学校的理想装备

电子图书·学校专集

校园网上的最佳资源

天才引导的历程

BOOK

内部资料 非卖品

“科学与人译丛”出版说明

英国著名科学专栏作家布赖恩·阿普尔亚德在其《理解现在——科学与现代人的灵魂》一书中有这样一段话：

“1609 年，加利莱奥·伽利略使用一架望远镜观看月亮。这一时刻，对世界的意义如此重大，以至人们将它与耶稣的诞生相提并论。因为，就像在伯利恒，自这一时刻，人类生活中的不可能成为可能。”

阿普尔亚德据此将科学划分为伽利略之前的科学，或称“智慧”，以及从 1609 年开始的现代科学。前一科学建立在推理基础上，后一科学建立在观察与实验基础上。经过如此划分，我们习以为常的科学，竟然只有 400 年的历史。

但人类就在这 400 年内经历了飞速发展。

我们有了蒸汽机，有了轮船，有了电话、电报，有了飞机、火箭，有了电视、电脑、互联网络，我们还有重力场理论、元素周期表、量子力学、相对论乃至被称为“自然中最基本物体”的超弦。工业革命、农业革命、信息革命使人类的社会生活发生了前人难以想象的变化。

人类改造了自然，也改造了人类自己。回顾这一切，人类完全有理由感到自豪。因为，人类就像上帝，也有自己的“创世纪”。人说，要有科学，就有了科学。科学是好的，它行之有效。

然而，“创世纪”中写道“到第七日，上帝造物的工已经完毕，就在第七日歇了他一切的工，安息了”。而人类的工却没有完毕，400 年后的今天仍然不能安息。

就像有光必有影，人在发现、发明、创造、拥有上述一切的同时，还得到了原子弹、氢弹、核泄漏、酸雨、温室效应、臭氧层空洞乃至伴随科学技术而来的种种风险。

人类曾以为已找到了通往自由王国的必由之路，他将乘着科学的飞船，摆脱一切束缚，重新确立自己在宇宙中的位置。但在科学爆炸的二十世纪，人类终于开始反思：

科学行之有效，但它是否就是真理？

为此，我们编辑了这套《科学与人译丛》，陆续分辑推出。其中，有对信息崇拜的批判，有对生命起源的求索，有对技术所导致风险的分析，有对世界最新科学动态和研究方向的展望。数学家用对策论证明，完全的民主实际上并无可能；物理学家提出全新的超弦理论，试图统一描述所有的力、物质的所有基本粒子和时空，继量子力学和相对论之后，成为“第三次物理学革命的重要标志”……《译丛》汇集了物理学家、数学家、生物学家、天文学家、哲学家、人类学家、伦理学家……自本世纪后半期、尤其是在本世纪末打通自然科学与社会科学之间的隔膜，对科学这一决定人类命运的工具的深刻思索。通过这套丛书，我们期望读者可以对科学的现状、科学的未来、科学的正面与负面效应，有一个较为全面的了解，更好地认识科学、掌握科学、利用科学。

中国对外翻译出版公司

1997 年 2 月

自序

伯特兰·罗素在自传中回忆了他青年时的危机：

“有一条小路，穿过田野，通向新南盖特，我经常独自一人到那里去观看落日，并想到自杀。然而，我终于不曾自杀，因为我想更多地了解数学。”

诚然，很少有人能够如此虔诚地皈依数学，但是，确有许多人懂得数学的力量，特别是懂得数学之美。本书谨献给那些愿更深入地探索漫长而壮丽的数学史的人们。

对于文学、音乐和美术等各种学科，人们一向以考证杰作——“伟大的小说”、“伟大的交响乐”、“伟大的绘画”，作为最适宜和最有启发性的研究课题。人们就这些题目著书立说，授课讲学，使我们能够了解这些学科的某些里程碑和创造这些里程碑的伟人。

本书运用类似的方法来研究数学，而书中大师们创造的不是小说或交响乐，而是定理。因此，本书不是一本典型的数学教材，没有一步一步地推导某些数学分支的发展，也没有强调数学在确定行星运行轨道、理解计算机世界，乃至结算支票等方面的应用。当然，数学在这些应用领域取得了惊人的成就，但并非这些世俗功利促使欧几里得、阿基米德或乔治·康托为数学殚精竭虑，终生不悔。他们并不认为应借功利目的为自己的工作辩解，正如莎士比亚不必解释他何以要写十四行诗，而没有写菜谱，或凡高何以要画油画，而没有画广告画一样。

我将在本书中从数学史的角度来探讨某些最重要的证明和最精巧的逻辑推理，并重点阐述这些定理为什么意义深远，以及数学家们是如何彻底地解决了这些紧迫的逻辑问题的。本书的每一章都包含了三个基本组成部分：

第一部分是历史背景。本书所述及的“伟大定理”跨越了 2300 多年的人类历史。因而本人在论述某一定理之前，将先介绍历史背景，介绍当时的数学状况乃至整个世界的一般状况。像其他任何事物一样，数学也是在一定的历史环境中产生的。因此，有必要指明卡尔达诺三次方程的解法出现在哥白尼日心说公布后两年和英格兰国王亨利八世死前两年，或强调青年学者艾萨克·牛顿 1661 年进入剑桥大学学习时，王政复辟对剑桥大学的影响。

第二部分是传说性的。数学是有血有肉的实实在在的人的造物，而数学家的生平则可能反映出灵感、悲剧或怪诞。本书所涉定理体现了许多数学家的勤奋努力，从交游广阔的李昂纳德·欧拉到生性好斗的约翰·伯努利和带有最市俗的文艺复兴特征的赫罗拉莫·卡尔达诺，不一而足。了解这些数学家的不同经历，有助于我们更好地理解他们的工作。

第三部分，也是本书的重点，是在这些“数学精萃”中所表现出的创造性。不读名著，无从理解；不观名画，无从体味，同样，如果不去认真地、一步一步地钻研这些证明方法，也不可能真正掌握这些著名的数学定理。而要理解这些定理，就必须全神贯注。本书各章仅仅意在为理解这些定理梳理线索。

这些数学里程碑还具有一种永世不灭的恒久性。在其他学科，今天流行的风尚，往往明天就遭人遗忘。一百多年前，沃尔特·司各脱爵士还是

当时英国文学中最受尊重的作家之一，而今天，人们对他已淡然。20 世纪，超级名星们匆匆来去，转瞬即成历史，而那些旨在改变世界的观念，最终却常常变成思想垃圾。

诚然，数学也必须时常改变其趣味。但是，受严格逻辑限制而证明的数学定理则是永恒的。公元前 300 年欧几里得对毕达哥拉斯勾股定理的证明，并未因时光的流逝而丝毫丧失它的美与活力。相反，古希腊时期的天文学理论或医术却早已变成陈旧而有点儿可笑的原始科学了。19 世纪的数学家赫尔曼 · 汉凯尔说得好：

“ 就大多数学科而言，一代人摧毁的正是另一代人所建造的，而他们所建立的也必将是另一代人所破坏的。只有数学不同，每一代人都在旧的结构上加进新的内容。 ”

在这种意义上，我们探讨伟大数学家历久弥新的成果，就能够从中体会奥利弗 · 亥维赛精辟的论说：“ 逻辑能够很有耐性，因为它是永恒的。 ”

对体现数学精髓的这些定理的选择，是由许多因素决定的。如前所述，我主要的考虑是找到具有深刻见解或独创性的论题。当然，这里有一个个人好恶的问题。我承认，不同的作者可能会选取不同的定理。然而，能够直接看到数学家通过巧妙的演绎，将看似深奥的问题变得清晰易懂，确实是一种不同寻常的经历。据说，聪明人可以战胜困难，而天才则可以战胜不可能。本书将展示许多天才。这里有真正的经典——数学中的《蒙娜 · 丽莎》或《哈姆雷特》。

当然，选择这些定理也有其他的考虑。其一，我希望本书能够包容历史上主要数学家的定理。例如，欧几里得、阿基米德、牛顿和欧拉必不可少。忽略这些数学家，犹如研究美术史而不提伦勃朗或塞尚的作品一样。

其二，为求全面，我兼顾到数学各个分支。本书的命题涉及平面几何、代数、数论、解析和集合等内容。各种命题，以及它们之间的偶然联系和相互影响，为本书增添了一些生动的气息。

我还希望能在本书中容纳各种重要的数学定理，而不仅仅是一些小巧的游戏或机变。实际上，本书的大部分定理或者解决了长期存在的数学问题，或者为将来提出了意义深远的问题，或者二者兼而有之。每一章的结尾处有后记，一般记述伟大定理提出的问题及其在数学史上的影响。

这里有一个定理难度的问题。显然，数学有许多伟大的里程碑，其深度和难度除专家外，其他人都会感到莫测高深。在一本针对一般读者的书中引入这些命题，是十分愚蠢的。本书所涉定理，仅要求具备高中代数和几何知识即可。但有两处例外，一是第 9 章在讨论欧拉的定理时使用了三角学的正弦曲线，二是第 7 章在讨论牛顿的定理时应用了初等积分；许多读者已经掌握了这些知识，而对于那些尚未掌握这些知识的读者，本书做了一些解释，以帮助他们克服阅读中的困难。

应当指出，本书不是一本学术著作。一些重大而微妙的数学或历史问题当然不可能在这种书中一一述及。但我尽力避免编入一些不正确或历史上不准确的材料，因为这不是对所有问题的所有方面追根问底的时间和场合。总之，本书是一本大众读物，不是科学著作或新闻报道。

就此，我必须对论证的确切性说几句。在准备写这本书的时候，我发现不可避免地要在定理创始人最初使用的符号、术语和逻辑思维与现代读者对数学资料的理解要求之间作出某些折衷。完全照搬原作会使人感到非

常难于理解；但严重偏离原作又与我的历史目标相冲突。总之，我实际上尽力保留了定理原作的全部要旨和大量细节。而我所作的某些修改并不严重，不过就像是用现代乐器演奏莫扎特的乐曲一样。

现在，我们即将开始穿越二千年数学里程的旅行。这些定理虽然古老，但在历经多少个世纪之后，却依旧保持着一种新鲜感和灿烂的魅力。我希望读者能够理解这些论证，并能够领会这些定理的伟大之处。对于那些做到了这一点的读者，我希望他们不仅会对他人的伟大之处肃然起敬，还会因为能够理解大师著作而怡然自得。

威廉·邓纳姆
俄亥俄州，哥伦布

鸣谢

我在编写本书时，曾得到过许多机构和个人的帮助，谨在此表示感谢。首先，我应感谢私人和公共部门提供的宝贵赠款：利利捐赠基金有限公司提供的 1983 年夏季津贴和全国人文学科基金会为 1988 年题为“ 历史上的数学经典定理 ”夏季研讨会提供的资金。利利捐赠基金有限公司和全国人文学捐赠基金会的支持，使我得以归纳以往对数学史的散乱兴趣，形成在汉诺威学院和俄亥俄州立大学教授的课程。

我衷心感谢俄亥俄州立大学，特别是数学系，在我作为客座教员编写本书时所给予我的热情支持。数学系主任约瑟夫·费拉尔以及琼·莱泽尔和吉姆·莱泽尔，在我任客座教员的两年期间，给予我可靠的帮助和支持，对此，我永志不忘。

许多个人也为本书提供了帮助。感谢图书馆管理员鲁思·埃文斯在我 1980 年休假年期间为我提供了 1900 年以前的数学资料汇编；感谢全国人文学科基金会的史蒂文·泰格纳和迈克尔·霍尔对本书之前夏季研讨会提出的良好建议；感谢卡罗尔·邓纳姆的热情和鼓励；感谢俄亥俄州立大学的艾米·爱德华兹和吉尔·鲍默—皮纳为我介绍麦金托什文字处理系统的细节；感谢威利公司编辑凯瑟琳·肖沃尔特、劳拉·卢因和史蒂夫·罗斯对一个初出茅庐的作者的宽容；感谢全国最有权威的发言人之一，鲍灵格林州立大学的 V．弗雷德里克·里基提出的关于数学也像其他学科一样具有不容忽视的历史的观点；感谢巴里·A．西普拉和韦斯特蒙特学院的拉塞尔·豪厄尔对本书手稿所作的非常彻底而有益的审查；感谢汉诺威学院的乔纳森·史密斯在出版前的最后阶段提出的编辑意见。

我应特别感谢彭尼·邓纳姆，她为本书绘制了插图，并就书的内容提出了许多宝贵建议。彭尼是一位非凡的数学教师，是我们共同主办全国人文学科基金会研讨会的不可替代的同仁，是支持者、顾问、夫人和可以想象到的最好朋友。

最后，我要特别感谢布伦丹和香农两位大师。

第一章

希波克拉底的求新月形面积定理

（公元前约 440 年）

论证数学的诞生

我们对人类远古时代数学发展的认识，在很大程度上依靠推测，是根据零星的考古资料、建筑遗迹和学者的猜测拼凑而成的。显然，随着公元前 15000 至 10000 年间农业的发明，人类不得不应付两个最基本的数学概念（至少是以初步形式）：量和空间。量的概念，或“数”的概念是在人们数羊或分配粮食时产生的，经过历代学者几百年的推敲和发展，量的概念逐渐形成了算术，后来又发展成代数。同样，最初的农夫也需要认识空间关系，特别是就田地和牧场的面积而言，随着历史的发展，这种对空间的认识就逐渐形成了几何学。自从人类文明之初，数学的两大分支——算术和几何，就以一种原始的形式共存。

然而，这种共存并非永远和谐。数学史上一个持续的特征就是在算术与几何之间始终存在着紧张关系。有时，一方超过了另一方，有时，另一方又比这一方在逻辑上更占优势。而一个新发现，一种新观点，都可能会扭转局面。也许，有人会感到奇怪，数学竟然像美术、音乐或文学一样，在其漫长而辉煌的历史进程中，同样存在着激烈的竞争。

我们在古埃及文明中，发现了数学发展的明显迹象。古埃及人研究的重点是数学的应用方面，以数学作为工具，促进贸易、农业和日益复杂的日常生活其他方面的发展。根据考古记载，在公元前 2000 年以前，埃及人已建立了原始数系，并具备了某些有关三角形和棱锥体等的几何概念。例如，据传说，古埃及建筑师用一种非常巧妙的方法确定直角。他们把 12 段同样长的绳子相互连成环状（如图 1．1 所示），把从 B 到 C 之间的五段绳子拉成直线，然后在 A 点将绳子拉紧，于是就形成了直角 BAC。他们将这种构形放在地上，让工人们按照这个构形在金字塔、庙宇或其他建筑的拐角处建成标准的直角。

这种构图表明，古埃及人已对直角三角形的毕达哥拉斯勾股弦关系有所认识。他们似乎懂得，边长为 3、4 和 5 的三角形肯定会含有直角。当然，$3^2+4^2=9+16=25=5^2$，我们从中可以看到在所有数学关系中最重要的关系之一——勾股关系的早期曙光（见图 1.2）。

从技术角度说，古埃及人的这种认识还不是勾股定理本身。勾股定理申明，“如果△BAC 是直角三角形，则 $a^2=b^2+c^2$”。而古埃及人的认识则是勾股定理的逆定理，“如果 $a^2=b^2+c^2$，则△BAC 是直角三角形”。也就是说，关于命题“如果 P，则 Q”，对其相关命题“如果 Q，则 P”，我们称之为逆命题。我们将会看到，一个完全正确的命题，其逆命题可能是错误的，但著名的勾股定理则不然，不论正命题，还是逆命题，都是正确的。实际上，这些就是我们将在下一章讨论的“伟大定理”。

虽然古埃及人对 3-4-5 直角三角形的几何性质有所认识，但他们是否具有更广义的理解，例如，对于同样含有直角的 5-12-13 三角形或 65-72-97 三角形（因为在这两个三角形中，都是 $a^2=b^2+c^2$），则还是个疑问。更重

要的是，没有迹象表明，古埃及人是如何证明这些关系的。也许，他们掌握某些逻辑论证，以支持他们对 3-4-5 三角形的观察；也许，他们仅仅是靠反复试验。但无论如何，在埃及的文字记载中都没有发现通过严密的逻辑推理，证明一般数学规律的迹象。

下面的古埃及数学例子也许可以给人以启发：这是他们发现截棱四棱锥体积的方法——即一个用平行于底面的平面截去顶部的四棱锥体（见图 1．3）。这种几何体如今叫做正四棱台。发现这种棱台体积的方法在公元前 1850 年的所谓“莫斯科纸莎草书”中有所记载：

“如果你被告知：一个截棱棱锥体，垂直高为 6，下底边长为 4，上底边长为 2。则你取 4 的平方，得结果为 16。你将 4 加倍，得结果 8。你取 2 的平方，得结果 4。你将 16、8 和 4 相加，得 28。你取 6 的三分之一，得结果 2。你取 28 的 2 倍得 56。看，是 56。你会发现答案是正确的。”

这段描述十分精彩，确实得出了棱台体积的正确答案。但是，请注意，它的计算方法却不是普遍适用的。这种方法没有导出一个一般公式，以适用于其他尺寸的棱台。古埃及人为计算不同尺寸棱台的体积，或许不得不比照这个例子重新演算一番，而这个计算过程又让人感到有点儿混乱不清。我们现代的计算公式就简单明了得多：

$$V = \frac{1}{3}h\left(a^2 + ab + b^2\right)$$

公式中，a 为正方形下底边长，b 为正方形上底边长，h 是棱台的高。更令人遗憾的是，没有任何资料证明古埃及人的方法为什么会得出正确的答案，他们仅仅留下了简单的一句话“你会发现答案是正确的”。

从一个特殊例子引出包罗万象的结论，很可能是危险的，而历史学家注意到，在法老统治下的埃及这种独裁社会，必然会产生这种武断的数学方法。在古埃及社会，民众无条件地服从他们的君主。由此推断，当时，如果提出一种官方的数学方法，并断言“你会发现答案是正确的”，则埃及臣民是不会要求对这种方法为什么正确作出更详尽的解释的。在法老统治的土地上，民众只能惟命是听，让你怎么做就怎么做，不论是建筑巨大的庙宇，还是解答数学题，一概如此。那些敢于怀疑现体制者必然不得善终。

另一处伟大的古代文明（或者更准确地说，另几处文明）在美索不达米亚蓬勃发展，并产生了比古埃及先进得多的数学。例如，巴比伦人已能解出带有明显代数特征的复杂数学题。现存称为“普林顿”的楔形文字泥版书 322 部（写作年代大约在公元前 1900 至 1600 年之间）表明，巴比伦人已明确理解了毕达哥拉斯勾股定理，其理解深度远远超过了古埃及人。他们懂得 5-12-13 三角形或 65-72-97 三角形（或更多）都是直角三角形。除此以外，他们还为他们的数系创造了一种复杂的进位系统。当然，我们都习惯于十进位数系。显然，十进位制是从人类有十个手指引申出来的。所以，似乎有点儿奇怪的是，巴比伦人选择了 60 进位制。当然，没有人会认为这些古巴比伦人长有 60 个手指，但他们选定的 60 进位制却仍然用于我们今天的时间（一分钟 60 秒）和角度测量（在一个圆中，6×60°=360°）。

然而，美索不达米亚人的所有成就也同样只是“知其然”，而回避了更为重要的“其所以然”的问题。看来，论证数学（一种重点放在证明判

定关系上的理论演绎体系）的出现还在别一时间和别一地点。

论证数学诞生的时间是公元前 1000 年,诞生地点是小亚细亚半岛的爱琴海岸和希腊。那里出现了最伟大的历史文明，其非凡的成就对西方文化进程产生了永久性的影响。随着希腊国内和跨越地中海贸易的勃兴，希腊人逐渐成为一个流徙不定，热中冒险的民族，他们比较精明和富裕，在思想和行动上都比以往看到的西方世界更具独立性。这些充满好奇心，且思想自由的商人对权威是不会言听计从的。实际上，随着希腊民主的发展，公民自己就已成为权威（但必须强调指出，公民的定义在古希腊是非常狭隘的）。在这些人看来，对任何问题都可以自由争论，都应该加以分析，对任何观点都不能被动地、无条件地服从和接受。

到公元前 400 年时，这一卓越文明已经能以其丰富的（或许可以说是无与伦比的）智力遗产而自豪。史诗诗人荷马，历史学家希罗多德和修昔底德，剧作家埃斯库罗斯、索福克勒斯和欧里庇得斯，政治家伯里克利和哲学家索克拉蒂斯——所有这些人都在公元前四世纪初叶留下了自己的足迹。在现代社会，名望会很快衰落。因而，现代人可能惊讶，这些古希腊人的名声何以在经历了 2000 多年之后依然保持其辉煌。直至今日，我们仍然钦佩他们以深邃的理性烛照自然与人类状况的勇气。其理性虽然不乏迷信与无知，但古希腊思想家确实取得了极大的成功。即使他们的结论并非永远正确，但这些希腊人仍旧感到，他们的道路将引导自身从野蛮的过去走向梦想不到的未来。人们在描述这一特别的历史阶段时，常常使用“觉醒”一词，这是十分贴切的。人类的确已从千百万年的沉睡中醒来，以大自然最强大的武器——人类思维，勇敢地面对着这一陌生而神秘的世界。

数学当然也是如此。公元前约 600 年，在小亚细亚西海岸的小镇米利都，生活着一位伟人，即古代“七贤”之一——泰勒斯（公元前约 640—546 年）。米利都的泰勒斯是第一个在“知其然”的同时提出“其所以然”的学者，并被公认为论证数学之父。因此，泰勒斯是最早的著名数学家。

关于他的生平，我们掌握的确切资料很少。他实际上是作为一个半神话式人物从历史的薄雾中显现的，归于他名下的那些发现是否属实，人们仅仅是猜测而已。传记作家普卢塔克（公元 46-120 年）回顾了 700 年前的史迹，他写道：“……当时，泰勒斯独自将纯粹基于实践的哲学上升到理性的高度。”泰勒斯作为著名的数学家和天文学家，以某种方式预言了公元前 585 年发生的日蚀，他像所有古板的科学家一样，常常心不在焉或长时间的出神——据传说，有一次，他一边散步，一边仰望星空，竟然掉进了一口深井中。

泰勒斯虽然被公认为论证数学之“父”，但实际上，他却从未结过婚。当同代人梭伦向他追问原因时，他竟开了一个刻薄的玩笑。泰勒斯让人带给梭伦一个消息说你的儿子死了。据普卢塔克记载，梭伦当时：

“……捶胸顿足，痛不欲生，像人们遭遇不幸时惯常所做的那样。但泰勒斯拉着他的手，笑了笑说：‘梭伦，就是这些事情让我不想结婚，也不想生儿育女，这实在太难了；不过，你不必太过伤心，因为这都是假的。’”

显然，泰勒斯不是那种心地善良之辈。从农夫的故事中，我们也可以得到同样的印象。一个农夫常常要将沉重的盐袋驮在驴背上，赶着驴去集市卖盐。聪明的驴子很快就学会了在涉过一条小河时打滚，把许多盐溶化在水里，大大减轻盐袋的重量。农夫非常生气，就去请教泰勒斯。泰勒斯

建议农夫在下次赶集时，给驴驮一袋海绵。

当然，泰勒斯对人或动物的不友善，并不妨碍他在数学领域赢得很高的声望。正是泰勒斯曾极力主张，对几何陈述，不能仅凭直觉上的貌似合理就予以接受，相反，必须要经过严密的逻辑证明。这是他留给数学界的一笔相当可观的遗产。

确切地说，泰勒斯的定理究竟是什么呢？传统上认为，泰勒斯第一个证明了下列几何性质：

■ 对顶角相等。

■ 三角形的内角和等于两个直角之和。

■ 等腰三角形的两个底角相等。

■ 半圆上的圆周角是直角。

虽然我们没有任何有关泰勒斯对上述命题证明的历史记载，但我们可以推断它们的本来面目，例如上述的最后一个命题。下列证明方法选自欧几里得的《原本》第三篇第 31 命题，但它简单明了，完全可以看作是泰勒斯自己最初的证明。

定理　半圆上的圆周角是直角。

证明　以 O 为圆心，以 BC 为直径作半圆，选半圆上任意一点 A 作圆周角 BAC（图 1.4）。我们必须证明∠BAC 是直角。连接 OA，形成△AOB。由于 OB 和 OA 都是半圆的半径，长度相等，所以△AOB 是等腰三角形。因此，根据泰勒斯先前所证明的定理，∠ABO 与∠BAO 相等（或用现代术语，迭合）；我们称这两个角为α。同样，在△AOC 中，OA 与 OC 相等，因此，∠OAC=∠OCA；我们称这两个角为β。而在大三角形 BAC 中，我们看到，

2 个直角=∠ABC+∠ACB+∠BAC

=α+β+（α+β）

=2α+2β=2（α+β）

因此，一个直角 $=\frac{1}{2}$〔2个直角〕 $=\frac{1}{2}$〔2（α + β）〕 = α + β = ∠BAC。

这正是我们要证明的。证讫。[①]

泰勒斯之后，希腊又一位伟大数学家是毕达哥拉斯。毕达哥拉斯公元前约 572 年出生于萨摩斯，并在爱琴群岛东部生活和工作，甚至，据说，他还曾师从泰勒斯。但当暴君波利克拉特斯僭取这个地区的政权之后，毕达哥拉斯逃到了现今意大利南部的希腊城镇克洛托内。他在那里创办了一个学术团体，今称为毕达哥拉斯兄弟会。毕达哥拉斯哲学认为，“整数”是宇宙的要素，万物的元质。不论是音乐、天文学，还是哲学，“数”的中心地位是随处可见的。关于物理可以“数学化”地理解的现代观点在很大程度上也源自于毕达哥拉斯学派的观点。

在严格意义的数学领域，毕达哥拉斯学派为我们提供了两个伟大发现。一个当然是无与伦比的毕达哥拉斯定理。像所有远古时代的其他定理一样，我们没有关于毕达哥拉斯原论证的历史资料，但古人却一致将这一

[①] 习惯上，在证明完毕后，要写明“证讫”，其原为拉丁文 Quod erat demonstrandum（Q．E．D．），提醒读者注意，论证完毕，我们可以转向新的方向了。

定理的发现归于毕达哥拉斯的名下。据说，毕达哥拉斯曾向上帝献祭一头牛，以庆祝他的论证带给各方的喜悦（大概这头牛除外）。

但毕达哥拉斯学派的另一个重要贡献却没有得到人们的热情支持，因为它不仅公然蔑视直觉，而且还冲击了整数的优势地位。用现代说法，他们发现了无理量，但他们的论证方法却有点儿几何学的味道：

两条线段，AB 和 CD，如果有一条可均匀分割 AB 和 CD 的小线段 EF，我们就说线段 AB 和 CD 是可公度的。也就是，对于整数 p 和 q 来说，AB 是由 p 段相等于 EF 的线段组成 ;而 CD 是由 q 段同样的线段组成(见图 1.5)。因而，$\dfrac{\overline{AB}}{\overline{CD}}=\dfrac{p(\overline{EF})}{q(\overline{EF})}=\dfrac{p}{q}$（我们在这里使用了符号$\overline{AB}$，表示线段AB的长度）。由于$\dfrac{p}{q}$是两个正整数的比，我们说，可公度线段的长度比是“有理”数。

凭着直觉，毕达哥拉斯学派认为，任何两个量都是可公度的。给定两个线段，必有另一条线段 EF，可以均匀地分割这两个线段，即使取非常小的 EF，也是如此。怀疑 EF 的存在，似乎是十分荒谬的。线段的可公度性对毕达哥拉斯学派至关重要，这不仅因为他们利用这一观点证明相似三角形，而且还因为这一观点似乎可以支持他们关于整数中心作用的哲学态度。

但是，据说，毕达哥拉斯的弟子希帕萨斯发现正方形的边长与其对角线（见图 1.6 中的 GH 与 GI）却不可公度。因为不论划分多小，都没有一个 EF 量可以均匀地分割正方形的边长和对角线。

这一发现产生了许多深远的结果。显然，这个发现粉碎了毕达哥拉斯那些建立在所有线段都可公度的假设基础之上的证明。几乎 200 年之后，数学家欧多克索斯才设法在不基于可公度概念的基础上，修补了相似三角形理论。其次，这一发现还动摇了整数至高无上的地位，因为如果并非一切量都可公度，那么，整数对于表示所有线段长度的比就显得不充分了。因此，这一发现在其后的希腊数学中，建立了几何对算术的绝对优势。例如，如图 1.6 所示，正方形的边长和对角线无疑属于几何问题。如果作为数字问题来计算，则会出现严重的问题。因为，如果我们设上图正方形的边长为1，根据毕达哥拉斯勾股定理，则对角线长度为$\sqrt{2}$；由于边长与对角线不可公度，因而我们看到，$\sqrt{2}$不能写成$\dfrac{p}{q}$形式的有理数。就数字而言，$\sqrt{2}$是“无理的”，其算术性质非常神秘。希腊人认为，最好回避完全的数字处理，而全神贯注于通过简明的几何体来表达量。这种几何对算术的优势将支配希腊数学一千年。

无理数的发现所带来的最终结果是，毕达哥拉斯的信徒们对希帕萨斯引起的所有混乱大为恼怒，据说他们把希帕萨斯带到地中海深处，然后掀下水中。如果故事属实，则自由思想的危险性，由此可见，即使是在比较严肃的数学领域，也不例外。

泰勒斯和毕达哥拉斯，虽然在传奇和传统中神乎其神，但他们都是远古时代模糊而朦胧的人物。我们下面将介绍的希俄斯的希波克拉底（约公元前 440 年）则是一位比较确实的人物。事实上，我们把有据可查的最早

的数学论证归于他的名下。这将是我们所要介绍的第一个伟大定理的主题。

希波克拉底公元前 5 世纪生于希俄斯岛。当然，这是产生上述他的杰出前辈的同一个地方。（顺便提请读者注意，希俄斯岛距科斯岛不远，当时那里出生了另一位“希波克拉底”；科斯的希波克拉底（不是我们所说的希波克拉底）乃希腊的医学之父和医生遵循的《希波克拉底誓言》的创始人。）

关于数学家希波克拉底，我们对他的生平知之甚少。亚里士多德曾写过，希波克拉底虽然是一位天才的几何学家，但他“……看起来在其他方面却显得迟钝又缺乏见识”。身为数学家，却难以应付日常生活，他即是早期的这样一类人。传说，希波克拉底是在被强盗骗去钱财后出名的，显然，他被人当作了容易受骗的傻瓜。为了挽回损失，他前往雅典，并在那里教学，他是少数几位为挣钱而开始教学生涯的人之一。

无论如何，我们都不会忘记希波克拉底对几何学作出的两个非凡的贡献。其一是他编写了第一部《原本》，第一次阐述了从几个已知公理或公设中精确而有逻辑性地推导出几何定理的过程。至少，人们相信是他写了这部著作，但遗憾的是，这部书没有能够流传至今。然而，这部书不论多么有价值，与 100 年后欧几里得的煌煌巨作《原本》相比，也不免黯然失色。欧几里得的《原本》从根本上宣判了希波克拉底著作的过时。即使如此，我们仍有理由认为，欧几里得借鉴了他前辈的思想，因此希波克拉底失传的大作无疑使我们受益良多。

然而，令人欣慰的是，希波克拉底的另一个伟大贡献——求新月形面积——却流传至今，虽然大家公认，其流传是无意的和间接的。我们未能得到希波克拉底的原作，而只传有欧德摩斯公元前约 335 年对希波克拉底著作的转述；即使就转述而言，事情也不乏含混之处，因为实际上，我们也没有真正找到欧德摩斯的原著。相反，我们只看到了辛普利西乌斯于公元 530 年写的概要，他在这本书中论述了欧德摩斯的著作，而欧德摩斯则概括了希波克拉底的著作。实际上，从希波克拉底到辛普利修斯，其间经历了近一千年之久，差不多等于我们与莱弗·埃里克松之间的时间跨度，这说明历史学家在考证古代数学时遇到了多么大的困难。尽管如此，我们没有理由怀疑我们所探讨的著作基本上是可靠的。

有关求面积问题的一些评论

在探讨希波克拉底的新月形面积之前，我们先要介绍一下“求面积”的概念。显然，古希腊人被几何的对称性，视觉美和微妙的逻辑结构吸引住了。尤其令人感兴趣的是以简单和初步的东西作为复杂和纷繁问题基础的方式。这一点在下章我们探讨欧几里得定理时，就会显得十分明了。欧几里得从一些基本的公理和公设开始，一步步地推导出一些非常复杂的几何命题。

这种以简单构筑复杂的魅力还表现在希腊人的几何作图上。他们作图的规则是，所有作图都只能使用圆规和（没有刻度的）直尺。几何学家利用这两种非常简单的工具，便能够作出完美、一致的一维图形（直线）和完美、一致的二维图形（圆）——这必定出自于希腊人对秩序、简明性和

美的感受。并且，这种作图方法也是当时的技术水平所力所能及的，例如，当时还不可能画出抛物线。也许，准确地说，是直线和圆的审美魅力加强了直尺和圆规作为几何作图工具的中心地位，同时，直尺和圆规的物质可用性又转过来增进了直线和圆在希腊几何中的作用。

古代数学家利用直尺和圆规绘制了许多几何图形，但同时也受制于这两种工具。正如我们所看到的，圆规和直尺这两种似乎并不复杂的工具，掌握在聪敏的几何学家手中，便可以绘制出丰富多采和各式各样的几何图形，从平分线段和角，绘制平行线和垂直线，到创造优美的正多边形，不一而足。但是，公元前 5 世纪，更加严重的挑战却是平面图形的求面积或求方。确切地说：

□ 一个平面图形的求面积（或化其为方）就是只用圆规和直尺作出面积等于原平面图形的正方形。如果一个平面图形的求面积能够实现，我们就说这个图形是可用等价平方表示的（或可为平方的）。

求面积问题能够引起希腊人的兴趣并不奇怪。从纯粹实践的观点看，确定一个不规则图形的面积当然不是一件易事。但如果这个不规则图形能够用一个等面积的正方形替换，那么，确定原不规则图形面积的问题就变成了确定正方形面积的简单问题。

毫无疑问，希腊人对求面积问题的强烈爱好已超出了实践范围。因为如果求方能够实现，就用规则的对称性正方形替换了不规则不对称的平面图形。对于那些寻求以理性和秩序支配自然世界的人来说，这在很大程度上是一个由不对称到对称，变缺陷为完美，以有理性取代无理性的过程。在这种意义上，求面积问题就不仅是人类理性的象征，而且也是宇宙本身所固有的和谐和美的象征。

对于希腊数学家来说，探讨求面积问题是一个特别具有吸引力的课题，为此，他们作出了许多巧妙的几何图形。解数学问题，答案常常是一步一步推导出来的，求面积也是如此。第一步先要求出一个大体“规则”的图形的面积，然后再以此为基础，继续推导出更不规则，更稀奇古怪图形的面积。在这一过程中，关键性的第一步是要求出长方形面积。长方形面积的解法在欧几里得《原本》第二篇的命题 14 中就有所阐述，但我们确信，在欧几里得之前，人们便已熟知这种解法。下面，我们先从长方形面积的解法讲起。

第 1 步　求长方形面积（图 1.7）

作任意长方形 BCDE。必须只用圆规和直尺作出与 BCDE 面积相等的正方形。用直尺将线段 BE 向右延长，再用圆规在延长线上截取长度等于 ED 的线段EF——即，$\overline{EF}=\overline{ED}$。然后，等分BF于G（用圆规和直尺的一种简单作图），再如图所示，以G为圆心，以$\overline{BG}=\overline{FG}$为半径作半圆。最后，过 E 点作线段 EH 垂直于 BF，这里，H 是垂线与半圆的交点，据此，作出正方形 EKLH。

据此可以说，我们刚刚作出的图形——边长为$\overline{EH}$的正方形（图中阴影部分）与原长方形 BCDE 面积相等。

但要证明这一结论，还需要花点儿力气。为计算方便，我们设 a、b、c 分别等于线段 HG、EG 和 EH。由于所作△GEH 是直角三角形，根据勾股定理，$a^2 = b^2 + c^2$，或等价移项，$a^2 - b^2 = c^2$。显然，$\overline{FG} = \overline{BG} = \overline{HG} = a$，

因为所有这些线段都是半圆的半径。因此，$\overline{EF} = \overline{FG} - \overline{EG} = a - b$，而$\overline{BE} = \overline{BG} + \overline{GE} = a + b$。所以

面积（长方形BCDE） = （底）×（高）

$= (\overline{BE}) \times (\overline{ED})$

$= (\overline{BE} \times \overline{EF})$，因我们作图$\overline{EF} = \overline{ED}$

$=(a+b)(a-b)$，据以上推理

$=a^2-b^2$

$=c^2$=面积（正方形 EKLH）

这样，我们就证明了原长方形面积等于我们用圆规和直尺所作正方形（图中阴影部分）的面积，并以此完成了长方形的求方。

求出长方形面积后，我们很快便可进入下一步，求更加不规则图形的面积。

第 2 步　求三角形面积（图 1.8）

已知△BCD，经D点作BC的垂线，与BC相交于E。当然，我们称$\overline{DE}$为三角形的“顶垂线”或“高”。已知三角形面积等于$\frac{1}{2}$（底）×（高）$= \frac{1}{2}(\overline{BC}) \times (\overline{DE})$。如果我们平分DE于F，并作长方形，使$\overline{GH} = \overline{BC}, \overline{HJ} = \overline{EF}$。我们知道，长方形的面积等于$(\overline{HJ}) \times (\overline{GH}) = (\overline{EF}) \times (\overline{BC}) = \frac{1}{2} = (\overline{DE}) \times (\overline{BC})$ = 面积（△BCD）。然后，我们按照第一步的步骤，作正方形，并使之面积等于该长方形

的面积，因而，该正方形的面积也等于△BCD 的面积。至此，三角形的求方完成。

下面，我们将讨论一个非常一般的图形。

第 3 步　求多边形面积（图 1.9）

我们首先讨论一个非常一般的多边形，如图所示。我们通过作对角线，将这个多边形划分为三个三角形，即 B、C 和 D。因此，整个多边形的面积就等于 B+C+D。

在第 2 步中，我们已知道三角形是可用等价平方表示的，因此，我们可以分别以边长 b、c 和 d 作正方形，并得到面积 B、C 和 D（图 1．10）。然后，以 b 和 c 为直角边，作直角三角形，其斜边长为 x，即 $x^2=b^2+c^2$。我们再以 x 和 d 为直角边，作直角三角形，其斜边为 y，因而，$y^2=x^2+d^2$。

最后，我们便可以以 y 为边长作正方形（见图 1.11 阴影部分）。

综合我们的推论，就得到

$$y^2=x^2+d^2=(b^2+c^2)+d^2=B+C+D$$

因此，原多边形的面积就等于以 y 为边长的正方形的面积。

显然，这一推导过程适用于任何可作对角线将其划分为四个、五个或任何数量三角形的多边形。不论什么样的多边形（见图 1．12），我们都可以将其划分为若干三角形，并依照第 2 步的方法，作每个三角形的等面积正方形，然后，根据勾股定理，利用每一个正方形，作出大正方形，其面积即等于原多边形的面积。总而言之，多边形是可用等价平方表示的。

利用类似方法，如果一个图形的面积为两个可用等价平方表示的面积之差（而不是其和），我们可将其化为正方形。假设已知面积 E 等于面积 F 与 G 之差，并且，我们已作出边长为 f 和 g 的正方形，如图 1．13 所示。然后，我们可作直角三角形，使其斜边等于 f，直角边等于 g 和 e。最后，以边长 e 作正方形。即

$$\text{面积（正方形）}=e^2=f^2-g^2=F-G=E$$

因此，面积 E 也同样可用等价平方表示。

希波克拉底时代的希腊人利用上述方法可以将杂乱无章的不规则多边形变为等面积正方形。但是，这一成就却因一个明显的事实而减色不少，即这些图形都是直线图形——它们的边虽然数量众多，并构成各种奇形怪状的角度，但都只是直线。而更严重的挑战是，曲边图形（即所谓曲线图形）是否也可以用等价平方表示。起初，人们认为，这似乎是根本不可能的，因为显然没有办法用圆规和直尺将曲线拉直。因此，当希俄斯的希波克拉底于公元前 5 世纪成功地将一种称为“新月形”的曲线图形化为正方形时，世人惊得目瞪口呆。

伟大的定理：求新月形面积

新月形是一种边缘为两个圆弧的平面图形——即月牙形。希波克拉底并没有作出所有新月形的等面积正方形，而只求出了一种他所精心构造的特定新月形的面积。（犹如“后记”中所述，这种区别似乎造成了后人对希腊几何的误解。）希波克拉底的论证是建立在三个初步公理之上的：

■ 勾股定理

■ 半圆上的圆周角是直角

■ 两个圆形或半圆形面积之比等于其直径的平方比。

$$\frac{\text{面积（半圆1）}}{\text{面积（半圆2）}}=\frac{d^2}{D^2}$$

前两个公理在希波克拉底之前很久便已为人所知。而最后一个命题却十分复杂。两个圆形或半圆形面积之比是基于以其直径为边长所作的两个正方形面积之比的（见图 1.14）。例如，如果一个半圆的直径是另一个半圆的 5 倍，则第一个半圆的面积是第二个半圆面积的 25 倍。然而，这一命题却给数学史家提出了一个问题，因为人们普遍怀疑希波克拉底是否确曾对此作出过正确的证明。他尽可认为他能够证明这一命题，但现代学者普遍认为，这一定理（后来被列入欧几里得《原本》第七篇的第二命题）所

提出的逻辑难题远不是希波克拉底所能够解决的。（这一定理的求导过程见第四章。）

我们暂且抛开这个问题不谈，先来看一看希波克拉底的证明。首先，以O为圆心，以$\overline{AO}=\overline{OB}$为半径作半圆，如图1．15所示。作OC垂直于AB，且与半圆相交于C，并连接AC与BC。平分AC于D，然后，以D为圆心，以AD为半径作半圆AEC，这样，就形成了新月形AECF，如图中阴影部分所示。

希波克拉底的证明方法既简单又高明。首先，他必须证实所论证的新月形与图中阴影部分的△AOC面积完全相等。这样，他就可以应用已知的三角形能表示为等价平方的公理来断定新月形也可用等价平方表示。这一经典论证的详细过程如下：

定理：新月形AECF可用等价平方表示。

证明；由于∠ACB内接于半圆，所以，∠ACB是直角。根据“边角边”全等定理，三角形AOC和BOC全等，因此，$\overline{AC}=\overline{BC}$。然后，我们应用勾股定理，就得到

$$(\overline{AB})^2=(\overline{AC})^2+(\overline{BC})^2=2(\overline{AC})^2$$

因为AB是半圆ACB的直径，AC是半圆AEC的直径，所以，我们可以应用上述第三条原理，即得到

$$\frac{\text{面积（半圆AEC）}}{\text{面积（半圆ACB）}}=\frac{(\overline{AC})^2}{(\overline{AB})^2}=\frac{(\overline{AC})^2}{2(\overline{AC})^2}=\frac{1}{2}$$

也就是说，半圆AEC的面积是半圆ACB面积的一半。

我们现在来看扇形AFCO（“扇形”是圆的四分之一）。显然，这一扇形也是半圆ACB面积的一半，据此，我们可直接得出

面积（半圆AEC）=面积（扇形AFCO）

最后，我们只需从这两个图形中各自减去它们共同的部分AFCD，如图1.16所示，即

面积（半圆AEC）—面积（AFCD部分）

=面积（扇形AFCO）—面积（AFCD部分）

我们从图中可以很快看出，剩下的部分就是

面积（新月形AECF）=面积（△ACO）

我们已知，我们可以作一个正方形，使其面积等于三角形ACO，因而也等于新月形AECF的面积。这就是我们所寻求的化新月形为方的问题。证讫。

这的确是数学上的一大成就。评注家普罗克洛斯（公元410—485年）以他五世纪的眼光，认为希俄斯的希波克拉底“……作出了新月形的等面积正方形，并在几何学中做出过许多其他发现，是一位作图的天才，如果曾经有过这种天才的话。”

后记

由于希波克拉底求新月形面积的成功，希腊数学家对求最完美的曲线

图形——圆的面积充满了乐观。古希腊数学家为解决化圆为方问题付出了大量的时间和精力，一些后世作家认为希波克拉底自己曾尝试解决这一难题，尽管接二连三的评论、注释把事情弄得扑朔迷离，要确定这点很困难。五世纪的辛普利西乌斯在其著作中引述了他的前辈——阿弗罗狄西亚的亚历山大（约公元 210 年）的话说，希波克拉底曾声称他能够求出圆的面积。将这些蛛丝马迹连缀起来，我们推测亚历山大考虑的是这样一种论证：

首先作任意圆，其直径为 AB。以 O 为圆心作大圆，使其直径 CD 等于 AB 的两倍。利用已知方法，在大圆中作内接正六边形，即使六边形的每一条边都等于半径。也就是

$$\overline{CE}=\overline{EF}=\overline{FD}=\overline{DG}=\overline{GH}=\overline{HC}=\overline{OC}$$

重要的是，我们应注意到，这六条边，每一条边都等于大圆的半径，也就是说，每一条边都等于$\overline{AB}$。然后，以这六条边为直径，分别作六个半圆，如图 1.17 所示。这样，就形成了六个新月形和一个以 AB 为直径的圆（见图中阴影部分）。

然后，我们想象将右边的图形按两种方式分解：其一，看作是一个正六边形 CEFDGH 加上六个半圆；其二，看作一个大圆加六个新月形。显然，这两种方式所得出的总面积是相等的，因为都是从同一个图形中分解出来的。但是，六个半圆可以合成三个整圆，而且，每个圆的直径都等于$\overline{AB}$。

因此

面积（正六边形）+3 面积（以 AB 为直径的圆）

=面积（大圆）+面积（6 个新月形）

由于大圆的直径等于小圆的两倍，因而，大圆的面积必定等于小圆面积的 $2^2=4$ 倍。即

面积（正六边形）+3 面积（以 AB 为直径的圆）

=4 面积（以 AB 为直径的圆）+面积（6 个新月形）

从等式两边分别减去“3 面积（以 AB 为直径的圆）”，我们就得到

面积（正六边形）=面积（以 AB 为直径的圆）+面积（6 个新月形）

或

面积（以 AB 为直径的圆）=面积（正六边形）—面积（6 个新月形）

据亚历山大所说，希波克拉底作了如下推论：正六边形作为多边形，可以用等价平方表示；根据前边论证，每一个新月形也同样可以用等价平方表示。利用加法，我们可以作出一个面积等于六个新月形面积之和的正方形。因此，以 AB 为直径的圆的面积可以按照我们前面所列等式，用简单的减法即可得到。

但是，正如亚历山大随即指出的那样，这一论证有一个明显的缺点：希波克拉底在这一定理中求其面积的新月形不是沿着内接正六边形的边长作的，而是沿着内接正方形的边长作的。也就是说，希波克拉底从来没有提出过求本例这种新月形面积的方法。

大多数现代学者都怀疑像希波克拉底这样水平的数学家会犯这种错误。相反，很可能是亚历山大，或辛普利西乌斯，或任何其他转述者在介绍希波克拉底最初的论证时，在某种意义上曲解了他的原意。我们也许永远不会知道全部真相。然而，这种推理方法似乎也支持了一种看法，即化

圆为方应该是可能的。如果说上述论证没有完成此事，那么，只要再付出一点儿努力，再多一点儿洞察力，也许就可以成功了。

但是，情况却并非如此。一代又一代的人经过数百年的努力，始终未能化圆为方。历经种种曲折，人们提出了无数的解法，但最后发现，每一种解法都有错误。逐渐地，数学家们开始怀疑，也许根本不可能用圆规和直尺作出圆的等面积正方形。当然，缺乏一种正确的证明方法，即使经过了 2000 年的努力，也依然不表明化圆为方是不可能的；也许，数学家只是不够聪明，还没有找到一条穿越几何丛林的道路。此外，如果化圆为方不可能的话，就必须借助其他逻辑严密的定理来证明这一事实，而人们亦不清楚如何作出类似证明。

应当指出一点，没有人会怀疑，已知一个圆，必然存在着一个与之面积相等的正方形。例如，已知一个固定的圆和圆旁一个正方形投影小光点，并且，正方形投影的面积大大小于圆的面积。如果我们连续移动投影仪，使之距离投影屏面越来越远，并以此逐渐扩大正方形投影的面积，这样，我们最终会得到一个面积超过圆面积的正方形。根据“逐渐扩大”的直观概念，我们可以得出正确结论，在过程中的某一瞬间，正方形面积恰好等于圆形面积。

但是，这毕竟有点儿离题。不要忘记，关键的问题不是是否存在这样一个正方形，而是是否可以用圆规和直尺作出这个正方形。这就出现了困难，因为几何学家只限于使用这两种特定工具；而移动投影光点显然违反这一规则。

从希波克拉底时代直到一百多年前，化圆为方问题始终未能解决。终于，1882 年，德国数学家费迪南德·林德曼（1852—1939 年）成功而明确地证明了化圆为方是根本不可能的。其证明的技术性细节非常高深，远远超出了本书的范围。但是，从下面的概要中，我们仍然可以看到林德曼是如何解答这一古老问题的。

林德曼解决这一难题的方法是将问题从几何王国转向数字王国。只要我们想象所有实数的集合（如图 1．18 中大长方形所包括的范围），我们就能够将它们再划分为两个穷举且相互排斥的类型——代数数和超越数。

根据定义，如果一个实数满足下述代数方程

$$a_n x^n + a_{n-1} x^{n-1} + \cdots\cdots + a_2x^2 + a_1x + a_0 = 0$$

那么，这个实数是代数数。方程中所有系数，a_n，a_{n-1}，……，a_2，a_1和a_0都是整数。因此，有理数$\frac{2}{3}$是代数数，因为它是多项式方程$3x-2=0$的解；无理数$\sqrt{2}$也是代数数，因为它可以满足方程$x^2-2=0$；甚至$\sqrt[3]{1+\sqrt{5}}$也是代数数，因为它可以满足方程$x^6-2x^3-4=0$。注意，在这几个例子中，每一个多项式的系数都是整数。

用不太正规的话说，我们可以认为，代数数是我们在算术和初等代数中遇到的“容易”或“熟悉”的量。例如，所有整数都是代数数，所有分数，及其平方根、立方根等，也都是代数数。

相反，如果一个数不是代数数，那么，就必然是超越数——也就是说，这个数不是任何带有整数系数的代数方程的解。超越数与其比较简单的代

数数亲族相比，要复杂得多。根据定义，显然，任何实数不是代数数，就是超越数，但不可能两者兼之。这就是严格的二分法，犹如一个人不是男的，就是女的，决没有中性可言。

下面，我们将先讨论单位长度（即代表数字“１”的长度），并以此为基础，进一步讨论我们能够用直尺和圆规作出的其他长度。情况表明，所有可构造线段长度的总和，虽然庞大，但却不可能包括每一个实数。例如，从长度 1 开始，我们可以作出长度 2、3、4，等等，也能作出有理长度，如$\frac{1}{2}$、$\frac{2}{5}$、$\frac{13}{711}$，甚至还能作出无理长度，且不只限于平方根，如$\sqrt{2}$或$\sqrt{5}$。而且，如果我们能够作出两个长度，我们就能够作出这两个长度的和、差、积和商。把所有这些作图集合在一起，我们可以看到，更加复杂的表达式，如

$$\sqrt{\frac{6-2\sqrt{2}}{1+\sqrt{4+\sqrt{23-\sqrt{7}}}}}$$

就是实际的可构造长度。

这些大量的可构造数就构成了代数数的子集，就像所有秃头男人的集合构成了所有男人的子集一样。如图 1．18 所示，这些可构造数严格隶属于代数数。重要的问题是，没有一个超越数能够用圆规和直尺作出。（如果把我们的比喻再扩大一步，那么，这后一句话的意思就是，没有一个女人会隶属于秃头男人之列。）

在林德曼开始着手研究化圆为方的问题时，所有这些知识都已为人所知。在其前辈、特别是在法国卓越数学家夏尔·埃尔米特（1822—1901 年）努力的基础上，林德曼攻克了著名的数字π。（在初等几何中，我们见到的π是作为圆的周长与直径的比；我们在第四章中将详尽论述这一重要的常数。）林德曼的成就是证明了π是超越数。也就是说，π不是代数数，因此，不可构造。同时，也告诉我们，$\sqrt{\pi}$也不可构造，因为如果我们能够构造$\sqrt{\pi}$，那么，只要我们再努一把力，也就能够用圆规和直尺作出π的图形。

乍看之下，这一数字上的发现对于化圆为方的几何问题似乎没有多大关系，但是，我们将看到，这一发现为这一古老难题补上了缺失的一环。

定理　化圆为方是不可能的

证明　为了形成最后的矛盾，让我们先假设圆能够化为方。我们可以很容易地用圆规作一个圆，使半径 r=1。因此，这个圆的面积就是$\pi r^2=\pi$。如果按照我们的假设，圆能够化为方，于是，我们便非常兴奋地用圆规和直尺猛砍圆弧，并画上直线。我们只需经过这样有限的几次，最后就终于得到了一个面积也是π的正方形，如图 1.19 所示。在这一过程中，我们构造了正方形，当然也就构造了它的四条边。我们设正方形的边长为 x。于是，我们看到

π=圆面积=正方形面积=x^2

因此，长度$x=\sqrt{\pi}$就应该能够用圆规和直尺构造。但是，如前所述，$\sqrt{\pi}$是无法构成的。

究竟错在哪里了呢？我们再回头看一看整个论证过程，以找出产生这一矛盾的原因。我们发现，问题只能出在最初的假设上，也就是圆能够化为方的假设，结果，我们必须否定这个假设，并据此得出结论，化圆为方在逻辑上是根本不可能的！证讫。

林德曼的发现表明，从希波克拉底时代直到现代数学家对化圆为方这一难题的刻意探索，实际上是徒劳的。从化新月形为方开始，所有有启发性的证明，所有有希望的线索，到头来都成了虚幻镜影。只使用圆规和直尺是不足以化圆为方的。

那么，历史对新月形求方又作如何评价呢？上述伟大定理表明，希波克拉底成功地作出了一种特定新月形的等面积正方形，并努力探求另外两种新月形的求方。因而，到公元前 440 年时，三种类型的新月形化方，已为众人所知。但从此便停滞在这一水平，两千多年没有进展。直到 1771 年，伟大的数学家莱昂哈德·欧拉（1707—1783 年）（我们将在第九章和第十章中详细介绍）才发现了另外两种可以用等价平方表示的新月形。此后，直到 20 世纪，N．G．切巴托鲁和 A．W．多罗德诺才证明出这五种新月形是唯一可用等价平方表示的新月形！所有其它类型的新月形，包括我们前面讲到的曾引起亚历山大尖锐批评的那种新月形，都像圆形一样，不可能化为等价正方形。

因此，希波克拉底及其新月形的故事便就此划上了句号，而且，这是一个相当曲折反复的故事。起初，直觉认为，不可能用圆规和直尺作出曲线图形的等价正方形。但是，希波克拉底通过新月形求方将直觉颠倒过来，并继续寻求更多可用等价平方表示的曲线图形。然而，最后，林德曼、切巴托鲁和多罗德诺的否定结论表明，直觉并非一无是处。曲线图形的求方远非规范，而必定永远只是例外。

第二章
欧几里得对毕达哥拉斯定理
（勾股定理）的证明
（公元前约 300 年）

欧几里得的《原本》

从希波克拉底到欧几里得，其间经历了 150 年。在这 150 年间，希腊文明发展并臻于成熟，因柏拉图、亚里士多德、阿里斯托芬和修昔底德的著作而光大。甚至在伯罗奔尼撒战争的动乱中和在亚历山大大帝统治的希腊帝国全盛时期，希腊文明都在发展。到公元前 300 年时，希腊文化的发展已跨越地中海，并扩展到更遥远的世界。在西方，希腊统治至高无上。

在从公元前 440 年到公元前 300 年期间，许多伟人都曾为数学的发展作出过不朽的贡献，其中有柏拉图（公元前 427—347 年）和欧多克索斯（公元前约 408—355 年），虽然只有后者才是真正的数学家。

柏拉图，雅典的伟大哲学家。我们之所以提到他，主要不是因为他对数学的创造，而是因为他对数学的热情和高度评价。柏拉图年轻时在雅典师从苏格拉底，我们对他那位值得尊敬的老师的了解，主要也由此而来。柏拉图曾漫游世界多年，认识了许多伟大的思想家，并形成了他自己的哲学思想体系。公元前 387 年，他返回他的出生地雅典，并在那里建立起学园。学园聚集了不少饱学之士来此献身于学习和研究。在柏拉图的引导下，希腊学园成为那个时代一流的思想中心。

在学园众多的学科中，没有一个学科能比数学更受重视。数学的美感和条理与秩序吸引了柏拉图，代表了他心目中未受单调日常生存需求污染的理想的抽象世界。柏拉图认为，数学是锻炼思维的最佳途径，其严密的逻辑推理要求人们极度专注、机敏和谨慎。据说，穿过拱形门楼，进入这一久负盛名的学园，首先映入眼帘的是一行大字：“ 不懂几何的男子请勿入内 ”。尽管这一警句带有明显的性别歧视，但却反映了一种观点，即只有那些首先证明自己在数学上成熟的人才有能力面对学园的智力挑战。可以说，柏拉图把几何学看作理想的入学要求，看作一种当时那个时代的学术能力测验。

虽然现在人们很少把当初的数学发现归于柏拉图的名下，但希腊学园的确培养了许多颇有才华的数学家，其中一个无可争辩的伟大数学家就是尼多斯的欧多克索斯。欧多克索斯在学园创建初期就来到雅典，并直接聆听过柏拉图的演讲。欧多克索斯的贫困迫使他不得不居住在雅典的郊区比雷埃夫斯，每日往返于学园和比雷埃夫斯之间，成为最早的通勤者（虽然我们不能确切知道，他是否需要支付远郊车费）。在他后来的生涯中，他曾到过埃及，后来又返回他的出生地尼多斯。在这期间，他注意吸收新的科学发现，并不断扩充科学的疆界。欧多克索斯对天文学尤其感兴趣，他对月球和行星的运动做出了深入的解释，在 16 世纪哥白尼革命之前，其学说颇有影响。他从不接受对自然现象的天命的或神秘的解释，相反，他主张对自然现象进行观察，并作出理性的分析。因此，托马斯 · 希思爵士曾

称道说：“ 如果当时有科学家的话，他称得起是其中一个。 ”

据认为，欧多克索斯对数学作出了两大贡献，其一是比例论，其二是穷竭法。毕达哥拉斯派曾因发现不可通约量而陷入绝境，而欧多克索斯的比例论则对走出这种绝境提供了逻辑依据。毕达哥拉斯派的绝境在有关相似三角形的几何定理中尤为明显，这些定理最初是根据一种假设论证的，即任何两个量都是可公度的。当这一假设被推翻后，几何学中一些最重要的定理也随之瓦解。这就是人们有时所谓的希腊几何的“ 逻辑耻辱 ”。也即，人们虽然相信这些定理是正确的，但他们却拿不出有力的证据来支持他们的观点。正是欧多克索斯发明的比例论为人们提供了这一长期寻觅的证据。他的理论自然使希腊数学界人士如释重负。我们如今可以在欧几里得的《原本》第五篇中找到欧多克索斯的理论。

欧多克索斯的另一个伟大贡献，即穷竭法，可以直接应用于确定更加复杂几何图形的面积和体积。他所采用的一般方法是，用一系列已知的基本图形不断逼近不规则图形，而每一次逼近，都比前一次更加近似于原图形。例如，我们可以认为，圆形是包含在一曲线里的图形，因而也是一种非常难于解出的平面图形。但是，如果我们在圆内作一个内接正方形，然后再把正方形的每条边一分为二，使之成为八边形，再把八边形的每条边平分，使之成为 16 边形，等等，依次进行，我们就可以得到一个非常近似于圆形面积的比较简单的多边形。用欧多克索斯的话说，这个多边形从内部“ 穷竭 ”了圆。

实际上，这个过程就是阿基米德确定圆面积的过程，我们将在第四章看到。阿基米德不仅将这一基本逻辑理论归功于欧多克索斯，而且还认为他用穷竭法证明了“ 任何锥体的体积都等于与之同底同高柱体体积的三分之一 ”，这决不是一个无足轻重的定理。熟悉高等数学的读者都会承认，穷竭法是现代“ 极限 ”概念的几何先驱，同时也是微积分的中心。欧多克索斯的贡献意义十分深远，人们一般认为他是仅次于最伟大的数学家阿基米德本人的古希腊卓越数学家。

公元前四世纪的最后 30 年，马其顿国王亚历山大大帝即位，并出发征服世界。公元前 332 年，亚历山大大帝征服埃及，随之在尼罗河口建亚历山大城。这座城市发展极为迅速，据说在其后 30 年间，人口已达 50 万。而更为重要的是，在这座城市中建立了宏大的亚历山大图书馆，这座图书馆很快便取代了希腊学园，成为世界的学术重镇。亚历山大图书馆藏有 600，000 多部纸莎草纸文稿，其藏书之丰超过了当时世界上的任何一个图书馆。的确，在整个希腊和罗马统治时期，亚历山大城始终是地中海地区的思想中心，直到公元 641 年被阿拉伯人摧毁。

公元前约 300 年，在亚历山大城吸引的众多学者中，有一位名叫欧几里得。他来到亚历山大城，创办了一所数学学校。我们对欧几里得的生平和他到达非洲海岸前后的情况都知之甚少，但他似乎曾在希腊学园接受过柏拉图弟子的训导。不管情况是怎样的，欧几里得的影响十分深远，实际上，所有后来的希腊数学家都或多或少地与亚历山大学校有过某种联系。

欧几里得在数学史上声名显赫，得益于他编纂的《原本》。这部著作对西方思想有着深远的影响，人们一个世纪又一个世纪地研究、分析和编辑此书，直至现代。据说在西方文明的全部书籍中，只有《圣经》才能够与欧几里得的《原本》比美。

得到人们高度评价的《原本》是一部大型汇编书籍，全书分为 13 篇，465 个命题，其涉及范围，从平面几何、立体几何到数论，无所不包。今天，人们一般认为，在所有这些定理中，只有比较少的一部分是欧几里得本人创立的。尽管如此，但从整个希腊数学体系来看。他毕竟创造了一个数学宝库，它是如此的成功，如此的受人尊崇，以致于所有前人的类似著作都相形见绌。欧几里得的著作很快就成为了一种标准。如此一来，如果一个数学家说到 1．47，就只能意为《原本》第一篇第 47 命题，而无须解释我们所说的是《原本》，犹如人们一提到“Ⅰ《列王记》7：23”，就知道说的是《圣经》一样。

实际上，这种比较是非常恰当的，因为没有一本书能像欧几里得的大作那样被人看作“数学的圣经”。几百年来，《原本》已出版了 2000 多个版本，这个数字足以使今天数学教科书的编写家羡慕不已。众所周知，即使在当时，《原本》也获得了巨大的成功。罗马帝国崩溃后，阿拉伯学者将《原本》带到了巴格达。文艺复兴时期，《原本》再度出现于欧洲，其影响十分深远。16 世纪的意大利著名学者及 100 年后年轻的剑桥大学学生艾萨克·牛顿都曾拜读过这部巨著。下面，让我们从卡尔·桑德堡著的亚伯拉罕·林肯传中摘录一段，看一看没有受过什么正规教育的年轻律师林肯是如何磨砺他的推理技能的：

“……购买一部欧几里得的《原本》，这部书已有 2300 年的历史……（他）在外出巡回出庭时，把书装在他的旅行袋里。晚上……别人都已入睡了，他还在借着烛光研读欧几里得。”

人们屡屡提及，林肯阅读莎士比亚和《圣经》，文风受到很大影响。同样，他的许多政论文也明显地反映出欧几里得命题的逻辑发展。

伯特兰·罗素（1872—1970 年）对《原本》情有独钟，他在自传中写下了这样一段引人注目的回忆：

“11 岁时，我开始学习欧几里得的书，并请我的哥哥当老师。这是我生活中的一件大事，犹如初恋般的迷人。”

我们在本章和下一章讨论《原本》时，应该知道，我们是在沿着一条其他许多人业已走过的道路前进。只有极少数的一些经典著作，如《伊利亚特》和《奥德赛》，才有资格共同组成这一文化遗产。我们将要讨论的命题，阿基米德、西塞罗、牛顿、莱布尼兹、拿破仑和林肯都曾研究过。侧身于这一长长的学生名单之中，不免令我们有些忐忑不安。

欧几里得天赋超人，与其说他创造了一种新的数学，不如说他把旧数学变成一种清晰明确、有条不紊、逻辑谨严的新数学。这绝不是无足轻重的小事。必须认识到，《原本》绝不仅仅只是数学定理及其证明；早至泰勒斯时代，数学家就已对命题作出过论证，而欧几里得对命题作了辉煌的公理化演绎，这是一个根本的区别。在《原本》中，他首先给出要素：23 条定义，5 条公设和 5 个公理。这些都是基础，是欧几里得体系的“已知”。他可以在任何时候应用这些要素。利用这些要素，他证明了他的第一个命题。然后，以第一个命题为基础，他可以将他的定义、公设、公理与第一个命题都融合进对第二个命题的证明。如此循序渐进，直至逐条证明所有的命题。

因此，欧几里得不仅仅作出了证明，更重要的是，他是在这种公理结构中作出的证明。这种论证方法的优越性十分明显，其一就是可以避免循

环推理。每一个命题都与前一个命题有着十分清晰而明确的联系，并可直接导回原来的公理。懂得计算机的人甚至还能够画出一张流程图，准确显示证明一个特定定理可以应用哪些推导结果。这种证明方法比“投入”法优越得多，因为使用“投入”法，人们总是不清楚以前的哪些推导结果可以应用，哪些不可以应用。而且，在推导过程中，还有一个很大的危险，就是，如果要证明定理 A，可能需要应用结果 B，但反过来，如果不应用定理 A 本身，可能又无法证明结果 B。这样，就出现了自我相关的“怪圈”，犹如一条蛇吞吃了自己的尾巴。在数学上，显然徒劳无益。

除此以外，公理化还有另一个优点。由于我们能够明确判别任何命题的前一个命题，因此，如果我们需要改变或消除某一基本公设，我们就能够立即觉察出可能会出现哪些情况。例如，如果我们没有应用公设 C 或根据公设 C 证明的任何结果，就证明出了定理 A，那么，我们可以断言，即使消除公设 C，定理 A 依然正确。这看起来似乎有点儿深奥，但在存有争议的欧几里得第 5 公设中，恰恰出现了这样的问题，引起了数学史上一次持续时间最长、意义最深远的辩论。我们将在本章的“后记”中详细讨论这一问题。

因此，《原本》的公理化演绎方法是非常重要的。虽然欧几里得没有使之尽善尽美，但它的逻辑极为严密，而且，欧几里得成功地将零散的数学理论编为一个从基本假定到最复杂结论的连续网络，所有这些，都使之成为其后所有数学著作的范本。时至今日，在神秘的拓扑学、抽象代数或泛函分析领域，数学家们还是首先提出公理，然后，一步一步地推导，直至建立他们奇妙的理论。而这正是欧几里得谢世 2300 年后的再现。

第一篇：序

在本章中，我们只重点讨论《原本》的第一篇；其后几篇，我们将在第三章讨论。第一篇一开始就提出了一系列互不连贯的平面几何定义。（欧几里得的全部引文均摘自托马斯·希思编辑的百科全书中“欧几里得《原本》十三篇”。）其中一些定义如下：

□ 定义 1　点是没有部分的一种东西。

□ 定义 2　线是没有宽度的长度。

□ 定义 4　直线是其上各点无曲折地排列的线。

欧几里得今天的学生会发现这些定义的措词都是不可接受的，而且，还多少有点儿古怪。显然，在任何逻辑系统中，并非每个名词都是可以定义的，因为定义本身又是由其它名词组成的，而那些名词也必须定义。如果一个数学家试图对每个概念都给出定义，那么，人们一定会批评他在制造一个庞大的循环论证的怪圈。例如，欧几里得所说的“没有宽度”究竟是什么意思？而“各点无曲折地排列”的技术性含义又是什么？

从现代观点来看，一个逻辑系统总是始于一些未经定义的词，而以后所有的定义都与这些词有关。人们肯定会尽力减少这些未定义词的数量，但这些词的出现却是不可避免的。对于现代几何学家来说，“点”和“直线”的概念就始终未经定义。像欧几里得所用的陈述，有助于我们在头脑中形成某些图像，并非完全没有益处；但是，作为准确的逻辑定义来说，这最初的几个词是不能令人满意的。

所幸他后来的定义却比较成功。其中一些在我们第一篇的讨论中非常突出，值得予以评述。

□ 定义 10　一条直线与另一条直线相交，如果两个邻角相等，则这两个邻角都是直角，而与另一条直线相交的直线叫做那条直线的垂线。

现代读者可能会对此感到奇怪，欧几里得并没有将直角定义为 90°角；实际上，在《原本》中，也没有任何一个地方讲到“度”是角的测量单位。在这部书中，唯一有意义的角测量是直角。正如我们所看到的那样，欧几里得将其定义为一条直线上两个相等的邻角之一。

□ 定义 15　圆是包含在一条线里的平面图形，因此，从圆内某一点出发连到该线的直线都相等。

显然，圆内的“某一点”是指圆心，而他所说的相等的“直线”则是半径。

欧几里得在定义 19 至 22 中，定义了三角形（由三条直线包含的平面图形）、四边形（由四条直线包含的平面图形）和一些特定的子类，如等边三角形（三条边都相等的三角形）和等腰三角形（“只有两条边相等”的三角形）。他最后的定义是十分重要的：

□ 定义 23　平行直线是两条在同一平面且向两个方向无限延伸的直线，这两条直线在两个方向上不相交。

请注意，欧几里得避免了用“处处等距”的术语来定义平行线。他的定义更为简单，而且少有逻辑陷阱：平行线只是在同一平面且不相交的直线。

基于这些定义，欧几里得提出了五个几何公设。请不要忘记，这些都是欧几里得体系中的“已知”，是不言自明的真理。他当然对此必须审慎地选择，以避免重叠或内在的不一致。

公设 1　从任一点到任一点〔可〕作一条直线。

公设 2　有限直线〔可〕沿直线无限延长。

我们即刻可以看出，这前两个公设恰好可以允许我们用无刻度直尺作图。例如，如果几何学家想用一条直线连接两点（这正是可以用直尺完成的作图），则公设 1 为此提供了逻辑依据。

公设 3　给定中心和距离（半径），〔可以〕作一个圆。

这样，公设 3 就为以已知点为圆心，以已知距离为半径，用圆规作圆提供了相应的逻辑根据。因此，我们可以说，这前三个公设加在一起，就为欧几里得作图工具的全部用途奠定了理论基础。

是否确实如此呢？人们只要回想一下自己的几何作图训练，就会想起圆规的另一个用途，即用以将平面上某一部分的固定长度转移到另一部分。具体做法是，已知一条线段，拟在另一处复制其长度。将圆规的尖端放在线段的一端，并将圆规的铅笔端对准线段的另一端；然后，将圆规固定，并拿起圆规，放在需要复制线段的位置。这是一种非常简便，又非常

有用的方法。但是，按照欧几里得的规则，这种方法却是不允许的，因为在他的著作中，没有一个地方提出一种公设，允许用这种方法转移长度。因此，数学家们常常称欧几里得的圆规是“可折叠的”。就是说，虽然圆规完全有能力作圆（如公设 3 所保证的），但只要把圆规从平面拿起，圆规就闭拢了，无法再打开。

造成这种情况的原因究竟是什么？欧几里得为什么不再增加一条公设，以支持这一非常重要的转移长度的方法呢？答案十分简单：他不需要假定这样一种方法作为公设，因为他证明出了这种方法，并将其作为第一篇的第三命题。也就是说，虽然欧几里得的圆规一从纸上拿起来变成“可折叠”的了，但他的确提出了一种十分巧妙的转移长度的方法，并证明了他的方法为什么奏效。欧几里得令人仰慕之处就在于，他尽力避免假定他实际上能够推导出来的公设，因而使他的公设的数目少而且精。

公设 4　所有直角都相等。

这一公设与作图无关，它提供了一个贯穿于整个欧几里得几何体系的统一的比较标准。定义 10 引入了直角概念，而现在，欧几里得则假定任何两个直角，不论在平面的什么位置，都相等。基于这一公设，欧几里得提出了一个在希腊数学界引起最大争议的公设：

公设 5　如果一条直线与两条直线相交，且如果同侧所交两内角之和小于两个直角，则这两条直线无限延长后必将相交于该侧的一点。

如图 2.1 所示，这一公设的意思是说，如果 $\alpha+\beta$ 小于两个直角，则直线 AB 与 CD 相交于右侧。公设 5 常常被人们称为欧几里得的平行线公设。这显然有点儿用词不当，因为实际上这一公设规定了使两条直线相交的条件，因此，根据定义 23，更准确的名称应

该叫不平行公设。

显然，这一条公设与其它公设完全不同。它的行文较长，而且需要有图帮助理解，似乎远不是那种不证自明的真理。这条公设看来过于复杂，与泛泛而谈的“所有直角都相等”显然不属同一类。实际上，许多数学家都直觉地感到这第 5 条公设实际上是一个定理。他们认为，正如欧几里得不需要假定可用圆规转移长度，他也不需要假定这样一条公设，他完全可以借助更基本的几何性质证明这一点。有证据表明，欧几里得自己也对这个问题感到有点儿不安，因为他在第一篇的演绎中一直尽力避免应用这一平行线公设。也即，在最初的 28 个命题中，既然他感到完全可以首先和经常使用其他公设，他就放弃了使用第 5 条公设。但诚如“后记”中所表明，怀疑是否需要这一公设是一回事，作出实际证明则是另一回事。

根据这一有争议的公设，欧几里得提出了五个公理，从而完了他的序篇。这五个公理也都是不证自明的真理，但具有更一般的性质，不仅仅只对几何学有效。这些公理是

- ☐ 公理 1　与同一个东西相等的东西，彼此也相等。
- ☐ 公理 2　等量加等量，总量仍相等。
- ☐ 公理 3　等量减等量，余量仍相等。
- ☐ 公理 4　彼此重合的东西相等。
- ☐ 公理 5　整体大于部分。

在这五个公理中，只有第 4 个公理有点儿让人费解。显然，欧几里得的意思是，如果一个图形能够严格不变地从纸上某一位置拿起，放到第二个图形上，两个图形完全重合，则两个图形在各个方面都相等——即它们有相等的角，相等的边，等等。长期以来，人们认为，公理 4 具有某种几何特征，应该归入公设的范围。

所有这些就是整个《原本》大厦建筑其上的假设陈述的基础。现在，让我们再来看一看青年伯特兰·罗素在其自传中的另一段回忆：

“我听说欧几里得证明了一些定理，但看到他从公理入手，感到非常失望。起初，我拒绝接受这些，除非哥哥讲明这样做的道理，但他说，‘如果你不接受它们，我们就无法继续。’我为了能继续学习，勉强接受了它们。”

第一篇：早期命题

在《序》的基础上，欧几里得开始证明他第一篇中的前 48 个命题。我们在此只讨论那些特别有趣或特别重要的命题，目标是要到达命题 Ⅰ. 47 和Ⅰ. 48，因为这两个命题是第一篇的逻辑顶峰。

如果一个人想从一些特定公理开始演绎几何，那么，他的第一个命题应该是什么呢？对于欧几里得来说，这第一个命题就是

命题Ⅰ.1　在已知有限直线上作等边三角形。

证明　欧几里得开始先作已知线段 AB，如图 2. 2 所示。然后，他以 A 为圆心，以 AB 为半径，作圆；再以 B 为圆心，以 AB 为半径，作第二个圆。当然，这两个圆都应用了公设 3，而且，在从纸上拿起圆规时，不要求圆规保持打开状态。设 C 为两圆交点。欧几里得根据公设 1 作直线 CA 和 CB，然后，宣布△ABC 是等边三角形。因为根据定义 15，由于 AC 和 BC都是它们各自圆的半径，所以，$\overline{AC}=\overline{AB}$，$\overline{BC}=\overline{AB}$。由于公理1称，与同一个东西相等的东西彼此也相等，所以，我们说，$\overline{AC}=\overline{AB}=\overline{BC}$，因此，根据定义，△ABC是等边三角形。

这是一个非常简单的证明，只应用了两个公设，一个公理和两个定义，乍一看，似乎很完美。但遗憾的是，这个证明是有缺点的。即使古希腊人，不论他们对《原本》评价多高，也都看出了欧几里得最初论证的逻辑缺陷。

问题出在 C 点上。欧几里得如何证明两个圆实际上一定会相交呢？他怎么知道这两个圆不会以某种方法相互通过而不相交呢？显然，由于这是他的第一个命题，他以前并没有证明过这两个圆必然相交。而且，在他的公设或公理中，也都没有提到这个问题。对 C 点存在的唯一证明就是图中的明确表示。

但问题就在这里。因为如果说欧几里得想从他的几何中排除什么，那就是代替了证明的对图的依据，根据他自己的基本规则，证明必须建立在逻辑基础上，必须建立在依据公设和公理所做的谨慎的推理基础上，一切结论最终都必须来源于此。欧几里得“让图说话”，就违背了他给自己制定的规则。并且，如果我们想从图中得出结论，我们完全可以根据观察来判明命题 1.1，即所作三角形看起来是等边三角形。如果我们求助于这种

视觉判断，那么，一切都不再成立。

现代几何学家认为，需要增加一个公设，以作为判定这两个圆必定相交的理论根据，这一公设有时称之为“连续性公设”。他们还引入了其他公设，以弥补《原本》中这里或那里出现的类似缺陷。本世纪初，数学家戴维·希尔伯特（1862—1943 年）依据 20 个公设演绎出他自己的几何学，堵塞了欧几里得的许多漏洞。因而，1902 年，伯特兰·罗素对欧几里得的著作给予了否定的评价：

“他的定义并非总是确定的，他的公理也不是都无法证明，他的论证需要许多公理，而他自己却没有意识到。严谨的证明应在没有图形辅助时依然保持其论证的力量，但欧几里得的许多早期证明却不能如此……他的著作作为逻辑名作的价值在很大程度上被夸大了。”

大家公认，欧几里得在以图像、而不是以逻辑为先导时，他不过是没做应该做的事。而在他全部 465 个命题中，并没有一处做了不该做的事。他的 465 个定理，没有一个是虚假的。只要对他的证明作一些小小的改动，并增加一些遗漏的公设，他的全部命题就能够经受住时间的考验。那些赞同罗素观点的人不妨首先将欧几里得的著作与希腊天文学家、化学家或物理学家的著作作一番比较。用现代标准来看，那些古希腊科学家真正是处于原始状态，今天，没有一个人会依据这些古代科学家的著作来解释月球的运动或肝脏的功能。但与此相反，我们经常可以请教欧几里得。他的著作是一项永恒的成就。它无须依赖收集数据或创造更精密的仪器。一切只需敏锐的理性，而欧几里得恰恰高于理性。

命题 I.2 和 I.3 巧妙解决了前面提到的在没有移动圆规的明确公设情况下转移长度的问题；而命题 I.4 则是欧几里得的第一个全等命题。用现代话说，这一命题就是“边角边”或“SAS”三角形全等模式，对此，读者应回想起中学几何课上学过的知识。命题 I.4 设定，如果有两个三角形，其中一个三角形的两条边及其夹角分别与另一个三角形的两条边及其夹角相等，则这两个三角形全等（图 2.3）。

欧几里得在证明中，首先假设$\overline{AB}=\overline{DE}$，$\overline{AC}=\overline{DF}$，$\angle BAC=\angle EDF$。然后，他拿起△DEF，放到△ABC 上，并证明，两个三角形完全重合。这种用叠加方式证明的方法早已不受欢迎。并且，谁能说当图形在纸上移动的时候，它们不会变形或扭曲呢？希尔伯特认识到了这种危险，他实质上已将 SAS 作为他的公理Ⅳ.6。

命题 I.5 确定等腰三角形的两个底角相等。这一定理以“笨人不过桥”著称。之所以有此说法，一则是因为欧几里得的图形有点儿像一座桥；再则是因为，许多差些的学生都难于理解这一定理的逻辑，因此，也就无法跨过这座桥，进入《原本》的其它部分。

接下来的命题，即命题 I．6，是命题 I．5 的逆命题。该命题确定，如果一个三角形的两个底角相等，则这个三角形是等腰三角形。显然，逻辑学家对定理及其逆定理极感兴趣，所以，欧几里得在证明一个命题后，常常会插入逆命题证明，即使省略或延迟这一证明都不致损害他著作的逻辑。

欧几里得的第二个三角形全等模式——“边边边”或“SSS”，写入了命题 I.8。这一命题确定，如果有两个三角形，其中一个三角形的三条边分别与另一个三角形的三条边相等，则这三条边所对应的两个三角形全

等。

随后的几个命题是作图命题。欧几里得演示了如何用圆规和直尺平分一个已知角（命题Ⅰ.9）或一个已知线段（命题Ⅰ.10）。紧跟其后的两个命题则演示了如何作已知直线的垂线，其一是过直线上已知点作垂线（命题Ⅰ.11），其二是过直线外已知点作垂线（命题Ⅰ.12）。

欧几里得下面的两个定理是关于邻角∠ABC 和∠ABD 的，如图 2.4 所示。他在命题Ⅰ.13 中证明，如果 CBD 是一条直线，那么，上述两个角之和等于两个直角；在命题Ⅰ.14 中，他证明了这一定理的逆定理，即，如果∠ABC 与∠ABD 之和等于两个直角，则 CBD 是直线。接着，他应用这一角与直线的性质，证明了更为重要的命题Ⅰ.15。

命题Ⅰ.15　如果两条直线相交，则所形成的对顶角相等（图 2．5）。

证明　因为 AEB 是一条直线，所以，命题Ⅰ．13 保证了∠AEC 与∠CEB 之和等于两个直角。同样，我们可以说，∠CEB 与∠BED 之和也等于两个直角。公设 4 称，所有直角都相等，并根据公理 1 和 2，得出∠AEC＋∠CEB＝∠CEB＋∠BED。然后，根据公理 3，从等式两边各减去∠CEB，欧几里得即得出结论，对顶角∠AEC 和∠BED 相等，与命题一致。　证讫。

这一定理又为我们引出了命题Ⅰ.16，即所谓外角定理，这是《第一篇》中最重要的定理之一。

命题 Ⅰ.16　在任何三角形中，一角的外角大于其他两角中的任何一角。

证明　已知△ABC，延长 BC 到 D，如图 2．6 所示，我们必须证明∠DCA 大于∠CBA 或∠CAB。欧几里得先根据命题Ⅰ．10，平分 AC 于 E，然后，根据公设 1，作线段 BE。公设 2 使他可以延长 BE，并根据命题Ⅰ．3，作$\overline{EF}=\overline{EB}$。最后，连接FC。

我们来看三角形AEB和CEF，欧几里得指明，根据平分作图，$\overline{AE}=\overline{CE}$；根据命题I.15，对顶角∠1和∠2相等；根据作图，$\overline{EB}=\overline{EF}$。因此，根据命题 Ⅰ.4（即“边角边”或 SAS），这两个三角形全等，所以，∠BAE=∠FCE。∠DCA 显然大于∠FCE，因为根据公理 5，整体大于部分。因此，外角∠DCA 大于内对角∠BAC。用同样方法也可以证明∠DCA 大于∠ABC。证讫。

外角定理是一个几何不等式。《原本》中随后的几个命题也是如此。例如，命题Ⅰ．20 确定，三角形任何两边之和必大于第三边。但据我们所知，古希腊伊壁鸠鲁派对这一定理很不以为然，因为他们认为这条定理太通俗，犹如不证自明的公理，甚至连驴子也会明白。也就是说，如果有一头驴站在 A 点（图 2.7），而它的食物放在 B 点。这头驴肯定本能地懂得，从 A 直接到 B，路程比沿两条边走，即从 A 到 C，再从 C 到 B 要短。人们曾认为，命题Ⅰ.20 确是一条不证自明的真理，因此应属于公设。然而，如果能够作为一条命题证明这一定理，犹如前文中圆规的例子一样，欧几里得当然不愿再去假定一条公设，而他对这一定理所做的证明又是非常富有逻辑性的。

欧几里得接着又提出了几条不等式命题，随后提出了他最后一条全等定理，即重要的命题 Ⅰ．26。在这一命题中，他首先证明了“角边角”或

ASA 的全等模式，并以此作为命题 I．4“边角边”或 SAS 全等定理的推论。然后，在命题 I．26 的第二部分，欧几里得又提出了第四个，也是最后一个全等模式，即“角角边”。对此，他证明，如果∠2=∠5，∠3=∠6，并且，$\overline{AB}=\overline{DE}$，如图2.8所示，则三角形ABC和DEF全等。

开始，人们会认为这只是“角边角”模式的直接推论而不予考虑。我们可以很清楚地看到，∠2+∠3=∠5+∠6，据此，我们可以得出

∠1=2 个直角-（∠2＋∠3）=2 个直角-（∠5＋∠6）=∠4

然后，我们可以再回复到“角边角”（ASA）的全等模式，因为我们可以把等式中的角放在 AB 与 DE 的任何一端。

这是一个简短的证明；但遗憾的是，这个证明同样不能令人满意。在这里欧几里得不能引用这一证明，因为他还必须证明一个三角形三个角的和等于两个直角。的确，如果没有这一关键性的证明，似乎完全不可能证明“角角边”（AAS）的全等定理。但是，欧几里得却确实证明出了这一定理，他用反证法作了如下精彩的证明。

命题 I.26（角角边或 AAS）　已知两个三角形，如果其中一个三角形的两个角分别与另一个三角形的两个角相等，一条边，即……相等角中的一个对边，等于另一个三角形相应的一条边，则其余的边和其余的角也相等。

证明　如图2.9，假设∠2 = ∠5，∠3 = ∠6，并且，$\overline{AB}=\overline{DE}$。欧几里得宣称，边长BC与EF也必定相等。但为证明这一点，他假设其中一条边比另一条边长；例如，假设$\overline{BC}>\overline{EF}$。因此，我们就可以作线段BH，使之等于 EF。然后，作线段 AH。

根据假设，已知$\overline{AB}=\overline{DE}$，∠2 = ∠5；并且，根据作图，$\overline{BH}=\overline{EF}$，因此，根据“边角边”定理，△ABH 与△DEF 全等。所以，∠AHB=∠6，因为它们是两个全等三角形的对应角。

然后，欧几里得将注意力集中于小△AHC，并注意到，其外角 AHB 和相对内角∠3 都等于∠6，因此，∠AHB 与∠3 也应该相等。但是，欧几里得在重要的命题 I．16 中已然证明，外角必定大于内对角。这一矛盾表明，最初的假设$\overline{BC}\neq\overline{EF}$是不正确的。因此，他断定，这两条边实际上相等，因而，根据“边角边”定理，原三角形 ABC 与 DEF 全等。　证讫。

我们再来看一看这一巧妙论证的重要意义：这四种全等模式（边角边、边边边、角边角和角角边）都成立，但无须涉及三角形三个角之和等于两个直角的问题。

命题 I.26 结束了第一篇的第一部分。回顾这一部分的内容，我们看到，欧几里得在几何上已很有造诣。即使他还不得不应用他的平行线公设，但他已经确立了四种全等模式，研究了等腰三角形、对顶角和外角，并进行了各种作图。但是，他并未就此止步，仍在尽力走得更远。《原本》随即提出了平行线的概念。

第一篇：平行线及有关命题

命题 Ⅰ.27 一条直线与两条直线相交，如果内错角相等，则这两条直线平行。

证明 见图 2.10，假设∠1=∠2，欧几里得必须证明直线 AB 与 CD 平行——即，根据定义 23，他必须证明这两条直线不会相交。他采用间接证法，先假设这两条直线相交，然后找到所涉的矛盾。假设直线 AB 与 CD 延长后，相交于 G。那么，图形 EFG 就是一个延伸很长的三角形。但是，△EFG 的外角∠2 等于这同一个三角形的内对角∠1。根据命题 Ⅰ. 16 外角定理，这种情况是不可能的。因此，我们断定，AB 与 CD，不论延长多长，也不会相交，而这恰恰是欧几里得的平行线定义。证讫。

命题 Ⅰ.27 打破了有关平行性的坚冰，但是，欧几里得依然避免应用平行线公设。这一争议很大的公设在欧几里得在命题 Ⅰ.29 中证明 Ⅰ.27 的逆命题时，终于出现了。

命题 Ⅰ.29 一条直线与两条平行线相交，则内错角相等。

证明 这次，欧几里得假设 AB 与 CD 平行（见图 2.11），并须证明∠1=∠2。他再次使用间接法，即，假设∠1≠∠2，然后引出逻辑上的矛盾。因为，如果这两个角不相等，那么，其中一个角必定大于另一个角，我们不妨假设∠1＞∠2。根据命题 Ⅰ.13

2 个直角=∠1＋∠BGH＞∠2+∠BGH

在此，欧几里得终于引用了公设 5，这一公设恰恰适合于这种情况。由于∠2+∠BGH＜2 个直角，根据公设 5，他可以断定，AB 与 CD 必定相交于右侧，这显然是不可能的，因为已知这条直线是平行的。因此，根据反证法，欧几里得表明，∠1 不能大于∠2；同样，∠2 也不能大于∠1。总而言之，平行线的内错角相等。证讫。

根据这一证明，欧几里得很容易地便推断出同位角也相等，即，在图 2.11 中，∠EGB=∠2，因为∠EGB 与∠1 是对顶角。

在最终引用了平行线公设之后，欧几里得发现，实际上不可能打破以往的习惯。在第一篇余下的 20 个命题中，几乎没有一处再直接应用平行线公设或基于这一公设的命题，唯一的例外是命题 Ⅰ.31，在这一命题中，欧几里得演示了如何通过直线外一点作已知直线的平行线。但是，平行线公设当然是被嵌入了一个人人都在等待出现的定理之中：

命题 Ⅰ.32 在任何三角形中……三个内角之和……等于两个直角。

证明 已知△ABC，如图 2.12 所示，他根据命题 Ⅰ.31，作 CE 平行于三角形的边 AB，并延长 BC 到 D。根据命题 Ⅰ.29（平行线公设的推论），他知道，∠1=∠4，因为它们是两条平行线的内错角；并且，还知道，∠2=∠5，因为它们是同位角。因此，△ABC 三个内角的和就是∠1＋∠2＋∠3=∠4+∠5＋∠3=2 个直角，因为这些角构成了直线 BCD。这样，这一著名的定理即证明完毕。 证讫。

自此，欧几里得开始将注意力转向更复杂的问题。他接下来的几个命题提出了有关三角形和平行四边形的面积问题，其中最精彩的是命题 Ⅰ.41。

命题 Ⅰ.41 如果一个平行四边形与三角形同底，且位于同两条平行线之间，则这个平行四边形的面积是三角形面积的两倍。

希腊人以此说法表示，如果一个三角形与任意平行四边形同底同高，则这个三角形的面积等于平行四边形面积的一半。由于这种平行四边形的面积与同底同高的矩形面积是一致的，而矩形的面积是（底）×（高），我们由此可以看到，在命题Ⅰ.41 中包含着一个现代公式，即，面积（三角形）$=\frac{1}{2}bh$。

但是，欧几里得并未用这种代数语言思维。相反，他想象，△ABC 确与平行四边形 ABDE 具有同一条边，且同位于两条平行线 AB 与 DE 之间，如图 2.13 所示。然后，欧几里得证明，面积（平行四边形 ABDE）=2 面积（△ABC）。

间隔几个命题之后,欧几里得在命题Ⅰ.46 中演示了如何在已知线段上作正方形图形。当然，正方形是一种规则四边形，因为它的所有边和所有角全等。最初，人们可能会以为这一命题只是一个普通的命题，特别是他们会回忆起第一篇一开始就介绍了等边三角形这种规则三边形的作图。我们只要看一看他对正方形作图的证明就会明白，正方形作图何以延迟了这么长时间，因为对正方形作图的论证，很多要根据平行线的性质，而这当然只能等到关键的命题Ⅰ.29 之后。因此，虽然欧几里得在第一篇的一开始就介绍了规则三角形的作图，但他不得不等到接近第一篇的尾声时才作规则四边形的图形。

第一篇除了证明这 46 个命题之外，还有最后两个命题需要证明。看来，欧几里得是将最好的留在了最后。在作好所有这些准备之后，他开始冲击毕达哥拉斯定理，这一定理显然是所有数学定理中最重要的定理之一。

伟大的定理：毕达哥拉斯定理（勾股定理）

众所周知，在欧几里得之前，毕达哥拉斯定理即已闻名遐迩，因此，欧几里得决不是这一数学里程碑的发现人。然而，我们下面看到的证明为他赢得了声誉，许多人都相信，这一证明最初是由欧几里得作出的。这个证明的美妙之处在于其先决条件的精练；毕竟，欧几里得为作出证明，只能依赖他的公设、公理和最初的 46 个命题，可谓捉襟见肘。我们不妨考虑一下他尚未涉及的几何论题：他以前唯一探讨过的四边形是平行四边形；对于圆，基本上尚未探索；而对于特别重要的相似性，则直到第六篇才开始阐述。虽然可以确信，如果应用相似三角形，可以对毕达哥拉斯定理作出非常简短的证明，但是，欧几里得不愿把这一重要命题的证明推迟到第六篇以后进行。显然，他希望尽可能早地直接涉及毕达哥拉斯定理，因此，他创立了一个证明，并以此作为《原本》的第 47 个命题。从这个命题中，我们可以看到，在此之前的许多命题都指向了伟大的毕达哥拉斯定理，因此，我们可以说第 47 命题堪称第一篇的高潮。

在我们详细介绍欧几里得的证明之前，我们不妨先来看一看用欧几里得语言阐述的这个命题，从中可以窥见其论证方法之巧妙。

命题 Ⅰ.47　在直角三角形中，斜边上的正方形面积等于两个直角边上的正方形面积之和。

请注意，欧几里得的命题不是关于代数方程式 $a^2=b^2+c^2$，而是述及了

一种几何现象，涉及到以直角三角形的三条边为边所作的实在的正方形。欧几里得必须证明，以 AB 和 AC 为边的两个小正方形面积之和等于以斜边 BC 为边的大正方形面积（见图 2.14）。为证明这一点，他采用了一个非常奇妙的方法，从直角顶点开始作线段 AL，使之与大正方形的边平行，并将大正方形分割为两个矩形。现在，欧几里得只要证明左边矩形（即以 B 和 L 为对角的矩形）的面积等于以 AB 为边的正方形面积；同样，右边矩形的面积等于以 AC 为边的正方形面积即可。由此可直接导出，两个矩形面积之和等于大正方形面积，同样也就等于两个小正方形面积之和。

这一普通方法非常巧妙，但还需要补充一些细节。幸好，欧几里得在他的早期命题中已完成了全部准备工作，因此，现在的问题是如何将它们谨慎地组合起来。

证明　根据假设，欧几里得已知∠BAC 是直角。他应用命题 I.46，在三条边上作正方形，并应用命题 I.31，过 A 点作 AL 平行于 BD，然后，连接 AD 与 FC。初看起来，这些辅助线似乎显得很神秘，但它们很快就会变得浅显易懂了。

对于欧几里得来说，关键的问题是要证明 CA 与 AG 在同一条直线上。欧几里得指明，根据正方形作图，∠GAB 为直角，而根据假设，∠BAC 也是直角。由于这两个角的和等于两个直角，命题 I.14 保证了 GAC 是一条直线。有趣的是，在这一显然只涉及到很少的技术性问题的证明中，欧几里得唯一一次应用了∠BAC 是直角这一事实。

现在，欧几里得开始将目光转向两个细长的三角形 ABD 和 FBC。这两个三角形的短边（分别为 AB 和 FB）相等，因为它们是一个正方形的两条边；同理，两个三角形的长边（BD 和 BC）也相等。那么，它们的对应夹角是否相等呢？由于∠ABD 是∠ABC 与正方形直角∠CBD 之和，而∠FBC 是∠ABC 与正方形直角∠FBA 之和。公设 4 规定，所有直角都相等。公理 2 则保证了等量之和相等。因此，∠ABD=∠FBC。根据“边角边”定理（即命题 I.4），欧几里得证明狭长三角形 ABD 与 FBC 全等；因此，这两个三角形的面积相等。

到目前为止，一切顺利。接着，欧几里得指明，△ABD 与矩形 BDLM 具有同一条边 BD，并且，位于同两条平行线（BD 与 AL）之间。因此，根据命题 I.41，BDLM 的面积等于△ABD 面积的 2 倍。同样，△FBC 与正方形 ABFG 也具有同一条边 BF。并且，欧几里得已证明 GAC 是一条直线，因此，△FBC 与正方形 ABFG 也同位于平行线 BF 与 GC 之间 根据命题 I.41，正方形 ABFG 的面积也等于△FBC 面积的 2 倍。

欧几里得综合这些结果和先前证明的三角形全等，得出：

面积（矩形 BDLM）=2 面积（△ABD）
=2 面积（△FBC）
=面积（正方形 ABFG）

至此，欧几里得完成了一半使命。下一步，他需证明矩形 CELM 的面积等于正方形 ACKH 的面积。对此，他可以用同样的方法证明。首先，连接 AE 与 BK，然后，证明 BAH 是一条直线，并根据“边角边”定理，证明△ACE 与△BCK 全等。最后，引用命题

I.41，欧几里得推论：

面积（矩形 CELM）=2 面积（△ACE）

=2 面积（△BCK）

=面积（正方形 ACKH）

至此，毕达哥拉斯定理呼之欲出，因为：

面积（正方形 BCED）

=面积（矩形 BDLM）+面积（矩形 CELM）

=面积（正方形 ABFG）+面积（正方形 ACKH）。证讫。

至此，欧几里得完成了数学中最重要的证明之一，而他所应用的图形（图 2.14）也因此成为了非常著名的图形。人们常常称欧几里得的图形为“风车”，因为它的外形看起来很像风车。从附图中我们可以看到 1566 年版《原本》所刊载的“风车”图形，图中的文字为拉丁文。显然，400 多年前的学生便已开始研究这一图形，犹如我们刚才所做的那样。

当然，欧几里得的证明并不是证明毕达哥拉斯定理的唯一方法。实际上，证明方法有数百种之多，有的非常巧妙，有的极其平庸。（其中包括俄亥俄州众议员詹姆斯·加菲尔德的证明，他后来成为美国总统。）读者如果对其他证明方法感兴趣，可以参考 E.S.卢米斯所著《毕达哥拉斯命题》一书，其中收录了对这一著名定理的千百种证明方法，令人眼花缭乱。

虽然命题Ⅰ.47 标志了第一篇的高潮,但欧几里得还有最后一个命题要证明，这就是毕达哥拉斯定理的逆定理。欧几里得对这一逆定理的证明，其巧妙和精练，依然是显而易见的。但遗憾的是，这一证明本该同样著名，却始终湮没不彰。实际上，大多数学生在其一生中，总会在某一时刻见到过对毕达哥拉斯定理的证明，但是见过对其逆定理证明的人就少得多，即使见到，也不敢肯定其正确性。

欧几里得对这一逆定理的证明有两个特点值得我们特别注意。其一是它非常短，将其与我们刚看到的论证相比，则尤其如此。其二是欧几里得在证明这一逆定理时，应用了毕达哥拉斯定理。这种逻辑方法虽然并非没有前例，但至少值得注意。让我们回想一下，欧几里得在证明有关平行线的两个重要命题（命题Ⅰ.27 及其逆命题Ⅰ.29）时，并没有用其中一个命题去证明另一个命题。但是,他对毕达哥拉斯逆定理的证明,却将命题Ⅰ.48 牢固地建立在命题Ⅰ.47 的基础之上，使这两个命题成为一个明确的序列单位。

命题Ⅰ.48　在一个三角形中，如果一边上的正方形面积等于其他两边上的正方形面积之和，则这两边的夹角是直角。

证明　欧几里得首先作△ABC，并假设$\overline{BC}^2=\overline{AB}^2+\overline{AC}^2$，如图2.15 所示。他必须证明∠BAC 是直角。

为此，欧几里得首先根据命题Ⅰ.11，作 AE 垂直于 AC，并交 AC 于 A。然后，作$\overline{AD}=\overline{AB}$，并连接CD。现在，欧几里得求证的中心问题是要证明三角形 BAC 与 DAC 全等。

显然，这两个三角形有一条共同边AC，并且，根据作图，$\overline{AD}=\overline{AB}$。

虽然我们显然不能断言∠BAC是直角（实际上，这正是该定理所要确定的），但根据垂线作图，我们知道∠DAC 是直角。因此，欧几里得完全有理由应用毕达哥拉斯定理于直角三角形 DAC，并根据假设，推导出

$$\overline{CD}^2=\overline{AD}^2+\overline{AC}^2=\overline{AB}^2+\overline{AC}^2=\overline{BC}^2$$

在此，$\overline{CD}^2$与$\overline{BC}^2$相等也意味$\overline{CD}^2$与$\overline{BC}^2$相等，因此，根据“边边边”定理，△DAC 与△BAC 全等。因而，∠BAC 与∠DAC 也必然全等。而根据作图，后者为直角，所以，∠BAC 也是直角。证讫。

命题Ⅰ.47 和Ⅰ.48 相得益彰，揭示了直角三角形的全部特征。欧几里得表明，一个三角形，如果，也只有当其斜边的平方等于两条侧边的平方和时，这个三角形才是直角三角形。这些证明过去是，现在依然是最佳几何例证。

这两个毕达哥拉斯命题在另一种意义上也是卓越非凡的。欧几里得以一种巧妙的方式证明这两个命题是一回事，而这两个命题是正确的则是另一回事。对于直角三角形与平方和的密切关系，没有直觉的推论。例如，它不像命题Ⅰ.20 那样，是一种甚至连驴子都能懂得的不证自明的真理。相反，毕达哥拉斯定理证明了一个非常奇特的事实，其奇特性之所以不被认识，仅仅是因为其结果太著名了。理查德·特鲁多在他的《非欧几里得革命》一书中精彩地描述了毕达哥拉斯定理这种固有的奇特。特鲁多注意到，直角是一种人人都熟悉的日常存在，它不仅存在于人为世界，而且也存在于自然界本身。还有什么能比直角更“普通”或更“自然”的呢？但特鲁多又说：

“毕达哥拉斯定理使我感到非常惊奇……‘$a^2=b^2+c^2$’……无论如何引不起我本能的记忆……因为这个方程抽象，精确，异乎寻常。我想象不出这样一种东西与日常生活中所见的直角有什么关系。因此，当偶然揭开‘熟悉’的帘幕，重新审视毕达哥拉斯定理，我不禁感到目瞪口呆。”

后记

纵观历史，《原本》第一篇基础中最令人困惑的是引起争议的平行线公设。困惑的产生并非因为有人怀疑平行线公设的真理性，相反，人们普遍认为这个公设是逻辑的必然。几何毕竟是一种抽象描述世界的方式，是一种“物理的抽象”，而物理现实又确实决定了平行线公设的真理性。

因此，受到质疑的不是欧几里得的陈述，而是他将其列为公设。古代作家普罗克洛斯一言以蔽之，“它（公设 5）完全应从公设中剔除，因为它是一条定理……”

对平行线公设的这种认识并不奇怪。首先，可能确实使古代几何学家感到迷惑的是，这一公设看起来的确十分像一条命题，因为它的陈述性语句就占了大半段。加之，欧几里得似乎不仅尽可能避免应用这一条公设，而且在证明一些相当深奥微妙的结论时，也尽可能设法绕过它。“如果说他的其他公设和公理的内容都非常丰富，足以产生诸如命题Ⅰ.16 或Ⅰ.27 这样的定理或四种全等格式的话，那么，它们当然也应该同样包容平行线公设的含义。”

出于似乎非常充分的理由，数学家们开始寻求公设 5 的推导根据。他们在寻求这一证明的过程中，可以自由地应用除公设 5 以外的任何其他公设或公理，以及欧几里得从Ⅰ.1 到Ⅰ.28 的全部命题。无数数学家都曾为此做出过不懈的努力，但非常遗憾的是，他们几年、几十年，甚至几百年的努力都失败了。这一证明至今依然是一个难解的谜。

几何学家在这一过程中，只发现了许多在逻辑上等同于平行线公设的

新的命题。为证明公设 5，常常需要数学家们去假设一种看来很明显，但迄今为止尚未得到证明的命题。然而，遗憾的是，为引出这样一个命题，平行线公设本身又是必不可少的，而问题就在这里。对于逻辑学家来说，这表明，两者实际都在表达同一个概念，而对公设 5 的“证明”，如果要求假设它的逻辑等价命题，自然就什么也没有证明。

比较著名的四个平行线公设等价命题记叙如下。应该指出的是，假如可以根据公设 1 至 4 证明下述任何一项，则公设 5 便是顺理成章的了。

■ 普罗克洛斯公理：如果一条直线与两条平行线中的一条相交，也必定与另一条平行线相交。

■ 等距公设：两条平行线之间距离处处相等。

■普莱费尔公设：经过已知直线外一点，可以作一条，而且只能作一条与已知直线平行的直线。

■ 三角形公设：三角形三个内角和等于两个直角。

尽管文艺复兴时期产生了这四个逻辑等价命题，但却依然未能解决平行线公设的性质问题。无论谁推导出平行线公设证明，都会在数学史上享有永久的声望。有时，这一证明似乎已近在咫尺，唾手可得，但世界最优秀数学家的努力却一次又一次落空。

19 世纪初叶，有三个数学家几乎同时爆发灵感，发现了解决这一难题的真正曙光。第一位数学家就是举世无双的卡尔·弗里德里希·高斯（1777—1855 年），有关他的生平，我们将在第十章中介绍。高斯立足对三角形角度的测量，重新设计了这个问题。为了证明三角形的内角和必定等于 180°，他先假设三角形内角和不等于 180°。这样，就使他面临两种选择：三角形内角和或者大于 180°，或者小于 180°。他进而研究了这两种情况。

高斯依据直线是无限长的事实（欧几里得也同样含蓄地提出过这样的假设，对此，没有人提出异议）发现，如果三角形的内角和大于 180°会导致逻辑矛盾。因此，这种情况显然应予排除。如果他能够同样排除另一种情况，他就可以间接地证明平行线公设的必然性。

高斯首先假设三角形的三个内角和小于 180°，然后便开始进行推理。但推理的结果非常奇怪，似乎有点儿不可理解和违背直觉（一种瞬间出现的现象）。但是，高斯却怎么也找不到他所寻求的逻辑矛盾。1824 年，他总结这种情况说：

“……一个三角形的内角和不能小于 180°……这是……一块暗礁，所有的船只都会在它面前撞得粉碎。”

随着高斯对这一特殊几何问题越来越深入的探讨，他逐渐相

信这其中不存在逻辑矛盾。相反，他开始感觉到，他所发展的不是一种不相容的几何学，而是一种选择几何学，用他的话说，是一种“非欧几里得”几何学。高斯在他 1824 年的一封私人信件中详细阐述了他的观点：

“三角形三个内角和小于 180°的假设导致了一种非常古怪的几何学，与我们现在的几何学不同，但又完全讲得通，对此，我感到非常满意。”

这是一段激动人心的话。高斯虽然被公认为是当时最优秀的数学家，但却没有公布他的发现。也许是为声名所累，因为他深信，对他见解的争议可能会损害他的崇高名望。1829 年，高斯在写给他一位知己的信中说，

他没有打算：

“……把我的深入研究公诸于众，也许终生都不会公布，因为我惧怕在我大声讲出我的观点之后，会引起维奥蒂亚人的鼓噪。”

今天的读者可能不明白维奥蒂亚人是何方神圣，对此，我们只需稍加解释，所谓的“维奥蒂亚人”是指那些缺乏想象力而又不开化的愚钝之人。显然，高斯忽略了数学界对他新观点的接受能力。

接下来是匈牙利数学家约翰·鲍耶（1802—1860 年）。约翰的父亲沃尔夫冈曾是高斯的密友，而且，他自己也曾为证明欧几里得的平行线公设空付出大半生的心血。当时的年代，儿子常常继承父亲的事业，成为牧师、皮匠或厨师……，而小鲍耶则继承了他父亲推导欧几里得平行线公设的深奥事业。但沃尔夫冈深知个中的难处，对他的儿子提出了强烈的警告：

“你不能再去论证平行线公设。我深知这条路会带来什么结果。我曾力图穿越这无尽的黑夜，并因此葬送了我生活的全部光明与欢乐……我恳求你，不要再去管平行线公设。”

但是，年轻的约翰·鲍耶并未理会父亲的忠告。像高斯一样，约翰也逐渐认识到了有关三角形内角和的关键性的三分法，并试图排除与平行线公设不符的所有情况。当然，同高斯一样，他也没有成功。随着鲍耶对这一问题越来越深入的研究，他同样得出结论，认为欧几里得几何在逻辑上遇到了强有力的对手，他十分惊讶地就他独待而显然论据确凿的命题写道，“从空无中，我创造了一个奇怪的新世界。”

约翰·鲍耶不像高斯，他毫不犹豫地公布了自己的发现，他将自己的论文作为附录载于他父亲 1832 年的著作之中。老鲍耶兴高采烈地将自己的著作给他的朋友高斯寄去一本，但高斯的回信却使鲍耶父子十分意外：

“如果坦言我不敢夸奖（令郎的）大作，你必然会感到吃惊：但是我别无选择；夸奖令郎就等于夸奖我自己；因为书中全部内容，他的思路，以及他所推导的结果，都与我自己的发现几乎同出一辙，这些发现在我脑子里已经存在了 30 至 35 年之久。”

显然，高斯给他年青的崇拜者泼了一瓢冷水。值得称道的是，高斯非常谦和地讲到他自己“……非常高兴，恰恰是老友的儿子以这种非凡的方式超过自己”。但是，约翰得知他最伟大的发现已经躺在高斯的抽屉里几十年了，这对他的自尊心，当然是一个沉重的打击。

然而，约翰的自尊心还要再经受一次打击，因为人们不久便得知，俄国数学家尼古拉·罗巴切夫斯基（1793—1856 年）不仅与高斯和鲍耶作了同样的工作，而且，于 1829 年就发表了他关于非欧几何的论文——比约翰早了整整三年。但罗巴切夫斯基的论文是用俄文写的，显然无声无息地传到了西欧。这种现象在科学界并不奇怪，一个发现有时会有许多人同时独立作出。沃尔夫冈·鲍耶讲得好：

“……的确，许多事物似乎都自有其时令，会在多处同时显现，犹如紫罗兰在春季到处开放。”

但是，这些发现还不能算是切中要害，另一位创新家乔治·弗里德里希·伯恩哈德·黎曼（1826—1866 年）对几何直线的无限长度别有一种见解。正是这种几何直线的无限性才使高斯、鲍耶和罗巴切夫斯基得以排除三角形内角和大于 180° 的情况。但是，是否确有必要假设这种无限性呢？欧几里得的公设 2 称，有限直线可沿直线无限延长，但这难道不是在说，

人们永远也达不到直线的尽头吗？黎曼完全可以想象，这些直线有几分像圆，长度是有限的，但却没有“尽头”。他说：

“……我们必须区别无界与无限延长的概念……，空间的无界具有一种比外部感受更强的经验确实性。但无限延长绝不是从这个意义上推导出来的。”

黎曼根据直线无界但长度有限的假设，重新检讨几何学，则三角形内角和大于 180°时所产生的逻辑矛盾消失了。结果，他发展了另一种非欧几何，在这种几何中，三角形的内角和大于两个直角。黎曼的几何学虽然与欧几里得和鲍耶的几何不同，但却显然同样严谨。

今天，我们承认所有这四位数学家为非欧几里得几何的创始人。他们理应享受先驱者的同等荣耀。但是，他们的发现也没有完全解决平行线公设的根本问题。因为，虽然他们把几何发展到了新的高度，但是，能够支持他们的新几何学与欧几里得几何并驾齐驱的，仅仅是一种知其然而不知所以然的直觉感受，并非白纸黑字的逻辑推理。尽管高斯、鲍耶、罗巴切夫斯基和黎曼的发现都有很强的说服力，但在将来的某一刻，仍有可能会出现一位天才数学家，从他们关于三角形内角和小于或大于 180°的假设中找出矛盾。

因而，这个古老故事的最后一章由意大利的欧金尼奥·奥尔特拉米（1835—1900 年）在 1868 年写完。他清晰地证明了非欧几何与欧几里得几何同样具有逻辑上的一致性。奥尔特拉米表明，如果说在高斯、鲍耶、罗巴切夫斯基或者黎曼的几何中，可能存有某种逻辑矛盾的话，那么，在欧几里得几何中也同样存在这种矛盾。既然人人都认为欧几里得几何逻辑严谨一致，因此可以断言，非欧几里得几何也同样无懈可击。换言之，非欧几何在逻辑上并不比先者——欧几里得几何低下。

为了理解高斯／鲍耶／罗巴切夫斯基派非欧几何（即三角形内角和小于 180°的那种几何）的某些古怪论点，我们不妨看一看非欧几何对某些命题的证明。首先，让我们从另一个角度看一看三角形全等问题。当然，欧几里得的全等定理是在他初次应用公设 5 之前确立的，并在非欧几何中依然有效，因为这些全等定理的证明只需应用欧几里得的其他公设和公理，而无需参考其他任何东西。但在鲍耶几何中，令人感到惊奇的发展是，还有另外一种表示全等的途径，即“角角角”。

在欧几里得几何中，我们知道，如果两个三角形的三个角分别相等，则这两个三角形相似。也就是说，它们形状相同，但无须全等。例如，一个小等边三角形和一个大等边三角形，尽管三个角都完全相等，但却是不全等图形。然而，我们下面将要讲到的非欧定理却表明，在非欧几何这个奇怪的世界里，这种情况却是不可能的。如果鲍耶的两个三角形形状相等，其面积也必定相等！

定理（角角角）如果一个三角形的三个角分别与另一个三角形的三个角相等，则这两个三角形全等。

证明　如图 2.16 所示，在三角形 ABC 和 DEF 中，假设∠1=∠4，∠2=∠5，∠3=∠6。我们断言，边长 AB 与 DE 必定相等。为证明这一点，我们先假设这两条边长度不等，以造成最后的逻辑矛盾，为了不失却一般性，我们不妨假设$\overline{AB}<\overline{DE}$。

作$\overline{DG}=\overline{AB}$，然后，根据命题Ⅰ.23，作∠DGH＝∠2。显然，根据“角边角”定理，△ABC 与△DGH 全等，因此，∠DGH＝∠2＝∠5，同理，∠DHG＝∠3＝∠6。

现在，我们来看四边形 EFHG。由于 DGE 和 DHF 是直线，根据命题Ⅰ.13，我们得知，∠EGH＝（180°－∠DGH）＝（180°－∠5），∠FHG＝（180°－∠DHG）＝（180°－∠6）。因此，四边形 EFHG 四个角的和等于

（180°－∠5）＋（180°－∠6）＋∠6＋∠5＝360°

现在，我们作四边形 EFHG 的对角线 GF，将四边形分为两个三角形。根据非欧几何的基本性质，这两个三角形，每个三角形的内角和都小于 180°；因此，两个三角形所有角的和必定小于 360°。而这两个三角形所有角的和恰恰就是四边形 EFHG 四个角的和。我们刚才已推导出，四边形 EFHG 的四角和等于 360°。

这样，就出现了矛盾。这就表明，第一步，即我们假设的$\overline{AB}\neq\overline{DE}$，是不正确的。总之，这两条边长度相等。然后，我们根据“角边角”定理，即根据命题Ⅰ.26，可以直接推导出原来的两个三角形 ABC 与 DEF 全等——而这正是我们所要证明的定理。　证讫。

从这一命题中，可以很容易地得出一个令人吃惊的推论：在非欧几何中，并非所有三角形的内角和都相等！欧几里得几何中这一最基本的性质（突出表现在许多几何推理中），在我们步入非欧几何领域时，却必须予以抛弃。因为假设有两个三角形，如图 2.17 所示，每个三角形的底角都是α和β，但是，AB 边显然小于 DE 边。因此，我们断言，∠1 不能等于∠2。因为如果它们相等，根据我们刚才证明的“角角角”全等定理，则这两个三角形全等，但由于$\overline{AB}\neq\overline{DE}$，因而这显然是不可能的。所以，我们看，一个三角形的内角和（∠1＋α＋β）不等于另一个三角形的内角和（∠2＋α+β）。总之，在非欧几何中，已知三角形的两个角，还不足以确定第三个角。从这一命题和许多其他类似命题中可以看出，为什么鲍耶说他创造了一个“奇怪的新世界”，以及为什么有那么多人在非欧几何刚刚露出地平线的时候就认为，非欧几何必然要出现逻辑矛盾。但是，正如我们刚才所证明的那样，他们全都错了。

那么，这些 19 世纪的发现者们究竟要将欧几里得置于何地呢？一方面，欧几里得几何作为对空间的唯一逻辑上一致的描述的地位不复存在。实际上，每个人都会感到非常吃惊的是，非欧几何证明了平行线公设不是逻辑所训示的。欧几里得假设了这一条公设，但在数学上却没有这种必然性。存在对立的几何，而且同样正确。

但另一方面，欧几里得的声誉得到了加强，而不是损毁。因为他没有像许多追随者那样落入陷阱，用其他不证自明的真理去证明平行线公设，我们现在知道，这种证明是注定要失败的。相反，他把他的假设理所应当地列为公设。欧几里得当然不可能知道两千年后会发现另一种几何学。但是，他凭着数学家的直觉，一定知道平行线的这一特性是一种个别的和独立的概念，它需要自己的公设，不论多么罗嗦和复杂。两千二百年后，数学家们证明了欧几里得始终是正确的。

第三章
欧几里得与素数的无穷性
（公元前约 300 年）

《原本》第二—六篇

《原本》第一篇中的 48 个命题为欧几里得的数学与组织才能树立了一座丰碑。作为第一篇，它当然是《原本》中最著名和人们研究最多的部分，但是，第一篇毕竟还只是《原本》13 篇中的一篇。本章将浏览一下这部古典巨著的其他部分。

《原本》第二篇探讨了我们如今称之为“几何代数学”的问题，即以几何概念构成的一定关系，今天，我们可以很容易地将这些关系转变为代数方程式。当然，代数概念对于古希腊人是陌生的，因为代数形成体系是几百年以后的事。我们不妨引用一条有代表性的命题，以使读者能够对第二篇有一个大致的了解。这条命题的行文初看起来似乎非常复杂，但仔细研究以后便会发现，这一命题乃是一个十分简单而熟悉的代数公式。

命题 II .4　如果把一直线，在任意一点截开，则以整条直线为边长的正方形面积等于两段上的正方形面积之和加上两个以这两段为边的矩形之面积。

证明　欧几里得首先设线段 AB，并于任意点 C 截开，如图 3.1 所示。如果我们设 $\overline{AC}=a$，$\overline{BC}=b$，则按照几何概念，“以整条直线为边长的正方形”面积（即 $(a+b)^2$）等于两段上的正方形面积之和（即 a^2+b^2）加上两个以这两段为边的矩形之面积（即 2ab）。也就是，

$(a+b)^2=a^2+b^2+2ab=a^2+2ab+b^2$　证讫。

当然，这就是我们初一代数所学的著名恒等式。欧几里得并未将此作为某种代数式，而是作为一种严格的几何表述，将以 AB 为边长的正方形分解为两个小正方形和两个全等矩形。然而，这一几何表述与其代数式显然是等价的。第二篇中绝大部分内容都是这种性质。第二篇最后以命题 II .14 结束，这条命题提出了一般多边形的求积问题，其证明已在第一章中介绍过。

第三篇包括有关圆的 37 个命题。我们在第一篇的作图中已经涉及到圆，但尚未集中讨论过。欧几里得在第三篇中证明了有关圆的弦、切线和角的标准命题。命题III.1 介绍了如何确定已知圆的圆心。根据定义 15，每个圆当然都有一个圆心，但对于一个画在纸上的圆，却并非一眼就能看出圆心所在。因此，欧几里得提供了一个非常必要的作图方法。

欧几里得在命题III.18 中明确地证明了圆的切线与经过切点的半径成直角。在其后的一个命题中，我们发现了一个重要的定理，“在一个圆中，同一弓形上的角相等”。如图 3.2 所示，∠BAD 与∠BED 相等，因为它们都是圆 BAED 中同一弓形上所形成的角。用现代术语说，这两个角都截取了同一条弧，即弧 BD。

在证明了这一定理之后，欧几里得又开始探讨圆内接四边形的问题，这种图形常常被人们称为“联圆四边形”。虽然这一定理可能有些专门，

但因其将在第五章的伟大定理中特别涉及，因此，我们将欧几里得对这个定理的简单证明放在本章介绍。

命题Ⅲ.22　圆内接四边形的对角和等于两个直角。

证明　我们首先作联圆四边形 ABCD，并作对角线 AC 与 BD，如图 3.3 所示。请注意，∠1＋∠2＋∠DAB＝2 个直角，因为它们都是△ABD 的内角。并且，∠1=∠3，因为它们截取同一条弧 AD；同理，∠2=∠4，因为它们截取同一条孤 AB。因此

2 个直角=（∠1+∠2）+∠DAB

=（∠3+∠4）+∠DAB=∠DCB＋∠DAB

换言之，联圆四边形的对角和等于两个直角，证讫。

在其后的命题Ⅲ.31 中，欧几里得确定了半圆上的圆周角是直角，其证明已在第一章中介绍。在这一方面，我们注意到，欧几里得在其关于圆的篇章中没有一处讲到半月形的问题，《原本》第三篇也没有讲到我们所熟悉的圆的周长（C=πD）或面积（$A=\pi r^2$）定理。对这些问题的全面探讨，只能等待阿基米德的出现，如第四章所述。

欧几里得的第四篇探讨了某些内接和外切几何图形的问题。像《原本》一书的所有作图一样，他在第四篇的作图中也只限于使用圆规和无刻度直尺。尽管存在着这种限制，但他的确推导出了一些非常复杂的结果。

例如，命题Ⅳ.4 介绍了如何作已知三角形的内切圆，其关键是将三角形角平分线的交点作为内切圆的圆心。在后面的命题中，他又介绍了如何作已知三角形的外接圆；这一次，他将圆心的位置确定在三角形边的垂直平分线的交点上。

由此，欧几里得着手考虑正多边形的作图，此正多边形的所有边长相等，而且，所有角也都相等。这些图形是“完美的”多边形，其对称美无疑吸引了古希腊人的想象力。

让我们回忆一下，欧几里得在《原本》的一开篇便提出了正三角形，或“等边”三角形的作图，在命题Ⅰ.46 中，他在已知线段上作正方形。在命题Ⅳ.11 中，欧几里得扩大了他的作图范围，作了一个圆内接正五边形，而在命题Ⅳ.15 中，他又作了一个圆内接正六边形。本篇的最后一个作图是正十五边形，其推理过程很值得读者一阅。

在一个已知圆中，欧几里得以 AC 为边长作内接等边三角形，并以 AB 为边，作内接正五边形，其顶点均位于 A（图 3.4）。欧几里得注意到，弧 AC 等于圆周长的三分之一，而弧 AB 则等于圆周长的五分之一。因此，这两条弧的差就等于圆周长的$\frac{1}{3}-\frac{1}{5}=\frac{5}{15}-\frac{3}{15}=\frac{2}{15}$。如果我们平分弦 BC，并从 BC 弦的中点作垂线，交圆于 E，这样，我们就平分了弧 BC。因此，弦 BE 等于圆周长的十五分之一，弦 BE 就是正十五边形的边长。沿着圆周复制 15 条 BE 弦，我们就完成了正十五边形的作图。

欧几里得在《原本》中没有再讲正多边形的作图问题，但他显然知道，如果一个人作出这样一个多边形，利用上述平分方法，就一定能够作出边多一倍的正多边形。例如，作出一个等边三角形后，古希腊几何学家就能够据此作出正 6、12、24、48……边多边形；据正方形，他们就能够作出

正 8、16、32、64……边形；据正五边形，可以作出正 10、20、40……边形；据欧几里得的最后一个作图（即正十五边形），可以产生正 30、60、120……边形。

如此说来，这类正多边形似乎数目颇多，但显然不是所有的正多边形都可以列入其中。例如，欧几里得在《原本》中没有一处讲到有关正 7 边形、正 9 边形或正 17 边形的作图,因为这些正多边形不适合上述规整的‘翻番”模式。人们相信，古希借人一定付出了许多时间和努力，试图解决其他正多边形的作图问题，但是，他们的努力显然没有成功。实际上，虽然欧几里得没有明确表示，但其后的大多数数学家都认为，他所提到的是仅有的可以作出的正多边形，而其他任何正多边形都超出了圆规和直尺的作图能力。

所以，当十几岁的卡尔·弗里德里希·高斯于 1796 年发现了正十七边形的作图方法时，无疑引起了巨大震憾。这一发现标志着年轻的高斯不愧为第一流的数学天才。前一章曾介绍过高斯在非欧几何方面的工作，我们还将在第十章详细介绍这位天才的数学家。

总而言之，《原本》第一至四篇讲述了有关三角形、多边形、圆及正多边形的基本定理。截止到目前，欧几里得不曾借助非常有用的相似性概念，尽其可能探讨了几何领域。我们在第一章曾讲过，由于毕达哥拉斯派发现了不可通约量，使对相似性的论证及其所产生的比例问题受到了致命的打击，最后，是欧多克索斯以其完美的比例理论堵塞了这一逻辑漏洞。欧几里得的《原本》第五篇则致力于发展欧多克索斯的思想，其意义非常深远，甚至影响到 19 世纪对无理数的思考。但是，第五篇中的许多定理现在都归入了实数系统，这一系统，不管怎样，我们都认为是天经地义的。这样，我们再去讨论第五篇中艰涩的论证，就显得有点儿多余，因而，我们转向第六篇。

欧几里得在第六篇中研究了平面几何的相似形问题。他的相似形定义是非常重要的。

■ 定义 Ⅵ.1　相似直线图形的对应角相等且对应边成比例。

这一定义具有双重意思，既要求对应角相等，又要求对应边成比例，这样才能保证图形相似。用非技术性语言说，就是在我们说这两个图形形状相同的时候，就已包含了我们所说的这两个条件。总之，这两个条件显然都是必要的。例如，在图 3.5 中，长方形与正方形的角都相等，但它们的边不成比例，所以，形状不相同。另一方面，正方形与菱形的边成同一比例，即 1∶1，但它们的角不相等，所以，形状也完全不同。

有趣的是，如果我们的注意力只局限于三角形的时候，相似性的这两个条件就突然消失了。欧几里得利用根据第五篇中的欧多克索斯理论，在命题Ⅵ.4 中证明，如果两个三角形的对应角相等，则其对应边必成比例；反之，他在命题Ⅵ.5 中证明，如果两个三角形的对应边成比例，则它们的对应角也必定相等。总之，对于三边图形来说，整个问题都极大地简化了，因为两个相似条件中的任何一个条件都保证了另一个条件的成立。因此，三角形占了欧几里得相似图形的大部分是不足为怪的。

其中一个重要的结果是命题Ⅵ.8。

命题Ⅵ.8　一个直角三角形，如果从直角作斜边的垂线，则垂线两边的三角形分别与整个三角形相似，并互相相似。

证明　根据前面第六篇中的命题，事情是很简单的。在图 3.6 中，△BAC 与△BDA 分别含有直角∠BAC 与∠BDA,并共同拥有∠1。根据命题Ⅰ.32，它们的第三个角也相等。因此，根据命题Ⅵ.4，其边成比例，所以，△BAC 与△BDA 相似。同理，三角形 BAC 与 ADC 相似，并由此证明两个小三角形 BDA 与 ADC 相似。证讫。

欧几里得以第六篇的第 33 个命题和最后一个命题,基本完成了他对平面几何的论述。然而，令那些将《原本》仅仅看作几何教科书的人们常常感到奇怪的是，他竟洋洋洒洒地又写出了后面的七篇。接下来的主题堪称后代数学家的一座金矿，犹如任何数学分支一样，丰富而辉煌。这就是数论，我们将在这里发现他的下一个伟大定理。

欧几里得数论

乍看之下,人们可能会以为整数完全无足轻重。毕竟,像 1+1=2 或 2+1=3 这类问题的确不是什么难事，与错综复杂的平面几何相比，更显得平淡无奇。但是，对数论的任何肤浅认识必定很快被摈弃，因为这一数学领域产生了许多富于刺激性的难题，向一代又一代的数学家提出了挑战。我们在欧几里得《原本》的第七至第九篇中发现了对数论最古老的重要阐述。

第七篇首先提出了有关整数性质的 22 个新定义。例如，欧几里得定义偶数为可以平均分为两部分的数，奇数则不可平均分为两部分。第七篇中一个重要的定义是素数定义，即，一个大于 1，且只能被 1 和其自身除尽的数。例如，2、3、5、7 和 11 都是素数。大于 1 的非素数叫合数；每一个合数都有除 1 和其自身以外的整数因子。排在前面的几个合数有 4、6、8、9、10 和 12。顺便说一句，数字 1，既不是素数，也不是合数。

除此以外，欧几里得还定义完全数为等于其各“部分”（即真因数）之和的数。所以，数字 6 是完全数，因为它的真因数是 1、2 和 3（我们排除 6 作为其自身因数，因为我们只要求真因数），显然，1+2+3=6。另一个完全数是 28，因为其真因数的和为 1+2+4+7+14=28。另一方面，像数字 15 就不符合要求，因为其真因数的和为 1+3+5=9≠15，因此，数字 15 显然是不完全数。完全数问题很久以来就一直对数字学家和其他伪科学家有一种特别的吸引力，他们没完没了地找寻 6 和 28 一类的数字，哗众取宠。幸好，欧几里得将他对完全数的研究只限于它们的数学性质。

欧几里得在对他的术语定义之后，随即确立了后人所称的“欧几里得算法”，并据此提出了第七篇的前两个命题。这是一种在两个整数的所有公约数中发现最大公约数的可靠方法。为简要说明欧几里得算法，让我们来确定数字 1387 和 3796 的最大公约数。

首先，用大数除以小数，并记下余数。本例即

$$3796=(1387\times 2)+1022$$

然后，用第一个余数 1022 去除第一个因数 1387，得

$$1387=(1022\times 1)+365$$

以此类推，这次用第二个余数 365 除 1022：

$$1022=(365\times 2)+292$$ 然后

$$365=(292\times 1)+73$$ 最后

$$292=(73\times 4)$$

最后，余数等于 0。

在余数等于 0 以后，欧几里得即断言，前一个余数（本例即 73）就是我们最初两个数字 1387 和 3796 的最大公约数,他对此作出了圆满的证明。请注意，他的这种演算方法最后必然会终止，因为余数（1022、365、292 和 73）越来越小。由于我们所研究的是整数，所以，这种演算过程当然不可能永远进行下去。实际上，当第一个余数等于 1022 时，我们就可以绝对肯定地说，最多还有 1023 步，余数就可以为 0（当然，实际上只用了 5 步）。

显然，欧几里得算法有其具体应用，而且，完全是机械性的。它不需要特别的知识和灵感，就能够确定两个数的最大公约数；当然，不难编一套程序，用计算机来进行这一运算。或许不那么明显的是，欧几里得算法在数论方面有其极大的理论重要性，至今仍被尊为数论的奠基石。

欧几里得对数论的探讨贯穿第七篇始终。在此过程中，他提出了极重要的命题 VII.30。这一命题证明，如果一个素数 p 能够整除两整数 a 与 b 的乘积，则素数 p 至少必能整除两因子之一。例如，素数 17 可以整除 2720=34×80，显然，17 也可以整除第一个因子 34。相反，合数 12 可以整除 48=8×6，但 12 却不能整除 8 或 6 这两个因子中的任何一个。当然，问题就在于 12 不是素数。

命题 VII.31 极大影响了后来的伟大定理。欧几里得的证明与现代数论教科书中的证明完全一致。其证明如下：

命题 VII.31　任一合数均能为某一素数量尽（即可被素数所除）。

证明　设 A 为合数。根据“合数”的定义，一定有一个小于 A，且能整除 A 的数字 B，即 1＜B＜A。这里，B 可能是素数，也可能不是。如果 B 是素数，那么，就正如命题所论断，原数字 A 确实有一个素数因子。另一方面，如果 B 不是素数，那么，B 就一定有一个因子，比如 C，而 1＜C＜B。如果 C 是素数，那么，按照上述推论，C 能够整除 B，B 又能够整除 A，因此，素数 C 自己也能整除 A。但是，如果 C 是合数又会如何呢？那么，它就一定会有一个真因数 D，然后，我们继续下去。

在最坏的情况下，我们会得到一系列降值排列的非素数因子：

$$A>B>C>D>\cdots\cdots>1$$

但是，所有这些数字都是正整数。欧几里得正确指出，我们一定会达到一点，在这一点上，我们所发现的因子是素数，因为“……如果找不到（一个素数因子），那么，就会有无穷多一系列越来越小的数量尽 A，这对于（整）数来说是不可能的。”当然，这种不可能的原因很简单，因为一条降值排列的正整数数字链中的数字是有限的。因此，我们可以非常肯定地说，这种推算过程一定会终止，数字链上的最后一个数字一定是一个素数，同时也是它之前所有数字的因子，特别也是原数字 A 的一个因子。证讫。

无论是在这个命题中，还是在他的算法中，欧几里得都提出了一个重要的概念，即对于任意整数 n 来说，一个降值排列且小于 n 的正整数序列一定是有限的。但是，如果我们把范围扩大到分数，这种概念当然就不正确了，因为降值排列的正分数序列，即 $\frac{1}{2}>\frac{1}{3}>\frac{1}{4}>\frac{1}{5}>\cdots\cdots$ 是无限的。并且，如果我们允许负整数出现，那么，一个降值排列的数序也是无限的，即：

32＞22＞12＞2＞-8＞-18＞-28＞……

但是，如果我们将注意力只限于正整数，如同欧几里得那样，那么，这种降值排列的数序就必定会在有限的步骤之后结束，并且，由此形成了欧氏数论演绎的许多奥妙。

欧几里得完成第七篇的最后一个证明，便毫不犹豫地转向第八篇。实际上，对于欧几里得何以不将他的三部数论著作合成《原本》中的一篇（虽然会很长），人们说不出什么好的理由。终于，欧几里得提出了重要的命题 IX.14。

命题 IX.14　如果一个数是能为一些素数量尽的最小的数，那么，除了原来能量尽它的这些素数以外，不能再为别的素数所量尽。

用现代话说，这条命题的意思就是，一个数只能以唯一的方式分解成素数的乘积。也就是说，我们只要将一个数分解（“量尽”）为素因子，那么，再去寻找不同素数组成的因式已毫无意义，因为其它任何素数都不能量尽原来的数。今天，我们称这一命题为“唯一析因定理”或“算术基本定理”。这后一个名称表明了它在数论中的中心作用，因为数论有时也称“高等算术”。

例如，可应用唯一析因定理解决下述小问题。我们首先设数字 8，然后作升冥排列：$8^2=64$；$8^3=512$；$8^4=4096$；$8^5=32,786$ 等等。我们将继续排列，直到找到一个尾数为“0”的数字为止。问题是，需要经过多少步骤才能得出这一数字，一百步，一千步，还是一百万步？

应用唯一析因定理，我们就会知道，这个问题是完全没有希望的。因为假设这一过程最后能够得出一个尾数为 0 的数字 N。一方面，由于 N 是从一系列 8 的连乘中得出的，所以，我们就能够把它分解成一长串 2 的乘积，因为 8=2×2×2。但是，如果 N 的尾数是 0，它就一定能够被 10 整除，因而也一定能够被素数 5 整除。但这样就出现了矛盾，因为欧几里得在命题 IX.14 中证明，如果 N 分解为一系列因数 2，那么，其它任何素数（也包括素数 5）都不能整除 N。总之，即使我们连续乘以 8，乘上一亿年，也永远得不出一个尾数为 0 的数字。

我们从前面的许多命题中可以清楚地看出，素数在数论中起着一种中心作用。尤其是，因为任何大于 1 的数，或者本身就是素数，或者可以以唯一的方式写成素数的乘积，所以，我们可以很恰当地将素数看作建筑整数大厦的砖石。在这个意义上，数学之素数犹如基础化学之原子，都是同样值得认真研究的。

在欧几里得之前，曾有许多数学家列出过最初的一些素数，以寻找素数的分布模式或其它分布线索。为便于参考，前 36 个最小的素数列举如下：

2，3，5，7，11，13，17，19，23，29，31，
37，41，43，47，53，59，61，67，71，73，
79，83，89，97，101，103，107，109，113，
127，131，137，139，149，151

没有什么特定的分布模式，但有一个明显的特点，即除 2 以外，所有素数都是奇数（因为所有大于 2 的偶数都有因子 2）。然而，仔细观察一下便会发现，随着数值的增大，素数之间的“跨度”似乎越来越大，或者说素数越来越稀少。例如，在 2 与 20 之间有 8 个素数，但在 102 与 120 之间却

只有 4 个素数。注意在 113 与 127 之间连续有 13 个合数，而在 100 之内却没有这样大的素数间隔。

对于素数明显趋于“稀少”的现象，不难作出解释。显然，在我们观察比较小的数字（十几或二十几）时，由于小于这些数字的数很少，所以可能因子也很少。但对于比较大的数字（如上百、上千或上百万）来说，则有很多小于它们的数字可以充当潜在因子。一个数作为素数，必须没有更小的因子，一个大数很难作到这一点，因为它具有许多小于自己的可能因子。

实际上，我们如果一直追踪下去，就会发现在素数之间的巨大间隔。例如，在从 2101 到 2200 这 100 个数字中，只有 10 个数是素数，而在从 10，000，001 到 10，000，100 这 100 个数字中，只有两个数是素数。也许，古希腊人曾像今天的学生一样，想到过素数最终可能会有尽头。也就是说，最后，素数会变得如此稀少，以致完全消失不见，而后边的所有数字都成了合数。

即使在这一方面存在某些迹象，也不足以动摇欧几里得的论断。相反，他在命题 IX.20 中证明，尽管素数数量越来越稀少，但任何限定的素数集合都不可能将所有的素数囊括无遗。他的论断常常被称为对“素数无穷性”的证明，因为他的确证明了全部素数是无限的。如果世界上确实有经典性的伟大定理的话，那么，欧几里得的论证即是一例。实际上，他的论证常常被人们作为数学定理的典范，因为这一定理简洁、优美，又极为深刻。20 世纪英国数学家 G.H.哈代（1877—1947 年）在其精彩的专著《一个数学家的自辩》中称欧几里得的证明“……自发现之日至今，永葆其生机与效力，两千年岁月没有使它产生一丝陈旧感。”

伟大的定理：素数的无穷性

现在，我们已讨论了欧几里得作出他巧妙证明所需要的几乎全部概念，唯有一点还未讨论。这尚缺的一点是一个非常简单的概念，即如果一个整数 G 可以整除 N 和 M，且 N＞M，那么，G 就一定能够整除这两个数的差 N—M。显然，因为 G 能够整除 N，即 N=G×A，A 为整数；又，G 能够整除 M，即 M=G×B，B 为整数。所以，N-M=G×A-G×B=G×（A-B）。由于 A-B 为整数，因此，G 显然能够整除 N-M。也就是说，两个 5 的倍数相减，其差等于一个 5 的倍数；两个 8 的倍数之差等于一个 8 的倍数；等等。

根据这一明显的原理，我们就可以进而研究欧几里得的经典命题。

命题 IX.20　素数的数目大于任何指定的素数集合。

欧几里得的特别术语又一次模糊了命题的含义。他所说的是，已知任何有限的素数集合（即任何“指定的素数集合”），然而，我们总能找到一个不包括在这一素数集合之中的素数。简言之，任何有限的素数集合都不可能包括全部素数。

证明　欧几里得首先设一有限的素数集合，即 A、B、C、……，D。他的目的是要找到一个不同于所有这些素数的素数。为此，第一步，他先设数字 N 一（A×B×C×……×D）＋1。N 大于原素数集合中所有素数的乘积，

显然也大于其中的任何素数。如同任何大于 1 的数字，N 或是素数，或是合数，对这两种情况，需要分别加以讨论。

情况 1　设 N 为素数。

因为 N 大于 A、B、C、……，D，所以，N 是原素数集合中不包括的新素数，至此，证明完毕。

情况 2　如果 N 是合数，情况又会如何呢？

根据命题 VII.31，N 肯定有一个素数因子，我们设其为 G。然后，欧几里得即断定（这是他推理的核心），G 为原“指定的素数集合”之外的素数。为便于论证，设 G=A，那么，G 当然能够整除 A×B×C×……×D 的积，并且，（如我们在情况 2 中所设，）G 同时也能够整除 N。因此，G 肯定还能整除这些数的差，即应该能整除

N-（A×B×C×……×D）

=（A×B×C×……×D）+1-（A×B×C×……×D）=1

但是，这是不可能的，因为素数 G 最小也必须等于 2，而且，根本没有能够整除 1 的数字。即使我们假设 G=B，或 G=C，等等，结果也都一样。因此，欧几里得宣称，素数 G 不包括在 A、B、C、……，D 之中。

所以，不论 N 是否是素数，我们都能够找到一个新的素数。因此，任何有限的素数集合永远会被素数集合之外的又一个素数所补充。　证讫。

对欧几里得证明的要点可以用两个具体的数字来说明。例如，假设我们原来“指定的素数集合”是{2，3，5}。那么，数字 N=（2×3×5）+1=31，N 为素数。31 显然大于我们开始时所设的三个素数 2、3 和 5，因此，31 是不包括在原素数集合中的新素数。这就是我们上面所证明的第一种情况。

另一方面，我们还可以设原素数集合为{3，5，7}，因而，N=（3×5×7）+1=106。106 显然大于 3、5 或 7，但它不是素数。然而，犹如第二种情况所证明的那样，106 肯定有一个素数因子，在本例中，106=2×53，而 2 和 53 都是不包括在集合｛3，5，7｝之中的新的素数。所以，即使 N 是合数，我们也能够证明有限的素数集合之外尚有其它素数存在。

这一证明将永远是数学论证的经典之作。但欧几里得却未能很好地处理他的数论研究。他证明了几个平淡的命题，如两个奇数之差是偶数等，然后，便以关于完全数的命题结束了第九篇。其实，他在第七篇的开始就曾对完全数的概念作出过定义，但后来似乎完全忘记了。终于，在第九篇的结尾处，完全数又重现了。

命题 IX.36　如果从某一单位开始有任意多个数连续成倍比，直到各数之和成为素数；并用此和乘以最末一个数，则乘积一定是完全数。

我们可以用现代术语更准确地说明欧几里得的意思：如果从 1 开始，一些 2 的几何级数项之和 $1+2+4+8+\cdots\cdots+2^n$ 是素数，则数字 $N=2^n(1+2+4+8+\cdots\cdots+2^n)$，即“最末”一个被加数 2^n 与这些数的和 1+2+4+8+……+2■的乘积，一定是一个完全数。

我们不必看欧几里得对这一命题的证明，只需看一两个具体数例即可。例如，1+2+4=7 是素数，根据欧几里得定理，数字 N= 4×7= 28 是完

全数。当然，我们已经证实了这一点。另一个例子， 1+2+4+8+16=31，是一个素数，那么，N=16×31=496 也应该是完全数。为证明这一点，我们先列出 496 的真因数，即 1、2、4、8、16、31、62、124 和 248，它们相加的和等于 496，完全符合定义。

顺便提请读者注意，1+2+4+……2^n 式中的数字不一定都是素数。例如，1+2+4+8=15 或 1+2+4+8+16+32=63 就都是合数。欧几里得的完全数定理只能应用于那些其和恰好是素数的特殊情况。这样的素数，如 7 和 31，我们今天称之为“梅森素数”，以纪念法国教士马兰·梅森（1588—1648 年），他曾在 1644 年的一篇论文中讨论过这一题目。梅森素数因其与完全数的关联，时至今日，仍对数论学家有着特别的吸引力。

总之，欧几里得以其对命题 IX.36 的证明，为我们提供了一个构造完全数的绝好方法。我们将在后记中回到这一问题上来，并讨论其发展现状。

《原本》的最后几篇

从第七篇至第九篇，欧几里得共证明了 102 个有关整数的命题。然后，他在第十篇中突然改变方向，使第十篇成为《原本》十三篇中篇幅最长，并且，许多人认为，在数学上也是最复杂的一篇。欧几里得在第十篇的 115 个命题中，彻底阐述了不可通约量的问题，这个问题，我们今天可以用实数的平方根来表示。这些十分微妙的问题，有许多在技术上都是非常复杂的，涉及到许多需要慎重定义和验证的概念。试举一例：

命题 X.96　如果一个面是由一个有理线段和一个第 6 余线构成的，那么，与此面相等的正方形的边是一个两中项面差的边。

显然，弄清欧几里得诸如“余线”和“中项”这些术语的含义，进而理解他的这一命题，已经需要花费一些功夫，更不要说去弄清其后的证明了。对于现代读者来说，他的许多命题都已过时，因为这些问题现在用有理数与无理数系统就都可以很容易地解决。

《原本》的第十一至第十三篇探讨了关于立体几何或三维几何的基本原理。例如，第十一篇有 39 个命题研究了有关相交平面与相交平面角一类的立体几何问题。其中一个重要的命题是命题 XI.21，在这一命题中，欧几里得提出了“立体角”（即三维角）的概念，例如，棱锥的顶角就是由三个或三个以上平面角会聚于一点形成的。欧几里得证明，会聚于棱锥顶点的所有平面角的和小于四个直角。虽然我们不必验证欧几里得巧妙的证明，但我们可以完全相信其命题，因为我们知道，一个由四个直角（用现代术语说，即 360°）组成的立体角，其平面角可以“压扁”成一个平面，因而也就完全没有角了。命题 XI.21 将在《原本》最后一篇的最后一个命题中起到重要作用。

如果说第十一篇只涉及立体几何的基本命题，则第十二篇就进行了更深入的探讨。在第十二篇中，欧几里得应用了欧多克索斯的穷竭法来阐述锥体体积等问题。

命题 XII.10　任何圆锥体的体积都等于与其同底等高柱体体积的三分之一（图 3.7）。

今天，我们可以用公式来表示这一命题。我们知道，一个半径为 r，高为 h 的圆柱体，其体积等于πr^2h，因此，欧几里得所说的圆锥体的体积就是$\frac{1}{3}\pi r^2 h$。他的精彩论证不仅证实了欧几里得的论证技巧，而且，也证实了最初发现者欧多克索斯的正确。许多年以后，阿基米德将这一命题归于欧多克索斯的名下，并评述说：

“……虽然这些性质始终是这些图形自然固有的，但在欧多克索斯之前，众多有才华的几何学家实际上并不知晓，而且也没有任何人去注意这些性质。”

在第十二篇中，还有另外两个非常重要的定理值得一提。其一是命题 XII.2，令人惊奇的是，这是一个关于平面图形圆的定理。

命题 XII.2　圆与圆的面积之比等于其直径平方之比。

我们在前面讨论希波克拉底求新月形面积时曾见过这个命题。如前所述，这一命题提供了一个比较两个圆面积的方法，而不是已知直径或半径求一个圆的面积。

现在，我们从略为不同的另一角度来看命题 XII.2。设两个圆，一个圆的面积为 A_1，直径为 D_1；另一个圆的面积是 A_2，直径是 D_2，我们得出

$$\frac{A_1}{A_2}=\frac{D_1^2}{D_2^2} \text{ 或等价式 } \frac{A_1}{D_1^2}=\frac{A_2}{D_2^2}$$

这一等式告诉我们，无论多大的圆，圆的面积与其直径的平方比总是一定的，数学家称这一比例为“常数”。这是一个非常重要的性质。但欧几里得未能对这一常数做出数值估计，也未能确立这一常数与我们在研究圆的过程中所遇到的其他重要常数之间的关系。总之，命题 XII.2 尽管给人深刻印象，仍有许多改进的余地。我们后面将介绍经阿基米德改进的这一命题，也即第四章的伟大定理。

与其相似的是第十二篇的最后一个命题，这一命题通过穷竭法，证明了“两个球体的体积之比等于其直径的三重比”。用现代术语说，这一有关球体体积的相对关系可以简单地表示为

$$\frac{V_1}{D_1^3}=\frac{V_2}{D_2^3}$$

（注意，所谓“三重比”是古希腊人的说法，我们今天称之为立方。）欧几里得在这一命题中又提出了另一个重要常数——这一次是球体积与其直径的立方比，但欧几里得依然未能提出这一常数的数值估计。阿基米德于公元前 225 年在其无可争议的杰作《论球体和圆柱体》一书中再次解决了这一问题，对此，读者或许不应感到惊奇。

最后，我们来讨论第十三篇，也是欧几里得《原本》的最后一篇。他在这一篇的 18 个命题中探讨了所谓三维几何的“正立体”及其相互之间绝妙的联系。一个正立体，其所有构成平面应当都是全等的正多边形。我们最熟悉的正立体是立方体，即六面立体，其中每一个平面都是一个正四边形——即正方形。对于古希腊人来说，正立体体现了一种三维的美与对称，因此，认识正立体显然是他们优先考虑的问题。

在欧几里得时代，有五种正立体已为人们所认识——四面体（四面都是等边三角形的角锥体）、立方体、八面体（八面都是等边三角形）、十二面体（十二面都是正五边形）和五种中最复杂的二十面体（由等边三角形构成的二十面立体）。

对这些给人以美感的正立体（如图 3.8 所示），柏拉图于公元前约 350 年在其《梯迈乌斯篇》中曾刻意描述。柏拉图在书中特别考虑了构成世界的四大“元素”——火、气、水、土的性质。柏拉图认为，这四种元素显然都是物体，而所有物体都是立体。由于世界只能由完美的物体构成，所以，显然（至少对柏拉图是这样），火、气、水和土都一定是正立体，而剩下的问题只是要确定哪种元素呈哪种形状。

柏拉图在他的论证中提出了一个引人发笑的伪数学论断：“……气与水之比等于水与土之比”。对此，他的最后说明如下：

火是四面体形状，因为火是四大元素中最小、最轻、最活跃和最锐利的物体，而四面体正适合火的这些特性。柏拉图说，土一定是立方体，因为立方体是五种立体中最稳定的形体；而水是四大元素中最活动的流体，其形状，或“种子”，一定是二十面体，因为这种形状最接近于球体，完全有可能轻松滚动。气在大小、重量和流动性方面都居中，所以，是由八面体构成的。柏拉图说：“我们必须想象所有这四种物体，每一种单位都非常小，肉眼看不见，只有大量聚集在一起时，我们才能分辨。”

然而，使柏拉图感到为难的是，这四大元素一一讲完，却还剩下一个正立体——十二面体没有去处。他强辩说，十二面体是“……上帝用来安排满天星座的”。换言之，十二面体代表了宇宙的形状。《梯迈乌斯篇》中的这一理论，即使算不得荒诞，也不免纯属空想，而这些正立体也从此被称为“柏拉图立体”。想到欧几里得据认为曾在柏拉图的雅典学园中学习过，就可以推测出这五种正立体对欧几里得有多么大的吸引力，以致要用它们来结束《原本》。

众所周知，几何学家很久以来即已知晓这五种正立体的存在。在《原本》的第 465 个命题和最后一个命题中，欧几里得证明了再没有其他正立体，几何学限定这些优美的立体形状只有五种，不多也不少。对此的简单证明乃是依据命题 XI.21。他只要依据构成任何立体角的平面角之和一定小于四个直角或（用现代术语）小于 360°的限制，来考虑构成正立体平面的多边形形状就可以了。

假设正立体的每一个平面都是等边三角形，因此，每一个平面角都是 60°。当然，立体角一定是由三个或三个以上的平面相交形成的，因此，最小的立体角是由三个等边三角形构成立体的每一个顶角时形成的，因为三个角的和为 3×60°=180°。这就是四面体。

我们还可以考虑立体的每一个顶角由四个等边三角形组成，因为四个角的和是 4×60°=240°（八面体）；或者，每一个顶角由五个等边三角形组成，因为五个角的和为 5×60°=300°（二十面体）。但是，如果我们让每一个项角由六个或六个以上的等边三角形组成，则平面角之和至少等于 6×60°=360°，而这样就违反了命题 XI.21 的限定。所以，以等边三角形为平面，不可能构成其他类型的正立体。

那么，正方形平面的正立体又如何呢？当然，正方形的每一个角等于 90°，所以，三个正方形相交组成一个立体角，其平面角之和为 3×90°

=270°；这就是立方体。但是，如果由四个或四个以上正方形组成一个立体角，则平面角的和又至少等于 4×90°=360°，而这又是不可能的。因此，没有其他正立体能够具有正方形平面。

同样，正立体的平面还可能是正五边形。因为正五边形的每一个内角等于 108°，所以，可以有三个正五边形组成一个立体角

（3×108°=324°<360°），但不能更多。这种正立体就是十二面体。

如果我们试图用正六边形、正七边形、正八边形等等正多边形平面组成正立体，则每一个平面角至少 120°，即使由三个角组成一个最小的立体角，每一个立体角也会等于或大于 360°。用欧几里得的话说，“由于同样不合理，一个立体角不能由其他多边形（超过正五边形的正多边形）构成。”

总而言之，欧几里得证明，除了这五种正立体外，不可能再有其他正立体存在——这五种正立体，有三种的平面是等边三角形，一种是正方形，一种是正五边形。任何努力和机巧都不可能产生此外的任何正立体。

至此，《原本》全部结束。2300 年来，《原本》始终是一部卓越的数学文献。《原本》像所有经典著作一样，即使一读再读，作者的天才依然令人玩味。时至今日，读者仍能从其精妙的数学推理技巧中获得无穷的乐趣。我们最好还是引述托马斯·希思爵士的话来说明，他简洁明了地评价说：《原本》“……现在是，并无疑将永远是一部最伟大的数学教科书。”

后记

本章的伟大定理涉及到数论问题，因此，现在，我们不妨看一看在这一迷人的数学分支中一些十分重要而又常常引起争议的问题。数论的一个真正诱惑是它的猜想简单得甚至连小学生都能看懂，然而却使一代又一代世界一流数学家为它付出了艰苦的努力。这似乎就是这一数学分支看似反常的特点。

例如，“孪生素数”现象曾引起过数学家们极大的兴趣。所谓孪生素数，即两个相差 2 的相邻素数，如 3 和 5，或 11 和 13，或 101 和 103 等。像素数本身一样，随着数值的不断增大，孪生素数也变得越来越稀少，人们随之必然会提出这样一个问题，“孪生素数的数量是有限的吗？”

这是一个十分简单的问题。并且，它于欧几里得 2300 年前在命题 IX.2。中解决的问题很相似，似乎并不难回答。但是，直至今天，还没有一位数学家知道这个孪生素数问题的答案。也许，正如大多数数学家所猜想的那样，孪生素数的数量是无限的，但是，迄今为止，还没人能够证明这一点。或者，也许，在某一点之后，我们可以找到最大的一对孪生素数，但对此，同样没人能够证明。总之，情况之复杂，即使对欧几里得本人来说，也不过如此。这不免令人沉重而又沮丧。

数论中还有其他一些饶有兴趣，但却未能解决的难题。我们曾介绍过欧几里得的一个证明：如果括号中的项是素数，那么，任何可以写成

$$2^n（1+2+4+8+\cdots\cdots+2^n）$$

形式的数字都是完全数。然而，他没有说这是唯一的完全数形式（但也没有说不是）。因此，许多数学家都曾试图发现不同于欧几里得公式的完全数。

迄今为止，人们仍然劳而无功。18 世纪数学家莱昂哈尔德·欧拉在其遗作中证明，任何偶完全数都一定适合欧几里得的公式，即，如果 N 是一个偶完全数，那么，就一定存在一个正整数 n，使得 $N=2^n$（1+2+4+8+……+2^n）

这里，括号中的项必定是（梅森）素数。

欧几里得与欧拉合力，已完全解开了偶完全数的谜。剩下的全部问题是确定奇完全数的形式。遗憾的是，至今还没有人发现奇完全数。时至今日，究竟是否存在奇完全数，还是一个难解的谜。当然，这并不等于说，人们没有去寻找。几百年来艰苦的理论研究，特别是最近，利用高速计算机进行的理论研究，都未能发现一个既是奇数，又是完全数的整数，但这当然并不意味着，这种难以想象的奇完全数根本不存在。

数学家进退两难。他们既不能发现奇完全数，又无法证明奇完全数不存在。然而，这种困境却也产生了一种诱人的可能性。也许，有一天，有人会证明奇完全数根本不存在，这样，所有完全数就都是偶完全数，犹如欧几里得所示，因而，所有完全数都适合欧几里得的公式。如果是这种结果，则伟大的欧几里得在公元前 300 年时便已确立了囊括世界全部完全数的公式。果真如此，那将是一个非常了不起的转折。

本章以所有数论问题中一个最棘手的问题作为结束，即所谓“哥德巴赫猜想”。这一猜想最初出现于 1742 年一个数学迷克里斯蒂安·哥德巴赫（1690—1764 年）的一封信中。哥德巴赫名声大振的主要原因就是他寄给欧拉的这封信。他在信中猜想，任何大于或等于 4 的偶数都可以表示成两个素数之和。欧拉倾向赞同哥德巴赫的猜想，但却苦于不知如何去证明。

像我们研究其他许多数论难题那样，我们不难用小数字来验证哥德巴赫的猜想。例如，4=2＋2，28=23＋5，以及 96=89＋7。哥德巴赫猜想特别引人注目，原因是它只涉及一些极为简单的概念，其仅有的几个技术性术语只是“偶数”、“素数”和“和”，而这几个术语的意思几分钟之内就可以给小孩子们讲明白。但是，自从哥德巴赫 250 年前寄出这封信以后，他的猜想却至今未能得到证明。

曾对哥德巴赫猜想作出了特殊贡献的是苏联数学家 L.什尼尔里曼。据数学史家霍华德·伊夫斯记载，什尼尔里曼于 1931 年证明，任何偶数都可以写成不多于 300，000 个素数和的形式。鉴于哥德巴赫猜想要求仅用两个素数相加即得到任何偶数，什尼尔里曼的证明实际上整整多了 299，998 个素数。

在某种意义上说，什尼尔里曼的 300，000 个素数似乎是数学家的败绩，但同时也表明，虽然历史上曾经有过欧几里得与欧拉，但如今仍然有大量伟大的定理以其永恒的耐心在等待着证明。

第四章
阿基米德的求圆面积定理
（公元前约 225 年）

阿基米德生平

从欧几里得到我们将要介绍的下一位伟大数学家——叙拉古城举世无双的阿基米德（公元前 287—212 年）之间，经历了两三代人之久。阿基米德在其辉煌的数学生涯中，将数学疆界从欧几里得时代向前推进了一大步。实际上，此后将近两千年，数学界再没有出现过像阿基米德这样伟大的数学家。

我们有幸了解一些阿基米德的生平，但因为历经沧桑，其细节的真伪往往受到怀疑。同时，他的一些数学著作也有幸流传下来，而且有他自己的注解。所有这些资料，为我们描绘了这位曾经统治古代数学界，受人尊敬，但又有点儿古怪的数学天才的一生。

阿基米德出生于西西里岛的叙拉古城。据说，他的父亲是一位天文学家，阿基米德从小就萌发了研究宇宙的兴趣，终生乐此不疲。阿基米德青年时代也曾到过埃及求学，并在亚历山大图书馆学习。这里曾是欧几里得治学之处，阿基米德自然也会受到欧几里得的影响，这一点在阿基米德的数学著作中可以很清楚地看出。

据说，阿基米德在尼罗河谷期间，曾发明了所谓‘ 阿基米德螺旋水车 ”，这种装置可以用来把水从低处提到高处。有趣的是，这一发明，直至今日仍在使用。他的发明证明了阿基米德的双重天才：他既可以脚踏实地地研究实际问题，又能够在最抽象、最微妙的领域中探索。亚历山大显然适合发挥他的才干，但阿基米德还是返回了他的故乡叙拉古城，据我们所知，就在那里度过了他的后半生。叙拉古城虽然十分闭塞，但阿基米德一直保持着与全希腊，特别是与亚历山大学者们的通信联系。这种书信往来，使得阿基米德的许多著作得以保存。

阿基米德能够在一段时间里非常专注地研究任何问题，更加提高了他令人敬仰的数学才能。他在进行研究时，常常会忽略日常的生活问题。我们从普卢塔克的著作中得知，阿基米德

“ ……忘记了吃饭，甚至忘记了他自己的存在，有时，人们会强制他洗浴或敷油，他都浑然不知，他会在火烧过的灰烬中，甚至在身上涂的油膏中寻找几何图形，完全进入了一种忘我的境界，更确切些说，他已如醉如痴地沉浸在对科学的热爱之中。 ”

这一段文字描绘了这位数学家心不在焉的形象，对于阿基米德来说，整洁似乎已与他无关。当然，有关阿基米德“ 心不在焉 ”的故事，最著名的还是关于叙拉古城国王希伦的王冠的故事。国王怀疑金匠用一些合金偷换了他王冠上的黄金，就请阿基米德来测定王冠的真正含金量。正如故事所说，阿基米德一直解不开这道难题，有一天（在他少有的一次洗浴时），他忽然找到了答案。他兴奋得从浴盆里跳出来，跑到叙拉古城的大街上，边跑边欢呼：“ 我找到啦！我找到啦！ ”但遗憾的是，他完全沉浸在他的新发现之中，竟然忘记了还没穿衣服。很难想象街上的人们看见他一丝不

挂地招摇过市，会说些什么。

这个故事也许是杜撰的，但阿基米德发现流体静力学的基本原理却是千真万确的。他留给我们一篇题为《论浮体》的论文，阐述了他在这一方面的思想。除此以外，他还发展了光学，创立了机械学，他不仅发明了水泵，而且还发现了杠杆、滑轮和复式滑轮的工作原理。普卢塔克记叙过这样一个故事：多疑的希伦国王怀疑这些简单机械装置的能力，就请阿基米德实际演习一下。阿基米德以一种戏剧般的方式满足了国王的要求，他选择了国王一艘最大的船只，

“……如果不花费巨大人力，是无法把这艘大船拖离船坞的，况且，船上还满载乘客和货物。阿基米德坐得远远的，手里只握住滑轮的一端，不慌不忙地慢慢拉动绳索，船就平平稳稳地向前滑动，就像在大海里航行一样。”

不用说，国王对此留下了深刻印象。或许，他从这件事中察觉了这位天才科学家的某种宝贵才能，遇有危难关头，这样的工程天才可以派上用场。公元前 212 年，罗马人在马塞卢斯率领下，进攻叙拉古城，危难关头来临。面对罗马的威胁，阿基米德奋起保卫自己的家园，他设计了许多杀伤力很强的武器。他的这项事业，或许只能称为个体军工企业。

我们继续引用普卢塔克的《马塞卢斯生平》一书，这本书是这位伟大的罗马传记作家在事件发生后约 300 年时写的。普卢塔克虽然是在为马塞卢斯作传，但他对阿基米德的钦敬心情却显而易见。这些描述使我们看到了一个非常引人，栩栩如生的阿基米德形象。

“马塞卢斯率领大军向叙拉古城进发，”普卢塔克写道，“并在离城不远处安营扎寨，又派使者进城劝降。”但叙拉古城人拒绝投降，马塞卢斯便凭借陆上的兵士和海上 60 艘装备精良的战船猛扑叙拉古城。马塞卢斯“……有备而来，历年征战，声威赫赫”，但事实却证明他敌不过阿基米德和他凶狠的守城器械。

据普卢塔克记载，罗马军团进逼城垣，自信战无不胜。

“但是，阿基米德开始摆弄他的器械，他对地面部队启动各种弹射武器，无数大小石块带着惊人的呼啸，猛烈地倾泻下来；乱石之中，无人能够站立，士兵乱了阵脚，纷纷被击中，成堆倒下。”

而罗马水师的情况也不见佳，

“……从城墙上伸出了长长的杆子，在船上方投下重物，将一些船只击沉；而其他船只则被一只只铁臂或铁钩钩住船头，提升起来……然后又船尾朝下，投入海底；同时，另一些船只在其引擎的拖动下，团团乱转，最后撞碎在城下突起的尖峭岩石上，船上的士兵死伤惨重。”

这种巨大的伤亡，用普卢塔克的话说，是“一件可怕的事情”，人们不会不同意他的说法。在这种情况下，马塞卢斯认为最好还是先撤退。他撤回了他的地面和海上部队，重新部署。罗马人经过认真研究，决定进行夜袭。他们以为，只要在夜幕掩盖下，贴近城墙，阿基米德的武器就没有用武之地了。然而，罗马人再次遭到了意外的打击。原来，不知疲倦的阿基米德已经为应付这种偷袭作好了充分的安排。罗马士兵一靠近城防，“石头就劈头盖脸地砸下来，同时，城内又射出飞箭”。结果，罗马人失魂落魄，不得不再次撤退，但又受到阿基米德远程武器的攻击，“损兵折将”。这次，自负的罗马军团“看到无形的武器给他们造成的重大伤亡，开始以

为他们是在与诸神作战。”

或许，说马塞卢斯的军队士气低落亦不为过。他希望他这支受到重创的军队能够重振勇气，继续进攻，但是，以前自认为无敌于天下的罗马人却不愿再打了。相反，士兵们“只要看到城墙上伸出一小段绳索或一片木头，就立时大哗，以为阿基米德又对他们使用什么武器了，并转身落荒而逃。”马塞卢斯明白，小心即大勇，于是，他放弃了直接进攻。

马塞卢斯想以断粮逼迫叙拉古城人投降，所以，罗马军团开始长期围困叙拉古城。时间一天天过去，军事态势没有什么变化。后来，在狄安娜节日期间，叙拉古城居民“完全放松了警惕，他们纵酒狂欢”，松懈下来。一直在窥测时机的罗马人乘其不备，一举攻破了防守懈怠的一段城防，怀着一腔怨毒涌入叙拉古城。据说，马塞卢斯环视着这座美丽的城市，为他的士兵不可避免地要对叙拉古城泄怒施暴雨落下了眼泪。的确，据历史记载,罗马人对叙拉古城人的做法完全不亚于他们在 66 年后对迦太基人的暴行。

但是，阿基米德的死使马塞卢斯极为悲伤，因为他对这位天才的对手至为尊敬。据普卢塔克记载：

“……也许是命该如此，（阿基米德）正在专心研究几何图形，他全神贯注地思考，完全没有注意到罗马人的入侵，也没有注意到城市的陷落。正在他聚精会神地研究和思考的时候，没想到一个士兵前来，命令他立刻去见马塞卢斯；但阿基米德在没有解出他的几何证明题之前，拒绝跟他走。士兵大怒，拔出佩剑，一剑刺死了阿基米德。”

就这样，阿基米德走完了他的一生，他死了，像他活着时一样，执着于他所喜爱的数学。我们可以认为他是一位科学研究的殉难者，也可以认为他是自己无暇它顾的牺牲者。总之，古往今来，数学家不知有多少，但像阿基米德这样结局者，却是绝无仅有的。

阿基米德尽管发明了许多利器和工具，但他真正喜爱的还是纯数学。与他发现的美妙定理相比，他的杠杆、滑轮和石弩都不过是雕虫小技。我们还是引用普卢塔克的话来说明：

“阿基米德具有高尚的情操，深刻的灵魂和丰富的科学知识，虽然这些发明使他赢得了超乎常人的名望，但他并未屈尊留下任何有关这些发明的著述；相反，他却鄙薄工程学这一行当，以及任何仅仅出于实用和赢利目的的技艺，他将他的全部情感与理想寄托于与尘世无涉的思索之中。”

数学是阿基米德的最大遗产。在这一领域，阿基米德无可争议地被公认为古代最伟大的数学家。他的那些幸存下来的十几部著作及一些零散的文稿是最高质量的。其逻辑上的严谨与复杂，令后人惊叹不已。毫不奇怪，他一定非常精通欧几里得的理论并不愧为欧多克索斯穷竭法的大师；借用牛顿的名言，阿基米德一定是站在巨人的肩上。但是，过去的影响虽然很大，却不能充分解释阿基米德带给数学学科的巨大发展。

伟大的定理：求圆面积

公元前约 225 年，阿基米德发表了一篇题为《圆的测定》的论文，这篇论文中的第一个命题对圆面积作了十分透彻的分析。但是，在我们讲述这一不朽之作之前，我们有必要先介绍一下在阿基米德探讨这一问题时，

有关圆面积问题的发展状况。

当时的几何学家已知，不论圆的大小如何，圆的周长与直径之比总是一定的。用现代术语，我们可以说

$$\frac{C_1}{D_1}=\frac{C_2}{D_2}$$

如图 4.1 所示，公式中的 C 代表周长，D 代表直径。

换句话说，圆的周长与直径之比是一个常数，现代数学家定义这一比率为π。（注意：古希腊人在这里不使用符号。）因此，公式

$\frac{C}{D}=\pi$　或其等价式$C=\pi D$

正是表明了常数π的定义，即两个长度（圆的周长与直径）的比。

那么，圆的面积又如何呢？我们已经知道，《原本》的命题Ⅻ.2 证明了两个圆的面积之比等于两圆直径的平方比，因此，圆面积与其直径的平方比是一个常数。用现代术语说，欧几里得证明了常数 k 的存在，因而

$$\frac{A}{D^2}=k\quad\text{或等价式}A=kD^2$$

至此，一切顺利。但是，这两个常数之间相互有什么关系呢？也就是说，人们是否能够发现在这“一维”常数π（表示圆周长与直径的关系）与“二维”常数 k（表示面积与直径的关系）之间存在着一种简单的联系？显然，欧几里得没有发现这种联系。

然而，阿基米德在其短小精炼的论文《圆的测定》中证明了有关结果，而这相当于现代涉及π的求圆面积公式。在证明中，他在圆周长（及因此产生的π）与圆面积之间建立了重要联系。他的证明需要两个非常直接的初步定理和一种非常复杂的逻辑方法，称为双重归谬法（反证法）。

我们先来看这两个初步定理。一个是关于正多边形面积的定理，正多边形的中心为 O，周长为 Q，边心距为 h。这里，边心距是指从多边形的中心引向任何一条边的垂线长度。

定理 正多边形的面积等于$\frac{1}{2}hQ$。

证明 设正多边形（图 4.2）有 n 条边，每条边长 b。作从 O 到每个顶点的连线，将多边形划分为 n 个全等三角形，每个三角形的高为 h（边心距），底为b。因此，每个三角形的面积为$\frac{1}{2}bh$，面积（正多边形）

$=\frac{1}{2}bh+\frac{1}{2}bh+\cdots\cdots+\frac{1}{2}bh$，共有n项

$=\frac{1}{2}h(b+b+\cdots\cdots+b)=\frac{1}{2}hQ$

因为（b+b+……+b）是周长。　证讫。

简单明快。阿基米德的第二个定理当时也非常著名，而且显然是不证自明的。这一定理称，如果给我们一个已知圆，我们可以作圆内接正方形；欧几里得在命题 Ⅳ.6 中已证明过这种作图。当然，正方形的面积肯定小于其外接圆的面积。我们通过平分正方形的每条边，就可以确定圆内接正八

边形的顶点位置。当然，正八边形比正方形更接近于圆的面积。如果我们再平分八边形的每条边，就可以得到圆内接正 16 边形，这当然比八边形又更接近圆的面积。

这一过程可以无限继续。实际上，这种方法的实质就是前面曾提到过的著名的欧多克索斯穷竭法。显然，内接正多边形的面积永远不会等于圆的面积；不论内接正多边形产生多少条边，都永远小于圆的面积。但是（这是穷竭法的关键），如果预先给定任一面积，不论其多小，我们都能作出一个内接正多边形，而使圆面积与其内接正多边形的面积之差小于这一预先给定的面积。例如，如果预先给定的面积为 $\frac{1}{500}$ 平方英寸，我们可以作一个内接正多边形，而使

面积（圆）- 面积（正多边形）$<\frac{1}{500}$ 平方英寸

这一正多边形也许有几百条边或几千条边，但这并不重要，重要的是它存在。

外切正多边形也具有类似的规律。我们可以用一句话来概括这两种正多边形的规律，即，对于任何已知圆，我们都可以作出它的内接正多边形或外切正多边形，其面积可任意接近圆的面积。正是这句“可任意接近”成为了阿基米德成功的关键。

以上就是阿基米德的两个初步命题。下面，我们有必要就他论证两个面积相等时所采用的逻辑方法作一个简单的介绍。在某种意义上，这种逻辑方法比我们以往所见到的任何方法都更复杂，或者说，至少更曲折。例如，我们可以回想一下，欧几里得是如何证明直角三角形斜边上正方形的面积等于两条直角边上正方形面积之和的：他直接推理，证明了问题中的面积相等。他的证明方法虽然非常巧妙，却只是正面论证。

然而，阿基米德在论证更为复杂的圆面积问题时，采用了一种间接证明的方法。他认为，任何两个量 A 与 B，一定只能属于下列三种情况中的一种：A＜B，或 A＞B，或 A=B。为了证明 A=B，阿基米德首先假设 A＜B，并由此推导出逻辑矛盾，因而排除这种情况的可能性。然后，他再假设 A＞B，并再次推导出逻辑矛盾。排除了这两种可能性后，就只剩下了一种可能性，即 A 等于 B。

这就是阿基米德极为精彩的间接证明方法——“双重归谬法”，将三种可能性中的两种引入逻辑矛盾。这种方法初看起来似乎有点绕圈子，但细想一下就会觉得非常合理。排除了三种可能性中的两种，就迫使人们得出结论，只有第三种可能性是正确的。当然，没有人能比阿基米德更熟练地应用双重归谬法了。

依据这两个初步定理，我们就可以来看一看这位几何大师是如何证明《圆的测定》一书中的第一个命题的。

命题 1　任何圆的面积都等于这样一个直角三角形的面积，该直角三角形的一条直角边等于圆的半径，另一条直角边等于圆的周长。

证明　阿基米德首先作两个图形（图 4.3）：圆的圆心为 O，半径为 r，周长为 C；直角三角形的底边等于 C，高等于 r。我们用 A 代表圆的面积，

用 T 代表三角形的面积，而前者就是阿基米德求证的对象。显然，三角形的面积就是$T=\frac{1}{2}rC$。

命题宣称 A=T。为证明这一点，阿基米德采用了双重归谬法证明，他需要考虑并排除其他两种可能性。

例 1　假设 A＞T。

这一假设表明，圆面积以一定量大于三角形面积。换言之，其超出量 A-T 是一个正量。阿基米德知道，通过作圆内接正方形，并反复平分正方形的边，他就可以得到一个圆内接正多边形，其面积与圆面积不等，且小于正量 A-T。即

A-面积（内接正多边形）＜A-T

在不等式的两边各加上“面积（内接正多边形）+T-A”，得

T＜面积（内接正多边形）

但是，这是一个圆内接正多边形（图 4.4）。因此，多边形的周长 Q 小于圆周长 C，其边心距 h 当然也小于圆的半径 r。我们据此得出

面积（内接多边形）$=\frac{1}{2}hQ<\frac{1}{2}rC=T$

至此，阿基米德推导出了预期的矛盾，因为他已得出 T＜面积（内接多边形）和面积（内接多边形）＜T 两种结论。这在逻辑上是不成立的，因此，我们得出结论，例 1 是不可能的；圆面积不能大于三角形面积。

现在，他来考虑第二种可能性。

例 2　假设 A＜T。

这次，阿基米德假设圆的面积小于三角形面积，因而，T-A 代表三角形面积对圆面积的超出量。我们知道，我们可以作一个圆外切正多边形，其面积大于圆面积，但小于 T-A。也就是

面积（外切正多边形）-A＜T-A

如果我们在这一不等式两边各加上 A，则

面积（外切正多边形）＜T

但是，外切正多边形（图 4.5）的边心距 h 等于圆的半径 r，而正多边形的周长 Q 显然大于圆的周长 C。因此

面积（外切正多边形）$=\frac{1}{2}hQ>\frac{1}{2}rC=T$

这样，就再次出现了矛盾，因为外切多边形的面积不可能既小于、又大于三角形的面积。因此，阿基米德推断，例 2 也是不可能的；圆面积不能小于三角形面积。

最后，阿基米德写道：“由于圆的面积既不大于、也不小于（三角形面积），因此，圆面积等于三角形面积。”证讫。

这就是阿基米德的证明，这颗小小的明珠出自一位无可争议的伟大数学家之手。阿基米德用圆面积既不大于、也不小于三角形面积的方法来证明这两个面积一定相等，这种证明方法使一些人感到甚为奇特。一些人感到这种论证方法太绕圈子，对他们我们不妨引述《哈姆雷特》中大臣波洛

涅斯一句话的大意：“这虽则是疯狂，却有深意在内。”人们可能会感到奇怪，这么简短的证明方法，希波克拉底或欧多克索斯或欧几里得怎么会忽略了呢？事后聪明总是不难。这里，我们再次引述普卢塔克关于阿基米德数学的描述：

“在全部几何学中很难找到比这更困难更复杂的问题，以及更简洁更清楚的证明。有些人将此归于他的天赋；而另一些人则认为是他惊人的努力和勤奋产生了这些显然十分容易而又未被他人证明的定理。你费尽力气，仍然一无所获，可一旦看到他的证明，立刻就会认为，自己本来也能够推导出这些结论。他引导你，沿着一条平坦的捷径，得出预定的结果。”

既然阿基米德已证明圆的面积与三角形面积相等，那么，他是否可以解决我们曾在第一章中讨论过的人们长期探索的求圆面积问题呢？答案当然是否定的，因为要成功地解决圆的求积问题，就必须要作出与圆面积相等的直线图形。但是，阿基米德的证明没有，也没有声称能够提供任何有关如何作这种等面积三角形的线索。当然，作出三角形的一条直角边等于圆的半径并不难，难的是作出三角形的另一条直角边，使之等于圆的周长。因为 $C=\pi D$，所以，要作出圆的周长，就必须作出 π。我们已知，这种作图是根本不可能的。阿基米德的证明决不能被解释成他试图据此作出圆的等面积正方形；情况不是这样。

尽管如此，读者从阿基米德的定理中，也许仍然看不出我们所熟悉的求圆面积公式，因为他所证明的毕竟只是圆面积等于一定三角形的面积。但是，我们将看到，这正是典型的阿基米德方法——使一个未知图形的面积与一个更简单的已知图形面积相联系。然而，问题不仅如此。因为我们所说的三角形，其底边等于圆的周长，这里有两层重要含义。其一，阿基米德不像欧几里得那样,不是将一个圆的面积与另一个圆的面积相联系(这基本上是一种“相对性”的方法)，而是将圆的面积与它自己的周长和半径相联系，并反映在它的等面积三角形之中。这样，通过证明 $A=T=\frac{1}{2}rC$。

阿基米德就在一维概念的周长与二维概念的面积之间建立了联系。因为 $C=\pi D=2\pi r$，我们可把阿基米德的定理重新写成

$$A=\frac{1}{2}rC=\frac{1}{2}r（2\pi r）=\pi r^2$$

这就出现了几何学中我们最熟悉，也是最重要的公式之一。

还应指出，阿基米德的命题显然包含了欧几里得关于两个圆面积之比等于其直径的平方比这一比较平淡的命题。也就是，如果我们设一个圆的面积为 A_1，直径为 D_1，设第二个圆的面积为 A_2，直径为 D_2，则阿基米德证明

$$A_1=\pi R_1{}^2=\pi(\frac{D_1}{2})^2=\pi\frac{D_1{}^2}{4} \quad \text{和} \quad A_2=\pi r_2{}^2=\pi(\frac{D_2}{2})^2=\pi\frac{D_2{}^2}{4}$$

因此

$$\frac{A_1}{A_2}=\frac{\pi\frac{D_1^{\ 2}}{4}}{\pi\frac{D_2^{\ 2}}{4}}=\frac{D_1^{\ 2}}{D_2^{\ 2}}$$

这就概括了欧几里得的定理。所以，阿基米德的命题足以表明欧几里得的命题是一个不甚重要的系定理。因而，阿基米德的命题标志着数学上的一个真正进步。

回头再看以前的讨论，现在我们能够确定“欧几里得”表达式 $A=KD^2$ 中常数 k 的数值。因为根据阿基米德的发现，我们知道，

$$\pi r^2=A=kD^2=k(2r)^2=4kr^2$$

因此，$4k=\pi$，$k=\frac{\pi}{4}$。换言之，欧几里得的“二维”面积常数恰好等于“一维”圆周长常数π的四分之一。所以，阿基米德的命题带给我们一个好消息，我们无需计算这两个不同的常数。如果我们能够从圆周长问题中确定π的值，就能够将其应用于圆面积公式。

这后一点也没有难倒阿基米德。实际上，在《圆的测定》一书的第三个命题中，他就推导出了常数π的值。

命题3　任何圆的周长与其直径之比都小于$3\frac{1}{7}$，但大于$3\frac{10}{71}$。

用现代符号表示，即：$3\frac{10}{71}<\pi<3\frac{1}{7}$。将这些分数化为等值小数，阿基米德的命题就成为 3.140845……＜π＜3.142857……：这样，就确定了常数π的值，精确到两位小数，为 3.14。

阿基米德得出π的估计值，又一次显示了他的才能。他准备再次应用他非常有用的圆内接和外切正多边形，不同的是，这次他不再求面积，而将注意力集中在多边形的周长上。他首先作圆内接正六边形（图 4.6）。已知正六边形的边长等于圆的半径，其长度我们称之为 r。因此，

$$\pi=\frac{\text{圆周长}}{\text{圆直径}}>\frac{\text{六边形周长}}{\text{圆直径}}=\frac{6r}{2r}=3$$

当然，这是对π值的非常粗略的估计，但阿基米德刚刚迈出第一步。接下来，他将这一内接多边形的边数加倍，得到一个正 12 边形。他必须计算出这个 12 边形的周长。正是在这个问题上，他使现代数学家惊叹不已，因为要确定十二边形的周长，就要算出 3 的平方根。当然，我们今天使用计算器或计算机，这已不是什么难事，但在阿基米德时代，不仅这些先进设备无法想象，而且，甚至没有帮助进行这种计算的适当数系。阿基米德对$\sqrt{3}$的数值作了如下估计：

$$\frac{265}{153}<\sqrt{3}<\frac{1351}{780}$$

这是非常接近的估计。

随后，阿基米德继续平分内接多边形的边，得到正 24 边形，然后是正 48 边形，最后得到正 96 边形。在这一过程中，每一步他都要估算复杂的

平方根，但他从不动摇。当他得到 96 边形时，他的估算值为

$$\pi=\frac{圆周长}{圆直径}>\frac{正96边形周长}{圆直径}>\frac{6336}{2017\frac{1}{4}}>3\frac{10}{71}$$

阿基米德似乎意犹未尽，又转向外切正 12 边形、24 边形、48 边形和96边形，作类似的估算，并由此得到π值的上限$3\frac{1}{7}$。虽然他面对的是糟糕透顶的数系，而且没有估算平方根的简单方法，但他的估算证实了他令人敬畏的才华。这些计算采用了笨拙的算术方法，犹如一个人戴着沉重的镣铐参加高栏赛跑。然而，阿基米德凭借他的智慧和毅力，成功地计算出了重要常数π的第一个科学近似值。犹如本章后记所述，自此，科学家再不曾停止过寻求高精确度的π近似值。

《圆的测定》一书流传到我们手中，只有三个命题，不过薄薄几页。而且，第二个命题也不恰当，难以令人满意。毫无疑问，这是阿基米德谢世后多少年来低劣的抄写、编辑和翻译造成的。表面看来，这样短的论文似乎不太可能产生这样大的影响。但试想在第一个命题中，阿基米德就证明了关于圆面积的著名公式；在最后一个命题中，他又出色地给出了π的近似值，这篇短论文何以得到历代数学家高度评价，就显而易见了。论文的优劣不在于篇幅长短，而在于其数学质量。根据这一标准，《圆的测定》一书不愧是一部经典之作。

阿基米德名作：《论球和圆柱》

上述三个命题仅仅讨论了阿基米德数学遗产的一部分，除此以外，他还写过有关螺线，劈锥曲面和椭球体的几何论文，并发现了通过求一无穷几何级数之和来确定抛物线弓形面积的方法。这后一个问题（求曲线面积），现在属于微积分领域，由此可见，阿基米德超越他所处的时代有多么远。

然而，相对于所有这些成就，他无可争议的代表作则是一部内容广泛的两卷本著作，题为《论球和圆柱》。在这部著作中，阿基米德以其近乎超人的智慧，确定了球体及有关几何体的体积和表面积，从而像在《圆的测定》中对二维图形的研究一样，解决了三维立体的问题。这是一项伟大的成就，阿基米德自己似乎也认为，这标志着他数学事业的顶峰。

我们首先应回顾一下古希腊人对三维立体表面积和体积的认识。如前一章所述，欧几里得证明了两个球体体积之比等于其直径的立方比；换言之，这里有一个“体积常数”m，因而，

$$体积（球体）=mD^3$$

这是欧几里得对球体体积的认识，但对于球体的表面积，他却始终保持沉默。因而，对这个问题的成功解决，再次有赖于阿基米德《论球和圆柱》的出现。

这一部两卷本著作运用了一种大家熟悉的论述方式，首先是一系列定义和假设，然后从中推导出复杂的定理。总之，还是欧几里得的模式。书中的第一个命题平平淡淡：“已知一个圆外切正多边形，则其周长大于圆的周长。”但是，阿基米德很快就转向了更复杂的问题。通观他的全部论

述，（至少用现代人的眼光来看），一个很大的缺憾是，由于缺乏简明的代数符号，他无法用简单的公式来表示体积和表面积，而只能依靠陈述，例如：

命题 13　任一正圆柱除上下底面以外的表面积等于一圆的面积,该圆的半径是圆柱的高与底面直径的比例中项。

乍一看，这一命题似乎非常深奥而陌生。但实际上，我们所感到陌生的，只是其语言，而不是其内容。由于没有代数，阿基米德只好用这种方式来表示他所求证的面积（本例为正圆柱体的侧面积）等于一个已知图形的面积（本例为一个圆）（图 4.7）。但是，是一个什么样的圆呢？显然，阿基米德必须要指定他的等面积圆，这就是命题中所说的以比例中项为半径的圆。

用现代术语表示，阿基米德的命题就是

侧面积（半径为 r，高为 h 的圆柱）

=面积（半径为 x 的圆）

这里，$\frac{h}{x}=\frac{x}{2r}$。我们可以很快从中推导出$x^2=2rh$，这样，我们就得到了著名公式：

侧面积（圆柱）=面积（圆）=$\pi x^2=2\pi rh$

阿基米德通过一系列测探性的命题，接近了他第一个主要目标——球体的表面积。由于篇幅所限，我们不能详细介绍他对这个问题的推理过程，但我们承认他推理的巧妙。前面我们已介绍过阿基米德数学的特点，读者对他再次应用穷竭法就不会感到奇怪了。他利用以前曾证明了其表面积的圆锥体和圆锥台，从内外双向逼近，“穷竭”了球体。待尘埃落定后，他已证明下面这一非凡的

命题 33　任一球体的表面积等于其最大圆之面积的四倍。

阿基米德运用他最喜欢的逻辑方法——双重归谬法完成了对这个命题的证明，即，他先证明球体表面积不可能大于、然后又证明了也不可能小于其最大圆之面积的四倍。如果我们注意到球体“最大圆”的面积（即通过球体“赤道”的圆之面积）正等于πr^2，那么，我们就可以把阿基米德对本命题的陈述（“球体的表面积等于其最大圆之面积的四倍”）转化成现代公式

表面积（球体）=$4\pi r^2$

这是一个非常复杂的数学命题。阿基米德对其概念的熟练驾驭和他所表现出的深刻洞察力，似乎成为现代积分学思想的先声。显然这就是阿基米德被公认为古代最伟大数学家的原因所在。

关于这一命题，还有另外一个事实值得一提，那就是它的奇特性。因为没有任何直觉能让我们感到球体的表面积恰好等于其最大横截面面积的 4 倍。为什么就不能等于 4.01 倍呢？究竟是什么使这不可思议的数字“4’能够保证球体的曲面表面积恰好等于穿过球心的大圆面积的四倍呢？

阿基米德在《论球和圆柱》一书的导言中讲到了球体这一奇异而固有的特性。导言是写给叫做“多西修斯”的某人，此人可能是亚历山大的一

位数学家，而阿基米德曾将论文寄给他。他写道：“……我想到了某些迄今为止尚未被证实的定理，我已经作出了这些定理的证明。”其中，他首先提到“……任一球体的表面积等于其最大圆之面积的四倍”，然后，他又继续写道，这一性质

“……始终是所述图形固有的，那些在我之前从事几何研究的人们只是尚未得知而已。但是，现在我发现了这些性质为真……，我毫不犹豫地将它们与我以前的研究和欧多克索斯关于立体的定理并列在一起，后者因此得到了无可辩驳的证明……”

这一段话提供了一个有趣的史料，我们得以看到，阿基米德是如何评价自己的工作及其在数学发展中的地位的。他毫不犹豫地将自己与伟大的欧多克索斯并列，因为他完全懂得他的发现的非凡性质和分量。但是，他还特别强调，他没有发明或创造 $S=4\pi r^2$。他只是幸运地发现了球体固有的性质，这一性质始终存在，只是以前没有被几何学家所发现。阿基米德认为，数学关系的客观存在与人类能否解释它们无关。他自己只是能够瞥见这些永恒真理的幸运者。

即使《论球和圆柱》除了前面的定理外，再没有其它内容，它也将永远是一部数学经典。然而，阿基米德随即将目光转向了球的体积问题。他再次应用双重归谬法，成功地证明了

命题 34　任一球体的体积等于底面积为球体最大圆面积,高为球体半径的圆锥体体积的四倍。

请注意，阿基米德依然没有将球体体积表述为简单的代数公式，而是借助了一个更简单的立体（本例为圆锥体）体积（图 4.8）。我们只要略加努力，就可以把他的文字陈述转变为现代等价公式。

设 r 为球体半径。那么，“底面积为球体最大圆面积，高为球体半径的圆锥体”就等于

$$\text{体积（圆锥体）} = \frac{1}{3}\pi r^2 h = \frac{1}{3}\pi r^2 r = \frac{1}{3}\pi r^3$$

但是，阿基米德的命题 34 证明，球体的体积等于这种圆锥体体积的四倍，由此得出著名公式

$$\text{体积（球体）} = 4\text{体积（圆锥体）} = \frac{4}{3}\pi r^3$$

这个公式的优点之一是阐明了π与欧几里得命题Ⅻ.18 提出的“体积常数”m 之间的联系。参照我们以上的讨论，我们可以直接得出

$$\frac{4}{3}\pi r^3 = \text{体积（球体）} = mD^3 = m(2r)^3 = 8mr^3$$

和一个代数式$m = \frac{\pi}{6}$。这样，阿基米德前关于圆周长、圆面积和球体积的谜就都得到了解决。不再需要三个不同的常数来强调这三个不同的问题；所有这三个常数都建立在π的基础之上。阿基米德已展示了它们之间惊人的统一。

阿基米德在完成了对命题 33 与 34 的证明之后，立即以一种引人注目的方式重述了这两个命题。他假设一个圆柱体外切一个球体，如图 4.9 所示。然后，他宣称，圆柱体的表面积和体积都等于球体的一倍半！在某种

意义上，这是他全部工作的高潮。他用一种简单的方式描述了两个伟大定理，用相应比较简单的圆柱体的表面积和体积来表示复杂的球体表面积和体积。本节将以对阿基米德这一惊人判断的证明作为结束。

首先，我们看一个圆柱体外切半径为 r 的球体，其自身半径也等于 r，并且，高 h=2r。圆柱体的全部表面积等于侧面积（见命题 13）与顶面积及底面积之和。因此，

圆柱体全部表面积$=2\pi rh+\pi r^2+\pi r^2$

$$=2\pi r(2r)+2\pi r^2=6\pi r^2$$

$$=\frac{3}{2}(4\pi r^2)$$

$$=\frac{3}{2}$$ （球体表面积）

这一公式准确地表达了阿基米德所说圆柱体的表面积等于球体表面积“一倍半”的意思。

那么，其相应体积如何呢？我们知道，一般圆柱体的体积公式是 $V=\pi r^2h$，在本例，则 $V=\pi r^2(2r)=2\pi r^3$。因此，

圆柱体体积$=2\pi r^3$

$$=\frac{3}{2}(\frac{4}{3}\pi r^3)$$

$$=\frac{3}{2}$$ （球体体积）

所以，圆柱体的体积等于球体体积的一倍半。

这样，阿基米德以一种简明而非凡的陈述表明了球体与圆柱体之间的联系。正是这种联系使他将这部论著题为《论球和圆柱》。阿基米德对他的这一发现深感自豪，这一点可以从普卢塔克所提到的阿基米德的墓志铭中看出：

“他的发现数量众多，令人钦佩；但据说他在弥留之际，曾请求他的朋友和亲属在他的墓上置放一个内盛球体的圆柱体，并且要使球体按照二者之间的比例（即，3：2）内接于圆柱体。”

有趣的是，西塞罗到叙拉古城时确曾拜谒过阿基米德的墓，并在其《图斯库卢姆谈话录》一书中作了记载。阿基米德的墓地上长满了“杂乱的荆棘与灌木”，遮盖了一切。但西塞罗知道他要找的是什么，所以，不难理解，当他发现“灌木丛中显露出来一个小圆柱顶上安放着内盛球体的圆柱体”时，心情该有多么激动。西塞罗发现了阿基米德墓地遗址后，曾想按照原貌尽力修复。如果这个故事真实，那么，西塞罗就发现了这位古希腊最伟大数学家的最后长眠地。他想修复这一被人忘却了的墓地遗址，不仅是为了向阿基米德罗表示敬意，也许还为了补赎他的残暴的马祖先的罪愆。

人们常常听说有人走在时代的前面。这一般是说，一个人超于世上其他人十年，或者，整整一代。但是，阿基米德对数学的贡献，其辉煌却是千百年无人能出其右！直到 17 世纪发展了微积分，人们对立体体积和表面积的理解才超出了阿基米德的基本原则。可以肯定地说，无论数学学科将来会有怎样的荣耀都永远不会再有人先于时代两千年了。

最后，我们不如引用伏尔泰对这位伟大数学家的成就所作的恰如其分

的高度评价：“阿基米德比荷马更富有想象力。”

后记

阿基米德《圆的测定》一书所留下的遗产之一是求我们称之为π的重要常数的更精确近似值。这一比率的重要性早在阿基米德之前很久便已为人所知，虽然阿基米德是科学地研究这一常数的第一人。在阿基米德之前，人们对π值的估算可以从《圣经》关于圆“海”（即一个盛水的大容器）的一段有趣的引文中推断出来：“……他又铸一个铜海……径十肘、围三十肘”（《列王记》，上，7：23）。

从这里，我们可以推导出π的近似值，即$\pi=\frac{C}{D}=\frac{30}{10}=3.00$，由于当时尚属远古时代，所以，这一近似值还是非常合理的。（当然，那些认为《圣经》在一切方面都十分精确的人，会在此吹毛求疵，因为3.00大大低估了π值。）

古代对π值的更精确计算是古埃及人作出的。据赖因德古本记载，他们用$(\frac{4}{3})^4=\frac{256}{81}=3.1604938$……作为C与D的比值。这些及其他“前科学”近似值代表了对π值估算的第一阶段。如我们所知。阿基米德开创了第二阶段。他所应用的圆内接或外切正多边形周长的几何方法一直为17世纪中叶前的数学家所采用（这是阿基米德走在时代前面的又一个证明）。

约公元150年，亚历山大著名的天文学家和数学家克劳迪厄斯·托勒密在其《天文学大成》这部巨著中提出了π值的一个近似值。这部巨著集天文学信息之大全，从太阳、月球及行星的运动，到恒星的性质，无所不包。显然，对天体的精确观测需要复杂的数学基础，为此，早在《天文学大成》中，托勒密就作出了弦值表。

他首先作一个圆，将其直径分为120等份。如果每一等份的长度为p，则我们可以确定其直径为120p，如图4.10所示。对于任何圆心角a来说，托勒密希望能发现与这一角所对的弦AB的长度。例如，一个60°角所对的弦恰好等于半径的长度，即60p。

这是一个很简单的例子，但是，要发现$42\frac{1}{2}$°角所对弦的长度却远非易事。然而，托勒密运用巧妙的推理和阿基米德的一个计算技巧，精确地计算出了范围从$\frac{1}{2}$°到180°，以半度递增的所有角度的弦值表。

但是，与我们的讨论有关的问题是，他发现了1°弦的值（用现代十进制记数法）为1.0472p。因此，内接于这个圆的正360边形的周长就等于1°弦长的360倍，即376.992p。虽然利用正多边形的思想显然是阿基米德的思想，但托勒密的360边形却比其前辈的96边形推算出的π近似值精确得多。即，

$$\pi=\frac{C}{D}\approx\frac{360\text{边形的周长}}{\text{圆的直径}}=\frac{376.992p}{120p}=3.1416$$

几百年后，π值计算的发展集中于非西方文化的中国和印度，在这两个国家的文化中，都有其自己光辉的数学史。中国的科学家祖冲之（430

—

501年）于公元约480年计算出π的近似值$\frac{355}{113}=3.14159292\cdots\cdots$，而印度数学家婆什迦罗第二（1114—约 1185 年）则于约公元 1150 年计算出$\frac{3927}{1250}=3.1416$。

当欧洲人终于从中古时代的数学停滞中再度崛起的时候，发现精确π值的速度大大加快了。16 世纪末，随着西蒙 · 斯蒂文（1548—1620 年）等数学家的艰苦探索，现代十进制问世了，人们可以更方便、更准确地计算平方根。因而，当法国天才数学家弗朗索瓦 · 韦达（1540—1603 年）试图利用阿基米德的方法计算π近似值时，他可以用正 393，216 边形推算出精确到 9 位小数的π值。他先按照阿基米德的方法作出正 96 边形，然后将正多边形的边数翻倍十几次，得到正 393，216 边形。即使阿基米德，在韦达的数系面前，也会惊讶，而十进制记数法则为韦达提供了用武之地。他所采用的基本方法仍然是阿基米德的方法，但韦达却拥有比阿基米德更先进的工具。

17 世纪初叶，一位德国数学家超越了所有前人，发现了精确到 35 位小数的π值。他的名字叫卢道尔夫 · 冯瑟伦，他用了几年时间钻研这个问题。像韦达一样，卢道尔夫也将新的十进制与旧的阿基米德方法结合起来，但他不是从正六边形开始并将其边数翻番的，卢道尔夫是从正方形开始的。到他完成的时候，他已推导出了有 2^{62} 条边的正多边形——或约 4，610，000，000，000，000，000 边形！不用说，这个多边形的周长与其外接圆的周长相差无几。

计算π近似值的古典方法已引导数学家们走得很远。然而，17 世纪末叶发生了一次计算这一不寻常比值的数学大爆炸，其中一个进步就是最终取代了阿基米德的方法，并将对π值的探索推向了第三阶段。17 世纪 60 年代末，年轻的艾萨克 · 牛顿应用其广义二项式定理和新发明的流数法（即微积分），比较轻松地计算出了非常精确的π近似值；这就是我们将在第七章中介绍的伟大定理。1674 年，牛顿的对手戈特弗里德 · 威廉 · 莱布尼兹发现级数

$$1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\frac{1}{11}+\frac{1}{13}-\frac{1}{15}+\cdots\cdots$$

随着我们计算的项数不断增加，接近于数字π/4。至少从理论上说，我们可以无限扩展这一级数的项，以得到更加精确的π/4，并且由此得到π本身的近似值。重要的问题是，我们必须求和的这个级数，按其方式，是完全可以预知的；也就是说，无论我们将这一级数扩展多长，都不难确定下一项。这样，求π近似值的问题就从以阿基米德正多边形为代表的几何问题突然变成了加减数字的简单算术问题。这是一个视角的重大的变化。

实际上，在这一点上，情节渐趋复杂，因为莱布尼兹的级数虽然确能接近π/4，但计算起来也实在太慢了。例如，即使我们用这一级数的前 150 项计算，我们所得到的π近似值也仅为 3.1349……，想到工作量之大，则这一计算的准确性实在令人失望。据估计，如果我们要利用这一级数得到精确到 100 位小数的π近似值，我们就需要计算

100，000，000，000，000，000，000，000，000，000，000，000，000，000，000，000，000

多项！因此，虽然莱布尼兹的级数预示了一种计算π近似值的新的算术方法，但它显然没有实用价值。

但是，无穷级数的计算很快有了前途。数学家亚伯拉罕·夏普（1651—1742 年）和约翰·梅钦（1680—1751 年）等对无穷级数做了巧妙的修改，并产生了计算速度快得多的收敛级数。利用这些修改，夏普于 1699 年发现了精确到 71 位小数的π近似值，而 7 年后，梅钦则计算出 100 位小数的π值。并且，他们的方法远比可怜的卢道尔夫用大半生时间才抠出 35 位小数的方法简便得多。显然，级数方法宣告了古典方法的过时。

同时，数学家还试图理解这一特殊常数的其他特点，在这一方面也有了进展。其中主要是约翰·海因里希·朗伯（1728—1777 年）于 1767 年证明π是无理数。我们回想一下，所谓无理数，就是那些不能写成两个整数商的实数——即，无理数是不能写成分数的数。我们可以很容易地指出，像$\sqrt{2}$或$\sqrt{3}$这些常数都是无理数，但是，直到18世纪，朗伯才证明出π是无理数。我们只要回想一下有理数的小数展开式不是终止，就是循环，便能理解朗伯的发现特别重要。例如，有理数$\frac{1}{8}$小数是0.125，而有理数$\frac{1}{7}$的小数却是无限的，但至少它是以六位小数循环：

$$\frac{1}{7}=0.142857142857142857142857\cdots\cdots$$

如果π是有理数，那么，它也应该能够表现出这两种形式中的一种——或者终止，或者循环，因而，确定其小数展开式的努力，经过一定时间后，势将最终完成。朗伯证明π是无理数，就决定了其小数计算是无穷无尽的。

似乎这种无理性还不够糟，1882 年，费迪南德·林德曼又证明出π实际上是超越数，如我们在第一章中所述。对超越数的发现不仅解决了化圆为方的问题，而且，也说明了π不能形成涉及有理数平方根、立方根等等的任何一种初等表达式。朗伯和林德曼的研究表明，π不是那种易于进行数学分析的“佳”数。然而，公元前 225 年的阿基米德明白无误地指出π是所有数中最重要的数之一。

π的历史引入了本世纪一位杰出的数学家斯里尼瓦萨 拉马努扬(1887—1920 年）。拉马努扬出生在印度，家境贫寒，未受过任何正规数学教育。他主要是根据几本教科书自学成才的。拉马努扬对数学的热中，严重影响了他的其他学业，由于他未能通过其他课程的考试，不得不中止了他的正规教育。1912 年，他在马德拉斯做文书工作，靠 30 英镑年薪勉强维持他夫妻二人的生活。作为失败者，他是很容易被社会所淘汰的。

尽管存在着这种种障碍，但这位孤立的天才却在从事着极富独创力，极有深度的数学研究。在别人的怂恿下，他把自己的发现

写成了一封长信，分别寄给了英国三个最著名的数学家。其中两人把信退回了拉马努扬。显然，他们认为有比给一个不知名的印度小职员回信更紧迫的事情要做。

第三位数学家是剑桥大学的 G.H.哈代，在他于 1913 年 1 月 16 日打开

早晨的邮件时，本来或许也会这样做。拉马努扬用蹩脚英文写的信中有 100 多个奇怪的公式，但没有任何证明，仿佛一个妄想狂随心所欲地在漫游世界。哈代随手把信丢在一旁。

但是，据说，这些数学公式一整天萦绕在他的脑海中。这些公式，有许多都是哈代这位世界上最优秀的数学家从来没见过的。逐渐地，他醒悟了，认识到这些公式“……一定是真实的，因为如果它们不真实，就不会有人能有这种想象力，发明出它们来。”哈代回到自己的房间，重新审看早上的信件，他意识到，这是一个大数学天才的杰作。

这样，就开始办理拉马努扬到英国来的手续。对于一个从小受到严格宗教熏陶的人来说，这是一件很复杂的事情，因为他在旅行、饮食等方面都有许多限制。但这些困难最后都被克服了，1914 年，拉马努扬终于来到了剑桥大学。

拉马努扬与哈代从此开始了长达五年之久的非凡合作——后者是受过世界上最好数学教育的温文尔雅的英国人；而前者却是一位“未经雕琢的天才”，虽具有令人难以置信的能力，但数学知识却有很大局限性。有时，哈代只好像对待一个普通大学生一样指导这位年轻伙伴。而拉马努扬也常常提出一些从未见过的数学定理令他惊奇。

在拉马努扬的公式中，有许多都能够迅速而精确地计算出π的近似值。这些公式，有的编入了 1914 年他的一篇重要论文；也有一些则潦草地涂写在他的私人笔记本里（急切的数学界直到现在才有幸目睹这些文献）。即使其中最简单的公式也使我们受益匪浅。其实，只要说一句话就够了，他的见解开通了更有效地计算π近似值的思路。

然而，不幸的是，拉马努扬的事业，开始得如此奇特，结束得又如此仓促。第一次世界大战期间，拉马努扬在远离家乡的剑桥大学累垮了身体。有些人认为原因在于疾病，而另一些人认为，原因在于严格的饮食限制造成了严重的维生素缺乏症。为了恢复健康，他于 1919 年返回了印度，然而，他家乡的温馨却无法阻止他病情的恶化。1920 年 4 月 26 日，拉马努扬与世长辞了，年仅 32 岁。从此，世界失去了一位数学奇才。

现在，让我们来引述英国人威廉·谢克斯（1812—1882 年）的惊人计算，至此，我们的故事迅速进入现代结局。谢克斯于 1873 年计算出 707 位小数的π值。他利用梅钦的一系列方法，达到了如此惊人的精确度，成为此后 74 年的标准。但是，1946 年，他的同胞 D.F.弗格森却令人吃惊地发现谢克斯在其非凡计算的第 527 位小数之后出现了错误。弗格森善意纠正了这些错误，并得到了 710 位小数的π值。对于那些少有计算兴趣的人来说，简直难以想象能够对一个带有 707 位小数的数字进行验算，而且，更令人难以置信的是，在验证了 100 位、200 位，甚至 500 位小数都没有发现错误的时候，竟然还能坚持验算下去！弗格森的惊人毅力终于有了回报。

1947 年初，美国 J.W.伦奇将自己的成就加入到这一历史之中，公布了 808 位小数的π值。这似乎是一个辉煌的新胜利——但后来，不屈不挠的弗格森又开始检验这一数值。并且，他果真发现伦奇计算的第 723 位小数有错误。然后，弗格森与伦奇两人通力合作，终于于一年后公布了精确到 808 位小数的π值。

自此，故事进入了第四阶段，也是最后一个阶段。我们已看到，人们

最初是如何“数着指头”估算π值的；后来，阿基米德引入了圆内接和外切多边形的方法，这种方法一直盛行到微积分的出现，并被无穷级数的算术技术所取代。最后，1949 年，计算机的出现引起了计算方面的根本革命。同一年，美国陆军的电子数字积分计算机计算出精确到 2037 位小数的π值。应当指出，用现代眼光来看，这台计算机是一部非常原始的机器，密密麻麻的电线和电子管占满了好几间房屋，其运算速度之慢，简直让人难以忍耐。但即使是这样一台古怪的老式计算机，也超越了前人的所有计算，其中一个重大的飞跃是将人类两千两百年计算的π近似值的小数位扩大了一倍半，而且，即便是 D.F.弗格森也找不出一个错误。随着计算机技术的发展，π值的小数位数以令人难以置信的速度飞快增加。1959 年，计算出 16，000 多位小数；1966 年，就发展到 25 万位小数，而到 80 年代末，巨型电子计算机已将π值的小数位数猛增到 5 亿位上下。

然而，我们人类脆弱的自我无须感到大受伤害。虽然计算机的计算速度超出任何人的想象，但毕竟还需要由数学家去编制程序，指导计算机正确运算。π的故事讲述了人类的胜利，而不是机器的胜利。即使在 20 世纪末叶，我们也绝不能忘记，这一数学历程最初是由最优秀的数学家、叙拉古城的阿基米德的一篇题为《圆的测定》的短文引发的。

第五章
赫伦的三角形面积公式
（约公元 75 年）

阿基米德之后的古典数学

阿基米德在数学景观上投下了长长的影子。其后的古代数学家虽然都有自己的建树，但却没有一个人能够比得上叙拉古城这位伟大的数学家，随着希腊文明的衰落和罗马的同时兴起，事情益发明显。阿基米德死于罗马人之手，预示了以后所发生的事情，这种看法也许有点儿简单化，但并非没有道理。希腊人专注于自己的理念世界，在罗马强大的军事力量面前，确实不堪一击，而罗马人则忙于建立政治秩序和征服世界，完全无视希腊人热中的抽象思维。如同对阿基米德一样，罗马新秩序同样也不能容许希腊传统的存在。

一些资料也许有助于我们的认识。我们已看到，叙拉古城于公元前 212 年陷落于罗马的马塞卢斯之手。三次残酷的布匿战争最终以公元前 146 年罗马消灭迦太基而告终，罗马人从此确立了对中地中海两岸的控制。同一年，希腊的最后一座重要城邦科林斯向罗马军投降。一百年后，尤利乌斯 凯撒征服了高卢；公元前 30 年，在安东尼与克娄巴特拉的统治失败后，埃及落入屋大维之手。甚至野蛮的不列颠也于公元 30 年臣服于罗马。自此，罗马正式成为帝国，对西方世界行使着史无前例的统治。

随着罗马的征服，他们复杂的工程项目也随之发展起来：桥梁、道路和沟渠遍布欧洲大陆。然而，曾强烈吸引希波克拉底、欧几里得和阿基米德的纯粹数学却未能像以前那样兴盛。

但是，依然保持辉煌的是亚历山大图书馆。这座环境优美的图书馆吸引了地中海地区最优秀的学者，是一个最令人兴奋的地方。阿基米德的一位同时代人，著名数学家厄拉多塞（公元前约 284—192 年）就曾大半生在这里担任馆长。厄拉多塞身居学术要职，是一位阅读广泛、著作等身的学者，许多关于纯数学、哲学、地理学，特别是天文学的著作都出自他的手，这最后一项，不仅包括许多学术论文，而且还包括一部题为《赫耳墨斯》的长诗，将天文学的基本知识写成了诗歌！像众多的古代著作家一样，厄拉多塞的著作大部分散失了，我们只能依靠后来注释者的描述来了解他。但他身为当时的学界名流，似乎是没有疑问的。阿基米德至少有一篇著作是题献给厄拉多塞的，并视其为一个伟大的天才。

厄拉多塞的一大贡献是他著名的“筛法”，这是一种寻找素数的简便方法。为了用厄拉多塞筛法选出素数，我们首先写下从 2 开始的连续正整数。请注意，2 是第一个素数，然后我们依次划掉后面所有 2 的倍数，即 4、6、8、10 等。越过 2，下一个没有划掉的整数是 3，这一定是第二个素数。我们现在再划掉所有 3 的倍数——6

（虽然它已经被划掉了）、9、12、15 等。下面我们来看，4 已经被划掉了，于是，下一个素数是 5；我们再划掉表中所有 5 的倍数——10、15、20、25 等。如此循序渐进。显然，我们划掉的数字都是较小整数的倍数，它们都不是素数，因而，都被筛掉了。而另一方面，素数却永远不会被筛

掉，它们将成为我们表中唯一剩下的数字：

2, **3**, ~~4~~, **5**, ~~6~~, **7**, ~~8~~, ~~9~~, ~~10~~, **11**, ~~12~~, **13**, ~~14~~, ~~15~~, ~~16~~, **17**, ~~18~~,

19, ~~20~~, ~~21~~, ~~22~~, **23**, ~~24~~, ~~25~~, ~~26~~, ~~27~~, ~~28~~, **29**, ~~30~~, **31**, ~~32~~, ~~33~~, ~~34~~, ~~35~~,

~~36~~, **37**, ~~38~~, ~~39~~, ~~40~~, **41**, ~~42~~, **43**, ~~44~~, ~~45~~, ~~46~~, **47**, ~~48~~, ~~49~~, ~~50~~, ~~51~~,

~~52~~, **53**, ~~54~~, ~~55~~, ~~56~~, ~~57~~, ~~58~~, **59**, ~~60~~, **61**, ~~62~~, ~~63~~, ~~64~~, ~~65~~, ~~66~~, **67**,

~~68~~, ~~69~~, ~~70~~, **71**, ~~72~~, **73**, ~~74~~, ~~75~~, ~~76~~, ~~77~~, ~~78~~, **79**, ~~80~~, ~~81~~, ~~82~~, **83**,

~~84~~, ~~85~~, ~~86~~, ~~87~~, ~~88~~, **89**, ~~90~~, ~~91~~, ~~92~~, ~~93~~, ~~94~~, ~~95~~, ~~96~~, **97**, ~~98~~, ~~99~~

用厄拉多塞筛法，可以自然而然地产生 100 以下的所有素数。但要找出，比如说，100 万亿以下的所有素数，用这种方法显然就非常困难了，但现代计算机运用这一古老方法，却有极大收获。

厄拉多塞最著名的科学成就也许是他对地球周长的测定。虽然有许多文字描述了他的这一计算，但由于找不到他的原始论文《论地球的测量》，我们还不能肯定厄拉多塞究竟用的是什么方法。但是，据说，他运用了一些地理数据和一个非常简单的几何图形，具体如下：

在埃及亚历山大以南，今天的阿斯旺附近，有一座城市叫赛伊尼。在夏季第一天的中午，赛伊尼处的太阳直射地面。如果此刻有人往井里看，就会感到水面反射的太阳非常刺眼，从而证实了太阳的直射。但在同一天的同一时刻，亚历山大处的竹竿却投下一个短影。厄拉多塞注意到，从竹竿顶端到阴影端点之间形成的夹角a恰好等于整个圆周角度的$\frac{1}{50}$（见图 5.1）。假设亚历山大位于赛伊尼的正北（这大体正确），而且，因太阳距离地球十分遥远，假设阳光射到地球是平行的（又一个合理的假设），厄拉多塞根据《原本》的命题Ⅰ.29，判定内错角∠AOS 等于∠a，而 0 则代表球状地球的球心，如图 5.1 所示。最后的已知条件是测得这两座城市之间的距离为 5000 斯达地。因此，我们得到比例

$$\frac{\text{从赛伊尼到亚历山大的距离}}{\text{地球周长}}=\frac{\text{角}a}{\text{全圆角度}}$$

也就是，$\frac{5000\text{斯达地}}{\text{地球周长}}=\frac{1}{50}$，所以，地球的周长等于$50\times5000=250{,}000$斯达地。至此，读者肯定会问：“一斯达地是多长？”厄拉多塞所用单位究竟多长已无从考订，只能冒险地引用估计值，即一斯达地约等于 516.73 英尺。利用这一数字，可以得出厄拉多塞计算的地球周长为 129，182，500 英尺，或约 24，466 英里。目前公认的地球周长为 24，860 英里，所以，厄拉多塞的计算结果非常接近此值。实际上，由于这个数字太精确了，以致一些学者怀疑其真实性，或者至少同意托马斯·希思爵士的观点：厄拉多塞给了我们“一个令人惊奇的近似值，不论其在多大程度上归因于计算中的偶然。”

暂且抛开这些怀疑不谈，厄拉多塞的推理方法还是值得我们注意的，这不仅是因为其巧妙，而且还因为，无论如何，他坚信我们的地球是一个球体。而 1500 年后的欧洲水手却还惧怕从扁平大地的边沿掉下去。我们有时忘记了古希腊人已完全意识到地球的球体形状，如果后来的水手还睁大

双眼，搜寻地平线的边缘，这与其说是学问太少，不如说是记性不佳。

另外，还有两位阿基米德之后的数学家值得介绍。其中一位是阿波罗尼奥斯（公元前约 262—190 年），他也是阿基米德同时代人，也曾到过亚历山大，在那充满学术气氛的环境里学习、工作。他在那里完成了他的代表作《圆锥曲线》，这是一部广泛讨论所谓圆锥曲线的巨著，涉及椭圆、抛物线和双曲线（图 5.2）。虽然古希腊数学家曾深入研究过这些曲线，但阿波罗尼奥斯重新整理了前人的工作，使之系统化、条理化。这种情形，很像欧几里得编著《原本》的情况。《圆锥曲线》共有八篇，前四篇系统叙述了圆锥曲线的基本原理，后三篇讨论了更专业化的问题，第八篇现已失传。

即使在古代，阿波罗尼奥斯的著作也被公认为是圆锥曲线问题的权威论述，而且，在文艺复兴期间被重新发现后，亦得到了很高的评价。当约翰·开普勒（1571—1630 年）提出他关于行星以椭圆形轨道围绕太阳运动的独创性理论时，圆锥曲线的重要性得到了证实。椭圆绝不仅是古希腊数学家手中把玩的珍品，它已成为地球，乃至地球上我们全体人类运行的轨道。大约一百年后，因发现彗星而声名大噪的英国科学家埃德蒙·哈雷用了几年时间来编定《圆锥曲线》的最后定本，并对这一古典数学著作推崇备至。今天，这部巨著与欧几里得的《原本》和阿基米德的著作并列，成为古希腊数学的里程碑。

最后一位古代数学家与本章的伟大定理有关，他就是亚历山大（还能是哪里呢？）的赫伦。一些现代书把他的名字写成“希罗”（Hero，即英雄之意——译注），这主要是因翻译造成的，而不是他自命不凡。遗憾的是，我们对他的生平知之甚少，甚至对他的生卒年月也颇多争议。不过可以肯定地说，赫伦是在阿波罗尼奥斯之后，但更确切的日期就有待于一位天才像侦探小说里经常描写的那样去推断了。我们采用了霍华德·伊夫斯的观点，确定赫伦的活动时期为公元 75 年前后。

尽管我们对赫伦的生平知之甚少，甚至不知是否相差了 150 年，但学者们却拥有大量关于赫伦数学的资料。赫伦的兴趣主要是在实践方面，而不是理论，他的许多著作都涉及了非常有用的实用科学，如机械学、工程学和测量学。他的这种侧重反映了希腊人与罗马人兴趣的截然不同。例如，赫伦在其《经纬仪》一书中介绍了挖掘穿山隧道及计算泉水流量的方法。在另一部著作中，他回答了一些日常生活问题，例如“为什么用膝盖在一根木棍的中间用力顶，木棍会折断？”或者“为什么人们用钳子而不用手拔牙？”

然而，有趣的是他关于三角形面积的命题。这一命题像赫伦的其他许多课题一样，明显地带有实用性，但他对这一命题的证明却是一篇出色的几何抽象推理。这条命题是赫伦《量度》一书中的命题Ⅰ.8，这一重要著作的发现还有一段有趣的历史。数学家们早就知道有这样一篇论文存在，因为评注家欧托休斯早在公元 6 世纪时就曾提到过这部著作，但人们却找不到它的下落。它就像恐龙一般神秘地消失了。1894 年，数学史家保罗·坦纳利在一个 13 世纪巴黎人的手稿中偶然发现了这篇论文的片段。更幸运的是，两年后，R.舍内在君士坦丁堡发现了这部著作的全部手稿。这样，现代人才有幸目睹《量度》一书的全貌。

伟大的定理：赫伦的三角形面积公式

我们已知，赫伦的公式涉及三角形面积。这个公式似乎完全不必要，因为众所周知，三角形面积的标准公式十分简单——面积 $=\frac{1}{2}$（底）×（高），而且已得到了广泛的应用。但是，如果用这个公式去求图 5.3 中三角形的面积就没有什么意义了，因为我们还不知三角形的高。

首先，应当指出，已知一个三角形的三条边，则其面积一定是确定的。这可以直接从“边边边”全等定理（欧几里得，命题 1.8）中推导出来，因为我们知道，任何边长等于（比如）17、25 和 26 的其他三角形，一定与图 5.3 中的三角形全等，因此，其面积也完全相等。所以，如果我们知道三角形的三条边，我们也就知道一定有一个，并且只有一个面积值。

但是，如何确定这一面积值呢？今天仍像两千年前一样，最简便的方法是应用赫伦的公式，其公式用现代符号表示，就是：

如果 K 是边长等于 a、b、c 的三角形的面积，那么，

$$K=\sqrt{s(s-a)(s-b)(s-c)}$$

这里$s=\frac{1}{2}(a+b+c)$是三角形的所谓“半周长”。

在图5.3中，$s=\frac{1}{2}$（$17+25+26$）$=34$，因此，

$$K=\sqrt{34(34-17)(34-25)(34-26)}=\sqrt{41616}=204$$

请注意，在应用赫伦的公式时，我们只要知道三角形的三条边就足够了；我们无须求出三角形的高。

这是一个非常特殊的公式，乍看之下，人们会感到是不是印刷有误。公式中出现的平方根和半周长似乎非常奇怪，这个公式完全没有直观感染力。然而，作为一项伟大的定理，引起我们注意的不仅有它的奇特，还有赫伦为此所作的证明。他的证明既非常曲折，令人惊叹，又非常巧妙。在某种意义上说，他的证明是很初等的，因为他只用了一些简单的平面几何概念——也就是说，只用了一些“元素”。但是，赫伦展示了他精湛的几何技巧，将这些元素组合成一个非常丰富而漂亮的证明，堪称数学中一个令人叹为观止的结论。赫伦的证明就像阿加莎·克里斯蒂的侦探小说一样，读者一直读到最后几行可能还弄不清问题如何解决。但我们不必着急，他最后的几步推理，将这一证明推向了高潮。

在介绍这一证明之前，我们有必要先介绍一些赫伦论证所依据的初步命题。前两个初步命题出自欧几里得的《原本》。

命题 1　三角形的角平分线交于一点，这个交点是三角形外接圆的圆心。

这一命题出自欧几里得《原本》的命题Ⅳ.4。三角形三条角平分线的交点（即三角形外接圆的圆心）恰当地叫做内心。

命题 2　一个直角三角形，如果从直角作斜边的垂线，则垂线两边的

三角形分别与整个三角形相似，并互相相似。

读者将会发现，这一命题出自《原本》的命题IV.8，我们在第三章中曾讨论过这一命题。

下面的定理虽然也非常著名，但没有编入欧几里得的《原本》。为了保持完整，我们同时附加了定理的简单证明。

命题 3　在直角三角形中，斜边的中点与三个角的顶点距离相等。

证明　首先设直角三角形 BAC（图 5.4），平分 AB 边于 D，作 DM 垂直于AB。连接MA。我们说，△MAD与△MBD全等，因为$\overline{AD}=\overline{BD}$，$\angle ADM=\angle BDM$，当然，$\overline{DM}=\overline{DM}$。所以，“边角边”全等定理保证了$\overline{MA}=\overline{MB}$和$\angle MAD=\angle MBD$。由于我们最初作的是直角三角形，`因此，

∠ACM=1 个直角-∠MBD=1 个直角-∠MAD=∠MAC 所以，△MAC 是等腰三角形，因而，$\overline{MC}=\overline{MA}$。因为线段MA、MB和MC都相等，所以，我们断定，斜边的中点 M 与直角三角形三个角的顶点距离相等。　证讫。

我们最后的两个初步命题涉及到联圆四边形，也就是圆内接四边形。

命题 4　已知 AHBO 是一个四边形，作对角线 AB 和 OH，如果∠HAB 与∠HOB 是直角（如图 5.5 所示），则可以过四个顶点 A、O、B、和 H 作一个圆。

证明　这是从前一个命题中直接推导出来的一个特殊命题。如果我们平分 BH 于 M，我们注意到，M 是直角三角形 BAH 与直角三角形 BOH 的共同斜边上的中点。所以，M 与 A、O、B 和 H 各点的距离相等，因而，以 M 为圆心，以$\overline{MH}$为半径，可以过四边形所有四个顶点作一个圆。证讫。

命题 5　联圆四边形的对角和等于两个直角。

这个命题出自《原本》的命题III.22，其证明见第三章。

这五个命题不妨看作一个特殊的工具箱，带给我们一个关于一般三角形面积的证明。但是，它们连同高度的技巧，只是赫伦在证明现在以他的名字命名的公式时所需要的“元素”。

定理　已知一个三角形，其边分别为 a、b 和 c，面积为 K，我们得知$K=\sqrt{s(s-a)(s-b)(s-c)}$，其中，$s=\frac{1}{2}(a+b+c)$，是三角形的半周长。

证明　设任意三角形 ABC，使 AB 边至少不小于其他两条边。为了使赫伦的论证清晰易懂，我们将他的证明分成三大部分。

第一部分　赫伦的第一步就令人非常震惊，因为他首先作了一个三角形的内接圆。他用三角形内心作为确定其面积的关键因素，大大出人预料，因为圆的性质与三角形这种直线图形的面积没有直观联系。尽管如此，我们还是设O为内接圆的圆心，用r表示半径，我们看到，$\overline{OD}=\overline{OE}=\overline{OF}=$ r，如图 5.6 所示。

现在，我们应用简单的三角形面积公式，得到：

面积（△AOB）$=\frac{1}{2}$（底）×（高）$=\frac{1}{2}$（$\overline{AB}$）×（$\overline{OD}$）$=\frac{1}{2}cr$

面积（△BOC）$=\frac{1}{2}$（底）×（高）$=\frac{1}{2}$（$\overline{BC}$）×（$\overline{OE}$）$=\frac{1}{2}ar$

面积（△COA）$=\frac{1}{2}$（底）×(高）$=\frac{1}{2}$（$\overline{AC}$）×（$\overline{OF}$）$=\frac{1}{2}br$

所以，K=面积（△ABC）=面积（△AOB）+面积（△BOC）+面积（△COA），或者，

$$K=\frac{1}{2}cr+\frac{1}{2}ar+\frac{1}{2}br=r(\frac{a+b+c}{2})rs$$

我们看到，赫伦在三角形的面积 K 与其半周长 s 之间建立了联系。这说明我们的方向走对了，但后面还有许多事情要做。

第二部分　我们继续参照图 5.6，并回想一下第一个初步命题，即利用三角形三个角的平分线作内接圆。因此，△ABC 可以分解为三对全等三角形，即

△AOD≌△AOF，△BOD≌△BOE，和△COE≌△COF，

这三对全等三角形，每一对都是根据“角角边”全等定理确定的（欧几里得，命题 I.26）。然后，将各部分相对应，我们得到

$\overline{AD}=\overline{AF}$，$\overline{BD}=\overline{BE}$，和$\overline{CE}=\overline{CF}$

而∠AOD=∠AOF，∠BOD=∠BOE，和∠COE=∠COF

现在，赫伦延长三角形的底边AB到G，并使$\overline{AG}=\overline{CE}$。然后，他推断

$\overline{BG}=\overline{BD}+\overline{AD}+\overline{AG}=\overline{BD}+\overline{AD}+\overline{CE}$根据作图

$=\frac{1}{2}(2\overline{BD}+2\overline{AD}+2\overline{CE})$

$=\frac{1}{2}$[（$\overline{BD}+\overline{BE}$）+（$\overline{AD}+\overline{AF}$）+（$\overline{CE}+\overline{CF}$）]根据全等

$=\frac{1}{2}$[（$\overline{BD}+\overline{AD}$）+（$\overline{BE}+\overline{CE}$）+（$\overline{AF}+\overline{CF}$）]

$=\frac{1}{2}[\overline{AB}+\overline{BC}+\overline{AC}]=\frac{1}{2}(c+a+b)=s$

因而，赫伦的线段 BG 的长度等于三角形的半周长，虽然“成了直线”。显然，赫伦是想得到成为一条直线的半周长。已知$\overline{BG}=s$，据此，我们可以很容易地得出

$s-c=\overline{BG}-\overline{AB}=\overline{AG}$

$s-b=\overline{BG}-\overline{AC}$

$=$（$\overline{BD}+\overline{AD}+\overline{AG}$）-（$\overline{AF}+\overline{CF}$）

$=$（$\overline{BD}+\overline{AD}+\overline{CE}$）-（$\overline{AD}+\overline{CE}$）$=\overline{BD}$

因为$\overline{AD}=\overline{AF}$，$\overline{AG}=\overline{CE}=\overline{CF}$。同样，

$$\begin{aligned} s-a &= \overline{BG}-\overline{BC} \\ &= (\overline{BD}+\overline{AD}+\overline{AG})-(\overline{BE}+\overline{CE}) \\ &= (\overline{BD}+\overline{AD}+\overline{CE})-(\overline{BD}+\overline{CE})=\overline{AD} \end{aligned}$$

因为$\overline{BD}=\overline{BE}$，$\overline{AG}=\overline{CE}$。

总之，半周长 s 与 s-a、s-b 和 s-c 三个量都等于图中的线段。这也是富有启发性的结论，因为这些量都是我们所求证公式的组成部分。赫伦剩下的工作就是要把这些“零件”组合成一个完整的证明。

第三部分　我们仍然设△ABC 及其内接圆，但我们现在需要一个延伸图，以说明赫伦的推理过程（图 5.7）。他先作 OL 垂直于 OB，并交 AB 于 K，然后作 AM 垂直于 AB，交 OL 于 H，最后，连接 BH。

由此形成的四边形 AHBO，我们应该很熟悉。根据命题 4，它实际上是一个联圆四边形；并且，根据命题 5，我们知道，四边形的对角和等于两个直角。即，

∠AHB+∠AOB=两个直角

现在，我们来看围绕内心 O 的各角。根据第二部分的全等，这些角可以分解为三对相等角，所以，

$$2\alpha+2\beta+2\gamma=\text{四个直角}\quad\text{或等于}$$

$$\alpha+\beta+\gamma=\text{两个直角}$$

但是，$\beta+\gamma=\angle AOB$，因此，$\alpha+\angle AOB=$两个直角$=\angle AHB+\angle AOB$。所以，$\alpha=\angle AHB$，这一点似乎是细枝末节，但在以下的推论中，将十分重要。

因为∠CFO 与∠BAH 都是直角，而且，根据上述推理$\alpha=\angle AHB$，所以，赫伦推断△COF 与△BHA 相似。据此，我们可以推出比例

$$\frac{\overline{AB}}{\overline{AH}}=\frac{\overline{CF}}{\overline{OF}}=\frac{\overline{AG}}{r}$$

因为$\overline{CF}=\overline{AG}$，$\overline{OF}=r$。我们还可以从这一比例式中推导出下列等式，并称之为（*）。

$$\frac{\overline{AB}}{\overline{AG}}=\frac{\overline{AH}}{r}\qquad(*)$$

赫伦还注意到，由于∠KAH 与∠KDO 都是直角，且对顶角∠AKH 与∠DKO 相等，因而，△KAH 与△KDO 也相似，并据此得出：

$$\frac{\overline{AH}}{\overline{AK}}=\frac{\overline{OD}}{\overline{KD}}=\frac{r}{\overline{KD}}\text{因而}\frac{\overline{AH}}{r}=\frac{\overline{AK}}{\overline{KD}}$$

将这一等式的最后结果与上述等式（*）合在一起，就得出了一个重要等式，我们称之为（**）。

$$\frac{\overline{AB}}{\overline{AG}}=\frac{\overline{AK}}{\overline{KD}}\qquad(**)$$

至此，读者难免会对这位数学家在这些无休无止的相似三角形中漫无目的地遨游感到不解。这种感觉到下一步时依然不会消失，因为赫伦在下一步又证明出了另外一对相似三角形。

赫伦注意到，在△BOK中，其高$\overline{OD}=r$。根据初步命题2，我们知道，△KDO 与△ODB 相似，因此，

$$\frac{\overline{KD}}{r}=\frac{r}{\overline{BD}}$$ 中简化为 $$(\overline{KD})(\overline{BD})=r^2 \qquad (**)$$

（希腊人会说，r是$\overline{KD}$与$\overline{BD}$这两个量之间的“比例中项”。）

现在，赫伦在等式（**）的两边分别加 1，得

$$\frac{\overline{AB}}{\overline{AG}}+1=\frac{\overline{AK}}{\overline{KD}}+1$$

化为公分母，成为

$$\frac{\overline{AB}+\overline{AG}}{\overline{AG}}=\frac{\overline{AK}+\overline{KD}}{\overline{KD}}$$ 或简化为 $$\frac{\overline{BG}}{\overline{AG}}=\frac{\overline{AD}}{\overline{KD}}$$

在这最后一个等式的左边乘以分数 $\frac{\overline{BG}}{\overline{BG}}$，右边乘以 $\frac{\overline{BD}}{\overline{BD}}$，等式当然依然成立，得

$$\frac{(\overline{BG})(\overline{BG})}{(\overline{AG})(\overline{BG})}=\frac{(AD)(\overline{BD})}{(KD)(\overline{BD})}$$ 因而

$$\frac{(\overline{BG})^2}{(\overline{AG})(\overline{BG})}=\frac{(AD)(BD)}{r2}$$ 根据（***G）

交叉相乘，得

$$r^2(\overline{BG})^2=(\overline{AG})(\overline{BG})(\overline{AD})(\overline{BD})$$

最后，赫伦将这大量“零件”组合，迅速而巧妙地达成他所求证的结论。我们只需注意到，这最后一个等式的组成部分恰恰是第二部分所推导出的线段。将第二部分的结果代入，便得到

$$r^2s^2=(s-c)(s)(s-a)(s-b)=s(s-a)(s-b)(s-c)$$

因而 $$rs=\sqrt{s(s-a)(s-b)(s-c)}$$

让我们再回忆一下第一部分的结论，如果 K 代表我们三角形的面积，则 rs=K。因此，最后代入上列等式，就得到了赫伦的公式：

$$K=\sqrt{s(s-a)(s-b)(s-c)}$$ 证讫。

这样，我们就完成了初等几何中一个最巧妙的证明。在证明过程中，他看似随意地漫游，实际上始终朝着预定目标前进。这无疑是我们迄今为止所见到的最曲折的证明。很难想象，脑力的回旋竟然引导赫伦得出了这样一个迂回曲折、令人惊叹的证明。

后记

关于这一著名公式，历史学家发现了一个离奇的事实。在一部写于赫伦身后几百年的古阿拉伯手稿中，伊斯兰学者阿布·赖汉·穆罕默德·比鲁尼认为，这一公式的发明不应归功于赫伦，而应归功于杰出的阿基米德。我们虽然没有阿基米德的论文来支持这种观点，但凭他的智力，本来完全可以推导出这样一个定理。

另一方面，抛开历史的本来面目不谈，出于情感原因，我们不妨让赫伦享有这一殊荣。如果将这个定理归功于阿基米德，而不是赫伦，似乎对前者过于慷慨，对后者又过于残酷，因为阿基米德的名声在古代数学家中已经无与伦比；而赫伦的名声在很大程度上却依赖于这个定理。

众所周知，赫伦的公式有其各种实用价值。测量员只要知道一块三边形地三条边的长度，就很容易计算出这块地的面积。对于四边形或更多边形的地，也不难分解成三角形的组合，进而求出面积。并且，利用赫伦的公式，还能够推导出一个我们早已熟悉的定理，下面我们来看。

假设有一个直角三角形，其斜边长度为 a，两个直角边分别为 b 和 c，如图 5.8 所示。因而，其半周长为

$$s=\frac{a+b+c}{2}$$ 我们发现

$$s-a=\frac{a+b+c}{2}-a=\frac{a+b+c}{2}-\frac{2a}{2}=\frac{-a+b+c}{2}$$

同样

$$s-b=\frac{2-b+c}{2} \quad \text{和} \quad s-c=\frac{a+b-c}{2}$$

代数运算进一步证明

$$(a+b+c)(-a+b+c)(a-b+c)(a+b-c)$$
$$=[(b+c)+a][(b+c)-a][a-(b-c)][a+(b-c)]$$
$$=[(b+c)^2-a^2][a^2-(b-c)^2]$$
$$=a^2(b+c)^2-(b+c)^2(b-c)^2-a^4+a^2(b-c)^2$$

简化为 $2a^2b^2+2a^2c^2+2b^2c^2-(a^4+b^4+c^4)$。

现在，我们再回到赫伦的公式，就得到三角形的面积为

$$K=\sqrt{s(s-a)(s-b)(s-c)}$$
$$=\sqrt{(\frac{a+b+c}{2})(\frac{-a+b+c}{2})(\frac{a-b+c}{2})(\frac{a+b-c}{2})}$$
$$=\sqrt{\frac{2a^2b^2+2a^2c^2+2b^2c^2-(a^4+b^4+c^4)}{16}}$$

另一方面，上述三角形的面积还可以简写成

$$K=\frac{1}{2}\text{（底）}\times\text{（高）}=\frac{1}{2}bc$$

将这两个 K 等式合成一个方程式，并将方程式两边分别平方，得

$$\frac{b^2c^2}{4}=\frac{2a^2b^2+a^2c^2+b^2c^2-(a^4+b^4+c^4)}{16}$$

然后，交叉相乘，得到

$$4b^2c^2=2a^2b^2+2a^2c^2+2b^2c^2-(a^4+b^4+c^4)$$

现在，把所有各项都移到方程左边，并合并同类项，得

$(b^4+2b^2c^2+c^4)-2a^2b^2-2a^2c^2+a^4=0$ 或简化为

$(b^2+c^2)^2-2a^2(b^2+c^2)+a^4=0$ 或再简化为

$$[(b^2+c^2)-a^2]^2=0$$

从这一长串演算中，最后，我们可以得出 $(b^2+c^2)-a^2=0$，并可以化为我们所熟悉的形式 $a^2=b^2+c^2$。这样，赫伦的公式就为我们提供了另一个毕达哥拉斯定理的证明。当然，这个证明极其复杂，很像从波士顿绕道斯波坎到纽约，似乎有点儿不必要。尽管如此，能够从赫伦古怪公式中发现对毕达哥拉斯定理的证明，虽然是间接的，毕竟值得注意。

欧几里得、阿基米德、厄拉多塞、阿波罗尼奥斯、赫伦以及其他许多

数学家都与亚历山大学派有关，这一古代的科学中心历久不衰。但是，即使罗马帝国也不是永恒的，亚历山大学派同样如此。

亚历山大图书馆从公元前约 300 年建立时起一直很活跃,直到公元 529 年被基督徒关闭（他们嫌恶图书馆收藏大量异教书籍）。公元 641 年，阿拉伯人最终将它付之一炬。虽然人们抢救出许多文献，但古代文明仍在这场劫难中元气大伤。其他古代遗迹，例如奇阿普斯（胡夫）的大金字塔、耶路撒冷的大卫庙和图书馆附近的法罗斯岛灯塔，也有类似的命运。今天的考古学家面对这种知识与美的永久毁灭，只能叹息再三。

数学活动从此结束了以亚历山大为中心的历史。从公元 641 年起，在以后的许多世纪中，阿拉伯数学家充当了古代数学的保护人，而他们本身，也是数学的创新者。阿拉伯帝国的故事当然应从穆罕默德（公元 570—632 年）开始，他最初默默无闻，继而成为世界史中的重要人物。穆罕默德死于耶路撒冷之后 150 年，他所创立的宗教已从印度穿过波斯湾和中东，横跨北非，一直扩展到西班牙南部。随着地理版图的扩大，伊斯兰学者急切地从他们接触到的许多文明中汲取知识。

这些知识之一是印度的数学，并从中产生了所谓“ 印度-阿拉伯 ”数系。这一数系远比罗马数系更先进，并取代了后者的地位，罗马数字的应用只在表盘、版权日期和超级橄榄球赛中还保留下来。即使阿拉伯人再没做其他事情，人们也会永志不忘他们传播这一最有用数系的历史功绩。

当然，他们还有其他许多贡献。早在 9 世纪初叶，阿拉伯人便开始翻译希腊名著，并对这些著作做了有益的评注。他们于公元 800 年翻译了《原本》，几十年后，又翻译出版了托勒密的名著《天文学大成》。这后一部著作写于约公元 150 年，是古代天文学的总汇。它很像欧几里得的《原本》，也是由 13 卷组成，包括论述日月蚀、太阳、行星和恒星的篇章，以及我们在第四章后记中所说到过的“ 弦值表 ”。托勒密还详细阐明了他的太阳系模式，这是一个以地球为中心的模式，这一地心学说，在波兰思想家哥白尼出现之前的 1400 年间，一直符合当时的科学与人类自尊的需要。阿拉伯人高度评价托勒密的著作，称其为“ Al magiste ”（阿拉伯语，“ 最伟大 ”的意思），因而，我们今天称这部巨著为“ 大成 ”（Almagest）。

后来，一位伟大的学者塔比特 · 伊本 · 戈拉（826—901 年）精译了阿基米德和阿波罗尼奥斯的著作，提供了非常忠实于原文的《原本》译本。阿拉伯学术中心位于今天伊拉克的巴格达市，那里建有“ 智慧宫 ”，是进行学术活动的重地，其成员中有许多天文学家、数学家和翻译家。数学界的中心以前在柏拉图的学园和亚历山大图书馆，现在转移到巴格达，在那里保持了很长一段时间。

在最重要的阿拉伯数学家中，有一位数学家名叫穆罕默德 · 伊本 · 穆萨，胡瓦里兹米（公元约 825 年），他借鉴东西方的数学成就——包括印度数学家婆罗摩笈多和我们所知的古希腊数学家，写出了一篇非常有影响的论代数和算术的论文。在这篇论文中，胡瓦里兹米不仅阐明了线性方程（一次方程）的解法，还阐明了曲线方程（二次方程）的解法。对于二次方程 $ax^2+bx+c=0$ 来说，其解为

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

但是，胡瓦里兹米完全是用文字，而不是用现今的简明的代数符号来

表达他的公式。然而，即使他没有发明代数符号，但却至少间接赋予了代数以名称。胡瓦里兹米的主要论文题为“Hisâb al-jabrw ’almuqâbalah”。400 年后，这篇论文被翻译成拉丁文，题目变为“Ludus algebrae et almucgrabalaeque”，最后，简称为“alge-bra”（代数）。

关于阿拉伯数学家的最大贡献，始终存有争议。一方面，他们在研究诸如欧几里得和阿基米德这些巨人的著作时，始终未能跳出其窠臼。在阿拉伯数学家的著作中，我们看不到那种数学知识的巨大飞跃，而这乃是古希腊数学家历代相沿的特点。特别是，阿拉伯数学家根本不认为“证明”是其数学的核心，在这个意义上，他们的数学酷似近东的希腊前文明。由于阿拉伯数学家不大重视将其成果归纳为一般原则，因而，本书没有阿拉伯数学家的伟大定理。

但另一方面，阿拉伯数学家确曾普及了非常有用的数系，并对解各次方程问题作出了重大贡献。此外，用霍华德·伊夫斯的话说，他们在欧洲沉睡的几百年间，充当了“大量世界智力财富的监管人。”如果没有这种伟大的服务，我们的许多古代文化知识，特别是古代数学知识，就有可能永远湮没无闻。

终于，阿拉伯人完成了他们作为欧几里得和阿基米德知识监管人的使命，这些知识又逐渐返回了欧洲。当然，其中一个主要的动力是从 11 世纪末到 13 世纪中叶十字军的一系列远征，在这些远征中，比较落后的西方基督教徒遭遇了比较先进的东方阿拉伯人。欧洲人虽然没能从穆斯林教徒手中夺取圣地，但都大开眼界，认识了敌手的高度学养。

也许，更重要的是基督教徒对西班牙摩尔人和对西西里地区的征服。1085 年，西班牙大城市托莱多陷落于基督教徒之手，几年后，西西里也被欧洲人征服。欧洲人在进入这些被占领的土地后，发现了落败的阿拉伯人的书籍和文献。欧洲人进入了一个以往难以想象的知识王国，闲暇时钻研一番，他们不仅领略了伊斯兰敌手的学识，而且也发现了自己祖先的学术成就。影响是巨大的。

这些经典（柏拉图和亚里士多德，当然还有欧几里得的著作）的巨大影响始见于意大利的大学。1088 年在博洛尼亚开办了第一所大学，随后又相继在帕多瓦、那不勒斯、米兰和其他地方办起了大学。此后一二百年，意大利的学术氛围从中世纪的深渊跃升至我们今天所称的文艺复兴的高点。

正是在 16 世纪的意大利,随着阿拉伯人对古代文化的传播和意大利学者的觉醒，产生了我们下一章的伟大定理：米兰的杰罗拉莫·卡尔达诺及其奇异而令人难以置信的三次方程解。

第六章
卡尔达诺与三次方程解
（1545 年）

霍拉肖代数的故事

毫无疑问，15 世纪的最后几十年标志着欧洲的文化骚动。西方文明显然已从中世纪的沉睡中觉醒。1450 年，约翰内斯 · 谷登堡发明了活字印刷术，从此，书籍大量流通。博洛尼亚、巴黎、牛津和其他地方的大学成为高等教育和学术活动的中心。在意大利，拉斐尔和米开朗基罗开创了非凡的艺术事业，而他们的前辈列奥纳多 · 达 · 芬奇则成为文艺复兴时期艺术家的杰出代表。

不仅仅是知识王国的疆域在扩展。1492 年，热那亚人克里斯托弗 · 哥伦布发现了大西洋彼岸的新世界。像其他事情一样，对美洲大陆的发现证明了当代文明的认知能力是能够超越辉煌的古代文明的。15 世纪结束时，欧洲无疑正处于出现伟大事变的前夕。

数学也是如此。1494 年，意大利数学家卢卡 ·帕西奥利（约 1445—1509 年）撰写了一部题为《算术大全》的书。在这部著作中，帕西奥利讨论了当代的标准数学，并重点讨论了一次方程和二次方程的解法。有趣的是，他在方程中用字母 co 代表未知量，无意中创造了原始的符号代数。co 是意大利语 cosa（意为“ 事物 ”）一词的缩写——即求解的事物。虽然 100 多年以后，代数才有了我们今天这样的符号系统，但《算术大全》却朝着符号代数方向迈出了一步。

然而，帕西奥利对三次方程（即一种形式为 $ax^3+bx^2+cx+d=0$ 的方程）的认识却是极其悲观的。他不知道应怎样解一般三次方程，并认为在现时的数学中，求解三次方程，犹如化圆为方一样，是根本不可能的。这种观点，实际上是对意大利数学界的一个挑战，并引出了关于下一个伟大定理的故事，即 16 世纪意大利代数学家和他们求解三次方程的故事。

故事是从博洛尼亚大学的希皮奥内 · 德尔 · 费罗（1465—1526 年）开始的。天才的费罗接受了帕西奥利的挑战，他发现了一个解所谓“ 缺项三次方程 ”的公式。所谓缺项三次方程，就是一个没有二次项的三次方程，其表现形式为 $ax^3+cx+d=0$。通常，我们习惯于用 a 去除方程的各项，并将常数项移到方程右边，这样，我们就可以将这一缺项三次方程转变为其标准形式

$$x^3+mx=n$$

出于明显的理由，文艺复兴时期的意大利人称这一方程为“ 立方加未知量等于数字 ”。虽然费罗只掌握了这种特殊形式的三次方程，但他对代数的推进却意义深远。人们或许会以为他将广为传播自己的成功，但实际上，他却完全不动声色。他对三次方程的解法绝对保密！

这种做法在“ 不发表即发霉 ”的今天，简直不可思议。为了能够理解费罗这种奇怪的做法，我们必须考虑到文艺复兴时期大学的特性。那时，大学里的学术职位没有安全感可言。除了保护人的庇护和政治方面的影响外，继续任职还取决于能否在任何地点任何时间赢得公开质疑。因而，像

费罗这样的数学家就必须随时准备与人进行学术论争，而当众出丑对于一个人的事业来说，可能是灾难性的。

因此，一个重要的新发现就是一件有力的武器。如果有一个对手提出一系列求解的问题，费罗就可以用一系列缺项三次方程来应付。即使费罗被他对手的某些问题难住了，他也可以相信，只有他一人掌握的三次方程注定了他那倒霉的对手必然失败。

希皮奥内对他的三次方程解法终生保密，直到弥留之际才将其传给了他的学生安东尼奥 · 菲奥尔（约 1506— ？）。虽然菲奥尔的才华比不上他的老师，但他利器在握，不禁心高气傲，于 1535 年向布雷西亚的著名学者尼科洛 · 丰塔纳（1499—1557 年）提出了挑战。

幼年时期一次不幸的灾难伴随了丰塔纳一生。1512 年，法国人进攻他的家乡时，一名士兵，手持利剑，在年幼的尼科洛脸上凶残地砍了一刀。据传说，这孩子能够活下来，完全是因为一条狗经常舔他脸上可怕的伤口。虽然狗的唾液挽救了他的性命，但却无法挽救他说话的能力。尼科洛 · 丰塔纳面目全非，以致再也不能清晰发言。于是，塔尔塔利亚（意为“结巴”）便成了他的绰号，而他今日正是以这一残忍的绰号而著称。

我们暂且抛开他的残疾不谈，塔尔塔利亚确是一位天才的数学家。实际上，他自称能够解出 $x^3+mx^2=n$ 形式的三次方程（即没有一次项的三次方程），但菲奥尔怀疑他是否真找到了这种解法。塔尔塔利亚受到菲奥尔挑战之后，便给菲奥尔寄去 30 道涉及各种数学问题的问题。而菲奥尔则回敬他 30 道“缺项三次方程”，使塔尔塔利亚陷于困境。显然，菲奥尔是在孤注一掷，塔尔塔利亚究竟能得 0 分，还是 30 分，就取决于他是否发现了解三次方程的秘密。

毫不奇怪，塔尔塔利亚开始夜以继日地疯狂研究缺项三次方程。日子一天天过去，他越来越沮丧。眼看最后期限就要到了，终于，1535 年 2 月 13 日夜，塔尔塔利亚发现了三次方程的解法。他的努力终于得到了回报。他现在可以轻而易举地解出菲奥尔的所有问题，而他的平庸的挑战者则成绩平平。塔尔塔利亚光荣地战胜了对手。作为酬报，倒霉的菲奥尔应以丰盛的酒宴款待塔尔塔利亚 30 次；但塔尔塔利亚却以一种宽宏的姿态，免却了这一约定。与受到的羞辱相比，省下的钱财对于菲奥尔来说实在微不足道；于是，菲奥尔从此销声匿迹。

接着出现的也许是整个数学史中最奇特的人物——米兰的杰罗拉莫 · 卡尔达诺（1501—1576 年）。卡尔达诺听说了有关这一挑战的故事后，就想更多地知晓塔尔塔利亚这位三次方程大师奇妙的技巧。卡尔达诺大胆地要求塔尔塔利亚这位布雷西亚学者公开他的秘密，从此，故事发生了意想不到的重大转折。

在继续叙述之前，我们先来看一看杰罗拉莫 · 卡尔达诺不平凡的一生。我们有幸在他写于 1575 年的自传《我的一生》中读到他第一人称的叙述。这本书充满了卡尔达诺的回忆、怨恨和迷信，还有大量奇闻轶事。虽然在几乎全部自传中，这一本自传是最不可信的，但我们从中却可以窥见他动荡不安的一生。

卡尔达诺首先追述了他的先祖。在他的家谱中可能包括教皇切莱斯廷四世，还有他的一个远亲安焦洛。安焦洛在 80 岁高龄时

“才得儿子——孩子像他们父亲一样衰弱……他的长子活了 70 岁，我

听说他的子女中有些成了伟人。”

接着，在《我的诞生》一章中，卡尔达诺写道：“我听说，虽然用了各种堕胎药，但都无效”，他活下来了，严格地说，只是“从我母亲的子宫里拖出来了”。这种方式使他近于夭折，用温酒洗浴才活了下来。卡尔达诺似乎是一个私生子，这才能解释他何以不受欢迎。伴随而来的耻辱影响了他的一生。

由于先天不足，卡尔达诺终生经受疾病的折磨就不足为怪了。在他的自传中，他坦率描述了这些痛苦，常常刻划入微，甚至到了令人厌恶的地步。他告诉我们，他患有严重的心率过速，胸腹部流出液体，还患有脱肛和痔疮，以及一种“排尿过多”的疾病，每天排尿多达 100 盎斯（约一加仑）。他惧怕登高和前往“据说疯狗出没过的地方”。他多年患有阳萎，直到临近结婚时才痊愈（无疑正是时候）。卡尔达诺常常连续八个夜晚失眠，这种时候，他只能“起床下地，绕着床转圈，一遍又一遍地数数，数到一千。”

偶尔不受这些疾病折磨时，卡尔达诺就自己折磨自己。他这样做是因为“我觉得快乐存在于强烈痛苦之后的放松”，而且，当身体上不受痛苦的时候，“精神上的痛苦就必然会来压迫我，没有什么能比这种痛苦更强烈的了”，所以，

“我想出了一个办法，用力咬我的嘴唇，拧我的手指，掐我左臂的肌肉，直到疼得流出眼泪为止。”卡尔达诺说，这种自我折磨还算值得，因为一旦停止下来，就会感到非常惬意。

然而，身体（和精神）上的疾患还不是他唯一的问题。卡尔达诺在帕多瓦大学以优异成绩完成他医学学业之后，却不能获准在他的家乡米兰行医。究其原因，可能是因为他是人人皆知的私生子，也可能是因为他那讨厌而古怪的个性，但不论什么原因，这在他一生的沉浮中标志着一个低潮。

在米兰遭到拒绝，卡尔达诺就转移到帕多瓦附近的一个小镇萨科，在乡间行医，那里不乏田园风光，但多少有些闭塞。在萨科的一天夜里，他梦见了一个身穿白衣的漂亮女人。他很信梦，因此，当有一天，他遇到了一个与他梦中所见完全一样的女人，不免受到极大震动。起初，贫困的卡尔达诺因为不能向她求爱而深感绝望：

> “如果我，一个穷人，娶一个女人，没有嫁妆，只有一大群弟妹需要供养，那我就完蛋了！我甚至连自己也养活不起！如果我试图诱拐她，或勾引她，周围又会有多少双眼睛在监视我！”

但终于，他的爱赢得了婚姻。1531 年，他娶了梦中的女人卢西亚·班达里妮为妻。

这段小插曲表明了梦、先兆和预兆在卡尔达诺的一生中所起的突出作用。他是一位热心的占星术士，一位护身符佩戴者，一位从雷雨中预卜未来的预言家。并且，他还常常感到守护神的存在，他在自传中写道：

“据说守护神……常常对某些人特别垂青——苏格拉底、柏罗丁、辛纳修斯、戴奥、弗莱维厄斯·约瑟夫斯，我觉得自己也包括在内。所有这些人，除了苏格拉底和我之外，都生活得非常幸福……”

显然，他很乐意与他的守护神热烈交谈。卡尔达诺 20 世纪的传记作家奥伊斯坦·奥尔说：“由于这类故事，无怪他的一些同时代人认为他精神不正常。”

他的另一个终生爱好是赌博。卡尔达诺经常沉湎于赌博，他常常能赢许多钱，贴补收入。他在自传中以忏悔的心情承认：

“……我过度沉溺于轮盘赌和掷骰子，我知道，我应该受到最严厉的批评。我染上这两种赌瘾有许多年了；不仅年年赌，而且，我羞愧地承认，是天天赌。”

幸好，卡尔达诺将这一恶习提到科学研究的高度。他为此撰写《论赌博》，死后于 1663 年出版，这是第一部论及数学概率的重要论文。

这样，杰罗拉莫·卡尔达诺从 1526 年至 1532 年，在萨科生活了许多年，他在那里算命、赌博，并成了家。但是，不论是他的收入，还是他的自尊，都使他不能长期忍受小镇的环境。1532 年，卡尔达诺携其妻子卢西亚与儿子詹巴蒂斯塔一道返回米兰，但他仍然被禁止行医，最后不得不依靠贫民院的救济过活。

终于，好运降到了他的头上。卡尔达诺开始讲授大众科学，这种讲演特别受到有教养的人和贵族的欢迎。他撰写了许多有趣的论文，论题从医学、宗教到数学，内容极为广泛。特别是 1536 年，他发表了一篇论文，攻击意大利医生中的腐败和不称职现象。这篇文章无疑得罪了医学界，但却受到公众的欢迎，医学界再不能将卡尔达诺拒之门外。1539 年，米兰医师协会勉强接收他为会员，不久，他就赢得了行业的最高声誉。到 16 世纪中叶，卡尔达诺已成为也许是欧洲最著名和最受欢迎的医生。他曾为教皇治过病，也曾越洋去苏格兰（这在当时是一个漫长而艰难的旅程）为圣安德鲁的大主教治病。

但是，好景不长，不久就接连发生了悲剧。1546 年，他的妻子去世了，年方 31 岁，留给卡尔达诺两个儿子、一个女儿。在这些子女中，长子詹巴蒂斯塔是卡尔达诺的希望与欢乐。这个孩子非常聪明，他在帕维亚大学获得了医学学位，子承父业，前途不可限量。但是，灾难像“疯女人”（卡尔达诺语）一般袭来。他写道，1557 年 12 月 20 日晚，“……我正当睡意朦胧之际，床突然抖动起来，继而整个卧室都在震动。”第二天早上，卡尔达诺从询问中得知，全城没有任何其他人感觉到了夜里的震动。卡尔达诺认为这是一个凶兆。他刚一得出这个结论，仆人就带来一个意想不到的消息：詹巴蒂斯塔娶了一个“平庸或没有任何可取之处”的女人为妻。

后来证明这果然是一桩不幸的婚姻。詹巴蒂斯塔的妻子生了三个孩子，她自称，没有一个是詹巴蒂斯塔的。她的不贞，乃至不知羞耻，令詹巴蒂斯塔失去了理智。为了报复，他在给妻子的糕点里下了砒霜。砒霜果然有效，而詹巴蒂斯塔自己也以谋杀罪被捕。卡尔达诺凭借他的声望，作了不懈的努力，但一切都无济于事；他的爱子罪名成立，并于 1560 年 4 月初被推上断头台。

“家门不幸，以此为甚。”极度悲痛的卡尔达诺写道。他心如死灰，失去了他的朋友、事业，甚至生活的兴趣。与此同时，他的另一个儿子阿尔多也成了罪犯，实际上，卡尔达诺‘不得不一次又一次地将他送进监狱”。令人心碎的事情似乎一件接着一件。1562 年，他离开米兰这座记载着他的成功与不幸的城市，接受了博洛尼亚大学的一个医学教职。陪同他一起的是他的孙子，詹巴蒂斯塔的儿子法齐奥。在他垂暮之年，这位老人与孩子之间也许发展了一种强烈的友爱关系，使他享受到了他自己的子女未能给予他的天伦之乐。

但是，年幼的孙子和新城市也未能给他动荡的生活带来宁静。1570年，卡尔达诺以异端罪被捕入狱。当时，意大利教会对宗教改革运动的异端采取了强硬态度，卡尔达诺曾为耶稣占星，并写了一本《尼禄颂》，记述这位可恨的反基督教的罗马皇帝，教会当然大为不快。

监禁和羞辱似乎使年迈的卡尔达诺名誉扫地。然而，一些有名望的朋友们为他讲情，加上教会的宽恕，卡尔达诺不久即被释放出狱，他来到罗马，不知怎么竟得到了教皇颁发的养老金！所谓否极泰来，大约就是这样的了。卡尔达诺恢复名誉后，与他心爱的孙子一道，度过了他的晚年。他在自传中骄傲地写道，虽然他年事已高，但仍有“十四颗好牙和一颗有点儿松动的牙；但我想，这颗牙会存在很长时间，因为它还好用。”卡尔达诺在比较平静的气氛里度过了他的晚年，并于 1576 年 9 月 20 日安祥地死去，结束了他充实的一生。

对于现代读者来说，卡尔达诺是一个自相矛盾但却依然十分迷人的人物。他的著述多得令人难以置信，累计达 7000 页，广泛涉及从科学到其他领域的各种主题。但他虽然一只脚站在现代理性世界，另一只脚却站在中世纪迷信的非理性世界。就在他谢世一百年后，伟大的哲学家兼数学家戈特弗里德·威廉·莱布尼兹恰当地概括了他的一生：“卡尔达诺是一个有许多缺点的伟人；没有这些缺点，他将举世无双。”

我们现在再回到三次方程的问题，卡尔达诺对解三次方程作出了重大贡献。1535 年，布雷西亚的塔尔塔利亚发现了某些类型三次方程的解法，从而战胜了安东尼奥·菲奥尔。卡尔达诺极感兴趣，他一次又一次地写信给塔尔塔利亚，请求塔尔塔利亚告诉他三次方程的解法，当然，他一次又一次地遭到拒绝，因为塔尔塔利亚决心趁势写一部解三次方程的书。卡尔达诺起初非常生气，但终于好言好语将塔尔塔利亚请到米兰作客。1539 年 3 月 25 日，塔尔塔利亚向卡尔达诺公开了他解缺项三次方程的秘密，但他是用密码书写的。卡尔达诺为此庄严宣誓：

“谨对着神圣的福音书，以君子的信义向你发誓，如果你把你的发现告诉我，我不仅绝不发表，而且还以我一个真正基督教徒的忠诚保证并发誓也用密码记录，这样，在我死后，就没有人能够读懂这些密码。”

现在，这场戏剧中的最后一个人物出现了。这就是年轻的卢多维科·费拉里（1522—1565 年），他敲开卡尔达诺的门，要求找一份工作。那天，卡尔达诺听到喜鹊不停地叫，知道是个吉兆，便急忙收下这个孩子为仆。小卢多维科很快显现出是一个绝顶聪慧的神童。他们的关系很快便从主仆关系发展为师生关系，最后，在费拉里不到 20 岁的时候，他们的关系又转变为伙伴关系。卡尔达诺将塔尔塔利亚的秘密告诉给了他聪明而年青的弟子，两人共同努力，取得了惊人的进展。

例如，卡尔达诺发现了如何求解一般三次方程 $x^3+bx^2+cx+d=0$

在这里，系数 b、c、d 可以是 0，也可以不是 0。但遗憾的是，卡尔达诺的工作是立足于将一般三次方程化为缺项三次方程，这样就遇到了为塔尔塔利亚保守秘密的问题。与此同时，费拉里也成功地发现了解四次多项式方程的方法。这是代数上的一个重大发现，但它也是依据化四次方程为相关的三次方程的方法，同样也受制于卡尔达诺的誓言而不能发表。他们两人都作出了当时代数学中最大的发现，但却都陷入了困境。

后来，1543 年，卡尔达诺与费拉里一起来到博洛尼亚，他们仔细查看

了希皮奥内·德尔·费罗的论文。对于费罗来说，这一整个故事早在三十年前就已开始了。他们在论文中看到了费罗亲手写的缺项三次方程的解法。它对卡尔达诺的含义是十分清楚的：他不必再受限制而不能发表这一解法了，因为这是费罗，而不是塔尔塔利亚发现的，他当然可以接受费罗的启示。急切的卡尔达诺才不管费罗与塔尔塔利亚的解法其实完全相同。

1545 年，卡尔达诺出版了他的数学名作《大衍术》。对于卡尔达诺来说，代数是一门“ 伟大的艺术 ”，而他的著作代表了代数学中一个惊人的突破。《大衍术》共包括 40 章，开始几章只讨论了一些简单的代数问题，而在题为“ 论三次方加一次方等于常数 ”的第十一章中，最终展现了三次方程的解式。值得注意的是，卡尔达诺为这关键的一章写了如下的序言：

“ 博洛尼亚的希皮奥内·费罗在近三十年前便已发现了这一规则，并将其传给了威尼斯的安东尼奥·马里亚·菲奥尔；而菲奥尔与布雷西亚的尼科洛·塔尔塔利亚的竞赛使尼科洛有机会发现了这一解法。后来，塔尔塔利亚应我的恳求，向我公开了他的发现，但保留了对这一解式的证明。在这一帮助下，我发现了（各种）形式的证明。这是极为艰难的。 ”

卡尔达诺在此赞誉了许多人，这种赞誉是公正的。除了塔尔塔利亚以外，人人都感到满意。而塔尔塔利亚则相反，他对卡尔达诺的欺骗和背叛行为大为恼怒。在塔尔塔利亚看来，卡尔达诺违背了他神圣的誓言，他曾以一个“ 真正基督教徒 ”的忠诚发誓，但他却是一个不折不扣的恶棍。塔尔塔利亚提笔问罪，但回答他的却不是卡尔达诺（他想凌驾于这场争斗之上），而是顽强忠诚的费拉里。费拉里以其脾气暴躁著称（他曾在一次恶性争斗中失去了几个手指），他激烈地驳回了塔尔塔利亚的指责。一时间，在布雷西亚与米兰之间，火药味十足的信件飞来飞去。例如，在 1547 年的一封信中，费拉里斥责塔尔塔利亚是一个

“ ……整天忙于……斤斤计较的人。如果要我报答你，我就给你肚里塞满草根和胡萝卜块，让你一辈子再也咽不下别的东西。 ”

（最后一句话是双关语，暗指在解三次方程问题中随处可见的数学乘方根。）

1548 年 8 月 10 日，塔尔塔利亚与费拉里在米兰的一次公开论战使冲突白热化。塔尔塔利亚后来指责卡尔达诺的缺席，说他“ 避免在论战中露面 ”是一种怯懦的表现。但是，这场论战是在费拉里的家门口进行的，最后以客座一方的失败而告结束。塔尔塔利亚埋怨观众的喧闹和偏见，而费拉里则当然把事情的结局归功于他自己的智力。但不管怎么说，塔尔塔利亚败下阵去，而费拉里则大获全胜。数学史家霍华德·伊夫斯注意到观众的敌意和费拉里暴躁鲁莽的名声，他说，塔尔塔利亚能够活着逃回去，还算是他的造化。

这些就是围绕着三次方程解所发生的故事，激烈、复杂，而又不免荒唐。现在我们所要做的，就是要讨论作为这一奇特故事核心的伟大定理。

伟大的定理：三次方程的解

现代读者在阅读《大衍术》第十一章时，会有两点感到意外。其一是卡尔达诺并没有给出解一般三次方程的证明，只列举了一种特殊形式的缺项三次方程，即

$$x^3+6x=20$$

我们在以下的讨论中，将采用更一般的形式

$$x^3+mx=n$$

其二是卡尔达诺的论证是一种纯几何式的，涉及真正的立方体及其体积。实际上，我们只要想一想当时代数符号的原始状态和文艺复兴时期数学家对古希腊几何的看重，疑团便会烟消云散。

本书用卡尔达诺自己的语言阐述了《大衍术》第十一章的重要命题，并附上了他对三次方程的巧妙分析。他用文字叙述的解三次方程的“ 法则 ”初看非常混乱，但如果用一种更常见的代数方法重新演算一遍，就会发现卡尔达诺的规则是正确无误的。

定理　解 $x^3+mx=n$ 的法则：

用 x 系数三分之一的三次方加上方程常数一半的平方；求这整个算式的平方根。复制（重复）这一算式，并在第一个算式中加上方程常数的一半，从第二个算式中减去同一数的一半……然后，用第一个算式的立方根减去第二个算式的立方根，其差即为 x 的值。

证明　卡尔达诺设想了一个大立方体，其边 AC 的长度，我们用 t 来表示，如图 6.1 所示。AC 边于 B 点截取线段 BC，其长度为 u，则线段 AB 的长度为 t—u。这里的 t 和 u 都是辅助变量，我们必须确定它们的值。如图所示，大立方体可以分为 6 部分，各部分的体积我们确定如下：

- 前下角小立方体的体积为 u^3
- 后上角较大立方体的体积为 $(t?\ u)^3$
- 两个垂直板块，一个沿 AB 向前，另一个沿 DE 向右，每一个长方体的边长分别为 t? u、u 和 t（大立方体的边长），因而，每一个长方体的体积分别为 $tu(t?\ u)$
- 前上角细长的长方体，其体积为 $u^2(t?\ u)$
- 在后下角，即较大立方体的下面，有一个扁平的立方体，其体积为 $u(t?\ u)^2$

显然，大立方体的体积 t^3 等于这 6 个小立方体的体积之和，即

$t^3=u^3+(t?\ u)^3+2tu(t-u)+u^2(t?\ u)+u(t?\ u)^2$

对方程式中的各项做一些整理，即得到

$(t-u)^3+\ [2tu(t?\ u)+u^2(t?\ u)+u(t?\ u)^2]\ =t^3-u^3$

从方括号中提取公因数 $(t?\ u)$，得

$(t?\ u)^3+(t?\ u)\ [2tu+u^2+u(t?\ u)]\ =t^3-u^3$ 或简化为

$$(t?\ u)^3+3tu(t?\ u)=t^3-u^3 \qquad (*)$$

（现代读者会注意到，这一方程式可以用简单的代数方法直接推导出来，而无需借助于神秘的几何立方体和板块。但 1545 年的数学家们还不可能采用这种方法。）

（*）方程式很容易使我们联想起最初的三次方程式的形式 $x^3+mx=n$。也就是说，如果我们设 $t?\ u=x$，则（*）方程就变为 $x^3+3tux=t^3-u^3$，然后，我们再设

$$3tu=m \text{ 和 } t^3-u^3=n$$

现在，如果我们能用原三次方程式中的 m 和 n 来确定 t 和 u 的值，那么，x

＝t？u 就能够推导出我们所求证的定理。

但是，《大衍术》没有推导这些量的值。相反，卡尔达诺直接提出了求解前述“三次方加一次方等于常数”的法则。要明确译解他纯用文字表达的解题方法绝非易事，这就使人更加赏识现代代数公式这种简明而直接的解题方法。卡尔达诺在这一段文字中究竟讲的是什么意思呢？

首先，我们来看他对 t 和 u 所规定的两个条件，即

$$3tu=m \text{ 和 } t^3-u^3=n$$

从第一个等式中，我们可以导出$u=\frac{m}{3t}$，将其代入第二个等式，即得到

$$t^3-\frac{m^3}{27t^3}=n$$

将方程两边分别乘以 t^3，经整理后，就得到方程

$$t^6-nt^3-\frac{m^3}{27}=0$$

初看似乎并没有什么改进，因为我们把原来 x 的三次方程变成了 t 的六次方程。然而，后者却可以看作变量 t^3 的二次方程：

$$(t^3)^2-n(t^3)-\frac{m^3}{27}=0$$

而数学家早已掌握二次方程的解法，我们在前一章的后记中也讲到过这一点。现在，我们可以解出这个二次方程：

$$t^3=\frac{n\pm\sqrt{n^2+\frac{4m^3}{27}}}{2}$$

$$=\frac{n}{2}\pm\frac{1}{2}=\sqrt{n2+\frac{4m^3}{27}}=\frac{n}{2}\pm\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}$$

然后，只选取正平方根，我们就得到

$$t=\sqrt[3]{\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}}$$

我们还知道 $u^3=t^3-n$，据此，我们得出

$$u^3=\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}-n \quad \text{或}$$

$$u=\sqrt[3]{-\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}}$$

最后，我们就得到了用代数式表达的卡尔达诺解缺项三次方程 $x^3+mx=n$ 的法则，即

$$x=t-u$$

$$=\sqrt[3]{\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}}-\sqrt[3]{-\frac{n}{2}+\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}}$$

这一方程式就叫做缺项三次方程的“根式解”或“代数解”。也就是说，这一解式只涉及了原方程的系数（即 m 和 n），而且，代数运算一般

也只限于加、减、乘、除和开方。再深入一些的研究表明，这一公式与卡尔达诺用文字阐述的“法则”结果完全相同。

卡尔达诺论证中的最精彩之处在于他用相关的（t^3的）二次方程解替代了三次方程解，并从而发现了将方程降低“一次”的方法，这样，他就从生疏的三次方程进入了熟悉的二次方程。这一非常巧妙的方法开辟了解四次、五次和更高次方程的道路。

例如，卡尔达诺用这种方法解出了他的原型方程 $x^3+6x=20$。按照卡尔达诺的方法，他首先求出x系数三分之一的三次方，即$(\frac{1}{3}\times 6)^3=8$；然后，他求出常数项一半（即 20 的一半）的平方，得 100，再加上 8，其和为 108，求出这个数的平方根。他再用这个平方根加上和减去常数项的一半，得到$10+\sqrt{108}$和$-10+\sqrt{108}$，他最后的解是这两个数立方根的差：

$$x=\sqrt[3]{10\sqrt{108}}-\sqrt[3]{-10+108}$$

当然，我们可以简单地用 m＝6 和 n＝20 代入有关代数式，就得到

$$\sqrt{\frac{n^2}{4}+\frac{m^3}{27}}=\sqrt{108}$$ 因而

$$x=\sqrt[3]{10+\sqrt{108}}-\sqrt[3]{-10+108}$$

显然，这是一个“根式解”。令人感到意外的是，正如卡尔达诺所正确指出的那样，这一貌似复杂的方程式实际上只不过是数字“2”的伪装而已，用计算器不难验证这一点。人们已经看出，x＝2 确是 $x^3+6x=20$ 的解。

有关解方程的其他问题

我们注意到，在知道了三次方程的一种解法后，就可以据此去发现一些其他类型的解法。例如，因为 x＝2 是上述方程的解，而且，我们知道 x? 2 是 x^3+6x? 20 的一个因式，经过长除后，就可以得到另一个二次方的因式。因而，x^3+6x? $20=(x-2)(x^2+2x+10)$。这样，解原三次方程的问题就变成了解一次方程和二次方程

$x-2=0$ 和 $x^2+2x+10=0$

这样简单的问题。（因为此二次方程无实数解，所以，原三次方程只有一个实数解 x＝2。）

对于现代读者来说，《大衍术》接下来的两章似乎是多余的。卡尔达诺第十二章的标题是“论三次方等于一次方加常数”（即 $x^3=mx+n$），第十三章的标题是“论三次方加常数等于一次方”（即 $x^3+n=mx$）。今天，我们认为，这两种形式的三次方程完全可以包括在上述方程式中，因为我们可以使 m 和 n 为负数。但是，16 世纪的数学家却要求方程的所有系数都必须是正数。换句话说，他们认为，$x^3+6x=20$ 与 $x^3+20=6x$ 不仅形式不同，而且是本质上完全不同的两种方程。由于卡尔达诺是以三维立方体的概念来看待三次方程的，所以，在他看来，立方体的边长为负数是没有意义的，因而，他们对负数项持否定态度就不足为怪了。当然，避免采用负数项就会使方程的种类增多，按照我们今天的看法，不必要地拉长了《大衍术》的篇幅。

这样，卡尔达诺能够解三种形式的缺项三次方程中的任何一种。但是，对于$ax^3+bx^2+cx+d=0$这种一般形式的三次方程又当如何呢？卡尔达诺的伟大发现在于，通过适当的置换，可以将这一方程转换为相关的缺项三次方程，当然，必须要符合他的公式。在讨论三次方程的这一“缺项”过程之前，我们不妨浏览一下一种更熟悉的解题方法——即应用于解二次方程的方法：

我们首先设二次方程的一般形式为

$$ax^2+bx+c=0 \text{ 这里 } a\neq 0$$

为了使之缺项——即消去一次项，我们引入一个新的变量 y，用 $x=y-\frac{b}{2a}$来替换x，就得到

$$a(y-\frac{b}{2a})^2+b(y-\frac{b}{2a})+c=0$$ 并由此得出

$$a\left(y^2-\frac{b}{a}y+\frac{b^2}{4a^2}\right)+by-\frac{b^2}{2a}+c=0$$ 或

$$ay^2-by+\frac{b^2}{4a}+by-\frac{b^2}{2a}+c=0$$

然后，消去 by 项，就得到缺项二次方程

$$ay^2=\frac{b^2}{2a}-\frac{b^2}{4a}-c=\frac{2b^2}{4a}-\frac{b^2}{4a}-\frac{4ac}{4a}=\frac{b^2-4ac}{4a}$$

因此

$$y^2=\frac{b^2-4ac}{4a^2} \qquad y=\frac{\pm\sqrt{b^2-4ac}}{2a}$$

最后

$$x=y-\frac{b}{2a}=\frac{\pm\sqrt{b^2-4ac}}{2a}-\frac{b}{2a}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

这样就再现了解二次方程公式。

这个例子说明，多项式的降次方法是非常有用的。了解了这种方法以后，我们再回到卡尔达诺解一般三次方程的问题上来。在这里，关键的替换量是$x=y-\frac{b}{3a}$，由此得出

$$a(y-\frac{b}{3a})^3+b(y-\frac{b}{3a})^2+c(y-\frac{b}{3a})+d=0$$

展开后，成为

$$(ay^3-by^2+\frac{b^2}{3a}y-\frac{b^3}{27a^2})+(by^2-\frac{2b^2}{3a}y+\frac{b^3}{9a^2})$$
$$+(cy-\frac{cb}{3a})+d=0$$

对这一堆字母，我们需要做的一件重要事情就是消去 y^2 项。这样，新的三次方程（正如我们所希望的那样，）就没有了二次项。如果我们用 a 去除各项，就得到 $y^3+py=q$ 这种形式的方程。我们可以用卡尔达诺的公式求出y的值，因而，也就不难确定$x=y-\frac{b}{3a}$的值了。

为了更清楚地说明这一过程，我们来看三次方程

$$2x^3-30x^2+162x-350=0$$

代入 $x=y-\frac{b}{3a}=y-(-\frac{30}{6})=y+5$，得

$$2(y+5)^3-30(y+5)^2+162(y+5)-350=0$$

整理后，成为

$$2y^3+12y-40=0 \text{ 或简化为 } y^3+6y=20$$

显然，这就是我们前面所解过的缺项三次方程，因而我们知道 $y=2$。所以，$x=y+5=7$，并可以此验证原方程。

但是，《大衍术》在论证解一般三次方程问题时，却远非我们这样简洁。由于卡尔达诺要求所有系数都只能是正数，他就必须跨越一连串艰难的障碍，诸如，“三次方加二次方加一次方等于常数”、“三次方等于二次方加一次方加常数”、“三次方加常数等于二次方加一次方”，等等。终于，他在解出缺项三次方程后，又用了 13 章的篇幅才完成了这一论证，从而解决了解三次方程的问题。

但果真解决了吗？虽然卡尔达诺的公式似乎是一个惊人的成就，但它却带来了一个重大的谜。例如，我们来看缺项三次方程 $x^3-15x=4$。

用 $m=-15$ 和 $n=4$ 代入上述公式，我们就得到

$$x=\sqrt[3]{2+\sqrt{-121}}-\sqrt[3]{-2+\sqrt{-121}}$$

如果说 16 世纪的数学家对负数持怀疑态度,则负数的平方根显然就是绝对荒谬的，当然可以将其作为不可解的三次方程而予以排除。然而，对于上述三次方程来说，却可以很容易验证出它有三个不同的和完美的实数解：$x=4$ 和 $x=-2\pm\sqrt{3}$。究竟是什么原因使得卡尔达诺产生了这种情况——所谓“三次方的不可约情形”呢？他也曾对我们今天称之为“虚数”或“复数”的情况进行过几次不太认真的研究，但最终还是全部放弃，因为它们“既捉摸不透，又没有用处”。

大约又经过了一代人的时间，拉斐罗·邦贝利（约 1526－1573 年）出现了，他在 1572 年的论文《代数》中迈出了勇敢的一步，他将虚数看作是运载数学家从实数三次方程到达其实数解的必要工具；也就是说，我们从熟悉的实数领域出发并最终回到实数，但中途却不得不进入一个我们所不熟悉的虚数世界以完成我们的旅程。对于当时的数学家来说，这似乎是不可思议的。

现在，我们来简要讨论一下邦贝利的论述。我们暂且忽略对 $\sqrt{-1}$ 的任何潜在的偏见，求出 $2+\sqrt{-1}$ 的三次方，得到

$$(2+\sqrt{-1})^3=8+12\sqrt{-1}-6-\sqrt{-1}$$
$$=2+11\sqrt{-1}=2+\sqrt{-121}$$

但是，如果 $(2+\sqrt{-1})^3=2+\sqrt{-121}$，我们当然就可以说

$$\sqrt[3]{2+\sqrt{-121}}=2+\sqrt{-1}$$

同样，我们还可以看到$\sqrt[3]{-2+\sqrt{-121}}=-2+\sqrt{-1}$。那么，我们再来看三次方程$x^3-15x=4$，邦贝利求出其解

$$x=\sqrt[3]{2+\sqrt{-121}}-\sqrt[3]{-2+\sqrt{-121}}$$
$$=(2+\sqrt{-1})-(-2+\sqrt{-1})=4$$

答案正确！

大家公认，邦贝利方法所提出的问题远远超出了他所解的问题。例如，怎样才能预先知道$2+\sqrt{-1}$就是$2+\sqrt{-121}$的立方根呢？直到18世纪中叶，莱昂哈尔德·欧拉才找到了一个发现复数根的可靠方法。此外，究竟什么是虚数，虚数的性质是否与实数相同呢？

诚然，复数的重要性直到 200 多年以后的欧拉、高斯和柯西时代才充分地显现出来，我们将在第十章的后记中详细介绍这个问题。尽管如此，邦贝利承认了复数在代数中的作用，应当得到赞誉，他因此成为 16 世纪最后一位伟大的意大利代数学家。

这里应强调一点。与人们普遍认为的相反，虚数不是作为解二次方程工具，而是作为解三次方程的工具进入数学王国的。当然，在$\sqrt{-121}$作为$x^2+121=0$（这一方程显然没有实数解）的解时，数学家可以轻易地将其排除。但是，在解上述三次方程时，对于$\sqrt{-121}$在导出$x=4$时所起的关键作用，就不能如此漠然置之了。因此，是三次方程，而不是二次方程，给了复数以原动力和它们今天无可争辩的合法地位。

我们还应对《大衍术》作最后一点评论。在其第三十九章中，卡尔达诺用文字说明了解四次方程的方法：

“还有另外一个法则，并且，比前一个法则更为壮观。这就是卢多维科·费拉里提出的法则，他应我的要求，将其发现交给了我。根据费拉里法则，我们可以求出所有四次方程的解。”

这是一个非常复杂的程序，其中两个关键性的步骤很值得一提：

1.设一般四次方程 $ax^4+bx^3+cx^2+dx+e=0$，代入 $x=y-\frac{b}{4a}$，使之缺项，并用a去除方程各项，就得到一个y的缺项四次方程：

$y^4+my^2+ny=p$

2.通过巧妙地引入辅助变量，就可以用相关的三次方程替代原四次方程，然后，可以用上述方法解出这个三次方程。在这里，费拉里再次采用了经验的做法，即用降幂的方法解出一定次数的方程。

那些有能力阅读这一定理及《大衍术》中所有其他发现的读者，掩卷之后势必感慨万端。解方程的艺术达到了新的高度，而卢卡·帕西奥利当初认为代数不能解三次方程（更不要说四次方程了）的观点已被彻底粉碎。无怪乎卡尔达诺在《大衍术》结尾时热烈而动情地写道：“用五年时间写就的这本书，也许可以持续几千年。”

后记

卡尔达诺－费拉里著作中一个悬而未答的问题是五次方程的代数解。他们的努力显然表明，五次方程的根数解是可能的，并且，他们对如何开

始解五次方程给了一个明显的提示。即，对于五次方程

$$ax^5+bx^4+cx^3+dx^2+ex+f=0$$

代入$x=y-\frac{b}{5a}$，即得到缺项五次方程

$$y^5+my^3+ny^2+py+q=0$$

然后，寻找某些辅助变量，使之降为四次方程，而我们已经知道求四次方程根数解的方法。这一论证之所以特别引人注目，不仅因为它酷似成功地解三次方程和四次方程的方法，而且还因为，众所周知，任何五次（或任何奇次）多项式方程都必定至少有一个实数解。这是因为奇次方程的曲线看起来很像图 6.2 中所示五次方程的曲线。也就是说，这些曲线随我们沿 x 轴方向移动而不断升高，但当我们向相反方向移动时，则曲线不断下降。因此，这种函数必定在某些点上为正值，而且，必定在另外一些点上为负值。所以，利用一种称为介值定理的方法，我们可以说，这条连续曲线一定会在某一点上与 x 轴相交。在上述五次方程曲线图上，c 就是这样一点，因此，$x=c$ 就是方程 $x^5-4x^3-x^2+4x-2=0$ 的解。同样的道理，任何奇次多项式方程都（至少）有一个实数解。

然而，虽然介值定理表明了五次方程实数解的存在，但却不能明确地确定它们的值。因而，费拉里之后的代数学家们所努力寻求的就是这样一个解五次方程的标准公式。

但是，在这方面的所有努力都失败了。一个世纪过去了，又一个世纪过去了，仍然没有一个人能够求出五次方程的“根式解”。尽管后来的数学家们发现，可以将一般五次方程变换成这样一种形式

$$z^5+pz=q$$

如果我们称以前的方程为“缺项方程”的话，则这一个方程就应称作“完全缺项”方程。甚至就是这样一个高度简化了的五次方程，也同样无人能够攻克。这即使算不得难堪，至少令人沮丧。

1824 年，年青的挪威数学家尼尔斯·阿贝耳（1802—1829 年）发现，不可能用代数方法求出五次或更高次方程的“根式解”，他的发现使数学界为之震惊。总之，寻找五次方程根式解从一开始就注定了必然失败。我们可以在 D.E.史密斯的《数学史料集》中找到阿贝耳的证明，这一证明非常复杂，很难读懂，但它确实是数学史上的一座里程碑。

值得注意的是，阿贝耳的证明是模棱两可的。他并没有说，所有五次方程都是不可解的，因为我们显然可以解出像 $x^5-32=0$ 这样的方程，其解无疑是 $x=2$。并且，阿贝耳并没有否认我们可以有不同于加、减、乘、除和开方这些代数方法的方法解出五次方程。的确，一般五次方程能够用一种称为“椭圆函数”的方法解出，但这种方法比初等代数要复杂得多。而且，阿贝耳的证明也没有排除我们按照我们（或计算机）所要求的精度求出五次方程近似解的可能性。

阿贝耳的论文只是证明了不存在一种代数公式，可以只用原五次方程的系数作为方程解的可靠生成元。同样，解二次方程类似的二次公式和卡尔达诺解三次方程的公式也都不存在——不可能找到一种普遍有效的方法来确定五次方程的根式解。

这种情况不由使人联想起化圆为方的问题，在这两个问题上，数学家都受到了他们所用工具的局限。对于我们在第一章中所讲到过的化圆为方

问题，圆规和直尺显然无力完成这一重任。同样，“根式解”这一限制也阻碍了数学家寻求五次方程解的努力。我们所熟悉的代数算法没有能力驯服像五次方程这样的猛兽。

我们似乎已处于一种矛盾的边缘，虽然数学家们知道五次方程一定有解，但阿贝耳却又证明用代数方法不可能找到方程解。而正是“代数”这一修饰词使我们免于逸出这一边缘，跌入数学混乱之中。实际上，阿贝耳展示的正是代数这种非常明确的局限性，就在我们从四次方程转向五次方程的时候，这种局限性凭空出现了。

因此，实际上，我们绕了一个大圈，又回到了原处。卢卡·帕西奥利的悲观看法，虽然因 16 世纪的发现而遭人冷淡，但却不幸而言中。一旦我们越出四次方程的范围，代数便丧失了它的显赫。

第七章
艾萨克·牛顿的明珠
（17 世纪 60 年代后期）

英雄世纪的数学

如果说 16 世纪是数学活动迅速发展的时期，17 世纪就是震撼人心的革新和发现时期。17 世纪在数学史上称为英雄世纪，因为在这一多产的年代，有众多的知识巨人往来其间。

17 世纪，科学活动的中心从我们前一章所介绍的天才的意大利代数学家向北转向了法国、德国和英国思想家。当然，造成这种北移的原因是多方面的，除了人的努力外，还有纯粹的机遇问题。但是，对于这种现象，某些学者认为，一个重要的原因是欧洲北部学术气氛比较自由，恰与意大利教会的严厉限制形成了鲜明对照。伽利略的命运就是一个最著名的例子,一个科学家根据科学研究所得出的结论却被 17 世纪强有势力的罗马天主教宗教机构视为不可接受的洪水猛兽。伽利略遭监禁，被迫否认自己的观点使知识界甚为寒心。整个事件构成了科学史上最不光彩的一页。

虽然北方并非一切都很自由和开放，但宗教改革运动的影响却似乎有利于消除对科学研究的种种禁锢，从此才有开普勒、笛卡儿和牛顿脱颖而出。而很有可能，由于教会试图推行僵化的正统观念，意大利才沦为科学上的二等公民。

在 16 世纪与 17 世纪交汇之际，赢得繁荣发展的不仅仅是数学。1607 年，英国殖民詹姆斯敦，同时，欧洲人涌向新大陆。就在英国殖民詹姆斯敦之前几年，伽利略认真而巧妙地研究了落体运动规律，从而永远改变了物理学的性质。英国在詹姆斯敦建立殖民统治后两年，同一个伽利略又将发明不久的“小望远镜”指向天空，开创了现代天文学，同时也开始了他个人的苦难历程。当然，我们还不应忽略艺术的发展，1605 年，塞万提斯写出了不朽的名著《堂吉诃德》；1601 年，英国剧作家威廉·莎士比亚写出了《哈姆雷特》。

当然，文化的新纪元并不是以整整 100 年为间隔， 16 世纪末叶，数学革命的最初迹象便已出现。“英雄世纪”需要英雄，下面，我们将简要介绍其中的一些英雄。

16 世纪 90 年代，法国数学家弗朗索瓦·维埃特出版了他颇有影响的著作《分析术引论》（通常译作《分析术》）。我们在第四章中曾讲到过维埃特对π近似值的计算,而他 1591 年的这部著作则成为他的代表作。《分析术引论》对发展符号代数作出了很大贡献,成为高等数学的“奠基之作”。众所周知，维埃特的代数符号与现代符号相去甚远，对于习惯于现代数学的读者来说，维埃特的符号似乎显得过于繁冗，而且还附有过多的文字说明。例如，对于现代方程式 DR? DE＝A^2，维埃特则写成

D in R-D in E aequabitur A quad

尽管如此，但他的确朝着用字母表示方程的方向迈出了重要的一步。后来，又经过了几十年的改进与发展，代数符号体系终于在新的世纪中改革了数学的外观与实质。

17 世纪初叶，不列颠群岛的两位数学家约翰·纳皮尔（1550—1617 年）与亨利·布里格斯（1561—1631 年）共同引入、完善和开发了“对数”，这是一个具有重大实际意义和理论意义的概念。对数具有简化诸如乘、除和开方这些繁冗计算的非凡性质，以至此后任何头脑健全的科学家在计算像$\sqrt[7]{234.65}$的值时都会想到利用对数。下一个世纪的皮埃尔 - 西蒙·拉普拉斯评论说，纳皮尔和布里格斯的对数“通过简化计算，使天文学家的生命延长了一倍”。当然，布里格斯与纳皮尔的合作也是值得称道的，这与后来某些损害数学发展的激烈争吵与妒忌恰恰形成了鲜明的对照。

随着时代的发展，三位法国数学家引起了人们的注意。第一位是哲学家兼数学家勒内·笛卡儿（1596—1650 年），他 1637 年的著作《方法论》成为哲学史上的一座里程碑。这部关于“一般科学”的论著不但预示而且促进了成为时代特征的科学大爆炸。《方法论》中的哲学内容引起了人们的广泛讨论和热烈争辩，而其题为“几何学”的附录部分则最直接地影响了数学的发展。笛卡儿在此第一次将我们今天所谓的解析几何形诸笔墨。如同维埃特的代数符号一样，笛卡儿的解析几何与现代解析几何也相去甚远，但它毕竟宣告了代数与几何的结合，成为其后所有数学著作中不可或缺的内容。

在《方法论》问世的时候，布莱兹·帕斯卡（1623—1662 年）只是一个 14 岁的少年，却已出席了法国高级数学家的聚会。他已开始步入其虽然短暂，但却辉煌的数学生涯。帕斯卡是一个聪慧过人的孩子，是我们有时在数学史中见到过的那种神童。他在 16 岁时所撰写的数学论文就给数学巨匠笛卡儿留下了极深的印象，笛卡儿简直难以相信这篇论文的作者竟是如此年少的孩子。两年后，帕斯卡发明了第一架计算机，这就是我们现代计算机的始祖。并且，帕斯卡还对概率论作出了重大贡献，推动了概率论在一百年前卡尔达诺创立的基础上向前发展。

尽管帕斯卡显然具有数学天才，但他成年后的大部分时间却致力于神学研究，他的神学著作至今仍然是人们经常研究的课题。帕斯卡常常从他周围的事物中感觉到种种预兆，他认为在上帝对他的安排中没有包括数学，于是，他便完全放弃了数学。但是，他在 35 岁的时候，有一次，因牙疼难忍，便去思索数学问题以排遣，而疼痛竟然消失了。他觉得这是上天的启示，随即重操旧业，研究数学。虽然帕斯卡这次对数学的研究还不足一个星期，但他已发现了旋轮类曲线的基本性质（我们将在下一章讨论旋轮类曲线的问题）。此后，帕斯卡再次放弃了数学。1662 年，年仅 39 岁的帕斯卡与世长辞。

在三位法国数学家中，也许最值得注意的要数图卢兹的皮埃尔·德·费马（1601—1665 年），他统领了 17 世纪中叶的数学发展。费马在数学的许多领域中都享有盛名，并作出过重大发现。他独立于、甚至早于笛卡尔创立了自己的解析几何，而且，费马的方法在某些方面比他这位同时代的名人更“现代化”。当然，笛卡儿是第一位发表解析几何著作的数学家，并因而获得了崇高的荣誉，但是，费马的工作同样应当受到推崇。并且，帕斯卡与费马在 17 世纪 50 年代的书信往来还奠定了我们前面所讲到过的概率论的基础。除此以外，费马还在我们今天称之为微分学的发展上作出过重大贡献。在一些地方，特别是在法国，人们有时认为他是微积分的共同创立者之一，而大部分数学史家虽然承认费马的巨大成就，但却认为这

种看法未免失之偏颇。

然而，在数论领域，费马留下了他不可磨灭的足迹。我们在欧几里得的《原本》第 7 篇至第 9 篇中曾见到过这个论题。有关数论的一部古代名著是丢番图的《算术》（约公元 250 年？）。在文艺复兴时期，这部著作被重新发现，并翻译成多种文字，证明是一部非常有影响的论文。费马得到了一本丢番图的著作，并深深地沉溺于其中，不久便在有关整数性质方面作出了他自己的惊人发现。

费马常常提出一些诱人的命题，有时又宣称已得出了确凿的证明，但又很少将这些证明写下来。因此，后代数学家（时常是欧拉）就不得不去补上这些欠缺的证明。结果，数学史学家在确定荣誉究竟应该归于谁时，常常感到左右为难——归于费马，是他第一个阐述了这些命题，而且也可能作出过证明；或者，应归于欧拉，因为事实上毕竟是他写下了这些论证。

显然，费马的大部分“定理”（我们颇费踌躇地使用“定理”一词，因为他的许多命题都过分自信，但缺少证明）都是受到丢番图著作的启发而提出的。费马在丢番图那本《算术》中命题Ⅱ.8 的书页边上写下了一条批语。命题Ⅱ.8 提出，一个整数平方可分解为另外两个整数平方之和，例如，■52＝32＋42 或 252＝72＋242。在丢番图这一定理旁边，费马写下了他著名的批语：

“但是，不可能将一个三次方数分解为两个三次方数之和，或将一个四次方数分解为两个四次方数之和。总之，高于二次方的任何次乘幂都不可能分解为两个同样次幂之和；对此，我已发现了极巧妙的证明，但页边空白太小，写不下了。”

用现代话说，他的批语表明，我们不能找到整数 a、b、c 和指数 $n\geq3$，并使 $a^n+b^n=c^n$。如果他的论点是正确的话，那么，一个整数平方分解为两个整数平方之和就完全是一种侥幸；费马说，除了平方以外，任何次幂的整数都不能写成两个较小整数的同次幂之和。

像往常一样，费马没有留下证明。他把其证明缺漏的原因仅仅归结于丢番图书页空白的狭小。费马似乎在说，只要有一张白纸，他会很高兴为他的发现作出精彩的证明。而实际上，就像他的大部分命题一样，他把寻求证明的重任留给了后人。

对于费马的这一论断，后人依然在寻求证明，因为他的论断至今依然未能解决。甚至连曾解开过许多费马“定理”之谜的欧拉，对他这一论断也只证明出 $n=3$ 和 $n=4$。也就是说，欧拉证明，一个三次方数的确不能写成两个三次方数之和，或者，一个四次方数也同样不能分解为两个四次方数之和。但是，就人们普遍称之为“费马大定理”的一般情况而言，问题仍然悬而未决。如同费马没有给出证明的其他许多命题一样，他的这一命题很可能也是正确的。尽管如此，迄今尚无一位数论学家证明这个命题；同样，也没有任何人提出反例，否定这个命题。所以，在这个意义上说，称其为费马的大“定理”，确实有些草率。即使在 20 世纪末叶，人们对这个问题的兴趣越来越高，如果能有人攻克这道难题，他肯定会在今后的数学史上留下光辉的一页。

现在，倘若我们能够回到 1661 年夏季，检点 17 世纪的数学遗产，我们将会注意到许多重要事情。代数符号、对数、解析几何、概率和数论——所有这些都已初具规模，而维埃特、纳皮尔、笛卡儿、帕斯卡和费马这

些名字将受到应有的尊崇。他们的确是英雄。当然，在 1661 年夏天，丝毫没有人注意到一个正在悄悄开始的数学旅程，这一旅程很快将使所有这些伟人黯然失色。这一数学旅程开始于美丽的剑桥大学三一学院。1661 年夏，来自学院附近乌尔索普的一位少年开始了他的大学生涯。他已经显露出他的才华，而与他一起进入三一学院读书的十几位同学，虽然同样无声无臭，却也同样才华横溢。然而，这位年青人日后将成为英雄世纪的最伟大的英雄，并同时永远改变了人类观察世界的方法。他的名字当然就是艾萨克·牛顿。

解放了的头脑

1642 年的圣诞节，一个早产儿危险地降生了，这就是牛顿，他瘦小得简直可以放进“一夸脱容量的杯子”中。而更为不幸的是，他的父亲已于 10 月初故去，只撇下母亲一人独自抚养这羸弱的婴儿。但是，他却终于绝处逢生，并顺利地度过了林肯郡严寒的冬天，最后，艾萨克竟活到了 84 岁高龄。

身体得到复原，苦难却仍未结束。牛顿三岁的时候，他的母亲汉纳·艾斯库·牛顿嫁给了邻村一个 63 岁的教长巴纳巴斯·史密斯。史密斯虽然急切地希望娶一个年青的妻子，但却不愿接受一个三岁的孩子。所以，牛顿的母亲再婚之后，小艾萨克就被留下来与他的祖母一起生活。骨肉分离使小牛顿感到万分痛苦。母亲就住在附近，这对他无疑是一种残酷的折磨，因为他只要爬到树上，就可以眺望田野对面村庄中教堂的尖顶，他的母亲和继父就住在那座教堂里。艾萨克从来没见过父亲，现在又失去了母亲，他的痛苦不是由于疾病，而是由于亲情的冷漠。我们将看到，牛顿成人后变得有些神经过敏和愤世嫉俗，因为他很少感受到人类友情的温暖。完全可以认为，他的这种性格是因为遭受亲人遗弃而造成的。

艾萨克长大后，进入了一所当时很不错的中学读书，也就是说，它主要是教授拉丁语和希腊语。课下，牛顿很少与人来往，他大部分课余时间都用来读书和制做各种精巧的小器械。传说他曾做过一个由小老鼠在踏车上驱动的小风车；还做过日晷，并将它们放在住处周围的各个主要方位上；他也曾将一个点燃的灯笼系在风筝上，高高放入春天的夜空中，想必曾使平静的英国村民们感到大为恐惧。这些活动显示了一个异常灵巧的年青人的智慧，他可不想只顾埋头于拉丁语复杂的动词变位中。这些活动还预示了一位天才实验物理学家的出现，他的实验小发明对他后来理论的发展具有无可估量的意义。

1661 年夏，艾萨克·牛顿离开家乡，去剑桥大学三一学院求学。当时，卡姆河畔这座平静的小镇作为高等教育中心已有 400 年的历史，是一个声誉卓著的古老学府，牛顿在这里有了用武之地。17 世纪初叶，随着英格兰清教主义和宗教改革运动的兴起，剑桥大学得到了蓬勃的发展。剑桥大学有许多值得骄傲的事情，从詹姆士王钦定本英文《圣经》、国王学院小教堂的建筑杰作，到清教革命的领袖奥利弗·克伦威尔，他出生于附近的亨廷顿，1617 年前就读于西德尼·萨赛克斯学院。

但当牛顿进入剑桥大学时，剑桥大学已失去了往昔的荣耀。其原因与英国历史的兴衰变迁密切相关。1642 年，也就是牛顿出生的那一年，在克

伦威尔领导下的清教徒胜利结束了他们与君主制的长期斗争。克伦威尔亲自主政，1649 年，国王查理一世在伦敦白厅被处死后，克伦威尔政府成为不容置疑的权威。其时，清教的剑桥大学正处于鼎盛之际，而保皇党的大本营牛津大学则相形见绌。

然而，好景不长。清教徒的共和国并不比被推翻的君主制好多少，也许还更糟。1658 年，克伦威尔死后，没有一个清教领导人能够填补这一空缺，英国民情汹汹，要求恢复君主制。因此，1660 年，断头国王的儿子查理二世登上王位，这段时期，历史上称作王政复辟时期。无庸赘言，局势发生了根本的变化。剑桥大学自然成了新当权的保皇党怀疑和敌视的目标。王政复辟的第二年，牛顿进入剑桥大学，而这时的剑桥大学充斥着政治阴谋，成了庸才的庇护所，到处死气沉沉，完全不是一个理想的学习场所。

我们今天尊崇剑桥大学为少数几个真正的教育中心之一，但我们很难想象 17 世纪 60 年代剑桥大学衰败的情形。那时，学校任命教授，完全是出于政治或教会的原因，其中有许多教授，完全与学术无关。据记载，甚至有人 50 年中竟然没有教过一个学生，没有写过一本书，或没有讲过一次课！实际上，有些教师根本不住在剑桥一带，他们只是偶尔来此一游。

教授对学术尚且如此冷漠，学生自然也就不求进取。表面上，剑桥大学维持了学术生活的虚假繁荣，为好学的青年人开设了大量人文科课程。但实际上，剑桥大学的学生更多地热中于到遍布校园的小酒店里开怀畅饮一类事情。学生乃至教授当然可以毫不费力地在剑桥大学中混日子。

起初，艾萨克 · 牛顿慕名而来，对学校寄予了很高的期望。他开始学习规定的拉丁文学和亚里士多德哲学课程，但他逐渐放弃了这类学业，或者是因为他感到老师无能，或者是因为他意识到这些课程的迂腐和无用，也或者只是因为显然没有任何人真正关心他的学习情况。

他在三一学院的同学们可能也有同感，他们晚上纷纷跑到小酒店去纵酒狂欢，而牛顿却与众不同。他贪婪地博览群书。人们常常看到他一边散步，一边沉思。当牛顿的注意力被一个想法所吸引时，他能以异于常人的专心，废寝忘食地进行研究，尤其是对一个特别有趣的难题。牛顿初到剑桥大学的时候，还表现出一种老式的负罪感，他有一个笔记本，里面记录了他的各式各样的罪孽，从他不经常祈祷，在教堂做礼拜时漫不经心，到他“ 不洁的思想、语言、行为和梦境 ”。诚然，清教主义的思想对他影响很大，但是，人们也会想到，生活的孤独也必定会在很大程度上对一个性格内向的青年人产生深刻的影响。

如果一时没有罪孽可以记录，这一永远好奇的学生便忙着对光、颜色和视觉的性质做各种实验。例如，他曾长时间地凝视太阳，然后，详细地记录他视觉中所出现的斑点和闪光，这个实验影响他的视力长达几天之久；实际上，他不得不将自己关在暗室中，让眼中的影象慢慢消退。又有一次，他对眼球的形状如何扭曲和改变形象感到好奇，便以自己为对象设计了一个十分可怕的实验。据牛顿记载，他用一根小棍，或“ 粗针 ”，

“ 在我的眼睛与眼骨之间扎，并尽可能地扎到眼球的后部，然后用粗针的顶端压迫眼球……于是便出现了许多白的、黑的和彩色的光环，当我用粗针头继续在眼睛上摩擦的时候，这些光环便显得分外清晰…… ”

牛顿亲手画了一张图来说明这个可怕的实验，他画出了用小棍在他扭曲了

的眼球下部和后部摩擦的情形，并用从 a 到 g 的字母一一标明。显然，这可不是一位普通的大学生。

王政复辟时期的剑桥大学，虽然有种种弊端，但它拥有一个很大的图书馆，对于这个充满好奇心的一流学生来说，这确是一个非常必要的知识宝库。说到书，这里还有一段故事。1663 年，牛顿在斯特布里奇集市上碰到一本关于占星术的书。为了弄懂书中的几何图，他决定阅读欧几里得的《原本》。有趣的是，他初次阅读，就发现这本古代教科书中充满了无关紧要和不证自明的定理（顺便说一句，成年后的牛顿抛弃了这种观点）。

牛顿读书，有一个特点就是他不满足于只读希腊的经典著作。他还花费了很大气力，阅读笛儿尔的几何学。他后来回忆说，在他开始阅读这部著作的时候，刚刚读过几页，就被完全难住了。然后，他再翻回第 1 页，重新读一遍，这一次会有所进展，但继续下去又会感到难以理解，这样，他就再翻回来重读。如此循环往复。他就这样，一点儿一点儿地独自啃完了这部《几何学》，没有任何导师或教授帮助。当然，考虑到教师庸庸碌碌，所规定的课程又厚古薄今，他也很难找到任何可以帮助他的人。

然而，在剑桥大学教授中，毕竟有一位教授堪当此任，他就是卢卡斯讲座数学教授艾萨克 · 巴罗（1630—1677 年）。虽然在现代意义上，巴罗算不上牛顿的老师，但他无疑曾与这位初露头角的学者有过接触，并曾指导过牛顿阅读当代主要的数学著作。通过不断的阅读与思考，牛顿在普通的科学与数学背景下一跃掌握了当代大部分的新发现。牛顿既已进入前沿，便开始向未开垦的领域进军。

1664 年，牛顿荣获三一学院奖学金，为他硕士学位的学习赢得了四年的经济资助。他有了更多的自由去探索自己感兴趣的问题。这种自由加上他通过博览群书打下的坚实基础，将解放一个历史上最伟大的天才。从此，牛顿开始着手解决摆在他面前的问题，其精神之专注，简直令人难以置信。20 世纪剑桥大学著名的经济学家约翰 · 梅纳德 · 凯恩斯曾对牛顿的能力做过如下评价：

“ 他的非凡天才在于他能够长时间地连续思考一个纯智力问题，直至解决……任何研究过纯科学或纯哲学问题的人都知道，一个人只可能短时间地集中思考一个问题，并且，集中全部精力思考，但过不久，注意力就会逐渐分散和转移，你会发现，你的思想成为一片空白。但我相信，牛顿能够连续几小时、几天和几星期地集中思考一个问题，直到解开其中的奥秘为止。 ”

牛顿对他如何解决难题做了同样的说明，只不过更加简洁，就是 “ 通过持续不断的思考 ” 。

其后几年，牛顿带着对新发现的极度兴奋，更加勤奋地工作。人们常常看到他在微弱的烛光下一直工作到深夜。据说，他的猫因为常常饱餐牛顿碰也没碰一下的饭菜，竟然长得十分肥胖。这位年轻人认为错过吃饭、耽误睡觉与取得的巨大进展相比，实在是微不足道的。

这两年，也许是任何思想家，当然是任何一位 23 岁的思想家可能有过的最多产的两年。他的成功，一部分是在剑桥大学，还有一部分是在他的家乡乌尔索普取得的，因为爆发了可怕的瘟疫，学校被迫关闭。1665 年初，他发现了我们现在所称的 “ 广义二项式定理 ” ，并成为他以后数学著作中

的重要部分。不久后，他提出了“流数法”（即我们今天所称的微分学）。1666 年，他发明了“逆流数法”（即积分学）。在这期间，他还创造性地提出了他的颜色理论。但据牛顿回忆，他还有更多的发现：

“……同一年，我开始思考重力与月球运行轨道的问题……我推算出保持星体绕其轨道运动的引力一定与它们球心距的平方成反比；比较保持月球绕轨道运动的引力与地球表面的重力，发现二者的答案非常接近。”

50 年后，年迈的牛顿所做的这些回忆准确地阐述了万有引力理论的雏形，这一理论远胜于牛顿其他任何成就，为他赢得了崇高的科学声望。面对这些发现，他以一种非常坦率而冷漠的笔触写道：

“所有这些发现都是在 1665—1666 这两年瘟疫期间做出的。因为这两年是我发现力最盛时期，我对数学和哲学的研究比其它任何时期都要多。”

因此，这两年瘟疫期间称为牛顿的“高峰年”，情况确实如此。据说，他所有的理论都是在这段时间内形成、完善和成熟的。这不免有点儿夸大其辞，因为在这之后的年月里，牛顿仍在继续推敲和改进这些理论。然而，牛顿在这短暂的两年里所表现出来的创造力不仅规定和指导了他自己一生的研究方向，而且在很大程度上规定和指导了科学的未来。

今天，人们很容易忘记牛顿做出这些非凡的发现时只是剑桥大学的一个无名之辈。R.S.韦斯特福尔也许是我们今天最出色的牛顿传记作家，他对这一明显的事实做了如下的精彩记载：

“（牛顿的成功）已经显示出一代宗师的风范，足以使欧洲所有的数学家由衷地羡慕、妒忌和敬畏。但实际上，欧洲只有一位数学家，即艾萨克·巴罗知道牛顿的存在，据说，1666 年，巴罗对牛顿的成就也仅仅略知一二。但牛顿的不为人知，并不影响这一事实，即这位不足 24 岁的青年人，虽然没有受过正规教育，却已成为欧洲最出色的数学家。真正举足轻重的人物，也就是牛顿自己，非常清楚自己的地位。他曾研究过诸位大师。他知道，他们各自都有其局限性。而他自己，却已远远地超过了他们所有人。”

纵观历史，我们已看到，数学的中心不断地从一个地方转移到另一个地方，从毕达哥拉斯学派所在的克罗托内先后转移到柏拉图的雅典学园、亚历山大、巴格达，然后又转移到文艺复兴时期卡尔达诺和费拉里所在的意大利。然而，令人难以相信的是，17 世纪 60 年代中期，数学中心又转移到了三一学院一个学生简朴的房间里，而此后，不论牛顿住在哪里，哪里就是世界的数学中心。

牛顿二项式定理

对于牛顿非凡的发现，我们在此只能略窥一斑。我们首先介绍牛顿的第一大数学发现——二项式定理。虽然按照欧几里得或阿基米德的概念来说，这不是一条“定理”，因为牛顿没有提供完整的证明。但是，他的见识和直觉足以使他发明出这一恰当而准确的公式，并且，我们将看到，他是如何以一种最奇妙的方式应用这一公式的。

二项式定理论述了$(a+b)^n$的展开式。人们只要有初步的代数知识和足够的毅力，便可以得到如下公式，

$(a+b)^2=a^2+2ab+b^2$

$(a+b)^3=a^3+3a^2b+3ab^2+b^3$

$(a+b)^4=a^4+4a^3b+6a^2b^2+4ab^3+b^4$

等等。对于$(a+b)^{12}$，人们显然希望不必经由(a+b)十几次自乘的冗长计算，就能够发现其展开式中a^7b^5的系数。早在牛顿出生之前很久，人们便已提出并解决了二项式的展开式问题。中国数学家杨辉早在13世纪就发现了二项式的秘密，但他的著作直到近代才为欧洲人所知。维埃特在其《分析术引论》前言的命题XI中也同样论证了二项式问题。但这一伟大发现通常是以布莱兹·帕斯卡的名字命名的。帕斯卡注意到，二项式的系数可以很容易地从我们现在称为"帕斯卡三角"[①]的排列中得到：

$$\begin{array}{ccccccccccccccc}
&&&&&&&1&&&&&&& \\
&&&&&&1&&1&&&&&& \\
&&&&&1&&2&&1&&&&& \\
&&&&1&&3&&3&&1&&&& \\
&&&1&&4&&6&&4&&1&&& \\
&&1&&5&&10&&10&&5&&1&& \\
&1&&6&&15&&20&&15&&6&&1& \\
1&&7&&21&&35&&35&&21&&7&&1
\end{array}$$

等等

在这个三角形中,每一个新增数字都等于其上左右两个数字之和。因此,根据帕斯卡三角,下一行的数值为

1 8 28 56 70 56 28 8 1

例如，表值56就等于其上左右两个数字21+35之和。

帕斯卡三角与$(a+b)^8$展开式之间的联系是非常直接的，因为三角形的最后一行数值为我们提供了必要的系数，即

$$(a+b)^8=a^8+8a^7b+28a^6b^2+56a^5b^3$$
$$+70a^4b^4+56a^3b^5+28a^2b^6+8ab^7+b^8$$

我们只要将三角形的数值再向下延伸几行，就可以得到$(a+b)^{12}$展开式中a^7b^5的系数为792。所以，帕斯卡三角的实用性是非常明显的。

年轻的牛顿经过对二项展开式的研究，发明了一个能够直接导出二项式系数的公式，而不必再繁琐地延伸三角形到所需要的那行了。并且，他对模式的持续性的固有信念使他认为，能够正确推导出诸如$(a+b)^2$或$(a+b)^3$这类二项式系数的公式，也应该适用于像$(a+b)^{\frac{1}{2}}$或$(a+b)^{-3}$这种形式的二项式。

关于分数指数和负数指数问题，在此还需多说一句。我们知道，在初等代数中，$a^{\frac{1}{n}}=\sqrt[n]{a}$，而$a^{-n}=\frac{1}{a^n}$。虽然牛顿可能不是认识到这些关系的第一人，但他在$\sqrt{1+x}$和$\frac{1}{(1-x^2)}$一类二项式展开式中，确实充分利用了这些关系。

以下所列牛顿的二项展开式公式是他在1676年写给其同时代伟人戈特弗里德·威廉·莱布尼兹的一封信中阐明的(此信经由皇家学会的亨利·奥

[①] 这个三角形早在我国南宋数学家杨辉1261年所著的《详解九章算法》一书里就已出现，因此，在我国，我们称它为杨辉三角。——译者注。

尔登伯格转交）。牛顿写道：

$$(P+PQ)^{\frac{m}{n}}=P^{\frac{m}{n}}+\frac{m}{n}AQ+\frac{m-n}{2n}BQ$$

$$+\frac{m-2n}{3n}CQ+\frac{m-3n}{4n}DQ+\cdots\cdots$$

公式中的$P+PQ$是所讨论的二项式；$\frac{m}{n}$是指数，对此，我们将提出二项式的“指数是整数还是（比如说）分数，是正数还是负数”的问题。公式中的A、B、C等表示展开式中该字母所在项的前一项。

对于那些见过现代形式的二项展开式的读者来说，牛顿的公式可能显得过于复杂和陌生。但只要仔细研究一下，就可以解决读者的任何疑问。我们首先来看，

$$A=P^{\frac{m}{n}}$$

$$B=\frac{m}{n}AQ=\frac{m}{n}P^{\frac{m}{n}}Q$$

$$C=\frac{m-n}{2n}BQ=\frac{(m-n)m}{(2n)n}P^{\frac{m}{n}}Q^2=\frac{\left(\frac{m}{n}\right)\left(\frac{m}{n}-1\right)}{2}P^{\frac{m}{n}}Q^2$$

$$D=\frac{m-2n}{3n}CQ=\frac{\left(\frac{m}{n}\right)\left(\frac{m}{n}-1\right)\left(\frac{m}{n}-2\right)}{3\times 2}P\frac{m}{n}Q^3$$ 等等

然后，应用牛顿的公式，并从方程两边分别提取公因数$P^{\frac{m}{n}}$，我们得出

$$P^{\frac{m}{n}}(1+Q)^{\frac{m}{n}}=(P+PQ)^{\frac{m}{n}}=P^{\frac{m}{n}}\Big[1+\frac{m}{n}Q$$

$$+\frac{\left(\frac{m}{n}\right)\left(\frac{m}{n}-1\right)}{2}Q^2+\frac{\left(\frac{m}{n}\right)\left(\frac{m}{n}-1\right)\left(\frac{m}{n}-2\right)}{3\times 2}Q^3+\cdots\cdots\Big]$$

消去$P^{\frac{m}{n}}$，得

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

$$+\frac{\left(\frac{m}{n}\right)\left(\frac{m}{n}-1\right)\left(\frac{m}{n}-2\right)}{3\times 2}Q^3+\cdots\cdots$$

也许，这种形式看起来就比较熟悉了。

我们不妨应用牛顿的公式来解一些具体例题。例如，在展开$(1+x)^3$时，我们用x替换Q，用3替换$\frac{m}{n}$，于是得到

$(1+x)^3$

$$=1+3x+\frac{3\times2}{2}x^2+\frac{3\times2\times1}{3\times2}x^3+\frac{3\times2\times1\times0}{4\times3\times2}x^4+\cdots\cdots$$

$$=1+3x+\frac{6}{2}x^2+\frac{6}{6}x^3+\frac{0}{24}x^4+\frac{0}{120}x^5+\cdots\cdots$$

$$=1+3x+3x^2+x^3$$

这恰恰就是帕斯卡三角的排列系数；并且，由于我们的原指数是正整数 3，所以，展开式到第四项结束。

但是，当指数是负数时，又有一个完全不同的情况摆在牛顿面前。例如，展开 $(1+x)^{-3}$，根据牛顿公式，我们得到

$$1+(-3)x+\frac{(-3)(-4)}{2}x^2+\frac{(-3)(-4)(-5)}{6}x^3+\cdots\cdots$$

或简化为

$(1+x)^{-3}=1-3x+6x^2-10x^3+15x^4-\cdots\cdots$

方程右边永远没有终止。应用负指数定义，这一方程就成为

$\frac{1}{(1+x)^3}=1-3x+6x^2-10x^3+15x^4-\cdots\cdots$**或其等价方程**

$$\frac{1}{1+3x+3x^2+x^3}=1-3x+6x^2-10x^3+15x^4-\cdots\cdots$$

牛顿将上式交叉相乘并消去同类项，证实

$(1+3x+3x^2+x^3)(1-3x+6x^2-10x^3+15x^4-\cdots\cdots)=1$

而当牛顿展开像$\sqrt{1-x}=(1-x)^{\frac{1}{2}}$这种形式的二项式时，问题变得更加奇特。在这一例子中，$Q=-x$，$\frac{m}{n}=\frac{1}{2}$，于是，我们得到

$$\sqrt{1-x}=1+\frac{1}{2}(-x)+\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)}{2}(-x)^2$$

$$+\frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{6}(-x)^3+\cdots\cdots$$

$$=1-\frac{1}{2}-\frac{1}{8}x^2-\frac{1}{16}x^3-\frac{5}{128}x^4-\frac{7}{256}x^5-\cdots\cdots \quad (*)$$

牛顿用等式右边的无穷级数自乘，也就是求这无穷级数的平方，以检验这一貌似奇特的公式，其结果如下：

$$\left(1-\frac{1}{2}x-\frac{1}{8}x^2-\frac{1}{16}x^3-\frac{5}{128}x^4-\cdots\cdots\right)$$

$$\left(1-\frac{1}{2}x-\frac{1}{8}x^2-\frac{1}{16}x^3-\frac{5}{128}x^4-\cdots\cdots\right)$$

$$=1-\frac{1}{2}x-\frac{1}{2}x-\frac{1}{8}x^2+\frac{1}{4}x^2-\frac{1}{8}x^2-\frac{1}{16}x^3+\frac{1}{16}x^3+\frac{1}{16}x^3-\frac{1}{16}x^3-\cdots\cdots$$

$=1-x+0x^2+0x^3+0x^4+\cdots\cdots$

$=1-x$

所以

$$\left(1-\frac{1}{2}-\frac{1}{8}x^2-\frac{1}{16}x^3-\frac{5}{128}x^4-\cdots\cdots\right)^2=1-x$$

这就证实了

$$1-\frac{1}{2}x-\frac{1}{8}x^2-\frac{1}{16}x^3-\frac{5}{128}x^4-\cdots=\sqrt{1-x}$$

与牛顿原推导结果相同。

牛顿写道：“用这一定理进行开方运算非常简便。”例如，假设我们求$\sqrt{7}$的小数近似值。首先，我们看到，

$$7=9\left(\frac{7}{9}\right)=9\left(1-\frac{2}{9}\right)$$

所以 $\sqrt{7}=\sqrt{9\left(1-\frac{2}{9}\right)}=3\sqrt{1-\frac{2}{9}}$

现在，将等式右边的平方根代入前面标有（*）符号的二项展开式中的前 6 项，当然，此处要用$\frac{2}{9}$替换原公式中的x，因而，我们得到

$$\sqrt{7}\approx 3\left(1-\frac{1}{9}-\frac{1}{162}-\frac{1}{1458}-\frac{5}{52488}-\frac{7}{472392}\right)=2.64576\cdots\cdots$$

这一结果与$\sqrt{7}$的真值仅相差0.00001，这当然是非常精确的，因为我们只取了前 6 个常数项。如果我们取二项展开式中更多的项，我们就会得到更加精确的近似值。并且，我们还可以用同样的方法求出三次根、四次根，等等，因为我们可以应用二项式定理展开$\sqrt[3]{1-x}=(1-x)^{\frac{1}{3}}$，然后按照上述方法继续演算。

在某种意义上说，用6个分数的和可以求出$\sqrt{7}$的近似值并没有什么特别奇怪的。而真正令人吃惊的是，牛顿的二项式定理精确地告诉我们应该采用哪些分数，而这些分数则是以一种完全机械的方式得出的，无须任何特殊的见解与机巧。这显然是一个求任何次方根的有效而巧妙的方法。

二项式定理是我们即将讨论的伟大定理的两个必要前提之一。另一个前提是牛顿的逆流数，也就是我们今天所说的积分。但是，对逆流数的详尽说明属于微积分问题，超出了本书的范围。然而，我们可以用牛顿的话来阐述其重要定理，并举一两个例子来加以说明。

牛顿在 1669 年中撰著的《运用无穷多项方程的分析学》一书中提出了逆流数问题，但这部论著直到 1711 年才发表。这是牛顿第一次提出逆流数问题，他将他的这部论文交给几个数学同事传阅。比如，我们知道，艾萨克·巴罗就曾看到过这部论文，他在 1669 年 7 月 20 日给他一个熟人的信里写道：“……我的一个朋友…… 在这些问题上很有天分，他曾带给我几篇论文。”巴罗或《分析学》一书的任何其他读者遇到的第一个法则如下。

设任意曲线 AD 的底边为 AB，其垂直纵边为 BD，设 AB=x，

牛顿求曲线下面积法则，摘自《运用无穷多项方程的分析学》1745 年译本

BD=y，并设 a、b、c 等为已知量，m 和 n 为整数。则：

法则1：如果$ax^{\frac{m}{n}}=y$，那么$\frac{an}{m+n}x^{\frac{(m+n)}{n}}=$面积ABD

在图7.1中，牛顿所求的是横轴之上，曲线$y=ax^{\frac{m}{n}}$之下，右边到x点之内的图形的面积。根据牛顿法则，这一图形的面积为$\frac{an}{m+n}x^{\frac{(m+n)}{n}}$。例如，如果我们取直线$y=x$（图7.2），则$a=m=n=1$，按照牛顿公式，面积为$\frac{1}{2}x^2$，对这一结果，可以很容易地用三角形面积公式$=\frac{1}{2}$（底）×（高）进行检验。同样，在原点与x点之间，$y=x^2$之下图形的面积为$\frac{x^{2+1}}{(2+1)}=\frac{x^3}{3}$。

牛顿又进一步说明了《分析学》一书的法则2，“如果y值是由几项之和组成的，那么，其面积也同样等于每一项面积之和。”例如，他写道，曲线$y=x^2+x^{\frac{3}{2}}$下图形的面积为

$$\frac{1}{3}x^3+\frac{2}{5}x^{\frac{5}{2}}$$

那么，牛顿所采用的两个工具就是：二项式定理和求一定曲线下面积的流数法。他运用这两个工具，可以得心应手地解决许多复杂的数学与物理问题，而我们将要看到的是牛顿如何应用这两个工具，使一个古老的问题获得了全新的生命：计算π的近似值。我们在第四章的后记中，追溯了这一著名数字的某些历史，确认了某些学者，如阿基米德、韦达和卢道尔夫·冯瑟伦在计算更精确的π近似值方面所作出的贡献。1670年左右，这个问题引起了艾萨克·牛顿的注意。他运用他奇妙的新方法，对这一古老问题进行研究，并取得了辉煌的成就。

伟大的定理：牛顿的π近似值

牛顿当然精通解析几何的概念，他用解析几何的方法研究π近似值问题。他首先作半圆，其圆心位于C点$(\frac{1}{2}, 0)$，其半径$r=\frac{1}{2}$，如图7.3所示。他知道这个圆的方程是

$$\left(x-\frac{1}{2}\right)^2+(y-0)^2=\left(\frac{1}{2}\right)^2 \quad \text{或} \quad x^2-x+\frac{1}{4}+y^2=\frac{1}{4}$$

经化简并求解y得到上半圆方程为

$$y=\sqrt{x-x^2}=\sqrt{x}\sqrt{1-x}=x^{\frac{1}{2}}(1-x)^{\frac{1}{2}}$$

（他为什么选择这样一个半圆也许完全是个谜，但其效用在论证结束时自会明了。）

前面带（*）标记的方程表明，$(1-x)^{\frac{1}{2}}$可以用其二项展开式替换，因此，半圆的方程可演变为

$$y = x^{\frac{1}{2}}(1-x)^{\frac{1}{2}}$$
$$= x^{\frac{1}{2}}(1-\frac{1}{2}x-\frac{1}{8}x^2-\frac{1}{16}x^3-\frac{5}{128}x^4-\frac{7}{256}x^5-\cdots\cdots)$$
$$= x^{\frac{1}{2}}-\frac{1}{2}x^{\frac{3}{2}}-\frac{1}{8}x^{\frac{5}{2}}-\frac{1}{16}x^{\frac{7}{2}}-\frac{5}{128}x^{\frac{9}{2}}-\frac{7}{256}x^{\frac{11}{2}}-\cdots\cdots$$

下面，艾萨克·牛顿的天才就明显地表现出来。他设B点于$(\frac{1}{4}, 0)$，如图 7.3 所示。并作 BD 垂直于半圆的直径 AE。然后，他用两种完全不同的方法，求阴影部分 ABD 的面积：

1.用流数法求面积（ABD） 我们已经看到，牛顿知道如何求一条曲线下起点为O，右端点为$x=\frac{1}{4}$的图形的面积。根据《分析》一书的法则 1 和法则 2，阴影部分的面积为

$$\frac{2}{3}x^{\frac{3}{2}}-\frac{1}{2}\left(\frac{2}{5}x^{\frac{5}{2}}\right)-\frac{1}{8}\left(\frac{2}{7}x^{\frac{7}{2}}\right)-\frac{1}{16}\left(\frac{2}{9}x^{\frac{9}{2}}\right)-\cdots\cdots$$
$$=\frac{2}{3}x^{\frac{3}{2}}-\frac{1}{5}x^{\frac{5}{2}}-\frac{1}{28}x^{\frac{7}{2}}-\frac{1}{72}x^{\frac{9}{2}}-\frac{5}{704}x^{\frac{11}{2}}-\cdots\cdots \qquad (**)$$

赋值$x=\frac{1}{4}$。这种方法的天才之处就在于，当我们赋值计算的时候，其方程式就会变得极为简单，因为

$$\left(\frac{1}{4}\right)^{\frac{3}{2}}=\left(\sqrt{\frac{1}{4}}\right)^3=\frac{1}{8}，\left(\frac{1}{4}\right)^{\frac{5}{2}}=\left(\sqrt{\frac{1}{4}}\right)^5=\frac{1}{32}，等等$$

所以，我们只要应用（**）方程式中的前 9 项级数，就可以计算出阴影部分（ABD）面积的近似值，得

$$\frac{1}{12}-\frac{1}{160}-\frac{1}{3584}-\frac{1}{36864}-\frac{5}{1441792}\cdots\cdots$$
$$-\frac{429}{163208757248}=0.07677310678$$

2.用几何方法求面积（ABD） 牛顿接着用纯几何方法验算阴影部分的面积。他首先求直角三角形△DBC的面积。我们知道，BC的长度为$\frac{1}{4}$，而CD是半径，其长度$r=\frac{1}{2}$。我们可以直接应用勾股定理，得出

$$\overline{BD}=\sqrt{\left(\frac{1}{2}\right)^2-\left(\frac{1}{4}\right)^2}=\sqrt{\frac{3}{16}}=\frac{\sqrt{3}}{4}$$

所以，

面积（ΔDBC）$=\frac{1}{2}$（$\overline{BC}$）×（$\overline{BD}$）$=\frac{1}{2}\left(\frac{1}{4}\right)\left(\frac{\sqrt{3}}{4}\right)=\frac{\sqrt{3}}{32}$

到目前为止，一切顺利。下一步，牛顿要求出楔形或扇形部分 ACD 的面积。为此，他再次利用△DBC。由于 BC 的长度恰好是斜边 CD 的一半，他认为，这就是我们所熟悉的 30°-60°-90°直角三

角形；特别是，∠BCD 是 60°角。

至此，我们再一次为他深邃的见识所折服，因为如果他在 B 点以外的其它地方作垂线，那么，在他最需要的时候，就不会恰好形成 60°角。现在，已知扇形的角度为 60°，也就是说，等于构成半圆的 180°角的三分之一，牛顿就可以断定，扇形的面积也等于半圆面积的三分之一。即

$$\text{面积（扇形）}=\frac{1}{3}\text{面积（半圆）}$$

$$=\frac{1}{3}\left(\frac{1}{2}\pi r^2\right)=\frac{1}{3}\left[\frac{1}{2}\pi\left(\frac{1}{2}\right)^2\right]=\frac{\pi}{24}$$

机敏的读者会想起，这一伟大定理是牛顿的π近似值，他们会很着急，不知道这个常数何时和怎样才能进入他的论证。终于，π在牛顿的推理链中出现了，现在只剩下最后一两步，就可以巧妙地计算出他的π近似值。

因此，用几何方法求出的阴影部分的面积为

$$\text{面积(ABD)}=\text{面积（扇形）}-\text{面积（}\Delta DBC\text{）}=\frac{\pi}{24}-\frac{\sqrt{3}}{32}$$

我们将牛顿用流数/二项式定理方法所计算出的同一阴影部分面积的近似值与上述结果列为方程，就得到

$$0.07677310678\approx\frac{\pi}{24}-\frac{\sqrt{3}}{32}$$

解π，就得到π的近似值：

$$\pi\approx24\left(0.7677310678+\frac{\sqrt{3}}{32}\right)=3.14159268\cdots\cdots$$

牛顿π近似值的惊人之处在于他只用了二项展开式中的前 9 项，就使其π值精确到 7 位小数，而且，我们发现，牛顿的π近似值与π的真值相差不足 0.000000014。牛顿的π近似值比我们在第四章中所讲到过的韦达或卢道尔夫的惊人的计算又前进了一大步。实际上，应用这一方法，唯一真正的困难在于精确地计算$\sqrt{3}$的近似值。但是，我们前面已经讲到，用牛顿的二项式定理就可以很容易地计算出平方根的值。总之，这一结果清楚地表明了他的数学新发现在解决这一古老问题时的显著效能和巨大成功。

牛顿的π近似值直接引自他的《流数法和无穷级数》，这篇论文写于 1671 年，但几十年没有发表。这篇论文发展了他几年前撰著的《分析学》一书中的流数思想。在《流数法和无穷级数》一书中，牛顿利用$\sqrt{1-x}$二项展开式中的 20 项，计算出 16 位小数的π值。一次，在讲到这一π近似值时，他有几分羞愧地承认，“我真不好意思告诉你我计算到了多少位小数，因为当时我没有其他事情好做。”

尽管牛顿感到羞愧，但那些能够细致入微地欣赏数学美的人却会对当时没有其他紧迫问题占据他那智慧的大脑而感到由衷的高兴，因为当时正是“……我创造力的全盛时期，而且对数学和哲学的关心超过其他任何时候。”

后记

这些就是牛顿在 17 世纪 60 年代中叶瘟疫期间就读于三一学院时所取得的成果。此后的六十余年，这个英格兰小乡村的不幸孩子逐渐名扬天下。本章的结尾部分将介绍他不平凡一生中的其他篇章。

1668 年，牛顿完成了他硕士学位的学习，并被选为三一学院的研究员。这就意味着他只要庄严宣誓，并保持独身，就可以无限期地保留他的学术职位，并得到额外津贴。不仅如此，第二年，艾萨克·巴罗辞去卢卡斯讲座数学教授的席位，并力荐牛顿接替他担任教授。据说，巴罗去职是因为他承认牛顿在数学上更胜于他，因而不能心安理得地占据教授一席。但实际上，巴罗的去职并非出于高尚动机，他在希腊文和神学方面也是一位出色的学者，当时正在角逐其他领域的更高职位。巴罗辞去卢卡斯讲座教授职位后，不久就担任了御前牧师。尽管如此，巴罗在牛顿担任教授一事中，毕竟起了很大作用。巴罗当然有知人之明，因而，他诚心诚意地推荐牛顿是“……我们学院的一位研究员，……非常年青……但却是一位非凡的天才和大师。”

牛顿作为卢卡斯讲座的教授，事情并不多。他既不必教学生，也不必作指导教师，他除了领取丰厚的薪酬，保持道德上的清高之外，主要工作就是定期做数学演讲。如果有人以为学生一定会蜂拥而至，聆听这位伟人的讲座，那么，他们就会感到非常吃惊。不要忘记，牛顿在他那非常狭小的圈子之外尚无名望，而且，当时剑桥大学的学生也不必勤奋向学。一位同时代人曾对牛顿的卢卡斯讲座作过如下记载：

“……听他讲座的人很少，而且，能够听懂的人就更少，由于缺少听众，他几乎常常对着墙壁宣讲。”

他还说，牛顿的讲座一般持续半个小时，除非一个听众也没有，而在这种情况下，他只在那里呆 15 分钟。

如果说牛顿口才不佳，那么，他的科学研究成果却非常丰硕。他很少交朋友，很少与人来往，在三一学院中成了一个离群索居而有几分奇特的人物。一位多年的同事回忆说，他只看到牛顿笑过一次。他唯一的那次笑是由他一位熟人引起的。当时，这位熟人正在读一本欧几里得的书，他问牛顿这部老朽的旧书有什么价值。对此，牛顿不禁放声大笑。

牛顿的侄子汉弗莱·牛顿对他的教授生活做了最形象的描述，他写道：

“他总是把自己关在屋子里研究，很少去拜访别人，也没有人来拜访他……我从来没见他有过任何消遣或娱乐，不论是骑马出去呼吸新鲜空气，散步、打保龄球，还是任何其他运动。他认为所有这些活动都是浪费时间，不如利用这些时间去作学问……他很少到餐厅用饭……如果没人关照他，他会变得非常邋遢，鞋子拖在脚上，袜子不系袜带，穿着睡袍，而且，几乎从来也不梳头。”

然而，随着他未发表的论文，如《分析学》和《流数法和无穷级数》等等的流传，牛顿的名望与日俱增。1671 年，他的第一部大作终于公诸于世，他在伦敦皇家学会的一次会议上展示了他新发明的反射望远镜。这一完美的光学仪器，是牛顿的光学理论和他实际动手能力相结合的产物。科学界高度赞扬他的努力，他的反射望远镜依靠底部反射镜而不是依靠顶部沉重而不稳定的透镜，直至今日，这种望远镜依然是光学天文学的首选仪器。

在这一成功发明的激励下，牛顿不久向皇家学会递交了一篇论光学的论文。但是，这一次，他的激进思想受到了某些著名学者，如罗伯特·胡克的质疑与嘲笑。论争本是学术界一个很普通的现象，但牛顿却深为厌恶。一旦面对批评，他就会退回到他个人的小世界中，拒绝发表或与人交流他的思想，以免再次与那些不开化的同事发生冲突。他的这一决定意味着有许多辉煌的科学论文将几十年地躺在他的抽屉里，不为世人所知。我们在下一章将看到，他的这种做法造成了灾难性的后果，几年后，他的发现，特别是微积分，被别人首先发表，他不得不要求优先权。

随着 17 世纪 70 年代的发展，牛顿的兴趣从数学与物理学转移到了其他方面。他将大量时间用于炼金术的研究，但我们从中可以看到一个现代化学家的头脑。然而，也有些事情未免迂腐，例如他在研究《圣经》时着眼于计算各位先知的年代与时期，计算约柜的尺寸大小，等等。他用了大量时间，如此这般地对《圣经》作了慎密的分析，其结果是他拒绝接受三位一体之中圣子耶稣的概念。想一想聘用他的三一学院这个名字，事情真有些古怪。艾萨克的观点过于激进，他不得不保持沉默，至少在他任卢卡斯讲座教授期间是如此。

这样到了 1684 年。后来以其名字命名彗星的著名天文学家埃德蒙·哈雷拜访了牛顿，并力劝牛顿公布他的惊人发现。犹如往常一样，牛顿仍不情愿，但哈雷的劝说（更不必说哈雷答应负担出版费用）使牛顿相信该是发表他的著作的时候了。牛顿狂热起来，开始勤奋地工作，整理他的论文。这部著作日后成为他的科学代表作，其中阐述了他对运动定律和万有引力原理的研究。1687 年，这部著作终于问世了，题为《自然哲学的数学原理》。展现在我们面前的，是一个宇宙体系，是对月球和行星运动的精确的数学推导，它使天地万物的严整性得到了解释，并与牛顿奇妙的方程正相吻合。自《原理》发表后，科学的面貌为之一变。

《原理》获得了巨大的成功。虽然很少有人能够洞晓书中全部奥妙，但人们普遍认为牛顿近乎超人。许多年以后，法国数学家皮埃尔—西蒙·拉普拉斯记载了他对牛顿科学发现的尊崇、敬畏和羡慕之情：

“牛顿是迄今为止最伟大的天才，也是最幸运的人，因为只有一个世界体系可供我们发现。”

《原理》发表后的第二年是英国历史上重要的一年。1688 年底，斯图亚特王朝的最后一位国王詹姆斯二世被赶下王位，逃往法国，威廉三世和玛丽二世即位。在随后称为“光荣革命”的政治改革中，国会的影响越来越大，而君主的权势则日趋衰落。有趣的是，1689 年，从剑桥派往威斯敏斯特的国会议员不是别人，正是卢卡斯讲座教授艾萨克·牛顿。

作为新君王的支持者，牛顿显然未能以其国会议员的身份给英国政府留下什么印象。尽管如此，他的生活却的确因此而发生了新的转机。如今，他不再是一个落落寡合，离群索居的学者，却以一种几年前甚至不可想象的方式登上了社会舞台。伴随《原理》一书的巨大成功，这位剑桥大学的教授成为伦敦的官员。他似乎很喜欢这种变化，并与许多知名人士交上朋友，如约翰·洛克和塞缪尔·佩皮斯。但是，1693 年，牛顿生了一场大病，几乎精神崩溃，他生病的原因在一定程度上是因为他在做炼金术实验时常常品尝化学药品。1695 年，牛顿身体康复，第二年，他辞去了卢卡斯讲座教授的职务，并离开了三一学院。自从牛顿作为一名普通的大学生从乌尔

索普进入了三一学院以来，已经度过了 35 个春秋。35 年的时光已将这个年青人变成了任何人都未曾预想到的伟人。

那么，这位前教授后来又做了些什么呢？由于社会公职给牛顿留下了良好的印象，也许还由于他越来越认识到自己科学发明的顶峰时期已经过去，牛顿准备尝试一条完全不同的道路。因而，1696 年，他接受了造币局局长的职位。当时，英国的货币是在伦敦塔上铸造的，而那里也是造币局局长生活和工作的地方。据说，牛顿在造币局干得很不错，他监管了英国币制的全面改革，并且与伦敦市的金融家和银行家们相处得十分融洽。

牛顿任造币局局长的这些年，还有机会从事科学撰著。1704 年，他出版了《光学》，在这部巨著中，牛顿奠定了他的光学理论，就像《原理》阐明了他的万有引力定律一样。有趣的是，牛顿是在《光学》的附录中，第一次发表了他的流数法理论，这篇论文题为《曲线求积术》。虽然牛顿早在 40 年前就已提出了这些思想，但是，直到 1704 年，他的这些理论才公诸于世。遗憾的是，他发表得太晚了。几年前，欧洲大陆的数学家已经发表了他们自己对微积分研究的论文。牛顿宣称他现在发表的这些理论其实已经诞生 40 年之久，对此，欧洲大陆的一些数学家即使没有公开表示怀疑，至少反应十分冷漠。

就在《光学》发表前不久，牛顿当选为皇家学会主席。他在这一职位上同样显示出非凡的管理才能，他的这些才能在他任职造币局局长期间便已有目共睹。牛顿担任皇家学会主席一职直到逝世。

1705 年，卓异的科学家、杰出的数学家、公务员和皇家学会主席艾萨克 · 牛顿被安妮女王封为爵士，倍极恩宠。授爵仪式恰当地选在剑桥大学三一学院举行。牛顿以“ 艾萨克爵士 ”的头衔又生活了 22 年。

在这最后 22 年里，牛顿一直生活在伦敦，他将时间分别用于造币局和皇家学会的公务、科学撰著和参加首都一些有影响的活动。这些年肯定是艾萨克爵士春风得意的时期，他的权势和名望（更不要说他的个人财产）与日俱增。

牛顿一直活到 84 岁高龄，于 1727 年逝世。其时，艾萨克 · 牛顿已被国人视为国宝，他也的确不愧这一崇高的赞誉。牛顿显然是欧洲最优秀的科学家，他的影响不亚于一次革命。他死后安葬在威斯敏斯特教堂，享受到与国王和英雄同等的殊荣。今天，牛顿的塑像醒目地矗立在威斯敏斯特教堂唱诗班大屏饰左面入口处，所有进入这一圣地的人都会一眼看到。

全世界有许多赞颂牛顿的诗篇。例如，英国大诗人亚历山大 · 蒲柏曾写道：

宇宙与自然的规律藏匿在夜空，
上帝说“ 要有牛顿 ”，于是一切都变得光明。

另一位著名诗人威廉 · 华兹华斯，性格有些压抑，他描写了诗人在三一学院度过的一夜：

遥借星月之光，
伏枕远望，
教堂前矗立着牛顿雕像，
看那默然无语，
却棱角分明的脸庞。
这大理石幻化的一代英才，

永远在神秘的思想大海中，
独自远航。

对这位孤独的远航家的影响，怎么估计也不过分。我们只需回忆一下 100 年前卡尔达诺的世界观，就足以理解牛顿影响的深远意义——卡尔达诺的世界观是一种将科学与最古怪的迷信混合在一起的大杂烩。那时，世界在很大程度上被看作一个无理性的地方，一种超自然的力量渗透在世间一切事物之中，从彗星的形状到日常生活中的灾难，无所不包。而牛顿却以其极有规律的世界，从自然界排除了超自然的力量。他的理论阐述了一个理性的世界，一个有其基本法则（这在牛顿的遗产中占有很大比重），凡人能够解释的世界。

有趣的是，就在牛顿进入剑桥大学 166 年后，另一位英国青年在剑桥大学基督学院开始了他的大学生涯，而且，他的住处离牛顿在三一学院的旧居仅隔几个街区。年轻的查尔斯 · 达尔文肯定常常走在许多年前牛顿所熟知的剑桥大学同一条街道上。像牛顿一样，达尔文也不愿公布他的发现，但是，1859 年，他动笔写出了经典性的《物种起源》，这部巨著对生物学的影响，一如牛顿的《原理》之于物理学。犹如牛顿创造了物理“ 自然 ”世界一样，达尔文也同样创造了生物“ 自然 ”世界，他解释了地球上生命冲动的似乎无法解释的机制。他们两人的影响都十分深远，远远超出了科学本身。他们两人的理论都使人类对现实世界的认识产生了一场深刻的革命。今天，达尔文也同样长眠在威斯敏斯特教堂，与牛顿墓只相距几英尺——两个科学巨人，两个登峰造极的剑桥大学学生。

艾萨克 · 牛顿在其晚年，回忆他不平凡的智力探索过程，谦和地承认，如果他比别人看得更远些，那是因为他站在巨人的肩膀上。这里，他当然是指维埃特、伽利略、笛卡儿和英雄世纪的其他伟人。现在，他自己的肩膀也将托起后代学人。在一段常常被人引用的非常著名的话里，牛顿写道：

“ 我不知道世人怎样看我；可我自己认为，我好像只是一个在海边玩耍的孩子，不时为拾到更光滑些的石子或更美丽些的贝壳而欢欣，而展现在我面前的是完全未被探明的真理之海。 ”

但是，也许我们应当以下面这段墓志铭，祷祝他在威斯敏斯特教堂安息：

“ 生民们，曾有如此一位伟人为人类而生，你们应当感到庆幸。 ”

第八章
伯努利兄弟与调和级数
（1689 年）

莱布尼兹的贡献

在剑桥大学孤独的艾萨克·牛顿改变数学面貌的同时，欧洲大陆的其他数学家们也并非无所用心。17 世纪后半叶，在笛卡儿、帕斯卡和费马的影响下，欧洲的数学蓬勃发展，其中最伟大的数学家就是戈特弗里德·威廉·莱布尼兹（1646—1716 年）。

人们常常称莱布尼兹为全才，他精通多种学科，并在每个学科中都有所建树。他的父亲是一位伦理学教授。莱布尼兹堪称神童，很小的时候就可以到他父亲庋藏丰富的书房中去读书。利用这一机会，小莱布尼兹幼年时便自学了拉丁文和希腊文。他如饥似渴地读书，15 岁就进入了莱比锡大学。他的学业进展神速，不足 20 岁时就在阿尔特多夫大学完成了他的博士论文。

虽然莱布尼兹的学术生涯很有前途，但他却离开了大学，去为美因茨的选帝侯工作。当时的德国划分为许多小的邦国，选帝侯是这些小邦国中的当权者。莱布尼兹在工作中审查了一些非常复杂的法律问题，包括神圣罗马帝国的重大改革。在业余时间里，他设计了一台计算机。这台计算机的乘法运算是通过快速地重复相加进行的，同样，其除法运算是通过快速重复相减进行的。虽然莱布尼兹努力宣传其计算机的高效率，但当时的技术条件限制了这种计算机的推广使用，这使莱布尼兹不免困恼。尽管如此，他的理论却是可靠的，而最终也是可行的。

1672 年，莱布尼兹作为高级外交使节被从德国派往巴黎。法国首都的文化生活令他深深地陶醉，在他附带出访伦敦和荷兰时，这位年轻的天才又有幸结识了一些著名的学者，如胡克、博伊尔、列文虎克和哲学家斯宾诺莎。莱布尼兹发现自己处于一种活跃的学术环境之中。然而在 1672 年，甚至他也只得承认他的数学教育只限于阅读了一些古典名著。具有强烈好奇心和很高天资的莱布尼兹感到自己需要一个“速成班”，以把握当代的数学趋势和方向。

幸运的是，他在巴黎遇上了绝好的机会。有一位荷兰科学家名叫克里斯蒂安·惠更斯（1629—1695 年），他享受太阳王路易十四的津贴，一直住在巴黎。惠更斯的研究成果给人印象至深。在理论方面，他对数学曲线，特别是对“旋轮线”作了广泛的研究。所谓旋轮线，就是一个圆沿一定直线滚动时圆周上的一个定点所产生的轨迹曲线（见图 8.1）。他的发现在他设计钟摆时起了很大作用，钟摆的工作原理与旋轮曲线密切相关。

这一发明表明，惠更斯不仅仅关注纯数学。实际上，他的声誉或许主要建立在物理学和天文学方面，他研究了运动定律和离心力，并提出了高明的光波理论。而且，惠更斯还借助望远镜第一个解释了土星周围古怪的附属物实际上是光环。

既然在巴黎有这样一位科学家，那么，莱布尼兹向惠更斯请教，以提高自己的数学水平，就毫不奇怪了。如果说惠更斯是莱布尼兹的老师，也

许有些夸大其词，但他在研究当代数学方面，的确给了这位年轻的外交家许多指导。当然，在历史上，老师也很少能有像戈特弗里德·威廉·莱布尼兹这样优秀的学生。

惠更斯指导莱布尼兹研究的一个问题是求三角形数的倒数和。所谓三角形数，就是对应于三角形阵列的数字，如图 8.2 所示。第一个三角形数是1，第二个是3，第三个是6。总之，第k个三角形数等于$\frac{k(k+1)}{2}$。在保龄球运动中，如果将滚道尽头楔形排列的木柱改为 10 个一组，那么，就构成了标准的“三角形数”。三角形数

惠更斯要求莱布尼兹求出的不是三角形数的和，而是三角形数的倒数和。总之，他要求他年轻的学生求出 S 的值，在这里，

$$S=1+\frac{1}{3}+\frac{1}{6}+\frac{1}{10}+\frac{1}{15}+\frac{1}{21}+\cdots\cdots$$

莱布尼兹想了一会儿，就把方程的所有各项全都除以 2，得

$$\frac{1}{2}S=\frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots\cdots$$

在这个方程式中，莱布尼兹发现了一个突出的特点，也就是说，他可以用等价式$1-\frac{1}{2}$替换方程右边的第一项$\frac{1}{2}$，然后依次用$\frac{1}{2}-\frac{1}{3}$替换$\frac{1}{6}$，用$\frac{1}{3}-\frac{1}{4}$替换$\frac{1}{12}$，等等。这样，就将上述方程变为

$$\frac{1}{2}S=\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\left(\frac{1}{4}-\frac{1}{5}\right)+\cdots\cdots$$

然后，莱布尼兹去掉括号，并约消化简，得

$$\frac{1}{2}S=1-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}+\frac{1}{4}-\frac{1}{5}+\cdots\cdots=1$$

如果说，S 的一半等于 1，那么，S 本身（即三角形数的倒数和）就显然等于 2。总之，莱布尼兹非常巧妙地解决了惠更斯的挑战，并发现

$$1+\frac{1}{3}+\frac{1}{6}+\frac{1}{10}+\frac{1}{15}+\frac{1}{21}+\cdots\cdots=2$$

虽然现代数学家对莱布尼兹解无穷级数的方法持有一定的保留意见，但谁也不能否认他的方法的基本独创性。

这仅仅是莱布尼兹对数学超凡洞察力的开始。不久，他又以其巨大的才智，研究牛顿在 10 年前论述的关于切线与面积的同一问题。1676 年，莱布尼兹离开了巴黎，这时，他已经发现了微积分的基本原理。在巴黎生活的四年，使他从一个数学上初出茅庐的新手成长为一个数学巨人。

这四年虽然奠定了莱布尼兹永久名望的基础，但同时也奠定了一场持久争论的基础。我们回想一下，艾萨克·牛顿的流数理论只有几个英国数学家知道，只有他们几个人见到过牛顿论流数法的手稿。1673 年，莱布尼兹在访问伦敦期间，被接受为英国皇家学会的外籍会员。在此，他见到了牛顿的一些文献，并留下了很深的印象。后来，莱布尼兹通过皇家学会的秘书亨利·奥尔登伯格转交给牛顿一封信，他在信中进一步询问了牛顿的发现。伟大的英国科学家牛顿则以一种含混的方式作了答复。牛顿 1676

年这两封著名的复信，我们今天称之为“前书”和“后书”。莱布尼兹认真地阅读了这两封信。

因而，当戈特弗里德·威廉·莱布尼兹首次发表他的论文，宣布这一惊人的数学新方法时，他的英国对手则大叫“卑鄙！”莱布尼兹这篇论文的题目十分冗长，题为《一种对分式和无理量也适用的求极大值、极小值和切线的新方法以及非常规类型的有关计算》（简称《求极大值和极小值的新方法》）。这篇论文刊登在 1684 年的学术杂志《博学者学报》上，而莱布尼兹恰恰是这本杂志的编辑。

因此，世界是通过莱布尼兹，而不是通过牛顿得知微积分的。实际上，微积分的名称就取自莱布尼兹一篇论文的题目。但是，袒护其同胞的英国人则转弯抹角地说，莱布尼兹剽窃了牛顿的全部发明。莱布尼兹访问过英国，他熟知牛顿手稿私下流传的情况，而且，他还与牛顿通过信——所有这一切都使英国人相信，是恶棍莱布尼兹窃取了牛顿的荣誉。

随后的争执构成了数学史上不光彩的一页。起初，两位主角都企图置身事外，而让他们的支持者去为自己作战。但是，最后双方都卷了进去，当然，这种争吵最后总是没有好结果的。莱布尼兹坦率地承认，他通过通信和阅读牛顿的手稿接触过牛顿的思想，但是这些只给了他某些提示，而不是明确的方法；这些新的计算方法是莱布尼兹自己发现的。

与此同时，英国人变得越来越愤怒。而且，（从英国人的观点来看）更糟糕的是，莱布尼兹的微积分很快便被欧洲所接受，并且，他的弟子还在努力扩大其影响；而孤独的牛顿却仍然拒绝发表任何有关微积分的论文。我们回想一下，牛顿早在 1666 年 10 月就写出了他第一篇论流数法的论文，比莱布尼兹发表的论文早了将近 20 年；但是，直到 1704 年，牛顿才在其《光学》的附录中专门论述了他的有关方法。1673 年，在莱布尼兹访问伦敦时，牛顿的一部更详尽论述流数法的著作《分析》还在英国数学界中非正式流传，直至 1711 年才正式付印出版。牛顿为提供一部“供学人使用的完整提要”，认真撰写了一部专著，全面阐述了其已经成熟的思想，但这部著作直到 1736 年才问世，而这时艾萨克爵士已经逝世整整 9 年了！实际上，牛顿发表他数学论文的速度太慢了，以致莱布尼兹的一些狂热的支持者可以宣称是牛顿剽窃了莱布尼兹已出版的著作，而不是相反。

显然，情况混乱不堪。鲁珀特·霍尔在其《争斗中的哲学家》一书中对英吉利海峡两岸纷纷扬扬的指责与反驳作了详尽而生动的描述。今天，飘荡了近三百年的迷雾终于散去，人们公认，牛顿和莱布尼兹两人实际上各自独立地发展了同一种思想体系。在科学发展中，两人或几人同时发现某一重要概念的现象并不罕见，如我们在第二章中曾介绍过的非欧几何的产生即是如此。自牛顿/莱布尼兹争论 150 年后，生物界又出现了英国科学家艾尔弗雷德·拉塞尔·华莱士与查尔斯·达尔文同时创立自然选择理论的问题。在这一事例中，达尔文《物种起源》产生了巨大影响，而华莱士的著作却默默无闻，这可能就是达尔文流芳百世的原因。并且，进化论的两位发现者都是英国人，因而排除了牛顿/莱布尼兹论争中存在的民族情绪。

莱布尼兹一旦从有关微积分发明权的争论中脱出身来，便致力于多种学科的研究。他在不伦瑞克公爵处谋得一个职位，着手追溯公爵的古老家世。他成为梵语和中国文化的专家。并且，他还继续进行哲学研究，哲学

一直是他最热衷的学科。莱布尼兹根据“人类思维字母化”的设想，运用一种谨慎规定的“有理微积分”，寻求发展一种完善的形式逻辑体系。莱布尼兹希望人类能够应用这一逻辑工具，摆脱充斥日常生活中的不准确和无理性。当然，这一切只能称为伟大规划，从未能够实现，但他在这一方面的努力却是朝着我们今日所谓“符号逻辑”的方向迈出的第一步。特别是，他应用代数公式替代逻辑叙述的方法是从古希腊文字推理的逻辑理论向前发展了一大步。

戈特弗里德·威廉·莱布尼兹

1700 年，莱布尼兹成为创建柏林科学院的主要推动者。这一学者、作家和音乐家云集的机构意在为柏林吸引欧洲最伟大的思想家，使柏林跻身于思想中心之列。莱布尼兹荣幸地担任了科学院的院长，直至逝世。

尽管柏林科学院的工作十分繁忙，但莱布尼兹并未因此而放弃研究。他继续钻研逻辑和哲学，并同时倡导世界宗教和政治体制的改革，希望能够因此给人类带来真正的和平与和谐。有趣的是，他最后几年的保护人是汉诺威的一名贵族，1714 年英国女王安妮逝世后，这位贵族竟然一跃成为英国国王乔治一世。莱布尼兹非常希望能够跟随乔治国王去英国，并担任宫廷史学家，但乔治从未给他这种机会。如果微积分之战的两位主角——牛顿和莱布尼兹——同时都住在伦敦，事情一定会很精彩，但遗憾的是，情况并未如此。

莱布尼兹死于 1716 年。当时，他的许多朋友和汉诺威宫廷的同僚都去了英国；他自己的地位也已衰落；据说，只有一位忠实的仆人参加了这个伟人的葬礼。这与牛顿在英国的巨大威望形成了鲜明的对照。如我们在前一章所述，牛顿的崇高名望使他得以安葬在威斯敏斯特教堂。牛顿的崇高声望无疑是当之无愧的，但莱布尼兹也应享有同样的荣誉。

比较一下这两位微积分的伟大发明者，就可以看到一个突出的事实。在一定意义上说，牛顿把他的流数法带入了坟墓。孤独、厌世的艾萨克爵士直到他最后的时日都始终未能有一群聪敏的弟子环伺左右，渴望学习、完善、并传播他的著作。相形之下，莱布尼兹的幸运之处就在于，他有两个最热心的弟子，即瑞士的雅各布·伯努利和约翰·伯努利兄弟，他们成为在欧洲传播和推广微积分的主要人物。他们的努力，也许和莱布尼兹自己的努力一样，令微积分呈现了保留至今的韵味与面貌。

伯努利兄弟

雅各布·伯努利（1654—1705 年）在两兄弟中居长，是一位天才的数学家，他对微积分、无穷级数的求和，也许最重要的，是对概率论的形成作出了重要的贡献。我们已知道，概率这一数学分支是如何在 16 世纪经卡尔达诺首先提出的,又是如何在 17 世纪中叶经费马和帕斯卡的共同努力而发展的。1713 年，雅各布死后出版的巨著《猜度术》为概率论的发展建立了又一个里程碑。这部巨著不但巩固了前人的发现，而且还把概率论研究提到了新的高度。这部巨著是雅各布·伯努利的名作。

同时，弟弟约翰（1667—1748 年）在数学上也自成一家。约翰·伯努利以其坦诚的热情，承担起在欧洲传播莱布尼兹微积分的重任。约翰经常

与他的德国老师通信，在与牛顿派英国人的论争中随时准备捍卫莱布尼兹的名望。我们可以回想一下，19 世纪中叶，托马斯·赫胥黎面对宗教界的攻击，勇敢地保卫了伟大的博物学家达尔文的学说，并由此赢得了“达尔文的斗牛犬”的称号，我们也可以出于同样的理由称约翰·伯努利为“莱布尼兹的斗牛犬”。像赫胥黎一样，约翰有时也以一种近于惊人的执着支持莱布尼兹；同样，他与赫胥黎一样，最终也完成了他的这一使命。

约翰的一个最重要的贡献是通过他与洛必达侯爵（1661—1704 年）的联系完成的。洛必达侯爵是一个法国贵族和数学爱好者，他非常希望学习这一革命性的新的微积分理论。因此，侯爵聘用约翰·伯努利来为他提供各种有关微积分及任何数学新发现的论文。在某种意义上说，洛必达似乎购买了伯努利的数学研究权。1696 年，洛必达汇编了伯努利的论著，出版了他第一部论微积分的书，题为《无穷小分析》。这部书是用本国语言，而不是用拉丁文写的，除书名外，书中内容几乎全部都是伯努利撰著的。

纵观历史，可以看到许多杰出的兄弟组合。从特洛伊战争中的阿伽门农和墨涅拉俄斯到航空先驱威尔伯·莱特和奥维尔·莱特兄弟，历史上有许多兄弟为实现崇高目标而并肩努力。雅各布与约翰写出了数学史中最重要的兄弟成功的故事，但我们也必须看到，他们两人的关系并不和谐。恰恰相反，在数学中，他们两人中每一个人都是另一个人的强劲竞争对手，两人为了胜出对方一筹而斗力，甚至到了可笑的地步。

例如，有关悬链曲线的问题。所谓悬链曲线，就是一根链条，两端固定，依其本身重量下垂的曲线。1690 年，久负盛名的哥哥雅各布在一篇论文中提出了确定悬链曲线性质（即方程式）的问题。实际上，这一问题已存在多年，伽利略就曾推测过悬链曲线是一条抛物线，但问题依然悬而未决。雅各布觉得，应用奇妙的微积分新方法也许可以解决这一难题。

但遗憾的是，他的努力没有取得结果。一年后，雅各布恼恨地看到他的弟弟约翰发表了这个问题的正确答案。而自命不凡的约翰，却很难算是一个谦和的胜利者，他后来回忆说：

“我哥哥的努力没有成功；而我却幸运得很，因为我发现了解开这道难题的全部方法（我这样说并非自夸，我为什么要隐瞒真相呢？）……为研究这道题，我整整一晚没有休息……第二天早晨，我兴冲冲地去见哥哥，他还在苦思这道难题，但一无进展。他像伽利略一样，始终以为悬链曲线是一条抛物线。行了！行了！我对他说，不要再折磨自己，去证明悬链曲线是抛物线了，因为这是完全错误的。”

有趣的是，约翰成功地解出这道难题所需要的时间：“整整一晚”，而雅各布却花费了整整一年的时间，这实在算得上是一种“奇耻大辱”。

我们将在本章讨论一个由伯努利兄弟两人共同创立的伟大定理（也许是在少有的休战期间创立的）。这个定理所涉及的是关于调和级数的性质问题，所谓“调和级数”，是一种具有特殊性质的无穷级数。虽然我们已见到过莱布尼兹所研究的一种特殊级数，但我们还是应首先对无穷级数问题作一番概述。

17 世纪时，无穷级数仅仅被看作是无穷项的和。当然，不能保证这种级数一定会有一个有限和；例如，像 1+2+3+4+5+……这样的级数，如果我们继续进行下去，其和显然会不断增大，并超过任何有限量。我们说，这种级数为“发散无穷级数”。

另一方面，也存在一种无穷多项的级数，其和为有限数。这种现象，初步看来，似乎自相矛盾，但仔细想一想，就会发现非常合理。例如，在我们写出大家所熟悉的小数展开式 $\frac{1}{3}=0.3333333\cdots\cdots$ 时，我们准确的意思是

$$\frac{1}{3}=\frac{3}{10}+\frac{3}{100}+\frac{3}{1000}+\frac{3}{10000}+\cdots\cdots$$

我们前面所介绍的莱布尼兹级数就显示了同样的性质，级数的无穷多项的和等于一个（有限的）数 2。我们说，这种级数为“收敛级数”，也就是说，不太正规地讲，当我们增加更多的项时，它的和越来越接近某一特定的值。

无疑，数学中最重要的收敛级数是几何级数，其形式为

$$a+a^2+a^3+a^4+\cdots\cdots+a^k+\cdots\cdots$$

在此，我们设 $-1<a<1$。因此，几何级数就是 a 及其所有高次幂的和。我们用一个“17 世纪式”的论证方法来证明这种级数的收敛性，其证明如下：

设 $S=a+a^2+a^3+a^4+\cdots\cdots$ 为我们所求的和。将方程两边同乘以 a，得 $aS=a^2+a^3+a^4+a^5+\cdots\cdots$，然后，将这两个方程相减，得

$$S-aS=(a+a^2+a^3+a^4+\cdots\cdots)-(a^2+a^3+a^4+a^5+\cdots\cdots)=a$$

除第1项外，所有项都消掉了。所以，$S(1-a)=a$，因而，$S=\frac{a}{(1-a)}$。

由于 S 是原几何级数的和，因此，我们可以认为

$$a+a^2+a^3+a^4+a^5\cdots\cdots=\frac{a}{1-a}$$

例如，如果 $a=\frac{1}{3}$，则

$$\frac{1}{3}+\frac{1}{9}+\frac{1}{27}+\frac{1}{81}+\cdots\cdots+\frac{1}{3^k}+\cdots\cdots=\frac{\frac{1}{3}}{1-\frac{1}{3}}=\frac{\frac{1}{3}}{\frac{2}{3}}=\frac{1}{2}$$

就数学的精密性而言，对无穷级数收敛性的这一证明显得十分幼稚，相比之下，现代数学对这个问题的论证就精妙得多。并且，这个证明还掩盖了我们最初为什么要设 $-1<a<1$ 的原因，虽然 a=2 这一几何级数已经表明了这一假设的必要。在这种情况下，我们直接应用公式，就得到

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

也就是说，$2+4+8+16+\cdots\cdots=-2$。这是一个“双重荒谬”的结果，一方面因为这个级数显然是发散无穷级数，另一方面还因为人们无法想象一系列正数相加的结果竟然得出一个负数。因而，几何级数的求和公式要求 α 必须位于 -1 与 1 之间。（对这一问题的更详尽分析，通常需要应用微积分。）

上述两个无穷级数说明了一般收敛级数的一个重要条件。对于第一个 $\alpha=\frac{1}{3}$ 的几何级数来说，其一系列相加的项（$\frac{1}{9}$，$\frac{1}{27}$，$\frac{1}{81}$ 等）越来越接近于零；因此，后面的各项可以越来越忽略不计。而另一方面，对于 α=2

的几何级数来说，我们相加的各项则离零越来越远——4，8，16，等等，其愈益增大的数值使其和不能等于一个有限数。

根据这两个例子，我们可以非常合理地提出下述推测：在无穷级数 $x_1+x_2+x_3+x_4+\cdots\cdots+x_k+\cdots\cdots$ 中，如果，并且只有当通项 x_k 的值趋向于零时，其和才能够收敛为有限数。正如结果所示，这个推测有一半是正确的。即，如果级数收敛于一个有限数，则级数中的通项一定趋向于零。换句话说，除非通项趋向于零，否则，我们不能将一个无穷级数表示为一个有限数。

然而，遗憾的是，其逆命题却是错误的。也就是说，有的无穷级数，即使其通项趋向于零，但其和却趋向于无穷。这一事实并不是显而易见的，但却是我们即将讨论的伟大定理的内容。在调和级数 $1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots\cdots+\frac{1}{k}+\cdots\cdots$（即正整数的倒数和）中，约翰·伯努利发现，虽然其通项趋向于零，但它的和却是无穷大。

伯努利发现了当今数学家称之为“病态反例”的现象——即一个似乎违反直觉的特定例子，其古怪之至，堪称“病态”。这一调和级数非常麻烦：要使其和不大于 5，就必须将级数的前 83 项相加，因为

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots\cdots+\frac{1}{82}=4.990020\cdots\cdots<5.00$$ 而

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots\cdots+\frac{1}{83}=5.002068\cdots\cdots>5.00$$

请注意一个突出的事实，在这一调和级数中，超过第 83 项以后的每一项都小于 $\frac{1}{83}$，因而对其和的增值作用不大。所以，要使和大于 6，就必须再加 144 项。由于其和增值非常缓慢，因此，要使级数和等于 10，就必须将前 12,367 项相加，而要使级数和等于 20，就要加 2.5 亿项！人们似乎根本难以想象调和级数最终可能会超过一百，一千，甚至一万亿。

但事实的确如此！而这正是其之所以被称为病态和伯努利的定理之所以值得我们注意的原因。

伟大的定理：调和级数的发散性

虽然这个证明是约翰·伯努利作出的，但却刊载在哥哥雅各布 1689 年的《论无穷级数》一书中。出于少见的兄弟情谊，雅各布甚至在书的序言中承认了弟弟对这一证明方法的优先权。

约翰必须要证明调和级数向无穷发散。他的证明是以莱布尼兹的收敛级数 $\frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots\cdots=1$ 为基础，莱布尼兹的这一级数，我们在本章前面已经讨论过。此事本身就很奇怪，因为，人们不清楚，这一清晰易解的收敛级数怎么会成为古怪的调和级数的论证基础呢？无论如何，约翰·伯努利作了如下推理。

定理 调和级数$1+\frac{1}{2}+\frac{1}{3}+\cdots\cdots+\frac{1}{k}+\cdots$**的和是无穷的。**

证明 引入$A=\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots\cdots+\frac{1}{k}+\cdots\cdots$**，这是一个缺少第一项的调和级数。将这一级数"……变为分子是 1、2、3、4 等等的分数"，就得到**

$$A=\frac{1}{2}+\frac{2}{6}+\frac{3}{12}+\frac{4}{20}+\frac{5}{30}+\cdots\cdots$$

约翰将这一级数作为后面的参考。

然后，他设定前述莱布尼兹的级数为C**，并从这一级数中连续减去**$\frac{1}{2}$**，**$\frac{1}{6}$**，**$\frac{1}{12}$**，**$\frac{1}{20}$**，等等，由此构成一系列相关级数：**

$$C=\frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots\cdots \qquad =1$$

$$D=\quad \frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots\cdots=C-\frac{1}{2}=1-\frac{1}{2}=\frac{1}{2}$$

$$E=\qquad \frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots\cdots=D-\frac{1}{6}=\frac{1}{2}-\frac{1}{6}=\frac{1}{3}$$

$$F=\qquad\quad \frac{1}{20}+\frac{1}{30}+\cdots\cdots=E-\frac{1}{12}=\frac{1}{3}-\frac{1}{12}=\frac{1}{4}$$

$$G=\qquad\qquad \frac{1}{30}+\cdots\cdots=F-\frac{1}{20}=\frac{1}{4}-\frac{1}{20}=\frac{1}{5}$$

$$\cdot\ \cdot \qquad \cdot \quad \cdot \qquad \cdot \quad \cdot$$

$$\cdot\ \cdot \qquad \cdot \quad \cdot \qquad \cdot \quad \cdot$$

$$\cdot\ \cdot \qquad \cdot \quad \cdot \qquad \cdot \quad \cdot$$

约翰接着将这一方程阵列的最左边两列相加，得到

C+D+E+F+……

$$=\frac{1}{2}+\left(\frac{1}{6}+\frac{1}{6}\right)+\left(\frac{1}{12}+\frac{1}{12}+\frac{1}{12}\right)$$

$$=\left(\frac{1}{20}+\frac{1}{20}+\frac{1}{20}+\frac{1}{20}\right)+\cdots\cdots$$

$$=\frac{1}{2}+\frac{2}{6}+\frac{3}{12}+\frac{4}{20}+\cdots\cdots=A \qquad \text{根据前面设定}$$

另一方面，如果将这一方程阵列的最左边和最右边的两列相加，他发现，

$$C+D+E+F+G+\cdots\cdots=1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\cdots\cdots$$

$$=1+A$$

由于 C+D+E+F+G+……既等于 A，又等于 1＋A，因而，约翰只能得出结论：1+A=A。正如他所说的那样，"整体等于部分"。但是，显然没有一个有限数会等于大于自己的数。约翰·伯努利认为，这只能说明一个问题：即 1+A 是无穷大。而 1+A 则是调和级数的和，所以，他的证明完毕。

今天的数学家可以对这一证明提出一些公正的批评意见。伯努利是以

一种“整体论”的态度来对待无穷级数的，他将其作为一个独立个体而随意处置。我们现在懂得，在处理这些数学问题时，必须特别慎重。并且，他证明调和级数发散性的方法与现代方法形成了鲜明的对照。今天的数学家采用下述方法证明：首先确定正整数 N（不论其数值多大），并证明该级数必定大于 N；那么，既然该级数大于任意正整数 N，则这个级数一定趋向无穷。但是，约翰没有这样证明。相反，他用更加简明的 A=1+A 来证明级数的发散性，对于现代读者来说，这是证明量的无穷性的一个最独特的方法。

我们必须承认，伯努利作出这一论证之后 150 年，才有真正精确的级数理论出现，考虑到这点，或许可以不致过分挑剔。并且，尽管有种种异议，但谁也无法否认约翰论证方法的巧妙。约翰的证明恰似数学王冠上的一颗明珠。

雅各布在其《论无穷级数》一书中就他弟弟的证明强调了一个非直观的重要推断，他写道：“一个最后一项为零的无穷级数之和也许是有限的，也许是无穷的。”现代数学家称赞他提出了无穷级数的“最后一项”问题，因为这些无穷级数的性质的确排除了任何最后项；然而，他的意思非常明确。他所强调的是，在无穷级数中，即使其中的某些项接近于零，其和仍然可能是无穷的。调和级数就是这种现象的首要例子，已如约翰所证明。

也许是因为这一结果太出乎意料了，雅各布情不自禁，挥笔写下了一首数学短诗：

有限环绕无穷级数朝夕相伴，
在无限的王国中也存在着有限；
至大寓于细微之所，
而最狭小的有限中却见到无限。
在无限中认识细微是多么快乐，
巨大存在于细小之中，啊，神秘的上天！

最速降线的挑战

伯努利兄弟在他们时代的数学中留下了深刻的印记，其中包括调和级数和许多其他贡献。但是，关于这对相互竞争，难以相处的兄弟，还必须要告诉读者另一个故事，它肯定是在整个数学史中最引人入胜的一则故事。

故事开始于 1696 年 6 月，其时，约翰·伯努利在莱布尼兹的杂志《教师学报》上刊登了一个挑战问题。显然，公开挑战的传统是从菲奥尔和塔尔塔利亚时代开始的。虽然现在的论争是在学术杂志上安静地进行笔战，但却依然有力量成就或摧毁一个人的声望，正如约翰自己所述：

“……肯定地说，正是摆在我们面前的那些困难同时也是有用的问题，激发着出类拔萃之辈为丰富人类的知识而奋斗，他们也因此一举成名，流芳百世。”

约翰提出的挑战很精彩。他设想在地面上不同高度的两个点 A 和 B，并且，不要让其中一个点直接位于另一点的上方。连接这两个点，当然可以作出无限多的不同曲线，从直线、圆的弧线到无数种其他曲线和波浪线。现在设想有一个球沿着一条曲线从 A 点滚向较低的 B 点。当然，球滚完全

程所需要的时间取决于曲线的形状。伯努利向数学界提出的挑战是，找出一条曲线 AMB，使球沿这条曲线滚完全程所用的时间最短（见图 8.3）。他称这条曲线为“最速降线”，这个词是从希腊文的“最短”和“时间”两个词合成而来的。

显然，第一个猜想是连接 A、B 两点作直线 AMB。但是，约翰对试图采用这一过于简单化的方法提出了警告：

“……不要草率地作出判断，虽然直线 AB 的确是连接 A、B 两点的最短线路，但它却不是所用时间最短的路线。而曲线 AMB 则是几何学家所熟知的一条曲线。如果在年底之前还没有其他人能够发现这一曲线，我将公布这条曲线的名称。”

约翰定于 1697 年 1 月 1 日向数学界公布答案。但是，到最后期限截止时，他只收到了“著名的莱布尼兹”寄来的一份答案，并且，莱布尼兹

“谦恭地请求我延长最后期限到复活节，以便在公布答案时……没有人会抱怨说给的时间太短了。我不仅同意了他的请求，而且还决定亲自宣布延长期限，看看有谁能够在这么长时间之后最终解出这道绝妙的难题。”

然后，为确保不会使人误解这道难题，约翰又重复了一遍：

“在连接已知两点的无限多的曲线中……选择一条曲线，如果用一根细管或细槽代替这条曲线，把一个小球放入细管或细槽中，放手让它滚动，那么，小球将以最短的时间从一点滚向另一点。”

此时，约翰开始热心鼓吹奖励解出他的最速降线问题的人。不要忘记，他自己是知道答案的，如此一来，他关于数学荣誉的一段话就不免有自诩之嫌：

“但愿有人能够迅速摘取桂冠。当然，奖品既非金，也非银，因为这些东西只能引起卑贱者的兴趣……相反，由于美德本身就是最好的奖励，而名望又是最强的刺激，所以，我们为高贵的得胜者所颁发的奖励是荣誉、赞颂和认可……”

在这段话中，似乎约翰认为自己面对他可怜的哥哥雅各布，又一次赢得了胜利。但是，在他心里还有另外一个目标。约翰写道：

“……很少有人能够解出我们独特的问题，即使那些自称通过特殊方法……不仅深入探究了几何学的秘密、而且还以一种非凡的方式拓展了几何学的疆域的人。这些人自以为他们的伟大定理无人知晓，其实早已有人将它们发表过了。”

还有谁能怀疑他所说的“定理”就是指的流数法，他所蔑视的目标就是艾萨克 牛顿呢？牛顿曾宣称早在莱布尼兹 1684 年发表微积分论文之前就已发现了这一理论。无疑，约翰的挑战目标非常明确，他把他的最速降线问题抄了一份，装进信封，寄往英国。

当然，1697 年，牛顿正在忙于造币局的事务，而且，正如他自己所承认的那样，他的头脑已不似全盛期时那样机敏了。当时，牛顿与他的外甥女凯瑟琳 · 康迪特一起住在伦敦。凯瑟琳记述了这样的故事：

“1697 年的一天，收到伯努利寄来的问题时，艾萨克 · 牛顿爵士正在造币局里忙着改铸新币的工作，直到四点钟才精疲力尽地回到家里，但是，直到解出这道难题，他才上床休息，这时，正是凌晨四点钟。”

即使是在晚年，并且，是在经过一天紧张的工作而感到精疲力竭的情况下，艾萨克 · 牛顿仍然成功地解出了众多欧洲人都未能解出的难题！由

此可见这位英国伟大天才的实力。他清楚感觉到他的名望与荣誉都受到了挑战；而且，伯努利和莱布尼兹毕竟都还在急切地等待着公布他们自己的答案。因此，牛顿当仁不让，仅仅用几个小时就解出了这道难题。然而，牛顿有些被激怒了，据说他曾言道：“ 在数学问题上，我不喜欢……给外国人……戏弄。 ”

我们再回到欧洲。复活节将近的时候，几份答案寄到了约翰 · 伯努利的手里。他们每个人所寻求的曲线都是一条颠倒了的旋轮线，而这的确“ 是几何学家所熟知的一条曲线 ”。我们注意到，帕斯卡和惠更斯就曾研究过这一重要曲线，但他们谁也没有认识到旋轮线还是一条最快的下降曲线。约翰以一种夸张的口吻写道：“ ……如果我明确说出惠更斯的……这一旋轮线就是我们所寻求的最速降线，你们一定会惊呆了。 ”

到复活节时，挑战期限截止。约翰一共收到了五份答案。其中包括他自己的答案和莱布尼兹的答案。他的哥哥雅各布寄来了第三份答案（这也许会使约翰感到沮丧），而洛皮塔侯爵则寄来了第四份答案。最后寄来的答案，信封上盖着英国的邮戳。约翰打开后，发现答案虽然是匿名的，但却完全正确。他显然遇到了他的对手艾萨克 · 牛顿。答案虽然没有署名，但却明显地出于一位绝顶天才之手。

据说（或许不尽可靠，但却非常有趣），约翰半是羞恼，半是敬畏地放下这份匿名答案，会意地说：“ 我从他的利爪认出了这头狮子。 ”

后记

在讲到约翰对调和级数发散性的证明时，雅各布曾说过，“ 是我弟弟首先发现的 ”。如果雅各布认为是约翰第一个掌握了调和级数的奇特性质，他就完全错了，因为至少有两个前辈数学家曾证明过调和级数的发散性。这两个数学家的证明各不相同，而且，也不同于上述约翰的证明，但每个人的证明都显示了自己独特的智慧。

最早对调和级数发散性作出证明的是 14 世纪的法国学者尼科尔 ·奥雷姆（约 1323—1382 年）。约 1350 年，奥雷姆写出了一部非凡的著作，题为《欧几里得几何问题》。当然，这是一部非常古老的文献，比卡尔达诺的《大衍术》还早了整整 200 年。尽管这部著作产生于我们也许可以称之为欧洲数学的“ 石器时代 ”，但奥雷姆的著作中的确含有一些非常精彩的论题。

特别是，他论述了调和级数的性质。实际上，他的全部论证如下：

“ ……将1英尺与$\frac{1}{2}$、$\frac{1}{3}$、$\frac{1}{4}$英尺等相加，其和无穷。实际上，可以构成一个其和大于$\frac{1}{2}$的多项无穷数。因此：$\frac{1}{3}+\frac{1}{4}$大于$\frac{1}{2}$；$\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}$大于$\frac{1}{2}$；$\frac{1}{9}+\frac{1}{10}+\cdots\cdots+\frac{1}{16}$大于$\frac{1}{2}$，等等。 ”

读者感到有点儿迷惑是可以理解的。这个论证毕竟完全是用文字阐述的，是在符号代数出现之前几百年写出的。然而，只要经过一点儿“ 净化 ”处理，这段文字就变成了一个非常简单而巧妙的发散性证明了。实际上，

奥雷姆的意思是用其和等于$\frac{1}{2}$的较小分数替换调和级数中的一组分数。

即，他说：

$$1+\frac{1}{2}>\frac{1}{2}+\frac{1}{2}=1$$

$$1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)>1+\left(\frac{1}{4}+\frac{1}{4}\right)=\frac{2}{3}$$

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)$$

$$>\frac{3}{2}+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)=\frac{4}{2}$$

$$1+\frac{1}{2}+\cdots\cdots+\frac{1}{8}+\left(\frac{1}{9}+\frac{1}{10}+\cdots\cdots+\frac{1}{16}\right)$$

$$>\frac{4}{2}+\left(\frac{1}{16}+\cdots\cdots+\frac{1}{16}\right)=\frac{5}{2}$$

这一方程可以扩展为适用于任何整数 k 的一般公式：

$$1+\frac{1}{2}+\frac{1}{3}+\cdots\cdots+\frac{1}{2^{k}}>\frac{k+1}{2}$$

例如，如果 k=9，我们看到，

$$1+\frac{1}{2}+\frac{1}{3}+\cdots\cdots+\frac{1}{512}=1+\frac{1}{2}+\frac{1}{3}+\cdots\cdots+\frac{1}{2^{9}}>\frac{9+1}{2}=5$$

如果 k=99，则

$$1+\frac{1}{2}+\frac{1}{3}+\cdots\cdots+\frac{1}{2^{99}}>\frac{100}{2}=50$$

如果 k=9999，我们得到

$$1+\frac{1}{2}+\frac{1}{3}+\cdots\cdots+\frac{1}{2^{9999}}>\frac{9999+1}{2}=5000$$

这样，只要取调和级数中足够多的项，我们就能够保证其和大于 5、50 或 5000，或一般地说，大于任何有限量。这种方法保证了所有调和级数都大于任何有限量，并因而趋向无穷。奥雷姆的证明巧妙、简洁和易记，已写入了现代大部分数学教科书中。但是，伯努利兄弟似乎不知道有这样一个证明存在。

先于约翰·伯努利作出证明的还有另外一位数学家——意大利数学家彼得罗·门戈利（1625—1686 年）。门戈利的论证作于 1647 年，因而比伯努利的证明早 40 年。门戈利的证明非常简单，他首先提出了一个初步命题。

定理　如果$a>1$，那么，$\frac{1}{a-1}+\frac{1}{a}+\frac{1}{a+1}>\frac{3}{a}$。

证明　首先提出一个明显的论据，即 $2a^3>2a^3-2a=2a(a^2-1)$，将这个不等式的两边分别除以 $a^2(a^2-1)$，得

$$\frac{2a^3}{a^2(a^2-1)} > \frac{2a(a^2-1)}{a^2(a^2-1)}$$ 或简化为 $$\frac{2a}{a^2-1} > \frac{2}{a}$$

因此

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

$$=\frac{1}{a}+\frac{2a}{a^2-1}$$ 根据代数运算

$$>\frac{1}{a}+\frac{2}{a}$$ 根据上述不等式

$$=\frac{3}{a}$$

证讫。

这一命题保证了在三个连续整数的倒数相加时，其和一定大于中间数字的倒数的三倍。我们可以用数字来检验，例如，

$$\frac{1}{8}+\frac{1}{9}+\frac{1}{10}>\frac{3}{9}=\frac{1}{3}$$ 或 $$\frac{1}{32}+\frac{1}{33}+\frac{1}{34}>\frac{3}{33}=\frac{1}{11}$$

这就是门戈利在他 1647 年对调和级数的简短证明中所需要的初步命题。

定理　调和级数趋向无穷。

证明　设 H 为调和级数的和。通过对级数各项归组和反复应用上述不等式，我们发现：

$$H=1+\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}\right)+\left(\frac{1}{8}+\frac{1}{9}+\frac{1}{10}\right)$$

$$+\left(\frac{1}{11}+\frac{1}{12}+\frac{1}{13}\right)+\cdots\cdots$$

$$>1+\left(\frac{3}{3}\right)+\left(\frac{3}{6}\right)+\left(\frac{3}{9}\right)+\left(\frac{3}{12}\right)+\left(\frac{3}{15}\right)+\cdots\cdots$$

$$=1+1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\cdots\cdots$$

$$=2+\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}\right)+\left(\frac{1}{8}+\frac{1}{9}+\frac{1}{10}\right)$$

$$+\left(\frac{1}{11}+\frac{1}{12}+\frac{1}{13}\right)+\cdots\cdots$$

$$>2+\left(\frac{3}{3}\right)+\left(\frac{3}{6}\right)+\left(\frac{3}{9}\right)+\left(\frac{3}{12}\right)+\left(\frac{3}{15}\right)+\cdots\cdots$$

$$=2+1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\frac{1}{9}+\cdots\cdots$$

$$=3+\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}\right)+\left(\frac{1}{8}+\frac{1}{9}+\frac{1}{10}\right)$$

$$+\left(\frac{1}{11}+\frac{1}{12}+\frac{1}{13}\right)+\cdots\cdots$$

等等。门戈利证明的精彩之处就在于它的自我复制性质。他每次将他的初步定理应用于调和级数，他就再一次遇到调和级数，但这一次则增加了一个单位。我们来看以上的不等式，我们发现，H 大于 1，大于 2，大于 3；而且，如果我们继续重复这一过程，H 大于任何有限量。因此，我们可以与门戈利一道得出结论，调和级数的和一定是无穷大。　　证讫。

所以，约翰的伟大定理虽然证明方法有所不同，但奥雷姆和门戈利都确实先于他发现了调和级数的这一性质。并且，雅各布在《论无穷级数》一书中载入约翰的证明之后，便直接提出了他自己对调和级数发散性的证明。他的证明虽然带有兄弟间争强好胜的味道，但确是一个非常精彩的证明。然而，雅各布的证明似乎过于复杂，不便在本书介绍。

在《论无穷级数》一书中，雅各布在论证了调和级数之后，又进一步阐述了整数平方的倒数和问题。他发现，

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots\cdots+\frac{1}{k^2}+\cdots\cdots=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\cdots$$

他注意到，$\frac{1}{4}<\frac{1}{3}$，$\frac{1}{9}<\frac{1}{6}$，$\frac{1}{16}<\frac{1}{10}$，一般来说，

$$\frac{1}{k^2}<\frac{1}{\frac{k(k+1)}{2}}$$

他由此推理，

$$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\cdots+\frac{1}{k^2}+\cdots\cdots<1+\frac{1}{3}+\frac{1}{6}+\frac{1}{10}+\frac{1}{15}+\cdots\cdots$$

$$+\frac{2}{k(k+1)}+\cdots\cdots=2\left(\frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\cdots\cdots\right)=2(1)=2$$

我们在这里再次应用了本章开始所介绍的莱布尼兹求和定理。在此，雅各布表明，上述级数趋向某一小于 2 的有限数。鉴于明显的原因，这一证明收敛性的方法现在称为“比较判别法”。雅各布的证明提供了一个实际应用比较判别法的早期例子。

虽然伯努利兄弟知道这一无穷级数是收敛级数，但他们未能找到其和的精确数值。雅各布带着几分绝望的恳求宣告了他的失败：“如果有人能够发现并告知我们迄今为止尚未解出的难题的答案，我们将不胜感谢。”

求级数$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\cdots$的值是一个非常难的难题，看来，只能由一

个胜过伯努利兄弟的天才来解出这一级数的和了。

有趣的是，1734 年，一位师从约翰 · 伯努利的青年人终于解出了这道难题。在求这一级数和的过程中，犹如在数学的许多其他领域一样，这个青年人最终超过了他的老师。实际上，他超过了曾经就数学研究写过些什么的所有人。这个青年学生就是我们下一章伟大定理的创始人李昂纳德 ·欧拉。

第九章
李昂纳德·欧拉非凡的求和公式
（1734 年）

通晓数学的大师

在漫长的数学史中，李昂纳德·欧拉的遗产是无与伦比的。他博大精深和空前丰富的著述令人叹为观止。欧拉厚厚的 70 多卷文选，如此深远地改变了数学的面貌，足以证明这位谦和的瑞士人的非凡天才。实际上，面对他数量奇多，质量极高的著述，人们的第一个感觉便是，他的故事似乎是一部天方夜谭，而不是确凿的史实。

这位伟人于 1707 年出生在瑞士的巴塞尔。毫不奇怪，他在年轻时即表现出超人的天赋。欧拉的父亲是一个加尔文教派的牧师，他设法安排年轻的李昂纳德师从著名的约翰·伯努利。欧拉后来常常回忆起与他的老师伯努利在一起的这段时光。小欧拉经过一星期的学习准备，然后在每个星期六下午的指定时间里，去向伯努利请教一些数学问题。伯努利并非总是仁慈和蔼，最初常常为了学生的不足而发火；而欧拉则更加勤奋，尽可能不以琐事去烦扰老师。

不论约翰·伯努利的脾气是否很坏，他很快就发现了他学生的非凡天才。不久，欧拉就开始发表高质量的数学论文。19 岁时，欧拉以其对船上安装桅杆的最佳位置的精彩分析而荣获了法国科学院颁发的奖金。（值得注意的是，那时，欧拉还从未见过海船！）

1727 年，欧拉成为俄国圣彼得堡科学院的成员。当时，俄国建立科学院，是为了与巴黎和柏林的科学院相匹敌，以实现彼得大帝的梦想。在移居俄国的学者中，有一位是约翰的儿子丹尼尔·伯努利。通过丹尼尔的影响，欧拉谋得了职位。但奇怪的是，由于自然科学方面没有空缺，欧拉只能就任医学和生理学方面的职位。然而，职位毕竟是职位，欧拉欣然领受。开始的路程十分艰难，他甚至在俄国海军当了一段医官。终于，1733 年，数学教授丹尼尔·伯努利辞职返回瑞士，欧拉接替了丹尼尔的职位。

当时，欧拉已显示出后来成为他整个数学生涯鲜明特征的过人精力和巨大创造力。虽然在 18 世纪 30 年代中期，欧拉的右眼开始失明，而且，不久就完全失明，但是，伤残并没有影响他的科学研究。他不屈不挠，解决了各个数学领域（如几何学、数论和组合）及应用领域（如机械学、流体动力学和光学）中的种种疑难问题。只要想象一下一个人在失明后还要向世界揭示光学的奥秘，我们就会受到强烈的感染和激励。

1741 年，欧拉离开了圣彼得堡科学院，并应腓特烈大帝的邀请，成为柏林科学院院士。在一定程度上，他离开俄国是因为他不喜欢沙皇制度的压抑。但遗憾的是，柏林的情况也并不理想。腓特烈嫌他太单纯、太文静、太谦和。这位德国国王在一次提到欧拉的视力问题时，竟称欧拉为“数学独眼龙。”由于腓特烈的这种态度，以及科学院内部的一些明争暗斗，欧拉于俄国叶卡捷琳娜二世在位期间应邀返回了圣彼得堡。他后来一直住在俄国，直到 17 年后逝世。

欧拉的同时代人称他是一个善良和宽宏大量的人，他喜欢自己种菜和

给他 13 个孩子讲故事。在这一方面，欧拉是一个受人欢迎的人物，恰与孤僻、缄默的艾萨克·牛顿形成鲜明对照，而牛顿确是少有的一位可与他比肩而立的数学大师。我们从中欣慰地看到，这一等天才并非个个都是神经质。甚至在 1771 年，欧拉的另一只眼睛也失明后，他仍然保持着这种温良的性格。尽管欧拉双目全盲，而且经常疼痛，但他依然坚持向他的助手口授他奇妙的方程和公式，在助手的帮助下，继续从事数学著述。正如失聪没有阻碍下一代的路德维希·冯·贝多芬的音乐创作一样，失明也同样没有阻碍李昂纳德·欧拉的数学探索。

欧拉的整个数学生涯，始终得益于他惊人的记忆力，对此，我们只能称他为超人。他在进行数论研究时，不但能够记住前 100 个素数，而且还能记住所有这些素数的平方、立方，甚至四次方、五次方和六次方。欧拉可以很轻松地背诵出诸如 241^4 或 337^6 的数值,而其他人却要忙着查表或笔算。但这还只是他显示非凡记忆力的一些小把戏。他能够进行复杂的心算，其中有些运算要求他必须要记住 50 位小数！法国物理学家弗朗索瓦·阿拉戈说，欧拉计算时似乎毫不费力,“ 就像人在呼吸，或鹰在翱翔一样轻松。”除此以外，欧拉还能够记住大量的论据、引语和诗歌，包括维吉尔的《埃涅阿斯纪》全篇，这部史诗是欧拉幼年时诵读的，时隔 50 年后，他依然能够一字不差地背出全文。任何一位小说作家都不敢编造出一个具有如此惊人记忆力的人物。

欧拉无与伦比的名望是与他的数学论著密不可分的。他的笔下，既有一些高难度的数学著作，也有一些初级数学书，而且，他并不认为如此就降低了身份。也许，他最著名的著作是他 1748 年发表的《无穷小分析引论》。这部不朽的数学论著可以与欧几里得的《原本》相比美。欧拉在这部著作中评述了前辈数学家的发现，组织并清理了他们的论证，其论著之精妙，使得绝大部分前人著作都显得陈腐。除《引论》一书外，1755 年，欧拉出版了一卷本的《微分学原理》，1768—1774 年，又相继出版了三卷本的《积分学原理》，从此确定了数学分析的方向，并一直延续至今。

欧拉所有著作的论述都非常清楚易懂，并且，他所选用的数学符号，都是为了将他的意思表达得更加清晰明了，而不是含混不清。对于今天的读者来说，欧拉的数学著述堪称是最早一些具有现代数学意味的著述；这当然不仅是因为他使用了现代数学符号，而且，还因为他的影响十分深远，所有后来的数学家都采用了他的文体、符号和公式。并且，欧拉在写作时，想到了并非所有读者都能像他那样，具有惊人的学习数学的能力。欧拉不是以往那类数学家，他们虽然对问题有深邃的见解，但却无法把自己的意思传达给旁人。相反，他深深地喜爱教学。法国数学家孔多塞在谈到欧拉时有一句精辟的话：“ 他喜欢教诲他的学生，而不是从炫耀中求取满足。”这正是对一个人的高度赞美，因为欧拉如果喜欢炫耀，他的数学才干确实足以令任何人吃惊。

任何人在谈到欧拉的数学时，都会提到他的《全集》，这是一部 73 卷的文集。这部文集汇编了他一生分别用拉丁文、法文和德文撰著的 886 卷书和文章。他的著作数量极多，产出速度极快，甚至在他完全失明后也是如此，据说，他的著作直到他谢世后 47 年才出版完毕。

如前所述，欧拉并未将他的工作局限于纯数学领域。相反，他广泛涉猎声学、工程学、机械学、天文学等许多领域，甚至还写有三卷著作，专

门论述光学仪器，如望远镜和显微镜。虽然听来令人难以相信，但据估计，如果有人清点 18 世纪后 70 几年中的所有数学著作，那么，其中大约有三分之一出自李昂纳德·欧拉之手！

如果你在图书馆里，站在收藏欧拉著作的书架前，一个书架一个书架地看去，其著述洋洋大观，令人惊叹。这成千上万页文字，涉及从变分法、图论，到复变函数和微分方程等数学的所有分支，它们指引了数学各个领域的新方向。实际上，数学的每个分支都有欧拉创立的重要定理。因此，我们可以在几何学中找到欧拉三角，在拓扑学中找到欧拉示性函数，在图论中找到欧拉圆，还不要说使人目不暇接的欧拉常数、欧拉多项式、欧拉积分等等名目了。即使这些还只是故事的一半，因为人们一向记于他人名下的许多数学定理，实际上却是欧拉发现的，并深藏于他卷轶浩繁的著述中。有一则似假还真的趣话说道：

“……法则和定理的命名，常有喧宾夺主的事情，否则，有半数应署上欧拉的名字。”

1783 年 9 月 7 日，李昂纳德·欧拉溘然长逝。尽管他已双目失明，但直至他逝世前，他一直在进行数学研究。据说，在他生命的最后一天，他还在与他的孙子们一起游戏并讨论有关天王星的最新理论。对于欧拉来说，死神来得非常突然，用孔多塞的话说，“他终止了计算和生命”。欧拉被埋葬在他曾居住过的圣彼得堡，他曾断断续续地在那里度过了许多美好的时光。

伟大的定理：计算$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots\cdots+\frac{1}{k^2}+\cdots\cdots$的值

要从欧拉庞大的数学体系中找出一两个有代表性的定理是很困难的。我们之所以选定这一定理，是出于以下几个原因。第一，从历史上看，这一定理提出了一个十分重要且引起争论的命题。第二，这个定理是欧拉的早期成就之一，是他于 1734 年到圣彼得堡的第一年宣布的，这一定理在很大程度上巩固了他数学天才的名望。最后，这一定理不仅证明了欧拉解决前人的难题的才干，而且还证明他有能力将一个个别解法转变为一连串同样深刻和出人意料的解法。没有一个定理能够包容李昂纳德·欧拉的全部天才，但我们即将讨论的这一定理却清楚显示了他的数学才华。

这个问题就是我们在第八章结束时所提到的问题。回想一下，伯努利兄弟刚刚攻克了调和级数问题时，曾经探讨了级数$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots\cdots$。虽然他们知道这一级数的和一定小于 2，但他们却不能确定这个和的精确数值。显然，这一级数的计算不仅难倒了雅各布·伯努利和约翰·伯努利兄弟俩，而且也难倒了莱布尼兹，更不要说世界上的其他数学家了。

欧拉显然从他的老师约翰那里听说过这道难题。他曾谈到开始研究这个级数的时候，只是简单地把级数的项越来越多地加起来，希望能够找到级数的和。他这样一直计算到 20 位（在计算机时代之前，这种计算绝非易事），发现级数的和趋向于数字 1.6449。但遗憾的是，这个数字看起来似

乎很陌生。欧拉没有被吓倒，他继续研究，最后终于发现了解开这个谜的钥匙。他兴奋地写道，“……完全意想不到，我发现了基于π的……一个绝妙的公式。”

欧拉导出这个公式需要两个工具。其一是初等三角学中的所谓“正弦函数”。对这一重要数学概念（通常写作“sin x”）的充分讨论会使我们离题太远。无论如何，任何接触过三角学或微积分预备知识的人都肯定知道具有无限振荡性质的著名的正弦波。函数 f（x）=sinx 的图形见图 9.1，这一函数是欧拉思想的核心。

我们看到，正弦曲线与 x 轴相交处的点 x，其函数值等于零，

因而，当 x=0，±π，±2π，±3π，±4π，等等时，sinx=0。这种使 sinx 等于零的无穷多的 x 值反映了正弦函数周期性重复变化的特性。

关于正弦的许多知识，我们都可以从初等三角中学到。但是，如果我们在其中引入微积分的概念，我们就会得出下列公式：

$$\sin x = x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\frac{x^9}{9!}-\cdots\cdots$$

同样，我们没有必要细述这一公式的推导过程，但是，凡是学过微积分泰勒级数展开式的读者都会立刻认出这个公式。这一公式的重要性在于它为欧拉提供了一种将 sin x 表达为“无限长多项式”的方式。

对 sinx 的级数展开式，我们需要作一点儿说明。第一，分母中使用了阶乘符号，这种符号在一些数学分支中是很常见的。根据定义，3！表示 3×2×1=6；5！=5×4×3×2×1=120，等等。并且，这一 sinx 表达式还将永远达不到终点，随着 x 的指数按奇整数序列增大，分母表现为相应的阶乘，而正负号则一正一负交替出现。这就是我们所说将 sin x 写成一个无限长多项式的意思。这也是欧拉解开他的难题所需要的线索之一。

另一条线索不是出自三角学或微积分，而是出自单代数。由于正弦函数的泰勒级数展开式是一个无穷多项式，欧拉即转而研究普通的有限多项式，并将它推广到无穷多项式。

设 P(x)为 n 次多项式，其 n 个根为 x＝a，x=b，x=c，……，及 x=d；换言之，P(a)＝P(b)=P(c)=……=P(d)=0。我们再设 P(0)=1。然后，欧拉知道可以将 P（x）分解为如下 n 个一次项乘积的形式：

$$P(x)=\left(1-\frac{x}{a}\right)\left(1-\frac{x}{b}\right)\left(1-\frac{x}{c}\right)\cdots\cdots\left(1-\frac{x}{d}\right)$$

不妨考虑一下这一一般公式的合理性。我们可以用直接代入的方法得到

$$P(a)=\left(1-\frac{a}{a}\right)\left(1-\frac{a}{b}\right)\left(1-\frac{a}{c}\right)\cdots\cdots\left(1-\frac{a}{d}\right)=0$$

因为第一个因子恰好是 1-1=0。同样，

$$P(b)=\left(1-\frac{b}{a}\right)\left(1-\frac{b}{b}\right)\left(1-\frac{b}{c}\right)\cdots\cdots\left(1-\frac{b}{d}\right)=0$$

这次是因为第二个因子为 1-1=0。正如我们所期望的那样，P(x)的方程式非常清楚地表明，P(a)=P(b)=P(c)=……=P(d)=0。

但是，对 P(x)还有另外一个条件：我们要求 P(0)=1。幸好，从这里也可以得出我们的公式，因为

$$P(0)=\left(1-\frac{0}{a}\right)\left(1-\frac{0}{b}\right)\left(1-\frac{0}{c}\right)\cdots\cdots\left(1-\frac{0}{d}\right)$$

$$=(1)(1)(1)\cdots\cdots(1)=1$$

总之，

$$P(x)=\left(1-\frac{x}{a}\right)\left(1-\frac{x}{b}\right)\left(1-\frac{x}{c}\right)\cdots\cdots\left(1-\frac{x}{d}\right)$$

具有我们所寻求的性质。

例如，假设 P(x)是一个三次多项式，在这里，P(2)=P(3)=P(6)=0，并且，P(0)=1。然后，我们进行因式分解，得到

$$P(x)=\left(1-\frac{x}{2}\right)\left(1-\frac{x}{3}\right)\left(1-\frac{x}{6}\right)=1-x+\frac{11}{36}x^2=\frac{1}{36}x^3$$

我们可以很容易地验证这一三次方程符合我们所要求的条件。

欧拉仔细研究了这一方程后认为，同样的法则肯定也适用于“无穷多项式”。他像牛顿一样，也特别相信模式的推广，既然这一模式对有限多项式是正确的，为什么就不能适用于无穷多项式呢？现代数学家都知道，这种做法是十分危险的，而且，将适用于有限多项式的公式推广为适用于无穷多项式的公式，肯定会遇到巨大的困难。这种推广当然要比欧拉想象得更微妙，也需要更多的谨慎。也许是因为欧拉走运；也许是因为他那强有力的数学直觉。无论如何，他的努力没有落空。

这些初步结果似乎离我们最初求$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots\cdots$之和的问题太远。但欧拉用他超凡的洞察力作为纽带将全部零散的部件组合在一起。

定理 $1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots\cdots+\frac{1}{k^2}+\cdots\cdots\frac{\pi^2}{6}$

证明 欧拉首先引入函数

$$f(x)=1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\cdots\cdots$$

欧拉认为他有充分理由把 f(x)看成是无穷多项式，并且 f(0)=1（从直觉上说，这是显而易见的）。因此，可以利用上述方法，对这一函数方程作因式分解，以确定方程 f(x)=0 的根。为此，规定 x≠0，并得出

$$f(x)=x\left[\frac{1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\cdots\cdots}{x}\right]$$

$$=\frac{x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\frac{x^9}{9!}-\cdots\cdots}{x}$$

$$=\frac{\sin x}{x}$$ 根据$\sin x$的泰勒展开式

因此，只要x不等于0，解$f(x)=0$就等于解$\frac{\sin x}{x}=0$，而后者又可以（通

过简单的十字相乘方法）简化为解 sin x=0。我们在前面已看到，当正弦函数等于 0 时，x=0，x=±π，x=±2π，等等。当然，我们必须从 f（x）=0 的解中排除 x=0，因为我们已规定 f（0）=1。也就是说，f（x）=0 的解只是 x=±π，x=±2π，x=±3π，……

基于这些考虑，欧拉将 f(x)分解因式为：

$$f(x)=\left(1-\frac{x}{\pi}\right)\left(1-\frac{x}{-\pi}\right)\left(1-\frac{x}{2\pi}\right)\left(1-\frac{x}{-2\pi}\right)\left(1-\frac{x}{3\pi}\right)\left(1-\frac{x}{3\pi}\right)\cdots\cdots$$

$$=\left[\left(1-\frac{x}{\pi}\right)\left(1+\frac{x}{\pi}\right)\right]\left[\left(1-\frac{x}{2\pi}\right)\left(1+\frac{x}{2\pi}\right)\right]\left[\left(1-\frac{x}{3\pi}\right)\left(1+\frac{x}{3\pi}\right)\right]\cdots\cdots$$

也就是

$$1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\cdots\cdots$$

$$=\left[1-\frac{x^2}{\pi^2}\right]\left[1-\frac{x^2}{4\pi^2}\right]\left[1-\frac{x^2}{9\pi^2}\right]\left[1-\frac{x^2}{16\pi^2}\right]\cdots\cdots$$

我们称这一方程为核心方程。这是一个最非凡的方程，因为它使一个无穷和等于一个无穷乘积。也就是，最初确定 f(x)的无穷级数等于方程右边的无穷乘积。对于欧拉一类数学家来说，这是非常有启发性的。实际上，他现在即将完成他的证明，但许多读者也许还完全茫然不解。

欧拉所做的是设想“乘出”上述方程右边的无穷乘积，然后合并 x 的同类项。这样，第一项就将是所有 1 的乘积，当然，等于 1。为得到 x^2 项，我们就必须依次用剩余因子中的 x^2 项去乘所有的 1，而不是与其它因子相乘。这样，欧拉的“无穷乘法”问题就得到了下列方程：

$$1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\cdots\cdots$$

$$=\left[1-\frac{x^2}{\pi^2}\right]\left[1-\frac{x^2}{4\pi^2}\right]\left[1-\frac{x^2}{9\pi^2}\right]\left[1-\frac{x^2}{16\pi^2}\right]\cdots\cdots$$

$$=1-\left(\frac{1}{\pi^2}+\frac{1}{4\pi^2}-\frac{1}{9\pi^2}+\frac{1}{16\pi^2}+\cdots\cdots\right)x^2+(\cdots\cdots)x^4-\cdots\cdots$$

终于，迷雾开始散去。欧拉只要计算出无穷乘积，并得出两个相等的无穷和，那么，同指数的 x 项也就当然相等。请注意，两个级数的第一项都是 1。因而，两个级数中的 x^2 项，其系数也一定相等。即，

$$-\frac{1}{3!}=-\left(\frac{1}{\pi^2}+\frac{1}{4\pi^2}+\frac{1}{9\pi^2}+\frac{1}{16\pi^2}+\cdots\cdots\right)$$

然后，在方程两边同乘以-1，左边即得到 3！=6，而右边则提取公因

数 $\frac{1}{\pi^2}$，于是，欧拉得出

$$\frac{1}{6}=\frac{1}{\pi^2}\left(1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\cdots\right)$$

最后，应用十字相乘法，即得到了令人震惊的结果：

$$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}\cdots\cdots=\frac{\pi^2}{6}$$ 证讫。

这样，李昂纳德·欧拉就发现了其他数学家几十年未能发现的答案。完全肯定地说，$\frac{\pi^2}{6}$ 的数值就等于1.6449……，这一近似值恰好就是欧拉最初所推算的数值。我们还注意到，这一无穷和也恰如雅各布·伯努利于 1689 年所正确推断的那样，的确小于 2。

然而，在欧拉之前，人们对于这一级数的和恰好等于 π^2 的六分之一，完全一无所知。这是一个多么古怪的答案。由于数学本身的种种神秘原因，这一级数的和竟然产生了一个关于 π 的公式。因为 π 当然是与圆密切相关的，而 1、4、9、16 这些数字则与正方形密不可分，所以，很难想象这二者会联系在一起。甚至欧拉自己也对他的答案感到吃惊。他的公式过去是，至今依然是所有数学问题中最独特和最令人吃惊的公式之一。这一以极巧妙的方法推导出来的公式，其意外性使欧拉的证明成为第一流的经典定理。

后记

这一定理帮助奠定了李昂纳德·欧拉在整个数学界的崇高声望。这是一个无可争议的成功，许多稍许平庸一些的人肯定会因这些巨大成就而心满意足，不求进取，但欧拉却不然。欧拉数学的特点就是努力探索一切值得探索的问题。至于欧拉这一绝妙的求和公式，不过只是牛刀小试而已。

欧拉注意到，他在核心方程中所计算出的无穷乘积在 $x\neq0$ 的情况下，等于 $\frac{\sin x}{x}$。我们从图9.1的正弦函数图形中可以看到，当 $x=\frac{\pi}{2}$ 时，$\sin x$ 得到最大值1。如果我们将 $x=\frac{\pi}{2}$ 代入无穷乘积，我们发现，

$$\frac{\sin\frac{\pi}{2}}{\frac{\pi}{2}}=\left[1-\frac{\left(\frac{\pi}{2}\right)^2}{\pi^2}\right]\left[1-\frac{\left(\frac{\pi}{2}\right)^2}{4\pi^2}\right]\left[1-\frac{\left(\frac{\pi}{2}\right)^2}{9\pi^2}\right]\left[1-\frac{\left(\frac{\pi}{2}\right)^2}{16\pi^2}\right]\cdots\cdots$$

整理后，得

$$\frac{1}{\frac{\pi}{2}}=\left(1-\frac{1}{4}\right)\left(1-\frac{1}{6}\right)\left(1-\frac{1}{36}\right)\left(1-\frac{1}{64}\right)\cdots\cdots$$

或简化为

$$\frac{2}{\pi}=\left(\frac{3}{4}\right)\left(\frac{15}{16}\right)\left(\frac{35}{36}\right)\left(\frac{63}{64}\right)\cdots\cdots$$

通过取倒数并将方程右边因式分解，欧拉偶然发现了以下公式：

$$\frac{\pi}{2}=\frac{2\times2\times4\times4\times6\times6\times8\times8\cdots\cdots}{1\times3\times3\times5\times5\times7\times7\times9\cdots\cdots}$$

这一方程式使$\frac{\pi}{2}$等于一个很大的商，其分子是偶数的乘积，而分母则是奇数的乘积。实际上，这一方程式早已为英国数学家约翰·沃利斯（1616—1703 年）所知，他早在 1650 年就已用完全不同的方法推导出了这一公式。所以，欧拉并非发现了一个新公式，他只是在对无穷和与无穷乘积的新奇而相当有力的使用过程中再次发现了它。

但是，欧拉心中还有更多的打算。他认为，他所发现的计算级数$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\cdots$的方法是打开计算“病态”更明显的级数之谜的钥匙。例如，假设我们要求偶数平方的倒数和：

$$\frac{1}{4}+\frac{1}{16}+\frac{1}{36}+\frac{1}{64}+\frac{1}{100}+\cdots\cdots+\frac{1}{(2k)^2}+\cdots\cdots$$

欧拉首先提取公因数$\frac{1}{4}$，然后，应用“伟大定理”，得出

$$\frac{1}{4}+\frac{1}{16}+\frac{1}{36}+\frac{1}{64}+\frac{1}{100}+\cdots\cdots$$
$$=\frac{1}{4}\left(1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots\cdots\right)=\frac{1}{4}\left(\frac{\pi^2}{6}\right)=\frac{\pi^2}{24}$$

于是，欧拉也可以毫不费力地计算出所有奇数平方的倒数和，因为

$$1+\frac{1}{9}+\frac{1}{25}+\frac{1}{49}+\cdots\cdots=\left(1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\cdots\cdots\right)$$
$$-\left(\frac{1}{4}+\frac{1}{16}+\frac{1}{36}+\frac{1}{64}+\frac{1}{100}+\cdots\cdots\right)=\frac{\pi^2}{6}-\frac{\pi^2}{24}=\frac{\pi^2}{8}$$

欧拉显然对他的发现感到振奋，他再接再厉，提出了求整数四次方的倒数和问题：

$$1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\frac{1}{625}+\frac{1}{1296}+\cdots\cdots\frac{1}{k^4}+\cdots\cdots$$

他能解出这道题吗？

欧拉感到他应该回到核心方程上来，但这一次是要确定方程两边 x^4 项的系数。但是，怎样才能从核心方程右边的无穷乘积中找到 x^4 项呢？这并非一个平凡的小问题。在求解这一问题的过程中，欧拉再次得益于他对模式认辨敏锐的感觉和他关于任何有限乘积都可以推广到无限乘积的信念。

为了理解欧拉的推理过程，我们先举两个十分简单但却富有启发性的例子，说明他推导 x^4 项系数的方法。第一个例子是

$$(1-ax^2)(1-bx^2)=1-(a+b)x^2+abx^4$$
$$=1-(a+b)x^2+\frac{1}{2}[(a+b)^2-(a^2+b^2)]x^4$$

第二个例子是

$(1-ax^2)(1-bx^2)(1-cx^2)$

$=1-(a+b+c)x^2+(ab+ac+bc)x^4-(abc)x^6$

$$=1-(a+b+c)x^2+\frac{1}{2}[(a+b+c)^2-(a^2+b^2+c^2)]x^4-(abc)x^6$$

这些方程可以简单地通过乘出右边方括号中的各项来直接验算。

请注意，模式已经出现——即，将$(1-ax^2)$、$(1-bx^2)$、$(1-cx^2)$等一系列因式相乘后，x^4 的系数就等于$(a+b+c+\cdots\cdots)$之和的平方与平方和$(a^2+b^2+c^2+\cdots\cdots)$二者的差的一半。如果这一模式对于两个或三个这种因式的乘积是正确的话，那么，为什么不能推广到四个、五个，甚至无穷多个因式的乘积呢？欧拉回到核心方程，热心地推导出：

$$1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\cdots\cdots$$
$$=\left[1-\frac{x^2}{\pi^2}\right]\left[1-\frac{x^2}{4\pi^2}\right]\left[1-\frac{x^2}{9\pi^2}\right]\left[1-\frac{x^2}{16\pi^2}\right]\cdots\cdots$$

我们看到，在这里，$\frac{1}{\pi^2}$相当于a，$\frac{1}{4\pi^2}$相当于b，$\frac{1}{9\pi^2}$相当于c，等等。因而，应用我们对 x^4 系数的这一认识，就得到：

$$1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\frac{x^8}{9!}-\cdots\cdots$$
$$=1-\left(\frac{1}{\pi^2}+\frac{1}{4\pi^2}+\frac{1}{9\pi^2}+\cdots\cdots\right)x^2$$
$$+\frac{1}{2}\left[\left(\frac{1}{\pi^2}+\frac{1}{4\pi^2}+\frac{1}{9\pi^2}+\frac{1}{16\pi^2}+\cdots\cdots\right)^2\right.$$
$$\left.-\left(\frac{1}{\pi^4}+\frac{1}{16\pi^4}+\frac{1}{81\pi^4}+\frac{1}{256\pi^4}+\cdots\cdots\right)\right]x^4-\cdots\cdots$$

现在，欧拉开始同时考虑这一方程两边的 x^4 的系数。方程左边的 x^4 项的系数为$\frac{1}{5!}=\frac{1}{120}$。而右边的相应系数就较为复杂，但我们可以用代数方法进行整理，先提取π的同指数公因数，然后再应用上述伟大定理。所以，方程右边的 x^4 项的系数是

$$\frac{1}{2}\left[\left(\frac{1}{\pi^2}+\frac{1}{4\pi^2}+\frac{1}{9\pi^2}+\frac{1}{16\pi^2}+\cdots\cdots\right)^2\right.$$
$$\left.-\left(\frac{1}{\pi^4}+\frac{1}{16\pi^4}+\frac{1}{81\pi^4}+\frac{1}{256\pi^4}+\cdots\cdots\right)\right]$$
$$=\frac{1}{2}\left[\frac{1}{\pi^4}\left(1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots\cdots\right)^2\right.$$
$$\left.-\frac{1}{\pi^4}\left(1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\cdots\cdots\right)\right]$$
$$=\frac{1}{2\pi^4}\left[\left(\frac{\pi^2}{6}\right)^2-\left(1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\cdots\cdots\right)\right]$$
$$=\frac{1}{72}-\frac{1}{2\pi^4}\left(1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\cdots\cdots\right)$$

现在，答案就在眼前。欧拉将以上两个 x^4 的系数列为等式，并解出所得方程如下：

$$\frac{1}{120}=\frac{1}{72}-\frac{1}{2\pi^4}\left(1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\cdots\cdots\right)$$

所以

$$\frac{1}{2\pi^4}\left(1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\cdots\cdots\right)=\frac{1}{72}-\frac{1}{120}=\frac{1}{180}$$

最后，通过十字相乘，就得出了欧拉的公式：

$$1+\frac{1}{16}+\frac{1}{81}+\frac{1}{256}+\cdots\cdots+\frac{1}{k^4}+\cdots\cdots=\frac{2\pi^4}{180}=\frac{\pi^4}{90}$$

欧拉发现了一个真正奇特的结果，他将完全四次方的倒数与π的四次方联系在了一起。然后，他像一个孩子发现了新玩具一样，兴高采烈地应用他非凡的方法，去求更奇特级数的和，如：

$$1+\frac{1}{64}+\frac{1}{729}+\frac{1}{4096}+\cdots\cdots+\frac{1}{k^6}+\cdots\cdots=\frac{\pi^6}{945}$$

他不断对偶次幂级数进行推算，并得出了

$$1+\frac{1}{2^{26}}+\frac{1}{3^{26}}+\cdots\cdots\frac{1}{k^{26}}+\cdots\cdots=\frac{1315862}{11094481976030578125}\pi^{26}$$

最后，甚至连欧拉自己也对此感到厌倦了。无须赘言，历史上没有一个人曾踏上这一数学之旅。从实践的观点来看，无论事情有多么琐碎，但它们确实是人类知识的一大进步，是对有关整数乘方倒数与最重要常数π之间关系的发现，而这些关系，以前人们从未想到过。

至此，人们会立即想到一个问题：整数奇次幂的倒数和又将如何呢？例如，我们能否计算出无穷级数

$$1+\frac{1}{2^3}+\frac{1}{3^3}+\frac{1}{4^3}+\cdots\cdots=1+\frac{1}{8}+\frac{1}{27}+\frac{1}{64}+\cdots\cdots$$

对此，甚至欧拉也缄默不语，的确，过去 200 年的数学研究对这些奇次幂级数的认识进展甚少。但是，我们可以很容易地推测，整数奇次幂的

倒数和一定对应于分数 $\frac{p}{q}$ 的 $\left(\frac{p}{q}\right)\pi^3$ 的形式，但时至今日，仍没有一个人能够肯定这一推测是否正确。

今天，我们认识到，欧拉关于无穷级数的推理并不十分严格。他相信有限级数所产生的模式和公式可以自动推广到无穷级数，这与其说是科学，不如说是一种信念。其后的数学家提供了大量的例证，证明了欧拉的这种推广是十分愚蠢的。总之，欧拉未能为他的推理提供充分的逻辑依据。然而，这些批评丝毫无损于他的声望。即使他推论无穷级数的方法还十分幼稚，但他所有这些奇妙的级数和都已通过了今天高标准的严密逻辑证明。

这些成就在欧拉的 70 余卷著作中只占了几页。下一章，我们将讨论欧拉在一个完全不同的数学分支——数论领域中的辉煌贡献。

第十章
欧拉对数论的贡献
（1736 年）

费马的遗产

我们已经了解了欧拉在计算复杂的无穷级数方面的成就。他的这些研究属于称作“解析法”的数学分支，他的发现在这一数学分支中显得特别重要和意义深远。但是，如果不介绍他在数论领域中的贡献，就不免是一大疏忽。欧拉在数论这一数学分支中也是当行出色的。我们前面曾讲到过一些有关数论的问题，在第三章，我们介绍了欧几里得关于素数无穷性的巧妙证明；我们还在第七章里介绍了费马关于数论的卓有见地的评论和猜想。如前所述，费马没有能够提供证明，而且，从费马到欧拉的 100 年间，数学界在证明费马猜想方面进展甚微。造成这种停滞的原因很多，一方面是由于 17 世纪末对微积分的新发现垄断了数学研究的方向,另一方面是由于数论对任何实际问题缺乏实用性，还有一部分原因是因为费马的猜想对于许多数学家来说，难度太大了。

欧拉对数论的兴趣是由克里斯蒂安·哥德巴赫引起的。关于哥德巴赫猜想，我们在第三章的后记中已作过简要介绍。哥德巴赫被数论问题深深地吸引住了，但是，他的热情远远超过了他的才能。他与欧拉一直保持着密切的通信联系，最初，是哥德巴赫告诉欧拉许多有关费马未证明的猜想，并引起了欧拉的注意。开始，欧拉似乎无意研究这些问题，但是，由于他自己无止境的好奇心和哥德巴赫的坚持，欧拉终于涉足其间。不久，他就被数论，特别是被费马一系列未证明的猜想深深地迷住了。正如现代作家兼数学家安德烈·韦尔所述，“……在欧拉（有关数论）的著作中，有相当一部分旨在证明费马的猜想。”在此之前，欧拉的数论著作在他的《全集》中已占了整整四大卷。人们认为，在他的科学生涯中，即使没有其他成就，这四卷著作也足以使他跻身于历史上最伟大的数学家之列。

例如，费马曾推测，某些素数可以写成两个完全平方数之和，欧拉对此作出了证明。显然，除 2 以外，其它所有素数都是奇数。当然，如果我们用 4 去除一个大于 4 的奇数，我们一定会得到余数 1 或 3（因为 4 的倍数或 4 的倍数加 2 是偶数）。我们可以更简明地说，如果 $p>2$ 是素数，那么，或则 $p=4k+1$，或则 $p=4k+3$（k 是整数）。1640 年，费马曾猜想，第一种形式的素数（即 4 的倍数加 1）可以并且只能以一种方式写成两个完全平方数之和的形式，而形如 $4k+3$ 的素数则无论以什么方式都不能写成两个完全平方数之和。

这是一个独特的定理。例如，素数 $193=(4\times48)+1$ 可以以一种唯一方式写成两个平方数之和。对本例，我们可以很容易地证明，$193=144+49=12^2+7^2$，而其他任何形式的平方和都不能等于 193。另一方面，素数 $199=(4\times49)+3$ 绝对无法写成两个平方数之和的形式，这同样可以通过列出所有可能的形式来证明其不可能性。因此，我们在这两种形式的（奇）素数之间，就其表达为两个平方之和而言，发现了根本的差别。这是一个无法预料或凭直觉预测的性质。但欧拉在 1747 年对此作出了证

明。

我们再来看另一个例子。我们在第三章后记中曾讨论过所有偶完全数的问题，欧拉对此也表现出了他的数论天才。与这个问题有关的是他对所谓亲和数的研究。亲和数是一对具有下列性质的数字：一个数字的所有因数之和恰好等于第二个数字，而第二个数字所有因数之和也同样等于第一个数字。亲和数早在古代就引起了数学家的兴趣，他们认为亲和数具有神秘的“超数学”色彩。即使在现代，亲和数也因其独特的互逆性质游弋在数字学的伪科学中。

古希腊人已知道数字 220 和 284 是亲和数。即，220 的所有因数是 1、2、4、5、10、11、20、22、44、55 和 110，这些因数加起来恰好等于 284；同样，284 的所有因数是 1、2、4、71 和 142，它们加起来等于 220。但遗憾的是，当时的数字学家们还不知道有其他的亲和数，直至 1636 年，费马才证明出 17，296 和 18，416 构成了第二对亲和数。（实际上，这对亲和数早已为阿拉伯数学家班纳（1256—1321 年）所发现，比费马早 300 多年，但是，在费马时代，西方人还不知道这一对亲和数的存在。） 1638 年，笛卡儿或许是为了与费马争胜，骄傲地宣布他发现了第三对亲和数：9，363，584 和 9，437，056。

在欧拉开始研究这个问题之前的一百年间，亲和数的研究一直停滞不前。1747 年至 1750 年期间，欧拉发现了 122，265 和 139，815 以及其他 57 对亲和数，这样，他独自一人就使世界已知亲和数增加了近 20 倍！欧拉之所以能够取得这样的成果，是因为他找到了生成亲和数的方法，并用这种方法生成了亲和数。

费马最重要的猜想之一见于他 1640 年的另一封信中。他在信中说，如果 a 是任意整数，p 是与 a 互质的素数，那么，p 就一定是数字 $a^{p-1}-1$ 的因数。费马按照他令人厌烦的习惯，宣称他已经发现了这一奇特现象的证明，但却没有写在信中。并且，他告诉他的收信人，“如果不是怕这个证明太长的话，我就写给你了。”

此后，这一性质便以“费马小定理”而知名。例如，素数 p=5 和数字 α=8，定理宣称，5 可以整除 $8^4-1=4096-1=4095$；显然，这是正确的。同样，如果素数 p=7，数字α=17，根据费马定理，7 能够整除 $17^6-1=24，137，569-1=24，137，568$；这个数字虽然很不明显，但却同样是正确的。

费马是如何作出证明的，我们只能去猜想了。直到 1736 年，才有欧拉提供了一个完整的证明。我们后面将要讨论欧拉的证明，但在此之前，我们应先介绍一下欧拉作出证明所需要的数论依据：

(A) 如果素数 p 能够整除 a×b×c×……×d 的乘积，那么，p 就一定能够整除 a，b，c，……，d 这些因数中的（至少）一个因数。用通俗的话讲，就是，如果一个素数能够整除一个乘积，那么，它就一定能够整除其中的一个因数。正如我们在第三章中所述，欧几里得早在欧拉之前二千年就已在其《原本》的命题Ⅶ.30 中对此作出了证明。

(B)如果 p 是素数，a 是任意整数，则下式

$$a^{p-1}+\frac{p-1}{2\times1}a^{p-2}+\frac{(p-1)(p-2)}{3\times2\times1}a^{p-3}+\cdots\cdots+a$$

也表示一个整数。

我们将不去证明这个论断，但要通过一两个例子来验证其正确。例如，

如果 a=13，p=7，那么，我们发现，

$$13^{6}+\frac{6}{2\times1}13^{5}+\frac{6\times5}{3\times2\times1}13^{4}+\frac{6\times5\times4}{4\times3\times2\times1}13^{3}$$

$$+\frac{6\times5\times4\times3}{5\times4\times3\times2\times1}13^{2}+\frac{6\times5\times4\times3\times2}{6\times5\times4\times3\times2\times1}13=4826809$$

$$+1113879+142805+10985+507+13=6094998$$

的确是一个整数，因为在原算式中所有貌似分数都约掉了，我们只剩下求整数的和。当然，这种约消不一定必然存在。实际上，如果我们在 p 的位置采用一个非素数，我们就会遇到麻烦。例如，如果 a=13，p=4，我们得到

$$13^{3}+\frac{3}{2\times1}13^{2}+\frac{3\times2}{3\times2\times1}13=2197+253.5+13=2463.5$$

这当然不是整数。所以，只有 p 是素数，才能保证这一算式得到整数值。

欧拉需要的最后一件数学武器是应用于 $(a+1)^p$ 的二项式定理。幸好，他在牛顿的著作中读到过这个定理，所以，他已经准备停当。我们将分四步来探讨他的证明，每一步都直接推导出下一步，最后一步将以费马小定理结束：

定理 如果 p 是素数，a 是任意整数，则 p 可以整除 $(a+1)^p-(a^p+1)$。

证明 应用二项式定理，展开第一个表达式，得到

$$(a+1)^{p}=\left[a^{p}+pa^{p-1}+\frac{p(p-1)}{2\times1}a^{p-2}\right.$$

$$\left.+\frac{p(p-1)(p-2)}{3\times2\times1}a^{p-3}+\cdots\cdots pa+1\right]$$

我们将这一展开式代入 $(a+1)^p-(a^p+1)$，然后，合并同类项，并提取公因数 p，即得到

$$(a+1)^{p}-(a^{p}+1)$$

$$=\left[a^{p}+pa^{p-1}+\frac{p(p-1)}{2\times1}a^{p-2}+\frac{p(p-1)(p-2)}{3\times2\times1}a^{p-3}+\cdots\cdots pa+1\right]$$

$$-(a^{p}+1)$$

$$=pa^{p-1}+\frac{p(p-1)}{2\times1}a^{p-2}+\frac{p(p-1)(p-2)}{3\times2\times1}a^{p-3}+\cdots\cdots+pa$$

$$=p\left[a^{p-1}+\frac{p-1}{2\times1}a^{p-2}+\frac{(p-1)(p-2)}{3\times2\times1}a^{p-3}+\cdots\cdots+a\right]$$

根据上述(B)，我们知道方括号中的项是一个整数。因而，我们证明了 $(a+1)^p-(a^p+1)$ 可以分解因式为素数 p 与一个整数的乘积。换言之，如定理所称，p 可以整除 $(a+1)^p-(a^p+1)$。　　证讫。

我们利用这一结果即可直接导出第二个定理。

定理　如果 p 是素数，并且，p 可以整除 a^p-a，那么，p 也可以整除 $(a+1)^p-(a+1)$。

证明　前一个定理保证了 p 可以整除 $(a+1)^p-(a^p+1)$，并且，我们已知 p 也可以整除 a^p-a。所以，p 显然可以整除这两者的和：

$$[(a+1)^p-(a^p+1)]+[a^p-a]=(a+1)^p-a^p-1+a^p-a$$
$$=(a+1)^p-(a+1)$$

而这正是我们所要证明的。　　证讫。

上面的结论为欧拉提供了证明费马小定理的钥匙，这一证明过程被称作“数学归纳法”。归纳法是适于包含整数在内的一些命题的完美的证明技巧，这种方法利用了整数一个紧跟一个的“阶梯”牲质。归纳法证明很像攀登一个（非常高的）梯子。我们最初的工作就是要踏上梯子的第一级。然后，我们必然能从第一级登上第二级。这两步完成后。我们需要的是从第二级登上第三级，然后再从第三级登上第四级。如果我们掌握了从一级登上更高一级阶梯的方法，那么，这个梯子就属于我们了！我们确信，没有我们达不到的阶梯。欧拉应用归纳法作了如下证明：

定理　如果 p 是素数，a 是任意整数，那么，p 能够整除 a^p-a。

证明　因为这一命题涉及到所有整数，所以，欧拉开始先证明第一个整数，即 a=1。但是这种情况极为简单，因为 $a^p-a=1^p-1=1-1=0$，p 当然可以整除 0（实际上，任何正整数都能够整除 0）。这使欧拉踏到了梯子上。

现在，欧拉将 a=1（即我们刚刚证明的 p 是 1^p-1 的因数）应用于前面的定理，据此，欧拉就可以推断，p 同样也能够整除

$$(1+1)^p-(1+1)=2^p-2$$

换句话说，欧拉证明了 a=2。如果我通过前面的命题再循环一次，我们就发现，这表明，p 能够整除

$$(2+1)^p-(2+1)=3^p-3$$

我们只要不断重复这个过程，就会看到 p 能够整除 4^p-4、5^p-5，等等。这样，欧拉就像从梯子的一级不断爬向更高一级那样，能够一直爬向整数梯子的顶端，从而保证了对于任意整数 a 来说，p 是 a^p-a 的因数。　证讫。

最后，欧拉准备证明费马小定理。由于已完成了上述艰苦的准备工作，所以，他最后一步证明极为轻松：

费马小定理　如果 p 是素数，而 a 是与 p 互质的整数，那么，p 能够整除 $a^{p-1}-1$。

证明　我们已证明，p 能够整除

$$a^p-a=a[a^{p-1}-1]$$

根据上述(A)，因为 p 是素数，p 就一定能够整除 a 或 $a^{p-1}-1$（或两者）。但是，我们已假设 p 不能整除 a，因而我们推断，p 能够整除后者，即，p 能够整除 $a^{p-1}-1$。这就是费马小定理。　　证讫。

欧拉的论证是一颗明珠。他需要的仅仅是一些比较简单的概念；他溶入了归纳法这种关于整数的典型证明方法；并运用了远及欧几里得，近自

二项式定理的命题。对于这些知识，他大量注入了自己的天才，这样就出现了费马以前提出但未能证明的费马小定理的第一个证明。

顺便说几句题外的话，令人惊奇的是，欧拉的这一命题最近被应用于一个实际问题——即设计某些高度复杂的密码系统，以发送机密信息。纯数学抽象定理亦有其非常实际的用途，在这方面，这不是第一例，当然也不是最后一例。

伟大的定理：欧拉对费马猜想的反驳

就我们本章的目的来说，前面的论证只是序曲。费马／欧拉的另一个命题将作为本章的伟大定理。毫不奇怪，正是欧拉的笔友哥德巴赫的信引起了欧拉的兴趣。在 1729 年 12 月 1 日的一封信中，哥德巴赫带着几分天真问道，“费马认为所有形如$2^{2^n}+1$的数字都是素数，你知道这个问题吗？他说他没能作出证明；据我所知，也没有其他任何人对这个问题作出过证明。”

费马声称发现了一个始终能生成素数的公式。显然，就 n 的最初几个值而言，他的公式是对的。也就是说，如果$n=1$，$2^{2^1}=2^2+1=5$是素数；如果$n=2$，则$2^{2^2}+1=2^4+1=17$是素数；同样，$n=3$和$n=4$生成素数 $2^8+1=257$ 和 $2^{16}+1=65537$。按序列，下一个数字 n=5 就生成了一个巨大的数字

$$2^{2^5}+1=2^{32}+1=4294967297$$

费马同样认为这是一个素数。如果沿着费马的思路，从直觉上没有理由怀疑他的推断。另一方面，任何数学家如果想要否定费马的猜想，就必须要找到一种方法，以将这一 10 位数字分解为两个较小的因数；而这种研究可能需要几个月的时间，当然，如果费马对这个数字的素数性的推断是正确的话，则这种探索将是徒劳无益的。总之，我们有种种理由接受费马的推测，转而去忙别的事情。

但这不是欧拉的性格。他开始对数字 4，294，967，297 进行研究，最后，欧拉终于成功地分解出这个数字的因数。费马的猜想是错误的。无需赘言，欧拉发现这个数字的因数并非偶然。他就像侦探一样，首先从一个案件的真正嫌疑犯中排除无辜的旁观者。按照这种思路，欧拉设计了一个非常巧妙的检验方法，从一开始就排除掉所有无关的数字，只留下 4，294，967，297 的几个潜在因数。他的非凡观察力使摆在他面前的任务变得格外简单。

欧拉首先提出一个偶数 a（但如果能够知道真相的话，他心里实际想的是 a=2）和一个素数 p，且 p 不是 a 的因数。然后，他希望能确定对素数 p 的限制，看其能否分别整除 $a+1$、a^2+1、a^4+1，或其一般式$a^{2^n}+1$。依照费马猜想的性质，欧拉特别注意$n=5$的情况。也就是说，他能否发现 $a^{32}+1$ 的素因数呢？

命运似乎对费马开了一个不大不小的玩笑，欧拉用以否定费马猜想$2^{2^n}+1$的主要定理不是别的，恰恰正是费马小定理本身。换句话说，费马自己种下了埋葬自己的种子。的确，随着我们介绍欧拉推导下述定理的

过程，我们不能不承认，费马小定理起了关键性的作用。

定理　A 设 a 为偶数，p 为素数，且 p 与 a 互质，但却能够整除 a+1。那么，对于某一整数 k，p=2k＋1。

证明　这是一个非常简单的定理。如果 a 是偶数，那么，a＋1 就是奇数。因为我们假设 p 能够整除奇数 a＋1，所以，p 自身也一定是奇数。因而，p-1 是偶数，并且，对于某一整数 k 来说，p-1=2k，也就是 p=2k＋1。证讫。

我们来看一个具体数例。如果我们先设偶数 a=20，那么，a＋1=21，并且，21 的两个素因数（即 3 和 7）都符合 2k＋1 的形式。

下一步是更具挑战性的一步：

定理 B　设 a 为偶数，p 为素数，且 p 与 a 互质，但却能够整除 a^2+1。那么，对于某一整数 k 来说，p=4k＋1。

证明　因为 a 是偶数，所以 a^2 也是偶数。根据定理 A，我们知道，a^2+1 的任何素因数（特别是数字 p）都一定是奇数。也就是说，P 等于 2 的倍数加 1。

但是，如果我们用 4 去除 p，结果如何呢？显然，任何奇数都一定等于 4 的倍数加 1 或者 4 的倍数加 3。使用符号，p 可以表示为 4k＋1 或 4k＋3 的形式。

欧拉想消除后一种可能性，为了造成最后的矛盾，他必须先假定 p=4k＋3，其中 k 为某一整数。由于定理设 p 与 a 互质，所以，根据费马小定理，p 能够整除

$a^{p-1}-1=a^{(4k+3)-1}-1=a^{4k+2}-1$

另一方面，定理给出 p 是 a^2+1 的因数，所以，p 也是下列乘积的因数：

$(a^2+1)(a^{4k}-a^{4k-2}+a^{4k-4}-\cdots\cdots+a^4-a^2+1)$我们可以用代数方法对这一乘式进行计算，通过乘出上式，并合并同类项，就可以将这一复杂的乘积简化为 $a^{4k+2}+1$ 的形式。

现在，我们可以断定，p 既能够整除 $a^{4k+2}+1$，也能够整除 $a^{4k+2}-1$。所以，p 一定能够整除这两者的差

$$(a^{4+2}+1)-(a^{4k+2}-1)=2$$

但是，这是一个明显的悖论，因为奇素数 p 不能整除 2。它表明，p 不能象我们在开始时所假设的那样具有 4k＋3 的形式。由于只剩下了一种选择，所以，我们可以断定，对于某一整数 k 来说，p 一定等于 4k＋1。证讫。

与前面一样，我们现在来举几个具体数例。如果 a=12，那么，$a^2+1=144+1=145=5\times29$，5 和 29 都是 4k＋1 形式（即 4 的倍数加 1）的素数。同样，如果 a=68，则 $a^2+1=4625=5\times5\times5\times37$，其中的每一个素因数都等于 4 的倍数加 1。

接着，欧拉提出了数字$a^{2^2}+1=a^4+1$的素因数问题。

定理 C　设 a 为偶数，p 为素数，且 p 与 a 互质，但 p 能够整除 a^4+1。

那么，对于某一整数 k 来说，p=8k+1。

证明　首先说明，$a^4+1=(a^2)^2+1$。所以，我们可以应用定理 B，将 p 写成 4 的倍数加 1 的形式。据此，欧拉提出，如果不用 4，而用 8 去除 p，结果又会如何呢？起初，我们可能会遇到 8 种可能性：

p=8k（即，p 是 8 的倍数）

p=8k＋1（即，p 等于 8 的倍数加 1）

p=8k＋2（即，p 等于 8 的倍数加 2）

p=8k＋3（即，p 等于 8 的倍数加 3）

p=8k＋4（即，p 等于 8 的倍数加 4）

p=8k＋5（即，p 等于 8 的倍数加 5）

p=8k＋6（即，p 等于 8 的倍数加 6）

p=8k＋7（即，p 等于 8 的倍数加 7）

幸运的是（而这正是欧拉分析的核心），我们可以消除其中 p 的某些可能形式。首先，我们知道，p 一定是奇数（因为 p 是奇数 a^4+1 的因数），所以，p 不可能呈现 8k、8k＋2、8k＋ 4 或 8k＋6 的形式，因为它们显然全都是偶数。

并且，8k+3=4(2k)＋3 等于 4 的倍数加 3，根据定理 B，我们知道， p 不可能具有这种形式。同样，数字　8k ＋7=8k＋4＋ 3= 4(2k+1)＋3 也等于 4 的倍数加 3，所以，也不在考虑之列，应予以消除。

这样，a^4+1 的素因数就只剩下了　8k＋1 和 8k＋5 两种可能形式。但是，欧拉按下述方法成功地排除了后者：

为了造成矛盾，必须先假定 p=8k＋5，其中 k 为某一整数。那么，由于 p 与 a 互质，所以，根据费马小定理，p 能够整除

$$a^{p-1}-1=a^{(8k+5)-1}-1=a^{8k+4}-1$$

另一方面，由于 p 能够整除 a^4+1，所以，p 也肯定能够整除

$$(a^4+1)(a^{8k}-a^{8k-4}+a^{8k-8}-a^{8k-12}+\cdots\cdots+a^8-a^4+1)$$

这一乘积可以用代数方法简化为 $a^{8k-4}+1$。但是，如果 p 既是 $a^{8k-4}+l$ 的因数，又是 $a^{8k-4}-1$ 的因数，那么，p 也就应该能够整除它们的差

$$(a^{8k-4}+1)-(a^{8k-4}-1)=2$$

这样，就出现了矛盾，因为 p 是奇素数。所以，我们看到，P 不可能有 8k＋5 的形式，因而，正如定理所断定的那样，p 的唯一可能形式只能是 8k+1。证讫。

我们再来举一个简单的例子。如果偶数　a=8。那么，$a^4+1=4097$，这个数字可以分解为　17×241，而　17 和　241 都可以分解为　8 的倍数加 1 的形式。

据此，欧拉证明了更多同样形式的情况，但是，为了我们的目的，我们应将这一模式整理一下，使之条理更加清晰。我们可以概

括前面的所有工作如下。若 a 为偶数，p 为素数，那么，

如果 p 能够整除 a＋1，则 p 为 2k＋1 的形式（定理 A）

如果 p 能够整除 a^2+1，则 p 为 4k＋1 的形式（定理 B）

如果 p 能够整除 a^4+1，则 p 为 8k＋1 的形式（定理 C）

如果 p 能够整除 a^8+1，则 p 为 16k＋1 的形式

如果 p 能够整除 $a^{16}+1$，则 p 为 32k＋1 的形式

如果 p 能够整除 $a^{32}+1$，则 p 为 64k＋1 的形式

总之，如果p能够整除$a^{2^n}+1$，那么，对于某一整数k来说，$p=(2^{n+1})k+1$。

终于，我们可以回到费马关于 $2^{32}+1$ 的素数性的猜想。然而，我们是带着一种关于这一数字可能具有素因数的特别信息回到这个问题上来的。欧拉不是盲目地探索这个数字的素因数，相反，他很快地触及到问题的核心。

定理 $2^{32}+1$ 不是素数。

证明 由于 a=2 当然是偶数，前面的探索告诉我们，$2^{32}+1$ 的任何素因数都一定为 p=64k+1（k 为整数）的形式。因而，我们可以一个个地检验这些极特殊的数字，看它们是否(1)是素数，(2)能整除 4，294，967，297（欧拉用长除法检验后者，而现代读者则可望使用计算机）：

如果 k=1，64k＋1=65，这当然也不是素数，因而无须检验；

如果 k=2，64k＋1=129=3×43，当然也不是素数；

如果 k=3，64k＋1=193，这是一个素数，但却不能整除 $2^{32}+1$；

如果 k=4， 64k＋1=257，这是一个素数，但同样不能整除 $2^{32}+1$；

如果 k=5，64k＋1=321=3×107，这不是素数，无须检验；

如果 k=6，64k＋1=385=5×7×11，也可以略过；

如果 k=7，64k＋1=449，这是一个不能整除 $2^{32}+1$ 的素数；

如果 k=8，64k＋1=513=3×3×3×19，略过；

如果 k=9，64k＋1=577，是一个素数，但却不是 $2^{32}+1$ 的因数；

但是，当欧拉试算 k=10 的时候，他就击中了要害。在这种情况下，p=（64×10）＋1=641，这是一个素数，而且，看哪！恰好能够整除费马的数字。即，

$$2^{32}+1=4294967297=641\times 6700417$$

欧拉仅仅试算了 5 个数字，就发现了因数 641，意义十分深远。他通过谨慎地排除 $2^{32}+1$ 的可能性因数的方法，穷竭了可疑数字，使他的任务变得几乎轻而易举。这是一个数学检测的辉煌范例。

关于欧拉上述定理，还有一则有趣的补遗，即 4k＋1 形式的素数只能分解为一种形式的两个平方数之和。首先，我们来看，

$$2^{32}+1=(2^2)(2^{30})+1=4(1073741824)+1$$

所以，$2^{32}+1$ 的确具有 4k＋1 的形式。我们可以直接用数字来检验，

$$2^{32}+1=4294967297=4294967296+1=65536^2+1^2$$

同时，

$$\begin{aligned}2^{32}+1&=4294967297=418161601+3876805696\\&=20449^2+62264^2\end{aligned}$$

这样，我们就以两种不同的方式，将数字 $2^{32}+1$ 分解为两个完全平方数之和。根据欧拉的准则，这证明了 $2^{32}+1$ 不可能是素数，因为 4k+1 形式的素数只能有一种分解方式。因此，虽然我们没有找到确定的因数，但我们仍然可以非常巧妙地间接证明，这一巨大的数字是合数。

费马关于对所有整数n来说，$2^{2^n}+1$是素数的猜想，在n = 5时是

错误的。但是，如果取更大的 n 值，结果又会如何呢？例如，如果 n=6，我们得出

$$2^{2^6}+1=2^{64}+1=18446744073709551617$$

这个数字是能够被素数 p=274，177 除尽的。毫不奇怪，按照欧拉发现的模式，p 具有 128k＋1 的形式；即，p=（128×2142）＋1。费马又错了。

接下来的情况更糟糕。1905 年，一个非常复杂的论证表明，费马的下一个数字——$2^{2^7}+1=2^{128}+1$，也是合数，但是，这一证明未能提供这一巨大数字的具体因数。直到 1971 年，人们才发现了这个长达 17 位的因数。

到1988年时，数学家已知道，$2^{2^8}+1$，$2^{2^9}+1$，……，$2^{2^{21}}+1$都是合数。显然，费马关于$2^{2^n}+1$形式的数字的总括性猜想，可以说，只不过是以偏盖全面已。尽管他宣称所有这些数字都是素数，但当 n≥5 时，却从来也没有发现过这种形式的素数。实际上，现在许多数学家都在猜测，除了费马已发现的当 n=1、2、3、4 时的四个素数以外，根本就没有这种形式的素数存在。这样，费马猜想就不仅是错误的，而且是大错特错了。

至此，我们可以就我们对欧拉数论的简要评述作一个总结。如前所述，本章的这些定理最直接地表明了欧拉在数论领域的巨大影响。诚然，他是站在天才的前辈、特别是站在费马的肩膀上。但是，欧拉的研究，不可估量地丰富了这一数学分支，并使他自己跻身于第一流的数论学家之列。

后记

当欧拉逝世的时候，卡尔·弗里德里希·高斯刚刚 6 岁。然而，这个德国男孩超常的智力已经给其长者留下极深的印象。几十年后，他继承欧拉的衣钵，成为世界上最优秀的数学家。

我们在第三章曾介绍过高斯最初期的成就,他在 1796 年发现可以用圆规和直尺作出正 17 边形的图形。这一证明在数学界引起了一场轰动，因为自古以来，没有任何人想到过有可能作出这一图形。我们还是让年轻的高斯自己来加以说明：

“每一个略通几何的人都清楚地知道，许多正多边形，即三角形、五边形、15 边形以及它们的 2^n 倍的正多边形，都可以用几何方法作出。远在欧几里得时代，人们就已懂得这一点，并且，从那时起，人们似乎就已相信，初等几何的范围是不可能扩大的……然而，我认为，除了这些常规多边形之外，更非凡的是可以同样用几何方法作出的其他一些图形，如正 17 边形。”

高斯虽然当时尚不足 20 岁，但在正多边形的几何作图方面，却比欧几里得、阿基米德、牛顿或其他任何人都看得深远。

然而，高斯所做的，还不仅仅是证明了正 17 边形几何作图的可能性，他还证明了如果 N为形如$2^{2^n}+1$的素数，则正N边形可用几何方法作图。当然，我们已知道，这种形式的素数正是费马所称的素数。由于某种原因，这一数论问题与几何作图有着内在的联系。正如数学史上有时出现的那样，一个数学分支（本例为数论）中的发现和研究会对另一看来无关的分

支（正多边形的几何作图）产生深刻影响。当然，这里的关键是“看来无关”。但实际上，高斯的研究表明，这两者之间确有着不可否认的关系。

他的发现不仅向世界揭示了正 17 边形几何作图的可能性（因为 $2^{2^2}+1=17$ 是素数），而且还证明了正 $2^{2^3}+1=257$ 边形，甚至巨大的 2^{2^4} +1=65537 边形也可以用几何方法作出！当然，这些作图绝对没有任何实际意义，但它们的存在再次表明，在我们所熟悉的欧氏几何下面，隐藏着一个奇怪的、令人意想不到的神密世界。高斯自己对这一发现也颇感自豪，甚至在他毕生取得非凡的数学成就之后,他还是要求将一个正 17 边形铭刻在他的墓碑上。（令人遗憾的是，这点没有做到。）

卡尔·弗里德里希·高斯于 1777 年出生在德国的不伦瑞克，他很小时就显示出聪明过人的迹象。三岁时，这个还没有桌子高的小家伙就能够核算他父亲的帐目，偶尔还能够改正其中的错误。有一个高斯在小学时的脍炙人口的故事，讲的是，一次，他的一个老师显然是上课太累了，想稍事休息，就要求全班同学静静地计算前一百个整数的和。这些孩子们无疑得费一会儿功夫。但是，老师刚刚把题目讲完，卡尔就站起来，把答案放在了老师的桌上，而这时，其他同学几乎刚刚计算出“ 1＋ 2+ 3+4+ 5= 15 ”。面对这一意想不到的情况，大家可以想象老师脸上那种交织着怀疑与沮丧的表情，但当他瞥了一眼高斯的答案时，却发现答案完全正确。高斯是怎样计算出来的呢？

首先，这不是魔法，也不是那种能够以闪电般的速度累加一百个数字的能力。确切地说，是高斯甚至在如此小小年纪就已表现出来的敏锐的洞察力，这种洞察力贯穿了他的一生。据说，他只是想象他所求的和（我们用 S 表示）可以同时写成递升次序和递降次序：

$$S=1+2+3+4+\cdots\cdots+98+99+100$$
$$S=100+99+98+97+\cdots\cdots+3+2+1$$

高斯没有横向去加这两行数字，而是竖向将各列相加。由于每一列的和都恰好等于 101，这样，他就得到

$$2S=101+101+101+\cdots\cdots+101+101$$

但是，要有 100 列相加，因而，2S=100×101=10100，所以，前 100 个整数的和等于

$$S=1+2+3+\cdots\cdots+99+100=\frac{10100}{2}=5050$$

所有这些，在高斯脑子里都是瞬间完成的。显然，他的前途无量。

高斯学业进展神速，受到不伦瑞克公爵的赏识，15 岁时，在公爵资助下，进入卡罗林学院。三年后，他进入了久负盛名的哥廷根大学深造。1796 年，在哥廷根大学，他作出了有关正 17 边形的非凡发现。显然这是他投身于数学研究的一个决定性因素；他以前曾想成为一个语言学家，但 17 边形的发现使他相信，也许，他天生是个数学家。

1799 年，高斯因其对现在称之为代数基本定理的命题作出了第一个合理而完整的证明，在黑尔姆施泰特大学获得博士学位。仅从名称我们就能够感觉到这一定理的重要性。这一命题涉及到解多项式方程问题，显然，这是代数学上的一个基本课题。

虽然早在 17 世纪就有关于代数基本定理的论述,但真正使其著名的是

法国数学家让·达朗贝尔（1717—1783 年），他曾于 1748 年试图作出证明，但失败了。他所论述的定理是：任何实系数多项式都可以分解为实系数一次因式和二次因式的乘积。例如，因式分解

$$3x^4+5x^3+10x^2+20x-8=(3x-1)(x+2)(x^2+4)$$

即说明了达朗贝尔所论的分解方式。本例中的实系数多项式分解为几个简单的因式：两个一次因式和一个二次因式。

并且，我们注意到，我们还可以用复数来分解这一二次因式。虽然我们在讨论三次方程时曾接触过许多复数问题，但实际上，复数是在其后确立代数基本定理的过程中才日渐突出起来的。我们可以对下列方程进行验算：如果 a、b、c 是实数，并且 a≠0，那么，

$$ax^2+bx+c=\left(ax-\frac{-b+\sqrt{b^2-ac}}{2}\right)\left(x-\frac{-b-\sqrt{b^2-4ac}}{2a}\right)$$

这样，实系数二次多项式 ax^2+bx+c 就分解为两个有点不太赏心悦目的一次因式。（反应快的读者会看到，在这一因式分解过程中应用了二次方程公式。）

当然，我们不能保证这些一次因式都由实数组成，因为如果 $b^2-4ac<0$，我们就进入了虚数王国。例如，在上述例子中，我们可以进一步分解二次项，以得到完全国式分解：

$$3x^4+5x^3+10x^2+20x-8$$
$$=(3x-1)(x+2)(x-2\sqrt{-1})(x+2\sqrt{-1})$$

这样，随着一个四次实系数多项式分解为四个一次因式的乘积，我们当然会意识到有希望进行任何次多项式的完全因式分解。据此，代数基本定理称，任何 n 次实系数多项式都可以分解为 n 个（也许是复数）一次因式的乘积。

如前所述，达朗贝尔认识到了这一定理的重要性，并曾试图作出证明。但遗憾的是，他的努力远未成功。尽管他实际上未能证明这一定理，但也许是为了对他的努力表示敬意，这一定理长期以来一直称为“达朗贝尔定理”。这很有点儿像用拿破仑的名字命名莫斯科，只是因为拿破仑曾试图到达莫斯科。

18 世纪中叶，对这一定理的研究一直处于停滞不前的状态。关于这一定理是否正确，数学家们众说纷纭，例如哥德巴赫就曾怀疑其正确性，而那些相信其正确的数学家们也未能提出证明。也许最接近于作出证明的是李昂纳德·欧拉 1749 年的一篇论文。

欧拉的“证明”显示了他特有的机敏和独创性。他开始时论述得十分出色，漂亮地证明了实系数四次或实系数五次方程可以分解为实系数的一次或二次因式的乘积。但是，当他依据这一定理，论证更高次多项式时，他发现自己陷入了极度复杂的混乱之中。例如，对于他事先引入的一个辅助变量 u，首先要证明某一特定方程可解。欧拉不无遗憾地写道，“确定未知的 u 值，必须要解一个 12870 次方程。”他试图采用间接方法证明这一点，但他未能使他的评论家们信服。总之，他做出了令人钦佩的努力，但代数基本定理最终仍击败了他。而欧拉落败可能会给那些缺少数学才能的人（实际上包括每一个人）带来某些心灵上的安慰。

代数基本定理确立了以复数进行多项式因式分解的原则，但这一定理

一直处于非常不确定的状态。达朗贝尔未能作出证明；欧拉也仅仅证明了一部分。显然，需要极大的毅力，才能彻底证明其正确与否。

这样，我们再回到高斯的划时代论文，这篇论文的题目很长，且富有描述性：《关于任何有理代数整函数（即每一个带有实系数的多项式）都能够分解为一次或二次实因子的定理的一个新证明》。他首先就其前辈对这一定理的研究提出了自己批判性的评论。在论及欧拉的证明时，高斯认为，欧拉证明的缺陷是没有作出“数学所要求的一般性证明”。他不仅在这篇论文中，而且还在他 1814、1816 和 1848 年发表的对这一定理的三个别证中都成功地作出了对一般情况的证明。

今天，我们比 19 世纪初叶时更认识到这一重要定理的普遍性。我们现在可以在以下意义上将这一定理完全转化为复数问题：我们不再要求所论多项式必须具有实系数。总之，我们认为 n 次多项式既可以有实系数，也可以有复系数，例如，

$$z^7+(6\sqrt{-1})z^6(2+\sqrt{-1})z^2+19$$

尽管这种修改使其明显地变得更加复杂，但基本定理保证了即使是这种类型的多项式也能够分解为 n 个一次因式（当然带有复系数）的乘积。

高斯下一步的主要工作是对数论的研究，在这方面，他继承了欧几里得、费马和欧拉的传统。1801 年，他发表了他的数论名作《算术研究》。顺便说一句，他在这本书的最后广泛讨论了正多边形的作图（出人意料是，他将这一问题的讨论与复数密切联系起来）及这种作图与数论的关系问题。高斯一生始终关注这个问题，他曾说过，“数学，科学的皇后；数论，数学的皇后。”

卡尔·弗里德里希·高斯虽然尚不足 30 岁，但他已在几何、代数和数论领域作出了划时代的发现，并被任命为哥廷根天文台台长。他后来一直担任这一职务，直至逝世。这一工作要求他必须努力将数学应用于现实世界，这些问题与他所热爱的算术有天渊之别，然而他依然干得十分出色。在确定谷神星的运行轨道中，高斯起了很大作用；他还细心地描绘了地球磁场图，高斯与威廉·韦伯一起，是最早的磁学家。高斯还与韦伯合作，发明了电磁电报，几年后，美国科学家塞缪尔·F.B.莫尔斯在此基础上，发明了更大规模的电报通讯，并因此声誉雀起。高斯在数学应用方面的成就堪与他在纯数学领域的贡献相匹敌。像牛顿一样，他在这两个领域都获得了辉煌的成就。高斯与艾萨克爵士不仅在数学方面，而且在心理上，也有许多相似之处。他们两人都以冷淡、孤僻的个性及甘愿孤立从事研究而著称。他们都不大喜欢教学，但高斯却曾指导过 19 世纪一些最优秀的数学家进行博士研究。

并且，他们两人都尽力避免学术论争。我们回想一下，牛顿年青时似乎宁肯下油锅，也不愿将他的研究成果交给社会评判。高斯同样对与流行的科学观点相左而感到不安，最明显的是他在发现非欧几何时所表现的那样。我们在第二章的后记中曾提到，他担心自己如果在这个问题上提出命性见解，会遭到“蠢人的讥笑”。19 世纪初叶，高斯已成为世界最优秀的数学家。对此，他似乎特别意识到他的思想的影响及其必将受到的严格评判。对代数基本定理作出绝妙的证明是一回事，但要告诉世界三角形内角和可能会小于 180°则又是另一回事。高斯断然拒绝采取这种立场。他也像牛顿那样，把自己奇妙的发现收藏起来，锁进了抽屉深处。

然而，不应忽略，高斯这位刻板而内向的数学家还有其另外一面，令人意想不到。事情涉及他对法国女数学家索菲 · 格尔曼（1776—1831 年）的鼓励。索菲 ·格尔曼克服了重重障碍，终于成为 19 世纪初的杰出数学家。她的故事明确地揭示了一种社会态度，即认为数学学科不适于妇女。

格尔曼幼年时在他父亲的书房里发现了一些数学书，这些书深深地迷住了她，尤其是普卢塔克关于阿基米德之死的描述，对于阿基米德来说，数学甚至比生命更重要。但是，当她表示有志学习数学时，却遭到了她父母的反对。他们禁止她读数学书，索菲 · 格尔曼就只好把书偷偷拿进自己的房间，在微弱的烛光下苦读。后来，家里人发现了她的这些秘密，就拿走了她的蜡烛，并且，还拿走了她的衣服，让她无法在阴冷的屋子里读书。但是，这些极端的措施都没有能够使她屈服，这足以证明了格尔曼对数学的热爱，也许还证明了她身体的耐力。

当格尔曼掌握了更多的数学知识后，她就准备向更高级的程度进军。但是，她想进入学院或大学学习的想法在当时看来似乎十分荒谬，于是，她就只好在教室门外偷听，尽可能地记住老师讲课的内容，然后向富有同情心的男学生借来课堂笔记。很少有人是经过这样一条崎岖小路才进入高等数学殿堂的。

然而，索菲 · 格尔曼获得了成功。1816 年，她的工作已经给人以深刻印象，她对弹性片振动性质的透彻分析，为她赢取了法兰西研究院奖金。在这期间，她用假名安托万 · 勒布朗隐瞒了自己的身份，以免暴露身为女人这一不可宽恕的罪过。并且，她还以这一笔名与世界最优秀的数学家保持通信联系。

高斯从一开始时就对他的法国笔友印象极佳。勒布朗显然曾认真读过《算术研究》，并就书中定理有所概括和发展。1807 年，卡尔 · 弗里德里希 · 高斯终于知道了索菲 · 格尔曼的真实身份。格尔曼显然对这一消息所产生的影响甚为担忧，她写给高斯的信简直就像是一封忏悔书：

“ ……我以前曾用勒布朗的名字与您通信，这些信件无疑不值得您答复……我希望今天向您吐露的真情不会剥夺您曾经给予我的荣幸，并恳请您抽出几分钟时间向我介绍一些您自己的情况。 ”

也许出于格尔曼意外的是，高斯的回信充满了慈爱与理解。他承认，他在看到勒布朗“ 变成 ”索菲 · 格尔曼的时侯，确实感到“ 吃惊 ”，并且，他对数学界中的不公正表示了自己深刻的见解：

“ 人们很少对一般抽象科学，尤其是对数的奥秘发生兴趣。我们不会对此感到惊讶。这门卓越的科学，只向那些有勇气深入探索的人，展现它迷人的魅力。由于我们的习惯和偏见，女性要熟悉这些棘手的研究，必定会遇到比男性多得多的困难。但是当一个女性成功地越过了这些障碍，深入到其中最难解的部分时，那就毫无疑问，她必定具有最崇高的勇气，非凡的才能和超人一等的天才。 ”

高斯以同样的热情赞扬了格尔曼的数学著作，称其“ 给了我无比的快乐 ”。然后，他又继续写道，“ 如果我冒味对你的上封信作一点儿评论，请你把这看作是我对你关心的证明 ”，并进而指出了她推理中的错误。虽然索菲 · 格尔曼的数学能够给高斯以无穷的快乐，但这封信清楚表明了，在高斯心目中，究竟谁是大师。

应当指出，即使在她的身份暴露后，格尔曼的数学生涯依然很有成果。

1831 年，经高斯的大力推荐，哥廷根大学准备授予她名誉博士学位。这在19 世纪初叶的德国，对一个女人来说，是极大的荣耀。但非常遗憾，未及授予，格尔曼已逝世。

那么，卡尔·弗里德里希·高斯又是如何呢？他一直活到 78 岁高龄，最后死于他任台长近 50 年的哥廷根天文台。到他逝世的时侯，他的声望已达到近乎神话的程度，人们只要一提到数学家之王，就知道是指高斯，而非其他人。

然而，高斯自己却遵循着一句不同的格言：“ 少些，但要成熟 ”，这句格言贴切地反映了他的生活和工作。高斯在有生之年发表的著作比较少。但他大量未发表的著作却足以使众多数学家成名。他特别注意他的著作可能产生的影响，并尽可能达到尽善尽美的程度才予以发表。高斯的著作虽然不如欧拉数量多，但一旦下笔，就会引起数学界的注意。他身后留下的成果（从正 17 边形的作图，到《算术研究》和辉煌的代数基本定理），具备了任何数学著作所应具备的成熟。

第十一章
连续统的不可数性
（1874 年）

19 世纪的数学

每一个世纪都以一种奇特的方式，显示不同的数学重点和数学思维方向。18 世纪显然是“欧拉世纪”，因为他在学术领域没有任何对手，始终居于统治地位，并为后代留下了珍贵的数学遗产。相比之下，19 世纪虽然没有一位特别出类拔萃的数学家，但却有幸拥有许多优秀数学家，他们将数学疆界推向新的令人意想不到的方向。

如果说 19 世纪不属于某一位数学家，那么，它确实呈现出几个重要的主旋律。19 世纪是抽象与广义化的世纪，是对数学的逻辑基础进行深入分析的世纪，这种逻辑基础曾构成牛顿、莱布尼兹和欧拉的理论基础。数学不再受“物理实在性”的局限而变得越来越独立，而在此之前，这种“物理实在性”始终明显地将数学束缚于自然科学。

这种脱离实在世界的倾向可以说是以 19 世纪前 30 年出现的非欧几何作为其独立宣言的。我们在第二章的后记中曾说过，当欧几里得的平行线公设被舍弃而代之以另一命题的时候，出现了一个“奇怪的新世界”。突然间，通过直线外一点，至少可以画两条直线与之平行；相似三角形变成了全等；而三角形的内角和也不再等于 180°。然而，对于非欧几何中所有这些似乎矛盾的性质，没有一个人能够从中找出逻辑矛盾。

欧金尼奥·贝尔特拉米证明了非欧几何与欧氏几何一样，在逻辑上是成立的。从而在这两种几何之间架起了一座桥梁。我们可以设想，比方说，数学家甲致力于研究欧氏几何，而数学家乙则专攻非欧几何。双方的工作具有等效的逻辑正确性。然而“实在的世界”却不可能既是欧氏几何的又是非欧几何的；其中的一位数学家必定要付出终生的努力去探索一种并非“实在的”体系，那么，他或她是否在虚掷年华呢？

19 世纪，数学家越来越感到对这个问题的答案应该是否定的。当然，物质世界是否如欧几里得所述，这个问题应留待物理学家去探讨。这是一个经验性问题，是通过实验与严格的观测来确定的，但却与这两种几何体系的逻辑发展无关。对于一个热中非欧几何优美定理的数学家来说，美就足够了。无需物理学家去告诉数学家哪一种几何是“实在”的，因为在逻辑王国里，两者都是正确的。听以，几何学的这一根本问题带有一种解放的性质，将数学从只依赖于实验室的实验结果中解放出来。在这个意义上，我们看到，这与当时美术摆脱对现实的依赖的情形十分相似。19 世纪初期，画家的画布还像以往一样，仅仅充当了一扇窗户，人们通过这扇窗户，可以看到有趣的人和事。当然，画家可以自由设定基调，选择颜色，确定明暗，强调某一局部而弱化其他部分；但无论如何，画家的作品就像一幅屏幕，让大家看到瞬间静止的事物。

19 世纪后半期，情况发生了明显的变化。在一些美术大师如保罗·塞尚、保罗·高庚和樊尚·凡高的影响下，美术作品获得了自己的生命。画家可以视画布为发挥自己绘画技能的二维战场。例如，塞尚认为，可以任

意将静物苹果与梨变形，以增强整体效果。他批评伟大的印象派画家克劳德·莫奈只有“一只眼睛”，他的意思是说，画家的艺术不仅仅限于记录眼睛所看到的事物。

总之，美术宣告了从视觉现实中的独立，同时，数学也显示出其脱离物质世界的倾向。这种并行的情况很有趣，以塞尚、高庚和凡高为代表的绘画，连同以高斯、鲍耶和罗巴切夫斯基为代表的数学，其哲学内涵意义深远，影响持久，至今不衰。

当然，我们也必须看到，这些发展并非得到了人们的一致认可。20 世纪末，任何一个到美术馆参观的人，随时都能听到种种议论，人们对视觉艺术的现状，对在大幅画布上毫无意义地胡乱涂抹，对那些自称并不反映现实的作品（这些作品常常争议很大，而又十分昂贵）颇有微词。艺术家的赞助人则常常抱怨当代艺术家的解放走得太远了。他们渴望看到他们所熟悉的肖像画和令人赏心悦目的风景画。

在这一方面，数学与美术也十分相似。在现代数学界中也有一种对当今数学状况不满的情绪。20 世纪的数学家不但偏好非欧几何革命所带来的思想解放，而且还推动数学越来越远地脱离与实在世界的联系，直到把他们的逻辑结构变得抽象而神秘，以致使物理学家和工程师都如堕烟海，不知其所云。在许多人看来，这种趋势已把数学变成了一种毫无意义的符号游戏。数学史家莫里斯·克兰对这种倾向提出了最畅言无忌的批评，他写道：

“随着深奥晦涩的原理被系统地阐述，已远离了最初的应用领域，而专注于抽象的形式。通过引入上百个分支概念，数学雨后春笋般地扩张为琐细而庞杂的一个个小门类，它们相互之间很少联系，且与最初的应用领域很少关联。”

克兰认为，数学在其争取独立于物理学的来之不易的自由的过程中，走得太远了，以致成为枯燥而任意的纯粹形式主义体系。对他的严厉批评，数学界确应认真考虑。

作为对克兰批评的回答，令人感兴趣的是，数学理论无论有多么抽象，却常常出人意料地应用于非常确实的实际问题。甚至将数学与实在断然分开的革命的非欧几何，也可以在现代物理书籍中找到它的足迹，现代相对论宇宙学就在很大程度上依据非欧几何建立了宇宙的模型。当然，19 世纪的数学家是不可能预见到这种应用的，他们对于非欧几何，只是为了研究而研究；如今，非欧几何已成为应用数学的一部分，并成为物理学家的必要工具。数学有时会在最不可思议的地方出现。

论争还在继续。最后，历史学家可能会看到，今天的数学虽然已远远地脱离了实在世界的桎梏，但令人难以置信的是，数学总能在其他学科的研究与发展中承担不可替代的角色。数学的抽象化将永远是 19 世纪留给人类的一笔财富。

除了非欧几何的产生所提出的这些问题以外，另一个主要论争是关于微积分的逻辑基础。我们可以回想一下，微积分是 17 世纪末由牛顿和莱布尼兹奠定基础，而后在 18 世纪由李昂纳德·欧拉进一步完善的。然而，这些先驱者及其同时代的数学天才，都未能对微积分的基础给予充分注意。这些数学家如履薄冰，基础上的裂痕随时可能招致灭顶之灾。

长期以来，人们始终感到，微积分有其问题。问题存在于对“无穷大”

和“无穷小”概念的使用上，在牛顿的流数术和莱布尼兹的微积分中，这是必不可少的。微积分的一个核心思想是“极限”。无论微分，还是积分（还不要说级数收敛性和函数连续性的问题），都以这种或那种形式依据于这一概念。“极限”一词很有启发性，并有很强的直感。我们常常说，“我们的耐性或耐力到了极限”。然而，如果我们要从逻辑上准确地说明这一概念，就立刻出现了困难。

牛顿曾对此作过尝试。他的流数概念要求他必须观察两个量的比，并确定当这两个量同时趋向于零时，它们的比将会怎样。用现代术语来说，他讲的正是两个无穷小量的比例极限，但他使用了一个更具特色的词“最后比”。对于牛顿来说，所谓两个正在消失的量的最后比

“……应当理解为，既不是在两个量消失之前，也不是在它们消失之后，而是正当它们消失时的瞬间比。”

当然，作为数学定义，这没有什么意义。我们可能赞同牛顿关于不应将极限概念基于两个量消失之前的比这一观点，但他所说两个量消失之后的比又是什么意思呢？牛顿考虑的似乎是当分子和分母刚好同时成为零时的瞬间比。可是在那一瞬间，分数$\frac{0}{0}$是无意义的。牛顿陷入无法自圆其说的逻辑困境。

那么，莱布尼兹如何走出这一泥淖呢？他同样需要阐明极限过程中发生的变化，但他倾向于通过对“无穷小量”的讨论来探索这一问题。莱布尼兹所谓的无穷小量尽管不是零，但却小于任意有限量。他的无穷小量，犹如化学中的原子一样，是不可再分的数学单元，是最接近于零的量。但与此相关的哲学问题显然使莱布尼兹感到困惑，他不得不作出如下晦涩的说明：

“当我们谈及无穷小量……（即在我们的知识中是最小的），它可以被看作是……无限小……如果有人想理解这些（无穷小），可以想象它们是最终的东西……这就足够了……如此假设是充分的……即使认为这样的东西是不可能的，也完全可以利用它们作为计算的手段，就像代数中用虚根有极大好处一样。”

在这里，除了莱布尼兹对复数的偏见以外，还可以看到他关于数学的令人莫名其妙的陈述。显然，概念的含糊不清（特别是构成微积分基础的概念）使莱布尼兹犹豫不定。

当数学家们正因微积分遗留的逻辑基础问题而深感不安时，又受到来自上帝的仆人——乔治·贝克莱大主教（1685—1753 年）的强有力的攻击。贝克莱大主教在他刻薄的文章《精神分析学家或神学家致不信教的数学家》中嘲弄那些批评神学基础是一种虚幻信仰的数学家，攻击他们所信奉的微积分，其逻辑基础同样十分脆弱。贝克莱采取以子之矛陷子之盾的策略：

“可以说，所有这些（来自数学的）观点都是那些对宗教过于苛求的人设想和信奉的，他们自称只相信亲眼所见……那么如果他们能消化二阶或三阶流数和微分，就不会因为某一神学观点而反胃。”

如果说这些挖苦还不够刻薄的话，贝克莱又发出了更加无情的嘲笑：

“所谓流数是什么？数学家们说，是瞬时增量的速度。那么，这些瞬时增量又是什么？它们既不是有限量，也不是无穷小量，然而又不是虚无。

我们难道称它们为消失量的幽灵吗……？”

这真糟透了，微积分的基础居然成了“消失量的幽灵”。可以想象，对于数以百计的数学家们来说，贝克莱的冷嘲热讽会使他们多么焦躁不安。

数学界逐渐认识到，他们必须正视这一令人头痛的问题。纵观 18 世纪，数学家们对微积分在实际应用上的巨大成功过于乐观，以致阻碍了对其基础理论的研究。但是数学界内部日益增多的关注及外界贝克莱的傲慢无礼，已使他们别无选择。这个问题已经迫在眉睫，不能不解决了。

这样，我们看到一个又一个才华横溢的数学家开始探讨这一基础理论。建立严格的“极限”理论是一个困难的漫长的过程，因为这一概念的内涵非常深奥，需要精确的推理和对实数系性质的深刻理解，这绝非易事。但数学家们对这个问题的研究已逐渐有所突破。1821 年，法国数学家奥古斯坦-路易·柯西（1789—1857 年）提出了如下定义：

当一个变量逐次取的值无限趋近一个定值时，如果最终使变量的值与该定值的差要多小就有多小，那么，这一定值就称为所有其他值的极限。

我们看到，柯西的定义避免了使用像“无穷小”样含糊不清的词，他没有将自己束缚于确定变量达到极限时的瞬间会如何如何。因而，这里也就不会出现消失量的幽灵。相反，他只是说，如果我们能够使变量的值与某一定值的差要多小就有多小，那么，这一定值就是该变量的极限。这就是所谓“极限回避”，柯西的定义绕开了关于达到极限的瞬间会发生什么这一哲学上的障碍。在柯西看来，最后瞬间的结局是完全不相干的，重要的是我们已经尽可能地澄清了极限这一概念，这才是我们所需要的。

柯西的定义产生了深远的影响，以这一定义为基石，他继续阐明了微积分的许多重要概念。数学家们经过漫长的道路，进一步完善了基于这一极限定义的微积分，有力地反击了贝克莱大主教的“关心”。然而，柯西的陈述尚有一些不足之处。首先，他讲到，一个变量“趋近”某一极限，仅凭幻想就提出了一个关于运动的不明确的概念；如果我们必须依靠直觉来阐述关于点的移动和相互接近的概念，那么，我们仅仅依赖直觉提出“极限”概念难道就会更好些吗？其次，柯西使用的“无限”这一措词看起来也有点儿不确定；其意义需要进一步明确。最后，柯西的定义完全是文字叙述，有必要代之以简洁、明确、清晰的数学符号。

于是，便出现了德国数学家卡尔·维尔斯特拉斯（1815—1897 年）及其追随者。他们使用一种读来有些拗口的方法，即“微积分的算术化”，支撑起微积分的基础。维尔斯特拉斯学派的语言是“当 x 趋近于 a 时，函数 $f(x)$ 以 L 为极限”，可以严格地表述为：

对于任意给定的 $\varepsilon>0$，总存在着一个 $\delta>0$，所以，如果 $0<|x-a|<\delta$，那么，$|f(x)-L|<\varepsilon$ 能够成立。

不必全面理解这一定义，我们就可以清楚地看出，这个定义与柯西的定义明显不同。维尔斯特拉斯的定义几乎全部使用了数学符号，而且无一处暗示某一量向其他一些量的移动。总之，这是一个极限的静态定义。另外，维尔斯特拉斯的定义与前面所引牛顿和莱布尼兹的含糊不清、几乎引人发笑的陈述相比，大相径庭。维尔斯特拉斯逻辑严谨的定义虽然缺乏其前辈的某些趣味和魅力，但在数学上却是无懈可击的。在此基础上建立起的微积分大厦一直矗立至今。

康托与无穷的挑战

科学中常常会出现这种情况，一个问题的解决打开了解决另一个问题的大门。随着越来越少地依赖直觉构造概念而越来越多地依靠维尔斯特拉斯数学中的 ε 和 δ，数学家们开始从更高的视角严格地审查微积分。他们得到了一些非常奇特和令人不安的发现。

例如，考虑有理数与无理数两者之间的区别。有理数全都可以写成分数的形式，可以表示为整数的比。如果把有理数化为小数，则很容易确定：它们或为有限小数（如，$\frac{3}{8}=0.375$），或为无限循环小数（如，$\frac{3}{11}=0.27272727\cdots\cdots$）。与之相反，无理数则是不能写成分数形式的实数。大家非常熟悉的$\sqrt{2}$和 π 就是无理数。无理数既不是有限小数，也不是循环小数，而是无限不循环小数。

我们可以说，不论有理数，还是无理数，在实数轴上是处处稠密的，即：在任意两个有理数之间，分布着无穷多个无理数；反之亦然，在任何两个无理数之间也分布着无穷多个有理数。自然而然，我们会放心地推断，实数轴上一定均匀地分布着两个基本相等的巨大的有理数族与无理数族。

然而，19 世纪，随着时间的推移，越来越多的数学发现表明，与上述认识相反，这两个数族并不相等。这些发现一般需要非常高深的技巧和精妙的推理。例如，要证明函数在每一个无理点连续（直觉上不间断），并在每一个有理点不连续（间断），就必须证明在每一个有理点不存在连续的函数，而在每一个无理点不存在不连续的函数。这里有一个明显的指标，即在有理数族与无理数族之间不存在对称或平衡。这就表明，从某种根本意义上说，有理数与无理数是不可交换的数族，但当时的数学家对这两个数族的根本性质，尚不十分明了。

因而，对实数系性质的深刻理解就促成了我们本章将要讨论的定理的产生。虽然柯西、维尔斯特拉斯及其同事们成功地用“极限”概念建立了微积分大厦，但数学家们越来越清楚地认识到，最重要、最基本的问题是将微积分最终置于集合的严格基础之上。探索这个问题，并单枪匹马地创立了奇妙的集合论的是一位时而被人恶意中伤，又曾一度精神崩溃的天才，他的名字叫乔治·费迪南德·路德维希·菲利普·康托。

康托 1845 年出生于俄国，但他 12 岁的时候，随家移居到德国。宗教是康托家庭的重要组成部分。康托的父亲原是犹太教徒，后来皈依了新教，而他的母亲则生来就是罗马天主教徒。由于家庭中这种混合的宗教信仰，所以，毫不奇怪，小乔治对神学产生了一种终生的兴趣，特别是那些与无穷性质有关的神学问题对成年康托的数学产生了很大的影响。

并且，康托的家庭还显示了明显的艺术素质。在康托家庭中，音乐受到特别的尊崇。康托有几个亲戚在大交响乐团演奏。乔治本人是一个很不错的素描画家，他留给后人一些很能表现他天才的铅笔画。总之，我们可以说康托具备了“艺术家”的天性。

这位敏感的年轻人特别擅长数学，1867 年，他在柏林大学获得博士学位。在此，他从师于维尔斯特拉斯，并完全掌握了前面所介绍的有关微积

分的严谨的推理方法。康托对数学分析的深入研究使他越来越多地考虑各种数集之间的本质区别。特别是，他开始认识到，创立一种比较数集大小的方法是十分重要的。

表面看来，比较数集大小似乎轻而易举：只要会数数，就会比较。如果有人问你，“你左手与右手的手指一样多吗？”你只要分别数一数每只手的手指，确认每只手都有 5 个手指，然后，就可以作出肯定的回答。看来，原始的“数数”方法似乎对于确定更复杂的“同样大小”或“相同基数”概念也是必要的。然而，乔治·康托以一种貌似天真的方法，颠倒了前人传统的观念。

我们来看一看他是如何论证的。首先假设我们生活在一种数学知识非常有限的文化中，人们最多只能数到“3”。这样，我们就无法用数数的方法来比较左手与右手的手指数目，因为我们的数系不能使我们数到“5”。在超出我们计数能力的情况下，是否就无法确定“相同基数”了呢？完全不是。实际上，我们不必去数手指，而只需将两手合拢，使左手拇指与右手拇指，左手食指与右手食指……一一对齐，就能够回答这个问题了。这种方法展示了一种纯粹的一一对应关系，然后，我们可以回答，“是的，我们左手与右手的手指一样多”。

我们再来看另一个例子。假设许多观众涌入一个大礼堂。那么，观众与座椅是否一样多呢？要回答这个问题，我们可以分别数一数观众与座椅，然后将两个数字加以比较，但这种方法过于繁琐。我们其实只需要求礼堂中的所有观众坐下。如果每个人都有座位，或者，每个座位都有人，那么答案就是肯定的，因为坐下这个过程已显示了一种完全的一一对应关系。

这些例子阐明了一个关键的论据，我们无须去数集合中元素的个数，以确定这些集合是否具有同样数值。相反，根据一一对应关系来确定同等数量的概念已成为一种更原始和更基本的概念；相形之下，数数的方法却成了更复杂和更高级的方法。

乔治·康托对这一概念作出了如下定义：

如果能够根据某一法则，使集合 M 与集合 N 中的元素建立一一对应的关系……那么，集合 M 与集合 N 等价。

如果集合 M 与集合 N 符合上述康托的等价定义，那么，按现代数学家的语言，集合 M 与集合 N“等势”或具有“相同基数”。然而，我们暂且抛开这些术语不谈。这一定义之所以重要，就在于它并未限定集合 M 与集合 N 必须包含有限个元素；因此它同样适用于那些包含无限多个元素的集合。

据此，康托进入了一片未开垦的处女地。在数学发展的历程中，人们始终以一种怀疑的眼光（即使不是敌对的眼光）看待无穷，并尽可能回避这一概念。从古希腊时期直到康托时代，哲学家和数学家们都只承认“潜无穷”的存在。也就是说，他们能够在如下意义上同意整数集是无穷的：对于整数集中的任何一个数我们都能找到下一个比它更大的整数，但我们决不可能穷举所有整数。例如，可以想象把每一个整数都写在一张纸条上，然后把这些纸条放进一个（非常大的）袋子里，那么，即使地老天荒我们的工作也永远不会终止。

但是，康托的前辈们反对“实无穷”的概念——即，他们反对认为这

一过程能够结束或袋子能够装满的观点。用卡尔·弗里德里希·高斯的话说：

“……我首先反对将无穷量作为一个实体，这在数学中是从来不允许的。所谓无穷，只是一种说话的方式……”

康托不同意高斯的观点。与其他无穷集相比，他极愿意将这个装有所有整数的袋子看作一个自足的和完整的实体。与高斯不同，他不是将“无穷”仅仅看作一种说话的方式而不予考虑。对于康托来说，“无穷”是一个应予以高度重视的确实的数学概念，值得我们对其进行严格的理性论证。

这样，乔治·康托仅仅依据这两个基本前提（即可以通过一一对应的方法来确定相同基数和实无穷是一个确实的概念），就创立了最令人兴奋和意义十分深远的理论。这一理论使我们进入了一个难以捉摸的奇特世界，虽然一些数学权威时时嘲笑他的努力，但康托没有因此而气馁。终于，凭着天才和勇气，康托以完全前所未有的方式，正面探讨无穷。

我们首先设自然数集 N＝｛1，2，3，……｝，并设偶数集 E＝｛2，4，6，……｝。注意到这两个数集都是完全集，而不必顾忌他们的无穷性质。根据康托的定义，我们可以很容易看出集合 N 与集合 E 具有“相同基数”，因为我们可以列出这两个数集之间单纯的一一对应关系：

$$\begin{array}{lcccccc} N: & 1 & 2 & 3 & 4 & 5 & \cdots\cdots n\cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ E: & 2 & 4 & 6 & 8 & 10 & \cdots\cdots 2n\cdots\cdots \end{array}$$

这种对应关系明确地显示出，N 集中的每一个元素都被一个、并且只被一个偶数（即其 2 倍）所指定；反之，每一个偶数也都被一个、并且只被一个自然数（即其一半）指定。康托认为，这两个无穷数集显然是等价的。当然，乍一看，似乎很矛盾，这里人们本来会以为，偶数的个数应该是整数个数的一半。那么，我们依据什么才能够非难康托的演绎推理呢？我们或者抛弃实无穷的概念，甚至否认自然数集是一个自足的实体；或者拒绝承认简明的相同基数定义，而把它看作是荒谬的。但只要我们承认这两个前提，那么，就不可避免地会得出结论：偶数的个数绝不少于自然数。

同样，如果设 Z＝｛……－3，－2，－1，0，1，2，3，……｝，即所有整数（正数、负数和零）的集合，那么，我们会看到，N 与 Z 也有相同的基数，因为它们可以构成如下一一对应关系：

$$\begin{array}{lcccccccccc} N: & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ Z: & 0 & 1 & -1 & 2 & -2 & 3 & -3 & 4 & -4 & \cdots\cdots \end{array}$$

对于这一对应关系，我们可以进行检验，集合 N 中的每一个自然数 n 都与集合 Z 中的

$$\frac{1+(-1)^n(2n-1)}{4}$$

相对应。

据此，康托迈出了勇敢的一步。他说，任何能够与集合 N 构成一一对应关系的集合都是可列或可数无穷集。特别是，他引进了“超限”基数的新概念，用以表示可数集中元素的个数。他选用希伯来文的第一个字母$\aleph_0$

（读作“阿列夫零”）来表示超限基数。

康托通过对无穷集的研究，创造了一种新的数字和一种新的数字类型。我们可以想象，他的许多同代人都会对这个异想天开的可怜虫摇头叹息。然而，不要忘记在我们所假设的原始数学文化中，人们只能数到三。在这种文化中，一个富有革新精神的天才也许会突发灵感，通过引入一个新的基数五来扩大原有数系：如果一个集合的元素能够与她右手的手指一一对应，那么，她就可以说，这个集合包含了五个元素。

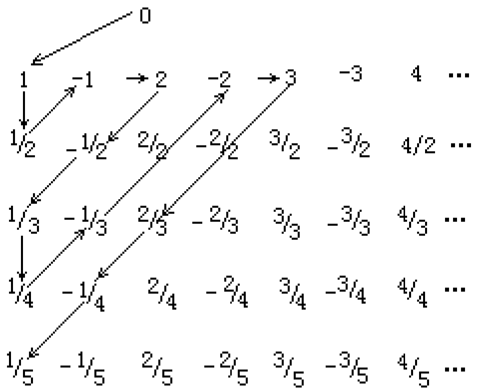
这样一个定义是非常有效的，它提供了一个明确的方法，以确定一个集合在什么情况下具有五个元素（只要手指不受损伤）。在这个意义上，她的手指就成为确定集合是否具有五个元素的标准参考点。这一切看起来是非常合理的。

而这恰恰是康托的证明方法，所不同的只是他采用自然数集合 N 作为扩大我们数系的基准。对于他来说，N 是基数为 $\aleph_0$ 的原型集合。引入符号 $\overline{M}$表示“集合M的基数”，我们看到，

$$\overline{N}=\overline{E}=\overline{Z}=\aleph_0$$

如果我们接下来讨论有理数集合 Q，情形又会如何呢？如前所述，有理数是处处稠密的。在这个意义上说，有理数与整数不同，整数是一个紧跟一个，循规蹈矩地分布在数轴上的，其中的每一个数字都与前一个数字保持相同的距离。实际上，在任何两个整数之间（比如在 0 与 1 之间），都有无限多的有理数。因此，任何人都会猜想，有理数的个数远远超过自然数。

但是，康托证明，有理数集是可列的；也就是说，$\overline{Q}=\aleph_0$。他的证明方法是在有理数集与自然数集之间构成一一对应的关系。为了弄清他是怎样构成这种对应关系的，我们把有理数排列成如下形式：



注意第一列中所有数字的分子是 1，第二列所有数字的分子是 −1，等等；而第一行中所有数字的分母为 1，第二行所有数字的分母为 2，依此类推。总之，任何分数，都能够在这一排列中找到它的固定的归宿。例如，分数$\frac{133}{191}$就存在于第191行，第265列（包括正数和负数）。当然，这一排列包含了集合 Q 中的所有元素。

现在，我们按照这一排列中箭头所示方向，列出集合 Q 的元素，由此便产生了以下对应关系：

$$
\begin{array}{cccccccccccccccc}
N: & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & \cdots\cdots \\
 & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\
Q: & 0 & 1 & \frac{1}{2} & -1 & 2 & -\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & -\frac{1}{3} & -2 & 3 & \frac{2}{3} & -\frac{1}{4} & \frac{1}{5} & \cdots\cdots
\end{array}
$$

细心的读者会发现，这里我们已去掉了重复的分数（如，$1=\frac{2}{2}=\frac{3}{3}$，等等）。于是，按照这一方案，每一个自然数必与一个且仅与一个有理数相对应；更令人吃惊的是，每一个有理数也将被一个且仅被一个自然数所指定。根据康托的定义，我们可以直接得出结论：有理数与自然数一样多。

康托 1874 年论文中关于连续统不可数性的最初证明

至此，似乎所有的无穷集都是可列的，也就是说，每一个无穷集都能与正整数构成一一对应的关系。但是，在看到康托 1874 年的一篇论文后，数学界彻底放弃了这个一相情愿的念头。这篇论文有一个平铺直叙的题目：《论所有代数数集合的性质》。在这篇论文中，康托明确地提出了不可数无穷集的问题。

仅从文章平凡的标题来看，人们丝毫不会感到这篇论文的革命性。这恰恰与美术界的根本变革形成了鲜明的对照，美术作品常常明显地表现出它的革新。1874 年，任何人，即便是门外汉，只要在巴黎看到过莫奈的作品，都会对他“印象派”的绘画方法感到震惊。只需随意看一眼，也会从莫奈表现光的手法中看出他的作品与其前辈，如德拉克洛瓦或安格尔，有着明显的区别。显然，莫奈作了某些根本的变革。同是 1874 年，乔治·康托在其划时代的数学论文中，开创了同样不乏革命性的事业。然而，这一惊人的数学思想恰恰缺乏美术作品那样的直接冲击。

康托发现的不可数集是所有实数的集合。实际上，他 1874 年的论文指出，没有任何实数区间（不论其长度多么小）能够与自然数集构成一一对应的关系。他最初的证明使他进入了分析的王国，同时，这一证明需要借助某些相关的比较先进的数学工具。然而，1891 年，康托再次回到这个问题上来，提出了一个非常简单的证明。我们下面将讨论这个证明。

伟大的定理：连续统的不可数性

这里“连续统”一词的意思是指某一实数区间，我们可以用符号（a，b）来表示（图 11.1），即

（a，b）表示满足于不等式 $a<x<b$ 的一切实数 x 的集合

在以下的证明中，我们将要证明的不可数区间是（0，1），即所谓“单位区间”。在这一区间的实数都可以写成无穷小数。例如，

$\frac{1}{2}=0.50000000\cdots\cdots$，$\frac{3}{11}=0.27272727\cdots\cdots$，和 $\frac{\pi}{4}=0.78539816\cdots\cdots$出于技术上的原因，我们必须谨慎地避免采用两个不同的小数来表示同一个数字。例如，$0.50000\cdots\cdots=\frac{1}{2}$，还可以写成$0.49999999\cdots\cdots$。

在这种情况下，我们选择以一连串 0 结尾的小数展开式，而不选择以一连串 9 结尾的小数，这样，在(0,1)区间中的任何实数都只有一种小数表示。

我们现在来看康托关于区间(0,1)不可数的证明。康托的证明采用了反证法,他从假定自然数集合 N 与区间(0,1)内的实数存在一一对应关系这一前提出发，然后，由此推导出逻辑矛盾。这一漂亮的证明可以当之无愧地排在伟大的定理之列。

定理 0 与 1 之间的所有实数不可数。

证明 我们首先假定区间(0,1)内的实数能够与自然数一一对应，然后，从这一假定出发最终推出逻辑矛盾。为了讲清楚康托的论证，我们假定存在如下的对应关系：

$$
\begin{array}{lll}
\underline{N} & \underline{(0,1)\text{内的实数}} & \\
1 & x_1 = 0.371652\cdots\cdots & \\
2 & x_2 = 0.500000\cdots\cdots & \\
3 & x_3 = 0.142678\cdots\cdots & \\
4 & x_4 = 0.000819\cdots\cdots & \\
5 & x_5 = 0.987676\cdots\cdots & \\
\vdots & \vdots & \\
n \leftrightarrow & x_n = 0.a_1a_2a_3a_4a_5\cdots\cdots a_n\cdots\cdots & \\
\vdots & \vdots &
\end{array}
$$

如果这是真正的一一对应关系，那么，右边一列区间(0,1)内的每一个实数都应该唯一地与左边一列中的一个自然数相对应。康托定义了一个区间(0,1)内的实数 b，令 $b=0.b_1b_2b_3b_4b_5\cdots\cdots b_n\cdots\cdots$。其中：

选择 b_1（b 的第一位小数）为与 x_1 的第一位小数不同且不等于 0 或 9 的任何数字。

选择 b_2（b 的第二位小数）为与 x_2 的第二位小数不同且不等于 0 或 9 的任何数字。

选择 b_3（b 的第三位小数）为与 x_3 的第三位小数不同且不等于 0 或 9 的任何数字。

一般地，选择 b_n（b 的第 n 位小数）为与 x_n 的第 n 位小数不同且不等于 0 或 9 的任何数字。

为便于理解这一过程，我们可以参照上述的对应表。x_1 的第一位小数为“3”，因而，我们可以选择 $b_1=4$；x_2 的第二位小数是“0”，我们可以选择 $b_2=1$；x_3 的第三位小数是“2”，我们选择 $b_3=3$；x_4 的第四位小数是“8”，所以，我们选择 $b_4=7$；等等，依此类推。所以，我们的数字 b 就是

$b=0.b_1b_2b_3b_4b_5\cdots\cdots=0.41378\cdots\cdots$

现在，我们只需要来看两个十分简单，但却是相互矛盾的事实：

(1)因为 b 是一个无穷小数，所以，b 是实数。由于我们禁止选择 0 或 9，因而，数字 b 既不可能是 0.00000……＝0，也不可能是 0.99999……＝1。换言之，b 一定严格地位于 0 与 1 之间。所以，b 一定会在我们上述对应表的右边一列中出现。但是，

(2)b 不可能出现在数字 x_1，x_2，x_3，……x_n，……中的任何位置，因为 b 与 x_1 的第一位小数不同，$b\neq x_1$；b 与 x_2 的第二位小数不同，$b\neq x_2$；总之，b 与 x_n 的第 n 位小数不同，$b\neq x_n$。

这样，(1)告诉我们 b 一定位于上表的右列，而同时(2)又告诉我们，b 不可能列入上表，因为它已被明确地“设计”为不与 x_1、x_2……x_n 等等数字中的任何一个数字相同。这一逻辑矛盾说明，我们最初的假定，即单位区间内的所有实数与自然数之间存在一一对应的关系是不正确的。可以断定，这种对应关系是根本不可能的，所以，0 与 1 之间的所有实数是不可列的。证讫。

我们在选择 b 的数值时之所以避免采用“9”，还有一个原因。我们再来看看上述对应表，但这一次我们选用 9 作为 b_n 的数值（当然，按规定，它们必须与 x_n 的第 n 位小数不同）。那么，我们可以选择 $b_1=4$，$b_2=9$，$b_3=9$，$b_4=9$，等等。因而，我们最后选定的数字是 b＝0.49999……。但是，这个数字恰恰等于 $\frac{1}{2}$，是在表的右列中已经存在的一个数字 x_2。这样，我们所寻求的矛盾（确定一个不能列入表右列的实数 b）就消失了。但是，如果我们在选定 b 值时避免采用“9”，我们就可以消除因无尽小数的这一双重表示所造成的技术陷阱，从而使证明有效。

康托自己显然对这个证明感到非常满意，他称这一证明“……很不寻常……因为证明方法非常简单”。证明中他把焦点集中在右边一列小数的某些位置特殊的数字上，这些数字恰好连成一条下降的对角线——第一个实数的第一位小数，第二个实数的第二位小数，等等。这一方法因此被称为康托的“对角线法”。

应特别注意的是，在证明中，我们并没有依赖上述假定的对应关系中的具体数字去说明问题。仅仅通过抽象的讨论就证明了这种一一对应的关系是不可能存在的。

持怀疑态度的人常常一方面承认康托找出的数字 b 不能出现在原始对应表中，一方面又提出以下补救方法：为什么不将 b 与自然数 1 对应，并将表中右边的每一个数字都下移一个位置呢？这样，2 将与 x_1 对应，3 与 x_2 对应，等等。因而，康托所推出的矛盾似乎也就消失了，因为 b 出现在表中右边一列的最上端。

然而，对于这些怀疑论者，遗憾的是，康托可以悠闲地坐等他们将最初的对应表调整完毕，然后再次应用对角线法找出一个新表中没有的实数 b'。如果我们多疑的朋友又将 b' 插入了表的最上端，那么，我们可以如法炮制，得出一个表中不存在的 b"。总而言之，在 N 与（0，1）之间是不可能存在一一对应关系的。至此，我们神经过敏的朋友心中的疑团一定会烟消云散了。

这样，康托证明了许多无穷集合（特别是有理数集合）都具有基数 $\aleph_0$。

然而，尽管同是无穷，0 与 1 之间的实数似乎是“更高一级的”无穷。这一区间内的点如此之多，其数量绝对超过了正整数。

在这一意义上，单位区间(0,1)不失一般性。对于任意给出的有限区间(a,b)，我们可以引入函数 y＝a＋(b－a)，使区间(0,1)内的点（x 轴上的点）与区间(a,b)内的点（y 轴上的点）之间建立起一一对应的关系，如图 11.2 所示。这种一一对应的关系保证了区间(0,1)与(a,b)具有相同的（不可数）基数。也许会令人感到吃惊的是，区间的基数与其长度无关；0 与 1 之间的所有实数并不比 2 与 1000 之间的所有实数少（在这种情况下，函数 y＝998x＋2 提供了必要的一一对应关系）。初一看，这似乎是违反直觉的，但当人们熟悉了无穷集合的性质，便不再相信幼稚的直觉。

在此基础上，再向前迈一小步，我们便可以证明，所有实数的集合同样具有与区间(0,1)相同的基数。这一次，确定一一对应关系的函数是

$$y=\frac{2x-1}{x-x^2}$$

如图 11.3 所示，区间(0,1)内的每一个点 x 都有唯一的一个实数 y 与它相伴，反之，每一个实数 y，也都有一个且仅有一个区间(0,1)内的点 x 与之相对应。总之，这就是必要的一一对应关系。

现在，我们可以跟随康托，再向前迈出勇敢的一步。正像我们曾把 N 作为基本集合而引入了第一个超限基数 $\aleph_0$ 一样，区间(0,1)也将作为定义一个新的、更大的超限基数的标准。也就是说，我们可以规定这一单位区间的基数为 c（英文“连续统”一词的第一个字母）。我们前面的讨论表明，不仅区间(0,1)有基数 c，而且，任何有限长的区间，以及所有实数集合本身，都具有这一相同的基数。另外区间(0,1)的不可数性说明，c 是一个与 $\aleph_0$ 不同的基数。这样，康托就用他的方法建立了超限数的序列。

所有这些讨论在认识有理数集与无理数集的内在区别方面开始显示出它的重要意义。有理数集与无理数集的区别绝不仅仅是前者可以写成有限小数或无限循环小数而后者则不能的问题。为了更清楚地说明这一点，康托只需要再增加一个结果：

定理 U　如果集合 B 与 C 是可数的，而集合 A 的所有元素属于 B 或者属于 C（或者属于两者），那么，集合 A 是可数的。（在这种情况下，我们说 A 是 B 与 C 的并集，记作 A＝B∪C。）

证明　所设的 B 与 C 的可数性保证了它们各自与自然数的一一对应关系：

$$\begin{array}{llllll} N: & 1 & 2 & 3 & 4 & \cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ B: & b_1 & b_2 & b_3 & b_4 & \cdots\cdots \end{array} \quad 和 \quad \begin{array}{llllll} N: & 1 & 2 & 3 & 4 & \cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ C: & c_1 & c_2 & c_3 & c_4 & \cdots\cdots \end{array}$$

在集 T 合 B 的元素中均匀地插入集合 C 的元素，我们可以在 N 与 A＝B∪C 之间建立起一一对应的关系：

$$\begin{array}{llllllllll} N: & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ A: & b_1 & c_1 & b_2 & c_2 & b_3 & c_3 & b_4 & c_4 & \cdots\cdots \end{array}$$

所以，集合 A 也是可数的。这一定理说明，两个可数集的并集也是可数的。证讫。

现在我们可以证明有理数集与无理数集的一个较重要的区别：我们已证明前者是可数的，对于后者，我们将断定它不可数。因为，假设无理数集是可数集。那么，根据定理 U，所有有理数（我们已证明其可数性）与所有无理数（我们假设其可数）的并集也应该同样是可数集。但是，这个并集恰恰是全部实数的集合，是一个不可数集。用反证法，我们可以断定，无理数过于丰富，以致无法与集合 N 构成一一对应关系。

不太正规地说，这意味着无理数在数量上大大超过有理数。实数远比有理数多的原因恐怕只能解释为实数轴几乎被漫无边际的无理数所淹没。数学家有时说“大部分”实数，常常是对无理数而言；至于有理数集，公认是一个非常重要的无穷集，尽管有理数在数轴上处处稠密，然而与无理数相比不过是沧海一粟。数不胜数的有理数当初是如此丰富，现在在实数集中却突然变得似乎无足轻重了。其实，有理数果真那样多吗？并非如此，对于康托来说，从基数的意义上讲，有理数的确非常稀少，而无理数则占据着统治地位。

为深入探索微积分的奥秘，康托证明了一些奇特的定理。他的研究对于认识实数集之间的内在区别当然有着重要意义，并有助于解释某些迄今为止尚不能解释的现象。如果说康托开始研究的还只是微积分的算术化问题的话，那么，他由此创立的集合论则呈现出极大的活力，对此，我们将在下章进行讨论。

后记

所有这一切已足以震古烁今,但康托 1874 年的论文中还包含着一个更加令人震惊的结果。康托不但证明了实数的不可数性，而且还把这一性质应用于一个长期困扰数学家的难题——超越数的存在。

我们已经注意到，所有实数的集合可以再细分为相对稀有的有理数集和比较丰富的无理数集。然而，让我们回忆一下，在第一章的后记中曾提到，实数还可以详尽无遗地分为两个相互排斥的数系——代数数和超越数。

代数数看来可以构成一个庞大的集合。所有有理数都是代数数，因为它们都是多项式方程的根，大量无理数（诸如$\sqrt{2}$或$\sqrt[3]{5}$）也属于这一集合。相比之下，超越数就极难得到。虽然欧拉最早猜测超越数的存在（即，并非所有实数都是比较驯顺的代数数簇），但第一个具有某种形式的超越数实例却是由法国数学家约瑟夫·刘维尔于 1844 年给出的。1874 年，当康托开始研究这个问题的时候，林德曼关于π是超越数的证明还没有出台。直到将近 10 年以后，这个证明才问世。也就是说，在康托发展他的无穷论时，人们还只发现了非常少的超越数。也许，这些超越数只是实数中的一种例外，而不是一种常规。

然而，乔治·康托已习惯于将例外转变为常规，在超越数问题上，他又一次成功地实现了这种转变。他首先证明了全部代数数的集合是可数的。基于这一事实，康托开始探索看似稀有的超越数问题。

首先设任意区间(a,b)。他已证明在这一区间中的代数数构成了一个可

数集；如果超越数也同样可数，那么，根据定理 U，(a,b)本身也应该可数。但是，他已经证明，区间是不可数的。这就表明，

无论在任何区间，超越数在数量上都一定大大超过代数数！

从另一个角度说，康托认识到，在 (a,b)区间实数远远多于代数数，这也许就是代数数相对比较少的原因。然而，所有这些实数是从哪里来的呢？它们必定是极大量存在的超越数。

这是一个真正引起争论的定理，因为人们毕竟只知道极少数几个非代数数的存在。而康托却十分自信地说，绝大多数实数是超越数，但他在作出这种推断的时候却没有展示出任何一个具体的超越数实例！相反，他只是“数”区间中的点，并由此认为，区间中的代数数只占很小一部分。这种证明超越数存在的间接方法真是令人吃惊。一位受人欢迎的数学史作家埃里克·坦普尔·贝尔以充满诗意的语言概述了这种情况：

“点缀在平面上的代数数犹如夜空中的繁星；而沉沉的夜空则由超越数构成。”

这就是康托 1874 年划时代的杰出论文所留下的宝贵遗产。许多数学家看到康托的结论，都惊异地摇头或干脆表示怀疑。在保守的数学家看来，比较无穷的大小简直就像是这位有点儿神秘兮兮的年青学者搞的一场浪漫而荒唐的恶作剧；断言有大量的超越数存在，却又举不出一个实例，真是十足的愚蠢。

乔治·康托听到了这些批评。但是，他坚信他的事业是正确的，他所做的还仅仅是开始。他后来的研究与他目前的这些发现相比，确实更见其辉煌。

第十二章
康托与超限王国
（1891 年）

无限基数的性质

乔治·康托究竟要去向何方呢？在他 1874 年的论文发表之后，康托对无穷点集的性质进行了更加深入的研究。他的研究向许多方向发展，并出人意料地打开了许多新的大门。但是，他在对有关无穷这一无人回答（准确地说，是无人提出）的问题的探索，却再清楚不过地表现出他那特有的勇敢和想象力。

康托一旦意识到他能够成功地定义许多超限基数，就立即感到需要使这种新基数的“小于”概念形式化。为此，他必然要再次依靠一一对应关系，但是，这一次，显然必须特别谨慎。我们在抽象地讨论这个问题之前，应再次提醒读者注意，在我们生活的原始社会里，人们只能数到 3。我们再回想一下，在这个原始社会里，有一位天才引入了 5，这是一个新的基数，是任何能够与她右手的手指构成一一对应关系的集合所具有的基数。那么，她怎样才能证明 3 小于 5 呢？（对于我们来说，这完全不费吹灰之力，因为我们惯于计算大于 3 的数。）我们假设她经过认真思考和艰苦寻觅之后，发现了一个右手只有三个手指的人，比如说，只有拇指、食指和无名指。这样，她就可以使那人右手的全部手指与她右手的部分手指构成一一对应的关系——即，使其拇指、食指和无名指一一相对。结果，她的右手还剩下两个手指无法求得对应，这多出来的两个手指就证明了 5 大于 3。

人们尝试将这一定义扩展到一般集合。如果集合 A 的全部元素能够与集合 B 的部分元素构成一一对应的关系，我们就说，集合 A 的基数小于集合B的基数，写作$\overline{\mathrm{A}}<\overline{\mathrm{B}}$。也就是说，如果A能够与B的子集构成一一对应关系，那么，A 就一定小于 B。

然而，非常遗憾，虽然这一定义在证明 3＜5 时十分完美，但应用于无穷集时，就不能令人满意了。例如，自然数集 N 和有理数集 Q。我们可以很容易地写出 N 集全部元素与 Q 的某一子集（即那些分子为 1 的正分数）之间的一一对应关系：

$$\begin{array}{cccccccccc} N: & 1 & 2 & 3 & 4 & 5 & 6 & \cdots\cdots & n & \cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ Q: & 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \cdots\cdots & \frac{1}{n} & \cdots\cdots \end{array}$$

我们无疑不希望利用这种对应关系推断出$\overline{\mathrm{N}}<\overline{\mathrm{Q}}$。事实上，我们已经知道，在集合 N 的所有元素与集合 Q 的所有元素之间存在着与此不同的另一种一一对应关系，所以，这两个集合具有相同的基数。乍一看，我们似乎陷入了进退维谷的尴尬境地。

康托发现，如果在开始时引入的不是“小于”，而是“小于或等于”的概念，就可以巧妙地摆脱这种困境：

□定义：设有集合 A 和 B，如果集合 A 的所有点与集合 B 某一子集的所有点存在一一对应关系，那么，我们就说，$\overline{A}\leq\overline{B}$。

注意到集合 B 的某一“子集”可能是集合 B 的全部点，在这种情况下，我们就得到$\overline{A}=\overline{B}$。当然，这与广义的$\overline{A}\leq\overline{B}$完全一致。并且，上述集合N的全部元素与Q集部分元素之间的一一对应关系也不过是证明了$\overline{N}\leq\overline{Q}$，这并不矛盾，因为两个集合都有基数$\aleph_0$。

现在，康托可以用严格的不等式给出一个定义来描述两个集合之间的基数性质：

□定义：设$\overline{A}\leq\overline{B}$（如前一个定义所述），如果A与B之间不存在一一对应关系，那么，$\overline{A}<\overline{B}$。

表面看来，这个定义似乎价值不大，但若仔细想一想就会发现，这个定义包含着一一对应关系的重要性质。因为，如果要证明$\overline{A}<\overline{B}$，我们首先必须找出集合A的所有点与B集的部分点之间存在着一一对应关系(因而证明$\overline{A}\leq\overline{B}$），然后，我们还必须证明，集合A的所有点与B集的所有点之间不存在一一对应的关系。问题很快就变得不那么简单了。

尽管如此，这个定义依然行之有效。例如，它为我们的原始朋友证明了 3＜5。也就是说，拇指、食指和无名指与右手五个手指中的三个手指所构成的一一对应关系证明了 3≤5；然而，却没有办法将她的全部五个手指与她伙伴的三个手指一一对应，所以，基数 3 与 5 不相等，结论只能是 3＜5。

至于无限基数，应用同一逻辑方法足以证明$\aleph_0<c$，因为我们可以很容易地发现集合 N 的全部点与区间(0,1)某一子集之间的一一对应关系：

$$\begin{array}{lcccccccc} N: & 1 & 2 & 3 & 4 & 5 & \cdots\cdots & n & \cdots\cdots \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ (0,1): & \frac{1}{\pi} & \frac{1}{2\pi} & \frac{1}{3\pi} & \frac{1}{4\pi} & \frac{1}{5\pi} & \cdots\cdots & \frac{1}{n\pi} & \cdots\cdots \end{array}$$

所以，$\overline{N}\leq\overline{(0,1)}$。但是，康托的对角线法证明，在这两个集合之间不存在一一对应关系。因而，$\overline{N}\neq\overline{(0,1)}$。综合这两个方面，我们得出结论，$\overline{N}<\overline{(0,1)}$——即，$\aleph_0<c$。

至此，康托已提出了一个比较基数大小的方法。请读者注意，这个定义的直接结果是一个在直觉上十分明显的事实，即，如果 A 是 B 的子集，那么，$\overline{A}\leq\overline{B}$。也就是说，我们肯定能够使集合A的每一点与它自己相匹配，在集合 A 的所有元素与集合 B 的某一子集之间建立起一一对应的关系。所以，一个集合的基数大于或等于其任何子集的基数。在一系列违背直觉的命题中，这一命题似乎还算令人安心。

康托在比较基数大小的基础上，又提出了一个非常重要，而且，据他认为，是一个非常关键的论断：

如果$\overline{A}\leq\overline{B}$并且$\overline{B}\leq\overline{A}$那么，$\overline{A}=\overline{B}$

如果我们只限于有限基数，这一论断看来非常明确。但是，如果我们将其应用于超限基数，就显得不那么明确了。让我们仔细想一想康托提出的问题：如果在 A 集的全部与 B 集的部分之间存在着一一对应关系（即，

$\overline{A} \le \overline{B}$），并且，在B集的全部与A集的部分之间同样也存在着一一对应的关系（即，$\overline{B} \le \overline{A}$），那么，我们就可以断定，在A集的全部与B集的全部之间必然存在着一一对应的关系（即，$\overline{A} = \overline{B}$）。但是，人们从哪里才能找到这最后的对应关系呢？稍加思考，我们会发现这个论断的确具有深远的意义。

乔治·康托虽然从未能对这一命题作出令人满意的证明（足以说明其复杂性），但却仍然认为这个命题是正确的，也许这恰恰表明了他对他的集合论的“合理性”始终抱有坚定的信念。幸运的是，这一定理由两位数学家——恩斯特·施罗德（于 1896 年）和费利克斯·伯恩斯坦（于 1898 年）各自独立地作出了证明。由于这个定理是由几个人共同创立的，所以，我们今天称之为“施罗德-伯恩斯坦定理”，但有时也称作“康托-伯恩斯坦定理”或“康托-施罗德-伯恩斯坦定理”，或把这些名字按其他方式排列。我们暂且抛开这些名称不谈，这一定理对于研究超限基数，是一个十分有用的工具。

虽然这一定理的证明超出了本书范围，但我们可以阐明这一定理在确定所有无理数的集合 I 的基数中所起的重要作用。我们在前一章中已看到，无理数集是不可数的；也就是说，无理数集的基数大于$\aleph_0$。但是，我们并没有明确地给出这个基数。要确定这个基数，我们就可以应用施罗德-伯恩斯坦定理。

首先，无理数集是实数集的一个子集，根据前一章的评述，我们知道，$\overline{I} \le c$。另一方面，考虑到每一个实数与一个无理数所构成的对应关系，我们定义如下：如果 $x=M.b_1b_2b_3b_4\cdots\cdots b_n\cdots\cdots$是一个小数形式的实数，M 是这个小数的整数部分，那么，我们就可以得到与 x 相伴的实数

$$y=M.b_10b_211b_3000b_41111b_500000b_6111111\cdots\cdots$$

也就是说，我们在第一位小数后面插入一个 0，在第二位小数后面插入两个 1，在第三位小数后面插入三个 0，等等，依此类推。例如，对应于实数 x=18.1234567……的是

$$y=18.1021130004111150000061111117\cdots\cdots$$

而与实数 x=－7.25=－7.25000……相对应的则是

$$y=-7.20511000001111000000111111\cdots\cdots$$

无论我们选用什么数值的实数 x，与之相对应的 y 都有一个既不终止，也不循环的小数展开式，因为我们在这个小数展开式中得到越来越长的连续 0 或连续 1 的数组。因此，每一个实数 x 都与一个无理数 y 相对应。

并且，这种对应是一一对应的。因为，如果我们已知一个 y 值，比如 5.304114000711111000002……，我们就能够从中“分解”出一个，并且，只有一个可能与之对应的 x 值，就本例而言，x=5.344712……。我们应该注意到，并不是每一个无理数最终都能够与一个实数对应。如，无理数 y=$\sqrt{2}=1.414159\cdots\cdots$，在其小数展开式中就没有这种形式的0和1数组能够与任何实数 x 相对应。

在全部实数与部分无理数之间的这种一一对应关系表明$c \le \overline{I}$。但是，我们前面已讲过，$\overline{I} \le c$，因而，根据施罗德-伯恩斯坦定理，我们可以断定，无理数集的基数是 c，与全部实数集的基数相同。

由康托提出，并由施罗德和伯恩斯坦证明的这个定理，使超限基数这一大难题成功地得到了解决，但是，康托的奇妙问题层出不穷。另一个问题是，是否存在任何大于 c 的基数。根据以前的对应关系，康托感到这个问题的答案应该是肯定的，并觉得他知道如何得到一个更充分的点集。

康托认为，发现一个大于一维区间(0,1)基数的关键是要在一个由 x 轴上的区间(0,1)与 y 轴上的区间(0,1)所构成的二维正方形中去寻觅，如图 12.1 所示。康托在 1874 年 1 月写给朋友理查德·狄德金的信中问道，区间和正方形这两个点的集合是否能够构成一一对应的关系？他近于肯定地认为，在二维正方形与一维线段之间不可能存在这种对应关系，因为前者似乎显然具有更多的点。虽然作出证明可能十分困难，但康托却认为作证明也许是“多余”的。

然而，有趣的是，这一几乎多余的证明却从未能够作出。康托尽管尽了最大努力，但始终未能证明在区间与正方形之间不可能存在一一对应的关系。后来，1877 年，他发现他原来的直觉是完全错误的。这种一一对应的关系确实存在！

为了证明这一令人吃惊的事实,我们令 S 表示由全部有序偶(x,y)构成的正方形，在这里，0＜x＜1，0＜y＜1。我们只要简单地将区间(0,1)中的点z与S中的一对有序实数$(z,\frac{1}{2})$相配，就可以很容易地构成单位区间全部与单位正方形 S 中部分之间的一一对应关系。根据我们前面的定义，我们推断，$\overline{(0，1)}\le\overline{S}$。

另一方面，对于 S 中的任何点(x,y)，设其横坐标 x 与纵坐标 y 都是无穷小数。即，$x=0.a_1a_2a_3a_4\cdots\cdots a_n\cdots\cdots$和 $y=0.b_1b_2b_3b_4\cdots\cdots b_n\cdots\cdots$。如我们在第十一章中所述，我们认为，这些小数展开式是唯一的——对于一个结尾可以用 0 的无限循环或 9 的无限循环表示的小数，我们采用前一种表示方法，而不采用后者，所以，我们用小数 0.2000……，而不用其等价小数$0.1999\cdots\cdots$来表示$\frac{1}{5}$。

我们采用这一约定,并根据以下定义,将 S 中的每一个点(x,y)与(0,1)中的点 z 相对应：

$$z=0.a_1b_1a_2b_2a_3b_3a_4b_4\cdots\cdots$$

例如，按上面规律对 x 和 y 的小数位重新组合，就可以使单位正方形中的一对有序实数$(\frac{2}{11}，\frac{\sqrt{2}}{2})=（0.18181818\cdots\cdots，0.70710678\cdots\cdots）$与单位区间中唯一的一个点 0.1780178110861788……相对应。没有比这再简单的了。同样，如果已知区间中的一点对应于 S 中的某一点，我们可以通过分解小数位的方法，使之回到唯一的有序数偶。也就是说，如果 z=0.93440125……，那么，在正方形中一定有由它规定的唯一的一对有序实数

$$(x,y)=(0.9402\cdots\cdots,0.3415\cdots\cdots)$$

我们注意到，根据这种对应关系，并非单位区间中的每一个点都能够与正方形中的某一点对应。例如，$(0,1)$区间中的点$z=\frac{6}{55}=0.109090909$

……可以分解为有序偶（0.19999……，0.0000……）。但是，我们已完全排除了应用 0.1999……，而代之以与其等价的小数 0.2000……；更糟糕的是，第二个坐标 0.0000……=0，严格地说，并不位于 0 与 1 之间，所以，分解这个小数的结果使我们超出了S的范围。换言之，$\frac{6}{55}=0.1090$ 9090……不对应于正方形内的任何一点。

然而，我们毕竟在 S 的全部点与(0,1)的部分点之间构成了一一对应关系，因此，我们认为，$\overline{S}\le\overline{(0,1)}$。将这一事实与前面的不等式$\overline{(0,1)}\le$ $\overline{S}$联系起来，就可以应用施罗德 - 伯恩斯坦定理，推断出$\overline{S}=\overline{(0,1)}=c$。

这一讨论表明，尽管维数不同，但正方形中的点并不比区间内的点多。这两个集合都有基数 c。至少可以说，这个结果是十分令人吃惊的。康托在 1877 年写给狄德金的信中报告了这一发现，并惊呼，“我发现了它，但简直不敢相信！”

那么，我们到哪里去找大于 c 的超限基数呢？康托可以很容易地证明，一个比较大的正方形，甚至整个平面中的全部点，都具有与单位区间(0,1)相同的基数。即使在三维立方体中，也没有发现更大的基数。这样看来，c 似乎是最高一级的超限基数。

但是，事实却证明并非如此。1891 年，康托成功地证明了更高一级超限基数的存在，而且，是令人难以置信地大量存在。他的研究结果，我们今天通常称之为康托定理。如果说他一生中证明了许多重要定理，那么，这个定理的名称就表明了它所得到的高度评价。这个定理像集合论的任何定理一样辉煌。

伟大的定理：康托定理

为了讨论这一证明，我们需要引入另一个概念：

□定义：　设有集合 A，A 的所有子集的集合称为 A 的幂集，记作 P[A]。

这个定义看来非常简单。例如，如果 A= {a，b，c} ，那么，A 有 8 个子集，因而，A 的幂集就是包含这 8 个子集的集合，即：

P[A]= { {} ，{a} ，{b} ，{c} ，{a,b} ，{a,c} ，{b,c} ，{a,b,c} }

请读者注意，空集 {} 和集合 A 本身是 A 的幂集的两个元素；无论我们采用什么样的集合 A，这都是正确的。我们还应注意，幂集本身也是一个集合。这一基本事实有时容易被忽视，但在康托的思想中，这个事实却有着重要意义。

显然，在我们上述例子中，幂集的基数大于集合本身的基数。也就是说，集合 A 包含 3 个元素，而它的幂集则包含 $2^3=8$ 个元素。我们不难证明，一个包含 4 个元素的集合有 $2^4=16$ 个子集；一个包含 5 个元素的集合有 $2^5=32$ 个子集；总之，一个包含 n 个元素的集合 A 有 2^n 个子集。我们可以用符号来表示，即，$P\overline{[A]}=2^n$。

但是，如果 A 是一个无穷集，又将如何？无穷集的幂集其基数是否同样大于集合本身的基数呢？康托定理回答了这一引起争论的问题：

定理 设A为任意集合，那么，$\overline{A} < \overline{P[A]}$。

证明 为证明这一定理，我们必须依据本章前面所介绍的康托关于超限基数之间严格不等式的定义。显然，我们可以很容易地发现在 A 与部分 P[A]之间存在着的一一对应关系，因为如果 A=｛a,b,c,d,e,……｝，我们就可以使元素 a 对应于子集｛a｝，使元素 b 对应于子集｛b｝，等等。当然，这些子集｛a｝，｛b｝，｛c｝……仅仅是 A 的全部子集中微不足道的一小部分，所以，这种一一对应关系就保证了$\overline{A} \leq \overline{P[A]}$。

到此为止，一切都很简单。但还有必要证明 A 与 P[A]没有相同的基数。我们采用间接证明的方法，首先假定它们的基数相同，然后从中导出逻辑矛盾。即，我们假设在 A 的全部与 P[A]的全部之间存在着一一对应关系。为便于论证，我们将采用一个与这一假定一致的例子，以备后面参照：

A的元素		P[A]的元素（即，A的子集）
a	↔	｛b，c｝
b	↔	｛d｝
c	↔	｛a,b,c,d｝
d	↔	｛ ｝
e	↔	A
f	↔	｛a,c,f,g,……｝
g	↔	｛h,i,j,……｝
·		·
·		·
·		·

于是，这个排列表示在 A 的全部元素与 P[A]的全部元素之间存在着假定的一一对应关系。请注意，在这种对应关系中，A 的某些元素属于它们所对应的子集；例如，c 是与之相对应的集合｛a,b,c,d｝的元素。而另一方面，A 的某些元素则不属于它们所对应的子集；例如，a 就不是其对应子集｛b,c｝的元素。

令人不可思议的是，这种将互不相容的东西一分为二的做法提供了导致本证明逻辑矛盾的线索，因为我们现在可以对集合 B 作出如下定义：

B 是原集合 A 中每一个不属于它所对应的子集的元素的集合。

参照上述假设的对应关系，我们看到，a 以及 b（因为 b 不是｛d｝的元素）、d（d 当然不属于空集）和 g（不是｛h,I,j,……｝的元素）都属于集合 B。但是，c、e 和 f 就不是集合 B 的元素，因为它们分别属于｛a,b,c,d｝、A 本身和｛a,c,f,g……｝。

因此，集合 B=｛a,b,d,g,……｝。这样构造的 B 是原集合 A 的子集。所以 B 属于 A 的幂集，因而必然会出现在上述对应关系右边一列的某个位置。但是，按最初假定的一一对应关系，我们同样断定在左边一列一定有 A 的某一元素 y 对应于 B：

·　　　·
·　　　·
·　　　·
y ⟷ B
·　　　·
·　　　·
·　　　·

到目前为止，一切顺利。但是现在我们提出一个致命的问题：“y 是 B 的元素吗？”当然，有两种可能：

第一种情况　假设 y 不是 B 的元素。

那么根据我们最初对 B 所下的定义，“……原集合 A 中每一个不属于它所对应的子集的元素的集合”，我们看到，y 理所当然是 B 的成员，因为在这种情况下，y 不是它所对应的集合的元素。

换句话说，如果我们首先假定 y 不属于 B，那么，我们就被迫得出结论，y 应当是 B 的元素。显然这是自相矛盾的，所以我们排除第一种情况，因为这是不可能的。

第二种情况　假设 y 是 B 的元素。

我们再次求助于 B 的定义。因为第二种情况假定 y 属于 B，那么，y 应当符合 B 的定义；即 y 不是它所对应的集合的元素。太遗憾了！与 y 对应的集合恰恰是 B，因此，y 不可能是集合 B 的元素。

这样，由于第二种情况假定 y 属于 B，我们不得不直接得出结论，y 不是 B 的元素。作为逻辑结构，我们再次走进了死胡同。

一定是什么地方出了毛病。第一种情况与第二种情况是仅有的两种可能，但这两种可能都导致了逻辑矛盾。我们断定，在论证中的某个地方一定有一个假设是错误的。当然，问题正是开始时我们所假定的在 A 与 P[A] 之间存在着一一对应关系。我们的悖论显然摧毁了这一假设：不可能存在这种对应关系。

最后，综合我们的结论，$\overline{A}\leq\overline{P[A]}$，但$A\neq\overline{P[A]}$，至此，我们已证明了康托定理：对于任意集合A，$\overline{A}<\overline{P[A]}$。　证讫。

也许，用一个有限集作为具体例子，我们可以清楚地看到康托的天才。令 A={a,b,c,d,e} ，并在集合 A 的元素与其幂集的某些成员之间建立对应关系：

A 的元素		A 的幂集的元素（即，A 的子集）
a	⟷	{a，c}
b	⟷	A
c	⟷	{a，e}
d	⟷	{d}
e	⟷	{a，b，c，d}

回忆一下 B 的定义，即，集合 A 中那些不属于它所对应的集合的元素，我们看到，B={c,e} 。

康托注意到问题的关键是，B 不可能出现在上述对应关系的右列，

因为逻辑证明，不存在可能与之对应的元素。对于 A 与 P[A]之间任何假设的对应，康托的证明令人击节称赞之处在于他巧妙地描述了 A 的幂集不可能的一个成员（也就是 B）对应于 A 的某一元素。这就直接否定了任何一个集合与其幂集之间存在着一一对应关系的可能。

我们有必要停下来思考一下康托定理更深一层的含义。康托证明，无论我们最初采用什么样的集合，其幂集严格地说具有更大的基数。用他自己的话说：

“……可以用另一集合 M 置换任何已知集合 L，构造一个新的集合，使其基数大于 L 的基数。”

这样，在寻找基数大于 c 的集合这一长期探索过程中，我们没有着眼于平面中的正方形或三维空间的立方体。而是代之以区间(0,1)内点的所有子集的集合P[(0,1)]。根据康托定理，$c=\overline{\overline{(0,1)}}<\overline{\overline{P(0,1)}}$，于是我们发现了更大的超限基数。

现在，我们来回忆一个基本事实，按照幂集的定义，它理所当然地也是一个集合。因此，我们可以构造 P[(0,1)]的幂集，即构造(0,1)的所有子集的集合的所有子集的集合。尽管 P[(0,1)]的确是一个非常惊人的集合，但上述推理表明，$\overline{\overline{P[(0,1)]}}<\overline{\overline{P[P[(0,1)]]}}$。

这个妖怪一旦逃出魔瓶，就再没有什么能够阻止乔治·康托了。因为我们显然能够无限地重复这个过程，并由此生成一个越来越长的不等式链：

$$\aleph_0<c<\overline{\overline{P[(0,1)]}}<\overline{\overline{P[P[(0,1)]]}}<\overline{\overline{P[P[P[(0,1)]]]}}<\cdots\cdots$$

简直没有喘息的时间。乔治·康托不仅打开魔瓶放出了第一个超限基数（$\aleph_0$），进而又发现了甚至更加漫无边际的无穷基数(c)，而且，通过反复应用康托定理，还给出了一个生成更大超限基数的永无尽头的不等式链。这是一个没有结尾的故事。

毫不夸张地讲，这个定理，以及康托所有关于无穷的深奥理论，都引起了反对派不绝于耳的喧嚣。的确，他推动数学进入一片未被开垦的处女地，在那里，数学开始并入哲学和形而上学的王国。值得注意的是，乔治·康托的数学中的形而上学含义并非对他无关紧要。据现代康托的权威传记作家约瑟夫·多邦记载，康托在他的超限理论中发现了一种宗教意义，并认为自己“不仅是上帝的信使，准确地记录、转述和传送新发现的超限数理论，而且，也是上帝的使节”。康托自己写道：

“我毫不怀疑超限数的正确性，因为我得到了上帝的帮助，而且，我曾用了二十多年的时间研究各种超限数；每一年，甚至每一天我在这一学科中都有新的发现。”

这一段文字表明，宗教在很大程度上已成为康托思想的中心。我们只要回想一下他父母的混合宗教背景，就可以想象出康托家庭中必定不乏各种各样的神学讨论。也许，这更增强了他对神学的兴趣。无论如何，不论是在数学，还是在其他领域，他的思想时时显示出宗教色彩。

这种态度使这位神秘的怪人不能见容于他的批评者。康托声称他的数学乃是上帝的信息，无怪那些反对者可以随意攻击他的激进的无穷论。康托不仅迷恋神学，而且热中于证明莎士比亚的剧作乃是由弗朗西斯·培根捉刀，不免更加损害了他的形象。也许，对这一切，他的同事只感到有点

儿古怪，而当他声称发现了有关第一位不列颠国王的资料，并且，“只要这些资料一公布，必然会使英国政府感到恐惧”时，许多人就惊得目瞪口呆了。这让人很难不把乔治·康托看作某种狂人。

还有他的数学。在他的祖国德国和其他一些地方，都有许多保守分子大叫大喊地反对他的理论，康托与某些很有影响的数学家逐渐交恶。当然，这些反对意见并非都是盲目的反动，因为康托的数学确实提出了一些令人莫名其妙的问题，即使那些善意的数学家也深感困惑。我们在后记中将对这样的一个问题进行讨论。

在批评康托的人中，有一位名叫列奥波德·克罗内克（1823—1891年），他是德国数学界很有影响的人物，并固定在颇具声望的柏林大学执教。这所大学曾培养出著名的维尔斯特拉斯和他的出色的学生（包括康托自己在内）。康托在哈雷大学执教，但哈雷大学的声望较之柏林大学，则大为逊色。因此，康托渴望能在柏林大学任职。他对于被“放逐”到二流大学，总有一种强烈的怀才不遇感，并常常把造成这种情况的原因归结为克罗内克的迫害。康托在与其对手间的相互攻忤中，明显地表现出一种妄想狂的倾向。在此过程中，康托既攻击了敌手，也得罪了朋友，也就更难有机会在柏林大学谋职。

毫不奇怪，乔治·康托由于生活的失意和对最神秘的无穷概念的拼命钻研，多次受到精神病的折磨。他第一次发病是在 1884 年，当时他正在狂热地研究一个称为“连续统假设”的猜想，力求对之作出简短的证明。一种流行的看法认为，除了克罗内克及其他人的迫害以外，数学的压力也是造成他精神崩溃的原因。现代对康托医学资料的分析认为这种看法夸大其词，因为有迹象表明，康托表现出一种双相（即狂郁性）精神病的症状，在任何情况下都有可能使他精神崩溃。也就是说，他在受到人身攻击或遇到数学困难时，都有可能发病，但他的疾病似乎还有更为深刻的原因。

不管怎样，他的病不断发作，而且，变得日益频繁。1884 年，康托经过一段时间的住院治疗之后，虽然情况好转，但仍有复发的可能。他除了在数学与职业方面颇感失意以外，1899 年，他的爱子鲁道夫的意外死亡又使他受到了一个沉重的打击。1902 年，康托再次住进了哈雷的神经病医院，后来，1904 年、1907 年和 1911 年，又多次住院治疗。然而，他出院后，又周期性地出现抑郁症状，他常常一动不动，静静地坐在家中。

康托的一生无疑是困苦的一生。1918 年 1 月 6 日，他在因精神病发作再次住院期间，不幸逝世。对于一位伟大数学家来说，这真是一个令人悲痛的结局。

回顾乔治·康托的生活和工作，人们不禁将他与其同时代的美术大师樊尚·凡高相比。二人颇有一些相似之处。康托的父亲笃信宗教，而凡高的父亲则是一位荷兰牧师。他们两人都深爱艺术，热中文学，并喜欢写诗。我们回想一下，凡高像康托一样，也有一种古怪而反复无常的个性，最后，他甚至疏远了像保罗·高庚这样的朋友。他们对自己的工作都有一种极强烈的献身精神。当然，他们两人也都患有精神疾病，因此住院治疗，而且给他们造成了沉重的思想负担，因为他们时时担心疾病的再次发作。

最重要的是，凡高和康托两人都是革命者。凡高在短暂而辉煌的生涯中，使美术超越了印象主义的范畴；同样，康托也推动数学沿着意义深远的新方向发展。无论人们对这位伟大而不幸的数学家如何评说，我们都不

禁对他勇敢地以一种崭新的方式探索无穷的性质肃然起敬。

康托尽管面对重重困境，却从未对他工作的价值丧失信心。他在谈到争议很大的有关无穷的观点时写道：

“我认为是唯一正确的这种观点，只有极少数人赞同。虽然我可能是历史上明确持有这种观点的第一人，但就其全部逻辑结果而言，我确信我将不是最后一人！”

的确，他不是最后一人。虽然多少代的数学家都曾探索过古老的几何、代数和数论的问题，但乔治·康托却开创了全新的境界。由于他既提出、又回答了前人不曾想到的问题，所以，将他的理论称之为自古希腊以来第一部真正具有独创性的数学，也许是最恰当不过的了。

后记

我们曾提到过集合论的某些问题，对于这些问题，即使像康托这样的伟大天才，也未能解决。其中最令人困惑的问题来自康托的发现中一些费解的矛盾——逻辑学家称之为“悖论”。也许，其中最简单者即来自康托的定理。

如果我们构造一个一切集合的集合，并称之为 U（即“泛集”）。这是一个令人难以置信的巨大集合。它包含一切概念集、全部数集、所有数集的子集的集合，等等。在集合 U 中，我们可以找到每一个存在的集合。在这个意义上说，U 不可能再扩大；因为它已经包含了全部可能存在的集合。

但是，现在我们把康托定理应用于U。康托已经证明，$\overline{U}<\overline{P[U]}$，这显然表明 P[U]远远大于 U 本身。这样，在康托集合论的核心出现了灾难性的矛盾。

1895 年，康托发现了这一悖论；其后几十年间，数学界一直在寻求一种方法，以弥补这一悖论所造成的逻辑缺陷。问题的最终解决看来需要建立集合论的形式公理系统（正如欧几里得建立了几何学的公理系统），通过精心地选择公理，合法地将上述悖论排除在外。从逻辑上说，这并非易事。但是，最后终于将集合论“公理化”，这一新体系谨慎地明确规定了什么是和什么不是“集合”。在这一体系中，“泛集”在任何意义上不再是集合；一切集合的集合就被从集合论公理所规定的集合中予以排除。于是，悖论也就魔术般地消失了。

这一解决方法显然是削足适履，即用一个公理，像外科手术一样，精确地切除集合论中使人困惑的部分，而保留康托理论中所有完美的部分。康托的集合论，现在被称为“朴素集合论”，以区别于公理化的集合论。后者尽管非常艰深，并需要专门的知识，但如今已成为集合论的坚实基础。它标志着数学家大卫·希耳伯特所表述的一种激情的胜利，他曾大声疾呼：“没有人能把我们从康托为我们创造的乐园中赶走。”

但是，还有另外一个问题，康托也未能完美地解决，这一问题至少像悖论的出现那样使他忧心忡忡。实际上，一些人认为，康托对这个问题年复一年地刻苦钻研，也是造成他精神崩溃的重要因素。这个问题现在称为康托的“连续统假设”。

这个问题讲起来非常简单。连续统假设断定，在 $\aleph_0$ 与 c 之间不存在别

的超限基数。在这一意义上，基数$\aleph_0$与 c 的性质很像整数 0 与 1。0 与 1 是前两个有限整数，在它们两者之间不可能插入任何其他整数。康托猜测，$\aleph_0$与 c 这两个超限基数也有相似的性质。

从另一个角度讲，连续统假设表明，实数的任何无穷子集或者可数（在这种情况下，它有基数$\aleph_0$），或者能够与(0,1)构成一一对应的关系（在这种情况下，它有基数 c）。没有中间的可能性。

康托在他的数学生涯中，用了很多时间来钻研这个问题。1884 年，即他的精神病第一次发作的那一年，他作出了一次重大努力。是年 8 月，康托认为他的努力已获成功，便写信给他的同事古斯塔夫·米塔格-列夫勒，宣称他对这个问题已作出了证明。但是，三个月以后，他在随后的信中不仅收回了他 8 月份的证明，而且还声称他现在已证明出连续统假设是错误的。这种观点的根本改变仅仅持续了短短的一天，之后，他又再次写信给米塔格-列夫勒，承认他的两个证明都有错误。康托不是一次，而是两次承认他所犯的数学错误，却仍然搞不清他的连续统假设究意是否正确。

我们在前一章的后记中曾讲过，如果康托对他的猜测作出了证明，那么，他就能够，比如，很容易地确定超越数的基数。如本章所述，超越数构成了实数的不可数子集，因此，它一定具有基数 c 只要康托能够证明他的连续统假设，一切都会变得如此简单。

然而，他却始终没有成功。尽管他几十年来付出了艰辛的努力，但是，直到逝世，也未能取得任何进展。这也许是他一生中无法摆脱的最大困扰和遭受到的最大挫折。

实际上，并非只有乔治·康托一人在探索这个问题的答案。1900 年，希耳伯特审视了大量未解决的数学问题，并从中选出 23 个问题作为对 20 世纪数学家的重大挑战。在这 23 个问题中，第一个就是康托的连续统假设，希耳伯特称之为一个“……似乎非常有理的猜想，然而，尽管人们竭尽全力，却没人能够作出证明。”

在对集合论这一貌似简单的猜想作出某些突破之前，数学家们还需要殚精竭虑，努力一番。进入 1940 年，一个重大的突破在 20 世纪最非凡的数学家库特·哥德尔（1906—1978 年）的笔下产生。哥德尔证明连续统假设在逻辑上与集合论公理系统彼此相容。也就是说，不可能用集合论公理系统证明连续统假设不成立。如果康托还活着的话，他一定会对这一发现感到无比的振奋，因为这似乎证明了他的猜想是正确的。

果真如此吗？哥德尔的结果无疑并没有证明这一假设。这个问题依然悬而未决。1963 年，美国斯坦福大学的数学家保罗·科恩（1934—）证明，我们同样不能用集合论公理系统证明连续统假设成立。综合哥德尔和科恩的工作，连续统假设以一种最奇特的方式得到了解决：这一假设不能用集合论公理系统判定其真伪。

这似乎再现了数学史上我们熟悉的一幕。两千多年前，欧几里得引入了平行线公设，随后数代人绞尽脑汁，试图从其他几何公设中推出这一公设。后来我们认识到，这是根本不可能的，因为平行线公设完全独立于其他几何原理；我们不能证明它是对的，也无法证明它是错的。它就像一个离开海岸的孤岛，形单影只地自成体系。

康托的连续统假设在集合论领域中处于类似的地位。是否采用连续统

假设完全取决于数学家的口味，这一假设变成了一种选用的理论，而不是必须采用的定理。如果我们要探索在$\aleph_0$与 c 之间不存在其他超限基数的集合论，我们毫无疑问会非常愿意接受连续统假设作为公设，从而满足我们的需要。相反，如果我们更喜欢另一种不同的集合论，我们同样可以按我们的需要抵制连续统假设。连续统假设这一性质与欧氏几何和非欧几何十分相似。这种异曲同工的结局把当代一个最著名的疑难问题与古希腊的一个经典难题出人意料地联系了起来。它表明，即使在数学中，事物越变化，越具有相同性。

那么，康托连续统假设的证明这一悬而未决的问题究竟如何呢？根据 20 世纪哥德尔和科恩的研究结果，我们看到，康托所面对的不是一项困难的工作，而是一项完全没有希望的工作。这一事实就像是对这位忧虑的数学家一生的一段辛辣写照。

然而，乔治 · 康托的失败丝毫无损于他数学遗产的光辉。1888 年，他对自己大胆闯入超限王国作出了评价，我们不妨摘录于此：

“ 我的理论坚如磐石；射向它的每一枝箭都会迅速反弹。我何以得知呢？因为我用了许多年时间，研究了它的各个方面；我还研究了针对无穷数的所有反对意见；最重要的是，因为我曾穷究它的根源，可以说，我探索了一切造物的第一推动力。 ”

结束语

随着康托的超限基数喧嚣着走向无限的无穷大，至此，我们结束了我们参拜大师的数学之旅。这是一个漫长的旅程——从希俄斯的希波克拉底一直到 20 世纪。我希望这一旅程能够以其强大的阵容和辉煌的表演给人留下深刻的印象。这是一段值得口口相传的故事。

我们在第四章讨论拉玛努扬时曾提到过 G.H.哈代，他对数学证明中的美学有一种强烈的感受。哈代认为，真正的伟大定理应该具有三个特点，即，精练、必然和意外。我认为，这些性质极恰当地概括了我们所讨论的定理的特征。欧几里得对素数无穷性的证明堪称简明、优雅和“精简”。约翰·伯努利的一系列无穷级数必然导致调和级数的发散性，犹如人们在讲到阿基米德数学时那样，“只要看上一眼，就立刻相信，本来你也能够发现它。”我们讨论的许多命题，从新月形的化方求积，到三次方程的可解，以及乔治·康托所发现的一切，都是令人感到非常意外的。总之，我希望哈代会认可我所选择的这些“伟大定理”。

最后，我将以两段引文结束本书。这两段引文虽然相距 1500 年，但却传达了几乎完全同一的思想。第一段引文出自 5 世纪希腊评注家普罗克洛斯之笔：

“所以，这就是数学：她赋予自己的发现以生命；她令思维活跃，精神升华；她烛照我们的内心，消除了我们与生俱有的蒙昧与无知。”

我们在本书的序言部分曾引述过 20 世纪伯特兰·罗素的一段话，最后，我再引述他的另一段话。罗素认识到数学中的美，他也像其他任何人一样，尽力描绘这种美。我最后引述他的一段评论，并希望它能够代表读者对本书中这些数学杰作的感受：

“正确地说，数学不仅拥有真理，而且，还拥有极度的美——一种冷静和朴素的美，犹如雕塑那样，虽然没有任何诱惑我们脆弱本性的内容，没有绘画或音乐那样华丽的外衣，但是，却显示了极端的纯粹和只有在最伟大的艺术中才能表现出来的严格的完美。”