序

1987年部颁的中学数学教学大纲把“培养学生对数学的兴趣”也列为教学目的之一，这完全正确，十分必要。因为，学生大都觉得数学又难学又“枯燥”，所以，兴趣是最好的老师这句名言对数学课特别重要：没兴趣就学不好数学！

我们看到，学生对数学课大都努力而为，勉力而为，但是，推动力主要是社会、学校、老师、家长所施加的压力，他们大都对数学没有多少真正的兴趣，师生因之往往有“力尽关山未解围”之类的苦恼感、压迫感。

我们也看到，即使被动的勉力而为，也会有好处：下了工夫，解题能力会增强些，考试成绩会好些，于是多少感到快慰，受到鼓励。如果继续努力下去，就会滋生或增进学习的积极性和主动性，相对地逐渐冲消被动性、畏缩心理和厌烦情绪，实现难而枯燥→难而颇有趣→难而很有趣的转变。

数学的中心活动就是解题。

解题，就是应用已经学过的知识来解决“前提—结论”之间或“已知—求解”之间的各种疑难和障碍。对此，正确而有效的办法就是每次都认真仔细地考察和尝试，哪些定理、定义、公理用得上？怎样用？这样，每次都会对所学知识的一部分之正确理解、深刻理解、正确运用、灵活运用有所增益和提高，同时，解题能力有所增益和提高，学习兴趣有所增益和提

高。这之间，教师课内外的主导作用应当是重要的外因。

解题能力提高必然导致成绩提高，困难减少，从而兴趣增加；随之，兴趣上升有助于学习效率和质量的改进，从而又反过来促进能力的发展和智力的发展。这样，解题能力的变化与学习兴趣的变化互为因果，互相促进，并且归结为教学质量和效果的不断提高。但是，这种互为因果的变化也可能是：

能力发展迟滞⇄兴趣衰退⇄效率和质量下降

结果究竟如何，主要看教师的"主导作用力"的方向和大小。

所以，有计划地在课内外致力于学生对数学兴趣的培养和解题能力的培养，使之互为因果地互相促进，是提高教学效率和质量的一个好办法。坚持不懈，迟早能达到教和学都能事半功倍的美妙境界，达到"师逸而功倍，又从而用之"的理想境界。

许多数学家、数学教育家因此提倡以培养解题能力为关键的教学思想，著名的波利亚（G.Polya）尤为热心，他在这方面的一系列著作，如

Mathematical Discovery—On understanding, learning, and teaching problem solving(2 vols)

How to Solve It?

Mathematic and Plausible Reasoning(2 vols)

等等，多年来蜚声世界，传诵不衰。我国从1981年起也陆续出版它们的中译本（数学的发现——对解题的理解、研究和讲授，共两卷；怎样解题，一卷；数学与猜想，共两卷），产生着越来越巨大而深远的积极影响。曾晓新老师这本《数学的魅力》也是这方面的一本佳作，它比较适合我国中等程度这个层

次的读者；并且有它自己的特点：

此书把数学智育和数学美育密切结合起来，互相推动；选取了一些饶有方法论、认识论意义和创造精神的数学史料和数学问题，加以深入浅出的阐述，夹叙夹议；展示数学美的真实性和千姿百态，有助读者培养和发展对数学的美感，从而滋生对"困难、但很有趣"的数学的兴趣；与此同时，协助读者扩展眼界，活跃思想、提高数学才能和创造精神。

但愿各个学科、各个层次的饶有兴味、饶有益智的这种读物多些，更多些。

麦雨农　1988.5.

前　言

本书叫做《数学的魅力》，顾名思义，是想让读者看到：数学美是怎样以她迷人的魅力，激起了众多的数学家为之奋斗，吸引了无数爱好者为之倾倒。

著名的英国数学家哈代（G.H.Hardy）曾说：“现在也许难以找到一个受过教育的人对数学美的魅力全然无动于衷。”“没有什么比数学更为‘普及’的学科了。大多数人都能欣赏一点数学，正如大多数人能欣赏一支令人愉快的曲调一样。”

本书试图通过数学史与数学方法论上的一些典型事例，向读者展现数学美的若干侧面。数学美不同于艺术美，艺术美主要是通过形象和感觉唤起人们情感的愉快和共鸣，而数学美则是通过抽象和深思给予人们理智的满足和神往。当你对数学所揭示的自然规律浮想联翩时，当你对数学本身的简洁与和谐回味无穷时，当你对数学家们的成就拍案叫绝时，当你对复杂深奥的数学问题豁然开朗时，你的内心会有说不出的惊奇、喜悦和陶醉，你也就领悟到了数学的美、数学的魅力。

数学美不仅仅具有欣赏的价值，而且还具有方法论的价值。马克思在分析人类劳动与动物本能活动的区别时，深刻地指出：人“是按照美的规律来造就东西的”。数学美既是自然美在人类精神上的反映，也是人类智慧的结晶。它给人以思维上的启迪，给人以创造发明的灵感，给人以科学美感的熏陶。

我国著名数学家徐利治先生认为："数学教育与教学的目的之一，应当让学生获得对数学的审美能力，从而既有利于激发他们对数学科学的爱好，也有助于增长他们的创造能力。"本书正是本着这样一个宗旨，为各方面初中文化程度以上的读者，（个别内容涉及到高中或高中以上的数学知识），特别是广大数学教师和大、中学生，提供了数学美育的素材。作者希望读者能够从中享受到学习和研究数学的无穷乐趣，并体验到在数学的创造性活动中，数学的美感所起的重要作用。

广西数学会理事长、广西师范大学数学系顾问麦雨农教授在百忙中仔细地审阅了原稿，并欣然提笔作序。桂林市数学会理事、广西师范大学数学系副教授祝炳宏老师为本书的写作提出了很好的意见和建议，作者一并在此表示衷心的感谢。

曾晓新　1988.5　于桂林

目　录

第一章

数学，科学的大门和钥匙

数学是人类历史上诞生得最早的一门科学，从古希腊文明时代起就受到人们的高度重视。古希腊人深信自然界是和谐、统一、有规律的，而数学则是人们了解自然规律的重要方法和手段。古希腊人的这种观念，集中地体现在毕达哥拉斯(Pythagoras)学派“万物皆数”的信条上。他们认为：“如果没有数和数的性质，世界上任何事物本身或其与别的事物的关系都不能为人所清楚地了解。”

另一位稍后些的古希腊著名哲学家柏拉图(Plato)不但试图通过数学来了解自然界，甚至还想用数学来取代自然界本身。他认为神永远按几何规律办事，只有通过数学才能领悟现实世界的实质和精髓；只有找出永恒不变的数学规律，才有真正的价值。

毕达哥拉斯-柏拉图的思想影响了一代又一代的古希腊学者，形成了古希腊崇尚数学、钻研数学、用数学带动其它科学研究的传统。古希腊人之所以能在众多的领域为人类文明做出了不可磨灭的贡献，在很大程度上正是得益于这一优良传统。

虽然在中世纪的漫长黑夜之中，毕达哥拉斯-柏拉图的思想沉寂了一段时间，但是数学却仍然是人们特别青睐的学科，始终象一盏明灯指引着科学研究的前进方向。随着文艺复兴曙光的到来，毕达哥拉斯-柏拉图的思想又重新活跃起来，哺育

出了一大批卓有成就的杰出学者。哥白尼（Copernicus）、第谷（Tycho）、开普勒（Kepler）、伽利略（Galileo）、帕斯卡（Pascal）、笛卡尔（Descartes）等人，把上帝塑造成了一个全能的数学家，上帝在创造世界时预先安排好的数学方案，理所当然地成为宇宙的真理，这样就使笃信宗教的科学家们的寻找自然规律的研究活动披上了合法的宗教外衣。在当时的科学家的心目中，科学研究就是探求大自然的数学规律，数学规律比起圣经中的信条更加神圣不可侵犯。

由哥白尼等人所开创的近代自然科学的道路，到了牛顿（Newton）、莱布尼兹（Leibniz）、欧拉（Euler）、康德（Kant）、拉普拉斯（Laplace）时期越走越宽，使毕达哥拉斯-柏拉图的传统不断发扬广大。虽然上帝创造世界的至高无上的权威已经动摇，但这并不妨碍人们继续把自然界看做是按数学方式设计的。几乎所有的学者都认为：各种不同的自然科学不过是从各个不同角度探索自然界的数学设计。他们始终遵循并倡导这样的原则：科学的最终目标是确立定量的数学规律。因而，数学成了当然的带头科学，在一切自然科学中享有至高无上的地位。德国大数学家高斯（Gauss）甚至尊数学为“科学的女皇”。

毕达哥拉斯-柏拉图的思想至今仍在影响着人们，现代各门科学不断数学化的发展趋势，实际上一直是自古希腊开始的历代科学家们所努力奋斗的目标。面对数学帮助人们所揭开的一个又一个的大自然的奥秘，我们不能不惊叹毕达哥拉斯“万物皆数”的断言的合理内核是何等的深刻和贴切！

当然，并不是所有的人都是从一开始就认识到了数学的重要性，甚至连本世纪最伟大的科学家、相对论的创始人爱因斯

坦（Einstein）也在如何对待数学的问题上走过弯路。爱因斯坦上大学时，曾在一定程度上忽视了数学。他在《自述》中反省道："数学在我求学时代并不太使我感兴趣，因为我天真地认为，对于一个物理学家来说，掌握好基本的数学概念就足够了。我认为数学里其余的部分对于认识自然是并不重要的奢侈品。这个错误，后来我只好痛心地承认了。"他的大学数学老师、著名数学家闵可夫斯基（Minkowski）在他成名后回忆说："噢，那个爱因斯坦，总是不来上课，我真想不到他能有这样的作为！"如果不是爱因斯坦的同学和好友——数学家格拉斯曼（Grassmann）在关键时刻助了他一臂之力，也许爱因斯坦真的不会有那样大的作为。当时，爱因斯坦在创立广义相对论的过程中，长期未能取得根本性突破的主要原因之一就是他还缺乏必要的数学工具。格拉斯曼及时帮助他掌握了后来成为广义相对论数学基础的黎曼几何和张量分析，使得他终于在经过艰苦的摸索之后打开了广义相对论的大门，完成了物理学的一场革命，宣告了核时代的来临。爱因斯坦从自己的切身体会认识到："在物理学中，通向更深的基本知识的道路是同最精密的数学方法联系着的。"他正确指出："数学给予精密的自然科学以某种程度的可靠性，没有数学，这些科学是达不到这种可靠性的。"

事实上，不仅仅是自然科学，社会科学和技术科学也都在朝着越来越精确或越来越定量的方向发展。数学理论和数学方法应用程度的高低，正在成为衡量科学技术现代化水平的最显著的标志之一。近代自然科学的先驱、著名的英国思想家培根（R. Bacon）的名言："数学是科学的大门和钥匙"，成了越来越多的科学家乃至普通人的座右铭。

生物过去一向是应用数学最少的自然科学，恩格斯上个世纪末还说数学在生物学中的应用等于零；本世纪50年代，数学在生物学中的应用也还主要限于初等数学，现在却已大大改观，数学理论和方法几乎渗透了生物学的每个角落。用微分方程可以描述生物链的相互制约现象；用非线性振荡模型可以说明生物的繁殖节律；进化论需要用数理统计；种群繁衍规律依赖于概率分布；遗传理论离不开抽象代数。怀特（White）提出的曲线环绕数的微分几何公式可以刻画出生命基础物质DNA的空间结构；哈代（Hardy）给出了群体遗传学的基本数学法则；托姆（Thom）的突变理论解释了胚胎生长、形态发生和动物行为中的突变现象；计算机的"蒙特卡洛"方法通过数值计算模拟和代替周期长、难度大、耗资多的生物实验。一门崭新的生物数学正在兴起，并迅速发展出了统计生物学、数学生态学、数学遗传学、数学生物分类学四大分支。人们估计，21世纪很可能是生物数学的黄金时代，生物学将会取代物理学成为使用数学工具最多的学科。

另外，过去被认为与数学毫无缘份的考古学、语言学等社会科学也成了数学大显身手的领域。英国的霍斯金，依靠一台IBM 7090 计算机，用数学模拟方法揭开了英国阿姆斯伯里的著名古迹——Stonehenge石柱群的秘密，发现该石柱群原来是古代居民用以确定二十四节气的一本石头天文历，而且居然还可以精确地预测当地的日食与月食时间。日本的平山朝治，用德国大数学家高斯所创造的"最小二乘法"，根据零碎的史料记载，从众说纷云的零乱线索中推算出早已被人遗忘的日本文化的发祥地——古邪马台国的旧址（在现今日本九州北部的筑紫平野），初步解决了日本考古学界长期争论不休的一个重大问

题。

我国数学家李贤平，利用数理统计原理和电子计算机技术，对我国古典名著《红楼梦》的成书进行了分析研究。他把《红楼梦》一百二十回做为一个整体，以回为单位，从中挑选出几十个常用字；由于字的使用频率与作品文字风格直接相关，用计算机进行统计，并将其使用频率绘成图形，从星云状和阶梯状的图形上可以直观地看出几大群落，而这就是不同作者的创作风格的形象反映。据此，李贤平对以往流行的"前八十回为曹雪芹所作，后四十回为高鹗所续"的看法提出了异议，并提出了《红楼梦》成书新说：轶名作者作《石头记》；曹雪芹"披阅十载，增删五次"，将自己早年所作《风月宝鉴》插入《石头记》，定名《红楼梦》；程伟元、高鹗是全书的整理，抄成者。尽管其结论尚值得商榷，但是这种用现代数学方法和电子计算机技术研究古典文学名著的研究方向，受到了国际红学界的赞赏。据说，国外也曾有人用类似的数学方法鉴别过新近发现的莎士比亚作品的真伪。

为什么数学在科学研究中会起到如此重要的作用呢？原因主要有以下三点：

首先，数学理论是一个严密的逻辑体系，为科学研究提供了进行推理和证明的坚实基础。用数学方法研究复杂的自然现象和社会现象，不论其推演过程如何冗长，也丝毫不会丧失其可靠性，避免了靠常识推理可能会发生的先入为主、似是而非、令人将信将疑的结论。

其次，数学语言在形式上具有高度的抽象性，简洁、精练准确，为科学研究提供了进行思考和表达的高效率工具。由数字、图形、符号按数理逻辑法则所组成的数学语言，在认识世

界的空间形式和数量关系方面，在描述物质运动的规律方面，具有无可比拟的优越性。

再次，数学方法在运算上还具有高度的准确性，为科学研究提供了进行定量分析和数值计算的不可缺少的手段。自十七世纪微积分创立以来，数学在很大程度上本来就是朝着为科学提供分析和计算的方法这个方向发展的。特别是系统科学、计算机科学和现代各个新数学分支的出现和发展，更是大大加快了这一进程。

现代数学犹如一棵硕果累累的参天大树，物质世界供它扎根，生产实践和社会实践促它发芽，科学技术滋养它成长。集合论、数理逻辑和数学哲学是它强大的根系；数论、函数论、概率论、代数学、几何学和拓扑学是它粗壮的躯干和主枝；统计学、运筹学、对策论、组合数学和计算数学是它茂密的分枝；模糊数学、非标准分析和突变理论是它旺盛的新枝；系统论、信息论、控制论、理论物理、经济数学、气象数学、生物数学、数学心理学、数学语言学和数学考古学是它与其它科学杂交的果实。由于数学在现代科学中越来越广泛的应用，在近代数学的发展中还形成了一套与所谓“纯数学”相区别的“应用数学”，专门致力于将数学成果应用于各个科学和技术部门。人类智慧的领域，没有受到数学的渗透和影响的已经寥寥无几了。仰望这棵参天大树，不禁使我们回想起马克思所说的至理名言：“一种科学只有在成功地运用数学时，才算真正达到了完善的地步。”

第二章

料事如神的数学预言

数学正确地反映了现实世界的数量关系和空间形式，因而在科学研究和生产技术中得到了广泛的应用，并且表现出惊人的准确性和预见性。

继16世纪意大利天文学家哥白尼创立划时代的“日心说”之后，17世纪的德国天文学家开普勒根据“日心说”和天文观测资料用数学方法发现了行星运动的三大定律。几乎与此同时，意大利物理学家伽利略提出了物理学中的数学——实验方法、相对性原理和天地统一观点。半个多世纪以后，英国物理学家和数学家牛顿在开普勒和伽利略研究成果的基础上，总结出了万有引力定律，使得开普勒的行星运动三大定律成了万有引力定律的必然的数学推论。

根据万有引力定律进行计算，可以精确地确定太阳系中的行星在任何一个时刻的位置。当时，人们只知道太阳系中有五颗行星，其运行规律与万有引力定律计算的结果几乎完全一致。1781年，人们发现了太阳系中的第六颗行星——天王星，并照例用数学方法为它编制了运行表。可是，实际的观测数据表明，天王星的运行与计算的结果差异很大。是万有引力定律本身不正确呢？还是有什么别的原因影响了天王星的正常运行？有人怀疑这是因为天王星附近还有一颗未知行星，其引力使天王星的运行受到摄动。1844到1845年，有两位青年天

文学家和数学家——英国的亚当斯（Adms）与法国的勒威烈（Leverrier），各自独立地根据万有引力定律和天王星的观测资料，进行了大量的计算，推算出那颗未知行星的运行轨道及其方位。1845年9月，亚当斯曾先于勒威烈将计算结果通知了英国天文台，可惜未能受到应有的重视，英国天文台白白失去了一次做出伟大发现的机会。一年以后，德国天文台按照勒威烈所预言的位置，果然找到了那颗未知行星，即现在的海王星。这是人类历史上第一次通过数学计算准确地预言未知行星，充分显示了数学方法的巨大威力。以后，人们又多次根据天文资料和数学运算成功地预言过另一些未知天体。

电磁现象是人们所熟悉的。早在远古的时候，人类就已经知道摩擦生电和磁石吸铁了，可是经过几千年的不断摸索，才开始找到它们之间的联系。19世纪30年代，英国物理学家法拉第（Faraday）发现电与磁可以互相转化，并且提出：带电体周围存在电场，带磁体周围存在磁场。但是，法拉第的思想并没有为当时大多数人所理解，这除了牛顿“超距离作用”的力学在当时的强大影响以外，法拉第自己未能用数学形式把电磁场理论表达出来也是个重要的原因。法拉第的物理思想是卓越的，实验研究是高水平的，他一直致力于把自己的物理思想和实验研究上升为完美的理论和法则，但由于缺乏系统的数学训练，始终未能如愿以偿。

法拉第的未竟事业，最后由英国物理学家麦克斯韦（Maxwell）所完成。麦克斯韦比较擅长数学，他用数学方法对法拉第的思想进行了分析、整理和推广，把电磁学中一系列基本定律表述成了一组偏微分方程组，即著名的麦克斯韦方程组：

$$
\begin{cases}
Div\vec{E}=Q & (1)\\
Div\vec{H}=0 & (2)\\
Rot\vec{E}=-\dfrac{1}{c}\cdot\dfrac{\partial\vec{H}}{\partial t} & (3)\\
Rot\vec{H}=\dfrac{1}{c}\cdot\dfrac{\partial\vec{E}}{\partial t}+\dfrac{1}{c}\vec{j} & (4)
\end{cases}
$$

其中方程(1)相当于库仑定律，电场强度$\vec{E}$的散度等于闭合体内所含总电量Q；方程(2)表明磁场强度$\vec{H}$的散度为零，是由磁力线的封闭性质所决定的；方程(3)表明电场强度$\vec{E}$的旋度正比于磁场强度$\vec{H}$对时间t的变化率。前面三个方程都是根据实验结果总结出来的，并且(1)与(2)对称；如果没有方程(4)，孤独的方程(3)就显得很不协调，从数学形式的优美与对称考虑，麦克斯韦大胆地假设了方程(4)。

方程(4)意味着不但磁场的变化能够产生电场（正如方程(3)所指出的那样)，反过来，电场的变化也能产生磁场。变化的磁场在其周围产生出变化的电场，变化的电场又在其周围产生出变化的磁场，如此不断循环，电磁场就波动地辐射开了。据此，麦克斯韦预言了电磁波的存在！麦克斯韦还从上述偏微分方程组出发，推导出了电磁波的波动方程，求出了其传播速度。他发现电磁波的速度恰好与当时由实验测得的光速（方程(3)与(4)中的常数c）相同。于是，麦克斯韦又推断说：光也是一种电磁波！

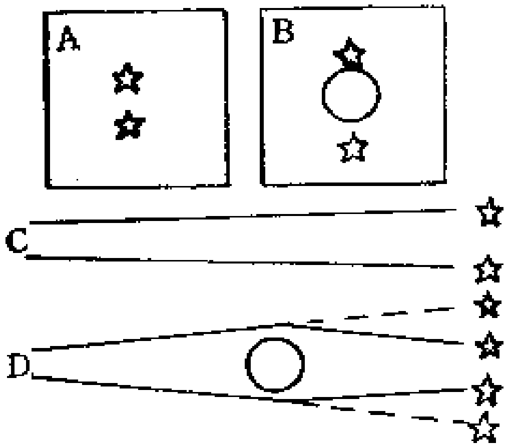
约30年后的1887年，德国物理学家赫兹（Hertz）用实验证实了麦克斯韦所预言的电磁波。麦克斯韦关于光也是一种电磁波的论断后来也被证实。人们无不惊叹麦克斯韦用数学所作

出的物理预言竟如此神奇！其实，最早理解麦克斯韦的还是法拉第本人。早在麦克斯韦刚刚完成用偏微分方程组表述电磁学定律的时候，法拉第就在惊讶之余及时给予了肯定。法拉第欣慰而又赞叹地说："数学的魅力竟与此有这样密切的关系，实在令我吃惊，想不到数学有这样大的用处。"

在法拉第的鼓舞和支持下，麦克斯韦用数学方法逐步建立起了统一、完美的经典电磁学理论，这一理论不但是现代无线电通讯和射电天文学的基础，而且直接导致了本世纪初的物理学革命，导致了爱因斯坦的相对论和普朗克（Planck）的量子论的诞生。

本世纪初，爱因斯坦在数学家格拉斯曼的帮助下，运用黎曼几何和张量分析等数学工具，建立起广义相对论，提出了革命性的崭新的时空观念。爱因斯坦根据广义相对论，从数学上推导出光线在经过强引力场附近时会发生弯曲。这对于自古以来就已习惯于光是沿直线传播的人们来说，简直象是天方夜谭式的预言。

1919年，有一次日全蚀，正是检验这一预言是否正确的大好时机。正常情形下的两颗星间的相对距离（图A和C），在日全蚀时（图B和D）由于光线通过太阳附近，受太阳引力场作用而弯曲，看上去就要分得更开。英国物理学家爱丁顿（Eddington）率领一支天文考察队专程前往西非几内亚湾的普林西比岛，观察星光经过全蚀的太阳时是否真的会发生弯曲。这对相对论是一次严峻的考验，如果观察到星光在太阳附近并不发生弯曲，相对论就要被推翻。结果，从日蚀前后的天文照片上计算所得的数值，以及星光在经过太阳时弯曲的程度，都与爱因斯坦的理论预言十分吻合。这件事立刻轰动了全世界，人们



称颂爱因斯坦的预言是人类思想的一曲凯歌，而爱因斯坦本人却认为这一切都是在意料之中的。据说，曾经有人问爱因斯坦："如果您的预言不灵，您会怎样？""我将会为上帝感到遗憾。"爱因斯坦不无幽默地回答。

量子论可以说是与相对论几乎同时诞生的孪生兄弟。本世纪20年代，正是量子力学初露头角、迅速发展的时代。1928年，英国物理学家狄拉克（Dirac）在研究量子力学的过程中寻出了一个描述电子运动的方程——狄拉克方程。在解这个方程的时候，由于开平方而得到了正负两个完全相反的解，也就是说，它不但描述了人们已知的带负电荷的电子的运动，还描述了另一种除电荷是正的以外，其它结构和性质与电子一模一样的人们尚未知道的反粒子。当时，谁也不了解反粒子，人们所知道的唯一带正电的粒子就是质子。但是，质子的质量过大而不符合狄拉克电子运动方程的"负能解。"比较合理的解释只有两种：要么量子力学的理论有错误，要么自然界确实存在与普通的负电子相反的正电子。

狄拉克经过一年多的犹豫不决和潜心研究，终于从尊重数学的对称美出发，作出了"存在与电子质量相等而电荷相反的

'负能粒子'——正电子"的预言。不过，相信这一预言的人并不多。1932年8月，实验物理学家安德逊（Anderson）在宇宙射线中发现了这种粒子，使人心悦诚服地证实了狄拉克的预言。这样一来，量子论和相对论就在狄拉克方程中完满地统一起来。正如狄拉克自己所说："从此，人们有了一个新的波动方程，它与量子力学的基本原理相一致，而且与狭义相对论的对称性要求相符合。"

自然科学当中，由于数学推导或考虑数学形式而作出的结论，可以先于经验事实而成为令人惊异的预言，这样的例子实在举不胜举，我们只好请读者在此"管中窥豹"吧。

社会科学方面的情况如何呢？谨请看如下一例：

美国新闻界历来有"总统竞选预测"的传统。过去常用"摸拟选举"，即在报刊上登模拟选票，让读者填好寄回，以此推测候选人中谁最有希望当选总统。1916至1932年，当时公认的全美权威性的《文学文摘》杂志社，先后为四届总统选举搞过模拟预测，结果相当灵验。1936年，《文学文摘》根据收回的模拟选票，再次发出预测：共和党候选人兰登将以57%的选票，战胜民主党在任总统罗斯福而成为新总统。与此同时，另一位名不见经传的乔治·盖洛普却告诫人们：罗斯福再度当选的可能性大于兰登。然而盖洛普人微言轻，谁也没把他的预言当回事。直到开箱验票，罗斯福再度当选，这位有先见之明的小人物才名扬天下，由他创办的"美国舆论研究所（AIPO）"顿时声名大振。盖洛普成功的秘诀，主要是基于数学上概率统计理论的"大量观察、随机抽样"的调查方式。从1936到1984年，美国举行过13届总统选举，由盖洛普领导的AIPO对这13次竞选的预测，实际误差平均仅2.6%，除1948和1980年两届

选举以外，AIPO的预测都是相当成功的。这种精确程度在社会科学研究史上是极为罕见的。虽然也有两次失败的不光彩记录，AIPO的专家们却从认真总结经验教训中，设计出了一种能够从调查对象中发现“临阵变卦”者，从而对预测值进行修正的数学方法。由于使用了这种“新式武器”，AIPO在1984年总统选举预测中取得了有史以来的最佳成绩，预测当选总统里根的得票数与实际结果几乎分毫不差。

第 三 章

形式化的数学语言

人们要研究和思考数学问题，要在科学技术中使用数学工具，就必须理解和掌握数学语言。伽利略（Galilei）说得好："哲学是写在这部永远摆在我们眼前的大书中的——我这里指的是宇宙。但是我们如果不首先学习用来写它的语言和掌握其中的符号，我们是不能了解它的。这部著作是用数学的语言写成的，……没有这些数学语言和数学符号的帮助，人们就不可能了解它的片言只语；没有它们，人们就会在黑暗的迷宫中徒劳地徘徊。"

古代的数学语言与普通的日常语言并没有多大差别（事实上，基本上也就是使用日常语言来叙述数学内容）。例如，古埃及人（公元前1700年左右）在叙述上底边长为2，下底边长为4，高为 6 的正四棱台的体积计算方法时说："你要取 4 的平方，得结果为 16；你要把 4 加倍，得结果为 8；你要取 2 的平方，得结果为4；你要把16、8和 4 相加起来，得28；你要取 6 的三分之一，得2；你要取28的两倍，得 56；你看，它等于 56，你可以知道它是对的。"古埃及人在此实际相当于使用了正四棱台的体积公式（现代形式）

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

其中h是高，a、b分别是上、下底边之长。

如果不是对照这个现代的正四棱台体积公式，你我恐怕都很难弄明白古埃及人究竟是按照什么样的一般法则来计算正四棱台体积的，当时文化不发达的人们恐怕就更难弄明白了。所以那时的数学只能掌握在祭师、僧侣、学者等少数“高级知识分子”手中。你可能会感到奇怪，为什么古埃及人的数学语言与普通的日常语言如此接近，但反而晦涩难懂呢？这主要是因为：数学上的概念、关系与法则都是十分严格的东西，不应该含混不清、模棱两可，以至随使用者本身的主观理解或意愿而有所不同；用普通的日常语言来表达就很难达到这样的要求。在日常语言中，一个字或一个词，往往是多义的，常常随使用的场合或上下文的衔接而有所不同。例如，“因材施教”中的“材”，既可以指所要传授的内容，也可以指接受教育的人，从而这句话就有不同的含义、不同的解释和不同的用法。

在各个文化较为发达的民族和国家的历史进程中，都产生过与日常语言文字区别较大的专用的数学语言和符号。例如，我国最古的经典数学著作之一——《九章算术》（成书于公元50—100年间），就曾用如下四句话来说明有理数的减法法则：

《九章算术》用语	现代符号表示
同名相除	$(+a)-(+b)=+(a-b)$
异名相益	$(+a)-(-b)=+(a+b)$
正无入负之	$0-(+b)=-b$
负无入正之	$0-(-b)=+b$

《九章算术》还对线性方程组的表示方法及其解法的术语做了规定，如现代的三元一次线性方程组

$$\begin{cases}3x+2y+z=39 & (1)\\ 2x+3y+z=34 & (2)\\ x+2y+3z=26 & (3)\end{cases}$$

就记为(相当于现代矩阵的雏形)：

左	中	右
𝍠	𝍡	𝍢
𝍡	𝍢	𝍡
𝍢	𝍠	𝍠
=𝍪𝍥	=𝍫𝍣	=𝍫𝍨
(3)	(2)	(1)

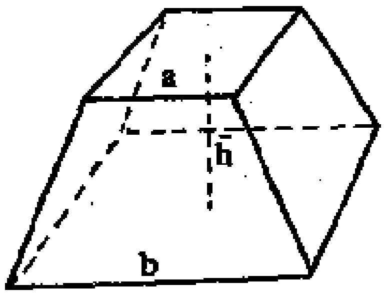
近代，由于数学语言越来越形式化、符号化，全世界已经基本上形成了一套统一的数学符号系统。从最简单最基本的阿拉伯数字到最复杂最深奥的公式，尽管各国的读法不同，然而含义却是相同的。例如，当我们画出正四棱台的图形，并且写下

$a=2$，$b=4$，$h=6$，

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

$$=\frac{1}{3}\times6\times(2^2+2\times4+4^2)$$

$$=56$$



时，几乎每个国家、每个民族稍受过中等数学教育的人都能毫无困难地明白其中的意思，而普通的日常语言恐怕是难以达到这种境地的。

现代的数学语言为什么要采用形式化的符号语言呢？

首先，是因为形式化的符号语言，可以更加简明、抽象，

使语言信息所包含的知识内容大为增加。

我国古代有个“朝三暮四”的故事。说的是有个人养了一群猴子，每天喂它们吃橡子。后来因为家里穷了，他不得不减少喂猴的橡子。他对猴子说，从今天起，每天早上给三颗橡子，晚上给四颗橡子。猴子听了很不高兴。于是他又说，那就早上给四颗，晚上给三颗吧。愚蠢的猴子以为这样一来，橡子就比原来多了，立刻高兴起来。它们不懂“三颗橡子加四颗橡子”与“四颗橡子加三颗橡子”其实是一样多的，结果被养猴人欺骗了。

可是，即使我们告诉猴子，“三颗橡子加四颗橡子”与“四颗橡子加三颗橡子”一样多(当然，要假定猴子能够听懂)，猴子所了解到的知识也十分有限，仍然可能上当受骗。“三个苹果加四个苹果”与“四个苹果加三个苹果”、“三条香蕉加四条香蕉”与“四条香蕉加三条香蕉”是否也一样多呢？它们并不知道。可见，“‘三颗橡子加四颗橡子’与‘四颗橡子加三颗橡子’一样多”这句话所含的知识内容少得可怜。

我们当然知道：三个同名的任意事物在与四个同样名称的事物相加时，其结果是与相加时的顺序无关的。这个事实可以用数学符号和式子表示成“$3+4=4+3$”。“$3+4=4+3$”代表了一种数量变化上的规律，橡子也好，苹果也好，香蕉也好，统统不过是其特例而已。

然而$3+4=4+3$所反映的规律仍然不够普遍，原因是这一式子中的“3”与“4”还是比较具体的数字。$2+5$是否等于$5+2$呢？$7+9$是否等于$9+7$呢？等等，都是$3+4=4+3$所不能回答的。如果我们要从$3+4=4+3$，$2+5=5+2$，$7+9=9+7$，等算式中总结出更为一般的普遍规律，就需要将其中的数字进一

步抽象化为字母，事实上，只需用一个与这些算式类似的字母算式“$a+b=b+a$”，我们就可以表达出“自然数的加法满足交换律”这一结论。虽然“$a+b=b+a$”比“$3+4=4+3$”更抽象，但是同时也包含了更多的知识内容，更带普遍性。随着a、b取值范围的不同，我们还可以用$a+b=b+a$表示有理数、实数、复数等的加法交换律。

不要小看这里有关“加法交换律”的数学知识，它似乎十分简单和明显，但却是近代数学中阿贝尔群的一个极为重要的基本性质之一。在群的“加法”意义下，并非所有的“加法”都是满足交换律的。

其次，形式化的符号语言更有利于揭示和说明客观物质世界在数量变化和空间关系上的一般规律。

例如，随便取一个三位数235与其反序数532相减得

$$532-235=297$$

很难看出所得差有什么特别值得令人注意之处，而一旦用字母表示一般的三位数与其反序数相减及其结果

$$(100a+10b+c)-(100c+10b+a)=99(c-a),$$

我们立刻看出：“任意的三位数与其反序数相减的差都能被99整除”这样一个普遍规律。

又如，观察以下算式：

$$1\times2\times3\times4+1=5^2$$
$$2\times3\times4\times5+1=11^2$$
$$3\times4\times5\times6+1=19^2$$
$$\cdots\cdots$$

我们猜测：“任意四个连续自然数之积与1的和都是一个完全平方数。”我们还可以再取一些实际例子来验证这一猜测。但

是，如果只停留在具体数字上，即使验证成千上万个例子，猜测也还只是猜测，总不能保证对下一个尚未验证过的例子一定正确。如果我们用字母表示这样的一般算式

$$n(n+1)(n+2)(n+3)+1=(n^2+3n+1)^2,$$

我们的猜测就从理论上被证实了。完全可以放心，它对任意的自然数n都正确。

要研究自然的奥秘，进行抽象思维，揭示客观规律，必须使用比日常语言远为优越的形式化的数学语言。伽利略说宇宙这部书是用数学语言写成的，并非言过其实。从质子、中子、电子这些基本粒子的运动，到月亮、地球、太阳这些天体之间的关系；从简单的机械运动，到复杂的化学变化；从逻辑思维的基本规律，到不容辩驳的推理和证明，无不可以用数学语言(符号、图形、式子和方程)清晰、准确地描述出来。

例如，牛顿力学的惯性定律、作用力与反作用力定律、万有引力定律可以用数学式子分别表示为

$$\vec{F}=m\vec{a},$$

$$\vec{F}_1=-\vec{F}_2,$$

$$F=f\cdot\frac{m_1\cdot m_2}{r^2}.$$

以此为基础，既可以精确地描述太阳系中的行星、卫星和彗星的运动，又可以完满地解释地球上潮汐和其它物体的运动。爱因斯坦相对论的质能转换关系，也可以用数学式子表示为

$$E=mc^2.$$

它揭开了原子内部的秘密，使人们发现了蕴藏在原子核内部的巨大能量，帮助人们发掘出了取之不尽、用之不竭的新能源——原子能。

近代的数学家们都十分重视数学的符号化，精心为数学创造了许多流传至今的符号，使数学本身以及其它以数学为主要工具的科学的面貌发生了革命性的巨大变化。例如，代数学创立，引进符号代表数就是关键的一步。是16世纪法国大数学家韦达(Viète)首先系统地使用符号表示数字及算式，并对字母符号进行运算，使"代数"开始正式成为一门学科而独立出来。著名的德国数学家克莱因(F. Klein)指出："代数学上的进步是引进了较好的符号体系，这对它本身和分析的发展比16世纪技术上的进展远为重要。事实上，采取了这一步，才使代数有可能成为一门科学。"

微积分理论的迅速发展和广泛应用，也是和它具有一套优越的符号体系分不开的。在这方面，德国大哲学家、数学家莱布尼兹功勋卓著。他提出的微分符号dx、dy，导数符号$\frac{dy}{dx}$，积分符号$\int$等，使用起来极为方便，进行微积分运算"简直就象是在自动进行似的"(数学史专家的评价)。

莱布尼兹还首先提出了仿照数学符号的方式，为人类的思维也建立一套普遍适用的符号体系。他认为，数学由于有一套独特的符号系统，因而能够有效地表达思想和进行推理。如果人的思维中的概念、判断和推理等也能用适当的符号表示，我们就可以借助这套符号来进行思维的运算，用计算来代替思考。"有了这种东西，我们对形而上学和道德问题就几乎能象几何学和数学分析一样进行推论。""两个哲学家也不需要辩论，因为他们只要拿起石笔，在石板前面坐下来，彼此说一声：我们来算算。也就行了。"不过，莱布尼兹只是设计了一个大体的计划，并未来得及动手进行更多的具体工作。19世纪中

叶，英国数学家布尔(Boole)沿着莱布尼兹的方向创立了一套新的符号和逻辑系统，最先开始部分地实现了莱布尼兹的未竟事业。与此同时，英国数学家德摩根(A. de Morgan)也做了大量的这方面的工作。本世纪初，英国数学家罗素(Russel)和德国数学家希尔伯特(Hilbert)进一步发展了莱布尼兹和布尔的思想，使符号化的逻辑思维规律成了数学的研究对象，同时也使得人们可以从逻辑的角度来研究数学基础的问题，创立了数理逻辑学。

在数理逻辑学中，概念、判断、推理和证明全都可以用符号和式子十分简洁地给予表示。例如，古典逻辑中的矛盾律(在同一推理过程中，对同一对象的相互矛盾的两个判断，至少有一个是假的)、排中律(在同一推理过程中，对同一对象互相否定的两个判断，有一个并且只有一个是真的)、传递律(如果命题p蕴含命题q，命题q又蕴含命题r，那么命题p蕴含命题r)可以分别表示为：

矛盾律　　$p\wedge\sim p\equiv 0$。

排中律　　$p\vee\sim p\equiv 1$。

传递律　　$(p\to q)\wedge(q\to r)\to(p\to r)$。

由于数理逻辑已经把推理和证明过程符号化、形式化了，从一些简单明了的公理出发，使用一些简单甚至机械的推理规则和步骤，便可以一步步地把从日常生活直到科学研究中所使用的推理模式推导出来。因此，用它来检验数学的基础，或探索人类逻辑思维的规律，是十分简便有效的（事实上，数理逻辑正是朝着这样两个方向不断发展的）。最近十几年来，已经有人（例如我国数学家吴文俊）将数理逻辑的研究成果应用于电子计算机，初步实现了数学证明的机械化。

我国古代的数学成就辉煌，明朝以前在许多方面一直遥遥领先于西方各国。我国是最早使用十进制、位值记数法和零的文明古国之一，最早发明二进制，最早使用分数和小数，最早系统地总结了几何学的基础知识，最早发现了著名的勾股定理，最早掌握线性方程组的一般解法和高次方程实根的近似解法，最早研究了不定方程、一次同余式方程组，并得出了著名的"中国剩余定理"，等等。但是，自明朝中期以后，我国的数学水平却一直停滞不前。到清朝后期，我国的数学水平就大大落后于西欧各国了。原因是多方面的，其中令人十分痛心的一个重要原因就是：我国古代数学忽视了数学符号的使用，致使先进的数学思想、精湛的数学方法受累于笨拙的语言表达方式，严重地阻碍了数学的普及、提高和创新。由此可见，形式化的符号语言，是现代数学发展的必要条件。没有比较完善的数学符号系统，就没有现代数学。

第四章

数学的砖瓦——素数

素数，又称质数，是指只具有两个约数的自然数。除了1及其本身以外，再也没有其它自然数能够整除素数。最初的几个素数是

2，3，5，7，11，13，17，19，23，…

自然数除了素数以外，还有两大类：

一类数只有唯一的一个，就是1，它只有自身是它的约数。

另一类数叫合数，合数不止两个约数，即除了1及其本身以外，还有其它自然数能够整除合数。最初的几个合数是

4，6，8，9，10，12，14，15，16，…

素数对于数学的重要性主要在于如下两个定理（读者可以在几乎任何一本初等数论或高级算术书中找到它们的证明）：

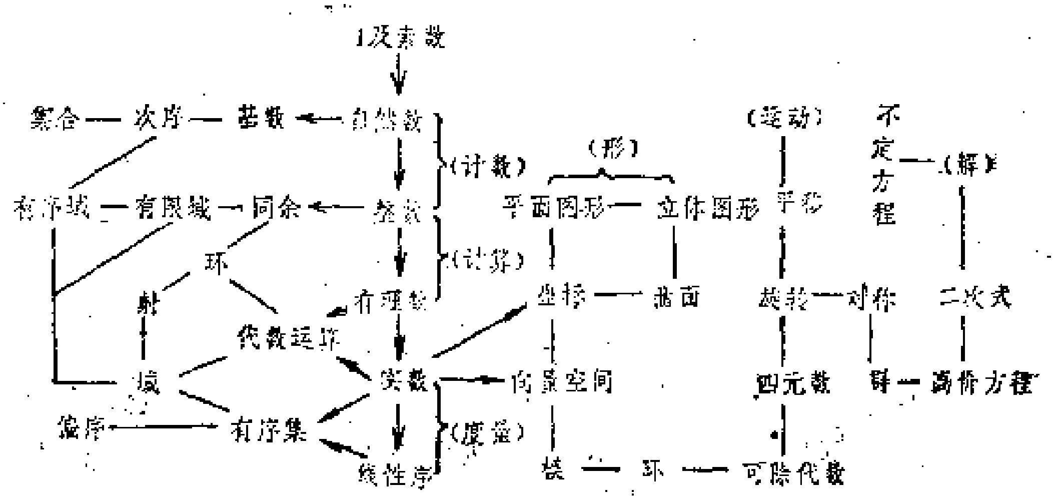
算术基本定理：任何一个大于1的自然数都可以唯一地分解成若干个素数的乘积（例如$10=2\times5$，$63=3\times3\times7$，$2695=5\times7\times7\times11$，等等）。

欧几里得定理：素数的个数是无穷的。

算术基本定理告诉我们：任何一个大于1的自然数都可以用若干个素数的乘积构造出来；欧几里得定理则告诉我们：用来构造自然数的材料——素数，是取之不竭、用之不尽的。

在自然数的基础上，再引进零及相反数，我们就有了整数；由整数的除商（除数不为零），我们又有了有理数；由有理

数的分割或基本有理序列的等价类，我们又可以构造出实数；由实数的有序对或引进虚单位i，还可以进一步构造出复数；……这样，便可以一步步地把整个数学构造出来。以下便是这一过程与抽象代数起源有关的一部分的大致情况：



素数既然如此重要，很自然地就产生了三个问题：1)能否用一个统一的公式将素数表示出来？2)素数在自然数中的分布规律如何？3)如何判断一个自然数是否是素数？

第一个问题的希望看来十分渺茫，但是人们一直在兴致勃勃地寻找那些能普遍产生素数的公式。

17世纪末，法国数学家费马(Fermat)曾宣称公式$F_n=2^{2^n}+1$对零及一切自然数n都给出素数。后人就把形如$F_n=2^{2^n}+1$的数称为费马数。不幸的是，虽然当$n=0$，1，2，3，4时确实有$F_0=3$，$F_1=5$，$F_2=17$，$F_3=257$，$F_4=65537$都是素数，但是18世纪30年代，瑞士数学家欧拉却证明了$F_5=2^{2^5}+1=641\times6700417$，并非素数。不久人们又发现，不仅$F_5$不是素

数，F_6、F_7、F_8等等也都不是系数，迄今为止，人们只知道前五个费马数是素数，却有近五十个费马数都已证明全是合数。其余的费马数究竟是素数还是合数尚不清楚。令人惊异的是，费马数竟与正多边形的尺规作图有关。德国数学家高斯1894年证明了正p边形（p是素数）可以用尺规作图的充要条件是p为费马素数。

1772年，欧拉指出，n的二次多项式$f(n)=n^2+n+41$对于n从-40到39这八十个相继的整数都给出素数值，$f(-40)=1601$，$f(-39)=1523$，……，$f(39)=1601$。当$n=40$时，$f(40)=1681=41^2$是合数；而当n为41以及41的倍数时，$f(n)$显然都是合数。如果限制n只取非负整数，上述n的二次多项式可以等价地改写为$f(n)=n^2-79n+1601$，则当$n=0, 1, 2, \cdots, 79$时都给出素数。(欧拉二次多项式$f(n)=n^2+n+41$还有一个奇妙的性质，即它能给出一千万以内素数中的大约49.5%，也就是说用它表示素数的成功率是很高的)。到目前为止，还不知道有哪个二次多项式能接连给出八十个以上的素数。

可以证明：n的任何多项式都不可能对n的所有非负整数值$n=0, 1, 2, \cdots$，全给出素数。然而可以证明：对于任何自然数m，都存在一个n的m次多项式$f(n)$，当$n=0, 1, 2\cdots, m$时，$f(n)$接连给出$m+1$个素数。现在还不知道是否存在n的次数大于1的多项式能够在n取（并不一定相继）无限多个自然数时都给出素数。

如果不只考虑n的多项式，那么也确实存在某些十分简单的n的函数，对n的一切自然数值都给出素数。例如数学家穆尔士（Mills）就曾证明，存在着一个实数θ，使得$[\theta^{3^n}]$（即

θ^{3^n}的整数部分）对所有的$n=1, 2, 3, \cdots$，都给出素数。可惜我们至今仍不知道这个实数θ究竟是多少。

注：最近有人宣称，素数的统一公式已经找到，即

$$f(m, n)=\frac{n-1}{2}\{|[m(n+1)-(n!+1)]^2-1|-[m(n+1)-(n!+1)]^2+1\}+2.$$

此公式对m、n的一切自然数值都给出素数，而且可以给出所有的素数。

若令$a=m(n+1)-(n!+1)$，则有

$$f(m, n)=\frac{n-1}{2}(|a^2-1|-a^2+1)+2$$

$$=\begin{cases}2 & 若a\neq 0\\ n+1 & 若a=0\end{cases}$$

例如，$m=1$，$n=2$时，$a=0$，$f(1, 2)=3$；

$m=3$，$n=4$时，$a\neq 0$，$f(3, 4)=2$；

$m=5$，$n=4$时，$a=0$，$f(5, 4)=5$。

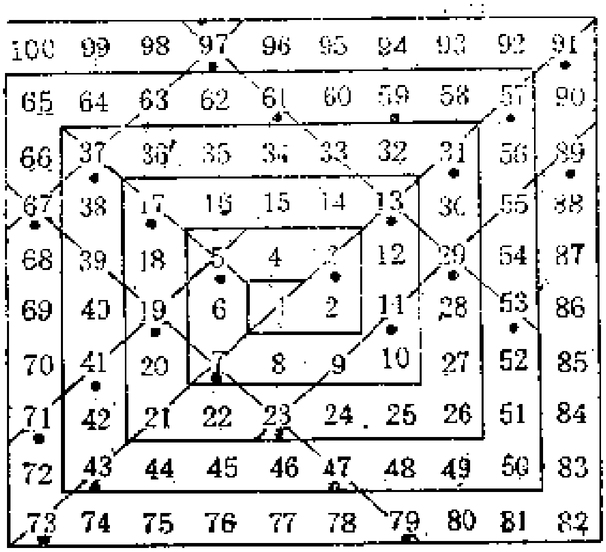
第二个问题：素数在自然数中的分布规律如何？研究起来十分有趣而又相当艰难。事实上，古典数论的大部分内容就是以此为中心展开的。有人形象地指出："素数乃是数学家所研究的对象中最任意和最不易处理的，它们在自然数中就象野草一样乱长，没有人能预报下一个素数会从哪儿冒出来。但是另一件事更加令人惊奇：素数呈现出美好的规律性，存在一些控制着它们整齐变化的规律，它们服从这些规律几乎就象军人服从命令一样准确。"

说素数象野草一样乱长，主要是素数的分布极不规则。例如，第一个一百中有25个素数，第二个一百中有21个素数，第三个一百中有16个素数，…，总的趋势似乎是素数的个数越来

越少，但是高斯却发现：虽然第26379个一百中没有一个素数，第27050个一百中却有17个素数，比第三个一百中还多一个。又如，孪生素数（p与$p+2$都是素数）究竟有多少对呢？是有限多对，还是无穷多对？我们至今还不知道。唯一比较确切的断定是：孪生素数要么有有限对，要么它们的平均分布极其稀疏，因为由全体素数的倒数所组成的级数是发散的，而由全体孪生素数的倒数所组成的级数却是收敛的。

说素数象军人一样服从命令，主要是素数的分布尽管表面上看起来杂乱无章，也仍然隐含着支配它们的内在规律。例如，虽然我们无法指出第n个素数p_n的准确位置，但是我们知道它必定落在一个范围内：$\alpha n\ln n<p_n<\beta n\ln n$（其中$\alpha$、$\beta$是与$n$无关的常数）；我们还知道，在$n$与$2(n-1)$之间至少存在一个素数。又如，用$\pi(x)$记不超过$x$的素数个数（$\pi(10)=4$，$\pi(20)=8$，$\pi(100)=25$，…），高斯和勒让德（Legendre，法国）曾通过计算，猜想当x越来越大时，$\pi(x)$就越来越接近$x/\ln x$。后来，法国数学家阿达玛（J. Hadmard）证明了高斯的猜想是正确的，即$\pi(x)$与$x/\ln x$的比值当x趋向无穷时的极限是1．这就是著名的素数定理。

美国著名数学家乌兰教授（S. Ulam），有一次偶然把从1到100的自然数按反时针方向排成了一种螺旋形式（如图），他突然惊奇地发现其中的素数（数字下打“·”者）具有直线分布的现象。超过100的自然数又如何呢？乌兰和他的同事们用当时具有最先进绘图功能的计算机Maninc Ⅱ，把从1到1千万的自然数按反时针方向排成螺旋形式，打印了出来，结果发现其中的素数仍有直线分布的特点。这种现象后来被称为“乌兰现象”，数学家们从“乌兰现象”中找出了素数的不少有趣性，



质。这一事例也说明素数的分布是有一定规律的。

素数分布的更多规律还有待我们进一步去发现。

第三个问题：如何判断一个数是素数？实际上是早已解决了的。不过，方法亟待改进。

判别一个自然数p是否是素数，最原始的方法就是根据素数的定义，即检查p是否只能被1及p本身整除。这个方法看起来十分简单，但是由于需要用所有不超过$\sqrt{p}$的自然数（尽管也可以改进为所有不超过$\sqrt{p}$的素数）去一一试除，实际运算量往往很大，以至在判别较大的素数时简直无法使用。

也可以用制作素数表的筛法来判别素数。筛法就象是使用一面巨大的筛子，将自然数中的合数逐步筛选掉而最后只剩下素数。最古老的筛法是古希腊数学家厄拉多塞(Eratostheues)公元前200年左右发明的。例如，要求50以内的素数，先将2到50的所有自然数按顺序排好；然后在2上打圈，划去所有2的倍数；第一个未打圈而又未划去的是3，将3打圈，划去所有 3 的倍

数；下一个未打圈而又未划去的数是5，将5打圈，划去所有5的倍数；…。如此继续下去，直到所有不超过$\sqrt{50}$（≈7）的

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~~					

各数都打了圈或划去为止（有些数在这一过程中被划了多次）。最后，2到50各数中打了圈的以及未划去的数：2，3，5，7，11，13，17，19，23，29，31，37，41，43，47就是50以内的素数。

过去和现在，许多数学家都曾经或正在致力于“筛法”的改进。今天，经过不断改进后的筛法已经成了近代解析数论中的重要工具（例如，我国数学家陈景润运用改进的筛法，在证明著名的“哥德巴赫猜想”上取得了重大进展）。但是，即使用改进的筛法来判别素数仍然不够简便。

有关目前判别素数的最佳方法，请参看本书“成果诚可贵，方法价更高”一章。

还有一种从理论上判别素数的方法，叫做威尔逊(Wilson)判别法：p为素数的充要条件是p能整除$(p-1)!+1$（这一充

要条件也经常表示为 $(p-1)!\equiv -1 \pmod{p}$，其中 $(p-1)!=1\cdot 2\cdot 3\cdots\cdots(p-1)$。）这一法则几乎没有任何实用价值，然而它却解决了不少理论问题。

我国学者张春暄认为，素数反映了自然界在量的方面的某种稳定性。他提出了以两条原理为基础的"多体稳定性理论"：

素数原理 由素数个基本单位构成的多体结构是有极性的稳定系统（如电荷的正、负）。

对称原理 由双奇素数个基本单位构成的多体结构是中心对称的稳定系统（如原子结构）。

这个理论可以很好地解释从微观世界的基本粒子到宏观世界的星系所组成的结构的稳定性，也可以推测大量的自然现象和科学规律。例如，它可以解释：为什么人有两双手共十个指头；为什么大多数昆虫每侧各有三足；为什么植物叶子多为三片或五片，而植物花瓣多为五瓣或六瓣；为什么惰性气体那么稳定；……。又如，1956年以前，生物界普遍认为人的染色体为24对，这与"多体稳定性理论"矛盾。后来实际测得人的染色体为23对，是与"多体稳定性理论"相符的。再如，按照"多体稳定性理论"，太阳系作为一个稳定的结构系统，至少应有10颗行星，而不是目前已知的九颗行星，这与用其它方法推测的结果完全相同。张春暄的"多体稳定性理论"在国际上引起了很大反响，得到了各国科学家的高度评价。

第五章

古老而又年轻的二进制

就数的进位制而言，我们通常使用得最广泛的是十进制，也就是“逢十进一”。有很多证据表明，十进制跟我们的一双手有十个指头有关。在不发达的原始社会，双手便是人类随身携带的最方便的计算工具。至今我们仍然能够在幼儿以及尚未开化的落后民族身上，看到人类是怎样使用手指做为计数器的。英文中的“digit”一词，既可以当“数字”讲，也可以当“手指”讲，显然是因为古代人们长期使用手指来表示数字和进行运算。如果我们赤足的祖先再把双脚的十个指头也用上，那么便产生出二十进制，这种进位制曾被美洲古代的玛雅人所普遍使用。欧洲一些国家的文字，也留下了使用二十进制的痕迹，如法文的“quatre-vingts（八十）”，原意即“四个二十”；又如英文的“three-score and seven years ago（六十七年以前）”，原意是“三个二十又七年以前”。

我们也常采用六十进制。例如，对于时间有一小时等于六十分钟，一分钟等于六十秒钟；对于角度有一度等于六十分，一分等于六十秒。这种进位制也产生得很早，四、五千年以前古巴比仑人就已经使用这种进位制了。古巴比仑人采用六十进制，可能是因为一年有12个月，一只手有5个指头，而12×5＝60；另一个可能的原因是60的约数比较多，2、3、4、5、6、10、12、15、20、30等都是，在很多情况下使用起来较十进制

方便。

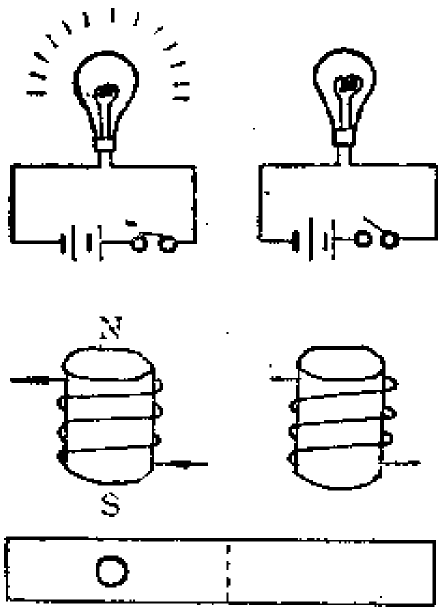
还有些不大常用的进位制，例如英国有些计量单位就采用十二进制（1英呎＝12英吋；1先令＝12便士；1罗＝12打）；我国、英国和俄国还曾采用过十六进制（1斤＝16两；1英磅＝16英两；1俄尺＝16俄寸）。许多人认为，我们沿用祖宗习惯的十进制并不科学，十二进制就比十进制优越。10只含有1、2、5 10四个约数，而12却含有1、2、3、4、6、12六个约数。用12做除法，整除的机会就比用10做除法要多得多，这在应用上是很重要的。瑞典国王查理十二世曾大力推行过十二进制；美国则至今仍有一个"美国十二进制学会"在致力于十二进制的普及工作。

然而，电子计算机中却需要采用二进制。二进制"逢二进一"，只用0和1两个数码就可以表示一切数目，下表即可说明这一点：

十进制数	二进制数	二进制基的表示形式
0	0	0×2^0
1	1	1×2^0
2	10	$1\times 2^1+0\times 2^0$
3	11	$1\times 2^1+1\times 2^0$
4	100	$1\times 2^2+0\times 2^1+0\times 2^0$
5	101	$1\times 2^2+0\times 2^1+1\times 2^0$
6	110	$1\times 2^2+1\times 2^1+1\times 2^0$
7	111	$1\times 2^2+1\times 2^1+1\times 2^0$
8	1000	$1\times 2^3+0\times 2^2+0\times 2^1$ $+0\times 2^0$
9	1001	$1\times 2^3+0\times 2^2+0\times 2^1$

		$+1\times2^0$
10	1010	$1\times2^3+0\times2^2+1\times2^1$ $+1\times2^0$

电子计算机为什么要采用二进制呢？最主要的一个原因就是二进制只需0和1两个数码就可以表示一切数目。这对于机器来说最为有利：只要找到一种具有两种稳定状态的元件就可以用二进制表示数了。而具有两种稳定状态的元件是很多的，例如：电灯的"开"与"关"，磁蕊的"充磁"与"消磁"，纸带的"穿孔"与"无孔"，晶体管的"通导"与"截止"等等。如果在计算机上采用其他进位制，则需要具有多种稳定状态的元件，譬如十进制就需要具有十种稳定状态的元件，这在技术上是困难的，而且线路复杂，造价昂贵，维修不便，可靠性低。



电子计算机采用二进制的另一个原因是二进制的运算比较简单，可以大大提高运算速度，这对于电子计算机来说是至关重要的。以下是二进制的加法表与乘法表：

+	0	1
0	0	1
1	1	10

·	0	1
0	0	0
1	0	1

从这两个运算表可以看出：对每一位而言，二进制的加、乘运算都只各有四种情形，但是十进制的加、乘运算却都各有一百

种情形，由此即可见运算的繁简。

电子计算机采用二进制的第三个重要原因是：二进制不但可以进行数值运算，而且只要将其运算法则稍做修改就可以进行二值逻辑运算，从而使计算机能够进行判断和推理。

我们已经知道，二进制一位数的加法是

$$
\begin{array}{r} 0 \\ +\ 0 \\ \hline 0 \end{array}
\qquad
\begin{array}{r} 0 \\ +\ 1 \\ \hline 1 \end{array}
\qquad
\begin{array}{r} 1 \\ +\ 0 \\ \hline 1 \end{array}
\qquad
\begin{array}{r} 1 \\ +\ 1 \\ \hline 10 \end{array}
$$

如果不计进位，只看本位，就叫"本位加"或"模 2 加"为了与普通的加法相区别，可将这种加法记为⊕，则有

$$
\begin{array}{r} 0 \\ \oplus\ 0 \\ \hline 0 \end{array}
\qquad
\begin{array}{r} 0 \\ \oplus\ 1 \\ \hline 1 \end{array}
\qquad
\begin{array}{r} 1 \\ \oplus\ 0 \\ \hline 1 \end{array}
\qquad
\begin{array}{r} 1 \\ \oplus\ 1 \\ \hline 1 \end{array}
$$

如果我们把"0"看做是假命题（如"$2 \nless 3$"）的逻辑值，把"1"看做是真命题（如"$2<3$"）的逻辑值，则运算⊕就可以看做是命题的逻辑加。用"或"将两个命题 p 与 q 连接起来（析取），得到一个新命题"p或q"，其逻辑值（"真"或"假"）就可以由逻辑加来决定。例如，$0\oplus1=1$，意味着假命题与真命题的析取（如"$2\nless3$ 或 $2<3$"）仍是真命题。

类似的还有逻辑乘，可用⊙表示，其法则几乎与二进制的乘法法则完全相同。用"且"将两个命题 p 与 q 连接起来（合取），得到一个新命题"p 且q"，其逻辑值可由逻辑乘来决定。例如，$0\odot1=0$，意味着假命题与真命题的合取（如"$2\nless3$且$2<3$"）是个假命题。

现将逻辑加与逻辑乘的法则列表如下，请与二进制加，乘运算表做个比较：

⊕	0	1
0	0	1
1	1	1

⊙	0	1
0	0	0
1	0	1

以逻辑加、逻辑乘，还是一种叫逻辑非的运算做为基础，电子计算机就能够进行逻辑判断，机器也就具有了人类思维的某些功能。

非常凑巧，生物神经系统的兴奋与抑制也是以与二进制类似的“全或无”方式表现出来的。所谓“全或无”方式，说得通俗些，是指神经元在接受外界刺激时，只要刺激量达到了神经元做出反应的必需强度，它就会做出反应，否则它就完全没有任何反应，而且做出反应的输出信号的大小总是一样的。因而神经元的“全”与“无”两种状态，就可以用二进制的“1”与“0”表示，神经系统则可与电子线路中的二进制网络类比。以离散的二进制元件为基础的电子计算机，就成了神经系统某些功能机制的一种理想模型。

可以证明，在计算技术和编码理论中，用r进制表示十进制数，取$r=e=2.7183\cdots$（即自然对数的底），效率最高。而与e最接近的整数就是2与3。考虑到机器的结构，在计算机上自然就应采用二进制了。

二进制是17世纪德国大哲学家和数学家莱布尼兹最早明确提出的，距今已近三百年了，不过，与具有几千年历史的其它进位制比较，二进制还“年轻”得很呢！虽然如此，二进制的思想萌芽却可以一直追溯到公元前一千年左右成书的我国圣典

著作《易经》。《易经》中说："易有太极，是生两仪，两仪生四象，四象生八卦。"可以理解为：未分是一，一分为二，二分为四，四分为八。用现代数学式子表示就是：$2^0=1$，$2^1=2$，$2^2=4$，$2^3=8$。"八卦"则是整个《易经》符号系统的基础，传说由我国古代伏羲氏首创（如此，八卦便有五万多年的历史）。八卦由两种基本的卦画——阳爻"——"（肯定）与阴爻"— —"（否定）的不同排列所组成，恰好与二进制数码相对应：

名称	符号	对应的二进制数	十进制数
坤	☷	000	0
震	☳	001	1
坎	☵	010	2
兑	☱	011	3
艮	☶	100	4
离	☲	101	5
巽	☴	110	6
乾	☰	111	7

因此，《易经》中的符号系统，实际上就是一个二进制的符号系统。这一点还是二进制的发明人莱布尼兹首先看出来的。莱布尼兹对《易经》的评价极高，当他发现几千年前的《易经》中的符号系统竟与二进制不谋而合的时候，心情非常激动，甚至表示愿意加入中国籍。他说："易图是流传于宇宙间科学中之最古的纪念物。""伏羲是古代的君王，世界知名的哲人，中华国家和东洋科学的创造者。""《易经》也就是变易之书，在伏羲许多世纪以后，文王和他的儿子周公以及在他们五个世纪之后著名的孔子，都曾在这六十四个图形（指易经中的六十四卦）中寻找过哲学的秘密，……这恰是二进制算术，……在这个

算术中，只有两个符号：0和1。用这两个符号可以写出一切数字。"

第六章

哥仑布的鸡蛋——0

数“零”的最初意义是表示“无”，幼儿在计算5－5＝0时，用的就是这一含义。但在位值记数制中，“零”已经不是简单的“无”了。十进制中的“10”，其“0”一方面表示个位上的“无”，另一方面又指示出左边的“1”在十位，代表的数值正好是原来的十倍。引入正、负数以后，“零”成了既不正又不负的唯一的中性数，既是正与负的分界点，也是数轴上我们观察事物数量关系的出发点——即原点。解方程时，我们经常需要把方程整理成所有各项都在等号一边而另一边为零的形式。研究多项式的性质时，多项式的零点是最值得我们注意的对象之一。群论中的零元，是构成加群不可缺少的具有特殊意义的中心元素。而有无零因子，则是判断环中消去律是否成立的标准，……。所以，恩格斯说：“零比其他一切数都有更丰富的内容。”

在我们今天看来，没有“零”的数学简直是不可想象的，可是历史上“零”的出现，特别是“零”的符号“0”的出现，却经历了非常曲折的过程。正因为如此，数学史家曾形象地把“0”比做“哥仑布鸡蛋”。

哥仑布（Columbs）是15世纪末西班牙著名的航海家，历尽千辛万苦发现了美洲新大陆。他返回西班牙以后，受到了群众的欢迎和王室的优待，同时也招致了某些王公贵族的忌妒。

在一次宴会上，有人鄙夷地说："到那个地方(指美洲新大陆)去有什么了不起？只要有船，谁都能去"。哥仑布并未正面反驳，而是拿起一个熟鸡蛋问道："谁能把这只鸡蛋用小的那头竖起来？"许多人试了又试，都说不可能。只见哥仑布将鸡蛋在桌上轻轻敲破了一点壳，就竖了起来。于是又有人不服气地说："这谁不会！"哥仑布回答："在别人没有做之前，谁都不知道怎么做，一旦别人做了之后，却又认为谁都可以做。"这就是流传了近五百年的"哥仑布鸡蛋"的故事。

新事物刚刚诞生的时候，总会遇到各式各样的困难和挫折，一旦有人开了头，仿效起来也就容易了。"0"的出现就是这样。

"零"的最初意义既然是表示"无"人们开始时也就未曾感到需要用某种特殊的符号来表示它。遇到此类数量时，常常不写或空位，非说明不可的时候，也只是使用具有"无"、"空"等含义的文字（如印度的 sūnya，阿拉伯的 as-sifr，希腊的τξιφρα，意大利的 zero，以及英文的cifer）。但是随着位值记数法的完善，人们越来越感到了创造表示"零"的符号的迫切需要。

人们所使用的记数法，主要有两种，一种是以现今通用的阿拉伯数码记数法为代表的位值制；另一种是以罗马数码记数法为代表的非位值制。所谓位值制就是一个数码所代表的数量级，视其所在的数目中的位置而决定。例如"2"放在个位(如132中)就只代表"二"，放在十位（如123中）就代表"二十"，放在百位（如213中）则代表"二百"。非位值制就不一样，数码所代表的数量级法则要比位值制复杂。例如罗马记数法中虽然只有七个基本数码I（1）、V（5）、X（10）、L（50）

C（100）、D（500）、M(1000)，但用它表示即使是一千九百八十五这样简单的数，也要写成MDCCCCLXXXV，显得冗长，相比之下，阿拉伯数码记数法中的1985就要简明多了．位值记数法堪称数学史上的一项伟大发明，多亏了这种优越的记数法，数学才有了今天的面貌。

“零”的符号“0”可以说是位值记数法的必然产物，否则人们便难以区分和表示诸如307，3007，370，3700这样的数目。在古印度，307曾被表示成“3　7”，中间所空的格表示其十位上的数字没有，但这很容易与3007等数字混淆不清．人们逐渐发现了这一缺点，开始在空格处加上小圆点“·”，把307表示成“3·7”。后来小圆点才慢慢演变成“0”，并随着其它印度数码传入阿拉伯和欧洲，最后形成了现今世界通用的印度–阿拉伯数码。古印度对数学的贡献不仅仅在于最早创造了“零”的符号“0”，而且最早承认“零”是一个数。在公元三世纪，在古印度数学家的眼里，“零”就已经不是简单的“无”或空位了，而是象其它实在的数一样可以参加各种运算。

我国是最早使用位值记数法的文明古国之一。在漫长的历史进程中，我们中华民族独创了不同于印度–阿拉伯数码“0”的“〇”。我国古代也曾采用空格来表示位值记数法中的“零”，例如唐代的《立成算经》上，208就记为“‖　Ⅲ”，中间空一格，后来为了明确起见，借用“□”表示脱落文字的习惯，用“□”来表示空位，这样208就可以表示成“‖□Ⅲ”。古人书写时，字体常成行书，方块“□”也就顺笔画成了圆圈“〇”。至迟在金代的《大明历》中已经看到“〇”的使用，如将403写成“四百〇三”。到秦九韶的《数书九章》（1247年）就普遍使用“〇”了。中国的“〇”虽然比印度的“0”晚出现了两

个世纪，但在当时的世界上仍处于领先的地位，欧洲直到13世纪初，才由斐波那契（Fibonacci）正式引进了包括0在内的印度-阿拉伯数码，又过了两三百年，才广泛流传开来。

有趣的是，汉字“零”公元前就已出现，比数码“〇”早一千多年，但其原义并不含有“空”和“无”的意思，而是指雨后的小水滴。后来引申做“零头”解，如“零丁”、“零碎”都有这种意思。我国古代早就把105读作“一百零五”，意思是说除了一百以外还有个“零头”五。以后，因为105又写作“一百〇五”，“〇”也就随之读作了“零”。巧得很，“〇”外形也颇象个小水珠，恰好与“零”的原义不谋而合。

尽管印度-阿拉伯的“0”从东方传入古罗马的时间很早，但是罗马数字中至今没有相应的表示“零”的符号。保守的罗马教皇认为数是神圣的上帝创造的，任何人不得随意添加和更改，因而禁止使用从东方传入的“0”。一位介绍和提倡使用“0”的古罗马学者，竟因违反了罗马教皇的禁令而残遭拶刑，永远失去了握笔写字的能力。罗马数字曾在13世纪以前盛行欧洲各国，然而由于使用起来极不方便，上帝所创造的笨拙的罗马数字最终不得不让位给人类所创造的灵巧的印度-阿拉伯数字。

第七章

令人着迷的π

我们知道，圆的周长和直径的比是个与圆的大小无关的常数，其值为3.1415926……，通常用希腊字母π来表示，称为圆周率。

圆周率π是数学和其它自然科学中经常使用的一个重要常数。计算圆的周长和面积，计算球、球缺、球台、圆柱、圆锥、圆台等各种旋转体的表面积和体积等等，只要一涉及到与圆有关的各种几何体的计算，就几乎不可避免地要求助于π。如果用弧度做单位，那么从最简单的三角形直到比较复杂的多边形的内角和，都可以用π来表示。概率积分以及许多无穷级数之和，结果中都含有π。此外，许许多多的函数也都与π关系密切，其中最有代表性的就是大家所熟知的三角函数，要研究这些函数的性质，少不了要与π打交道。

正因为π在数学中扮演着这样一个“不可一日无此君”的重要角色，而π的求值又有相当的难度，需要有效的方法和娴熟的技巧，所以世界数学史上许多国家都十分重视圆周率π的计算。一位德国数学家评论道：“历史上一个国家所算得的圆周率的准确程度，可以做为衡量这个国家当时数学发展水平的一个标志。”

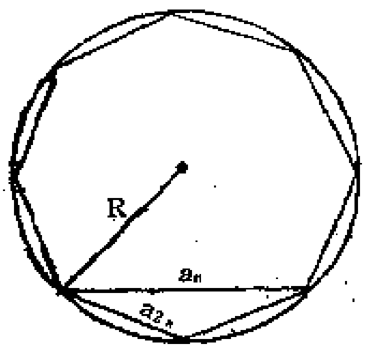
最早定出π的比较准确的值的是古巴比仑尼亚人，公元前2 000年左右，他们就得出了$\pi=3+1/8$（即3.125）。

其后是希腊人，公元前五世纪的古希腊数学书中，π取3.1416，这在当时是领先的。

首先从理论上给出π值的正确求法的，是公元前200年间的古希腊数学家阿基米德（Archimedes）。阿基米德用圆外切与内接正多边形的周长从大、小两个方向上同时逐步逼近圆的周长，并用穷竭法巧妙地避开了当时还难以掌握的极限，求得$3\frac{10}{71}<\pi<3\frac{1}{7}$。

公元前150年左右，另一个古希腊数学家托勒密（Ptolemaeus）用弦表法，以1°的圆心角所对的弦长的360倍再除以圆的直径，给出了六十进制的π值为3.8′30″（即$\frac{377}{120}=3.1416$）。

我国则迟至东汉初年（公元前一世纪）仍在普遍使用"周三径一"的古率，即取$\pi=3$，这是相当粗略的。虽然西汉末年刘歆就已经开创了不沿用古率的先例，取π为3.1547，但并未受到足够的重视。后来，东汉时代的著名科学家张衡又求得两个π值，一个是$\pi\approx\frac{92}{29}$（即3.17241…），另一个是$\pi\approx\sqrt{10}$（即3.1622…）。直到三国时代（公元200多年间），数学家刘徽首创划时代的"割圆术"，我国才在π值的计算上取得了长足的进步。



"割圆术"的基本思想是用圆内接正多边形的面积来逐步逼近圆的面积，体现了一种十分先进的极限观点，是我国古代数学贡献给人类文明的一件珍宝。

利用勾股定理和面积公式，刘徽证明了如下三个关系式：

$$a_{2n}=\sqrt{2R^2-R\sqrt{4R^2-a_n^2}} \quad (1)$$

$$S_{2n}=\frac{1}{2}nRa_n \quad (2)$$

$$S_{2n}<A<2S_{2n}-S_n \quad (3)$$

其中a_n、S_n分别表示圆内接正n边形的边长和面积，A为圆面积。

这些关系式今天证明起来都很容易（有兴趣的读者不妨一试），但在当时却是很了不起的重大发现。特别是(3)式，只要令圆半径$R=1$，立即得到

$$S_{2n}<\pi<2S_{2n}-S_n \quad (3')$$

就成了“割圆术”中的精华——割圆不等式。

刘徽从单位圆内接正六边形开始，逐次将边数加倍，一直计算到圆内接正192边形，得

$$3.1410<\pi<3.1427$$

取精确到小数点后两位数字，$\pi\approx3.14$，这就是历史上有名的“徽率”。

继刘徽之后，我国南北朝时期的数学家祖冲之，又在世界上第一次准确地把π值推算到小数点后七位，得到

$$3.1415926<\pi<3.1415927.$$

这一惊人的记录，直到一千多年以后才被打破。

祖冲之所用的方法，记载在他的著作《缀术》中，可惜已经失传。如果他仍然是沿用刘徽的“割圆术”的话，就一直要计算到圆内接24576边形，总共需要进行23次16位有效数字的开方。而一个16位有效数字的开方用今天的笔算也要足足花一整天，在当时用筹算（用竹棍摆成数字和算式）就更为艰难了。由此

可见祖冲之的成就是来自不易的。

祖冲之还使用了两个分数来表示π的近似值，一个是$\frac{22}{7}$（即$3.\dot{1}4285\dot{7}$），另一个是$\frac{355}{113}$（即3.1415926…），分别称为“疏率”和“密率”。密率$\frac{355}{113}$是个有趣的分数，分母、分子排在一起是113355，恰好是三对最小奇自然数1、3、5的顺次重复排列，它与π的准确值相差还不到10^{-6}。欧洲直到一千多年以后，才由荷兰数学家安托尼兹（Anthonisz）重新发现密率。为了纪念祖冲之的贡献，后人又称密率为“祖率”。

15世纪，伊斯兰的数学家阿尔•卡西（Al-kāshj），通过分别计算圆内接和外切正3×2^{28}边形的周长，把π值的计算推到小数点后16位，打破了保持了上千年的祖冲之记录。

16世纪的法国数学家韦达（Viéta），沿用阿基米德的思想方法，推导出了历史上第一个用无穷乘积来准确表达π的公式。1593年，韦达用这一公式把π值计算到小数点后17位。

1615年，德国数学家鲁道夫（Ludolph）求出了π的35位准确数字。德国人因此至今仍自豪地把π称为“鲁道夫数”。

1717年，英国数学家阿布拉罕，求出了π的72位准确数字。

接着，马青（Marchin）、勒让德（Legendre）、高斯（Gauss）等数学家相继发现了一系列有关π与反正切函数间的关系式，利用反正切函数的无穷级数展开，使得π值的计算迅速突破了100位大关。

1873年，英国数学家威廉·向克斯（W.Shanks）花了毕生精力，终于把π算到小数点后707位，曾被传为佳话。遗憾的是，七十二年后的1945年，他的同胞福格逊（Ferguson）发现，其结果从小数点后528位开始出现错误。

1944年，福格逊对向克斯的π值发生怀疑。他统计了π的608位小数中0–9各数码出现的次数，结果如下：

数码	0	1	2	3	4	5	6	7	8	9
出现次数	60	62	67	68	64	56	62	44	58	67

大多数数码出现的频率都在$\frac{1}{10}$左右，偏差不很大，而数码7却出现特别少。福格逊认为，π的小数展开式中，各个数码出现的机会应该大致均等，不会对一两个数码持有“偏见”。因此他怀疑向克斯算错了，决心自己重算一次。经过一年的认真核算，终于发现向克斯的π值果然有错。福格逊本人于1947年把π值计算到了小数点后710位。

1949年，有两个美国人，斯密司和伦奇，把π值计算到了小数点后1120位，成为用笔算求π值的世界冠军。

本世纪50年代以后，圆周率π的计算开始借助电子计算机。首次使用的计算机叫ENIAC，花了70小时，计算到第2037位小数；五年以后，用NORC计算到第3089位小数；1957年，又用IBM7000计算到第100265位小数。1973年，法国数学家用IBM7600花了23个多小时，突破了π的百万位有效数字大关。1983年，日本东京大学的电子计算机用了30个小时，把π值计算到了千万位以上的有效数字，1984年又用速度更快的电子计算机证实了其正确。目前已经有人宣称把π计算到了一亿位甚至

十亿位以上的有效数字。

π的前一百位有效数字为

$$\begin{aligned}\pi=3.&14159\quad 26535\quad 89793\quad 23846\quad 26433\\ &83279\quad 50288\quad 41971\quad 69399\quad 37510\\ &58209\quad 74944\quad 59230\quad 78164\quad 06286\\ &20899\quad 86280\quad 34825\quad 34211\quad 70679\end{aligned}$$

对于实际应用来说，现代科学技术只要用到π的十位小数也就足够。用祖冲之的密率计算地球赤道长度，误差仅3米左右，高精度的人造卫星的测量也不过如此。人们之所以还在借助电子计算机对π的准确值穷追不舍，主要是为借助这一典型的带挑战性的课题，不断提高计算机的硬件与软件的功能，不断创造新的更强有力的计算方法。

古代数学家总是千方百计地企图把π值表示成分数的形式，大概是认为π是有理数。然而，始终没有一个分数能把π准确表示出来。1766年，德国数学家兰伯特（Lambert）证明了π是无理数。一百多年以后，又一位德国数学家林德曼(Lindemann)进一步证明了π是超越数，就是说π不仅不能表示成分数的形式，也不是任何代数方程的根。π被证明是超越数，宣告了古希腊三大尺规作图难题中的“化圆为方”是不可能的。

由于π是无限不循环小数，所以，尽管电子计算机把π值计算到了上亿位的有效数字，也没有发现其数字变化上有什么规律。不过，仍有些值得注意的统计性质，例如，π值中从0到9十个数码所出现的频率都大致相同，平均值大约为4.5左右。

π值从小数点后第710100位起连续出现3333333，第3204765位起又连续出现3333333，小数点后一千万位中，有八十七回同一数字连续重复六次，例如从第762位开始，就连续出现

999999；从小数点后第995998位起，出现23456789连续上升的情形；而从第2747956位起，又出现876543210连续下降的情形；小数点后一千万位中，π的起首六个有效数字314159出现过六次，更令人惊异的是从小数点后第52638位起，出现了14142135，恰好是$\sqrt{2}$的前八位有效数字。

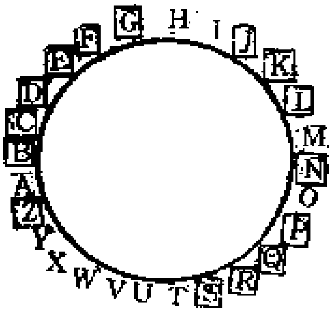
π值从开头第一个有效数字起，恰好能组成素数的有效数字组迄今为止已发现四个；头三个是3、31和314159；如果倒过来看，则素数组更多，头六个是3、13、51413、951413、2951413、和53562951413。

取π的前六位有效数字314159，是个素数，把它反过来还是素数（即951413是素数）；如果取314159各位数的补数，得796951，又是个双向素数（即796951与159697都是素数）。314159恰好是三个素数31、41、59连写而成，它们都是孪生素数中的一个；不仅如此，这三个素数的和$31+41+59=131$，它们的立方和$31^3+41^3+59^3=304091$，以及五次方和$31^5+41^5+59^5=859409651$也都是素数。

1965年，美国数学家马丁·伽德纳(M. Cardner)研究了π的前32位小数：在这32位小数中，出现了两个26，以第二个26为中心，其两侧有三对数79、32、38处在对称的位置上；并且第一个26前五个数码与后五个数码之和$1+4+1+5+9+5+3+5+8+9=50$，恰是第二个79后的两位数；而第二个26前两个数码所组成的两位数46与后两个数码所组成的两位数43，它们的和是89，又恰是第一个79前的两位数。马丁·伽德纳为什么要研究π的前32位小数呢？据说是因为很多自然现象与数字32有关，例如：人有32颗牙齿，水晶体共有32类，水在华氏32度结冰，地球的重力加速度为32英尺/秒²，原子的第四级轨道可以

3. 14159 26 535 89 79 32 38 46 26 43 38 32 79 50

充满32个电子而成为稳定状态，基本粒子中有32种长命粒子等等。更有趣的是：π的前32位小数中出现的两个26与群论中的有限单群的总数完全一致。当时，这一点尚未证实，只是有人猜想有限单群的总数可能是26。马丁•伽德纳敏锐地看到了这点，他认为这绝非偶然。果然，1983年，数学家们证实了有限单群的总数确是26。



如果把26个英文大写字母按顺序排成一个圆圈，把其中左、右对称的字母去掉，而只留下并非左、右对称的字母（图中打框者）则留下的字母依次形成一个个字母群，从J开始，各个字母群所包含的字母数，正是准确到第四位小数的π值3.1416。

还有人发现，若将下图左边所示的五阶幻方（其纵、横、斜每五个数字之和都是65）中的各数换成π的相应位上的有效数字（1换成π的第1位有效数字3，2换成π的第2位有效数字1，3换成π的第3位有效数字4，……，25换成π的第25位有效数字3），就得到下图右边所示的样子，其各行之和（标在每行右边）恰对应于各列之和（标在每列下面）：

π与数学上的其它重要常数e、i、1、0也有十分密切的联系：

1719年，意大利数学家法革纳（Fagnano）发现：

17	24	1	8	15
23	5	7	14	16
4	6	13	20	22
10	12	19	21	3
11	18	25	2	9

五阶幻方

$\rightleftarrows$

2	4	3	6	9	24
6	5	2	7	3	23
1	9	9	4	2	25
3	8	8	6	4	29
5	3	3	1	5	17
17	29	25	24	23	

$$\pi=4\ln\left(\frac{1-i}{1+i}\right)^{\frac{i}{2}}$$

1740年，瑞士数学家欧拉（Euler）发现

$$e^{i\pi}+1=0.$$

第八章

神通广大的e

e是自然对数$\ln x$（即$\log_e x$）的底，其值为

$$e = 2.71828\ 18284\ 59045\ 23536\ 02874\ 71352\ 66249\ 77572\ 47093\ 69995\cdots\cdots$$

是无限不循环小数（即无理数）。今天，人们用电子计算机已经把e值计算到了上万位的有效数字。

从数学理论上的定义来说，通常把e定义为无穷递增有界序列

$$\left(1+\frac{1}{1}\right)^1,\ \left(1+\frac{1}{2}\right)^2,\ \left(1+\frac{1}{3}\right)^3,\ \cdots,$$

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

当其项数n趋向无穷大时的极限，即

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

先让我们粗浅地解释一下这个定义：

所谓序列，是指按一定规律排列的一列数，此处$\left\{\left(1+\frac{1}{n}\right)^n\right\}$是按$n$从小到大的顺序排列的。

所谓无穷，指的是序列没有尽头，每一项后面都还有紧跟

它的后一项。

该序列$\left\{\left(1+\frac{1}{n}\right)^n\right\}$从第二项起，每一项都比其前一项要大，这就是"递增"的含义。虽然如此，但其所有各项却都小于3，这就是"有界"的含义。

从下表

n	$\left(1+\frac{1}{n}\right)^n$	n	$\left(1+\frac{1}{n}\right)^n$
1	2	50	2.6915
2	2.25	100	2.7051
3	2.3074	200	2.7164
4	2.4414	1000	2.71692
5	2.4883	10000	2.71815
10	2.5936	↓	↓
20	2.6534	∞	2.71828…

可以看出，$\left(1+\frac{1}{n}\right)^n$是个随$n$的增大而不断增大的变量，而$e=2.71828\cdots$却是个与$n$无关的常量。当项数$n$越来越大时，变量$\left(1+\frac{1}{n}\right)^n$越来越接近常量$e$，我们把这一事实用数学式子表示成

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

并称常量e是变量$\left(1+\frac{1}{n}\right)^n$当$n$趋向无穷大时的极限。

人们是通过制作对数表开始认识数e的。

我们知道，利用对数简化繁难计算，必须使用专门的对数表。由于我们习惯于十进制计数系统，当然以10为底的对数对于我们就比较方便。但是，历史上最早的对数表却并不是以10为底的。无论是苏格兰数学家纳皮尔（Napier），还是瑞士力学家彪奇（Biirgi），他们的第一张对数表的底都与e有关（事实上，他们几乎是同时各自独立地制作出了第一张对数表，纳皮尔的对数表以$\frac{1}{e}$为底，彪奇的对数表以e为底）。

为什么呢？原来，如果一开始就直接编制以10为底的对数表，哪怕只精确到万分位，也会碰到要把10、100、1000等数开10000次方的难以克服的困难。看看下表就可以明白这一点：

以10为底的对数$\lg N$	相应的真数N
0.0000	$10^0=1$
0.0001	$10^{0.0001}=\sqrt[10000]{10}$
0.0002	$10^{0.0002}=\sqrt[10000]{100}$
…	…

而如果不以10为底，改以某个常数的10000次幂为底，就可以避免制表时开方的困难。下表是以2^{10000}为底的情形：

$\log_a N(a=2^{10000})$	相应的真数N
0.0000	$(2^{10000})^{0.0000}=1$
0.0001	$(2^{10000})^{0.0001}=2$
0.0002	$(2^{10000})^{0.0002}=2^2=4$
0.0003	$(2^{10000})^{0.0003}=2^3=8$
…	…

但是，如果底很大，相应的真数的间隔也就很大，很多数的对数我们就无法取得。例如，在上面以$a=2^{10000}$为底的对数表中，3、5、7…等等一连串数的对数就都没有。

经过一段探索之后，容易发现，以$a=r^{10000}$为底制作（四位）对数表，r越接近1，相应的真数的间隔也就越细密，所制成的对数表也就越精确。下面是以1.1^{10000}、1.01^{10000}、1.0001^{10000}为底的对数表的比较：

$\log_a N$	$a=1.1^{10000}$时相应的N	$a=1.01^{10000}$时相应的N	$a=1.0001^{10000}$相应的N
0.0000	1.0000	1.0000	1.0000
0.0001	1.1000	1.0100	1.0001
0.0002	1.2100	1.0201	1.0002
0.0003	1.3310	1.0303	1.0003
0.0004	1.4641	1.0406	1.0004
…	…	…	…

就这样，纳皮尔和彪奇在制作对数表的过程中不谋而合地采用了形如$a=\left(1+\frac{1}{n}\right)^n$的底。事实上，$a=1.0001^{10000}=\left(1+\frac{1}{10000}\right)^{10000}$就是彪奇对数表的底，纳皮尔对数表的底则相当于是$a=1.0000001^{10000000}=\left(1+\frac{1}{10^7}\right)^{10^7}$，它们都是数$e$的近似值。

因此可以说，正是数e帮助人们制作了世界上第一个比较完整、实用的对数表。

不过，人们对于数 e 的真正认识，还是在17世纪中叶，数学家们发现了双曲线下的面积与自然对数之间的关系 $\left(\int_{1}^{x}\frac{1}{x}dx=\ln x\right)$ 之后。从那以后，人们逐渐了解到，很多重要的函数、重要的极限、微分与积分，很多物理现象和生物现象，都与数 e 有着极为密切的关系。所以，几乎没有哪一本数学分析教科书的开头几章不专门介绍数 e 和 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}=e$ 这一数学中最重要的基本极限之一。

现代科学技术的迅猛发展，使人类跨入了太空领域，其中也有 e 的一份汗马功劳。

大家知道，发射人造卫星或宇宙飞船，必须使用火箭。俄国的星际航行专家齐奥科夫斯基早就发现，火箭的运动速度可以用方程

$$v=w\ln z \qquad (*)$$

表示，其中 v 是火箭的理想速度，w 是火箭发动机的排气速度，z 是火箭的最初质量与发动机停止时的质量之比（简称质量比）。

从火箭的运动方程（*）可以看出，要想提高火箭的运动速度 v，要么增加喷气速度 w，要么增加质量比 z。而火箭的最终速度至少要达到第一宇宙速度（约8千米/秒），才能将人造卫星送至运行轨道。

由于目前技术水平的限制，固体燃料的最高喷气速度只能达到2.5千米/秒，液体燃料一般为3千米/秒，最好的混合液体燃料也只能达到4千米/秒。同时，质量比 z 也无法提高很多，$z=6$ 已经相当困难，最大也难于超过10。即使按 $w=4$ 千米/秒，

$z=100$的理想情形计算，火箭的理论速度虽然可以达到9千米/秒，但除去重力和空气阻力等造成的损失（约为2千米/秒），实际速度只有7千米/秒，人造卫星还是无法送至运行轨道。

科学家们通过研究火箭运动方程，想到了采用多级火箭来解决火箭的速度问题，由接连的几个火箭发动机组成一个动力系统推动火箭前进：第一级发动机使火箭达到一定的速度，燃料耗尽以后就自动脱落；第二级发动机立即接着工作，使火箭达到更高的速度，燃料耗尽以后又自动脱落；第三级发动机再接着工作，……。

假设各级发动机的喷气速度和质量比都相同，即

$$w_1=w_2=\cdots\cdots=w_n=w,$$

$$z_1=z_2=\cdots\cdots=z_n=z,$$

则火箭的最终速度

$$\begin{aligned}v&=v_1+v_2+\cdots\cdots+v_n\\&=w_1\ln z_1+w_2\ln z_2+\cdots+w_n\ln z_n\\&=nw\ln z\end{aligned}$$

也就是在这种情况下，每增加一级火箭，就可以使火箭的速度提高一倍。

现代的宇航技术，使发动机的喷气速度达到$w=3$千米/秒，质量比达到$z=3$还是不难的。只需三级火箭，就可以使最终的理论速度达到$v=3w\ln z=3\times3\ln3\approx10$（千米/秒），即使实际损失2千米/秒，也足以达到8千米/秒的第一宇宙速度。

数e帮助人们揭示出火箭运动的规律，启发人们设计出多级宇航火箭，终于实现了人们梦寐以求的遨游太空的愿望。我们说e神通广大，其实并不为过。

对于数学家们来说，e的另一个功劳也许更值得赞誉，那

就是它帮助人们证明了π的超越性。

超越数是相对于代数数而言的。

我们把整系数代数方程

$$a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$$

（其中a_0、a_1、…、a_n是整数，n是自然数，且$a_0\neq0$）的根称为代数数。

例如，所有的有理数都是一次代数数（事实上，有理数$\frac{q}{p}$就是一次代数方程$px-q=0$的根）。又如，$\frac{1+\sqrt{5}}{2}$是二次代数数（它是二次代数方程$x^2+x-1=0$的根）等等。

一个数如果不是任何整系数代数方程的根，就叫做超越数。"超越数"这一术语是欧拉首先引入的，意思是这种数"超越了代数方法的能力"。不过，欧拉并未给出任何具体的超越数，直到一百多年后的1844年，法国数学家刘维尔(Liouville)才证明了超越数确实存在。现在我们已经知道，超越数甚至比代数数还要多得多。

人们早就猜测π和e可能是超越数。1873年，法国数学家厄尔米特（Hermite）证明了e是超越数。1882年，德国数学家林德曼（Lindemann）在厄尔米特证明e是超越数的基础上，借助欧拉公式$e^{i\pi}+1=0$，终于证明出了π也是超越数。而在此之前，π的超越性一直是数学家们几乎啃不动的硬骨头。

知道了π是超越数，古希腊遗留下来的三大尺规作图难题之一——"化圆为方"也就迎刃而解了。由解析几何可以知道，一个几何图形可以用尺规作出，必需并且只需其中所要作出的几何量可以由给定的几何量经过有限次的四则运算和开平方获得。在"化圆为方"的问题中，涉及到$\sqrt{\pi}R$（R是给定

圆的半径）这样的几何量，因为π是超越数，不可能由R经有限次四则运算和开平方而获得$\sqrt{\pi}R$，所以用尺规作图来“化圆为方”也是不可能的。

“化圆为方”问题的解决，用到了π的超越性，而π的超越性又是通过e的超越生来证明的。归根结底，还是e帮助我们解决了这个有两千多年历史的不知使多少数学家感到头疼的难题！怎能不令人对e刮目相看呢！

第九章

“虚无飘渺”的i

i是什么？用比较现代的数学术语来说，i是虚单位，它满足于$i^2=-1$，是-1的一个平方根。有时候我们也可以把i不太严格地表示成$\sqrt{-1}$。

16世纪以前，人们对于数的认识和使用还只限于实数的范围。由于所有实数的平方总是非负的（即大于或等于零），人们找不到任何一个实数x能够满足$x^2=-1$，因而也就认为-1开平方（即$\pm\sqrt{-1}$）没有意义。不仅如此，任何一个负数的偶次方根也都是没有意义的。事实上，当时不要说负数的偶次方根，甚至连负数本身在欧洲数学界也尚未站稳脚跟（这与我国公元一世纪就已经普遍地使用负数成了鲜明的对照）。

有文字记载最早接触负数开平方的，是12世纪的印度数学家培斯卡拉（Bhāskara）。他在研究二次方程

$$x^2-px+q=0$$

时发现，当$p^2<4q$时，就会在解方程过程中出现负数开平方。可惜他并不承认负数有平方根，他认为：“正数和负数的平方恒为正数；正数的平方根有二，一正一负；负数无平方根，因其非平方之数。”

16世纪中叶，意大利数学家卡丹（Cardan）在研究二次与三次方程的解法时又遇到了负数开平方的问题。特别使他惊奇的是，按照他从数学家塔尔塔里亚（Tartaglia）那里学来的

著名的三次方程求根公式（现称卡丹公式），某些三次方程的实数根竟可以用含负数平方根的代数式表示出来，例如$x^3=15x+4$的三个实数根$x_1=4$、$x_{2,3}=-2\pm\sqrt{3}$就可以用含负数平方根的代数式统一表示成

$$x=\sqrt[3]{2+\sqrt{-121}}+\sqrt[3]{2-\sqrt{-121}}.$$

因此，尽管感到十分为难，他还是"不管受到多大的良心责备"承认负数的平方根仍然是"数"。例如，他曾把方程$x(10-x)=40$的两根表示成$x_1=5+\sqrt{-15}$和$x_2=5-\sqrt{-15}$。虽然卡丹认为负数的平方根是"虚构的""超诡辩的量"，但他毕竟首次将负数的平方根引进了数学领域，允许它们作为某些方程的根，允许它们参加某些数学运算，而不再把它们拒之门外。

16世纪末叶，另一位意大利数学家邦别利（Bombelli）也在解三次方程的过程中遇到了负数的平方根。比卡丹更进一步的是：他不但理直气壮地承认负数的平方根是实实在在的数，而且大胆地建立起了一整套有关$\sqrt{-1}$甚至$a+b\sqrt{-1}$（其中a、b是实数）的运算法则，邦别利的工作实质上是复数理论的开端。

但是，18世纪以前，包括一些相当杰出的数学家在内的绝大多数数学家对于负数开偶次方仍持怀疑态度。就连17世纪的法国大数学家笛卡儿（Descartes）也认为负数开平方是"不可思议的"，称之为"虚数"，这一名称就一直沿用到现在。

从18世纪初，情况开始有所转变，不少数学家，如德国的莱布尼兹（Leibniz），瑞士的约翰·伯努利（Johann Bernoulli），法国的棣美弗（De Miovre），尽管对复数的本质还不甚理解，仍然干劲十足地使用虚数来解决某些数学问题，甚至还将这种使用扩展到了积分和对数的邻域。这与微积分刚刚诞

生前后的数学界大胆创新的风气很有关系。微积分刚刚诞生的时候，理论基础尽管十分薄弱，且混乱而不严密，数学家们仍然信心十足地使用微积分成功地解决了当时所遇到的大多数数学和物理问题，这些成功极大地鼓舞了数学家们毫无顾忌地涉及未知的数学领域。

18世纪70年代，瑞士数学家欧拉（Euler）引进了虚单位i，系统地建立起了以形如$a+bi$的数（其中a、b是实数）为运算对象的复数理论（不过，"复数"这一名称直到19世纪才由德国数学家高斯（Gauss）给出）。欧拉利用法国数学家棣美弗1722年所得的著名公式

$$(\cos x+i\sin x)^n=\cos nx+i\sin nx$$

发现了三角函数与e的虚指数幂之间的关系（欧拉公式）

$$e^{ix}=\cos x+i\sin x.$$

并由此揭示出数学上两个极为重要的常数π和e可以通过虚单位i联系起来，即

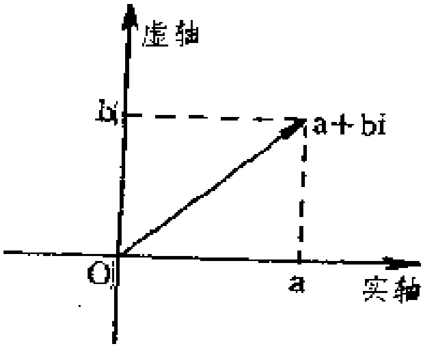
$$e^{i\pi}+1=0.$$

欧拉还由他所发现的三角函数与e的虚指数幂之间的关系，推导出了今天被称为复变函数这一数学分支中的一些基本定理。

18世纪末，挪威的一位测量学家威塞尔（Wessel）提出可以把复数$a+bi$看做平面上坐标为(a, b)的点，从而使复数的全体与平面上的点的全体发生了一一对应的关系，比较完整地给出了复数意义的几何解释。从此以后，复数在大多数数学家的眼里，总算有了一席立足之地。

19世纪初，德国数学家高斯又进一步发展了威塞尔的思想，用互相垂直的实轴和虚轴建立起了复平面，并正式引进了复数的概念。高斯正确地指出："这样的几何表示使人们对虚数

真正有了一个新的看法。"由于"看得见"、"摸得着"，虚数已不再是虚无飘渺没有实际意义的数了，它广泛渗透到了各个数学分支，并在流体力学、热力学、地图学上得到了应用。



以后，又经过法国数学家柯西(Cauchy)，德国数学家黎曼(Riemann)和魏尔斯脱拉斯(Weierstrass)等许多数学家的努力，有关复数理论的一系列重要成果终于形成了堪称现代数学基础之一的复变函数论，成为理论物理、天体力学、现代电子和航天技术不可缺少的数学工具。

在此只要略举一例就够了。自从19世纪初德国物理学家欧姆(Ohm)提出电路中的电流强度I与加在电路两端的电压U成正比，与电路本身的电阻R成反比的欧姆定律$I=\frac{U}{R}$以来，人们曾对它的适用范围有过较长时间的争论。尽管欧姆定律在直流电路上得到了十分精确的验证，不少人仍怀疑它能否适用于交流电路的情形。19世纪末，英国电学家奥利夫·亥维赛(Oliver Heaviside)终于用复数分析的方法澄清了这个长期遗留的问题。他从理论上证明了，只要假设感抗为$i\omega L$，容抗为$1/i\omega C$，则欧姆定律在交流电路上也同样成立。以此为基础，人们逐渐建立起了交流电路的电学理论。

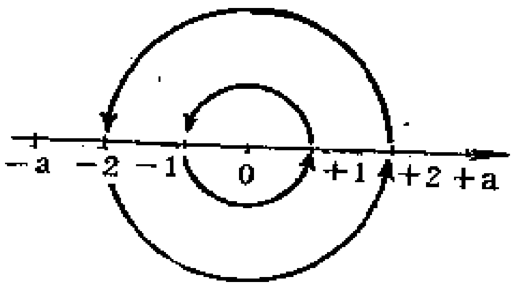
值得注意的是，我国18世纪末到19世纪初的两位数学家汪莱和李锐，也曾独立地探讨和研究过虚数的概念及其性质（不过，他们把虚数称为"无数"），并且还得到过实系数方程的虚

根成双的重要结果。但是由于历史的原因，当时我国正处在封建社会末期，又长期闭关自守，科学技术和数学十分落后，与先进国家差距很大，整个数学界万马齐喑，汪莱与李锐孤立的研究工作未能继续深入下去。

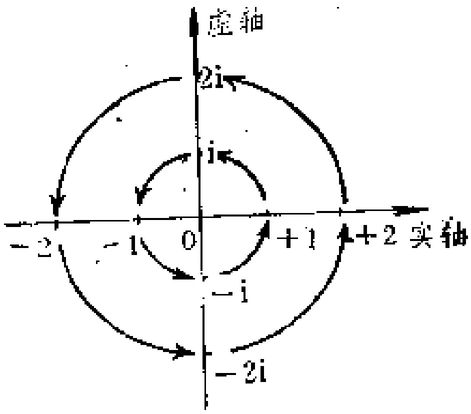
历史上，虚数的产生是由于负数的开平方引起的，这对于早已习惯于实数世界的人们来说是很难接受的，所以虚数从产生到它在数学中不可动摇的地位的确立，足足经历了两百多年漫长的历程。其实，不要说古人，就是我们现代人，刚刚接触虚数时，也对为什么$i^2=-1$普遍感到不好理解。

这个问题，如果用旋转变换的观点，就比较容易得到直观而自然的回答。

观察如图所示的数轴。我们知道，实数与数轴上的点是一一对应的，而绝对值相等，符号相反的一对实数，如±1、±2、±a等，总是位于原点两侧关于原点对称的。正因为如此，我们可以把任何一个负实数点看做是相应的正实数点（严格地说是该点所代表的位置向量）按反时针方向绕原点旋转180°而得到的；任何一个正实数点也可以看做是相应的负实数点按反时针方向绕原点旋转180°而得到的。这就是说，任何一个实数点按反时针方向绕原点旋转180°就得到其相反数。另一方面，任何一个实数乘−1也同样得到其相反数。显然，把一个实数点按反时针方向绕原点旋转180°与把这一点所代表的实数乘上−1所起的效果完全相同。换句话说，把一个实数乘上−1，就相当于把这个实数点按反时针方向绕原点旋转180°。



那么，如果把数轴上的点按反时针方向绕原点旋转90°又怎样呢？完全类似地可以规定：任何一个实数乘上某个数i以后，就相当于把相应的实数点按反时针方向绕原点旋转90°。我们暂且不去追究数i到底是什么，而仅仅认为它的作用就是使与它相乘的实数所对应的点按反时针方向绕原点旋转90°。把数轴上所有的点都按反时针方向绕原点旋转90°以后，我们就得到了过原数轴的原点而与其垂直的另一新数轴。但这根新数轴上的点所代表的数已经与原来的实数不同，而是相应的实数乘i了，如$\pm i$对应于± 1，$\pm 2i$对应于± 2，$\pm ai$对应于$\pm a$等等。为了与原数轴相区别，我们不妨把原数轴称为实轴，而把新数轴称为虚轴，虚轴上任何一点都是由实轴上相应的点按反时针方向绕原点旋转90°得来的。我们把虚轴上的点所代表的数称为虚数，为相应的实数乘上i。



任何一个实数乘i之后就成为相应的虚数，任一虚数再乘i呢？无非相当于把相应的虚数点又按反时针方向绕原点再旋转90°，结果又落回到实轴上，成为最初那个实数的相反数。特别是，1乘i变为i，i再乘i变为-1，恰好意味着$i^2=-1$。事实上，按反时针方向接连旋转两个90°，等于按同一方向旋转180°，因此也有$i^2=-1$。就是说，i是-1的一个平方根。现在我们终于知道i是什么了！由于虚轴上的点所代表的虚数都是由相应的实数乘i得来的，因此我们又把i叫做虚单位。

第十章

神奇的对数与RMI方法

我们知道，在中学数学里，“对数”是这样定义的：

如果一个不等于1的正实数a的b次幂等于N，则幂指数b就叫做以a为底的N的对数，并记为$b=\log_a N$，即

$a^b=N \Longleftrightarrow b=\log_a N$（$a$、$b$、$N\in R$，$a>0$，$N>0$且$a\neq 1$）。

对数有几个重要的性质：

$$\log_a(N\cdot M)=\log_a N+\log_a M$$

$$\log_a \frac{N}{M}=\log_a N-\log_a M$$

$$\log_a N^n=n\log_a N$$

$$\log_a \sqrt[n]{N}=\frac{1}{n}\log_a N$$

对数可以将乘、除、乘方、开方运算降低一级，转化成相应的加、减、乘、除运算。

历史上，对数的出现是为了满足人们对于繁难计算化简的要求。当时（16世纪）的欧洲正处在文艺复兴时期，科学技术得到了空前的解放，天文、测绘、航海事业蓬勃发展，同时也带来了大量的令人头疼的复杂的计算问题。面对这种咄咄逼人的形势，不少数学家在考虑：能否将乘除运算转化成加减运算，以减轻科学家们的计算负担？有人试图用三角函数的积化和差来达到这一目的；有人则试图通过协调等差数列与等比数

列之间的关系来解决这一问题。

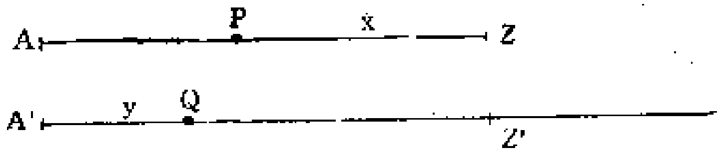
1544年，德国数学家史蒂福（Stifel）发现，如果把整数序列（以1为公差的等差数列）与2的各次幂序列（以2为公比的等比数列）一一对应地排列如下：

$$\cdots,\ -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ \cdots$$

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

则下排的等比数列中数的乘除运算就可以通过相应的上排的等差数列中数的加减运算来实现。史蒂福将上排数称为相应的下排数的“指数”，意即“代表者”。举例来说，要求下排数中$\frac{1}{4}$与32的乘积，可以先查出其相应的指数分别为-2与5，然后将此二指数相加得3，则与指数3对应的下排中的数8就是$\frac{1}{4}$与32的乘积；要求下排数中8除以64的结果，可以先查出其相应的指数分别为3与6，然后将此二指数相减得-3，则与指数-3对应的下排中的数$\frac{1}{8}$就是$8\div64$的商。史蒂福的这一发现虽然已经初步接触到了“对数”的实质，但是由于划分不够细密（当时还没有分数指数的概念）而无法付诸实用。

苏格兰数学家纳皮尔（Napier）于16世纪末进一步研究了等差数列与等比数列之间的对应关系。纳皮尔假设有两个质点P与Q分别沿着线段AZ和射线$A'Z'$以同样的初速度运动，沿线段AZ运动的质点P的速度正比于它离开Z的距离，沿射线$A'Z'$运动的质点Q的速度则始终保持不变。若$PZ=x$，$A'Q=y$，则x在变化时可以看成一个无穷递缩等比数列，而y在变



化时可以看成一个无穷递增等差数列。通过 y 即可把有关 x 的乘除运算转化成加减运算，纳皮尔称 y 为 x 的对数（事实上，由微积分易知 $y=\log_{1/e}x$，e 是自然对数的底）。纳皮尔据此制作了世界上第一张对数表，并于1614年正式出版。

由于纳皮尔对数是以 $\frac{1}{e}$ 为底的，使用起来仍有些不大方便。与纳皮尔同时代的英国数学家布立格斯（Briggs）根据他在牛津大学讲授纳皮尔对数的经验和体会，将对数的底改为10，制作了使用更为简便的常用对数。

18世纪法国著名的大数学家拉普拉斯（Laplace）高度评价了纳皮尔发明对数的伟大意义，他称赞说："对数的发明，等于将科学家的寿命延长了两倍。"

对数的出现，不仅引起了计算技术上的进步，而且激发了指数函数与对数函数等基本概念的产生和发展，有力地推动了数学前进的步伐。

在科学技术高度发达、电子计算机日益普及的今天，对数在计算技术上的重要作用已经日渐消失，但是纳皮尔利用对数将乘除运算转化成加减运算，以简化繁难计算的思想方法却仍然闪烁着千古不朽的光辉，并逐步形成了现代数学中的"关系—映射—反演"（RMI）原则。

什么叫"关系—映射—反演"呢？以日常生活中的梳头为例：我们常常喜欢对着镜子梳头，镜子里映出我们的头发和梳

子的象，我们对照镜子中的象，不断调整梳子与头发间的关系，当镜中头发的象梳理好了以后，我们也就梳理好了真正的头发。这一过程可以用框图表示如下：

$$
\begin{tikzcd}
\boxed{\text{头发和梳子}} \arrow[r, "\text{镜面映射}"] \arrow[d] & \boxed{\text{头发和梳子的象}} \arrow[d] \\
\boxed{\text{实际梳好的真实头发}} & \boxed{\text{镜内梳好的头发的象}} \arrow[l, "\text{镜面反射}"']
\end{tikzcd}
$$

当纳皮尔用对数来简化繁难计算的时候，思想方法也是类似的。例如，用对数计算

$$\sqrt{84.62 \times (14.61)^3}$$

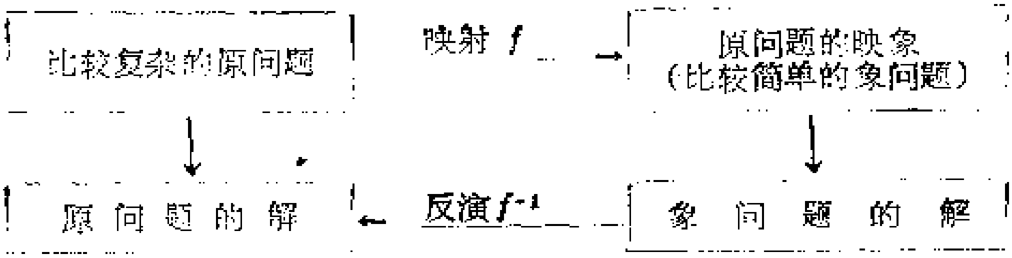
的过程可以用框图表示如下：

$$
\begin{tikzcd}
\boxed{\sqrt{84.62 \times (14.61)^3}} \arrow[r, "f(=\lg)", "\text{(映射：取对数)}"'] \arrow[d] & \boxed{\frac{1}{2}\lg 84.62+3\lg 14.61} \arrow[d] \\
\boxed{28687} & \boxed{\begin{array}{c}\frac{1}{2}\times 1.92747+3\times 1.16465 \\ =4.45763\end{array}} \arrow[l, "f^{-1}(=\lg^{-1})"', "\text{(逆映射：取反对数)}"]
\end{tikzcd}
$$

这一过程的实质是通过取对数这一映射，把比较繁难的原数之间的乘、除、乘方、开方等运算，转化成相应的对数之间的加、减、乘、除等运算，从而化繁为简，降低难度，再通过取反对数的逆映射，得到原数的计算结果，这样就大大提高了计算的效率。

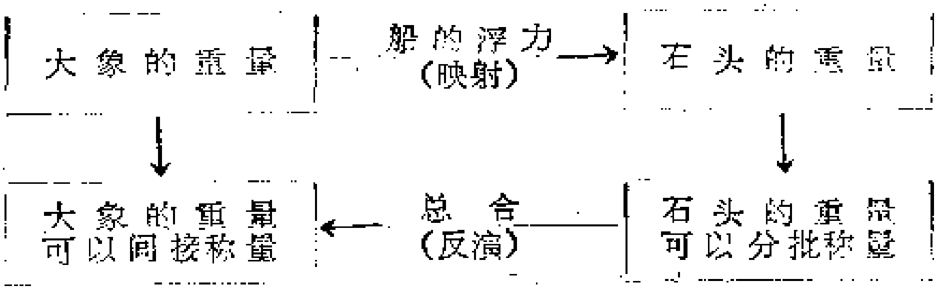
一般地说，“关系—映射—反演”(RMI) 原则就是：通过某种映射与反演之间的对应关系，将比较复杂的问题转化成比较简单的问题来加以解决。其过程可以用框图表示成(见69框图)

这就是纳皮尔用对数来简化繁难计算的思想方法的精髓，也是纳皮尔留给我们的最宝贵的历史遗产。难怪恩格斯把它与



笛卡尔的解析几何，牛顿-莱布尼兹的微积分并列为“最重要的数学方法”。

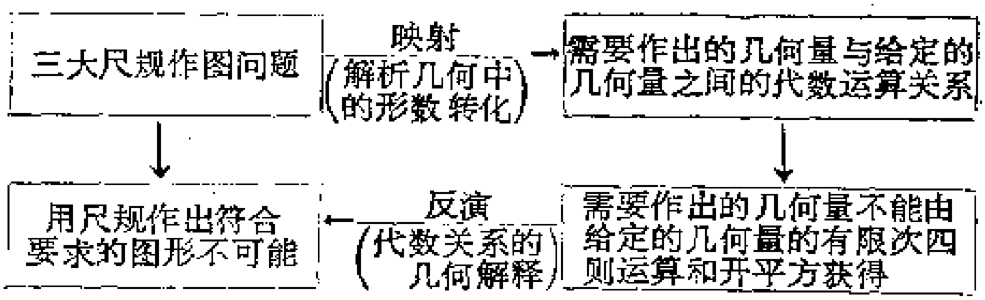
RMI原则的思想方法，其实也是我国人民所早已熟悉的。妇孺皆知的发生在一千七百多年以前的“曹冲称象”的故事，就生动地体现了这一思想方法。年幼的曹冲之所以聪明，就在于他敏锐地看出了“称量大象的重量”这一问题可以利用船的浮力转化成“称量石头的重量”，从而轻易地解决了满朝文武都未能解决的问题。“曹冲称象”的过程可以用框图表示成：



可惜，我们的祖先并未从中总结、提炼出一个一般的原理或方法，这自然和历史条件的限制有关，也和我国传统的只重实用、不重基础理论研究的风尚有关。

数学史上很多疑难问题的解决都是人们自觉或不自觉地在RMI原则的指导之下完成的。古希腊三大尺规作图难题就是如此。早在公元前三、四百年间，古希腊数学家就提出了用无刻度的直尺和圆规作一个正方体，使其体积为给定正方体体积的两倍（倍立方体）；将任意给定的角三等分（三等分角）；作一

个正方形，使其面积为给定的圆面积（化圆为方）的问题。但是，十八世纪解析几何产生之前，尽管许多才华横溢的数学家为之付出了巨大的心血，也仍然未能彻底解决这些问题。直到法国数学家费马和笛卡尔创立了解析几何以后，数学家们用代数方法对尺规作图的局限性进行了分析，发现一个几何图形可以用尺规作出的充要条件是：其中所需要作出的几何量，可以由给定的几何量经过有限次的四则运算和开平方获得。而三大尺规作图难题中所需要作出的几何量，却不能满足这样的条件，因此要用尺规作出符合要求的图形是不可能的。这一问题的RMI框图如下：



数学家们在几何基础的研究中，为了证明非欧几何与欧氏几何的相对相容性（无矛盾性），也使用了 RMI 方法。大致过程是：首先订出一套解释规则，规定非欧几何中的点、直线、平面等各对应于欧氏几何中的什么概念，非欧几何中的相交、介于、合同关系等各对应于欧氏几何中的什么关系（这相当于建立起非欧几何与欧氏几何之间的一个——映射），使得非欧几何的公理和定理，经过解释和翻译后能够成为相应的欧氏几何中的公理和定理。如果非欧几何出现矛盾，相互矛盾的非欧几何定理经过解释和翻译就成了欧氏几何中相互矛盾的定理，从而欧氏几何也就有矛盾了。人们既然承认欧氏几何没有矛

盾，当然也就得承认非欧几何也没有矛盾。更进一步，用笛卡尔的解析几何所建立的“形”与“数”之间的对应，欧氏几何中的公理和定理还可以转化成实数理论中相应的公理和定理，如果欧氏几何有矛盾，那么实数理论也就有矛盾；如果承认实数理论无矛盾，那么也就得承认欧氏几何无矛盾。

RMI原则已经成为数学中最重要最基本的原则之一。数学中所经常采用的变量替换、三角代换、以及现代运算微积中的拉氏（Laplace）变换和积分方程中的弗雷德霍姆(Fredholm)基本公式，都可以看做是RMI原则的具体运用。

第十一章

从万能代数模型到MM方法

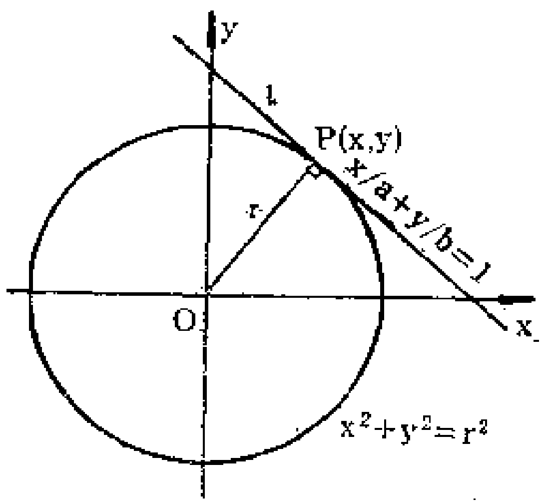
解析几何的主要创始人——17世纪法国大哲学家、数学家笛卡尔在他生前未能来得及写完的遗著《指导思维的法则》中，设计了一套旨在用于解决一切问题的“万能代数模型”方案，这套方案的梗概可以大致描述如下：

首先，把任何类型的问题，都归结为数学问题；然后，把任何类型的数学问题，都归结为代数问题；最后，把任何类型的代数问题，都归结为单一方程的问题。

从《指导思维的法则》一书的写作宗旨及其背景来看，这部遗著应该说是笛卡尔毕生宝贵经验的总结。事实上，“万能代数模型”的想法，是笛卡尔在创建解析几何的过程中逐步发展和成熟起来的。笛卡尔之前，几何和代数这两门学科早就发展起来了。笛卡尔认为，欧几里得几何过分强调证明的技巧，过分依赖图形，不利于提高人们的想象力；而代数又太受法则和公式的束缚，影响人们思维的灵活性。笛卡尔觉得应把代数和几何结合起来，取长补短。由于代数方法比几何方法更便于处理，代数语言较几何语言更富于启发，笛卡尔开始用代数方法来研究几何作图问题，并萌发了用方程表示曲线的想法。更进一步，笛卡尔试图创造一种方法，以便解决所有的几何问题。为此，笛卡尔建立了坐标系，通过点与坐标（有序数对）之间的对应关系，把几何中的“形”与代数中的“数”有机地

联系起来，从而使人们可以用代数方法来研究几何问题。

例如，在直角坐标系中，横截距为a、纵截距为b的直线l可以用方程$\frac{x}{a}+\frac{y}{b}=1$来表示；而以原点为圆心、以$r$为半径



的圆O时可以用方程$x^2+y^2=r^2$来表示；判断直线l与圆O是否相切这一几何问题就可以归结为方程$\frac{x}{a}+\frac{y}{b}=1$与$x^2+y^2=r^2$是否有并且只有唯一的一个公共解的问题。

笛卡尔通过坐标系用代数方法来研究几何问题的天才思想，开创了一门新的数学理论体系——解析几何学（这当中还有其他一些数学家，如法国的费马的功劳）。法国数学家拉格朗日（Lagrange）说得好："代数与几何分道扬镳的时候，它们的进展就缓慢，应用也有限；但是，这两门学科一旦联袂而行，它们就互相从对方吸收新鲜活力，从而大踏步走向各自的完美。"十七世纪以来数学的迅速发展，特别是微积分的出现，在很大程度上应归功于解析几何的创立。

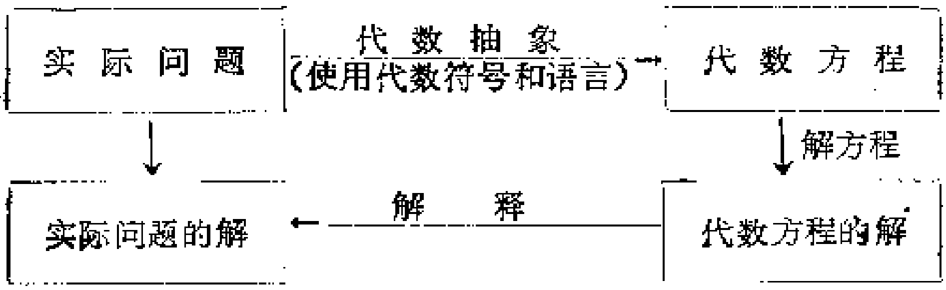
笛卡尔十分重视科学方法论的研究，他说："我所解决的每个问题都将成为一个范例，用以解决其它问题。"正因为如此，笛卡尔的思想并没有只停留在用代数方法研究几何问题上，而是继续向前发展到用数学解决一切科学问题的高度。笛卡尔认为，科学的本质就是数学，数学能解释一切自然现象，一切现象都可以用数学描述出来。在这样的思想指导下，笛卡尔提出了"万能代数模型"。

在笛卡尔那个时代，由于科学水平的限制，人们所遇到的可以用数学方法处理的大多数问题，要么可以直接归结为一个代数问题，要么可以先归结为一个几何问题，再通过"坐标"转化为代数问题。而当时的所谓"代数"，主要还是囿于方程论的古典代数。因此，当时能用数学方法处理的问题，确实有相当大的一部分可以按照笛卡尔的"万能代数模型"来解决。时至今日，当我们列方程解应用题，求曲线的交点位置，以及用数学方法处理初等的物理、化学等问题时，仍然基本上是按照这一套方案行事的。

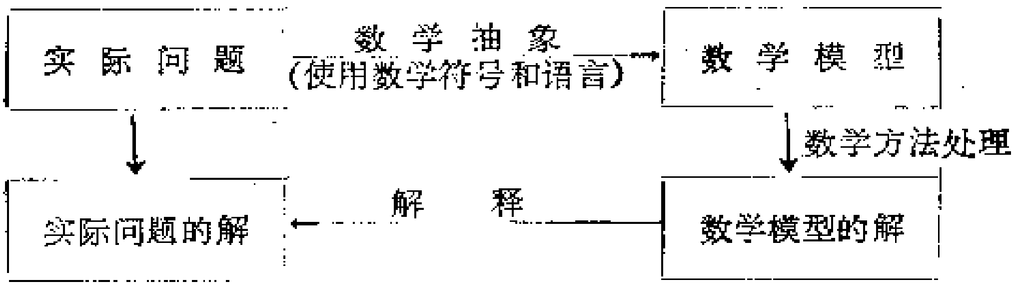
当然，若用"万能"这一尺度来衡量，笛卡尔的方案是失败了。美国著名的数学家和数学教育家波利亚（G. Polya）指出："在笛卡尔方案的基本意图中，有些内容看起来是很正确的，但是要把这种意图付诸实施，其困难、障碍和错综复杂的情况远比笛卡尔最初热情想象的要多得多。虽然笛卡尔的方案失败了，但仍不失为一个了不起的方案。纵然失败，它对科学的影响也比那些偶然成功的许多微不足道的方案大得多。"

认真分析一下笛卡尔的"万能代数模型"方案的实质，可以清楚地看到现代"数学模型（MM）"方法的雏型。

笛卡尔"万能代数模型"的框图如下：



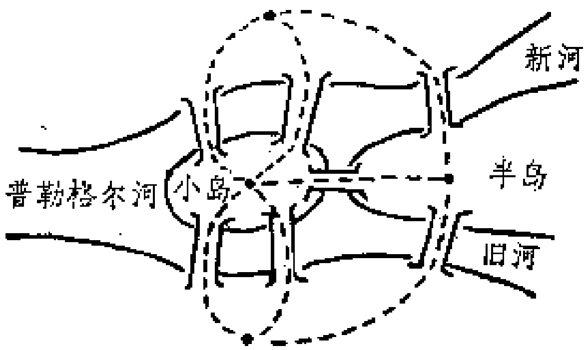
现代“数学模型（MM）”方法的框图如下：



可见，现代“数学模型（MM）”方法完全是笛卡尔“万能代数模型”方案的继续延伸和直接推广。

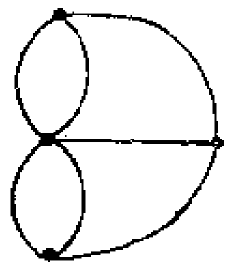
欧拉解决哥斯尼堡七桥问题，堪称是数学史上运用MM方法的典范。

在风景如画的东普鲁士的古都哥斯尼堡（现苏联的加里宁格勒），普勒格尔河横贯全城，新河与旧河两条支流，在市中心的一座公园汇合。汇合处有一个小岛和一个半岛，人们建造了七座桥把河的两岸、半岛及小岛连接起来。



附近的居民空闲时都喜欢上那座公园散步。年长日久，很自然地就产生了这么一个问题：能不能在一次散步中把所有的桥都走遍，而每座桥又都只走一次，最后回到原来的出发点？很多人都考虑过这个问题，但是都未解决。于是就有人写信向当时已颇负盛名的年仅二十多岁的大数学家欧拉请教。

欧拉知道这个问题以后，进行了仔细的分析。欧拉想：既然岛与半岛无非是桥梁的连接点，两岸陆地也是桥梁的通往地点，那就不妨把岛与半岛及两岸陆地抽象地看成缩小了的四个点，而把七座桥抽象地看成七条线。于是，人们试图一次无重复地通过七座桥的问题，就转化成了一个等价的一笔画问题。

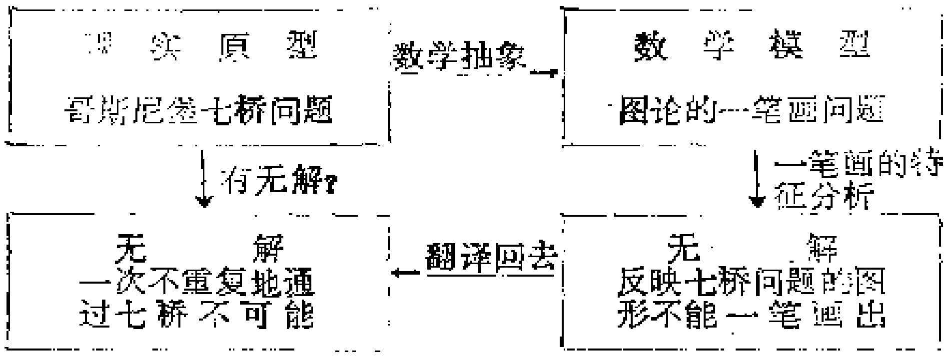


接着，欧拉又从数学的角度考察了一笔画的结构特征：一笔画总有个起点和终点（当起点与终点重合时就成为封闭图形）；除起点和终点外，一笔画中线路的交叉点每有一条曲线进，跟着就要有一条曲线出，故在交叉点处汇合的曲线数必为偶数。换句话说，一笔画中至多只有起点与终点有可能交汇奇数条曲线，除这两个点外的其它各点都只能交汇偶数条曲线。然而反映哥斯尼堡七桥问题的图中，却有四个点交汇了奇数条曲线，根据上述分析，该图不可能一笔画出。因此，人们也就不可能在一次散步中无重复地通过七座桥。哥斯尼堡七桥问题便得到了解答。

欧拉对哥斯尼堡七桥问题的分析与解答过程可以用框图表示如下：（见77页框图）

欧拉的高明之处就在于：

首先，不拘泥于特殊问题，而是从特殊问题提出和解决一



般问题，特别是用点和线来代表陆地和桥，用抽象的形式记录过桥的方式，使得七桥问题通过漂亮、清晰的数学模型而归结为网络的一笔画问题。

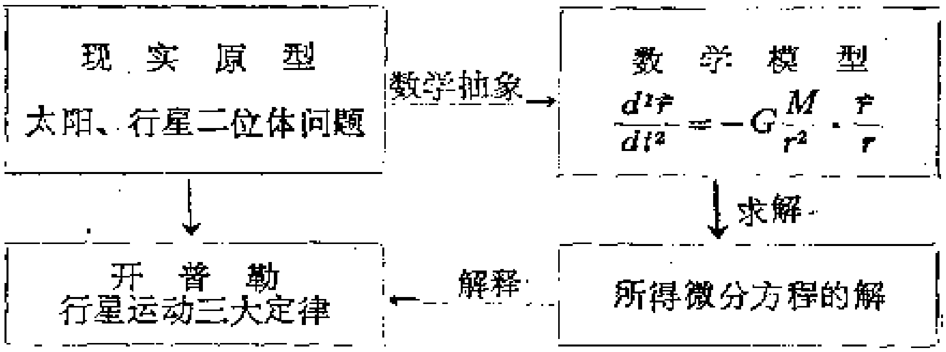
其次，欧拉所用的分析方法也很独特。他的着眼点不在桥，也没有试图一次次走或一次次画来穷尽所有的可能性；他把着眼点放在陆地上，即网络的点上，注意各个点交汇的曲线的数目，分析一笔画对各点交汇的曲线数目奇偶性的要求，从而求出了一笔画的判别准则，不但使七桥问题迎刃而解，而且一般网络的一笔画问题也迎刃而解。

运用MM方法需要注意两条原则：简单化原则与相似性原则。

简单化原则就是：在对照现实原型构造数学模型的过程中，必须用抽象分析法抛弃那些与问题的本质关系不大的次要因素，化繁为简，化难为易，以突出主要的关系结构，便于进行数学处理。

欧拉在解决哥斯尼堡七桥问题时，就把七座桥抽象地看成七条线，而把连接桥梁的河岸与小岛、半岛抽象地看成连接七条线的四个点，能否一次不重复地通过七座桥的问题就转化成了网络的一笔画问题。

17世纪的英国物理学家、数学家牛顿研究行星绕日运动的规律（太阳行星二位体问题），用的也是MM方法。他忽略了其它各种天体的影响，并把太阳与行星都看成是质量集中在中心的理想质点，从而根据万有引力定律得出了行星绕日运动的微分方程；再通过微分方程的解，推导出了开普勒（Kepler）的行星运动三大定律。



相似性原则就是：用数学符号和数学语言所建立起来的数学模型，必须能够如实地反映决定问题本质的主要关系结构。数学模型应该是与现实原型相似的简化摹本，数学模型的解能够返回原型，解决实际问题。

哥斯尼堡七桥问题中，桥的宽窄与长短，岛与半岛的面积大小，以及两岸的地理状况等，并不影响问题的实质。欧拉把桥抽象成线，把岛、半岛与河岸抽象成点，所得的图形完全保持了原问题的特征，是与现实原型相似的简化摹本。

太阳行星二位体问题中，牛顿证明了均匀的球状物体的质量确实可以看成集中于球心，而且太阳和行星本身的大小与它们之间的距离相比，其它天体对行星的作用与太阳对行星的作用相比，都微乎其微。同时，万有引力定律又是从经典力学的理论推导出来的经过实践反复检验的普遍规律，因此，牛顿所

得的行星运动微分方程，确实反映了行星绕日运动的规律，也是与现实原型相似的简化摹本。

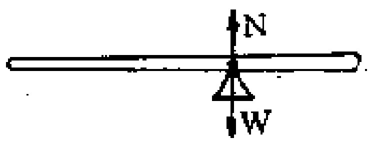
科学理论的数学化已经成了现代科学发展的基本特点之一，而科学理论数学化的关键就在于建立适当的数学模型。例如，在生物学中，我们可以用微分方程模型（$dN=kNdt$）来描述细胞的繁殖和人口的增长；可以用Bernoulli二项分布模型（$P\{X=k\}=C_n^k p^k q^{n-k}$）来描述生物群落在某个区域内的分布；可以用线性代数网络模型来描述生物神经系统中信号的传递等等。这样，就使得对这些生物现象的研究上升到了精确可靠的定量分析的水平。故有人说，这种借助数学模型来显示生物现象变化规律的方法，使生物学获得了突飞猛进的第二次生命。

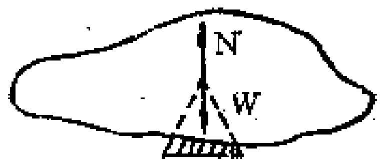
第十二章

维妙维肖的模拟方法

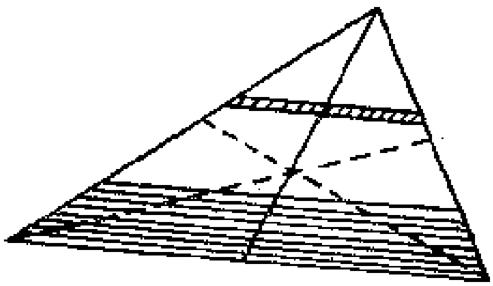
两千多年以前，古希腊数学家、物理学家阿基米德(Archimedes)，写了一部有趣的著作，叫做《一些几何命题的力学证明》。你也许会感到奇怪：难道几何命题还可以用力学方法来证明吗？确实可以。请看以下如何用力学方法来证明"三角形的三条中线相交于一点"这一几何命题：

我们知道，任何一个物体都有一个重心，物体所受的重力可以看做集中在这一点上，每个物体的重心都是唯一的。如果在一根细棒或一块平板的重心处放一个支撑物，由于物体本身的重力恰好与支撑物的弹力相抵消，细棒和平板就会保持水平上的平衡。





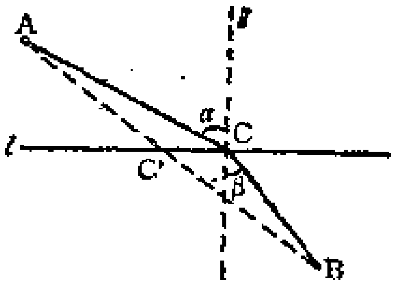
现在让我们把几何上的三角形看成一块质量均匀的薄板，用平行于底边的线段把它划分成许多很细的狭条。由于每个狭条都细而均匀，其重心都在各自的中点，而所有这些中点又恰好连接成了三角形底边上的中线。换句话说：所有这些狭条的重心都在三角形底边的中线上。因此，整个三角形的重心也必在底边的中线上。同理，三角形的重心也一定会在另两条中线



上。又，三角形的重心是唯一的。所以，三角形的三条中线必相交于三角形的重心这一点上。

再请看用光学实验的方法来解决极值问题的一例：

设河流l的两岸分别有居民点A与B。由于河岸两边地理条件的限制，汽车在两岸行驶的速度分别为v_A与v_B。如果要使汽车从A到B所花时间最少，应在l上何处架桥？



让我们把河两岸看成是以l为界面的两种不同介质，有一束光线由A出发，经界面l上的C点最后到达B；光线在这两种不同介质中传播的速度分别为v_A与v_B，汽车则可以看成是这束光线中的光子。

当光线通过两种不同介质的交界处C时，由于传播速度发生了变化，光通路的方向也会发生相应的变化，光线不再沿直线AC前进，而是走出一条折线ACB。这种现象称为光的折射现象。表面上看来，光线由A经C到B，走的是弯路；然而实际上，光线由A沿直线$AC'B$到B所花时间反而比沿折线ACB所花时间要长。不仅如此，光线沿其它任何一条路线由A到B都比沿折线ACB所花时间要长。简言之，光线在其传播过程中总是沿着时间最短的路线前进的，这是“光行最速原理”。十七世纪的法国数学家费马(Fermat)最早用微分法给予了证明。

根据“光行最速原理”，汽车的行驶路线应与光通路ACB重合，才能使从A到B所花时间最少。因此，我们完全可以用

光学实验的办法来决定这样一条费时最少的路线，并由此找出所要求的架桥位置：首先选择绝对折射率的比值为$\frac{v_A}{v_B}$的两种不同光学材料做为介质来表现河l两岸不同的地理环境，并按适当的比例在这两种介质中确定出A、B两个居民点相对于l的位置；然后以A为光源向B射出一束光线，再调整这束光线的方向，使其进入另一介质后刚好能够达到B；此时，光线束射在两种介质交界处C的位置，就相当于最佳的架桥位置。

阿基米德用力学方法来证明几何命题，以及我们用光学实验方法来解决极值问题，都属于现代科学方法论中"模拟"的范畴。在"模拟"中，我们所要真正研究的事物（如以上所举两例中的三角形及其中线；河两岸的居民点、汽车及架桥位置）叫做"原型"。但是，我们并没有直接去研究"原型"，而是构造了一个与"原型"相似的"模型"（如以上所举两例中的三角形薄板；两种不同介质的光学材料组成的光学系统）。通过对"模型"的观察、实验和分析（如将三角形薄板划分成许多很细的狭条；让光束通过两种不同的介质由A射到B），我们掌握了有关"原型"的知识，最终解决了我们原来所要解决的问题。

在科学研究中，常常有一些对象很难、或者简直不可能对它进行直接的实验和研究（例如对于生命起源的研究，我们就不可能回复到原始的地球环境中去；又如对于尚未施工的大型水库的设计是否合理，也不能等水库建成以后再说），这就需要构造与原型相似的模型（例如在实验室内模拟原始的地球环境；按某种比例"缩小"大型水库的结构），再通过模型来间接地对原型进行实验和研究。

由于科学发展水平的限制，早期的“模拟”主要是建立在模型与原型的（物理、化学、生理等）基本结构和运动过程相似基础上的“物理模拟”。这种模拟方法虽然也为科学技术的发展立下了卓越的功勋，但是有时却是相当困难或要付出昂贵代价的。例如，对于飞机的设计，过去常采用风洞实验的办法，即把按比例缩小的飞机模型置于风洞实验室中，用人造的强大气流吹向飞机来模拟飞机在空中飞行的过程，以取得各种有关的数据。这样做需要消耗大量的人力、物力和财力，仅实验功率就要达几十万千瓦，而且最多只能模拟超音速一、二倍的飞机。对于超过音速20倍的远程导弹，风洞实验就无能为力了。

随着科学的不断进步，人们逐渐突破了“物理模拟”的局限而跨入了“数学模拟”的新天地。“数学模拟”着眼于模型与原型在数学形式和数量关系上的相似，而不管其物理、化学结构和运动过程如何。例如，一个由机械零件组成的产生机械振荡的机械系统，和一个由电子元件组成的产生电磁振荡的电子系统，虽然其物理结构和运动过程都不相同，但是描述机械振荡与描述电磁振荡的微分方程却是完全一致的，这就使得我们可以用易于制作和便于控制的电子系统来模拟相应的机械系统，通过电磁振荡实验来研究机械振荡的规律。

近几十年以来，由于电子计算机的出现和飞速发展，使得“数学模拟”这一方法又获得了更强大的活力。由于电子计算机具有高速运算、大量存贮、自动处理的突出特点，已经在许多“模拟实验”中逐步取代了“模型”的角色。

例如，现代高速飞行器的设计，已经不再象过去那样依赖于风洞实验或实际实验，而可以采用计算机模拟实验。美国宇

航科学家在用计算机模拟宇宙飞船运行的实验过程中，出乎意料地发现飞船返回大气层时烧蚀最严重的是尾部，而不是人们通常认为的头部，后来的实际实验证实了计算机的结论。根据这一结果，人们改进了宇宙飞船的设计。

不仅如此，现在连对飞行驾驶员的训练也可以通过计算机模拟来完成。飞行员用不着上天，就可以掌握高超的飞行驾驶技术（只在训练的最后阶段，用少量时间上天试飞一下）。计算机为飞行员在地面的飞行模拟器中模拟出十分逼真的在万里长空飞翔的环境和感觉，教练员可以根据计算机输出的飞行员的各种反映及时给予指导。计算机模拟的飞行训练，不仅缩短了训练周期，保证了训练安全，而且还大大提高了训练水平。

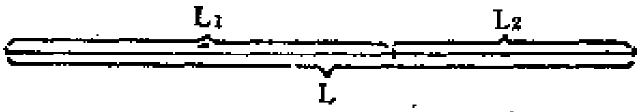
又如，法国自然科学家布丰(Buffon)早在1777年就已经提出了用投针实验来确定圆周率π值的方法：在一张白纸上画满等距离间隔的平行线，把一根长度为平行线间距离一半的小针随机地抛到白纸上；如果在m次抛针中有n次与平行线相交，则当实验次数m足够大时，m与n的比值$\frac{m}{n}$就是π的近似值。从理论上说，用这种统计实验的办法不但可以计算π值，而且还可以计算几何图形的面积，求解线性方程组和微分方程等等。但是由于需要做大量的实验，在很长一个时期内根本无法付诸实施，只好将其束之高阁。高速电子计算机出现以后，模拟统计实验的数值计算方法——现在称为"蒙特卡洛方法"，正方兴未艾，数学、物理、化学、生物等需要通过大量的统计实验才能解决的问题，都可以交给电子计算机通过模拟用"蒙特卡洛方法"解决。一个十分成功的例子是，过去用传统的实验方法来寻找合成维生素B_{12}的方案，工作量要超过1000人年，现在用

一部极其普通的微电子计算机，只用6分钟就发现了6种不同的合成方案。由于计算机模拟的出现，很多复杂的化学反应根本无需实验，就可以通过计算机了解反应的全过程及其结果，使化学从实验科学走向了计算科学的阶段。

第 十三 章

奇妙的黄金分割

黄金分割就是把一条长L的线段分为一长L_1与一短L_2两部分，并使L_1为L与L_2的比例中项（即$L:L_1=L_1:L_2$）。容易算出，



L_1与L之比为

$$\frac{L_1}{L}=\frac{\sqrt{5}-1}{2}=0.618\cdots$$

人们把这一比值及其倒数

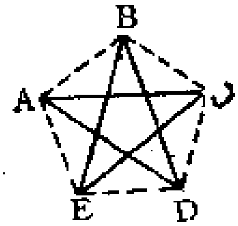
$$\frac{L}{L_1}=\frac{\sqrt{5}+1}{2}=1.618\cdots$$

统称为黄金比（又称中外比或黄金数）。

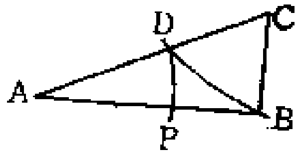
黄金分割与黄金比冠以的“黄金”二字，说明了它们在艺术和科学上的重要价值。十七世纪著名的德国天文学家开普勒曾把黄金分割与勾股定理并列，誉为古希腊几何学的两颗明珠。

黄金分割最早由古希腊毕达哥拉斯学派所发现。毕达哥拉斯学派是古希腊数学家与哲学家毕达哥拉斯所创建的秘密学术团体，正五角星是该学派的标志，他们所最先找到的正五角星形的尺规作图法，就是以黄金分割为基础的。正五角星相邻两

个顶点的距离（如 AB）与其边长（如AC）之比，或说正五边形的边长与其对角线之比，恰好就是黄金比。



以下是古希腊学者欧多克斯(Eudoxus)所给出的黄金分割的尺规作图法：



（1） 过已知线段 AB 的端点B，作 $BC\perp AB$，且使 $BC=\frac{1}{2}AB$；

（2） 连接AC，以C为圆心，CB为半径作圆弧交 AC 于 D；

（3） 以A为圆心，AD为半径作圆弧交AB于P，则P即线段AB的黄金分割点（$\frac{AP}{AB}$是黄金比）。

有兴趣的读者可以自行证明。

文艺复兴时期，有好几位颇具几何学修养的艺术大师，如丢勒(Dürer)、达·芬奇(da Vinci)等人，把几何学上对图形的定量分析，应用于一般的绘画艺术，给绘画艺术建立起了科学的理论基础。在这一过程中，他们发现了黄金分割与人们的审美观点之间的联系，例如：把一条线段分为长短不匀的两部分时，黄金分割是一种最为和谐的分法；而当一个矩形的长与宽之比为黄金比时，则是优美的矩形；绘画的表现主题如果置于画面的黄金分割处，就更能吸引观赏者的注意等。丢勒与达·芬奇等人把黄金分割与黄金比成功地用于绘画艺术，创造出了许多千古不朽的世界名画。

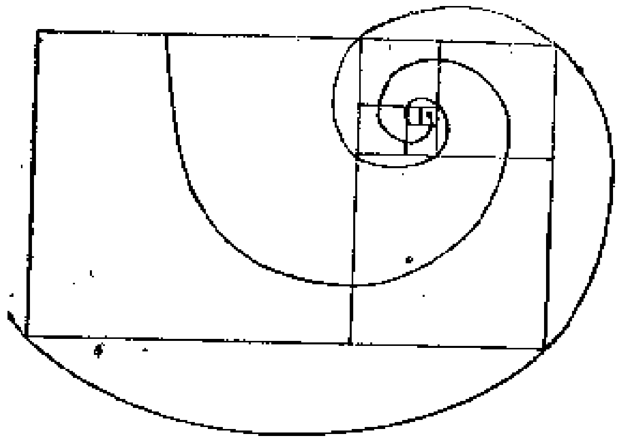
事实上，古代的建筑大师和雕塑家们早就巧妙地利用黄金比创造出了雄伟壮观的建筑杰作和令人倾倒的艺术珍品：世界七大奇迹之一的胡夫金字塔（公元前3000年）——原高度与底部边长之比约为1:1.6，庄严肃穆的雅典巴特农神殿（公元前五世纪）——正面高度与宽度之比约为1:1.6，风姿妩媚的爱神“维纳斯”和健美潇洒的太阳神阿波罗的塑象——下肢与身高之比也都近乎1:1.6。

随着时间的推移，黄金分割与黄金比越来越显示出其绚丽多彩的美学价值。建筑、雕塑、绘画、音乐、舞台艺术、工艺装饰、家俱摆设、服装款式等，只要是与人的形象审美有关的领域，几乎无处不见黄金分割与黄金比的踪迹。

然而更令人惊异的是，大自然似乎比人类更懂得黄金分割与黄金比的奥秘。例如：很多植物的叶片是按空间螺旋线自下

而上的顺序逐个萌出的，每相邻两张叶片中线的夹角都是约137°28′，正好是将360°的圆周角按黄金分割法分割的结果。这样一来，每片叶子都不会刚好长在另一片叶子上方，而总会为下方的叶子留出一点空隙，保证了所有的叶子都能得到尽可能多的阳光、雨露和空气。

在长与宽的比为黄金比的“黄金矩形”中分割出一个正方形以后，余下的部分还是一个“黄金矩形”；在余下的“黄金矩形”中又分割出一个正方形，余下的部分仍然是“黄金矩形”；如此可以不断地继续分割下去。如果把这些相继分割出

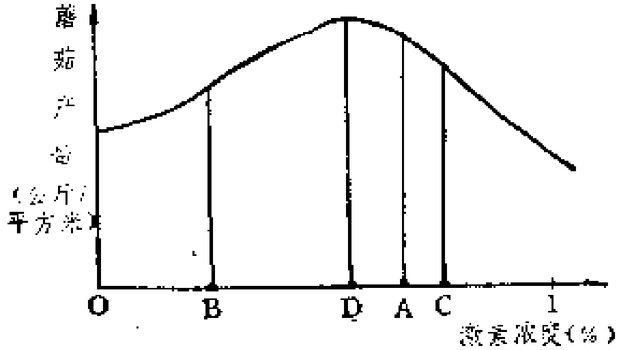


来的正方形的中心用光滑的曲线连接起来，就形成了一个可以称为“黄金螺线”的美丽曲线；而如果把这些一个套一个的“黄金矩形”的顶点顺次用光滑曲线连接起来，又形成了另一个称为“对数螺线”的美丽曲线。海螺等甲壳动物身上的螺纹、向日葵等盘状的植物花朵、松果与菠萝等植物的果实，都具有明显的黄金螺线的特征，甚至连宇宙中星系的分布也常呈黄金螺线状。

当初，人们之所以珍视黄金分割与黄金比，主要在于其美学上的价值。岂料，在现代应用数学的最优化理论中，居然也有黄金分割与黄金比大显身手的用武之地。

在科学实验和生产实践中，人们常常要通过试验对各种有关的条件进行选择，以便达到省时、省料、省钱的目的。为了尽可能少地试验，而又能尽快选择出达到人们预期结果的最佳条件，人们创造了许多强有力的数学方法，黄金分割优选法（又称0.618法）就是其中较为广泛应用的一种方法。

例如，某种可以提高人工栽培蘑菇产量的植物生长激素，如果喷施浓度过低，则增产效果并不明显；如果喷施浓度过高，又会抑制蘑菇生长，反而引起减产；只有当喷施浓度恰到



好处时，才能达到最高的增产效果。激素浓度与蘑菇产量之间的关系可以用如图所示的单峰曲线来表示。如果已知浓度超过1%时，蘑菇就要减产，那么我们就应在0～1%的范围内对最佳浓度进行优选。

按黄金分割法，第一个试验点A应取在试验范围0～1%这段的0.618处(即按0.618%的浓度喷施)；第二个试验点B则取在与第一个试验点A关于试验范围中点对称的0.382处(即按0.382%的浓度喷施)。假如A点增产效果比B点好，则由单峰曲线的特点可知，最佳点必在B～1这段范围内，因此保留从B到1这段而把从0到B那段去掉（否则，就保留0～A而去掉A～1）。第三个试验点C又取在B～1这段范围的0.618处（即按$(1+0.382)\times 0.618\%\approx 0.842\%$的浓度喷施），若$A$比$C$好，则去掉$C$～1而保留$B$～$C$；若C比$A$好，则去掉$B$～$A$而保留$A$～1。如此继续下去，每次都保留“好”点所在的一段、去掉“坏”点所在的一段，新的试验点都取在所保留的那段范围的0.618处。由于最佳点始终在保留段上，而保留段的范围又越来越小，就可以使我们经过若干次试验后，至少找到一个与最佳点十分接近的较好的喷施浓度。

为什么要按照黄金分割法来取试验点呢？

为了简便起见，仍然假定试验范围是0～1。在这段范围选取一点x_1做第一次试验，再选取一点x_2做第二次试验。x_1与x_2哪一点较好预先是不知道的，因此，比较试验结果的好坏而去掉$0\sim x_2$一段或去掉$x_1\sim 1$一段的可能性相同，这就要求这两段长度一样，即

$$x_2-0=1-x_1 \tag{1}$$

由此知，x_1与x_2两点应关于0～1的中点对称。

此外，如果x_2比x_1的试验结果好，去掉了$x_1\sim 1$一段而保留了$0\sim x_1$一段，则x_2在留下的范围内所处的位置应与x_1原来在0～1范围内的位置相仿，即所占比例相同

$$x_2:x_1=x_1:1 \tag{2}$$

由（2）得

$$x_1^2=x_2 \tag{3}$$

把（3）代入（1）得关于x_1的一元二次方程

$$x_1^2+x_1-1=0$$

解此方程得

$$x_1=\frac{-1\pm\sqrt{5}}{2}$$

舍掉没有实际意义的负根，就有

$$x_1=\frac{-1+\sqrt{5}}{2}$$

0.618不过是x_1精确到小数后第三位的一个近似值而已。

可见，第一个试验点确应安排在整个试验范围的黄金分割处。以后各个试验点的选取原则与第一、二个试验点并无本质区别（仅仅是试验范围不断缩小而已），所以每次所取的试验点都应象第一、二个试验点那样按黄金分割法确定。

黄金分割优选法中包含着数学上一个十分重要的解题方法——"逐步逼近法"，即：首先根据问题的条件确定问题的解存在的大致范围，然后一步步地缩小解的存在范围，直至最后求得问题的解答。

黄金分割（0.618）优选法只适合试验范围由连续点所组成的情形(如上例中的区间(0,1)),如果试验范围由一些离散点所组成(例如,某个关于温度条件的试验中，只有1℃、2℃、…、8℃共八个不同的温度供选择)，就需要用到与黄金比密切相关的斐波那契数列。

斐波那契数列源于13世纪意大利数学家斐波那契(Fibonacci)划时代的名著《算经》（又名《算盘书》）中的“兔子问题”：

假定一对一雌一雄的大兔，每月能生一雌一雄的一对小兔，每对小兔过两个月就能长成大兔。若年初时在兔房里放入一对小兔，并且一年内没有发生死亡，那么这对兔子一年内将繁殖成多少对？

容易算出，一年十二个月内兔子对数的变化如下：

1,1,2,3,5,8,13,21,34,55,89,144

斐波那契以后的数学家发现，若用F_n表示上述数列的第n项,则数列$\{F_n\}$满足初始条件

$$F_1=F_2=1$$

和递推关系

$$F_n=F_{n-1}+F_{n-2} \qquad (n\geqslant 3)$$

(即从第3项起，数列$\{F_n\}$中的每项都等于其前两项之和)。数列$\{F_n\}$具有许多有趣的性质和重要的应用，为了纪念斐波那契，人们将此数列称为斐波那契数列。

随着项数 n 的不断增大，斐波那契数列相邻两项的比值越来越接近黄金比：

$$\frac{1}{2}=0.5,\ \frac{2}{3}=0.66\cdots,\ \frac{3}{5}=0.6,\ \frac{5}{8}=0.625,\ \frac{8}{13}=0.615\cdots,$$

$$\frac{13}{21}=0.619\cdots,\cdots\cdots$$

事实上，我们有

$$\lim_{n\to\infty}\frac{F_n}{F_{n+1}}=\frac{\sqrt{5}-1}{2}$$

即斐波那契数列相邻两项的比值，当项数 n 无穷增大时以黄金比为极限. 又，在所有分母不大于 F_{n+1} 的分数中，以 $\frac{F_n}{F_{n+1}}$ 最接近黄金比。所以，分数 $\frac{F_n}{F_{n+1}}$ 又称为黄金比的渐近分数。

分数 $\frac{F_n}{F_{n+1}}$ 能帮助我们解决试验点为离散量的优选问题，实际上也就是用适当的分数值 $\frac{F_n}{F_{n+1}}$ 来代替黄金比 0.618……而已。

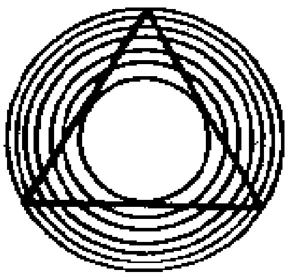
第十四章

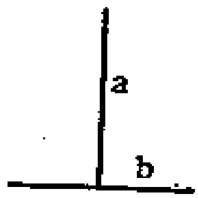
数学中的逻辑思维与非逻辑思维

数学的理论体系是演绎的，数学上的任何研究成果都必须经过逻辑推理的严格证明才能得到承认。

人们由直观和经验所得到的知识并不一定可靠。

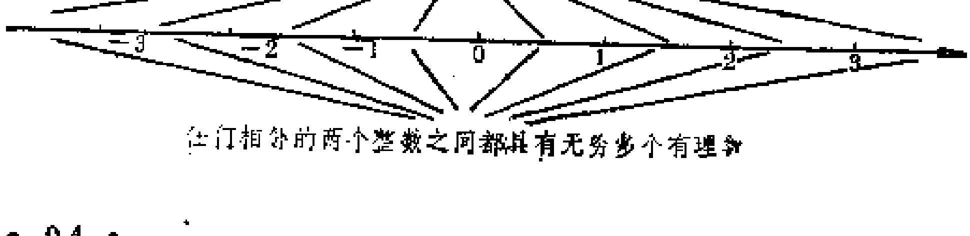
例如，从直观上看，图中三角形的三边似乎是向内弯曲





的，互相垂直的两条线段之长似乎是$a>b$，而事实上都不是。当然，只要用一把直尺检验一下，你马上就会发现这一点。不过，更多的知识却是很难、甚至无法象这样去检验的。

又如，整数点在数轴上的分布是离散的，任何相邻的两个整数之间都不再具有其它整数；而有理点的分布却是稠密的，



任何两个有理数之间都还存在无穷多个有理数（当然，任何相邻的两个整数之间也都具有无穷多个有理数）。如此看来，似乎有理点要比整数点“多”得“多”。而实际上，全体有理数却可以按一定的顺序用自然数编号，这说明有理点的“个数”其实只不过与自然数“一样多”罢了（整数点的个数也与自然数“一样多”）。

很多数学问题只靠直观和经验也是无法解决的。

例如，三角形的内角和是180°吗？直观回答不了。如果用量角器去测量，由于绘图和测量的误差，一般只能得到一些十分接近而又彼此不完全相同的结果；即使每次都可以极其精确地测量，但由于我们只能测量有限个三角形而无法测量一切三角形，同样也无法保证所有的三角形的内角和都是180°。

又如，$\sqrt{2}$是有理数还是无理数呢？不定方程$x^n+y^n=z^n$当$n>2$时是否存在整数解呢？处处连续的曲线是否处处都有切线呢？等等，这些问题也都是只凭直观和经验无法解决的。

因此，数学必须依靠演绎推理，数学的结论必须经过逻辑论证。

演绎推理是一种从一般到特殊的推理方式，即根据某个带普遍性的一般原理推求一定条件下的个别的特殊的事物所具有的性质。

逻辑论证则是通过演绎推理的方式为正确的结论寻找逻辑依据，说明其为何成立的理由。

数学上如何通过演绎推理建立理论体系，读者从本书下一章“宏伟的欧几里得大厦”可见一斑。从另一方面来看，演绎推理也是获取数学知识的一个重要途径和方法。

例如，根据平行四边形的两组对边分别平行这一前提条

件，我们可以推出平行四边形一连串的重要性质：对边相等、对角相等、邻角互补、对角线互相平分等等。

又如，数学家欧拉在研究二阶微分方程

$$y''+y=0$$

时，得到了两个不同的解

$$y=\cos x 和 y=e^{ix}+e^{-ix}.$$

当他采用无穷级数展开的办法来比较这两个解时，得到了著名的“欧拉公式”

$$e^{ix}=\cos x+i\sin x.$$

只要令其中$x=\pi$并移项，又可以得到联系数学上0、1、i、π、e五个重要常数的公式

$$e^{i\pi}+1=0.$$

法国数学家阿达玛（Hadamard）在谈到逻辑论证的作用时说：“数学严密性的目的就是肯定直觉的成果，并使之有根据。”

例如，从下列等式

$$\underbrace{1}_{1个}=1^2$$

$$\underbrace{1+3}_{2个}=4=2^2$$

$$\underbrace{1+3+5}_{3个}=9=3^2$$

$$\underbrace{1+3+5+7}_{4个}=16=4^2$$

……

我们可以直观地看出：前n个奇数的和等于n^2。为了肯定这个结论的正确，我们可以用数学归纳法证明如下：

（Ⅰ）当$n=1$时，有$1=1$，命题成立；

（Ⅱ）假设当$n=k$时，命题成立，即

$$\underbrace{1+3+\cdots\cdots+(2k-1)}_{k\text{个}}=k^2 \quad (*)$$

则在（$*$）式两边同时加上（$2k+1$），就有

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

即 $$\underbrace{1+3+\cdots+(2k-1)+(2k+1)}_{k+1\text{个}}=(k+1)^2 \quad (**)$$

这说明当$n=k+1$时，命题也成立。

综合（Ⅰ）与（Ⅱ）可知，对一切自然数n，原命题都成立。

其中的（Ⅰ）验证了$n=1$时命题成立，（Ⅱ）则在$n=k$时命题成立的前提下推出了$n=k+1$时命题也成立，如此我们便可以从$n=1$出发不断地递推下去：既然$n=1$时命题成立，则由（Ⅱ）知$n=2$时命题也成立；既然$n=2$时命题成立，又由（Ⅱ）知$n=3$时命题也成立；既然$n=3$时命题成立，再由（Ⅱ）知$n=4$时命题也成立……谁还能说对哪个自然数命题不成立吗？这就是逻辑论证的力量。

又如，我们认为$\sqrt{2}$是无理数，有什么根据呢？可以用反证法证明如下：

如果$\sqrt{2}$不是无理数，那么$\sqrt{2}$就必是有理数；而非零有理数一定可以表示成既约分数$\frac{b}{a}$（a、b是互质的整数）的形

式。所以在$\sqrt{2}$不是无理数的假设前提下，可以设$\sqrt{2}=\frac{b}{a}$（a、b是互质的整数）。

这样，就将得到$\sqrt{2}a=b$，即$2a^2=b^2$；而若一个数的平方为偶数，则该数本身也必偶数，所以b为偶数，可设$b=2q$；于是$2a^2=4q^2$，由此得$a^2=2q^2$，可见a也是偶数；a、b都是偶数显然是与a、b互质矛盾的，这个矛盾的产生来源于$\sqrt{2}$不是无理数的假设。

所以，$\sqrt{2}$不是无理数的假设不能成立，$\sqrt{2}$只能是无理数。

这样的逻辑推理过程不容反驳，令人心悦诚服。

除了数学以外，我们在日常的生活、学习、工作中也处处要用到逻辑推理。因此，一个人无论从事何种职业，逻辑推理能力总是必备的基本素质。著名的美国数学教育家波利亚(G·Polya）正确地指出："严格的证明是数学的标志，这是数学对于一般文化修养所提供的不可缺少的素养。一个学生若对数学证明从未留下印象，那他就缺少了一种基本的思维经历。"

然而，尽管数学本身是一个通过逻辑推理建立起来的演绎的理论体系，但是在数学的发现过程中却始终离不开非逻辑的创造性思维。数学的逻辑严密性往往掩盖了数学的发现过程，给人造成一种数学只是逻辑思维驰骋领域的错觉。事实上，在数学的发现过程中，非逻辑的思维（归纳、类比、直感等）常常占主导地位。著名的法国数学家拉普拉斯指出，甚至在数学里，发现真理的主要工具也是归纳和类比。波利亚也说，在数学中，发现常常从观察、类比和归纳开始。

让我们来看看，数学家是怎样发现“素数定理”的：

我们已经从“数学的砖瓦——素数”一章中知道，素数在自然数中的分布是极不规则的，虽然随着自然数x的增加，不超过x的素数的个数$\pi(x)$也在增加，我们却无法具体地把握$\pi(x)$究竟怎样随x的增加而增加。因此，人们就想寻找$\pi(x)$的一个渐近表达式，使得当x比较大时，用渐近表达式所计算出的$\pi(x)$的近似值与$\pi(x)$的实际值误差尽可能小（而当x趋于无穷大时，该误差就趋于0）。

寻找$\pi(x)$的渐近表达式的工作，开始时只能靠观察自然数表中$\pi(x)$的变化情况“拼凑”一些与$\pi(x)$的变化大致吻合的经验公式，然后再不断修改公式使之尽可能符合$\pi(x)$的变化。

观察下表

自然数 x	不超过x的素数个数$\pi(x)$	x与x的自然对数之比$x/\ln x$
10000	1229	1086
100000	9592	8686
1000000	78498	72382
10000000	664579	620417
100000000	5761455	5428613
1000000000	50847478	48254630

我们看到$\frac{x}{\ln x}$还是比较接近$\pi(x)$的，但略嫌小些。很自然地我们就会想，能不能把$\frac{x}{\ln x}$的分母缩小些，使其比值增大而更接近$\pi(x)$？沿着这样一个方向，法国数学家勒让德(Legendre)终于在1830年首先提出了$\pi(x)$的一个渐近公式（～表示等价

于）

$$\pi(x)\sim\frac{x}{\ln x-1.08366},\quad(x\to\infty).$$

后来，德国数学家高斯又以一千个相继的自然数为单位，统计每个单位中素数的个数，发现可以用$\frac{1}{\ln x}$表示充分大整数x附近的素数分布的平均密度（单位区间内所含素数的百分比）。由此，高斯提出了$\pi(x)$的另一个渐近公式（其中$\mathrm{li}x$表示$\lim\limits_{\delta\to0}\left(\int_0^{1-\delta}+\int_{1-\delta}^{x}\right)\frac{dt}{\ln t}$）

$$\pi(x)\sim \mathrm{li}x,\quad(x\to\infty).$$

可以从下表看出$\mathrm{li}x$对$\pi(x)$的渐近程度

自然数 x	不超过x的素数个数 $\pi(x)$	对数积分函数 $\mathrm{li}x$ 的值
10000	1229	1246
100000	9592	9630
1000000	78498	78628
10000000	664579	664918
100000000	5761455	5762209
1000000000	50847478	50849235

如果只考虑主阶，则勒让德与高斯的渐近公式可以统一成

$$\pi(x)\sim\frac{x}{\ln x},\quad(x\to\infty).$$

这就是素数分布理论的中心定理——素数定理。

但是，无论是勒让德还是高斯的渐近公式，都只是一种经

验归纳的结果，没有任何逻辑根据说明它们一定正确。所以，尽管勒让德和高斯发现了素数定理，素数定理在他们手上也只不过是一种直觉的推测（而并非逻辑的推导）。直到1896年，阿达玛才用复变函数论的方法证明了素数定理。

与此类似，数学家欧拉首次求出无穷级数$\sum_{k=1}^{\infty}\frac{1}{k^2}=1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots+\frac{1}{n^2}+\cdots$的和为$\frac{\pi^2}{6}$的思维过程，也是非逻辑的推导。

17世纪的瑞士数学家伯努利（Bernoulli）曾公开征求无穷级数$\sum_{k=1}^{\infty}\frac{1}{k^2}$的求和的解答。伯努利去世以后，这个难题引起了欧拉的极大兴趣。当时微积分学正处在刚刚发展起来的最初阶段，欧拉不可能象我们今天这样用无穷级数理论做为工具去求和。经过一段时间的深思熟虑，欧拉大胆地仿照多项式理论，用类比法求出了$\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$。虽然他未能给予严格的证明，但他检验了等式两边的值直到第七位数字都一致（都等于1.644934…），终于使他相信自己的结果正确。

让我们来看看欧拉发现上述结果的过程：

首先，如果a_1、a_2、…、a_n是n次代数方程$a_0+a_1x+a_2x^2+\cdots+a_nx^n=0\ (a_0\cdot a_n\neq 0)$的$n$个非零根，则多项式$\sum_{k=0}^{n}a_kx^k$可以分解因式为

$$\sum_{k=0}^{n} a_k x^k = a_0\left(1-\frac{x}{\alpha_1}\right)\left(1-\frac{x}{\alpha_2}\right)\cdots\left(1-\frac{x}{\alpha_n}\right)$$

同样，如果$2n$次只含偶次项的代数方程$b_0-b_1x^2+b_2x^4-\cdots+(-1)^n b_n x^{2n}=0(b_0\cdot b_n\neq 0)$的$n$对非零根为$\pm\beta_1$、$\pm\beta_2$、$\cdots$、$\pm\beta_n$，则有

$$\sum_{k=0}^{n} b_k x^k = b_0\left(1-\frac{x^2}{\beta_1^2}\right)\left(1-\frac{x^2}{\beta_2^2}\right)\cdots\left(1-\frac{x^2}{\beta_n^2}\right) \qquad ①$$

由多项式的恒等定理，①式两边的x的同次幂的系数相等，特别有两边x^2项的系数相等，于是

$$b_1=b_0\left(\frac{1}{\beta_1^2}+\frac{1}{\beta_2^2}+\cdots+1-\frac{1}{\beta_n^2}\right) \qquad ②$$

假如能找到一个只含偶次项的$2n$次代数方程恰以±1、±2、$\cdots\pm n$为根，则由②式立即可得$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}=\frac{b_1}{b_0}$，就可以求出有限和$\sum\limits_{k=1}^{n}\frac{1}{k^2}$。因此，容易想到，要求无限和$\sum\limits_{k=1}^{\infty}\frac{1}{k^2}$，就要设法找一个只含偶次项的方程，它以$\pm1$、$\pm2$、$\cdots$、$\pm n\cdots$等无穷多对整数为根。这样的方程的次数当然不可能再有限（因而也就不可能再是代数方程），但是不妨设想它在一定条件下仍保持着与②式所表示的性质类似的性质。

现在的困难是以±1、±2、$\cdots$、$\pm n\cdots$为根的方程不容易找，但若把它们全部分别乘上因子π，成为$\pm\pi$、$\pm2\pi$、$\cdots$、$\pm n\pi\cdots$并不妨碍我们的工作，因为只要求得出无穷和$\sum\limits_{k=1}^{\infty}\frac{1}{(k\pi)^2}$，

也就可以求得无穷和$\sum_{k=1}^{\infty}\frac{1}{k^2}$。而找以$\pm\pi$、$\pm2\pi$、…、$\pm n\pi$…为根的方程很容易，熟悉三角方程$\sin x=0$的人，当然知道它有无穷多对根0、$\pm\pi$、$\pm2\pi$、…、$\pm n\pi$…去掉它唯一为零的根可得三角方程$\frac{\sin x}{x}=0$，将此方程左边$\frac{\sin x}{x}$展开成幂级数，就找到了我们所需要的，只含偶次项的、以$\pm\pi$、$\pm2\pi$、…、$\pm n\pi$、…为根的方程

$$1-\frac{1}{3!}x^2+\frac{1}{5!}x^4-\frac{1}{7!}x^6+\cdots+(-1)^n\frac{1}{(2n-1)!}x^{2n}+\cdots=0 \quad ③$$

③式中左边第一项1即相当于②式中的b_n，第二项$-\frac{1}{3!}x^2$系数的相反数$\frac{1}{6}$即相当于②式中的b_1，仿照②式的样子令其中的n趋向无穷，就得

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

两边再乘π^2，$\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$的结果就出来了

$$\frac{\pi^2}{6}=1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}+\cdots.$$

欧拉的方法虽然是不严格的，但是却是富有启发性的。欧拉还用另外的例子试验了他的方法，并成功地用类似的方法来

出了另一个重要级数——莱布尼兹级数的和。近七年之后，欧拉给出了他的结果的严格证明。

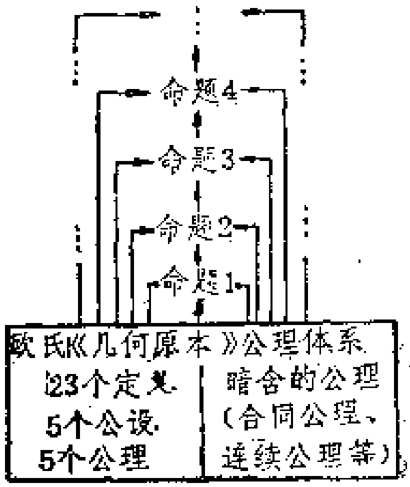
可见，在数学上，逻辑思维与非逻辑思维相辅相成，都是不可缺少的；在“数学与实验”一章中，我们将进一步探讨这个问题。

第十五章

宏伟的欧几里得大厦

几何是研究图形（平面或空间点集）性质的一门学科。我们在中学里所学的平面几何与立体几何知识，俗称"初等几何"，其主要内容来源于古希腊数学家欧几里得（Euclid，约公元前三世纪）的杰出著作《几何原本》，因此又称"欧几里得几何"。

几何学知识，最初是人们从实践经验中归纳总结出来的。"几何学"这一名词的希腊文（γεωμετρτια）原义指的就是"测地术"，据说是从古埃及传入的。当时，横贯埃及的尼罗河每年都要泛滥，冲毁地界，人们在水退之后必须重新丈量、分配土地，几何学便在这种年复一年的测量中得以萌发、成长起来。我国的情形也是类似的。早在公元前4世纪，我国的古数学书《周髀算经》中就特别提到：夏禹治服洪水的方法是使洪水按一定河道东流入海，这就需要测量地形和地势高低，从而导致了几何知识的产生。各种土木建筑工程的设计和施工、天文学和力学的研究，都极大地推动了几何学的发展。



到古希腊文明盛期，几何经验知识的积累已经相当丰富，迫切需

要系统地加以整理和提高。从公元前7世纪到公元前3世纪，许多古希腊数学家和哲学家着手进行了这一工作。欧几里得在前人工作的基础上，用逻辑演绎的方法，从精心选择的少数不加证明的命题（公设和公理）出发，建立起了一整套可以十分自然地概括当时所有几何知识的优美的理论系统，写出了《几何原本》这一流芳百世的巨著。

《几何原本》共十三卷，除了平面几何和立体几何的主要内容以外，还包括了数论、比例等内容，几乎包括了古希腊数学的全部。整个著作从23个定义开始，设置了5个公设和5个公理，以此为基础，用严格的逻辑顺序展开命题的序列，同时给予演绎的证明（做为严密逻辑推理的根据，欧氏的公设和公理事实上是不够完善的，因此在这一过程中欧氏常常不自觉地借助直观引入了一些暗含的新公理）。就这样，在《几何原本》中所展现出来的几何知识体系，成为了一座堪称集当时人类几何学知识之大成的雄伟大厦。

《几何原本》自公元前300年左右问世以来，曾以手抄本的形式广泛流传了近一千八百多年；印刷术出现以后，又被翻译成多种文字，先后出版了一千多种版本，成为人类历史上发行得最多、影响最大的一部数学书。在近两千年的时间里，它一直是标准的教科书，哺育出一代又一代的数学家。

几乎没有哪一个数学家不曾被欧氏几何迷住过。

阿基米德（Archimedes，公元前200年间杰出的数学家和物理学家，至今仍以他所创立的阿基米德公理和他所发现的浮体定律而闻名于世。）是欧几里得的学生卡农的学生，他对几何学寄予了最大的热忱。“他为这个女妖神的甜蜜引诱而陶醉，而精神恍惚。他简直可以说是与她形影不离，以至他常常

废寝忘食。”阿基米德发现了正圆柱内切球的表面积和体积恰好分别是正圆柱表面积和体积的2/3，他是如此为这个发现自豪，立刻嘱咐将来把一个有内切球的正圆柱体图案铭刻在他的墓碑上。据传，阿基米德被攻陷叙拉古斯城的罗马士兵杀死之前，仍在专心致志地研究画在沙盘上的几何图形，“别动我的图！”是他所说的最后一句话。

牛顿（Newton，17世纪以来誉满全球的数学家和物理学家，万有引力的发现者，微积分和经典力学的创始人。）年少时学习欧氏几何，不明白人们为什么要煞费苦心地去证明那些简直是显而易见的事实。但当他成长为一个科学家以后，逐渐明白了其中的奥秘，对欧氏几何称颂备至。牛顿赞叹说：“几何学的辉煌之处就在于只用很少的公理而能得到如此之多的结果。”

高斯（Gauss，18世纪被人们敬畏地称为数学王子的数学家和物理学家。）17岁时发现了正n边形可以用尺规作图的充要条件，并用尺规实际作出了正17边形，他非常激动，竟放弃了原来想当哲学家的念头而决心献身数学。高斯逝世以后，他曾经长期工作过的哥廷根为他塑了铜象，铜象的底座就是正17边形，以此来象征他的荣誉。

罗素（Russell，本世纪著名的数学家和哲学家，逻辑主义学派的鼻祖。）在他的自传中写道：“我十一岁时，开始学习欧几里得几何，并请哥哥教我，这是我一生中的一件大事，它使我象初恋一样入了迷，我当时没有想到世界上还会有这样有趣的东西。”

爱因斯坦（Eienstein，本世纪最伟大的科学家之一，划时代的相对论的创始人）说得好：欧几里得几何“这一思维

的辉煌成绩赋予人类的理智以自信，这种自信为以后的活动所必需。”“如果欧几里得未能激起你少年时代的热情，那么你就不是一个天生的科学思想家。”

为什么欧几里得几何学具有如此迷人的魅力呢？这不仅在于其内容丰富、论述流畅和推理严谨，更重要的还在于其方法论上的强大力量。欧几里得首创的“公理化”方法，在《几何原本》中，为后来的数学家和其他科学家树立了如何建立科学理论体系的光辉典范。正如著名的美国数学家波利亚(G. Polya)所指的那样，“欧几里得几何——不是一种普通的逻辑体系。它是这一体系的第一个最伟大的典范，其他学科过去与现在都试图模仿它。”

我们知道，数学上证明一个命题（例如三角形的内角和等于180°）的成立，总需要某些已经证明过的命题（例如三角形的外角等于不相邻两内角的和）做为依据，而那些做为依据的业已证明过的命题又需要另一些已经证明过的命题（例如两条平行直线与第三条直线相交所成同位角相等）做为更原始的依据，如果这样一直追根溯源，我们将永远找不到证明的最初依据，证明也就永远不能完成，因而必须在某个或某些比较显而易见的命题（例如过直线外一点可以并且只可以作一条直线与原直线平行）处停止下来，对于这些命题不加证明地承认其正确，称其为公理（欧几里得又把几何学专用的公理称为公设，因此《原本》中既有公理又有公设，实际上都是公理）。

数学上的推理过程又需要使用各种不同的概念（如三角形的内角），而定义每个概念又需要其它已经定义过的概念（如三角形和角），追根溯源，必然有些概念（如点和直线）是无法定义的，至多只能粗略地加以描述，这些概念称为初始概

念。

公理事实上是有关初始概念的各自存在性、基本性质及其相互关系的规定，通常认为应该尽可能满足三个条件：一是相容性，即从所选取的公理出发推导出的任何结论都不会彼此矛盾（绝对的相容性事实上是无法证明的，只能要求相对相容，例如在承认实数理论无矛盾的前提下证明欧氏几何无矛盾，或在承认欧氏几何无矛盾的前提下证明实数理论无矛盾）；二是完备性，即从所选取的公理出发无需进一步扩充就可以推导出所需要的全部理论体系（这样的完备性事实上是办不到的，因此现代对完备性的理解是：同一公理系统的所有模型都是同构的。即对同一公理系统而言，不存在在某一模型中成立而在另一模型中不成立的命题）；三是独立性，即所选取的每一公理都不能由其它的公理用逻辑方法推导出来，这一要求使公理的数目减少到最低限度，不允许出现多余的公理。

所谓公理化方法就是：以尽可能少的不加定义的初始概念和不加证明而直接采用的命题（公理）做为基础，利用逻辑推理法则来产生演绎的科学理论。

从现代公理化方法的观点来看，欧几里得《几何原本》中的公理系统是不够完善的。例如，它除了所列出的五条公设和五条公理以外，还暗用了合同公理、连续公理等其它一些公理，因而是不完备的；它还含有可以从其它公理推出的多余公理（公设4“凡直角皆相等”），因而也是不互相独立的。但是，就当时的历史条件而言，其严谨性却是公认的。《几何原本》的问世标志着数学领域（也是科学史上）公理化方法的诞生，欧几里得所取得的成功至今仍是令人敬佩的。

《几何原本》的不足之处早在古代就已被一些敏锐的学者

所察觉。例如，阿基米德就曾提出过阿基米德公理（任给非负数量a和b，$a<b$，存在n使$na>b$）对欧氏公理系统做必要的扩充。19世纪后半叶，由于非欧几何的出现导至几何学的革命，促使很多数学家，如康托尔（Contor）、戴德金(Dedekind)、巴士（M. Pasch）、皮亚诺（Peano）、庇爱里（Pieri）、希耳伯特（Hilbert）等人深入地研究了几何基础问题，使得欧氏几何公理系统日臻完善。其中希尔伯特的工作最为出色，他于1899年出版的《几何基础》一书，最终解决了欧氏几何的欠缺问题，进一步发展了几何学的公理化方法，成为近代公理化思想方法的开端。

数学公理化的深刻意义主要表现在如下三个方面：

（1） 公理化方法具有系统地归纳总结数学知识的作用，把各命题按逻辑演绎的关系连接起来，学习和使用起来方便、可靠。

（2） 公理化方法清楚地揭示出数学理论的基础，有利于比较各种数学分支本质的异同，并能促进和推动新数学理论的创立和发展。

（3） 公理化方法在科学方法论上也占有举足轻重的地位，是从理论上探索事物发展的逻辑规律，做出新的预见和发现的重要方法之一。

欧氏几何学中的公理化方法，不仅为后世的数学家们推崇和仿效，而且还影响到其他科学（特别是物理学）领域。牛顿的经典力学、克劳修斯（Clausius）的热力学、爱因斯坦的相对论等都不同程度地使用了公理化方法。对于现代自然科学，公理化已经成了科学数学化的一个重要特征，它明确地揭示出建立该门学科所必需的基本概念与原始假设，从而在阐明新概

念与证明结论的过程中排除模棱两可、含混不清的表达，以及直观的、想当然的推理因素，把科学观点奠定在切实可靠的逻辑基础之上。

最后需要指出的是，早在欧几里得前一百多年，我国战国时期一位著名的学者墨翟（墨家学派的创始人）与他的弟子合著的《墨经》一书，就已经比较系统地整理记录了当时的几何学、力学、光学、逻辑学等方面的知识。其中有关几何学知识的内容有二十来条，其定义的确切、立论的精辟，均不亚于《几何原本》。特别难得可贵的是，《墨经》中还叙述了一条《几何原本》中所没有的、后来才为阿基米德所补充的阿基米德公理（确切地说，是与阿基米德公理等价的“或不容尺公理”）。可惜的是，春秋战国时期曾经兴旺一时的墨家学派，由于其政治主张违背当时统治阶级的利益，又受到儒家的排挤，秦汉以后逐渐灭绝，《墨经》一书也几乎长期无人问津。其后中国数学的发展，并未受到《墨经》思想的影响，而是走上了一条与公理化方法截然相反的道路。直到清末，由于西方科学文化、技术的大量引进，已经相当落后的中国数学的发展才汇入了世界数学发展的潮流。

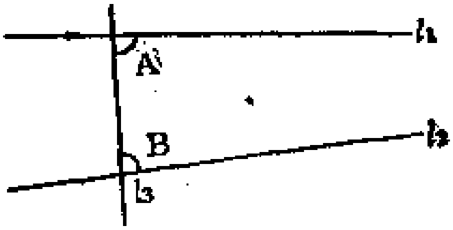
第十六章

离经叛道的非欧几何

欧几里得在《几何原本》中所建立起来的欧氏几何学，虽然被人们誉为雄伟壮丽的大厦，但也不是毫无瑕疵的。例如，欧氏对“线”的定义是“线有长度没有宽度”，其中使用了未加任何说明的“长度”、“宽度”等概念，使定义本身含糊不清；而且，这个定义除对“线”这一几何形象进行了不甚严格的直观描述以外，并未起到任何逻辑作用，欧氏在后来的讨论中再也没有使用这个定义。对于使用得很多的“介于”这一重要概念，欧氏不恰当地将其列为不加任何说明的初始概念，而当证明中需要用到这一概念时，就不得不利用图形的直观。欧氏在证明两个三角形全等时，采用了把一个三角形重叠到另一个三角形上去的手法，这必须假定在重叠过程中图形的长度和角度都保持不变，可是《原本》中却没有相应的合同公理。欧氏的尺规作图，多处用到了两段圆弧在一定条件下相交，交点的存在以圆的连续性为前提，《原本》中也没有关于连续性的公理。

数学家们早就注意到了《几何原本》的美中不足，历代都有不少人试图把欧氏几何的公理系统加以改造，以消除这一雄伟大厦基础上的隐患。在这一过程中，第五公设的试证占有极其重要的地位，由此而发现了非欧几何，引起了人们几何学观点的根本改变。

第五公设　若两直线和第三直线相交，且在同一侧所构成的两个同侧内角之和小于两直角，则把此二直线向这一侧适当延长之后一定相交。



现代的欧氏平面几何教科书中，已经使用了一条与之等价的公理取代了它：过已知直线外一点，可以并且只可以作一条直线与已知直线平行。由于这条公理与平行线有关，所以现代又常称欧氏第五公设为欧氏平行公理。

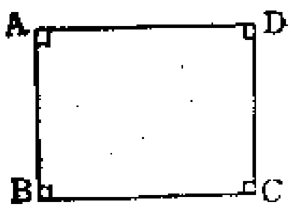
由欧氏第五公设可以得出一个极为重要的结论（也是与第五公设等价的命题）：三角形三个内角的和等于180°。这是欧氏几何最显著的基本特征之一。

数学家们对欧氏第五公设特别感兴趣，主要是因为：首先，第五公设与欧氏的其它公理、公设相比显得冗长、复杂，缺乏其它公理、公设那样的显而易见性；其次，欧氏在使用第五公设时非常小心谨慎，尽量避免使用，直到非用不可的时候（第29命题）才用了一次，而且以后绝不再用，可见欧氏本人对他的第五公设似乎也有所不满和怀疑。然而，第五公设又是不能缺少的，否则欧氏几何中将有一系列命题失去逻辑基础。于是，数学家们就考虑能不能用其它的公理和公设把第五公设证明出来，换句话说，也就是试图取消第五公设的公理地位，把它由公理变为定理；另一条途径，则是考虑能不能用更为简单明了的命题来代替第五公设，在这方面并未取得什么积极的成果，所以我们只着眼于前者，即第五公设的试证。

19世纪以前，很多大数学家都曾涉足第五公设的试证问题，结果都未能达到原来预期的目的。往往只得到与第五公设

等价的某个命题，承认该命题就能证明第五公设，实际上只相当于将第五公设改头换面，并未真正给予证明。也有人想到用反证法，假定第五公设或者它的某个等价命题（如三角形内角和为180°）不成立，试图由此推出矛盾，结果除了导致一系列与人们的现实经验不相符合的奇怪结论以外，并未产生任何逻辑矛盾。这使某些比较敏锐的数学家开始意识到：几何学不一定只有欧几里得所建立起来的那一种，从任何一组不会导致矛盾的前提出发，都有可能产生出一种新几何学。

最早明确发表这一观点的，是18世纪的瑞士数学家兰伯特（Lambert）。他对第五公设的研究，是从三个角A、B、C都是直角的四边形$ABCD$开始的。第四个角D有三种可能，或是锐角、或是直角、或是钝角。可以证明直角的假设与第五公设等价，而钝角的假设则导致逻辑矛盾，因此只需再否定锐角的假设就使第五公设得证。于是，他在锐角的假设下展开了讨论，引出了一个与通常的直观相悖但并无逻辑矛盾的几何系统。在他之前，另一位意大利数学家萨克切里（Saccheri）也曾通过类似的手法得到过类似的结果，却认为太不合情理而主观否定了锐角的假设，自以为就此证明了第五公设。兰伯特则不然，他没有轻易否定锐角的假设。他注意到钝角的假设虽然导致矛盾，但却给出了在球面图形上成立的定理，因此他猜想锐角假设给出的定理可能适用于虚半径球面上的图形（多么大胆而又天才的想法！）并由此预见到不同于欧氏几何的新几何学的存在。



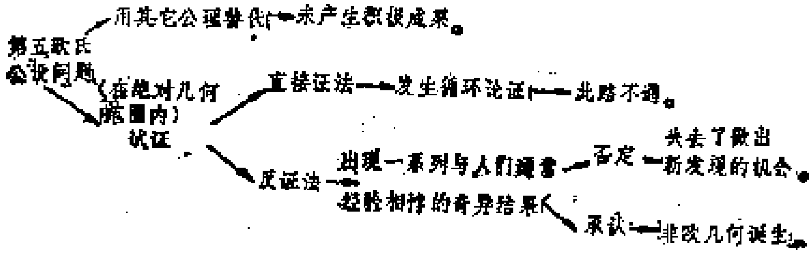
新几何学的雏形终于首先在德国大数学家高斯手中完成了。他在1824年的一封信中透露：他已经发现，如果假定三角

形的内角和小于180°，就会导致一种与传统几何学根本不同的独特的几何学。他能在这个新几何学中解决任何问题，只是某个常数必须预先确定。高斯起初把这种新几何学称为反欧几何，以后又改称星空几何，最后称为非欧几何，历史上便第一次出现了"非欧几何"这一名词。高斯深知，在一般人眼里非欧几何无异于离经叛道的革命。由于"害怕引起愚人的喊声"，高斯终生没有公开发表他对非欧几何的研究成果。甚至，当他的匈牙利挚友伏尔刚·鲍耶（Wolfgang Farkas Bolyai，一个为研究第五公设耗尽一生的数学家）之子约翰·鲍耶（Johaun Bolyai），继他之后独立地创建非欧几何，他也没有给予应有的支持。

与此同时，俄国数学家罗巴契夫斯基（Н. И. Лобачовский）也从研究欧氏第五公设发现了非欧几何。他引进了一条与欧氏第五公设截然相反的平行公理："过已知直线外一点至少可以作两条直线与已知直线不相交"，同时保留了欧氏几何的其它公理。从这样一个公理系统出发，他推出了一个与欧氏几何同样严谨、同样具有内在根据的新几何学，现代称为罗氏几何学。罗氏几何中有很多从习惯于欧氏几何的人看来十分荒谬的奇谈怪论，例如：三角形的内角和小于180°；同一直线的垂线与斜线不一定相交；不存在相似形；过不共直线的三点不一定可以作圆；等等。因此，正如高斯所担心的那样，罗氏几何不但没有被当时绝大多数人所理解和承认，反而受到了嘲笑、讽刺和谩骂。更有甚者，俄国总主教竟宣布它是异端邪说。直到罗巴契夫斯基逝世以后，高斯生前的通讯录出版，其中有高斯私下对罗巴契夫斯基的高度赞扬，罗氏几何才得到了人们的普遍承认和重新评价。1826年2月11日，罗巴契夫斯基

发表新几何学报告的这一天，也就被认为是非欧几何的诞生日。

现在，让我们用图式把从欧氏第五公设开始到非欧几何产生的过程，扼要地概括一下：

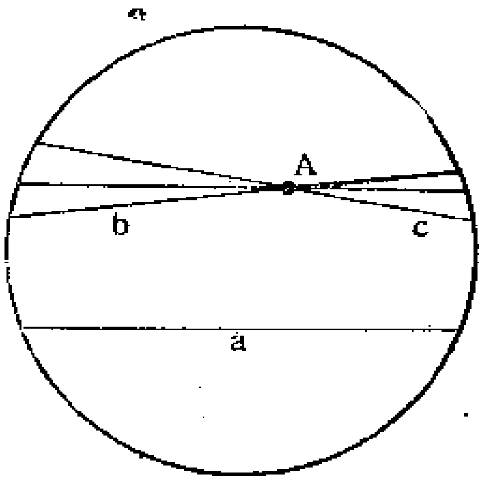


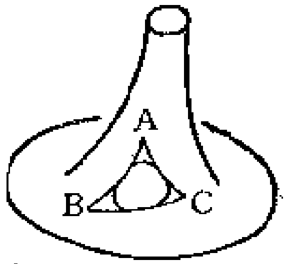
罗氏几何与欧氏几何的根本区别，就在于二者的平行公理不同。通常我们认为，过直线外一点只可以引一条直线与原直线平行，仅在有限的平面范围内与我们的直观相吻合，当范围扩大到超出我们的经验之外，无法直接观察与验证的时候，是否仍然如此就不一定了。在这一点上，欧几里得的处理是非常聪明的，他把平行公理拐弯抹角地叙述成第五公设的那种说法，实际上是恰好回避了无限远空间的事情，体现了一个数学大师的高瞻远瞩和巧妙手法。

现在我们知道，客观物质世界是极其丰富的，物质的空间形式也多种多样，因此，没有理由认为欧氏空间就是物质空间形式的唯一模型。

让我们设想圆内有一弦a及弦外一点A，过A而不与a相交的弦显然至少有两条b与c（事实上有无穷多条）。当圆的半径无穷增大时，可以把这样的圆看做是一个平面；此时，弦a及过A而又不与a相交的弦b、c也无穷增大，可以把它们看做是直线。这样，过直线a外一点A，就至少有两条直线b、c与a平

行了。如此，我们便构造出了一个罗氏几何模型。



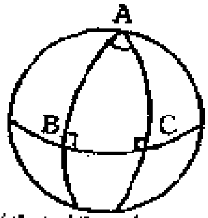


$\angle A+\angle B+\angle C<180°$

有一种形状接近高音喇叭的负曲率曲面，叫做伪球面。1868年，意大利数学家柏尔特拉米（Beltrami）证明了伪球面的内蕴几何是与罗氏几何相一致的（或者说，伪球面可以做为罗氏几何的一个现实模型），这说明罗氏几何确实反映了客观物质世界空间形式的某些性质和规律。接着，德国数学家克莱因（F. Klein）和法国数学家庞加莱（Poincaré）又在欧氏平面上构造出了罗氏几何模型，将罗氏几何的无矛盾性归结为欧氏几何的无矛盾性。人们既然承认欧氏几何学，当然也就要承认罗氏几何学了。这种将罗氏几何的无矛盾性归结为欧氏几何的无矛盾性的想法也十分巧妙，欧、罗两种几何体系同真、同假的结果令人情不自禁地拍案叫绝！

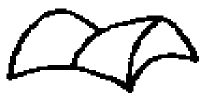
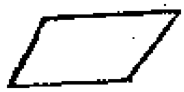
罗氏几何的创立一方面解决了欧氏第五公设的独立性问题，另一方面，也是更重要的一方面，它开阔了人们的视野，加深了人们对几何本身意义的认识。从此，几何学的发展便开始了繁荣的新时期，同时对物理学中时空观念的彻底改变起了巨大的推动作用。

1854年，德国数学家黎曼（Riemann）提出了另一种非欧几何——黎曼几何。在黎曼几何中，有关直线的公理是：(1)凡直线都相交；(2)直线不能无限延长。黎曼几何中的三角形内角和大于180°．黎曼几何模型类似于球面。本世纪初，爱因斯坦所创立的广义相对论，就是以黎曼几何做为其时空的数学基础。



$\angle A+\angle B+\angle C>180°$

1872年，德国数学家克莱因在射影几何的基础上给出了欧氏几何、罗氏几何和黎曼几何的新解释。克莱因把欧氏几何称为“抛物几何”，因为欧氏直线有一个无穷远点，犹如抛物线有一个无穷远点；把罗氏几何称为“双曲几何”，因为罗氏直线有两个无穷远点，犹如双曲线有两个无穷远点；把黎曼几何称为“椭圆几何”，因为黎曼直线没有无穷远点，犹如椭圆没

几何名称	空间类型	远距离状况	模型简图	三角形内角和
罗氏几何	负曲率	无穷伸展	 类似马鞍面	<180°
欧氏几何	零曲率	无穷伸展	 类似平面	=180°
黎曼几何	正曲率	自行封闭	 类似球面	>180°

有无穷远点。三种几何空间的区别在于“曲率”的不同：罗氏空间曲率为负；欧氏空间曲率为零；黎曼空间曲率为正。三种几何学的简要比较如表所示。（见118页）

第十七章

从公理化到结构主义

非欧几何的建立向人们表明，公理对于数学是何等重要：欧氏几何与罗氏几何的天壤之别，其根源仅仅在于一条平行公理的不同！这就大大提高了公理方法的声誉。同时，以欧氏公理系统为代表的老的公理化方法的缺陷也逐渐暴露无遗，特别是老公理化方法没有很好考虑公理系统的独立性、无矛盾性和完备性，难以为数学提供坚实可靠的逻辑基础。

非欧几何诞生以后，许多数学家参加了重建几何公理系统的工作。1899年，德国大数学家希耳伯特的经典著作《几何基础》出版，宣告了重建欧氏几何公理系统的工作已经基本完成，演绎科学公理化的新时期正在到来。

希耳伯特的公理化方法与欧氏公理化方法的根本区别在于：

1．希氏没有对几何上的"点"、"线"、"面"等初始概念做任何直观的描述，而是通过公理来定义概念。不管什么样的数学对象，只要它们能够满足相应的公理，都可以叫做"点"、"线"、"面"。正如希氏本人所形象地指出的那样："我们可以用桌子、椅子、啤酒杯来代替点、线、面。"意思是，不必理会点、线、面这些概念所反映的具体事物究竟是什么。正因为如此，希氏的几何公理系统可以赋予多种多样的具体解释，因而它也就可以成为多种多样几何空间的理论基础，特别是当把它

解释成普通欧氏几何中的对象时，我们就得到了普通欧氏几何的全部理论。

2. 希氏证明了其公理系统的无矛盾性。这一点十分重要，因为如果从同一公理系统推出了互相矛盾的结果，那么这一公理系统也就失去了存在的意义。严格地说，希氏的证明是不彻底的，他只把几何公理系统的无矛盾性归结到了实数系统的无矛盾性。也就是说，他只证明了其几何公理系统的相对相容性。

3. 希氏考虑了其公理系统的独立性，其公理系统中每个公理都不能由其它公理推证出来（粗浅地说，其公理系统中没有任何多余的公理）。希氏证明独立性的方法是建造模型，使得除了要证独立性的公理之外，其余公理都成立，而要证独立性公理的否定也成立。

4. 希氏考虑了其公理系统的完备性，确保从其公理系统可以推出全部的几何命题。不过，现在一般认为，对公理系统完备性的要求应视需要而定。公理系统所包含的公理越少，适用的范围就越广；反之，公理系统所包含的公理越多，适用的范围就越窄，而推导出的理论也就越丰富。例如，取消希氏几何系统中的平行公理，我们就得到绝对几何，它普遍适用于欧氏几何与罗氏几何；但若添加希氏（实际也是欧氏）平行公理，则其公理系统就只能适合于欧氏几何。

以希耳伯特的几何公理化系统为楷模，数学界掀起了一股公理化的热潮，许多数学分支都在本世纪初先后逐步公理化，奠定了比较坚实可靠的逻辑基础。以希耳伯特为首的一大批数学家甚至雄心勃勃试图将整个数学公理化。

本世纪二十年代，是希耳伯特的公理化思想占统治地位的

时期，整个数学的基本理论都已经归结到了公理化的自然数算术和实数理论，而自然数算术和实数理论又进一步归结到了公理化的逻辑与集合论。数学界非常乐观，以为只要能证明这些公理系统的相容性和完备性(独立性相对来说不那么重要)，便可大功告成！然而可惜的是，奥地利数学家哥德尔 (Gödel) 1930年所发现的不完全性定理，却使希耳伯特全盘公理化的计划转眼成了泡影。

哥德尔不完全定理说：在任何包含初等数论的相容的形式系统中，存在着不可判定的命题（即命题本身及其否定在该系统中都不可证)。这一方面意味着包含初等数论的形式系统是不完备的（因为存在着无法证明的真命题；请注意任何命题及其否定二者之中必有一真)；另一方面又意味着包含初等数论的形式系统的绝对相容性是无法证明的（因为存在着不可判断真假的命题)。这对希耳伯特的理想真可谓是致命的打击！尽管如此，希耳伯特在数学公理化方面的贡献仍然是伟大的，他在这一研究过程所形成的卓越的思想方法至今仍在影响着全世界研究数学基础的“元数学”家们。

本世纪30年代，在希耳伯特公理化思想方法的影响下，法国出现了一个以布尔巴基为笔名的青年数学家集团。他们试图用公理化方法从整体上对数学的各个分支的基础进行深入细致的整理、比较和重建，在这一过程中，他们逐步产生了“结构”的观念。

布尔巴基认为，全部数学（或至少是大部分数学）都可以按照内在结构的不同而加以分类。最基本的数学结构有三类：代数结构、序结构和拓扑结构，称为母结构。每类母结构又有很多子结构，如代数结构有群、环、体、线性空间等等；序结

构有半序、全序、良序等等；拓扑结构有距离、紧致、连通等等。这些子结构彼此之间可以交叉、又形成各式各样的分支结构。正象几十种原子构成丰富多彩的物质世界一样，以上所谈的各种结构也组成了千姿百态的数学世界。数学的发展，无非是各种结构的建立和发展。如果用公理化方法抽象出各个数学分支的基本结构，找出它们的联系和差别，就能获得各个数学分支内在联系的清晰图景。

为了理解布尔巴基的“结构”观念，让我们试对大家最熟悉的数学对象——全体实数 R（即数直线或一维欧氏空间）的结构稍做一番剖析：

Ⅰ. R中定义了“+”、“·”两种运算，满足

（Ⅰ.1） $x+(y+z)=(x+y)+z$（加法结合律）

（Ⅰ.2） $x+y=y+x$（加法交换律）

（Ⅰ.3） 存在$0\in R$，对任意的$x\in R$，都有$0+x=x$（加法单位元）

（Ⅰ.4） 对任意的 $x\in R$，存在$-x\in R$，使得 $x+(-x)=0$（加法逆元）

（Ⅰ.5） $x(yz)=(xy)z$（乘法结合律）

（Ⅰ.6） $xy=yx$（乘法交换律）

（Ⅰ.7） 存在$1\in R$，对任意的$x\in R$，都有 $1\cdot x=x$（乘法单位元）

（Ⅰ.8） 对任意的$x\in R$且$x\neq 0$，存在$x^{-1}\in R$，使得$x\cdot x^{-1}=1$（乘法逆元）

（Ⅰ.9） $x(y+z)=xy+xz$（乘对加的分配律）

（Ⅰ.1）到（Ⅰ.4）表明 R对加法成一交换群，（Ⅰ.5）到（Ⅰ.8）表明 R对乘法也成一交换群；而对加、乘两种运算 R成一交

换环；全部九条公理表明R恰好构成一个域。这些都是所谓代数结构。

Ⅱ. R中定义了关系"$\leqslant$"，满足

(Ⅱ.1) $x\leqslant y$且$y\leqslant z\Longrightarrow x\leqslant z$（传递性）

(Ⅱ.2) $x\leqslant y$且$y\leqslant x\Longrightarrow x=y$（对称性）

(Ⅱ.3) 对任何x、y，或$x\leqslant y$，或$y\leqslant x$（可比性）

(Ⅱ.4) $x\leqslant y\Longrightarrow x+y\leqslant y+z$（加法保序性）

(Ⅱ.5) $x\leqslant y$且$0\leqslant z\Longrightarrow x\cdot z\leqslant y\cdot z$（乘法保序性）

(Ⅱ.1)到（Ⅱ.5）反映了R的全序结构，(Ⅱ.4)与(Ⅱ.5)表明代数结构和序结构之间具有协调性。如此便形成了新的结构——有序域。

Ⅲ. 对任意的x、$y\in R$且$x>0$，$y\geqslant 0$，存在自然数n，使得$y\leqslant nx$。此即阿基米德公理，它保证了R的连续性结构。

Ⅳ. 对R中的元素可以定义绝对值

$$|a|=\begin{cases} a & （当a\geqslant 0时）\\ -a & （当a<0时）\end{cases}$$

从而引出距离的概念和邻域的概念。此即拓扑结构。同时还可以进一步由此得出极限的概念和基本序列的概念，并可以证明凡基本序列的极限值也都含于R内，也就是说R的拓扑结构具有完备性。可以在R上展开全部数学分析的理论。

以上Ⅰ、Ⅱ、Ⅲ、Ⅳ中的代数结构、序结构和拓扑结构的有机结合便形成了数直线R的数学结构——完备的阿基米德有序域。这样，我们对于数直线R就有了一个十分清楚的本质的认识。

容易看出：代数结构是从数量关系的四则运算、序结构是

从数量的大小关系中抽象概括出来的；而拓扑结构则是距离、邻域等概念的自然延伸。因此，可以认为"结构"是现实世界中的数量关系与空间形式的进一步抽象和发展。

布尔巴基认为，数学主要是研究结构的，只要结构确定了，采用什么形式是无关紧要的。因此，抽象地看，我们可以把数学结构完全相同的两个数学系统不加任何区别，称它们"同构"，并认为它们在本质上是一致的。

例如对于复数，在代数中可以看做是形如 $a+bi$ （a、$b\in R$，$i^2=-1$）的数，或者形如$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$的实二阶反对称方阵；在几何中可以看做是平面上坐标为（a，b）的点，或者是以原点为起点，以$\sqrt{a^2+b^2}$为模、$\operatorname{arc\,tg}\dfrac{b}{a}$为幅角的位置向量。无论其形式如何，数学结构总还是一样的，我们可以根据具体问题的性质决定采用哪种形式。

从同构的观点来看，结构上相同的数学对象可以互相转化，这种转化并不改变这些数学对象的实质，然而却对我们研究问题的难易程度有很大的影响。一个比较复杂而又难于求解的数学问题，经过同构变换，可能会变得十分简单明了，非常便于处理。

例如，笛卡尔的解析几何，建立起了有序实数对的代数系统与欧氏几何之间的同构对应关系。通过解析几何，我们可以用代数方法来研究几何问题，也可以用几何方法来研究代数问题。在第七章（神奇的对数与RMI方法）中，我们曾谈过，数学家们是通过解析几何，用代数方法彻底解决了古希腊时代遗留下来的、在两千多年漫长时间里令人感到束手无策的三大尺

规作图问题。如果不是把三大尺规作图问题转化成了相应的代数问题，恐怕人们至今仍然无能为力。

数学家们把欧氏几何的无矛盾性化归为实数系统的无矛盾性，从而证明了欧氏几何与实数系统相对相容，思想方法也是类似的。以下是这一工作的若干片断：

首先建立基本概念的解释字典，即规定几何中的点、直线、角等等对应于实数系统中的什么：

基本概念解释字典

基本概念	实数解释
点 p	有序实数偶 (x, y)
直线 l	有序实数比 $(a:b:c)$，其中 $a^2+b^2\neq 0$
p 在 l 上	$ax+by+c=0$，其中 $a^2+b^2\neq 0$
⋮	……

然后，按照解释字典把几何公理逐条翻译成相应的实数命题，并说明该命题成立的理由：

几何公理	翻译与证明	
……	……	
结合公理 $I_{3.1}$ 每条直线上至少有两点	已知实数比 $(a:b:c)$，$a^2+b^2\neq 0$ 则方程 $ax+by+c=0$ 至少有两组解	理由：事实上，此乃不定方程，有无穷多解
结合公理 $I_{3.2}$ 至少有三点不在一直线上	至少有三对数偶 (x_i, y_i) 使方程组 $ax_i+by_i+c=0 (i=1,2,3)$ 无解 $(a:b:c)$，$a^2+b^2\neq 0$	理由：事实上，三对数偶 $(0,0)$，$(1,0)$ 与 $(0,1)$ 即符合条件
……	……	

数学结构的观念，使我们的眼界扩大了许多。例如，函数

不再限于实变数的实函数和复变数的复函数，一般集合上也可以定义函数、函数空间上也可以定义函数（即所谓“泛函”）；点、线、面的概念也不再局限于普通的一、二、三维空间，而可以推广到高维、甚至无穷维空间。布尔巴基之所以能在代数数论、代数几何、代数拓扑、泛函分析和李群等众多的领域做出引人注目的成就正是得力于他们“结构”的观念。

“结构”也是一种非常有效的工具。我们一旦从所研究的客体中认识到某种满足已知系统公理的性质，立刻可以随心所欲地使用相应结构的“工具库”里的任何一件“库存”：概念、方法或定理，可以节省大量的思维劳动。

诚然，布尔巴基用“结构”对整个数学基础进行条理化、系统化，与各门数学分支的公理化相比，这是更高一步、更深一层的抽象和概括，也是一个新的飞跃；但是，物极必反，过分强调结构与形式，就会扼杀数学的创造精神和蓬勃发展的活力。事实上，尽管布尔巴基的结构主义对数学做出了巨大的贡献，但是布尔巴基本身却已由盛到衰。这是因为，现代数学的发展已经渗透到几乎所有的科学领域，而在这一过程中数学所遇到的实际问题往往与计算有关，常常要借助计算机才能解决，使得数学的发展由布尔巴基时代的抽象的、结构主义的道路转向了具体的、构造主义的、结合实际的、结合计算机的道路。

第十八章

从常量到变量的飞跃——微积分的建立

17世纪以前，数学的研究对象基本上还停留在常量的范畴(即今天我们所谓的初等数学)，固定数量之间的运算、固定图形之间的关系是其主要内容。当时的自然科学和生产技术所提出的数学问题，常量数学大都可以解决。

17世纪前后，正是欧洲封建主义日趋没落，新兴资本主义急剧发展的时期。由于生产力的不断提高、科学技术的不断进步，航海、天文、力学、军事、机械都向数学提出了大量迫切需要解决的涉及变量的数学问题。例如：如何确定航行中的船舶在大海中的准确位置；如何掌握行星运行的规律；如何分析物体的受力情况及变速运动情况；如何精确计算炮弹运行的轨迹；如何设计更有效的实用机械。这些问题归结到数学上主要有四类：

1. 求变速运动物体路程的时间函数、瞬时速度和加速度；

2. 求曲线切线的斜率和方程；

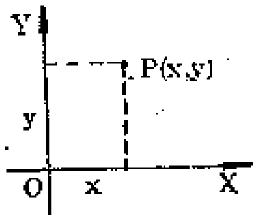
3. 求曲线长度、曲边梯形面积；

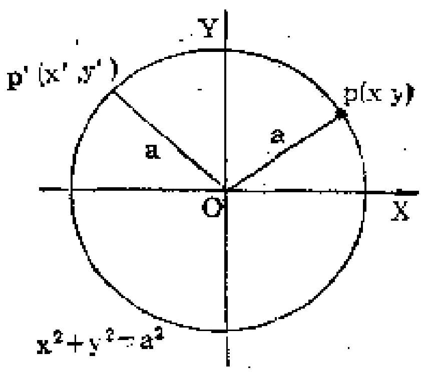
4. 求函数的最大值与最小值。

这四类问题的共同特征都要求把“变量”做为研究对象。对“变量”的分析和研究，促使了微积分的产生和发展。

变量思想的萌芽可以一直追溯到古代。例如，古希腊哲学家德谟克利特（Democritus）求面积和体积的“数学原子法”；欧多克斯（Eudoxus）和阿基米德求圆面积、抛物线拱形面积的“穷竭法”；我国春秋战国时期庄子的“一尺之棰，日取其半，万世不竭”，以及魏晋时期刘徽求圆周率的“割圆术”；都包含有朴素的变量和极限的思想。但是，那时的常量数学已经足够解决大部分数学问题，变量数学的研究仅是零星地偶尔为之，未能形成系统的方法和理论。

17世纪中叶，法国数学家笛卡尔和费马在研究上述四种类型的变量问题中创立了解析几何。他们通过引进点的坐标，将几何图形与代数方程联系了起来，把代数方程不仅看做是联系已知量与未知量的等式，而且还看做是反映动点的变化坐标相互依赖的关系式。笛卡尔指出，如果方程的个数少于未知数的个数（如$x^2+y^2=a^2$，一个方程两个未知数），这样的方程是不定方程（即由所给的方程不能唯一确定其中未知量的数值）；如果把不定方程中的未知数看做是代表点的坐标的不断变化的量，则它就代表了一条由相应的动点所描绘的曲线（如$x^2+y^2=a^2$所描出的曲线就是一个半径为a的圆）。



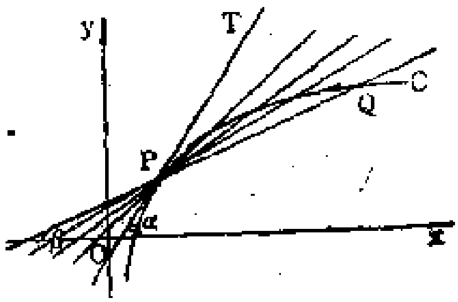


笛卡尔和费马把以往对立的“数”与“形”统一了起来，

在数学中明确引进了变量的思想。恩格斯在评价这一伟大功绩时指出："数学中的转折点是笛卡尔的变数。有了变数，运动进入了数学，有了变数，辩证法进入了数学，有了变数，微分和积分也就立刻成为必要了，而它们也就立刻产生。"

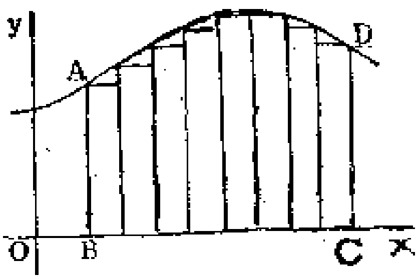
微积分的主要理论基础是极限论。什么是极限呢？粗浅地说，如果一个变量在其变化过程中越来越接近、乃至无限接近于一个常量，我们就把这个常量称为该变量在这个变化过程中的极限。极限的思想方法是：为了确定某个用常量无法或难以确定的数量，我们可以先设法去确定该数量的某些容易把握的近似值，并且通过一系列这样的越来越精确的近似值去逐步逼近该数量本身；从而不但可以最终求出该数量的准确值，而且还剖析出了运动的过程。

例如，曲线C上有邻近的两点P、Q，PT是曲线C在P点的切线，PQ是曲线C过P点的割线。切线PT与水平方向的夹角α、割线PQ与水平方向的夹角β，其大小反映了PT、PQ二直线的倾斜程度，我们把这两个夹角的正切$tg\alpha$、$tg\beta$分别称为PT、PQ的斜率，并分别用K_{PT}、K_{PQ}表示。当P点固定而Q点沿曲线向P点运动时，割线PQ的斜率K_{PQ}（由于倾角β的不断变化）是一个不断变化的变量，而切线PT的斜率K_{PT}（由于倾角α保持不变）则是一个保持不变的常量。在Q点向P点运动的过程中，割线PQ越来越接近切线PT的位置，变量



K_{PQ}也就越来越接常量K_{PT}（要多么接近就有多么接近），K_{PQ}成为一系列越来越准确的K_{PT}的近似值。K_{PT}就叫做K_{PQ}在Q

沿曲线向 P 运动这一过程中的极限，并记作 $\lim\limits_{P\to Q} K_{PQ}=K_{PT}$。换句话说，切线 PT 的斜率 K_{PT}，可以通过割线 PQ 的斜率 K_{PQ} 来逐步逼近，通过求变化着的割线 PQ 的斜率 K_{PQ} 的极限，我们就可以求出用常量难以确定的切线 PT 的斜率 K_{PT}。这一思想的进一步发展，就产生出了"导数"的概念和"微分"的运算。



又如，我们把如图由三条线段 AB、BC、CD 和曲线段 AD 所围成的封闭图形 $ABCD$ 称为曲边梯形。为了求这个曲边梯形的面积 S，我们可以把它分割成许多小的曲边梯形，对每个小曲边梯形，其面积都可以用与其十分接近的内接矩形的面积近似地代替，设所有这些内接矩形的面积之和为 σ，则整个曲面梯形的面积 $S\approx\sigma$。S 当然是个常量，而 σ 却是个随我们对曲边梯形 $ABCD$ 分割方法不同而改变的变量，小曲边梯形分割得越多越细，其内接矩形面积之和 σ 就越来越接近常量 S（要多么接近就有多么接近），σ 成为一系列越来越准确的 S 的近似值。S 就叫做 σ 在曲边梯形 $ABCD$ 的分割无限变细（用 $\lambda\to 0$ 表示）的过程中的极限，记作 $\lim\limits_{\lambda\to 0}\sigma=S$。换句话说，曲边梯形 $ABCD$ 的面积 S，可以通过分割得越来越多、越来越细的内接小矩形面积之和 σ 来逐步逼近，通过求变化着的内接小矩形面积之和 σ 的极限，我们就可以求出用常量难以确定的曲边梯形的面积 S。这一思想的进一步发展，就产生出了"定积分"的概念和"积分"的运算。

现在公认的微积分的创始人是德国的莱布尼兹(Léibniz)

和英国的牛顿。他们首先在代数概念的基础上建立起了比较系统的微积分方法，并意识到了微分与积分互为逆运算。在他们手中，微积分成了分析和研究各种运动和变化、以及变量之间相互依赖关系的数学工具。形象地说，微积分把复杂的运动变化过程分解成一段段较为简单的相对静止状态，犹如一部描述运动的电影用“慢镜头”放映出来，使人们能够清楚地看到一个连续的运动是如何由许多微小的“瞬时”变化合成的，从而通过“瞬时”变化规律把握住连续运动的规律。

牛顿称他的微积分方法为“流数术”，并成功地将这一方法用到解决大量的几何问题与力学问题当中去。特别值得一提的是，牛顿用微积分方法为他所发现的著名的万有引力定律提供了数学依据，并使开普勒的行星运动三大定律成了万有引力定律在太阳系中作用的自然推论。从此，微积分方法及牛顿本人声名大振。但是，牛顿的“流数术”是经验的、具体的和谨慎的，带有物理观点。莱布尼兹的微积分方法则是富于理性、比较抽象和大胆的、带有哲学观点。莱布尼兹更关心用运算公式创造出广泛意义下的微积分，并精心为微积分选择了诸如$\frac{dy}{dx}$（导数或微商）及$\int$（积分）这样流传至今的符号。由于莱布尼兹的微积分比牛顿的“流数术”更简明易懂，便于普及，因而在欧洲大陆的传播上占了优势。

在牛顿、莱布尼兹的时代，刚刚诞生的微积分的理论基础还很薄弱，甚至连“极限”、“无穷小”、“连续”等最基本的主要概念都含混不清，因而遭到了以英国大主教贝克莱(Berkeley)为首的、害怕微积分这一新生事物动摇宗教神权的一批人的反

对和攻击，酿成了数学史上的第二次危机（详见第十九章“数学的三次危机”）。这促使数学家们深入地探讨和研究了微积分的基础，进一步巩固和发展了微积分理论。到19世纪上半叶，经过欧拉、伯努利、泰勒、柯西、波尔察诺、高斯、维尔斯特拉斯、阿贝尔等好几辈数学家们的共同努力，微积分（数学分析）理论终于可以说是比较严密和基本完备了。

微积分创立以后，立刻显示出了强大的生命力，在数学和其他科学中得到了极其广泛的应用。数学的面貌从此焕然一新，进入了变量数学时期，并由此逐步形成了现代数学的主要分支，如微分方程、无穷级数、变分法、微分几何、复变函数等等。恩格斯在上一世纪评价说：“在一切理论成就中，未必再有什么象17世纪下半叶微积分的发现那样被看作人类精神的最高胜利了。”就在我们今天看来，微积分虽然已经可以算做比较古老的数学分支，可是仍然是富有生命力的。它仍然是解决实际问题的强有力的数学工具之一，也仍然是孕育新的数学分支的源泉之一。微积分所包含的深刻的数学思想，永远值得我们每个人很好地去领会。

第十九章

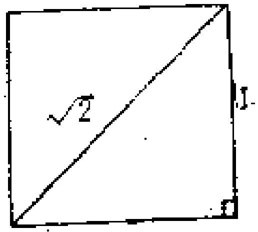
数学的三次危机

数学中充满了矛盾，比如从加与减、乘与除、正与负、有理数与无理数、实数与虚数、微分与积分，一直到有限与无限、连续与离散、存在与构造、逻辑与直观等。在数学发展的过程中，始终都贯穿着矛盾的斗争与解决。当矛盾激化到涉及整个数学的基础，非解决不可的时候，就产生了数学危机，数学史上已经出现了三次较大的危机。

公元前500年左右，正是古希腊毕达哥拉斯学派兴旺发达的时期。毕达哥拉斯学派深信：数学知识可以由纯粹思维获得，是绝对可靠、准确，用于四海而皆准的。他们认为，宇宙中的一切现象和规律都可以归结为整数与整数之比，这相当于用离散的观点来看待世界中的数量关系。当时的数学还处在刚刚从自然数中脱胎出来而形成有理数概念的阶段，数学所处理的数量关系还没有超出有理数的范围，毕达哥拉斯学派的观点符合当时人们的认识水平，代表了当时人们的普遍见解。

勾股定理（直角三角形斜边的平方等于两直角边的平方和）是毕达哥拉斯学派在数学上的一项重大发现。恰恰是这个毕达哥拉斯学派引以为自豪的发现，无情地推翻了他们自己奉若神明的信条，——勾股定理揭示了某些直角三角形的斜边（例如两直角边长为1的等腰直角三角形的斜边，或说单位正方形的对角线，其长为$\sqrt{2}$）就不能用整数或整数之比来表示

(这表明线段是连续的而非离散的)。传说，这一事实是毕达哥拉斯学派的门徒希帕索斯(Hippsuss)公元前400年左右发现的。希帕索斯的这一发现是如此令毕达哥拉斯学派惊恐不安，以至希帕索斯竟被活活投入海中淹死！



希帕索斯的伟大发现向人们表明，绝不能只由整数与整数之比来一统天下，几何上还存在着不能由整数与整数之比来表示的数量。毕达哥拉斯学派的离散的数量观点，在连续的现实世界的数量面前碰了壁，人们不得不考虑如何才能解决这一矛盾。这就是历史上第一次数学危机。它极大地冲击了古希腊传统的数学观念，使古希腊数学走上一条重演绎、重逻辑推理的道路。

由于整数与整数之比不能表示所有的几何量，而反过来，任何整数与整数之比却都能由几何量给予表示，这就使几何开始在古希腊数学中逐渐取代算术占据了统治地位，逻辑的推理和论证也就立刻显得十分必要了。古希腊人在解决第一次数学危机时，并没有采取承认无理数的地位，把它们当做普通的数来处理的直接办法，而是采取了用几何上的量与量之比避开无理数的间接办法(欧洲直到17世纪初才开始承认无理数)。这种小心谨慎似乎有点回避矛盾的做法，反倒促使了古典逻辑与欧氏几何的孕育和产生，引起了数学上的一场深刻变革，数学被奠基到了古典逻辑和演绎公理系统的基础之上。

第二次数学危机是由于18世纪初微积分的诞生而引起的。当时微积分发展十分迅速，成功地解决了许多问题，广泛地渗透到了各个领域。但是，其理论基础却是含混不清的无穷小。

请看微积分的创始人之一——牛顿当时是怎样用无穷小来求 $y=x^2$ 的导数的：

先取无穷小量 Δx，则

$\Delta y=(x+\Delta x)^2-x^2=2x\Delta x+\Delta x^2$

两边除以 Δx 即得

$$\frac{\Delta y}{\Delta x}=2x+\Delta x$$

因为 Δx 是无穷小量可以忽略不计

所以 $y'=2x$。

在这里，无穷小量 Δx 既不是0（可以用 Δx 做除数）、而又是0（最后 Δx 忽略不计，消失了），简直有点象变魔术！数学家们可以根据需要随心所欲地把它们看做异于零进行各种运算，也可以把它们看做是零而忽略掉。召之即来，呼之即去，总带有某种主观主义的神秘色彩。因此，以无穷小做为基础的微积分遭到了某些人的诽谤和攻击，导致了一场激烈的争论。

攻击和争论的高潮是由英国大主教贝克莱挑起的。原因是微积分的创始人牛顿的好友哈雷，曾劝说贝克莱的病中朋友拒绝宗教祈祷，使贝克莱感到了在微积分帮助下的自然科学的发展对于宗教信仰日益增长的威胁。贝克莱气得大骂微积分“招摇撞骗，把人们引入歧途”，嘲讽无穷小是“消失了量的鬼魂”，挖苦微积分方法是“分明的诡辩”。许多数学家对此进行了反驳。但是，由于微积分本身的理论基础确实并非无可指责，致使来自各方面的非难似乎言之有理。虽然贝克莱之流的嚎叫并没有怎么影响数学家们大胆地运用微积分不断扩展新的领域，然而大本营的内患却始终有些令人担心。

这种状况是数学家们不能长期容忍的，很多数学家深入地

探讨了微积分的理论基础。法国数学家柯西首先甩开了直观的无穷小，给出了包括极限、连续、导数、微分、积分以及无穷级数等微积分基本概念的比较严格的定义，并初步提出了极限的 ε-δ 描述方法，使极限概念算术化；而无穷小则定义成极限为0（即绝对值越来越小、无限趋于0）的变量。德国数学家维尔斯特拉斯进一步改进了柯西的理论和方法，同时确定了一致收敛的概念。到19世纪70年代，戴德金和康托尔等人各自独立地建立起了实数理论，并以此重新整理了极限理论，才使微积分有了一个比较完整、比较严密的基础。（需要指出的是：在牛顿、欧拉的时代，无穷小是一种具有特殊性质的“数”，虽然有些神秘，但使用起来直观、方便；到了柯西、维尔斯特拉斯的时代，无穷小被剥掉了神秘的外衣，成为一种以 0 为极限的变量，虽然比较严密，但使用起来繁琐、艰难。本世纪60年代，数理逻辑专家罗宾逊（A. Robinson）证明了在扩大的实数系R^*中，无穷小可以做为一种数而存在，这就为牛顿的无穷小提供了逻辑基础。微积分的理论基础又有了新发展，是令人高兴的事！）

第二次数学危机总算勉强平息下去了。法国大数学家庞加莱在1900年巴黎国际数学家大会上高兴地宣称：“今天，我们可以说，绝对的严格已经取得了。”

但是，一波刚平，一波又起。微积分理论以实数理论做为基础，实数理论又以集合论做为基础，而做为基础的基础的集合论，偏偏在刚诞生不久，就出现了暴露其内部矛盾的悖论，从而构成了更大的危机——第三次数学危机。从某种意义上说，第三次数学危机乃是前两次危机的延续和深化，所涉及的问题更为严重、范围更为广泛，影响更为深刻。

集合论中最著名的悖论是由英国数学家罗素1902年提出的，只涉及集合论的最基本概念——元素和属于（$\in$）。大家知道，当我们把某些事物做为一个整体来看待时，这些事物就构成一个集合S，组成集合的事物x称为集合的元素，并用$x \in S$来表示。可以把集合分成两大类，第一类是集合本身也可以看做是该集合的元素那样的集合（例如由图书馆构成的集合M仍是图书馆，即$M \in M$）；第二类是集合本身不能看做是该集合的元素那样的集合（例如由自然数构成的集合N不再是自然数，即$N \overline{\in} N$）。现在考虑不以其自身做为其元素（即$X \overline{\in} X$）的第二类集合X，显然，所有第二类集合的全体也构成一个集合A，试问是否$A \in A$呢？如果$A \in A$，则由A的构成（A中全是第二类集合）知应有$A \overline{\in} A$；而如果$A \overline{\in} A$，同样由A的构成，又应有$A \in A$；既有$A \overline{\in} A$，又有$A \in A$，这是个矛盾。

罗素后来又用一个比较浅显的"理发师悖论"来形象地说明这一矛盾：一个乡村理发师，宣称他要给所有那些不给自己刮脸的人刮脸，而绝不给那些给自己刮脸的人刮脸。但是，有一天发生了疑问：他该不该给自己刮脸呢？如果他给自己刮脸，按照他声明的后半部分，他就不能给他自己刮脸；而如果他不给自己刮脸，按照他声明的前半部分，他又得给自己刮脸。无论如何，理发师总使自己左右为难。

罗素悖论犹如一颗炸弹，动摇了整个数学的基础，数理逻辑学的先驱弗雷格（Frege）在他的《论数学基础》卷二的书后写道："对一个科学家来说，没有一件事如下列事实更令人扫兴，当他的工作刚完成的时候，突然一块奠基石崩塌下来了。当本书的印刷快要完成时，罗素先生给我的一封信就使我陷入这样的境地。"当时最伟大的数学家之一希耳伯特也忧心

忡忡地说："必须承认，由于悖论的出现而造成的形势是难以忍受的。只要设想一下，每个人都曾学过、教过，并在数学中加以应用的定义和演绎的方法，一直都被认为是真理和必然的典范，而现在却导致了荒谬。如果连数学思维都不可靠，那么到哪里还能找到真理和必然性呢？"

既然集合论出现了矛盾，人们当然可以抛弃集合论，而把数学建立在其它更为严密可靠的理论之上。但是，经过一段时间的探索，数学家们发现别的理论更不好建立和运用，远不如集合论方便有力。于是很多人致力于集合论的改造，开始了集合论公理的研究。1908年，策莫罗(Zermelo)首先创立了集合论的一个公理系统，后来经费兰克尔（A.Fraenkel）改进，形成了今天著名的ZF系统。同时，罗素也发表了他的一个集合论公理系统——类型论。以后，冯·诺依曼（von Neumann）、贝尔奈斯（P.Bernays）、哥德尔等人也相继建立了其它类型的公理系统。在研究这些公理系统的过程中，数学家们不仅为整个数学奠定了比较坚实的基础，而且也取得了极为丰富的成果，推动了数学基础理论的进一步发展和完善。现在，尽管数学大厦的基础依然存在一丝裂缝，但是总的说来，第三次数学危机已经趋于平缓。

每次数学危机的产生，都促使数学家们认真地检查数学的基础，创造出各式各样解决危机的方案和方法。矛盾的消除、危机的解决，给数学理论带来了变革和飞跃。沉舟侧畔千帆过，病树前头万木春；数学本来就是在解决各种问题（生产实践和其它科学对它所提出的问题、以及它本身的问题）的斗争中不断发展壮大的，任何危机也不能阻挡数学大踏步地前进！

第二十章

从三分天下到趋于统一

第三次数学危机震撼了整个数学界，使得很多卓越的数学家深入地考虑了数学基础的问题。由于哲学观点不同，数学家们对数学究竟应该奠基于什么的看法也有很大不同。于是在本世纪初，数学上相继产生了逻辑主义、直觉主义和形式主义三大流派。三大流派各持己见，争论不休，互不相让，形成了三分天下的局面。

逻辑主义学派的代表人物是英国数学家兼哲学家罗素(Russel)和怀特海德(Whitehead)。逻辑主义学派的主要宗旨是把数学划归为逻辑，正如罗素所指出的那样："应当从一些普遍承认是属于逻辑的前提出发，再经过演绎而达到那些明显地属于数学的结果。"在逻辑主义学派看来，数学不过是"由命题p推出命题q"这种演绎的总和，全部数学都可以从基本的逻辑概念和逻辑规则推导出来，只要不使用"集合的集合"这类逻辑语言，悖论就不会发生。这样一来，数学就成为逻辑学的分支了。

罗素和怀特海德合著的《数学原理》是逻辑主义学派的代表著作。在这一巨著中，罗素和怀特海德企图实现他们将数学划归逻辑的理想，但是问题和困难程度远远超出他们的想象。要想由逻辑推出数学，第一步是推出"数"，他们虽然成功地用"类"来定义"1"，但过程极为繁琐、费力，用了两三百页的

篇幅。以致著名数学家庞加来（Poincare）讥讽说："这是一个可钦可佩的定义，它献给那些从来不知道1的人。"其他的诸如$n\times m=m\times n$等的证明也都十分复杂冗长。尽管如此努力，他们也未能彻底地将数学划归为逻辑。他们过于夸张了数学与逻辑在演绎结构上的同一性，而抹杀了数学与逻辑的质的差异。

逻辑主义遭到了很多数学家的批评，主要还不是其教条式的繁琐和不能自圆其说，而是因为如果数学只是一套逻辑演绎系统，就不再具有创造性，不能产生新的观念和新的方法，也不会给出任何新的知识，最终总会走到山穷水尽的地步。当然也应该看到，逻辑主义学派的工作并非一无是处。相反地，他们仍然对数学与逻辑的发展做出了不可磨灭的贡献：逻辑主义者把古典数学纳入了一个统一的公理系统，成为数学公理化方法近代发展过程中的一个重要里程碑。虽然这个公理系统不是纯逻辑的，但是它揭示了数学与逻辑之间的关系，对于计算机的人工智能研究具有很大的理论价值；同时，由于逻辑主义的努力，基本上完成了从传统逻辑到数理逻辑的过渡和演变，开创了数理逻辑的新纪元。

直觉主义的代表人物是荷兰数学家布劳威尔（Brow-wer）。直觉主义者的基本出发点不是集合论，而是自然数论，即"数学开始于自然数及自然数相等的概念形成之后"。自然数来源于人们对于时间顺序的"原始直觉"，因而是可信的。其他的数学概念和数学事实必须是在自然数的基础之上，按照某种方式经过有限步骤构造出来，才是可信的。

例如，直觉主义者认为，复系数一元二次方程$ax^2+bx+c=0$（$a\neq 0$）在复数范围内有两个根，是可信的，因为这两个

根可以通过求根公式

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

被具体地构造出来；但是高斯关于“复系数一元n次方程$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0$ $(a_n\neq0)$在复数范围内至少有一个根”的论断，是不可信的，因为高斯的证明是“存在性”的（亦即假如没有一个这样的根存在的话，便会产生矛盾），并未指出用何种方法可以在有限步内确定出其中的一个根。

又如，直觉主义者认为，对有限个自然数而言，可以下结论说：“每个自然数要么是偶数，要么是奇数”；但是对全体自然数而言，则不能下这样的结论。因为对有限个自然数，能够给出一个切实可行的过程在有限步之内把它们一一检查完毕；而对全体自然数来说，就无法做到这一点，只能检查到哪里才算到哪里。至于下一步检查结果如何，只能等做出来以后才知道，检查之前不能随便下任何结论。由于自然数是无穷的，我们永远也不可能完成这一检查工作。

直觉主义学派与逻辑主义学派正好相反，非但不把逻辑看成是数学的基础，反而把逻辑视为数学的一部分。直觉主义学派发展了自己的逻辑，是与传统逻辑很不相同的逻辑，特别是在关于否定性质的推理（如排中律、反证律、量词的解释）上与传统逻辑有很大不同。

我国著名的数理逻辑专家莫绍揆指出：“直觉主义因反对古典逻辑，从而需要把整个逻辑及数学全盘改造，连人们日常认为最简单、最明白无讹的部分也需要重新审查，这显然是一件非常艰巨的工程。再由于直觉主义逻辑强调能行性，反而变

得啰嗦、不方便起来，从而这个数学改造极慢，几乎可以肯定难以成功。”

直觉主义学派对数学的限制过严，抛弃的合理因素过多，只承认一部分最保险的数学，尤其是他们对本来在很多方面行之有效的非构造性数学和传统逻辑的排斥，更使得人们难以接受，因而遭到了大多数数学家的反对。然而，直觉主义者强调构造性的观点还是正确的，有穷性的观点也有其合理的一面，例如在使用电子计算机时就不能不考虑这两点。

形式主义学派的创始人是德国数学家希耳伯特，但他与法国的布尔巴基这个狂热、彻底的形式主义作者完全不同。希耳伯特及其追随者的基本主张是：古典数学应形式化地奠定，即抛开数学对象与形式无关的内容和解释，使数学能从一组符号化的公理出发，构成一个纯形式的演绎系统。他们一方面希望保存古典数学的基本概念和经典逻辑的推理原则，特别是那些与实无限性有关的概念和方法（诸如无穷集合的概念和排中律在无限论域上的使用等等）；另一方面，与直觉主义同样地认为，可信性只存在于有限之中，无限性概念不过是理性的规定而已。为了实现这一主张，他们提出了如下的形式公理化方案：

（1） 证明古典数学的每一分支都可以公理化；

（2） 证明每个这样的公理系统都是完备的（即任意一个系统内命题的真假均可以在该系统内得到判定）；

（3） 证明每个这样的公理系统都是协调的（即公理系统是无矛盾的，不会出现一个命题及其否定能同时成立的现象）；

（4） 证明每个这样的公理系统所对应的数学模型都是同

构的；

（5） 寻找这样一种方法，借助于它，可以在有限步之内判定任何一个命题的可证明性。

正当希耳伯特雄心勃勃地企图实现这一计划的时候，奥地利数学家哥德尔（Gödel）1 931年发现了不完全性定理：在包含初等数论的一致的形式系统中，存在一个不可判定真假的命题，该命题或其否定都无法成为该系统内的定理。换句话说：任何包含初等数论的一致的形式系统都是不完备的。这无疑给形式主义学派当头一瓢冷水，打破了希耳伯特的美梦。

希耳伯特的形式公理计划虽未全部实现，但他开创的“元数学”（又称“证明论”）却不断成长壮大，今天已经成为把命题证明方法做为研究对象的一门基础数学分支。正如波兰数学家塔斯基（Tarski）所说：“称希耳伯特为元数学之父，他是当之无愧的，他创始了作为独立学科的元数学；他为它的存在而奋斗，以一个伟大数学家的威信全力支持它；他策划出它的前进途径并寄托以重任。诚然，婴儿未能实现慈父的全部希望，并未成长为一个神童，但是它健康地发展了，已经成为数学大家庭的正式成员。”

逻辑主义、直觉主义和形式主义三大流派的基本观点互相对立，分歧很大，各有所偏、各有所见。他们对整理和重建数学系统的实际做法，互有相反相成之处，他们对于数学的发展都起了程度不同的积极影响和推动作用。现在研究数学基础的数学家们，对三大流派的观点和方法取长补短，互相渗透和融合，由三大流派三分天下的局面已经不复存在，数学家们对数学基础的态度正在逐步趋于统一。

第二十一章

"来路不正"的概率论

什么是"概率"？可以粗浅地把"概"理解为"大概"或"可能"，"率"则是两个相关数的比值；所谓"概率"就是用一个比值来表示某个事件出现（发生）的可能性大小的一种度量。

自然界中所发生的各种现象以及现象中的各种事件，按其发生的可能性来分，无非有以下三种：

一种是在一定条件下必然发生的现象，例如：铁块在水中下沉；木炭燃烧时发热；导体通电后产生磁场；标准气压下的水在100℃沸腾；等等。这些现象称为必然现象，必然现象中出现的事件称为必然事件。概率论创立以前，数学以及其它自然科学主要研究的是必然现象中必然事件的变化规律。18世纪与19世纪的大多数自然科学家都毫不怀疑地相信，由于微积分方法成功地分析了事物的运动与变化的规律，因而可以绝对精确地揭示出任何一个事物发展的过去和未来。18世纪著名的法国数学家、天文学家拉普拉斯就曾说："只要给出一个能包括一切使自然界受到激发时的力的瞬时信息，并且……这种信息足够使这些数据得以提交分析，那么上至宇宙间最庞大的天体的运动，下到最微小的原子的运动，都将包括在一个公式中，对此公式，没有任何一样东西会是不确定的，无论将来还是过去，在它眼里都是现在。"

然而事实上，却有大量的事件在一定条件下有多种变化的可能，而究竟发生哪种变化事先又不能完全肯定，例如：随手抛出一枚质量均匀的硬币，硬币落到地面上时可能这面朝上也可能那面朝上；炮手瞄准某个目标射击，炮弹可能击中目标也可能未击中目标；在一张白纸上画满一条条等距离的平行线，用长度为平行线间距离一半的小针随便扔到纸上，小针可能与直线相交也可能不相交；一年后的今天的天气，可能天晴、也可能天阴或下雨。这些现象统称随机现象。随机现象中那些可能发生、也可能不发生的事件称为随机事件。

还有一种是在一定条件下根本不可能发生的现象，例如：钢铁在常温下熔化；橡胶在低压下导电；种子没有水分而发芽；等等。这些现象称为不可能现象，不可能现象中的事件称为不可能事件。

随机现象是自然界中普遍存在的现象。由于随机现象的发生具有多种可能、我们无法确切地预见会出现什么样的结果，但是我们常常希望了解出现某个事件的可能性到底有多大？

例如，某公司计划向市场投入某种新产品，必须了解该产品在一段时期内保持较好销路的可能性、以及短期内市场上出现与其竞争的同类产品的可能性，才能比较有把握地做出该产品是否投入市场的决定。

又如，在无线电通讯工作中，由于各种原因的随机干扰，往往会发生信号误判，为了提高通讯工作的可靠性，就必须了解通讯系统中由于随机干扰而发生信号误判的可能性有多大。

某个随机事件在一定条件下出现的可能性的大小，就叫做这个随机事件在这一定条件下的概率。概率的值可以用百分比、也可以用小数或分数表示。比如，某种大炮在自动瞄准系

统的控制下，平均每射击100次击中目标93次，可以认为该种大炮击中目标的概率（命中率）为93%或0.93；某个字母在英文中平均每100个字符出现1.2次，可以认为该字母在英文中出现的概率（使用频率）为1.2%或0.012。必然事件也可以看做是概率为1（即发生的可能性为100%）的事件；不可能事件也可以看做是概率为0（即发生的可能性为0）的事件；而随机事件则是概率介于0与1之间的事件。

随机事件的发生带有某种偶然性。所以，尽管某种大炮的命中率为93%，却有可能在某个接连十次的射击中，只击中目标六、七次；尽管某个字母在英文中的使用频率为1.2%，却有可能在某个10000字符的短文中，出现了130多次。

随机事件的发生虽然带有偶然性，但是并非没有规律可循。正如恩格斯所指出的："在表面上是偶然性在起作用的地方，这种偶然性始终是受内部隐蔽着的规律支配的，而问题只是在于发现这些规律。"某种随机现象中各个随机事件发生的规律，只要这个随机现象充分（大量）地多次重复就会呈现出来。

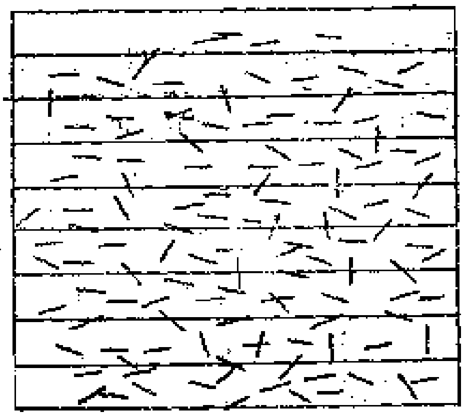
例如，硬币有两个面，规定一个面是正面，另一个面就是

实验序号	每抛掷5次		每抛掷50次		每抛掷500次	
	正面朝上次数	频率	正面朝上次数	频率	正面朝上次数	频率
①	2	0.4	22	0.44	251	0.502
②	3	0.6	25	0.50	249	0.498
③	1	0.2	21	0.42	256	0.512
④	5	1.0	25	0.50	253	0.506
⑤	4	0.8	24	0.48	251	0.502

反面。抛出的硬币落到地面上，正面朝上还是反面朝上纯属偶然。从下表可以看出：当抛掷次数较少时，正面朝上的频率是不稳的；而当抛掷次数逐渐增多时，正面朝上的频率就越来越稳定（大致在0.5附近摆动）。十八世纪，有三位数学家做了成千上万次的抛掷硬币的实际试验，结果如下：其中出现正面与反面的次数几乎都各占总次数的一半（事实上，如果硬币的质量是均匀的，那么哪一面朝上的机会都是均等的）。

试验者	抛掷次数	出现正面的次数	频率(所占百分比)
德·摩根	2048	1061	51.8%
布丰	4040	2048	50.69%
皮尔逊	12000	6019	50.16%
皮尔逊	24000	12020	50.05%

又如，在白纸上画满一条条等距离的平行直线，用长度恰好为平行线间距离一半的均匀小针随便抛到纸上，小针是否与直线相交也是偶然的。可是1777年，法国数学家布丰(Buffon)



却用数学方法证明了这样的小针与直线相交概率为 $\frac{1}{\pi}$。

换句话说，如果m次抛针中有n次小针与直线相交，则当抛针次数m充分大时，比值$\frac{m}{n}$就稳定在圆周率π值的附近做微小的摆动。曾有不少人亲自做了实验来证实这一结论：

年　份	试验者	抛针次数m	比值$\frac{m}{n}$
1850	乌尔夫	5000	3.1596
1855	史密斯	3204	3.1553
1894	福克斯	1120	3.1419
1901	拉萨里尼	3408	3.1415929

可见，随机现象的发生的确是有其内在规律的。

随机现象在大量的观测或试验中呈现出来的规律，叫做随机现象的统计规律。统计规律反映了随机现象的数量性质。概率论就是研究随机现象统计规律的一门数学分支。

要研究随机现象的统计规律，离不开对观测与试验所得的数据进行整理和分析，由此构成了数理统计学的基本内容。数理统计学研究，如何根据观测与试验的统计资料，对随机现象的一般概率特征（如概率、分布律、数学期望等）做出正确的推断。

我们又常把概率论与数理统计学统称为概率统计；有时也把数理统计学视为概率论的一个部分。

我们说概率论"来路不正"，主要是概率论最初来源于赌博问题。16世纪的意大利数学家卡丹首先察觉到：赌博中的输赢虽带偶然性，但是在较多的赌博次数中会呈现出一定的规律。卡丹还专门为此写了一本《论赌博》的小册子。但是，真正为概率论奠定了基础的，还是17世纪的两位法国数学家帕斯卡

(Pascal)和费马，他们对当时一些赌徒们所提出的古怪问题进行了认真的讨论，发现这种带偶然性现象的规律用以往的数学方法无法解决，必需开创和发展新的方法；并且预见到这种对偶然性现象的研究将会对自然科学和哲学产生深刻的影响。以后，又经过18世纪的数学家欧拉、伯努利、达朗贝尔(D'Alembert)、布丰、拉普拉斯，以及19世纪的数学家高斯、泊松、契比雪夫(Цебышев)等人的不断努力，终于使概率论逐步形成了一门充满生命力的独立学科。

最近几十年以来，生产过程的日益机械化、自动化与电子化，和自然科学的研究越来越进入到尖端、边缘的领域，以及电子计算机的出现，不但促使概率论本身更加突飞猛进地发展，而且促使它在生产、科研、国防等各个方面得到了越来越广泛的应用。

举例来说，牛顿的经典力学是决定论的。这种理论使人们相信：只要知道了物体在某一时刻的位置和速度、以及作用于该物体上的力，则根据牛顿运动定律列方程求解，就可以精确地算出该物体以后任何时刻的位置和速度。然而这种理论在研究微观世界的粒子运动时却碰壁了。近代量子力学中的测不准原理指出：人们不可能同时准确地测定微观粒子的位置和速度，也不可能同时准确地测定其能量和时间。但是，却可以通过用统计方法确定粒子在运动过程中速度、能量等物理量的概率分布，来把握粒子的运动。可见，微观现象并不遵从经典力学刻画的规律，而遵从于随机现象的统计规律。于是，概率论就成了近代量子力学的主要数学工具之一。

为了更清楚地比较牛顿经典力学的决定论观点与近代量子力学的概率论观点，让我们稍微详细地讨论一下原子模型：按

经典力学的观点，原子结构与太阳系相似，核外的电子按一定轨道绕核转动。只要知道了电子的轨道和电子的速度，我们就可以精确地描述电子的运动规律。但是这种模型却面临难以克服的巨大困难：由于我们无法同时准确地测定电子的初始位置和速度，我们也就无法确定电子的运行轨道（事实上，实验物理学的电子衍射等观测也表明电子并非沿固定轨道绕核转动）。而按量子力学的观点，尽管电子的绕核转动并不遵循固定轨道，表面上看起来似乎杂乱无章，一时在这里出现、一时又在那里出现，电子的运动仍然是服从一定的统计规律的，它在核外任一确定位置的微小空间出现的可能性（即概率）不随时间而改变。如果我们用点的分布密度来形象直观地反映行踪飘忽的电子在不同位置处出现的概率，就成了量子力学的电子云模型。电子云模型可以很好地描述电子的运动状态，也能够有效地解释各种原子的许多物理、化学性质。

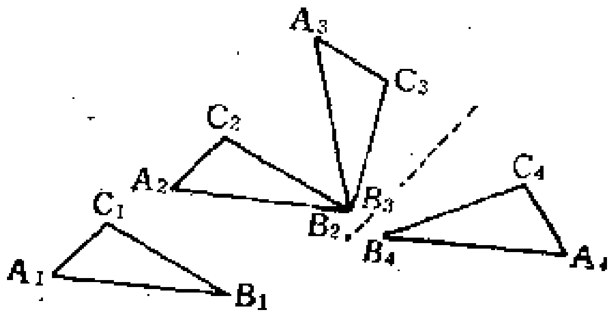
第二十二章

橡皮膜上的几何学——拓扑学

读者可能还记得，我们在“从万能代数模型到MM方法”一章中，曾叙述过大数学家欧拉是如何运用MM方法解决了哥斯尼堡七桥问题的（请参阅该章有关的叙述）。欧拉把七座桥抽象地看成七条线，而把连接桥梁的河岸与小岛、半岛缩小成为连接七条线的四个点，从而使哥斯尼堡七桥问题这样一个现实原型转化成了图论中一笔画问题的数学模型。欧拉在建立哥斯尼堡七桥问题的数学模型这一过程中，实质上是对原型进行了一次“拓扑变换”。

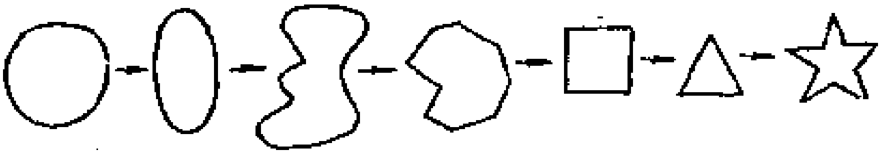
什么是“拓扑变换”呢？

在以欧氏几何为代表的普通几何学里，我们也常常要对图形进行各种变换，如：平移、旋转和反射。这些变换有个共同的性质，那就是图形经过变换之后，大小和形状均不改变（其中最重要的是任何两点间的距离和任何两条直线间的夹角不变，从而建立在距离和夹角基础上的性质都不改变），这种变换称为刚体变换。例如，图中的$\triangle A_1B_1C_1$经平移变换成为



$\triangle A_2B_2C_2$然后经过旋转变换成为$\triangle A_3B_3C_3$，再经反射变换成为$\triangle A_4B_4C_4$，始终有$\triangle A_1B_1C_1 \cong \triangle A_2B_2C_2 \cong \triangle A_3B_3C_3 \cong \triangle A_4B_4C_4$。

而拓扑变换则是弹性的，可以把它想象成是在橡皮膜上进行的：一个画在橡皮膜上的图形F，经过将橡皮膜拉伸、压缩、扭折（但既不拉破也不粘合），变成了另一个图形F'，从F到F'的变换就是拓扑变换。例如，一个圆经拓扑变换可以成为如下一些图形。但绝不会成为“8”字形（有了粘合）或一条线



段（有了断裂）：又如，一个皮球在拓扑变换下可以变为鸡蛋形、月饼形、罐头形、苹果形、葫芦形、香蕉形，……，但绝不会变成口盅形或救生圈形。

图形F经拓扑变换成为F'，大小、形状可以很不相同，然而F上的每个点都只变成F'上的一个点，反过来，F'上的任意一点也只由F上的某一个点所变成（即不会有一个点变成几个点，也不会有几个点变成一个点），这叫一一对应；此外，F上邻近（非常接近）的两点，变成F'上也是邻近（非常接近）的两点，这叫双方连续。因此，我们也可以说拓扑变换是一一对应的双方连续的变换；而把F与F'称为拓扑同胚的两个图形。图形在拓扑变换下保持不变的性质称为图形的拓扑性质，拓扑学就是研究图形拓扑性质的一门数学分支。

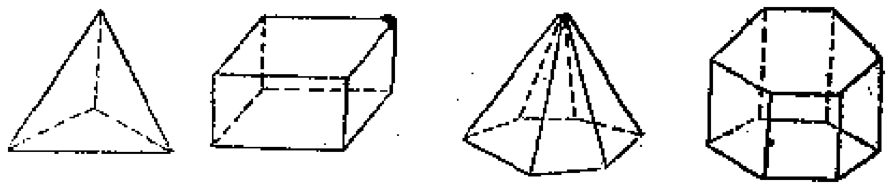
哥斯尼堡七桥问题可以说是数学史上第一个著名的拓扑学问题，也可以说是拓扑学诞生的起点。欧拉给出了哥斯尼堡七桥问题的否定性答案以后，又进一步研究了平面网络以及与之密切相关的多面体的拓扑性质，发现了后来以他的名字命名的

拓扑学中的第一个基本定理——欧拉定理：

任何一个闭凸多面体的顶点数V、面数F与棱数E之间有

$$V+F-E=2。$$

以下给出了欧拉定理的几个特例：



	四面体	六面体	六棱锥	六棱柱
V	4	8	7	12
F	4	6	7	8
E	6	12	12	18

读者可能会注意到：这条拓扑学定理在风格和内容上与我们所熟知的普通几何学中的定理是如此不同，它丝毫也没有涉及到图形的度量性质（如线段的长度、夹角的大小、面积的多少，等等）；读者可能还注意到了：所有的闭凸多面体都是拓扑同胚的（即它们在拓扑变换下可以由一个变为另一个；事实上它们都可以经拓扑变换成为球），关系式$V+F-E=2$确在拓扑变换下保持不变。

拓扑学的起源虽然可以追溯到欧拉时代，但真正成为一门独立的学科，还是近百年的事。19世纪末叶，由于天文学和物理学关于动力系统相空间的研究，引起了数学家们对空间连续性、连通性等拓扑性质的高度重视，拓扑学才蓬勃发展起来，并在电磁学理论、无线电技术、自动控制、基本粒子研究和电

子计算机设计中发挥了巨大的作用。

不动点定理堪称是拓扑学中最辉煌的成就之一。什么是"不动点"呢？举例来说，实数经变换$f(x)=x^2-2$后，绝大多数点都变成了另外的点（如$f(-2)=2$，$f(0)=-2$，$f\left(\frac{1}{2}\right)=-\frac{7}{4}$，$f(\sqrt{2})=0$，……），但是却有两个点$-1$与2经变换$f(x)$后仍然保持不变（$f(-1)=-1$，$f(2)=2$），这种在变换$f(x)$下保持不变的点，就称为变换$f(x)$的不动点。

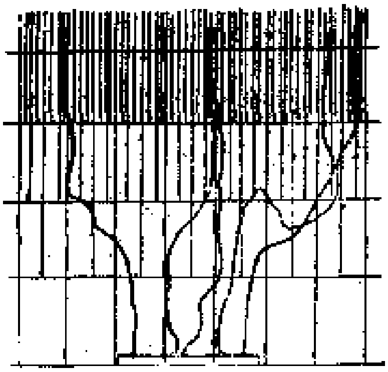
并非所有的变换都有不动点，例如变换$f(x)=x+1$就没有不动点。1912年，荷兰数学家布劳维尔（L.Brouwer）首先证明了满足一定条件的连续变换（不一定是拓扑变换，但拓扑变换一定是连续的）至少有一个不动点。自那以后，数学家们对上百种比较重要的具体的连续变换找出了它们各自的不动点。有关不动点问题的一系列成果，已经发展成了一整套理论。数学家们还在努力发掘新的成果。数学家们之所以如此热衷于不动点理论，是因为有很多数学理论问题和应用问题，可以用不动点理论来判断所要求的解是否存在。虽然它不能帮助我们把具体的解求出来，但是在着手进行求解工作之前，事先判断解是否存在还是很有必要的；如果解并不存在，就避免了徒劳无功的损失；更何况对很多理论问题来说，我们仅仅需要知道解是否存在就够了。

"不动点"和方程的根有很密切的联系。例如变换$f(x)=x^2-2$的不动点，实质上就是方程$x=x^2-2$（即$x^2-x-2=0$）的根。数学家们已经在不动点理论的基础上创造了一些求不动点（或方程的根）的有效算法，可以按照预定精度的要求把不动点（或方程的根）求出来。其中以美国普林斯顿大学哈德

罗·库恩教授1976年所提出的求 n 次复系数方程

$$Z^n+a_1Z^{n-1}+\cdots+a_{n-1}Z+a_n=0$$

的根的方法最为奇特：准备一个培养皿和一个立体大篱笆，篱笆越往上越密；把所要求解的方程的有关数据输入培养皿，培养皿内便吐出几只芽，芽转眼变成藤，飞快攀上篱笆，一层层往上穿，每根藤恰好指向方程的一个根，于是所求方程的根很快就被找出来了。这种求方程根的"盆栽技术"并非神话，完全可以通过电子计算机模拟来实现。

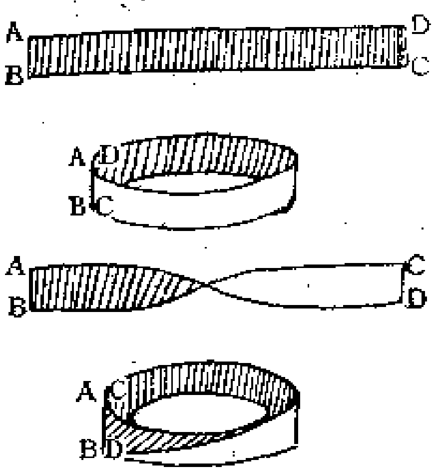


拓扑学中还有很多令人惊叹的奇珍异宝。

"四色定理"便是其中之一。早在一百多年以前，人们就发现：无论多么复杂的地图，只需用四种颜色着色就足以使相邻的不同区域分开。这个问题貌似简单，其实非常复杂，不少优秀数学家涉足其中，陷入困境，始终无法给予严格的证明。虽然"五色定理"(即任何一幅地图最多只需要五种颜色就足以区分邻国)在上一世纪末就已获证，但是"四色定理"却直到1977年，才由美国数学家汉肯(W.Haken)和阿佩尔(Appel)借助高速电子计算机IBM360，花费了1200个机器小时，检验了1482种的全部可能情形，才算完成证明(后来很快又有人宣布了一个只用50个机器小时的更为简单的证明)。专家们评论说："四色定理"的计算机证明，不仅仅在于解决了一个纠缠了人们一百多年的难题，而且很可能成为数学思想发展史上一系列新想法的起点。

"默比乌斯带"也是拓扑学中的瑰宝之一。拿一张纸条

$ABCD$，使A与D、B与C重合地粘接起来，我们就得到了一个普通有两个面的曲面。为了区分这两个面，我们不妨把朝里的一面涂成黑色，而朝外的一面涂成白色。如果让一只蚂蚁在这个曲面的某一面上爬行，不让它绕过曲面的边缘，也不让它穿过曲面，那么无论它怎么爬，它也爬不到另一面上去。现在，把纸条从粘接处分开，扭转180°，再使A与C、B与D重合地粘接起来，我们就得到了只有一个面的曲面，已经无所谓里外了。蚂蚁在这种曲面上爬行，不用绕过曲面的边缘、也不用穿过曲面，就可以从原来白的一面爬到原来黑的一面，或者反过来从原来的黑面爬到原来的白面。19世纪的德国数学家默比乌斯（Möbius）首先研究了这种单面曲面的拓扑性质，故称之为"默比乌斯带"。由于"默比乌斯带"性质奇特，在工程技术上得到了广泛的应用。例如：电话无人自动回答器上的磁带若用默比乌斯带，磁带的两面都可以自动录音，比长度相同的普通磁带，其信息存储量可以提高一倍；又如：电子计算机中的环形磁芯若扭成默比乌斯带的形状，其内部磁场的分布就比普通磁芯均匀得多，反复磁化的过程便可以得到加速，从而提高计算机信息的存取效率。



第二十三章

从精确到模糊的转变——模糊数学的诞生

本世纪60年代以前，众多的数学分支虽然五花八门，但是不外乎可以分为两大类：一类是以微积分、代数学为首的经典数学，一类是以概率论、统计学为首的随机数学。

经典数学研究的是事物发展变化的精确数学规律。例如，由万有引力定律可以推导出行星的运行轨道并准确地指出某一时刻行星的相应位置。随机数学研究的是随机现象中随机事件的统计数学规律。例如，尽管抛掷立方体骰子哪一面朝上是偶然的，但是在大量的投掷中每一面朝上的次数在总次数中所占的比例都逐渐趋于稳定在1/6左右。

经典数学与随机数学的共同特点是：它们所研究的对象都是确定的。例如：金星而不是火星的运行轨道，这个时刻而不是那个时刻的运行位置；立方体的均匀骰子而不是其它骰子，各次投掷中每个面朝上的可能性而不是在接连的两次投掷中同一个面连续朝上的可能性；等等。

确定的数学对象可以用分明集合来表示。所谓集合，可以理解为具有某种性质的事物的全体，例如：太阳系中的九大行星；立方体骰子的六个面；所有的自然数；某平面上所有的点；等等，就都是集合。分明集合的界限是明确的，一个事物要么属于给定的集合，要么不属于给定的集合，绝不能模棱两

可．如果用N表示全体自然数组成的集合，$\in$表示属于，$\bar{\in}$表示不属于，则有$1\in N$，$2\in N$，$10\in N$，$4756\in N$，$0\bar{\in}N$，$\frac{1}{2}\bar{\in}N$，$-1\bar{\in}N$，$\sqrt{2}\bar{\in}N$，等等。

但是，在日常生活、生产实践和科学技术中，并不是所有的事物和现象都是确定的．例如："高个子的人"，究竟多高才算是"高个子"呢？又如："现在"，什么样的时间范围才属于"现在"呢？其它的诸如："快慢"、"长短"、"轻重"、"冷热"、"干湿"、"晴阴"…也都没有可以区别、划分的绝对标准。这些事物和现象就称为模糊事物和模糊现象。其发展变化的规律，用传统的经典数学与随机数学的办法是难以研究的。

不过，在电子计算机出现以前，人们并没有感觉到有专门研究模糊事物和模糊现象的必要。电子计算机出现以后，机器的人工智能问题很自然地提到了议事日程上。计算机虽然在运算速度和精确性上令人类望尘莫及，但是在识别、判断和推理能力上却还远远落后于人类。主要是因为人脑非常善于认识和处理模糊事物和模糊现象（例如辨别极为潦草的笔迹，听懂不完整甚至不合常规的语言，跳过推理的某些中间环节直接得出正确结论），而计算机目前还只能处理相对来说比较简单的确定事物和确定现象（例如辨认标准规范的印刷体文字，听懂语音和语法都极其严格的语言，进行按步就班的机械性的推理），稍遇需要灵活处理的模棱两可的情况就会使计算机左右为难，无法正常工作。由于传统的数学方法只适用于确定的研究对象，因而在建立计算机的人工智能方面遇到了难以克服的巨大障碍。这就迫使数学家们不得不考虑传统的精确数学的反面——模糊，模糊数学也就应运而生了。

1965年，美国加里福尼亚大学的控制论专家查德（L．A．Zadeh）首先提出用模糊集合来表示模糊的数学对象。模糊集合与分明集合的根本区别在于：模糊集合的界限并不是绝对分明的，而是相对模糊的。对于分明集合A来说，一个元素x要么属于A，要么不属于A，二者必居其一，用A的特征函数表示就是

$$\mu_A(x)=\begin{cases}1 & \text{当}x\in A\text{时}\\ 0 & \text{当}x\notin A\text{时}\end{cases}$$

其中的1与0可以看做是元素x从属于集合A的程度（隶属度）：1表示x有百分之百的资格属于A，0则表示x没有任何资格属于A。而对于模糊集合来说，元素的隶属度就不再只是1与0两种可能，而可以是0与1之间的任何一个实数。例如，“高个子”的集合A，按照我国目前公认的标准（以厘米为单位），其元素的隶属度可用A的特征函数表示成

$$\mu_A(x)=\begin{cases}1 & 180<x\\ 0.9 & 178<x\leqslant 180\\ 0.8 & 175<x\leqslant 178\\ 0.5 & 170<x\leqslant 175\\ 0.3 & 165<x\leqslant 170\\ 0.1 & 160<x\leqslant 165\\ 0 & x\leqslant 160\end{cases}$$

其含义是：180厘米以上的人有百分之百的资格看做“高个子”，178到180厘米的人有百分之九十的资格看做“高个子”，175到178厘米的人有百分之八十的资格看做“高个子”，而160厘米以下的人就再也没有任何资格可以看做“高个子”了。

对于历史的分期问题，用模糊集合的方法表示也比较科学。有人提出可以把我国的奴隶社会看做是如下一个模糊集合

$$奴隶社会=\frac{1}{夏}+\frac{1}{商}+\frac{0.9}{西周}+\frac{0.7}{春秋}+\frac{0.5}{战国}$$

$$+\frac{0.4}{秦}+\frac{0.3}{西汉}+\frac{0.1}{魏晋}$$

其含义是：夏与商是百分之百的奴隶社会；西周时奴隶社会开始衰落，出现了封建社会的萌芽；战国时期是半奴隶半封建社会；魏晋时封建社会已经基本确立并巩固但还剩下一点奴隶制的残余。这种历史分期的好处是，它使我们一目了然地看到了历史发展的动态和历史时期的交迭。

正如分明集合是传统的精确数学的基础一样，模糊集合也是模糊数学的基础。自查德提出模糊集合以来，数学对象之间的各种模糊关系、模糊运算也相继产生，模糊数学迅速发展起来，理论不断完善，应用日益广泛。模糊数学的方法已经渗透到自动控制、图象识别、故障诊断、信息检索，以及心理学、生物学、历史学、语言学、气象学、地理学、经济学等各方面。

例如，天气的好坏与人民的生活和社会生产是息息相关的，几乎没有人不关心天气的变化，然而人们又常用“天有不测风云”来形容天气活动的复杂和预测的困难。尽管今天人们在电子计算机的帮助下已经将数值预报方法付诸实施，短期内的天气预报准确率逐步提高，但是传统的数值预报方法仍有计算量太大、长期预报以及局部预报的准确率较低的缺点，特别是一些设备较差的小型气象站几乎无法使用。我国安徽省长

丰县气象站，利用模糊数学方法对当地世代相传的民间测天谚语“冻大水大”，结合当地的有关气象资料进行了模糊化处理，建立起了汛期降水预报的模糊数学模型，从1980年起正式投入业务使用。以这个模糊数学模型为主并参考其他预报工具，连续好几年相当准确地预报了当地的旱涝趋势，为当地农业生产抗旱、防涝保丰收做出了很大的贡献。我国中央气象台最近几年也开展了用模糊数学模型预报天气的试验，取得了令人鼓舞的成果，在国际气象学界受到了好评。

又如，疾病的很多症状，如“头疼”、“鼻干”、“口苦”、“心烦”、“食欲不振”等等，都是一些模糊性的概念，某种疾病的表现实际上是若干症状所组成的模糊集合的交。而中医传统的诊断方式“切、望、闻、问”，也都带有很大的模糊性，很多中医的宝贵经验甚至只能意会，无法言传。如果用电子计算机来模拟中医诊断，模糊数学模型就要优于概率统计模型或其它模型。北京市中医院正是以模糊数学模型为主，总结了著名肝病中医关幼波的辨证施治经验，研制了一套电脑诊断肝病的专家系统。经过门诊试验，电脑所开处方的正确率已达98%左右，受到了关大夫本人的赞许，并荣获了北京市科技成果奖。

模糊数学还是一门非常年轻的新兴学科，暂时还没有一个坚实的理论基础；而且虽然有着广泛的应用，却又缺乏重大深刻的结果；因此引起了许多数学家的焦虑和不安。但是，模糊数学的创始人查德始终信心十足，他在展望模糊数学时说：“在即将到来的年代里，我相信近似推理的模糊逻辑将发展成为一个重要领域，而变成研究哲学、语言、心理、管理、社会、医学诊断等领域的新方法的基础。与此同时，我们同样也会看

到古典逻辑基础上的模糊集数学理论的许多重要发展，将对纯粹及应用数学做出显著贡献。无须多说，对模糊集合论今后十年的发展，只能有个朦胧的想法，但可以肯定的是，必将使模糊集论的重要性、影响力和应用都得到迅速发展，最后还必将在人类知识和科学方法论的宝库中占有一席之地。”

第二十四章

"人尽其才、物尽其用"的运筹学

在实际工作中，我们常常会遇到这样的问题：在现有的人员配备下，怎样根据每个人不同的工作能力安排不同的工作任务，以充分发挥每个人的特长，使所花费的代价最小或完成的任务最多？这就是一种"人尽其才"的问题。

说得具体一点：假定有甲、乙、丙、丁、戊五名工人，要分派他们去做A、B、C、D、E五种工作，每人完成各种工作所需要的时数如左表所示，应怎样安排，才能使得他们每人都担当一种工作，而所用的总时数又最少？

工作 / 时数 / 工人	A	B	C	D	E
甲	3	1	8	4	1
乙	7	5	7	6	6
丙	4	3	5	5	3
丁	4	3	1	6	3
戊	6	5	6	5	5

显然，这五名工人完成五种工作所用的总时数，是随着安排方案的不同而不同的，其中必定存在一个总时数最少的方案。以下列出了两个不同的方案，用数学方法可以证明，方案

Ⅱ是一切所有可能的方案中总时数最少的一个：

方案Ⅰ 总时数21	A	B	C	D	E
甲		1			
乙				6	
丙			5		
丁	4				
戊					5

方案Ⅱ 总时数16	A	B	C	D	E
甲					1
乙		5			
丙	4				
丁			1		
戊				5	

在生产管理中，我们还常常会遇到这样的问题：在现有的生产条件下，如何合理安排生产，才能以最少的物资消耗去获取最大的经济效益？这就是一种“物尽其用”的问题。

也说得具体一点：假定某厂生产 A、B 两种产品，产品 A 每个耗电8度，用银 8 克，利润五元；产品B每个耗电6度，用银16克，利润六元；该厂每月用电限制为 36000 度以下，用银限制在64千克以内，两种产品各生产多少才能获得最大的利润？

根据要求，若每月生产产品 Ax 个、产品By个，则上述问题相当于在约束条件

$$\begin{cases} 8x+6y\leqslant 36000 & \text{(用电限制)} \\ 8x+16y\leqslant 64000 & \text{(用银限制)} \\ x\geqslant 0,\ y\geqslant 0 & \text{(}x\text{、}y\text{的意义)} \end{cases}$$

下，求目标函数

$$f(x、y)=5x+6y \qquad \text{(所获利润)}$$

的最大值。

象这样，约束条件和目标函数都是线性关系（即变量 x、y 等的一次式，其图象为一直线）的最大、最小问题，称为线性规划问题，一般需要用线性规划中的专门方法（如单纯形法）

求解。不过，上述问题比较简单，用图解法即可求得当$x=2400$，$y=2800$时，$f(x,y)$ 取得最大值28800。即，每个月产品 A 应生产2400个，产品B应生产2800个，可获最大利润28800元。

"人尽其才"也好、"物尽其用"也好，一般都需要我们从多种可行的方案中选择一个最优方案（或付出的代价最少，或所用的时间最省，或产品的产量、质量最高，等等）。从多种方案中选择一个，叫做"决策"，运筹学就是研究决策问题的数学分支。"运筹"，二字，出自我国一句成语"运筹帷幄"，原指军事上的策划与指挥，后来泛指一切政治、经济、军事活动中的决策过程。

朴素的运筹思想，古人早已有之。在此仅举一例：北宋真宗年间，京城发生了一场大火，富丽堂煌的皇宫毁于一旦。灾后的修复工程十分艰难，既要清理断垣残壁，又要挖土烧砖，还要从外地调运大量建筑材料。搞得不好，旷日持久事小，劳民伤财事大。主持修复工程的丁谓冥思苦想，提出了一个"一举三得"的施工方案：先在宫前大街开挖一条运河，挖出来的土用来烧砖制瓦；然后使运河与汴水接通，大批建筑材料可以直接由水路运来；再把清理废墟的瓦砾填入运河，以恢复原来的大街。这一施工方案，巧妙地解决了就地取材、方便运输和处理废物的难题，一直为后人所称道。

古人虽已早懂"运筹"，但并没有形成一套完整的理论体系，也没有自觉地使用数学工具。最早用数学方法来研究运筹问题的，当推19世纪法国数学家傅里叶（Fourier）。1823年，傅里叶首次提出了在线性约束条件下的线性目标函数极小值的算法，虽然比较粗糙，毕竟是定量研究的开端。不过，由于当时生产力水平还不高，人们并没有感到用数学方法进行运筹的

必要，因而运筹学的发展十分缓慢。

直到20世纪30年代，运筹学才在大规模生产的刺激下蓬勃发展起来。特别是第二次世界大战期间，由于物资调运、物资库存、武器部署、军事对策等大量实际问题迫切需要解决，数学家与其他方面的科学家和工程师密切配合，终于在解决这些实际问题中提炼出了规划论、对策论、排队论等系统的数学理论，运筹学才正式脱胎出来成为一门独立的数学分支。

除了许多具体的有效算法以外，运筹学还包含了一些十分可贵的思想方法，其中最重要的两条——整体性观点和有序性原理，后来成了一般系统论的思想基础。

为了说明整体性观点和有序性原理，让我们来看如下的例子：

某工厂有甲、乙两种不同型号的车床，要安排加工两种零件A、B的任务。其中加工零件A是重点任务，470个零件必须当天加工完毕；加工零件B是非重点任务，加工得越多越好。两种车床的数目与加工两种零件的工效（每台每天所加工的零件个数）如表所示。应如何安排才最优？

零件 车床　工效	零件A	零件B	车床数
甲车床	15(个/天)	5(个/天)	30台
乙车床	12(个/天)	3(个/天)	25台
任　务	470个	越多越好	（工效表）

如果只从保证重点任务的角度出发，由于甲车床干重点活的工效比乙高，自然应先安排甲车床加工零件A，30台可加工

	零件 A	零件 B	车床数
甲车床	15×30	0	30台
乙车床	12×2	3×23	25台
完成任务	$474>470$	69	（方案Ⅰ）

$15\times30=450$个，还差20个零件A安排乙车床加工，实际需要2台乙车床（假定每台车床只加工一种零件）；剩下的23台乙车床再加工零件B，共可加工$3\times23=69$个。于是，不但完成了470个零件A的重点加工任务，还完成了一部分非重点加工任务（69个零件B）。这就是方案Ⅰ。

然而，方案Ⅰ的安排并不十分合理。如果从甲车床中抽调10台去加工零件B，这10台甲车床原来加工零件A 150个的任务交给13台乙车床去做，不但同样可以完成470个零件A的重点加工任务，而且比方案Ⅰ还多加工11个零件B。这就是方案Ⅱ。

	零件 A	零件 B	车床数
甲车床	15×20	5×10	30台
乙车床	12×15	3×10	25台
完成任务	$480>470$	80	（方案Ⅱ）

可见，只着眼于保证重点任务的考虑并非最佳。重点任务虽然应该优先考虑，但也绝不能忽视非重点任务对整个工程安

排的影响。表面上看起来，甲车床加工零件A的工效确比乙车床要高；可是，如果兼顾加工非重点零件B的工效，情况就有所不同了。事实上，甲车床加工A、B两种零件的工效比是$15:5=3:1$，即甲车床每加工1个零件B相当于加工3个零件A；而乙车床加工A、B两种零件的工效比是$12:3=4:1$，即乙车床每加工1个零件B则相当于加工4个零件A。相比之下，让乙车床去加工零件B就显得有点“大材小用”了，因为乙车床每多加工1个零件B就要少加工4个零件A，甲车床每多加工1个零件B才少加工3个零件A。不如干脆反过来，尽量先安排乙车床加工零件A。由此便得到方案Ⅲ

在方案Ⅲ中，25台乙车床全部用来加工零件A，共可加工$12\times25=300$个，还差170个零件A由甲车床完成，实际需要12台甲车床（同样假定每台车床只加工一种零件）；剩下的18台甲车床再去加工零件B，共可加工$5\times18=90$个。也同样完成

	零件A	零件B	车床数
甲车床	15×12	5×18	30台
乙车床	12×25	0	25台
完成任务	$480>470$	90	（方案Ⅲ）

470个零件A的重点加工任务，还比方案Ⅰ与方案Ⅱ都完成了更多的加工零件B的非重点任务。实际上，方案Ⅲ是所有可行方案中最好的。

上述例子告诉我们：考虑问题不能只看问题的一个方面，而必须统观全局，看到与问题有关的各个事物之间的相互联系

和影响。用系统论的话来说：系统是一个由各组成要素相互联系、相互作用并与周围环境发生关系的有机的整体，系统的功能并不是各组成部分功能的简单总和，还要取决于各部分之间的相互关系和作用。这就是系统论的"整体性观点"，也是系统论的核心观点。当然，一般系统论所研究的是开放系统，即与周围环境发生关系的系统；在上述例子中，为了简单起见，我们已经把所讨论的问题相对孤立了出来。

上述例子还告诉我们：在各组成要素不变的条件下，只需调整各要素之间的相互关系和有序结构，就可以得到最优化方案。用系统论的话来说：系统的功能不仅取决于组成系统的要素，还取决于各要素之间的相互关系和有序结构；在组成要素已经确定、环境影响不变的情况下，要改变系统的功能，关键就是改变组成要素间的相互关系和有序结构。这就是系统论的"有序性原理"，它是"整体性观点"的进一步延伸和发展。

本世纪40年代以来，以运筹学为先导，系统工程、控制论、信息论、一般系统论等新兴学科相继出现，发展很快，并且互相渗透、互相结合，在工程技术、社会生产、经济管理等各方面发挥着日益明显的效益，并对整个人类的思维方式和科学方法产生了极其深远的影响。

第二十五章

千奇百怪话悖论

悖论由来已久，早在两千多年以前的古希腊和我国先秦哲学时期，就陆续发现了不少令人惊奇而有趣的悖论。下面是古希腊哲学家喜欢谈论的"鳄鱼二难"悖论：

一条鳄鱼从一位母亲手中抢走了一个小孩。鳄鱼对孩子的母亲说："请你回答，我会不会吃掉你的孩子？答对了，我就把孩子不加伤害地还给你；否则，就别怪我不客气了！"聪明的母亲机智地回答说："啊，你是要吃掉我的孩子的。"这下鳄鱼可左右为难了，如果它交回小孩，母亲就说错了，它当然可以吃掉小孩；可是如果它吃掉小孩，母亲又说对了，它又得把孩子毫无伤害地交还出来；无论如何，都与它自己的允诺相矛盾。拙劣的鳄鱼搞懵了，结果把孩子交给了母亲。母亲一把抱过孩子，跑掉了。

古代的一些哲学家们虽然喜欢谈点悖论，但是在很长一个历史时期内，悖论并未引起普遍的足够重视。大多数数学家和哲学家们认为悖论不过是巧妙编造的貌似有理而又自相矛盾的谬论，与数学和哲学本身的理论并无多大关系。直到集合论中罗素悖论（请参看"数学的三次危机"一章）的出现，整个数学的基础为之动摇，人们才感到悖论不得不需要认真对待了。

现在一般把悖论分为两大类，一类是逻辑悖论，只涉及逻辑问题；另一类是语义悖论，除了涉及逻辑以外，还涉及讨论

问题时所用的语言。

逻辑悖论又分两种：第一种，从确实正确的前提出发，得到了人们的直觉或当时的认识水平看来是错误的，而实际上却又是正确的结论；第二种，前提貌似正确，实则隐含矛盾，因而得到确实错误的结论。

做为第一种逻辑悖论的典型例子，比较著名的有伽利略悖论。

可以表示成某个自然数平方的形式的那样的自然数，例如1（$=1^2$）、4（$=2^2$）、9（$=3^2$）、16（$=4^2$）、25（$=5^2$），叫做完全平方数。完全平方数（以下打框者）只在自然数中占很少的一部分：

$\boxed{1}$ 2 3 $\boxed{4}$ 5 6 7 8 $\boxed{9}$ 10 11 12 13 14

15 $\boxed{16}$ 17 18 19 20 21 22 23 24 $\boxed{25}$ 26 27

28 …

10以内的完全平方数只有3个，100以内的完全平方数只有10个，1000以内的完全平方数只有31个，一般地，n以内的完全平方数不会超过$\sqrt{n}$个。从直观上看，完全平方数的“个数”似乎要比自然数的“个数”少得多，是符合人们从欧几里得时代起就已经普遍接受的“部分小于全体”这一公理的。

可是，自然数有无穷多个，完全平方数也有无穷多个，我们不能象平常数数那样去比较自然数与完全平方数的多少。除了一个个地数以外，还有什么办法能比较它们的多少呢？那就是设法在完全平方数与自然数之间建立起某种“对应”关系。正好象我们要比较某个教室中的座位与某个班学生的多少一样，也不必一个个地数，只要让学生坐到教室里的座位上去，

每个座位只准坐一个学生，每个学生也只准坐一个座位。坐好以后，如果全体学生都坐下了，还有座位剩，座位就比学生多；如果座位不够坐，还有学生没有座位，学生就比座位多；如果全体学生都有座位，而又没有任何一个空位留下，座位就和学生一样多，最后这种情形叫做座位与学生之间的“一一对应。”

伽利略通过映射$n^2 \stackrel{\varphi}{\longleftrightarrow} n$（即令$n^2$与$n$对应）：

$$\begin{array}{cccccccc}
1, & 4, & 9, & 16, & 25, & 36, & \cdots, & n^2\cdots \\
\updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow \\
1, & 2, & 3, & 4, & 5, & 6, & \cdots, & n\cdots
\end{array}$$

建立起了完全平方数集与自然数集之间的一一对应关系。每个完全平方数，都有唯一确定的自然数与之对应；反过来，每个自然数，也都有唯一确定的完全平方数与之对应。这意味着完全平方数的“个数”与自然数的“个数”是一样多的（读者不妨把1、2、3、4、…、n、…这些自然数看做“座位”，而把1、4、9、16、…、n^2、…这些完全平方数看做是坐在相应座位上的“学生”，就容易理解为什么说是“一样多”了）。部分居然可以“等于”全体，在当时的人们看来简直荒谬透顶。这就是著名的伽利略悖论。

直到19世纪末，德国数学家康托尔（Cantor）和戴德金（Dedekind）建立起无穷集合论，人们才认识到：部分可以“等于”全体，正是无穷与有限的本质区别。今天看来，伽利略悖论已经不成其为“悖论”，只是由于历史的原因，我们仍然沿用它原来的名称。

“无穷”相对于“有限”而言，已经由量的积累产生了质

的飞跃。因此，有人将“无穷”形容为危险的陷井，以告诫人们千万不要用“有限”的眼光来看待“无穷”。

请看下面的“无穷和”：

$$S=1+(-1)+1+(-1)+1+(-1)+\cdots\cdots$$

如果从第一项起，先把每相邻的两项分别相加再求和，就得

$$\begin{aligned}S&=\underbrace{1+(-1)}+\underbrace{1+(-1)}+\underbrace{1+(-1)}+\cdots\cdots\\&=0+0+0+\cdots\cdots\\&=0.\end{aligned}$$

如果从第二项起，先把每相邻的两项分别相加再求和，又得

$$\begin{aligned}S&=1+\underbrace{(-1)+1}+\underbrace{(-1)+1}+\underbrace{(-1)+1}+\cdots\cdots\\&=1+0+0+0+\cdots\cdots\\&=1.\end{aligned}$$

之所以会如此奇怪，原因就在于“无穷和”的运算不能象“有限和”那样随意结合或交换。事实上，可以证明上述“无穷和”S是发散的（即其“和”根本就不存在）。可见，“无穷”必须慎重对待。

做为第二种逻辑悖论的比较简单的例子，可以举所谓“秃头悖论”：

先假定两个事实做为前提：

（Ⅰ）一个人头上有20万根头发则不是秃头；

（Ⅱ）如果一个人不是秃头，则掉一根头发以后仍然不是秃头。

这两个事实乍看都可以承认，但是不难由此推出“一根头

发也没有的光头不是秃头”的荒谬结果（请读者试试）。原因在于“秃头”的定义不明确，何谓“秃头”？何谓“不秃”？两个前提都没讲清楚。特别是（Ⅱ），混淆了“秃”与“不秃”的界限。即使按模糊数学的办法，把“秃头”看做是个界限不分明的模糊集，（Ⅱ）的叙述仍有问题，因为当一个人掉了一根头发以后，其隶属于“秃头”集的程度不可能总没有丝毫的变化。

前面提到过的集合论中的罗素悖论，属于比较复杂的第二种逻辑悖论。“理发师悖论”不过是罗素悖论的通俗说法。还可以将罗素悖论改头换面成如下的“机器人悖论”：

对某个机器人，规定它必须修理并且只修理一切不修理自身的机器人，请问谁修理该机器人？如果该机器人不修理自己，按照规定它又必须修理自己；如果该机器人要修理自己，则同样按照规定它又不能修理自己。无论如何，总会导致矛盾。

罗素悖论的成因十分复杂，究竟是什么样的前提导致了悖论的出现，众说纷纭，各种解释都有，并因此形成了各种各样的学派，争论的焦点主要集中在“一切集合汇集在一起”是否仍然构成一个集合。这一问题迄今仍未得到彻底解决。

做为语义悖论的典型例子，有“说谎者悖论”：“我正在说的这句话是谎话。”试问这句话是否真是谎话呢？若是谎话，则它陈述的事实为假，因而“我”说的就不是谎话；若不是谎话，则它陈述的事实为真，因而“我”说的又是谎话。总之，无法自圆其说。

“自谓悖论”也属于语义悖论。一个形容词能够自己形容自己，则称为“自谓的”，否则就是“非自谓的”。例如，“中文的”这个形容词就是“自谓的”，因为它不但可以形容一切用

中文书写或印刷的东西，还可以形容它自己，它自己也是"中文的"；而"英文的"这个形容词则是"非自谓的"，因为它本身并非是"英文的"。现在问"非自谓的"这个形容词是否"自谓"呢？若它是自谓的（即可自己形容自己），则按词义它又应该是"非自谓的"；若它是非自谓的，则正好又与它本身的词义相同，按照"自谓的"定义，它又应该是自谓的。

语义悖论的成因是由于自然语言比较含糊。按塔斯基（Tarski）和罗素的观点，语言应该层次分明，同一层次的语言不能在其自身中讨论真假，只能在高一层的语言中加以讨论。例如以下两句话

（A） 下一命题B是假的。

（B） 苹果有毒不能吃。

这两句话就是不同层次的语言，句（A）是抽象语言，句（B）是实际语言，在比实际语言（B）高一层次的抽象语言（A）中指出了（B）的真假，就不会出现任何悖论。

我国著名数学家徐利治先生指出："由于人的认识在各个历史阶段中的局限性和相对性，在人类认识的各个历史阶段所形成的各个理论体系中，本来就具有产生悖论的可能性，但是在人类认识世界的深化过程中同样具备排除悖论的可能性和现实性。人类认识世界的深化没有终结，悖论的产生和排除也没有终结。""悖论问题的研究，对于数学基础理论、逻辑学、语言学和哲学的研究都有重要意义。试看语义学、类型论、多值逻辑、公理化集合论的几个重要系统，直到近代数学的三大流派的形成和发展，无不直接来自悖论问题的研究；还有公理化方法论，数理逻辑、证明论和模型论的形成和发展的一个切近原因，除了非欧几何的产生之外，仍然是悖论问题的研究。"

第二十六章

数学与实验

欧拉早就正确地指出："数学这门科学，需要观察，还需要实验。"然而，观察与实验在数学教育的主流中却长期得不到应有的重视，以至在大多数人的心目中似乎观察与实验是与数学无缘的。人们一般认为：做为理论科学的数学，与做为实验科学的物理、化学、生物等在研究方法上有很大的差别。理论科学的发展形式主要是演绎，即从一些自明的和公认的事实（公理）出发，经过演绎推理得出一批结论，在此基础上经过进一步的演绎推理又得出更多的新的结论，如此继续下去，直到最终形成一个完整的具有严格逻辑顺序关系的理论系统。实验科学的发展形式则主要是假说，即首先通过对物质世界的观察与实验收集一批有关的事实，以此为基础提出某个足以解释已知事实的假说，再从假说出发建立相应的理论系统，并做出一定的预见；然后通过对物质世界的进一步观察和实验来检验预见，修正假说，如此继续下去，不断形成越来越符合物质世界客观规律的新的假说和理论。观察与实验对于数学这样的理论科学来说似乎无足轻重、可有可无，而只对物理、化学、生物等实验科学来说才是必不可少的。

但是，为什么历史上卓有贡献的大数学家，从古代的阿基米德、刘徽，近代的欧拉、高斯，直到现代的G·波利亚、华罗庚等人，都一再强调和推崇数学中的观察与实验呢？让我们

<table>
<tr><th></th><th>欧 氏 几 何 学</th><th>牛 顿 动 力 学</th></tr>
<tr><td rowspan="2">前奏</td><td>约公元前600年以前。</td><td>17世纪中叶以前。</td></tr>
<tr><td colspan="2">以广泛收集事实、不断积累经验为主；观察与实验基础上的归纳是主要手段。</td></tr>
<tr><td>形成</td><td>约公元前300年左右。
由欧几里得完成了几何学的公理化；演绎的分析与综合是主要手段。（欧氏所提出的公理、公设也可以看做是某种假说。）</td><td>17世纪80年代。
由牛顿完成了动力学的公理化；演绎的分析与综合是主要手段。（牛顿三大运动定律可以看做是假说，也可以看做是公理。）</td></tr>
<tr><td>确立</td><td>尽管人们早已感觉到欧氏公理系统不完善，并对欧氏第五公设提出了质疑，但欧氏几何却能与当时人们的直观和实践很好相符，这就确立了欧氏几何的统治地位。
19世纪20年代，高斯、罗巴契夫斯基等人观测大尺度光线三角形，没有发现异常，也可以看做是对欧氏几何在某种程度上的证实。</td><td>19世纪40年代，亚当斯、勒威烈等人根据万有引力定律成功地推算出海王星的轨道，并精确地预言海王星的位置，确认了牛顿动力学的统治地位。</td></tr>
<tr><td>后续</td><td>19世纪末到20世纪初，希耳伯特总结了前人对欧氏几何公理系统修正的大量工作，提出更为完善的希耳伯特公理体系。
爱因斯坦的相对论指出：非欧几何的时空模型更符合客观物质世界的实际情况，得到了天文观测和高能物理实验的证实。</td><td>20世纪初，爱因斯坦指出了牛顿力学的缺陷，并提出了以光速不变原理和相对性原理为公理（或假说）的演绎理论体系——相对论。牛顿力学成了相对论中物体在低速运动情形下的一种特例。</td></tr>
</table>

从宏观与微观两个不同的角度来谈谈这个问题。

从宏观的角度来看，尽管数学理论的发展与物理、化学、生物等实验科学有很大的不同，然而也不乏相似之处，请看表中(178页)欧氏几何学与牛顿动力学的发展简明而有趣的比较。

我们清楚地看到：数学与物理的发展在很多方面是十分相似的。数学与物理在研究方法上也是可以互相借鉴、互相补充的。事实上，不少大数学家同时又是杰出的物理学家（如高斯、庞加莱、闵可夫斯基等）；也有不少大物理学家同时又是杰出的数学家（如牛顿、麦克斯韦、洛伦兹等）。庞加莱正确地指出："物理科学不仅给我们（数学家）以解决问题的机会，而且也帮助我们发现解决问题的方法，它把这贯穿于两种途径之中：引导我们去预测问题的解，以及启示适当论证的线索。"

有无数生动的事例可以从微观的角度说明观察与实验在数学中的作用。当代著名的数学家、数学教育家G．波利亚(G.Ploya)在其经典性的世界名著《数学与合情推理》(中译本又名《数学与猜想》)中，对大量的这类事例进行了详尽的剖析，用以说明如何通过观察与实验、归纳与类比做出数学上的发现。其中最著名的典型事例有：哥德巴赫猜想（每个不小于6的偶数都可以表示为两个奇素数的和）、凸多面体的欧拉公式（在凸多面体中，面数加顶点数等于棱数加2）、阿基米德关于球体体积公式的力学证明、以及等周问题（周长相等的一切平面图形中，圆的面积最大）等等。波利亚一再把数学的研究方法与其他自然科学的研究方法做比较，指出它们在收集材料、进行观察与实验方面是完全类似的。波利亚强调数学来源于实际观察，不仅概念、定理、公式是由观察与实验归纳出来的，就连证明的方法也是如此。事实上，不仅仅是波利亚，很

多数学家都认为，数学既是演绎体系又是归纳体系，既是证明科学又是实验科学。

为了说明问题起见，让我们来看如下一个实例：

假设u是正奇数，研究x、y、z、w的不定方程

$$4u=x^2+y^2+z^2+w^2 \qquad (*)$$

有多少组正奇数解；换句话说，有多少种不同的方式可以把形如$4u$的自然数表示成（或分解成）四个正奇数的平方和？

先让我们来考察一个具体的例子——当$u=25$时，有多少种不同的方式可以把$4u=100$表示成四个正奇数的平方和？

由于这些正奇数的平方都必须小于100，所以可能的平方数只有如下五个

$$1,\ 9,\ 25,\ 49,\ 81.$$

若最大的平方数是81，则其余三个平方数的和就是

$$100-81=19$$

而小于19的奇数的平方只有9与1，且容易看出$19=9+9+1$，于是

$$100=81+9+9+1$$

只要变换其排列顺序（注意：我们把$100=81+9+9+1$与$100=9+81+9+1$等等只有排列顺序不同的也都看做是不同的分解方式，因为它们对应于x、y、z、w的不同取值），可以知道对于100而言，最大平方数为81的分解方式共有12种。

同样可知，对于100而言，最大平方数为49的分解方式共18种；最大平方数为25的分解方式只有1种。

如此，把100分解成四个正奇数的平方和的方式总共有$12+18+1=31$种。

用类似的办法，可以找出当$u=1$时分解方式有1种，当u

$=3$时分解方式有4种，当$u=5$时分解方式有6种等等．所得结果可以列成下表

u值	1	3	5	7	9	11	13	15	17	19	21	23	25
分解方式数	1	4	6	8	13	12	14	24	18	20	32	24	31

以上是属于数学研究中的"实验"，目的是为了从实验中找出与问题有关的数据。

然后，就要对所得的数据进行仔细的观察，以便发现数据之间的若干联系、总结出一定的规律．

我们看到，数据表中除1以外，下行的数总比上行相应的数大，而且有些下行的数比上行相应的数恰好只大1。再仔细看看那些下行数比上行数恰好大1的各栏，我们突然惊奇地发现这些栏中的u值居然都是素数（质数）！于是，我们会很自然地猜测：当u为奇素数p时，$4u$分解成四个正奇数平方和的方式有$p+1$种（为了更有把握，可以取u的另一些奇素数值如29、31等，来对这一猜测进一步加以验证，此处从略）。

数据表中其余各栏情况又如何呢？由于我们刚刚注意到了素数，容易看到下行数中也有两个素数，13与31；相应的上行数（u值）为9与25，恰是两个平方数．于是，我们又会猜测：当u为平方数时，$4u$的分解方式数为素数（不过，由于观察数目太少，又不方便验证，不宜对这一猜测抱太大的希望；事实上，这一猜测是错误的）。

现在只还剩下两栏，上、下行的数都是合数，试将它们分解因数，并观察对应的因数分解式间的关系，我们又看到了一种引人注目的关系

下行数中的每个因子都比相应的上行数中的素因子恰好大1！

u 值	……	3×5	……	3×7	……
分解方式数	……	4×6	……	4×8	……

联系到我们原先看到的p与$p+1$之间的对应关系，这里又看到了pq与$(p+1)(q+1)$间的对应关系（其中p、q均为素数），这难道仅仅是巧合吗？

如果我们把$(p+1)(q+1)$展开成 $pq+p+q+1$，上述对应关系的实质就更为明显了：pq、p、q、1 这些数不正好就是pq的所有的约数么！而p与1也正好就是p的所有约数！也就是说，$4u$的分解方式数似乎恰好是u的所有约数之和。

再检查一下两个完全平方数，9与25，所对应分解数 13 与 31也能写成类似的形式

$$13=9+3+1 \qquad 31=25+5+1$$

综合上述，我们得到了如下结果：

u 值	1	p	pq	9	25
$4u$分解方式数	1	$p+1$	$pq+p+q+1$	$9+3+1$	$25+5+1$

根据这一结果，我们可以比较大胆地猜测：当 u 为正奇数时，$4u$分解成四个正奇数平方和的方式数，等于 u 的所有约数之和。

只要再经过严格的证明，上述漂亮的猜测就可以成为一个坚实的定理（从略）。我们也就完成了数学研究中的一个小小的、但也颇为令人鼓舞的“发现”！当然，并非所有的猜测（或猜想）都能成为正确的命题（或定理），请读者参看本书“数学与猜想”一章。

“为什么数学真理如同物理科学领域中的定律和原理那

样，有时可以通过实验与归纳方法去发现呢？”我国著名的数学家、数学方法论的倡导者徐利治先生回答说，“原因很简单，因为数学对象本身（如数量关系与空间形式等）也具有客观实在性。”

在数学基础的研究中，主张数学应奠基于观察与实验的呼声也越来越高。希耳伯特企图在形式系统内部来解决数学理论真实性的方案的失败，说明数学理论的真实性归根结底还得依靠实践的检验。以著名的科学哲学家拉卡托斯（I.Lakatos）为首的一大批数学哲学家和数理逻辑学家认为，数学只有摆脱纯粹演绎观点的束缚，扎根于实践和经验之中，才能更加生机勃勃地不断向前发展。

近年来，由于电子计算机的出现和普及，就更为观察与实验方法在数学中的应用开辟了广阔的新领域。过去，观察与实验在数学领域中受到一些客观条件的限制，主要是数学中缺乏象物理、化学中那样帮助人们进行观察与实验的各种仪器设备。现在电子计算机扫清了这一障碍，为人们在数学中进行观察与实验提供了强有力的手段。人们已经通过电子计算机的数值试验发现了不少激动人心的数学事实，如非线性映象分支点的一些普遍规律和普适常数等等。一个崭新的数学分支——实验数学正在兴起。有人认为，用计算机做为数学实验、甚至数学证明的工具，可以使数学家们从一支笔、一张纸的“手工劳动”下解放出来，从而进入更高级的“机器生产”时代。

第二十七章

数学与猜想

数学上的猜想指的是由人们的直观或直觉上的初步判断认为可能成立，而又未经严格证明的命题；以及人们依据某些数学事实建立这种尚待证明的命题的创造性的思维活动过程。

当代世界著名的数学家和数学教育家波利亚（G.Polya）指出："数学被人看做是一门论证的科学，然而这仅仅是它的一个方面。以最后确定的形式出现的定型的数学，好象是仅含证明的纯论证性的材料，然而，数学的创造过程是与任何其它知识的创造过程一样。在证明一个数学定理之前，你先得猜测这个定理的内容，在你完全作出详细证明之前，你先得推测证明的思路。你先得把观察到的结果进行综合然后加以类比。你得一次又一次地进行尝试。数学家的创造性工作成果是论证推理，即证明；但是这个证明是通过合情推理，通过猜想而发现的。"

一个好的深刻的猜想，往往会成为数学家们长期研究的课题，成为推动数学不断向前发展的源泉和动力。数学史上曾经出现过许许多多的著名猜想，如费马猜想、哥德巴赫猜想、欧拉猜想、黎曼猜想、比勃巴赫（Bieberbach）猜想、四色猜想、连续统猜想、庞加莱猜想、希耳伯特猜想和魏尔（Weyl）猜想等等。这些猜想，有的已被证实，直接转化成了数学理论；有的已被推翻或部分地被推翻，但也在研究过程中产生出了许

多富有成效的新成果和新方法；有的尚未证实也未推翻，仍在激励着人们为之奋斗。

18世纪40年代，德国数学家哥德巴赫在把一些自然数拆成其它自然数之和的过程中发现：一个不小于6的偶数拆成两个奇数之和时，其中至少有一组是两个奇素数；一个不小于9的奇数拆成三个奇数之和时，其中至少有一组是三个奇素数，如

$6=3+3$，　$8=3+5$，　$10=5+5$，　$12=5+7$，

$14=7+7$，$16=3+13$，$18=5+13$，$20=7+13$，$22=3+19$，

$24=5+19$，　$26=3+23$，　$28=5+23$，　……

$9=3+3+3$，　$11=3+3+5$，　$13=3+5+5$，

$15=3+5+7$，　$17=3+7+7$，　$19=3+5+11$，

$21=5+5+11$，　$23=5+7+11$，　$25=5+7+13$，……

这是否是个普遍规律呢？哥德巴赫不敢肯定。1742年6月7日，他写信向当时世界上最有名的数学家欧拉请教，提出了以下两个猜想：

（1）每个不小于6的偶数都可以表示成两个奇素数之和；

（2）每个不小于9的奇数都可以表示成三个奇素数之和。

欧拉接信以后，经过实际的验算和证明的尝试，于同年6月30日回信给哥德巴赫说："任何大于6的偶数都是两个奇素数之和，虽然我还不能证明它，但我确信无疑认为这是完全正确的。"容易看出，由第一个猜想可以推出第二个猜想。这就是历史上著名的哥德巴赫猜想。这一猜想简单得说起来几乎人人都懂，欧拉也确信它无疑完全正确，但是却对它的证明无能为力，便引起了许多数学家的好奇和注意。不过，这个貌似简单的问题其实如此困难，以至直到19世纪结束时，数学家们甚至不知道如何着手进行研究。

1900年，德国大数学家希耳伯特把哥德巴赫猜想列为具有重大意义的23个问题中的第8个问题。1921年，英国数学家哈代（Hardy）曾说，哥德巴赫猜想的困难程度可以和任何没有解决的数学问题相比。从本世纪20年代起，随着数论上重要工具“筛法”的改进，哥德巴赫猜想逐步取得了一些进展和突破。目前最好的成果是由我国数学家陈景润1973年所取得的。陈景润证明了：每个充分大的偶数都可以表示为一个奇素数与另一个不超过两个奇素数乘积的奇数之和（即命题（1+2））。这距整个哥德巴赫猜想（命题（1+1））的彻底解决已经十分接近，有人甚至形容是“只差一步之遥”。尽管如此，猜想在未经严格证明之前毕竟也还只是猜想而已。

猜想不受现成事实的束缚，它包含着可贵的大胆想象和推测的成分，正因为如此，它才会对人们有着经久不衰的强大吸引力；也正因为如此，哪怕是最伟大的数学家的猜想也有可能会发生错误。前面我们曾经提到过，法国数学家费马曾断言形如$F_n=2^{2^n}+1$的费马数对$n=0$及一切自然数都给出素数，就是一个错误的猜想（请参看“数学的砖瓦——素数”一章）。这一错误是由欧拉所首先指出的。后来，欧拉本人也在正交拉丁方的问题上犯了类似的错误。

1779年，欧拉应普鲁士腓特烈大帝的请求，研究了一个由阅兵式产生的问题：有6个不同的师团，各选出上校、中校、少校、上尉、中尉、少尉6种不同军衔的军官各一人，能否把这36名军官排成一个6行6列的方阵，使得每行每列都有各个师团各种军衔的军官代表？

这个问题的一般提法也可以叙述为：有n个不同的拉丁字母和n个不同的阿拉伯数码，能否将这些字母和数码排成一个

n行n列（即n阶）的方阵，使得每行每列中的n个字母和n个数字都互不相同。因此，这个问题也称为正交拉丁方问题。

欧拉证明了任何奇数阶的正交拉丁方和任何4的整数倍阶的正交拉丁方都是可排的。以下是3阶与4阶的正交拉丁方：

$$\begin{array}{|ccc|}\hline A_1 & B_2 & C_3 \\ B_3 & C_1 & A_2 \\ C_2 & A_3 & B_1 \\ \hline\end{array} \qquad \begin{array}{|cccc|}\hline A_1 & B_2 & C_3 & D_4 \\ B_3 & A_4 & D_1 & C_2 \\ C_4 & D_3 & A_2 & B_1 \\ D_2 & C_1 & B_4 & A_3 \\ \hline\end{array}$$

但是对于6阶的正交拉丁方（如上述腓特烈大帝的36军官问题），欧拉却始终找不到答案。于是，他猜想2（2k+1）（其中k是自然数）阶的正交拉丁方是不可排的。欧拉说："我毫不犹豫地认为人们不可能造出一对6阶的正交表。同时对于10阶、14阶、……也不可能造出。一般地说，对任何奇数的2倍，都不可能造出。"直到1900年，塔里（Tarry）才用完全归纳法非常吃力地证明了欧拉的6阶正交拉丁方猜想是正确的。除此以外，在长达一百五十多年的时间里，欧拉的其它正交拉丁方猜想始终未获解决。

本世纪30年代，数理统计专家费歇耳（Fisher）发现：原来以为纯粹是来自智力游戏的正交拉丁方问题却可以用于试验设计的合理安排。例如，在三块不同土质的试验田上对三种不同品种的水稻，用三种不同的方法施加氮肥，所有可能的方案如果都做试验，共需要$3^3=27$次；但是如果按3阶正交拉丁方（其中的三列可以看做三块不同土质的试验田，拉丁字母A、B、C可以看做三种不同品种的水稻，阿拉伯数码1、2、3可以看做三种不同的施肥方法）来安排试验，只需要9次，就可以把最佳的水稻品种和最佳施肥方法确定下来。

这样一来，欧拉的正交拉丁方猜想就具有了直接的重大现实意义，促使人们更加努力地想方设法加以解决。第一个较为重大的突破是由印度几何学家玻斯（R·C·Bose）所首先取得的，他用伽罗瓦域GF（P^n）为坐标的有限射影几何，轻易地证明了n是素数时有n-1阶两两正交的拉丁方。1958年，美国数学家帕克（E.T.Parker）用群论和有限几何构造了21阶正交拉丁方。玻斯及其学生接着又证明了$n=2\times11$时正交拉丁方存在，此时的欧拉猜想不成立。帕克跟着又证明了$n=2\times5$时正交拉丁方存在，欧拉猜想也不成立。最后，玻斯及其学生终于证明了除$n=2$及$n=6$以外，都存在正交拉丁方，——欧拉对于正交拉丁方的猜想事实上仅当$n=6$时是正确的。但是这丝毫也没有降低欧拉做为一个大数学家的声誉，毕竟还是欧拉最先开始了对正交拉丁方的研究，而且还是他所做出的猜想激励和启发了后来的数学家们在这一问题上的发现。

伟大的科学家，微积分的创始人之一牛顿曾说："没有大胆的猜测就做不出伟大的发现。"数学家们正是在不断地提出猜想、研究猜想、解决猜想的过程中，推动着数学向前发展。

第二十八章

数学的简洁美

简洁是一条重要的美学标准，同时也是一条重要的科学标准，正如一句古老的拉丁格言所说："简单是真理的标志。"

那么，数学的简洁美表现在哪里呢？

一、数学问题的简洁美

数学问题为我们进行数学研究提供了源泉和动力，有人形象地指出：数学问题是"数学的心脏"。

对于爱好数学的人来说，数学问题有一种难以抗拒的诱惑力，有时甚至会影响和改变一个人的一生。最典型的一个事例就是，高斯17岁时解决了什么样的正p边形（p为素数）可以用尺规作图的难题，他是如此激动，以至立刻放弃了打算当哲学家的念头，而决心献身数学；否则，高斯也许就永远不会得到"数学王子"的光荣称号了。

笛卡尔的情况也有些类似。1618年，笛卡尔服役的军队正驻扎在荷兰的布莱达，使他有机会了解到该地公开征解的一个带挑战性的数学问题。笛卡尔被这个问题深深吸引，并且很快做出了正确的解答，使他确信了自己的数学能力，从此开始专心致力于数学研究。

康托尔认为在数学中设问的艺术比起解法更为重要。希耳伯特也说："问题的完美提法意味着问题已经解决了一半。"

一个好的数学问题为了突出主要和本质的东西，必然是简

洁的。而一个数学问题越提得简洁、清晰易懂，也就越容易引起人们的兴趣。所以希耳伯特把这一点做为衡量一个数学问题是否有价值的第一个标准。

翻开数学的历史，凡是经久不衰引人入胜的数学问题，如三大尺规作图问题、费马猜想问题、七桥问题、哥德巴赫猜想问题、黎曼猜想问题，连续统问题，科克曼女生问题、货郎担问题、……等等，除了理论上的重要与方法上的困难以外，无不是以极其简明而又深刻的表述方式吸引着人们的注意。多么象是引人垂涎欲滴的美丽果实，在诱使人们向它们伸出手来！而一旦伸出手去便欲罢不能。且不要说解决一个数学问题对于数学本身的推动如何，仅仅为了体验奋斗的愉快和胜利的喜悦，就已经足以使人们陶醉其中了。

二、数学概念的简洁美

“数学的语言不但是最简单和最容易理解的语言，而且也是最精炼的语言。”（H.L.Brougham）而数学概念则是数学语言的精髓。正是凭借由简洁的数学概念所组成的精炼的数学语言，才使得我们仅用寥寥数语，就能刻划出复杂现象的数学规律。

不少基本的数学概念已历经沧桑，内涵不断发生着深刻的变化。一个数学概念内涵的每一次变化，都使得这个概念本身更为清晰、准确，因而也就更为简洁。

请看“函数”这一概念内涵的不断发展变化。（见191页表）

“函数”的内涵一步步深入，外延也就一步步扩充。在现代函数概念这个精美的艺术品面前，17世纪的最初的函数概念简直成了粗陋的胚胎。怀特（White）指出：“复杂的东西是用简单的东西表示出来的。在某种观点下，数学可以定义为相

年　代	函　数　定　义
1673年 莱布尼兹	函数就是象曲线上的点的坐标那样随点的变动而变动的量。
1698年 伯努利	由变数 x 和常数所构成的式子，叫做 x 的函数。
1748年 欧　拉	函数为由一个变量与一些常量，通过任何方式形成的解析表达式。
1755年 欧　拉	如果某些量这样地依赖于另一些量：当后者改变时它经受变化，那么称前者为后者的函数。
1821年 柯　西	对于x的每一个值，如果y有完全确定的值与之对应，则y叫做x的函数。（注：此处的值系指数值）
近　代 维布伦	在变量y的集合与另一变量 x 的集合之间，如果存在着对于x的每一个值，y都有确定值与之对应这样的关系，那么变量y叫做变量x的函数。（注：此处的值可以是数、也可以是点或其它东西）
现　代	设A、B为非空集合，f是从A到B的一个对应法则，则A到B上的映射 $f:A\xrightarrow{f}B$称为A到B上的函数。

继用简单概念来代替复杂概念的科学。”

三、数学符号的简洁美

数学语言的一个重要特点就在于通过符号化而具有形式上的高度的抽象性，为我们认识客观现实世界的空间形式和数量关系提供了比普通的自然语言更为简明、准确的思维与表达的工具。

当我们看到欧拉用等式

$$e^{i\pi}+1=0$$

把数学上重要的常数0、1、i、π、e如此简洁而巧妙地联系在一起的时候；当我们看到爱因斯坦用公式

$$E=MC^2$$

把茫茫宇宙中的质能互换这样深奥复杂的关系如此简洁地揭示出来的时候，我们都会不由自主地赞叹简洁的数学符号帮助我们把人类智慧的精华凝炼成了光芒四射的美丽结晶。

广泛地使用数学符号是现代数学的特征。我们对于数学的"情感"，很大程度上取决于我们是否能够灵活自如地运用数学符号进行思维、推理和运算。

苏联数学家鲁金（Лузин）在谈到欧拉时说："欧拉的洞察力是那样深邃，不论多么复杂深奥的公式，在他强有力的手里，都变得服服贴贴，宛如柔软的蜂蜡一样；在他的威力面前都得规规矩矩地献出一切。……他可以本能地直接感觉到公式里的真理与虚假，他调注公式的技巧，对公式进行数量上的评估与变换的功夫，对结果的本质瞬间猜中的本领——这些都令人叹为观止。可以毫不夸张地说，在欧拉眼里，数学公式本身自始至终都充满了生命力，讲述着有关自然现象的最深刻的东西。只要他一碰到公式，就能使公式由'哑巴'变成会说话的人，并能作出饱含深邃含义的回答。"

当然，象欧拉这样的数学大师，毕竟只是少数。但是，掌握数学符号、理解数学公式却是我们每个学习或研究数学的人应该做到、而且是可以做到的，只不过掌握得不一定象欧拉那么娴熟、理解得不一定象欧拉那么透彻罢了。

如果我们体会不到数学符号在数学中所起的巨大作用，我们也就体会不到数学的美。

四、数学证明的简洁美

如果说数学概念是数学上的砖瓦，那么数学证明就是用逻辑的"粘合剂"将数学概念有机地结合起来而修筑成的数学宫殿。虽然数学上的发现并不直接依赖数学证明，但是，一个数学发现要想得到人们的普遍承认，就不能不通过证明使之成为既定的事实。

马丁·伽德纳（Martin Cardner）指出："数学的真谛在于不断寻求用越来越简单的方法证明定理和解答问题。"简洁的证明，看上去思路自然、条理清楚，显示出数学证明不容辩驳的逻辑力量，给人带来愉快和美的享受。而冗长复杂的证明，则使人感到牵强附会，难以接受。当然，冗长复杂的证明，在人们认识过程中的某些阶段，几乎是难以避免的。数学上不少举足轻重的著名定理的第一个证明，就都比较复杂。但是，人们对于简洁美的追求，绝不会满足冗长复杂的证明。因此，每个较为重要的证明出来之后，总有不少人致力于它的改进。而一个重大的改进往往需要在理论上和方法上有所突破，这就极大地推动了数学的蓬勃发展。

希耳伯特在谈到证明的简洁性与严格性之间的关系时说："把证明的严格化与简单化绝然对立起来是错误的。严格的方法同时也是比较简单、比较容易理解的方法。"

最近有个很有趣的事例。一位美国数学家德·贝兰治（Louis de Branges）花了三十多年功夫，好不容易证明出了连最优秀的数学家也感到为难的比勃巴赫（Bieberbach）猜想，可是却在美国数学界受到了冷遇。部分原因是因为他过去曾给出一个错误的证明，以致有些人仍对他抱有偏见；不过最主要的原因还是他的证明长达350多页，很难找到一个人愿意

仔细地审核。而且由于证明太长，难免会有些小错误，审核人往往遇到一两个哪怕无关紧要的小错误就再也没有耐心了。幸好德·贝兰治有机会到了苏联，并找到了知音。在苏联数学家的帮助下，证明终于简化为12页，他的成果才得到了广泛的承认和热烈的反响。

五、数学理论的简洁美

从欧几里德时代起，演绎的数学理论就成了大多数理论科学的典范，除了数学是发展得最早、最为成熟的带头科学以外，数学理论的高度简洁美也是一个极为重要的原因。无数的学者曾为数学理论的简洁美所倾倒。

牛顿认为，做为一个力学家必须具备一定的数学才能，这对于力学的数学化是必不可少的。力学家可以从数学形式的完美性中吸取丰富的营养，使力学理论从一开始就严格遵循健康的道路成长。牛顿所创立的经典力学，就是一个仿照欧几里得《几何原本》、并可以与《几何原本》相媲美的逻辑演绎体系。牛顿在《自然哲学的数学原理》中，首先提出了四条法则、八个原始定义、三条定律，然后利用演绎法简单明了得出了有关大至宇宙、小至沙粒运动状况的全部推论，完满地概括了物质世界极其丰富的机械运动现象，翻开了科学史上伟大的一页。

著名的德国物理学家、量子力学的奠基人之一海森堡(Heisenberg)曾对爱因斯坦说："我相信自然规律的简单性具有一种客观的特征，它并非只是思维经济的结果。如果自然界把我们引向极其简单而美丽的数学形式——我所说的形式是指假设、公理等等贯彻一致的体系——引向前人所未见过的形式，我们就不得不认为这些形式是真的，它们显示出自然界的

真正特征。……我坦白承认，我被自然界向我们显示的数学体系的简单性和美强烈地吸引住了。”

数学理论是丰富多彩的物质世界的变化规律的反映和概括。数学理论的简洁美正是物质世界自然的简洁美的表现，深刻地揭示出宇宙万物多样性的统一，令人神往和深思。

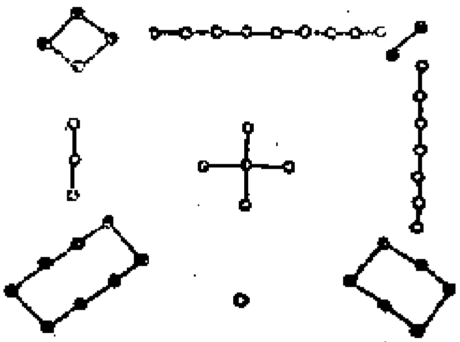
从另一方面来看，数学理论的简洁美也是人们把越来越丰富的数学知识不断地精简、提炼的结果。著名的英国数学家阿蒂亚（Atiyah）强调指出：“数学的统一性及简单性都是极为重要的，因为数学的目的，就是用简单而基本的词汇去尽可能多地解释世界。归根结底，数学仍然是人类的活动而不是计算机的程序。如果我们积累起来的经验要一代一代传下去的话，我们就必须不断地努力把它们加以简化和统一。”

第二十九章

数学的和谐美

自然界充满了美。从浩瀚无际的宇宙、灿烂多姿的天体到变化无穷的基本粒子，从奥秘的生命起源、神奇的生物进化到充满斗争的人类社会发展，万事万物运行有序，和谐统一，显示出一幅幅美妙动人的自然图景。数学美是自然美的反映，也以其和谐、有序而令人陶醉。

我们中华民族是最早懂得数学的和谐美的民族之一。标志着中华民族远古文化的“洛书”（传说最早记载于五、六千年以前，大禹治水时在洛水一带发现的刻有图纹的龟甲）至今仍



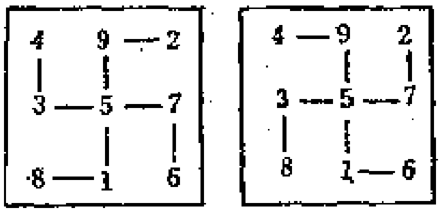
洛书

4	9	2
3	5	7
8	1	6

洛书今译

象一颗明珠一样闪烁着光辉，堪称数学史上一绝。“洛书”又称三阶幻方。幻者，奥妙无穷也。把它用当今的阿拉伯数码翻译过来，就成了由1到9的九个自然数按一定次序排成的三行三列的方阵。这个方阵中，数的大小分布与奇偶分布十分均称，其每行、每列以及两对角线上的三个数之和都是15；卍形两分支以

及卍形两分支上每五个数之和都是25；从2出发按逆时针方向可以依次得到2^1、2^2、2^3、2^4，从3出发按顺时针方向又可以依次得到3^1、3^2、3^3、3^4。洛书是组合数学的发端，是位置解析的先河，体现了许多数学、物理的原理和功能。有人用它研究博奕对策，有人用它来研究非均匀物质分布的平衡，也有人用它来研究自动控制的网络设计。早在四千多年以前，我国就有洛书包含了治国之道的说法。



15	10	3	6
4	5	16	9
14	11	2	7
1	8	13	12

洛书是唯一的三阶幻方（不包括它的旋转和反射象），还有四阶幻方（共880种）、五阶幻方（共275305224种之多），甚至任意阶的幻方。有一种被人称为“具有灵魂”的四阶幻方，不但每行、每列以及两对角线上的四个数之和都是34，而且随便任意圈出一个包含四个数的正方形，其中的四个数之和也都是34。

数学上的和谐首先意味着秩序和规律。控制论的创始人维纳说：“数学的使命就是在混沌之中去发现秩序。”乍看起来，大千世界变幻无穷，很多事物令人感到眼花缭乱、难以捉摸。可是一经用数学去分析、研究，最终总能发现支配这些事物发展变化的极其简单明了的规律。从这个意义上说，数学正是以自己的和谐美展现了自然的和谐美。如玻尔茨曼定律，以极其简明的数学式子

$$S=K\ln W$$

表述了一个深刻的物理规律：熵与状态几率的对数成正比。这个数学式子把气体分子杂乱无章的随机运动，纳入了整个自然和谐有序的图式。

数学本身有时也会遇到秩序和规律似乎受到破坏的情况。请看两个典型的无理数：

$$\sqrt{2}=1.414213562373095\cdots\cdots$$

$$\pi=3.141592653589793\cdots\cdots$$

它们是无限不循环小数，无论到小数后多少位，都不会呈现整体上的规律。难怪深信“万物皆数”和“宇宙和谐”的毕达哥拉斯学派对无理数是如此惊慌失措，以致要把它的发现者匆匆投入大海；毕达哥拉斯学派虽然还不懂什么无限不循环小数，仅从无理数不能表示成两个整数之比这一点上，他们就已经感到了它是不祥之物。然而，只要我们把$\sqrt{2}$、π这些无理数展开成无穷级数或连分数，它们就立刻呈现出令人惊奇的简单规律：

$$\sqrt{2}=1+\frac{1}{2}-\frac{1}{2\times4}+\frac{1\times3}{2\times4\times6}-\frac{1\times3\times5}{2\times4\times6\times8}+\cdots\cdots$$

$$\sqrt{2}=1+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cfrac{1}{2+\cdots\cdots}}}}$$

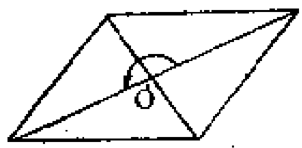
$$\pi=4\left(1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\cdots+(-1)^{n+1}\frac{1}{2n-1}+\cdots\cdots\right)$$

$$\pi=\cfrac{4}{1+\cfrac{1^2}{2+\cfrac{3^2}{2+\cfrac{5^2}{2+\cfrac{7^2}{2+\cdots}}}}}$$

这足以使我们相信，表面上看起来似乎无规律的东西，只不过是暂时还没有找到可以表现出其规律的数学表达式罢了。

其次，数学上的和谐也意味着对称。

所谓对称，就是在某种变换下保持不变的性质。例如平行四边形绕其中心旋转180°后与其原位置重合，看上去就好象根本没有旋转一样（此即"中心对称"）；又如，代数式$a^3+b^3+c^3-3abc$，把a换成b、b换成c、c换成a以后，所得的新代数式$b^3+c^3+a^3-3bca$仍与原代数式恒等(此即"轮换对称")。



数学上的不少研究对象，如对称图形、对称多项式、对称矩阵；对称变换、奇偶函数、周期函数、向量内积、等价关系、等等，性质都直接与对称性有关；数学上的许多成对相反概念，如奇与偶、正与负、实与虚、有限与无限、常量与变量、连续与间断、收敛与发散、等等，也来源于或体现了对称性；数学上的大量公式和定理，如海伦公式、平均不等式、赫尔德不等式；二项式定理、韦达定理、莱布尼兹公式、分部积分公式等等，本身的形式也都具有某种对称性。从研究代数方程的根的对称性而发展起来的群论，已经成了数学、物理、化学以及其他科学研究事物对称性的有力工具。

数学上的对称反映了大自然与宇宙的对称：基本粒子、原子、分子的结构，晶体、蛋白体的空间点阵排列，雪花、露珠、花瓣、树叶的形状，动物躯体、地球和其它天体的外观，天体的运行轨迹，电荷的吸引与排斥，化学反应的化合与分解，生物的新陈代谢，……，无不显示出优美和谐的对称。自然界中的很多对称性被科学家们表述成了这样或那样的守恒定律，而这些众多的守恒定律又与数学上的某种对称性相联系，例如：动量守恒与坐标平移的对称性相联系，角动量守恒与坐标旋转的对称性相联系，能量守恒与时间运动的对称性相联系，等等。

当然，与事物原始、自然的对称相比，数学上的对称已经经过了严格的多层次的抽象，更能揭示出事物对称的实质，所以倍受人们青睐。在“料事如神的数学预言”一章中，我们谈到过英国物理学家狄拉克曾根据数学的推导预言了正电子的存在。狄拉克开始也非常犹豫，对于提出一种尚未发现的新粒子感到没什么把握。当时狄拉克唯一的依据就是从电子运动的波动方程得出的正、负两个解，正解对应于当时已知的带负电荷的电子，那么负解对应于什么呢？必须找到一种合情合理的解释。数学家魏尔支持他提出正电子的假说，认为数学的对称美是应当遵守的。这使狄拉克从数学的对称美的角度重新考虑了自己所得的正、负两个解，终于做出了存在正电子的大胆预言，几年之后就得到了实验物理学家的证实。狄拉克后来在谈到理论物理学的方法时说：“理论物理学家把数学美的要求当作信仰的行为。它没有什么使人非相信不可的理由，但过去已经证明了这是有益的目标。”

数学上的和谐还意味着数学本身理论的协调、统一。

早从欧几里得时代起，追求数学本身理论的协调、统一就已经是数学研究的目的之一。在布尔巴基鼎盛时期，这种追求更是到了登峰造极的地步。除了刚刚开辟的新的数学领域，数学家们总是想方设法把数学理论整理得井井有条。试看欧氏几何、罗氏几何与黎曼几何（参看《离经叛道的非欧几何》一章），哪一个不具有坚实而又简单的逻辑基础？哪一个不具有硕果累累的丰富内容？虽然其公理系统互不相同，个别公理甚至截然相反，但是它们却又是相对相容的（例如，可以在欧氏空间中构造出罗氏几何模型，在罗氏空间中也可以构造出欧氏几何模型）。其脉络的清晰、结构的严谨、相反相成的互为补充，不正是整个数学理论协调、统一的缩影吗？

布尔巴基从结构主义的立场出发（参看"从公理化到结构主义"一章），用全局的观点分析和比较了各个数学分支结构的差异和内在联系，并按照结构的不同对数学重新加以分类和安排，力图将数学整理成博大精深、井然有序、渊源统一的理论体系。他们之所以要这样做，一方面是想创造出一批有实用价值的理论工具，另一方面也是为了追求数学理论的内在和谐、形式的漂亮。尽管受到某些人的责难，但是应该承认，布尔巴基的大部分工作还是干得十分出色、相当成功的。

法国大数学家庞加莱说："数学家非常重视他们的方法和理论是否优美，这并非华而不实的做作。那么到底是什么使我们感到一个解答、一个证明优美呢？那就是各部分之间的和谐、对称，恰到好处的平衡。一句话，那就是井然有序、统一协调。"

美国著名的数学家和数学教育家G.波利亚也认为："和谐的简单次序不会骗人这样一种感觉，在数学及其它科学领域指

引着做出发现的人们。"他结合一个十分浅显的例子说道：

"均匀棒的重心与其两端的重心相重合。均匀三角形的重心与其三顶点的重心相重合。为什么我们不应设想均匀四面体的重心与其四顶点的重心相重合呢？

还有，均匀棒的重心按比例1:1来划分其端点间的距离。均匀三角形的重心按比例2:1来划分任何顶点与其对边中点间的距离。为什么我们不应该猜测均匀四面体的重心是按比例3:1来划分任何顶点与其对面的重心间的距离呢？

说上述问题所提出的猜测是错误的，说这样美妙的一种规律性竟遭破坏，这点总叫人觉得极不可能。"

"那里有数学，那里就有美。"美国数学家哈尔莫斯把数学视为创造性的艺术，"数学是创造性的艺术，因为数学家创造了美好的概念；数学是创造性的艺术，因为数学家象艺术家一样地生活，一样地工作，一样地思索；数学是创造性的艺术，因为数学家这样对待它。"

第三十章

成果诚可贵　方法价更高

数学的成果，是人类知识宝库中难以比拟的巨大财富；而获得这些成果的方法，则更是人类智慧领域中的无价之宝。

色格勒（C.Segre）指出："一个结果的重要性在很大程度上是相对的，对它的判断因人而异，并随着时间和环境发生变化，往往由于问题困难而认为它非常重要。的确，如果说为了解决问题必须发明新的方法；优越的技巧等等，那么使科学受益较大的，正是这些方法和技巧，而不是最终的结果。"

前面我们曾经谈到过古希腊数学遗留下来的"三等分角"、"化圆为方"和"倍立方体"的三大尺规作图问题，这些问题已经被证明是不可能的了(请参看"神奇的对数与RMI原则"一章)。单就这些尺规作图问题本身而言，无论是可能还是不可能，对我们并无多大影响。即便现在已经知道是不可能了，充其量也不过是避免了一部份人毫无必要地再把时间和精力白白花在对这些本来就不可能的问题的钻研上。但是从解决这些问题的方法上来看，却使我们受益不小，至少我们明白了以下三点：

（1）几何作图问题的可能与不可能，取决于作图时的工具和规则，如果允许改变作图工具和规则，那么很多原来不可能的作图问题就将变成可能的。人们正是在钻研三大尺规作图问题的过程中，创造出了很多巧妙的作图工具。当然，如果当

初古希腊数学家没有对几何作图提出尺规的限制，那么我们今天就要失去很多富有价值的成果了。

（2） 当我们长期考虑一个问题始终得不到解答的时候，很可能是因为我们的思路不知不觉地陷入了某种固定、僵化的模式，在这种情况下就应该换一个角度重新考虑，甚至从相反的角度来考虑问题是否真的有解。起初谁也没有怀疑三大尺规作图的可能性，经过一代又一代人的漫长摸索之后，人们才开始不情愿地领悟到这些作图问题也许是根本不可能的。虽然在这一摸索过程中也产生了一些值得怀念的副产物，但与人们付出的代价相比毕竟有些得不偿失。

（3） 几何问题可以转化成代数问题求解（反过来，我们自然也应该想到代数问题也可以转化成几何问题求解）。在不改变问题实质的条件下，促使问题的形式发生转化，从繁杂的难于处理的形式，转化成简单的便于处理的形式，这就是RMI原则的精髓。不仅仅对数学，对其它科学也是同样适用的。如果不是笛卡尔创立了解析几何，使得人们能够把三大尺规作图问题转化为代数方程的求根问题，也许这些作图问题至今仍然无法解决呢！

正是因为我们从方法论上得到了这些足以铭记在心的启示，才使得我们津津乐道于这一话题，借以不断告诫我们自己并教育子孙后代。

数学上一个重大而困难的问题的解决，往往首先取决于方法上根本性的突破。

大家知道，解代数方程是古典代数的主要内容。早在很古的时代，数学还刚刚萌芽不久，人们就已经懂得了一次与二次代数方程的解法。公元前两千年左右的古巴比伦人解二次方程

的配方法与我们今天用求根公式求二次方程的根并无多大的本质区别。16世纪，意大利数学家又给出了三、四次代数方程的求根公式。这使人们看到：五次以下的代数方程的根总可以根据所给方程各项的系数，经过有限次的加、减、乘、除、乘方和开方运算精确求得（即存在所谓公式解或代数解）。很自然地，人们试图继续寻找五次和五次以上的代数方程的公式解。从16世纪中叶到18世纪末叶，许多数学家经过两百多年的努力，始终毫无结果。特别令人惋惜是，尽管很多人（包括一些有名的大数学家）都在这个问题上碰了壁，然而却从未怀疑过公式解的存在（这与三大尺规作图问题是何等相似！所以我们要把“从相反的角度考虑问题是否真的有解”列为我们必须铭记在心的教训）。

直到18世纪70年代，五次和五次以上代数方程的公式解长期得不到解决的问题引起了法国数学家拉格朗日（Lagrange）的深思。拉格朗日认为，这主要是因为当时的人们受配方、代换等传统代数方法的影响太深，没能从根本上抓住代数方程的根与系数的关系，就是著名的韦达定理也未能揭示出根与系数更深一层的本质联系。拉格朗日深入地研究了代数方程的根的对称多项式在置换下保持不变的性质，发现前人解二、三、四次代数方程的方法尽管看起来千差万别，却都可以在根的对称多项式的基础上用置换理论统一起来。这个统一的理论在寻找五次方程的公式解时所遭受的巨大困难，开始使拉格朗日意识到一般的五次代数方程的公式解很可能并不存在。1824年，挪威青年数学家阿贝尔（Abel）沿着拉格朗日的方向，用根的置换理论证明了五次代数方程的一般公式解的确不存在。稍后，法国青年数学家伽罗瓦（Galois）又进一步发展了拉格朗日与

阿贝尔的思想。伽罗瓦发现：每个代数方程都有一个反映其根的对称性的置换群存在，判断一个代数方程是否存在公式解的问题，可以归结为研究相应的置换群性质的问题。1831年，伽罗瓦终于通过对置换群性质的研究，彻底解决了代数方程的公式解问题。伽罗瓦给出了代数方程公式解的判别法则，可以判断任何一个具体的代数方程什么情况下存在公式解，什么情况下不存在公式解。

至此，代数方程的公式解（或代数解）的问题，用戏谑的说法就是已经"死"了，不值得我们再去耗费任何精力了。然而，为了解决这一问题而在置换理论上发展起来的群论，却成为数学家们研究事物对称性的有力工具而迅速渗透到了数学的各个方面，今天已经成了近世代数最重要的基本内容之一。

再让我们来看一个与电子计算机的算法有关的素数判别问题：

在很多纯数学和实际问题中（例如破译一种借助电子计算机编制的数论密码），需要判别一些较大的自然数是否是素数。然而，判别一个自然数是否是素数，历来是数学中最困难的问题之一。本世纪80年代以前，即使借助高速电子计算机，判别一个较大的（例如几十位或上百位）自然数是否是素数，都有可能要花成千上万年的时间才能做到。这并不是因为判别一个数是否是素数缺乏一个一般方法，而是因为方法虽然从理论上行得通，但在实际运用时困难太大。

以往判别素数的方法主要有两种：

一种是"试除法"。要判别自然数n是否是素数，需要用不超过$\sqrt{n}$的所有除1以外的自然数（也可以改进为不超过$\sqrt{n}$的所有素数）来试除n，如果它们都不能整除n，则n就是

素数；而只要它们当中有一个能整除n，则n就不是素数。

一种是"筛法"。要判别自然数n是否是素数，需要将从2到n的全部自然数列出，先筛去其中所有2倍数，再筛去其中所有3的倍数，……，如此不断地往下筛，如果在哪一步n被筛掉了，它就不是素数；如果直到最后n仍未被筛掉，则n就是素数。

试除法的计算量十分庞大，当n是接近40位素数时，用相当高速的计算机昼夜不停地工作，试除时间都要长达100万年以上；而筛法则要计算机拥有难以想象的巨大贮存装置，速度也快不了多少。

为了解决这一困难，数学家们经过长期的研究之后，决定抛开传统的试除法和筛法，采用一种新的两步走的方案：首先判断n是否整除$2^{n-1}-1, 3^{n-1}-1,$ …$(n-1)^{n-1}-1$。根据法国数学家费马（Fermat）所发现的一个定理（费马小定理），若n是素数的话，n应该能够整除$a^{n-1}-1(a=2, \cdots, n-1)$；而$n$是否整除$a^{n-1}-1$并不需要去做真正的除法，用数论上的同余式和模的理论，可以把它转化成十分简单的运算。其次，当n通过了整除$a^{n-1}-1(a=2, \cdots, n-1)$这一关时，还不能立即判断$n$就是素数，因为其中还可能混有诸如$n=561$这样的合数，这种合数称为"伪素数"，幸而伪素数所占比例极小（大约只占三万分之一），数学家们又找到一种迅速剔除伪素数的方法。于是，借助电子计算机来实现这两个步骤，便可以很快地判别一个给定的自然数n是否是素数。

新方法问世以前，用高速电子计算机判别一个百位数是素数可能要花百万年以上的时间（这在实际上是根本办不到的），而用新方法只需要15秒；由此可见方法的改进是何等的

重要：

汉克尔（H.Hankel）形象地比喻说："如果我们把一个数学问题，比做一个我们要了解其内部秘密的巨石，那么，在我们看来，古希腊数学家的工作就好象是一个不屈不挠的健壮石工，利用锤子和凿子从外部慢慢地剖析岩石，而今天的数学家却象一个熟练的矿工，先在岩石上打些眼，然后用炸药把它炸开，从而获得其内部的珍宝。"