来表示。

课程，为了给人留下深刻印象，你必须承认学校数学中有多少是基于未经证实的断言。

通常写作 $t = c^{1/n}$ 或 $t = \sqrt[n]{c}$ 表示 n 次方根 c。在指数表示法中，规则

$$a^{x+y} = a^x \cdot a^y \quad \text{and} \quad a^{xy} = (a^x)^y$$

对于所有有理指数，都成立。到目前为止，我们只有适合整数指数的指数定义，我们还没有定义非有理或非实指数的指数。

多项式的实根

寻找多项式的根是数学中最古老的问题之一；早在4000多年前，巴比伦人就知道如何解决我们现在所说的通用二次方程。直到15世纪才找到了立方和四次方程的类似公式。19世纪初，N. H. Abel表明，不存在“五次公式”，也就是说，五次或更高次的一般多项式的根不能一般地写成系数的根式表达式。

即使对于三次和四次多项式，代数公式也很混乱，因此希望有更简单（如果不够精确）的工具来获取关于根的信息。可以使用数值方法来近似根，但如果没有知道存在实根，就很难开始。介值定理可以用来证明每个奇数次多项式（具有实系数）至少有一个实根。这是可以期望的最好结果，因为一般偶数次多项式可能没有实根，而奇数次多项式可能恰好有一个实根。

设 $p : \mathbf{R} \to \mathbf{R}$ 为一个次数为 $n \geq$ 的多项式函数 1. 在一个非常数多项式中，最高次项“主导”其他项之和，在以下意义上：

$$p(x) = \sum_{k=0}^{n} a_k\, x^k = x^n \left(\sum_{k=0}^{n} a_k\, x^{k-n} \right), \tag{5.1}$$

并且括号内的项有极限值 a_n 为 $|x| \to \infty$，因为其他所有项都趋于0（注意指数 $k - n$ 是非正的）。由于根据假设 $a_n \neq$ 为0，括号内的项具有相同的符号

如 a_n 提供 $|x|$ 足够大。从“渐近”的角度来说，非常数多项式的行为类似于其最高次项。

这个对多项式根的存在有什么影响？首先，我们不妨假设 $a_n > 0$，因为乘以 -1 并不会改变 p 的根集。如果 $a_n > 0$，那么对于足够大的 x，$p(x) > 0$；实际上，方程 (5.1) 断言 $p(x) \to +\infty$ 随着 $x \to +\infty$。然而，如果 $a_n > 0$，那么 $p(x)$ 在大负 x 时的行为取决于次数是偶数还是奇数。如果 n 是奇数，那么 $p(x) \to -\infty$ 随着 $x \to -\infty$。在这种情况下，存在实数 $x_1 \ll 0$ 和 $x_2 \gg 0$，使得 $p(x_1) < 0 < p(x_2)$。根据介值定理，存在一个 $x_0 \in (x_1, x_2)$，使得 $p(x_0) = 0$。换句话说，奇数次多项式至少有一个实根。

与之一比，如果 n 是偶数，那么 $p(x) \to +\infty$ 作为 $x \to -\infty$，因此不能保证 p 的符号改变；对于大的 $|x|$，$p(x)$ 的符号与 a_n 的符号相同，无法得出符号改变的结论。这说明我们用于奇数次多项式的论点失败了，但可能存在一个“更好的”证明。为了彻底解决这个问题，观察多项式 $p(x) = 1 + x^2$ 的次数是偶数，并且没有实根，因为域 $\mathbf{R}$ 是有序的，有序域不能包含 -1 的平方根。总之，奇数次多项式至少有一个实根（并且通常没有更多，参见练习5.8），而一般偶数次多项式没有实根。

二分法

中间值定理是寻找连续函数根的数值方法——二分法——的基础。这种方法仅适用于值容易计算的函数（如多项式）。为了了解该方法的工作原理，我们将用它来近似 {v*}^2。考虑定义域为 [1{v*}^2] 的函数 {v*}^2。通过直接计算，我们发现 {v*}^1{v*}^1 {v*}^0 和 {v*}^2 {v*}^2 {v*}^0。中间值定理表明，在开区间 (1{v*}^2) 中存在一个实数零点。现在我们进行二分；中点是 3{v*}^2，我们发现 {v*}^3{v*}^2{v*}^0 9{v*}^4 {v*}^2 {v*}^1{v*}^4 {v*}^0。由于 {v*}^1{v*}^0 {v*}^3{v*}^2{v*}^0，中间值定理断言 {v*} 在开区间 (1{v*}^2 3{v*}^2) 中有一个零点。我们再次进行二分；中点是 5{v*}^4，{v*}^5{v*}^4{v*}^0 25{v*}^16{v*}^2 {v*}^0，因此我们得出结论，{v*} 在区间 (5{v*}^4 3{v*}^2) 中有一个零点，或者用小数表示，(1{v*}^25 1{v*}^5)。搜索模式应该是清晰的；你应该自己执行另一个步骤，以确保你理解。

这里是一般设置和程序的算法形式，适用于编写计算机程序。假设f是$[a, b]$上的连续函数，并且$f(a)$和$f(b)$具有相反的符号。根据介值定理，区间(a, b)中存在一个根。在区间中点$(a+b)$处评估$f/2$；如果值为零，则停止。否则，在半区间的函数值之间有一个符号差异，这表明该半区间中存在一个根。重复直到获得所需的精度。这是一个简单的算法，但精度只能每次迭代翻倍，因此需要大约三到四次迭代才能获得额外的每一位小数精度。其他算法可以给出更高的精度；回想一下，根据练习4.23，示例4.47中的序列以这种方式收敛到$\sqrt{b}$，即每次迭代精度的小数位数翻倍。通过选择良好的起始值，后一种方案的六次迭代可以给出31位精度，这需要大约100次二分法。十次迭代将给出近500位精度，需要超过1600次二分法。然而，二分法有其用途，尽管在实际计算中通常有更好的方法。

练习

练习5.1 证明如果 $f : \mathbf{R} \to \mathbf{R}$ 是连续的并且 ℓ-周期的，那么 f 在 $\mathbf{R}$ 上是一致连续的。$\diamond$

练习5.2 我们看到，将符号函数限制在 $\mathbf{R}^\times$ 上是局部常数的，但不是一致连续的。证明如果 $a > 0$，那么限制到 $\mathbf{R} \setminus [0, a]$ 的是一致连续的。注意移除一个点和移除任意小区间之间的区别。$\diamond$

练习5.3 找到一个有界、连续的函数 $f : (0, 1) \to \mathbf{R}$，使得 f 不是一致连续的。
建议：令 g 为一个非常数、连续、周期函数，并设 $f(x) = g(1/x)$。$\diamond$

练习5.4 设$p : \mathbf{R} \to \mathbf{R}$是一个至少2次的多元多项式函数。证明$p$不是一致连续的。$\diamond$

练习5.5 设$I \subset \mathbf{R}$为一个开区间（可能是不界的），并且设$f : I \to \mathbf{R}$是有界、连续且递增的。证明f是一致连续的。$\diamond$

练习5.6 设$I \subset \mathbf{R}$为一个区间，且$f : I \to \mathbf{R}$是递增的。证明f^{-1}是连续的。
提示：画一个草图，并询问 f^{-1} 连续意味着什么。$\diamond$

练习5.7 给出一个连续且递增的函数，其逆函数是不连续的例子。

提示：根据前面的练习，定义域不能是一个区间！ $\diamond$

练习5.8 证明多项式$p(x) = x^5 + x - 1$有一个唯一的实根，并使用二分法（以及计算器！）将这个根近似到小数点后两位。如果你的计算器是可编程的，那么你应该找到你的计算器允许的小数位数。无论如何，编写一个算法流程图以获得逐次更好的近似值是一个很好的练习。$\diamond$

练习5.9 设$p : \mathbf{R} \to \mathbf{R}$为一个单次且次数为$\geq 2$的偶数次多项式函数。

(a) 证明 $\lim(p, \pm\infty) = +\infty$.

(b) 证明 p 在 $\mathbf{R}$ 上有绝对最小值，即存在一个实数 x_0，使得 $p(x_0) \leq p(x)$ 对于所有实数 x 成立。

请注意，您不能立即在(b)部分应用极值定理；然而，通过仔细利用(a)部分的信息，您可以把寻找最小值的过程减少到一个闭区间。$\diamond$

练习5.10 定义 $f : (0, +\infty) \to \mathbf{R}$ 通过 $f(x) = 5x^7 - 2x^4 - x^2 + 1 \; x^7 + 2$ $x^5 + 1$ 。
(a) 证明存在一个 $x_0 \in (0, +\infty)$ 满足 $f(x_0) = 1$.

(b) 证明 f 有一个绝对最小值。

提示：如果你不使用本章的工具尝试完成任何一部分，你将会遇到麻烦。$\diamond$

练习5.11 本练习涉及连续函数的定点存在性。

(a) 设 $f : [a, b] \to [a, b]$ 为一个连续函数。证明存在一个 $x_0 \in [a, b]$ 满足 $f(x_0) = x_0$。换句话说，每个将闭区间映射到自身的映射至少有一个不动点。

建议：考虑由 $g(x) = f(x) - x$ 定义的函数 g。

(b) 如果 f 将某个开区间映射到自身，是否得出相同的结论？如果 f : $X \to X$，但 X 不是一个区间呢?

说明部分（b）的答案。 $\diamond$

练习5.12 假设 $f: X \to \mathbf{R}$ 在 X 上是一致连续的，并且让 $(x_n)_{n=0}^{\infty}$ 是 X 中的一个柯西序列。证明像序列 $\big(f(x_n)\big)$ 是柯西序列。$\diamond$

练习5.13 假设 $f: X \to \mathbf{R}$ 在 X 上仅仅是连续的，并且让 $(x_n)_{n=0}^{\infty}$ 是 X 中的一个柯西序列。你能推断出像序列 $\big(f(x_n)\big)$ 是柯西序列吗?

建议：首先尝试扩展你为前一个练习找到的证明。如果你不能推广这个证明，尝试找出不工作的步骤，并使用这些信息去寻找反例。这个问题的最困难的部分是决定是否存在一个证明或反例！$\diamond$

练习5.14 设 $f: (a, b) \to \mathbf{R}$ 是一致连续的。

(a) 证明存在 f 到 $[a, b]$ 的连续扩展，即存在一个连续函数 $F: [a, b] \to \mathbf{R}$ 使得 $F|_{(a,b)} = f$。提示：使用练习 5.12 和定理 4.45 定义 $F(a)$ 和 $F(b)$。

(b) 逆命题是否成立？（“如果 f 不是一致连续的，那么不存在一个连续扩展。”）给出证明或反例。

(©) 证明在(a)部分中找到的扩展是唯一的，即如果F_1和F_2是f到$[a, b]$的连续扩展，那么$F_1 = F_2$。

这是一个分析中一般原理的实例：一个一致连续的函数在其定义域的极限点集上有唯一的连续扩展。$\diamond$

练习5.15 设$I \subset \mathbf{R}$为一个区间。如果存在一个实数M，使得函数$f: I \to \mathbf{R}$是Lipschitz连续的，则称其为Lipschitz连续

$$|f(x) - f(y)| \le M\,|x - y| \qquad \text{for all } x, y \in I.$$

(a) 证明如果 f 是Lipschitz连续的，那么 f 在 I 上是均匀连续的。

(b) 证明 $\sqrt{\cdot}: [0, \infty) \to$ **R** 是一致连续的，但不是Lipschitz连续的。提示：问题发生在接近0的地方。

(©) 证明 f 是Lipschitz连续的当且仅当

$$f(x+h) = f(x) + O(h) \quad \text{near } h = 0 \text{ for all } x \in I.$$

给出Lipschitz连续性的几何解释。 ⋄

练习5.16设cb为查理·布朗函数，且对于$x \in (0, 1)$，有$f(x) =$ = cb(1/x)。

(a) 绘制 f 的图形。证明对于每个 $\delta > 0$，f 将区间 (0, δ) 映射到区间 [0, 1]。 (b) 定义 $g : (0, 1) \to$ **R** 为 $g(x) = (1/x)f(x)$。证明对于每个 $\delta > 0$，g 将区间 (0, δ) 映射到区间 [0, ∞)。 (c) 找到一个连续函数 h : (0, 1] → **R**，其像为开区间 (−1, 1)。是否存在一个连续函数 H : [0, 1] → **R** 满足 $h = H|_{(0,1]}$? 提示：构造 h 并不像看起来那么容易，但与这个练习相关。

查看练习5.3，如果您还没有的话。 ⋄

练习5.17 存在双射函数 $f : [0, 1] \to (0, 1)$ 吗？存在连续双射函数 $g : [0, 1] \to (0, 1)$ 吗？像往常一样，根据需要给出证明或反例。⋄

第6章

什么是微积分?

微积分是研究变化率的数学。这个名字指的是计算程序——微分、无穷小和积分——而不是我们主要关心的理论基础。微积分自然分为两半：微分——研究变化率，积分——研究总变化。如果 f 是一个“合适的”函数，那么询问“当 x 变化时 $f(x)$ 变化的速度有多快”是有意义的。使这个想法精确的非平凡部分是确定哪些函数是“合适的”，并定义在单一点上的变化率意味着什么。相反，如果知道 f 在某个区间的每个点的变化率，人们可能希望确定 f 在区间内的总变化。直观上，人们希望“加总”瞬时变化率以得到总变化。

微积分运算的定义源于简单的想法，但其中隐藏了许多理论和实践上的复杂性。本章的目的是通过一些直观的论据来激发即将到来的理论。

6.1 变化率

考虑沿着一条直线高速公路的汽车旅行。汽车的位置 {v*} 是时间 {v*} 的函数，比如说 {v*}。为了明确起见，我们同意旅行从 {v*} 0 开始，并且 {v*}0{v*} 0。在时间 {v*} 的 {v*} 值是里程表读数（因为我们出发时将里程表归零了）。里程表读数作为时间的函数的图表是旅行的数学理想化。

假设我们想用数学术语描述我们的速度，只使用里程表读数和秒表。速度被定义为位置随时间变化的速率，因此我们在两个不同时间观察里程表，t 和 $t + \Delta t$，并计算

$$\text{Average speed over } [t, t + \Delta t] = \frac{X(t + \Delta t) - X(t)}{\Delta t}. \tag{6.1}$$

实验上，我们认为我们的测量“从 t 开始，持续 Δt”。你可能已经进行了这个实验：一些高速公路上有“测量里程”——每隔一英里放置的标志。如果你以恒定速度（通常是每小时40或60英里）驾驶，并计算通过标志所需的时间，你可以校准你的速度表。

如果 Δt 很大，那么如果我们的速度不恒定，(6.1) 可能相当不准确。一对测量只能产生平均变化率，而小于 Δt 的时间尺度上的波动往往会“平均掉”。为了解决更短的时间间隔，我们使 Δt 更小。如果我们想象 dt 代表一个“无穷小”的时间间隔，方程 (6.1) 变为

$$\text{Speed at time } t = \frac{X(t + dt) - X(t)}{dt}.$$

几何上，我们关注的是越来越小的矩形内X的图形，即“放大”。说变化率存在，相当于说放大使得X的图形越来越像一条线。这种讨论适用于任何作为其他量函数变化的量。

对于某些现实世界现象，例如汽车或行星的速度、生态系统中物种的种群、电路中的电压、试管中化学物质的浓度，或飞机机翼前缘的空气压力，一个量的变化率在某种意义上是“良好行为”的，即当输入变量的增量变得较小时，平均变化率有一个极限。从数学上讲，我们说这些量的函数是可微的；它们具有定义良好、有限的速率变化。粗略地说，可微函数的图像由无限多个无穷小线段组成。（这个断言是歌德引言的极端例子，但足够真实，足以作为启发。）

经典物理学（尤其是力学和电磁学）几乎完全涉及可微函数。可以通过频闪摄影精确测量子弹的速度

在两个紧密分离的时间点对其进行测量，并测量其在其间行进的距离。如果已知行星在几个时间点的位置，则可以非常准确地计算出其位置，甚至可以预测几个月或几年后的位置。微积分和统计分析的早期惊人成功之一发生在19世纪30年代，当时天文学家发现了小行星谷神星，然后又在太阳的眩光中失去了它。基于对其过去位置的测量，C. F. 高斯准确地预测了谷神星在最后一次被观测到几个月后可能的位置。

可微函数的局限性

现在广泛认为，许多现象不能用可微函数很好地建模；例如，股票价格、互联网流量、地球沿地震断层运动、山脉和云的形状，以及一杯水中单个分子的位置。这些现象仍然由连续函数建模，但缩小时间或空间尺度并不能提高变化率的测量精度。相反，随着增量减小，差商以复杂的方式变化，导致变化率的不稳定数值。从几何上看，放大图形并不能“平滑”出变化。如果所测量的量具有大范围的、小规模的波动，则在测量变化率时存在权衡：增量必须足够大以平滑“噪声”，但又足够小，以免平均掉想要测量的变化率。

这个讨论略微过于简化。例如，地球（人类）人口在分钟的时间尺度上变化是混沌的，但在年的尺度上却非常规律。同样，单个汽车事故或家庭火灾很难预测，但保险公司可以非常准确地预测每个月会发生多少事故和火灾，以及总索赔金额是多少。单个水分子运动是混沌的，但一杯水在我们眼中看起来却是平滑和均匀的。（通过仔细加入一滴食用色素可以揭示小规模的复杂性。如果水的行为像理想化的模型，颜色会立即在整个杯子中扩散。）所有数学模型都有一个“特征尺度”，在此尺度之外，模型无法很好地工作。经典物理学的现象在它们适用的时空尺度范围内是显著的。

6.2 总变化

返回我们的汽车之旅，假设汽车的速度 {v*} 是时间的函数，比如说 $s = S(t)$。S 的图表表示速度表读数。在 t 和 $t + dt$ 之间的无穷小时间间隔内行驶的距离是 $S(t)\,dt$，总行驶距离是通过“累加”这些无穷小距离，从 0 到 t_0 的时间范围内得到的。

形式代数计算令人信服：瞬时速度满足

$$S(t) = \frac{X(t + dt) - X(t)}{dt} = \frac{dX}{dt}, \quad \text{or } S(t)\,dt = X(t+dt) - X(t) = dX.$$

再次，dt 应该是“无穷小”，大于零但小于每一个正实数。当我们对所有 $t \in [0, t_0]$ 的这些项“求和”时，我们发现，令我们非常满意的是，X 的增量构成一个形式上的望远镜求和，参见练习4.29，而这个“和”是 $X(t_0) - X(0)$，即汽车行驶的距离！

6.3 符号和无穷小

我们肯定是在一些有趣的数学的轨道上。然而，对于这个论点，人们至少应该有一些令人烦恼的怀疑，因为 dt 不能被视为一个实数，因为它违反了三分法属性。正是关于这个问题产生了大量的争议：“什么是无穷小 dt？”微积分的标准处理通过将 dt 降级为一种方便的记账工具来回避这个问题。然而，强调谨慎地将 dt 作为代数量使用——就像上面的论点一样——既迅速地“证明”了许多微积分的基本定理，又阐明了这些定理的陈述意义。术语“微积分”本身指的就是正确操作无穷小的计算程序——以一种不与普通代数相矛盾的方式。

在最终分析中，逻辑一致性而非直观可能性是判断一个数学思想的根本标准。证明逻辑一致性的方法之一是定义所有考虑的对象，并以已知（或假定）一致的形式公理来表述所有的断言。我们已经将函数的连续性概念降低到实数的公理，而这反过来

在公理的基础上构建了有理数，这些公理又建立在自然数的算术之上，最终用集合来定义。假设集合论是一致的。为了证明美丽的“验证”即位置可以通过速度恢复（或者更一般地说，通过累加无穷小增量可以恢复一个量的总变化），我们必须执行以下之一:

- 定义无穷小量以实数的形式，并验证它们满足适当的操作形式规则，例如有序域公理。这是非标准分析的方法。

- 使用实数公理来定义某些出现无穷小量的表达式（例如商和适当的无限和），并验证这些表达式可以像它们包含实际无穷小量一样进行操作，而实际上从未使用过包含“裸露”无穷小量的表达式。例如，如果无穷小量的商在实数系统中定义，我们需要证明如下规则:

$$\frac{dy}{dx}+\frac{dz}{dx}=\frac{dy+dz}{dx}\quad\text{and}\quad\frac{dz}{dy}\frac{dy}{dx}=\frac{dz}{dx}.$$

第一种方法是一个令人惊讶的困难技术任务，需要向集合论本身添加一个新公理。此外，可以证明没有任何收获，即每个“非标准”定理都可以通过“标准”方法证明。（尽管如此，可以很好地论证非标准分析更直观，因此使用非标准技术比不使用它们更有可能找到定理。）

看似复杂的第二个程序实际上是我们要走的路径，这解释了论点往往具有的间接味道。我们需要的绝大多数技术工具——主要是极限的概念及其操作手段——已经开发出来，因此在这个阶段，间接路径无疑是更经济的。无穷小量的符号称为莱布尼茨符号，以戈特弗里德·威廉·冯·莱布尼茨命名。莱布尼茨符号中的每个表达式都可以写成一种不那么挑衅性的（即非无穷小量）形式，这种形式以艾萨克·牛顿爵士命名，称为牛顿符号。例如，牛顿符号将 $S(t) = X'(t)$ 写作上述汽车的速率，而莱布尼茨符号的拥护者会写成 $s = dx/dt$。数学家倾向于更喜欢牛顿符号，它不太容易滥用，

当科学家倾向于偏好莱布尼茨符号，因为它允许他们将实际问题转化为符号形式，并以相对容易的方式解决它们时，你应该努力掌握两种语言的流畅性，因为它们的优点是互补的。

行动计划

积分和微分的定义可能看起来有点复杂，但只是对上述讨论的思想的正式化。直接使用定义进行操作通常很困难，但有一些定理可以简化对一大类函数的导数和积分的计算。

下一章探讨了这些想法的主要方面，以下为结果：

- 无限多个瞬时变化率可以相加，得到一个确定的数值结果，表示总变化。
- 变化率可以像无穷小量的比率一样被操纵。
- 变化率和总变化量的运算本质上互为逆运算。

集成和微分也可以被视为从已知函数创建新函数的操作。我们将使用微积分运算来构造对数、指数函数和圆三角函数。导数和积分之间的逆关系将使我们能够确定这样定义的函数的性质。

第7章

集成

我们以积分运算开始对微积分的研究。直观的想法是从一个定义在闭区间 $[a, b]$ 上的函数 f 开始，并对每个 $t \in [a, b]$ 的无穷小项 $f(t)\,dt$ “求和”。然而，这个目标目前是没有意义的，因为我们不知道“dt”代表什么，也不知道如何将无限多个量相加。我们希望捕捉的量有一个很好的几何解释。首先考虑一个非负函数 f，并想象将定义域 $[a, b]$ 划分为大量（但有限）长度相等的子区间。按照图7.1中的方式构造矩形。这些矩形的面积之和“近似”于我们想要定义的量。

面积之和是一个精确的近似，因为每个矩形都有正的（实数）宽度，而不是“无穷小”的宽度。为了得到“更好的”近似，我们应该使宽度更小，即把区间细分为更多的子区间，见图7.2。

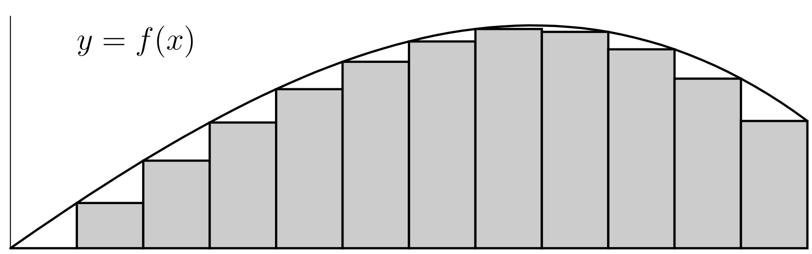


图7.1：近似求和$f(t)\,dt$的矩形面积，对于$t \in [a, b]$。

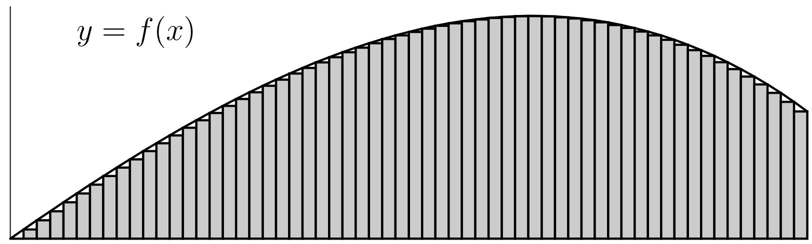


图7.2: 更多矩形给出更好的近似 近似

当然，无论我们取多少个矩形，得到的总和可能都不会恰好是我们想要定义的数量。正是在这里，**R**的完备性公理拯救了局面；一旦事情被正确设置，就会很容易看出我们的每一个总和都小于一个固定的实数。根据完备性公理，总和的集合有一个实数上确界。如果我们幸运的话，这个上确界将正好是我们想要的量！

这是重新考虑歌德名言的另一个好时机。取更窄的矩形可以得到更好的近似，所以在“极限”情况下（取上确界），我们是在“无限多个无限薄的矩形的面积之和”。注意，如果我们能使这些想法生效，那么在某种意义上，我们将给形式为$0 \cdot \infty$的表达式分配了一个实数。

应该清楚，刚刚概述的过程与“f的图像下的面积”有关。这部分的偶然性在于，我们假设f是非负的，并且因为图7.1和7.2描绘了这些想法起作用的情况。碰巧的是，如果f是连续的，那么刚刚概述的过程，即“积分”，在真正代表“面积”的意义上运作良好。尽管如此，正如我们将看到的，使用定义来计算特定的积分相当困难。幸运的是，“将无限多个无穷小矩形的面积相加”的想法被两个显著的事实所挽救：

- 可以推导出许多积分的性质，而无需计算单个积分。

- 使用积分的抽象性质，我们发现积分与“微分”操作密切相关，这更有利于计算。

这些项目是本章及以下章节的主题，在结尾时，我们将拥有一套强大的工具来描述和解决涉及变化率的广泛问题。

您应该记住，本章基本上没有计算结果。它是通过一系列谨慎的定义和事后观察的长链，最终将积分锻造成为一种有用的计算工具。现在暂且搁置不切实际的反对意见。积分在概念上并不困难，但如果您担心定义在实际中的应用，可能会显得令人畏惧。请放心，您很快就会拥有强大而灵活的计算技巧。

7.1 分区与和

集成以有界函数 $f\colon [a, b] \to \mathbf{R}$ 为输入，并返回一个实数。正如前述章节所述，我们将通过精确定义我们希望定义的量的上下界，然后声明当这些界限重合时，该量存在，间接地追求这个目标。

让 $I = [a, b]$ 是一个闭区间，有界。I 的一个划分是一个有限点的集合 $P = \{t_i\}_{i=0}^{n}$，使得

$$a = t_0 < t_1 < t_2 < \cdots < t_n = b.$$

$I_i = [t_{i-1}, t_i]$ 的区间被称为划分的第 i 个子区间，其长度为

$$\Delta t_i = t_i - t_{i-1}.$$

我们不假设所有子区间长度相同。分割 P 的网格是最大的 Δt_i，即最长子区间的长度。P 的网格表示为 mesh(P) 或 $\|P\|$。如果 P 和 Q 是 I 的分割，并且如果 $P \subset Q$，那么 Q 被称为 P 的细化。换句话说，通过向 P 中添加有限多个点来获得 P 的细化。注意，如果 $P \subset Q$，则 mesh(Q) $\leq$ mesh(P)；你不能通过添加点来增加网格！

现在令 $f\colon I \to \mathbf{R}$ 为一个有界函数。（如果 f 在 I 上连续，那么根据极值定理，f 是有界的；然而，在这个阶段，函数 f 可能处处不连续。）给定 I 的一个分割 P，我们在 f 上取“最佳”的上界和下界。

每个子区间:

$$m_i := \inf\{f(t) \mid t \in I_i\}$$
$$M_i := \sup\{f(t) \mid t \in I_i\} \tag{7.1}$$

对于 $i = 1, \ldots, n$。直观上，m_i 是 f 在第 i 个子区间上的“最小值”，而 M_i 是“最大值”，但 f 在 I_i 上可能既没有最小值也没有最大值，因为我们知道 inf 和 sup 存在，因为 f 是有界的。利用 f 的这些下界和上界，我们通过以下方式在划分 P 上形成 f 的下和与上和:

$$L(f, P) = \sum_{i=1}^{n} m_i \,\Delta t_i, \qquad U(f, P) = \sum_{i=1}^{n} M_i \,\Delta t_i. \tag{7.2}$$

在引言中，我们只考虑了下和；由于即将明显的技术原因，我们必须考虑上界和下界。

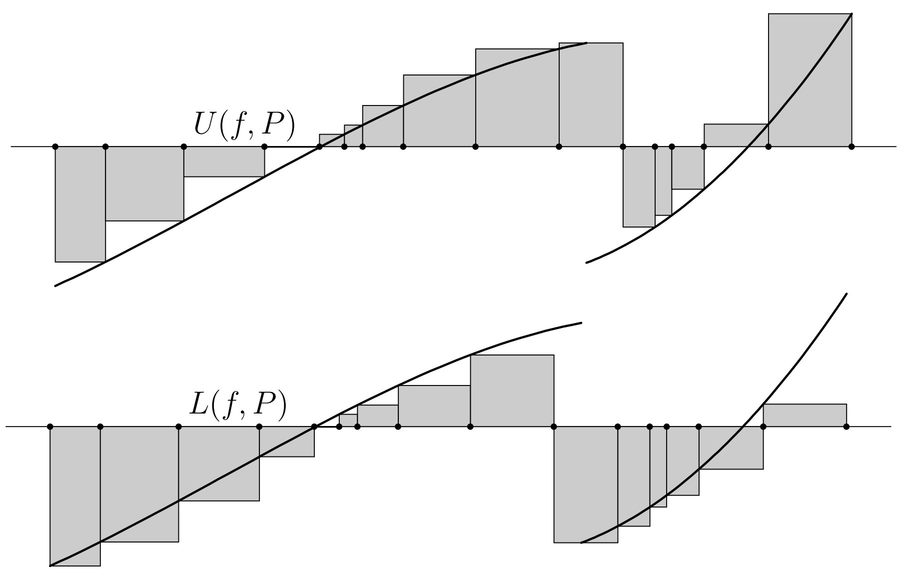


图7.3：与一个划分相关的f的上和（顶部）与下和。

因为对于每个P的子区间$m_i \le M_i$，显然对于每个f和每个P，有$L(f, P) \le U(f, P)$。进一步，对于每个i，$m_i \le f(t) \le M_i$对所有$t \in I_i$成立，因此$m_i \Delta t_i$是“$f(t)\,dt$在$t \in I_i$上的和”的合理下界，同样地，$M_i \Delta t_i$是合理上界。

上和下和可以看作是矩形面积的求和，前提是 f 非负。一般来说，下和是位于水平轴上方的矩形面积之和减去轴下方的面积之和，上和同理，见图7.3。细化分割 P 只能改善界限。我们将这个有用的观察正式化如下：

引理7.1. 设$f : [a, b] \to \mathbf{R}$为一个有界函数，并且设$P$和$Q$是具有$P \subset Q$的$[a, b]$的划分。那么

$$L(f, P) \le L(f, Q) \le U(f, Q) \le U(f, P).$$

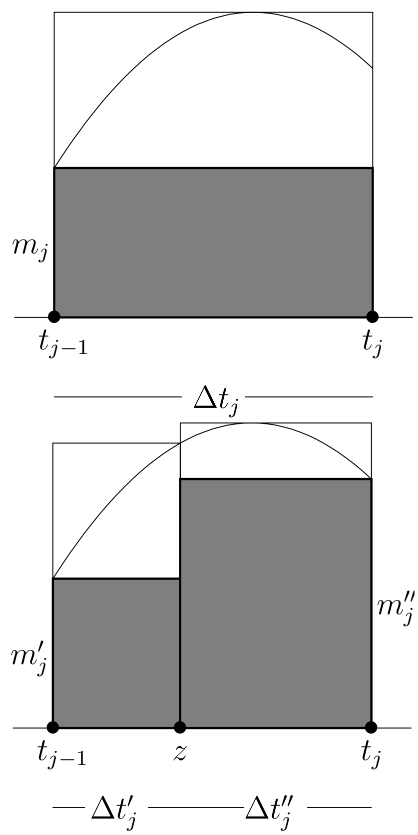


图7.4：分区细化——前后对比。

证明。由于每个细化都是通过将有限个点添加到 P 中得到的，因此对 Q 中额外点的数量进行归纳

将索赔降低到$Q = P \cup \{z\}$比P恰好多一个点的情况。为了明确起见，假设$z \in I_j$；子划分$I_j = [t_{j-1}, z] \cup [z, t_j]$将$m_j \Delta t_j$在$L(f, P)$中的项分割为

$$m'_j \Delta t'_j + m''_j \Delta t''_j, \tag{7.3}$$

查看图7.4。但m'_j是f在$[t_{j-1}, z] \subset I_j$上的下确界，这肯定不小于m_j，同样$m''_j \geq m_j$。因此，(7.3)中两个项的和大于或等于$m_j \Delta t_j$。这从图7.4中几何上很清楚。

由于否则 $L(f, P)$ 和 $L(f, Q)$ 是相同的，这个论点证明了 $L(f, P) \leq L(f, Q)$。一个完全类似的论点表明 $U(f, Q) \leq U(f, P)$。 □

命题7.2. 设 $f: I \to \mathbf{R}$ 是一个有界函数，并且设 P 和 P' 是 I 的任意划分。那么

$$L(f, P) \leq U(f, P').$$

每个下和都小于或等于每个上和。

证明。划分 $Q = P \cup P'$ 是 P 和 P' 的细化。根据引理7.1，$L(f, P) \leq L(f, Q)$ 和 $U(f, Q) \leq U(f, P')$。 □

给定一个有界函数 $f: I \to \mathbf{R}$，其定义域是一个闭区间，我们对其所有划分取下和与上和，命题7.2表明下和集有上界；任何特定的上和都是一个上界！因此，下和集有一个实上确界 $\mathbf{L}(f, I)$，称为 f 在 I 上的下积分。类似地，上和集有下界，因此有一个下确界 $\mathbf{U}(f, I)$，称为 f 在 I 上的上积分。命题7.2暗示了 $\mathbf{L}(f, I) \leq \mathbf{U}(f, I)$，幸运的是这与我们的直觉相符：我们试图定义的量——f 的积分——肯定不小于下积分，也不大于上积分。

定义7.3 设I为一个闭、有界区间。如果f: $I \to \mathbf{R}$在I上可积，则称$\mathbf{L}(f, I) = \mathbf{U}(f, I)$。在这种情况下，这个公共值被称为$f$在$I$上的积分。

积分 f 关于 I 表示为 $\int_I f$，或者当 $I = [a, b]$ 时表示为 $\int_a^b f$。直观上看，上下积分似乎相等，尽管证明并不明显，上下积分

积分确实在 f 连续时相等。然而，正如我们将通过例子所看到的，f 的下积分通常严格小于上积分。在这种情况下，上下积分不指定一个唯一实数，我们说 f 在 I 上不可积。

积分的定义依赖于一个数集的上确界与另一个数集的下确界相等。为了证明定理，通常更方便使用以下标准，它使用一个划分而不是所有划分的集合。

命题7.4. 有界函数 $f : [a, b] \to \mathbf{R}$ 在 $[a, b]$ 上可积，当且仅当以下条件成立：
对于每个 $\varepsilon > 0$，存在一个 $[a, b]$ 的划分 Q，使得

$$U(f, Q) - L(f, Q) < \varepsilon.$$

证明。假设 f 是可积的。固定 $\varepsilon > 0$，并选择分割 P 和 P' 使得

$$\left(\int_a^b f\right) - L(f, P) < \frac{\varepsilon}{2}, \qquad U(f, P') - \left(\int_a^b f\right) < \frac{\varepsilon}{2}.$$

这样的划分根据上确界和下确界的定义存在。正如命题7.2的证明中那样，取$Q = P \cup P'$，并得出$U(f, Q) - L(f, Q) < \varepsilon$。

相反，假设 f 不可积。令 $2\varepsilon = \mathbf{U}(f, I) - \mathbf{L}(f, I)$。那么
$U(f, Q) - L(f, Q) > \varepsilon >$ 对于每个划分 Q 都大于 0。 □

7.2 基本示例

函数 f 的积分只依赖于积分区间；因此，写出 $\int_I f$ 是有意义的。然而，特定的函数通常由公式给出，如 $f(t) = t^2$，并且写出 $\int_I t^2$ 会很方便。问题是，“t^2” 这个表达式除非我们同意 t^2 是 f 在 t 处的值，否则不定义一个函数，而坚持这样的惯例要求过多，这一点将变得明显。所需的是一个“占位符”来表示 t 是被积函数中的“变量”。在这种情况下，标准记法是写“$\int_I t^2\,dt$”，其中“dt”表示 t^2 是被积函数在 t 处的值。这种特殊的记法将在下面更详细地讨论，但如果你是字面意思，这里的解释就足够了。在

表达式 $\int_I t^2\,dt$，t 是一个虚拟变量，并且（就像极限一样）可以用任何其他方便的符号替换，而不会改变表达式的含义。

有教育意义的是看到可积性的定义是如何独立工作的。这里给出了几个例子来说明这个定义如何捕捉几个简单情况下的面积概念，并展示一个函数如何无法可积。

示例7.5设$c \in \mathbf{R}$，设C：$[a, b] \to \mathbf{R}$表示相应的常值函数。对于$[a, b]$的每一个划分以及每一个子区间，$m_i = c = M_i$。因此，每一个下和以及每一个上和都等于$c(b-a)$，所以

$$\int_a^b C = \int_a^b c\,dt = c(b-a).$$

观察当 $c >= 0$ 时，积分的值是 t-轴、C 的图像以及直线 $t = a$ 和 $t = b$ 所围成的矩形的面积。当 $c <= 0$ 时，积分是矩形的面积的相反数。

积分恒等函数 f 的计算相对痛苦，从定义出发。然而，初等几何暗示了答案应该是什么，因此我们可以对我们的结果进行合理性检查。

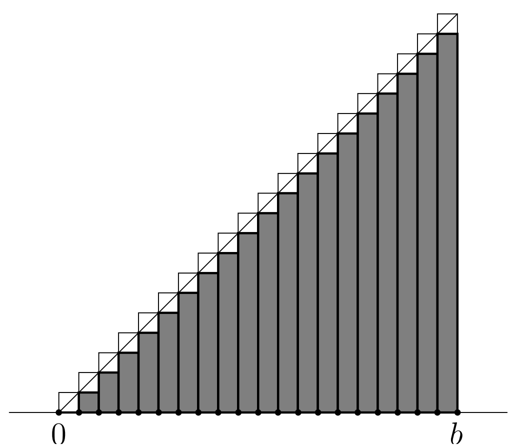


图7.5：恒等函数的上下和，$n = 20$。

考虑一个形式为$[0, b]$的区间以简化问题，并设$P_n = \{t_i\}$为具有($n+1$)等间距点的划分：$t_i = ib/n$，和$\Delta t_i = b/n$对于所有i。f在$[t_{i-1}, t_i]$上的下确界和上确界

t_{i-1} 和 t_i 分别是，因此下和上和分别是（注意求和的极限）

$$L(f,P_n)=\sum_{i=1}^{n} t_{i-1}\,\Delta t_i=\sum_{i=0}^{n-1}\frac{ib}{n}\,\frac{b}{n}=\frac{b^2}{n^2}\sum_{i=0}^{n-1} i,$$

$$U(f,P_n)=\sum_{i=1}^{n} t_i\,\Delta t_i=\sum_{i=1}^{n}\frac{ib}{n}\,\frac{b}{n}=\frac{b^2}{n^2}\sum_{i=1}^{n} i.$$

这些和已在练习2.6中计算：

$$\sum_{i=1}^{n} i=\frac{(n+1)n}{2},\qquad \sum_{i=0}^{n-1} i=\frac{n(n-1)}{2},$$

所以

$$L(f,P_n)=\frac{b^2}{2}\left(\frac{n-1}{n}\right),\qquad U(f,P_n)=\frac{b^2}{2}\left(\frac{n+1}{n}\right).$$

从这些公式中的第一个可以看出，下和（针对这个特定的分割族）的上确界等于$b^2/2$；因此，实际的下积分至少为$b^2/2$。同样，从第二个公式可以得出，上和（再次针对这个分割族）的下确界为$b^2/2$，因此实际的积分上界不超过$b^2/2$。用符号表示，

$$\frac{b^2}{2}\le \mathbf{L}(f,I)\le \mathbf{U}(f,I)\le \frac{b^2}{2}.$$

但是这意味着恒等函数是可积的，并且对$[0,b]$的积分等于$b^2/2$，正如从积分的几何解释所预期的那样。□

示例7.6设$f=\chi_{\mathbf{Q}}$为$\mathbf{Q}$的特征函数。那么f在区间$[a,b]$上不可积，无论如何选择a和$b>a$。实际上，设P为$[a,b]$的一个划分。在每个子区间中，都存在有理数和无理数，因此f在每个子区间中取0和1的值。但这意味着对于每个i，有$m_i=0$和$M_i=1$，所以

$$L(f,P)=\sum_{i=1}^{n} 0\,\Delta t_i=0,\qquad U(f,P)=\sum_{i=1}^{n} 1\,\Delta t_i=(b-a),$$

无论 P。因此，下限积分为0，而上限积分为 $b - a > 0$。由于这些不相等，f 不可积。□

示例7.7 最后但同样重要的一个例子将说明在牛顿之前如何计算曲线下的面积。这既为我们提供了一个示例库，又强调了直接使用定义的难度。（尽管如此，结果的价值证明了付出的努力是值得的。）假设 $0 < a < b$，并令 k 为一个正整数。考虑定义在 $[a, b]$ 上的单项式函数 $f(t) = t^k$。我们希望计算积分 $\int_a^b t^k\,dt$。我们不是使用所有子区间长度相等的“算术”划分，而是使用“几何”划分，其中连续区间的长度比是相同的，如图7.6所示。其理由是，由于被积函数是幂函数，连续矩形的面积将呈几何级数，因此我们可以使用有限几何级数公式来计算下和与上和。

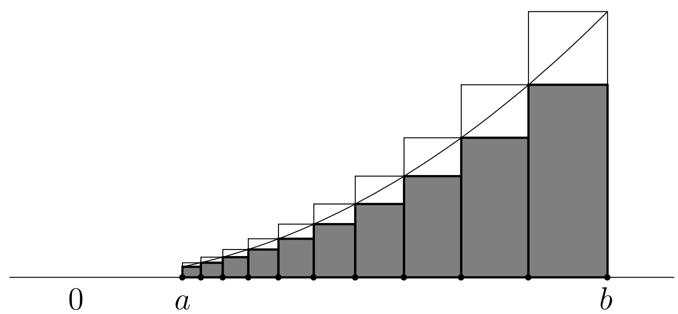


图7.6：与几何划分相关的下和与上和。

设 $n > 0$ 为子区间的数量，并将 $\rho = \sqrt[n]{b/a} > 1$，以便 $b = a\rho^n$。该分割为 $P = \{t_i\}_{i=0}^n$，具有 $t_i = a\rho^i$。被积函数是递增的，因此区间 [t_{i-1}, t_i] 上的极值在端点处取得：

$$m_i = (a\rho^{i-1})^k = a^k(\rho^k)^{i-1}, \quad M_i = (a\rho^i)^k \qquad \text{for } i = 1, \dots, n.$$

因为对所有 i，$M_i = \rho^k m_i$ 的上和是下和的 ρ^k 倍。我们将计算上和，这从形式上稍微简单一些。通项是

$$f(t_i)\,\Delta t_i = (a\rho^i)^k\,a(\rho - 1)\rho^{i-1} = a^{k+1}\,\frac{\rho - 1}{\rho}(\rho^{k+1})^i,$$

所以几何级数公式$\sum_{i=1}^{n} r^i = r \; r^n - 1 \; r - 1$意味着

$$U(f,P) = \sum_{i=1}^{n} f(t_i)\Delta t_i = a^{k+1}\,\frac{\rho-1}{\rho}\cdot\rho^{k+1}\cdot\frac{(\rho^{k+1})^n-1}{\rho^{k+1}-1}$$

自 $\rho^n = b/a$ 以来，我们已有 $a^{k+1}\big((\rho^{k+1})^n - 1\big) = b^{k+1} - a^{k+1}$，因此

$$U(f,P) = \rho^k\,(b^{k+1}-a^{k+1})\frac{\rho-1}{\rho^{k+1}-1}.$$

分数的倒数本身就是一个几何级数，$S(\rho) := \sum_{i=0}^{k} \rho^i$。现在，当$n \to +\infty$时，比率$\rho = (b/a)^{1/n}$趋近于1。因为$S$是$\rho$的多项式，我们有lim$(S, 1) = \sum_{i=0}^{k} 1^i = k+1$。因此，

$$\mathbf{U}(f,I) \le \lim_{\rho\to 1} \rho^k\,(b^{k+1}-a^{k+1})\frac{\rho-1}{\rho^{k+1}-1} = \frac{b^{k+1}-a^{k+1}}{k+1}.$$

为了证明这个数确实是积分，只需回忆上和是下和的 ρ^k 倍。正如$n \to +\infty$，下和与上和趋向于相同的极限，所以

$$\frac{b^{k+1}-a^{k+1}}{k+1} \le \mathbf{L}(f,I).$$

与之前一样，这同时证明了 f 是可积的，并计算了积分。我们很快就会看到可积性可以更容易地得出；这里需要做的是计算积分。

□

差异形式 $F(b) - F(a)$ 出现得足够频繁，足以需要特殊的符号表示：$F(x)\big|_{x=a}^{b} := F(b) - F(a)$。在这个符号表示法中，例7.7表明

$$\int_a^b t^k\,dt = \frac{t^{k+1}}{k+1}\bigg|_{t=a}^{b} \qquad \text{for } 0 < a < b. \tag{7.4}$$

$\int f(t)\,dt$的符号选择是为了它所蕴含的直观，即积分是无限多个无穷小项$f(t)\,dt$的和。无穷小dt可以看作是一个“重整化”因子，加权后$\int_a^b dt = b - a$。这种符号如此引人入胜，以至于它拥有了生命，并导致了诸如“什么是‘dt’？”这样的合理但难以回答的问题。在这本书中，积分下的无穷小是一个占位符和记忆辅助工具，但仅此而已。

7.3 积分的抽象性质

积分在固定区间 $[a, b]$ 上可以看作是一个定义域为 $[a, b]$ 上可积函数的实值函数。这种“积分泛函”与有限和有多个共同特征，这对我们希望积分对应（至少直观上）于求无穷小和的愿望来说是个幸运的事。首先，它在第3章的意义上是线性的，见下文定理7.8。其次，非负函数的积分也是非负的。第三，积分满足三角不等式的类似形式，见定理7.14。最后，从定理7.15的意义上讲，积分是“平移不变的”。

定理7.8. 设f和g是在区间I上的可积函数，且c是一个实数。那么$f+g$和cf是可积的，并且

$$\int_I (f+g) = \int_I f + \int_I g, \qquad \int_I (cf) = c\int_I f.$$

证明。设$P = \{t_i\}_{i=0}^n$是I的一个任意划分，并设

$$m_i' = \inf_{t\in I_i}\{f(t)\}, \quad m_i'' = \inf_{t\in I_i}\{g(t)\}, \quad m_i = \inf_{t\in I_i}\{(f+g)(t)\},$$

$f+g$ 在 I_i 上的下确界至少与 f 的下确界加上 g 的下确界一样大，即 $m_i \geq m_i' + m_i''$，并且用类似记号，$M_i \leq M_i' + M_i''$。将这些不等式相加，

$$\begin{aligned} L(f,P) + L(g,P) &\leq L(f+g,P) \\ &\leq U(f+g,P) \quad \leq \quad U(f,P) + U(g,P) \end{aligned}$$

对于每个划分 P。根据适当的情况取上确界或下确界，可以得出

$$\int_I f + \int_I g \leq \mathbf{L}(f+g, I) \leq \mathbf{U}(f+g, I) \leq \int_I f + \int_I g. \tag{7.5}$$

这同时表明 $f+g$ 在 I 上可积，并且积分具有所述值。断言 $\int_I (cf) = c\int_I f$ 是练习 7.5。

□

示例7.9 积分，像极限一样，不能区分在有限多个点之外相等的函数。精确地说，如果 $f : I \to \mathbf{R}$ 是可积的，并且如果 $g : I \to \mathbf{R}$ 在有限多个点之外等于 f，那么 g 是可积的，并且 $\int_I g = \int_I f$。

为了证明这一点，考虑函数 $h = g - f$，它在有限多个点之外为零。如果我们能证明 h 是可积的并且积分等于零，那么根据定理7.8，结论将随之而来，因为 $g = h + f$。由于 h 在有限多个点之外为零，我们可以写出

$$h = \sum_{j=1}^{k} c_j \, \chi_{\{x_j\}}$$

对于某些实常数 c_j 和 I 中的不同点 x_j。因此，只需证明每个函数 $\chi_{\{x_j\}}$ 都是可积的，并且积分等于零。这很容易从定义中得出，参见练习7.6。□

积分在以下意义上是单调的：如果 f 是 $[a, b]$ 上的非负可积函数，则 $\int_a^b f \geq 0$。以下是一个有用的改写；证明留给你们（练习7.7）。

定理7.10。设f和g是I上的可积函数。如果对于所有$t \in I$，则 $\int_I f \leq \int_I g$。

在文字上，不等式在固定区间上的积分得到保留。一个特殊情况，是定理7.10和例7.5的直接结果，被反复使用：

推论7.11。如果$f : [a, b] \to \mathbf{R}$可积，并且对于所有$m \leq f(t) \leq M$ $[t \in]$，则

$$m(b - a) \leq \int_a^b f \leq M(b - a).$$

积分的“修补性质”在定理7.12中给出。直观内容是，为了对一个区间上的函数进行积分，我们可以将区间分成有限多个子区间，并分别计算这些子区间上的积分之和。

定理7.12. 设$f : [a, b] \to$ $\mathbf{R}$为一个有界函数，并设$a < c < b$。那么f在$[a, b]$上可积当且仅当f在两个区间$[a, c]$和$[c, b]$上均可积，在这种情况下 $\int_a^b f = \int_a^c f + \int_c^b f$。

证明。假设 {v*} 在 [{v*}] 上可积。固定 {v*} 0，并选择一个 [{v*}] 的划分 {v*}，使得 {v*}。如果需要，通过添加点 {v*}，我们可以假设 {v*}。令 {v*} 为 [{v*}] 中点的集合。从定义中可以看出 {v*}。

根据命题7.4，f在$[a,c]$上可积。一个类似的论证表明f在$[c,b]$上可积。

相反，假设 f 在 $[a,c]$ 和 $[c,b]$ 上分别可积。固定 $\varepsilon >$ 为 0，并分别对 $[a,c]$ 和 $[c,b]$ 选择相应的分割 P' 和 P''，使得
$U(f,P')-L(f,P') < \varepsilon$/^2 和 $U(f,P'')-L(f,P'') < \varepsilon$/^2。并集
$P=P'\cup P''$ 是 $[a,b]$ 的一个分割，对于它有 $U(f,P)-L(f,P)<\varepsilon$。这表明根据命题 7.4，$f$ 在 $[a,b]$ 上可积。

在任一情况下，$L(f,P)=L(f,P')+L(f,P'')$以及对于上和的类似情况，这证明了$\int_a^b f=\int_a^c f+\int_c^b f$。 □

受此结果启发，我们对可积函数 f: $I\to \mathbf{R}$ 做出以下定义：

$$\int_a^a f=0,\quad \int_b^a f=-\int_a^b f \qquad \text{for all } a,\, b\in I. \tag{7.6}$$

使用这些定义，积分的以下“余性质”很容易验证。

命题7.13. 设$f:I\to\mathbf{R}$可积，并设a，b，$c\in I$。那么

$$\int_a^b f+\int_b^c f+\int_c^a f=0.$$

积分满足三角形不等式的类似形式。至于有限和，这个结果是一种用被积函数的绝对值来估计积分的工具。

定理7.14。如果 $f:I\to\mathbf{R}$ 可积，那么 $|f|:I\to\mathbf{R}$ 可积，并且

$$\left|\int_I f\right|\le\int_I|f|.$$

证明。反向三角不等式表明

$$\Big||f(x)|-|f(y)|\Big|\le|f(x)-f(y)| \quad \text{for all } x,\, y\in I. \tag{7.7}$$

选择一个任意的分割 P 的 I，并令 m_i 和 M_i 分别为 f 在第 i 个子区间上的下确界和上确界。令 m_i' 和 M_i' 分别表示 $|f|$ 在第 i 个子区间上的下确界和上确界，方程 (7.7) 意味着

$$M_i'-m_i'\le M_i-m_i. \tag{7.8}$$

现在固定 $\varepsilon > 0$ 并选择一个分割 P 使得 $U(f,P)-L(f,P) < \varepsilon$。方程 (7.8) 表明对于这个分割，$U(|f|,P)-L(|f|,P) < \varepsilon$ 也成立。由于 $\varepsilon > 0$ 是任意的，$|f|$ 是可积的。

第二部分现在很容易：对所有 $-|f(x)| \le f(x) \le |f(x)|$；由引理7.11，

$$-\int_I |f| \le \int_I f \le \int_I |f|, \quad \text{or} \qquad \left|\int_I f\right| \le \int_I |f|,$$

正如将要展示的那样。 □

几何上显然，如果我们“平移”f的图形左右，然后对适当平移的极限进行积分，积分的值是相同的。这个性质是积分的平移不变性。

定理7.15。如果 $f:[a,b] \to \mathbf{R}$ 是可积的，并且 $c \in \mathbf{R}$，那么

$$\int_{a+c}^{b+c} f(s-c)\,ds = \int_a^b f(t)\,dt.$$

证明。字母 s 和 t 用于区分，尽管它们也暗示了“变量替换”。关键观察是，如果 $P = \{t_i\}_{i=0}^n$ 是 $[a,b]$ 的一个划分，那么 $P_c = \{t_i + c\}_{i=0}^n$ 是 $[a+c, b+c]$ 的一个划分。如果我们设置 $g(s) = f(s-c)$，那么显然 g 在 $[t_{i-1}+c, t_i+c]$ 上的下确界等于 f 在 $[t_{i-1}, t_i]$ 上的下确界，对所有 i 都成立，上确界同理。因此，

$$L(f,P) = L(g,P_c) \text{ and } P(f,P) = U(g,P_c)$$

对于每个 P；定理立即成立。 □

黎曼和

设 $P = \{t_i\}_{i=0}^n$ 是 $[a,b]$ 的一个划分，设 f 是 $[a,b]$ 上的一个有界函数。从 P 中取出的黎曼和是一个形如的表达式

$$\sum_{i=1}^n f(x_i)\,\Delta t_i, \quad x_i \in [t_{i-1}, t_i] \text{ for } i = 1, \dots, n. \tag{7.9}$$

自对于所有 i 和所有 $x_i \in [t_{i-1}, t_i]$ 的 $m_i \le f(x_i) \le M_i$，每个黎曼和 P 都位于 $L(f,p)$ 和 $U(f,P)$ 之间。如果对有界

下限和上限和，然后可以通过任何方便的黎曼和来近似积分。实际优点是我们可以选择x_i的任何方便方式，并且不需要知道f在子区间上的下确界或上确界。典型的黎曼和由$x_i = t_{i-1}$ (左端和)，$x_i = t_i$ (右端和)，以及$x_i = (t_{i-1} + t_i)/2$（中点和）给出。

7.4 积分与连续性

本节包含两个重要的技术结果。第一个结果是关于连续函数可积的，它给出了一类大的可积函数，尽管它并没有直接给出关于特定积分评估的信息。第二个结果断言定积分是上限的连续函数。将积分视为上限的函数的想法是基本的，下面将详细讨论。

定理7.16. 设$f : [a, b] \to \mathbf{R}$为一个连续函数。则$f$在$[a, b]$上可积。

证明。我们将展示通过适当选择分割，上和下和可以任意接近，从而证明根据命题7.4，f是可积的。关键事实是定理5.5：在闭区间上的连续函数是均匀连续的。

修复 $\varepsilon > 0$。由于 f 的均匀连续性，存在一个 $\delta > 0$，使得 $|x - y| < \delta$ 蕴含 $|f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}$。选择一个最多包含 P 的网格的任意划分 δ。对于这样的划分，上和下和是 ε-接近的；实际上，如果 x 和 y 在子区间 I_i 中，那么 $|x - y| < \delta$，因此 $|f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}$。由此可得

$$M_i - m_i = \sup\{f(x) \mid x \in I_i\} - \inf\{f(x) \mid x \in I_i\} \leq \frac{\varepsilon}{2(b-a)}.$$

由于 i 是任意的，这个不等式对所有 i 都成立，因此我们有

$$U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)\Delta t_i \leq \frac{\varepsilon}{2(b-a)} \sum_{i=1}^{n} \Delta t_i = \frac{\varepsilon}{2} < \varepsilon.$$

根据命题7.4，f是可积的。 □

定理7.16就像极值定理一样，是一张“狩猎许可证”：它断言某些函数是可积的，但并没有说明如何求特定函数的积分。然而，定理7.16确实为积分的数值近似提供了一个理论基础，前提是积分函数f是明确已知的。如果证明的“获胜策略”可以计算实现，那么f的积分将通过对一个方便的网格划分的上和或下和进行近似，其误差在ε以内。下面的两个推论是应用中有用的重述。第一个通常被称为黎曼定理。

推论7.17。设$f:[a,b]\to\mathbf{R}$是连续的。固定$\varepsilon>0$，并选择$\delta>0$使得

$$|x-y|<\delta \implies |f(x)-f(y)|<\frac{\varepsilon}{2(b-a)}.$$

如果 P 是一个具有 $\|P\|<\delta$ 的划分，那么

$$\left|S-\int_a^b f\right|<\varepsilon \quad \textit{for every Riemann sum } S \textit{ taken from } P.$$

推论7.18.设$f:[a,b]\to\mathbf{R}$是连续的，且(P_n)是一个分划序列——不一定嵌套——使得$\|P_n\|\to 0$当$n\to\infty$。如果S_n是从P_n中取出的黎曼和，那么$\lim\limits_{n\to\infty} S_n=\int_a^b f$。

例如，$P=\{x_i\}_{i=0}^n$可能是等间距点的分区。对于评估积分，有更好的数值方案，但它们通常有效，因为可以证明该方法比定理7.16的证明给出的方法更好。

积分作为上限的函数

设 $f:[a,b]\to\mathbf{R}$ 可积。对于每个 $x\in[a,b]$，函数 f 在 $[a,x]$ 上由定理7.12 可积。定义 $F:[a,b]\to\mathbf{R}$ 为

$$F(x)=\int_a^x f=\int_a^x f(t)\,dt. \tag{7.10}$$

一些思考可以确认，F接受一个数字作为输入，并返回一个数字作为输出。潜在地，这个过程将产生新的、有趣的功能。然而，F的定义可能看起来

奇怪；如果给定 x，那么（从实际角度来说）如何评估 $F(x)$? 这是我们第一次遇到一个不以代数公式形式呈现的函数；我们必须求助于积分的定义。为了从单个 x 的定义评估 $F(x)$，我们必须计算当 P 在 [a, x] 的划分集合上变化时，f 的下和集合的上确界。正如我们在例子7.5和7.7中已经看到的，即使 f 是一个单项式，这也是一个相对繁琐且非算法化的任务。为了了解我们希望得到什么，让我们回顾一下例子7.7的结果：

$$\int_a^x t^k\,dt = \frac{x^{k+1} - a^{k+1}}{k+1} \qquad a,\ x > 0,\ k \in \mathbf{N}. \tag{7.11}$$

根据这个方程的视角，结果要么令人失望，要么引人入胜。也许，我们希望发现奇异的新函数，但遗憾的是，我们只恢复了一个单项式的积分作为多项式。然而，我们可能会发现，左边的积分（一个复杂对象）等于右边的多项式（一个简单对象），并意识到这是一条重要且非平凡的信息。如果你不明白为什么，回顾第3章中关于函数的哲学观点可能是个好主意。特别是，一个函数可能由两个完全不同的“规则”来描述。方程（7.11）左侧的规则很复杂，但有一个有趣的解释（一个非多边形区域的面积）。右侧的规则很简单，但本身没有特别的内在意义。两个规则定义了同一个函数，这确实很有用！假设我们希望找到由t轴、抛物线$y = t^2$和直线$t = 1$和$t = 2$所围成的区域的面积。很容易将这个面积表示为一个积分，但直接从定义中计算这个积分是困难的。鉴于方程（7.11），我们不需要使用定义；我们立即可以看出

$$\text{area} = \int_1^2 t^2\,dt = \frac{2^3 - 1^3}{3} = \frac{7}{3}.$$

如果我们有生产其他像这样的“魔法公式”的手段，我们就会有一个强大的计算工具可供使用。

结果表明，k次幂函数的积分不是一个新函数，实际上存在许多简单的积分可以产生“奇异”函数，这些函数不能仅通过代数手段表达。最重要的非代数函数之一是

自然对数，定义为看似无害的积分

$$L(x) = \int_1^x \frac{dt}{t}, \qquad t > 0.$$

在练习7.17中，你被要求建立自然对数的一些基本性质。除了细节之外，你应该仔细注意如何使用积分的抽象性质来推断由积分定义的函数的事实。其他有趣的积分有

$$\operatorname{asin}(x) = \int_0^x \frac{dt}{\sqrt{1-t^2}}, \qquad |x| < 1,$$

和

$$\operatorname{atan}(x) = \int_0^x \frac{dt}{1+t^2}, \qquad x \in \mathbf{R}.$$

令人惊讶的是，这些函数与圆三角学有关，但要了解原因，以及更深入地理解定义为积分的函数，需要第8章的内容。

集成和 O 符号

我们的方程（7.11）的证明需要假设$0 < a \leq x$。为了研究O符号在积分下的行为，我们需要将我们的知识扩展到$a = 0$的情况。证明告诉我们（7.11）的两边在a上是连续的，因此我们可以通过求值在$a = 0$处取极限。

命题7.19。如果 $k \geq 0$ 是一个整数且 x 是实数，那么

$$\int_0^x t^k\,dt = \frac{x^{k+1}}{k+1}.$$

证明。我们首先假设 $x > 0$。固定 $a \in (0, x)$，并注意对于 $t \in [0, a]$，有 $0 \leq t^k \leq a^k$。使用“平凡”的划分 $P' = \{0, a\}$，我们有

$$0 = L(f, P') \leq U(f, P') = a^{k+1}.$$

由于精炼改善了界限，相同的非等式对$[0, a]$的每个划分P'都成立。现在令P''是$[a, x]$的一个划分；正如定理7.12的证明中那样，如果 $P = P' \cup P''$，那么

$$\begin{aligned} L(f, P'') &\leq L(f, P') + L(f, P'') = L(f, P) \leq U(f, P) \\ &= U(f, P') + U(f, P'') \leq a^{k+1} + U(f, P''). \end{aligned}$$

取下和的上确界和上和的下确界（但保持 a 不变），我们有

$$\frac{x^{k+1}-a^{k+1}}{k+1} \le L(f,P) \le U(f,P) \le a^{k+1} + \frac{x^{k+1}-a^{k+1}}{k+1}.$$

设 $a \to$ = 0，根据夹逼定理，我们得到 $\int_0^x t^k\,dt = \dfrac{x^{k+1}}{k+\ }$ = 1。

案例 x < 0 可以通过重复计算示例 7.7，在适当的地方改变符号来处理。或者，命题 7.13 和练习 7.13 意味着如果 $y > 0$，那么

$$\int_0^{-y} t^k\,dt = -\int_{-y}^0 t^k\,dt = -\int_0^y (-t)^k\,dt = (-1)^{k+1}\int_0^y t^k\,dt,$$

该操作将情况 $x < 0$ 降低到情况 $y > 0$。 □

我们接下来建立以下基本原理：“积分增加消失阶数一级”在以下意义上：

定理7.20。设$k \in \mathbf{N}$，并设f在包含a的某个区间上可积。如果 $f(x) = O\big((x-a)^k\big)$在包含$a$的某个区间$I$上，那么 $\int_a^x f = O\big((x-a)^{k+1}\big)$在$I$上。

证明。如果我们定义 $g(x-a) = f(x)$，那么 $g(u) = O(u^k)$，即存在一个实数 C，使得 $|g(u)| \le Cu^k$ 对于 u 在某个关于 0 的区间内成立。定理 7.15 意味着

$$\int_a^x f(t)\,dt = \int_0^{x-a} g(u)\,du,$$

因此，根据定理7.14、命题7.19和方程（7.11），我们有

$$\left|\int_a^x f(t)\,dt\right| = \left|\int_0^{x-a} g(u)\,du\right| \le C\cdot\left|\int_0^{x-a} u^k\,du\right|$$

$$\le \frac{C}{k+1}|u|^{k+1}\Big|_{u=0}^{x-a} = O\big((x-a)^{k+1}\big)$$

在某些区间关于 a。 □

定理7.20在理论和实践上都很有用；它允许我们研究积分，而无需直接处理ε和δ。为了给出

简单但重要的应用，我们将证明由积分定义的函数自动连续。证明预示了所谓的微积分基本定理，这是大量函数可以积分的关键。

推论7.21。设$f:[a,b]\to$ **R**可积。由以下公式定义的函数F:

$$F(x)=\int_a^x f,\qquad x\in[a,b],$$

是连续的。

证明。根据假设，f是有界的，即$f=O(1)$在$[a,b]$上。如果x和$x+h$在$[a,b]$中，那么

$$|F(x+h)-F(x)|=\left|\int_a^{x+h} f-\int_a^x f\right|=\left|\int_x^{x+h} f\right|=O(h)$$

根据定理。这不仅仅证明了F是连续的，还证明了F是Lipschitz（练习5.15）。此外，对$|f|$的每一个界都是F的Lipschitz常数。

□

您可能会想知道是否每个连续函数都是某个其他函数的积分。答案是“不”，因为存在不是Lipschitz的连续函数。在$[0,1]$上的平方根函数就是一个例子。

一个例子将说明如何在具体情况下找到 F。因为我们可用的计算工具相对较少，所以这个例子（在计算上）极其简单。

示例7.22设$f:\mathbf{R}\to\mathbf{R}$为符号函数，设$F$为积分，$F(x)=\int_0^x f$。分别考虑$x>0$和$x<0$的情况。假设$x>0$。符号函数在半开区间$(0,x]$上等于1，在0处等于0。由于积分在有限多个点改变$f$的值时不改变，我们可以假设$f$在$[0,x]$上等于1。因此

$$F(x)=\int_0^x f=\int_0^x 1\,dt=x\quad\text{for } x\ge 0.$$

同样，如果 $x<0$，那么 f 在区间 $[-1,x)$ 上等于 -1，在改变 0 的值后我们发现

$$F(x)=\int_0^x f=-\int_x^0(-1)\,dt=-(-1)(0-x)=-x\quad\text{for } x<0.$$

总结来说，$F(x)=|x|$ 对 $x\in\mathbf{R}$。 □

7.5 不定积分

集成本质上需要定义域为有界区间的有界函数。存在一些情况，人们希望放宽其中一项或两项要求。不定积分是将普通积分推广到特殊情况的手段，在这些特殊情况下，被积函数和/或积分区间是不确定的。一些例子将说明我们希望回答的问题类型。

函数 $f : [0, 1] \to \mathbf{R}$ 定义为

$$f(t) = \begin{cases} 1/t \text{ if } t > 0 \\ 0 \text{ if } t = 0 \end{cases}$$

在0附近无界，但在其他地方局部有界。假设我们希望计算$[0, 1]$上f的积分。无论我们选择什么分割，总有一些子区间上f是无界的，因此无法计算上和。一种可能的补救方法是取$0 < \delta < 1$，将积分

$$F(\delta) := \int_{\delta}^{1} \frac{1}{t}\,dt$$

作为一个关于 δ 的函数，并考虑 $\lim(F, 0^+)$。如果这个极限存在，那么 f 就被称为在 $[0, 1]$ 上“不适当可积”。根据练习7.17，在这种情况下，这个极限不存在，因此倒数函数在 $[0, 1]$ 上不是不适当可积的。（被积函数在0处的值无关紧要；不适当可积性完全由“被积函数在无界点附近增长的速度”决定。）

如果相反，我们希望不当地积分 $f(t) = 1/\sqrt{|t|}$ 在 $[-1, 1]$ 上，我们首先将积分区间拆分为 $[-1, 0] \cup [0, 1]$（以保证被积函数仅在每个子区间的端点附近无界），然后考虑两个单独的不当子积分。如果两个不当子积分都存在，那么原始函数是“不当可积的。”不幸的是，我们目前还没有手段来决定这个问题。

我们的最终例子涉及倒数函数，但是在无界区间$[1, +\infty)$上。被积函数是有界的，但无法对区间进行划分，因为划分只有有限个点。在这种情况下，想法是尝试定义

$$\int_{1}^{\infty} \frac{1}{t}\,dt = \lim_{R \to +\infty} \int_{1}^{R} \frac{1}{t}\,dt.$$

再次根据练习7.17，极限不存在。

一般来说，一个不定积分是一个积分表达式，其中被积函数或积分区间（或两者）是无界的。为了判断一个不定积分是否存在（或“收敛”），将区间分成有限多个子区间，使得在每个部分中，要么区间是无界的，要么被积函数在恰好一个端点无界，但不能同时无界。对每个结果不定积分分别考虑。如果它们都有极限，则极限之和被宣布为原始表达式的值。如果一个或多个子问题没有极限，则原始积分不存在，或“发散”。不难证明，在满足上述标准的情况下，可以对定义域进行任意方便的划分。记号 $\int_{\mathbf{R}}$ 有时会代替 $\int_{-\infty}^{+\infty}$ 使用.

不恰当的可积性很少通过精确计算近似“恰当”积分来决定；相反，通过适当估计近似积分来推断其存在性。存在一个有用的积分测试，该测试将级数的可和性与不恰当积分的存在性联系起来。

命题7.23。设$f:[0,\infty)\to\mathbf{R}$为一个非递增、正函数，并且对于$a_k=f(k)$有$k\in\mathbf{N}$。序列$(a_k)_{k=0}^{\infty}$可求和当且仅当$f$是不定可积的。换句话说，

$$\sum_{k=0}^{\infty} a_k \ \textit{converges iff} \int_0^{\infty} f(t)\,dt \ \textit{converges.}$$

在练习7.8中，你将证明非增函数是自动可积的，因此无需单独提出这个假设。求和/积分的下限无关紧要；不定积分的收敛，就像级数的收敛一样，完全取决于尾部的行为，与被积函数在有限区间上的行为或级数有限项的行为无关。

证明。设$k\in\mathbf{N}$。因为f是正的且非递增的，所以我们有

$$0<a_{k+1}=f(k+1)\le f(x)\le f(k)=a_k \quad \text{for all } x\in[k,k+1].$$

因为区间 $[k,k+1]$ 的长度为 1，前面的不等式积分得到

$$a_{k+1}\le\int_k^{k+1} f(t)\,dt\le a_k \quad \text{for all } k\in\mathbf{N}.$$

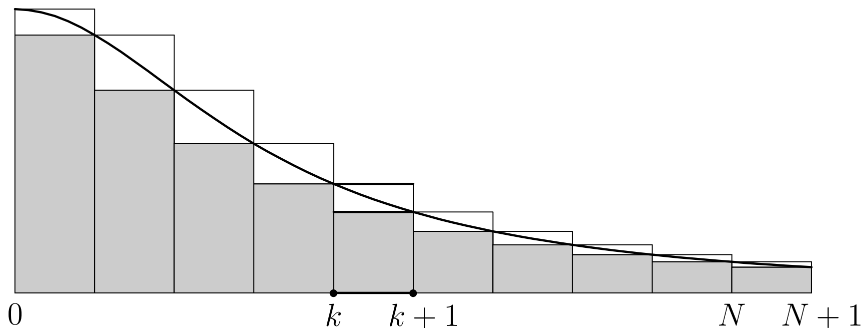


图7.7：用级数的部分和上下界积分。

这些不等式在 $k = 0, \ldots, N$ 上求和，我们得到

$$0 < \sum_{k=1}^{N+1} a_k = \sum_{k=0}^{N} a_{k+1} \leq \int_0^{N+1} f(t)\,dt \leq \sum_{k=0}^{N} a_k.$$

该系列是积分的下和与上和，见图7.7。因此，如果该级数的部分和是有界的，则积分是有界的，这正是要证明的。

□

使用“p-系列测试”（示例4.64）和命题7.23，你可以轻松确定当不定积分

$$\int_1^{\infty} x^{-r}\,dx$$

收敛，参见练习7.28。这些积分对于估计具有更复杂被积函数的不定积分非常有用。不定积分是许多令人愉快和巧妙的公式的来源，但这样的应用必须等到我们拥有更多的函数可供使用时才能进行。

练习

练习7.1设f : [−^2,^2] → **R**为定义的阶跃函数

$$f(x) = \begin{cases} -1 & \text{if } -1 \leq x < 0 \\ 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

制作 f 的仔细草图，然后在同一组坐标轴上绘制函数

$$F_0(x) = \int_0^x f, \qquad F_1 = \int_{-1}^x f.$$

找到一个关于 F_1 的代数公式。 $\diamond$

练习7.2 设 $f : [a, b] \to \mathbf{R}$ 为一个阶梯函数。证明定积分 F 是分段线性的。$\diamond$

练习7.3修改示例7.5中的论点以评估

$$\int_0^x t^2\,dt \qquad \text{for } x > 0$$

直接。给出两个一般原因，说明平方函数在[0, x]上是可积的。（抛物线下面积的计算归功于西西里的阿基米德。）$\diamond$

练习 7 .4 假设 f 在 $[a, b]$ 上可积，并且 $c \in$ $[a, b]$。

(a) 证明由以下定义的函数

$$F(x) = \int_a^x f(t)\,dt \quad \text{and} \quad G(x) = \int_c^x f(t)\,dt$$

不同一个常数。提示：使用命题7.13。

(b) 如果 $H(x) = \displaystyle\int_x^b f(t)\,dt$，那么 F 和 H 之间有什么关系?

$\diamond$

练习7.5 通过证明如果 f 在 I 和 $c \in \mathbf{R}$ 上可积，那么 $\int_I (cf) = c \int_I f$. 来完成定理7.8的证明
建议：如果 $c = 0$，则该主张是明显的。分别考虑 $c > 0$ 和 $c < 0$ 的情况。$\diamond$

练习7.6 证明点特征函数可积，积分为零。更确切地说，如果 $a \le x \le b$，那么 $\chi_{\{x\}}$: $[a, b] \to \mathbf{R}$ 是可积的，并且

$$\int_a^b \chi_{\{x\}} = 0.$$

$\diamond$

练习7.7 证明如果 f 是 $[a, b]$ 上的非负可积函数，那么 $\int_a^b f \geq \geqslant 0$。使用这个结果来证明定理7.10。$\diamond$

练习7.8 证明非递减函数 {v*} : $[a, b] \to \mathbf{R}$ 是可积的。

建议：有界性是明显的。考虑将 $[a, b]$ 划分为等长的小区间，并明确写出上下和的差。$\diamond$

练习 7.9

(a) 给出一个非减函数 {v*10}：[0{v*11} 1] {v*12} 的例子，该函数有无限多个间断点。

(b) 证明例3.11的分母函数在区间[0, 1]上可积。积分的值是多少?

$\diamond$

练习 7.10

(a) 证明函数 $f : [a, b] \to \mathbf{R}$ 可积当且仅当以下条件成立：对于每个 $\varepsilon > 0$，存在阶梯函数 s_1 和 s_2，使得 $s_1 \leq f \leq s_2$ 在 $[a, b]$ 上且

$$\int_a^b (s_2 - s_1) < \varepsilon.$$

直观上，一个可积函数可以被“夹”在积分任意接近的阶梯函数之间。

(b) 对于以下函数 f 在 $[0, 1]$ 上，按照 (a) 部分绘制一对阶梯函数的图像：恒等函数；原点的特征函数；“$1/q$”函数。

$\diamond$

练习 7.11

(a) 证明如果 {v*} 在 [{v*}] 上可积，那么存在连续函数 {v*} 和 {v*}，使得 {v*} 在 [{v*}] 上，并且

$$\int_a^b (h - g) < \varepsilon.$$

上一次练习的结果和草图应该有帮助。

(b)对于以下函数 f 在 $[0,1]$ 上，绘制与部分 (a) 中类似的两个连续函数的图像：恒等函数；原点的特征函数；“$1/q$”函数。

$\diamond$

练习7.12 设f和g是$[a,b]$上的可积函数。

(a) 证明 f^2 在 $[a,b]$ 上可积。建议：根据定理7.14，$|f|\geq 0$ 是可积的。使用 $f^2=|f|^2$ 来界定下和与上和。

(b) 证明 fg 在 $[a,b]$ 上可积。提示：2
$fg=(f+g)^2-f^2-g^2$。

代数技巧（部分（b））是一个“极化恒等式”。 $\diamond$

练习7.13 设a和b为实数，设f为在闭区间端点为ca和cb的某个实数c上的可积函数。证明

$$\int_{ca}^{cb} f(t)\,dt = c\int_a^b f(ct)\,dt.$$

建议：情况 $c=0$ 是明显的。依次考虑以下情况：$a<b$ 和 $c>0$；$a<b$ 和 $c=-1$；$a<c$ 和 $c<0$；$a>b$ 和 $c\in\mathbf{R}$。$\diamond$

练习7.14 设$a>0$，并设f在$[-a,a]$上可积。使用练习7.13来证明以下内容：

(a) 如果 f 是奇数，那么 $\displaystyle\int_{-a}^{a} f=0$。

(b) 如果 f 是偶数，那么 $\displaystyle\int_{-a}^{a} f=2\int_0^a f$。

(c) 设 $F(x)=\int_0^x f$ 对于 $x\in[-a,a]$。证明如果 f 是偶数，那么 F 是奇数，并且如果 f 是奇数，那么 F 是偶数。

这只是操作积分的练习；不需要任何技术。$\diamond$

练习7.15 使用示例7.7和练习7.14来证明

$$\int_0^x |t|\,dt = \frac{x|x|}{2} \qquad \text{for all } x\in\mathbf{R}.$$

提示：分别考虑 $x \geq = 0$ 和 $x < = 0$ 的情况。 ◇

练习7.16 证明{v*}

$$\int_a^b t^k\,dt = \frac{t^{k+1}}{k+1}\bigg|_{t=a}^{b} \qquad \text{for all } a,\, b \in \mathbf{R}.$$

建议：首先使用示例7.7和命题7.19处理$a < 0$，$b = 0$的情况；然后在0处拆分积分。◇

练习7.17 分析中最重要的函数之一是自然对数函数 log: $(0, \infty) \to \mathbf{R}$，定义为[1]

$$\log x = \int_1^x \frac{1}{t}\,dt.$$

(a) 使用命题7.13证明对数是一个增函数，并且log(1) = 0。

(b) 证明以下公式：{v*}

$$\log(ab) = \log(a) + \log(b) \qquad \text{for } a,\, b > 0.$$

建议：编写

$$\int_1^{ab} \frac{1}{t}\,dt = \int_1^a \frac{1}{t}\,dt + \int_a^{ab} \frac{1}{t}\,dt,$$

然后使用练习7.13。

(c) 使用部分(b)证明对于所有 $a > 0$，有 $\log(1/a) = -\log a$，并且更一般地，对于所有 $a > 0$ 和 $n \in \mathbf{Z}$，有 $\log(a^n) = n \log a$。得出结论：log 映射 $(0, \infty)$ 到 $\mathbf{R}$。

(d) 证明倒数函数在$[0, 1]$上不可积，也不在$[1, \infty$）上可积。

(e) 根据部分(c)，存在一个实数 $e > 0$，使得 $\log(e) = 1$，根据部分(a)这个数是唯一的。使用显式下和上和来证明 $2 < e < 4$。草图将非常有帮助。

[1]No mathematician calls this function ln outside a calculus course.

(f) 实际上，$e < 3$；从几何上看，曲线 $y = 1/t$ 在 $t = 2$ 处的切线位于图形下方，并在 $t = 1$ 和 $t = 3$ 之间包围一个单位面积。仅使用迄今为止开发的工具严谨地表达这个论点并不困难。首先证明对于所有 $t > 0$，有 $1 - \frac{1}{2}t \le \frac{1}{t}$。接下来，在 $[1, 3]$ 上对等式两边进行积分，使用练习 7.16 处理线性多项式。为了完成证明，调用本章中适当的一个定理。

数字 e 在数学中扮演着重要角色。在第12章中，我们将找到一个收敛速度快的级数表示。$\diamond$

练习7.18 通过命题7.23的证明，

$$\sum_{k=2}^{n} \frac{1}{k} < \int_1^n \frac{1}{t}\,dt < \sum_{k=1}^{n-1} \frac{1}{k} \qquad \text{for all } n \ge 2.$$

(a) 仔细绘制一张插图，说明这些不等式。

(b) 对于 $n \geq$^2，设置

$$\gamma_n = \int_1^n \frac{1}{t}\,dt - \sum_{k=2}^{n} \frac{1}{k} = \log n - \sum_{k=2}^{n} \frac{1}{k}.$$

证明序列$(\gamma_n)_{n=2}^{\infty}$是递增且有上界的。
您的草图应建议一种归纳方法。

(©) 根据部分(b)，$\gamma := \lim_n \gamma_n$ 存在。确定γ是否为有理数.[2]

欧拉引入了常数 γ。 $\diamond$

有限数字列表的平均值是列表总和除以条目数。类似地，可积函数 $f : [a, b] \to \mathbf{R}$ 的平均值定义为:

$$\text{(7.12)} \qquad \text{Average value of } f \text{ on } [a, b] = \frac{1}{b-a}\int_a^b f = \int_a^b f \Big/ \int_a^b 1.$$

如果 f 非负，则图形下的面积等于宽度为 $(b - a)$ 且高度等于 f 平均值的矩形的面积。如果区间固定，平均可能表示为 $\overline{f}$ 或 f_{avg}。

练习7.19 如果 f 在 $[a, b]$ 上可积，那么 $\displaystyle\int_a^b (f - \overline{f}) = 0$。 $\diamond$

[2]Resolving this open question will earn you an excellent publication and make you famous in mathematical circles.

练习7.20 设 $f\colon [a,b] \to \mathbf{R}$ 是连续的。对于每个正整数 n，设 $P_n = \{t_i\}_{i=0}^n$ 是将 $[a,b]$ 分成 n 个等长区间的划分。证明

$$\lim_{n\to+\infty} \frac{1}{n+1} \sum_{i=0}^{n} f(t_i) = f_{\text{avg}}.$$

这进一步证明了“平均值”的定义。 ◇

练习7.21 证明如果 $f\colon [a,b] \to \mathbf{R}$ 是连续的，那么存在一个 $c \in (a,b)$ 使得 $f(c) = \overline{f}$。这个结果被称为积分的平均值定理。

如果 $f\colon [0,1] \to \mathbf{R}$ 是平方函数，$f(x) = x^2$，求满足平均值定理的 (0, 1) 中 c 的值，并仔细绘制 f 及其平均值。◇

练习 7.22 设 f 在某个区间 ($c-\eta$ 上可积 $, c+\eta$）。

(a) 证明如果 $0 < |h| < \eta$，则

$$\frac{1}{h}\int_c^{c+h} f \tag{7.13}$$

是闭区间端点为 c 和 $c+h$ (的 f 在 $h<0$) 下的平均值。

(b) 假设 f 在 c 处连续。证明

$$\lim_{h\to 0} \frac{1}{h}\int_c^{c+h} f = f(c).$$

(c) 通过例子说明，如果 f 在 c 处不连续，则 (b) 的结果在一般情况下会失败。

◇

练习7.23 设 $f\colon [a,b] \to \mathbf{R}$ 可积。证明存在一个 $x \in [a,b]$ 使得

$$\int_a^x f(t)\,dt = \int_x^b f(t)\,dt.$$

通过示例说明通常不可能选择 $x \in (a,b)$。 ◇

练习7.24 设$f\colon [0,1] \to \mathbf{R}$ 是一个在对于每个δ属于(0, 1)上的可积函数。对以下每个命题给出证明或反例:

(a) 如果 $\lim_{\delta\to 0}\int_{\delta}^{1} f(x)\,dx$ 的极限存在，那么 f 在 [0, 1] 上可积。

(b) 如果lim(f, 0)存在，那么f在[0, 1]上可积。

注意，未假设 f 有界。 $\diamond$

练习7.25 设 $g : [0, \infty) \to \mathbf{R}$ 是非负且非恰当可积的。假设 $f : [0, \infty) \to \mathbf{R}$ 在区间 $[0, x]$ 上对所有 $x > 0$ 是可积的，并且存在一个 $R > 0$ 使得 $|f(t)| \le g(t)$ 对于 $t \ge R$ 成立。证明 f 在 $[0, \infty)$ 上是非恰当可积的。 $\diamond$

练习7.26 假设 f 在 $[0, x]$ 上对所有 $x > 0$ 可积。通常，让

$$f_+ = \max(f, 0), \qquad f_- = \min(f, 0)$$

正负部分为 f 。证明当且仅当 f_+ 和 f_- 在 $[0, \infty)$ 上是不规则可积的，$|f|$ 在 $[0, \infty)$ 上是不规则可积的。$\diamond$

练习7.27 使用练习7.25和练习7.17的第(d)部分，

(a) 证明 $t \mapsto t^{-r}$ 在 $[1, \infty)$ 上不是不适定可积的，对于 $r < 1$。提示：如果 $r < 1$ 且 $t \ge 1$，那么 $t^{-r} \ge t^{-1}$。

(b) 证明 $t \mapsto t^{-r}$ 在 [0, 1] 上对于 $r > 1$ 不是不适定可积的。

您不需要知道如何积分 $t^{-r}\,dt$，并且如果您有这方面的知识，也不应该使用它。如果您很挑剔，您可以将 r 视为有理的，因为我们还没有为无理的 r 定义 t^r。$\diamond$

练习7.28 证明如果 $r > 1$，则在先前的条件下 $\int_1^{+\infty} t^{-r}\,dt$ 收敛。使用这个结果来证明

$$\int_0^{+\infty} \frac{dt}{1+t^2} \quad \text{and} \quad \int_2^{+\infty} \frac{dt}{t^2 - t}$$

收敛。第一个应该很简单；第二个稍微有点棘手，但如果正确处理，则不难。$\diamond$

第八章

微分

积分区间是一个将无穷小 $f(t)\,dt$ “组合”在一起以获得实数的过程。通过改变区间，我们得到一个函数。微积分的另一个主要操作在某种程度上是相反的。微分是寻找函数变化率的过程，是将函数“拆分”成无穷小增量。通过改变取变化率的位置，我们得到一个新函数，该函数测量原始函数的变化率。

牛顿和莱布尼茨符号

对于我们来说，无穷小是一个方便的虚构，值得简要地重新讨论它们的状态。当数量 y 依赖于另一个数量 x 时，即当 y 是 x 的函数时，定义了“变化率”的概念。思考揭示，我们感兴趣的核心对象是函数本身，而不是我们为其输入和输出值所赋予的名称。微积分中有两种突出的符号：牛顿符号，其中强调函数，和莱布尼茨符号，其中强调输入和输出的名称。每种都有其优点和缺点：

- 牛顿符号更为紧凑，不会引入“自变量”的虚假符号，但也不暗示自变量的无穷小性质。

- 莱布尼茨记号通常更容易用于计算和现实世界建模，但将无穷小视为数字，并为符号赋予多重含义，让用户自行解读正确的解释。

牛顿不那么挑衅性的符号类似于建筑的框架。其重要性通常不是直接的实用性，而是它明确地用数字和函数表达微积分概念的方式，以及它因此给予更友好但更容易被滥用的莱布尼茨符号的支持。每个人都使用莱布尼茨符号，但数学家们不知不觉地将一切翻译回牛顿语言，尤其是在莱布尼茨符号失效时。为了充分利用微积分，你应该精通这两种语言，并且能够在它们之间自由翻译。因此，本书并行发展了这两种语言。

我们尚未定义无穷小量[1]，因此只能将其用于指导，但不能用于定义或证明。牢记这一点，必须承认，在传统意义上，“微积分”正是无穷小量的操作。本章和下一章中提出了几个证明这种操作的定理，并且通常无穷小量（莱布尼茨）的解释比基于极限（牛顿）的解释更有说服力——因此更容易记住和使用。仅从概念上讲（更不用说它们的计算价值），完全摒弃无穷小量是不明智的。然而，在最终分析中，我们必须确保我们的定义和论证不是建立在实数公理之外的东西上。如果对无穷小量操作不当，会导致明显的悖论和其他哲学难题。如有疑问，定义总是最后的决定。

8.1 导数

假设“y是x的函数”。y相对于x的变化率是函数值变化与输入值变化的比值。用牛顿语言来翻译，如果f是一个函数，并且如果$[a,b]$是包含在f定义域内的一个区间，那么

$$\text{Average rate of change of } f \text{ over } [a,b] = \frac{f(b)-f(a)}{b-a}. \tag{8.1}$$

当f的图像是一条直线（即，“f的变化率是恒定的”）时，上面的商给出了直线的斜率，符合直觉。

[1]More to the point, we have not shown that the existence of infinitesimals is logically consistent with the axioms of $\mathbf{R}$.

如何定义函数在一点的改变率? 直观的答案, “在上面的公式中将 $a=b$ 设置为”, 并不有帮助, 因为右侧变成了不确定表达式 0/0 。在微积分的早期, 所谓的“答案”是将 a 和 $b=a+dx$ 看作无限接近:

$$\text{Rate of change of } f \text{ at } a = \frac{f(a+dx)-f(a)}{dx} = \frac{dy}{dx}.$$

这个想法在实践中非常有效（如果明智地应用）, 但会受到合法的投诉。符号 dx 的意思是表示一个比每个正实数都小的正量。那么 dx 又是什么呢? 如果目标是证明微积分没有逻辑矛盾, 这个反对意见是致命的。如果目标只是用微积分来描述自然世界, 那么只要结论与现实没有显著差异, 这个反对意见就无关紧要了。

借助三个世纪的回顾（以及第4章的结果）, 我们可以巧妙地规避这一反对意见。设f为一个域包含某些$\eta>0$的区间$(x-\eta,x+\eta)$的函数。目前, 点x是任意但固定的。如果$0<|h|<\eta$, 那么在x处的f的牛顿商是一个表达式

$$\frac{\Delta y}{\Delta x}(x,h) := \frac{f(x+h)-f(x)}{h}, \tag{8.2}$$

查看图8.1。式(8.2)中的牛顿商是f在端点为x和$x+h$的区间上的平均变化率, 参见图(8.1)。你应该验证即使在$h<$为0的情况下, 这也是正确的。

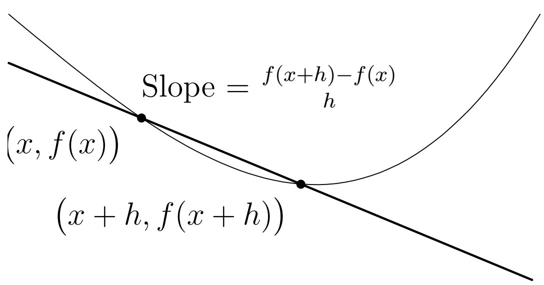


图8.1: 牛顿商作为平均变化率。

为了简单起见, 我们可能将 $\Delta_x f(h)$ 写作 $\frac{\Delta y}{\Delta x}(x,h)$。对于 f 的定义域中的每个 x, $\Delta_x f$ 是一个定义域包含所有非-

[2]The notation $\Delta_x f$ is not standard in this context.

零 h 使得 $x+h$ 在 f 的定义域内。根据假设，$\Delta_x f$ 在关于 0 的删除区间上定义，但在 0 处未定义；然而，$\Delta_x f$ 很可能在 0 处有极限.[3] 如果

$$\lim(\Delta_x f, 0) = \lim_{h\to 0} \frac{\Delta y}{\Delta x}(x, h)$$

存在，则极限分别用牛顿和莱布尼茨记号表示为 $f'(x)$ 或 $\frac{dy}{dx}(x)$，并称为 f 在 x 处的导数，图8.2。在这种情况下，称 f 在 x 处可微，导数被解释为 f 在 x 处的“瞬时变化率”。

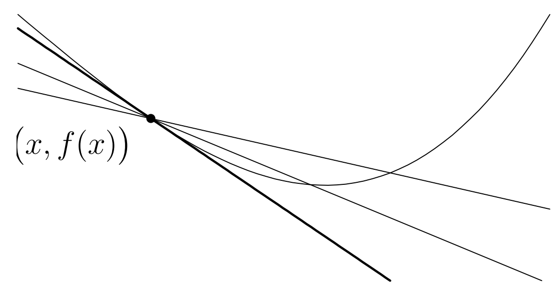


图8.2：切线线作为割线线的极限。

莱布尼茨记号 $\frac{dy}{dx}$ 对导数的表示暗示了无穷小量的商，即 y 的无穷小增量除以 x 中相应的无穷小增量。尽管我们没有单独定义符号 dy 和 dx，但我们确实为它们的“商”赋予了精确的含义：后者是比率的极限 $\frac{\Delta y}{\Delta x}$ 当 $\Delta x \to$ 为 0。然而，我们必须小心使用这些商的熟悉操作。在我们可能使用诸如

$$\frac{d(y+z)}{dx} = \frac{dy}{dx} + \frac{dz}{dx} \quad \text{or} \quad \frac{dz}{dy}\frac{dy}{dx} = \frac{dz}{dx},$$

我们必须回到这些表达式的定义，即实数商的极限，以验证这些方程是否确实成立。目前我们没有逻辑基础假设这些公式可以扩展到无穷小商。

[3] In this assertion is resolved one of Zeno's paradoxes of motion, as well as the heated debate between Newton and Bishop Berkeley on the nature of infinitesimals.

派生和 o 符号

第4章中我们引入的“粗略计算”在微分学中发挥其作用。在 o 符号表示法中，如果 f 在 x 处可微，那么

$$\frac{f(x+h)-f(x)}{h} = f'(x) + o(1) \quad \text{at } h = 0.$$

乘以 h 并将 $f(x)$ 加到两边，我们发现如果 f 在 x 处可微，导数为 $f'(x)$，那么

$$f(x+h) = f(x) + h\,f'(x) + o(h) \quad \text{at } h = 0. \tag{8.3}$$

这个论点也可以反过来运行；如果 f 在 x 的邻域内定义，并且如果 $f(x+h) = f(x)+h\,c+o(h)$ 在 $h =$ 处为 0，那么 f 在 x 和 $f'(x) = c$ 处可微。方程 (8.3) 是可微性定义的一个非常有用的重新表述：为了证明一个函数 ϕ 在 x 处可微，我们只需要证明存在一个数 c，使得 $\phi(x+h) = \phi(x) + h\,c + o(h)$ 在 $h =$ 附近为 0。此外，如果我们能够用已知量来表示 c，那么我们就找到了 $\phi'(x)$。

导数的定义简短（与积分的定义不同），但表面上看似简单。第4章的大部分技术工具都涉及其中，定义的深远后果看似直观可信，但需要第5章的结果。这些更深的性质收集在第9章中。本章关注微分的初步方面，这些方面往往表现为（8.3）的简单计算结果。与积分相比，微分从定义上来说是相对算法化的。可微函数的和、积、商和复合的导数可以用几个容易记住的公式来计算。在一点的可微性在几何上表现为与图形“相切”的直线的存在，而$f'(x)$的符号表明函数在x处是否在某种意义上增加或减少。

命题8.1. 如果 f 在 x 处可微，那么 f 在 x 处连续。

证明。根据假设，f的定义域包含关于x的一些区间，因此x是定义域的一个极限点，因此有意义的询问是否lim(f, x) $= f(x)$。但是，由于f在x处可微，我们有

$$f(x+h) = f(x) + h\,f'(x) + o(h) = f(x) + O(h) + o(h) = f(x) + O(h).$$

这表明 f 在 x 处是连续的。 □

命题8.1的逆命题是错误的。绝对值函数$f(x) = |x|$在0处连续，但在0处不可导。实际上，牛顿商是

$$\frac{f(h) - f(0)}{h - 0} = \frac{|h|}{h} = \text{sgn } h,$$

符号函数，在0处没有极限。一般来说，一个连续函数在无处可导，尽管我们将在第11章给出一个例子。

在最基本的函数中包括单项式，对于单项式有一个简单的微分公式：

命题8.2. 设$f(x) = x^n$对于n是一个正整数。那么f在每处都是可微的，并且$f'(x) = nx^{n-1}$对于每个$x \in \mathbf{R}$。在莱布尼茨记号中，$\frac{d(x^n)}{dx} = nx^{n-1}$。

特别地，单项式的导数是一个比原单项式低一度的单项式。如果我们同意对于所有 x，$0x^{-1} = 0$，则该公式也适用于 $n = 0$。我们很快就会看到，这个结果使我们能够轻松地求多项式函数的导数。

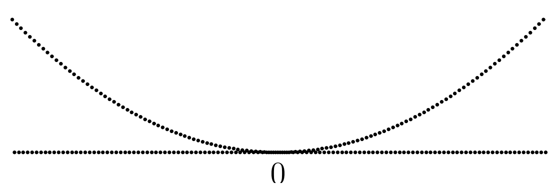
证明。二项式定理意味着

$$(x + h)^n = x^n + h\,nx^{n-1} + O(h^2) = x^n + h\,nx^{n-1} + o(h)$$

在 $h = 0$。命题随即得出。 □

示例8.3 一个函数 $f : \mathbf{R} \to \mathbf{R}$ 可以在恰好一个点可导，在其他每个点不可导。一个例子是

$$f(x) = \begin{cases} x^2 \text{ if } x \in \mathbf{Q}, \\ 0 \text{ if } x \notin \mathbf{Q}. \end{cases}$$



如果 $x \neq$ 为 0，则根据推论 4.21，f, x 的极限不存在。由于 f 在 $x \neq$ 为 0 处不连续，根据命题 8.1，f 在 x 处也不可导。要在 0 处对 f 求导，计算牛顿商：

$$\Delta_0 f(h) = \frac{f(h) - f(0)}{h} = \begin{cases} h \text{ if } h \in \mathbf{Q}, \\ 0 \text{ if } h \notin \mathbf{Q}. \end{cases}$$

自$0 \leq \Delta_0 f(h) \leq |h|$ 对所有 $h \neq 0$，压缩定理表明 $\lim(\Delta_0 f, 0)$ 存在且等于0；换句话说，$f'(0) = 0$。

一般来说，令 $f : (-\eta, \eta) \to \mathbf{R}$ 为一个函数，使得 $f(h) = O(h^2)$ 在 $h = 0$ 附近。从几何上看，f 的图像位于一对形式为 $y = \pm Cx^2$ 的抛物线之间。由于 $f(0)$ 必须为 0，我们有 $f(h) = f(0) + h\,0 + o(h)$，这表明 $f'(0)$ 存在且等于 0。□

导数的和、积和商

如上所述，存在对可微函数的和、积或商进行求导的计算规则。作为一个简单的翻译练习，你应该用莱布尼茨符号表达以下结果的结论。

命题8.4. 假设f和g在x处可微，并且$c \in \mathbf{R}$。那么函数$f + g$和cf在x处可微，其导数分别为$(f + g)'(x) = f'(x) + g'(x)$和$(cf)'(x) = c\,f'(x)$。

证明。根据假设，存在实数 $f'(x)$ 和 $g'(x)$，使得

$$\left.\begin{aligned} f(x+h) &= f(x) + h\,f'(x) + o(h) \\ g(x+h) &= g(x) + h\,g'(x) + o(h) \end{aligned}\right\} \quad \text{at } h = 0. \tag{8.4}$$

这些方程相加，我们得到

$$(f+g)(x+h) = (f+g)(x) + h\,\big(f'(x) + g'(x)\big) + o(h).$$

这同时证明了$f + g$是可微的，以及在x处的导数是$f'(x) + g'(x)$。对于常数倍的断言类似，留给你们。

□

推论8.5.如果$p(x) = \sum\limits_{k=0}^{n} a_k x^k$是一个多项式，那么对于所有$x \in \mathbf{R}$，$p$在$x$处可微，并且

$$p'(x) = \sum_{k=1}^{n} k a_k x^{k-1}$$

例如，

$$\frac{d}{dx}(1-x+x^2-x^3+x^4) = -1+2x-3x^2+4x^3,$$
$$\frac{d}{dx}(x+2x^2+3x^3) = 1+4x+9x^2.$$

设 $X = (a,b)$ 为一个开区间，设 $\mathscr{D}^1(X) \subset \mathscr{F}(X,\mathbf{R})$ 表示 X 上的可微函数集。命题 8.4 表示

- $\mathscr{D}$1(X) 是一个向量子空间（第3章），并且
- The mapping $f \in \mathscr{D}^1(X) \mapsto Df := f' \in \mathscr{F}(X,\mathbf{R})$ is linear.

我们将立即看到，可微函数的导数通常不可微，因此 D 技术上不是 $\mathscr{D}^1(X)$ 上的算子。实际上，D 的像在这本书中过于复杂，难以描述。

产品与商公式

存在类似的产品和商的公式，使用 o 符号计算很容易。首先建立一般倒数公式是方便的。

引理8.6. 如果 $f(x+h) = 1 + ah + o(h)$ 接近 $h = 0$，那么

$$\frac{1}{f(x+h)} = 1 - ah + o(h)$$

附近 $h = 0$。

证明。在关于 $h = 0$ 的某个区间内，我们有 $|ah + o(h)| < 1$。几何级数公式给出

$$\begin{aligned}\frac{1}{f(x+h)} &= \frac{1}{1+ah+o(h)} \\ &= 1 - \big(ah + o(h)\big) + \big(ah + o(h)\big)^2 - \cdots = 1 - ah + o(h)\end{aligned}$$

附近 $h = 0$。 □

定理8.7. 假设f和g在x处可微。那么fg在x处可微，并且

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

如果 $g(x) \neq 0$，那么 f/g 在 x 处可微，并且

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

证明。根据假设，公式（8.4）成立，所以

$$\begin{aligned}(fg)(x+h) &= \big[f(x) + h\,f'(x) + o(h)\big]\big[g(x) + h\,g'(x) + o(h)\big] \\ &= (fg)(x) + h\,\big[f'(x)g(x) + f(x)g'(x)\big] + o(h)\end{aligned}$$

该公式建立了乘积法则。我们将商的论证分为两步，首先处理简单的倒数。

为了简洁，写 $a_0 = g(x) \neq 0$ 和 $a_1 = g'(x)$。g 的可微性表明 $g(x+h) = a_0 + h\,a_1 + o(h) = a_0\big(1 + h\,(a_1/a_0) + o(h)\big)$ 在 $h = 0$ 附近。根据引理8.6,

$$\frac{1}{g(x+h)} = \frac{1}{a_0} \cdot \frac{1}{1 + h\,(a_1/a_0) + o(h)} = \frac{1}{a_0} - h\,\frac{a_1}{{a_0}^2} + o(h) \quad \text{near } h = 0.$$

因此，$1/g$在x处可微，并且

$$\left(\frac{1}{g}\right)'(x) = -\frac{a_1}{{a_0}^2} = -\frac{g'(x)}{g(x)^2}.$$

现在通过写出 $f/g = f \cdot (1/g)$ 并使用刚刚证明的结果，得出一般结果：

$$\begin{aligned}\left(\frac{f}{g}\right)'(x) &= f'(x)\frac{1}{g(x)} + f(x)\left(\frac{1}{g}\right)'(x) \\ &= f'(x)\frac{1}{g(x)} - f(x)\frac{g'(x)}{g(x)^2} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}\end{aligned}$$

如所述。 □

可以以两种不同的方式区分多项式，例如 $p(x) = x^{n+m} = x^n x^m$ 或 $q(x) = (1-x)(1+x^2)$，要么通过展开乘积并使用引理8.5，要么使用乘积法则。你应该验证这两种方法得到相同的结果。

定理8.7允许我们将单项式微分公式扩展到具有负指数的项。证明留作练习。

命题8.8. 如果 $f(x) = x^n$ 对于 n 是一个整数，那么 $f'(x) = nx^{n-1}$ 对于所有 $x \neq 0$。

从定理8.7和推论8.5可以得出，每个有理函数在其自然定义域内都是可微的。例如，

$$\frac{d}{dx}\frac{1}{1+x^2} = -\frac{2x}{(1+x^2)^2}, \qquad f(x) = 1,\ g(x) = 1+x^2,$$
$$\frac{d}{dx}\frac{x}{1+x^2} = \frac{(1+x^2)-x(2x)}{(1+x^2)^2} = \frac{1-x^2}{(1+x^2)^2}, \quad f(x) = x,\ g(x) = 1+x^2.$$

注意，代数变换可能使有理函数更容易求导。例如，

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

但是右侧比左侧更容易微分。

微分积分

在第七章中，我们看到了如何使用积分从已知函数构造新的函数：如果 f 在某个区间 $[a, b]$ 上可积，那么方程

$$F(x) = \int_a^x f, \qquad x \in [a, b],$$

定义在 $[a, b]$ 上的连续函数。尝试对 F 求导，并期望得到一个关于 f 的 F' 的公式是自然的。如果你已经完成了第7.22题，你已经知道了结果。

定理8.9。设$f : [a, b] \to \mathbf{R}$为一个可积函数，并且假设$f$在$c \in (a, b)$处连续。那么上面定义的函数$F$在$c$和$F'(c) = f(c)$处可导。

这个定理可能看起来像是一种有趣的奇思妙想，但正如我们将看到的，它确实配得上微积分基本定理这个名称。我们还没有达到理解其全部意义的地步，但无疑它表明了积分和微分之间存在着密切的关系。证明已在练习7.22中概述，但这里还有一些更多细节。

证明。积分的余项性质（命题7.13）表明，只要 $c+h$ 在 $[a,b]$ 中，$\int_a^{c+h} = \int_a^c + \int_c^{c+h}$ 就成立。换句话说，

$$F(c+h) - F(c) = \int_a^{c+h} f - \int_a^c f = \int_c^{c+h} f.$$

牛顿商在 F 处的 c 因此由以下公式给出

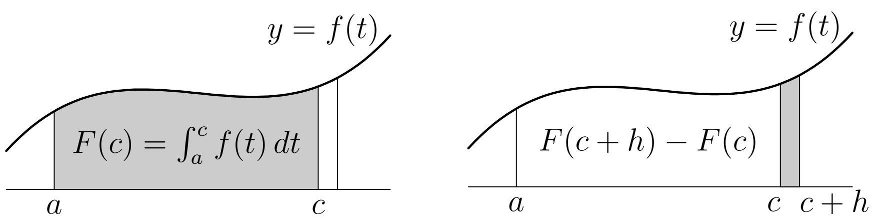


图8.3：定积分的增量。

$$\Delta_c F(h) = \frac{F(c+h) - F(c)}{h} = \frac{1}{h}\int_c^{c+h} f,$$

f在端点为c和$c+h$的区间上的平均值。现在我们使用连续性：$f = f(c) + o(1)$在c处，因此对于接近0的h，我们有

$$\int_c^{c+h} f = \int_c^{c+h} \big(f(c) + o(1)\big) = h\,f(c) + o(h),$$

参见定理7.20。因此，$\Delta_c f(h) = f(c) + o(1)$在$h = 0$附近。 □

如果 f 在 c 处不连续，则无法进行说明；示例表明 F 在 c 处可能或可能不可微。符号函数在 0 处有跳跃间断性，积分后得到绝对值函数，该函数在 0 处不可微。相比之下，如果 f 在除原点外处处为零，且 $f(0) = 1$，那么 f 在任意区间上的积分为零，因此 F 是零函数，显然在原点处也是可微的。

链式法则

链式法则是对两个可微函数复合的导数的公式。

定理8.10. 假设f在x处可微，且g在$f(x)$处可微。那么$g\circ f$在x处可微，并且

$$(g\circ f)'(x) = g'\big(f(x)\big)\cdot f'(x).$$

证明。根据假设，

$$\begin{aligned} f(x+h) &= f(x) + h\,f'(x) + o(h) \quad \text{at } h=0 \\ g(y+k) &= g(y) + k\,g'(y) + o(k) \quad \text{at } k=0 \end{aligned}$$

如果我们写出 $y=f(x)$ 和 $k=h\,f'(x)+o(h)$，那么 $k=O(h)=o(1)$ 在 $h=0$，因此 $o(k)=o(h)$。因此，

$$\begin{aligned} (g\circ f)(x+h) = g(y+k) &= g(y) + k\,g'(y) + o(k) \\ &= g\big(f(x)\big) + \big(h\,f'(x) + o(h)\big)\,g'(y) + o(h) \\ &= g\big(f(x)\big) + h\,g'\big(f(x)\big)\cdot f'(x) + o(h), \end{aligned}$$

这完成了证明。 □

链式法则是微分学中最强大的计算工具之一。考虑尝试对 $p(x) = (4+x-x^3)^{11}$ 求导。如果没有链式法则，唯一的方法是展开乘积，得到一个33次的多项式，然后使用命题8.4进行求导。假设没有出错，答案将以未因式分解的形式出现，而因式分解它并非易事。相比之下，链式法则一步就给出了因式分解的答案。通过 $f(x)=4+x-x^3$ 和 $g(y)=y^{11}$ 定义 f 和 g。（使用 y 完全是为了心理上的方便，所以我们可以立刻设置 $y=f(x)$。）上述关于多项式函数导数的公式意味着 $f'(x)=1-3x^2$ 和 $g'(y)=11y^{10}$。由于 $p=g\circ f$，链式法则给出

$$p'(x) = g'\big(f(x)\big)\cdot f'(x) = 11\big(4+x-x^3\big)^{10}(1-3x^2).$$

链式法则在莱布尼茨记号法中看起来特别有说服力。如果我们写出 $y=f(x)$ 和 $z=g(y)$，那么 $z=(g\circ f)(x)$，所以

$$\frac{dz}{dx}(x) = \frac{dz}{dy}(y)\cdot\frac{dy}{dx}(x), \quad \text{or even} \quad \frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

链式法则因此可以被视为一个定理，它为无穷小量的商的某种形式操作提供了正当性。以免这种解释使结果看起来显而易见（“只需取消 dy”），请记住

- 链式法则看起来像分数的约分，因为我们像分数一样表示导数，并不是因为它们是分数。对于像 dx 这样的无穷小量，对我们来说在孤立状态下是没有意义的。从逻辑上讲，“约掉 dy” 和“约掉 n”然后“推导”出 $\frac{\sin x}{\tan x} = \frac{\text{six}}{\text{tax}}$ 一样不合法。此外，我们修改了符号（通过省略函数的参数）来使结论看起来像分数约分。

- 左侧的“z”表示函数$g \circ f$在x处的值，或函数$g \circ f$本身。链式法则右侧的“z”表示g在y处的值，或函数g本身。这些z通常不是同一个函数！

通常，不必要的混淆源于在莱布尼茨符号（科学家们喜欢这样做）中编写函数和使用牛顿导数符号（数学家们喜欢这样做）；参见第8.4题，以了解这种陷阱的一个简单例子。然而，链式法则的“分数消去”解释可以是一个有用的记忆法，只要你记得刚才提到的要点。

8.2 导数与局部行为

如果对于其定义域中的每个 x，f 在 x 处可微，则称 f 可微。在这种情况下，存在一个函数 f'，其定义域与 f 的定义域相同，并且对于每个 x，由以下公式定义

$$f'(x) = \lim(\Delta_x f, 0) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}. \tag{8.5}$$

导数的符号

如果 f 在 x 处可微，那么为了简便，我们用 $f(x) = a_0$ 和 $f'(x) = a_1$ 表示，则有

$$f(x+h) = a_0 + a_1 h + o(h).$$

此条件断言 f 在 x 附近被线性函数近似；$f(x+h)$ 与线性函数之间的差异相对于 h 非常小。

假设 $f'(x) = a_1 > 0$。上述主导非常数项是 a_1h，这意味着 f 的值在某些位于 x 右侧的区间内大于 $f(x)$，而在某些位于 x 左侧的区间内小于 $f(x)$。形式上，存在一个 $\delta > 0$ 使得

$$0 < h < \delta \implies f(x-h) < f(x) < f(x+h).$$

这个条件可以通过说 f 在 x 处增加来表示。一个类似的论点表明，如果 $f'(x) < 0$，那么（在明显意义上）f 在 x 处减少。

注意8.11 如果 f 在 x 上增加，并不意味着存在一个 $\delta > 0$ 使得 f 在区间 $(x-\delta, x+\delta)$ 上增加。符号函数

$$\operatorname{sgn}(x) = \begin{cases} x/|x| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

在0处增加（仔细阅读定义！）但在0的任何邻域内都不增加。练习8.12描述了一个在0处连续但在0的任何邻域内甚至不是非递减的例子。在练习8.15中，你将找到一个可微函数g，其中$g'(0) > 0$，但在0的任何开区间内都不增加！□

关于导数符号的观察使我们能够陈述并证明与优化相关的一个重要性质。根据定理5.8，一个连续函数 $f : [a, b] \to \mathbf{R}$ 有最小值和最大值。上述论据表明，一个点 x 在其中 $f'(x) \neq 0$ 不能是 f 的极值点。

定理8.12. 设$f : [a, b] \to \mathbf{R}$为一个连续函数，并且设$x_0 \in [a, b]$为$f$取得最小值或最大值的一点。那么$x_0$是以下之一：

- 一个端点为$[a, b]$;
- 一个内点使得 $f'(x_0) = 0$;
- 一个内点，其中 $f'(x_0)$ 不存在，

一个点 $x \in (a, b)$，其中 $f'(x) = 0$ 是 f 的一个临界点。临界点之所以重要，是因为它们是 f 极值可能的位置。例8.13说明了在寻找极值时使用定理8.12的应用。

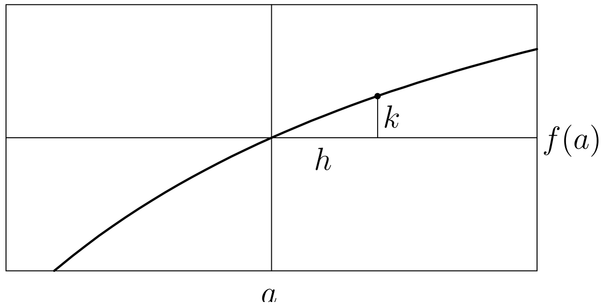


图8.4：放大图形。

切线与导数

几何上，对函数 f 在 a 处求导相当于“以无穷大因子放大图形。”为了进一步拓展这种莱布尼茨风格的隐喻，可导函数的图形由无限多个无穷小线段组成，a 处线段的斜率是 $f'(a)$，“a 的上升与下降比。”本节的目的在于强调这些观点。

设 f 是在 a 处可微的函数。通过点 $\bigl(a, f(a)\bigr)$ 的直线如果斜率为 $f'(a)$，则与图形相切。直观上，图形的切线是图形的无穷小“模型”。

要了解为什么“在 $\bigl(a, f(a)\bigr)$ 处以无穷大因子放大”相当于找到 a 处的切线，考虑一下在 $\bigl(a, f(a)\bigr)$ 处放大对平面的影响。如果 h 和 k 表示从放大中心的水平和垂直位移（图8.4），那么以 λ 的因子放大将 (h, k) 映射到 $(\lambda h, \lambda k)$。这用图 $k/\lambda = f(a + h/\lambda) - f(a)$ 替换了图 $k = f(a + h) - f(a)$，或者说

$$k = \left(\frac{f(a + h/\lambda) - f(a)}{(h/\lambda)}\right) \cdot h.$$

如果 f 在 a 处可微，那么当 $\lambda \to \infty$ 上面的方程趋近于 $k = f'(a) \cdot h$ 时，即为切线方程。

在 o 符号表示法中，有一个更简单（但不太严谨）的解释：由于 $f(a + h) = f(a) + h\, f'(a) + o(h)$，以无穷大因子放大消除了可忽略项 $o(h)$，留下了切线方程。

优化

如果 $f : [a, b] \to \mathbf{R}$ 是一个连续函数，那么根据极值定理，f 实现最大值和最小值。在实践中，人们通常希望找到函数的极值点，而不仅仅是证明它们的存在。定理8.12断言极值点必须是端点、临界点或不可微点，在这种情况下是一个有用的工具。

示例8.13 假设我们希望找到一个底边位于 x 轴上的最大面积的矩形，并且该矩形内切于抛物线 $y = 1 - x^2$，如图8.5所示。

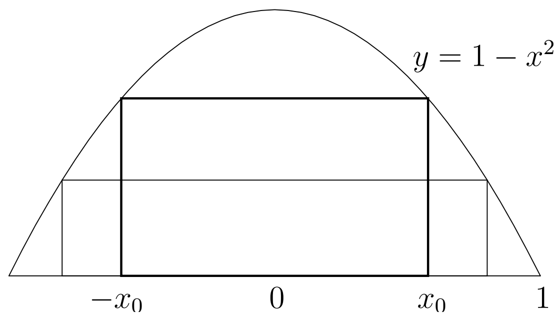


图8.5：抛物线内切的最大面积矩形。

如果我们令 $x \geq 0$ 为矩形的右侧坐标，那么面积是 $A(x) = 2x(1 -x^2) = 2x - 2x^3$ 对于 $x \in [0, 1]$。函数 A 在一个闭区间上连续，因此根据极值定理存在一个最大点。此外，函数 A 在 $(0, 1)$ 的每个点可导，且 $A'(x) = 2 - 6x^3 = 2(1 - 3x^2)$。只有一个临界点，$x_0 = 1/\sqrt{3}$，因此 A 的极值必须在列表 0，x_0，1 中找到。由于 $A(0) = A(1) = 0$，而 $A(x_0) > 0$，x_0 必须是最大点！我们作为额外收获得到最大面积，$A(x_0) = 4/(3\sqrt{3})$。仅用代数和几何找到最大面积的矩形不是一件容易的事。□

所给出的论点是排除法的过程。首先，我们知道A在$[0, 1]$区间内存在一个最大点。其次，我们知道如果$0 < x < 1$且$A'(x) \neq 0$，那么x不是极值点。这个事实排除了上述所有可能性，除了三种。端点不能是最大点，因为面积函数在端点消失，而在其他地方为正。唯一剩下的可能性是x_0是最大点。

在其他情况下，可能会有多个临界点，但只要要优化的函数在$(a,b$ $)$上可微，在$[a,b]$上连续，定理8.12仍然是有帮助的；和之前一样，极值必须是端点或临界点，如果只有有限个临界点，那么极值搜索就简化为有限搜索。

示例8.14 假设我们想了解多项式$f(x)=x-x^3/6$的最小值和最大值，在$-2\le x\le 3$的约束条件下。首先注意到f在$(-2,3)$上是可微的，并且$f'(x)=1-x^2/2$，因此f的临界点是$-\sqrt{2}$和$\sqrt{2}$。定理8.12保证极值必须出现在列表-2，$-\sqrt{2}$，$\sqrt{2}$和3中。直接计算给出

$$f(-2)=\frac{2}{3},\quad f(-\sqrt{2})=-\frac{2\sqrt{2}}{3},\quad f(\sqrt{2})=\frac{2\sqrt{2}}{3},\quad f(3)=-\frac{3}{2},$$

所以问题就转化为在这个列表中找到最大和最小的数字。

现在，$2\sqrt{2}=\sqrt{8}<\sqrt{9}=3$，所以最小值是 $f(3)=-3/2$；f 的唯一最小点是 3，f 的最小值是 $-3/2$。同样，$1<\sqrt{2}$，所以最大值是 $f(\sqrt{2})$：唯一最大点是 $\sqrt{2}$，f 的最大值是 $2\sqrt{2}/3$。□

8.3 导数的连续性

如果 $f:(a,b)\to\mathbf{R}$ 可微，则存在一个函数 $f':(a,b)\to\mathbf{R}$；然而，函数 f' 在一般情况下不连续。如果 f' 是一个连续函数，则称 f 为连续可微的，或 $\mathscr{C}^1$。这样的函数集合表示为 $\mathscr{C}^1(a,b)$。例如，有理函数在其自然域上是 $\mathscr{C}^1$，因为导数是具有相同域的另一有理函数。

有很大可能性您从未见过一个导数不连续的可微函数。自然的第一猜测，即绝对值函数，并不是一个例子，因为它在0处（不连续性“应该存在”的地方）不可微。事实上，我们必须更加狡猾：

示例8.15 设$\psi:\mathbf{R}\to\mathbf{R}$为一个非常数、可微、周期函数。（我们在示例9.11中构建此类函数，）

第13章。）定义

$$f(x) = \begin{cases} x^2\psi(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

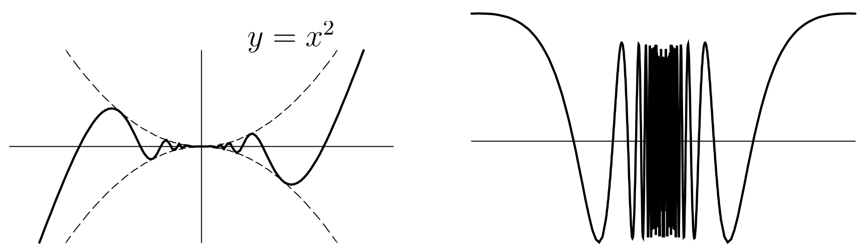


图8.6：一个非-$\mathscr{C}^1$函数及其导数。

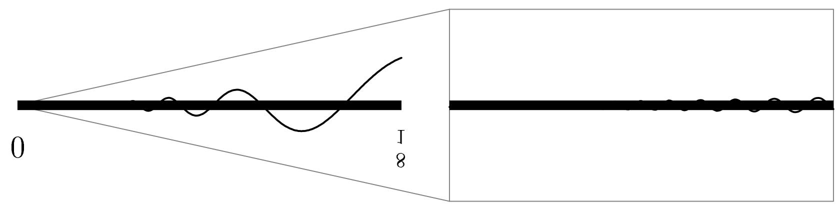
远离0，f是通过组合和乘以可微函数得到的，因此可以通过定理8.7和8.10进行微分：

$$f'(x) = 2x\psi(1/x) - \psi'(1/x) \qquad \text{for } x \neq 0.$$

在 $x = 0$ 这些定理不适用（它们的假设不成立），但可以从定义中计算出 0 处的导数；实际上，$f(h) = O(h^2)$ 在 $h = 0$ 附近，所以根据 8.3 例末的注释，$f'(0)$ 存在且等于 0。总之，f 在 $\mathbf{R}$ 的所有地方都是可微的。验证 f' 在 $x = 0$ 处不连续的任务留给你，参见练习 8.15。□

了解Example 8.15中0附近的精确情况非常重要。图8.6是一个起点，但更好的方法是使用可以缩放并高分辨率显示的绘图程序。图8.6和下面的图片都是使用$\psi = \sin$绘制的。

如果您在 $x = 0$ 处放大，图表会迅速变平成水平线；这反映了 $f'(0) = 0$ 的实际情况。然而，如果您在接近 0 的点处放大，图表首先会放大成 ψ 图表的近似，然后稳定到切线。这反映了切线斜率在 $x \searrow 0$ 时无限次振荡的性质。



当 f 在 0 附近绝对值较小时，其导数并不小。例 8.15 展示了这种情况发生的确切原因：一个接近水平轴的图形可以具有大斜率的小振荡。从线性映射的角度来看，D 可以将一对差值较小的函数映射到差值较大的函数。这应该让你想起不连续函数的作用。

8.4 高阶导数

如果 $f:(a,b)\to\mathbf{R}$ 可微，那么导数 f' 是定义域为 $X:=(a,b)$ 的函数，并且有意义的询问 f' 是否本身可微。如果是这样，我们说 f 是二阶可微的；f' 的导数表示为 $f''=D^2f$，称为 f 的二阶导数。为了展望未来，我们也用 $f^{(2)}$ 表示 f 的二阶导数。二阶可微函数的集合是一个向量子空间 $\mathscr{D}^2(X)\subset\mathscr{F}(X,\mathbf{R})$。

连续性考虑适用于二阶导数；具有连续二阶导数的函数被称为 {v*}，所有此类函数的集合是一个子空间。稍加思考就会使你信服这些包含关系

$$\mathscr{C}^2(X)\subset\mathscr{D}^2(X)\subset\mathscr{C}^1(X)\subset\mathscr{D}^1(X)\subset\mathscr{C}(X)\subset\mathscr{F}(X,\mathbf{R}).$$

如您所猜，模式继续向左延伸；由 k 乘以的连续可微函数的向量子空间定义为

$$\mathscr{C}^k(X)=\{f\in\mathscr{F}(X,\mathbf{R})\mid f^{(k)}\text{ exists and is continuous}\}.$$

在这本书中，我们对这些空间并不那么感兴趣，尽管我们将遇到它们交集的几个成员，即光滑函数空间：

$$\mathscr{C}^\infty(X)=\bigcap_{k=1}^{\infty}\mathscr{C}^k(X)=\{f\in\mathscr{F}(X,\mathbf{R})\mid f^{(k)}\text{ exists for all }k\text{ in }\mathbf{N}\}.$$

多项式和有理函数在其自然定义域上是光滑的，自然对数也是如此，其导数是有理的。例8.15中的函数f除了在0处外是光滑的，但当然在$\mathbf{R}$上甚至不是$\mathscr{C}^1$。练习8.9和8.11给出了更多非光滑函数的例子。

高阶差分

关于高阶导数，目前我们可以说得不多，因为我们还没有严格地建立导数的某些“显然”性质（例如，如果 $f' > 0$，那么 f 是递增的）。正如连续函数的“显然”性质一样，可微函数的熟悉性质比它们最初看起来要微妙，除非对假设进行一些注意，否则实际上并不成立！研究导数所需的技术工具是“平均值定理”，这是第9章的主题。

然而，初等代数让我们对函数的一阶和更高阶导数中编码的信息有所了解。在本节的其余部分，设 f 为定义域为区间 I 的实值函数；除非另有说明，所有点均假定是 I 的元素。

差商

$$\frac{\Delta y}{\Delta x}(a, b-a) = \frac{f(b) - f(a)}{b - a}$$

测量f在区间$[a, b]$上的变化率。在真实实验中，差商是我们所知道的全部，因为无法（甚至从哲学上讲有意义）收集定义域内所有点的数据。相反，科学家假设存在一个数学模型（一开始未知），并且测量数据作为模型的输出（包括实验误差）。

至少需要两个测量来确定一个函数是否（平均而言）是增加还是减少。f的两个测量相当于f'的一个测量，该测量通过在两个无限小分离点采样f来计算。

现在假设我们想要测量变化率的变化速度。变化率是 f'，其变化率为 f''。我们需要对 f' 进行两次测量，或者对 f 进行三次测量。想象一下等待穿越繁忙的街道，左右张望（就像网球比赛中的观众）看是否有车辆驶来。你必须进行两次观察来

确定一辆车辆行驶的速度有多快。此外，如果你看到一辆车从左侧驶来，然后观察到右侧没有车辆驶来，仍然谨慎地再次向左看，看看迎面而来的车辆是否在加速。如果在第三次看后，车辆行驶的距离比你的前两次观察中行驶的距离要远得多，你应该重新评估是否安全过马路.[4]

作为另一个例子，考虑一家公司在从2000年1月（比如说）开始计算的月份中，其净值为$V(t)$美元。在我们能够确定该公司是盈利还是亏损之前，我们必须至少看到两个季度报告（即获得两个V的值），并且必须看到至少三个报告才能知道盈利是上升还是下降。在商业界，如果公司的净值在增加，但净值增加的速度在下降，即净值的二阶导数为负，那么通常认为该公司“在亏损”。

练习

练习8.1 将命题8.4用莱布尼茨符号表示，并解释结果作为形式操作工具的有用性。◇

练习8.2 证明命题8.8。 ◇

练习8.3 设 n 为一个正整数，考虑以下有限几何级数

$$\sum_{k=0}^{n} x^k = 1 + x + x^2 + \cdots + x^n = \frac{x^{n+1} - 1}{x - 1} \qquad \text{if } x \neq 1.$$

(a) 使用定理8.7和命题8.8对当 $x \neq 1$ 时此方程进行微分。

(b) 使用(a)部分找到该级数的闭式表达式

$$\sum_{k=1}^{n} kx^k = x + 2x^2 + 3x^3 \cdots + nx^n, \qquad x \neq 1.$$

[4]This advice is distilled from an incident in which the author was nearly hit by a speeding cab at the intersection of Bloor and St. George streets in Toronto.

(c) 继续同样的思路，推导出

$$\sum_{k=1}^{n} k^2x^k = \frac{n^2x^{n+2} - (2n^2+2n-1)x^{n+1} + (n+1)^2x^n - (x+1)}{(x-1)^3}$$

对于 $x \neq 1$.

已知和的积分或微分技术是在适当情况下的一种强大技巧。$\diamond$

练习8.4 假设 $y = x^2$ 和 $z = y^2$，使得 $z = x^4$。然后 $z'(y) = 2y$，当 $x = 1$ 时我们得到 $z'(1) = 2$。然而，$z'(x) = 4x^3$，所以在 $x = 1$ 时我们有 $z'(1) = 4$。因此 $2 = 4$。有什么错误？ $\diamond$

练习8.5 证明在所有周长为 $P > 0$ 的矩形中，存在一个面积最大的矩形，并找出其尺寸。使用微积分和纯代数（完成平方）来解决这个问题。$\diamond$

练习8.6 证明在所有面积为 A 的矩形中，存在一个周长最小的矩形，参见练习8.5。与使用微积分相比，这个问题用纯代数方法更容易解决，因为不能使用极值定理来推导最小值的存在。教训是，微积分并不总是优化问题的最佳技术。$\diamond$

练习8.7 求内接于半径为r的半圆内面积最大的矩形的尺寸；你可以假设矩形的一边沿着直径。$\diamond$

练习8.8 考虑一组矩形，其左下角位于原点，其右上角位于图 $y = 1/(1+x^2)$ 上，且其边与坐标轴平行。证明在这个家族中存在一个面积最大的矩形，并找出其尺寸。$\diamond$

导数连续性

练习8.9 设k为一个正整数，并定义f：$\mathbf{R} \to \mathbf{R}$为$f(x) = x^k|x|$。求$f$的导数，并证明$f$是$\mathscr{C}^k$但不是$(k+1)$次可微的。换句话说，包含$\mathscr{C}^k(\mathbf{R}) \subset \mathscr{D}^{k+1}(\mathbf{R})$是适当的。

建议：对 k 进行归纳。 $\diamond$

练习8.10 设 $f : \mathbf{R} \to \mathbf{R}$ 可微但不是 $\mathscr{C}^1$，并设

$$F(x) = \int_0^x f(t)\,dt.$$

证明 F 是二阶可微的，但不是 $\mathscr{C}^2$。 $\diamond$

练习8.11 设$k > 1$为一个整数。继续上一个练习，证明存在一个函数f，它是k次可微的，但不是$\mathscr{C}^k$。换句话说，包含$\mathscr{D}^k(\mathbf{R}) \subset \mathscr{C}^k(\mathbf{R})$是适当的。$\diamond$

练习8.12 定义 $f : \mathbf{R} \to \mathbf{R}$ 为

$$f(x) = \begin{cases} x \text{ if } x \in \mathbf{Q} \\ 2x \text{ if } x \notin \mathbf{Q} \end{cases}$$

证明 f 在 0 处是增加的，但不存在 $\eta > 0$ 使得 f 在开区间 $(-\eta, \eta)$ 上是增加的。$\diamond$

练习8.13 设 $\psi : \mathbf{R} \to \mathbf{R}$ 是一个非常数、可微、周期函数，其导数在 -1 和 1 之间变化。

(a) 对于正整数 n，定义 $f_n : \mathbf{R} \to \mathbf{R}$ 为 $f_n(x) = \frac{1}{n}\psi(n^2 x)$。求

$$\max_{x \in \mathbf{R}} |f_n(x)| \quad \text{and} \quad \max_{x \in \mathbf{R}} |f_n'(x)|.$$

(b) 给定一个可微函数 f 的例子：$\mathbf{R} \to \mathbf{R}$，使得极限 $f, +\infty$ 存在，但极限 $f', +\infty$ 不存在。

部分（a）旨在提出构建 f 的想法，尽管当然你的（b）部分的答案不应依赖于 n。$\diamond$

练习8.14 设$\psi : \mathbf{R} \to \mathbf{R}$如练习8.13中所述，并定义

$$f(x) = \begin{cases} x^2\psi(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

我们在例8.15中找到了f的导数。证明f'在0处不连续。回顾第4章可能会有帮助。$\diamond$

练习8.15 设f如练习8.14中所述，且$g(x) = f(x) + (x/2)$。证明$g'(0) > 0$，但不存在一个关于0的开区间，使得g是递增的。为什么这不会与“具有正导数的函数是递增的”这一事实相矛盾？$\diamond$

第九章

平均值定理

上一章的结果主要依赖于一个函数在单一点可导，因此是点值或至多局部的性质。在本章中，我们将可导性的机制与第5章关于连续性的全局定理联系起来。在适当的假设下，平均值定理将区间 $[a,b]$ 上 f 的平均变化率与 f 在 (a,b) 的某一点的瞬时变化率等同起来。这个结果使我们能够在点值信息和全局信息之间进行转换，并被认为是微积分的技术基石。

9.1 均值定理

第五章中，我们假设 f 在 $[a,b]$ 上是连续的，并推导了 f 的全局性质。这里，我们进一步假设 $f:[a,b]\to\mathbf{R}$ 在开区间 (a,b) 上可微。（等价地，如果您对冗余感到不安，f 在 (a,b) 上可微，并在 a 和 b 处连续。）一个典型的例子是函数 $f(x)=\sqrt{1-x^2}$ 在 $[-1,1]$ 上，其图形是平面中单位圆的上半部分。

定理9.1. 设$f:[a,b]\to$ $\mathbf{R}$是在(a,b)上连续且可微的函数。那么存在一个$x_0\in(a,b)$使得

$$f'(x_0)=\frac{f(b)-f(a)}{b-a}.$$

在文字上，存在一个点（a,b）在该点处，f的瞬时变化率等于f在区间$[a,b]$上的平均变化率。图9.1以简单（但具有代表性）的方式描述了这一结论。

情况。当用速度和距离来表述时，结论相当合理：如果在一次汽车旅行中，你在某个一小时的时间内覆盖了60英里，那么在那个小时内的某个时刻，你的速度必须是每小时60英里。当然，这并不能证明什么，因为实际距离和速度并不完全对应于实数和函数，但这是一种记住定理结论的好方法。

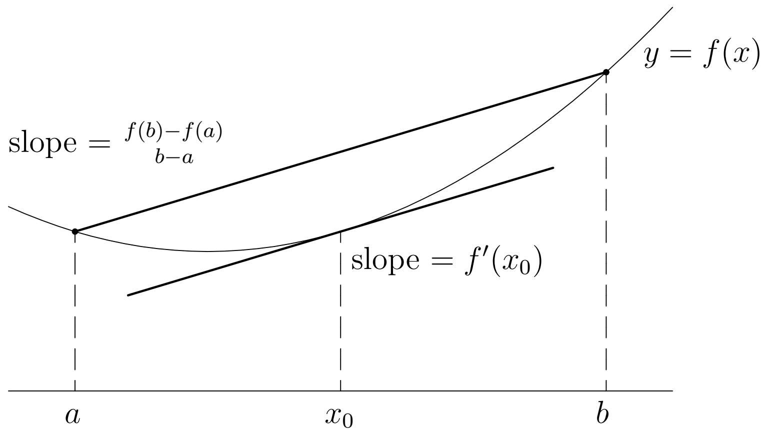


图9.1：平均值定理的结论。

证明。平均值定理的证明在概念上分为两步。第一步，称为罗尔定理，处理函数在端点值相等的情况，并使用极值定理和定理8.12。第二步通过代数技巧将定理简化为罗尔定理。

假设首先 $f(a) = f(b)$，因此平均变化率为0。我们希望证明 f 有一个临界点。根据极值定理，存在点 $x_{\min}$ 和 $x_{\max} \in [a, b]$，使得

$$f(x_{\min}) \leq f(x) \leq f(x_{\max}) \quad \text{for all } x \in [a, b].$$

(这些点在一般情况下不是唯一的.) 首先假设 $x_{\min}$ 和 $x_{\max}$ 中至少有一个在 (a, b) 中，并将其称为 x_0。根据定理 8.12，$f'(x_0) = 0$，我们就可以完成了。唯一的其他可能性是每个点 $x_{\min}$ 和 $x_{\max}$ 都是 $[a, b]$ 的端点。但由于 $f(a) = f(b)$，这意味着 f 是一个常数函数，然后对于每个点 $x_0 \in (a, b)$，$f'(x_0) = 0$。这证明了罗尔定理。

接下来考虑函数 $g:[a,b]\to\mathbf{R}$，通过“线性调整”f的端点值使其相等来定义：

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

函数 g 在 $[a,b]$ 上连续，在 (a,b) 上可导，作为具有这些性质的函数之和。直接计算表明 $g(a)=f(a)=g(b)$，因此 g 满足罗尔定理的假设。因此存在一个 $x_0\in(a,b)$ 使得 $g'(x_0)=0$。但是，再次直接计算给出

$$g'(x)=f'(x)-\frac{f(b)-f(a)}{b-a}\quad\text{for all } x\in(a,b),$$

并且定理随之而来。 □

我们现在可以推导出一些简单但重要的结论。你应该记住，没有平均值定理，以下定理的证明是多么困难。

9.2 同构定理

一个常数函数的导数恒为零。反之，一个导数在每一点都消失的函数似乎一定是常数函数。如果定义域是实数的一个区间，这确实是正确的。然而，你应该注意，没有平均值定理，证明是不可能的。

定理9.2. 设f和g是区间I上的可微函数。如果$f'=g'$，则存在一个实数c使得$f(x)=g(x)+c$对所有$x\in I$成立。

它无论如何都是至关重要的，该域必须是实数区间。考虑定义在所有 $x\neq 0$ 上的函数 sgn $x=x/|x|$，其导数在域上恒为零，但 sgn 不是常数。

证明。通过考虑可微函数 {v*}，只需证明对于所有 {v*}，如果 $h'(x)=$ 不为 0，那么 h 是一个常数函数。我们证明逆否命题：如果 h 在一个区间上不是常数，那么 h' 不恒为零。

假设 h 是 I 上的非常数函数，并选取 a，$b \in I$ 使得 $h(a) \neq h(b)$。根据平均值定理，存在一个 x_0 在 a 和 b 之间，使得

$$h'(x_0) = \frac{h(b) - h(a)}{b - a}.$$

右侧非零，因此 h' 不全为零。 □

身份定理最常用于证明两个函数相等。如果 f 和 g 是在某区间上具有相同导数的可微函数，那么 f 和 g 在该区间上仅相差一个常数。如果，此外，对于某些 x_0，$f(x_0) = g(x_0)$，那么 f 和 g 在该区间上相等。我们将充分利用这项技术，因为我们有许多有趣的方法来获取具有相同导数的函数对。

单调性与导数的符号

恒等定理告诉我们当函数的导数消失时应该期待什么。当导数处处为正或处处为负时，可以得出有趣的结论。你应该将以下结果与第8.2节的观察结果以及练习8.15进行比较。

定理9.3。如果 $f : (a, b) \to \mathbf{R}$ 可微，并且对于所有 $x \in (a, b)$，有 $f'(x) > 0$，那么 f 在 (a, b) 上是递增的。

证明。在定理的假设下，如果 $a < x < y < b$，则存在一个 $x_0 \in (x, y)$ 使得

$$\frac{f(y) - f(x)}{y - x} = f'(x_0).$$

根据假设，$f'(x_0) > 0$，并且由于 $x < y$，因此 $f(x) < f(y)$。一个完全类似的论据表明，如果对于某个区间内的所有 x，$f'(x) < 0$，那么 f 在该区间上是递减的。 □

指数函数

为了展示刚刚证明的定理的力量，这里有一个简短的旁白，展示了如何在没有具体函数表示的情况下研究函数的性质。

方程 $f' = f$ 是一个普通微分方程，或简称为 ODE 的例子。未知数 f 是一个可微函数，其定义域是某个未指定的开区间。这个微分方程是否有任何“有趣”的解（零函数是一个“无趣”的解）并不明显，如果有，它有多少个解。然而，我们可以确定微分方程的后果，这些后果告诉我们可能存在的任何解的性质。

在示例9.14中，我们将证明存在一个非零的、可微的函数exp: $\mathbf{R} \to \mathbf{R}$，即自然指数函数，使得$\exp' = \exp$和$\exp(0) = 1$。在本节的其余部分，我们假设exp的存在。从逻辑上讲，没有问题，因为我们不是在推导exp的存在，而是在推导exp必须具备的性质。我们对exp的了解源于它解决了微分方程$f' = f$，正如我们现在将看到的。

命题9.4。设$f : \mathbf{R} \to \mathbf{R}$满足$f' = f$。则对于所有$x \in \mathbf{R}$，有$f(x) = f(0)\exp(x)$。特别是，如果$f' = f$和$f(0) = 1$，则$f = \exp$。

证明。因为exp可微且处处非零，函数{v*}exp是可微的。商法则表明

$$q' = \frac{\exp\ f' - f\ \exp'}{\exp^2} = \frac{f' - f}{\exp},$$

由于 $f' = f$，它相同地消失。根据恒等定理，q 在 $\mathbf{R}$ 上是一个常数函数，并在 0 处求值显示

$$\frac{f(x)}{\exp(x)} = q(x) = q(0) = \frac{f(0)}{\exp(0)} = f(0)$$

对于所有 x，因此 $f = f(0)$ 指数如所声称。 □

命题9.5。设k为一个实数。如果$f : \mathbf{R} \to \mathbf{R}$是一个满足$f' = kf$和$f(0) = 1$的可微函数，那么对于所有$x$，$y \in \mathbf{R}$有$f(x + y) = f(x)f(y)$

证明。函数 $g(x) = \exp(kx)$ 通过链式法则求解 ODE $g' = kg$，并满足初始条件 $g(0) = 1$。命题 9.4 的证明表明这是唯一这样的函数。因此，只需证明

$$\exp(x + y) = \exp(x)\exp(y) \qquad \text{for all } x,\, y \in \mathbf{R}.$$

修复 y，并考虑由 $f(x) = \exp(x+y)$ 定义的函数 $f: \mathbf{R} \to \mathbf{R}$。根据链式法则，$f$ 可导，并且 $f' = f$。由于 $f(0) = \exp(y)$，命题 9.4 意味着对于所有 x，有 $\exp(x+y) = \exp(x)\exp(y)$。 $\square$

特殊情况 $\exp(x)\exp(-x) = \exp(x-x) = \exp(0) = 1$ 表示 exp 在任何地方都不为零。由连续性，对于所有实数 x，有 $\exp(x) > 0$。从定义性质 $\exp' = \exp$，我们得出 exp 是一个增函数。数 $e := \exp(1) > \exp(0) = 1$ 是数学的一个基本常数。尽管我们现在对 e 的数值知之甚少，但我们能证明“指数函数”这个名称是合理的。

推论9.6. 对于所有有理数 r，有 $\exp(r) = e^r$。

证明。通过归纳，对于所有 $p \in \mathbf{N}$，有 $\exp(p) = e^p$。如上所述，命题9.5意味着对于所有 p，有 $\exp(p)\exp(-p) = 1$，因此 $\exp(-p) = 1/\exp(p) = e^{-p}$。最后，如果 $q \in \mathbf{N}$，那么

$$\begin{aligned}\big[\exp(p/q)\big]^q &= \exp(p/q) \overset{q}{\cdots} \exp(p/q)\\ &= \exp(p/q + \overset{q}{\cdots} + p/q) = \exp(p) = e^p,\end{aligned}$$

因此 $\exp(p/q) = \sqrt[q]{e^p} = e^{p/q}$。 $\square$

在推论的基础上，为所有 $x \in \mathbf{R}$ 合理地定义 $e^x = \exp(x)$。命题9.5是熟悉的法则

$$e^{x+y} = e^x e^y \qquad \text{for all } x, y \in \mathbf{R}.$$

记住，我们尚未证明exp存在，但我们已经推导出它必须具有的一些属性，仅假设$\exp' = \exp$和$\exp(0) = 1$。

导数的介值性质

设 $f: I \to \mathbf{R}$ 是一个在区间 I 上的可微函数。Darboux[1] 的一个定理断言 f' 具有介值性质；特别是，导数的间断必须是野的。例 8.15 中的函数并不特别病态!

定理9.7。设f在包含$[a, b]$的某个开区间上可微。如果c是$f'(a)$和$f'(b)$之间的实数，那么存在$x_0 \in (a, b)$使得$f'(x_0) = c$。

1dar BOO

证明。不失一般性地假设 $f'(a) < f'(b)$，并考虑由 $g(x) = f(x) - cx$ 定义的可微函数 g。由于 g 在 $[a, b]$ 上连续，g 在某点 $x_0 \in [a, b]$ 处达到其最小值。然而，$g'(a) = f'(a) - c < 0$，所以 g 在 a 处是递减的。这意味着 g 的最小值不是在 a 处达到的。同样地，$g'(b) = f'(b) - c > 0$，所以 g 在 b 处是递增的，这意味着 b 不是 g 的最小值。因此，g 的最小值必须在某点 $x_0 \in (a, b)$ 处达到，并且根据定理 8.12 $g'(x_0) = 0$，或 $f'(x_0) = c$。 □

推论9.8. 设$f : I \to \mathbf{R}$是包含$[a, b]$的区间上的可微函数。如果f'在开区间(a, b)上非零，则f在闭区间$[a, b]$上是严格单调的——因此是可逆的。

证明。假设不失一般性，对于某个 $x \in (a, b)$，有 $f'(x) > 0$。Darboux定理表明 f' 在区间内处处为正，因为如果 f' 在某处为负，那么它必须在某处消失。定理9.3表明 f 在开区间 (a, b) 上严格递增。最后，如果 $x \in (a, b)$，那么

$$\frac{f(x) - f(a)}{x - a} > 0$$

正如在定理9.2的证明中一样，这暗示了$f(x) > f(a)$。同样地，$f(x) < f(b)$。 □

推论9.8给出了一个函数可逆性的充分条件。对于有理函数，这个条件通常非常容易检查，实际上通常是证明函数在某个区间上可逆的最简单方法。

示例9.9 如果 $f(x) = x - x^3/3$ 对于 $x \in \mathbf{R}$，那么 $f'(x) = 1 - x^2$，且临界点是 -1 和 1。导数——一个多项式函数——是连续的，因此推论9.8意味着 f 在每个区间 $(-\infty, -1]$，$[-1, 1]$，和 $[1, \infty$）上是单射的。实际上，$f'(x) > 0$ 当且仅当 $|x| < 1$，因此 f 在每个无界区间上是递减的，并在 $[-1, 1]$ 上是递增的。

这些区间共享端点，但不存在矛盾，如图9.2所示应很清楚。□

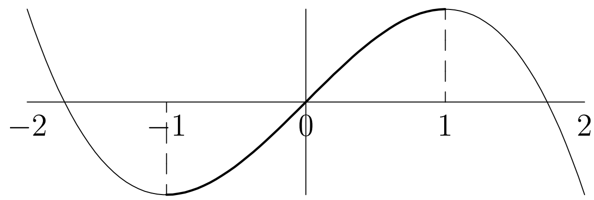


图9.2：多项式单调的区间。

修复

在应用中，可能希望通过给出在相邻区间上成立的两个或更多公式来表示一个函数，例如，

$$\psi(x) = 4x(1-|x|) = \begin{cases} 4x(1+x) & \text{if } -1 \le x \le 0 \\ 4x(1-x) & \text{if } 0 < x \le 1 \end{cases} \tag{9.1}$$

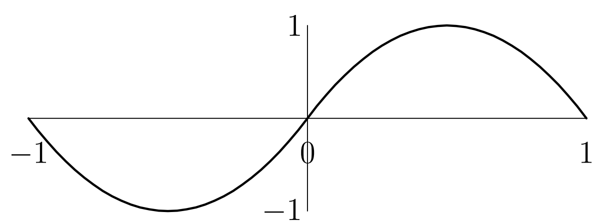


图9.3：函数 ψ。

我们想知道这样一个“修补”函数在公式“连接”的点（处）是否可微。下一个定理给出了一个对于许多应用来说是充分的充分条件。

定理9.10。设f是一个在x_0处连续且在关于x_0的某个删除区间内可微的函数。如果$\lim(f', x_0)$存在且等于ℓ，那么f在x_0处可微，并且$f'(x_0) = \ell$。

特别地，f' 在 x_0 处是连续函数。尽管这个结果看起来合理，但证明需要用到平均值定理。尝试一个“天真”的证明是有教育意义的；问题是 $\lim(f', x_0)$ 涉及双重极限，定理将其互换：

$$\lim_{x\to x_0}\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \stackrel{?}{=} \lim_{h\to 0}\lim_{x\to x_0}\frac{f(x+h)-f(x)}{h}$$

证明。根据假设，存在一个 $\delta > 0$，使得在闭区间 $[x_0-\delta, x_0+\delta]$ 上，函数 f 连续，且在 x_0 处可能不可导。特别是，f 满足每个区间 $[x_0-\delta, x_0]$ 和 $[x_0, x_0+\delta]$ 上的平均值定理的假设。对于每个满足 $0 < |h| < \delta$ 的数 h，存在一个 $x_h \in (x_0, x_0+h)$ 使得

$$f'(x_h) = \frac{f(x_0+h)-f(x_0)}{h}.$$

通过构造，$x_h \to x_0$作为$h \to 0$；取极限得到$\lim(f', x_0) = f'(x_0)$，正如所声称的。 □

编写最后一步的正式 ε-δ 证明是一个很好的练习。

示例9.11设$\psi : \mathbf{R} \to \mathbf{R}$为2周期的函数，其限制在$[-1, 1]$上由(9.1)给出。对于$x \in (0, 1)$，我们有$\psi'(x) = 4 - 8x$，而在$(-1, 0)$上我们有$\psi'(x) = 4 + 8x$。使用定理9.10，我们发现$\psi$在0处可微，且$\psi'(0) = 4$。同样，也利用周期性，我们发现$\psi$在$\pm1$处可微，且$\psi'(\pm1) = -4$。由于$\psi$在整个周期上可微，它在$\mathbf{R}$上也可微。特别是，我们构造了一个非常数、周期性的、属于$\mathcal{C}^1$类的函数。

我们称 ψ 为伪正弦函数。 □

9.3 逆函数的可微性

微积分提供了一个有效的工具，即导数的符号，用于确定函数是否单调。对于定义域为区间的可微函数，单调性与可逆性等价。我们现在转向探讨一个逆函数本身是否可微的问题，以及如果可以，如何计算导数。

设 $f : I \to \mathbf{R}$ 为定义在开区间上的单射函数，其像为 $J = f(I)$。存在一个函数 $g : J \to I$ 使得

$$(9.2)\qquad \begin{aligned} g\big(f(x)\big) &= x \qquad \text{for all } x \in I, \\ f\big(g(y)\big) &= y \qquad \text{for all } y \in J. \end{aligned}$$

换句话说，对于 $x \in I$，方程 $y = f(x)$ 和 $x = g(y)$ 是等价的。如果需要，可以将 f 替换为 $-f$，我们也可以假设 f 是递增的。

定理9.12。设$f : I \to \mathbf{R}$在区间I上是一一对应且可微的，并设$x_0 \in I$。函数$g = f^{-1}$在$y_0 = f(x_0)$处可微当且仅当$f'(x_0) \neq \neq 0$，在这种情况下

$$g'(y_0) = \frac{1}{f'(x_0)}.$$

证明。首先假设 $g = f^{-1}$ 在 $y_0 = f(x_0)$ 处可微。将链式法则应用于 (9.2) 的第一个式子，得到 $1 = g'(y_0)f'(x_0)$。我们推断出

$$g'(y_0) = g'\big(f(x_0)\big) = \frac{1}{f'(x_0)} \qquad \text{if } f'(x_0) \neq 0,$$

如果 $f'(x_0) =$ 为 0，则 g 在 y_0 处不可微。因此，我们知道 f^{-1} 的导数必须是多少，前提是导数存在。这个方程的莱布尼茨版本是看起来很自然的方程

$$1 = \frac{dy}{dx} \cdot \frac{dx}{dy},$$

通常情况下，x不相同，y也不相同。

为了证明当 $f' \neq \neq 0$ 时，g 确实是可微的，观察到一个 $\eta > \neq 0$，使得 $(x_0 - \eta, x_0 + \eta) \subset I$，并且对于 $|h| < \eta$，我们可以写成 $f(x_0 + h) - f(x_0) = k$，其中 k 由 h 唯一确定，见图 9.4。将此重写为 $f(x_0 + h) = f(x_0) + k = y_0 + k$ 并将 g 应用于两边，

$$\frac{g(y_0 + k) - g(y_0)}{k} = \frac{h}{f(x_0 + h) - f(x_0)}.$$

如果 $f'(x_0) > 0$，则右侧当 $h \to 0$ 时有极限，即 $1/f'(x_0)$，而左侧是 $g'(y_0)$ 的牛顿商。这证明了如果 $f'(x_0) \neq 0$，则 $g = f^{-1}$ 在 $f(x_0) = y_0$ 处可微。

$\square$

示例9.13设q为一个正整数，且对于$f(x) = x^q$，$x >$为0。函数f是递增的（因此可逆）且可微，其导数为$f'(x) = qx^{q-1} >$0。逆函数是q次根函数，$g(y) = y^{1/q}$。根据定理9.12，g在$y = x^q$处可微，并且

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{qx^{q-1}} = \frac{1}{q} y^{(1/q)-1}.$$

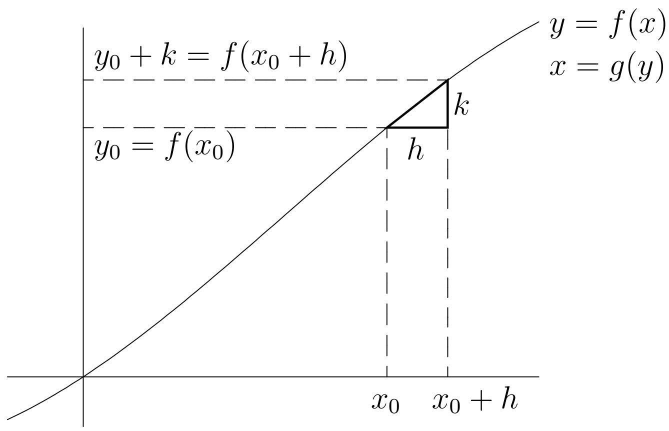


图9.4：逆函数的差商。

练习9.4将此结果扩展到具有任意有理指数的幂函数。□

示例9.14自然对数（练习7.17）定义为

$$\log x = \int_1^x \frac{1}{t}\,dt, \qquad t > 0.$$

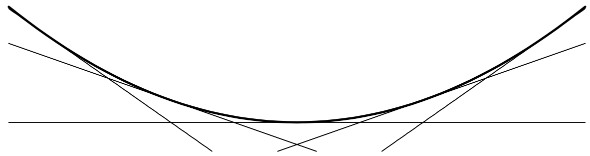
图像的对数是 **R**。定理8.9表明对数是可微的，并且对于 $x > 0$，有 lo g$'$ $x = 1/x$。推论9.8表明对数是递增的，因此是可逆的，这也是我们从练习7.17中知道的事实。令 exp : **R** $\to (0, \infty)$ 表示其逆函数。显然 l og 1 $= 0$，所以 exp(0) $= 1$，并且对于每个 $x > 0$，我们有

$$\exp'(x) = \frac{1}{\log'[\exp(x)]} = \frac{1}{1/\exp(x)} = \exp(x).$$

因此，正如我们之前所声称的，存在一个在 **R** 上等于其自身的导数的可微函数，并在 0 处取值为 1。□

9.4 二阶导数与凸性

让 $f\colon [a, b] \to$ **R** 连续，并假设 f 在 (a, b) 上是二阶可微的。值 $f'(x)$ 可以解释为 f 在 x 处切线图形的斜率，而 $f''(x)$ 是斜率随 x 变化的瞬时变化率。如果 $f'' >$ 在某个区间上为 0，几何直觉表明 f 的图形应该是“凸”或“向上凹”：



本节的目标是精确定义“凸性”，并证明具有正二阶导数的 $\mathscr{C}^2$ 函数在此意义上是凸的。

一个集合 $R \subset \mathbf{R}^2$ 被称为凸集，如果对于 R 中的所有点 p_1 和 p_2，连接 p_1 和 p_2 的线段完全位于 R 内。为了用代数方式表达这个标准，请注意，连接 $p_1 = (x_1, y_1)$ 和 $p_2 = (x_2, y_2)$ 的线段是形式为的点的集合

$$(1-t)p_1 + tp_2 = \big((1-t)x_1 + tx_2, (1-t)y_1 + ty_2\big) \qquad \text{for } t \in [0,1].$$

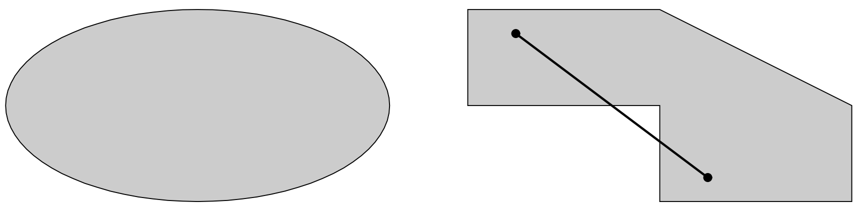


图9.5：凸集和非凸集。

凸函数

定义9.15 设$I \subseteq \mathbf{R}$为一个区间。若函数$f\colon I \to \mathbf{R}$满足“位于图形上方的区域”，则称其为凸函数。

$$\Gamma_f^+ := \{(x,y) \mid x \in I,\ f(x) \le y\},$$

平面上的凸集。

一个函数 f 是凹的，如果 $-f$ 是凸的，即如果“图形下方的区域”，

$$\Gamma_f^- := \{(x,y) \mid x \in I,\ y \le f(x)\},$$

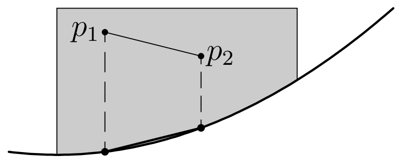
在平面上是一个凸集。术语“凹向上”和“凹向下”通常具有相同的（相应的）含义。

要测试 f 图像上方的区域是否凸，只需选择点 p_i 位于图像上即可：

命题 9.16. 函数 $f\colon I \to \mathbf{R}$ 是凸的，当且仅当每条割线都包含在 Γ_f^+ 中。

证明。“仅当”方向是显然的。

对于“如果”蕴含，想法是如果 p_1 和 p_2 是 R 的点，那么连接它们的线段位于由垂直投影得到的割线之上：



但是，根据假设，割线包含在 R 中，因此连接 p_1 和 p_2 的线段也包含在 R 中。一个代数证明留给你作为翻译练习。

□

正如一阶导数的符号通过平均值定理与单调性相关联一样，二阶导数的符号与凸性相关联，平均值定理提供了无穷小信息（f''）与有限差分（凸性）之间的联系。

引理9.17. 假设f在(a, b)上二阶可微，对于某个$x_1 < x_2$，有$f(x_1) = f(x_2) = 0$，并且$f'' \geq 0$在$[x_1, x_2]$上。那么$f \leq 0$在$[x_1, x_2]$上。

证明。与我们的单调性结果一样，逆否命题更容易证明：如果存在一个 x 在 (x_1, x_2) 中，使得 $f(x) > 0$，那么对于 (x_1, x_2) 中的某个 z，有 $f''(z) < 0$，图9.6。

应用平均值定理于区间 $[x_1, x]$ 上的 f，我们得出存在 $z_1 \in (x_1, x)$ 使得

$$f'(z_1) = \frac{f(x) - f(x_1)}{x - x_1} = \frac{f(x)}{x - x_1} > 0.$$

同样，存在一个点 $z_2 \in [x, x_2]$，其 $f'(z_2) <$ 为 0。

现在，应用应用平均值定理于 $[z_1, z_2]$ 上的 f'，我们得到 d 那个

$$f''(z) = \frac{f'(z_2) - f'(z_1)}{z_2 - z_1} < 0$$

对于某些 $z \in (z_1, z_2)$。 □

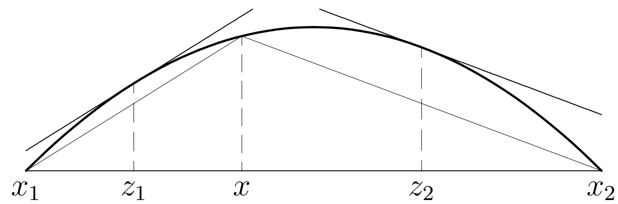


图9.6：根据f的值确定f''的符号。

显然存在一个逆序不等式的9.17引理版本。（"如果 $f'' \leq 0$，那么 $f \geq 0$。"）此外，如果 f'' 是连续的，并且在假设中存在严格不等式（即对于某些 x，$f''(x) < 0$），那么我们将在结论中得到严格不等式，因为一个在一点上为正的连续函数在区间上也是正的。

使用引理9.17，我们可以用二阶导数来描述$\mathscr{C}^2$函数的凸性。

定理9.18。假设f是区间I上的$\mathscr{C}^2$。对于每个闭区间$[a, b] \subset I$，f在$[a, b]$上是凸的当且仅当$f'' \geq 0$在(a, b)上。

证明。根据命题9.16，只需证明以下条件等价：

- $f'' \geq 0$ 在 I。
- "f的图像位于每条割线之下。"形式上，对于所有a，b属于I，连接$p_1 = \big(a, f(a)\big)$和$p_2 = \big(b, f(b)\big)$的线段位于Γ_f^+内。

假设第一个条件成立。固定元素 $a < b$ 在 I 中，令 ℓ 为其图是割线的线性多项式，并引入 $\mathscr{C}^2$ 函数 $g = f - \ell$。直接计算表明 $g'' = f''$，并且 $g(a) = g(b) = 0$。引理9.17表明 $g \leq 0$ 在 $[a, b]$ 上。但这意味着 $f \leq \ell$ 在 $[a, b]$ 上，这正是我们要证明的。注意，我们使用了二阶可微性，但没有使用 f'' 的连续性。

相反，假设上述第一个条件失败：存在 z 在 (a, b) 中，使得 $f''(z) <= 0$。由于 f'' 的连续性，存在一个 $\delta >> 0$，使得 $f'' <= 0$ 在区间 $[z - \delta, z + \delta]$ 上。像之前一样，让 ℓ 为割线，并让 $g = f - \ell$。上述提到的引理9.17（见上文）的一个明显修改表明，$g >$ 在 $[z - \delta, z + \delta] \subset (a, b)$ 的某一点上为0。这意味着第二个条件也失败了：存在一条割线不包含在 Γ_f^+ 中。 □

证明给出了比定理中所述更具体的信息。例如，如果对于某些 x ，$f''(x) > 0$，那么关于 x 的某个区间内的所有割线都位于 f 的图形“严格上方”。所有这些陈述对于二阶导数非正的函数都有明显的修改。注意，一个函数可以在不是两次可微的情况下是凸的。绝对值函数在 $\mathbf{R}$ 上是凸的，但甚至不是一次可微的，而练习9.15表明，即使是断续的函数也可以是凸的。

派生与绘图

在整个本节中，我们假设f是一个定义域包含$[a, b]$的实数区间上的$\mathscr{C}^2$函数。一阶和二阶导数可以用来获取关于函数图形的几何信息。图形计算器导致了手动绘图技术的淡化，但技术知识仍然有用，尤其是在计算机处理不佳的病态函数方面。

通过Darboux定理和平均值定理，如果 $f' \neq$ 在 (a, b) 中为0，则 f 在 $[a, b]$ 上是单调的。如果 f 只有有限个临界点，可以通过绘制每个临界点 x 的点 $\big(x, f(x)\big)$ 来绘制一个非常粗糙的图形，然后“连接点”；这样的图形包含关于 f 的单调性信息。

关于图凸性的信息通过计算 f'' 并确定符号来找到：当 $f'' >$ > 0 时，图是凸的，当 $f'' <$ < 0 时，图是凹的。图凸性发生变化的点是几何上有趣的:

定义9.19 设f在x_0处连续，在x_0的去除邻域内二阶可导。如果点$\big(x_0, f(x_0)\big)$是拐点或折点[2]，则满足

- 极限(f', x_0^-) = 极限(f', x_0^+) (可能为 $+\infty$ 或 $-\infty$),
- f'' 变化符号在 x_0 处，且 f'' 在关于 x_0 的某个已删除区间上非零。

几何上，图在每个以 x_0 为中心的邻域内都有切线，且在 x_0 处凸性发生变化，因此图呈“S”形。根据达布定理，一个拐点对应于 f'' 的零点

[2] In mathematics, “flex” is a noun.

或是一个f''不存在的点。然而，并非每个这样的点都是柔点。

使用临界点作为支架并利用凸性信息来完善图形通常可以得到一个准确的图形。如果需要更多定量信息，可以通过直接计算绘制几个点。如果方程 $f(x) = 0$ 可以精确求解，则应绘制这些点。

示例9.20 假设我们想要绘制 $f(x) = x^{2/3}(x^3 - 1)$ 的图形。首先我们展开并求导：

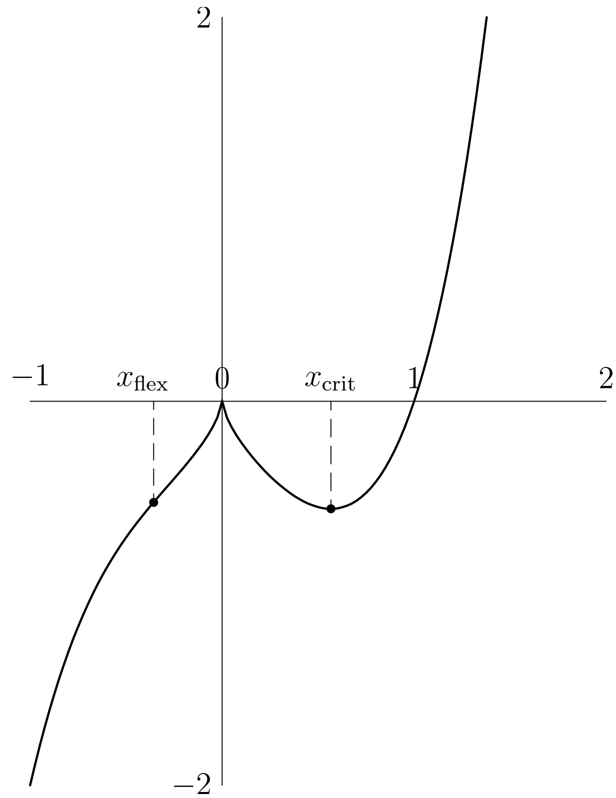
$$f'(x) = \tfrac{11}{3}x^{8/3} - \tfrac{2}{3}x^{-1/3} = \tfrac{1}{3}x^{-1/3}(11x^3 - 2)$$
$$f''(x) = \tfrac{88}{9}x^{5/3} + \tfrac{2}{9}x^{-4/3} = \tfrac{2}{9}x^{-4/3}(44x^3 + 2)$$

注意，展开后更容易求导，但因式分解后更容易找到临界点。在实践中，你可能想要计算前两个导数的展开和因式分解形式。

存在一个实临界点，$x_{\text{crit}} = \sqrt[3]{2/11}$，且$f'(0)$不存在。由于$\sqrt[3]{x}$的符号与$x$的符号相同，而在0附近$(11x - 2) < 0$，我们发现$\lim(f', 0^-) = +\infty$和$\lim(f', 0^+) = -\infty$。因此，当$x < 0$或当$x > \sqrt[3]{2/11}$时，$f' > 0$，且当$0 < x < \sqrt[3]{2/11}$时，$f' < 0$。

二阶导数在 $x_{\text{flex}} = -\sqrt[3]{1/22}$ 处消失，在 0 处未定义。由于 $x^{4/3} \geq 0$ 对于所有实数 x 都成立，我们看出 $\lim(f'', 0) = +\infty$。因此，有一个拐点，因为 f'' 在括号内的项为 0 时改变符号。图形在 $x < -\sqrt[3]{1/22}$ 时是凹的，在 $x > -\sqrt[3]{1/22}$ 时是凸的。

方程 $f(x) = 0$ 有两个实数解，$x = 0$ 和 $x = 1$。我们绘制这些点，以及临界点和拐点，然后利用单调性和凸性信息绘制曲线。



上图显示了真实比例的图表。 □

加速度

如果一个函数表示位置随时间变化，那么一阶导数代表速度的物理概念，二阶导数代表加速度。

这是一个值得注意的物理事实，加速度可以通过“局部实验”来测量，这些实验不涉及观察整个宇宙。想象你在一架飞机上，以恒定速度和高度飞行。除了引擎的噪音外，你没有物理证据表明你“在运动”。如果你把球扔过过道，它看起来就像飞机停在地面上一样移动。从罐子里倒出的液体将以预期的方式落入杯子中。用秤称重的你的体重与在地面上时相同。相比之下，如果飞机突然俯冲、上升或急剧转弯，将会发生可观察到的变化：球将沿着奇怪的路径移动，可能偏向一侧；饮料可能错过杯子，或者（如果飞机进入“自由落体”）

形成一个球并静止悬挂；你的体重可能会增加或减少，当你站在秤上时。这些影响对战斗机飞行员和宇航员来说很大，在极端情况下可能导致腿部和脚部血液积聚而出现昏厥。

航空公司建议乘客在整个飞行过程中保持安全带系紧，因为已经有过罕见的飞机急剧下降几秒钟的意外事故，在客舱中短暂地“逆转重力”，导致未系安全带的乘客从座位上向上跌落，并在“着陆”在头顶的行李舱时严重受伤。如果你曾经飞行过，你几乎肯定体验过“湍流”或“空气囊”的感觉。你的内耳是一个极其敏感的平衡仪；当飞机直线、水平且以恒定速度飞行时，你的内耳无法告诉你正在运动，但如果飞机加速，你的内耳会记录这种变化，通常表现为跌落或恶心的感觉。NASA使用特别改装的波音707在接近失重的环境中训练宇航员。这架飞机（显然是因为这个原因被昵称为“呕吐彗星”）上升到大约4万英尺的高度，然后进入一个抛物线弧，与自由落体的轨迹相匹配。在这条路径上飞行大约30秒后，飞行员必须拉起飞机的机头，这（顺便说一下）使飞机在几秒钟内体验到“大于正常重力”。

早在19世纪初，一些科学家认为以超过约15英里/小时的速度旅行会导致严重的身体伤害。蒸汽机车彻底驳斥了这种观点，我们现在知道，速度本身并不仅无害，而且在绝对意义上是物理上无意义的。相比之下，加速度确实具有绝对意义，并表现为一种力（在物理学意义上的力）。“加速度可以感觉到，但匀速运动却不能”这一事实是谚语“不是跌倒杀死你，而是突然停止在最后”的基础。

9.5 不定极限

存在一个强大的计算工具，洛必达法则，它利用导数的机制来评估不定极限。正如本章许多结果的情况一样，对于该结果有一个令人信服的莱布尼茨符号解释，但实际的证明依赖于一个看似无关的技术结果。

柯西均值定理

定理9.21。设f，g：$[a,b]\to\mathbf{R}$在(a,b)上连续且可微，存在一点$x_0\in(a,b)$使得

$$f'(x_0)\big(g(b)-g(a)\big)=g'(x_0)\big(f(b)-f(a)\big).$$

证明。普通平均值定理是特殊情况 $g=\mathrm{id}$，即恒等函数，柯西版本通过类似技巧得到证明。定义 $h:[a,b]\to\mathbf{R}$ 为

$$h(x)=f(x)\big(g(b)-g(a)\big)-g(x)\big(f(b)-f(a)\big).$$

它是立即的，h 满足罗尔定理的假设，因此存在一个 $x_0\in(a,b)$，使得 $h'(x_0)=0$，正如所声称的。 □

拉普拉斯法则

柯西平均值定理的推论，即微积分学生所熟知的洛必达法则，实际上是由约翰·伯努利证明的。洛必达侯爵是伯努利的富裕赞助人，但却是平庸的数学家。有时人们开玩笑说，洛必达法则是用金钱可以买到的最好的定理。

定理9.22。假设f和g在c的某个删除邻域内可微，且$\lim(f,c)=\lim(g,c)=0$，g'在c的某个删除区间内非零，并且$\lim(f'/g',c)$存在且等于ℓ。那么$\lim(f/g,c)$存在且等于ℓ。

在英语中：当尝试求一个商的极限时，如果答案形式上是0/0，那么对分子和分母进行求导（不要与商法则混淆！）并再次尝试求值。如果极限是ℓ，那么原始极限也是ℓ。

证明。根据假设，f'/g' 在关于 c 的某个删除区间上定义，因此存在一个 $\delta>0$，使得如果 $0<|x-c|<\delta$，则存在 $f'(x)$ 和 $g'(x)$，且 $g'(x)\neq 0$。如果需要，重新定义 f 和 g 在 c 处为 0，使得 f 和 g 在 ($c-\delta,c+\delta$) 上连续。

在删除的 δ-区间关于 c，分母 g 不为零，否则 g' 在某处将为零，罗尔定理。因此 f/g 在删除的区间 $N_\delta^\times(c)$ 上定义。根据柯西中值定理，存在一个点 $x_0\in(c,c+\delta)$，使得

$$f'(x_0)\big(g(c+\delta)-g(c)\big)=g'(x_0)\big(f(c+\delta)-f(c)\big),$$

或者，因为 $f(c) = g(c) = 0$ 和 $g(c+\delta) \neq 0$,

$$\frac{f(c+\delta)}{g(c+\delta)} = \frac{f'(x_0)}{g'(x_0)}.$$

当 $\delta \to 0$ 时取极限完成证明，因为 $x_0 \to c$ 当 $\delta \to 0$。

□

非常重要的是不要在未检查假设的情况下应用公理化；在英语中，如果极限不是形式上的0/0，不要尝试应用洛必达法则。为了了解原因，考虑以下错误的计算

$$\lim_{x\to 0} \frac{1}{x^2} \overset{\text{oops!}}{=} \lim_{x\to 0} \frac{0}{2x} = \lim_{x\to 0} \frac{0}{2} = 0.$$

此外，l'Hôpital法则的逆命题是错误的。如果$f'/g', c$的极限不存在，那么没有获得任何信息；原始极限可能存在也可能不存在。尽管如此，仍然可以说些东西，参见练习9.24。

示例9.23 设 n 为一个正整数。根据洛必达法则，

$$\lim_{x\to 1} \frac{x^n - 1}{x - 1} = \lim_{x\to 1} \frac{nx^{n-1}}{1} = n,$$

根据几何级数公式，练习2.16。根据练习9.4，结论可推广到任意的有理数 r。□

L'Hôpital的法则在事件 $f'(c) = g'(c) = 0$ 时可以反复应用。如果对于某个正整数 k，极限 lim$(f^{(k)}/g^{(k)}, c)$ 存在且等于 ℓ (，并且所有之前应用L'Hôpital法则的结果都是 0/0)，那么原始极限存在且等于 ℓ。

示例 e 9.24 对洛必达法则的一次应用 gi ves us

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

因为exp是它自己的导数。后者仍然是形式上的0/0，所以我们再次应用洛必达法则。通过将$x =$设为0，可以得到结果表达式，我们发现极限是1/2。□

L'Hôpital的法则在 $+\infty$

拉普拉斯法则在 $+\infty$ 的极限处有一个版本，将被反复使用。

定理9.25。设f和g是可微函数。假设g'在某个区间$(R,+\infty)$上非零，并且

$$\lim(f,+\infty)=0=\lim(g,+\infty).$$

如果 $\lim\limits_f'/g',+\infty = L$，那么 $\lim\limits_f/g,+\infty = L$。

注意结论是商 f/g 在 $+\infty$ 处有极限，并且这个极限与 f'/g' 的极限相同。形式上，这等同于有限点的拉格朗日法则。

证明。因为 g' 在某个区间 $(R,+\infty)$ 上非零，g 在 $(R,+\infty)$ 上单调。由于 $\lim(g,+\infty)=0$，因此 g 本身在该区间 $(R,+\infty)$ 上非零，所以商 f/g 在此区间上有定义。因为 f 和 g 在 $+\infty$ 处有有限极限，我们可以假设 f 和 g 在 $(R,+\infty)$ 上有界。

一个基于柯西平均值定理的论证是可能的，但稍微复杂，因为定理的假设不适合使用闭区间和有界区间。相反，我们使用“变量替换”$y=1/x^2$。这种选择是由希望x在0附近的删除区间中，如果y在$+\infty$附近的删除区间中。定义

$$F(x)=f(1/x^2),\quad G(x)=g(1/x^2)\qquad \text{for } x\neq 0.$$

您可以验证 F 和 G 满足定理9.22的假设。根据链式法则，我们有 $F'(x)=-(\text{^}2/x^3)f'(\text{^}1/x^2)$，对于 g 同样适用，所以

$$\frac{F'(x)}{G'(x)}=\frac{f'(1/x^2)}{g'(1/x^2)}.$$

定理9.22意味着$\lim(f/g,+\infty)=\lim(F/G,0)=L$，如所声称。

□

练习

练习9.1 考虑区间 $[-1,1]$ 上的绝对值函数 $f(x)=|x|$。是否存在一个 $x_0\in(-1,1)$ 使得

$$f'(x_0)=\frac{f(1)-f(-1)}{1-(-1)}=0?$$

为什么这不会与平均值定理相矛盾？为定义在 $x \neq 0$ 上的函数 $g(x) = x/|x|$ 提出相同的问题。

练习9.2 设 $f(x) = 1/x$ 对于 $x \neq 0$。证明对于 x 在 f 的定义域中的所有值，有 $f'(x) < 0$。f 是一个递减函数吗？当限制在它的定义域的一个区间时，f 是递减的吗？解释。$\diamond$

练习9.3 设 $f(x) = x^3$ 为 $x \in \mathbf{R}$。证明 f 有一个临界点。f 在 $\mathbf{R}$ 上是否递增？$\diamond$

练习9.4 使用例9.13的结果来证明，如果 $r = p/q$ 是有理数，那么 $\frac{d}{dx}x^r = rx^{r-1}$。$\diamond$

练习9.5 使用链式法则和练习9.4，求以下函数的导数，并绘制图形。确保给出每个函数及其导数的定义域。

(a) $f(x) = \sqrt{1-x^2}$

(b) $f(x) = 1/\sqrt{1-x^2}$

(c) $f(x) = \sqrt{1+x^2}$

(d) $f(x) = x/\sqrt{1+x^2}$

可能有助于找到适当的垂直和/或水平渐近线。$\diamond$

练习9.6 定义 f : [−^2,^2] → $\mathbf{R}$ 如下

$$f(x) = \begin{cases} (1-x^2)^2 & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

证明 f 可微（请注意点 $x = \pm 1$），并在同一组坐标轴上绘制 f 和 f' 的图形。$\diamond$

练习9.7设$\psi : \mathbf{R} \to \mathbf{R}$为示例9.11的伪正弦函数。

(a) 证明 ψ 是奇数，将导数用查理·布朗函数表示，并找到临界点和极值。

(b) 在 $[-3, 3]$ 上绘制 ψ 的图形。注意，(9.1) 仅在 $[-1, 1]$ 上成立。

$\diamond$

练习9.8 设 $f: \mathbf{R} \to \mathbf{R}$ 为 $\mathscr{C}^1$，并假设 f' 是非常数且周期为1的周期函数。

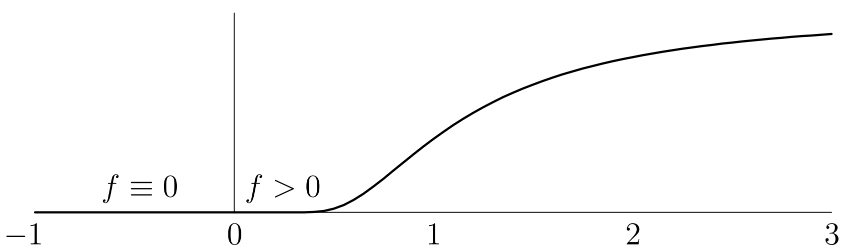
(a) 证明存在一个实数 c，使得对于所有 $x \in \mathbf{R}$，有 $f(x+1) = f(x) + c$。(b) 证明 $g(x) = f(x) - cx$ 是周期为 1 的周期函数。(c) 给出一个非周期函数的例子，其导数是非常数且是周期性的，并绘制 f 的图像。

$\diamond$

练习9.9 绘制方程 Z 的轨迹 $y^2 = x^2 + x^3$，并证明存在一个连续函数 $f: [-1, +\infty) \to \mathbf{R}$，在 $(-1, +\infty)$ 上可微，使得 Z 是 f 和 $-f$ 的图形的并集。（您需要在0处修补两个函数。）$\diamond$

练习9.10 令 $f: \mathbf{R} \to \mathbf{R}$ 定义为

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$



(a) 使用前一个练习来证明，如果 p 是一个多项式，那么

$$\lim_{x \to 0} p\left(\frac{1}{x}\right) f(x) = 0.$$

(b) 使用链式法则计算 $f'(x)$，如果 $x > 0$，并充分论证地评估 lim(f', 0)。使用你的答案证明 f 在 $\mathbf{R}$ 上可微。

(c) 使用对 n 的归纳法证明 f 是 n 次可微的，即使在 0 处。

建议：通过归纳法证明，如果 $x > \neq 0$，则存在某个多项式 p_k，使得 $f^{(k)}(x) = p_k\left(\frac{1}{x}\right) f(x)$。

因此 f 是光滑的 ($\mathscr{C}^\infty$)。 $\diamond$

练习9.11 在练习8.3中，你证明了

$$\sum_{k=1}^{n} kx^k = \frac{nx^{n+2} - (n+1)x^{n+1} + x}{(x-1)^2}$$

$$\sum_{k=1}^{n} k^2x^k = \frac{n^2x^{n+2} - (2n^2+2n-1)x^{n+1} + (n+1)^2x^n - (x+1)}{(x-1)^3}$$

对于 $x \neq 1$. 在 $x = 1$ 处使用洛必达法则评估 $x = 1$ 处左侧的和。 $\diamond$

凸性

练习9.12 设R_1和R_2是平面上的凸集。证明$R_1 \cap R_2$是凸集。 $\diamond$

练习9.13 设a和b为正。集合

$$(*) \qquad \{(x,y) \mid \tfrac{x^2}{a^2} + \tfrac{y^2}{b^2} \leq 1\}$$

这是一个椭圆，参见图9.5。

(a) 将 y 解为 x 的函数，以将椭圆的边界表示为图的并集。

(b) 证明椭圆上半部分的图形对应的函数是凹的，下半部分是凸的。

(© 证明椭圆({v*})在平面中是一个凸集。)

练习9.12应该有帮助。 $\diamond$

练习9.14 设$I \subset \mathbf{R}$为一个区间。证明当且仅当f在I上是凸的

$$f\big((1-t)a + tb\big) \leq (1-t)f(a) + tf(b), \qquad 0 \leq t \leq 1$$

对于所有 a 和 b 在 I 中。 $\diamond$

练习9.15 证明指示函数 $\chi_{\{0\}}$: [0, 1] $\to$ $\mathbf{R}$ 是凸的。

提示：您可以显式地找到割线。 $\diamond$

练习9.16 证明如果 $f:[0,1]\to\mathbf{R}$ 在 $(0,1)$ 上有一个跳跃间断点，那么 f 不是凸的。$\diamond$

练习9.17 证明如果 f 是二阶可微的，并且 $f'' \neq$ 在 $[a,b]$ 上为0，那么 f 在 $[a,b]$ 上最多消失两次。
提示：直接使用平均值定理；不要假设 f 是 $\mathscr{C}^2$。

直观原理是，如果 f'' 非零（无处为零），那么 f 在 f 的定义域内每个区间上至多消失两次。如果 f 的定义域不是一个区间，那么 f 可能消失超过两次。$\diamond$

练习9.18 设 $f:[0,+\infty)\to\ \mathbf{R}$ 是一个非递增、凸函数。f 是否可以有一个跳跃间断点？两个跳跃间断点？如果我们不假设 f 是非递增的会怎样？像往常一样，为你的论点提供证明或反例。$\diamond$

练习9.19 本练习建立了Hölder不等式：如果 {v*} 1，并且如果 {v*} 和 {v*} 在 [{v*}] 上可积，那么

$$(**)\qquad \Big|\int_a^b fg\Big| \le \int_a^b |fg| \le \Big(\int_a^b |f|\Big)^{1/p}\Big(\int_a^b |g|\Big)^{1/q}$$

(回忆一下，fg 可以通过练习7.12进行积分.)

(a) Let $\alpha\in(0,1)$. Prove that $t^\alpha \le \alpha t + (1-\alpha)$ for all $t\ge 0$.

(b) 令 $\beta = 1-\alpha$，并注意 $\beta\in(0,1)$。证明　　　那

$$u^\alpha v^\beta \le \alpha u + \beta v \qquad \text{for all } u,\ v>0.$$

建议：在(a)部分设置 $t=u/v$。

(c) 设 $p>$ = 1，并令 $q=p/(p-$ = 1)，使得 $\frac{1}{p}+\frac{1}{q}=$ = 1。证明

$$AB \le \tfrac{1}{p}A^p + \tfrac{1}{q}B^q \qquad \text{for all } A,\ B\ge 0.$$

建议：使用部分(b)，并对变量进行适当的更改。

(д) (有限序列的Hölder不等式) 证明如果p和q如(c)中所述，并且对于 $1\le k\le n$，a_k和b_k是实数，那么

$$\sum_{k=1}^n |a_k b_k| \le \Big(\sum_{k=1}^n |a_k|^p\Big)^{1/p}\Big(\sum_{k=1}^n |b_k|^q\Big)^{1/q}.$$

提示：设置 $A_k = a_k/(\sum_k |a_k|^p)^{1/p}$ 等，并使用部分 (c)。

(e) 证明方程 $(**)$。有几种方法可以继续；要么使用部分 (d) 和黎曼和，要么使用部分 (c) 以及 A 和 B 适当的积分。

Hölder的不等式在分析中具有极大的技术重要性。特殊情况 $p = q =$^ 2 是Schwarz不等式。$\diamond$

拉普拉斯法则

练习9.20 使用洛必达法则计算以下极限:

$$\lim_{x\to 1}\frac{\log x}{x-1}\qquad \lim_{x\to 0^+} x\log x\qquad \lim_{x\to+\infty}\frac{\log x}{x}\qquad \lim_{x\to+\infty}\frac{x}{\exp x}.$$

建议对第二个极限进行修改：写为 $x = 1/(1/x)$。 $\diamond$

练习9.21 使用前一个练习的结果，以及其他所需技术，来评估以下极限:

$$\lim_{x\to+\infty}\frac{x^n}{\exp x}\quad\text{and}\quad \lim_{x\to 0}\frac{e^{-1/x^2}}{x^n}\quad \text{for } n\in\mathbf{N};\qquad \lim_{x\to 0^+} x^x.$$

尽管有最后的限制，表达式 0^0 是未定义的。 $\diamond$

练习9.22 设f在x的邻域内可微。求值

$$\lim_{h\to 0}\frac{f(x+h)-f(x-h)}{2h}.$$

Find 一个不连续的函数，对于这个极限 ex ists. $\diamond$

练习9.23 设f在邻域x中是二阶可微的。求

$$\lim_{h\to 0}\frac{f(x+h)+f(x-h)-2f(x)}{h^2}.$$

为什么我们不使用这个极限来定义 f''? $\diamond$

练习9.24 假设 f 和 g 在 $(c-\delta, c+\delta)$ 上可微，且 $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} g(x) = 0$，以及 $\lim\limits_{x\to c} g'(x) = 0$ 但 $\lim\limits_{x\to c} f'(x) = \ell \neq 0$。证明 $\lim\limits_{x\to c} f(x)/g(x)$ 不存在。$\diamond$

第10章

基本定理

我们引入了一种操作（积分），它“将无限多个无穷小量相加”，以及另一种（微分），它“以无穷大因子放大”。除了技术问题之外，这些操作是线性映射 $f \mapsto I_a f$ 和 $f \mapsto Df$，由以下定义：

$$I_a f(x) = \int_a^x f(t)\,dt, \qquad Df(x) = f'(x).$$

我们非正式地论证了这些运算应该是彼此的逆运算，并在这一方向上严格地建立了一些部分结果。在均值定理可用的情况下，我们现在可以系统地、详细地研究积分与微分之间的关系。

10.1 积分与微分

定理10.1. 设$f\colon [a,b] \to \mathbf{R}$可积，且设$F\colon [a,b] \to \mathbf{R}$为$f$的定积分，定义为

$$F(x) = \int_a^x f(t)\,dt \qquad for\ x \in [a,b].$$

如果 f 在 $c \in (a,b)$ 处连续，那么 F 在 c 处可微，并且 $F'(c) = f(c)$。

定理10.1通常被称为微积分的第一基本定理，这个名字强调了它在微积分中的核心作用。在莱布尼茨记号中，结论写为

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x),$$

强调积分和微分的逆性质。

通常对于复杂的定理，人们倾向于只记住结论。但请注意，如果 f 不连续，结论通常是不正确的：积分的导数可能根本不存在，即使它存在，也可能具有“错误”的值。

证明。固定 $c \in (a,b)$。根据假设，$f = f(c) + o(1)$ 在 c，因此对于每个 $\varepsilon > 0$，存在一个包含在 (a,b) 中的开 δ-区间，使得 $f = f(c) + A(\varepsilon)$。如果 $0 < |h| < \delta$，那么牛顿商 $\Delta_c F(h)$ 被定义，并且等于

$$\Delta_c F(h) = \frac{1}{h}\Big(F(c+h) - F(c)\Big) = \frac{1}{h}\int_c^{c+h} f(t)\,dt.$$

如第7章所述，这是区间端点为c和$c+h$ (的f的平均值，即使$h < 0$)。根据定理7.20,

$$\Delta_c F(h) = \frac{1}{h}\big(f(c)h + A(\varepsilon)\,h\big) = f(c) + A(\varepsilon)$$

在开区间 δ 关于 c。由于 $\varepsilon > 0$ 是任意的，我们已经证明了 $\Delta_c F(h) = f(c) + o(1)$ 在 $h = 0$ 附近，或 $F'(c) = f(c)$。 □

第一条基本定理说明了当连续函数 f 进行积分后再求导时会发生什么：函数 f 被恢复！微积分的第二条基本定理处理相反的问题，即对导数进行积分。

定理10.2.如果 $f : [a,b] \to \mathbf{R}$ 是可积的，并且对于某个函数 $f = F'$，那么

$$\int_a^b f = F(b) - F(a).$$

证明。设$P = \{t_i\}_{i-0}^n$是$[a,b]$的一个划分。根据对F应用平均值定理，对于每个$i = 1,\dots,n$，存在一个点$x_i \in (t_{i-1}, t_i)$，使得

$$F(t_i) - F(t_{i-1}) = F'(x_i)(t_i - t_{i-1}) = f(x_i)\Delta t_i.$$

如果 m_i 和 M_i 是 f 在第 i 个子区间上的下确界和上确界，那么

$$m_i \Delta t_i \le F(t_i) - F(t_{i-1}) \le M_i \Delta t_i \qquad \text{for } i = 1,\dots,n.$$

对 i 求和表明

$$L(f,P) \le F(b) - F(a) \le U(f,P) \qquad \text{for every partition } P.$$

由于 f 可积且中间项不依赖于分割的选择，积分的值必须是 $F(b) - F(a)$。 $\square$

定理10.2的证明是第6章中给出的直观论证的严谨版本，其中包含了微分积分{v*}。

$$F'(t)\,dt = \frac{dF}{dt}dt = dF = F(t+dt) - F(t)$$

引起一个形式上可缩并的求和。在莱布尼茨记号中，第二基本定理的结论读作

$$\int_a^x dF = \int_a^x \frac{dF}{dt}\,dt = F(x) - F(a).$$

如符号所示，我们通常将 a 视为固定，将 x 视为变量。然而，在定理10.3中，a 或 x 都可以被视为“变量”，因为它们都是任意的。再次强调，你不应该只记住结论而忘记假设；假设 F' 可积是至关重要的。

这是一个回顾记号的好时机

$$F(x) - F(a) = F\Big|_a^x = F(t)\Big|_{t=a}^{t=x} = \Big[F(t)\Big]_{t=a}^{t=x}.$$

它也常见“$F(t)\big|_a^x$”，尽管严格来说这是错误的语法。

F的差 $\big|_a^x$ 可能被视为一个“0维积分”（又称“带方向的函数值之和”）的 F 在区间 $[a,x]$ 边界上的积分。基本定理断言这等同于微分 $dF = F'(t)\,dt$ 在 $[a,x]$ 上的“1维积分”。这些评论的适当背景是多变量微积分。

由于连续函数在每一个闭区间上都是可积的，定理10.2有一个直接的推论：

定理10.3. 如果 $F\colon(\alpha,\beta)\to\mathbf{R}$ 是 $\mathscr{C}^1$，那么

$$\int_a^x F'(t)\,dt = F(x) - F(a) \qquad \textit{for all } a,\ x \in (\alpha,\beta).$$

集成与反微分

一个 f 的不定积分是一个具有 $F' = f$ 的函数 F。定理 10.3 表示，在区间 $[a, x]$ 上对连续函数 f 进行积分相当于寻找 f 的一个不定积分并计算函数值的差。注意强大的含义：一个不定积分使我们能够评估某个区间内所有 x 的积分。寻找不定积分通常比计算下和并取上确界要容易得多，即使对于 x 的单个值也是如此。

基本定理的重要性有两方面：

- (实用)它极大地简化了可显式找到反导数的函数的积分计算。

- (理论)它表现出连续函数的积分，无论是否可以通过其他方法找到积分。

仅就实际重要性而言，就足以证明大多数微积分课程中的基本定理是合理的。然而，理论重要性也不应被忽视：我们已经在练习7.17中看到，如何研究定义为定积分的函数的性质（即如何将有趣的函数定义为积分）。

定理10.2描述了积分与反微分之间的紧密联系，许多微积分教材（以及工作科学家和数学家）使用积分符号来表示反导数，如下所示

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C,$$

无处不在的“$+C$”代表一个任意的加性常数.[1]许多学生留下的印象是积分是反导数，也许甚至可以说是定义上的。原因可能是人类心理学：积分的定义相对复杂（分割、求和和上确界），相比之下，反导数简单，而且“大多数时候”它们在功能上是等价的。然而，积分和反导数不是同一回事，这正是微分和积分计算的奇迹所在。反导数通常很容易找到，但没有明显的有趣解释。

[1]We avoid such notation; “x” is a dummy variable on the left but not on the right.

积分富含意义（总变化），但根据定义计算起来很费力。

> 微积分基本定理提供了一种轻松处理具有极大理论和实际兴趣的数量的方法。

一个更微妙的问题是，定理10.1和10.3是为连续被积函数陈述的。积分和微分是逆运算，即使是加性常数，这根本不是真的。确切地，

- 存在（不连续的）可积函数 f，其积分在每处都可微，但这样的

$$f(x) \neq \frac{d}{dx}\int_a^x f(t)\,dt$$

  对于无限多个 x。例3.11的分子函数具有此性质，参见练习7.9。

- 存在可微函数 F（，其导数）不连续，且 F' 不可积，因此定理10.2 甚至不适用。

尽管有这些警告，基本定理可以加强；例如，如果我们将“积分”的定义扩大到包括不定积分，定理10.3的结论对于导数在有限多个点处无界的可微函数F仍然成立，参见练习10.20。

10.2 反微分

因为积分和反导数对于连续函数关系非常密切，所以关于导数的每一个计算定理（特别是链式法则和乘积法则）都对应着计算积分的有用工具。换元法源于链式法则，而分部积分法对应于莱布尼茨法则。

令人瞩目（考虑到与导数的状况），不存在寻找乘积、商和复合函数的积分的一般公式或算法。这并非数学断言

无知，但却是生活的现实。人们能做的最好的事情就是遵循格言，[2] “先开枪，把打中的任何东西都称为‘目标’。”具体来说，策略是编制一个已知导数的表格；这个表格的条目是我们知道如何反求导的函数。

模糊原则“反导数是非算法的”总结了几个定理的教训。这个问题在第15章中进行了更深入的研究，但我们可以给出一个错误发生迹象。回忆一下，“有理函数”是多项式的商，而“代数函数”是由两个变量的多项式隐式定义的。每个有理函数都是代数的；形式上，如果$f(x) = p(x)/q(x)$与p和q是多项式，那么$F(x, y) = p(x) - yq(x)$隐式定义了f。第8章的定理表明，有理函数的导数是有理的，代数函数的导数是代数的。

自然对数，有理倒数函数的反导数，甚至不是代数的。这不仅仅是一个孤立的例子，而是“随机选择”的有理函数的一个特性。此外，在第13章中，我们将看到反三角函数，它们都不是代数的，都具有代数导数，甚至反正切函数还有有理导数。至于有理函数，一个“典型”的代数函数没有代数反导数。

尽管有这些严峻的警告，但许多函数可以明确地反微分。然而，这个主题的味道更像是艺术而非科学。在本书的这个阶段，该领域的现状并不令人印象深刻:

命题10.4. 设{v*}7为一个有理数，并定义{v*}8对于{v*}9 0。幂函数{v*}10 1{v*}11是{v*}12的一个不定积分。自然对数是倒数函数{v*}13的一个不定积分。

常数的这些有理幂函数（例如，多项式）的乘积之和可以简单地处理。基本定理使我们能够计算，例如，

$$\int_0^1 (t + 2\sqrt{t})\, dt = \left(\frac{t^2}{2} + 2\frac{t^{3/2}}{3/2}\right)\Bigg|_{t=0}^{t=1} = \frac{1}{2} + \frac{4}{3}.$$

然而，我们目前被如 $\sqrt{1+t^2}$、$t\sqrt{1+t^2}$ 和 $\sqrt{1+t^{1/2}}$ 这样的被积函数所困扰。当然，这些函数有不定积分-

[2]Due to Ashleigh Brilliant, used with permission.

tives（即它们在合适区间上的定积分），但我们还不知道如何将这些不定积分写成“封闭形式”。

替换

正如链式法则允许对复合函数进行微分一样，“换元法”或“变量替换定理”允许对合适的函数进行反微分。

定理10.5. 设$g : (\alpha, \beta) \to \mathbf{R}$是$\mathscr{C}^1$，$[a, b] \subset (\alpha, \beta)$，并假设$f$是一个定义域包含$g([a, b])$的连续函数。那么

$$\int_a^b (f \circ g)\, g' = \int_{g(a)}^{g(b)} f.$$

证明。设 F 是 f 的一个不定积分，并设 $G = F \circ g : [a, b] \to \mathbf{R}$。函数 G 是连续可微的，链式法则意味着 $G' = (f \circ g)\, g'$。根据微积分基本定理的第二部分，

$$\begin{aligned}
\int_a^b (f \circ g)\, g' &= \int_a^b G' = G\Big|_a^b \\
&= F\big(g(b)\big) - F\big(g(a)\big) = F\Big|_{g(a)}^{g(b)} \\
&= \int_{g(a)}^{g(b)} F',
\end{aligned}$$

自 $F' = f$ 以来，该定理已被证明。 □

在传统符号中，使用虚变量，变量替换定理的结论可写为

$$\int_a^b \big(f \circ g\big)(t)\, g'(t)\, dt = \int_{g(a)}^{g(b)} f(u)\, du. \tag{10.1}$$

这个公式在莱布尼茨记号表示的导数中特别引人入胜。“替换” $u = g(t)$被微分，得到

$$du = g'(t)\, dt = \frac{du}{dt}\, dt,$$

这看起来就像分数的约分.[3] 要改变积分的上下限，注意如果 $t = a$，那么 $u = g(a)$。同样 $t = b$

[3]Remember that we have not assigned meaning to isolated infinitesimal expressions.

对应于 $u = g(b)$。形式替换将 (10.1) 的左侧转换为右侧。再次，这意味着莱布尼茨记法是恰当选择的，并不是定理 10.5 是重言式。

示例10.6考虑

$$\int_0^2 (1+t-t^3)^{10}\,(1-3t^2)\,dt.$$

评估直接从定义出发是毫无希望的，因为这首先需要展开（得到一个32次的多项式），然后求上和下和，分别找到下确界和上确界。然而，$(1+t-t^3)$关于t的导数是$1-3t^2$。如果我们设$u=(1+t-t^3)$，那么当$t=0$时$u=1$，当$t=2$时$u=-5$，所以

$$\int_0^2 (1+t-t^3)^{10}\,(1-3t^2)\,dt = \int_1^{-5} u^{10}\,du = \frac{u^{11}}{11}\bigg|_{u=1}^{u=-5} = \frac{1}{11}\Big((-5)^{11}-1\Big).$$

如果寻找不定积分，可以通过设置 $u = (1+t-t^3)$ 而不是插入数值极限来进行反微分后找到。在莱布尼茨记号中，

$$\frac{d}{dt}\frac{1}{11}(1+t-t^3)^{11} = (1+t-t^3)^{10}\,(1-3t^2),$$

从链式法则中可以看出。 □

练习

“标准”微积分练习包括大量不定积分的计算。虽然这是一个重要的技能，但越来越有可能依赖符号操作程序；对于非数学家来说，积分技术可能会被忽视。了解给定函数何时可以期望有显式的不定积分是值得的。在许多情况下，代数可以将看似不可能的函数转换为容易积分的形式。

练习10.1 给定以下函数在 t 的值。找到一个原函数，并通过微分来检查你的答案。一些

它们可以有不止一种方式完成。

(a_1)	$(1+t)^3$	(a_2)	$t(1+t^2)^3$	(a_3)	$(1+t^2)^3$
(b_1)	$\sqrt{1+t}$	(b_2)	$2t\sqrt{1+t^2}$	(b_3)	$t/\sqrt{2+t^2}$
(c_1)	$(1+\sqrt{t})^2$	(c_2)	$(t+t^{-1})^3$	(c_3)	$(1-t^{-2})(t+t^{-1})^3$
(d_1)	$(t+\sqrt{t})^2$	(d_2)	$(\sqrt{t}+1/\sqrt{t})^2$	(d_3)	$(t^{1/2}+t^{3/2})^3(1+3t)/\sqrt{t}$
(e_1)	$1/(1-t^2)$	(e_2)	$t/(1-t^2)$	(e_3)	$1/\big((t-1)^2(1+t)\big)$

请注意，即使是单一部分，您可能也需要不同的技术。$\diamond$

练习10.2求$\displaystyle\int_0^x \frac{(1+t)2+2}{(}\,t+t^2)^3\,dt$对于$x\in\mathbf{R}$的值 $\diamond$

假设 $f:\mathbf{R}\to\mathbf{R}$ 是连续的。

(a) 定义 $G:\mathbf{R}\to\mathbf{R}$ 为

$$G(x)=\int_0^{x^2} f=\int_0^{x^2} f(t)\,dt.$$

证明 G 可微，并找到 $G'(x)$。
建议：将 G 写成你熟悉的两个函数的复合，并使用链式法则。

(b) 定义 $H:\mathbf{R}\to\mathbf{R}$ 为

$$H(x)=\int_x^{x^2} f.$$

证明 H 可微，并求 $H'(x)$。

(c) 假设一般地，ϕ 和 ψ 在 (α,β) 上可微，并定义

$$\Phi(x)=\int_{\psi(x)}^{\phi(x)} f \qquad \text{for } x\in(\alpha,\beta).$$

证明 Φ 可微，并求 $\Phi'(x)$ 关于 f、ϕ 和 ψ 的表达式。

$\diamond$

练习10.4 存在可积函数 $f : [0, 1] \to \mathbf{R}$ 使得

$$\int_0^x f = 1 \qquad \text{for all } x \in (0, 1]?$$

解释。 $\diamond$

练习10.5 存在可积函数 {v*} : $[-1, 1] \to \mathbf{R}$ 使得

$$\int_{-1}^x f = \sqrt{1-x^2} \qquad \text{for all } x \in [-1, 1]?$$

解释。如果f只需要是不恰当可积的会怎样? $\diamond$

练习10.6 假设 f 和 g 是 $[a, b]$ 上的可积函数，并且

$$\int_a^x f = \int_a^x g \qquad \text{for all } x \in [a, b].$$

这是否意味着对所有 x，$f(x) = g(x)$ 都成立？如果 f 和 g 是连续的怎么办？ $\diamond$

练习10.7 设 $u : \mathbf{R} \to \mathbf{R}$ 连续。证明对于每个 $c \in \mathbf{R}$，存在唯一一个 $\mathscr{C}^1$ 函数 $f : \mathbf{R} \to \mathbf{R}$ 满足初值问题

$$f' = u, \qquad f(0) = c.$$

$\diamond$

练习10.8 设$I \subset \mathbf{R}$为一个区间，且设$f : I \to \mathbf{R}$ 是 $\mathscr{C}^1$。证明存在非递减函数 g_1 和 g_2，使得 $f = g_1 - g_2$。

提示：考虑 f' 的正负部分。

让 $f(x) = x - x^3/3$, $|x| \leq 2$，见图 9.2。在同一组坐标轴上，绘制 f 和一对差为 f 的非递减函数。 $\diamond$

练习10.9 设 $A : (-1, 1) \to \mathbf{R}$ 为由以下 $\mathscr{C}^1$ 函数特征化的

$$A'(x) = \frac{1}{\sqrt{1-x^2}}, \qquad A(0) = 0.$$

(a) 证明 A 是单射的。

(b) By (a), A is invertible; let $S = A^{-1}$. Prove that $S' = \sqrt{1-S^2}$.

(c) 使用 (b) 证明 S 是 $\mathscr{C}^2$，并以 S 为基础求出 S''。

$\diamond$

练习10.10 使用变量替换定理重新做练习7.13。（你的计算应该非常简短。良好的理论工具可以是一个巨大的节省劳动力的设备。）$\diamond$

练习10.11 设 $f : \mathbf{R} \to \mathbf{R}$ 是连续的。

(a) 定义

$$g(x) = \int_0^x t f(t)\, dt, \qquad h(x) = \int_0^x x f(t)\, dt.$$

查找 $g'(x)$ 和 $h'(x)$。
建议：将 h 重新编写，使得积分内没有 “x”。

(b) 使用(a)部分的结果来证明

$$\int_0^x f(u)(x-u)\, du = \int_0^x \left(\int_0^u f(t)\, dt \right) du.$$

建议：对每边关于 x 求导。 $\diamond$

练习10.12 假设 u 和 v 是在 $[a, b]$ 上的连续可微函数。证明分部积分公式：

$$\int_a^b u\, v' = uv\Big|_a^b - \int_a^b v\, u',$$

在Leibniz记号法中写出。建议：整合导数的乘积法则。$\diamond$

练习10.13形式上“证明” $\int_{-1}^{1} 1\, {}_{x^2}\, dx = -2$。—

(a) 被积函数是正的，但积分是负的。出了什么问题?

(b) 解释为什么在要求评估一个不定积分时，你必须始终单独证明该积分存在。

课程教训是，在合法的情况下，正式计算是好的，但如果盲目应用，可能会导致错误。$\diamond$

练习10.14 使用练习10.12来评估以下内容：

$$\int_1^b t^n \log t\, dt \qquad \int_0^1 t^n \log t\, dt.$$

建议：让 $u(t)$ = 对 t 和 $v(t) = t^n$ 取对数。注意第二个积分是不恰当的，因此你必须建立收敛性。◇

练习10.15求 $\int_0^1 \frac{dt}{\sqrt{t}}$和 $\int_1^{+\infty} \frac{dt}{t^3}$。 ◇

练习10.16 这将推广前面的问题。设$r > 0$ 为有理数；计算以下不定积分：

$$\int_0^1 \frac{dt}{t^r} \quad (0 < r < 1) \qquad \int_1^{+\infty} \frac{dt}{t^r} \quad (1 < r).$$

◇

练习10.17 评估不定积分 $\int_0^1 \frac{t}{\sqrt[3]{1 - t^2}}\, dt$ ◇

练习10.18 确定以下哪些不定积分收敛；无需求值。

(a) $\int_0^{+\infty} \frac{t\,dt}{t^2+1}$

(b) $\int_{-1}^{1} \frac{dt}{\sqrt{1-t^2}}$

(c) $\int_{-1}^{1} \frac{t\,dt}{\sqrt{1-t^2}}$

确保如有必要将积分分成几部分。 ◇

练习10.19 设 $\psi : \mathbf{R} \to \mathbf{R}$ 为伪正弦函数，并考虑以下函数

$$F(x) = \begin{cases} x^2\,\psi(1/x^2) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

证明 F 在 $\mathbf{R}$ 上可微，但 F' 在 0 附近是无界的。是否可以写成

$$F(x) = \int_0^x F'(t)\,dt$$

在理解积分是不定积分的情况下? ◇

练习10.20 假设 F 在 $(\alpha,\beta)\supset[a,b]$ 上可微，并且 F' 只有有限个间断点。证明

$$\int_a^b F' = F(b) - F(a)$$

给定左侧被解释为不定积分。练习10.19给出了一个适用于此结果的应用示例。$\diamond$

第11章

函数序列

在这个阶段，我们已经发展了微积分的基本工具，即微分和积分，并看到了它们之间的关系以及如何使用它们来研究函数的行为。我们所缺乏的是一个函数库，特别是对数、指数和三角函数。（我们已经定义了自然对数，但尚未证明像 $\log(a^b) = b\log a$ 和 $e^{xy} = (e^x)^y$ 这样的恒等式，因此我们还没有熟练地操作对数和指数。）我们知识缺乏的原因是：为了精确地定义非代数函数（以 $\mathbf{R}$ 的公理为依据），我们必须使用非代数运算，例如极限和上确界。如果你怀疑（你应该怀疑！），尝试在不参考几何的情况下定义一般角度的余弦，或者仅使用 $\mathbf{R}$ 的公理来解释 $2^{\sqrt{2}}$ 的含义。

一种将极限纳入函数定义的最具体方法是通过幂级数，即“无限次多项式”。一个例子是

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{x}{1} + \frac{x^2}{2\cdot 1} + \frac{x^3}{3\cdot 2\cdot 1} + \cdots + \frac{x^k}{k!}\cdots. \tag{11.1}$$

尽管无限多个项的逐字相加是未定义的，但我们可能固定一个特定的 x，将 (11.1) 视为一个数列，并询问该数列在哪些 x 上收敛。这就是所谓的“逐点收敛”。由于将变得明显的技术原因，最好是将 (11.1) 的部分和视为一个函数序列，而不仅仅是询问数列在哪些 x 上收敛，而是要求数列在某些区间上“以相同的速率”收敛。

在详细研究收敛性之前，让我们看看我们从幂级数中可以得到什么。乐观地讲，如果(11.1)的右侧在某个区间内对所有x收敛，我们希望像处理“无限项多项式”一样处理得到的结果函数；具体来说，我们希望逐项微分和积分，通过插入x的具体值来近似级数的数值，等等。更有雄心的是，我们可能考虑更一般的函数序列，并询问我们是否可以使用近似函数（在许多情况下很简单）来积分或微分极限函数（这可能是困难的）。这些问题在一般情况下没有简短的答案，但幂级数的情况尽可能好：如果幂级数在其中心以外的某个地方收敛，那么它在某个区间内收敛，并且在其收敛区间内可以精确地像多项式一样处理。这就是本章的哲学道德。

序列和级数是描述同一事物的两种方式。虽然级数在应用中自然出现，我们在推导一般理论结果时使用序列表示法。此外，尽管我们的主要兴趣是幂级数，我们首先研究一般函数序列。这种对不必要的细节的初步避免澄清了几个技术问题。

11.1 收敛

让 $I \subset \mathbf{R}$ 非空（通常是一个区间）。非正式地，一个在 I 上的“函数序列”是一个有序的实值函数集合 $(f_n)_{n=0}^{\infty}$，其定义域为 I。正式地，一个在 I 上的函数序列是一个函数

$$F : \mathbf{N} \times I \to \mathbf{R}, \qquad f_n = F(n,\ \cdot\ ) : I \to \mathbf{R}.$$

几何地取极限的方法是将 f_n 的图像视为一部无限电影中的框架（在“时间 n”拍摄），然后观察电影播放时发生了什么。粗略地说，数列的收敛意味着图像“稳定”到一个平衡图像，即 $n \to \infty$。给出一个精确的定义并不困难，但要正确塑造直觉需要很多例子，并且需要一定的回顾来理解为什么“最佳”的收敛定义不是第一个出现在脑海中的定义。

逐点收敛

设(f_n)是I上的一个函数序列。对于每个$x \in I$，$f_n(x)$的值构成一个实数序列，并且对于每个x，可以询问极限是否存在。如果

$$f(x) := \lim_{n\to\infty} f_n(x) \tag{11.2}$$

存在所有 $x \in I$，则序列 (f_n) 被称为逐点收敛到 f，并且由方程 (11.2) 定义的函数 f 是序列 (f_n) 的逐点极限。

大多数学生，如果被要求定义函数序列的“收敛”，可能会选择一个与逐点收敛等价的定义。然而，正如以下例子所证明的，一个逐点极限可能不具有其逼近序列项的理想性质。因此，我们将被迫寻求更强的收敛标准。

示例11.1 设 $I = [0, 1]$，并设 $f_n : I \to \mathbf{R}$ 为分段线性函数

$$f(x) = \begin{cases} 1 - nx & \text{if } 0 \le x \le 1/n \\ 0 & \text{if } 1/n < x \le 1 \end{cases}$$

序列(f_n)逐点收敛到一个函数f；显然对于所有$n \in \mathbf{N}$，有$f_n(0) = 1$，因此$f(0) = 1$。如果相反，$x > 0$，那么对于每一个$n >1/x$，我们有$1/n < x$，因此$f_n(x) = 0$。由此可知，当$x > 0$时，$f(x) = 0$。

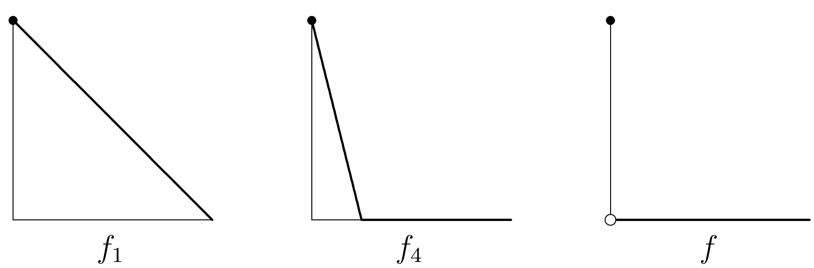


图11.1：具有不连续极限的连续函数序列。

总结来说，序列(f_n)逐点收敛到

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

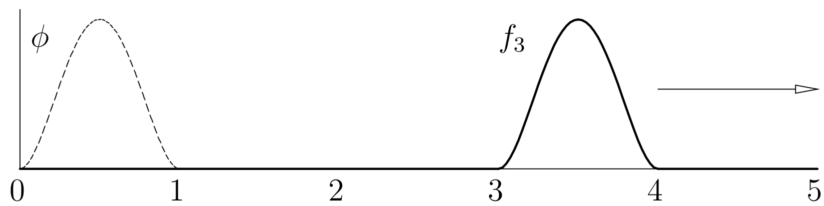


图11.2：在 $+\infty$ 处消失的隆起。

虽然每个函数 f_n 都是连续的，但极限函数不是。将点态收敛序列的极限传递可能导致图形“断裂”；极限的图形在直观几何意义上不是“图形的极限”。□

示例11.2 考虑由以下可微函数 $\phi : \mathbf{R} \to \mathbf{R}$ 定义：

$$\phi(x) = \begin{cases} 30(x - x^2)^2 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

并且定义 $f_n(x) = \phi(x-n)$，图11.2。ϕ 的图形有一个宽度为1的“隆起”，包围了1个单位的面积，f_n 的图形向右平移了 n。特别是，对于每个 $n \in \mathbf{N}$，f_n 在区间 $[-\infty, n]$ 上恒为零。

如果 $x \in \mathbf{R}$，则存在一个自然数 N 使得 $x < N$，这意味着对于所有 $n > N$，有 $f_n(x) = 0$。因此，

$$f(x) := \lim_{n\to\infty} f_n(x) = 0 \qquad \text{for all } x \in \mathbf{R};$$

序列 (f_n) 在每一点上收敛到零函数。 □

f_n（在这种情况下，是一个凸起）的特性通常不会被极限函数继承。在这个例子中，随着n的增加，凸起向右移动，其距离原点的上限不存在。对于每一个$x > 0$，无论有多大，凸起都像波浪一样从左侧进入，然后穿过x的右侧。之后，在x处没有任何变化。从某种意义上说，凸起“在无穷大处消失。”

修改前例可能看起来更加令人惊讶。序列$(nf_n)_{n=1}^{\infty}$逐点收敛到零函数，尽管随着向右移动，波峰可以任意增大。此外，波峰在空间中“在无穷远处”消失并非必要：

示例11.3 如上所述，设置 $h_n(x) = n\phi(nx)$。这留给你

(查看练习11.1)以证明序列(h_n)逐点收敛到零函数，尽管对于每个$n>0$，h_n的图像在0的右侧有一个高度为n的“尖峰”。

这个示例的一个附加特性最初似乎自相矛盾：对于每个 n，函数 h_n 在 $[0,1]$ 上是可积的，极限函数也是可积的。因此，人们可能会预期

$$(*) \qquad \lim_{n\to\infty}\int_0^1 h_n = \int_0^1 \Big(\lim_{n\to\infty} h_n\Big).$$

然而，h_1 在 $[0,1]$ 上的积分等于 1，并且因为 h_n 的整个“尖峰”都发生在同一个区间内，练习 7.13 显示

$$\int_0^1 h_n = 1 \qquad \text{for all } n \ge 1,$$

因此，$(*)$ 的左侧等于 1。但逐点极限是零函数，所以 $(*)$ 的右侧等于 0。方程 $(*)$ 在这个例子中是错误的！□

这个例子已经表明，逐点收敛是一个过于弱小的收敛概念。然而，情况可能更糟：一个可积函数序列的逐点极限甚至可能根本不可积。

示例11.4 设$(a_k)_{k=1}^{\infty}$是一个枚举区间$[0,1]$内有理数的序列，参见定理3.19。对于$n\in\mathbf{N}$，定义

$$A_n = \bigcup_{k=1}^{n}\{a_k\};$$

集合 A_n 包含序列 (a_k) 的前 n 项。最后，令 $f_n:[0,1]\to\mathbf{R}$ 为 A_n 的特征函数。将这些图看作一部电影，我们开始于零函数，在时间 n 将图上的一个点——位于 a_n 上的一点——移动到高度 1。每个函数 f_n 都是可积的，因为 f_n 除了有限多个点外恒等于零。然而，逐点极限是 $[0,1]$ 的特征函数 $\mathbf{Q}\cap$，正如我们在第 7 章中看到的，它不可积。□

均匀收敛

逐点收敛是不够的，因为微积分的基本运算——极限提取、积分和微分——具有局部性质：它们依赖于函数在一点的值，而不是依赖于函数在所有点的值。

但是关于函数在开区间上的行为。如果序列(f_n)在I上逐点收敛，那么对于每一个$x \in I$和每一个$\varepsilon > 0$，存在一个N使得

$$f_n(x) = f(x) + A(\varepsilon) \quad \text{for } n \geq N.$$

如果选择不同的点 y，那么对于相同的 ε 可能需要更大的 N，并且如果 x 范围甚至是一个任意小的区间，相应的 N 可能不存在上界。

回顾起来，我们当时不能期望连续性（例如）被逐点收敛的函数序列的极限所继承。为了纠正这种情况，我们引入了“收敛”的更强标准。

定义11.5 设(f_n)是I上的一个函数序列，该序列在[1]上逐点收敛到函数$f: I \to \mathbf{R}$，并设$X \subset I$。如果对于每个$\varepsilon > 0$，存在一个索引N，使得

$$f_n = f + A(\varepsilon) \quad \text{on } X \text{ for } n \geq N. \tag{11.3}$$

如果 I 是一个区间，并且如果 $(f_n) \to f$ 在 I 中每个闭的、有界的区间上均匀，那么我们说收敛在紧集上是均匀的。

如果 (f_n) 在 I 上 $\to f$ 均匀，那么显然收敛在 I 的每个非空子集上都是均匀的。均匀收敛包含了一致收敛，但是一个更强的条件。直观上，“相同的 N 对所有 x 都适用”，或者说“均匀收敛是一致收敛，就像均匀连续是一致连续一样。” 后者不是一个同义反复；相反，这意味着术语选择得很好。这些概念——而不仅仅是它们的名称——是相似的。

均匀收敛具有几何解释。如果 $f : I \to \mathbf{R}$ 是一个函数且 $\varepsilon > 0$，那么我们定义关于图形的 ε-管为

$$\{(x, y) \in \mathbf{R}^2 : |f(x) - y| < \varepsilon,\ x \in I\},$$

即与f的图形垂直距离为ε的点集。

要说明(f_n)一致收敛到f，意味着对于关于f的图上的每一个ε-管，都存在一个N，使得f_n的图对于所有的$n \geq N$都位于该管内。换句话说，序列

[1]It is the *sequence* that converges, not the functions!

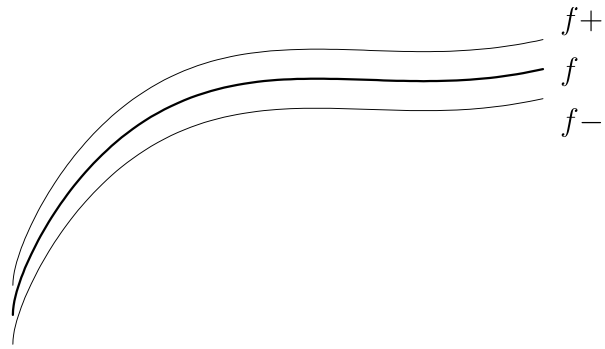


图11.3：关于函数图形的 ε-管

(f_n) 在 I 上对 f 一致收敛，如果 f_n 和 f 的图形之间的最大垂直距离可以任意小。为了后续使用，我们精确地表述如下：

命题11.6.设(f_n)是一个在I上逐点收敛到f的序列，并且对于每个
$n \in \mathbf{N}$设定

$$a_n = \sup_{x \in I} |f(x) - f_n(x)|.$$

然后 $(f_n) \to f$ 在 I 上均匀当且仅当 $(a_n) \to 0$。

证明直接从定义中得出，留作练习。命题11.6是有用的，因为a_n通常可以很容易地计算。

示例11.7 设(f_n)是示例11.1中的序列。正如我们所看到的，这个序列逐点收敛到$\{0\}$的特征函数，即

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

然而，即使在半开区间（$0, 1$]上，收敛性也不是统一的，其中f恒为零。在命题11.6的记法中，

$$a_n = \sup\{f_n(x) \mid 0 < x \le 1\} = 1 \qquad \text{for all } n,$$

所以 (a_n) 不收敛到 0。

与相比之下，考虑一个形式为 $[\delta, 1] \subset (0, 1]$ 的任意闭区间。由于 f_n 在区间 $[\frac{1}{n}, 1]$ 上恒为零，我们发现

$a_n = 0$ 当且仅当 $n > 1/\delta$。因此，序列 (f_n) 在每个区间 $[\delta, 1] \subset (0, 1]$ 上对零函数一致收敛。由于 (0, 1] 的每个闭子区间都包含在形式为 $[\delta,$ 1] 的区间中，我们注意到 (f_n) $\to 0$ 在 (0, 1] 中的紧致集上一致。注意，移除单个点不足以“修复”0处的问题，而移除任意小的区间则足够。□

示例11.8设$\phi : \mathbf{R} \to \mathbf{R}$为一个有界函数，并定义

$$f_n(x) = \frac{1}{n}\phi(nx), \qquad \text{for } x \in \mathbf{R}.$$

几何上，f_n的图是ϕ的图通过一个因子n“缩小”得到的。显然，(f_n) $\to$在$\mathbf{R}$上均匀为0，因为如果$|\phi| \le M$在$\mathbf{R}$上，那么对于每个n，$|a_n| \le M/n$。

假设此外 ϕ 是可微的。那么对于每个 n 和 $f_n'(x) = \phi'(nx)$，f_n 是可微的。除非 ϕ' 是一个常数函数，否则序列 (f_n') 即使在点值上也无法收敛。换句话说，可微函数序列的一致收敛甚至不意味着导数序列的点值收敛。□

尽管有最后一个例子，但序列（f_n）的一些性质，如连续性和可积性，是由一致极限继承的。

定理11.9. 假设(f_n)是定义在I上的连续函数序列，且一致收敛于f。那么f是连续的。

在词义上，连续函数的均匀极限是连续的。

证明。我们希望证明如果 $x \in I$，那么在 x 处有 $f - f(x) = o(1)$。技巧是写出

$$f - f(x) = \big(f - f_N\big) + \big(f_N - f_N(x)\big) + \big(f_N(x) - f(x)\big). \tag{11.4}$$

修复 $\varepsilon > 0$，并使用一致收敛来选择 N 使得 $f - f_N = A(\varepsilon/3)$ 在 I 上。因为 f_N 在 I 上是连续的，所以在 x 的一个邻域上存在 $f_N - f_N(x) = A(\varepsilon/3)$。在这个邻域中，(11.4) 右边的三个项都是 $A(\varepsilon/3)$，所以它们的和是 $A(\varepsilon)$。由于 $\varepsilon > 0$ 是任意的，我们已证明 $f - f(x) = o(1)$ 在 x 上。□

定理11.10. 设(f_n)是$[a, b]$上的可积函数序列，且一致收敛于f。则极限函数f在$[a, b]$上可积，并且

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f = \int_a^b \Big(\lim_{n\to\infty} f_n\Big).$$

证明。证明极限函数可积的策略类似于定理11.9的证明：极限函数在每一点都接近一个可积函数，因此其上、下积分差异不大。然后可以很容易地证明极限函数的积分具有“正确”的值。

修复 $\varepsilon > 0$，并选择 N 使得

$$|f(x) - f_N(x)| < \frac{\varepsilon}{b-a} \qquad \text{for all } x \in [a,b]. \tag{11.5}$$

通过将这个不等式写成

$$f_N(x) - \frac{\varepsilon}{b-a} \;<\; f(x) \;<\; f_N(x) + \frac{\varepsilon}{b-a} \qquad \text{for all } x \in [a,b],$$

并且根据下和上和的定义，很明显

$$L(f_N, P) - \varepsilon \le L(f, P) \le U(f, P) \le U(f_N, P) + \varepsilon \tag{11.6}$$

对于每个分割 P of $[a,b]$。因为 f_N 是可积的，存在一个分割 P 使得 $U(f_N, P) - L(f_N, P) < \varepsilon$。对于这个分割，方程 (11.6) 意味着 $U(f, P) - L(f, P) < 3\varepsilon$，由于 $\varepsilon > 0$ 是任意的，f 是可积的。

现在考虑序列$(f - f_n)$，它一致收敛到零函数。固定$\varepsilon > 0$并选择N如方程(11.5)所示。然后

$$\left|\int_a^b f - \int_a^b f_n\right| = \left|\int_a^b (f - f_n)\right| \le \int_a^b |f - f_n| < \varepsilon$$

对于 $n \ge N$。这意味着 $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f = \int_a^b \Big(\lim_{n\to\infty} f_n\Big)$。 □

存在，如已建议的，没有类似的结果对于导数。原因是绝对值小的函数可以有很大的导数。仔细考察下一个定理表明，假设与定理11.9和11.10的假设在本质上不同。下面的连续性假设（i）可以减弱，但这里给出的陈述对我们的目的来说是足够的。

定理11.11. 设(f_n)是在区间I上的可微函数序列，并且假设此外

(i) 每个函数 f_n' 在 I 上是连续的;

(ii) 导数序列(f_n')在紧集上一致收敛到函数g;

(iii) 原始序列(f_n)在单一点$x_0 \in I$收敛。

然后，原始序列(f_n)在紧集上均匀收敛到一个可微函数f，并且 $f' = g = \lim f_n'$。

证明。根据微积分基本定理二,

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)\, dt \qquad \text{for all } x \in I.$$

设置 $f(x_0)$ ＝的极限 $f_n(x_0)$。根据 (iii) 和定理 11.10，右边的极限逐点收敛到

$$f(x) := f(x_0) + \int_{x_0}^{x} g(t)\, dt.$$

函数 f 可由微积分基本定理一阶可导，以及 $f' = g$。

它还需要检查 $(f_n) \to f$ 在紧致集上均匀。固定 $\varepsilon > 0$ 并选择一个包含 x_0 的闭区间 $[a, b] \subset I$。注意，如果 $x \in [a, b]$，则 $|x - x_0| \leq (b - a)$。根据 (ii) 和 (iii)，存在 $N \in \mathbf{N}$ 使得 $n > N$ 推出

$$|f(x_0) - f_n(x_0)| < \frac{\varepsilon}{2} \text{ and } |g(x) - f_n'(x)| < \frac{\varepsilon}{2(b-a)} \quad \text{for all } x \in [a, b].$$

但是这意味着如果 $n > N$，那么

$$\begin{aligned} |f(x) - f_n(x)| &\leq |f(x_0) - f_N(x_0)| + \int_{x_0}^{x} |g(t) - f_n'(t)|\, dt \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2(b-a)} |x - x_0| < \varepsilon, \end{aligned}$$

对于所有 $x \in [a, b]$，这完成了证明。 □

11.2 函数系列

正如数值无穷级数是部分和序列的极限一样，函数的无穷级数是部分和序列的极限。假设 $(f_k)_{k=0}^{\infty}$ 是 I 上的一个函数序列，并定义序列 (s_n)

通过部分和

$$s_n(x) = \sum_{k=0}^{n} f_k(x). \tag{11.7}$$

如果序列(s_n)一致收敛到f，则称(f_k)在I上一致可和，我们记作

$$f = \sum_{k=0}^{\infty} f_k.$$

幂级数，其形式为

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k \quad \text{for some } x_0 \in \mathbf{R},$$

这些是典型例子，但绝非唯一有趣的例子。一个幂级数是从一个序列（f_k）中产生的，其中f_k是次数为k (的单项式，可能为零)，对于每个k。并非每个多项式级数都是幂级数；一个例子是

$$\sum_{k=0}^{\infty} (x^k - x^{2k}),$$

尽管这是两个幂级数的差。一般来说，一个多项式序列可以具有令人惊讶的性质。我们将很快看到，每个连续函数 $f \colon [0, 1] \to \mathbf{R}$ 都是多项式序列的极限。这是值得注意的，因为一般连续函数在无处可导，而多项式在每处都是可无限次导的。

均匀可求和与极限交换

定理11.10表明，如果(f_k)是某个区间I上可积函数的均匀可和序列，并且如果$[a, b] \subset I$，那么$\sum_k f_k$在$[a, b]$上可积，并且

$$\int_a^b \left(\sum_{k=0}^{\infty} f_k \right) = \sum_{k=0}^{\infty} \left(\int_a^b f_k \right).$$

类似地，定理11.11表明，如果(f_k)是一个在单一点可积分的函数序列，并且如果导数序列是均匀可积分的，那么

$$\left(\sum_{k=0}^{\infty} f_k\right)' = \sum_{k=0}^{\infty} f_k'.$$

当这些等式成立时，我们说级数 $\sum_k f_k$ 可以逐项积分或微分。对于有限和，没有问题，因为积分和微分是线性算子，但无穷和涉及第二个极限操作。这些等式断言，在适当的假设下，两个极限操作可以互换。记住，我们主要对幂级数感兴趣。上述等式保证，在幂级数一致可求和的区间上，它可以被计算上当作多项式来处理。

在我们可以充分利用这些结果之前，我们需要一个简单的标准来确定一个函数序列何时是均匀可和的，特别是当幂级数是均匀可和的时候。均匀可和性的理想简单、一般标准被称为魏尔斯特拉斯M测试，比较命题11.6。

定理11.12. 设(f_k)是I上的一个函数序列，并设

$$a_k = \sup_{x\in I} |f_k(x)| \geq 0.$$

如果 (a_k) 是一个实数可求和序列，那么 (f_k) 在 I 上是一致可求和的。

证明。固定 $x \in I$。序列 $\big(f_k(x)\big)$ 是绝对可和的，因为 $|f_k(x)| \leq a_k$ 和 (a_k) 是可和的。令 $f : I \to \mathbf{R}$ 为序列 (f_k) 的逐点之和。为了证明部分和一致收敛到 f，写出

$$\left| f(x) - \sum_{k=0}^{n} f_k(x) \right| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} a_k.$$

这是收敛级数的尾项，它独立于 x 而收敛到 0，当 $n \to \infty$ 时。

□

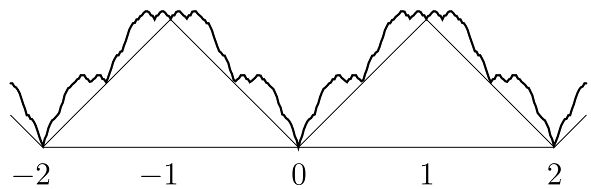


图11.4：一个魏尔斯特拉斯不可微函数。

在详细考虑幂级数之前，这里有一个长期承诺的例子，它是一个处处不可微的连续函数。这个函数是在19世纪中叶由魏尔斯特拉斯发现的，它震惊了当时习惯于将“函数”视为几乎处处可微的数学家。

示例11.13 设cb : $\mathbf{R} \to \mathbf{R}$为查理·布朗函数，并考虑由以下序列 $(f_k)_{k=1}^{\infty}$ 定义：

$$f_k(x) = 4^{-k}\,\mathrm{cb}(4^k x).$$

几何上，f_k的图像是cb“缩小”4^k倍后的图像；特别是，f_k以4^{-k}为周期，f_k的最大值a_k为4^{-k}。序列$(4^{-k})_{k=0}^{\infty}$是几何序列，公比为1/4，因此是可求和的。根据M测试，序列(f_k)在$\mathbf{R}$上是一致可求和的，并且由于每个函数f_k都是连续的，因此函数

$$F = \sum_{k=1}^{\infty} f_k$$

也是连续的。因为cb是偶数且周期为2，F是偶数且周期为2，所以只需证明F在[0, 1]上无处可微。F的递归结构是关键。

非可微分的简单几何原因：减去“第一模式”给出 $F(x) - \mathrm{cb}(x) = \frac{1}{4}F(4x)$，见图11.4。换句话说，直到加上一个分段线性项，Weierstrass函数的图像是自相似的。在第8章中，我们看到了放大可微函数的图像会得到切线。然而，图像的自相似性保证了无论我们放大多少，图像都不会变得更加线性。目前，这个论点并没有提供证明，因为我们没有考虑求和项“角落”发生的情况，尽管它

可能相当可信，F 在部分和有拐角的地方无法微分。

一个非可微分的解析证明并不困难，如果我们仔细组织我们的知识。只需证明牛顿商不存在即可：

$$\text{For every } x \in [0,1],\ \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} \quad \text{does not exist.}$$

确定一个正整数 n，并考虑该数列的第 n 部分和，以及第 n 尾项：

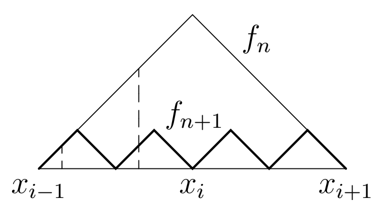
$$F_n(x) = \sum_{k=0}^{n} 4^{-k}\,\mathrm{cb}(4^k x), \qquad t_n(x) = \sum_{k=n+1}^{\infty} 4^{-k}\,\mathrm{cb}(4^k x).$$

函数 f_{n+1} 是周期为 $4^{-(n+1)}$ 的周期函数，因此 t_n 也是。因此

$$F(x \pm 4^{-n}) - F(x) = F_n(x \pm 4^{-n}) - F_n(x) \qquad \text{for all } n. \tag{11.8}$$

现在固定 x；我们将构造一个序列 $(h_n)_{n=0}^{\infty}$，其中包含 $h_n = \pm 4^{-n}$，使得相应的牛顿商在 $n \to \infty$ 时没有极限。

考虑点 $x_i = i2^{-(2n+1)}$ 对于 $0 \le i \le 2^{2n+1}$。第 n 项和 $(n+1)$ 项看起来如下：



f_n的周期为4^{-n}，分割相邻点之间的距离是这一距离的一半，是f_{n+1}周期的两倍。点x至少位于一个形式为$[x_{i-1}, x_i]$的子区间内，该区间的长度为$2 \cdot 4^{-(n+1)}$。因此，至少存在一种符号选择，使得$x \pm 4^{-n}$和x位于$[x_{i-1}, x_i]$内；这定义了序列$(h_n)_{n=0}^{\infty}$。这种典型对的位姿用虚线表示。

对于每个和项 f_k 与 $k \le n$，点 x 和 $x + h_n$ 位于一个区间上，在该区间上 f_k 是线性的，因此 f_k 在 x 和 $x + h_n$ 之间的牛顿商为 ± 1。根据（11.8），

$$\frac{F(x+h_n)-F(x)}{4^{-n}} = \frac{F_n(x+h_n)-F_n(x)}{4^{-n}} = \sum_{k=0}^{n} \frac{f_k(x+h_n)-f_k(x)}{4^{-n}}$$

对于每个 n 是一个整数。此外，将 n 增加 1 不会改变右侧任何加数的值，但会增加另一个项，这将商改变为 ± 1。

这表明每个牛顿商都是整数，但连续商不同。因此，牛顿商的序列没有极限，即对于所有 $x \in [0, 1]$，F 在 x 处不可导。□

我们选择了一种特定的缩放比例（每个加数是前一个的1/4倍）以便简化非可微性的证明。对于缩放不同的序列，可以使用类似的想法，获得更多例子。

11.3 幂级数

回忆　一个形式幂级数是表达式的　　　　　　　　　　　　表单

$$\sum_{k=0}^{\infty} a_k(x-a)^k;$$

系数构成一个序列(a_k)，点a是序列的中心。与一个形式幂级数相关联的是一系列逼近多项式，这些多项式是通过截断级数得到的：

$$p_n(x) = \sum_{k=0}^{n} a_k(x-a)^k.$$

一个形式幂级数定义了一个函数 f，其定义域是级数收敛的 x 的集合，这始终包含至少一个点 a，在该点处的值是 a_0。当 $x = a$ 时，$(x-a)^0 = 1$ 是一种惯例；这并不意味着 $0^0 = 1$。

定理11.14。假设幂级数$\sum_{k=0}^{\infty} a_k(x-a)^k$在$x_0$处收敛，并设$|x_0-a| = \eta$。如果$|x-a| < \eta$，则幂级数在$x$处收敛，且在$(a-\eta, a+\eta)$的紧集上收敛是一致的。

证明。只需假设 $a = 0$（因此定义 $\eta = |x_0|$）；这相当于进行变量平移变换。假设幂级数 $\sum_k a_k x^k$ 在某点 x_0 收敛。特别是，序列 $(a_k x_0^k)_{k=0}^{\infty}$ 收敛于 0，这反过来又意味着序列是有界的：存在一个实数 M 使得

$$|a_k x_0^k| \le M \quad \text{for all } k \in \mathbf{N}.$$

现在假设 $|x| < |x_0|$，即 x 比较接近序列的中心，比 x_0 更靠近。那么 $0 < \rho := |x|/|x_0| < 1$，并且

$$\begin{aligned}|f(x) - p_n(x)| &= \left|\sum_{k=n+1}^{\infty} a_k x^k\right| = \left|\sum_{k=n+1}^{\infty} a_k x_0^k \cdot \frac{x^k}{x_0^k}\right| \\ &\le \sum_{k=n+1}^{\infty} |a_k x_0^k| \cdot \rho^k \le \sum_{k=n+1}^{\infty} M\rho^k = \frac{M}{1-\rho}\,\rho^{n+1}\end{aligned}$$

对于每个 $n \in \mathbf{N}$。通过取 n 足够大，这个上界可以任意小；换句话说，幂级数在 x 处收敛。

它仍然需要在 $(a-\eta, a+\eta)$ 中的紧集上证明收敛性是一致的。固定一个正数 $\delta < \eta$，并设置 $\rho = \delta/\eta <$ 为 1。如果 $|x-a| \le \delta$，那么上面的论点表明

$$|f(x) - p_n(x)| \le \frac{M}{1-\rho}\,\rho^{n+1}$$

独立于 x，并且因为 $\rho < 1$，这可以任意小。

□

因此，如果一个幂级数在单一点 $x_0 \ne a$ 处收敛，那么它在开区间 $I = (a-\eta, a+\eta)$ 上收敛，并且在 I 的每个闭子区间上均匀收敛。在 I 上，该级数表示一个可微函数，其导数可以逐项计算。逆否命题是，如果一个幂级数在 x_0 处发散，那么它对于所有 x 也发散，其中 $|x-a| > |x_0 - a|$。

给定一个幂级数，实数集被划分为两个集合；级数收敛的点和级数发散的点。定理11.14表明前者是以 {v*} 为中心的区间；它自然被称为幂级数的收敛区间。令 {v*} 0 为收敛区间上的 {v*} 的上确界。幂级数的收敛区间必须是以下之一:

- 单例 $\{a\}$ (if $R = 0$);
- 一个以 a (为中心的有限区间，如果 $0 < R < \infty$);
- 所有 $\mathbf{R}$ (如果 $R = \infty$)。

示例11.16演示了收敛区可以是开区间、半开区间或闭区间。

幂级数的收敛半径仅取决于系数序列(a_k)，而与中心无关。由于Hadamard，有一个公式，它以系数的形式给出半径：

$$\frac{1}{R} = \lim_{n\to\infty} \sup_{k\geq n} \sqrt[k]{|a_k|}.$$

我们不需要这种普遍性，而将开发仅适用于某些系数序列的更简单的公式，但包括我们将遇到的所有这些。

定理11.15。设(a_n)为一个序列，并且假设

$$L = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

存在。如果 $L > 0$，那么 $R = 1/L$ 是幂级数的收敛半径

$$\sum_{n=0}^{\infty} a_n (x-a)^n.$$

如果 $L = 0$，则幂级数对所有 $x \in \mathbf{R}$ 收敛。

证明。我们使用假设来比较幂级数与几何级数。假设$0 < |x-a| < R$，R如定理中所述。然后

$$\lim_{n\to\infty} \left| \frac{a_{n+1}(x-a)^{n+1}}{a_n(x-a)^n} \right| = \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \cdot |x-a| < 1.$$

因为极限严格小于1，它可以写成$1-2\varepsilon$，其中 $\varepsilon > 0$。从收敛的定义来看，存在 $N \in \mathbf{N}$ 使得

$$\left| \frac{a_{n+1}(x-a)^{n+1}}{a_n(x-a)^n} \right| < 1 - \varepsilon =: \rho \qquad \text{for } n > N.$$

归纳法证明 m 成立

$$|a_{N+m}(x-a)^{N+m}| \leq |a_N(x-a)^N| \, \rho^m, \qquad m \geq 0.$$

除了前 N 项之外，因此幂级数主要由收敛的几何级数的项支配。

一个完全类似的论据表明，如果 $|x-a| > R$，则幂级数发散；因此 R 是半径。 $\square$

定理11.15被称为比值测试。当适用时，它通常是计算幂级数收敛半径的最简单方法。请注意，比值不提供关于$a \pm R$处收敛的任何信息。必须始终单独手动检查端点的收敛性。

示例11.16 这里给出的例子以0为中心，以简化符号。比值测试适用于每个级数，但半径的计算留给您。

- 幂级数 $\sum_{k=0}^{\infty} x^k$ 的收敛半径为 1。当 $x = \pm 1$（收敛区间的端点）时，该级数发散，因为其项没有极限为零。因此，收敛区间为 $(-1, 1)$。

- 该级数 $\sum_{k=0}^{\infty} \frac{x^k}{k}$ 的半径为 1。在 $x = -1$ 处，该级数通过交错级数测试收敛，而在 $x = 1$ 处，该级数是调和级数，因此发散。收敛区间是半开区间 $[-1, 1)$。

- 该级数 $\sum_{k=0}^{\infty} \frac{x^k}{k^2}$ 的半径为 1。在每个端点，该级数通过与 2-级数的比较而绝对收敛。收敛区间是闭区间 $[-1, 1]$。

- 序列 $\sum_{k=0}^{\infty} \frac{x^k}{k!}$，见方程（11.1），半径为 $+\infty$，因此收敛区间为 $\mathbf{R}$。
- 该序列 $\sum_{k=0}^{\infty} k^k x^k$ 的半径为 0，因此收敛“区间”是单元素集合 $\{0\}$。

这些例子表明，为了找到收敛区间，首先计算收敛半径，然后检查端点。唯一不需要检查端点的情况是当没有端点时，要么因为 $R = 0$，要么因为 $R = +\infty$。□

实分析函数

定理11.14表明，如果以a为中心的幂级数半径为$R > 0$，那么在区间($a - R, a + R$)上，该级数表示x的连续函数。实际上，如果 $x_0 \in (a - R, a + R)$，那么存在一个$\delta > 0$，使得

$$[x_0 - \delta, x_0 + \delta] \subset (a - R, a + R),$$

并且序列的偏和函数在 $[x_0-\delta, x_0+\delta]$ 上均匀收敛，因此在这个区间上表示一个连续函数，并且可以逐项积分。由于 x_0 是任意的，幂级数在这个区间 $(a-R, a+R)$ 上表示一个连续函数，并且可以逐项积分。

实际上，更多是真实的；通过逐项微分得到的幂级数具有与原始级数相同的半径，这可以用来表明原始幂级数在$(a-R, a+R)$上表示一个可微函数。这个无害的事实可以被自我增强，表明幂级数在其收敛开区间上表示一个可无限微分的函数。

我们将不会在完全一般性下证明这些主张，尽管很容易修改下面的论据，用Hadamard公式替换比率测试中的出现。(自然地，还必须证明Hadamard公式通常给出幂级数的半径。)

定理11.17. 设$\sum_{k=0}^{\infty} a_k(x-a)^k$为一个幂级数，使得

$$\frac{1}{R} = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$$

存在且为正，设 $f:(a-R, a+R)\to \mathbf{R}$ 为级数的和。那么 f 是可微的；事实上，逐项求导的级数

$$\sum_{k=1}^{\infty} k a_k(x-a)^{k-1} = \sum_{k=0}^{\infty}(k+1)a_{k+1}(x-a)^k \tag{11.9}$$

具有半径 R，表示 $(a-R, a+R)$ 上的 f'。

证明。通过比值测试，导出级数的收敛半径为

$$\lim_{n\to\infty}\left|\frac{(n+1)a_{n+1}}{(n+2)a_{n+2}}\right| = \lim_{n\to\infty}\left(\frac{n+1}{n+2}\right)\left(\left|\frac{a_{n+1}}{a_{n+2}}\right|\right) = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|,$$

在区间$(a-R, a+R)$上，序列(11.9)表示一个连续的

函数 g 可以逐项积分。对于每个 $x \in (a-R, a+R)$，

$$\begin{aligned}\int_a^x g(t)\,dt &= \int_a^x \left(\sum_{k=1}^{\infty} ka_k(t-a)^{k-1}\right) dt \\ &= \sum_{k=1}^{\infty}\left(\int_a^x ka_k(t-a)^{k-1}\,dt\right) = \sum_{k=1}^{\infty} a_k(t-a)^k\Big|_{t=a}^{t=x} \\ &= \sum_{k=1}^{\infty} a_k(x-a)^k = f(x) - f(a).\end{aligned}$$

由于 g 是连续的，第二基本定理意味着 $f' = g$ 在 $(a-R, a+R)$ 上。

□

逻辑稍微有些微妙，并且需要重复一次。我们从幂级数 $\sum_k a_k x^k$ 开始，其收敛半径（由比值测试给出）为 $R > 0$。逐项导出的级数 $\sum_k ka_k x^{k-1}$ 具有相同的收敛半径，并且通过单独的论证表示原始级数的导数。这确立了本章的主要目标，即证明收敛的幂级数可以像无限项的多项式一样进行操作。但如果这还不够，我们可以启动论证：如果 f 在一个开区间 I 上表示为幂级数，那么 f' 也在同一个区间上表示为幂级数。因此，f'' 在 I 上表示为幂级数，以此类推。形式上，归纳法表明函数 f 是无限可微的，并且可以通过逐项微分来找到连续的导数。

定义11.18 设$I \subset \mathbf{R}$为一个开区间。若函数$f: I \to \mathbf{R}$对于每个$a \in I$，都存在一个$\delta > 0$和一个以a为中心的幂级数，该级数对于所有$|x-a| < \delta$都收敛，则称该函数为实分析函数。

许多初等数学函数是实分析的。这种看似复杂的定义的原因是，一个单独的幂级数通常不足以在其整个定义域上表示一个解析函数。实分析函数将为我们提供一些关于无穷级数的显著评估，并在第13章中构建三角函数时至关重要。目前，我们满足于建立实分析函数的基本算术性质。

定理11.19。设f和g是区间I上的实分析函数。那么函数$f+g$和fg在I上是实分析的，且f/g在每个g非零的区间上是解析的。

证明。实解析性是一个局部性质，因此定理的真正内容是收敛幂级数的和、积或商可以表示为收敛幂级数（前提是分母不为零）。为了简单起见，假设所有级数都是以0为中心；用$x-a$代替x可以处理一般情况。写

$$f(x)=\sum_{i=0}^{\infty}a_i x^i, \qquad g(x)=\sum_{j=0}^{\infty}b_j x^j,$$

并且让 $p_n(x)$ 和 $q_n(x)$ 分别表示各自的 n 部分和。因为每个级数在某些非零 x 上收敛，存在一个 $\eta>0$，使得每个级数在区间 $[-\eta,\eta]$ 上可统一求和。

求和在练习11.3中完成。乘法运算使用绝对可求和级数的柯西乘积公式处理，参见第4章。形式上，我们乘以级数并收集k的项:

$$\left(\sum_{i=0}^{\infty}a_i x^i\right)\left(\sum_{j=0}^{\infty}b_j x^j\right)=\sum_{k=0}^{\infty}\left(\sum_{i=0}^{k}a_i b_{k-i}\right)x^k. \tag{11.10}$$

因为幂级数在其收敛区间内绝对收敛，所以右侧的级数等于左侧级数的乘积，在 $[-\eta,\eta]$ 上。

设 g 是解析的，并假设 $g(0)\neq 0$。为了证明 $1/g$ 在 0 处是解析的，我们将 g 写成幂级数，

$$g(x)=a_0+a_1x+a_2x^2+\cdots=\sum_{k=0}^{\infty}a_k x^k,$$

并且寻求一系列

$$h(x)=\sum_{n=0}^{\infty}b_k x^k$$

使得 $g(x)h(x)=1$ 对于所有 x 在 0 的某个邻域内。将乘积的系数写出来，并与 1 的幂级数相等，我们得到

$$b_0=\frac{1}{a_0}, \qquad \sum_{i=0}^{k}a_i b_{k-i}=0 \quad \text{for all } k\geq 1.$$

这个无限方程组可以通过递归求解。首先将其重写为

$$b_0 = \frac{1}{a_0}, \qquad b_k = -\frac{1}{a_0}\sum_{i=1}^{k} a_i b_{k-i} \quad \text{for } k \geq 1.$$

然后 $k = 1$ 得到 $b_1 = -(a_1/a_0)b_0 = -a_1/(a_0)^2$，$k = 2$ 得到

$$b_2 = -\frac{1}{a_0}(a_1 b_1 + a_2 b_0) = -\frac{1}{{a_0}^3}(a_0 a_2 - a_1^2),$$

等等。为了检查倒数级数的收敛性，使用“几何级数技巧”：用 $g(x) = a_0\big(1 - \phi(x)\big)$，其中 $\phi(x) = O(x)$。然后

$$\frac{1}{g(x)} = \frac{1}{a_0} \cdot \frac{1}{1 - \phi(x)} = \frac{1}{a_0}\sum_{k=0}^{\infty} \phi(x)^k,$$

在某些0的邻域内。参见例11.23以了解具体应用。

□

幂级数和 O 符号表示法交互得非常好，这解释了在许多符号计算机软件中使用 O 符号表示法的原因。我们只陈述以 0 为中心的幂级数的结果；对于以 a 为中心的级数，有一个明显的修改。

推论11.20.如果$f(x) = \sum\limits_{k=0}^{\infty} a_k x^k$在0处是解析的，那么对于每个正整数 n，我们有

$$f(x) = \sum_{k=0}^{n} a_k x^k + O(x^{n+1}).$$

证明。根据定理11.19，

$$f(x) - \sum_{k=0}^{n} a_k x^k = \sum_{k=n+1}^{\infty} a_k x^k = x^{n+1}\sum_{j=0}^{\infty} a_{n+j+1} x^j,$$

并且 他右侧的和是实分析的，因此 $O(1)$ 靠近0。 □

以下结果称为幂级数的恒等定理。其实际后果是，如果且仅若它们的系数相同，以 a 为中心的两个幂级数定义了相同的函数。再次强调，我们只对以 0 为中心的级数陈述此结果。

推论11.21.如果$\sum_{k=0}^{\infty} a_k x^k \equiv$在0附近为0，那么对于所有$k$，$a_k =$为0。

证明。我们证明逆命题。假设并非所有 a_k 都为零，并令 a_n 为第一个非零系数。根据定理11.19和推论11.20，

$$\sum_{k=0}^{\infty} a_k x^k = x^n \sum_{j=0}^{\infty} a_{n+j} x^j = x^n\big(a_n + O(x)\big).$$

自 $a_n \neq 0$ 起以来，括号内的项在某些 0 的邻域内非零，因此该级数在某些关于 0 的删除区间内非零。

□

定理11.19是寻找倒数幂级数计算技巧的基础。两个例子将用于说明有用的方法。

示例11.22设$g(x) = 1 - x$，它在0处显然是解析的且非零。g的系数是 $a_0 = 1$，$a_1 = -1$。其倒数有幂级数

$$h(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots ,$$

在所有 $k \geq 1$ 的条件下，$b_0 = 1$ 和 $0 = a_0 b_k + a_1 b_{k-1} = b_k - b_{k-1}$。我们立即得出结论，所有 $1/g$ 的系数都等于 1:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots = \sum_{k=0}^{\infty} x^k.$$

这只是几何级数公式的表达。 □

定理11.19证明中的几何级数技巧是一种很好的计算方法，如果只需要前几个系数，就可以找到倒数级数的系数。

示例11.23 假设 $f(x) = 1 - \frac{1}{2}x^2 + \frac{1}{4!}x^4 + O(x^6)$（一个有趣的级数以这种方式开始），并且我们希望找到 $1/f$ 的幂级数。写出 $f(x) = 1 - \phi(x)$，然后使用几何级数公式:

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

此过程是合理的，因为 $\phi(0)=0$，所以我们通过连续性在0的某个邻域内得到 $|\phi(x)|<1$。代数给出

$$\begin{aligned}
\phi(x) &= \frac{1}{2}x^2 - \frac{1}{4!}x^4 + O(x^6) \\
\phi(x)^2 &= \frac{1}{4}x^4 + O(x^6) \\
\phi(x)^3 &= O(x^6) \\
\frac{1}{f(x)} &= 1 + \frac{1}{2}x^2 + \frac{5}{24}x^4 + O(x^6).
\end{aligned}$$

我们无法使用给定的信息得到更多项，但如果 f 已知至 $O(x^n)$，则我们保证可以找到相同阶数的倒数。□

11.4 近似序列

幂级数并非唯一的有趣逼近序列。在本节中，我们介绍一些更专业的例子。

皮卡尔迭代

递归定义的函数序列在逼近微分方程解的过程中自然出现。一维空间中的通用一阶微分方程可以写成

$$y'(t) = f\big(t, y(t)\big), \qquad y(t_0) = y_0. \tag{11.11}$$

第二个方程称为初始值，通常被视为指定时间 t_0 时 y 的值。将此方程的左右两边从 t_0 到 t 进行积分，得到等价的积分方程

$$y(t) = y(t_0) + \int_{t_0}^{t} f\big(s, y(s)\big)\,ds. \tag{11.12}$$

右侧的这个方程式可以被视为一个关于 t 的函数，它还依赖于函数 y 。[2] 换句话说，存在一个

[2] The right-hand side depends upon f, but we regard f as fixed in this discussion.

操作符 P 将函数映射到函数；函数 y 映射到由以下定义的函数 Py

$$Py(t) = y(t_0) + \int_{t_0}^{t} f\big(s, y(s)\big)\,ds.$$

一个(11.12)的解恰好是一个函数y，使得$Py = y$，即P的定点。受递归定义的数值序列的启发，我们希望从初始猜测开始迭代P来找到P的定点。我们的初始猜测是常函数y_0，并将$y_{n+1} = Py_n$设为$n \geq 0$，即

$$y_{n+1}(t) = y_0 + \int_{t_0}^{t} f\big(s, y_n(s)\big)\,ds. \tag{11.13}$$

序列$(y_n)_{n=0}^{\infty}$的项被称为初值问题(11.11)的Picard迭代。在f的相当温和的约束下，Picard迭代的序列在t_0的某个邻域内一致收敛到P的固定点。

示例11.24 为了在特定例子中感受Picard迭代，考虑初值问题

$$y' = y, \qquad y(0) = 1, \tag{11.14}$$

其解为自然指数函数。这里 $f(t, y) = y$，因此 $f\big(s, y_n(s)\big) = y_n(s)$。我们对某个关于0的区间内的所有 t 进行初始猜测 $y_0(t) =$ 为1。方程（11.13）给出

$$\begin{aligned}
y_1(t) &= 1 + \int_0^t y_0(s)\,ds = 1 + t \\
y_2(t) &= 1 + \int_0^t y_1(s)\,ds = 1 + t + \frac{t^2}{2\cdot 1} \\
y_3(t) &= 1 + \int_0^t y_2(s)\,ds = 1 + t + \frac{t^2}{2\cdot 1} + \frac{t^3}{3\cdot 2\cdot 1},
\end{aligned}$$

等等。看来我们正在从（11.1）中恢复我们的老朋友，并且确实可以通过对 n 进行归纳来轻松检查

$$y_n(t) = \sum_{k=0}^{n} \frac{t^k}{k!} \qquad \text{for } n \geq 0.$$

形式上，如果我们迭代无限次，我们得到幂级数（11.1）。将这个级数当作多项式进行微分，我们发现（11.1）中每一项的导数是前一项，所以（至少在形式上）$y' = y$。□

近似恒等式

本节我们研究函数的“卷积积”，这是信号处理中一个相当重要的运算。卷积积在逼近理论中也有重要的应用。我们将证明一个惊人的事实：在闭区间上的连续函数可以用多项式序列进行一致逼近。

形式上，f 和 g 的卷积积定义为

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\,dt. \tag{11.15}$$

然而，这个广义积分通常不收敛。我们的第一个任务是限制对函数的适当向量空间的关注。

一个函数 $f \in \mathscr{F}(\mathbf{R}, \mathbf{R})$ 是紧支集的，如果存在一个 R 在 $\mathbf{R}$ 中，使得 $f(x) = 0$ 对于 $|x| > R$。换句话说，f 在某个闭区间之外恒等于零。在 $\mathbf{R}$ 上的连续、紧支集函数的集合表示为 $\mathscr{C}_c^0(\mathbf{R})$。这个集合是 $\mathscr{F}(\mathbf{R}, \mathbf{R})$ 的向量子空间：如果 f 和 g 是连续且紧支集的，那么 $f + g$ 和 cf 显然是连续且紧支集的。

引理11.25. 如果 f 和 g 在 $\mathscr{C}_c^0(\mathbf{R})$ 中，那么 $f * g$ 也在其中。

证明。假设 $f(x) =$ 为 0 对于 $|x| > R_1$ 和 $g(x) =$ 为 0 对于 $|x| > R_2$。我们断言，如果 $|x| > R_1 + R_2$，那么对于所有 t，$f(t)g(x-t) =$ 为 0，因此肯定 $(f * g)(x) =$ 为 0。但反向三角不等式表明，如果 $|t| \leq R_1$，那么

$$|x - t| \geq \big||x| - |t|\big| > (R_1 + R_2) - R_1 = R_2.$$

换句话说，如果 $f(t) \neq 0$，那么 $g(x-t) = 0$。 □

一个连续、紧支集的函数是“信号”的模型，仅在有限的时间间隔内非零，或“滤波器的响应曲线”。将信号 f 与滤波器响应 g 卷积给出输出信号。

示例11.26设$g = (b-a)^{-1}\chi_{[a,b]}$是$[a, b]$中的“单位脉冲”。那么 $g(x-t) =1/(b-a)$当$x - b \leq t \leq x - a$时，否则为零，所以

$$(f * g)(x) = \int_{-\infty}^{\infty} f(t)g(x-t)\,dt = \frac{1}{b-a}\int_{x-b}^{x-a} f(t)\,dt$$

对于所有 $f \in \mathscr{C}_c^0(\mathbf{R})$。右侧的积分是 $[x-b, x-a]$ 上 f 的平均值。

更一般地，如果 $g(x) >$ 在 $x \in (a, b)$ 处为 0，在其他地方为零，并且如果 g 包围一个单位面积，那么 $f * g$ 可以看作是“平均” f 在长度为 $b - a$ 的区间上的结果。□

尽管其看似奇怪的定义，卷积积满足两个美丽的恒等式:

命题11.27. 设f_1，f_2和f_3是连续且紧支集的。那么$f_1 * f_2 = f_2 * f_1$和($f_1 * f_2$)$*$ $f_3 = f_1 * (f_2 * f_3)$。

在文字上，卷积是交换的和结合的。交换性的证明留给你们，见练习11.14。结合性需要两个变量的积分结果（“富比尼定理”），仅从概念上提及；在这本书中我们不使用结合性。

狄拉克 δ-函数

狄拉克的 δ-函数是一个具有以下性质的虚构“函数”:

- $\displaystyle\int_{-\infty}^{\infty} \delta(t)\, dt = 1.$
- $\delta(t) = 0$ if $t \neq 0.$

例如，$\delta(x - t)$ 是一个“在 x 处集中的单位脉冲”，因此形式上

$$(f * \delta)(x) = \int_{-\infty}^{\infty} f(t)\delta(x - t)\, dt = f(x);$$

卷积与 δ 是恒等映射。

很遗憾，上述性质在逻辑上是不相容的：如果一个函数满足第二个性质，那么它的积分是0。然而，物理学家和电气工程师发现这个“函数”在他们的工作中非常有用，如果被数学家追问，通常会调和上述两个性质，说：“是的，但是δ(0) $=$ ∞”。工程师甚至愿意将δ-函数视为“Heaviside阶跃函数”的导数，该函数定义为：当x <0时为0，当x ≥0时为1！

δ-函数的实用性强烈暗示了一个精确的数学概念潜伏其中。物理学家在量子力学的早期就开始使用δ-函数，在20年内，数学家至少找到了三种严谨的解释。我们介绍其中之一

这些，“近似恒等式”。与其将狄拉克 δ 视为一个单一函数，我们构建一个近似满足上述条件的序列。

定义11.28 一个非负函数序列({v*})是近似恒等式，如果:

- For all $n \in \mathbf{N}$, δ_n is integrable, and $\displaystyle\int_{-\infty}^{\infty} \delta_n = 1$.
- For every $\beta > 0$, $\displaystyle\lim_{n\to\infty} \int_{-\beta}^{\beta} \delta_n = 1$.

第二个条件形式化了“积分集中在0”的想法。我们仅为了简化假设了$\delta_n \geq 0$；经过更多的工作，下面的证明可以扩展以去除非负性假设。$f * \delta = f$的正式计算采取以下严谨形式；再次，你应该想到歌德的引言！

定理11.29。设$f \in \mathscr{C}_c^0(\mathbf{R})$，并设$(\delta_n)$为一个近似恒等式。由$f_n = f * \delta_n$定义的序列$(f_n)$一致收敛于$f$。

证明。由卷积积的交换律，我们有

$$f_n(x) = \int_{-\infty}^{\infty} f(x-t)\delta_n(t)\,dt, \qquad f(x) = \int_{-\infty}^{\infty} f(x)\delta_n(t)\,dt$$

对于所有 n。积分的线性、δ_n 的非负性和三角不等式意味着

$$\begin{aligned}\bigl|f(x) - f_n(x)\bigr| &= \left|\int_{-\infty}^{\infty} \bigl[f(x) - f(x-t)\bigr]\,\delta_n(t)\,dt\right| \\ &\leq \int_{-\infty}^{\infty} \bigl|f(x) - f(x-t)\bigr|\,\delta_n(t)\,dt.\end{aligned}$$

我们希望表明，通过选择 n 足够大，可以独立于 x 使其变得很小。对于每一个 β，最后一个积分可以分解为

$$\int_{-\beta}^{\beta} \bigl|f(x) - f(x-t)\bigr|\,\delta_n(t)\,dt + \int_{|t|\geq\beta} \bigl|f(x) - f(x-t)\bigr|\,\delta_n(t)\,dt.$$

想法是对于小的 t，f 的增量很小（因此第一项很小），而对于大的 t，贡献很小，因为 δ_n 集中在 0。

修复 $\varepsilon > 0$。因为 f 是连续且紧支撑的，f 在 $\mathbf{R}$ 上是一致连续的：存在 $\beta > 0$ 使得 $|t| \le \beta$ 蕴含 $|f(x) - f(x-t)| < \varepsilon$。这也意味着 f 是有界的：存在 $M > 0$ 使得对于所有 x，$t \in \mathbf{R}$，有
$|f(x) - f(x-t)| \le M$。使用定义 11.28 中的第二个性质来选择 N 使得

$$\int_{|t|\ge\beta} \delta_n(t)\,dt < \frac{\varepsilon}{M} \qquad \text{for } n \ge N.$$

使用这些选择，如果 $n \ge N$，那么

$$\begin{aligned}\int_{-\beta}^{\beta} \underbrace{\left|f(x) - f(x-t)\right|}_{<\varepsilon} \delta_n(t)\,dt + \int_{|t|\ge\beta} \underbrace{\left|f(x) - f(x-t)\right|}_{\le M} \delta_n(t)\,dt \\ < \int_{-\beta}^{\beta} \varepsilon\delta_n(t)\,dt + \int_{|t|\ge\beta} M\delta_n(t)\,dt \le \varepsilon + M\cdot\frac{\varepsilon}{M} = 2\varepsilon\end{aligned}$$

独立于 x。 □

韦尔斯特拉斯逼近定理

我们已经看到两种使用多项式序列来逼近函数 f 的方法。第一种，拉格朗日插值（第3章），构造在指定点与 f 一致的多项式。然而，无法保证在整个 f 的定义域上逼近是好的。第二种，幂级数的部分和，仅适用于实分析函数。第14章致力于仔细研究解析函数及其逼近多项式。在本节中，我们提到了第三种逼近类型，这是由K.魏尔斯特拉斯提出的，其中使用多项式序列来统一逼近一般连续函数。该序列是通过与适当选择的近似恒等式卷积构建的。

定理11.30. 设 $f : [a, b] \to \mathbf{R}$ 是连续的。存在一个多项式序列 (p_n)，使得 $(p_n) \to f$ 在 $[a, b]$ 上一致。

证明。通过添加一个线性多项式（如平均值定理的证明中所示），我们可以假设 {v*} 0：如果我们能通过多项式逼近 {v*}，则我们可以逼近

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

通过多项式。其次，我们可以用 $x = a + (b - a)u$ 替换，简化为 $[a, b]$ = $[0, 1]$。更详细地说，如果我们写 $\widetilde{\phi}(u) = \phi(x)$，那么 $(p_n) \to f$ 在 $[a, b]$ 上均匀当且仅当 $(\widetilde{p_n}) \to \widetilde{f}$ 在 $[0, 1]$ 上均匀。因此，只需证明对于满足 $f(0) = f(1) = 0$ 的连续函数 $f\colon [0, 1] \to \mathbf{R}$ 的定理即可。

我们首先构建一个由分段多项式函数组成的近似恒等式。令 {v*} 定义为

$$c_n \int_{-1}^{1} (1 - x^2)^n \, dx = 1,$$

并且设置

$$\delta_n(x) = \begin{cases} c_n(1 - x^2)^n & \text{if } -1 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

第15章中我们将显式评估 c_n，但到目前为止以下估计就足够了：

$$\begin{aligned} \frac{1}{c_n} &= \int_{-1}^{1} (1 - x^2)^n \, dx = 2 \int_0^1 (1 - x^2)^n \, dx \\ &> 2 \int_0^{1/\sqrt{n}} (1 - x^2)^n \, dx > 2 \int_0^{1/\sqrt{n}} (1 - nx^2) \, dx = \frac{4}{3\sqrt{n}}; \end{aligned}$$

第二个不等式是练习2.5部分(b)的结果。我们立即得出结论，(δ_n)是一个近似恒等式，因为

$$0 \le \delta_n(x) \le \tfrac{3\sqrt{n}}{4}(1 - x^2)^n,$$

and 上界在 $[-\beta, \beta]$ 上对所有的 $\beta > 0$ 一致收敛到 0。现在设 $p_n = f * \delta_n$；因为 $f = 0$ 在 $[0, 1]$ 外，我们有

$$p_n(x) = \int_{-\infty}^{\infty} f(t)\delta_n(x - t) \, dt = \int_0^1 f(t)\delta_n(x - t) \, dt.$$

如果 $x \in [0, 1]$，那么 $x - t \in [-1, 1]$ 对于 $0 \le t \le 1$，因此 $\delta_n(x - t) = c_n\big(1 - (x - t)^2\big)^n$。因此，被积函数 $f(t)\delta_n(x - t)$ 是关于 x 的多项式，其系数是 t 的连续函数；在 t 从 0 到 1 的积分中，我们发现 p_n 是关于 x 的多项式。

根据定理11.29，$(p_n) \to f$在$[0, 1]$上均匀。 □

几个要点值得强调。首先，近似恒等式是明确的，因此对于特定的函数 f，存在一种有效的计算方法来找到逼近多项式。其次，

连续函数空间庞大而复杂，而多项式空间简单而明确。从概念上讲，魏尔斯特拉斯定理类似于一个简单的数字集合$\mathbf{Q}$ (在)一个庞大、复杂的集合$\mathbf{R}$ (中的密集性。最后，一个一般的连续函数在无处可导，但在每个闭、有界区间上被光滑函数均匀逼近。

练习

练习11.1 令 $\phi : \mathbf{R} \to \mathbf{R}$ 如例11.3中所述，并令 $(h_n)_{n=0}^{\infty}$ 是由 $h_n(x) = n\phi(nx)$ 定义的序列。仔细在同一组坐标轴上绘制 ϕ 和 h_n $(n = 1, 2, 3)$ 的图形。证明 (h_n) 在点wise上收敛到零函数。$\diamond$

练习11.2 设 $\phi : \mathbf{R} \to \mathbf{R}$ 连续，并定义 (f_n) 为

$$f_n(x) = \frac{1}{n}\phi(x), \qquad \text{for } x \in \mathbf{R}.$$

(a{v*20} 证明在紧集上 ({v*21}) {v*22} 0 均匀成立。

(b) 证明当且仅当 f 有界时，$(f_n) \to 0$ 一致成立。

(c) 假设

$$\sum_{k=0}^{\infty} a_k x^k$$

在 $\mathbf{R}$ (上逐点收敛，即具有无穷收敛半径)。证明如果且仅当和是有限的，则部分和在 $\mathbf{R}$ 上一致收敛。

$\diamond$

练习11.3 证明如果 f 和 g 在0的邻域内是实解析的，那么 $f + g$ 也是实解析的。提示：所有必要的估计都可以在文本的早期部分找到。$\diamond$

练习11.4 计算序列 $1/(1 + x)$ 和 $1/(1 + x^2)$ 的幂级数。收敛半径是多少？使用级数的和与积公式来验证

$$\frac{1}{1-x} + \frac{1}{1+x} = \frac{2}{1-x^2} \quad \text{and} \quad \frac{1}{1-x} \cdot \frac{1}{1+x} = \frac{1}{1-x^2}.$$

$\diamond$

练习11.5使用倒数技巧计算以下级数序列的{v*}

$$\frac{1}{1-x^2+x^4}$$

最高包括6次项。 $\diamond$

练习11.6 使用乘积和几何级数公式计算以下幂级数

$$\frac{1+x}{1-x}.$$

什么是半径? $\diamond$

练习11.7 让

$$f(x)=1+x+\frac{1}{2!}x^2+\frac{1}{3!}x^3+\cdots=\sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

计算收敛半径，并证明 $f'=f$ 在其收敛区间上。求 $\big(f(x)-1\big)/x$ 的幂级数，并使用你的答案计算倒数级数，$x/\big(f(x)-1\big)$，直到（包括）3次幂的项。

提示：在计算倒数技巧中的幂时，不需要携带次数大于3的项。

$\diamond$

练习11.8 设$f(x)$为前一个问题中的级数。使用级数的乘积公式和二项式定理来证明

$$f(x)f(y)=f(x+y)\quad\text{for all } x,\, y\in\mathbf{R}.$$

这证实了你已经知道的事情吗? $\diamond$

练习11.9 几何级数公式 sa ys

$$\frac{1}{1-x}=1+x+x^2+x^3+\cdots=\sum_{k=0}^{\infty}x^k.$$

这个方程在形式上如果 $x=1$ 时意味着什么？如果 $x=-1$ 呢？ $x=2$ 呢？这些公式中的任何一个有意义吗？解释一下。$\diamond$

练习11.10 对几何级数公式（如前一个练习中给出的）在开区间$(-1,1)$上求导。使用结果来找到幂级数的封闭形式表达式

$$f(x)=\sum_{k=1}^{\infty}kx^k,\quad g(x)=\sum_{k=1}^{\infty}k^2x^k,\qquad -1<x<1.$$

类似地，从0到x对几何级数进行积分，并将结果表示为闭式形式。对于哪个x，得到的公式是正确的？ $\diamond$

练习 1 1.11 找出所有实数 x，使得以下条件成立 rrect:

$$\frac{1}{x}+\frac{1}{x^2}+\frac{1}{x^3}+\cdots=\sum_{k=1}^{\infty}\left(\frac{1}{x}\right)^k=\frac{\frac{1}{x}}{1-\frac{1}{x}}=\frac{1}{x-1}.$$

我们可以得出结论吗? $-\sum_{k=0}^{\infty} x^k = -\frac{1}{1-x} = \sum_{k=1}^{\infty}\left(\frac{1}{x}\right)^k$? $\diamond$

练习11.12 找出所有实数 x，使得$\sum_{k=0}^{\infty} e^{kx}$ 收敛，并以闭式表达和。 $\diamond$

练习11.13 求以下幂级数的半径:

$$\text{(a)}\ \sum_{k=1}^{\infty}\frac{(-x)^k k!}{k^k} \qquad \text{(b)}\ \sum_{k=1}^{\infty} x^{k!}$$

$\diamond$

练习11.14 证明命题11.27的交换性部分。 $\diamond$

练习11.15 设$(f_n)_{n=1}^{\infty}$为非递减函数序列，并假设级数$\sum_n |f_n(0)|$和$\sum_n |f_n(1)|$收敛。

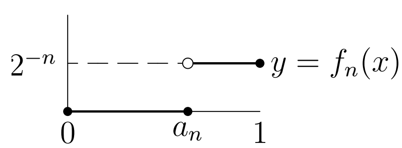
(a) 证明对于每个 $x \in [0, 1]$，级数$\sum_{n=1}^{\infty} f_n(x)$ 是收敛的。

(b) 证明函数 $f := \sum_n f_n$ 是非减函数。

(©) 证明部分(a)中的收敛性在[0, 1]上是一致的。

(d) 设 $(a_n)_{n=1}^{\infty}$ 是 [0, 1] 中的一个序列，并且假设项是不同的: $a_n \neq a_m$ 如果 $n \neq m$。定义

$$f_n(x)=\begin{cases}0 & \text{if } 0 \le x \le a_n\\ 2^{-n} & \text{if } a_n < x \le 1\end{cases}$$



证明 f 在 $x \in [0, 1]$ 处不连续当且仅当存在某个 k 使得 $x = a_k$。

(e) 证明存在一个增函数 {v*} : [0, 1] → [0, 1]，它在每个 $x \in \mathbf{Q} \cap [0, 1]$ 处都是不连续的。

$\diamond$

第12章

对数和指数

除了“病态”或分段定义的例子之外，自然对数函数log及其逆函数自然指数函数exp是本书中首先研究的非代数函数。它们的重要性无法用几句话概括，尽管它们在自然科学中的普遍性可以相当简单地解释：自然指数函数出现在任何某种数量的变化率与该数量本身成比例的情况中。人口的增长率大致与它们的规模成比例；金钱的积累利息率与本金成比例；在良好的近似下，许多化学反应的速率与反应物的浓度成比例；放射性核素以随机方式衰变，因此每秒钟的衰变数与样本中核素的数量成比例。在每种情况下，时间t时存在的“物质”量将由t的指数函数很好地近似。

对数在任何数量在大范围内变化的场合都是一种方便的语言，用比率来描述。声音波携带的能量或溶液中氢离子的浓度是那些在现实情况下跨越许多数量级的数量。对数单位（如分贝或pH）使这些数量更容易管理。（例如，人类耳朵能够忍受的最响亮的声音比耳语大数十亿倍，但我们方便地谈论130分贝与30分贝。）

对数和指数函数有公理化定义（见下文），这在历史上是它们发现的基础。我们的处理从逻辑上颠倒了历史顺序。

12.1 自然对数

历史上，“对数”是一个函数 L，非恒等于零，它将乘法转换为加法，在以下意义上：

$$L(xy) = L(x) + L(y) \qquad \text{for all positive, real } x \text{ and } y. \tag{12.1}$$

在学校课程中，我们假定这样的函数存在，并且这个等式被视为公理。我们现在可以引入并研究具有这种性质的函数，但对于我们来说，之前的等式将是一个定理。

自然对数是定义为函数 $\log : (0, \infty) \to \mathbf{R}$ 的对数

$$\log x = \int_1^x \frac{1}{t}\, dt \quad \text{for } x > 0. \tag{12.2}$$

积分的性质立即表明 $\log(1) = = 0$。第二个基本定理表明对数函数可导，并且

$$\log' x = \frac{1}{x}, \qquad x > 0. \tag{12.3}$$

因此，在$(0, \infty)$上对数是增加的，并且特别地，如果$x > 1$，则为正，见图12.1和12.2。

定理12.1. 对于所有 x，$y > \geqslant 0$，有 $\log(xy) = \log x + \log y$。

证明。如果 x 和 y 是正实数，那么

$$\begin{aligned}
\log(xy) &= \int_1^{xy} \frac{dt}{t} = \int_1^x \frac{dt}{t} + \int_x^{xy} \frac{dt}{t} \\
&= \int_1^x \frac{dt}{t} + \int_1^y \frac{du}{u} = \log(x) + \log(y).
\end{aligned}$$

第二积分的极限变化通过设定 $t = xu$ (记住“x 是一个常数”) 并调用练习7.13来证明。 □

在莱布尼茨记号中，对数性质（12.1）由被积函数 dt/t 的尺度不变性得出：

$$\frac{dt}{t} = \frac{d(xu)}{(xu)} = \frac{du}{u} \qquad \text{for all } x > 0.$$

很难证明（练习12.1）在乘以一个整体常数的情况下，dt/t是唯一具有所需不变性性质的连续积分函数。

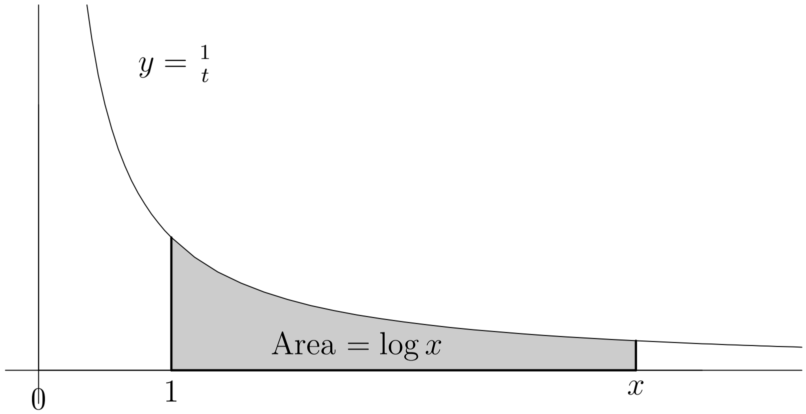


图12。 1: 自然对数的定义为 积分

方程（12.1）意味着对于所有 $x >> 0$，有 $\log(1/x) = -\log x$，并且（通过 p 的归纳）对于所有 $p \in \mathbf{N}$，有 $\log(x^p) = p \log x$。设定 $x = y^{1/q}$ 为 $q \in \mathbf{N}$ 并组装之前的观察结果，可以得出

$$\log(y^r) = r \log y \qquad \text{for all real } y > 0, \text{ all } r = \tfrac{p}{q} \text{ rational.} \tag{12.4}$$

实数对数log 2是正的，且对于所有整数n，$\log(2^n) = n \log 2$。根据$\mathbf{R}$的阿基米德性质，对数函数可以取任意大的（正的和负的）值。由于对数是连续的，它具有中间值性质，因此我们得出结论，对数将$(0, \infty)$映射到$\mathbf{R}$。因为对数是递增的，每个实数都是唯一正实数的对数。

注意仔细，对数函数的切线变得任意接近水平（因为 $\log' x = 1/x$），但该图形没有水平渐近线！

12.2 自然指数

历史上，指数函数是一个函数 $E : \mathbf{R} \to \mathbf{R}$，它满足

$$E(x + y) = E(x)E(y) \qquad \text{for all } x, y \in \mathbf{R}, \tag{12.5}$$

cf. (12.1). 回想一下，我们定义自然指数函数 $\exp : \mathbf{R} \to (0, \infty)$ 为对数（自然对数）的逆函数。这种定义的历史根源是：

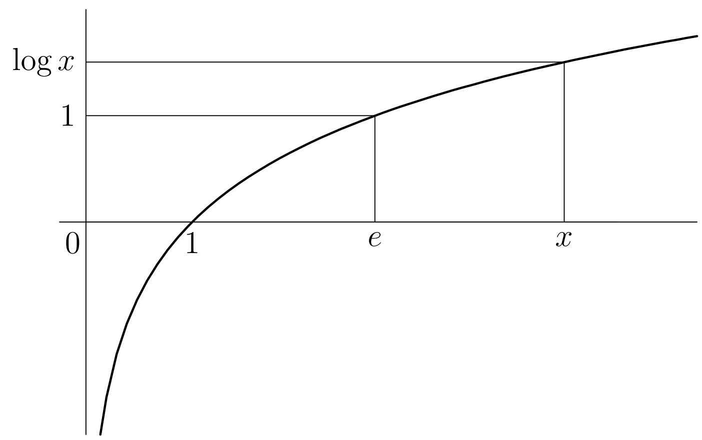


图12.2：自然对数函数的图像。

引理12.2. 如果 $L:(0,\infty)\to$ $\mathbf{R}$ 是对数函数，那么它的逆函数 E : $\mathbf{R}\to(0,\infty)$ 是指数函数。

证明。根据定义，对数满足对所有正数 {v*} 的恒等式 $L(xy)=L(x)+L(y)$。如果我们写出 $x=E(u)$，$y=E(v)$ 并将对数恒等式应用 E，我们发现

$$E(u)\cdot E(v)=xy=E\big(L(x)+L(y)\big)=E(u+v)\qquad\text{for all } u \text{ and } v,$$

这是指数函数的特征属性。 $\square$

数字 e

数字 e:= 指数 1 是数学中最突出的常数之一。请注意，根据定义，log $e=1$; e 是唯一满足图 12.1 中区域面积为单位的实数。练习 7.17 (d) 显示 $2<e<4$；实际值大致为

$$e=2.718281828459045\ldots,$$

参见引理12.8。

12.3 exp 和 log 的性质

方程（12.4）表示对所有实数 $y>0$，所有有理数 r，有 $y^r=\exp(r\log y)$。左侧的表达式有一个纯粹代数定义（涉及

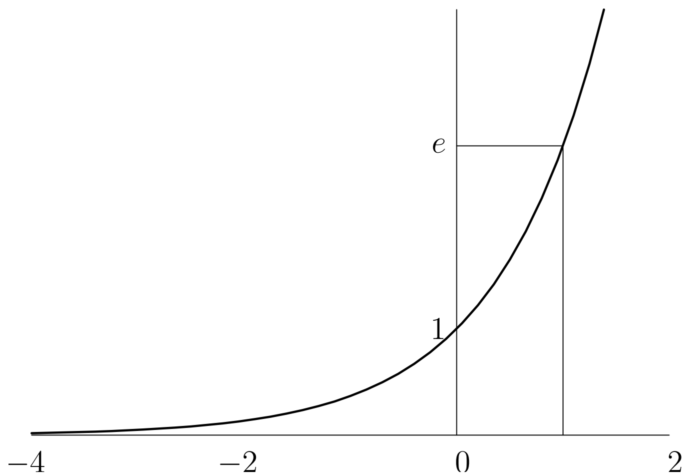


图12.3：自然指数函数的图像。

除了实数的乘法之外)，而右侧不是纯粹代数的，但具有对所有实数 r 都有定义的优点。这引导我们定义

$$x^r = \exp(r \log x) \qquad \text{for } x > 0,\, r \in \mathbf{R}. \tag{12.6}$$

特别地,

$$e^x = \exp x \quad \text{for all } x \in \mathbf{R}.$$

如果 $b >$ 为 0，则底数为 b 的指数函数，记作 $\exp_b$，定义为

$$\exp_b x = \exp(x \log b) \quad \text{for all } x \in \mathbf{R}.$$

名称将很快得到解释。

一个指数函数与其自身的导数成正比，指数函数的特点就是这一性质：{v*}

定理12.3. 设$b > 0$为固定。函数$\exp_b$是可微的，并且

$$\exp_b' x = (\log b) \exp_b x \quad \textit{for all } x \in \mathbf{R}.$$

相反，如果 $k \in \mathbf{R}$ 并且如果 f 是满足微分方程 $f' = kf$ 的可微函数，那么对于所有 $x \in \mathbf{R}$，有 $f(x) = f(0)e^{kx}$。

证明。回忆$\exp' = \exp$；证明是有教育意义的，此处重复。首先，对于所有 $x \in \mathbf{R}$，有$\log(\exp x) = x$。其次，log是可微的，并且其导数非零，因此其逆函数是

可微。因此，可以对方程 $x = \log(\exp x)$ 关于 x 求导。链式法则给出

$$1 = \log'(\exp x)\ \exp' x = \frac{1}{\exp x}\ \exp' x \qquad \text{for all } x,$$

证明 $\exp' = \exp$. 如果 $k \in \mathbf{R}$，链式法则意味着$\frac{d}{dx}e^{kx} = ke^{kx}$。由于 $\exp_b x = \exp(x \log b)$ 由定义，导数公式是直接的。

相反，假设 $f : \mathbf{R} \to \mathbf{R}$ 是一个可微函数，使得 $f' = kf$。通过 $g(x) = f(x)/e^{kx}$ 定义 g；这是合理的，因为 exp 在 $\mathbf{R}$ 上非零。函数 g 作为可微函数的商是可微的，商法则意味着

$$g'(x) = \frac{e^{kx} f'(x) - ke^{kx} f(x)}{(e^{kx})^2} = \frac{f'(x) - kf(x)}{e^{kx}} = 0 \quad \text{for all } x$$

由于 $f'(x) = kf(x)$ 对所有 x 根据假设。但这意味着 g 是一个常数函数，将 $x =$ 设置为 0 显示 $g(x) = f(0)$ 对所有 x。 □

推论12.4.设$r \in \mathbf{R}$为固定值。如果$f(x) = x^r$对于$x > 0$，那么f可微，且$f'(x) = rx^{r-1}$。在莱布尼茨记号中，

$$\frac{d}{dx}x^r = rx^{r-1} \qquad \textit{for all } r \in \mathbf{R}.$$

证明留给你们，练习12.6。定理12.3还暗示了exp的一些熟悉的代数性质：

定理12.5. 对于所有 x，$y \in \mathbf{R}$，$e^{x+y} = e^x e^y$ 和 $e^{xy} = (e^y)^x$。

证明。固定 $y \in \mathbf{R}$，并考虑由 $f(x) = e^{x+y}$ 定义的功能 f：$\mathbf{R} \to \mathbf{R}$。根据链式法则，$f$ 可微，且 $f' = f$。由于 $f(0) = e^y$，定理 12.3 对所有 x 都意味着 $e^{x+y} = e^x e^y$。

为了证明第二个断言，固定 y 并定义 g：$\mathbf{R} \to \mathbf{R}$ 通过 $g(x) = (e^y)^x$。应用定理12.3的第一部分，使用 $b = e^y$，我们注意到

$$g'(x) = \log(e^y)(e^y)^x = yg(x).$$

定理12.3的第二部分意味着 $g(x) = e^{xy}$，因为 $g(0) = 1$。作为额外的好处，我们已经证明了 $(e^y)^x = (e^x)^y$，因为每个项都等于 e^{xy}。

□

定理12.5证明了“以b为底数的指数函数”这个名称的合理性：

$$\exp_b(x) = \exp(x \log b) = \big(\exp(\log b)\big)^x = b^x \qquad \text{for all } x \in \mathbf{R}.$$

简单修改证明即可建立以下恒等式

$$b^{x+y} = b^x b^y, \qquad b^{xy} = (b^x)^y \qquad \text{for all } b > 0,\ x,\ y \in \mathbf{R}.$$

如果您试图直接从方程（12.6）证明这些恒等式，您将对刚才给出的论证的简单性印象深刻。微积分的一个优点是它能够将代数信息编码在具有唯一解的微分方程中。

您可能不止一次地想知道为什么我们不定义$0^0 = 1$。这里有一个形式为0^0的极限，它不等于1：

$$\lim_{x\to 0} \exp(\alpha/x^2)^{x^2} = \lim_{x\to 0}\Big(e^{(\alpha/x^2)}\Big)^{x^2} = e^{\alpha} \qquad \text{for all } \alpha < 0,$$

根据定理12.5。找到函数f和g的工作留给你，使得$\lim(f, 0) = \lim(g, 0) = \lim(f^g, 0) = 0$。

$_b$ 的逆函数是以 b 为底的对数，表示为 $\log_b$：

$$y = \log_b x \quad \text{iff} \quad x = b^y = \exp_b y.$$

定理12.5意味着对所有 $x > 0, y \in \mathbf{R}$，有 $\log_b(x^y) = y \log_b x$。

下一个结果表示以 b 为底的对数与自然对数成正比。因此，除了自然对数之外的对数函数在数学中只出现得很少。

命题12.6. 如果 $b > 0$，那么对于所有 $x \in \mathbf{R}$，有 $\log_b x = (\log x)/(\log b)$。函数 $\log_b$ 是可微的，并且

$$\log_b' x = \frac{1}{x \log b} \qquad \textit{for all } x > 0.$$

证明。说 $y = \log_b x$ 等于 $x = b^y = \exp(y \log b)$。对这个方程取自然对数得到 $y = (\log x)/(\log b)$。关于导数的陈述随即得出。

□

两种 exp 的表示

exp作为满足$f(0) = 1$的微分方程$f' = f$的解的特征，意味着两个引人注目的、非平凡的表示。

定理12.7. $e^x = \displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k\,!}$ 对所有 $x \in \mathbf{R}$ 成立。

证明。右侧的幂级数有 $a_k = 1/k!$，因此比值测试表明收敛半径是

$$\lim_{k\to\infty} \frac{a_k}{a_{k+1}} = \lim_{k\to\infty} \frac{(k+1)!}{k!} = \lim_{k\to\infty} k+1 = \infty.$$

相关函数 f 对所有 $x \in \mathbf{R}$ 定义，并且可微。通过逐项微分得到导数 f' :

$$f'(x) = \sum_{k=0}^{\infty} \frac{1}{k!}\, kx^{k-1} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = f(x)$$

对于所有 x。由于 $f(0) = 1$（除了第一个项之外的所有项都消失），定理12.3意味着 $f(x) = e^x$ 对于所有 x。 □

Corollary 12.8. $e = \displaystyle\sum_{k=0}^{\infty} \frac{1}{k!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots$

为了将这个推论的结论转化为有效的计算事实，如果我们通过将这个级数的有限项相加来近似 e，我们需要知道误差的大小。仅仅将前四项相加就表明 $2.6\bar{6} < e$，这已经比 $2 < e$ 有很大的改进。一个好的数值估计可以在练习12.23中找到。

定理12.9. $e^x =$ 当 $\displaystyle_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$ 时，对于所有 $x \in \mathbf{R}$。

证明。记住，在取极限时，x是固定的。我们首先明确允许“n”取任意正实数值，而不是整数值。变量替换 $h = 1/n$ 转换为

期望的极限

$$\begin{aligned}
\lim_{h\to 0^+}(1+xh)^{1/h} &= \lim_{h\to 0^+}\exp\left[\frac{1}{h}\log(1+xh)\right] \\
&= \lim_{h\to 0^+}\exp\left[x\frac{\log(1+xh)-\log 1}{xh}\right] \quad \text{since } \log 1 = 0 \\
&= \exp\left[x\lim_{h\to 0^+}\frac{\log(1+xh)-\log 1}{xh}\right] \quad \text{continuity of exp.}
\end{aligned}$$

设置 $\eta = xh$ 并注意到极限项是自然对数的牛顿商表明前面的表达式等于

$$\exp\left[x\lim_{\eta\to 0}\frac{\log(1+\eta)-\log 1}{\eta}\right] = \exp\left[x\ \log'(1)\right] = \exp x,$$

并且这是根据定义的 e^x。 □

定理12.9将自然指数函数描述为几何增长的极限。例如，如果x是储蓄账户的年利率，并且每年有n次复利，那么右侧的乘数给出了储蓄在一年内增加的因子。随着每年复利次数的无穷增长，余额不会在有限时间内变为无穷大。相反，如果允许1以连续复利的方式累积利息，那么在储蓄不进行复利的情况下翻倍所需的时间内，余额增加到2.72美元（四舍五入到最接近的美分）。

练习

练习12.1 设 $f:(0,\infty)\to \mathbf{R}$ 为一个连续函数，并定义

$$L(x) = \int_1^x f(t)\,dt.$$

证明如果 L 满足 (12.1)，则存在一个实数 k 使得对所有 $t>0$，有 $f(t) = k/t$。◇

练习12.2 设a，b和c为正实数。同意a^{b^c}等于$(a^b)^c$还是$a^{(b^c)}$更合理，或者这有关系吗？◇

练习12.3 证明对于所有 x，$y>0$，有 $x^{\log y} = y^{\log x}$。 ◇

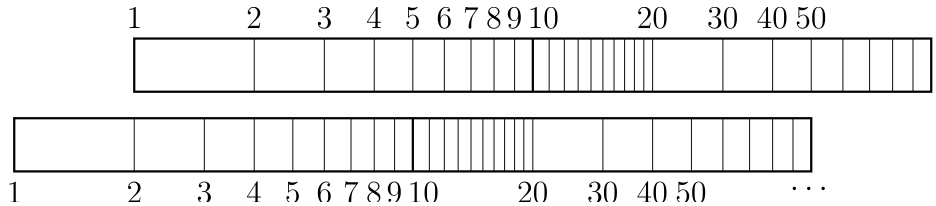


图12.4：标有常用（以10为底）对数的尺子。

练习12.4 解释如何使用两个对数刻度，如图12.4所示，来乘以数字。 ◇

练习12.5 设u为一个可微函数。求$\exp\circ u$的导数，以及$\log\circ|u|$在u非零点的导数。请用牛顿和莱布尼茨符号写出你的公式。◇

练习12.6 证明推论12.4。
提示：从实数 x^r 的定义开始。 ◇

练习12.7设$f(x) = e^{-x}(x^2-1)$对于$x \in \mathbf{R}$。

(a) 绘制 f 的图形，利用前两个导数的信息来确定 f 单调、凸和凹的区间。建议：引入符号常数以简化计算。(b) 方程 $f(x) = a$ 有多少个解？（你的答案将取决于 a。）

◇

练习12.8 如果f：$(0,\infty) \to \mathbf{R}$由$f(x) = x^x$定义，求$f'$。证明$f$有唯一的极小值，并找出极小值的位置和值。提示：为适当的函数u写$f(x) = e^{u(x)}$。◇

练习12.9 计算以下极限：

(a) $\lim\limits_{x\to 0^+} x\log x$ (b) $\lim\limits_{x\to 0^+} x^x$ (c) $\lim\limits_{x\to +\infty} x^{1/x}$ (d) $\lim\limits_{x\to +\infty} (1+x)^{1/x}$

◇

练习12.10 计算以下极限：

(a) $\lim\limits_{x\to +\infty} \dfrac{x}{\log x}(x^{1/x}-1)$ (b) $\lim\limits_{x\to 0^+} \dfrac{e-(1+x)^{1/x}}{x}$ ◇

练习12.11 设 n 为一个正整数。求值

(a) lim {v*}log $\frac{x)^n}{x}$ $x^n e^{-x}$ $x\to\infty$　　(b) $\lim\limits_{x\to+\infty} \frac{x^n}{e^x} = \lim$

使用您的结果证明对于所有 $\alpha > 0$，$\log x = o(x^\alpha)$ 近似 $+\infty$。找到一个类似的小-o 表达式用于幂函数和对数函数。$\diamond$

练习12.12 求解 $f_n(x) = x^n e^{-x}$ 在 $x \geq 0$ 时的最大值。（特别是，你必须证明最大值的存在。）$\diamond$

证明 $\displaystyle\int_0^\infty e^{-t^2}\,dt$ 收敛。 $\diamond$

练习12.14 让 $F_n(x) = \displaystyle\int_0^x t^n e^{-t}\,dt$。

(a) 使用分部积分法（练习10.12）找到以 F_{n-1} 为变量的 F_n 的递推公式。

(b) 使用部分(a)和关于 n 的归纳法来证明

$$F_n(x) = n!\Big(1 - e^{-x}\sum_{k=0}^{n}\frac{x^k}{k!}\Big).$$

(c) 评估不定积分 $\displaystyle\int_0^\infty t^n e^{-t}\,dt$

这个积分对所有实数 $n > -1$ 定义，表示为 $\Gamma(n+1)$。$\diamond$

练习 12.15

(a) 使用前一个练习的部分(b)和变量替换来找到一个公式，用于

$$\int_0^x t^n e^{-\alpha t}\,dt, \qquad \alpha \in \mathbf{R}.$$

(b) 证明广义积分 $\displaystyle\int_0^1 (\log u)^n\,du$ 收敛。

(c) 使用变量变换 $u = e^t$ 来评估部分 (b) 的不定积分。

◇

练习12.16 设 $f:(0,+\infty)\to\ \mathbf{R}$ 为一个增函数。回忆起 f 在 $[a,b]$ 上对所有 $0<a<b$ 都是可积的。

(a) 证明以下公式：{v*}

$$\sum_{k=1}^{n-1} f(k) \le \int_1^n f(t)\,dt \le \sum_{k=2}^{n} f(k)$$

对于所有 $n\in\mathbf{N}$。(一张草图会有帮助。比较命题7.23。)

(b) 在(a)部分取$f=$对数，证明

$$(n-1)! \le e\left(\frac{n}{e}\right)^n \le n!$$

对于所有 $n\in\mathbf{N}$。

(c) 求极限 $\lim\limits_{n\to\infty} \dfrac{(n!)^{1/n}}{n}$.

有一个关于 n 的更精确估计，称为斯特林公式。◇

练习12.17 确定以下哪些收敛（当然，需要证明）；不要尝试计算这些和！

(a) $\displaystyle\sum_{n=1}^{\infty}\frac{\log n}{n^{3/2}}$ (b) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{\log n}{n}$ (c) $\displaystyle\sum_{n=2}^{\infty}\frac{1}{n(\log n)^2}$

(d) $\displaystyle\sum_{n=1}^{\infty}\frac{(n+1)^n}{n^{n+1}}$ (e) $\displaystyle\sum_{n=1}^{\infty}\frac{n!}{n^n}$

提示：(a)：借用 n 的一小部分功率来消除对数。 ◇

练习12.18 设 n 为一个正整数。 ◇

练习12.19 定义 $f:\mathbf{R}\to\mathbf{R}$ 为

$$f(x)=\begin{cases}\exp(-1/x^2) & \text{if } x\ne 0\\ 0 & \text{if } x=0\end{cases}$$

证明对于所有 $k\in\mathbf{N}$，$f^{(k)}(0)$ 存在且等于零。
建议：首先使用对次数的归纳法证明对于每个多项式 p，

$$\lim_{x\to 0} p(1/x)\,f(x)=0.$$

然后通过归纳法证明 f 的每个导数都是这种形式。最后，证明 f 的导数在 0 处不能是断开的。$\diamond$

练习 se 12.20 Fix $b > 0$，并定义一个序列 $(x_n)_n^\infty$ =0 由

$$x_0 = 1, \quad x_{n+1} = b^{x_n} \text{ for } n \geq 0.$$

因此 $x_1 = b$，$x_2 = b^b$，$x_3 = b^{b^b}$ 以及如此等等。证明如果 $(x_n) \to \ell$，则 $b^\ell = \ell$。利用这个观察结果来确定序列收敛的 b 集合。然后证明 ℓ 是 b 的增函数，并找到 ℓ 的最大可能值。

两个人在争论。一个人说如果 $b = \sqrt{2}$，那么 $\ell = 2$，因为 $\sqrt{2}^2 = 2$；另一个人说 $\ell = 4$，因为 $\sqrt{2}^4 = 4$。谁——如果有人——是正确的，为什么？$\diamond$

练习12.21 设 n 和 x 为正整数。证明

$$n \leq \log_{10} x < n+1 \quad \text{iff} \quad 10^n \leq x < 10^{n+1},$$

如果 x 是一个有 $n+$ 位数字的整数。用文字来说，x 的以 10 为底的对数的整数部分比 x 的位数少一位。

哪个更大，$2^{2^{2^{2^2}}}$ 或者$10^{10^{100}}$？每个数字有多少位？$\diamond$

练习12.22 所有洛必达法则问题的源头：证明

$$\lim_{x\to\infty} e^{e^{e^{x+e^{-(a+x+e^x+e^{e^x})}}}} - e^{e^{e^x}} = e^{-a}$$

对于所有 $a \in \mathbf{R}$。 $\diamond$

练习12.23使用12.8推论中级数的前$n+1$项来估计e的错误是

$$(*) \qquad e - \sum_{k=0}^{n} \frac{1}{k!} = \sum_{k=n+1}^{\infty} \frac{1}{k!}.$$

(a) 证明 $(n+m)! \geq (n+1)^m n!$ 对于 $m \geq 1$.（如果你将这表示为 $m = 2$，归纳证明应该很清楚。）

(b) 证明 $(*)$ 中的误差不超过 $1/(n \cdot n!)$！
建议：使用部分(a)和几何级数。

（c）使用部分（b）来证明 $2.716\bar{6} < e < 2.7183\bar{3}$。你可以只使用计算器上的算术运算。

(d) 需要多少项才能给出20位数的精度？给出尽可能小的答案。（当然，10^{20}项就足够了！）

$\diamond$

证明 e 是无理数。提示：假设 $e = p/q$ 是有理数，以最简形式表示。根据12.23题，

$$0 < \frac{p}{q} - \sum_{k=0}^{n} \frac{1}{k!} < \frac{1}{n(n!)} \quad \text{for each } n \in \mathbf{N}.$$

取 $n = q$ 并推断存在一个小于 $1/q$ 的正整数。（记住：如果 $k \leq q$，那么 $k!$ 能整除 $q!$。）$\diamond$

第13章

三角函数

三角函数通常通过几何学引入，要么是直角三角形边长的比例，要么是单位圆上的点的术语。然而，这里采取的方法可能看起来很晦涩，甚至有些人为。然而，我们的目标是使用$\mathbf{R}$公理来定义一切，因此我们将给出三角函数的解析定义。为了与几何学建立联系，我们必须证明我们的定义与熟悉的几何定义一致。这些论证必然是几何的，但鉴于它们的目的是教学（而不是逻辑）的，对几何的依赖不会损害章节的逻辑结构。

13.1 正弦和余弦

指数函数exp由一阶微分方程表征，即它是唯一的可微函数 $f: \mathbf{R} \to \mathbf{R}$ ，使得

$$f' = f \quad \text{and} \quad f(0) = 1. \tag{13.1}$$

唯一性断言——方程（13.1）至多有一个解——仅使用了平均值定理，而存在性部分——（13.1）至少有一个解——需要一些额外的工具，要么是积分，要么是幂级数。我们对基本圆三角函数的方法类似。以下定义隐含地依赖于定理，即给定的标准确实唯一地定义了函数；这些定理将立即正式陈述并

已证明。唯一性将是一个简单的论证，使用平均值定理，而存在性将取决于幂级数。

定义13.1 正弦函数sin：$\mathbf{R} \to \mathbf{R}$是初值问题的解

$$f'' + f = 0, \qquad f(0) = 0, \quad f'(0) = 1. \tag{13.2}$$

余弦函数 cos : {v*} 是初值问题的解

$$f'' + f = 0, \qquad f(0) = 1, \quad f'(0) = 0. \tag{13.3}$$

切线、余切、正割和余割函数定义为（及其自然域）以通常方式为正弦和余弦的比:

$$\tan = \frac{\sin}{\cos}, \quad \cot = \frac{\cos}{\sin}, \quad \sec = \frac{1}{\cos}, \quad \csc = \frac{1}{\sin}.$$

唯一性

我们首先表明，最多只有一个二阶可微函数满足（13.2）和（13.3）。一个关键观察是，微分方程 $y'' + y = 0$ 是线性的，即如果 f 和 g 是解，c 是一个常数，那么 $(cf + g)$ 也是解。大多数微分方程都没有这个性质。

命题13.2. 设$y : \mathbf{R} \to \mathbf{R}$为一个二阶可微函数，满足$y''+y = 0$在$\mathbf{R}$上。如果$y(0) = y'(0) = 0$，那么对于所有$x$，有$y(x) = 0$。

证明。如果 $y'' + y =$ 在 $\mathbf{R}$ 上为 0，那么我们推断出

$$\begin{aligned}
\big((y')^2 + y^2\big)' &= 2y'y'' + 2yy' && \text{the chain rule} \\
&= 2y' \cdot (y'' + y) && \text{factoring} \\
&= 0 && \text{by hypothesis.}
\end{aligned}$$

这意味着函数$(y')^2 + y^2$在$\mathbf{R}$上是常数。在0处求值并使用$y'(0) = y(0) = 0$，我们发现$(y')^2 + y^2 = 0$在$\mathbf{R}$上，这反过来又意味着y恒为零。 □

推论13.3。设$f : \mathbf{R} \to \mathbf{R}$为一个满足$f'' + f = 0$, $f(0) = a$和$f'(0) = b$的函数。那么对于所有$x \in \mathbf{R}$，有$f(x) = a \cos x + b \sin x$。

证明。设$y = f - (a\cos + b\sin)$。微分方程$f'' + f$ =0的线性意味着y也是解。sin和cos的初始条件意味着$y(0) = y'(0)$ =0。根据命题13.2，y是零函数。

□

难以高估这个推论的重要性。三角函数的基本性质都是直接的结果，通过构造函数并证明它们满足定义的微分方程以及适当的初始条件。

存在

目前我们没有逻辑基础相信微分方程 $y'' + y = 0$ 有任何非平凡解。然而，我们确实拥有一个强大的工具来尝试猜测解的形式，即幂级数。让我们假设

$$y(x) = \sum_{k=0}^{\infty} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots$$

这是一个(13.2)的实分析解。使用微分方程，我们推导出系数。逐项微分（并移动求和的索引）给出

$$\begin{aligned} y'(x) &= \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k, \\ y''(x) &= \sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} x^k. \end{aligned} \tag{13.4}$$

初始条件确定了前两个系数：$y(0) = a_0 = 0$，和 $y'(0) = a_1 = 1$。因为 $y'' = -y$ 根据假设，方程（13.4）意味着

$$a_{k+2} = -\frac{a_k}{(k+2)(k+1)} \qquad \text{for} \quad k \geq 0. \tag{13.5}$$

我们立即发现 $0 = a_0 = a_2 = a_4 = \cdots$，而 $a_3 = -1/(3 \cdot 2)$，$a_5 = 1/(5 \cdot 4 \cdot 3 \cdot 2)$，以此类推。稍加思考，我们猜测

$$a_{2k} = 0, \qquad a_{2k+1} = \frac{(-1)^k}{(2k+1)!} \qquad \text{for} \quad k \geq 0,$$

这是通过在 k 上的归纳法容易证明的。因此

$$y(x)=\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}x^{2k+1}=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots \tag{13.6}$$

这是一个（13.2）的候选解。你应该迅速正式地检查这个级数的二阶导数是原函数的负值。这个论证仅仅证明了如果（13.2）有一个实分析解，那么这个解由（13.6）给出。现在我们验证（13.6）确实是（13.2）的一个实分析解。当连续非零项的比率的绝对值小于1时，（13.6）中的级数收敛。然而，对于所有 $x\in\mathbf{R}$,

$$\lim_{k\to\infty}\left|\frac{a_{2k+3}\,x^{2k+3}}{a_{2k+1}\,x^{2k+1}}\right|=\lim_{k\to\infty}\left|\frac{(2k+1)!\,x^2}{(2k+3)!}\right|=\lim_{k\to\infty}\left|\frac{x^2}{(2k+3)(2k+2)}\right|=0,$$

所以幂级数（13.6）对于所有实数 x 绝对收敛，因此定义了一个函数 $s:\mathbf{R}\to\mathbf{R}$。逐项微分表明 $s''+s=$ 在 $\mathbf{R}$ 上为0；系数的选择是出于希望这个方程成立，毕竟！初始值 $s(0)=0$ 和 $s'(0)=1$ 也被纳入了系数的选择；因此我们已证明（13.2）有一个解，实际上有一个实解析解。

方程（13.3）可以通过并行论证处理，参见练习13.1。从此以后，我们将使用（13.2）和（13.3）具有实分析解的事实，以及相应的幂级数在$\mathbf{R}$上的收敛性。

摘要

通过巧妙猜测和适当使用强大工具的结合，我们已证明存在实分析函数 sin 和 cos：$\mathbf{R}\to\mathbf{R}$，它们满足

$$\begin{aligned}\sin''&=-\sin, &\quad \sin 0&=0, &\quad \sin' 0&=1\\ \cos''&=-\cos, &\quad \cos 0&=1, &\quad \cos' 0&=0\end{aligned}$$

这些函数在 $\mathbf{R}$ 上由幂级数定义

$$\sin x=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k+1}}{(2k+1)!},\qquad \cos x=\sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!}. \tag{13.7}$$

此外，每个满足微分方程 $y''+y=0$ 的两次可微函数 $y:\mathbf{R}\to\mathbf{R}$ 是正弦和余弦的线性组合，因此是实解析的。最后，正弦和余弦由它们满足的初值问题来表征。为了证明某个函数 f 是正弦函数，只需证明 $f''+f=0$，以及 $f(0)=0$，$f'(0)=1$。

几个正弦和余弦的有用性质可以从这种特征推导出来。这些性质被收集在下面的定理13.4中。这种特征还将被用来将几何定义与我们的解析定义联系起来；我们将用扇形面积来定义函数，并证明（几何上）这些函数满足表征正弦和余弦函数的初值问题。

定理13.4.正弦函数是奇函数；余弦函数是偶函数。正弦和余弦的导数由以下公式给出

$$\sin'=\cos,\qquad \cos'=-\sin. \tag{13.8}$$

对于所有 $x\in\mathbf{R}$，$\sin^2 x+\cos^2 x=1$。最后，正弦和余弦满足以下加法公式:

(13.9)
$$\begin{array}{rcl}\sin(a+b)&=&\sin a\cos b+\sin b\cos a\\ \cos(a+b)&=&\cos a\cos b-\sin a\sin b\end{array}\qquad \textit{for all } a,b\in\mathbf{R}.$$

特别地，$\sin(2x)=2\sin x\cos x$ 和 $\cos(2x)=\cos^2 x-\sin^2 x$ 对于所有 $x\in\mathbf{R}$。

证明。显然，正弦函数是奇函数，从其幂级数表示中可以看出；然而，可以使用引理13.3给出一个直接的证明（符合定理的精神）。确实，由$g(x)=\sin(-x)$定义的函数$g:\mathbf{R}\to\mathbf{R}$满足微分方程$g''+g=0$和初始条件$g(0)=0$，$g'(0)=-1$，因此引理意味着$g=-\sin$。余弦函数的偶性可以通过类似的方式看出。

如果 $y=\sin'$，则 $y''+y=0$；这立即从方程 $\sin''+\sin=0$ 的微分中得出。但是，根据正弦的定义，$y(0)=1$，并且 $y'(0)=\sin''0=-\sin 0=0$。根据推论13.3，$\sin'=\cos$。类似的论证表明 $\cos'=-\sin$。

考虑函数 $f=\sin^2+\cos^2$。上一段的结果表明

$$f'=2\sin\sin'+2\cos\cos'=2\sin\cos+2\cos(-\sin)=0,$$

这意味着 f 是常数。由于 $f(0) = \sin^2 0 + \cos^2 0 = 1$，$f$ 在任何地方都等于 1。

要证明加法公式，固定 $b \in \mathbf{R}$ 并考虑由 $y(x) = \sin(x + b)$ 定义的函数 y。链式法则意味着 $y'' + y = 0$ 在 $\mathbf{R}$ 上，并且正弦的导数公式意味着对所有 $x \in \mathbf{R}$ 有 $y'(x) = \cos(x + b)$。将 $x = 0$ 代入得到 $y(0) = \sin b$ 和 $y'(0) = \cos b$，所以

$$\sin(a + b) = y(a) = \sin a \cos b + \sin b \cos a \qquad \text{for all } a \in \mathbf{R}$$

由推论13.3。余弦函数的加法公式被类似地证明。 □

某些标准极限是正弦和余弦幂级数表示的简单结果。这些极限通常从几何考虑中推导出来，并用于证明定理13.4中的导数公式。

$$\lim_{x\to 0} \frac{\sin x}{x} = 1; \qquad \lim_{x\to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}.$$

尽管每个极限都可以通过洛必达法则轻松推导出来，但如果计划使用结果来推导公式 $\sin' = \cos$，则这样做在逻辑上是不允许的，因为由此产生的论点将是循环的！在任何情况下，sin 和 cos 的幂级数允许直接计算这些极限。对于 $x \neq 0$,

$$\frac{\sin x}{x} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k + 1)!} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \cdots.$$

通过比值测试，右侧的级数在$\mathbf{R}$上表示一个连续函数，因此可以通过将$x =$设为0来在0处评估；显然这给出1。第二个极限在练习13.3中处理。

周期性

在这个部分中，我们证明正弦和余弦是周期性的。它们的共同周期将被定义为2π；这与通常的几何定义（如“单位圆的面积”或“单位圆周长的一半”）不同，这是一个非几何的基本常数π的定义。当前的定义适用于理论目的和数值评估。自然地，几何定义将作为定理被恢复。

一个物理学家会怀疑正弦和余弦函数在两个基础上表现出振荡行为：首先，方程 $y''+y=0$ 是在合适单位下弹簧上质量的运动方程。（在这种情况下，（y'）$^2+y^2$ 是常数，这正好是能量的守恒。）其次，方程 $y''=-y$ 从定性上说明，当 y 为正时，其图形向下凹，反之亦然。因此，y 的图形始终向水平轴弯曲。由于方程是“时间独立的”，每次解从下方穿过轴到上方时，解都处于与上次穿过此方向时相同的“物理状态”，因此其未来的行为重复其过去的行为。

在某种意义上，周期性的数学证明只是使物理直觉变得精确。第一步是证明余弦函数有一个最小的正零点，我们暂时用 α 表示。α 的存在是通过以下估计实现的，其证明推迟到本节的末尾，以保持概念上的连续性。

命题13.5. 对于所有实数 x，$1-\frac{x^2}{2!} \le \cos x \le 1-\frac{x^2}{2!}+\frac{x^4}{4!}$ 。

授予此结果，我们发现

$$0 \le \cos\sqrt{2}, \qquad \cos\left(\sqrt{6-\sqrt{12}}\right) \le 0,$$

自 $\sqrt{6-\sqrt{12}}$ 是四次上界的第一正根；参见图13.1。介值定理表明，cos 在 $\sqrt{2} \simeq 1.41421$ 和 $\sqrt{6-\sqrt{12}} \simeq 1.59245$ 之间有一个零点。我们现在定义 $\pi=2\alpha$，并观察到命题13.5表明

$$2.82842 \le \pi \le 3.1849.$$

这些粗略的界限与从 e 的定义直接获得的估计 $2 \le e \le 4$ 相似。我们最终将找到收敛相当快的数值级数，这允许获得更精确的界限。

返回主要论点，命题13.5表明cos有最小的正根：$\cos\alpha=0$，且$\cos'\alpha=-\sin\alpha$ 要么是1要么是 -1，因为$\cos^2+\sin^2=1$。因为$\cos 0=1$且 α 是最小的正零点，所以cos在区间$[0,\alpha]$上非负。因此，$\cos'\alpha=-1$，如果它是1，那么余弦在某些位于 α 左侧的区间上将是负的。cos的偶函数性质意味着cos的最大负零点是 $-\alpha$；特别是，存在一个长度为 π 的区间，即 $(-\alpha,\alpha)$，在该区间上cos是正的。

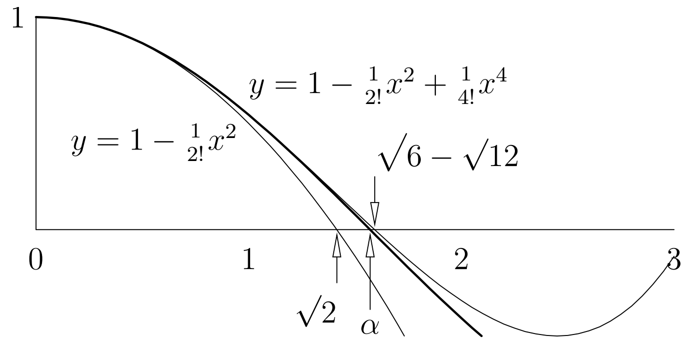


图13.1：余弦函数的最小正零点（粗体）。

通过正弦加法公式,

$$\begin{aligned}\sin(x+\alpha) &= \sin x\cos\alpha + \sin\alpha\cos x\\ &= \sin\alpha\cos x = \cos x \quad \text{for all } x\in\mathbf{R},\end{aligned}$$

自 $\sin\alpha = -\cos'\alpha = 1$. 从几何上看，cos 的图像是 sin 的图像向左平移 α。类似的论证表明，对于所有 $x\in\mathbf{R}$，有 $\sin(x+\pi) = -\sin x$ 。将此方程应用两次可得出

$$\sin(x+2\pi) = -\sin(x+\pi) = \sin x \qquad \text{for all } x\in\mathbf{R}. \tag{13.10}$$

余弦函数也是2π周期的，因为对于所有$x\in\mathbf{R}$，有$\cos x = \sin(x+\alpha)$。最后，没有更小的正周期，因为余弦在$(-\alpha,\alpha)$上是正的，在$(\alpha, 3\alpha)$上是负的。换句话说，2π是正弦和余弦的最小正周期的事实是α是余弦的最小正零点的事实的后果。

剩余的周期性证明是命题13.5。论证不过是重复积分一个基本不等式，但与上面的论证在性质上完全不同。首先，注意方程sin{v*}cos{v*} = 1意味着对于所有{v*}，有{v*}1 {v*}cos {v*} ⩽ 1。固定{v*} = 0，并使用微积分基本定理从0积分到{v*}：

$$-x \le \int_0^x \cos t\,dt = \sin t\Big|_{t=0}^{x} = \sin x \le x.$$

因此 $-t \le \sin t \le t$ 对于 $t \ge 0$。从 0 到 x 的积分给出

$$-\frac{x^2}{2} \le \int_0^x \sin t\,dt = 1 - \cos x \le \frac{x^2}{2},$$

并且，由于 $x \geq 0$ 是任意的，因此可以得出

$$1 - \frac{t^2}{2} \leq \cos t \leq 1 \qquad \text{for } t \geq 0.$$

另一种积分（和变量重命名）给出 $x - (x^3/6) \leq \sin x \leq x$ 对于 $x \geq 0$，第四种给出 $x^2/2 - x^4/24 \leq 1 - \cos x \leq x^2/2$，或者

$$1 - \frac{x^2}{2} \leq \cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24} \quad \text{for } x \geq 0. \tag{13.11}$$

这个最后的一组不等式也适用于 $x < 0$，因为每个项在 x 中都是偶数。这完成了命题13.5的证明。

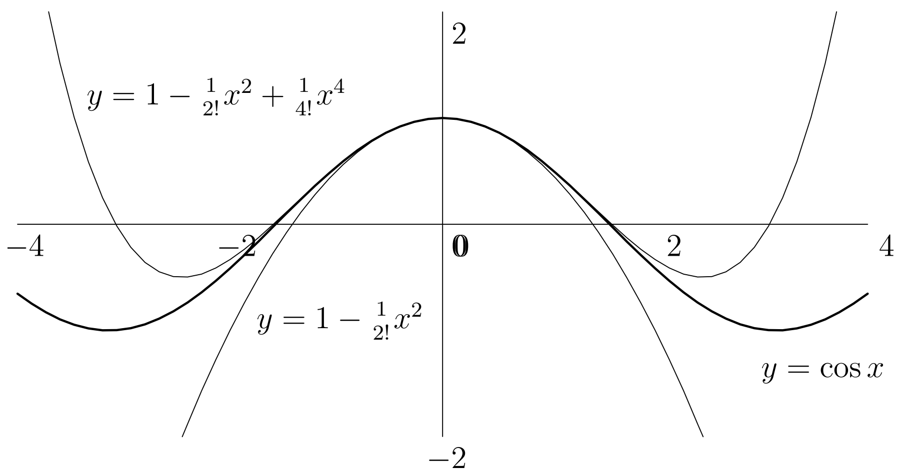


图 e 13.2：余弦函数的上界和下界 动作。

余弦的幂级数具有交替的正负项，方程（13.11）表明奇数部分和（以正项结束）是余弦的上界，而偶数部分和（以负项结束）是下界。视觉证据令人信服，见图13.2。正如通过归纳证明命题13.5的过程所展示的那样，所概述的论断确实是正确的。此外，随着逼近多项式的次数增加，逼近效果越好。然而，重要的是要强调，结论不能仅基于项的符号得出；必须考虑幂级数中的实际系数。第14章系统地研究了多项式逼近。

13.2 辅助三角函数

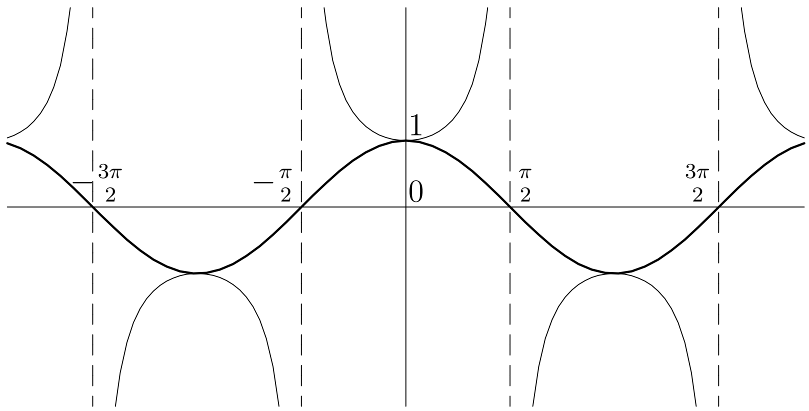


图13.3：余弦（粗体）和正割的图。

正弦函数恰好消失在 π 的整数倍处，而余弦函数消失在“奇半整数”倍 π 处，即 $k+\frac{1}{2}\pi$ 对于 $k\in\mathbf{Z}$。sin 和 cos 都是“反周期”的，周期为 π，即

$$\sin(x+\pi)=-\sin x,\quad \text{and}\quad \cos(x+\pi)=-\cos x\qquad \text{for all } x\in\mathbf{R}.$$

割线函数，sec $=1/$ cos，在cos消失处未定义，是偶函数，周期为2π，具有周期π的反周期性。因为对于所有 $x\in\mathbf{R}$，$|\sec x|\geq 1$ 对于sec的定义域内的所有实数 x。由于正弦和余弦之间具有类似的关系，正割函数，csc $=1/$ sin，满足 $\csc(x+\frac{\pi}{2})=\sec x$。

切线函数，tan $=$ sin / cos，在cos消失的地方未定义，并且具有周期 π (why?)。切线函数是奇函数，作为奇函数除以偶函数的商。其行为完全由其在基本区间 $(-\pi/2,\pi/2)$ 上的行为决定，我们现在转向这个区间。

切线函数在$(-\pi/\hat{}2,\pi/\hat{}2)$上可导，其导数通过商法则求得为

$$\tan'=\frac{\cos\sin'-\sin\cos'}{\cos^2}=\frac{\cos^2+\sin^2}{\cos^2}=\frac{1}{\cos^2}=\sec^2.$$

特别地，$\tan'>0$ 在 $(-\pi/2,\pi/2)$ 上，这意味着 tan 在这个区间上是递增的。正如已经指出的，cos 在 $(-\pi/2,\pi/2)$ 上是正的，而

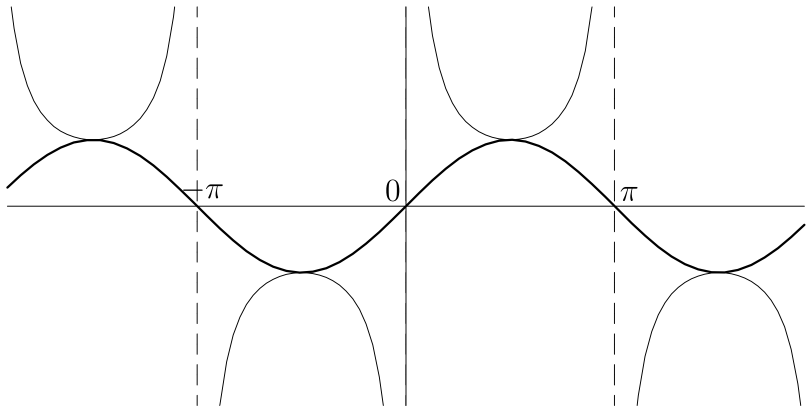


图13.4：正弦（粗体）和余割的图。

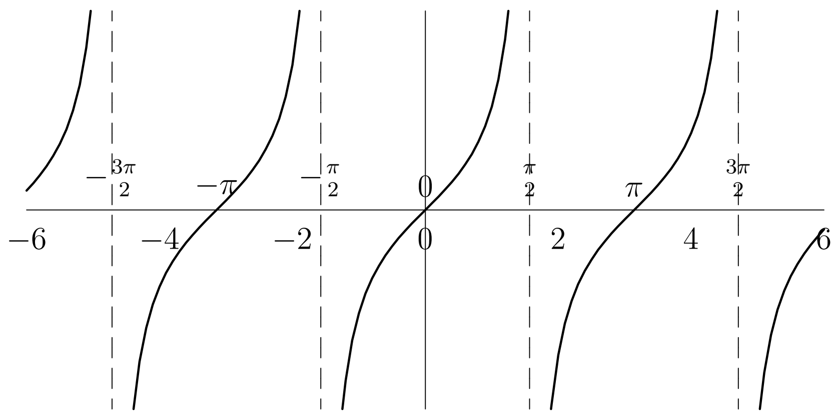


图13.5：正切函数的图像。

sin在$(0, \pi/2)$上为正，因此在$(-\pi/2, 0)$上为负。如下$x \to \pi$，tan $x \to +\infty$，由于tan是奇函数，

$$\lim_{x \to -\pi/2^+} \tan x = -\infty.$$

总结来说，tan 映射 $(-\pi/\hat{}2, \pi/\hat{}2)$ 双射到 $\mathbf{R}$，并且在形式为 $\big((k-\frac{1}{2})\pi, (k+\frac{1}{2})\pi\big)$ 的每个区间上单调递增，具有 $k \in \mathbf{Z}$。

余切函数cot = cos / sin不完全是tan的倒数，因为存在零点和极点（分母为零的地方）；然而，这两个函数在它们都定义的地方是互为倒数，tan的零点正好是cot的极点，反之亦然。

导数的cot是找到的为$-1/\csc^2$，这表明cot在其定义域的每个区间上都是递减的，特别是在形式为$\big(k\pi,(k+1)\pi\big)$的每个区间上，其中$k\in\mathbf{Z}$。

双曲三角函数

六种已提到的三角函数有时被称为圆三角函数，因为它们与圆的几何学有关。确实，恒等式$\cos^2+\sin^2=1$意味着点$(\cos t,\sin t)$对所有实数t都位于单位圆上。存在一个“双重”的函数族，称为双曲三角函数，它们具有类似的名字，并在后面附加了一个“h”，如cosh，sinh（有时发音为“cinch”或“shine”），以及tanh（与“ranch”押韵）。这些函数可能是隐晦地直接用自然指数函数定义的。双曲余弦和正弦函数分别是exp的偶数和奇数部分：

$$\cosh x=\frac{e^x+e^{-x}}{2},\qquad \sinh x=\frac{e^x-e^{-x}}{2}. \tag{13.12}$$

辅助双曲函数由类似方程定义，

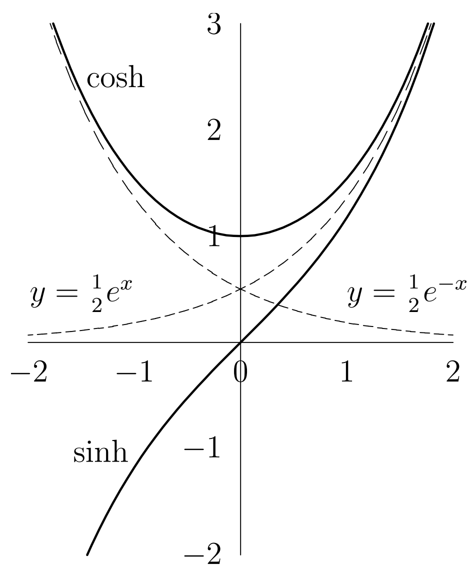


图13.6：双曲余弦和双曲正弦的图形。

例如

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad \operatorname{sech} x = \frac{1}{\cosh x}.$$

剩余的函数，双曲余切和双曲正割，很少遇到，但

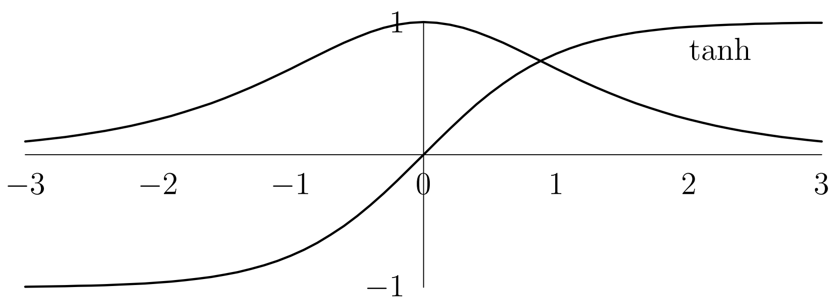


图13.7：tanh和sech的图形。

第一个四项在悬挂链条、肥皂膜、非欧几里得几何和孤立波等多样化的环境中出人意料地频繁出现。圆函数和双曲函数之间存在许多形式上的相似性，其中一些将在下面进行探讨。这些相似性的根本原因既深刻又简单，但如果不定义所有函数在复数集上的定义，就无法看到这一点，参见第15章。

简单计算（作为练习）验证了$\cosh^2 - \sinh^2 = 1$，并且

$$\cosh' = \sinh, \quad \sinh' = \cosh, \quad \tanh' = \operatorname{sech}^2. \tag{13.13}$$

方程 $\cosh^2 - \sinh^2 = 1$ 表示点 $(\cosh t, \sinh t)$ 在单位双曲线上（具有笛卡尔方程 $x^2 - y^2 = 1$），对于所有实数 t。导数表达式与圆三角函数公式类似，但不含符号，并且可以追溯到 sinh 和 cosh 被定义为微分方程的解这一事实：

$$\begin{aligned} \sinh: &\quad y'' - y = 0, &\quad y(0) = 0, &\quad y'(0) = 1, \\ \cosh: &\quad y'' - y = 0, &\quad y(0) = 1, &\quad y'(0) = 0. \end{aligned}$$

存在类似于（13.9）的sinh和cosh的加法公式，你可以通过一点毅力直接检查。为了完成

类比，sinh 和 cosh 的幂级数表示可以从 exp 的幂级数中找到；结果是
，

$$\sinh x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \qquad \cosh x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!},$$

清楚地显示了圆和双曲三角函数之间的相似性。

13.3 反三角函数

每个圆周三角函数都是周期的，因此没有逆函数。在双曲三角函数中，双曲正弦和双曲正切是单射的，因此有“全局”逆函数。双曲余弦和双曲余割是偶函数，因此不是一一对应的，但每个函数在限制到正实轴时都是单射的。在本节中，我们将研究各种三角函数的逆函数的分支。也许最引人注目的是，虽然逆函数不是代数函数，但它们的导数是代数的。这不是偶然，而是由描述初等三角函数的微分方程直接得出的结果。

反正弦和反余弦函数

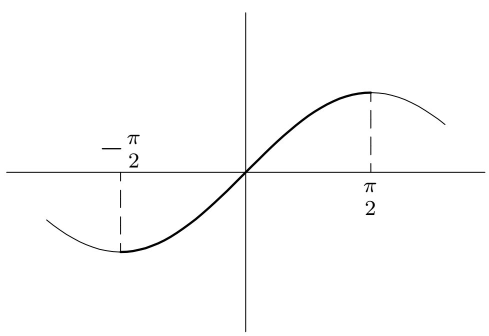


图13.8：正弦。

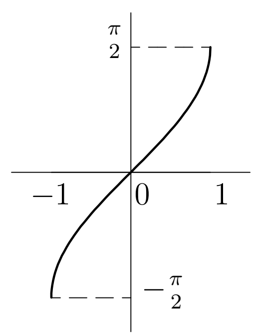


图13.9：arcsin

余弦在区间$(-\pi/2, \pi/2)$上为正；因此正弦函数在此区间上单调递增，因为$\sin' = \cos$。因为对于所有实数x，$\sin(\pi - x) = \sin x$，不存在更大的开区间使得sin是单射的。sin在闭区间$[-\pi/2, \pi/2]$上的限制表示为Sin。

The inverse function $\mathrm{Sin}^{-1} : [-1, 1] \to [-\pi/2, \pi/2]$, sometimes denoted a rcsin, is called the principal branch of arcsine. Thus

$$\begin{aligned} \sin(\mathrm{Sin}^{-1}x) &= x && \text{for all } x \in [-1,1], \\ \mathrm{Sin}^{-1}(\sin x) &= x && \text{for all } x \in [-\pi/2, \pi/2]. \end{aligned} \tag{13.14}$$

正弦函数在$[\pi/2, 3\pi/2]$上递减，因为$\sin(\pi - x) = \sin x$。周期性意味着对于每个整数 k，sin 在 $\big[(k-\frac{1}{2})\pi, (k+\frac{1}{2})\pi\big]$ 上是一一对应的。对于每个 k，都有一个相应的反正弦分支，即 sin 的限制的逆，即 $\big[(k-\frac{1}{2})\pi, (k+\frac{1}{2})\pi\big]$。在考虑反正弦的非主分支的罕见情况下，它表示为 $\sin^{-1}$，k 由上下文提供。

大致相同的评论适用于余弦函数。主分支是Cos^{-1}：$[-1, 1] \to [0, \pi]$，对于每个整数k，存在一个arccos分支，其值在$[k\pi, (k+1)\pi]$内。恒等式$\sin(x + \frac{\pi}{2}) = \cos x$ 变为

$$\mathrm{Cos}^{-1}x = \frac{\pi}{2} + \mathrm{Sin}^{-1}x \qquad \text{for all } x \in [-1,1]. \tag{13.15}$$

我们现在希望找到 Sin^{-1} 的导数。首先，$\sin' = \cos$ 在 $(-\pi/2, \pi, 2)$ 上非零，因此 Sin^{-1} 在 $(-1, 1)$ 上可微。这意味着我们可以对 (13.14) 中的第一个方程进行微分:

$$\cos(\mathrm{Sin}^{-1}x) \cdot (\mathrm{Sin}^{-1})'(x) = 1 \qquad \text{for all } x \in (-1,1).$$

函数 Sin^{-1} 在 $(-\pi/2, \pi, 2)$ 中取值，且在此区间上 cos 是正的。因此，在此区间上 $\cos = \sqrt{1 - \sin^2}$，所以前一个方程可以重写为

$$\begin{aligned} (\mathrm{Sin}^{-1})'(x) &= \frac{1}{\cos(\mathrm{Sin}^{-1}x)} = \frac{1}{\sqrt{1 - \sin^2(\mathrm{Sin}^{-1}x)}} \\ &= \frac{1}{\sqrt{1-x^2}} \qquad \text{for all } x \in (-1,1). \end{aligned} \tag{13.16}$$

因为 Sin^{-1} 和 Cos^{-1} 差一个加性常数，它们的导数相等:

$$(\mathrm{Cos}^{-1})'(x) = \frac{1}{\sqrt{1-x^2}} \qquad \text{for all } x \in (-1,1). \tag{13.17}$$

这些方程在我们将正弦和余弦的几何定义与我们的解析定义等同起来时将至关重要。

其他圆周三角函数

切线函数将 $(-\pi/\hat{}2, \pi/\hat{}2)$ 双射地映射到 $\mathbf{R}$。tan 在此区间的限制的逆是 arctan 的主分支，表示为 $\mathrm{Tan}^{-1}: \mathbf{R} \to (-\pi/\hat{}2, \pi/\hat{}2)$。因为 tan 是 π-周期的，arctan 的其他分支与主分支相差一个 π 的倍数。通过微分第一个得到 Tan^{-1} 的导数。

$$\begin{aligned} \tan(\mathrm{Tan}^{-1} x) &= x \qquad \text{for all } x \in \mathbf{R}, \\ \mathrm{Tan}^{-1}(\tan x) &= x \qquad \text{for all } x \in [-\pi/2, \pi/2]. \end{aligned} \tag{13.18}$$

计算简略，结果留作练习；结果是

$$(\mathrm{Tan}^{-1})'(x) = \frac{1}{1+x^2} \qquad \text{for all } x \in \mathbf{R}. \tag{13.19}$$

这甚至比方程（13.16）更引人注目；Tan^{-1}的导数是一个有理函数，而不仅仅是一个代数函数！为了强调一个哲学观点，有理函数的导数始终是有理函数，但不定积分不一定是。我们已经为倒数函数看到了这一点，但这个观点值得重复。

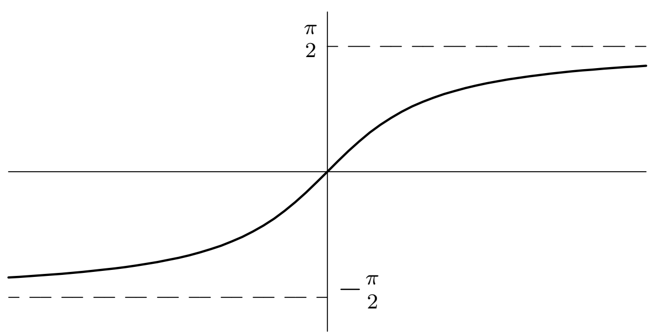


图13.10：反正切的主支

其他圆周三角函数的倒数在应用中不太突出，尽管反正割在计算某些积分时会出现。为了详细描述定义域和值域，考虑将余弦函数限制在$[0, \pi]$。其倒数，即正割的限制，定义在并集$[0, \pi/2) \cup (\pi/2, \pi]$上。反正割的主分支是这个限制的逆；其定义域是$(-\infty, -1] \cup [1, \infty)$。

双曲三角函数

反双曲三角函数可以直接从它们的定义中计算。为了解方程 $y=$ cosh $x=(e^x+e^{-x})/$^2 对于 x，将两边乘以 $2e^x$ 并重新排列，得到

$$(e^x)^2-(2y)e^x+1=0.$$

这是一个关于 e^x 的二次方程，可以使用二次公式求解:

$$x=\log\bigl(y\pm\sqrt{y^2-1}\bigr),\qquad y\geq 1. \tag{13.20}$$

我们期望有两个实分支，因为双曲余弦函数不是一一对应的。作为一致性检查，观察 $y-\sqrt{y^2-1}=1/(y+\sqrt{y^2-1})$ 对于 $|y|\geq 1$，并且这两个量在 $y\geq 1$ 时都是正的，所以

$$\log\bigl(y\pm\sqrt{y^2-1}\bigr)=\mp\log\bigl(y-\sqrt{y^2-1}\bigr),$$

并且这两个分支确实相差一个符号。类似的计算表明，$\sinh^{-1}$由以下公式定义

$$\sinh^{-1}x=\log\bigl(x+\sqrt{x^2+1}\bigr). \tag{13.21}$$

没有符号歧义，因为只有这个选择在 x 为实数时才导致一个实值函数。你也可以验证右边的表达式是 x 的奇函数。

tanh的逆函数更容易找到。简单的代数运算表明

$$y=\frac{e^x-e^{-x}}{e^x+e^{-x}}=\frac{e^{2x}-1}{e^{2x}-1}$$

仅当且仅当

$$\tanh^{-1}y=x=\frac{1}{2}\log\left(\frac{1-y}{1+y}\right),\qquad -1<y<1. \tag{13.22}$$

作为一致性检查，右侧的表达式是关于 y 的奇函数。

这些逆函数的导数是看起来与它们的圆函数非常相似的代数函数。它被留作一个

练习以证明

$$
\begin{aligned}
(\cosh^{-1})'(x) &= \frac{1}{\sqrt{x^2-1}}, \\
(\sinh^{-1})'(x) &= \frac{1}{\sqrt{x^2+1}}, \\
(\tanh^{-1})'(x) &= \frac{1}{x^2-1}.
\end{aligned}
\tag{13.23}
$$

13.4 几何定义

本节严格来说并不贡献于圆三角函数的逻辑发展。其意图更多的是将已引入的正弦函数与单位圆上弧长和圆扇形所围成的面积等熟悉的图形联系起来。该表述相对非正式，并自由地使用图片和几何直觉。为了强调所展示的内容非平凡，在证明它们与上面解析定义的函数相同之前，将使用大写字母表示三角函数的几何版本（例如，COS）。

三角学一词源于希腊语根，意为“三角形测量”。它是欧几里得几何的一个特征，即相似三角形存在；存在非全等的三角形，它们具有相同的内角.[1] 直角三角形的形状，在相似性上，由其边长的比例决定。它也在相似性上，由其一个锐角决定。圆三角函数是锐角作为边长比例的函数，见下方的方程（13.24）。许多三角学的学生学会了一种某种形式的助记法[2] 来记住哪个函数是哪个比例。这个定义只对锐角有意义；为了定义任意实数的三角函数，通过对称性和周期性进行扩展。为了激发这些扩展，引入一个以锐角 θ 为原点，斜边缩放的笛卡尔坐标系

[1] Strange as it may seem, not all geometries have this property. Think of measuring patches of the surface of a sphere; the sides of triangles are arcs of great circles. If the three internal angles of a triangle are known, then the side lengths may be deduced; consequently, two triangles with the same internal angles are actually congruent.

[2] Like sohcahtoa.

为了具有单位长度，如在图13.11的前半部分所示。此时，三角形本身降为次要角色，θ允许为任意值，甚至可以是负数（对应于“顺时针”角度）。正弦比、余弦比和正切比被定义为

$$x = \mathrm{COS}\,\theta, \qquad y = \mathrm{SIN}\,\theta, \qquad \frac{y}{x} = \mathrm{TAN}\,\theta. \tag{13.24}$$

如果一个2Π的角度对应一个完整旋转，那么正弦和余弦是2Π周期的。角度θ仅由点(x, y)确定，直到加上2Π的倍数。

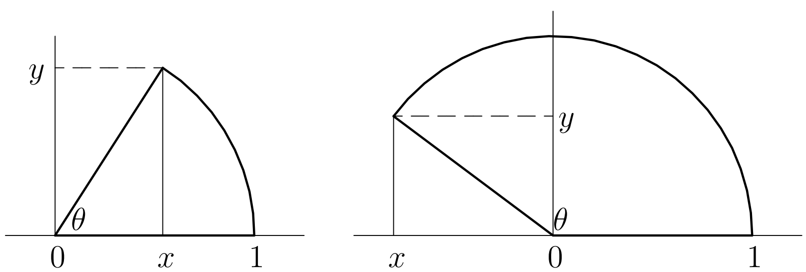


图13.11： 圆心角由射线通过所截取的圆弧 原始

它可能直观上很明显存在一个称为“角度”的数值量，但可能并不明显如何“自然地”测量它。由于历史原因，起源于天文学，巴比伦人将圆分成360个部分，称为度。即使是现代英语也使用基于这个系统的习语。[3]度作为角度的度量在数学上并没有什么自然性，就像十进制记数法不是自然地表示整数的手段一样。相反，自然界更喜欢[4]以几何方式测量角度，要么是圆的弧长，要么是圆扇形的面积。

圆单位弧长与其定义的扇形面积密切相关。在笛卡尔坐标系(u, v)中，圆的方程为 $u^2 + v^2 = 1$。通过原点的每条射线都与圆相交于一个唯一点(x, y)。（**R**的完备性公理在此是隐含的。）设θ为从(0, 0)到(x, y)的弧长，沿圆逆时针[5]测量，设2Π为圆周长。存在

[3]“No it doesn’t,” said the author, making a 180° reversal of his claim. “Wait, I was wrong. It does,” he added, coming around a full 360.

[4]This claim will be fully justified shortly.

[5]This is also a convention, but a harmless one.

一个与圆相对应的扇形区域，由正u-轴、弧和射线所围成，如图13.11所示。正如阿基米德所知，该扇形的面积是$\theta/$的平方。他用以下论据证明了这一点，有关整个圆盘的情况，请参阅图13.12。

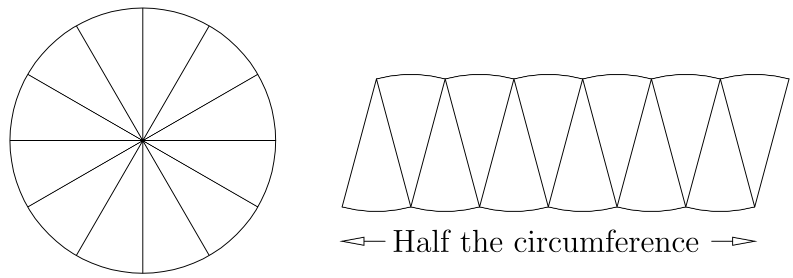


图13.12：阿基米德将圆盘分割成近似矩形。

设 A 为原点处角度为 θ 的扇形面积。将扇形分成 N 个全等的部分，每个部分近似为一个底边为 θ/N 高为 1 的等腰三角形。每个切片的面积约为 $\theta/^{2}N$，因此总面积 A 约为 $\theta/$^2。通过取 N 很大，可以使得近似任意精确，所以 $A = \theta/$^2.[6]。特别是，单位圆的面积是 Π。

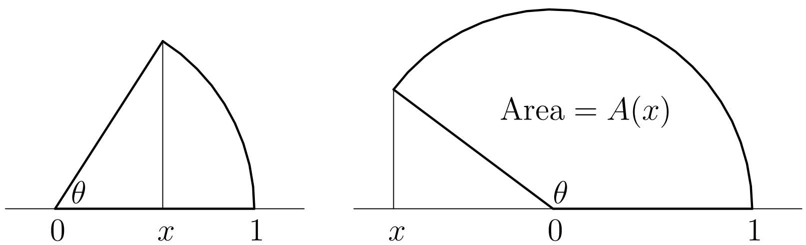
直观上，该区域可以被切割成无限多个无限薄的三角形，这些三角形被重新排列成一个高度为1、宽度为$\theta/$^2的矩形，但这种直觉在字面上并不正确。积分的语言适合使这个断言变得严谨，但你应该再次提醒自己书中开头的歌德名言。

命题13.6.下图中圆扇形的面积是

$$A(x) = \frac{x}{2}\sqrt{1-x^2} + \int_x^1 \sqrt{1-u^2}\,du$$

对于 $-1 \leq x \leq 1$。

[6]If two numbers are ε-close for all $\varepsilon > 0$, then they are equal.



证明。由于(x, y)位于单位圆的上半部分，$y = \sqrt{1-x^2}$。如果$x > 0$（左侧图片），那么$x\sqrt{1-x^2}/2$是直角三角形的面积，而积分是曲线区域的面积。另一方面，如果$x < 0$（右侧图片），那么$x\sqrt{1-x^2}/2$是负数，但绝对值等于直角三角形的面积，而积分是整个封闭区域的面积。再次，总和——即面积之差——是圆扇形的面积。

□

特定情况 $x = -1$ 很有趣:

Corollary 13.7. $\dfrac{\Pi}{2} = \displaystyle\int_{-1}^{1} \sqrt{1-u^2}\, du.$

然而，通过微分可以得到一个更为重要的结论。根据基本定理,

$$A'(x) = \frac{\sqrt{1-x^2}}{2} + x \cdot \frac{-x}{2\sqrt{1-x^2}} - \sqrt{1-x^2} = -\frac{1}{2\sqrt{1-x^2}}$$

对于 $-1 < x < 1$：函数 $2A$ 和 Cos^{-1} 具有相同的导数。此外，它们在 1 处相等，因此它们是同一个函数。这是圆三角函数与几何定义的函数 A 之间的第一个联系。

注意，$A(-1) = \mathrm{Cos}^{-1}(-1) = \pi$；推论13.7意味着$\pi = \Pi$；正弦和余弦的周期是单位圆的周长。为了解决剩余的悬而未决的问题，阿基米德关于扇形面积的定理（在图13.6的符号中）表示$x = \cos\theta$对于$\theta \in [0, \pi]$。这个等式在$\theta \in [-\pi, 0]$时成立，因为两个事实：（i）余弦是偶函数，并且（ii）命题13.6中的图形在水平轴的反射下是对称的，这交换了θ和$-\theta$。结合这些观察，$\cos = \mathrm{COS}$在区间$[-\pi, \pi]$上，由于这两个函数都是2π-周期的，它们在所有地方都相等。这意味着圆三角函数（定义为微分方程的解）与定义为水平和

圆上一点的垂直坐标。“变量”不是以度为单位测量的，而是以弧度为单位——圆的弧长单位。

之前有人断言弧度是“自然”的角度度量。主要理由是正弦$'$ $=$余弦和余弦$' = -$正弦。假设使用度来定义圆三角函数正弦$^\circ$和余弦$^\circ$。（函数余弦$^\circ$“就像余弦一样，但以度为单位输入”。）像余弦$^\circ$ $90 = 0$和正弦$^\circ$ $90 = 1$这样的等式将成立（这不会成问题），但等式（正弦$^\circ$）$'$ $=$余弦$^\circ$和（余弦$^\circ$）$' = -$正弦$^\circ$将是错误的（这将非常不方便）。为了看到将取代它们的等式，观察到余弦和余弦$^\circ$通过缩放域不同。精确地，

$$\mathrm{COS}\,(\pi \cdot \Theta/180) = \mathrm{COS}^\circ\Theta \qquad \text{for all } \Theta \in \mathbf{R},$$

自Θ度与$\pi \cdot \Theta/180$弧度相同的角度。从该方程中，很容易验证

$$(\mathrm{COS}^\circ)' = -\frac{\pi}{180}\mathrm{SIN}^\circ.$$

这个方程在美学上至多是不愉快的，因为它将一个任意数字（即180）构建到一个基本的三角函数关系中。

练习

练习13.1 详细模仿sin的构造来构造cos。唯一的不同在于初始条件。$\diamond$

练习13.2 使用逐项微分给出一个证明 sin{v^2/2} cos{v^2/3} 和 cos{v^2/2} sin{v^2/3} 的替代证明。

练习13.3 证明当使用洛必达法则和幂级数展开时，极限 $\lim\limits_{x\to 0} \dfrac{1 - \cos x}{x^2} =$ 12 成立。$\diamond$

练习13.4 求极限 $\lim\limits_{x\to 0} \displaystyle\int_0^{x^2} \frac{1 - cost}{x^6}\,dt$ $\diamond$

证明 $\sec' = \sec \cdot \tan$（别忘了显示定义域相同）。$\diamond$

练习13.6 使用恒等式$\sec^2$ $=$ = 1/cos+ = $\tan^2$来证明（13.19）。$\diamond$

练习13.7 建立以下恒等式:

(a) $\tan(x+y) = \dfrac{\tan x + \tan y}{1 - \tan x \tan y}$

(b) $\cot 2x = \frac{1}{2}(\cot x - \tan x)$

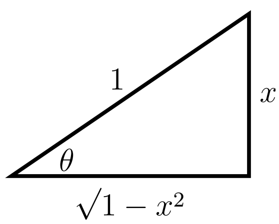
对于每个，确定满足恒等式的 x 和 y 的集合。$\diamond$

练习13.8 使用本章的结果评估以下内容:

(a) $\sin\frac{\pi}{4}$, $\cos\frac{\pi}{4}$, $\sec\frac{\pi}{4}$, $\tan\frac{\pi}{4}$.
(b) $\sin\frac{\pi}{6}$, $\cos\frac{\pi}{6}$, $\sec\frac{\pi}{6}$, $\tan\frac{\pi}{6}$.
(c) $\sin\frac{\pi}{3}$, $\cos\frac{\pi}{3}$, $\sec\frac{\pi}{3}$, $\tan\frac{\pi}{3}$.
(d) $\sin\frac{\pi}{8}$, $\cos\frac{\pi}{8}$

您的答案应仅涉及平方根和有理数。$\diamond$

练习13.9 设 $0 \leqslant < x < \leqslant 1$，并设 $\theta = = \text{Sin}^{-1}x$:



立即可以得出 $\cos\theta = \sqrt{1-x^2}$,

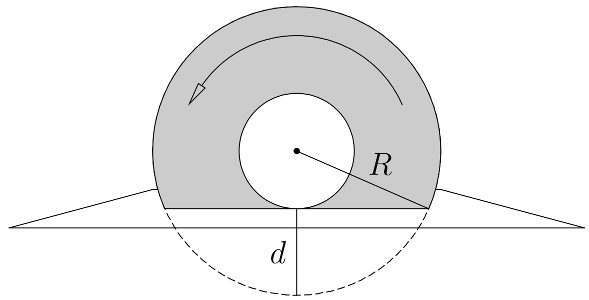
$$\sec\theta = \frac{1}{\sqrt{1-x^2}}, \quad \tan\theta = \frac{x}{\sqrt{1-x^2}}, \quad \cot\theta = \frac{\sqrt{1-x^2}}{x}.$$

同样，求 $\cos \text{Tan}^{-1}x$，$\sec \text{Tan}^{-1}x$ 和 $\sin \text{Tan}^{-1}x$。 $\diamond$

Exercise 13.10 Verify equation (13.13). $\diamond$

Exercise 13.11 Verify equation (13.23). $\diamond$

练习13.12 作为工业过程的一部分，一个薄圆形金属板在部分浸入聚合物槽中时绕其轴线旋转。沿轴线的水平视图如下:



如果轮子的半径为 R，求使暴露于空气中的聚合物量最大化的深度 d（阴影部分）。$\diamond$

以下多部分练习展示了伊万·尼文证明 π^2 是无理数的证明。由此可知，π 本身也是无理数。

练习13.13 定义 $f_n : [0, 1] \to \mathbf{R}$ 为

$$f_n(x) = \frac{x^n(1-x)^n}{n!}.$$

证明以下每个断言。

(a) 如果 $0 < x < 1$，那么 $0 < f_n(x) < \frac{1}{n!}$。

(b) 导数 $f_n^{(k)}(0)$ 和 $f_n^{(k)}(1)$ 对于所有 $k \in \mathbf{N}$ 都是整数。

假设 $\pi^2 = p/q$ 为最简形式，并令

$$F_n(x) = q^n \sum_{k=0}^{n} (-1)^k f_n^{(2k)}(x) \pi^{2n-2k}.$$

(c) $F_n(0)$ 和 $F_n(1)$ 是整数。

(d) $\pi^2 p^n f_n(x) \sin \pi x = \frac{d}{dx}\left(F_n'(x) \sin \pi x - \pi F_n(x) \cos \pi x\right)$.

(e) $F_n(1) + F_n(0) = \pi p^n \int_0^1 f_n(x) \sin \pi x \, dx$.

(f) 取 $n \gg 1$ 以推导出 $0 < F_n(0) + F_n(1) < 1$.

简而言之，如果 π^2 是有理数，那么在 0 和 1 之间存在一个整数。$\diamond$

第14章

泰勒近似

装备了一系列强大的数学工具（平均值定理、微积分基本定理和幂级数）以及丰富的示例库（代数、指数、对数和三角函数），我们转向对计算问题的系统研究。

14.1 数值逼近

您可能从小就了解到数学是“一门精确的科学”，并且问题只有一个正确答案。讽刺的是，这种印象被电子计算器所强化，只需按下一个按钮，计算器就能将许多常见函数计算到8位小数（或更多）。正如第二章中提到的，计算器返回的任何数值答案都不可能是无理数，因此大多数计算器结果只是近似值。有时，人们对此事实只是口头上表示尊重，例如写成$e = 2.71828\dots$或$e \simeq 2.71828$，而对省略号或波浪形等号的意义解释甚少。一些问题应该立即浮现在脑海中:

- 如何定义像 $\sqrt{2}$、e 和 π 这样的常数，如果不是无限小数呢?
- 当计算器返回一个可能的无理数的数值时，这意味着什么?
- 计算器是如何知道返回的值的? （或者，“编写计算器的人是如何知道的? ”进一步追溯一步。）

您已经知道第一个问题的答案（要么如此，要么是时候重新阅读第5、12和13章了！）。第二个问题的答案是第2章的A符号，我们简要回顾一下。第三个问题占据了本章的剩余部分。

当计算器显示 $e = 2.71828$ 时，它实际上意味着“e 的值四舍五入到小数点后五位是 2.71828”，这又意味着（按照惯例）$2.718275 \le e < 2.718285$，或者说（本质上）

$$e = 2.71828 + A(0.5 \times 10^{-5}). \tag{14.1}$$

一个四舍五入的答案表示知识的局限性；它断言所讨论的数字位于实数的一个特定区间内。每个额外的十进制位对应一个长度为十分之一的区间，因此更多的十进制位意味着更多的信息：

$$e = 2.718\,281\,828\,459\,045 + A(0.5 \times 10^{-15}).$$

相反，为了获得一个额外的十进制位，必须对近似值知道得更加准确。不难看出为什么工程师和科学家通常对4位数的精度感到满意，以及为什么9或10位数的精度大致是测量的极限。例如，到月球的距离大约是238,000英里，或者说大约是1.508×10^{10}英寸。使用阿波罗宇航员留下的镜子反射激光，科学家可以将到月球的距离测量到大约6英寸的精度，这正好是9位数的精度。20位数的精度将对应于一个原子直径量级的实验误差，而在这种情况下，50位数的精度在物理上是没有意义的，因为非常小的距离不能很好地用实数来建模，而是受到量子力学后果的影响。

与相比之下，一个纯粹数学家除非能找到任意多位小数，否则不太可能感到满意；任何少于这个数都是不确定的。一个引人注目的例子是C. Hermite[2]在1859年发现的那个数，其数值为

$$x_0 := e^{\pi\sqrt{163}} = 262,537,412,640,768,744 + A(10^{-12}).$$

是 x_0 精确地是一个整数吗？数值证据是压倒性的：误差项是 $\pm 0.000\,000\,000\,000\ldots$，因此 x_0 是一个整数到一位

[1] Really, the distance from a laser in a telescope to a mirror on the moon!

[2] air MEET

部分在10^{30}中，在科学测量中绝对无法达到的精度。然而，这种“推理”是愿望思考；实际上，这种巧合在许多情况下必须发生。误差项不是零，而是$A(0.75\times10^{-12})$！数学家有时被实验科学家视为繁琐（“他们让你证明明显的事情。”），但数学家的怀疑并非无端。

甚至科学家和应用数学家都对“数学精度”有既得利益，因为数值误差在计算中往往会迅速增长，每次乘法都会丢失固定数量的有效数字。许多学生在被要求评估 $e^{10\log 2}$ 时，会首先将对数四舍五入，log 2 $\simeq$ 0.693（到3位小数），然后乘以10并求指数，得到 $e^{10\log 2}\simeq$ 1022.494。然而，指数和对数的性质表明，确切值是 $2^{10}=1024$。这个例子表明了单步中的精度损失；即使是一个简化的文字问题也可能有三四步这样的步骤，而且粗心或过早地使用数值常数可能会导致答案完全错误。不幸的是，计算器助长了在计算开始时输入数字的坏习惯。理论科学家或数学家必须是一个熟练的符号计算器，但即使你渴望成为一个成功的实验科学家，你也必须培养计算符号的能力。

这些考虑解释了数学家坚持对像 e 和 π (这样的常数的程序定义，这些定义是精确的，可以转化为计算算法，以任意精度) 进行计算，而不是对数值“规格”（例如“$\pi\ \simeq$ 3.141 592 653 589 793 . . .”）进行定义，这些根本不是数学定义。

对于数学家来说，“评估”一个数值量通常意味着找到一个表达式，用初等函数和已知的超越常数来表示其确切值。例如，数字 π 被定义为余弦函数周期的1/2，而

$$\sum_{k=1}^{\infty}\frac{1}{k^2}=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots$$

被定义为序列部分和的极限。我们很可能无法精确评估给定的无穷级数，即使我们能够证明该级数存在。当然，可以通过任意数值精度来评估该级数，这在某种程度上是令人满意的，但在此基础上，任何数学家都不会说该级数已经被“评估”。如果，比如说，该级数被证明

等于 $\pi^2/6$（正如上述例子中所示），因为在这种情况下，总和的精确值可以用 π 来表示，而 π 在数学上与整数一样熟悉和普遍。在数值上，这样的信息是有帮助的，因为 π 的数值近似允许我们无需任何努力地近似总和。从级数或积分的精确评估中产生的这些宝石常常揭示了隐藏的现象，因此它们的重要性超越了其固有的美。

14.2 函数逼近

上述讨论适用于函数。类似于方程（14.1），我们希望找到如下表达式

$$e^x = \underbrace{\left(\sum_{k=0}^{n} \frac{x^k}{k!}\right)}_{\text{estimate}} + \underbrace{A\left(\frac{3|x|^{n+1}}{(n+1)!}\right)}_{\text{error}} \quad \text{for } |x| \le 1. \tag{14.2}$$

在这个方程中，e^x 是我们希望近似的数值。右侧的第一个项是我们对 e^x 的估计，而第二个项，误差项，是 e^x 与估计值之间差异的上界，并衡量估计的准确性。方程（14.2）的关键特性是估计和误差都是 x 的多项式，可以使用仅算术运算进行数值评估。取 $x = 1$ 和 $n = 6$ 得到

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + A\left(\frac{3}{7!}\right),$$

或 $e = 2.718 + A(0.0006)$。为了提高精度，我们会选择更大的 n。像（14.2）这样的方程封装了无限多个数值条件（每个 x 一个），具有明显的计算兴趣。

本节的目标是为具有足够多导数的函数建立类似于（14.2）的估计。策略是在点 x_0 附近通过选择一个至多 n 次的泰勒多项式 p_n 来近似函数 f，该多项式“匹配” f 及其在 x_0 处的至多 n 阶导数。当 $n =$ = 1 时，这种策略产生切线，这是一个导数为 $f'(x_0)$ 的线性函数。我们的目标是三方面的:

- 提供一个有效的计算过程来找到 p_n。

- 证明在适当的意义下，p_n 是度数不超过 n 的“最佳逼近多项式”。
- 找到一个误差项的有效计算界限。

一旦这些问题得到解决，关于“计算器如何知道返回哪个数值”的问题几乎得到了解答：计算器被编程为评估常见函数的泰勒多项式算法，返回的值与实际（数学）值相差小于0.5×10^{-8} (say)。泰勒多项式将计算e^x或$\cos x$简化为加法、减法、乘法和除法。这些算术运算本身由浮点单元（FPU）执行，这是一个在第二章中我们构建整数之上的级别仅略高的电路。

泰勒多项式

设 f 是定义在某些包含 $x_0 \in \mathbf{R}$ 的开区间上的函数，并假设 f 对于某个正整数 N 是 N 次可微的。我们将构造一个多项式序列 $\{p_n\}_{n=0}^{N}$，使得

(i) 对于每个 n，多项式 p_n 的次数最多为 n。

(ii) 对于$0 \leq k \leq n$，导数 $f^{(k)}(x_0)$ 和 $p_n^{(k)}(x_0)$ 一致。

序列由这两个性质唯一指定（正如我们将看到的），称为在 f 处的 x_0 的泰勒多项式序列，直到 N 次度。结果，p_n 和 p_{n-1} 之间的差是一个 n 次的单项式，因此 p_N 立即确定了具有 $k < N$ 的每个泰勒多项式 p_k。在实践中，这意味着我们不需要写下整个序列，而只需写下多项式 p_N。

备注14.1 泰勒多项式 p_n 依赖于函数 f，“中心”点 x_0 和次数 n，但在上下文中明确时，省略函数和中心点以简化符号。在应用中，中心通常是0，但理论上允许 x_0 为任意值。必要时，以 x_0 为中心的泰勒多项式表示为 p_{n,x_0}。 □

要找到泰勒多项式的系数，我们不是在 x 的幂次上展开，而是在 $(x - x_0)$ 的幂次上展开。设 f 是 n 次可微的

在 x_0 的一个邻域内，并写出

$$p_n(x) = \sum_{j=0}^{n} a_j(x - x_0)^j$$

具有 $\{a_j\}_{j=0}^n$ 未知；右侧的表达式是满足条件（i）的度数至多为 n 的一般多项式，接下来我们等于系数，如（ii）中所述。

引理14.2。以上以 p_n 为例，$p_n^{(k)}(x_0) = f^{(k)}(x_0)$ 当且仅当 $a_k = \dfrac{f^{(k)}(x_0)}{k}!$ 。

证明。我们通过累加和式的贡献来计算 $p_n^{(k)}(x_0)$，以 p_n 的系数表示，

$$(*) \qquad \frac{d^k}{dx^k}\bigg|_{x=x_0} a_j(x - x_0)^j.$$

如果 $j < k$，导数恒为零（每次求导都会使指数降低1），而如果 $j > k$，k 阶导数能被 $(x - x_0)$ 整除，因此 x_0 处消失。（参见例14.6。）$(*)$ 的唯一非平凡贡献来自包含 $j = k$ 的项；证明它通过归纳法留给你。

$$(**) \qquad \frac{d^k}{dx^k}\bigg|_{x=x_0} a_k(x - x_0)^k = k!\, a_k.$$

因此 $p_n^{(k)}(x_0) = k!\ a_k$。词元紧随其后。 □

根据引理，性质（ii）成立（p_n 和 f 的导数在 n 阶上相同）当且仅当

$$(14.3) \qquad p_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

这是我们要找的简单公式。只要我们能计算 f 的导数并在 x_0 处评估它们，我们就有泰勒多项式的显式表示。

TheTranslated Text: The 连续泰勒多项式之间的差 als is

$$p_n(x) - p_{n-1}(x) = \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n,$$

一个次数为 n 的单项式。这确立了先前的断言，即每个泰勒多项式 p_n “包含”所有具有 $k < n$ 的多项式 p_k。要从 p_n 获得到 p_k，只需删除所有次数大于 k 的项。要从 p_{n-1} 获得到 p_n，我们只需要添加一个（可能为零的）次数为 n 的单项式。

备注14.3在第11章中，我们看到了任意连续函数在每一个闭的、有界区间上都可以由多项式一致逼近。然而，这样的逼近序列通常不具有p_n和p_{n-1}之间差一个单项式的性质。我们在一般性上所获得的，是以失去简单性为代价的：Weierstrass多项式不需要可微性假设，但关于p_n的知识并不能告诉我们关于p_k与$k < n$的任何信息。□

示例14.4（指数函数）因为对于所有正整数k，$\exp^{(k)}(x) = \exp x$，对于所有k，$\exp^{(k)}(0) = 1$，所以exp在0处的泰勒多项式是

$$p_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

泰勒多项式恰好是exp的幂级数的部分和，由定理12.7给出。这是实分析函数的一般特性，参见推论14.11。□

示例14.5（圆三角函数）基本三角函数sin和cos有易于从定义计算出的泰勒多项式.[3]至于exp，其泰勒多项式是幂级数的截断。例如，sin的$2n+1$次泰勒多项式是

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

在这个例子中，二阶{v*}泰勒多项式等于一阶{v*}泰勒多项式，因为幂级数的偶数项为零。{v*}

示例14.6（幂函数）固定 $a \in \mathbf{R}$ 和一个正整数 N，并考虑多项式 $f(x) = (x-a)^N$。逐次微分给出

$$\begin{aligned} f'(x) &= N(x-a)^{N-1} \\ f''(x) &= N(N-1)(x-a)^{N-2} \end{aligned}$$

[3]The other trig functions do not, as it turns out.

并且通常

$$\begin{aligned} f^{(k)}(x) &= N(N-1)\cdots(N-k+1)(x-a)^{N-k} \\ &= \frac{N!}{(N-k)!}(x-a)^{N-k} \qquad \text{for } k=0,\ldots,N. \end{aligned}$$

将 (14.3) 代入得到 $x_0=0$ 处的泰勒多项式:

$$p_n(x) = \sum_{k=0}^{n} \binom{N}{k} (-a)^{N-k} x^k, \qquad n=0,\ldots,N.$$

当 $n=N$ 时，根据二项式定理，右侧为 $(x-a)^N$，对于 $n \geq N$ 也成立，因为 f 的高阶导数为零。N 次泰勒多项式恰好是 f 的 N 次多项式 f 本身，这并非偶然，见下文定理14.9。□

示例14.7（二项式级数）设α为一个非非负整数的实数，并定义 $f(x)=(1+x)^\alpha$对于$x>-1$。函数f在0处无限可微，并且与前面的例子一样,

$$f^{(k)}(x) = \alpha(\alpha-1)\cdots(\alpha-k+1)(1+x)^{\alpha-k} \qquad \text{for } x>-1.$$

n次f在0处的泰勒多项式是

$$p_n(x) = \sum_{k=0}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k; \tag{14.4}$$

系数 x^k 与组合二项式系数 [4] 完全类似，因此对于一般的 α，读作“从 α 中选择 k”。以下三种情况值得进一步说明:

- $\alpha=-1$. In this case, $f(x)=(1+x)^{-1}=1/(1+x)$, and the coefficient of x^k is $\big((-1)(-2)(-3)\cdots(-k)\big)/k! = (-1)^k$, so

$$p_n(x) = 1-x+x^2-x^3+\cdots+(-x)^n.$$

[4] That is, if α were a non-negative integer, then the coefficient of x^k would be a combinatorial binomial coefficient.

- $\alpha = 1/2$. 在这种情况下，$f(x) = \sqrt{1+x}$，并且 x^k 的系数是
$$\frac{(1/2)(-1/2)(-3/2)(-5/2)\cdots\big((3-2k)/2\big)}{k!} = (-1)^{k-1}\frac{(2k-3)!!}{2^k\,k!},$$
因此，泰勒多项式是
$$p_n(x) = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \frac{5x^4}{128} + \cdots - \frac{(2n-3)!!}{2^n\,n!}(-x)^n.$$
- $\alpha = -1/2$。这里，$f(x) = (1+x)^{-1/2} = 1/\sqrt{1+x}$，并且 x^k 的系数是
$$\frac{(-1/2)(-3/2)(-5/2)\cdots\big((1-2k)/2\big)}{k!} = (-1)^k\frac{(2k-1)!!}{2^k\,k!};$$
泰勒多项式是
$$p_n(x) = 1 - \frac{x}{2} + \frac{3x^2}{8} - \frac{5x^3}{16} + \frac{35x^4}{128} - \cdots + \frac{(2n-1)!!}{(2n)!!}(-x)^n.$$

当指数 α 是非负整数（因此幂函数是度数为 α 的多项式）时，泰勒多项式在度数为 α 时“稳定”；否则，我们得到一个无限序列的不同多项式。□

在先前的例子中，泰勒多项式是直接从定义中找到的。存在一些情况下，直接方法不可行，需要采取迂回策略。你不应该有这样的印象，即泰勒多项式必须从定义中计算出来；参见推论14.10。

接触顺序

设 f 和 g 是定义在某些 x_0 邻域上的函数。回忆一下，大-O 和小-o 符号给出了 f 和 g 之间距离的一个度量。设 ϕ 是定义在某些 x_0 邻域上的函数，设 n 是一个正整数。我们说 $\phi = O(|x-x_0|^n)$ 如果存在一个常数 C 和一个 $\delta > 0$ 使得
$$|\phi(x)| \le C|x-x_0|^n \qquad \text{for } |x-x_0| < \delta.$$
我们说 $\phi = o(|x-x_0|^n)$ 如果
$$\lim_{x\to x_0} \frac{\phi(x)}{(x-x_0)^n} = 0.$$

明确地

$$\phi = o(|x-x_0|^n) \implies \phi = O(|x-x_0|^n) \implies \phi = o(|x-x_0|^{n-1}).$$

两个蕴涵在一般情况下都是不可逆的：$|x|^2$ 是 $O(|x|^2)$ 但不是 $o(|x|^2)$，而 $|x|^{3/2}$ 是 $o(|x|)$ 但不是 $O(|x|^2)$。

两个 n-次可微函数 f 和 g 在 x_0 处等于 n 阶，即如果 $f-g = o(|x-x_0|^n)$，也就是说如果

$$\lim_{x\to x_0} \frac{f(x)-g(x)}{(x-x_0)^n} = 0.$$

有时人们说，在这种情况下，f和g的图形在x_0处具有阶-n接触。阶-0接触意味着图形相交，而阶-1接触意味着图形相切。更高阶的接触被视为“高阶切触”。

我们接下来量化这样一个说法：一个可 n 次微分的函数 f 的 n 次泰勒多项式是“最佳”逼近。换句话说，接下来的两个定理断言 f 和 p_{n,x_0} 在 x_0 处具有 n 阶接触，而泰勒多项式是唯一具有这种性质的至多 n 次的多项式。

定理14.8. 如果 f 在 x_0 的 $\mathscr{C}^n$ 附近，那么 $f - p_{n,x_0} = o\big(|x-x_0|^n\big)$。

证明。因为 k 阶导数 f 和 p_n 在 x_0 处连续且对 $k \le n$ 一致，我们可以应用拉格朗日中值定理 n 次：

$$\lim_{x\to x_0} \frac{f(x)-p_{n,x_0}(x)}{(x-x_0)^n} = \lim_{x\to x_0} \frac{f^{(n)}(x)-p^{(n)}_{n,x_0}(x)}{n!} = 0,$$

这完成了证明。 $\square$

定理14.9。如果p和q是次数不超过n的多项式，并且对于某个x_0，$p-q$是$o(x-x_0)^n$，那么$p=q$。

证明。假设意味着对于 $k = 0, \dots, n$，$p-q$ 是 $o(x-x_0)^k$。写出 $(p-q)(x) = b_0 + b_1(x-x_0) + \cdots + b_n(x-x_0)^n$。取 $k=0$ 表示 $b_0 = 0$。删除第一个项并取 $k=1$ 表示 $b_1 = 0$。依次进行，表明对于所有 $k \le n$，$b_k = 0$，这意味着 $p=q$。 $\square$

一个关于 n 的归纳证明是直接的，如果你对论证的不正式感到困扰，应该提供它。为了记录，这里是有用的结果，可以称为“泰勒多项式的唯一性”。

推论14.10.若f在x_0处是n次连续可微的，且若p是一个在x_0处与f接触的阶数为n的多项式，则$p = p_{n,x_0}$。

推论14.11。如果f是实分析，那么通过截断幂级数到n次幂，可以得到n次泰勒多项式。

证明。由引理11.20，{v*}

$$
\begin{aligned}
f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k &= \sum_{k=0}^{n} a_k(x-x_0)^k + \sum_{k=n+1}^{\infty} a_k(x-x_0)^k \\
&= \sum_{k=0}^{n} a_k(x-x_0)^k + O\big(|x-x_0|^{n+1}\big),
\end{aligned}
$$

因此，推论14.10表明有限和是泰勒多项式。 □

在第三章中，我们声称存在一种简单的计算程序来展开以$(x-a)$为幂的多项式。作为推论14.10的应用，我们描述了这一程序。

示例14.12 设$p(x) = \sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} b_k (x-a)^k$。根据14.10引理，右边的多项式是$f$在$a$处的$n$次泰勒多项式，其系数由$b_k = p^{(k)}(a)/k!$给出。

例如，假设我们想将 $p(x) = (x+1)^4$ 写成 $(x-1)$ 的幂次。取 $a = 1$，我们计算

$$
\begin{array}{lll}
f(x) = (x+1)^4 & f(1) = 2^4 = 16 & b_0 = f(1) = 16 \\
f'(x) = 4(x+1)^3 & f'(1) = 4 \cdot 2^3 = 32 & b_1 = f'(1)/1! = 32 \\
f''(x) = 12(x+1)^2 & f''(1) = 12 \cdot 2^2 = 48 & b_2 = f''(1)/2! = 24 \\
f'''(x) = 24(x+1) & f'''(1) = 24 \cdot 2 = 48 & b_3 = f'''(1)/3! = 8 \\
f^{(4)}(x) = 24 & f^{(4)}(1) = 24 & b_4 = f^{(4)}(1)/4! = 1
\end{array}
$$

我们立即读出

$$(x+1)^4 = 16 + 32(x-1) + 24(x-1)^2 + 8(x-1)^3 + (x-1)^4,$$

一个可能通过展开（费力地）检查的恒等式。 □

推论14.10可以间接用于计算泰勒多项式。

示例14.13 考虑计算 $f = \mathrm{Tan}^{-1}$ 在 $x_0 = 0$ 处的泰勒多项式的问题。直接方法是一条死胡同；前几个导数是

$$f'(x) = \frac{1}{x^2+1}, \quad f''(x) = -\frac{2x}{(x^2+1)^2}, \quad f'''(x) = \frac{6x^2-2}{(x^2+1)^3}, \ \ldots$$

并且既没有容易辨认的模式，也没有简化，使得可以找到一般导数。相反，我们按以下方式推理。根据有限几何级数公式，如果 n 是一个正整数，那么

$$\begin{aligned}\frac{1}{1+t^2} &= 1 - t^2 + t^4 - \cdots + (-1)^n t^{2n} + \frac{(-1)^{n+1}t^{2n+1}}{1+t^2} \\ &= \left(\sum_{k=0}^{n}(-1)^k t^{2k}\right) + \frac{(-1)^{n+1}t^{2n+2}}{1+t^2} \qquad \text{for all } t \in \mathbf{R}.\end{aligned}$$

从0到x的积分给出

$$\mathrm{Tan}^{-1}x = \int_0^x \frac{1}{1+t^2}dt = \left(\sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{2k+1}\right) + \int_0^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt.$$

括号中的项是次数为($2n+1$)的多项式，该积分是待估计的“误差项”。现在,

$$\left|\int_0^x \frac{(-1)^{n+1}t^{2n+2}}{1+t^2}\,dt\right| \le \int_0^{|x|}\left|\frac{t^{2n+2}}{1+t^2}\right|dt \le \int_0^{|x|}\left|t^{2n+2}\right|dt = \frac{|x|^{2n+3}}{2n+3},$$

这证明了误差项是 $o(|x|^{2n+2})$。我们将 Tan^{-1} 写成次数为 ($2n+1$) 的多项式和误差项之和，该误差项是 $o(|x|^{2n+1})$。根据推论14.10，该多项式是 Tan^{-1} 在0处的次数-($2n+1$) 泰勒多项式。□

这个例子有一个历史性的尾声。在先前的讨论中将 $x =$ 设置为 1 给出

$$\frac{\pi}{4} = \mathrm{Tan}^{-1}1 = \sum_{k=0}^{n}(-1)^k\frac{1}{2k+1} + A\Big(\frac{1}{2n+3}\Big)$$

当n增加时，误差项的绝对值减小到0，所以

$$\frac{\pi}{4} = \sum_{k=0}^{\infty}(-1)^k\frac{1}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots. \tag{14.5}$$

很遗憾，这个显著的方程没有提供计算 π 的良好数值技术；误差项太大。为了保证三位小数，误差必须不大于 $0.5 \times 10^{-3} = 1/2000$。上述级数无法达到这种精度

直到求和999项。[5] 更糟糕的是，每个额外的十进制位需要求和的项数增加十倍。即使对于这个序列，也有更好的数值技术来计算 π，参见练习14.13。一般思路是使用三角恒等式将已知常数与Tan^{-1}在多个小绝对值处的值联系起来；$|x|^{2n+3}$使得相对较小的 n 误差很小。

余项

函数 f 与其一阶泰勒多项式之间的差称为余项。记住，本章的一般问题是数值评估非有理函数，策略是将 f 写成一个多项式加上一个易于估计的误差项。“泰勒定理”允许系统地估计余项，前提是已知足够关于 f 的信息。余项有三种标准表达式，称为积分形式、拉格朗日形式和柯西形式。每种形式在特定应用中都很有用；我们只讨论积分和拉格朗日形式，因为它们对我们来说足够且易于记忆。我们也不尝试给出最弱的可微假设，因为我们感兴趣的大多数函数都是光滑的。

定理14.14. 假设f在某个邻域$N_\delta(x_0)$上属于类$\mathscr{C}^{n+1}$，并记$R_{n,x_0}(x) = f(x) - p_{n,x_0}(x)$为$|x - x_0| < \delta$。那么

$$R_{n,x_0}(x) = \int_{x_0}^{x} \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,dt \qquad \text{(Integral Form)},$$

存在 z，使得 $|z - x_0| < \delta$

$$R_{n,x_0}(x) = \frac{f^{(n+1)}(z)}{(n+1)!}(x-x_0)^{n+1} \qquad \text{(Lagrange Form)}.$$

从定性上看，如果要将一个($n+1$)-次可微函数 f 在 x_0 附近尽可能精确地用至多 n 次的多项式逼近，那么最佳策略是使用 n 次的泰勒多项式，在这种情况下

$$f(x) = p_{n,x_0}(x) + O\big(|x-x_0|^{n+1}\big).$$

定理14.14给出了O中常数的具体界限。

[5] We have proven only that 999 terms suffice to give 3 decimal places, but it is not difficult to show that in this example the error term is no *smaller* than $1/(4n+6)$, so *at least* 499 terms are needed.

证明。第二个基本定理表明

$$P(0): \qquad f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\,dt \qquad \text{for } |x - x_0| < \delta,$$

这正是 $n = 0$ 的余项的积分形式。假设归纳法如下:

$$P(n): \qquad f(x) = \underbrace{\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k}_{p_{n,x_0}(x)} + \underbrace{\int_{x_0}^{x} \frac{f^{(n+1)}(t)}{n!}(x - t)^n\,dt}_{R_{n,x_0}(x)}$$

通过分部积分法积分余项，使用

$$u(t) = f^{(n+1)}(t) \qquad v(t) = -\frac{1}{(n+1)!}(x - t)^{n+1}$$

$$u'(t) = f^{(n+2)}(t) \qquad v'(t) = \frac{1}{n!}(x - t)^n\,dt$$

自 $\int uv' = uv - \int v'u$ 以来，积分变为

$$\begin{aligned} R_{n,x_0}(x) &= \int_{x_0}^{x} \frac{f^{(n+1)}(t)}{n!}(x - t)^n\,dt \\ &= -\frac{f^{(n+1)}(t)}{(n+1)!}(x - t)^{n+1}\Bigg|_{t=x_0}^{t=x} + \int_{x_0}^{x} \frac{f^{(n+2)}(t)}{(n+1)!}(x - t)^{n+1}\,dt \\ &= \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} + \int_{x_0}^{x} \frac{f^{(n+2)}(t)}{(n+1)!}(x - t)^{n+1}\,dt. \end{aligned}$$

添加到 $P(n)$ 得到 $P(n+1)$。通过归纳，余项的积分形式得以建立:

$$R_{n,x_0}(x) = \int_{x_0}^{x} \frac{f^{(n+1)}(t)}{n!}(x - t)^n\,dt \qquad \text{for all } n \geq 0.$$

为了获得余项的拉格朗日形式，令 M 和 m 表示 $f^{(n+1)}$ 在 x_0 和 x 之间的最大值和最小值。积分的单调性意味着

$$\frac{m}{(n+1)!}(x - x_0)^{n+1} \leq R_{n,x_0}(x) \leq \frac{M}{(n+1)!}(x - x_0)^{n+1}.$$

因此，存在 c 满足 $m \le c \le M$ 和

$$R_{n,x_0}(x) = \frac{c}{(n+1)!}(x-x_0)^{n+1}.$$

根据介值定理，存在一个 z 在 x_0 和 x 之间，使得 $c = f^{(n+1)}(z)$。

□

对于基本初等函数（指数、正弦和余弦），泰勒定理给出了误差项的有效界限，这进而定量地说明了特定多项式逼近的好坏。

示例14.15设f =exp为自然指数函数。由于$f' = f$，对k的归纳表明$|f^{(k)}(x)| = |e^x| \le e$对所有$k \in \mathbf{N}$和$|x| \le 1$成立。示例14.4和拉格朗日余项形式表明，对于所有$n \in \mathbf{N}$和$|x| \le 1$，

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

我们从第7章的练习7.17中知道 $e < 3$。前一个方程立即意味着方程（14.2）。参见图14.1以了解这些误差界限的几何解释。请注意，估计误差（阴影区域的宽度）在 $x = 0$ 处急剧增加。

当 $|x| \le 1$ 时，指数级数收敛迅速；例如，$1/10! < 2.76 \times 10^{-7}$，因此使用直到9次方的项来估计 $e^{1/2}$ 给出的误差小于

$$\frac{3 \cdot (1/2)^{10}}{10!} < 8.1 \times 10^{-10}.$$

平方结果，精度优于5×10^{-9}，即8位小数。□

示例14.16 对于正弦函数，没有偶数次幂的项，并且所有导数的绝对值都受限于1，所以

$$\begin{aligned}|R_{2n+1}(x)| = |R_{2n+2}(x)| &= \left|\int_0^x \frac{\sin^{(2n+3)}(t)}{(2n+2)!}(x-t)^{2n+2}\,dt\right| \\ &\le \left|\int_0^x \frac{(x-t)^{2n+2}}{(2n+2)!}\,dt\right| \quad \le \quad \frac{|x|^{2n+3}}{(2n+3)!}.\end{aligned}$$

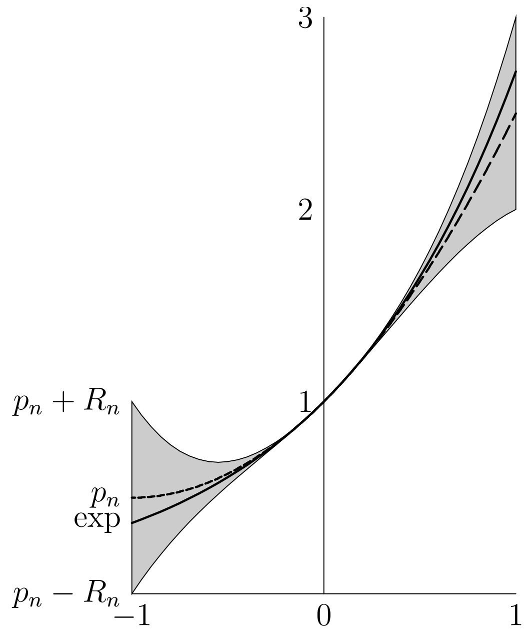


图14.1：在[−1, 1]上的exp的泰勒近似

因此

$$\sin x = \left(\sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{(2k+1)!}\right) + A\left(\frac{|x|^{2n+3}}{(2n+3)!}\right). \tag{14.6}$$

假设我们希望计算 sin x 对于 x 个实数。因为 sin 是 2π-周期的，所以我们只需要一个准确的 $|x| \leq \pi$ 表格，并且由于 $\sin(\pi - x) = \sin x$ 对于所有 x，实际上只需要一个准确的 $0 < x \leq \pi/2$ 表格，并且准确计算 π。回忆一下，根据第 13 章的一个结果，$\pi/2 < 1.6$。表 14.1 中的误差界限（$(1.6)^{2n+3}/(2n+3)!$）由泰勒定理保证。

作为一个实际操作，为了计算sin x 到5位小数，对于任意的 $|x| \leq \pi/2$，使用9次幂泰勒多项式就足够了，而13次幂的多项式几乎可以给出9位小数。要计算 x 在此区间外的 sin 值，首先添加或减去适当的2π的倍数，以得到 $[-\pi, \pi]$ 区间内的一个数，然后使用关系 $\sin(-x) = -\sin x$ 得到一个 $[0, \pi]$ 区间内的数，最后使用 $\sin(\pi - x) = \sin x$ 得到一个 $[0, \pi/2]$ 区间内的数。这些操作依赖于 π 本身的准确值（例如20位小数），这可以轻松地硬编码到计算器芯片或计算机程序中。

Degree $(2n+2)$	$p_{2n+2,0}(x)$	Error Bound
2	x	0.683
4	$x-\frac{x^3}{3!}$	0.0874
6	$x-\frac{x^3}{3!}+\frac{x^5}{5!}$	0.00533
8	$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}$	0.00019
10	$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\frac{x^9}{9!}$	0.00000441

表14.1：用泰勒多项式逼近 sin {v*}。

应注意的是，对于接近0的x，需要相当少的项来获得相似的精度。例如，

$$\sin(0.1) = (0.1) - \frac{(0.001)}{6} + \text{error}, \qquad |\text{error}| \leq \frac{(0.1)^5}{5!} < 10^{-7};$$

三次泰勒多项式提供了 sin(0.1) 的 6 位小数。 □

示例14.17 令 $f : [-1, 1) \to \mathbf{R}$ 由 $f(x) = \log(1-x)$ 定义。泰勒多项式可以通过直接方法（练习14.5）或与用于 Tan^{-1} 相同的技巧找到：

$$f'(x) = -\frac{1}{1-x} = -\left(\sum_{k=0}^{n} x^k\right) - \frac{x^{n+1}}{1-x},$$

当 $f(0) = 0$，所以

$$f(x) = \int_0^x f'(t)\,dt = -\left(\sum_{k=0}^{n} \frac{x^{k+1}}{k+1}\right) - \int_0^x \frac{t^{n+1}}{1-t}\,dt.$$

错误项在 $x \to 1$ 处越来越难以处理（正如预期的那样，因为 f 在 1 附近是无界的），但有一个简单的估计

$$\left|\int_0^x \frac{t^{n+1}}{1-t}\,dt\right| \leq \left(\frac{1}{1-a}\right)\frac{|x|^{n+2}}{n+2} \qquad \text{for } x \leq a < 1.$$

为了计算以2为底的对数，例如，我们首先设置 $a = 1/2$，然后通过在 $x = 1/2$ 处评估级数来计算 log 1/2 = − log 2；这利用了误差界中的因子 $(1/2)^{n+1}$。显然的方法，设置 $x = -1$，给出了误差界 $1/(n+1)$，它下降得非常慢，类似于 (14.5) 中的界限。然而，设置 $x = -1$ 给出了另一个著名的评估，即交错调和级数的和：

$$\log 2 = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

请注意，虽然

$$\log(1-x) = -\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} \qquad \text{for } -1 < x < 1,$$

它并不立即意味着在 $x = -1$ 处等式成立，尽管左边的函数是连续的，而右边的级数在 $x = -1$ 处是收敛的：右边的收敛在 $x = -1$ 附近不是一致的。□

二项式级数

设 α 为一个非正整数的实数。如方程（14.4）中计算所示，函数 $f(x) = (1+x)^\alpha$ 有泰勒多项式

$$p_n(x) = 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k.$$

比值测试表明，所得到的级数

$$g(x) = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k =: 1 + \sum_{k=1}^{\infty} a_k\, x^k$$

具有半径

$$\lim_{k\to\infty} \left|\frac{a_k}{a_{k+1}}\right| = \lim_{k\to\infty} \left|\frac{k+1}{\alpha-k}\right| = 1,$$

因此，g的定义域包含开区间$(-1, 1)$。除了与我们为exp和圆三角函数所做类似的技术细节之外，这证明了牛顿二项式定理：

定理14.18. 如果 α 不是一个非负整数，那么

$$(1+x)^\alpha = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k \qquad for\ |x| < 1.$$

剩余步骤在练习14中概述 .14.

练习

练习14.1 令 $a \in \mathbf{R}$。使用 a 作为展开中心，计算 exp 的泰勒多项式。得出结论 $e^x = e^a e^{x-a}$ $\diamond$

练习14.2 使用泰勒定理估计用其二阶泰勒多项式的值来逼近cos {v*} 的误差。需要多少项来估计cos 1到小数点后四位？{v*}

练习14.3 求cosh的泰勒多项式。在计算之前，你应该能够猜出它们是什么。$\diamond$

练习14.4 从定义直接求出tan的四阶和五阶泰勒多项式。

$\diamond$

练习14.5 设 $f(x) = \log(1-x)$ 对于 $|x| < 1$。从定义中找出 f 的 n 阶泰勒多项式。$\diamond$

练习14.6 设 $f(x) = \log(1+x^2)$ 对于 $|x| < 1$。求 f 的一般泰勒多项式。请注意，直接方法不是最容易的方法。$\diamond$

练习14.7比较在练习14.5中找到的泰勒多项式p_n与在练习14.6中找到的多项式q_n：$q_n(x) = p_n(-x^2)$。陈述并证明一个定理，将f和$g(x) = f(cx^2)$的泰勒多项式与$c \in \mathbf{R}$联系起来。$\diamond$

练习14.8 求以下函数的泰勒级数：

(a) $f_1(x) = e^{-x^2}$ (b) $f_2(x) = \sin(x^2)$ (c) $f_3(x) = \cos(x^2)$.

建议：使用之前的练习。 $\diamond$

练习14.9 使用级数乘积验证

$$\frac{1}{1-x^2} = \frac{1}{1+x} \cdot \frac{1}{1-x}$$

对于 $|x| < 1$. $\diamond$

练习14.10 求以下函数的泰勒级数：{v*}

$$\tanh x = \int_0^x \frac{1}{1-t^2}\,dt.$$

对于哪个 x 级数收敛? $\diamond$

练习14.11 让

$$f(x) = \int_0^x e^{-t^2}\,dt.$$

(a) 求解 f 的泰勒级数。该级数在哪些 x 下收敛?

(b) 仅使用计算器上的算术函数，将 $f(1)$ 计算到小数点后4位。提示：需要多少项的级数?

(©) 证明 $\lim\limits_{v*}(f,+\infty)$ 存在。这个问题与泰勒级数的内容无关。仔细解释为什么级数在计算不定积分时没有帮助。

一个与 f 密切相关的函数出现在概率中。 $\diamond$

练习14.12 使用泰勒定理来界定exp的n次余项。当$x = 1$时，这个估计比第12章中用级数得到的估计更好还是更差? $\diamond$

练习14.13 使用 $\tan\frac{\pi}{6} = \frac{1}{\sqrt{3}}$ 证明

$$\frac{\pi\sqrt{3}}{6} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)3^k}$$

需要多少项才能保证3位小数的精度? $\diamond$

练习14.14 本练习概述了定理14.18的证明。假设α不是一个非负整数，并设$f(x) = (1+x)^\alpha$和

$$g(x) = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}x^k \qquad \text{for } |x| < 1.$$

(a) 使用级数变换来证明 $(1+x)g'(x) = \alpha g(x)$ 在 $(-1, 1)$ 上。

(b) 使用(a)部分来证明g/f在$(-1, 1)$上是常数。

可能有助于回顾第12章中exp的唯一性证明。 ◇

练习14.15 证明“等于在 x_0 处的 r 阶导数”的性质是在定义在 x_0 的某个邻域上的光滑函数空间上的一个等价关系。一个等价类称为 x_0 处的 “r-jet”。◇

翻译练习

你如何吸收数学语言？例如，当你听到，“整个区间内项 f_n 的总贡献变得可以忽略不计”时，你应该想到：

> 设 $I \subset \mathbf{R}$ 为一个区间，设 (f_n) 为 I 上的一个函数序列。
> 对于每一个 $\varepsilon > 0$，存在一个正整数 N，使得
> $\sum_{n=N+1}^{\infty} |f_n(x)| < \varepsilon$ 对于 x 在 I 中的所有值都成立。

这是您与歌德一较高下的机会：

练习14.16 将以下非正式短语或句子翻译成精确的、量化的陈述。

- 如果 (a_n) 和 (b_n) 是正项序列，并且当 $n \to \infty$ 时 a_n/b_n 趋近于 1 /2，那么对于足够大的 n，有 $0 < a_n < b_n$。
- 函数值 $f(x)$ 在 x 接近 x_0 时接近 $f(x_0)$。
- 函数 $f(x) = x^{-(1+1/x)}$ 当 $x \to \infty$ 时渐近于 $1/x$。
- 对于 $x \simeq 0$, $1 - \cos x$ 远小于 x。
- 将足够多的项 a_n 相加，部分和可以接近到所需的和。
- 让内接矩形的宽度趋近于零，$f : [a, b] \to \mathbf{R}$ 的图形下的面积可以任意接近地近似。
- 项 a_n 可以任意排列而不改变级数的和。
- 术语 a_n 最终减少到零。

如果适当，则识别属性或定理。 ◇

第15章

初等函数

回忆一下，一个函数 f 在 x_0 处是实解析的，如果存在一个收敛半径为正的幂级数在某邻域 x_0 上等于 f。对于收敛区间内的每个 a，幂级数可以在 a 处展开，新级数的收敛半径为正。因此，在一点解析的函数在开区间内也是解析的。

我们的分析函数库包括

- 代数函数（由两个变量的多项式隐式定义的函数），例如 $f(x) = \sqrt[3]{1+x^2}/\sqrt{3-x}$，满足

  $$F(x,y) = (1+x^2)^2 - (3-x)^3y^6 = 0.$$

  一元多项式和有理函数属于这一类别。

- 指数和对数函数，以及双曲三角函数。

- 圆三角函数。

一个可以通过有限次算术运算和函数复合从这些族中获得的函数被称为基本函数。“典型”的基本函数包括 $f_1(x) = \log\bigl(x+\sqrt{x^2-1}\bigr)$
，

$$f_2(x) = x^x = e^{x\log x}, \qquad f_3(x) = e^{e^{\sin[1+\sqrt{\log(4+x^2)}]}}.$$

15.1 复分析简明课程

圆三角函数与双曲三角函数在形式上具有强烈的相似性，它们显然与指数函数有关。为了解释这种相似性背后的魔法，有必要在复数域内进行工作，特别是要考虑复数幂级数和复数微分。全面详细地介绍这一材料超出了本书的范围，但其中的简单方面几乎完全类似于我们已介绍的定义和定理。

复杂数学

设 a 和 b 为实数。α 的范数 $= a + bi \in \mathbf{C}$ 为

$$|\alpha| := \sqrt{\alpha\bar{\alpha}} = \sqrt{a^2 + b^2},$$

即从0到α的距离。$\mathbf{C}$上的范数函数具有与$\mathbf{R}$上的绝对值函数非常相似的性质；除了复范数扩展了实绝对值这一明显事实之外，范数是可乘的，并满足三角不等式：

定理15.1. 设α和β为复数。那么$|\alpha\beta| = |\alpha|\,|\beta|$，并且

$$\Big||\alpha| - |\beta|\Big| \le |\alpha + \beta| \le |\alpha| + |\beta|. \tag{15.1}$$

这些不等式是反三角形不等式和三角形不等式。

证明。写出 $\alpha = a + bi$ 和 $\beta = x + yi$，其中 a、b、x 和 y 是实数。根据定义，

$$\alpha\beta = (a + bi)(x + yi) = (ax - by) + (ay + bx)i,$$

因此，直接计算给出

$$\begin{aligned} |\alpha\beta|^2 &= (ax - by)^2 + (ay + bx)^2 \\ &= (ax)^2 - 2(axby) + (by)^2 + (ay)^2 + 2(aybx) + (bx)^2 \\ &= (a^2 + b^2)(x^2 + y^2) \quad = \quad |\alpha|^2\,|\beta|^2. \end{aligned}$$

三角形不等式也可以通过暴力计算来建立，但使用复杂数学形式更令人愉快。对于所有 $\alpha \in \mathbf{C}$，$|\alpha + \bar{\alpha}| = |2\,\mathrm{Re}\,\alpha| \le 2|\alpha|$。特别是，

$$|\alpha\bar{\beta} + \beta\bar{\alpha}| = |\alpha\bar{\beta} + \overline{\alpha\bar{\beta}}| \le 2|\alpha\bar{\beta}| = 2|\alpha|\,|\bar{\beta}| = 2|\alpha|\,|\beta|$$

对于所有 α，$\beta \in \mathbf{C}$。因此

$$\begin{aligned}|\alpha+\beta|^2 = (\alpha+\beta)\overline{(\alpha+\beta)} &= |\alpha|^2 + \alpha\bar{\beta} + \beta\bar{\alpha} + |\beta|^2 \\ &\le |\alpha|^2 + 2|\alpha|\,|\beta| + |\beta|^2 = (|\alpha| + |\beta|)^2.\end{aligned}$$

平方根证明复数的三角不等式。通过与证明实数版本相同的方法，即定理2.25所用的技巧，建立了反向三角不等式。

□

武装这些基本工具，原则上我们可以回到第4、8和11章，检查序列、极限、连续性、可微性、幂级数和收敛半径的定义和性质可以以完全相同的方式应用于复变复值函数，前提是我们将绝对值解释为复数的范数。以 $a \in \mathbf{C}$ 为中心的幂级数的收敛域是复平面上的一个真正的圆盘，解释了“收敛半径”这个术语。

集成明显缺失在上面的项目列表中，这些项目可以立即推广。基本原因是集成严重依赖于 $\mathbf{R}$ 的顺序属性；区间的划分使用顺序定义，而对于复值函数，上和下和没有意义，因为上确界和下确界是需要顺序的概念。存在一种不同的积分——“轮廓积分”，它在复分析中扮演着核心角色，但其定义和意义超出了本书的范围。

指数函数和三角函数

对于本节的剩余部分，我们限制关注具有无限收敛半径的复幂级数。与这种幂级数相关联的函数 $f : \mathbf{C} \to \mathbf{C}$ 被称为全纯函数；迄今为止我们所看到的例子是多项式、exp 和双曲正弦和双曲余弦函数、正弦和余弦函数，以及通过有限次加法、乘法和函数复合从这些函数构建的函数。基本初等函数的复幂级数与之前找到的实幂级数相同；唯一的区别是“变量”允许是复数（并且传统上用 z 而不是 x 表示）。

比较余弦和双曲余弦的幂级数:

$$\cos z = \sum_{k=0}^{\infty}(-1)^k \frac{z^{2k}}{(2k)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \cdots$$
$$\cosh z = \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \frac{z^6}{6!} + \cdots$$

两个系列都具有半径 $+\infty$，正如在实际情况中一样，因此每个都代表一个全函数。稍加思考即可发现，cos z = cosh(iz)，因为 $(iz)^{2k} = (-1)^k z^{2k}$。将变量乘以 i 的简单方法将双曲余弦转换为余弦！类似地，

$$\sin z = \sum_{k=0}^{\infty}(-1)^k \frac{z^{2k+1}}{(2k+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots$$
$$\sinh z = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \frac{z^7}{7!} + \cdots$$

它们之间的关系是 i sin z = sinh(iz)，通过类似的推理得出。

回忆圆和双曲三角函数被表征为某些初值问题的解。表征这些函数的微分方程相似性的原因不应难以理解。假设我们考虑某个 $k \in \mathbf{C}$ 的函数 $f(z)$ = sinh(kz)，求导得到 $f'(z) = k$ cosh(kz)，因此 $f'(0) = k$，然后 $f''(z) = k^2$ sinh(kz)；总之，

$$f'' - k^2 f = 0, \qquad f(0) = 0, \quad f'(0) = k.$$

现在，假设我们取 $k = i$；前面的方程变为 $f''+f = 0$，$f(0) = 0$，$f'(0) = i$。但是，将正弦和余弦作为方程 $f'' + f = 0$ 的解的特征（第13.3推论，其证明可以立即推广到复数域）表明 $f(z) = i$ sin z，并且我们再次发现 i sin z = sinh(iz)。对于余弦和双曲余弦也有类似的论述。

如果我们把前一段落的结论综合起来，我们得到一个通常被称为（当 z 为实数时）棣莫弗定理的恒等式:

定理15.2. e^{iz} = cos z + i sin z 对所有 $z \in \mathbf{C}$ 成立。

特殊情况 $z = \pi$ (通常写作 $e^{i\pi} + 1 = 0$) 被称为欧拉公式。说这个方程具有神秘意义太过分了，

但是它显著地与“数学中最重要的五个常数”相关。

定理15.2的证明是一行：

$$e^{iz} = \cosh(iz) + \sinh(iz) = \cos z + i \sin z.$$

通过直接检查各自的幂级数得出相同的结论。从实变量的角度来看，这个公式可以说是令人难以置信的。几何增长与测量圆有什么关系？答案隐藏在复数乘法的本质中，这一点将逐渐变得明显。举一个例子，观察一下，如果 x 和 y 是实数，那么

$$\begin{aligned} e^{ix}e^{iy} &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\sin x \cos y + \sin y \cos x) \\ &= \cos(x+y) + i \sin(x+y) \qquad \text{by equation (13.9)} \\ &= e^{i(x+y)}. \end{aligned}$$

正弦和余弦的加法公式不过是指数加法公式的扩展，扩展到复数上！一个与之密切相关且值得强调的公式。

推论15.3. 如果 x 和 y 是实数，那么 $e^{x+iy} = e^x \cos y + ie^x \sin y$。

复数乘法的几何

单位复数集 $S^1 := \{z \in \mathbf{C} \mid z\bar{z} = 1\}$ 在几何上是半径为 1，以 0 为中心的圆。

命题15.4.每个单位复数都是形如 e^{it} 的，其中 $t \in (-\pi, \pi]$ 是唯一的。

证明。由 $\gamma(t) = e^{it}$ 定义的函数 $\gamma : \mathbf{R} \to \mathbf{C}$ 是 2π-周期的，并且没有更小的正周期。由于 $|e^{it}| = \sqrt{\cos^2 t + \sin^2 t} = 1$，$\gamma$ 的限制是一个单射映射 $(-\pi, \pi] \to S^1$。

要证明 $\gamma|_{(-\pi,\pi]}$ 是满射的，回忆一下 cos 映射 $[0, \pi]$ 双射到 $[-1, 1]$。如果 $\alpha = a + bi \in S^1$ 且 $b \geq 0$，那么存在唯一的 $t \in [0, \pi]$，使得 $a = \cos t$。对于这个 t，

$$b = \sqrt{1 - a^2} = \sqrt{1 - \cos^2 t} = \sin t,$$

所以 $\alpha = e^{it}$。如果相反 $b < 0$，那么 $\bar{\alpha} = e^{it}$ 对于一个唯一的 $t \in (0, \pi)$，所以

$$\alpha = \overline{e^{it}} = \cos t - i \sin t = \cos(-t) + i \sin(-t) = e^{-it}.$$

这是，对于唯一的0) $\alpha = e^{it}$。注意，由于周期性，γ将长度为2π的每个半开区间双射地映射到圆上。 □

每个非零复数 z 都是正实数与单位模复数的乘积:
$z = |z| \cdot (z/|z|)$。根据命题15.4，我们可以写成

$$z = e^{\rho} \cdot e^{it} = e^{\rho+it} \qquad \text{for } unique\ \rho \in \mathbf{R} \text{ and } t \in (-\pi, \pi]. \tag{15.2}$$

这个表示称为 z 的极坐标形式；数字 t 被称为 z 的幅角，表示为 arg z，并解释为正实轴与从 0 到 z 的射线之间的角度。

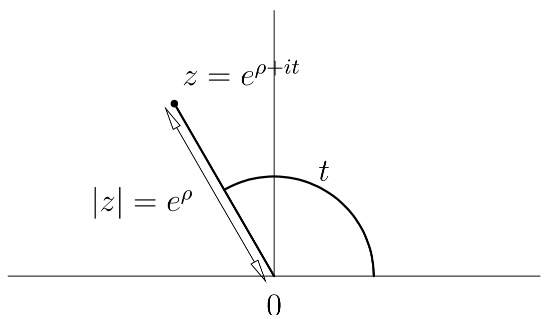


图15.1：非零复数的极坐标形式。

方程（15.2）立即给出了图2.12（第81页）中复数乘法的几何描述：如果 $z_j = e^{\rho_j+it_j}$ 对于 $j = 1, 2$，那么

$$z_1 z_2 = (e^{\rho_1} \cdot e^{it_1})(e^{\rho_2} \cdot e^{it_2}) = e^{\rho_1} e^{\rho_2} \cdot e^{i(t_1+t_2)}.$$

在文字上，我们通过乘以两个复数的范数并加上它们的辐角来相乘两个复数。如果 $\alpha = e^{\rho+it}$，那么映射 $z \mapsto \alpha z$ 通过 t 弧度旋转平面并按因子 e^{ρ} 缩放。特别是，乘以 $i = e^{i\pi/2}$ 会使平面逆时针旋转四分之一圈。

德莫弗公式在复数乘法方面具有美丽的几何解释。回忆一下

$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n \qquad \text{for } x \in \mathbf{R}.$$

假设我们尝试将此应用于 $x = i\theta$ 纯虚数。复数 $1 + (i\theta/n)$ 定义了一个直角三角形，如图 15.2 左侧所示，其原点处的角度几乎等于 θ/n。（情况 $\theta = \pi$ 已展示。）将此数提高到 n 次幂在几何上相当于迭代地缩放和旋转此三角形，如图 15.2 右半部分所示。最后一个三角形的对角顶点几乎位于与正水平轴成 θ 角的射线上。作为 $n \to \infty$，从几何上明显可以看出，序列的终点趋近于 $\cos\theta + i\sin\theta$，因此它等于 $e^{i\theta}$
。

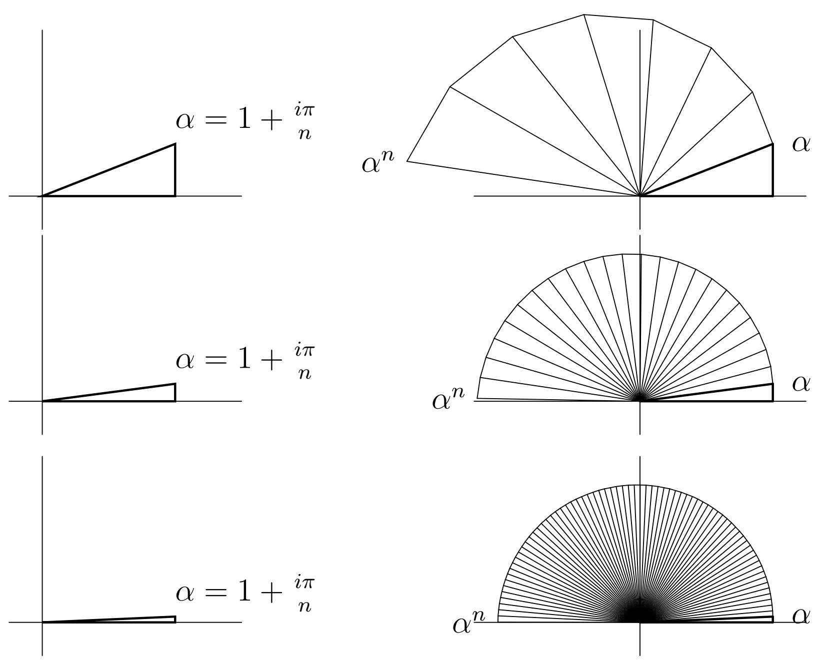


图15.2：de Moivre公式的几何解释。

为了使这个几何论证精确，写出 $\alpha = 1 + (i\theta/n) = |\alpha| \cdot e^{it}$ 并注意

$$|\alpha| = \sqrt{1 + (\theta/n)^2} = 1 + O(1/n^2), \quad t = \mathrm{Tan}^{-1}(\theta/n) = (\theta/n) + O(1/n^3).$$

估计值直接来自适当的泰勒多项式。取n次幂，我们发现

$$|\alpha^n| = \big(1 + O(1/n^2)\big)^n, \qquad \arg(\alpha^n) = nt = \theta + O(1/n^2).$$

作为 $n \to +\infty$，我们有 $|\alpha^n| \to 1$ 和 $\arg(\alpha^n) \to \theta$，所以 $\alpha^n \to e^{i\theta}$。

复对数

回忆一下，$\mathbf{C}^\times$ 表示非零复数集合，它是在复数乘法下的阿贝尔群。$\mathbf{C}^\times$ 中的每个 w 都可以唯一地表示为 $e^z = e^{\rho+it}$，其中 ρ 是实数，$t \in (-\pi, \pi]$。对数的原理分支是定义为 Log : $\mathbf{C}^\times \to \mathbf{C}$ 的映射。

$$\operatorname{Log} w = z = \rho + it, \quad -\pi < t \le \pi. \tag{15.3}$$

对数图像是一个高度为2π的水平条带，基本条带，见图15.3。$\mathbf{C}^\times$中负实轴上的点对应于图中的$t = \pi$点。

如果 $w \in \mathbf{C}^\times$，则 exp(Log $w) = w$)。相比之下，如果 $z \in \mathbf{C}$，则通常 Log(exp z) $= z$ 是错误的；对于所有 z，左侧在 $(-\pi, \pi]$ 有虚部，因此如果 z 不在基本条带中，则两边不能相等。相反，对于每个 z，存在一个唯一的整数 k，使得 $z - 2\pi ik$ 在 $(-\pi, \pi]$ 有虚部，并且 Log(exp z) $= z - 2\pi ik$ 对于这个 k。

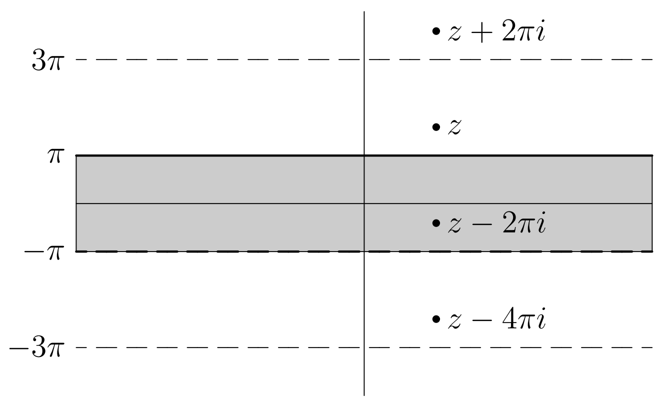


图15.3：复对数作为映射。

复指数函数是周期性的：对于所有 z 在 $\mathbf{C}$ 中，$e^{z+2\pi i} = e^z$。因此，exp 不能有逆函数；我们所能期望的最好的是限制 exp 在一个它是双射的集合上。基本带就是这样的一组。如图15.3所示，存在无限多个“周期带”，每个对应于对数的一个分支。所有显示的点在 exp 下映射到相同的 w，而 Log 的一个分支选择这样一个点作为 w 的函数。

存在一个关于exp可逆性的更微妙的问题。考虑由$z_n = (-\pi + \frac{1}{n})i$定义的序列$(z_n)$。每个项都位于

基本条带，但极限，$-\pi i$，则不然。如果我们对这个序列应用指数函数然后取对数，我们发现

$$\text{对数}(\exp z_n) = z_n, \text{但 } \mathrm{Log}\big(\exp(-\pi i)\big) = \pi i;$$

对数的原理分支是不连续的。从几何上容易看出，问题恰好发生在w平面上的负实轴上，这是$\mathrm{Im}\, z = \pi$线的像，是基本条带的上边缘。当点w“连续”地穿过负实轴时，点$z = \mathrm{Log}\, w$在基本条带的上边缘和下边缘之间“不连续”地移动。

这个讨论有点非正式，因为我们没有精确地定义复变函数的“连续性”概念。然而，应该可以推断在$\mathbf{C}^\times$上不存在对数连续分支。这个基本事实是数学和物理学中各种复杂而微妙现象的起源。练习15.1给出了一个（如果从数学角度看）恰当的例子。

复杂数多项式

第五章中我们看到了每个正实数都有一个唯一的 n 次方根，对于每个正整数 n，并且每个奇数次多项式都有一个实根。在本节中，我们将看到在复数域上类似的问题是如何变得如此简单美丽。

复数根

设 w 为复数，设 n 为一个正整数。w 的 n 次根是一个满足 $z^n = w$ 的复数 z。多项式方程 $z^n - w = 0$ 至多有 n 个解 z，因此如果我们能找到 n 个不同的 n 次根，我们就有一个完整的列表。

如果使用 z 和 w 的实部和虚部来写出 $z^n - w = 0$，结果过于复杂，无法得出任何明显的见解。然而，直接写出 $z = |z|e^{it}$ 和 $w = |w|e^{i\theta}$ 当 $w \neq 0$ 时，立即得到 n 个不同的根。两个极坐标形式的数相等当且仅当它们具有相同的范数，并且它们的辐角相差 $2\pi i$ 的整数倍。因此，$z^n = w$ 当且仅当

$$|z| = |w|^{1/n}, \qquad t = \tfrac{1}{n}(\theta + 2\pi ik), \quad k = 0,\ \dots,\ n-1.$$

第一个是一个正实数的方程，我们知道它有一个唯一解。第二个是列出 $nt - \theta =$^$2\pi ik$ 的解的显式方法，使得 e^{it} 这组数是不同的。

即使情况 $w = 1$ 也很有趣。数字 $e^{2\pi ik/n}$，$0 \le k < n$ 被称为单位根的 n 次根，因为 $(e^{2\pi ik/n})^n = e^{2\pi ik} = 1$。写成 $\zeta_n := e^{2\pi i/n}$ 是很常见的，这样单位根就是 ζ_n 的幂。每个单位根都是一个单位复数（即，范数为 1），相邻根之间的角度是 $1/n$ 圈。因此，这些点位于单位圆内正 n 边形的顶点上：

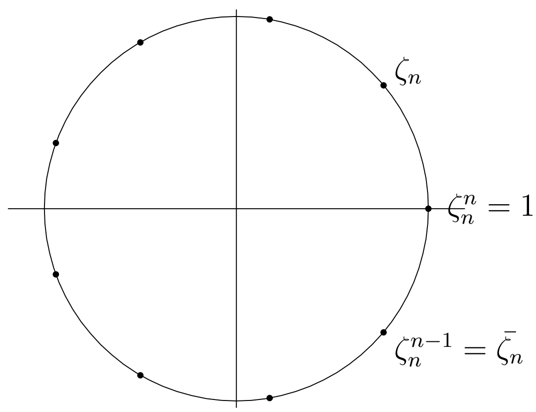


图15.4：n次单位根。

对于一个一般的非零 w，如果 z 是一个特定的 n 次根，那么其他根是 $z\zeta_n^k$，对于 $0 \le k < n$；特别是，n 次根 w 也位于一个半径为 $|w|^{1/n}$ 的圆内正 n 边形的顶点上。0 的 n 个根都是 0，但在某种意义上，仍然有 n 个!

代数基本定理

多项式 $p(z) = z^n - w$ 在 $\mathbf{C}$ 上完全分解：如果 z_1，...，z_n 是 w 的 n 次根，那么

$$z^n - w = \prod_{k=1}^{n} (z - z_k).$$

这是代数基本定理的一个特例：

定理15.5. 设$p : \mathbf{C} \to \mathbf{C}$是一个具有复系数的非常数多项式函数。那么存在一个$z_0 \in \mathbf{C}$，使得$p(z_0) = 0$。

非正式地说，每个复数多项式都有一个根。回忆一下，根据推论3.7，多项式的根对应于一个线性因子：$p(z_0) = 0$ iff $(z - z_0)$ divides p。定理15.5的反复应用意味着每个非常数多项式都可以完全分解为线性因子。

很遗憾，一个完整的证明需要我们尚未完全发展的技术，但这里展示的草图应该能给出涉及内容的概览。

证明。假设像往常一样，最高系数 a_n 是非零的，写出

$$p(z) = \sum_{k=0}^{n} a_k z^k = a_n z^n \big(1 + O(1/z)\big).$$

取绝对值，$|p(z)| = |a_n| \cdot |z|^n \big(1 + O(1/|z|)\big)$ 近 $|z| = +\infty$：一个多项式在绝对值上趋近于 $+\infty$，当 $|z| \to +\infty$。

考虑 $a_0 = p(0)$；根据前一段，存在一个实数 $R > 0$，使得 $|p(z)| > |a_0|$ 对于 $|z| > R$ 成立。由于 $|p| : \mathbf{C} \to \mathbf{R}$ 是连续的，根据极值定理的推广，在某个 $z_0 \in D_R$ 处，限制在圆盘 $D_R := \{z \in \mathbf{C} : |z| \le R\}$ 上的绝对最小值存在.[1] 点 z_0 是 $|p|$ 在整个 $\mathbf{C}$ 上的绝对最小值，因为通过选择 R，我们有 $|p(z)| > |p(0)| \ge |p(z_0)|$ 对于 $|z| > R$。如果我们能证明 $|p(z_0)| = 0$，我们就完成了。

展开 p 为 $(z - z_0)$ 的幂次 $= re^{it}$。设 b_ℓ 为最低次非常数项的系数，则有

$$p(z) = \sum_{k=0}^{n} b_k (z - z_0)^k = b_0 + b_\ell (z - z_0)^\ell \big(1 + O(z - z_0)\big).$$

现在，如果 $b_0 \neq 0$，那么

$$|p(z_0 + re^{it})| = |b_0| \cdot \left| 1 + \frac{b_\ell}{b_0} r^\ell e^{i\ell t} \big(1 + O(r)\big) \right|.$$

选择 $r \ll 1$ 使得 $r^\ell \big(1 + O(r)\big) > 0$，然后选择 t 使得 $(b_\ell / b_0) e^{i\ell t}$ 是实数且为负。将 b_ℓ / b_0 写成极坐标形式表明这是可能的。前一个方程意味着 $|p(z_0 + re^{it})| < |b_0| = |p(z_0)|$，这与 z_0 是 $|p|$ 的绝对最小值的事实相矛盾。必须满足 $p(z_0) = 0$。 $\square$

[1] This is the only step we fail to justify substantially. The crucial properties of the disk are that is is a bounded subset of the plane that contains all its limit points.

代数基本定理对实系数多项式有有趣的推论。

推论15.6.设$p:\mathbf{C}\to\mathbf{C}$是一个具有实系数的多项式。那么$z_0\in\mathbf{C}$是根当且仅当$\bar{z}_0$是根。此外，$p$在实数域中分解为一次和不可约二次项。

证明。回忆一下，一个复数 α 是实数当且仅当 $\bar{\alpha}=\alpha$，并且共轭在 (2.16) 的意义上与加法和乘法交换。设 $p(z)=\sum_k a_k z^k$ 中 a_k 对所有 k 都是实数。那么

$$p(\bar{z})=\sum_k a_k \bar{z}^k=\sum_k a_k \overline{z^k}=\sum_k \overline{a_k z^k}=\overline{p(z)}.$$

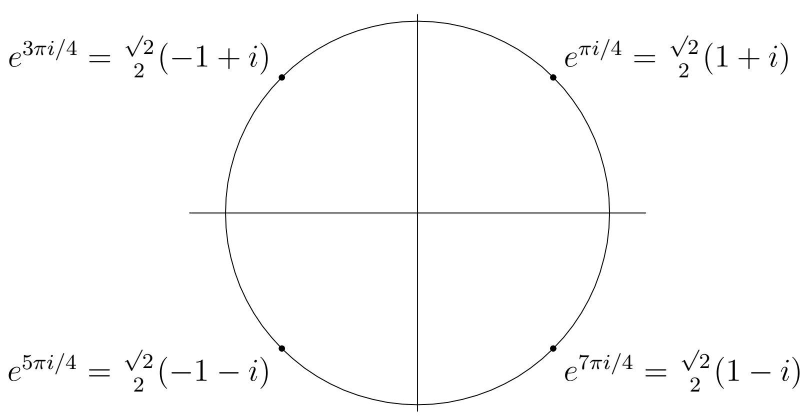
特别地，$p(z_0)=0$，则 $p(\bar{z}_0)=\overline{p(z_0)}=0$。由于两次共轭是恒等映射，第一个命题得证。

对于第二个，使用基本定理将 p 分解为线性项（具有复系数），然后对对应于共轭根的项进行分组。对于所有 $\alpha\in\mathbf{C}$,

$$(z-\alpha)(z-\bar{\alpha})=z^2-2\operatorname{Re}\alpha+|\alpha|^2$$

具有实系数，因此 p 可以写成具有实系数的不可约二次因子的乘积。 □

示例15.7 让 $p(z)=z^4+1$. p 的根是 -1 的四次根，它们是1的8次根，不是平方根或1的四次根：



乘以共轭对中的对应项给出

$$z^4+1=(z^2+\sqrt{2}z+1)(z^2-\sqrt{2}z+1).$$

回顾起来，我们可能已经写下了 $z^4 + 1 = (z^4 + 2z^2 + 1) - 2z^2$ 并使用平方差公式得到相同的结果。注意，这个多项式没有实根，但仍然在实数范围内有因式分解。□

15.2 基本反微分

微积分的传统主题之一是精确地用符号形式计算定积分。本节介绍了计算基本不定积分的几种技术，并陈述了一些（未证明）的准则，在这些准则下，基本函数没有基本不定积分。

符号计算不定积分与符号计算导数在本质上不同。乘积、商、链式和隐函数求导法则意味着基本函数的导数也是基本的，并共同提供了一种有效的符号计算工具，用于求导。相比之下，不存在符号反求导的算法；不存在类似于乘积、商和链式法则的反导数的一般公式。当然，如果 f 是连续的，那么

$$F(x) = \int_a^x f(t)\,dt$$

是 f 的一个不定积分，因此每个初等函数在某种意义上都是平凡可微的。然而，生活中有一个事实是，一个初等函数通常没有初等的不定积分，例如参见下面的定理15.10。

随着计算机代数系统变得普遍且价格低廉，手动使用积分技巧的能力可能会逐渐消失，就像现代学生不需要提取平方根或进行长除法的能力一样，这要归功于电子计算器。然而，本书的哲学是提供对微积分内部运作的观察，因此至少你应该了解计算机积分软件是如何工作的。一些公式可以用来评估一些好奇的不定积分或无穷级数，这是一个额外的益处。

符号

我们将写

$$F(x) = \int^x f \quad \text{or} \quad F(x) = \int^x f(t)\,dt$$

表示 F 是 f 的一般原函数。这种表示法与“定积分”的表示法很好地结合在一起；积分常数被吸收在缺失的下限中，一个方程中所有 x 的出现都表示相同的意思。通常，虚变量 t 可以被任何未出现的符号替换。必须谨慎行事；表达式 $\int^x f$ 并不表示一个函数，而是一个无限函数族，其中任意两个函数之间的差异是一个加性常数。在 a 处消失的特定原函数是

$$F(x) = \int_a^x f(t)\,dt.$$

这个观察在实际中非常有用。

替换

代换法源于链式法则。在定理10.5的指导下，我们试图找到一个“变量替换”，使得被积函数简化到可以通过观察进行反导的程度。使用我们对抗导的符号，定理10.5的结论可以写成

$$\int^x f\big(U(t)\big)U'(t)\,dt = \int^{U(x)} f(u)\,du, \tag{15.4}$$

在其中有使用替换 $u = U(t)$。没有虚变量，前面的方程变为

$$\int^x f(U)\,U' = \int^{U(x)} f.$$

在莱布尼茨记号中，由于习惯于写作 $u = u(t)$，虚变量转换得令人满意。正式来说，我们有 $du = \frac{du}{dt}\,dt = u'(t)\,dt$，方程（15.4）是通过（或多或少字面意义上的）符号替换得到的。

以下是一些示例；在整个过程中，n 是一个非负整数，$k \neq 0$，$\alpha \neq -1$ 是实数。假设积分已经取过

在积分函数定义的区间上。为了简洁，我们给出替换及其微分，然后是它适用的公式。每个“答案”都包含一个显式的加性常数。

- $u = kt$，$du = k\,dt$。

$$\int^x e^{kt}\,dt = \frac{1}{k}\int^{kx} e^u\,du = \frac{1}{k}e^{kx} + C$$

  相同的技巧处理 $\int^x f(kt)\,dt$，如果 f 是反可微的。

- $u = 1 + e^t$，$du = e^t\,dt$。

$$\int^x e^t\sqrt{1+e^t}\,dt = \int^{1+e^x} u^{1/2}\,du = \frac{u^{3/2}}{3/2}\bigg|^{1+e^x} = \frac{2}{3}(1+e^x)^{3/2} + C$$

- $u = k + t2$, $du = 2t\,dt$。

$$\int^x (k+t^2)^\alpha t\,dt = \int^{k+x^2} u^\alpha\,\frac{du}{2} = \frac{1}{2}\frac{u^{\alpha+1}}{\alpha+1}\bigg|^{k+x^2} = \frac{(k+x^2)^{\alpha+1}}{2(\alpha+1)} + C$$

- $u =$ 余弦 t，$du = -$ 正弦 $t\,dt$。

$$\int^x \cos^n t\,\sin t\,dt = -\int^{\cos x} u^n\,du = -\frac{u^{n+1}}{n+1}\bigg|^{\cos x} = -\frac{\cos^{n+1}x}{n+1} + C$$

- $u =$ 余弦 t，$du = -$ 正弦 $t\,dt$。

$$\int^x \frac{\sin t}{\cos t}\,dt = -\int^{\cos x}\frac{du}{u} = -\log|u|\bigg|^{\cos x} = -\log|\cos x| + C$$

  or $\int^x \tan t\,dt = \log|\sec x| + C$.

- 复指数与三角函数之间的关系可用于评估一些否则需要分部积分的积分。设 a 和 b 为实数，考虑

$$\int^x e^{at}\cos bt\,dt + i\int^x e^{at}\sin bt\,dt = \int^x e^{(a+bi)t}\,dt.$$

反导数公式对exp成立，即使对于复指数也成立，因此右边的积分是

$$\frac{1}{a+bi}e^{(a+bi)x} = \frac{a-bi}{a^2+b^2}\, e^{ax}(\cos bx + i\sin bx).$$

将实部和虚部分别相等，我们得到两个有用的公式

$$\int^x e^{at}\cos bt\, dt = \frac{e^{ax}}{a^2+b^2}(a\cos bx + b\sin bx) + C$$

$$\int^x e^{at}\sin bt\, dt = \frac{e^{ax}}{a^2+b^2}(a\sin bx - b\cos bx) + C$$

这些函数出现在阻尼谐振子研究之中。

代换法有点像一门艺术，并且只在有限的被积函数上有效；看似相似的积分之间有很大的区别

$$(*) \qquad \int^x e^{-t^2}\, dt \qquad \text{and} \qquad \int^x te^{-t^2}\, dt.$$

一般做法是让 u 成为根号或超越函数内的任意值，尤其是如果 du 在被积函数的其余部分中明显存在。有时一个“显然”的代换会导致需要进一步代换的积分；在这种情况下，通过单个（明智的）代换总能完成相同的变换，即单独代换的合成。

对于$(*)$中的两个积分，选择$u=-t^2$是自然的，但只有在第二种情况下替换才立即有用。结果，第一个积分没有初等反导数；非正式地说，“e^{-t^2}不能以封闭形式积分。”

三角积分

每个三角函数都有一个基本的原函数。显然

$$\int^x \sin\theta\, d\theta = -\cos x + C, \qquad \int^x \cos\theta\, d\theta = \sin x + C,$$

当我们在上面找到了tan的积分；cot与此完全类似。相比之下，正割函数更难。现在，我们只陈述公式，可以通过手工检查：{v*}

$$\int^x \sec\theta\,d\theta = \log\left|\sec x + \tan x\right| + C.$$

导数公式对于正切和余割给出了有用的积分：

$$\int^x \sec^2\theta\,d\theta = \tan x + C, \qquad \int^x \sec\theta\tan\theta\,d\theta = \sec x + C.$$

最后，双角公式

$$\cos^2\theta = \tfrac{1}{2}(1+\cos 2\theta), \qquad \sin^2\theta = \tfrac{1}{2}(1-\cos 2\theta)$$

可以反微分，得到

$$\int^x \cos^2\theta\,d\theta = \tfrac{1}{4}(2x+\sin 2\theta)+C, \qquad \int^x \sin^2\theta\,d\theta = \tfrac{1}{4}(2x-\sin 2\theta)+C.$$

积分法

积分法是微分乘积法则的类比。在莱布尼茨记号中，如果 u 和 v 是可微函数，那么

$$d(uv) = u\,dv + v\,du, \qquad \text{or} \quad u\,dv = d(uv) - v\,du.$$

如果导数是连续的，即如果 u 和 v 是 $[a,b]$ 上的 $\mathscr{C}^1$ 函数，则前一个方程可以积分，得到分部积分公式：

$$\int_a^b uv' = (uv)\Big|_a^b - \int_a^b vu'. \tag{15.5}$$

为了应用分部积分，所涉及的被积函数必须写成两个函数 u 和 v' 的乘积，使得 u' 容易计算，v 容易找到，并且 vu' 比较容易反微分于 uv'。这些是严格的要求，但对于某些函数类，分部积分效果很好。在选择 u 和 v' 时还有额外的技巧。正如上面所述，例子是最好的说明者。

- $u = t$ 并且 $dv = \sin t\,dt$，所以 $du = dt$ 和 $v = -\cos t$。
$$\int^x t\sin t\,dt = -t\cos t\Big|^x + \int^x \cos t\,dt = \sin x - x\cos x + C.$$
类似的选择处理积分函数 $t\cos t\,dt$ 和 $te^t\,dt$。如果涉及到 t 的高次幂，分部积分得到递归公式。将 u 视为 t 的幂，将 dv 视为三角函数，
$$\begin{aligned}\int^x t^n\sin t\,dt &= -t^n\cos t\Big|^x + n\int^x t^{n-1}\cos t\,dt\\ &= \big(-t^n\cos t + nt^{n-1}\sin t\big)\Big|^x - n(n-1)\int^x t^{n-2}\sin t\,dt\\ &= \big(-x^n\cos x + nx^{n-1}\sin x\big) - n(n-1)\int^x t^{n-2}\sin t\,dt\end{aligned}$$
此公式可以递归应用，无需进一步工作。
- $u =$ 对数 t，$dv = t^n\,dt$。
$$\begin{aligned}\int^x t^n\log t\,dt &= \frac{1}{n+1}t^{n+1}\log t\Big|^x - \frac{1}{n+1}\int^x t^n\,dt\\ &= \frac{x^{n+1}}{(n+1)^2}\big((n+1)\log x - 1\big)\end{aligned}$$

积分部分积分法可以以一种非平凡的方式将一个积分用自身表示。接下来的两个例子和练习15.4是典型的。

- $u = \sin^{n-1} t$, $dv = \sin t\,dt$,
$$\begin{aligned}\int^x \sin^n t\,dt &= -\cos x\,\sin^{n-1}x + (n-1)\int^x \sin^{n-2}t\,\cos^2 t\,dt\\ &= -\cos x\,\sin^{n-1}x + (n-1)\int^x \sin^{n-2}t\,dt\\ &\qquad - (n-1)\int^x \sin^n t\,dt.\end{aligned}$$
未知积分出现在两边，我们可以通过代数方法将其隔离：
$$\int^x \sin^n t\,dt = -\frac{1}{n}\cos x\,\sin^{n-1}x + \frac{n-1}{n}\int^x \sin^{n-2}t\,dt. \tag{15.6}$$
此公式将 $\sin^n t$ 的积分降低到 $\sin^{n-2} t$ 的积分。

- $u = \tan$ {v*}, $v' = \sec t \tan t\,dt$。

$$\begin{aligned}\int^x \sec t \tan^2 t\,dt &= \sec t \tan t\Big|^x - \int^x \sec^3 t\,dt \\ &= \sec t \tan t\Big|^x - \int^x \sec t(1+\tan^2 t)\,dt \\ &= \sec x \tan x - \log\big|\sec x + \tan x\big| - \int^x \sec t \tan^2 t\,dt.\end{aligned}$$

求解未知积分,

$$\int^x \sec t \tan^2 t\,dt = \frac{1}{2}\Big(\sec x \tan x - \log\big|\sec x + \tan x\big|\Big) + C.$$

三角代换

身份 $\cos^2\theta + \sin^2\theta = 1$ 及其变体 $1+\tan^2\theta = \sec^2\theta$ 可用于反微分包含平方根的许多函数。设 $a>0$。

Integrand Contains	Substitution	Identity Utilized
$\sqrt{a^2-x^2}$	$x = a\sin\theta$	$1-\sin^2\theta = \cos^2\theta$
$\sqrt{a^2+x^2}$	$x = a\tan\theta$	$1+\tan^2\theta = \sec^2\theta$
$\sqrt{x^2-a^2}$	$x = a\sec\theta$	$\sec^2\theta - 1 = \tan^2\theta$

同样可以使用恒等式$\cosh^2 t - \sinh^2 t = 1$ 和 $1-\tanh^2 t = \mathrm{sech}^2\, t$ 来表示这些被积函数,如果存在某种原因需要用指数形式而不是三角函数形式来表示结果。

第13章的阅读将揭示以下公式本质上是对定义:

$$\int^x \frac{dt}{\sqrt{a^2-t^2}} = \mathrm{Sin}^{-1}\frac{x}{a} + C \qquad \int^x \frac{dt}{a^2+t^2} = \frac{1}{a}\mathrm{Tan}^{-1}\frac{x}{a} + C$$

$$\int^x \frac{dt}{t\sqrt{t^2-a^2}} = \frac{1}{a}\mathrm{Sec}^{-1}\frac{x}{a} + C$$

其他积分被转换成我们已经找到的形式。

- $t = a$ 正弦 θ, $dt = a$ 余弦 $\theta\, d\theta$。
$$\int^x (\sqrt{a^2 - t^2})^n\, dt = a^{n-1} \int^{\mathrm{Sin}^{-1}(x/a)} \cos^{n+1}\theta\, d\theta.$$
这是通过一个缩减公式来处理的。

- $t = a\tan\theta$, $dt = a\sec^2\theta\, d\theta$。
$$\begin{aligned}\int^x \sqrt{a^2 + t^2}\, dt &= a^2 \int^{\mathrm{Sec}^{-1}(x/a)} \sec^3\theta\, d\theta \\ &= a^2 \int^{\mathrm{Sec}^{-1}(x/a)} \sec\theta(1 + \tan^2\theta)\, d\theta \\ &= \frac{a^2}{2}\Big(\sec\theta\tan\theta + \log\big|\sec\theta + \tan\theta\big|\Big)\Big|^{\mathrm{Sec}^{-1}(x/a)} \\ &= \frac{a}{2}\Big(x\sqrt{x^2 - a^2} + a\log\big|x + \sqrt{x^2 - a^2}\big|\Big) + C.\end{aligned}$$
细节留给你，练习15.5。

在将三角代换后的变量转换回原变量时，练习13.9中的“直角三角形技巧”很有帮助。

部分分式

每个具有实系数的有理函数都可以写成一定标准形式的项之和。这种标准形式，称为部分分式分解，将分阶段介绍。具体来说，我们希望系统地整理如下等式

$$\frac{1}{1-x^2} = \frac{1}{2}\left(\frac{1}{1-x} + \frac{1}{1+x}\right), \qquad \frac{x}{1-x^2} = \frac{1}{2}\left(\frac{1}{1-x} - \frac{1}{1+x}\right),$$
或者
$$\frac{2 + 2x + 4x^2 + x^3 + x^4}{(1+x^2)^2(1+x)} = \frac{x}{(1+x^2)^2} + \frac{1}{1+x^2} + \frac{1}{1+x}.$$

每个部分分式分解的项要么是一个多项式，要么是一个有理函数，其分母是线性或不可约二次多项式的幂，且其分子是分母中因子次数低于的多项式。符号上，部分分式项的形式为

$$\frac{b}{(x-c)^k}, \quad \text{or} \quad \frac{ax+b}{(x^2+cx+d)^k}, \quad a,\ b,\ c,\ d \text{ real},\ c^2 - 4d < 0.$$

推论15.6在证明部分分式分解的存在性中起着关键作用：每个具有实系数的多项式都可以分解为具有实系数的不可约线性多项式和二次多项式的幂的乘积。通过将常数吸收到分子中，我们可以假设分母的不可约因子是首一多项式。

想象一下尝试将特定有理函数 $f(x) = x/(1-x^2)$ 写成分式形式的难题。我们将逐步讲解这个过程，然后从一般的角度重新审视构建方法。第一步是分解分母：

$$f(x) = \frac{x}{(1-x)(1+x)}.$$

下一个选择分母的一个不可约因子，例如 $1-x$，并寻找一个常数 b 使得

$$\left(\frac{x}{(1-x)(1+x)} - \frac{b}{1-x}\right) = \frac{-b+(1-b)x}{(1-x)(1+x)}$$

具有有限的极限在 $x=1$。由于分母在 1 处消失，为了使商有极限，分子也必须在 $x=1$ 处消失。因此 $1-2b=0$，或 $b=1/2$。将这个值 b 代入右侧并简化后，我们发现我们的分数已经变得简单了：“新”的分母是原始分母中找到的不变量乘积，但我们选择的因子的指数，在这种情况下为 $1-x$，至少减少了一个。重复这个过程；最终分母变为 1，因为每一步都减少了分母的总次数。在本例中，我们只需一步就完成了。

应该合理地认为这样的论证在一般情况下是有效的，并且通过对分母次数进行归纳的正式证明是直接的。直观上，剥离一个部分分式项相当于减去与分母的一个因子相对应的“最高阶无穷”，留下一个更低阶的无穷。代替一个精确的定理（表述起来相当冗长），这里有一个代表性的例子：部分分式分解定理保证，对于每个次数最多为6的多项式p（小于下面分母的次数），存在常数a_i和b_i，使得

$$\begin{aligned}\frac{p(x)}{(x^2+x+2)^2(x-3)^3} &= \frac{a_1x+b_1}{(x^2+x+2)^2} + \frac{a_2x+b_2}{(x^2+x+2)} \\ &\quad + \frac{b_3}{(x-3)^3} + \frac{b_4}{(x-3)^2} + \frac{b_5}{(x-3)}.\end{aligned}$$

与初等积分的相关性如下：如果我们能证明每个部分分式项都有一个初等不定积分，那么我们就将证明每个有理函数都有一个初等不定积分。

一个形式为 $b/(x-c)^k$ 的项有不定积分

$$\frac{b}{(1-k)(x-c)^{k-1}} \quad \text{or} \quad b\log(x-c),$$

根据 $k \neq 1$ 或 $k = 1$。如果我们愿意与具有复系数的有理函数一起工作，我们就完成了，因为对于 $\mathbf{C}$ 中的每个有理函数都可以化简为这种项的和。然而，我们希望找到实函数的实反导数，因此我们被迫考虑形式为 $(ax+b)/(x^2+cx+d)^k$ 的项，其中 $c^2-4d<0$。

第一个简化是在分母中完成平方。设 $u = x + c/$^2，我们看到对于某些实数 α，有 $x^2+cx+d = u^2+\alpha^2$。因此，我们简化为反微分

$$\frac{au+b}{(u^2+\alpha^2)^k}, \qquad \text{or} \quad \frac{u}{(u^2+\alpha^2)^k}, \text{ and } \frac{1}{(u^2+\alpha^2)^k}.$$

请注意，常量 a 和 b 正在使用“通用”方式；它们不一定代表每一行相同的数字。

一个形式为 $u/(u^2+\alpha^2)^k$ 的项通过替换 $v = u^2+\alpha^2$，$dv = 2u\,du$ 来处理。形式为 $1/(u^2+\alpha^2)^k$ 的项使用三角替换 $u = \alpha\tan\theta$ 来处理，这导致

$$\int^x \frac{du}{(u^2+\alpha^2)^k} = \int^{\mathrm{Tan}^{-1}(u/\alpha)} \frac{\alpha\sec^2\theta\,d\theta}{\alpha^{2k}\sec^{2k}\theta} = \alpha^{1-2k}\int^{\mathrm{Tan}^{-1}(u/\alpha)} \cos^{2k-2}\theta\,d\theta.$$

这个积分通过反复分部积分来处理。最后我们追踪到了所有可能情况。总结如下：

定理15.8. 如果 $f:(a,b)\to \mathbf{R}$ 是一个有理函数，那么存在一个初等函数 $F:(a,b)\to\mathbf{R}$ 使得 $F' = f$.

注意，F本身可能不是有理数。使用进一步的三角代换来推导以下结论，这大致断言“每个正弦和余弦的有理函数都有一个基本的不定积分。”

推论15.9.如果R是两个变量的有理函数，那么

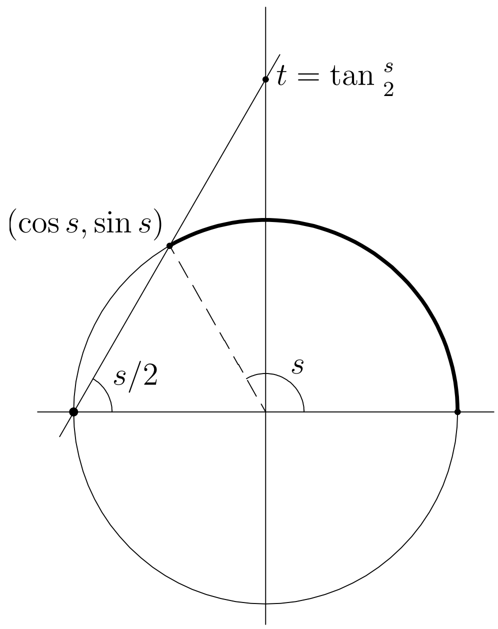
$$\int^x R(\cos s, \sin s)\,ds$$

是基本的。

证明。（草图）在替换 $t=\tan\frac{s}{2}$ 下，我们有

$$ds=\frac{2\,dt}{t^2+1},\qquad \cos s=\frac{t^2-1}{t^2+1},\qquad \sin s=\frac{2t}{t^2+1},$$

因此，被积函数 $R(\cos s,\sin s)\,ds$ 成为了 t 的有理函数。通过球面投影，人们得到了这个显著的替换：



细节留给你，练习15.6。 □

存在与不存在定理

如已反复强调，反微分是一种艺术，而不是一个定义明确的程序。即使是经验丰富的人，也很难一眼看出某个基本函数是否可以进行符号反微分。正如你可能已经注意到的，本节包含了一些专门的技术、一些小技巧和事后诸葛亮式的推理。例如，通过检查导数表并反向工作来评估了几个三角积分。其他积分通过试错法处理，使用各种替换和分部积分。这种情况反映了该学科的本质。计算机非常适合这种“搜索”计算，这也是它们作为积分工具越来越受欢迎的部分原因。

为了完成讨论，我们无需证明地陈述两个关于（非）存在基本不定积分的定理。证明依赖于“微分域”的概念，一个带有映射 $d:$ $\mathbf{F} \to \mathbf{F}$ 的域 $\mathbf{F}$，该映射满足形式乘法法则：$d(ab) = a\,db + b\,da$。一个例子是单变量有理函数域，其中 d 是普通导数。一般想法是研究 $df = g$ 在给定的 g 下是否有解，尽管细节有点复杂，并且需要比我们已发展的“抽象代数”更多的知识。我们选择了下面的版本，因为它们的陈述容易理解，并且它们导致了已知的例子。

定理15.10. 设f_1和f_2是单变量有理函数。如果函数

$$g(x) = f_1(x)e^{f_2(x)},$$

具有一个基本不定积分 G，则对于某个有理函数 F，有 $G(x) = F(x)e^{f_2(x)}$。

定理15.10表明以下不是初等的：

$$\int^x e^{-t^2}\,dt, \qquad \int^x \frac{\sin t}{t}\,dt, \qquad \int^x \frac{dt}{\log t}.$$

第一个积分与“误差函数”密切相关，该函数出现在概率中。我们将在下文更详细地研究这个积分。为了说明后两个不是初等函数，我们考虑函数 $g(t) = e^t/t$，对于它有 $f_2(t) = t$。根据定理15.10，

$$G(x) = \int^x \frac{e^t}{t}\,dt \quad \text{is elementary iff} \quad F(x) = \frac{1}{e^x}\int^x \frac{e^t}{t}\,dt \quad \text{is rational.}$$

然而，一系列计算表明 $F(x) = \log x + O(1)$ 在 $x = 0$ 附近，因此 F 不是有理数。由此可知

$$\int^x \frac{\sin t}{t}\,dt = \frac{1}{2i}\int^x \frac{e^{it} - e^{-it}}{t}\,dt$$

这不是基本的。替换 $t = \log u$ 表明 $\int^x du/\log u$ 也不是基本的。

我们的关于基本不定积分的最终结果归功于切比雪夫：

定理15.11. 设a，b，p，q和r为实数。反导数

$$\int^x t^p(a+bt^r)^q\,dt$$

是基本的，如果至少有一个 $\frac{p+1}{r}$， q 或 $\frac{p+1}{r}+q$ 是整数。

例如，

$$\int^x \sqrt{1+t^4}\,dt \quad \text{and} \quad \int^x t^2\sqrt{1+t^4}\,dt$$

它们不是基本的，而 $\displaystyle\int^x t\sqrt{1+t^4}\,dt$ 和 $\displaystyle\int^x t^3\sqrt{1+t^4}\,dt$ 是。

定积分

本节介绍了各种定积分及其在求和评估中的应用。在许多情况下，计算过程仅作概述，具体细节留作练习。

Translated Text: Proposition 15.12. 如果 n 是一个正整数，则 zh

$$\int_0^{\pi/2} \sin^n t\,dt = \begin{cases} \frac{(n-1)}{n}\frac{(n-3)}{(n-2)}\cdots\frac{2}{3} & \textit{if } n \textit{ is odd} \\[2ex] \frac{(n-1)}{n}\frac{(n-3)}{(n-2)}\cdots\frac{1}{2}\,\frac{\pi}{4} & \textit{if } n \textit{ is even} \end{cases}$$

换句话说，

$$\int_0^{\pi/2} \sin^{2k+1} t\,dt = \frac{(2^k k!)^2}{(2k+1)!}, \qquad \int_0^{\pi/2} \sin^{2k} t\,dt = \frac{(2k)!}{(2^k k!)^2}\,\frac{\pi}{4}.$$

替换 $u=(\pi/2)-t$ 将此积分转换为相同区间上余弦幂的积分。

证明。基本情形是

$$\int_0^{\pi/2} \sin t\,dt = -\cos t\Big|_{t=0}^{\pi/2} = 1$$

$$\int_0^{\pi/2} \sin^2 t\,dt = \frac{1}{2}\int_0^{\pi/2}(1-\cos 2t)\,dt = \frac{1}{2}\Big(t-\frac{1}{2}\sin 2t\Big)\Big|_{t=0}^{\pi/2} = \frac{\pi}{4}.$$

命题可从（15.6）和关于 n 的归纳得出。 □

伽马函数

练习12.14介绍了积分

$$(15.7) \qquad \Gamma(x)=\int_0^\infty t^{x-1}e^{-t}\,dt, \qquad 0<x<\infty,$$

满足 $\Gamma(n+1)=n!$ 通过对 n 的归纳。积分在非整数 x 上定义，并且是阶乘函数到正实数的扩展。本节根据鲁宾的研究，发展了 Γ 函数的一些性质。

第一个问题是建立对所给 x 的不定积分的收敛性。我们将积分分成 $\int_0^1+\int_1^\infty$ 并分别考虑各部分。

对于每个实数 α，根据练习9.21，有 lim $_{t\to+\infty}$ $t^\alpha e^{-t/2}$ $=$ = 0。因此，对于所有 $x>$ > 0,

$$0\le t^{x-1}e^{-t}\le \underbrace{(t^{x-1}e^{-t/2})}_{\to 0 \text{ as } t\to\infty}\cdot e^{-t/2}\le e^{-t/2} \qquad \text{for } t\gg 1.$$

因此，$\int_1^\infty t^{x-1}e^{-t}\,dt$ 对于所有 $x>0$ 收敛。现在，如果 $0<x<1$，那么 $0\le t^{x-1}e^{-t}\le t^{x-1}$，所以

$$\int_0^1 t^{x-1}e^{-t}\,dt<\int_0^1\frac{1}{t^{1-x}}\,dt,$$

收敛。如果 $x\le 0$，这些积分发散，因此表达式 (15.7) 是未定义的。

该节的主要结果是Bohr和Mollerup对Γ的简单抽象描述。

定理15.13. 函数 Γ 满足以下性质：

(a) $\Gamma(x+1)=x\Gamma(x)$ 对于所有 $x>0$。

(b) $\Gamma(n+1)=n!$ 对于所有整数 $n>0$。

(c) $\log\Gamma$ 是凸函数。

相反，如果 $f:(0,+\infty)\to\ \mathbf{R}$ 满足这些性质，那么 $f=\Gamma$。

证明。为了建立（a），使用 $u = t^x$ 和 $dv = e^{-t}\,dt$ 进行分部积分:

$$\begin{aligned}\Gamma(x+1) &= \int_0^\infty t^x e^{-t}\,dt = -t^x e^{-t}\Big|_{t=0}^{t=\infty} + \int_0^\infty x t^{x-1} e^{-t}\,dt \\ &= x\int_0^\infty t^{x-1}e^{-t}\,dt = x\Gamma(x).\end{aligned}$$

对于(b)，我们计算得到

$$\Gamma(1) = \lim_{b\to\infty}\int_0^b e^{-t}\,dt = \lim_{b\to\infty} -e^{-t}\Big|_{t=0}^{t=b} = \lim_{b\to\infty}(1-e^{-b}) = 1.$$

归纳 n 保证 $\Gamma(n+1) = n!$ 对于 $n > 0$。为了证明 (c)，观察如果 $\frac{1}{p}+\frac{1}{q}=1$，那么 $\Gamma\big(\frac{x}{p}+\frac{y}{q}\big)$ 的被积函数是

$$t^{\frac{x}{p}+\frac{y}{q}-1}e^{-t} = t^{\frac{x}{p}+\frac{y}{q}-\frac{1}{p}-\frac{1}{q}}e^{-t(\frac{1}{p}+\frac{1}{q})} = \big(t^{x-1}e^{-t}\big)^{1/p}\big(t^{y-1}e^{-t}\big)^{1/q}.$$

由于 $t^{x-1}e^{-t} > 0$ 对于 $t > 0$，Hölder 不等式（练习 9.19）意味着

$$\Gamma\big(\tfrac{x}{p}+\tfrac{y}{q}\big) \le \Big(\int_0^\infty t^{x-1}e^{-t}\,dt\Big)^{1/p}\Big(\int_0^\infty t^{y-1}e^{-t}\,dt\Big)^{1/q} = \Gamma(x)^{1/p}\Gamma(y)^{1/q}.$$

取对数,

$$\log\Gamma\big(\tfrac{x}{p}+\tfrac{y}{q}\big) \le \tfrac{1}{p}\log\Gamma(x) + \tfrac{1}{q}\log\Gamma(y),$$

这是所需的凸性陈述。

相反，假设 $f\colon (0,+\infty) \to \mathbf{R}$ 满足 (a)，(b) 和 (c)，并设 $\varphi = \log f$ 。性质 (a) 表示 $\varphi(x+1) = \varphi(x) + \log x$ 对于 $x > 0$，并且对 n 的归纳给出

$$\varphi(x+n+1) = \varphi(x) + \log\big((n+x)(n-1+x)\cdots(1+x)x\big). \tag{15.8}$$

属性（b）表示对于每个正整数 n，有 $\varphi(n+1) = \log(n!)$，而（c）表示 φ 是凸的。

修复 n，令 $0 < x < 1$，并考虑 φ 在区间 $[n, n+1]$，$[n+1, n+1+x]$，和 $[n+1, n+2]$ 上的差商。φ 的凸性意味着这些差商按非递减顺序排列。性质 (a) 意味着 φ 在 $[y, y+1]$ 上的差商是 $\log y$，所以

$$\log n \le \frac{\varphi(n+1+x)-\varphi(n+1)}{x} \le \log n+1. \tag{15.9}$$

乘以 x 并代入 (15.8),

x 对 $n \le \varphi(x) +$ 取对数 对 $\big((n+x)(n-$ 取对数 $1+x)\cdots x\big) - \log(n!) \le x\log(n+1)$,

或者

$$0 \le \varphi(x) - \log\frac{n!\,n^x}{x(x+1)\cdots(x+n-1)(x+n)} \le x\log\Big(1+\frac{1}{n}\Big) \tag{15.10}$$

当 $n \to +\infty$ 时，上界趋近于 0，因此夹逼定理给出

$$\varphi(x) = \lim_{n\to\infty} \log\frac{n!\,n^x}{x(x+1)\cdots(x+n-1)(x+n)} \tag{15.11}$$

对于$0 < x < 1$. 属性(a)意味着 φ 对所有 $x > 0$ 满足(15.11)。

方程（15.11）表明存在一个唯一的函数 f 满足性质（a）–（c），并且 $\log f$ 是右侧的极限；由于 Γ 满足（a）–（c），右侧的极限必须等于 $\log\Gamma(x)$。 □

一个论点的意外好处是以下等式

$$\Gamma(x) = \lim_{n\to\infty}\frac{n!\,n^x}{x(x+1)\cdots(x+n-1)(x+n)} \qquad \text{for } x > 0,$$

类似于将 exp 表征为几何增长极限的描述:

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

如同在exp的情况下，恰当地应用对Γ的这种描述会导致有趣的恒等式。例如，定义

$$\beta(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt \qquad \text{for } x,\ y > 0. \tag{15.12}$$

定理15.14. 对于所有 x，$y > 0$，有 $\beta(x,y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$。

证明。直接计算给出 $\beta(1,y) = 1/y$ 对于 $y > 0$。此外，$\log\beta(\cdot,y)$ 在某种意义上是凸的，即如果 $\frac{1}{p}+\frac{1}{q} = 1$，那么

$$\beta\big(\frac{x}{p}+\frac{z}{q}, y\big) \le \beta(x,y)^{1/p}\beta(z,y)^{1/q} \qquad \text{for all } x,\ y,\ z > 0.$$

TheTranslated TextThe 证明与Γ的对数凸性完全相同。 最初,

$$\beta(x+1,y) = \int_0^1 t^x(1-t)^{y-1}\,dt = \int_0^1\Big(\frac{t}{1-t}\Big)^x(1-t)^{x+y-1}\,dt.$$

积分部分，使用 $u = \big(t/(1-t)\big)^x$ 和 $v' = (1-t)^{x+y-1}$，得到

$$\beta(x+1, y) = -\frac{1}{x+y}\left(\frac{t}{1-t}\right)^x (1-t)^{x+y}\bigg|_{t=0}^{t=1} + \frac{x}{x+y}\int_0^1 t^{x-1}(1-t)^{y-1}\,dt.$$

由于“边界项”为零，$\beta(x+1,y) = \frac{x}{x+y}\beta(x,y)$。将这些部分组合起来，函数

$$f(x) := \frac{\Gamma(x+y)}{\Gamma(y)}\beta(x,y)$$

满足定理15.13的性质(a)–(c)，因此$f = \Gamma$。 □

推论15.15.2 $\displaystyle\int_0^{\pi/2} (\sin\theta)^{2x-1}(\cos\theta)^{2y-1}\,d\theta = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

由此可从替换 $t = \sin^2\theta$ 得出。从 $x = y = 1/2$ 得到一个有趣的定积分:

$$\pi = 2\int_0^{\pi/2} d\theta = \frac{\Gamma(\frac{1}{2})^2}{\Gamma(1)},$$

或者

$$\sqrt{\pi} = \Gamma(\tfrac{1}{2}) = \textstyle\int_0^\infty t^{-1/2}e^{-t}\,dt.$$

$u = \sqrt{t}$，$du = dt/2\sqrt{t}$的替换得到

$$\int_0^\infty e^{-u^2}\,du = \frac{\sqrt{\pi}}{2}, \qquad \text{or} \qquad \int_{-\infty}^{+\infty} e^{-u^2}\,du = \sqrt{\pi}. \tag{15.13}$$

取 $x = 1/2$ 或 $y = 1/2$ 表示以 Γ 为参数的正弦和余弦的幂的积分，参看命题15.12。

错误函数

让 $\mu \in \mathbf{R}$, $\sigma > 0$. 函数

$$\rho(x) = \frac{1}{\sqrt{2\pi\sigma}}e^{-(x-\mu)^2/2\sigma}$$

自然出现在概率中，作为著名的高斯密度或“钟形曲线”。变量 x 代表某种测量的结果，测量值位于 a 和 b 之间的概率是

$$P(a \le x \le b) = \frac{1}{\sqrt{2\pi\sigma}}\int_a^b e^{-(x-\mu)^2/2\sigma}\,dx$$

期望值和方差被定义为 是

$$E = \frac{1}{\sqrt{2\pi\sigma}} \int_a^b x e^{-(x-\mu)^2/2\sigma}\, dx, \quad V = \frac{1}{\sqrt{2\pi\sigma}} \int_a^b (x-\mu)^2 e^{-(x-\mu)^2/2\sigma}\, dx.$$

在概率方面，对 x 的重复测量预计将聚集在 E 附近，测量值到 E 的平均距离是标准差 $\sqrt{V}$。上面的高斯分布已归一化，使得 $E = \mu$ 和 $V = \sigma^2$，参见练习 15.21。

在变量线性变化范围内，我们不妨假设 $\mu =$ = 0 和 $\sigma =$ = 1。误差函数 erf : $\mathbf{R} \to \mathbf{R}$ 定义为如下不定积分

$$\text{(15.14)} \qquad \operatorname{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt,$$

查看图15.5。方程（15.13）表明，当$x \to +\infty$时，erf(x) → 1。x在a和b之间的概率是erf(b) − erf(a)。统计学

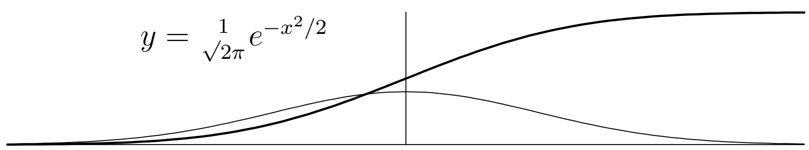


图15.5：高斯误差函数。

概率教材通常包含erf表。如上所述，erf不是一个基本函数。

黎曼ζ函数

考虑和 {v*}

$$\text{(15.15)} \qquad \zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}.$$

如果我们用 x 和 y 实数表示 $z = x + iy$，那么一般项的绝对值为

$$|n^{-z}| = |e^{-z\log n}| = |e^{-x\log n} e^{-iy\log n}| = |e^{-x\log n}| = n^{-x}.$$

与“p-系列”相比，系列（15.15）收敛当且仅当 $x = \operatorname{Re} z > 1$，对于所有此类 z 收敛是绝对的，并且对于每个 $\varepsilon > 0$,

收敛在集合Re $z > 1 + \varepsilon$ 上是统一的（在 z 中）。尽管上述级数在Re $z \leq 1$时并不收敛，但存在一个扩展到“有孔”平面 $\mathbf{C} \setminus \{1\}$，它在每个点 $z_0 \neq 1$附近是解析的，即存在一个具有正半径的复幂级数，其和为 ζ。这种解析扩展是黎曼ζ函数。

黎曼ζ函数，它不是初等的，在数学和物理学中有着深刻的渗透。已知（并且不难证明）ζ函数在负偶整数处消失，而所有其他零点都位于“临界带”$0 < \text{Re } z < 1$。数学中一个突出的未解问题是黎曼猜想:

> 每个位于Riemann zeta函数临界带中的零点都位于实部$z = 1/2$的直线上。

任何人对此问题取得重大进展都将赢得永恒的声誉；完全解决此问题将获得数学不朽。截至2000年5月，克莱数学研究所对黎曼猜想的解决提供100万美元的奖金。

我们的极端谦虚的目标是评估 $\zeta(2)$。我们根据技术将计算分为几个步骤。

引理15.16. $\displaystyle\int_0^1 \frac{\arcsin t}{\sqrt{1 - t^2}}\, dt = \frac{\pi^2}{8}$

证明。使用替换 $u = \arcsin t$，在 $t = 1$ 处注意极限。

□

引理15.17. $\displaystyle\int_0^1 \frac{t^{2k+1}}{\sqrt{1 - t^2}}\, dt = \frac{(2^k k!)^2}{(2k + 1)!}$

证明。使用替换 $t = \sin\theta$ 和命题 15.12。 □

引理15.18. $\displaystyle\arcsin x = \sum_{k=0}^{\infty} \frac{(2k)!}{(2^k k!)^2} \frac{x^{2k+1}}{2k + 1}$

证明。根据牛顿二项式定理（定理14.18），

$$\frac{1}{\sqrt{1 - x^2}} = 1 + \sum_{k=1}^{\infty} \frac{(2k)!}{(2^k k!)^2} x^{2k} \qquad \text{for } |x| < 1.$$

逐项积分，注意在 $x =$ 附近的非一致收敛问题。

□

引理15.19. $\int_0^1 \frac{\arcsin t 1}{\sqrt{}} - t^2 \, dt = \sum_{k=0}^{\infty} 1 \cancel{(2k+1)}$
2

证明。使用引理15.16展开被积函数:

$$\frac{\arcsin t}{\sqrt{1-t^2}} = \sum_{k=0}^{\infty} \frac{(2k)!}{(2k+1)(2^k k!)^2} \frac{t^{2k+1}}{\sqrt{1-t^2}},$$

然后逐项积分。引理15.17允许你评估得到的积分，并且发生了很多消去。 □

Lemma 15.20. $\sum_{n=1}^{\infty} \frac{1}{n^2} = \zeta(2) = \frac{\pi^2}{6}$

证明。将级数分为偶数项和奇数项。（为什么这是允许的？）你知道奇数项的和，偶数项的和可以用 $\zeta(\hat{}2)$ 本身来表示。

□

存在评估序列 $\zeta(2k)$ 的系统方法，尽管需要更好的技术，例如复数轮廓积分或傅里叶级数。结果发现 $\zeta(2k)$ 是 π^{2k} 的显式有理倍数。有趣的是，值 $\zeta(2k+1)$ 未知。1978年，A. Apèry 证明了 $\zeta(3)$ 是无理数。

练习

地球表面被划分为跨越约15°经度的时间带，当环绕地球一周时，总时间变化为24小时.[2] 对于这个问题，假设地球上每个人都保持太阳时间：“中午”是指太阳经过你正南方的时刻，假设你位于北半球。假设你的经度为0°，此时正是中午。（因为这是数学，我们只是在重新定义“经度”；你不需要去格林尼治！）

(a) 找出地球表面上某一点的时间作为经度的函数的公式。确保一致地调整你的时间角度单位。

[2] In reality, time zones obey political boundaries almost as much as geographic ones.

(b) 一天中的时间是位置的连续函数吗？如果不是，不连续性在哪里，如果你穿过不连续性，你的时间测量会发生什么变化？

(版权) 你的时间公式如何类似于对数的原理分支？根据你的公式，北极现在是什么时间？

(д) 假设存在一个没有间断的“太阳时函数”。如果你环绕地球一周，会发生什么？

实际上，不连续性是太平洋中的一条固定线，而不是“午夜点”，它随着地球的旋转而移动。$\diamond$

练习15.2评估以下内容：

$$\int^x \frac{dt}{1-t^2} \qquad \int^x \frac{t\,dt}{1-t^2} \qquad \int^x \frac{dt}{t-t^2} \qquad \int^x \frac{dt}{t-t^3} \qquad \int^x \frac{1+t^2}{t-t^3}\,dt$$

提示：一些部分分式已经为您完成。 $\diamond$

练习15.3 以文本中的示例为模型，找到以下递归公式

$$\int^x t^n \cos t\,dt \qquad \int^x t^n e^t\,dt$$

使用递归公式计算 $\int^x t^4 \sin t\,dt$ $\diamond$

练习15.4 找到 $\int^x \cos^n t\,dt$ 的递归公式 $\diamond$

练习15.5填写$\int^x$秒$^3\,\theta\,d\theta$评估的详细信息。$\diamond$

练习15.6. 如果15.9推论的证明缺少细节，请填写；具体来说，验证在替换 $t = \tan\frac{s}{2}$ 的情况下，圆三角函数成为 t 的有理函数。有关球面投影的详细信息，请参阅练习3.7。$\diamond$

练习15.7 使用以下提纲进行反微分sec：将分子和分母乘以cos，并使用$\cos^2 = 1 - \sin^2$。一个替换将其转化为一个有理函数，其部分分式分解是你所知道的。$\diamond$

练习15.8 使用文本中的一个定理来证明 $\int^x \sin t^2\,dt$ 和 $\int^x \cos t^2\,dt$ 不是初等函数。

$\diamond$

练习15.9求$\int_0^\infty e^{-t}\sin xt\,dt$在$x$实数下的值。 ◇

练习15.10 评估 $\int^1 x^2+1\ x^4+1\,dx$。
提示：将分子和分母都除以 x^2，然后使用（不恰当的）替换 $u=x-1/x$。◇

练习15.11 在这个练习中，你将评估 $\displaystyle\int_0^{\pi/2}\log\sin x\,dx$

(a) 证明这个广义积分收敛。建议：寻找sin在0附近的合适界限。

(b) 证明 $\displaystyle\int_0^{\pi/2}\log\sin x\,dx=\int_0^{\pi/2}\log\cos x\,dx$

(c) 使用代换和正弦的二倍角公式来评估原始积分。

如果您能，通过类似技巧评估 log cos 的积分，而不使用部分 (b)。◇

练习15.12 证明 $\Gamma(n+\frac{1}{2}) = (2n-1)!!\ 2_n\ \sqrt{\pi}$。 ◇

练习15.13 使用文本中的公式编写积分

$$\int_0^{\pi/2}\sin^n\theta\,d\theta,\qquad \int_0^{\pi/2}\cos^n\theta\,d\theta,$$

关于 Γ 函数。 ◇

练习15.14 评估

$$\int_{-1}^{1}(1-t)^n(1+t)^m\,dt,\qquad m,\ n\in\mathbf{N},$$

关于阶乘。
建议：首先使用 Γ 函数评估 $\displaystyle\int_0^1 t^n(1-t)^m\,dt$。 ◇

练习15.15 证明{v*}

$$\Gamma(x)=\frac{2^{x-1}}{\sqrt{\pi}}\Gamma(\tfrac{1}{2})\Gamma(\tfrac{x+1}{2})$$

对于所有 $x>0$。 ◇

练习15.16 填写引理15.16的细节。 $\diamond$

练习15.17 填写引理15.17的细节。 $\diamond$

练习15.18 填写引理15.18的细节。 $\diamond$

练习15.19 填写引理15.19的细节。 $\diamond$

练习15.20 填写引理15.20的细节。 $\diamond$

练习15.21 令 $\mu \in \mathbf{R}$ 和 $\sigma > 0$.

(a) 证明以下公式：{v*}

$$\frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} xe^{-(x-\mu)^2/2\sigma}\, dx = \mu.$$

建议：令 $z = x - \mu$。

(b) 证明以下公式：{v*}

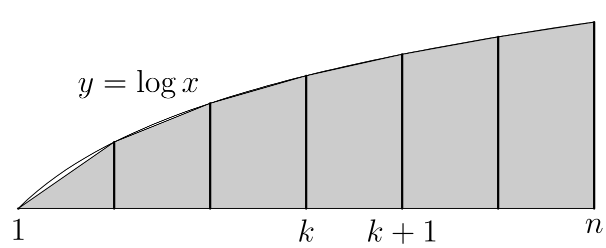
$$\frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-(x-\mu)^2/2\sigma}\, dx = \sigma.$$

建议：使用部分(a)的替换，然后通过分部积分，使用$u = z$。

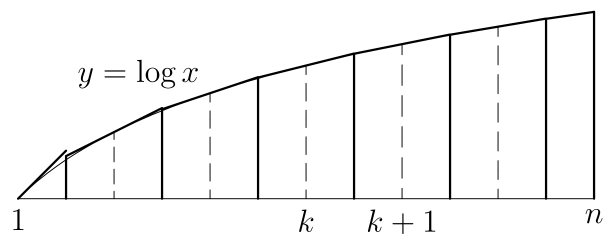
在每一部分中，你可以使用以下事实：lim(erf, $+\infty$) = 1。 $\diamond$

练习15.22 这个练习提供了一个对大 n 的 $n!$ 的粗略近似。出发点是观察到对数是凹的，因此图形位于每条切线之下并位于每条割线之上。我们将使用这些线来得到对数图形在1和 n 之间的面积的上界和下界。

(a) (下界) 将 $[1, n]$ 分成 $n -$ 个单位长度的区间。求梯形的总面积：



(b) (上界) 考虑到对数函数图像在整数 k 处的切线。位于此切线下方且在垂直线 $x = k \pm \frac{1}{2}$ 之间的梯形面积是对数从 k 到 $k+$ 积分的上界 1:



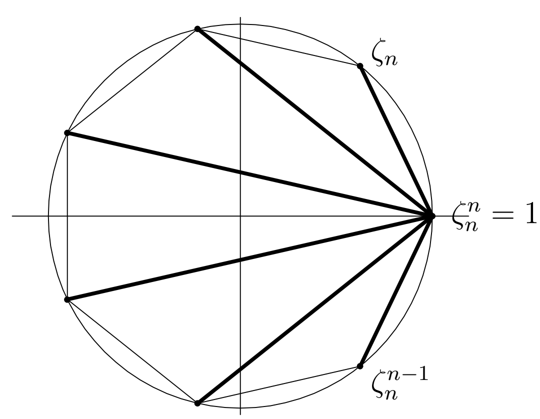
证明以 k 为中心，$1 < k < n$ 的梯形面积为 log k。当 $k =$ 为 1 或 $k = n$ 时，必须进行特殊考虑。求梯形和三角形的总面积。

（c）使用部分（a）和（b）中找到的界限来证明

$$e^{7/8}\frac{n^n\sqrt{n}}{e^n} < n! < e\frac{n^n\sqrt{n}}{e^n}.$$

换句话说，$e^{7/8} < \frac{n!}{(n/e)^n\sqrt{n}} < e$ 对于 $n \geq 1$。 ◇

练习15.23 设 $n \geq 3$ 为一个整数，并设 $\zeta_n = e^{2\pi i/n}$；回忆起 ζ_n 的幂构成单位圆内接正 n 边形的顶点。考虑连接 $1 = \zeta_n^n$ 到其他 $n-1$ 个顶点的弦集。



证明这些弦长度的乘积是 n，即边的数量！（只有少数特殊情况可以通过手工计算。）$\diamond$

练习15.24（动量变换）设 $\varphi : (-1, 1) \to \mathbf{R}$ 是连续且正的。

(a) 证明该方程

$$x = \int_0^{u(x)} \frac{dt}{\varphi(t)}$$

定义了一个递增的，$\mathscr{C}^1$ 函数 u，以及 $u' = \varphi \circ u$。u 的定义域和值域是什么？

(b) 证明该方程

$$f(x) = \int_0^{u(x)} \frac{t\,dt}{\varphi(t)}$$

定义一个 $\mathscr{C}^2$ 函数 f，并证明 $f' = u$。

(c) 假设 $\varphi(t) = 1 - t^2$。求 u 和 f 作为 x 的显式函数。

(д) 假设 $\varphi(t) = 1 + t^2$。求 u 和 f 作为 x 的显式函数。

提示：您将需要本书中的所有技术，但这个问题被保留到最后一个章节有其原因。$\diamond$

后置脚本

在里程碑式的论文《大教堂与市集》中，开源软件倡导者埃里克·S·雷蒙德写道：“每一个优秀的软件作品都是从满足开发者个人需求开始的。”无论这本书作为教学工具的质量如何，它都满足了作者的一个需求，即向那些背景是20世纪90年代高中数学课程的学生展示有趣、生动、严谨的数学。自从作者还是高中生以来，许多事情都发生了变化；如今，更加强调概念理解而非技术细节，通过实验发现，数值逼近和应用。

虽然这些目标是值得赞扬的，但将数学呈现为一种实验科学，以及依赖直观原则和可能性论证而不是精确的定义和逻辑、演绎证明，存在危险。在微积分课程的领域，风险可能并不那么明显，因为问题通常被仔细选择和陈述，以便适用于正在教授的概念方法。然而，现实生活永远不会像教科书那样干净简单，没有技术知识的纯粹直觉很容易走偏。这本书试图弥合技术和理解的世界，将数学呈现为一种由直观原则指导的精确语言，并引导到美丽、意想不到的目的地。

这本书最初是1997-8年在多伦多大学讲授的分析I（MAT 157Y）课程笔记的集合。虽然这门课程在所有衡量标准下并非完美无缺，但它是我所经历的最令人满意的长期教学经历之一。一位“UofT”的高级同事曾经说过，数学家应该向学生宣传自己是心智的重量训练师.[1] 在这个比喻中，分析I取得了巨大成功。问题集极具挑战性，但至少总有一位学生做出了一些

[1]“We are pumping brains!”

关于一个问题取得的进展。学生被激励去调查超出作业要求的问题，因此得以深入理解和培养独立思考能力。相当一部分学生继续在顶尖研究型大学攻读数学和物理学研究生课程。我最先要感谢的是《分析I》课程的学生，尤其是Jameel Al-Aidroos、Sunny Arkani-Hamed、Ari Brodsky、Chris Coward、Chris George、Fred Kuehn、Brian Lee、Cuong Nguyen、Caroline Pantofaru、Dave Serpa和Dan Snaith。

MAT 157Y如果没有助教加里·鲍姆加特纳和布莱尔·马多尔的赫拉克勒斯般的努力，也远不会那么成功。他们不知疲倦地激发学生的好奇心，创造性地构建例子和隐喻，并在辅导以及出测试和考试时，调和我的抽象飞跃。加里的一些锐利的极大值问题作为练习出现。

第一章的核心是Steve Hook的一份讲义，关于写作证明，他在1986年夏天将其分发给他在加州大学伯克利分校的真实分析课程。

Andres del Junco 在多伦多大学慷慨地提供了他的分析I课程材料。它们的影响在文本的几个地方都有所体现，无论是作为例子还是作为练习的灵感来源。

Weierstrass逼近定理的证明来自Serge Lang在20世纪80年代末于加州大学伯克利分校给出的一堂为期一小时的本科生讲座。

在给美国数学会《通报》的一封信中，Donald Knuth建议编写一本基于O-记法的微积分书籍。当我偶然看到他的信——在它被写下的多年之后，以及在这本书的大部分内容已经完成之后——将这本书从ε-δ语言“移植”到O-记法的可能性似乎既可行又具有教育意义。“粗心大意”的演算既具体到足以让学生在心理上感到满足，又强大到足以表达真正微妙的思想。这本书是否可以被认为是Knuth愿景的成功实现还有待观察。

我感谢几位在Usenet组sci.math上发帖的人，但尤其感谢马修·韦纳，他关于积分的简单表述促使我在第15章中包含一些一般定理，而不是仅仅散布民间传说结果。约翰·贝兹的数学物理网页“每周发现”也是灵感的来源。他关于魏尔斯特拉斯 $\wp$ 函数的描述，与使用代数积分定义圆三角函数的类比。

函数，非常诱人包含在第15章中，只是遗憾地省略了。

多年来，我收集了一些有趣的练习、趣闻和其他数学小知识，通常记录在杂乱的纸张或易出错的记忆中，我并不总是知道这些物品的来源。对于任何未注明出处的作品，我表示最诚挚的歉意。

这本书是在GNU/Linux系统上使用免费软件制作的。免费软件的概念对学者来说很重要，值得简要解释。计算机程序是一组执行某些任务的指令，实际上可以很容易地复制。软件是用人类可读的编程语言编写的源代码，并通过一个称为编译器的特殊程序转换为机器可读的可执行文件或二进制文件。将食谱（制作菜肴的指令）视为一个例子是有帮助的，它可以复制和分享而不丢失原始内容。在计算机的早期（大约在1980年之前），软件就像食谱一样编写和分享。在20世纪80年代，围绕软件作为商品的概念兴起，就像食谱中的成分一样。在这个新模型中，分享是被禁止的，成分（源代码）是保密的。不幸的是，由于技术上可以复制软件，因此有必要将所有软件用户视为潜在的窃贼。免费软件运动旨在在软件共享的精神下创建一个社区，每个人都可以自由地查看和修改程序的源代码以满足他们的特定需求。GNU项目由理查德·M·斯托尔曼于1984年启动，其目标是从头开始创建一个完整的计算机操作系统，并将其置于一个许可下，允许任何人阅读、修改和分发源代码，仅限于这些开放条款不得撤销。如今，GNU/Linux操作系统被全球数百万人所使用。“Linux”是林纳斯·托瓦兹的类Unix内核，该程序将硬件资源分配给所有其他运行的程序。在学术环境中，免费软件可以说是唯一可接受的软件；作为一名科学家或数学家，除非你确切知道计算机在做什么，否则你不能完全信任计算的结果。免费软件易于用户检查。虽然你可能永远不会阅读C编译器或Linux内核的源代码，但许多有知识的人已经审查了代码，并且你（作为个人用户）保留有选择审查代码的权利。在最实际的意义上，这项权利与准确

食品成分标签。本书使用 emacs，GNU 文本编辑器和 teTEX（Donald Knuth 的 Free TEX 排版引擎的实现）制作。图示是用 ePiX 创建的，这是一个用于创建数学精确图形的免费实用程序。

一些教科书对本书产生了明显的影响。最值得注意的是迈克尔·斯皮瓦克所著的《微积分》（第3版）[8]，这是MAT 157Y课程使用的教材。尽管该教材优秀，且非常适合MAT 157Y这样的课程，但我还是希望有稍微不同的重点、编排、覆盖范围和风格。沃尔特·鲁宾的经典著作《数学分析原理》[6]也极具影响力，一方面是因为我第一次通过这部杰作接触到了实变函数，另一方面是因为鲁宾在选择例子和练习时具有无可挑剔的品味，这些例子和练习突出了从初步观察中难以察觉的微妙技术细节。阿波斯特尔[2]和库朗与约翰[3]的较老版本的微积分教材是灵感的来源，但（遗憾的是）像斯皮瓦克的《微积分》一样，很少被认为适合现代学生。威廉·普里斯特利的《微积分：自由艺术》[5]是微积分的愉快入门课程，但针对的是不打算进一步学习数学的文科学生。最后，“哈佛”教材，由德博拉·休斯-哈莱特、安德鲁·格莱森等人编写，对本书的风格产生了独特的影响。这个可能令人惊讶的坦白值得更详细的评论。

在我的经验中，技术细节类似于骨骼，而概念直觉类似于肌肉和皮肤。没有肌肉，骨骼就会僵硬和惰性。没有骨骼，肌肉就无法有目的地移动。共同作用，协同工作，它们赋予我们力量和优雅的动作。在20世纪中叶，数学教育的天平已经偏向极端的形式主义（新数学）。从20世纪80年代末开始，数学教育工作者开始采用一种非技术性的概念理解作为一种补救措施，以应对未能触及大多数学生的“无意识的正式主义”。这种趋势目前正处于全盛时期，并在21世纪初的微积分教科书的内含和风格中广泛显现。虽然“相关性”和“包容性”是数学教学的理想目标，但如果完全忘记了数学的严谨基础，这种努力注定会失去其知识实质。20世纪初代数几何的灾难性失败浮现在脑海中：定理是通过直观论证证明的，并且常常是错误的。枚举几何的文献受到了损害，该领域的发展被推迟了几十年，直到建立了适当的基础。

无人将　　e well-served if generations of students, many of　他们

未来教师，在未接触真实分析的细节的情况下成长。这本书是对在微积分的概念理解肌肉中注入逻辑和技术形式主义骨骼的谦逊尝试，即拥抱概念理解的教育趋势，同时不失数学的智力基础。微积分及其阐述的基础是在几个世纪前奠定的，声称在材料或表述上有任何原创性都是荒谬的。尽管如此，我相信这本书填补了一个空白。虽然不是每个学生都期望按顺序从头到尾阅读这本书，但把细节放在一个地方是很重要的。微积分不是一次就能学会的科目。事实上，这本书几乎假设读者已经学过一年的微积分，就像MAT 157Y的学生一样。我希望这本书会随着读者的成长而成长，在多次阅读中保持可读性和信息性，并且为当前微积分教材和更高级的真实分析教材之间提供一个有用的桥梁。

Andrew D. Hwang
May 18, 2003
Sterling, MA

参考文献

[1] Lars V. Ahlfors，《复变函数论》，第3rd版，麦格劳-希尔，1979
。[2] Tom M. Apostol，《微积分》，Blaisdell（兰登书屋），1962。[
3] Richard Courant 和 Fritz John，《微积分与解析导论》，约翰威利
科学出版社，1975。

[4] Serge Lang，《微积分初步》，第3rd版，Addison-Wesley，1973年
。

[5] 威廉·麦戈文·普里斯特利，《微积分：一门自由艺术》，第2nd版，
斯普林格出版社，1998年。

[6] 沃尔特·鲁滨逊，《数学分析原理》，第3rd版，麦格劳-希尔，1976
年。

[7] George F. Simmons，《带有应用和历史注释的微分方程》，第2nd
版，麦格劳-希尔，1991年。

[8] 迈克尔·D·斯皮瓦克，《微积分》，第3rd版，出版或消亡，1994年。

索引