科学新悦读文丛

有脑力之人钻研之学问

数学不简单

从《最强大脑》发现思维乐趣

吴悦辰
编著

Mathematical Games

中国工信出版集团　人民邮电出版社 POSTS & TELECOM PRESS

科学新悦读文丛

数学不简单　从《最强大脑》发现思维乐趣

吴悦辰　编著

人民邮电出版社

图书在版编目（**CIP**）数据

数学不简单：从《最强大脑》发现思维乐趣/吴悦辰编著.--北京：人民邮电出版社，2019.2

（科学新悦读文丛）

ISBN　978-7-115-49982-0

Ⅰ.①数…　Ⅱ.①吴…　Ⅲ.①数学—思维方法—普及读物

Ⅳ.①01-0

中国版本图书馆CIP数据核字（2018）第253168号

◆编著　吴悦辰

责任编辑　刘朋

责任印制　陈犇

◆人民邮电出版社出版发行　　北京市丰台区成寿寺路11号

邮编　100164　　电子邮件　315@ptpress.com.cn

网址　http://www.ptpress.com.cn

大厂聚鑫印刷有限责任公司印刷

◆开本：700×1000　1/16

印张：9.25　　2019年2月第1版

字数：105千字　　2019年2月河北第1次印刷

定价：39.00元

读者服务热线：**(010)81055410**　印装质量热线：**(010)81055316**

反盗版热线：**(010)81055315**

广告经营许可证：京东工商广登字**20170147**号

目　录

内容提要

“数学是上帝用来书写宇宙的文字”蕴含在生活中的各个角落，越靠近它，你就越能体会到它的不简单之处。本书精选了《最强大脑》节目中的热门项目，详细剖析了这些烧脑问题背后的数学知识并加以扩展。数字华容道的排列问题，立体一笔画的解链，迷宫中的拓扑知识，繁花规图案的摆线方程，数独的设计与求解……这一系列有趣的问题不仅可以加深你对数学的了解，还能开发智力、活跃大脑。

本书适合喜欢数学的读者阅读。

前言

《最强大脑》是江苏卫视借鉴德国节目Super Brain推出的国内首档大型科学类真人秀节目，专注于传播脑科学知识，倡导脑力竞技，自2014年第一季开播以来便备受关注。

《最强大脑》节目组推出的一系列竞技项目，着重考察选手的六大能力：观察力、空间想象力、计算能力、推理能力、记忆力和创造力。不少竞技项目都蕴含着丰富的数学知识。为此，笔者选择了一些项目，来介绍它们的数学背景及数学思想。

要问数学是什么，我们可以给出许多不同的答案，诸如数学是定义、公理、定理、公式、法则等的集合，数学是关于数和形的科学，数学是符号游戏，数学是思维体操，等等。不过，上述每一种说法都有它的局限性。数学是一门抽象学科，具有多重性和复杂性，它源于人们对生活的思考，以解决实际难题。

拜占庭哲学家普罗克洛斯说过：“数学是这样一种东西：她提醒你有无形的灵魂，她赋予她所发现的真理以生命；她唤起心神，澄净智能；她给我们的内心思想添辉；她涤尽我们有生以来的蒙昧与无知。”所以，数学是一种文化，一门充满了人文科学风采和自然哲学味道的崇高科学。

在当今时代，大到卫星导航、无人驾驶，小到移动支付、人脸识别，它们的背后无不闪烁着数学的光彩，所以说“数学不简单”是真的不简单。愿读者们能真心地喜欢数学，品味数学，欣赏数学，应用数学。数学让大脑更灵活，让思维更敏捷，让生活更美好！

最后，要感谢本书的编辑李宁，没有她的鼓励与支持，我这近八旬的老者实难完成本书。

吴悦辰

2018年10月于湖北

第一章 数字华容道

2018年元月5日，《最强大脑》第五季开幕！第一期第一轮就展开了百人大战，竞赛项目是“数字华容道”。

在分成16个小方格的盒子里装着15块标有数字的小方块，还留下一个空格。按任意顺序把小方块放进盒子里，空格确定在右下角。要求通过移动，把它们按照自然顺序一一排好。图1-1是初始状态的一个例子，图1-2是要求达到的终局状态。

2	1	3	4
5	6	7	8
9	10	11	12
13	15	14	

图1-1

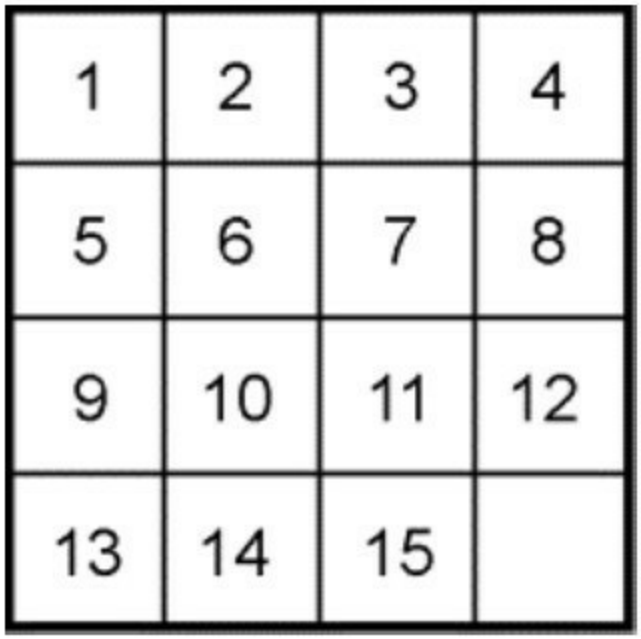


图1-2

在同一初始状态下，百位选手展开了激烈的角逐，最先完成并正确无误的选手得以晋级。

重排十五

数字华容道又名重排十五，是一个流传已久的单人智力游戏。这款游戏有两个基本问题：一是如何去解，二是是否一定有解。

这个游戏的主流玩法是降阶法（如图1-3所示），先还原1,2,3,4,5,9,13这7枚棋子，把它们分别安置在第一行和第一列上。

1	2	3	4
5	12	10	
9	8	14	7
13	11	15	6

图1-3

第二步是把6,7,8,10,14这5枚棋子安置在第二行和第二列的自然顺序位置上（如图1-4所示）。

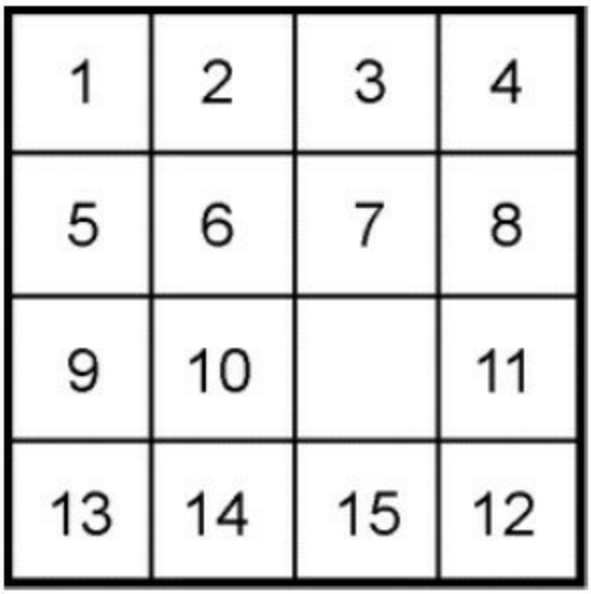


图1-4

现在，只剩下11,12,15这3枚棋子和右下角的1个空格了。这时会出现两种情况，一是11,12,15顺时针排列（不管空格在何处），二是11,12,15逆时针排列（也不管空格在何处）。如果是第一种情况，那么最多再走两步就可以到达终局了，如图1-5所示。如果是第二种情况，则对不起，你摊上大事了！即使你让临近的10、14两枚棋子一起参与进来，想尽法子，到何年何月也不可能达到终局。

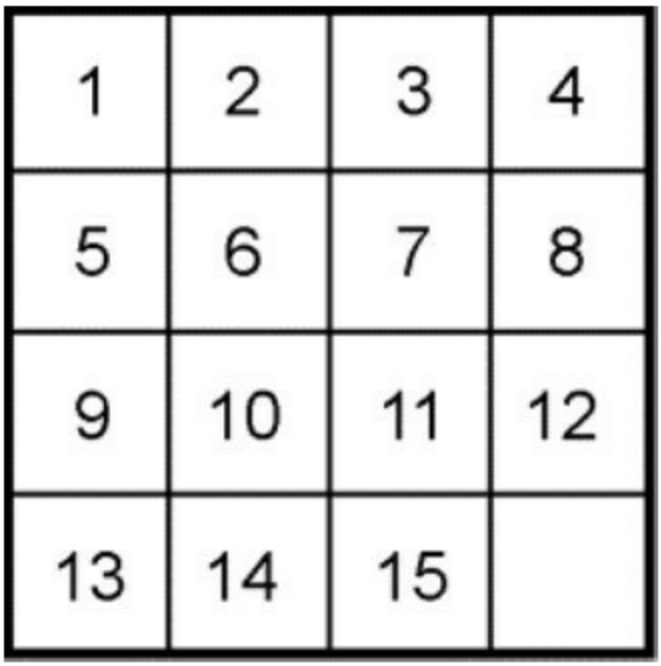


图1-5

下面用数学方法回答：重排十五什么时候是可解的，什么时候是不可解的。先看下面几个定义。

定义1：由1,2,…,*n* 组成的一个有序数组称为一个*n* 级排列。

例如，2431是一个四级排列，45321是一个五级排列。

定义2：在一个排列中，如果一对数的前后位置与大小顺序相反，即前面的数大于后面的数，那么它们就称为一个逆序。一个排列中逆序的总数就称为这个排列的逆序数。

例如2431中，21,43,41,31是逆序，2431的逆序数就是4；类似地，45321的逆序数是9。

定义3：逆序数为奇数的排列称为奇排列，相应地，逆序数为偶数的排列称为偶排列。

显然，任一排列都只有这两种可能。

现在回到重排十五上来，15枚棋子放在16个空格中。按从左到右、从上到下的顺序（空格不计）排成一列，因此也构成了一个排列，如图1-6所示的初局。如果写成排列的话就是

7	3	2	9	11	1	5	4	15	6	13	10	12	14	8
6	2	1	5	6	0	1	0	6	0	3	1	1	1	0

横线以下是对应数字的逆序数。例如7，在此排列中有3,2,1,5,4,6这6个数比它小，所以7的逆序数是6；又如5，在它之后仅有4比它小，所以5的逆序数就是1，以此类推。在判断这个排列是哪一类排列时，不必把这15个逆序数加起来，只需数一数有多少个奇数就可以了。在本例中，共有7个奇逆序数，所以它是一个奇排列。

7	3	2	9
11		1	5
4	15	6	13
10	12	14	8

图1-6

以下我们要研究棋子的移动对于排列种类的影响。棋子左右移动，排列的奇偶性不受影响；而上下移动的话，奇偶性将要发生变化。如图1-7所示，棋子丁向上面的空格移动，则由原来的排列“×甲乙丙丁××”变成了排列“×丁甲乙丙××”，也就是说丁跳过甲乙丙而成了新的排列。由此所导致的排列种类的变化很容易看出来，因为丁一次性跳过甲乙丙可以看作先跳过丙、再跳过乙、最后跳过甲的三级跳。每跳一次，就有一

个新的逆序产生，或一个旧的逆序消失。无论是哪种情况，每跳动一次的结果总是逆序数加1或减1。即跳动1次，排列的奇偶性变动一次；跳动3次，排列的奇偶性也连变3次。总结起来就是，一个棋子向上移动时，排列的奇偶性发生变化；反之，一个棋子向下移动时，奇偶性当然也要发生变化。

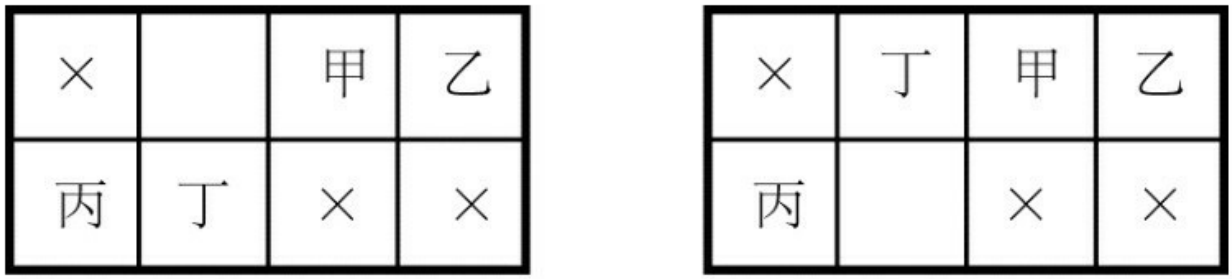


图1-7

现在假定游戏开始了，空格的位置在开始时照例放在棋盘的右下角。不管棋子移动多少次，空格的位置依旧在右下角才有可能达到终局。这样空格上移的次数必定与下移的次数相同，因此排列的奇偶性一定经过偶数次的变化，所以移动开始时初始排列的奇偶性必与其移动终了后的奇偶性是一样的。换言之，偶排列仍回到偶排列，而奇排列仍回到奇排列。但我们终局的要求是1,2,3,4,…,14,15的排列，逆序数为0，是个偶排列，这样我们就得到了一个重要结论：

初始开局是奇排列的重排十五一定无解，只有初始开局是偶排列的，才能有完美的终局。

在随机放置15枚棋子时，有多少情况是无解的呢？在一个n 级排列中，排列总数是n ！设全部的n 级排列中有s 个奇排列，t 个偶排列。现将s 个奇排列中的前两个数对换，得到s 个不同的偶排列，因此$s \leq t$ 。同样可证$t \leq s$ ，于是$s = t$ ，即奇、偶排列的总数相等，各有n ！/2个。所以，若15枚棋子随机放置的话，就可能有一半是无解的。

任何游戏之所以有趣，是因为无数次摸索尝试后，突然成功的那一

刻的兴奋。现在，通过数学上的排列论，可以事先预测其结果，于是一个有趣的游戏变得兴味索然了。有人说这是数学的缺点，但换个角度看，这正是数学的伟大。数学大的功用在于解释宇宙的奥秘，发掘真理，最后使得人类能充分利用这千变万化的宇宙现象；小的妙处就像我们可以用排列组合解决重排十五的问题一样。

最后有一点要声明：游戏终局时，空格的位置必须与初始状态一样在右下角，否则以上讨论的问题就完全改变了。

“14～15”游戏

历史上曾发生过一段有趣的故事。美国科学魔术师萨姆·洛伊德在1878年推出了一款著名的“14～15”智力玩具，这个游戏迅速风靡欧美大陆。在德国的马路上、工厂里、皇宫中、国会大厦，到处都有人在如痴如醉地玩这个游戏，以致许多工厂老板不得不出示公告，禁止人们在上班时玩游戏，否则开除！在法国，从比利牛斯山脉到诺曼底半岛的小山村和巴黎的林荫大道，也是处处可见玩这个游戏的人群。这个玩具在出售时的初始布局如图1-8所示，玩具的制作者洛伊德承诺，谁能够通过滑动其中的数块使错位的14和15换成正常次序，谁就能获得1000美元的奖金。1000美元可不是一个小数目，自然会引起轰动。事实上，这是永远无法达到的，只是洛伊德的一个“诡计”，因为这正是一个奇排列。

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

图1-8

但从图1-8所示的初始布局，还是可以得到一些有趣的终局。如图1-9所示，即把数字1～15的次序理顺，但空格不在右下角，而在左上角。目前已知这个玩法的最少步数是44步，如下所示：

14,11,12,8,7；6,10,12,8,7；4,3,6,4,7；14,11,15,13,9；12,8,4,10,8；4,14,11,15,13；9,12,4,8,5；4,8,9,13,14；10,6,2,1。

	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

图1-9

重排九宫

在分成9个小方格的盒子里，将1～8这8个数字任意安排，余下一个空格，与空格上、下、左、右相邻的数字可以移动。要求最后的数字从1至8按顺序排好，空格居于右下角。如图1-10所示，这就是重排九宫游戏。

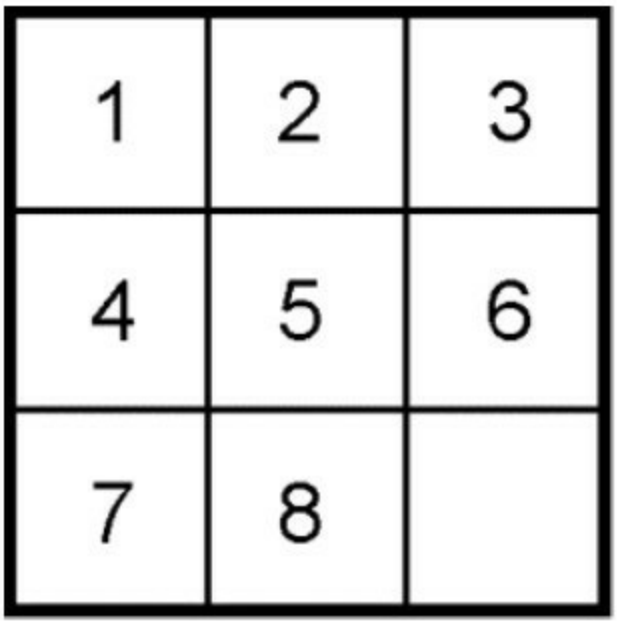


图1-10

重排九宫是否有解？其准则和重排十五是一致的。在重排九宫有解的情况下，人们又欲寻找以最少的步数完成从初始状态到终局状态的转换。这是相当有难度的问题，例如求欲使图1-11中的数排成图1-10的顺序，其最少步数是多少？

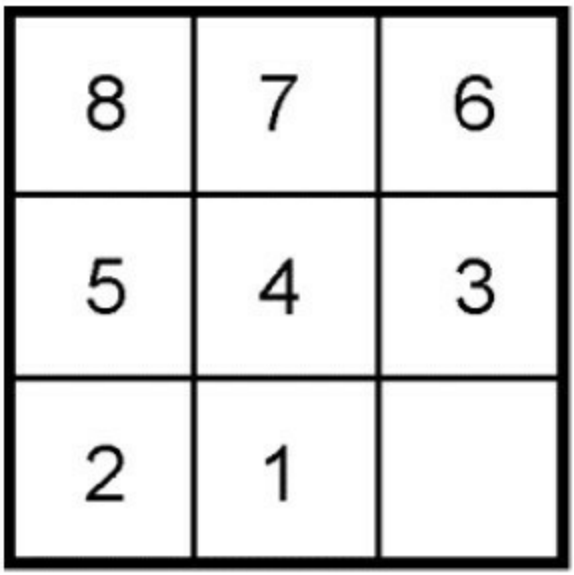


图1-11

19世纪，英国趣味数学家杜丹尼找到一种36步的解答方法：12543，　12376，12376，12375，48123，65765，84785，6。该方法长期以来被认为是最佳答案，没有人提出质疑。直到后来借助于电子计算机，一下子找到10种30步的解答方法，才打破了杜丹尼的记录。这些解法如下：

① 12587，43125，87431，63152，65287，41256；

② 14587，53653，41653，41287，41287，41256；

③ 14314，25873，16312，58712，54654，87456；

④ 12587，48528，31825，74316，31257，41258；

⑤ 14785，24786，38652，47186，17415，21478；

⑥ 34785，21743，74863，86521，47865，21478；

⑦ 34785，21785，21785，64385，64364，21458；

⑧ 34521，54354，78214，78638，62147，58658；

⑨ 34521，57643，57682，17684，35684，21456；

⑩ 34587，51346，51328，71324，65324，87456。

华容道

华容道游戏与重排九宫和重排十五类似，属于滑块类游戏。标准华容道是4×5的长方形棋盘，共有10枚棋子。其中，最大的一块是2×2，代表曹操；2×1的有5块，4块竖排的分别代表张飞、赵云、黄忠和马超，一块横排的代表关羽；还有4块1×1的代表卒。此外，盘面还空出2个单位面积，作为游戏时的滑动之路。游戏要求滑动棋子，使曹操移动到下端正中间的开口处逃脱（如图1-12所示）。这个故事来源于中国古代的文学名著《三国演义》。

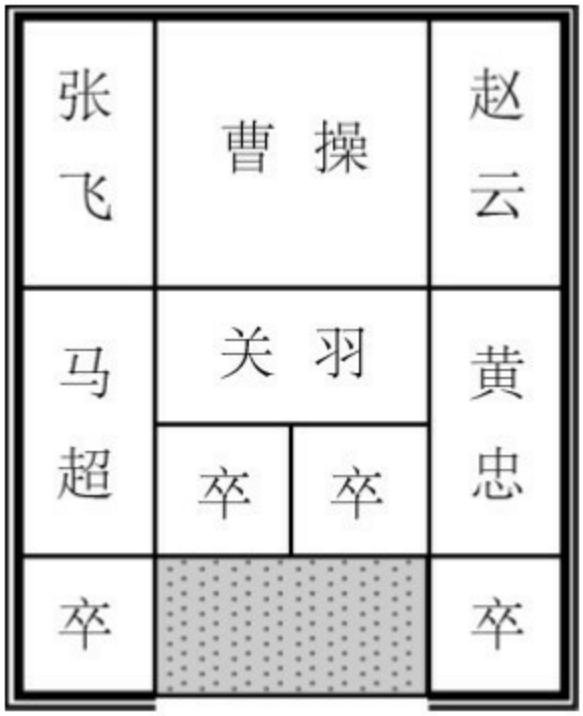


图1-12

图1-12所示的是典型的“横刀立马”开局，研究其解法的最小步数长期是热点问题。1964年，美国马丁·加德纳在《科学美国人》上公布了

81步的解法。后来经过计算机的验证，确认这是最少步数解。

为了介绍这个最少步数解，我们采用“横刀立马”开局式，然后以10个数字代表10个滑块（如图1-13所示）。

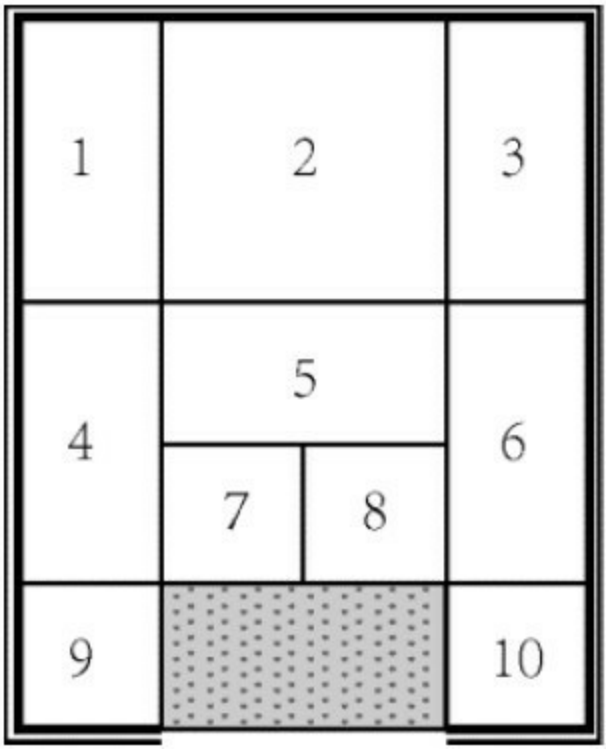


图1-13

解法如下，加下划线的表示只走半程（这里要解释一下“半程”的意思，如棋子9，它的右面有两个空格，“全程”是指9移到8的下面，而“半程”是指9移到7的下面）。

9 ,8,4,5,6；10 ,8,6,5,7；9,6,10,5,9；7,4,6,10,8；5,7,6,4,1；2,3,9,7,6；3,2,1,4,8；10,5,3,6,8；2,9,7,8,6；3,10,2,9,1；4,2,9 ,7,8；6,3,10,9,2；4,1,8 ,7,6；3,2,7,8,1；4,7,5,9,10；2,8,7,5,10；2。

在对华容道的典型开局式——“横刀立马”的解法进行介绍之后，让我们来讨论一下华容道的开局式问题。

造成华容道开局式多样性的原因是它有5个矩形块，矩形块可以横放，也可以竖放。可以只让一个横放，其他4个竖放；也可以只让一个

竖放，其他4个横放；或者取其他组合，甚至让5个都横放。只有让5个都竖放是不可取的。根据有几个矩形块横放，开局式可分一横式、二横式、三横式、四横式、五横式共5大类。在每一大类中，除了五横式只有一种布局外，其他大类每个又都有多种可能的开局式。我国华容道专家齐尧教授（原北京工业学院副院长）曾总结、提炼出180种开局式。

我们对开局式有下列约定俗成的规定：①一般都把曹操置于上方中央；②如果某一种布局是另一布局经移动若干棋子后形成的，则不认为它是一种合格的开局式，也就是说两种开局式原则上是“不互通的”；③对矩形块和小方块则要求它们布置得较有规律，使留下的空格对棋盘的纵轴是对称的。

不是所有的开局式都能解开，除了五竖式明显不可解、五横式只有一种开局式且可解以外，从一横式到四横式都有一些不能解的开局式。什么样的开局式可解？什么样的开局式不可解？有多少开局式可解？有多少开局式不可解？至今仍无人能明确回答，不像“重排九宫”和“重排十五”那样，可根据其排列的奇偶性判定。图1-14到图1-18是一些开局式的例子。

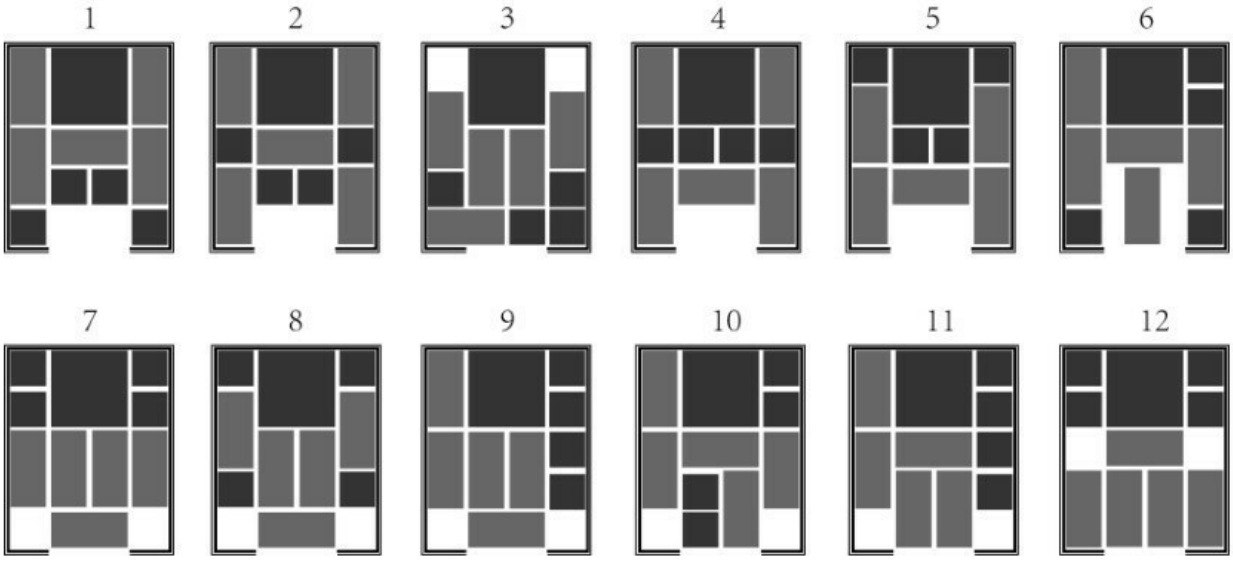


图1-14 一横类开局式

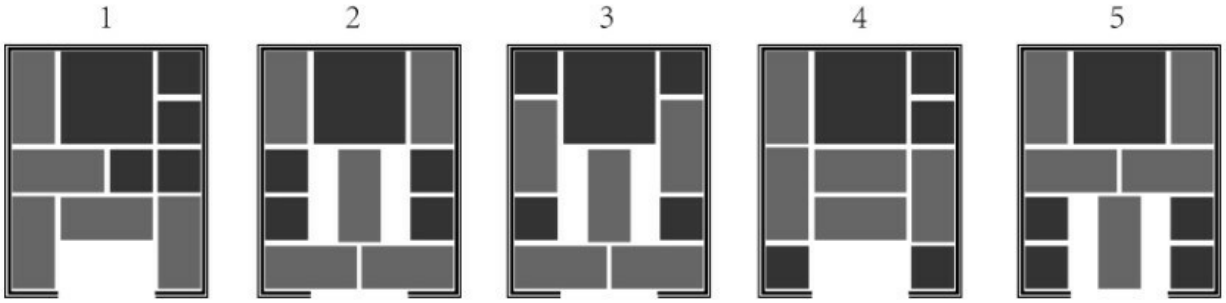


图1-15 二横类开局式

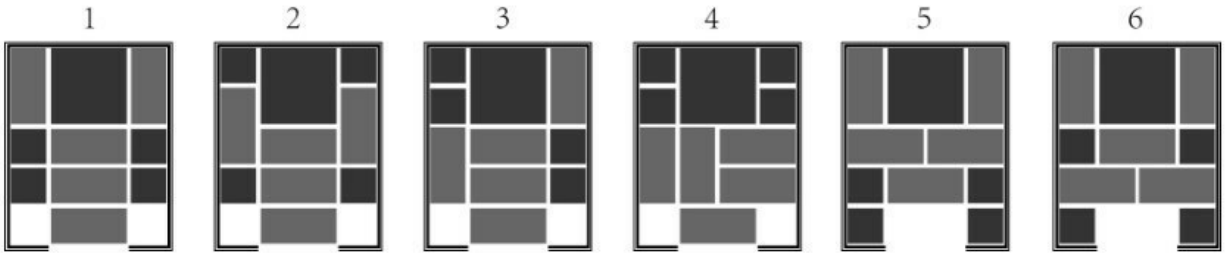


图1-16 三横类开局式

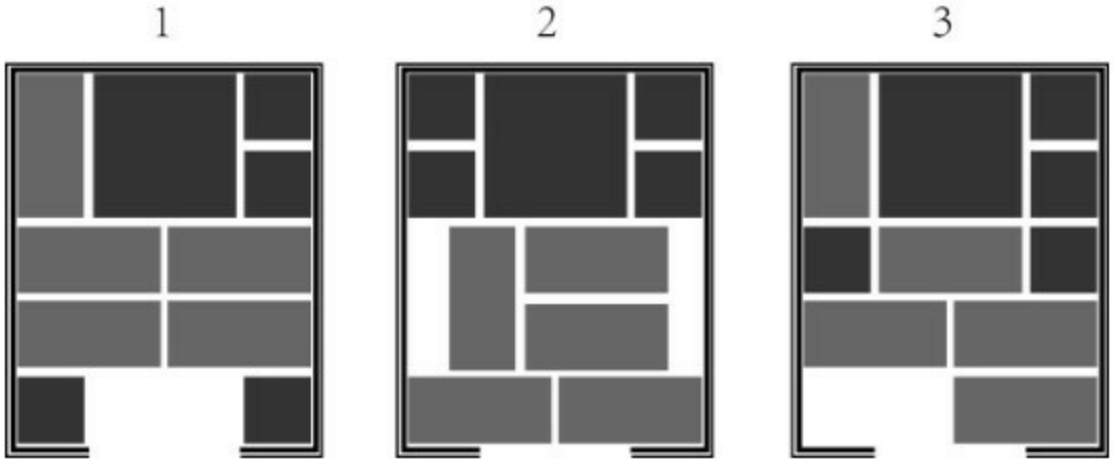


图1-17 四横类开局式

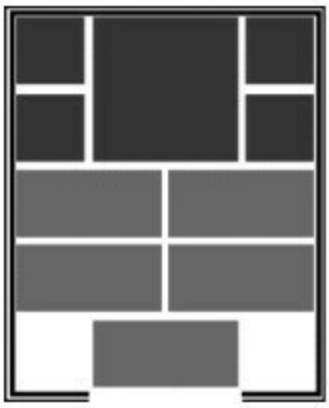


图1-18 五横类开局式

第二章 立体一笔画

《最强大脑》第五季在角逐三十强席位的收官之战时，导演组出示了类似图2-1中所示的多个多面体。要求选手迅速选出其中之一，满足从某一顶点开始用一笔画出多面体的多条棱，不准遗漏也不得重复。答对并用时较少的选手晋级。

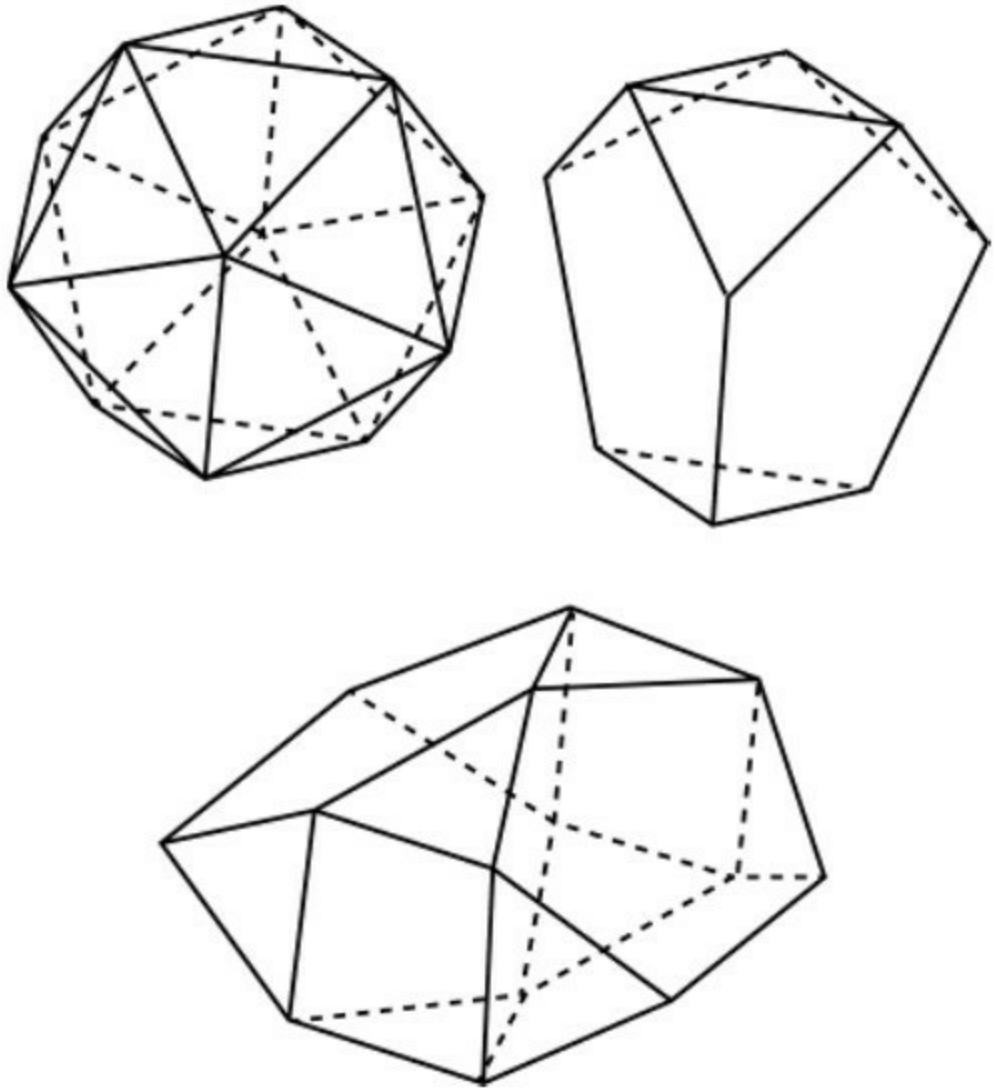


图2-1

平面图形一笔画

在普鲁士东部，濒临波罗的海有一座古老而美丽的城市叫作哥尼斯堡。昔日此为一座重要的工业城市，为东普鲁士的首府，并有一所历史悠久的大学。哥尼斯堡被新河、旧河及布勒格尔河贯穿全城，并将全城分成了4部分。于是人们建造了7座桥，以把哥尼斯堡连成一体（如图2-2所示）。

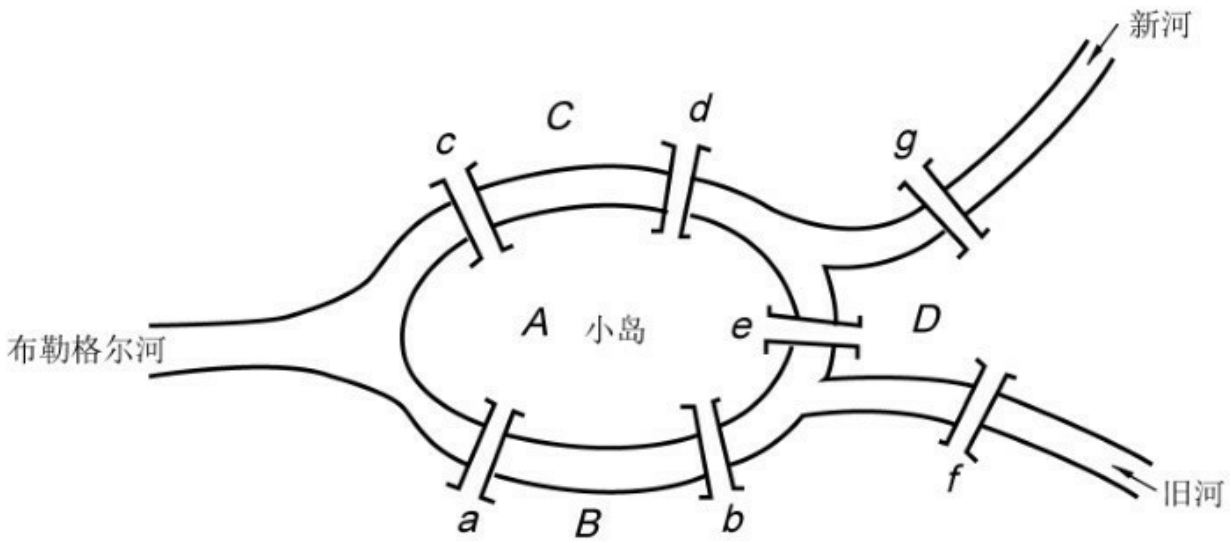


图2-2

每天城里的居民来往这7座桥，熙熙攘攘。望着淙淙流水，这里传出了一个有趣的问题：是否能够一次走遍所有这7座桥，而且每座桥只能走一次？

这个问题似乎不难，谁都乐意去试一下，只是日子一天天过去，也没有人做得到。随着此问题的传播，哥尼斯堡也因此出名。

1727年，欧拉受聘到俄国圣彼得堡科学院工作后，便听到了上述这个故事。他并不曾到过哥尼斯堡，但在听到这个问题后只花了几天的时

间，便解决了此“七桥问题”。他首先将七桥问题转化为一数学模型。他看出两岸陆地及河中的岛都不过是桥的连接点，其大小及形状与问题无关，因此可将它们视为4个点。至于7座桥是7条必经的路线，它们的长短曲直也与问题本身无关，因此可以用任意7条曲线来表示。他先想到如图2-3所示的简化图，然后得到如图2-4所示的数学模型。

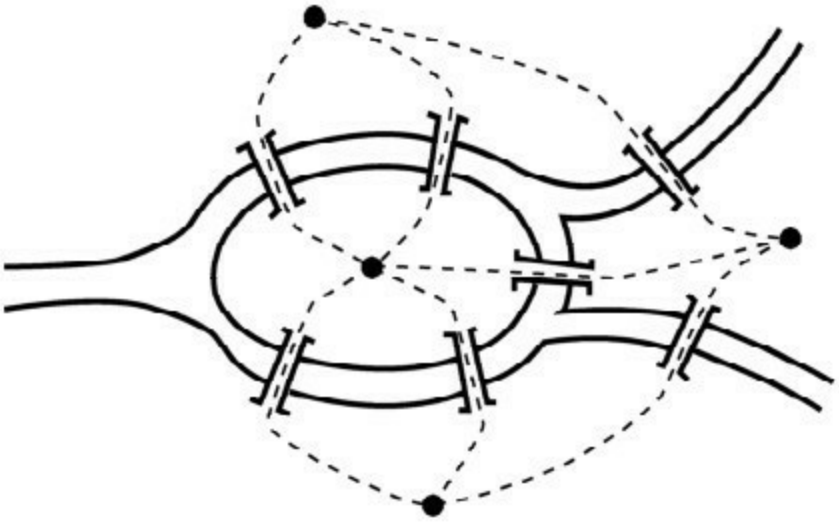


图2-3

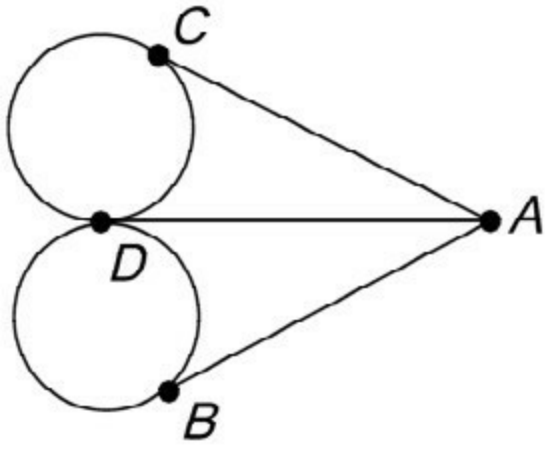


图2-4

如此，欧拉将七桥问题抽象成一个一笔画问题，即图2-4中的图形能否用一笔画出。接着欧拉发现，凡是能用一笔画出的图形，每当你用

笔画一条线（可以是曲线）进入其中的一个点（除了起点与终点）时，你还必须画一条线离开此点，否则图形便不能以一笔画出。也就是说，除了起点与终点，图中每一个点都应与偶数条线相连（这种点称为偶点，反之称为奇点）。若起点与终点重合，则此重合点也应与偶数条线相连（故此点亦为偶点）；若起点与终点不重合，则此两点皆与奇数条线相连（故皆为奇点）。因此，可以一笔画出的图形，其奇点数必为0或2。

现在图2-4中，共有A ,B ,C ,D 4个点，其中A ,B ,C 分别与3条线相连， D 与5条线相连。由于4点均为奇点，因此一笔画出此图形是不可能的，也就是想不重复地通过哥尼斯堡的7座桥是不可能的。

我们看到欧拉将一实际问题抽象化，虽然看不到河也看不到桥，但结果不但解决了原来的问题，连更一般的问题也一并解决了。

补充一点，上面这个答案尚不完整，这里还有一个从哪里开始画的问题。从上面的介绍中可以得到，凡奇点数为2的一笔画，图形必须以一个奇点作为起点，另一个奇点作为终点。

数学史上的一个错误

上节中我们关于七桥问题的论述是一种广为流传的说法（存在于各种图论教科书、数学史书籍和数学科普读物中），这样的描述看上去似乎非常合理，却与真实的历史不太相符。事实上，在欧拉1736年的论文中从来没有出现过任何具有现代意义的“图”，该问题与一笔画之间的联系直到19世纪末才被人们提及。图2-4是在1892年才首次出现于英国数学家罗兹·鲍尔的《数学游戏与古今问题》中，两者之间相差了150多年！

从数学史的角度来看，这可以算得上一个严重的错误。如果一段有关数学史实的记述连起码的真实性都做不到，那么它也就丧失了作为数学史内容而存在的价值。令人遗憾的是，这种以讹传讹的行为仍在延续。不过令人欣慰的是，在20世纪90年代出版的由卡兹所著的《数学史通论》一书中，关于欧拉解决七桥问题的记述已经回归了历史的本来面目。

笔者有幸读到北京化工大学理学院程钊先生发表于《数学传播》第36卷第4期上的一篇文章，它介绍了欧拉真实的解法。为此，略引于下，与广大读者分享。

欧拉写道：“我整个方法的根据是，以适当的并且简易的方式把过桥的行为记录下来。我用大写字母A ,B ,C ,D 表示被河分割开的陆地。当一个人从A 地经过桥a 或b 到B 时，我把这次过桥记作AB ，第一个字母代表他来的地方，第二个字母代表他过桥后所到的地方。如果步行者接着从B 经过桥f 到D ，这次过桥记作BD 。这接连的两次过桥AB 与BD 我就用3个字母ABD 来记录。”不难看出，按照欧拉的记法，记录过桥次

数的字符串中的字母个数总是比桥数多1。因此，如果表示过7座桥就需要用8个字母（欧拉标注的*A* ,*B* ,*C* ,*D* 及桥的标号已在图2-2中列出）。

于是，七桥问题可以重新表述成：在用4个字母*A* ,*B* ,*C* ,*D* 排成的含8个字母的字符串中，有没有可能使*AB* （或*BA* ）组合出现两次，*AC* （或*CA* ）组合也出现两次，而*AD* ,*BD* ,*CD* 这些组合各出现一次？接下来，欧拉想要寻找一个法则，使得对于这个问题或所有类似的问题都可以简易地判断出所要求的字母排法是不是行得通。欧拉注意到，如果去往某块陆地（比如说*A* ，如图2-5所示）经过一座桥*a* ，则不论步行者过桥前在*A* 还是过桥后到*A* ，字母*A* 一定出现一次。如果经过3座桥*a* ,*b* ,*c* ，那么不管他是不是从*A* 出发，字母*A* 都将出现两次。依次类推，如果通往*A* 的桥的个数*k* 是奇数，则字母*A* 出现的次数为（*k* +1）/2。

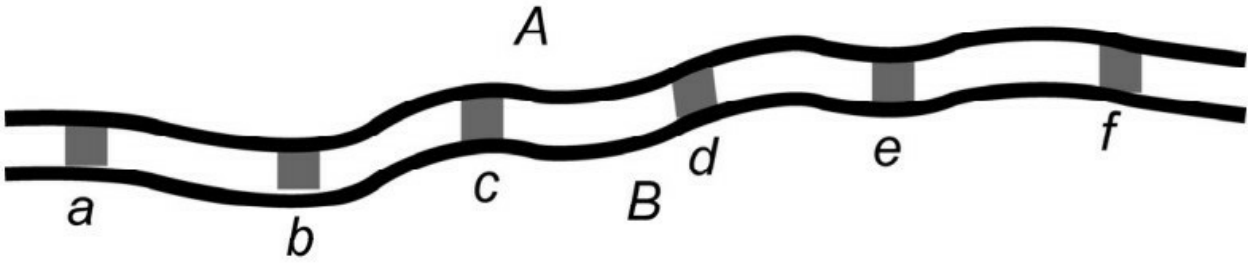


图2-5

本来至此为止，七桥问题已获得解决。因为通往*A* 的有5座桥，所以字母*A* 应出现3次，通往*B* ,*C* ,*D* 的各有3座桥，因此它们各出现2次，这样总的字母个数为3+2+2+2=9。但这在含8个字母的字符串中是不可能的，从而七桥问题无解。欧拉进一步寻找类似的一般问题的求解法则，因此需要找到当桥的个数*k* 是偶数时，字母*A* 出现的规律。他发现如果步行者从某块连通*k* （*k* 为偶数）座桥的陆地出发，则相应的字母出现$\frac{k}{2}+1$ 次；如果从别的陆地出发，则该字母出现$\frac{k}{2}$ 次。

现在，欧拉给出了他的法则。

①将各块陆地用*A* 、*B* 、*C* 等字母表示。

②记桥的总数为∧，将∧+1写在列表的上方。

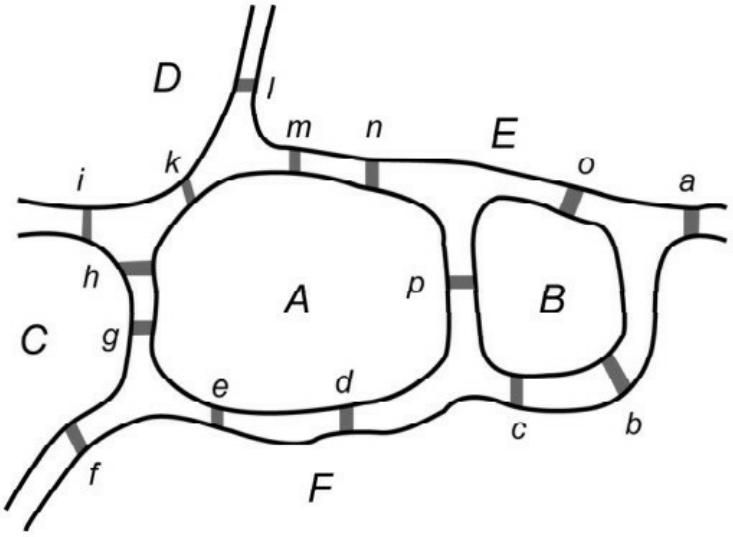
③表的第一列列出代表各陆地的字母A 、B 、C 等，第二列写下通往各陆地的桥数。

④在对应偶数的字母上打星号。

⑤如果桥数k 为偶数，则取$\frac{k}{2}$ 对应记入第三列；否则取$\frac{k+1}{2}$ 对应记入第三列。

⑥将第三列各数加起来，如果该和等于桥的总数∧，则所要求的路线便存在，但必须要从带星号的陆地出发；如果该和等于∧+1，则所要求的路线也存在，但必须从不带星号的陆地出发。

作为上述法则的应用，欧拉举了一个4条河、2座岛、15座桥的例子（如图2-6所示），右边给出的是它的解答。



		16
A*	8	4
B*	4	2
C*	4	2
D	3	2
E	5	3
F*	6	3
		16

图2-6

第三列的数字加起来，得到和数16，与最上方的数相同。所以这条路线是满足要求的，但按照法则要从不带星号的地区D 或E 开始。欧拉确实给出了这样的一条路线：

Ea Fb Bc Fd Ae Ff Cg Ah Ci Dk Am En Ap Bo El D

其中夹在大写字母之间的小写字母指经过的桥。

欧拉并没有满足于给出一个一般的解答，他想的是还有没有可能用更简单的方法来判断呢？欧拉以他特有的对于数字的洞察力看出，表中第二列的各数加起来一定是实际桥数的2倍。因此，如果这些数里有奇数的话，奇数的个数一定是偶数。据此，欧拉通过对表的分析得到了下列更简便的法则。

①如果通奇数座桥的陆地不止两个，则满足条件的线路是找不到的。

②如果只有两个地方通奇数座桥，则可以从这两个地方的其中之一出发，找出所要求的路线。

③如果没有一个地方通奇数座桥，则无论从哪里出发，所要求的路线总能实现。

这个法则就是图论中现在所称的“欧拉定理”的最初形态。

罗密欧急见朱丽叶

下列谜题出自H·杜登尼所著的《坎特伯雷谜题集》。在8×8的方格中，罗密欧必须找出一条通向朱丽叶的路（如图2-7所示），在此过程中，他必须通过所有的方格各一次，还要使转弯的次数最少。这是一个有限制条件的一笔画问题，你能找到这条路径吗？答案见附录二。

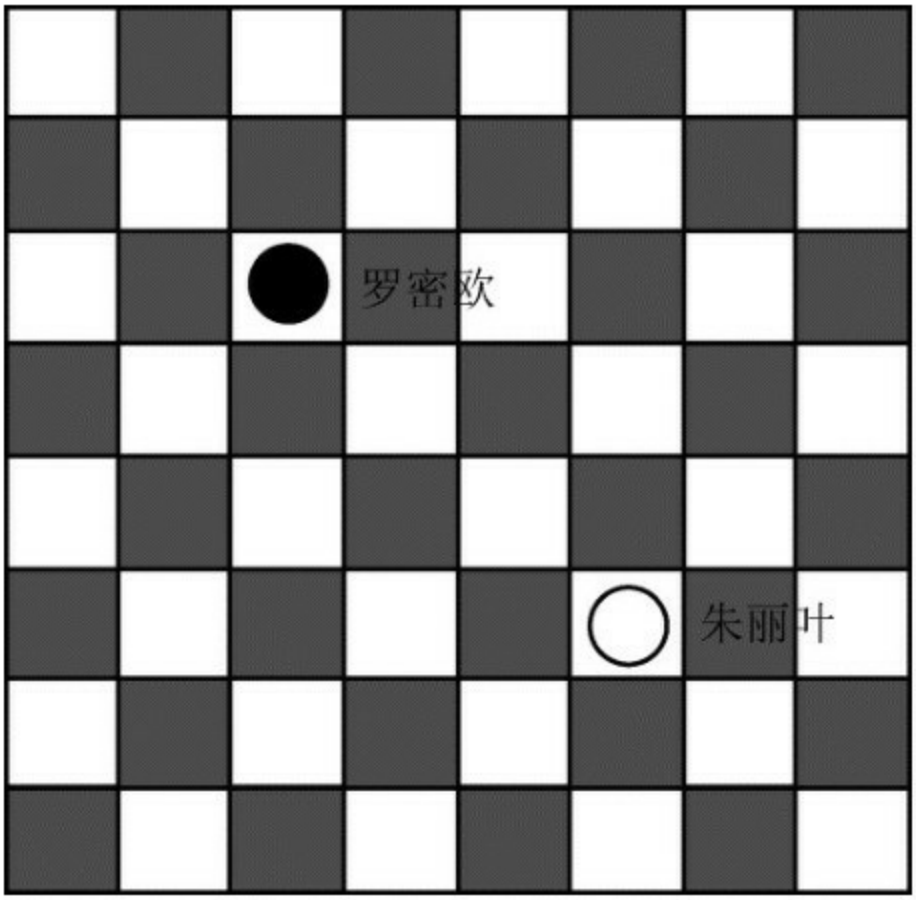


图2-7

多面体一笔画

拓扑学研究的是物体在弹性变形下保持不变的特征，你所考虑的物体就像一块橡皮泥或一个能够拉伸或拖曳的平面一样，拓扑学不考虑图形的大小、形状和刚直性。在拓扑学中，“多长”“多大”这类特征是没有意义的，它注意的是“哪里”“中间”“内部”或“外部”等性质。

对于一个简单多面体（指其中没有洞的多面体），我们可以把它去掉一个面，然后将其余部分展平，不让膜分裂，也不让它起褶痕。多面体的棱或许变成弯曲状的线，使图形变丑了，但利用伸缩自由性，它仍可以变成美丽的图形。如图2-8（a）的四面体*ABCD* ，在底面*BCD* 上开一个洞，然后向四周拉伸变成图2-8（b）的样式。外圈不整齐的一圈就是拉伸后小洞的轮廓，如果把开了洞的一面去掉，再把棱拉直或变曲就可以得到图2-8 （c）或（d）。所以在拓扑学的意义下，这4个图形是等价的。

有了这个概念，就可以把立体一笔画问题转化为平面一笔画问题。只要数一数这个多面体有多少个奇点，就可以得到正确的结论，选手们也可轻松晋级。

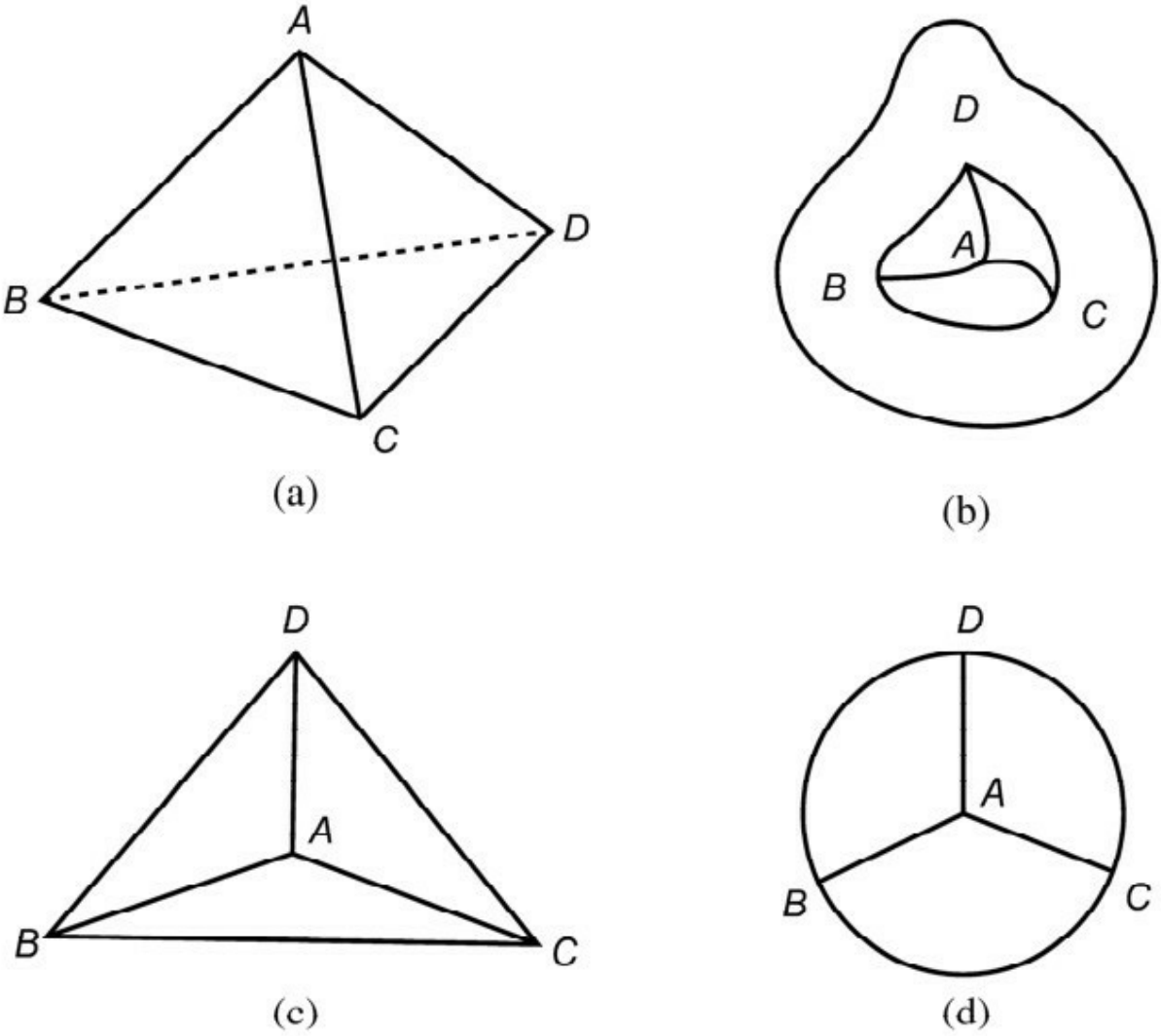


图2-8

蜘蛛与蚂蚁的比赛

如图2-9所示，蚂蚁和蜘蛛分别停在一个六面体的两个顶点上，蚂蚁得意地对蜘蛛说：“小蜘蛛，咱们来比赛吧。大家沿着棱爬，谁先爬过所有的棱到达顶点，谁就获胜。你敢和我比吗？”蜘蛛不声不响地点点头，比赛就这样开始了。你猜谁将取得胜利？为什么？

不妨假定六面体的棱长都相等，蚂蚁、蜘蛛的爬行速度也相同。问题的关键是谁能无重复地爬完所有的棱到达顶点A，这样问题就转化为空间一笔画问题。从图中可以看出点A、E均为奇点，而B、C、D为偶点，所以这个六面体是可以一笔画出的，但必须从奇点A或E开始。

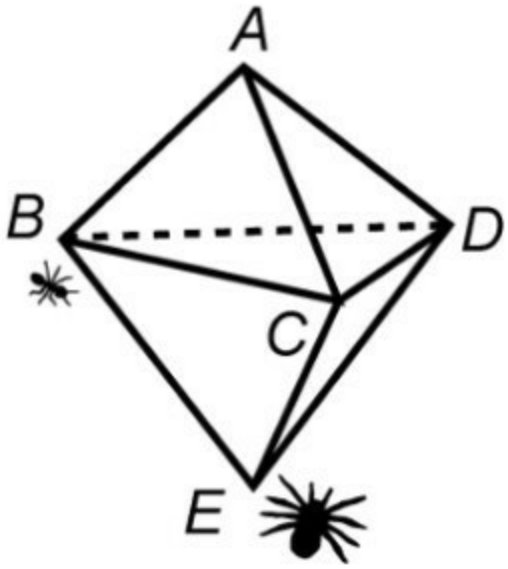


图2-9

蜘蛛的出发点是奇点（即有奇数条通路，这里是3条），终点A也是奇点，所以可以不重复地从E爬到A（E -D -C -B -D -A -B -E -C -A），每条棱只经过一次。而蚂蚁的出发点B是偶点，无法一次走完所有棱到达A点，必须走一些重复道路，因此落后于蜘蛛，正是“蚂蚁无知

夸大口，蜘蛛多想拔头筹”。

哈密顿周游世界问题

素有“19世纪牛顿”之称的英国数学家、物理学家哈密顿（1805—1865）于1859年发明了一种新奇的玩具。它是一个木制的正十二面体（如图2-10左所示），有20个顶点、30条棱和12个面，且每个面都是正五边形。哈密顿在所有的顶点处分别标记了世界上的一个重要城市名。游戏的要求是，从某一个城市出发，沿着这个多面体的棱游遍这20个城市各一次，再回到出发地。这也是个立体一笔画问题。为了玩起来方便，哈密顿做了一个如图2-10右图所示的木制棋盘。各顶点处都挖有小孔，并插入一面旗子。游戏者第一次可拔去任一面旗子，以后只能拔去与刚拔去的旗子相邻者，直到所有旗子都拔光。哈密顿并做了如下文字说明：“十二面遨游，单身周游列国游戏。本玩具系爱尔兰天文学博士威廉·罗恩·哈密顿爵士的发明。宴会上作为即兴表演，无比稀奇。”

图2-10

一般地，从图的某个顶点出发，沿着棱经过每个顶点各一次，这条路线称为哈密顿路；如果最终又回到原出发的顶点，则称为哈密顿回路。有哈密顿回路的图叫哈密顿图。上述哈密顿周游世界的问题，实际上就是在哈密顿图上寻找哈密顿回路。著名数学家苏步青教授曾经通俗

易懂地介绍过哈密顿回路的寻找方法。

周游路线显然是由这20个顶点连成的20角形的封闭折线，它的内部必由棋盘上的若干个五边形组成。因而，周游路线中的任一边不能为两个内部五边形的公共边，内部五边形中的任3个不会共顶点。进而，这些五边形不会围成环形（如图2-11所示），否则连贯的路线会被隔离成不连贯的两部分。综上所述，这些五边形必然是除两端外两两相邻的五边形链（如图2-12所示）。

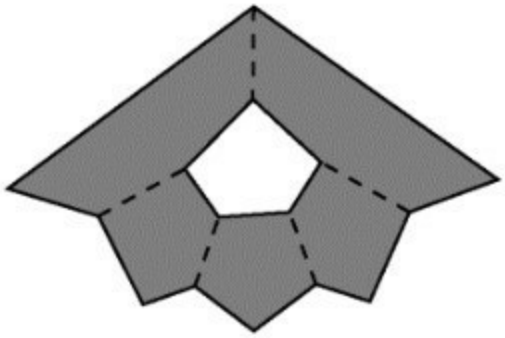


图2-11

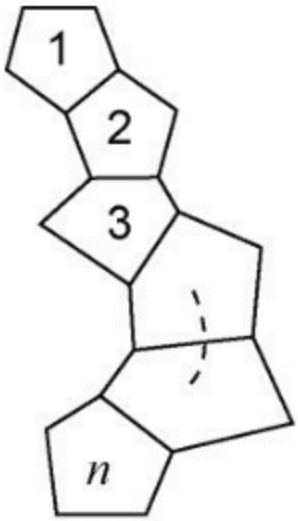


图2-12

设这个五边形链由n 个五边形组成，因而总计应有5n 条边，但在20

角形内部的公共边有（n -1）条。因为每条公共边都算了两次，于是有

$5n$ -2（n -1）=20

解之得n =6。这就是说，应设法找出6个五边形，除首尾两个不相连外，其余两两相邻且任3个五边形不共顶点。我们称之为五角形解链，简称解链。如图2-13即为两条解链，沿着这两条解链的边缘的路线即得所求的走法。显然，每个解链可得40种不同的走法。

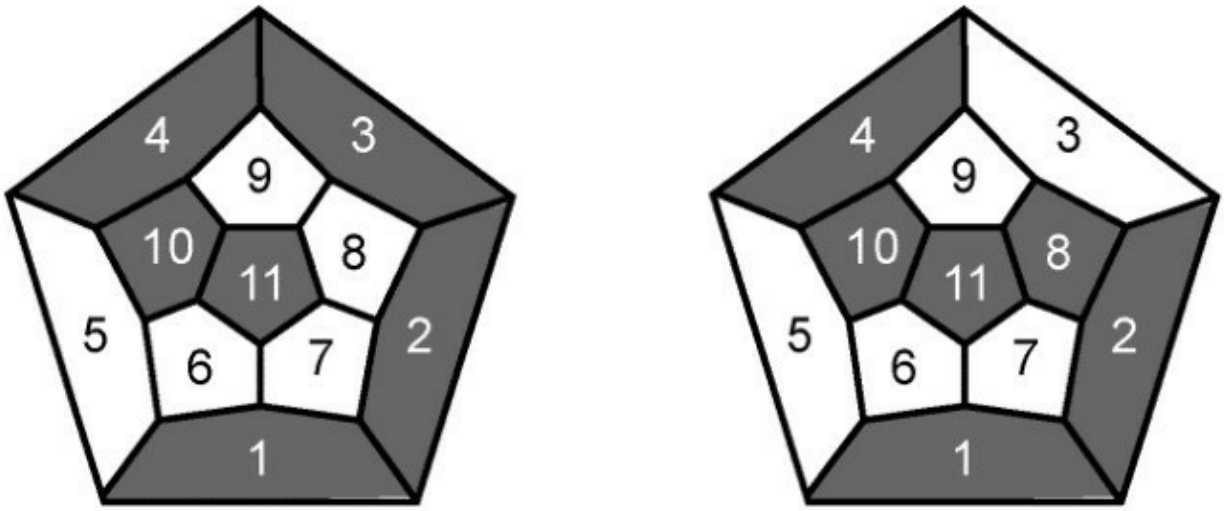


图2-13

为了确定解链的基本类型，我们将哈密顿棋盘上的每个五边形看作一个点，可得11个点；两个五边形相邻就用一条线将代表这两个五边形的点连起来，得到图2-14，这样一来所求的任一条五角形解链就成为图中联结6个顶点的5条边所组成的一条路。由于解链不成环，故此路的两个端点不相邻；又解链中任3个五边形不共顶点，故此路中任3点不两两相邻。符合上述条件的路线即为所求，简称解径。

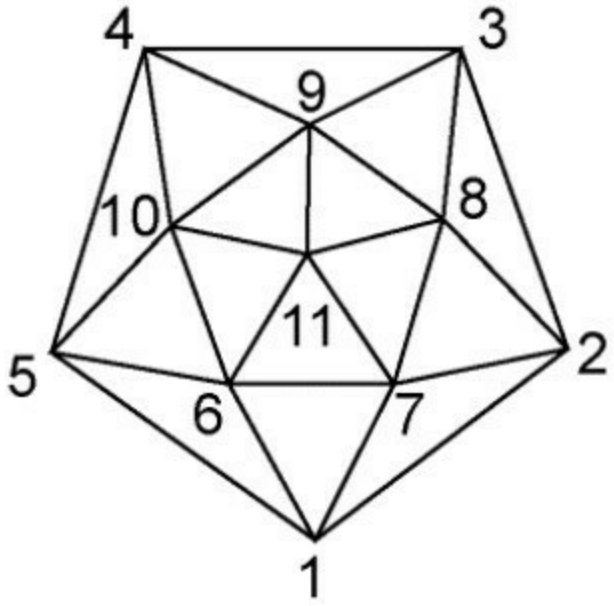


图2-14

显然，解径至多通过外边界的4个顶点，如顺次为1,2,3,4。第5个顶点不能取9（否则3,4,9两两相邻）只能取10，而第6个顶点不能取6（否则6与1相邻，6个数字成环）只能取11，于是得一条解径（1,2,3,4,10,11），其对应的五角形解链为图2-13左图。

如果外边界取3个顶点，便有两种方式。其一是这3点相邻，若取路径（1,2,3,…），接下来必取9，第5个顶点不论取10或11都得不到解径；若取路径（1,2,…,3），也组不成解径。其二是只有两点相邻，如取（1,2,…,4,…），则第3个顶点必取8，进而第4点必是11，第5点必为10，第6点为4，于是又得解径（1,2,8,11,10,4），其对应的解链为图2-13右图。其他情形均无解径。

如果把经过旋转或对称变换可化为上面两种解径的都当作同一类的话，那么从本质上来说，解径只有两条，因而解链也只有图2-13所示的两类。

游戏发明者哈密顿的方法很巧妙。当游戏者拔掉某面旗子（第一次除外）时，与其相邻的旗子有两面——向左走（即拔掉左面的旗子）或

向右走（即拔掉右面的旗子），二者必选其一。向左走记作z ，向右走记作y ，设在某处停止了记作1。然后规定这些动作的运算为积。例如，zy 表示这样一个动作：先向左走再向右走；$y^2=yy$ 表示连续向右走两次；而yz^2y^3 表示向右走一次再向左走两次，然后向右走3次。如果有两个动作，它们都从同一点出发而到达同一终点，则称这两个动作相等，用符号“=”表示。此外，在第一次拔旗子时，与其相邻的有3面旗子，应面对其中两个决定左与右。显然上述规定的运算不满足交换律，即$yz\neq zy$ ，但满足结合律，如（zy ）$z=z$ （yz ）。在哈密顿棋盘上，下列3组公式成立。

$$y^5=z^5=1 \qquad ①$$

$$yz^4=zy^4=1 \qquad ②$$

$$yzy=zy^2z,zyz=yz^2y,y^2=zy^3z,z^2=yz^3y \qquad ③$$

其中在第三组公式中，从左边到右边是升幂，次数分别升高1次或3次。

哈密顿巧妙运用上述公式，将1演变成20个y 或z 的积。举例说明如下：

$$1=z^5=z^2\cdot z^3=(yz^3y)z^3=(yz^3)(yz^3)=(yz^3)^2$$

$$=(yz^2\cdot z)^2=[y(yz^3y)z]^2=(y^2z^3yz)^2$$

$$=(y^2z^2zyz)^2=[y^2(yz^3y)zyz]^2$$

$$=(y^3z^3yzyz)^2=(y^3z^3yzyz)(y^3z^3yzyz)$$

于是构成一个走法：$y^3z^3yzyzy^3z^3yzyz$ ，即从某城市出发连续三次向右，再连续三次向左，一次向右，一次向左，再一次向右，一次向左；接着三次向右，三次向左，一次向右，再一次向左，再一次向右，一次向左。

在上述推演的过程中，要注意不要出现式①和式②的形式，即不要

出现次数高于3的幂。

由于从1演化成y 或z 的20项之积的方法有很多，因而哈密顿回路有很多，请读者仿效上述方法找出其他的走法。

这种“形数组合”的方法是解决数学难题的手段之一。下面来介绍周游世界问题的第三种解法，即拓扑学的方法。设想正十二面体的棱是由橡皮筋做成的，它的每个面都是正五边形。任取其中一个，把它向所在平面的多个方向扩张，其他棱受到这个张力的作用，则会变成形似图2-15所示的平面图形。尽管该图看上去一点也不像一个正十二面体，不过它却能正确反映正十二面体中顶点的相邻关系。我们只关心能否周游世界，至于走时所经路线的形状与长度就不是我们过问的事了。为此，只需思考图2-15中是否有含20个点的圈，我们已用粗实线构造出了一个例子，所以哈密顿周游世界的答案能用有限制条件的一笔画求出。

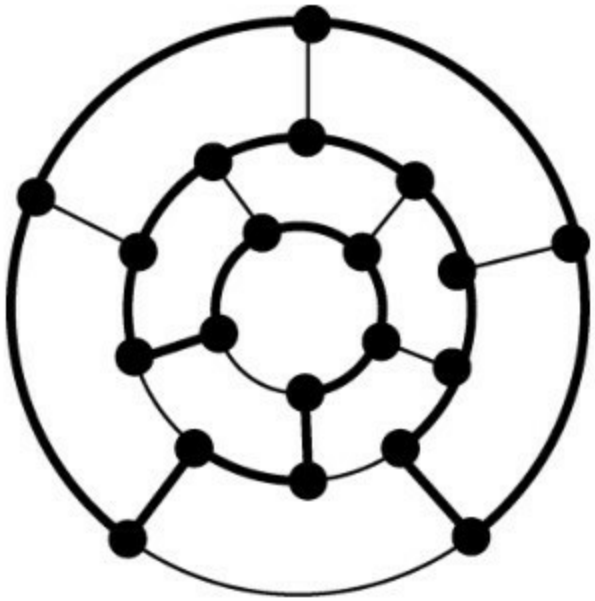


图2-15

由于正四面体、正六面体、正八面体和正二十面体都是哈密顿图，所以还可以把哈密顿周游世界的游戏玩具用所有的正多面体来制作。图2-16中的粗实线表示哈密顿回路。

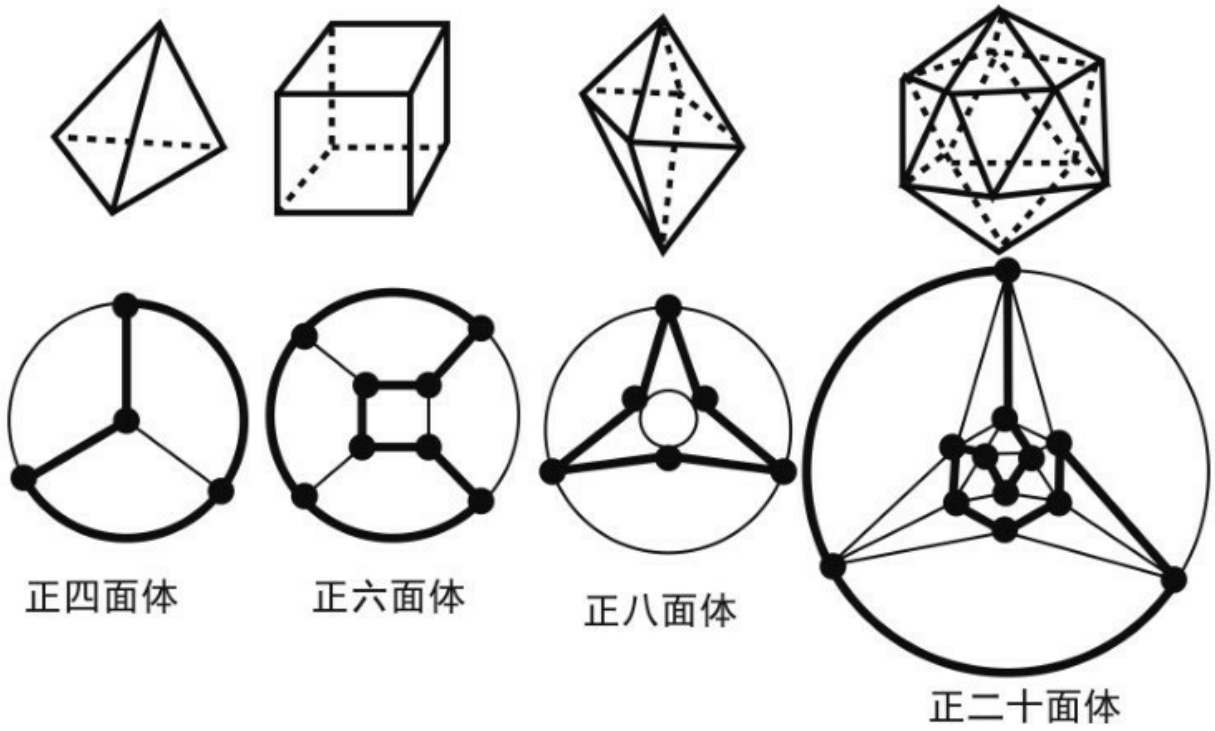


图2-16

类似周游世界的问题很多，解法多样。图2-17中的长方体框架有12个顶点，一只蚂蚁可否从A　　点出发、无重复地爬遍所有的顶点后到G点。如图2-17所示，将框架上的12个顶点相间地涂成黑白两色（相邻两点异色）。若存在这样的路线，则从A 点出发的路线经历的顶点颜色应为：

白→黑→白→黑→白→……

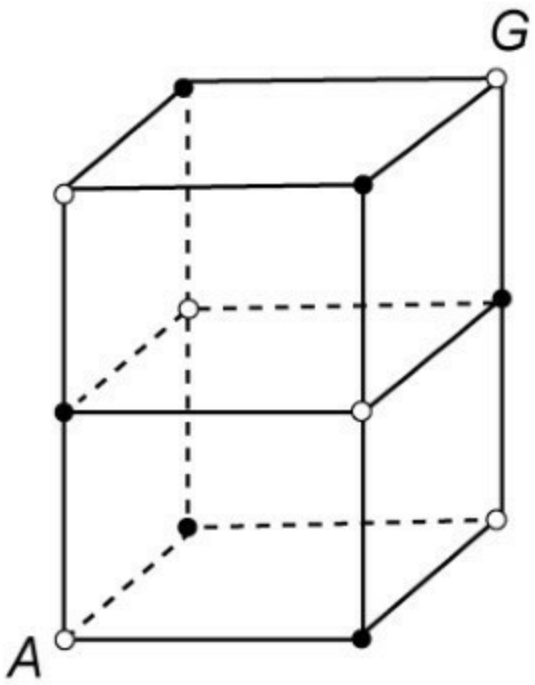


图2-17

这就是说双数步走过的点为黑色，单数步走过的点为白色。因到G点时为第12步，所以它的颜色应为黑色，然而在上述涂色中它却为白色，故本题无解。

上面讲的是在哈密顿图上寻找哈密顿回路的事，但给你一个图，你怎么判断它是不是哈密顿图呢？当然如果图的顶点不多，你可以通过找哈密顿回路进行判断，因为如果有哈密顿回路，这个图就是哈密顿图。这是最常见的逐一尝试的方法，更高效的判断方法仍在探索中。

棋盘上的马步哈密顿回路

通过棋盘上的每个方格有且仅有一次，并跳回到出发时的方格中。这就是棋盘上的马步哈密顿回路问题。

大数学家欧拉曾研究过这个问题，也得到了一个普遍适用的解法。但这个解法要多次调整，十分麻烦。1840年，数学家罗杰特提出了一种简单而巧妙的办法，可在棋盘上形成马步哈密顿回路。他把棋盘分为上下左右4部分，再把每个角一分为四，如图2-18（a）所示。在每个小组的4个方格中，分别标以同样的字母a,e,l,p，其位置在一个方角的4组中是互不相同的，而4个方角的相同位置应对应相同。从这种安排方式中我们看到，在每个方角中，元音a和e所在的方格构成一个环，辅音l和p所在方格构成了方角的对角线；而由同一字母标注的16个方格正好处在一个回路，整个棋盘形成了4个回路。这里需要解释一下“回路”的概念：例如标a的4个正方形，从左上角标a的方格可以跳到右上角标a的方格，接着可以跳到右下角标a的方格，以此类推。这样就形成了一个回路，4个字母就有4个回路。我们可以顺次在标字母p的方格中填入1～16，在标字母a的方格中填入17～32，在16个l中填入33～48，在16个e中填入49～64，分别形成4个回路。现在的问题是怎样把这4个回路合并成一个大回路。为了使合并的过程尽可能简单，要遵守如下两条规则。

l	e	a	p	l	e	a	p
a	p	l	e	a	p	l	e
e	l	p	a	e	l	p	a
p	a	e	l	p	a	e	l
l	e	a	p	l	e	a	p
a	p	l	e	a	p	l	e
e	l	p	a	e	l	p	a
p	a	e	l	p	a	e	l

(a)

34	51	32	15	38	53	18	3
31	14	35	52	17	2	39	54
50	33	16	29	56	37	4	19
13	30	49	36	1	20	55	40
48	63	28	9	44	57	22	5
27	12	45	64	21	8	41	58
62	47	10	25	60	43	6	23
11	26	61	46	7	24	59	42

(b)

图2-18

① 所选的起点方格和终点方格应分别标有辅音字母和元音字母。为此，要轮流取标辅音的回路和标元音的回路，以起始方格所标辅音的回路开始，以终止方格所标元音的回路结束。

②每个回路的旋转方向要一致，而且不要在方阵的四角或边缘方格上终止。

以图2-18为例，我们以标p的方格开始。第1个回路是标p的16个方格，第2个回路是标a的16个方格，第3个回路是标l的16个方格，最后是标e的16个方格，都取顺时针方向。这4个回路都是中心对称图形，且具有连贯性，仅从32到33是例外，没有转入下一方角，仅在本方角中接续，这样形成的哈密顿回路见图2-18（b）。

本章最后来介绍一下哈密顿其人。1805年，哈密顿生于爱尔兰都柏林的一个律师家庭。5岁通晓拉丁文，14岁时已经学会了12种语言，13岁时因为阅读牛顿的《普遍算数》一书而对数学产生了强烈兴趣。1823年，哈密顿进入都柏林三一学院学习，开始对天文学展露出特殊的天赋和偏爱。大学尚未毕业，他便成为天文学教授，并被任命为天文台台

长，时年22岁。1832年，哈密顿当选爱尔兰科学院院士，成为当时成绩卓著的科学大师。由于操劳过度，他于1865年去世，时年60岁。

哈密顿善于解决各种特殊问题，再把对具体问题的研究方法与结论总结为一般性理论，发表的著作有140余篇。此外，他在力学、数学和光学上都有杰出贡献，其中在数学上有深远影响的成就是四元数和哈密顿图。哈密顿为人谦虚诚恳，重视外语学习和科研。文章写得好，教书教得好，著作亦十分畅销，是19世纪青年人崇拜的典范。

第三章 迷宫中的数学

《最强大脑》第五季第三场是一对一的淘汰赛，挑战项目是立体迷宫。“迷宫”项目在《最强大脑》中并非第一次出现，在第二季晋级赛的第二场中就出现过室外蜂巢迷宫，其难度系数高达10。

我国的迷宫

关于迷宫的起源，有许多不同的说法，流传最广且为大多数人所接受的一种说法是基于一则古希腊神话故事的克里特迷宫，又名诺索斯迷宫。

在我国古典文学作品中也有关于迷宫的描述。如在《水浒传》“宋江三打祝家庄”的故事中，祝家庄本身就是一个迷宫，但它不叫迷宫而叫“盘陀路”。有诗为证：“好个祝家庄，尽是盘陀路；容易入得来，只是出不去。”宋公明一打祝家庄，先派石秀、杨林去探路。杨林不认路，“只拣大路走了，左来右去，只走了死路”，被活捉了去。幸好石秀机灵，从钟离老人口中套出了盘陀路的秘密：“只看有白杨树便可转弯，不问道路阔狭。否则都是死路，地下还埋着竹签、铁蒺藜。若是走差了，踏着竹签，准定吃捉了。待走那里去？”

比《水浒传》更早的《三国演义》中的“八阵图”也是一个迷宫。第84回的“陆逊营烧七百里，孔明巧布八阵图”中有描写，八阵图是刘备入川时，诸葛亮在长江边的鱼腹浦构筑的一个石头阵。当刘备为报东吴夺荆州、杀关羽之仇而大举兴兵伐吴时，东吴大将陆逊设计火烧连营七百里，大败蜀军并乘胜追击直至蚓关不远处，恰遇石头阵。陆逊引数十骑士上前观看，一刹那，飞沙走石，横沙立土，江声浪涌，杀气冲起。逊急欲回时，无路可出，正惊疑间，忽见一老人立于马前。逊问曰：“长者何人？”老人答曰：“老夫乃诸葛孔明之岳父黄承彦也。昔小婿入川时，于此布下石头阵，名‘八阵图’。反复入门，按循甲休、生、伤、杜、景、死、惊、开。每日每时，变化无端，可比十万精兵。老夫适于山岩之上，见将军从‘死门’而入，料想不识此阵，必为所迷，老夫平生

好善，不忍将军险没于此，故特自将汝自‘生门’引出也。”

更早的是《封神演义》第50回“三姑计摆黄河阵”中的“九曲黄河阵”。三仙岛上的云霄、琼霄、碧霄三位娘娘下凡摆九曲黄河阵的起因是她们的哥哥去战姜子牙，被西昆仑的道人陆龙用法术杀死。为报杀兄之仇，三仙姑摆了九曲黄河阵，把玉虚门人、杨戬、金吒、木吒等12弟子俱困于阵中，弄得姜子牙狼狈不堪，亏得元始天尊和老子两位祖师爷亲自出马才破了黄河阵，收了三仙姑。

上述情节撇去其中神怪因素，九曲黄河阵实际上是由600名精兵强将组成的一个活迷宫。陷入其中的人当然无路可寻，凶多吉少。

我国河北省井陉地区的板桥、长岗、南河头、北孤台、于家等村庄至今有在元宵节摆“九曲黄河灯”的习俗，其起源就是九曲黄河阵。据李苑、余音编著的《趣味迷宫畅游》一书，明确指出“九曲黄河阵”就是一种迷宫，而且绘出了它的图案，如图3-1所示。

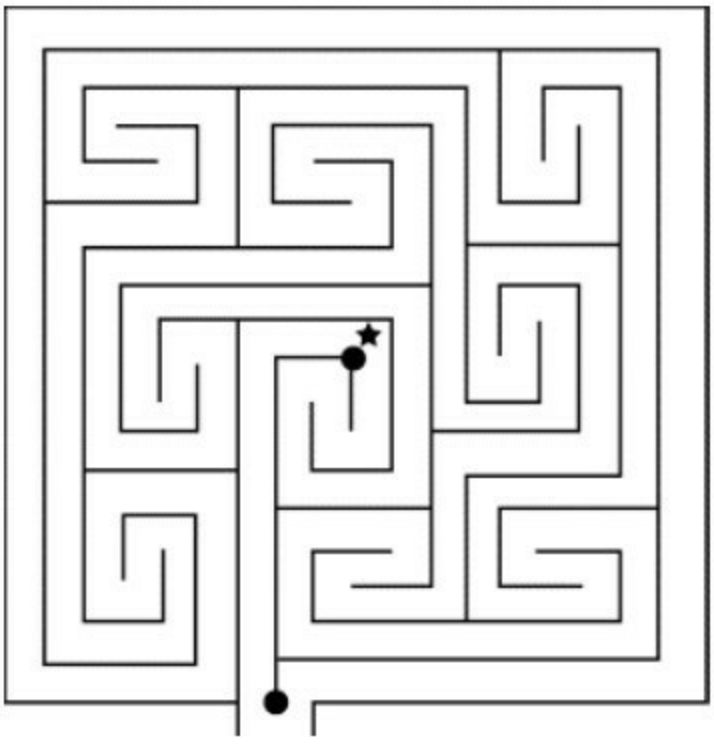


图3-1

由此可见，转九曲黄河阵的人从入口的右侧进入以后，必须按顺序转过9个方阵，到达标有星号的迷宫中央才能出阵。

在我国历史上最著名的迷宫则是圆明园中的万花阵，它是意大利传教士郎世宁在包括画家、建筑师、工程师和园艺师等9个助手的帮助下为清朝皇帝设计建造的。据说每年中秋节，皇帝、皇后会在其中赏月，同时命宫女们提着灯笼入迷宫，谁能到达中央的亭子谁就有赏。可惜这个迷宫连同整个圆明园已在1860年被英法联军所破坏。幸好，它的风姿被宫廷画家用铜版画的形式保存了下来。圆明园中现有的万花阵迷宫是后来复建的，游人可以入内一游。

如何走迷宫

在有关迷宫的所有数学问题中，大家最关心的恐怕就是如何走迷宫这个问题了。因为有些迷宫要求从入口到达中心，有些迷宫要求从一个口进去，从另一个口出来，所以我们用“走迷宫”这个说法泛指这两种情况。

如果你是在纸上走迷宫，那么最简单的办法就是用铅笔把所有死胡同和回路（如果有的话）都涂黑，这就把你应该走的路径突显出来了。例如图3-2（a）中的迷宫，先把不能通行的死胡同用字母标出，共有A、B、C、D、E、F 6处，把它们用阴影掩盖掉，掩盖的范围到死胡同口为止。如图3-2（b），我们看到分开A、B两个死胡同的结点相当于又一死胡同的底，于是把底也掩盖掉。在图3-2（c）中，G和H虽然不是死胡同，但进入里面转一圈后仍会回到原入口处，这是一条自身封闭的路，对到达目的地不起作用。于是把它们也掩盖掉，这样就得到图3-2（d）。在图3-2（d）中有两对黑点，每一对黑点之间也是一条迂回路线，此类路线仍须掩盖掉，这样便得出图3-2（e），这样就走出了迷宫。

如果你面对的是一个现实中的迷宫，当然无法看到它的整个布局，就不可能用上述方法决定你应该行走的路线。这时候右手法则（或者左手法则）就可以帮助你顺利地走出迷宫了。所谓右手法则（或者左手法则）就是在你进入迷宫以后，始终用右手（或左手）摸着墙前进，这就能确保你走到迷宫中心，并走出迷宫（到达另一个出口或者回到入口处）。图3-3就是用这个方法走汉普顿迷宫的示意图，用右手法则或左手法则所走的路线是完全一样的，但行进方向正好相反。

上述方法一般来说是有效的，但它有两个例外：

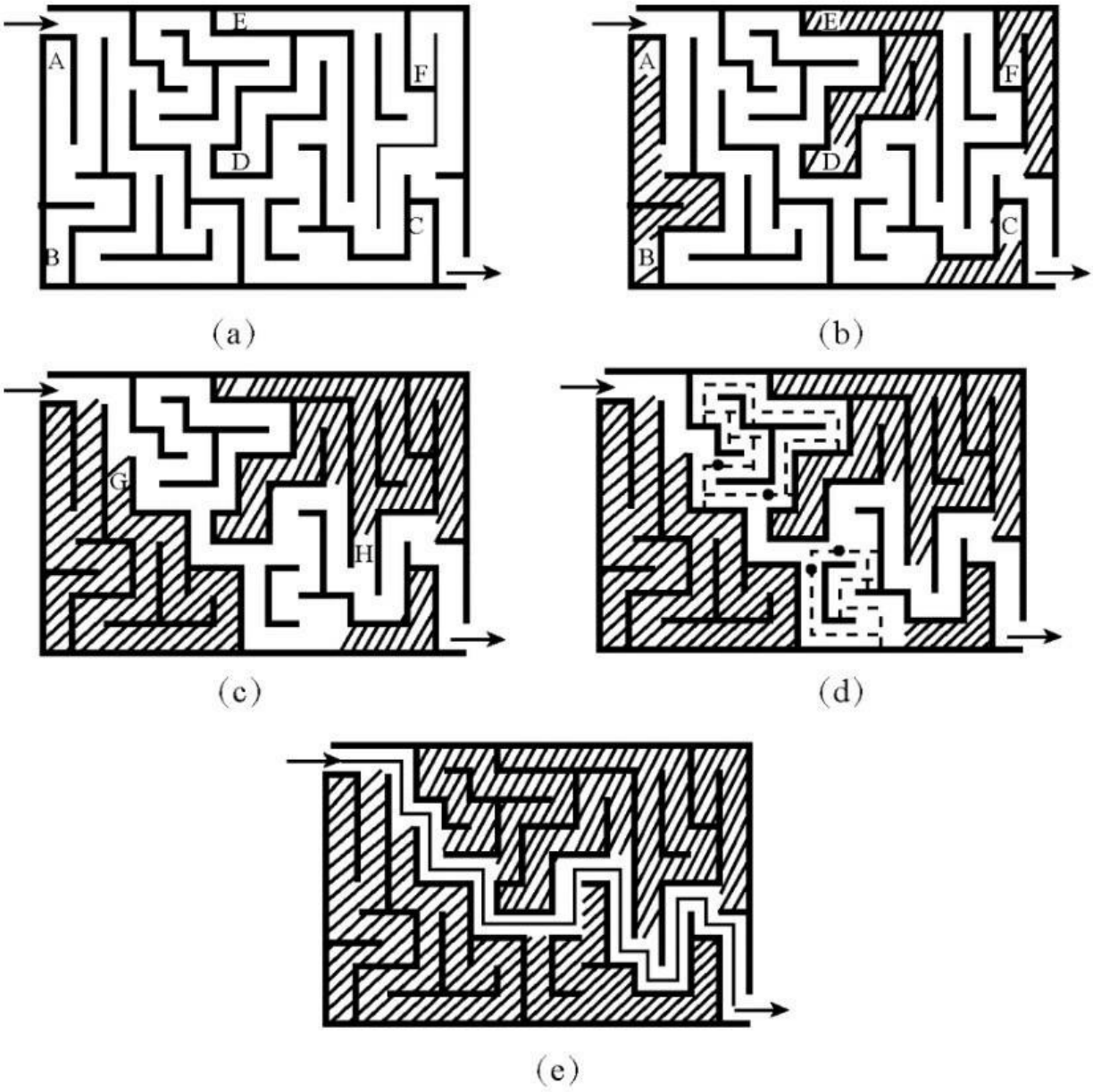


图3-2

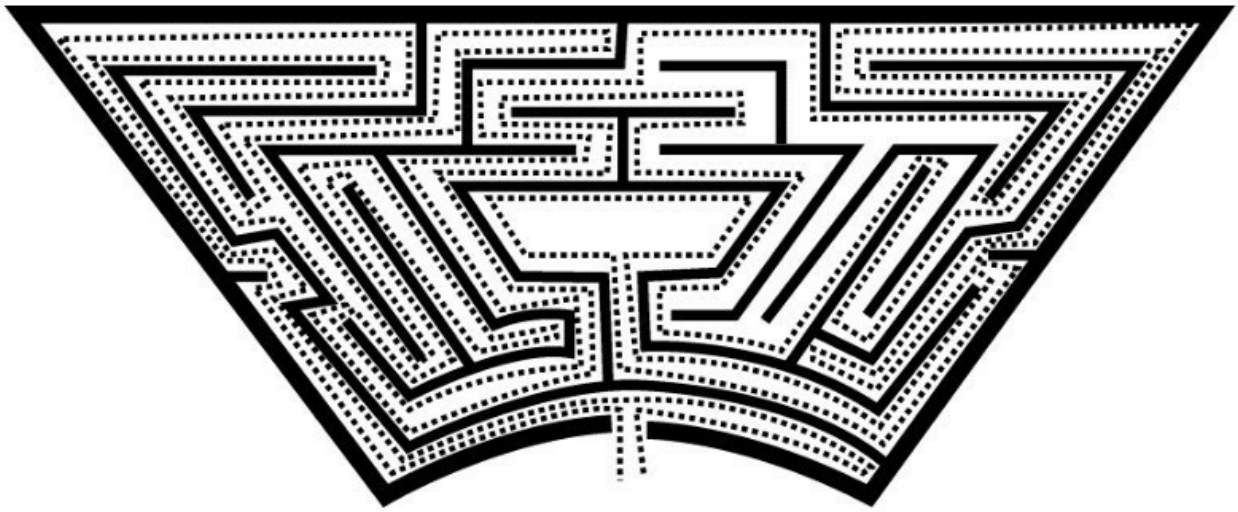


图3-3

①该迷宫有两个入口，而且有一条不通过终点的路线连接它们；

②迷宫的路中带有环绕终点的圈。

法国数学家M·特马克设计了一种解任意迷宫的普适性方法，程序如下。

①在你走过的路的右侧画一条线；

②当你走到一个新交叉点时，你可以任意选择一条你想走的路；

③如果你在新的路上又回到旧的交叉点或死胡同，那便往回走；

④如果你在旧路上走到一个旧的交叉点，那就任取一条新路（如果有的话），否则就取一条旧路；

⑤决不进入两侧都做了记号的路。

特马克法可以归纳为如下口诀：

新路新结点，

任选支路行；

走进死胡同，

掉头毋悻悻；

新路旧结点，

回头要清醒；

旧路旧结点，
新路最高兴；
若无新路选，
旧路也可行；
已走两遍路，
千万莫再行。
以上方法虽然通用，但可能要花费不少时间。

迷宫的拓扑结构

从数学上来说，各种各样的迷宫路线无非就是一个图，也就是由若干结点以及其中某些结点之间的连线即所谓的边所组成的。若以静态及定位方式研究迷宫，即不考虑各结点的具体位置而只考虑其相对位置，也不考虑各边的具体长度和形状的时候，我们可以把迷宫的拓扑结构画出来，这对于研究迷宫中的数学问题是十分必要的。

下面我们通过一个例子，来说明迷宫拓扑结构的画法。图3-4是英国的一座建于1690年的迷宫，名为“汉普顿迷宫”。我们把迷宫图中的每一个死胡同的底（绝点）及每一个分歧点都标上字母，那么图3-4的迷宫图可改画成图3-5的拓扑图。

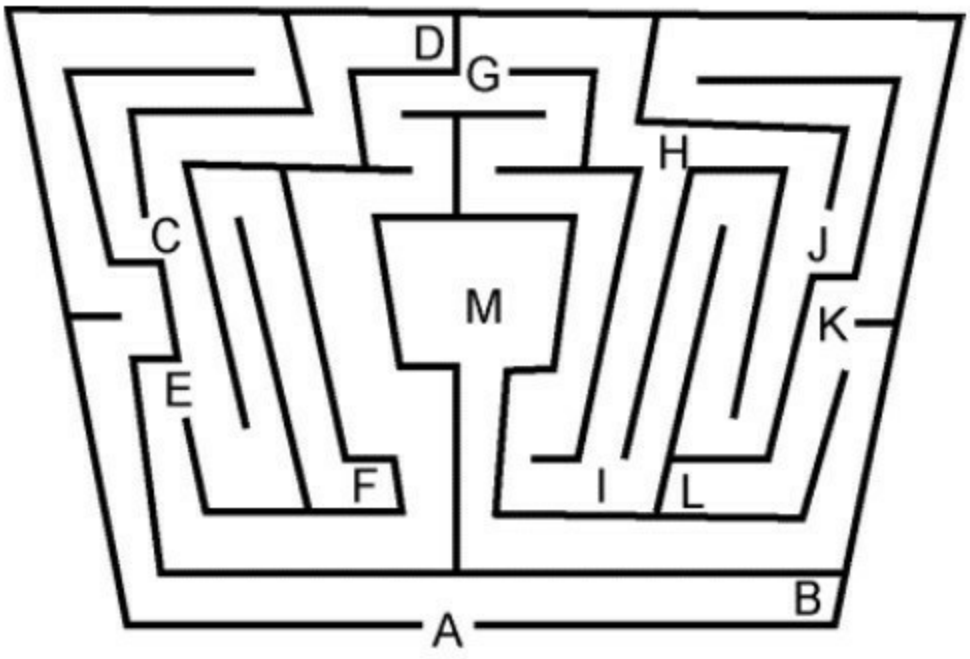


图3-4

在一个图中，若道路成环（如图3-5中的G-I-J-H、G-I-H等），则从

环上的某一点出发，循环行走，定可返回原地。若道路呈“枝”状，则来回走两遍也是可返回原地的。所以，如果把枝状路画成“重路”，环状路只画外圈，那么图3-5就成了图3-6的样子。在图3-6中，与每个顶点相关联的路都是偶数，满足一笔画的规则，定能满足从哪一点出发便回到哪一点结束，即可走出迷宫。

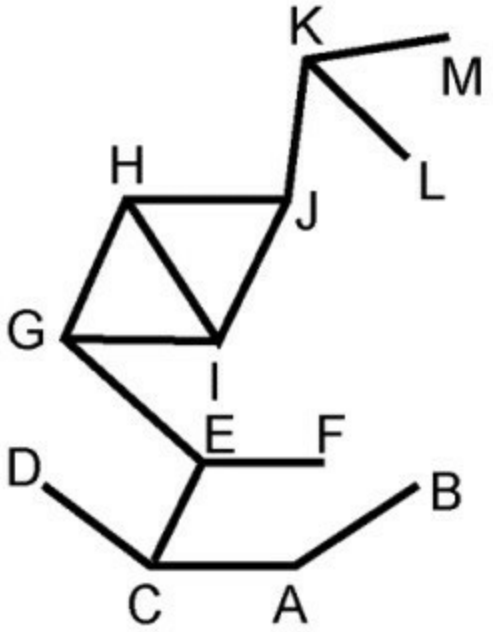


图3-5

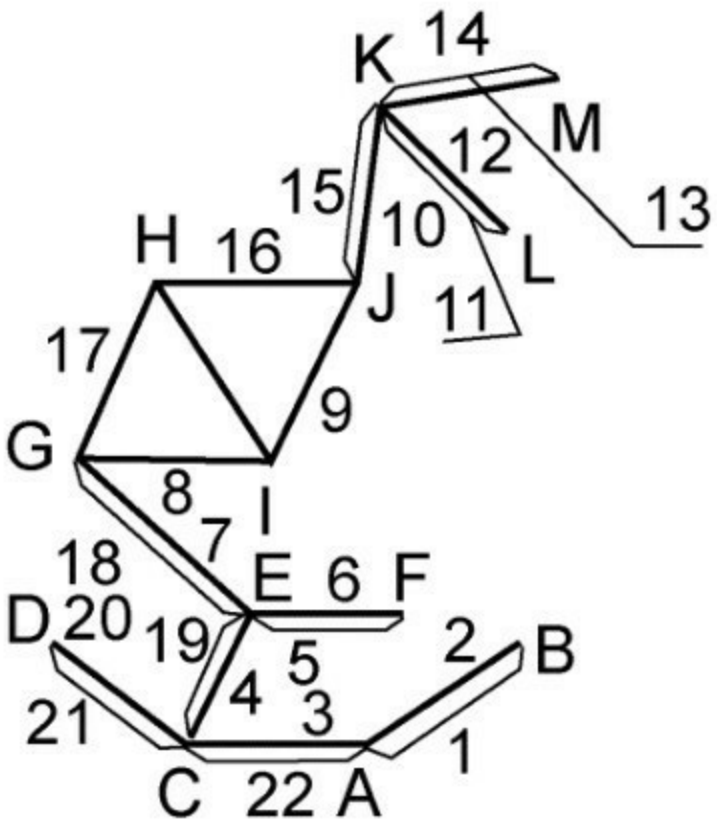


图3-6

对于图3-6，若从A出发，经路线1,2,3,…,21,22回到A，这就走出了迷宫，而且是紧贴右墙（即右手法则）行走的路线图。

在这个例子中，我们不仅画出了迷宫的拓扑结构，而且找到了走出迷宫的路线。对于复杂的拓扑结构自然要找更好的办法，甚至用计算机来求解。

计算机解迷宫

计算机解迷宫虽然没有实用意义，但有较高的学术价值。首先，要解决的问题是把由拓扑结构形成的迷宫转化为数字，毕竟计算机只认“数”而不认“图”。这里运用的主要手段就是邻接矩阵。我们把图中点的集合记成$V=\{v_1,v_2,\ldots,v_n\}$，线段组成的集合记成$E=\{e_1,e_2,\ldots,e_\varepsilon\}$这便构成如下矩阵。

$$A=\begin{array}{c} \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{array}\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \end{array}$$

其中第i行第j列处的元素$a_{ij}=\begin{cases}1, v_i\text{与}v_j\text{之间有线段相连}\\ 0, v_i\text{与}v_j\text{之间无线段相连}\end{cases}$

对于图3-7中的四边形G，它的邻接矩阵为

$$A(G)=\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array}\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

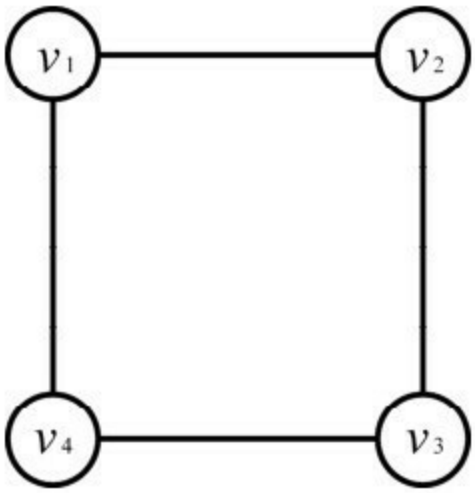


图3-7

对于有向图，也有邻接矩阵。例如在图3-8中，每条边皆有箭头指示方向，这种图称为有向图。

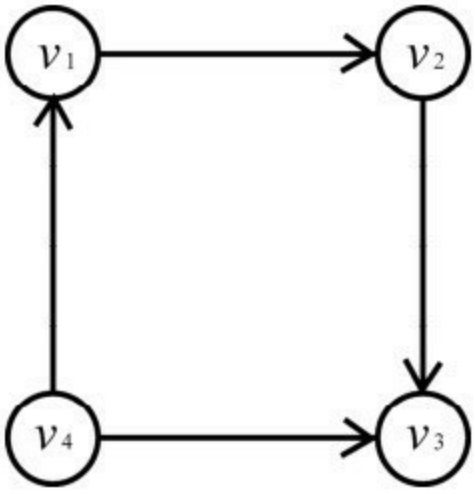


图3-8

有向图G'的邻接矩阵为

$$A(G')=\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array}\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

一般地，对于有向图G'，其邻接矩阵定义为

$$A(G')=\begin{array}{c} \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{array}\begin{array}{c} \begin{array}{ccccc} v_1 & v_2 & v_3 & \cdots & v_n \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ & & & \cdots & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \end{array}$$

其中$a_{ij}=\begin{cases}1, G'\text{中有}(v_i)\to(v_j) \\ 0, G'\text{中无}(v_i)\to(v_j)\end{cases}$

对于每个简单图或有向图，都可以轻易地写出它的邻接矩阵。具体到迷宫来说，如果已到了出口，则可以-1代替。这样迷宫拓扑结构中所含的一切信息都蕴含在了数字化的邻接矩阵当中，数学的量化功能在图的邻接矩阵上体现得淋漓尽致。

有了数据，就应寻找一种有效的算法。1970年，在斯坦福大学做访问学者的霍普克洛夫特和该校的研究生罗伯特·陶尔杨合作，提出了一

种适于解决平面图的新算法，即深度优先搜索算法。

由于迷宫也是图，因此适合采用深度优先搜索算法及回溯来求解。实际上这同人走迷宫是非常相似的：先任意选择一条路，只要能走下去就一直往前走；走不下去了才往回返，选一条新的路走。在通路的每一点上，按右、上、左、下的顺序搜索下一个落脚点，有路则进，无路则退，从上一点的下一个方向继续搜索。例如，从起点开始一路向右到达a 点，再向右搜索仍有路则进至b 点，若在b 点向右、上、下都没有路则退回a 点，重新向上、下搜索，哪里有路往哪走；若无路，则继续后退，如此等等。

解决了这两个基本问题，再加上一些技术手段，就可以顺利上机解迷宫了。本书只说个概要，有兴趣的读者可以参阅专业书籍。

迷宫与人工智能

当代杰出数学家、信息论的开山祖师克劳特·香农曾制造了一台能够自己学习的机器，称为“迷宫之鼠”。这是一个铝制的方盒子，用隔板分成了一个有25个方格的迷宫。这些方格分为5列，每列有5格。在一个格子里放上做成老鼠样子的永久磁铁，在另一个最难到达的格子里放上钢制的“奶油”。盒子背后设有控制装置，当机器一开动，“老鼠”就被迷宫下面的电磁铁控制，而电磁铁的运动由继电器电路控制，它是整个装置的核心部件。继电器电路共由110个继电器构成，起到控制、计算和存储的作用。迷宫的每个格子里都放着继电器，当老鼠碰到隔板时，隔板上的继电器便处于接通状态。它们可记存每一个格子内的4个运动方向。

为什么说这台机器能够自己学习呢？原来是这样的。在机器第一次开动时，老鼠会到处乱钻，沿着复杂的道路找到“奶油”，假设这时它所耗的时间是60秒。找到“奶油”以后，暂时把机器关闭一下，然后把老鼠放在原来的地方，再开动机器，这时老鼠会沿着最短的道路行走，不会钻到那些死胡同里去，在较短的时间（少于60秒）内就找到了“奶油”。用磁铁做的老鼠就像“活”的一样，它能记住路线，产生条件反射，并从已有的错误中吸取教训。

机器老鼠为什么有这种出色的学习本领呢？据说是因为香农为它设定了以下4个条件。

①老鼠没有到达的格子都亮绿光。

②老鼠移动时第一次到达的地方都亮黄光。

③第二次到达的地方亮红光。

④亮起红光的地方老鼠再也不会钻进去。

在这4个条件下，可以证明老鼠在第一次行动时一定会找到“奶油”，而在第二次行动时就会沿着最短的途径找到它。于是我们可以这样说，老鼠的行动已经被程序所控制。

1952年，在第5届控制论会议上香农做了一个报告，向与会代表演示了这台机器，顿时引起了广泛的兴趣和热烈的讨论。后来，香农在一篇综述性的论文《计算机和自动机》中总结说，这台机器在非常原始的水平上具备了下列能力：①通过试错解决问题；②重复同一过程时牢记教训，不再犯错；③对特定解增加新的信息；④原有解不再有用时便把它忘掉。

在香农研究学习型机器的同时，普林斯顿大学的博士研究生明斯基也在研究机器是否可以思考，是否可以具有学习能力的问题。1951年，他建造了一台学习机，名为Snarc。Snarc是世界上第一个神经网络模拟器，其目的也是学习如何穿越迷宫！Snarc虽然比较粗糙且不够灵活，但毕竟是人工智能研究中最早的尝试之一。1956年，明斯基、香农、麦卡锡、罗杰斯特等人一起发起了著名的达特茅斯会议。正是在这个会议上，才正式提出了“人工智能”这一名词，并宣告了这一崭新学科的诞生。由此可见，迷宫在催生人工智能这一研究课题方面功不可没。

扑克迷宫

讲了这么一大堆的数学难题，这一节大家来轻松一下吧。下面是一个单人游戏，相当有趣，可以随时随地玩一把。

如图3-9所示，52张扑克牌按6行放好，虚线位置不放牌（当然，发牌是随机的，此处为一个示例）。然后把4张K取走，这样在6×9的区域中共出现6个空位。下面开始调整扑克牌的位置，每次一张，最终目标是使整副牌重新理顺：红桃、黑桃、梅花、方块各自聚集在一起，并且从1点的“A”到王后“Q”按从小到大的次序排好（至于红桃、黑桃、梅花、方块这4者的次序是可以任意排的）。调整时应遵守以下规则：每次把一张牌移到任意一个空位上去，只要该空位左侧（或右侧）的牌同它属于一个花色，而且该牌的点数大于空位前牌的点数，且小于空位后牌的点数。例如，一张方块2可以移到空位前有一张方块A或空位后有一张方块3的空位上去。如果“Q”后有空位，那么不管什么花色的“A”都可以移进去，即使空位后的“2”同“A”的花色不同也没关系。

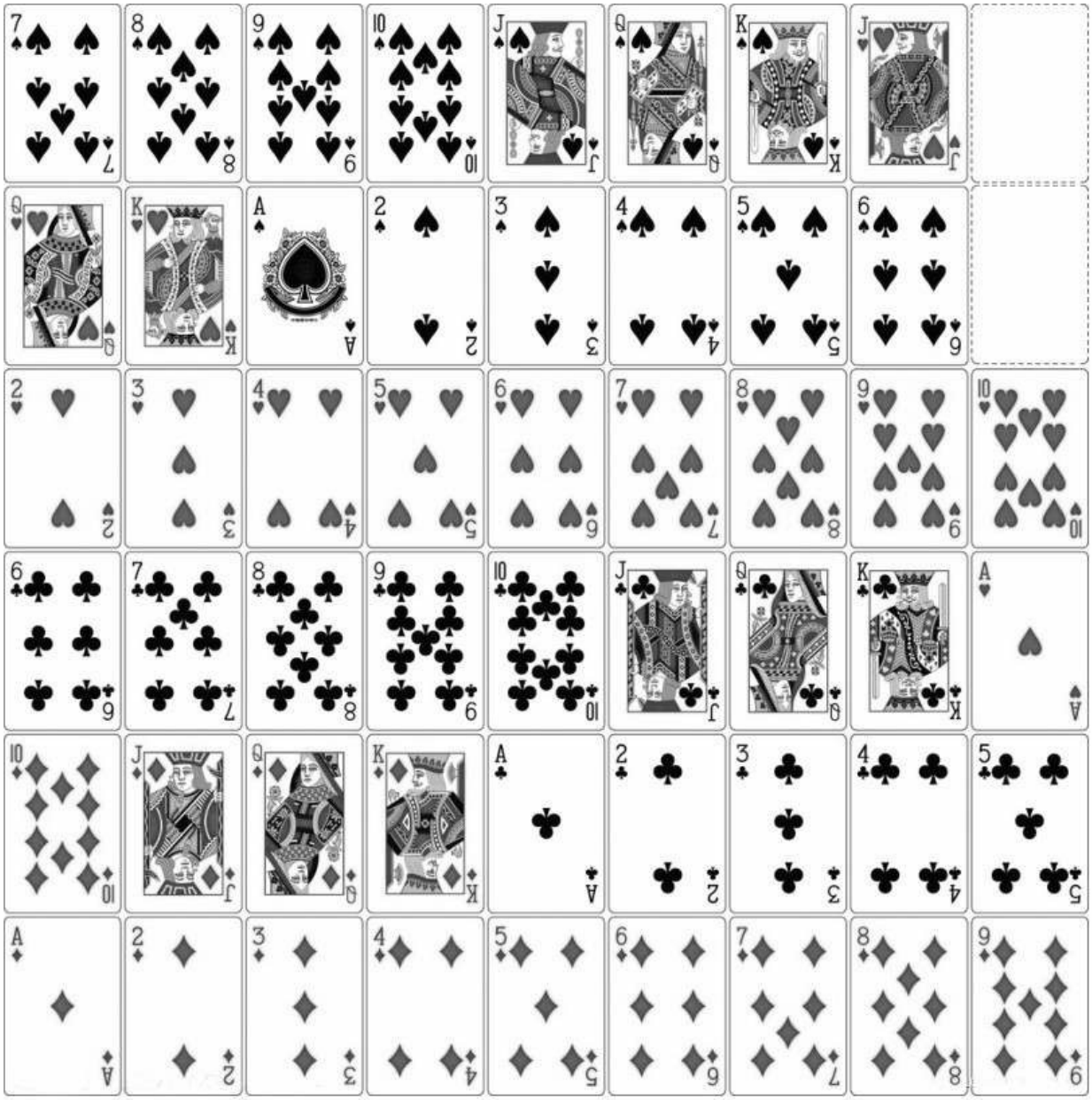


图3-9

在进行游戏时，你可以认为这副牌是首尾相接的。但在最后的布局中，第一张牌必须是“A”，最后一张牌必须是“Q”。

由于洗牌和移牌的随机性，这个扑克迷宫其实就是一个“动态迷宫”，其通路不但与初始布局有关，也同你每一步的走法有关，它时刻都在变化。所以，这个迷宫游戏虽玩起来方便有趣，但对人的智力考验更加苛刻。

立体迷宫简介

立体迷宫的首创者大概是英国著名的数学家和作家、《艾丽丝梦游仙境》的作者刘易斯·卡罗尔，这位多才多艺又爱思考的数学家设计的一个十分复杂的立体迷宫如图3-10所示。大家仔细看一下就会发现，这个迷宫虽然没有很强的立体感，但其中许多路径不是平面相交而是立体相交的，这开创了立体迷宫的先河。这个迷宫外围有7个入口，同中央的菱形区域相连的路径有8条，但实际上只有一个入口、一条路径能让你到达迷宫中心。你能找出这个入口和这条路径吗？

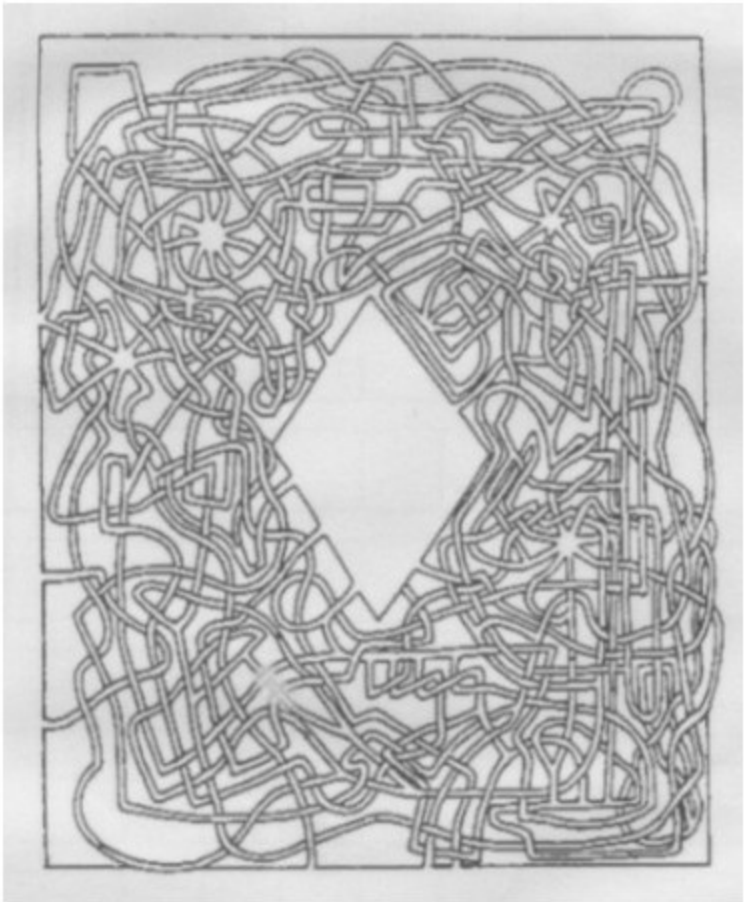


图3-10

当代数学家、IBM公司的研究员庇考夫设计过一个立体数学迷宫。在这个立方体的栅格上有带数字的小球，要求从标有21的这个球出发，在其中找出一条路径，满足从3这个球出来，且经过的小球上的数字之和为202 （如图3-11所示）。

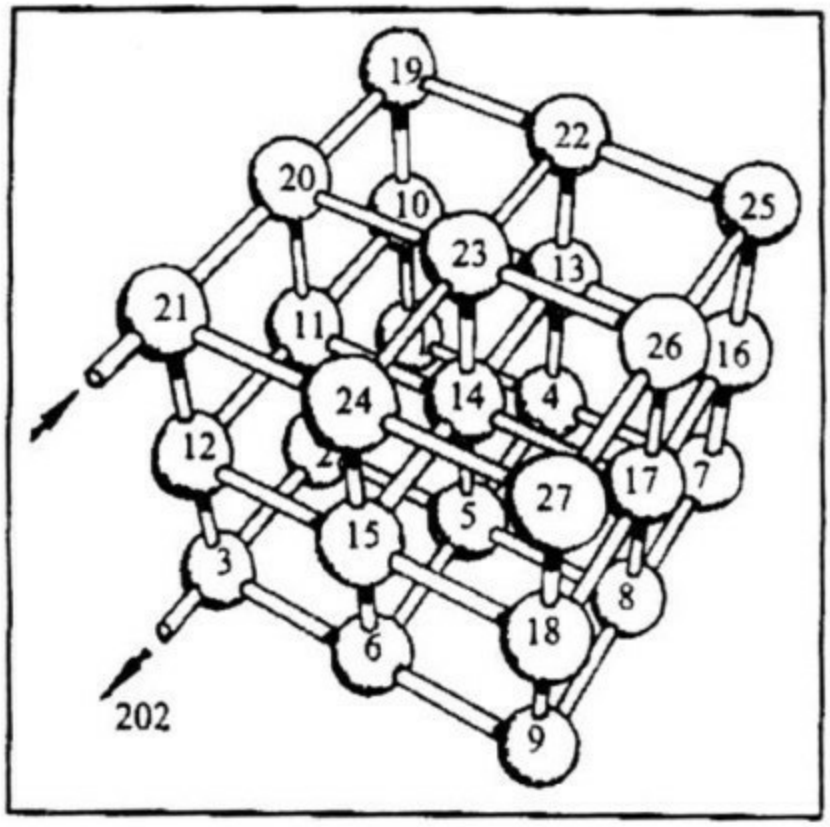


图3-11

马丁·加德纳设计了一个正立方体迷宫，如图3-12所示。这个迷宫由4×4×4（=64）个“小立方体”组成，分A、B、C、D四层。图中粗黑线表示有“墙”，相邻小方块被隔断，不能通行；涂黑部分表示小方块有“地板”，不能从下面的小方块进入到上面的小方块中；最顶上的A这一层当然是有“天花板”的，图上没有表示出来。现在请你在这个迷宫中找出一条通路来，满足从D的一角进去，从与它相对的一角A出来。

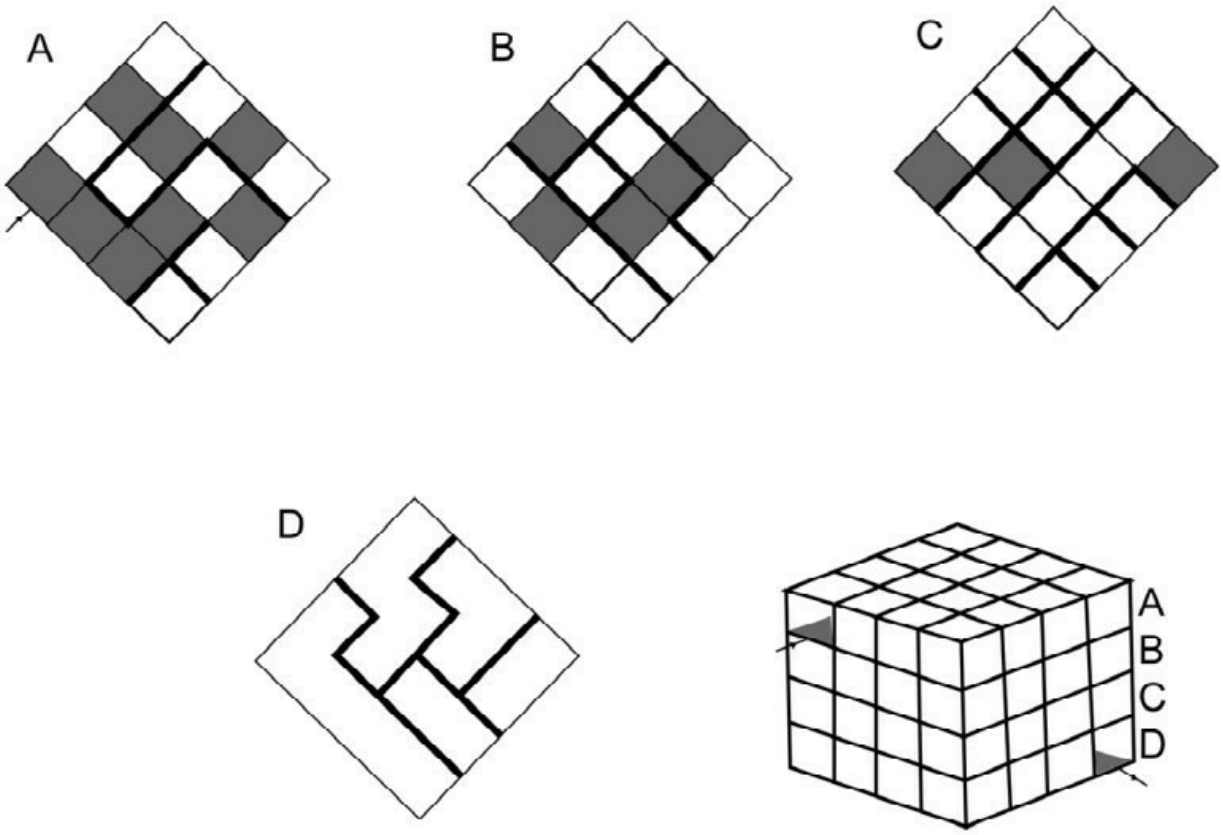


图3-12

以上这3个立体迷宫你能解出来吗？答案见附录二。最后要说明的是：立体迷宫不存在统一解法。

第四章 繁花规中的曲线

在19世纪70年代，市场上出现了一种引人入胜又具有教育意义的玩具，很快它便成为风靡一时的事物，那就是繁花规。它包含一组尺寸不同、边缘有锯齿的小塑料圆轮，以及两个在内边缘和外边缘都有锯齿的大圆环，且在每个小圆轮上离中心不同距离的地方都穿有小孔。如果把一个圆环放在纸上，再放上一个圆轮，使两者的锯齿吻合，然后在一个小孔中插入一支笔。当你沿着圆环移动轮子时，纸上就会出现一条曲线。图4-1所示的是国外的一个繁花规及操作示意图，国内的繁花规虽略有不同，但原理是相同的。

图4-1

《最强大脑》第五季第四期出现了6个由繁花规画出的不同图案，并且这6个图案的中心是合在一起的。屏幕上繁花似锦，眼花缭乱，选手们须逐一分辨出这6个图案的几何参数。

内摆线

已知半径为R的圆O和半径为r的动圆O_1（这里$r<R$），当圆O_1在圆O内无滑动地滚动时，圆O_1上任一点M的轨迹叫作内摆线。

现在来推导内摆线的方程。如图4-2所示，选取直角坐标系，设定圆心O为原点，并且使圆O与x轴正半轴的交点A是轨迹的起点。当圆O_1滚动到圆中的任意位置时，点M从初始位置A（R，0）移动到位置（x，y）。设此时两圆的接触点为N。因为$r<R$，点O_1必在线段ON上，记$\angle NO_1M=t$，则大圆的$\overset{\frown}{NA}$和小圆的$\overset{\frown}{NM}$的长度都是rt，$\angle AON=\dfrac{rt}{R}$。向量$\overrightarrow{O_1M}$的幅角$\alpha=t-\dfrac{r}{R}t$，由$\overrightarrow{OM}=\overrightarrow{OO_1}+\overrightarrow{O_1M}$可得，向量$\overrightarrow{OM}$的坐标可表示为

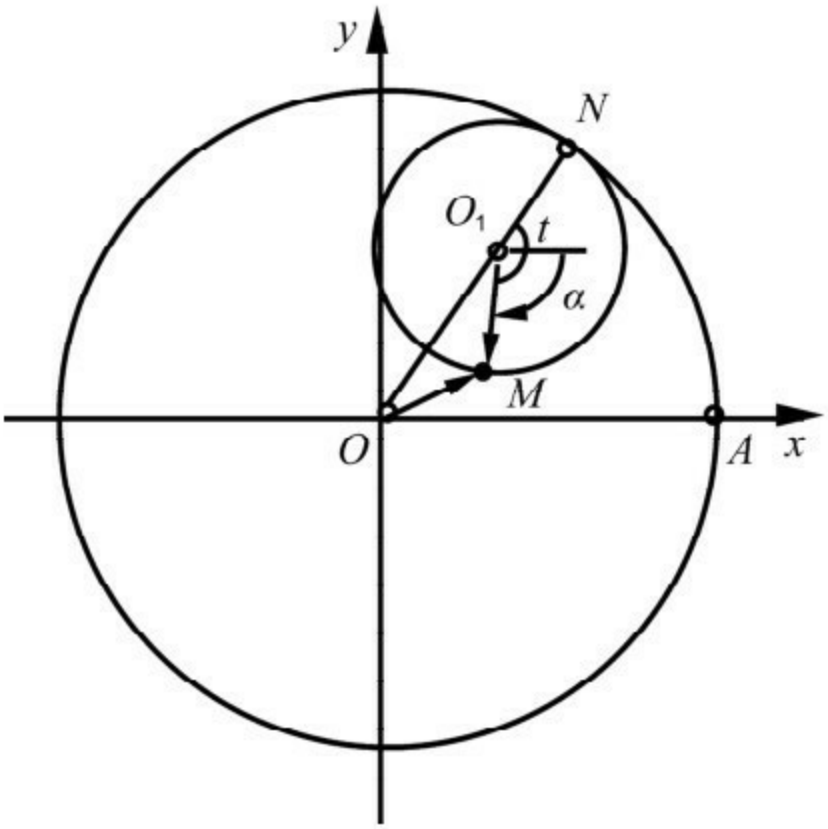


图4-2

$$(x,\ y)=\left[(R-r)\cos\angle AON,(R-r)\sin\angle AON\right]+(r\cos\alpha,r\sin\alpha)$$

$$=\left[(R-r)\cos\frac{r}{R}t+r\cos(t-\frac{r}{R}t),(R-r)\sin\frac{r}{R}t+r\sin(t-\frac{r}{R}t)\right]$$

令$\frac{r}{R}=m$，就得到内摆线的参数方程

$x = (R - mR)\cos(mt)+mR\cos(t - mt)$

$y = (R - mR)\sin(mt)+m R\sin(t - mt)$ $(0<m<1)$

当取0与1之间的不同数时，可以得到各种不同形状的内摆线，如图4-3所示。

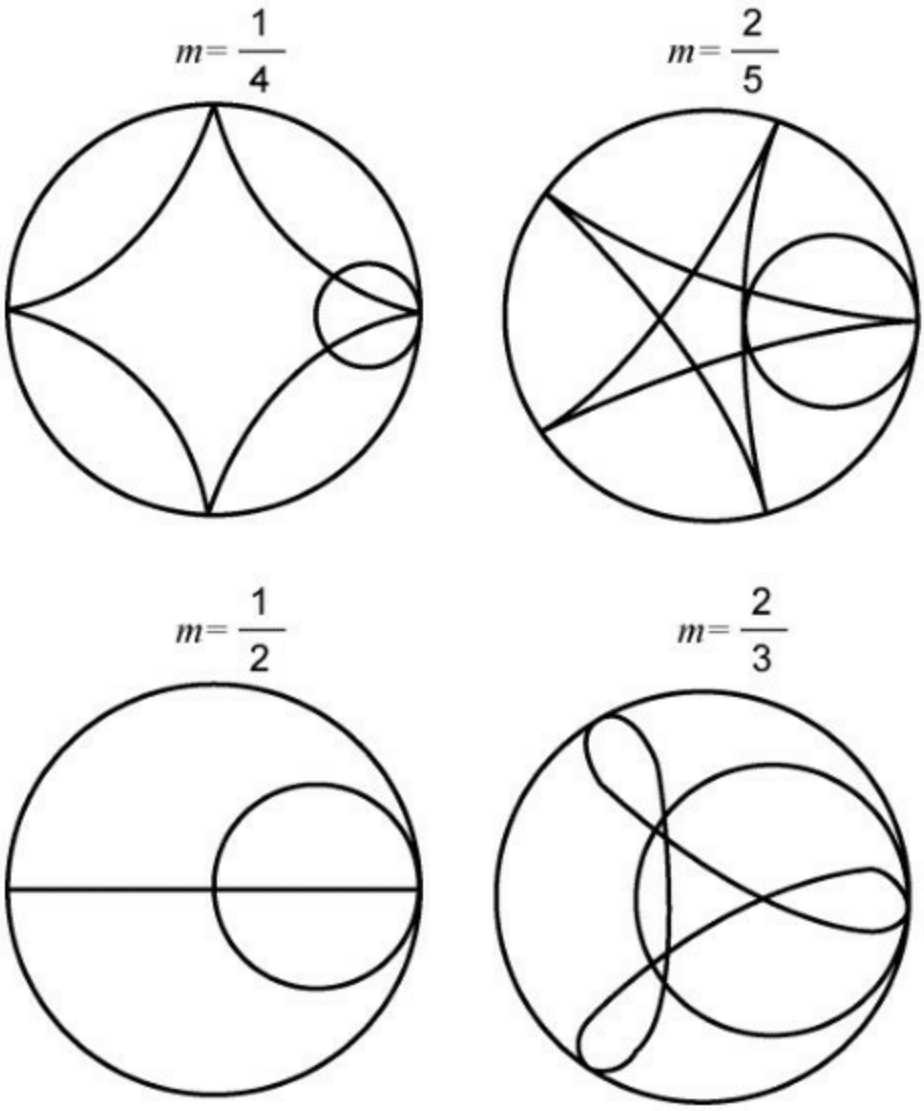


图4-3

细心的读者可能发现，我们在推导内摆线参数方程时，指的是动圆圆圈上点的轨迹。而在用繁花规画曲线时，总是要把铅笔放在动圆上的一个小孔里，所以它不可能是圆圈上的一点。这是不是内摆线轨迹呢？

设$r<R$，当半径为r的动圆O_1在半径为R的圆O内无滑动地滚动时，在圆O_1平面内并与圆O_1固定联结但不在圆O_1圆周上的一点M的轨迹叫作变幅内摆线。变幅内摆线可分成两类：当点M在圆O_1外部时，曲线称为长幅内摆线；当点M在圆O_1内部时，称为短幅内摆线。内摆

线和变幅内摆线统称为内摆线族曲线。

下面推导变幅内摆线的方程。如图4-4所示，取动圆的圆心O 为坐标原点，并且选取x 轴的方向，使动圆的圆心落在x 轴的正半轴上，此时点M 的位置离O 点最远。设动圆O_1 滚动到图4-4中的任意位置，这时点M 的坐标是（x, y），并且两圆的切点是N。若记$O_1M=h$，仿内摆线参数方程的求法，可得曲线的参数方程

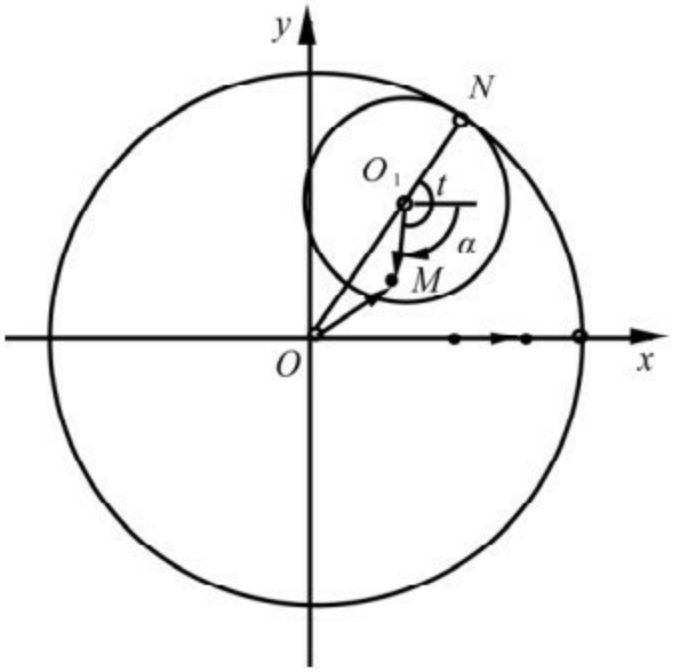


图4-4

$$x=(R-mR)\cos(mt)+hR\cos(t-mt)$$

$$y=(R-mR)\sin(mt)-hR\sin(t-mt)\quad(0<m<1)$$

上式中$m=\frac{r}{R}$。这里由于$0<r<R$，有$0<m<1$。当$h>mR$ 时，得到长幅内摆线；当$h<mR$ 时，得到短幅内摆线；若$h=mR$，则得到内摆线。所以，上述方程是内摆线族曲线的统一方程。图4-5中从左至右分别是$m=\frac{1}{4}$ 时的长幅内摆线、内摆线（星形线）和短幅内摆线。

由此知道，《最强大脑》比赛现场的屏幕上呈现的均是短幅内摆

线。

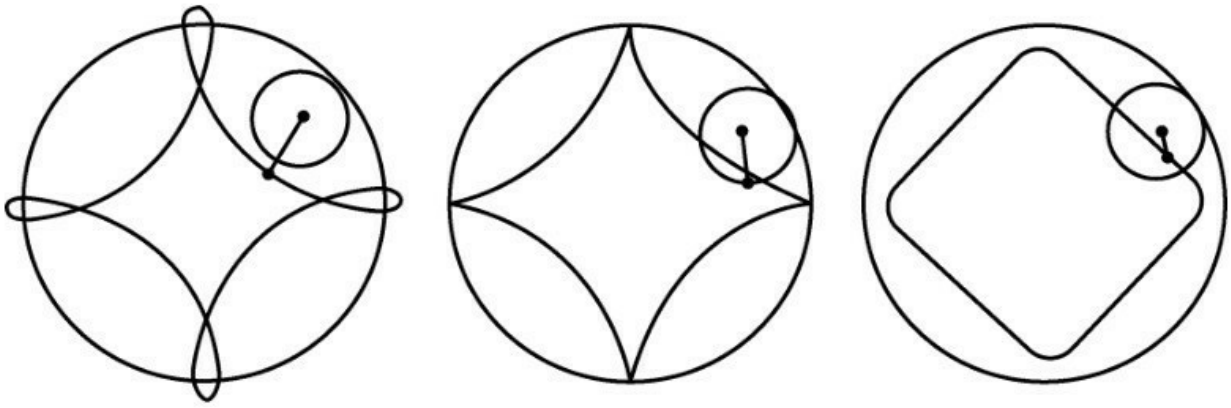


图4-5

当m （设$m=\frac{p}{q}$ ，且p 、q 互质）小于1即为真分数时，内摆线由q 支组成，动圆绕定圆p 周后即可返回初始位置，并描完q 支内摆线；当m 为小于1的无理数时，内摆线有无穷多支，此时无论绕多少圈，它都不能回到初始位置。

我们只需观察繁花有几个花瓣，即确定q ，因$\frac{p}{q}=\frac{r}{R}$ 且定圆直径R 是已知的，通过估算与q 互质的p ，r 也就得到了。这里还有个经验：花瓣越窄，则r 越大；花瓣越宽，则r 越小。所以只要选手们了解“短幅内摆线”的几何特性，就很容易胜出。例如图4-1中画出的曲线有8个花瓣（虽然未完整显示），花瓣较窄，可估计p =7。而定圆直径约为32mm，繁花曲线的直径为28mm，符合比例$\frac{7}{8}=\frac{28}{32}$ （读者可自己量一下）。

公共汽车的门

当$r=R$ 时的内摆线有4个尖角，如图4-5中间的图形，像夜空中光芒四射的星，因此叫作星形线。在内摆线的参数方程中，以$m=\frac{1}{4}$ 代入，就得到星形线的参数方程

$$x=\frac{3}{4}R\cos\frac{t}{4}+\frac{R}{4}\cos\frac{3t}{4}$$

$$y=\frac{3}{4}R\sin\frac{t}{4}-\frac{R}{4}\sin\frac{3t}{4}$$

令$t=4\varphi$，并利用三倍角公式

$$\cos(3\varphi)=4\cos^3\varphi-3\cos\varphi$$

$$\sin(3\varphi)=3\sin\varphi-4\sin^3\varphi$$

可将星形线的参数方程简化成

$$x=R\cos^3\varphi, y=R\sin^3\varphi$$

我们可以用画包络图的方式来得到星形线。如图4-6所示，任意作若干条长度为R 的线段，使它们的两端分别在x 轴和y 轴上；然后在每个象限里画一段光滑曲线弧，使它们与这些线段相切，就得到了所要画的星形线。通过简单分析，可知在第一象限内的包络图方程正是

$$x=R\cos^3\varphi, y=R\sin^3\varphi$$

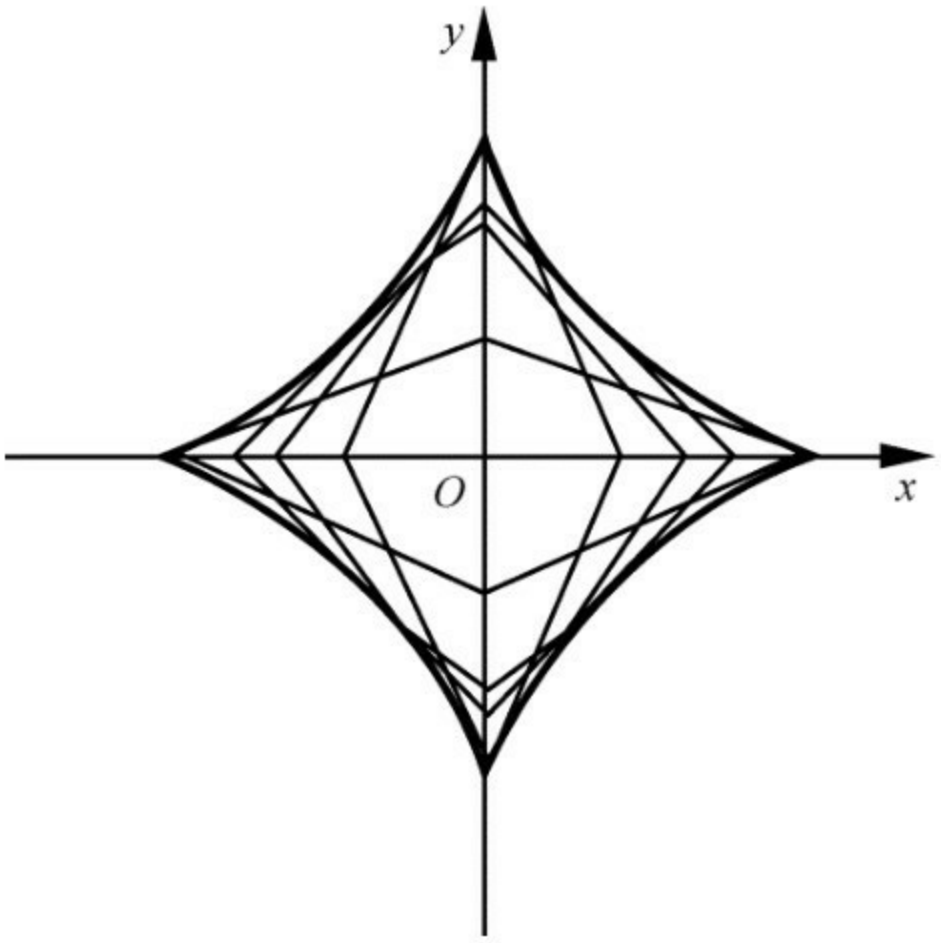


图4-6

大家知道，有一种公共汽车的门很特殊：它是对开的两扇，而且每一扇都由相同的两半用铰链铰接而成。在开门和关门时，靠近门轴的半扇绕着门轴旋转，另外半扇的外端沿着两个门轴的滑槽滑动，开门时一扇车门折拢为半扇，关门时又重新伸展为一扇。

公共汽车采用这种折叠式的车门有一个优点，就是车门开关时所需的活动范围比较小，在上下班高峰时能多载乘客。图4-7是该类车门的简化图， O 是门轴，$OA=AB=a$ 。取点O 为坐标原点，滑槽OB 为x 轴，延长BA 交y 轴于点C 。在△AOB 中，$AO=AB$ ，所以$\angle AOB=\angle ABO$，又因△BOC 是直角三角形，由此推出

$\angle AOC=90°-\angle AOB=90°-\angle ABO=\angle ACO$

故$AC=AO=a$，因而$CB=2a$（定长）。根据上述星形线的画法，当定长线段BC的两端分别沿x轴和y轴滑动时，其一切位置的包络正是星形线。由于车门只在第一象限活动，所以动线段BC的包络只是该星形线位于第一象限内的一段；这段弧又被第一象限的分角线分成上下两部分，根据对称性，半边车门AB活动时的包络只是该星形线弧的下半部分。

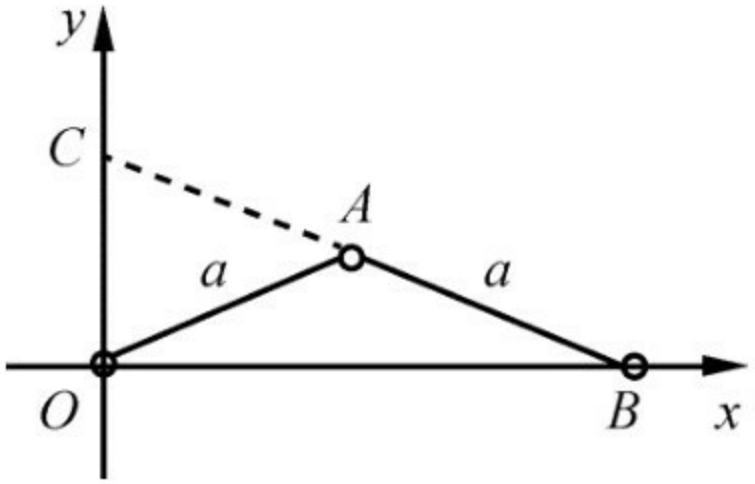


图4-7

由此知道，一扇折叠式车门所需要的活动范围在水平面内的投影是由图4-8中的圆弧$\overset{\frown}{MN}$、星形线弧$\overset{\frown}{NP}$和坐标轴围成的区域。其面积可分成3部分：扇形OMN、等腰直角三角形OQN和曲边三角形QNP。可顺次记为S_1、S_2、S_3，容易得到

$$S_1=\frac{1}{8}\pi a^2,\quad S_2=\frac{1}{2}\left(\frac{\sqrt{2}}{2}a\right)^2=\frac{1}{4}a^2$$

而S_3的计算则要用到积分的一些知识，我们直接写出它的结果$S_3=\frac{3}{16}\pi a^2-\frac{1}{4}a^2$。

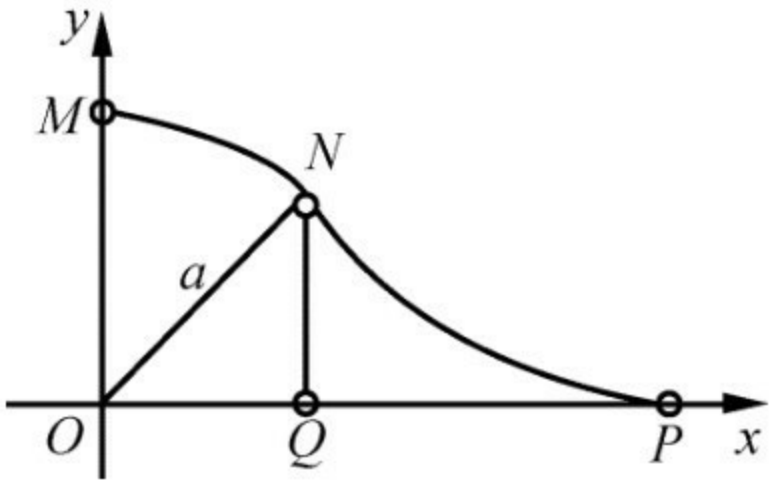


图4-8

总的有

$$
\begin{aligned}
&= S_1 + S_2 + S_3 \\
&= \frac{1}{8}\pi a^2 + \frac{1}{4}a^2 + (\frac{3}{16}\pi a^2 - \frac{1}{4}a^2) \\
&= \frac{5}{16}\pi a^2
\end{aligned}
$$

而一扇宽度为$2a$ 的普通门所需的面积为πa^2（圆面积的$\frac{1}{4}$），因而一扇折叠式活门所占的地方只有普通门的$\frac{5}{16}$。

外摆线

已知半径为R的定圆O和半径为r的动圆O_1，当圆O_1在圆O外无滑动地滚动时，圆O_1上一点M的轨迹叫作外摆线。

如图4-9所示，用与推导内摆线参数方程相类似的方法，可得外摆线的参数方程是

$$x=(R+mR)\cos(mt)-mR\cos(t+mt)$$

$$y=(R+mR)\sin(mt)-mR\sin(t+mt)$$

上式中$m=\dfrac{r}{R}$，且为正数。

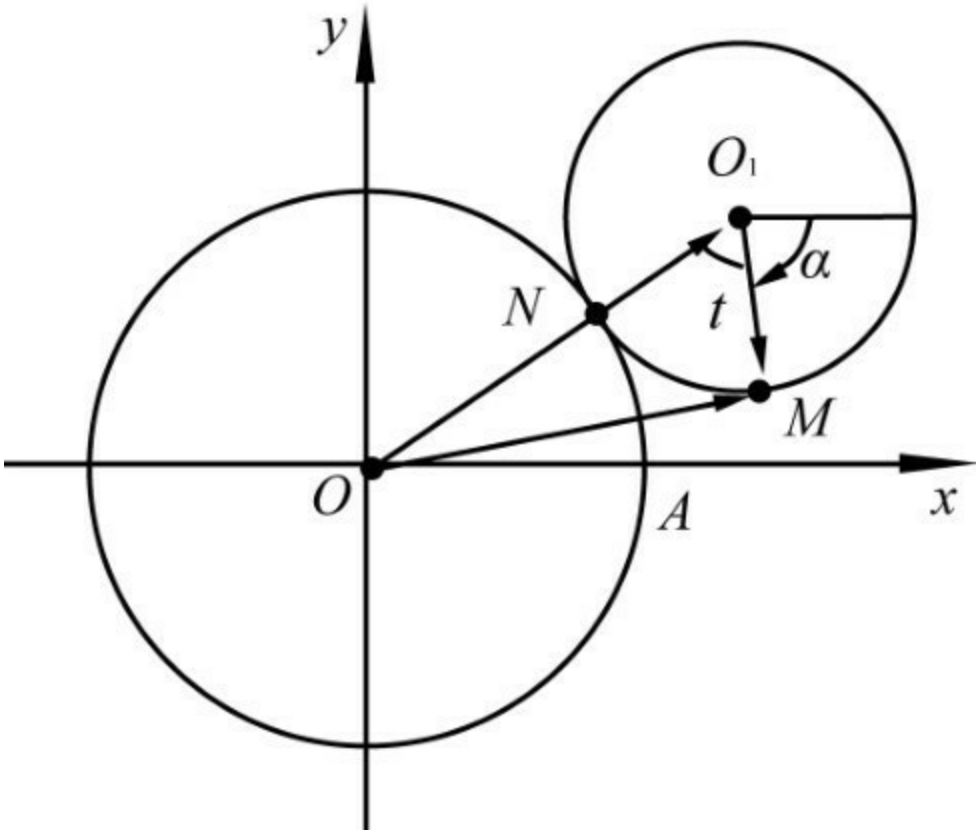


图4-9

类似地，若M 是固定在圆O_1 平面内的一点，但不在圆O_1 的圆周上。当圆O_1 在圆O 外绕圆O 无滑动地滚动时，点M 的运动轨迹叫作变幅外摆线。当点M 在动圆O_1 内部时，称为短幅外摆线；当点M 在动圆O_1 外部时，叫作长幅外摆线。

设点M 到点O_1 的距离为h，并令$\frac{r}{R}=m$，可得到曲线的参数方程为

$$x=(R+mR)\cos(mt)-h\cos(t+mt)$$

$$y=(R+mR)\sin(mt)-h\sin(t+mt)\ (m>0)$$

当$h>mR$ 时，曲线为长幅外摆线；反之，当$h<mR$ 时，曲线为短幅外摆线；

当$h=mR$ 时，上述方程可化为外摆线方程。故上述方程统称为外摆线族曲线方程。

在图4-10中，从左至右分别是$m=\frac{1}{4}$ 时的长幅外摆线、外摆线和短幅外摆线。

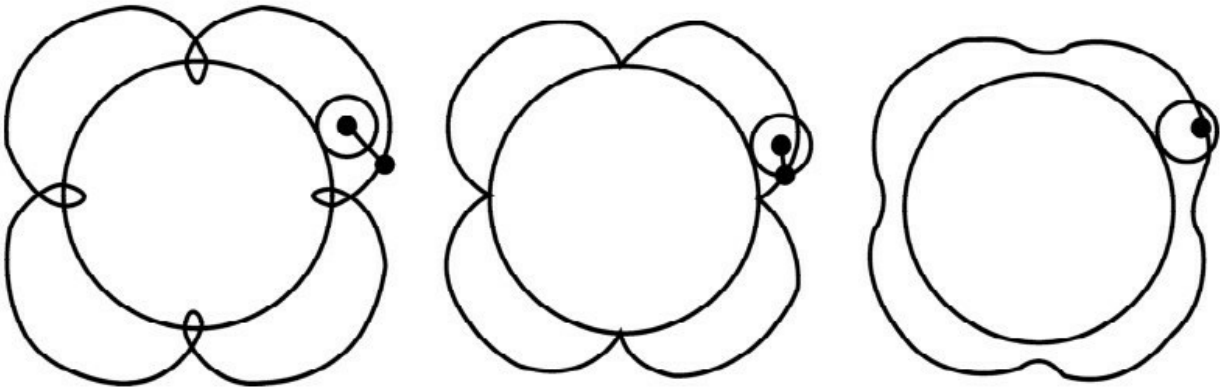


图4-10

内外摆线是一家

我们先回忆一下内摆线族曲线的参数方程

$x=(R-mR)\cos(mt)+h\cos(t-mt)$

$y=(R-mR)\sin(mt)-h\sin(t-mt)$

其中m 值应在0与1之间。现在我们倒过来考虑，从上面的参数方程出发，给出常数R 、h 和m 之后，就可用描点法画出对应的曲线。如果我们在这个方程中令m 取大于1的数，结果画出的曲线会是怎样的形状呢？

为此，我们限定$m>1$，并令$m_1=m-1$，$h_1=mR-R\cdot R_1=\frac{h}{m}$ ，则有$m_1>0$，$h_1>0$，$R_1>0$，并且$m=m_1+1$，$h=mR_1=R_1+m_1R_1$。将这些系式代入上面的参数方程中，得到

$x=-h_1\cos(t+m_1t)+(R_1+m_1R_1)\cos(-m_1t)$

$y=-h_1\sin(t+m_1t)-(R_1+m_1R_1)\sin(-m_1t)$

即

$x=(R_1+m_1R_1)\cos(m_1t)-h_1\cos(t+m_1t)$

$y=(R_1+m_1R_1)\sin(m_1t)-h_1\sin(t+m_1t)$

其中R_1、h_1和m_1都是正数。这正是上面我们讲过的外摆线族曲线方程。进而，当$h=mR$时，有

$$R_1=\frac{h}{m}=\frac{mR}{m}=R$$

$$h_1=mR-R=(m-1)R=m_1R_1$$

因而这时得到的是外摆线。当$h>mR$时，有

$$R_1=\frac{h}{m}>\frac{mR}{m}=R$$

$$h_1=(m-1)R=m_1R<m_1R_1$$

因而这时得到的是短幅外摆线。当$h<mR$时，有

$$R_1=\frac{h}{m}<\frac{mR}{m}=R$$

$$h_1=(m-1)R=m_1R>m_1R_1$$

因而这时得到的是长幅外摆线。

由此可见，内外摆线本质上是一家。内摆线族曲线和外摆线族曲线都属于一个更大的曲线族，它们的方程可以统一写成

$x=(R-mR)\cos(mt)+h\cos(t-mt)$

$y=(R-mR)\sin(mt)-h\sin(t-mt)$

其中R、h、m都是正数。当$0<m<1$时得到内摆线族曲线，$m>1$时得到外摆线族曲线，$m=1$时曲线退缩为一点（h，0）。

摆线

上述的定圆在极限情况下可变成直线。当圆C 沿着定直线L 无滑动地滚动时，动圆圆周上一点M 的运动轨迹所形成的曲线叫作摆线（如图4-11所示）。摆线在它与直线L 的两个相邻交点之间的部分叫作拱，摆线最高点到定直线的距离（$2r$ ）叫作拱高。

如图4-11所示，取定直线L 为x 轴，并使动圆位于x 轴上方。设动圆圆心在y 轴上时，点M 恰好在坐标原点。记圆的半径为r ，则当它滚动到图中位置时，由于$\overrightarrow{OM}=\overrightarrow{OC}+\overrightarrow{CM}$ ，$\alpha=\frac{3}{2}\pi-\theta$ ，所以点M 的坐标可由向量$\overrightarrow{OM}$ 的坐标表示或给出。

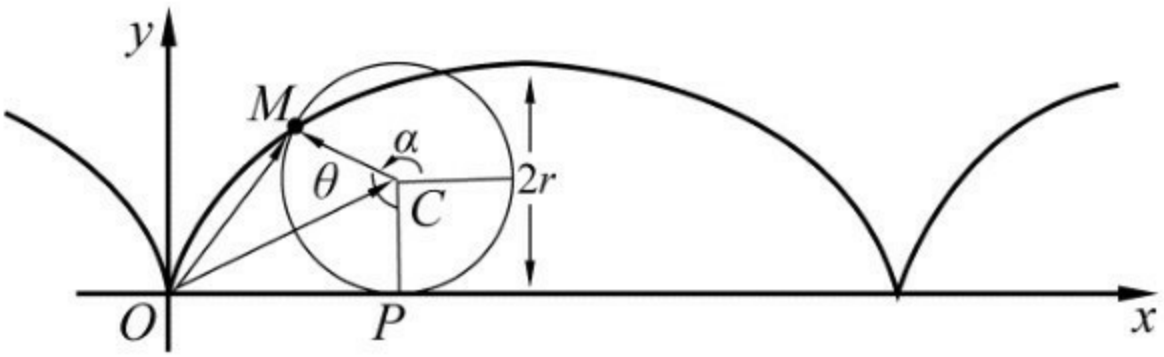


图4-11

$(x\,,y\,)=(OP\,,PC\,)+(CM\cos\alpha\,,CM\sin\alpha\,)$

$=(r\theta\,,r\,)+(-r\sin\theta\,,-r\cos\theta\,)$

$=(r\theta-r\sin\theta\,,r-r\cos\theta\,)$

即M 点的坐标是$x=r\,(\theta-\sin\theta\,)$，$y=r\,(1-\cos\theta\,)$。当$\theta$ 从0变化到2π时，动点M 描绘出摆线的一拱。

当点M 不在动圆的圆周上时，点M 所画出的曲线叫作变幅摆线。变

幅摆线可分成两类：当点M 在圆C 内部时，曲线称为短幅摆线（如图4-12上图所示）；当点M 在圆C 外部时，曲线称为长幅摆线（如图4-12下图所示）。

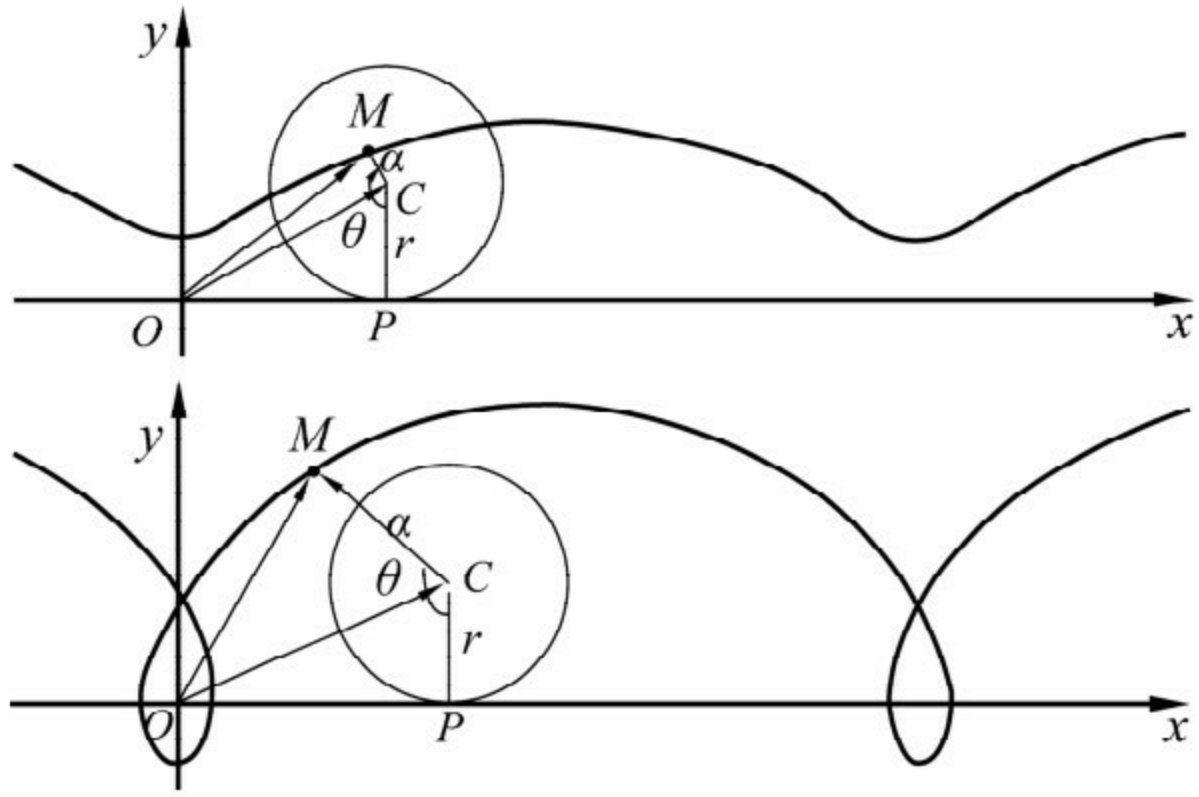


图4-12

变幅摆线的参数方程是

$x = rt - a\sin t$，$y = r - a\cos t$

当$a < r$ 时得到短幅摆线，$a > r$ 时得到长幅摆线，$a = r$ 时为普通摆线。

惠更斯的摆线时钟

图4-13中有一个光滑的摆线槽。取一颗钢珠，放在摆线槽中的任意位置（例如图中的M 处），当手指松开，滚珠就像荡秋千一样，沿着摆线槽来回摆动。试选择不同的高度松开滚珠，并注意观察滚珠连续两次通过摆线槽最低点K 的时间间隔，你就会发现一个奇怪的现象：尽管M 点位置的高低不同，但是滚珠每来回摆动一次所花的时间竟没有变！这个性质叫作摆线的等时性，故摆线又称为等时曲线，摆线的这一特性是17世纪荷兰物理学家惠更斯发现的。

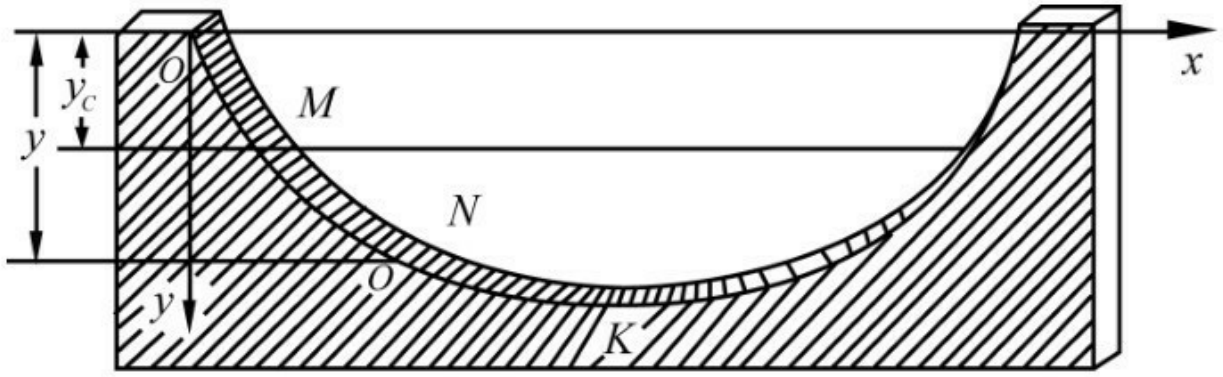


图4-13

通过计算可以得到，滚珠从初始位置点M 下降到摆线槽最低点所用的时间是常数π，与初始位置无关。因而不论摆动幅度有多大，滚珠沿摆线槽来回摆动一次所用的时间都相等。

钟摆在滴答声中来回摆动，每次摆动所用的时间必须严格相等才能保证计时的准确性。惠更斯在发现了摆线的等时性以后，就利用它设计出了一种摆线时钟，它的构造如图4-14所示。

把摆锤悬挂在长为$4r$ （摆线半拱的弧长）、能弯曲但不能伸长的细

线上，悬挂点的两边各放一块挡板，使挡板的曲线轮廓是拱高为$2r$ 的半拱摆线，这样就能保证摆锤沿着图中虚线所示的摆线弧摆动。无论摆幅大小如何，摆锤沿该摆线弧来回摆动一次所用的时间都是$4\pi\sqrt{\frac{r}{g}}$ 。

惠更斯发明的这种摆线时钟一问世便大受欢迎。这种曲线正是由于被应用于改进钟摆，才被正式命名为“摆线”，而惠更斯也因此芳名永留。

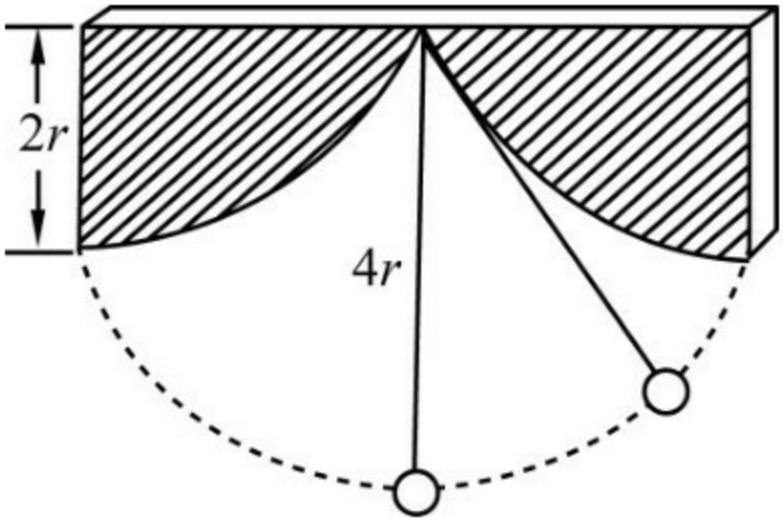


图4-14

速降线

1696年6月，瑞士数学家约翰·伯努利在《教师学报》上提出了脍炙人口的所谓“速降线”问题：

A 、B 两点不在同一直线上，且A 点较高，求一条曲线AMB ，使得质点仅在自身重力的作用下，沿此路径由A　点滑至B　点所耗的时间最少。

这个问题看上去是个很普通的极值问题，但只要仔细一分析，就知道它非同寻常。在人们熟知的极值问题中函数是已知的，而现在是不知道函数的表达式，倒过来要求出一种函数，使得整个式子达到极小值。很显然，这是一个崭新的问题。伯努利指出，这一问题暴露了普通几何学的局限性，应由此题为契机，创造出一种崭新的数学方法。

经过一番努力，这个问题还是被解出来了。1697年的春天，《教师学报》上出现了好几篇关于速降线的解答论文，其中伯努利兄弟有两篇，莱布尼兹和洛必达各一篇。答案就是倒放的摆线弧，当质点沿摆线弧滑下时，比任何其他路径都快。

当时牛顿也匿名发表了解答，约翰·伯努利开始并不知道是牛顿的解答，佩服得五体投地，对哥哥雅各布·伯努利说：“我们周围有一头科学的雄狮，这篇速降线的论文只是他露出的一条尾巴。”

正是由于牛顿的匿名，才有了后来关于这一问题发现权的争议和对剽窃者的指责。这样，作为一种结果，摆线一时被贴上了“几何的海玲”的标签。这是一则引自希腊神话的故事，海玲是Zeus与Leda之女，因被Panis所拐而引起了特洛伊战争，所以有“祸根”之意，这里暗指摆线是引发争议的祸根。

随着对速降线等问题的研究，“一种崭新的数学方法”被创造出来了，这种新的数学方法名曰“变分法”。

第五章 魔方与数学

魔方是一些竞技类节目中常用的道具，《最强大脑》也不例外。第一季中的“魔方墙找茬”“水下盲拧魔方”，第二季中的“魔方工厂”，第三季中的“魔方速拧”“魔方盲拧”等都无不与魔方有关。但这些都是在考察选手们的观察力、记忆力以及拧魔方的熟练程度。魔方与数学又有怎样的关系呢？

魔方简史

魔方是匈牙利布达佩斯应用艺术学院的建筑学教授鲁比克发明的。鲁比克最初想发明的并不是益智玩具，而是一个能演示空间转动，帮助学生直观理解空间几何的教学工具。经过一段时间的思考，他决定制作一个由小方块组成，各个面均能随意转动的3×3×3结构的立方体。

但如何才能让立方体的各个面既能随意转动，又不会散架呢？这一问题让鲁比克陷入了苦思。1974年一个夏日的午后，他正在多瑙河畔乘凉。当鲁比克的目光无意间落到河畔的鹅卵石上时，他忽然灵光乍现，想到了解决问题的方法，最后设计出了一个结构巧妙的十字轴作为旋转中心。十字轴有一个六向接头，每一头分别连接着6个中心块，8个角块和12个边块依次镶嵌在旋转中心上，便组成了一个完整的立方体。这时它就可以按横列或纵列绕中心任意旋转，构造变化无穷的图案了。

鲁比克在1975年为自己的发明申请了专利，并决心大量生产这种玩具。1980年，在纽伦堡的国际玩具展览会上，魔方（如图5-1所示）以它巧妙的结构、千变万化的图案和其中包含的深奥的数学原理而轰动一时，随后被评为1980年世界最佳玩具，不久便风靡世界。

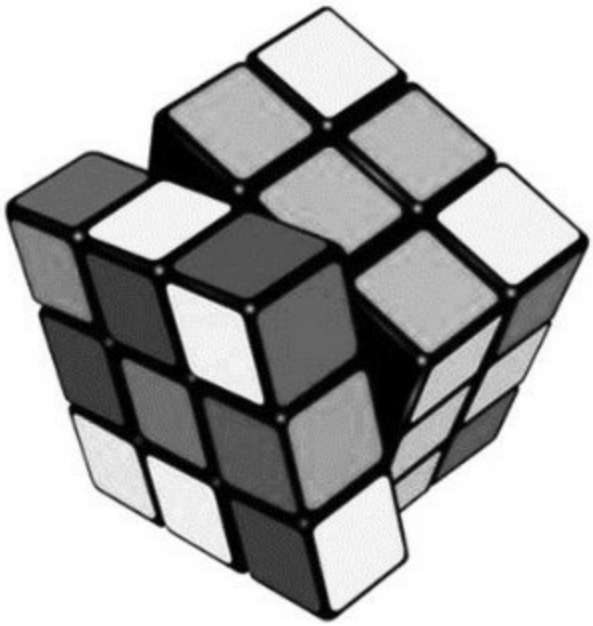


图5-1

不过对魔方的理论研究比这还要早两年，1978年在芬兰首都赫尔辛基召开的国际数学家大会上，一位匈牙利数学家带的几个魔方吸引了与会的专家、学者。其中有一位英国数学家辛马斯特，他对魔方更为着迷，于是就向匈牙利教授要了一个，回国后全力研究，后出版了一本《鲁比克魔方解法》的专著，这也是最早的关于魔方研究的理论专著。

据说德国数学家希尔伯特曾经表示，学习群论的窍门就是选取一个好的例子。自辛马斯特后，数学家们写了好几本通过魔方讲述群论的书。因此，魔方在一定程度上可以作为学习群论的“好的例子”。那么，什么是群论呢？

魔方与群论

群论是由法国数学家埃瓦里斯特·伽罗瓦（1811—1832）在研究高次代数方程的求解问题中创立的。在这之前，挪威数学家尼尔斯·阿贝尔（1802—1829）证明了4次以上代数方程不能通过根式求解，但他的证明方法十分复杂。伽罗瓦另辟蹊径，利用对每个方程构造一个置换群的思想与方法，彻底解决了代数方程通过根式求解的问题，这就是所谓的“伽罗瓦理论”，并从而开辟了一个新的代数领域——群论。

群有多种类型，但它们有一个基本定义：设G 是一个非空集合，*是它的一个运算，如果满足以下条件，则称G 对*构成一个群。若群中元素个数是有限的，则是有限群，否则为无限群。有限群的元素个数称为有限群的阶。

①封闭性：若a 、b $\in G$ （这里的记号∈代表“属于”），则存在唯一确定的$c \in G$ 使得$a*b=c$；

②结合律：即对G 中任意元素a,b,c 都有$(a*b)*c=a*(b*c)$;

③单位元存在：存在$e \in G$，对任意$a \in G$，满足$a*e=e*a=a$，e 称为单位元，也称幺元；

④逆元存在：任意$a \in G$，存在$b \in G$ 使$a*b=b*a=e$ （e 为单位元），则称a 与b 互为逆元素，简称逆元。b 记作a^{-1}。

我们用一个简单的例子来加深大家对群的理解：全体整数的集合G 对于加法“+”构成一个群，这是因为它满足以下条件。

①对于任何两个整数a 和b，它们的和$a+b$ 也是整数，也就是在加法运算下封闭。

②对于任何整数a,b,c，$(a+b)+c=a+(b+c)$，故符合结合律。

③对于任何整数a，有$0+a=a+0=a$，零叫作加法的幺元。

④对于任何整数a，存在另一个整数b，使得$a+b=b+a=0$，整数b叫作整数a的逆元，记为a^{-1}。

下面我们还必须将魔方的各部件及其基本操作予以命名。

魔方共有6个外表面，可分别用其英文名称第一个字母的小写形式来代替，即前（f）、后（b）、右（r）、左（l）、上（u）、下（d）。魔方有8个角，我们把位于各个角落上的方块称为“角块”。由于每个角块包含3个面的一部分（以下称为“小面”），我们可以用这3个小面的代号来命名这个角块，例如包含前、右、上这个小面的角块可以称为“fru”，如图5-2所示，其他的角块命名以此类推。魔方有12条处于外沿的边，我们把位于每条边中间位置上的方块称为“边块”。由于每个边块包含两个小面，可以用这两个小面的代号来命名这个边块，例如包含右、下这两个小面的边块便可以称为“rd”，如图5-2所示。此外，魔方的外表面的6个中心位置还各有一个方块，称为“中心块”，每个中心块只包含一个小面。我们可以用这个小面的代号来命名，例如包含右小面的中心块便可以称为（r），如图5-2所示。请注意，魔方的6个中心块起着识别6个外表面的作用，即在还原魔方的过程中，我们要设法令所有角块和边块包含的每个小面与该小面所在的外表面上的中心块同色。

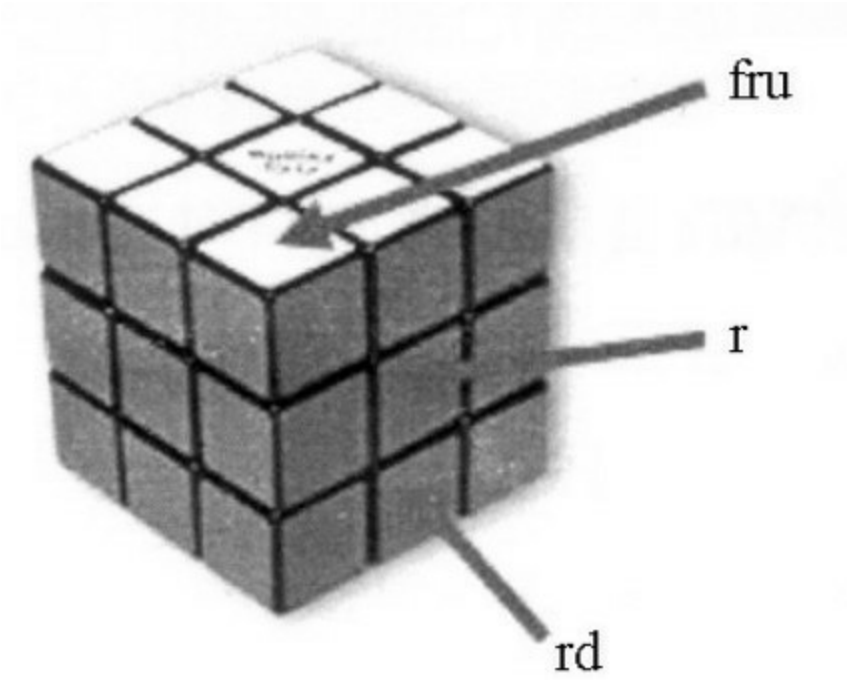


图5-2

魔方的每个外表面都可朝顺时针或逆时针（这里的顺/逆时针是相对于你把待旋转的外表面正对着你而言的）两个方向旋转，这也是玩魔方的基本操作。我们用魔方6个外表面英文名称的第一个字母的大写F,B,R,L,U,D分别代表把魔方的前、后、右、左、上、下等面顺时针旋转90°的动作。而对应的逆时针则是前述6个顺时针旋转的逆向操作，可用F^{-1} ,B^{-1} ,R^{-1} ,L^{-1} ,U^{-1} ,D^{-1} 表示。图5-3显示了上述12个操作。

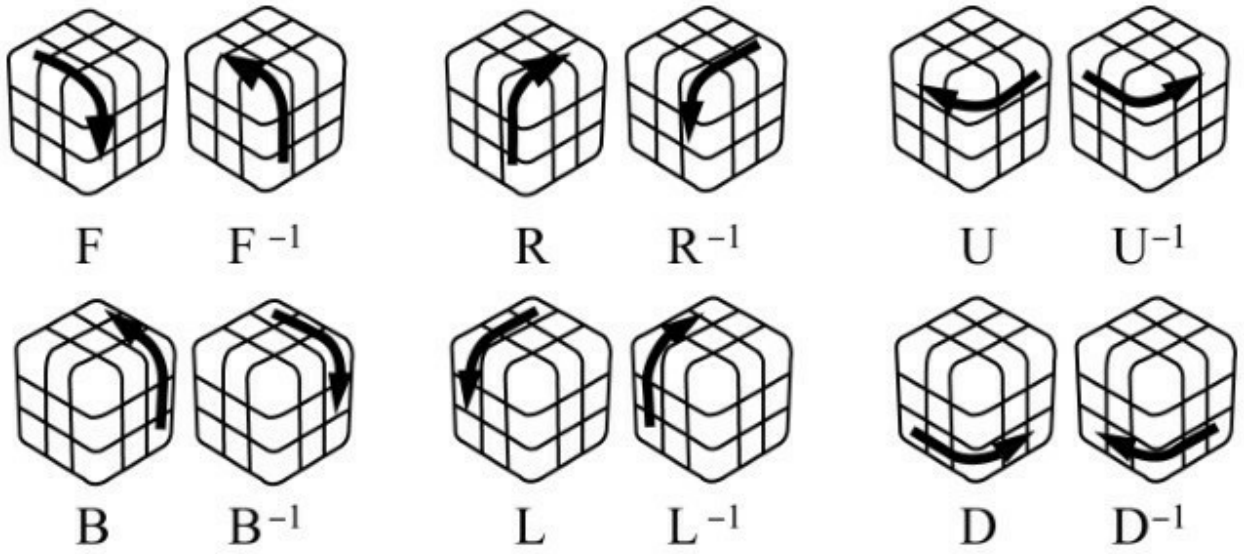


图5-3

知道了群的基本概念和魔方的基本操作后，下面把这两者结合起来讲一下魔方群。

群有许多类型，置换群是其中很重要的一种。我们知道，将n 个不同元素从一种排列变成另一种排列就叫一个置换。例如，把排列$a_1 a_2 a_4$变成$a_4 a_3 a_1 a_2$就是一个置换，记为

$$\begin{pmatrix} a_1 a_2 a_3 a_4 \\ a_4 a_3 a_1 a_2 \end{pmatrix}$$

也可以简单地记为（$a_1 a_4 a_2 a_3$），表示a_1变为a_4、a_4变为a_2、a_2变为a_3、a_3变为a_1。

如果在置换中把原排列的第一个元素换为第二个元素，第二个元素换为第三个元素……，最后一个元素换为第一个元素，则这种置换叫作“轮换”，也叫“循环换位”。如果在置换中只有两个元素互换了位置，其他元素位置不变，这样的置换叫作“对换”。显然，任何一个置换都可以看成连续施加若干个对换的结果。所谓置换群，就是由有限个集合元

素的置换所构成的群。

把数学上关于置换和置换群的概念用到魔方上来，我们看到，魔方任意面的一个q转就形成一个置换；对魔方的各种各样的操作的集合就形成一个置换群，被叫作“魔方群”。所谓1个q转就是：用左手握住魔方，用右手按顺时针方向或逆时针方向旋转一个外表面。旋转角度通常是90°，即一个象限（1 quarter），因此被叫作一个q转。如果外表面一次转180°，就叫2个q转。如顶面顺时针转90°记为U，那么转180°则记为U^2，以此类推。

为了让读者更直观地理解魔方群是置换群，我们用图解的方法介绍如下。先把原始状态的魔方展开成平面图，如图5-4所示，读者可以把小块的纸片粘在魔方小块上，则下述置换看起来就十分形象了。

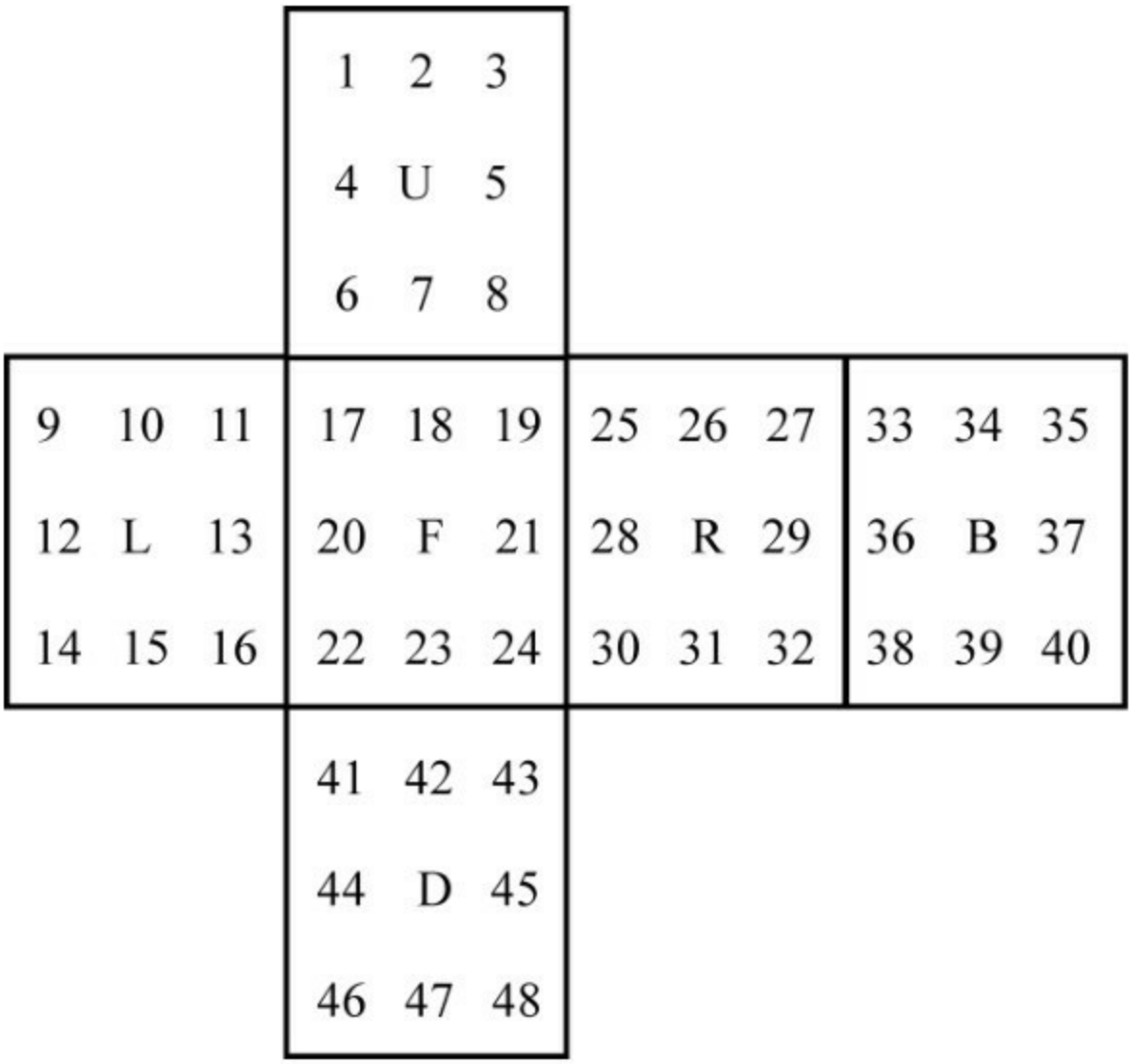


图5-4

现在实现操作FR（即前侧面顺时针旋转90°，接着将右侧面顺时针旋转90°），这时小面的布局发生了变化，如顶面上标号为3的小面跑到背面上原先被标记为38的小面的位置上去了，而标号为38的小面则跑到底面上原先被标号为43的位置，如此等等。其全部情况如表5-1所示，显然，这就是对48个元素（小面）的一个重新排列，即置换。

表5-1 魔方的一个置换

魔方 外侧面	原始状态	执行 FR 以后
U	1，2，3，4，5，6，7，8	1，2，17，4，18，16，13，19
L	9，10，11，12，13，14，15，16	9，10，41，12，42，14，15，43
F	17，18，19，20，21，22，23，24	22，20，25，23，45，24，21，48
R	25，26，27，28，29，30，31，32	8，7，6，31，26，32，29，27
B	33，34，35，36，37，38，39，40	11，34，35，5，37，3，39，40
D	41，42，43，44，45，46，47，48	30，28，38，44，36，46，47，33

由此可见，魔方上置换群的对象可视作对魔方的各种操作过程，也可以视作各小面的置换。类似FR的运算可称为一个“后随”，即把前后两次操作过程连接起来变成一个操作过程，或是说把前后两个操作过程所产生的置换连接起来，变成一个置换。

最后我们要验证上述对象和运算是否满足群的4个条件。

①封闭性。对魔方而言，置换X“后随”置换Y，一定是魔方上的另一个置换XY，这个条件是满足的。

②结合律。例如操作FRD，则（FR）D与F（RD）的结果是一样的，表示优先的括号是不起作用的，所以这个条件也满足。

③单位元存在。对魔方来说，在各种各样的置换中，哪个是幺元呢？不去拧魔方的任一面，其实就是不发生任何置换，就是幺元。

④逆元存在。魔方的任何过程都有一个逆过程，它产生的置换就是原过程产生的置换的逆。二者一先一后，就等于什么也没发生，所以魔方也是满足这个条件的。

综上所述，魔方满足群的基本要求（一个集合和一个运算），也满足群的4个条件。所以人们把这个群叫作“魔方群”，从而建立了魔方与群论的直接联系。

通过以上讨论，大家对群特别是魔方群已经有了一个大体上的认

识。在这个基础上，我们可以深入群的内部，讨论一下群的结构了。就魔方群而言，因为群中的元素就是置换，所以群的结构无非就是置换的结构，而置换的结构的核心问题就是它的奇偶性。

我们提到过，任何置换都可以看成连续施加若干个对换，即只有2个元素互换位置的置换。所谓置换的奇偶性指的是把它分解为对换时，对换的个数是奇数还是偶数。如果是奇数就说它是奇性的，是偶数就说它是偶性的。对于后随操作，可定义为置换的积。它的奇偶性类似于确定奇、偶整数相加的奇偶性。

下面以转动顶面为例，求解一个q转后产生了什么样的置换，以加深大家的印象。

转动顶面时，它产生如下置换：边块的循环换位是（uf,ul,ub,ur），用数字的简易写法是（7,4,2,5）；对角块的循环是（ufl,ulb, ubr,urf）。用边块的循环换位的置换P来分析，可以把P写成如下连续发生的3个对换P_1 ,P_2 ,P_3 。

$$\begin{array}{cccc} P_1 & P_2 & P_3 & \\ \downarrow & \downarrow & \downarrow & \\ 7 \rightarrow & 4 \rightarrow & 4 \rightarrow & 4 \\ 4 \rightarrow & 7 \rightarrow & 2 \rightarrow & 2 \\ 2 \rightarrow & 2 \rightarrow & 7 \rightarrow & 5 \\ 5 \rightarrow & 5 \rightarrow & 5 \rightarrow & 7 \end{array}$$

首先，边块7和4互相交换位置。然后，边块4（现在在逕垳7的位置）同边块2交换位置。接着边块2（现在在边块7的位置）同边块5互换位置。因此，P可以分解为3个对换：

$P=P_1\ P_2\ P_3=(7,4)(7,2)(7,5)$

同样地，角块的循环换位也可以分解为3个对换。显然，边块和角块的循环换位是两个不相交的过程，所以有奇奇得偶，因此这个置换是偶性的。类似地，任何外侧面的1个正转都造成偶置换。任何复杂的过程都可以分解为对6个外侧面的多次正转，因此必然都造成偶置换，也就是说，对魔方的任何操作都将产生偶置换。换句话说，要想在魔方上形成奇置换是不可能的。

联想我们在第一章中的“14～15”游戏，从初局到所要求的终局就是一个奇置换（实际上是2个元素的一次对换）。看来，这一类有排列性质的益智玩具都很可能遇上判断置换的奇偶性的问题。

必须指出，魔方中不仅有置换群，还有循环群、子群等概念，魔方中出现的操作也不只是6个面的旋转，例如夹心层操作及魔方作整体旋转。笔者只是选择了易被大家理解的置换群，以及涉及它的几个基本操作，以此将魔方与群论之间的关系做简单介绍，有兴趣的读者可参阅吴鹤龄先生的专著《魅力魔方》。

群论是从实践中发展起来的，但又是一门高度抽象的学科。鲁比克魔方问世之后，可通过魔方来学习群论，使它的理论变得十分具体、浅显易懂；反过来，在群论的指导下，魔方六面还原的方法也有规可循，不再难以捉摸。正是小玩具大学问！

上帝之数

魔方风靡的最大魔力就在于其惊人的颜色组合。一个魔方在出厂时每个面都仅有一种颜色，总共有6种颜色。但魔方被打乱后，所能形成的颜色组合却多达4.3×10^{19}。具体的计算过程是这样的。在组成魔方的小立方体中有8个是顶点，它们之间有8！种置换；这些顶点每个有3种颜色，从而在朝向上有3^7种组合（由于结构有限，魔方的顶点只有7个能有独立朝向）。类似地，魔方有12个小立方体是边，它们之间有12！/2种置换（之所以除以2，是因为魔方的顶点一旦确定，边的置换就只有一半是可能的）；这些边每个有两种颜色，在朝向上有2^{11}种组合（由于结构所限，魔方的边只有11个能有独立朝向）。因此，魔方的颜色组合总数为$8!\times3^7\times12!\times2^{11}/2=43,252,003,274,489,856,000$，即大约$4.3\times10^{19}$。这正是魔方群的阶。

但这里有一个疏漏，那就是未曾考虑到魔方作为一个立方体所具有的对称性。由此导致的结果是这些颜色组合中有很多其实是完全相同的，只是从不同的角度去看（比如让不同的面朝上或者通过镜子去看）而已。因此，4.3×10^{19}这个令人望而生畏的数字实际上是夸大了的。仅凭对称性一项，数学家们就可以把魔方的颜色组合减少至这个数的$1/96$，即约4.3×10^{17}种组合。

在对魔方的数学研究中，转动是指将魔方的任意一个面沿顺时针或逆时针方向转动90°或180°。对每个面来说，这样的转动共有3种。由于魔方有6个面，因此它的基本转动方式共有18种。那么，最少需要多少次转动，才能确保无论什么样的颜色组合都能被复原呢？这个问题引起

了很多人尤其是数学家们的兴趣。这个复原任意组合所需的最少转动次数被数学家们戏称为“上帝之数”，而魔方这个玩具宠儿也由于这个“上帝之数”一举侵入了学术界。

要研究“上帝之数”，首先当然要研究魔方的复原方法。在玩魔方的过程中，人们早就知道将任何一种给定的颜色复原都是很容易的，这一点已由玩家们的无数杰出记录所证明。不过，魔方玩家们所有的复原方法是便于人脑掌握的方法，不是转动次数最少的，因此无助于寻找“上帝之数”。

1992年，德国数学家科先巴提出了一种寻找魔方复原方法的新思路。他发现，在魔方的18种基本转动方式中有10种自成系列，通过这部分转动可以形成将近200亿种颜色组合。利用这近200亿种颜色组合，科先巴将魔方的复原问题分解成了两个步骤：第一步是将任意一种颜色组合转变为这200亿种颜色组合之一，第二步则是将这200亿种颜色组合复原。如果我们把魔方的复原比作让一条汪洋大海中的小船驶往一个固定的目的地，那么科先巴提出的那200亿种颜色组合就好比是一片特殊水域，一片比那个固定目的地大了200亿倍的特殊水域。他提出的两个步骤就好比是让小船首先驶往那片特殊水域，然后再从那里驶往那个固定目的地。在汪洋大海中寻找一片巨大的特殊水域，显然要比直接寻找那个小小的目的地容易得多，这就是科先巴新思路的巧妙之处。

但即便如此，要用科先巴的新思路对“上帝之数”进行估算仍不是一件容易的事。尤其是要想进行快速计算，最好是将复原那200亿种颜色组合的最少转动次数存储在计算机的内存中，这大约需要300兆字节（300MB）的内存。300兆字节在今天看来是一个不太大的数目，但在科先巴提出新思路的年代，普通计算机的内存连它的十分之一都远远不到。因此直到3年之后，才有人利用科先巴的新思路给出了第一个估算结果。此人名叫里德，是美国佛罗里达大学的数学家。1995年，里德通过计算发现，最多经过12次转动，就可以将魔方的任意一种颜色组合转

变为科先巴新思路中的200亿种颜色组合之一；而最多经过18次转动，就可以将那200亿种颜色组合中的任意一种复原。这表明最多经过30(=12+18）次的转动，就可以将魔方的任意一种颜色组合复原。

在得到上述结果后，里德很快对自己的估算作了改进，将结果从30减少为29，这表明“上帝之数”不会超过29。此后随着计算机技术的发展，数学家们对里德的结果又做出了进一步改进，但进展并不迅速。直到11年后的2006年，奥地利开普勒大学符号计算研究所的博士生拉杜才将结果推进到了27。第二年（即2007年），美国东北大学的计算机科学家孔克拉和库伯曼又将结果推进到了26。他们的工作采用了并行计算系统，所用的最大储存空间高达700万兆字节（7×10^6 MB），所耗的计算时间长达8000小时（相当于24小时不停歇计算将近一年的时间）。

这些计算表明，“上帝之数”不会超过26。但是，所有这些计算的最大优点（即利用科先巴新思路中的那片特殊水域）同时也是它们最致命的弱点，因为它们给出的复原方法都必须经过那片特殊水域。可事实上，很多颜色组合的最佳复原方法根本不需要经过那片特殊水域，比如紧邻目的地却恰好不在特殊水域中的任何小船，就没必要故意从那片特殊水域绕一下才前往目的地。因此，用科先巴新思路得到的复原方法未必是最佳的，由此对“上帝之数”所做的估计也极有可能是大了。

可是，如果不引进科先巴新思路中的特殊水域，计算量又实在太大，怎么办呢？数学家们决定采取折中手段，即扩大那片特殊水域的“面积”。因为特殊水域越大，最佳复原路径恰好经过它的可能性也就越大（当然，计算量也会相应地增加）。2008年，研究“上帝之数”长达15年之久的计算机高手罗基奇运用了相当于将科先巴新思路中的特殊水域扩大几千倍的巧妙方法，在短短几个月的时间内对“上帝之数”连续发动了4次猛烈攻击，将它的估计值从25一直压缩到了22。

由此我们进一步知道，“上帝之数”一定不会超过22。但是，罗基奇

虽然将科先巴新思路中的特殊水域扩展得很大，终究仍有些颜色组合的最佳复原方法是无须经过那片特殊水域的。因此，“上帝之数”很可能比22更小。那么，它究竟是多少呢？种种迹象表明，它极有可能是20。这是因为人们在过去这么多年的努力中，从未遇到过任何必须用20次以上转动才能复原的颜色组合，这表明“上帝之数”很可能不大于20；另一方面，人们已经发现了几万种颜色组合，它们至少要用20次转动才能复原，这表明“上帝之数”不可能小于20。将这两方面综合起来，数学家们普遍相信，“上帝之数”的真正数值就是20。2010年8月，这个由游戏与数学交织而成的神秘的“上帝之数”终于水落石出：研究“上帝之数”的元老科先巴、新秀罗基奇，以及另外两位合作者宣布了对“上帝之数”是20的证明。

第六章 数独和拉丁方

数独是《最强大脑》里的老面孔了，不过玩法一直在变，从盲填数独到立体数独，再到中外选手数独大对决，规则越来越复杂，难度系数越来越高。第五季《最强大脑》中又迎来了一个超级难题——翻滚数独。

现场有24种由12宫组成的异形数独，它们的外围全是4×4的16宫，空余的4宫是随机的，例如图6-1中不打斜线的12宫即组成一个异形数独。这12宫又均匀分成红、黄、蓝3种色块，且每宫少2个数字。嘉宾随机选取一个异形数独，在保持每宫的相对位置不变的情况下，将该数独题的每一宫随机放入现场另一个装置——三棱柱阵中。三棱柱阵由9根旋转柱体组成，每根柱体3个面，每面有9宫。

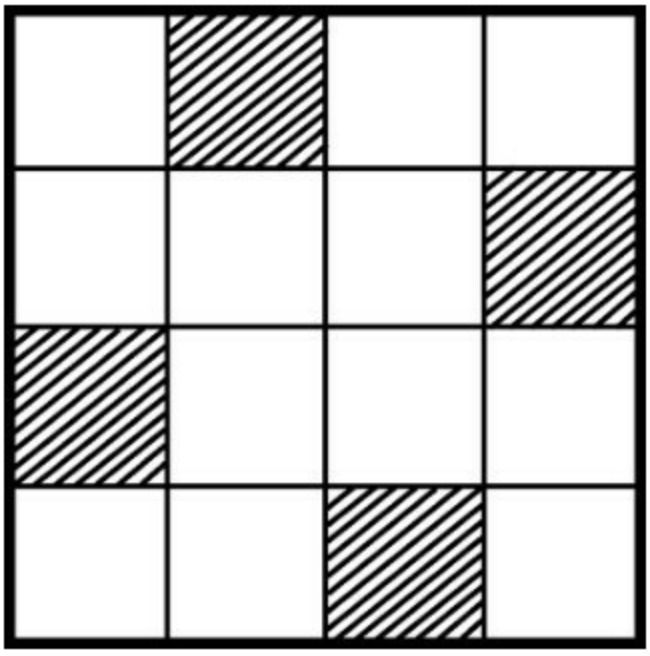


图6-1

竞赛开始时，三棱柱阵以一定转速（每4秒转一次）匀速翻转，两位选手应在这翻转的宫阵中找到嘉宾所选择的特定的异形数独题，并在答题板上写出每宫的坐标位置（三棱柱阵中每行每列都标有A~I的字母），并填上每宫中缺少的数字。

“翻滚数独”对选手综合实力，即空间想象力、推理计算力、观察力、快速记忆力等的考验都是前所未有的。

数独简介

数独是一种数学逻辑游戏。其盘面是一个九宫，每一宫又分为9个小格，在这81格中给出已知数字（称为提示数）作为解题条件。利用逻辑推理，在其他的空格上填入1~9的数字，并使每个数字在每一行、每一列和每一宫中都只出现一次，上述九宫形的数独亦称为标准数独。

数独起源于18世纪初瑞士数学家欧拉所研究的拉丁方阵。19世纪80年代，美国的一位退休建筑师格昂斯根据这种拉丁方阵发明了一种趣味填数游戏，这就是数独的雏形。20世纪70年代，人们在美国纽约的一本益智杂志上发现了这个游戏，当时被称为填数字，这是目前公认的数独最早的见报版本。1984年，一位日本学者将其带到了日本，发表在Nikoli公司的一本游戏杂志上，改名为“数独”。“数”是数字的意思，“独”是唯一的意思。后来，一位曾在中国香港工作的新西兰人高乐德在1997年3月到东京旅游时，无意中发现了它。他首先将其发表在英国的《泰晤士报》上，不久其他报纸也纷纷刊登，很快便风靡英国。在20世纪90年代，我国的部分益智类书籍开始刊登，从而得到推广。

影响数独难度的因素很多。就题目本身而言，包括提高难度的技巧、各种技巧所用的次数、是否用隐藏深度及广度的技巧组合，当前盘面可逻辑推导出的填数个数等；对于玩家而言，了解的技巧数量、熟练程度、观察力等自然也影响对一道题的难度判断。

如果一道题目的提示数少，那么题目就会相对难，提示数多则会简单，这是一般人判断问题难易的思维模式，但数独谜题提示数的多寡与难易并无绝对关系。多提示数比少提示数难的情况屡见不鲜，即使是相同的提示数（甚至相同谜题图形）也可以变化出难度各异的问题。提示

数少对于出题的困难度有比较直接的影响，以20~35个提示数而言，每少一个提示数，其出题难度会增加数倍。在制作谜题时，提示数在22以下非常困难，所以常见的数独题其提示数在23~30，其原因在于制作相对容易，且可以设计出比较漂亮的图形。

总之，数独游戏的难易程度很难给出一个客观的判断标准。下面就是一个例子。

“世上最难数独”

芬兰数学家因卡拉花费3个月的时间设计出了世界上迄今难度最大的数独游戏，而且它只有一个答案。因卡拉说只有思考能力最快、头脑最聪明的人才能破解这个难题。这是英国《每日邮报》2012年6月30日的一篇报道，其谜面如图6-2所示。

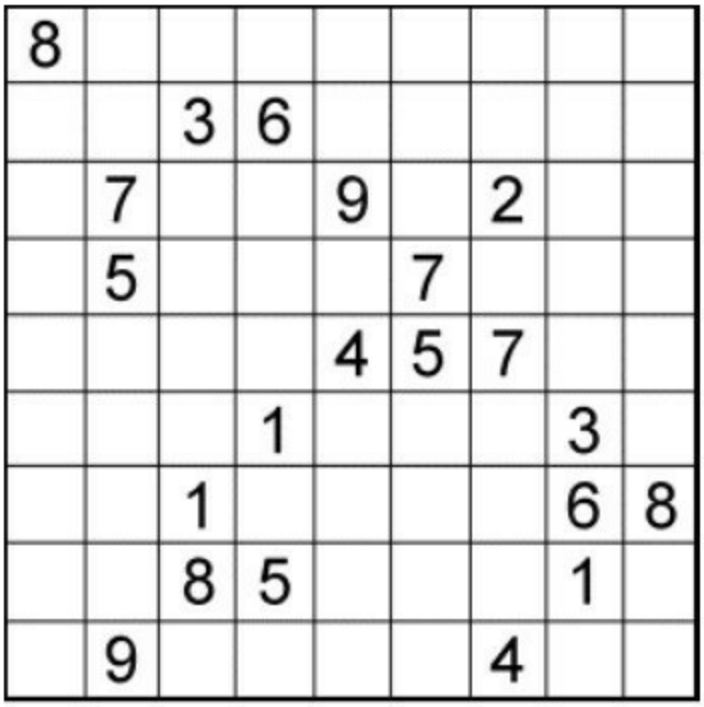


图6-2

2013年，国内不少媒体集中报道了一则消息：江苏扬州市一位69岁的老农用了3天时间破解了这一难题。后来又发现该农民更改了原谜面上的一个数字，当然这就不能算是已解决，但剧情又发生了翻转，媒体又陆续报道了几个答案。

图6-3是重庆八旬副教授的答案。据报道，这个答案也曾被温州的五旬老伯所获，另一位广东的学生用两小时也得到了同一答案。网

友“谷子大鬼”上传了他的解答，据称用时2小时20分钟，他的答案如图6-4所示。

8	1	2	7	5	3	6	4	9
9	4	3	6	8	2	1	7	5
6	7	5	4	9	1	2	8	3
1	5	4	2	3	7	8	9	6
3	6	9	8	4	5	7	2	1
2	8	7	1	6	9	5	3	4
5	2	1	9	7	4	3	6	8
4	3	8	5	2	6	9	1	7
7	9	6	3	1	8	4	5	2

图6-3

8	6	9	2	1	4	3	7	5
5	4	3	6	2	1	8	9	7
1	7	5	3	9	8	2	4	6
3	5	4	9	6	7	1	8	2
9	1	6	8	4	5	7	2	3
4	8	7	1	5	2	6	3	9
2	3	1	4	7	9	5	6	8
7	2	8	5	3	6	9	1	4
6	9	2	7	8	3	4	5	1

图6-4

显然，这是另一个正解，打破了因卡拉“它只有一个答案”的断言。笔者在前面提到过，数独难度并无科学的评判标准。制题者有他的思维，答题人又有各自的方法，难分伯仲。但有一点可以肯定，出题者自我评定是难度最大的，这未免有点“老王自夸”了！

任何时候、任何事情都不要轻视中国人的聪明才智！这倒是一句至理金句。

数独的数学问题

下面所谈数独中的数学问题，都是针对标准数独的。

（1）终盘的可能性

通常将一个完成了的数独题目称为终盘，在数独游戏盛行后，人们自然想知道这个游戏究竟存在多少个终盘形式。2005年，德国数学家费尔根豪尔给出了答案：数独的最大可能终盘数是6,676,903,752,021,072,936,960种，约为6.67×10^{21} 种组合。如果等价终盘（如旋转、翻转、行行对换、列列对换、数字对换等形式）不计入，则有5,472,730,538个组合，相当于全球人口总数。这一更准确的终盘数是英国数学家弗雷泽·贾维斯和拉塞尔给出的。

（2）最小初盘问题

与终盘相对应，一个数独游戏给出的初始条件称为初盘。一般常见的初盘数字个数在22~28，而数独爱好者们常问的一个问题是：最少给出多少个数字，数独游戏才能确保有唯一解。

此前发现这个数字很可能是17，这个最小唯一解初盘是由一名日本数独爱好者发现的。图6-5即是其初盘，但这只是一个例子，还未得到证明。澳大利亚数学家戈登·罗艾尔已经收集了36,628个有17个数字的唯一解初盘。

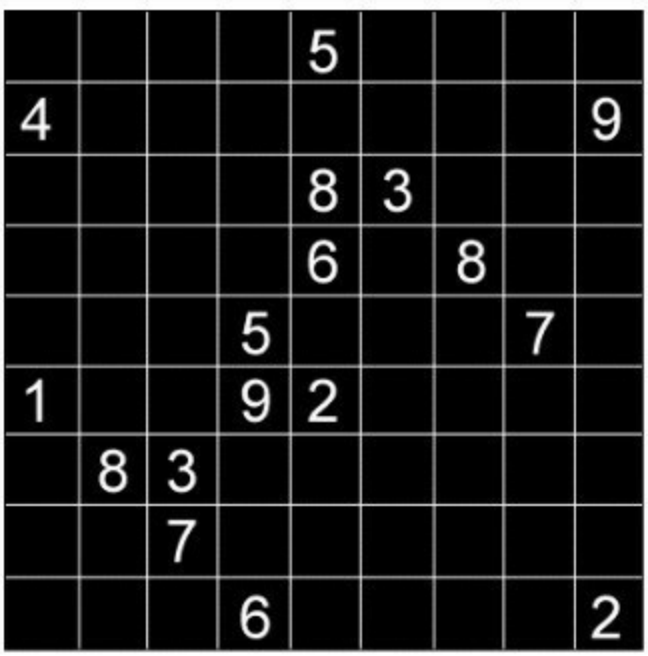


图6-5

刚刚迈入2012年，数学界就有一个不大不小的轰动。3位爱尔兰数学家发表了一篇论文，证明数独至少需要17个初始数字才有唯一解。

开始，几位数学家的想法非常简单：只要把每一种有16个初始数字的数独都尝试一遍，自然就知道答案了。但很可惜，数独的组合实在是太多了，即使是用现在运行最快的计算机，也不可能在我们有生之年穷尽所有组合。所以，必须用一些数学的方法来减少尝试的次数。他们发现，数独中的一些组合是等价的。如图6-6所示，交换第一列和第二列对整个数独并没有影响，而这只是其中一个例子。3位数学家总结了数独的4种等价交换。

5	3	4	6	7	8	9	1	2	3	5	4	6	7	8	9	1	2
6	7	2	1	9	5	3	4	8	7	6	2	1	9	5	3	4	8
1	9	8	3	4	2	5	6	7	9	1	8	3	4	2	5	6	7
8	5	9	7	6	1	4	2	3	5	8	9	7	6	1	4	2	3
4	2	6	8	5	3	7	9	1	2	4	6	8	5	3	7	9	1
7	1	3	9	2	4	8	5	6	1	7	3	9	2	4	8	5	6
9	6	1	5	3	7	2	8	4	6	9	1	5	3	7	2	8	4
2	8	7	4	1	9	6	3	5	8	2	7	4	1	9	6	3	5
3	4	5	2	8	6	1	7	9	4	3	5	2	8	6	1	7	9

图6-6

①列与列的重新排列；

②行与行的重新排列；

③数字1到9的重新排列，如把原先是1的位置都填上2，再把原先填2的位置都填上3……直到把原先是9的位置都填上1；

④网格的变换，如将整个数独顺时针旋转90°，做镜像对称等。

虽然等价类的数量已经降到可以接受的范围，但问题还远未解决。因为在选择了某一等价类的一种情况以后，还需要验证该情况的81个数字中是否可以选出16个，使得以这16个数为初始数字的数独有唯一解。

后来，几位数学家利用了“不可避免集”的概念。“不可避免集”是图论中的一个概念，如图6-7所示，圈出的5和9可以互相交换位置而产生两个不同的解。为了让这个数独的答案唯一，这4个数字里必须有一个数字是起始数字，这样我们才能限定5和9的最终位置。这4个数字也被称为一个不可避免集。

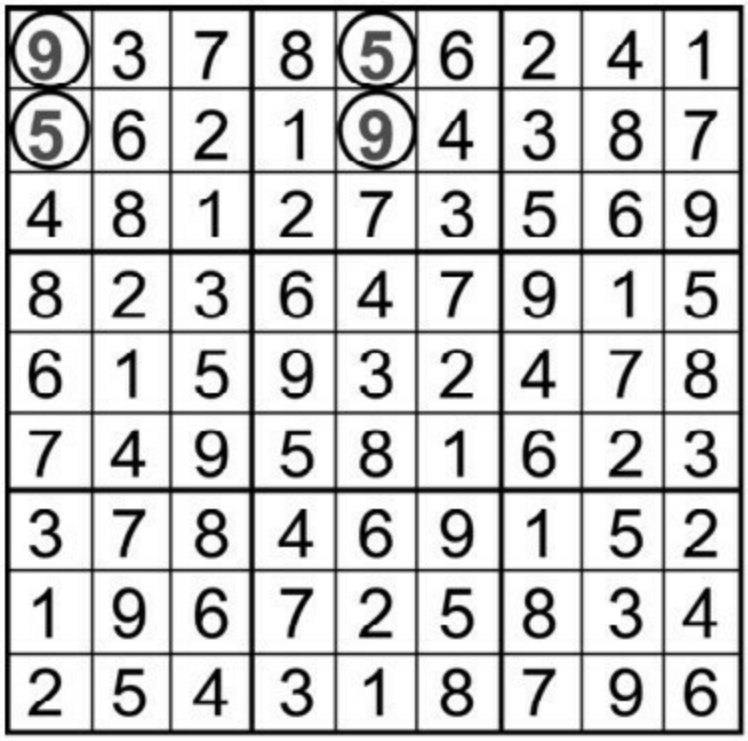


图6-7

数学家们总结了一套不可避免集的模板，总共记录了525种不同的不可避免集。因为一开始所说的4种等价变换对不可避免集也适用，所以他们对不可避免集进行了一些处理，以保证这525种不可避免集互相不能通过4种等价变换得到。

此时，枚举算法就被改造成这样：数学家给所有的不可避免集都设定了一个状态，分为“被击中”或“未被击中”两种。初始时九宫格81个方格内都没有填入数字，所有不可避免集的初始状态均为“未被击中”。之后每次选择一个最小的未被击中的不可避免集，枚举其中的每个格子。即每次选择不可避免集中的一个格子填充初始数字，直至试完不可避免集中的所有空格。同时将这个不可避免集标记为“被击中”状态，每次枚举都有4种可能。

①如果这个格子也出现在其他不可避免集中，那么将这些被涉及的不可避免集也标记为“被击中状态”；

②如果枚举了16个格子后还有不可避免集未被击中，说明以这16个格子为初始状态的数独一定没有唯一解；

③枚举了16个格子后恰好所有的不可避免集全部被击中；

④如果枚举了不到16个格子时所有的不可避免集已经全部被击中，则从剩下的所有格子中再枚举几个使填充了数字的格子达到16个。

在这一系列优化之后，算法的复杂程度大为降低，最后一步就是用解数独的程序来验证所有枚举的情况是否有唯一解了。幸而这个算法得到了一个确定的结果：所有仅含16个初始数字的数独都不存在唯一解！都柏林大学学院的麦奎尔于2012年1月1日在互联网上贴出了自己的证明。在1月7日美国波士顿召开的一次会议上，数学家们就此达成了共识，还表示这种“大集合算法”还可能在其他领域产生作用。

（3）最大初盘问题

与最小初盘相反，在一个数独初盘中，最多可给出多少个数字，使得再增加一个数字，该问题便没有唯一解了。

可以采用倒推法，在初始数字为80个的情况下无须说明，在初始数字为79个的初盘中也大约如此，因为考虑到每一个小九宫格必须满足每个数字出现且仅出现一次，这意味着所缺少的数字都必须出现在同一个九宫格内。所以，还可以依次推出78个数字的初盘也有唯一解，但当初盘中给定数字为77个的时候，该数独游戏便至少有两个解。

欧拉方阵

1735年，数学大师欧拉积劳成疾，右眼失明。他受普鲁士国王腓特烈大帝之邀，到了气候相对温和的德国，任柏林科学院物理数学所所长。

一次腓特烈大帝在阅兵仪式前问其指挥官：在一个由36名军官组成的方队里，若这些军官分别来自6支不同的部队，每支部队均有6种不同军衔的军官，他们能否排成一个6×6的方队，且满足每行、每列既有每支部队的军官，又有不同军衔的军官？指挥官试了许久之后感到无能为力。问题传到欧拉那里，他开始了对这个问题的研究。

如果我们用数字1,2,3,4,5,6表示部队的编号以及6种不同的军衔，而用数对（i，j）来代表从第i个部队挑选出的有军衔j的军官，当然i与j只能取1,2,3,4,5,6这6个不同的数。于是“36军官问题”可提炼成这样一个数学问题。

将36个数偶（1,1），（1,2），…，（1,6），（2,1），（2,2），…，（2,6），（3,1），（3,2），…，（6,6）排成6×6方阵，使得每一行的第一个分量为6个不同的数，第二个分量也两两不同；对于方阵中的每一列也是如此。

如果把“36军官问题”一般化就是：将n^2个整数偶（1,1），（1,2），…，（1,n），（2,1），（2,2），…，（2,n），…，（n,1），…，（n,n）排成$n \times n$方阵，使得方阵的每一行和每一列的两个分量两两不同。

欧拉很容易地就证明出n=2是不可能的，而n=3,4,5是可能的，即

n=3:

（1,1）（2,2）（3,3）

（2,3）（3,1）（1,2）

（3,2）（1,3）（2,1）

n =4:（1,1）（2,2）（3,3）（4,4）

（4,2）（3,1）（2,4）（1,3）

（2,3）（1,4）（4,1）（3,2）

（3,4）（4,3）（1,2）（2,1）

n =5:

（1,1）（2,2）（3,3）（4,4）（5,5）

（5,4）（1,5）（2,1）（3,2）（4,3）

（4,2）（5,3）（1,4）（2,5）（3,1）

（3,5）（4,1）（5,2）（1,3）（2,4）

（2,3）（3,4）（4,5）（5,1）（1,2）

但对于*n* =6的情况，既找不到合乎要求的方阵，又证明不了它不存在，但估计是不存在的。欧拉推测像*n* =2和*n* =6这样的形如4*m* +2（*m* 为非负数）的数，上述方阵是不存在的。直到他去世，这个猜想也没有解决，为后世的数学家留下了一个角逐的难题，后人称之为“欧拉猜想”，这种方阵也随之被称为“欧拉方阵”。

欧拉猜想引起了许多数学家的重视。为了研究方便，人们首先将欧拉方阵拆成两个简单方阵，用第一个分量构成一个方阵，第二个分量构成另外一个。例如，将前面的*n* =3的欧拉方阵分解成如图6-8所示的两个方阵。

1	2	3
2	3	1
3	1	2

1	2	3
3	1	2
2	3	1

图6-8

在这样的方阵中每一行和每一列都由两两不同的数组成，称为拉丁方，而把能合并成欧拉方阵的两个拉丁方称为一组正交拉丁方。一个欧拉方阵可分解成两个正交拉丁方，但是两个拉丁方不一定能组成欧拉方阵。例如，图6-9是两个三阶拉丁方，它们合并起来并不能组成欧拉方阵。因为合并后只能得到3个不同的数偶，即（1,1）、（2,3）、（3,2），而不是9个不同的数偶。

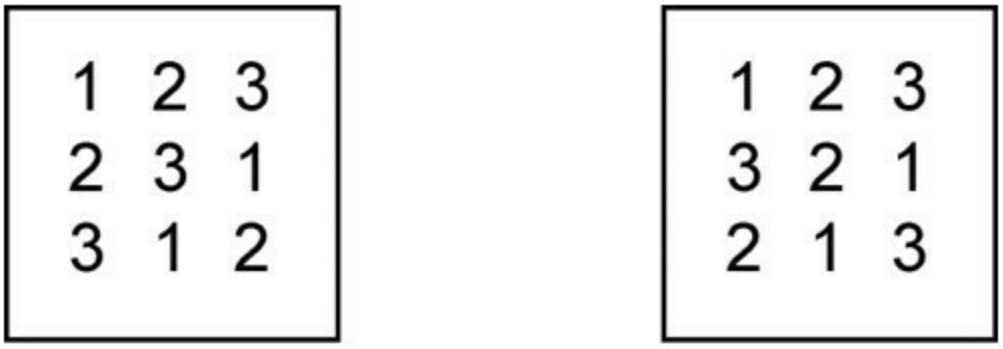


图6-9

正是拉丁方中每一行和每一列都由两两不同的数组成，这一特质催生了人们对数独的联想。把9个三阶拉丁方放在一起，参照拉丁方的特质，就成了当今风靡世界的逻辑推理游戏——标准数独。

由于构造六阶正交拉丁方困难重重，欧拉猜想经历了100多年也没有突破性的进展。直到1901年，法国数学家塔里证明n =6时欧拉猜想成立，即“36军官问题”没有解。此后，人们对欧拉猜想笃信不移。

半个多世纪后，“不幸”的事发生了，印度数学家玻色和美国数学家史里克汉德构造出了22阶正交拉丁方，而另一美国数学家帕克又造出了14阶和26阶的欧拉方阵。不久，他们又证明了当n 不等于2或6时，n 阶正交拉丁方皆存在。这个结果是人们在170多年的努力中未曾想到的。图6-10即为十阶正交拉丁方。

Aa	Eh	Bi	Hg	Cj	Id	If	De	Cb	Fe
Ig	Bb	Fh	Ci	Ha	Dj	Je	Ff	Ac	Gd
Jf	Ia	Ce	Gh	Di	Hb	Ej	Fg	Bd	Ae
Fj	Jg	Ib	Dd	Ah	Ei	He	Ga	Ce	Bf
Hd	Gj	Ja	Ic	Ee	Bh	Fi	Ab	Df	Cg
Gi	He	Aj	Jb	Id	Ff	Ch	Be	Eg	Da
Dh	Ai	Hf	Bj	Jc	Te	Gg	Cd	Fa	Eb
Be	Cf	Dg	Ea	Fb	Hc	Ad	Hh	Ii	Jg
Cb	Dc	Ed	Fe	Cf	Ag	Ba	Ij	Jh	Hi
Ec	Fd	Ge	Af	Bg	Ca	Db	Ji	Hj	Ih

图6-10

空间中的正交拉丁方

当人们把正交拉丁方的概念拓展到三维空间时为：$n \times n \times n$ 的立方体上分别写有0,1,2,…,n -1，使得n 个数字在每行、每列中恰好出现一次，这便是n 阶立体拉丁方。合并3个n 阶立体拉丁方时，若每个有序三重数对000,001,001,…,$\overline{n-1}\,\overline{n-1}\,\overline{n-1}$ 均出现一次，则称它为n 阶立体正交拉丁方。

已知平面上的“36军官问题”是无解的，但立体六阶正交拉丁方是确实存在的。1982年，阿尔金、史密斯和斯特拉斯给出了实例。这个六阶立体正交拉丁方的各层数字如下：

Ⅰ	313	435	241	522	000	154
	402	541	350	014	133	225
	534	050	423	105	242	311
	045	123	512	231	354	400
	151	212	004	340	425	533
	220	304	135	453	511	042
Ⅱ	201	353	415	134	542	020
	330	422	501	245	054	113
	443	514	030	351	125	202
	552	005	143	420	211	324
	024	131	252	513	200	445
	115	240	324	002	433	551
Ⅲ	455	221	333	040	114	502
	521	310	442	153	205	034

	010	403	554	222	331	145
	103	532	025	314	440	251
	232	044	111	405	553	320
	344	155	200	531	002	413
Ⅳ	120	504	052	315	431	243
	213	035	124	401	540	352
	302	141	215	530	053	424
	434	250	301	043	122	515
	545	323	430	152	214	001
	051	412	543	224	305	130
Ⅴ	032	140	524	203	355	411
	144	253	015	332	421	500
	255	322	101	444	510	033
	321	414	230	555	003	142
	410	505	343	021	132	254
	503	031	452	110	244	325
Ⅵ	544	012	100	451	223	335
	055	104	233	520	312	441
	121	235	342	013	404	550
	210	341	454	102	535	023
	303	450	525	234	041	112
	432	523	011	345	150	204

六阶幻方之王

中国人有句俗话“六六大顺”，这句话放在组合数学中却不灵。一是六阶正交拉丁方不存在，致使“36军官问题”无解；二是数学家已严格证明了用1～36作素材，不可能构造出完美幻方（完美幻方的含义见下文）。难道6是组合数学的魔咒吗？

阿尔金等3位数学家在空间中构造出了正交拉丁方，这是从二维中突围成功的实例。再后来，我国一位退休职工丁宗智先生也打破这一规则，突围成功。

丁先生打破思维束缚，先将其扩大到1~49，形成如图6-11所示的七阶方阵，然后再去掉中间一行和中间一列的数，剩下的数不多不少，恰恰是36个。经过这样一次手术，“癌细胞”就被彻底割除，起死回生，用这36个数不但可以构造出完美幻方，而且更加神奇，从而获得了“六阶幻方之王”的美名。

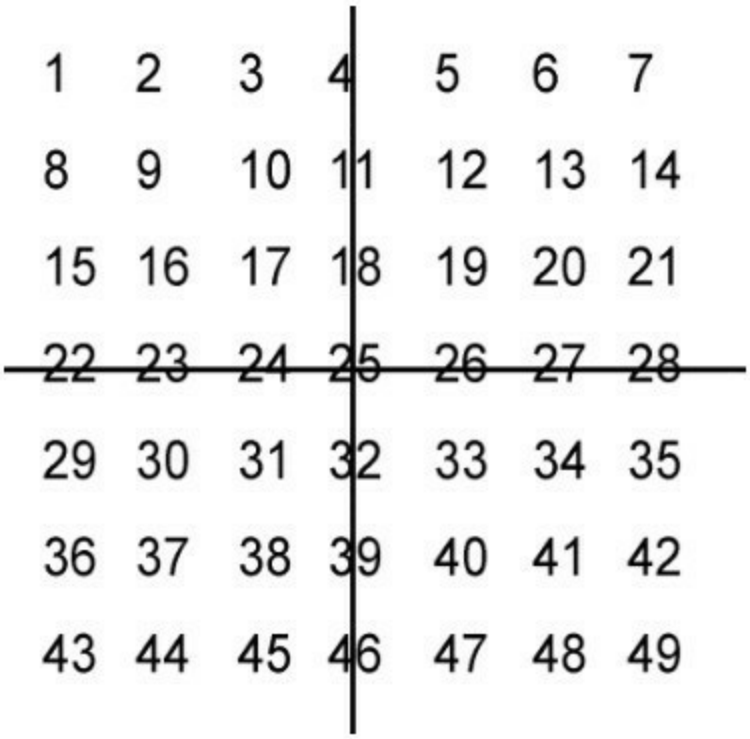


图6-11

用这36个数组成的六阶幻方如图6-12所示。丁先生的这一作品拔得头筹，竟与西方一些研究家所造出的六阶幻方一模一样。

1	42	29	7	36	35
48	9	20	44	13	16
5	38	33	3	40	31
43	14	15	49	8	21
6	37	34	2	41	30
47	10	19	45	12	17

图6-12

该幻方中每行、每列、每条对角线［包括主对角线、副对角线，以及“折断”了的对角线（所谓折断了的对角线，又名泛对角线，例如42+20+3+8+30+47，或如7+13+31+43+37+19等）］上的6个数之和都等于幻方常数150，这就是完美六阶幻方。该幻方还有8条重要性质，下面结合图形简要解释。在该六阶幻方中，按图示方法取数，所有数之和均相等，且等于25乘以☆的个数。

如图6-13（a）的意思是：在该六阶幻方中，任何一个三阶方阵的9个数之和是225（=25×9）。图6-13（b）的意思是：在该六阶幻方中，任何一个四阶方阵的16个数字之和是400（=25×16）。其余的图也是同样的意思。各种情况的验证请读者自行计算，在此就不浪费篇幅了。

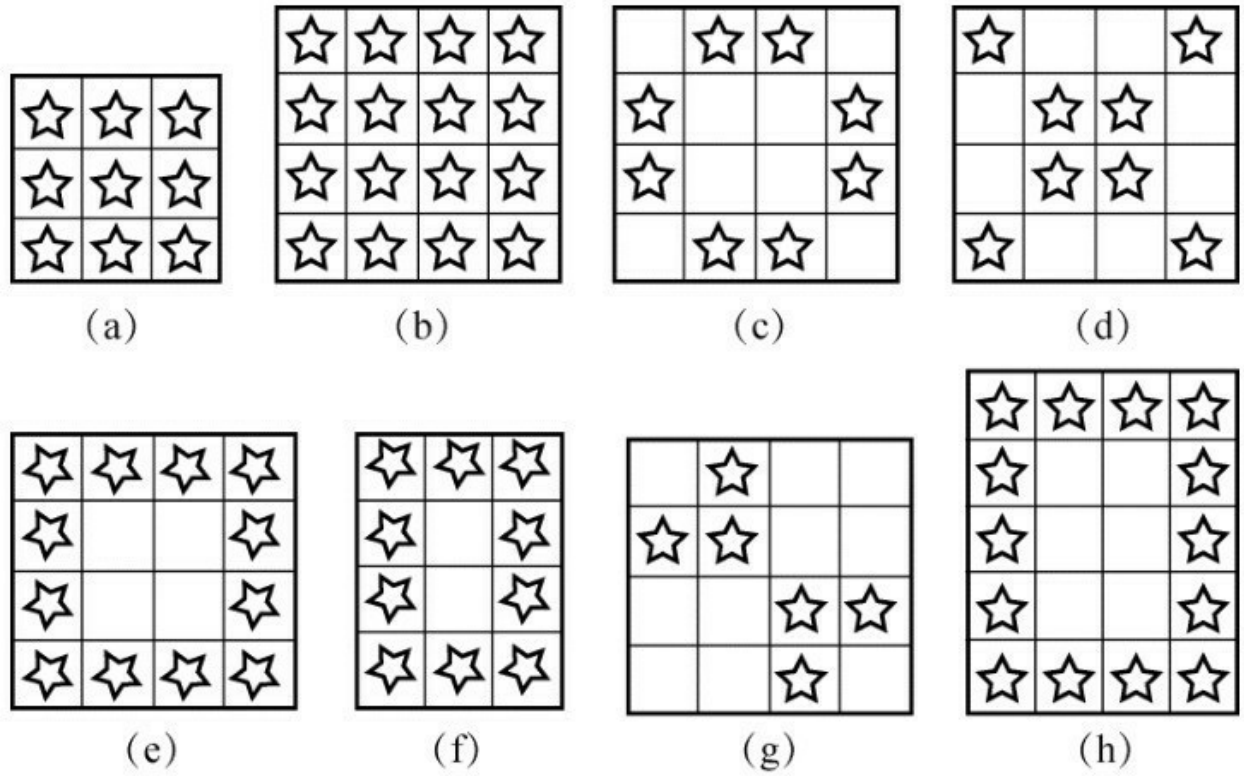


图6-13

第七章 幻立方与反幻方

幻立方是《最强大脑》第二季晋级赛第三场的挑战项目，道具一经亮相，剔透的“水晶立方”便震撼全场，评审人也禁不住惊叹于数学之美。项目要求在由343个小立方体组成的大立方体中先确认1的初始位置，随后在各方格中一一填入数字，使得大立方体中每行、每列、平面及空间对角线上的数之和全部等于1204。海量运算面前，选手在较短时间内完成了填数。

节目组特别邀请了中国幻方协会副主席进行验证，证明挑战成功，但对选手有质疑，称“挑战很简单，协会中有二十余人全能完成”。一语惊呆众人，现场剑拔弩张。最后评审团给出的难度系数仅为3，因总分低于80分而惨遭淘汰。

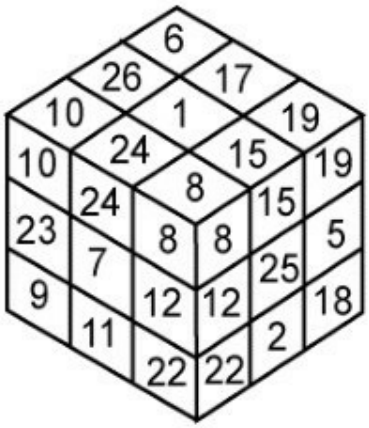
笔者无意加入口水战，只是介绍幻立方的构造方法，难易任人评说。

什么是幻立方

幻立方是指$n \times n \times n$的三维幻方，其n^2个行、n^2个列、n^2个纵列以及4条空间对角线上的n个数之和都相等，这叫“基本幻立方”或“半完美幻立方”；如果3个方向上的每个横截面本身都是幻方，即不但行、列上的n个数之和等于幻和，其2条主对角线上的数之和也都等于幻和，那么就叫“完美幻立方”。显然，不管是半完美的还是完美的，幻立方的幻和应等于$\frac{1}{2}n(n^3+1)$。半完美幻立方中有$3n$个平面幻方，满足幻和的n数之和有$3n^2+4$，完美幻立方则有$3(n+2)$个平面幻方（除了水平、垂直、纵切3个方向上各有n个外，通过立方体每一组对边的截面也是一个幻方），满足幻和的n数之和有$3n^2+6n+4$或$3n(n+2)+4$。

可以证明，不管是半完美幻立方还是完美幻立方，单偶阶即$n=2(2m+1)$的情况都是不存在的。奇数阶及双偶阶$(n=2\cdot 2m)$人们已经找出了构造的方法，并且从数学上给予了一般的证明。由于比较繁复，我们这里就不介绍了。

图7-1是一个最简单的三阶半完美幻立方，幻和为42，共31组。图中把幻立方分解为3个剖面进行展示。已经证明，三阶和四阶完美幻立方是不存在的。



10	24	8
23	7	12
9	11	22

26	1	15
3	14	25
13	27	2

6	17	19
16	21	5
20	4	18

图7-1

三阶幻立方不可能是完美的可用反证法证明如下。设有三阶完美幻立方，取平行于该立方体表面的任一方阵，设它第一行中的3个数分别是A 、B 、C ，第三行中的3个数是D 、E 、F ，第二行中间的一个数是X 。因为三阶幻立方的幻和为42，它又是完美的，所以必有以下等式成立。

$A+B+C=42$

$D+E+F=42$

$(A+X+F)+(C+X+D)+(B+X+E)=3\times42$

即有$3X+(A+B+C)+(D+E+F)=3\times42$。由此可得，$3X=42$，$X=14$。

注意，我们以上的讨论是针对幻立方中任意一个截面进行的，因此结论适用于幻立方中的任意截面。也就是说，幻立方中任意截面正中间的元素都应取14。但是按规定，幻立方中任意一个数都不能重复出现，这就说明三阶完美幻立方是不存在的。

连续摆线法

要制作奇数阶幻立方先要从普通幻方的连续摆线法开始。连续摆线法是一种适用于奇数阶幻方构造的比较古老的方法。据查，它于1687年由法国驻泰国大使洛贝利从泰国带回法国，从而传播开来，故这个方法在西方又称“暹罗法”。

连续摆线法的法则如下。

把1放在中间一列最上边的方格中，从它开始，按对角线方向（比如说按从左下到右上的方向）顺次把数按从小到大的顺序放入各方格中。如果碰到顶，则折向底；如果到达右侧，则转向左侧；如果进行中轮到的方格已有数或到达右上角，则退至前一格的下方。

按照这个法则建立的五阶幻方的示例如图7-2所示。

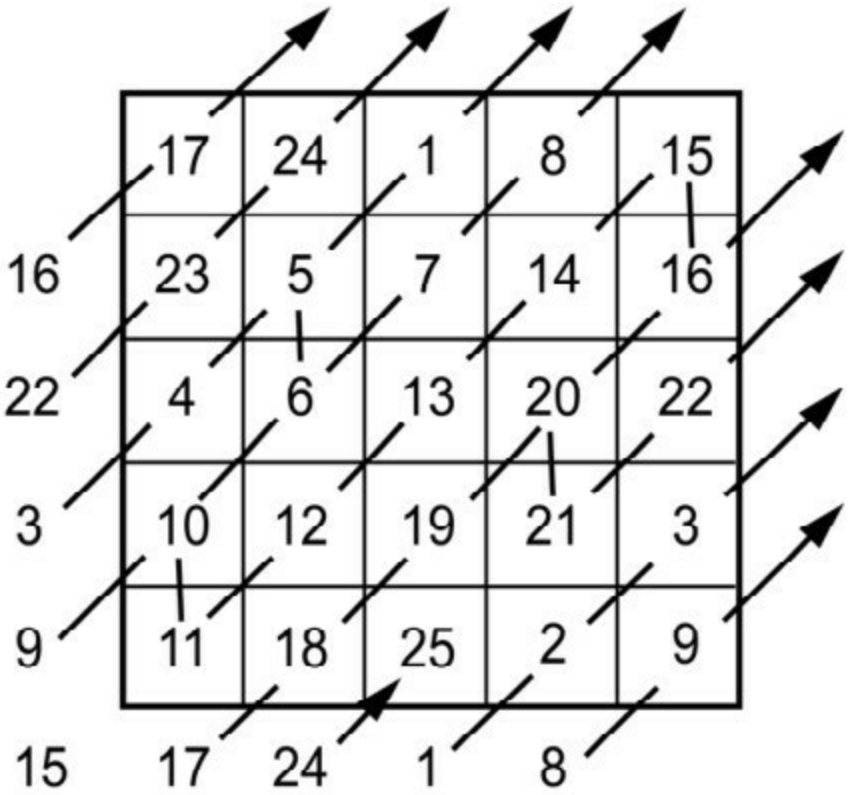


图7-2

这个方法可以推广到一般情况，即起始数不一定非摆在第一行中间，下一个数也不一定非摆在上一个数的右上方格中，比如可摆在上一个数的左下方格，即使图7-2中箭头的方向相反，或者一次跨2行或2列等。为此，我们定义一个“普通向量”（x ，y ），表示在正常走步情况下的偏置量。另外定义一个“中断向量”（u ，v ），表示发生冲突时的偏置量，即在异常走步下的摆数。

在上述的摆数法中，普通向量是（1，-1），表示右上方方格（可以想象图像的左上角为坐标原点，x 轴正方向仍向右，但y 轴正方向改为向下），中断向量是（0，1），即直接放在下方方格中。

17世纪的法国数学家菲利普·德·拉伊尔详细研究了连续摆线法的推广后发现：下面这一组“和差值”的绝对值$|u+v|, |(u-x)+(v-y)|, |u-v|, |u+y-x-v|$如果相对于阶数$n$ 都是互质的，则可以构成完美幻方。表7-1给出了某些特定的普通向量、中断向量和它们“和差值”的绝对值的集合所适用的幻方。

表7-1 连续摆线法的推广

普通向量	中断向量	“和差值”集合	适用幻方	是否完美幻方
（1，–1）	（0，1）	（1，3）	$2k+1$	否
（1，–1）	（0，2）	（0，2）	$6k\pm1$	否
（2，1）	（1，–2）	（1，2，3，4）	$6k\pm1$	否

续表

普通向量	中断向量	“和差值”集合	适用幻方	是否完美幻方
（2，1）	（1，–1）	（0，1，2，3）	$6k\pm1$	是
（2，1）	（1，0）	（0，1，2）	$2k+1$	否
（2，1）	（1，2）	（0，1，2，3）	$6k+1$	否

为了使大家更清楚地理解这两个向量的应用，图7-3给出了用推广

的连续摆线法构造的2个奇数阶幻方。其中图7-3（a）是一个五阶幻方，其普通向量为（2，1），中断向量为（1，-1），图7-3（b）是一个七阶幻方，普通向量和中断向量分别是（1，-1）及（0，1）。

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

(a)

32	41	43	3	12	21	23
40	49	2	11	20	22	31
48	1	10	19	28	30	39
7	9	18	27	29	38	47
8	17	26	35	37	46	6
16	25	34	36	45	5	14
24	33	42	44	4	13	15

(b)

图7-3

七阶完美幻立方的构造

学习了摆线法后，我们就可以制作幻立方了。奇数阶的幻立方可以用类似于普通幻方的连续摆线法构造，其普通向量是（1，-2），中断向量则有2个：小中断向量用于确定在一个面上摆7个数以后如何转到下一面摆数，如设向量值为（2，0）；大中断向量用于确定在7个面上摆好49 （=7×7）个数以后如何转到下一轮的49个数，如设向量值为（0，1）。在摆数过程中，假定行、列、面都是循环相接的。起始1置于中间一面（IV面）中间一列的最顶上一格以后，按普通向量（1，-2）在该面摆好7个数，然后按小中断向量（2，0）将8置于下一面（V面）的第5列第3格（因为7在上一面的第3列第3格），继续下一组7个数的摆放。按上述办法摆好49个数以后，按大中断向量（0，1）将50置于III面49的下方，开始下一轮的大循环，直至把343个数全部摆好，一个七阶完美幻立方就形成了。在这个完美幻立方中，共有193组的7数之和为1024。

按此方法及3个向量制成的七阶幻立方见图7-4。现在，读者们可以按上述方法来自制一个奇数阶的幻立方了。

I

322	87	153	261	33	141	207
29	144	210	318	90	149	264
86	152	260	32	147	206	321
143	209	317	89	148	263	35
151	266	31	146	205	320	85
208	316	88	154	262	34	142
265	30	145	204	319	91	150

II

100	215	323	95	161	269	41
157	272	37	103	211	326	98
214	329	94	160	268	40	99
271	36	102	217	325	97	156
328	93	159	267	39	105	213
42	101	216	324	96	155	270
92	158	273	38	104	212	327

III

277	49	108	223	331	54	162
334	50	165	280	45	111	219
48	107	222	330	53	168	276
56	164	279	44	110	218	333
106	221	336	52	167	275	47
163	278	43	109	224	332	55
220	335	51	166	274	46	112

IV

62	170	285	1	116	231	339
119	227	342	58	173	281	4
169	284	7	115	230	338	61
226	341	57	172	287	3	118
283	6	114	229	337	60	175
340	63	171	286	2	117	225
5	113	228	343	59	174	282

V

232	298	70	178	293	9	124
289	12	120	235	301	66	181
297	69	177	292	8	123	238
11	126	234	300	65	180	288
68	176	291	14	122	237	296
125	233	299	64	179	294	10
182	290	13	121	236	295	67

VI

17	132	240	306	71	186	252
74	189	248	20	128	243	302
131	239	305	77	185	251	16
188	247	19	127	242	308	73
245	304	76	184	250	15	130
246	18	133	241	307	72	187
303	75	183	249	21	129	244

图7-4

VII

194	253	25	140	199	314	79
202	310	82	190	256	28	136
259	24	139	198	313	78	193
309	81	196	255	27	135	201
23	138	197	312	84	192	258
80	195	254	26	134	200	315
137	203	311	83	191	257	22

图7-4（续）

偶数阶幻立方示例

八阶幻立方早在1875年就被发现了，但作者都没有留下姓名。它也不像奇数阶幻立方有一般化的构造法。此处只简单介绍如图7-5所示的八阶幻立方，这个幻立方是1970年由宾夕法尼亚州年仅16岁的高中生迈尔斯独立发现的。这个幻立方有许多奇妙的性质，在幻方界很有名气。

①它是对称的，即幻立方中任意一对在中心对称位置上的两数之和均为513。

②幻立方中每条正交线和对角线上的8数之和均为2052。

③幻立方8个顶角上的8数之和也是2052，且幻立方内任意对中心对称的矩形体的8角上8数之和也是2052。

④整个幻立方可分割成64个二阶的小立方体（2×2×2），每个小立方体内的8数之和也是2052。

⑤这个幻立方中的数是从1到512，如果从513开始取数，则顺序取其后的512个数，按照相同的规律可进一步组成63个幻立方。连同原始的幻立方，可以构成一个64阶完美幻立方。在此基础上，以相同方式又可构成512阶幻立方。依次类推，可构成任意8^n （n =1,2,3,…）阶幻立方。

I

19	497	255	285	432	78	324	162
303	205	451	33	148	370	128	414
336	174	420	66	243	273	31	509
116	402	160	382	463	45	291	193
486	8	266	236	89	443	181	343
218	316	54	472	357	135	393	107
185	347	85	439	262	232	490	12
389	103	361	139	58	476	214	312

II

134	360	106	396	313	219	469	55
442	92	342	184	5	487	233	267
473	59	309	215	102	392	138	364
229	263	9	491	346	188	438	88
371	145	415	125	208	302	36	450
79	429	163	321	500	18	288	254
48	462	196	290	403	113	383	157
276	242	512	30	175	333	67	417

III

306	212	478	64	141	367	97	387
14	496	226	260	433	83	349	191
109	399	129	355	466	52	318	224
337	179	445	95	238	272	2	484
199	293	43	457	380	154	408	118
507	25	279	245	72	422	172	330
412	122	376	150	39	453	203	297
168	326	76	426	283	249	503	21

IV

423	69	331	169	28	506	248	278
155	377	119	405	296	198	460	42
252	282	24	502	327	165	427	73
456	38	300	202	123	409	151	373
82	436	190	352	493	15	257	227
366	144	386	100	209	307	61	479
269	239	481	3	178	340	94	448
49	467	221	319	398	112	354	132

V

381	159	401	115	194	292	46	464
65	419	173	335	510	32	274	244
34	452	206	304	413	127	369	147
286	256	498	20	161	323	77	431
140	362	104	390	311	213	475	57
440	86	348	186	11	489	231	261
471	53	315	217	108	394	136	358
235	265	7	485	344	182	444	90

VI

492	10	264	230	87	437	187	345
216	310	60	474	363	137	391	101
183	341	91	441	268	234	488	6
395	105	359	133	56	470	220	314
29	511	241	275	418	68	334	176
289	195	461	47	158	384	114	404
322	164	430	80	253	287	17	499
126	416	146	372	449	35	301	207

VII

96	446	180	338	483	1	271	237
356	130	400	110	223	317	51	465
259	225	495	13	192	350	84	434
63	477	211	305	388	98	368	142
425	75	325	167	22	504	250	284
149	375	121	411	298	204	454	40
246	280	26	508	329	171	421	71
458	44	294	200	117	407	153	379

VIII

201	299	37	455	374	152	410	124
501	23	281	251	74	428	166	328
406	120	378	156	41	459	197	295
170	332	70	424	277	247	505	27
320	222	468	50	131	353	111	397
4	482	240	270	447	93	339	177
99	385	143	365	480	62	308	210
351	189	435	81	228	258	16	494

图7-5

反幻方

现在，我们从三维回到二维。你知道反幻方吗？实际上，数学里有正有负，生物界有阳有阴，和谐共生，正是大自然的法则。因此，反幻方的出现也自然不过了。

美国趣味数学家马丁·加德纳首先提出了反幻方的概念：

若把n^2个连续自然数1,2,…,n^2按照某种规则排成一个n阶方阵，使得每行、每列及两条对角线上n个元素之和都不相等，则称这个方阵为反幻方。

当n为不小于3的奇数时，马丁·加德纳给出了一个构造反幻方的方法。图7-6是他构造的三阶与五阶反幻方。

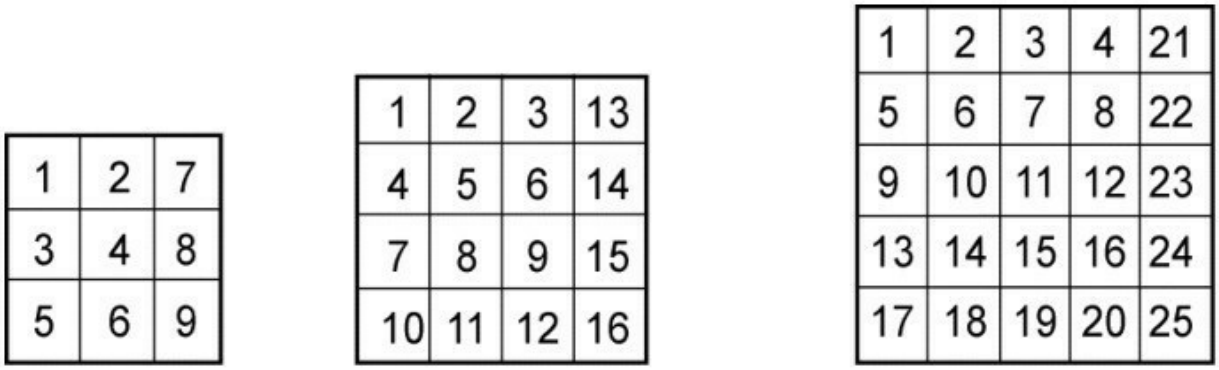
1	2	3
8	9	4
7	6	5

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

图7-6

大家可以看到，这两个反幻方都是将1填入左上角第1格，然后按顺时针螺旋形顺次填数。但马丁·加德纳的方法并不适用于偶数阶。我国数学家梁培基提出了一个构造法，可以构造出任意$n \geq 3$阶的反幻方。图7-7是n为3、4、5阶反幻方的实例，方法十分简单。这种构造方法可用

下述反幻方定理来说明，它也是由梁先生所创造的。



1	2	7
3	4	8
5	6	9

1	2	3	13
4	5	6	14
7	8	9	15
10	11	12	16

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

图7-7

反幻方定理：若n（$n \geq 3$）阶方阵为$A=[a_{ij}]$，

$$a_{ij}=\begin{cases}(i-1)(n-1)+j & i=1,\cdots,n,\quad j=1,\cdots,n-1\\ n(n-1)+i & i=1,\cdots,n,\quad j=n\end{cases}$$

则A　是一个n　阶反幻方。对于这个定理的证明，在此就不再叙述了，但特别要说的是，梁培基是一位农民数学家。

梁培基原来是农民，出生在黄河北岸的封丘县城关乡西孟庄村，中学未毕业就因家境贫困而辍学，开始务农，但他始终怀揣着一个学习数学的梦想。梁培基只读过初中二年级，他的老师说：“他的家距离学校很远，每天步行十几里路到学校。他吃的不如人，穿的不胜人，用的低于人，但是学习成绩特别是数学成绩突出，远远高于其他人。”

由于改革开放的政策，农民经过努力奋斗成为种田能手、劳动模范、农民企业家、农民艺术家的大有人在，而成为“数学家”的可谓凤毛麟角。梁培基这个面朝黄土背朝天、裤管上卷的“泥腿子”却迷上了幻方。他从事组合数学研究30余年，在数学领域取得了累累硕果，被破格录用为国家干部，破格晋升为副研究员，成为名副其实的“农民数学家”。梁培基敢于向国际难题叫板，并且屡屡攻克难关，不断创新，打

破多项“世界之最”的记录。他在这一领域中积极探索，取得了突出的成绩，发表了20余篇论文。

笔者在此插上一段梁先生的简介，就是鼓励大家要实干苦干，努力实现梦想！

第八章 泰森多边形

在《最强大脑》第三季中英对决赛中，出现了一个很难的竞技项目——泰森多边形。屏幕左侧出现一个多边形，右侧是几百个离散点，选手们要利用建立泰森多边形的规则，在脑海中找到一个与左侧一致的多边形，两个图形重合一致，并用时较少者获胜。

该场竞赛的国际评审托马斯表示：“这个项目是对参赛者眼力、脑力、空间想象力及对成像、几何图形等多方面知识熟练程度的考验，非常难。”主持人蒋昌建表示：“如果你能看懂这个规则，你就是半个最强大脑。”比赛持续了长达3个小时的时间，最后中方获胜。

肥皂泡的启示

在一个玻璃杯里配制比较浓的肥皂水，用一根吸管去吹，杯内会生出许多相互挤压的肥皂泡；上海人有个饮食习惯叫作吃泡饭，就是把吃剩的米饭倒入凉水中放在灶上煮，等开锅后就成了泡饭，在开锅时也会有类似于肥皂泡的泡沫挤压在一起；在平面上，也有相似的情况发生，例如极度干旱的大地上，土地龟裂成一块一块的，像一个个多边形，等等。

这些几何形象具有数学上的统一性。在平面中形似杂乱无章的凸多边形，在空间里形似杂乱无章的凸多面体，人们称诸如此类的结构为无序结构。系统研究这种结构的第一人是俄国数学家沃热诺伊（Voronoi），所以这种结构被称为Voronoi图。

数学家牛顿曾观测过肥皂泡。肥皂泡的形状受表面张力的控制，表面张力总是使表面积尽可能地小。每个肥皂泡里都包封有一定量的空气，结果使得表面积的减少有了一个最低限度，这就解释了为什么单个肥皂泡总是会变成球状，而一大堆肥皂泡集在一起便会产生不同的造型。牛顿发现，在肥皂泡沫中，肥皂泡的边缘相交成120°的角，这称为“三部结合”。在一个三部结合点处有3条线段相会，两两相交成120°角（如图8-1所示）。

3个 120°使人们不禁想起费马定理：在△ABC 内找一点 P ，使PA +PB +PC 的距离之和最小，答案就是P 点对三点的张角均为120°。距离是数学研究的基本对象之一，数学家指出：“Voronoi图实质是一种在自然界中宏观和微观实体间以距离 相互作用的普遍结构。”

我们无法揣测沃热诺伊创建Voronoi图的灵感和出发点，但上述的

肥皂泡结构是否能给大家以启示呢？

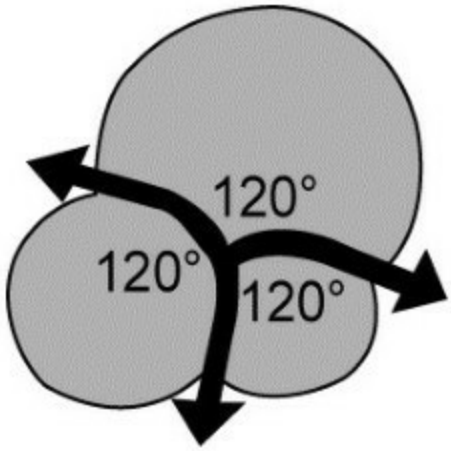


图8-1

Voronoi图的构造

Voronoi图可由下面的问题引出：

给定平面中的N个点，对于任意一点P_i，平面中距离点P_i比距离其他点更近的区域是什么？即区域内的任意一点（x，y）距P_i比距离平面中其他给定的点都近。

首先从最简单的情况入手，对于平面中的任意两点A、B，距离A点比距离B点近的区域是由线段A、B的垂直平分线确定的包含A的那个半平面V（A）。如果点集由N个点组成，距离点P_i比距离其他点更近的点的区域是包含点P_i的N-1个半平面的交集。这N-1个半平面是由点P_i与其他点的垂直平分线确定的。显然，V（i）是由一些垂直平分线构成的多边形。它将整个平面分成N个区域，每个区域包含一个点，称为种子点。

其次，Voronoi图的边是点集中某对点的中垂线上的一条线段或者射线。Voronoi图至多有2n-5个顶点和3n-6条边。在Voronoi图的构建中，首先要将离散的点构成三角网。对于给定的初始点集P，有多种三角网剖分方式，其中Delaunay三角网具有以下特征。

① Delaunay三角网是唯一的；

②三角网的外边界构成了点集P的凸多边形“外壳”；

③没有任何点在三角形的外接圆内部；

④形成的三角网总是具有最优的形状特征，任意两个相邻三角形形成的凸四边形的对角线如果可以互换的话，那么两个三角形的6个内角中最小的角度不会变大。

在实际操作构网时，总是选择最临近的点形成三角形，并且不与约束线段相交。图8-2是用8个点构造三角网的例子。

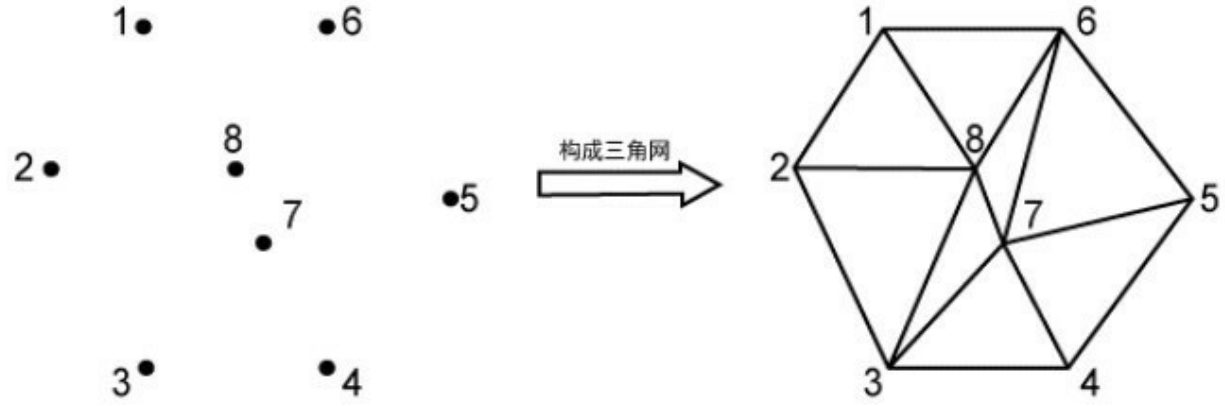


图8-2

构建Delaunay三角网后，求出各三角形的外心（外接圆圆心）并连接，就可以得到Voronoi多边形。在图8-3中，各小圆点即已知的离散点，实线三角形是Delaunay三角网，虚线多边形即Voronoi多边形。

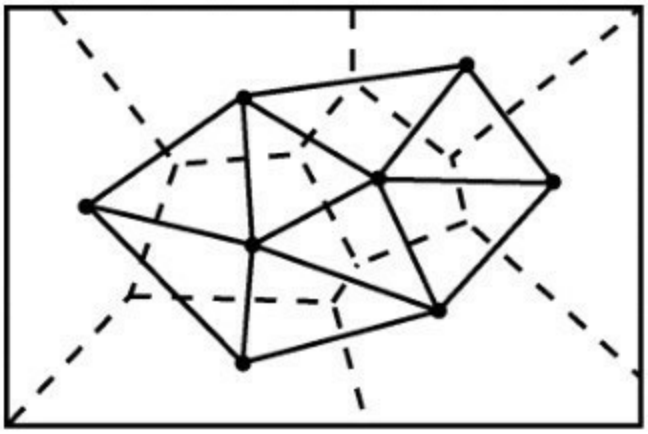


图8-3

Voronoi图的应用实例

● 气象学：荷兰气候学家泰森提出了一种根据离散分布的气象站的降雨量来计算平均降雨量的方法，即将所有相邻的气象站连成三角形，然后做出每个三角形各边的垂直平分线，于是每个气象站周围的若干垂直平分线便围成一个多边形。用这个多边形内所包含的唯一一个气象站的降雨量来表示这个多边形区域内的降雨强度，并称这个多边形为泰森多边形。所以，Voronoi图又称泰森多边形。

●人类学和考古学：考察由不同的部落、首领、堡垒等所确定的势力范围或影响范围。

●天文学：识别星群和星系群，比如由太阳和其他恒星所确定的星系。

●机器人技术：存在障碍物情况下的路径规划。

●统计学：分析统计聚类。

●动物学：动物疆域的分析。

最有趣的一个应用是近年来发展起来的细胞几何学。对于三维空间的情况，只需把平面情形的垂直平分线改成垂直平分面，凸多边形由凸多面体代替，面积由体积代替，上述论述即可推广到三维空间，并拥有一套相似的概念与结论。

上述数学模型反映的实际情况是：每个细胞核都以相同的速度向各个方向均匀生长，直到细胞间互相接触挤压而停止，于是构成了细胞的机体。

Voronoi图的概念源于生活，对它的定义、求解和研究都是为了应用于实际以产生价值，反过来应用又推进了研究。创新和应用正是数学

发展的根本动力。

第九章 巴克球模型

如果你在百度上搜索词条“巴克球”，它会告诉你两个答案：一是化学上碳60（C_{60} ）的俗称，也称为富勒烯；二是一种由带有磁性的金属实心圆球组成的益智玩具。

巴克球材料为由钕铁硼（NdFeB）磁矿石精细加工而成的球状强磁石，具有强磁特性，能组合出数亿种几何图案，有极高的娱乐性和创造性。大家可以通过发挥才智，创造出属于自己的作品。此外，它外观精致，光泽亮丽，不易褪色，既可以充当玩具，又可以作为饰品。

《最强大脑》第五季第一场国际赛中把巴克球作为一个挑战项目。在这项挑战中，由多个磁力球结合而成的巴克球模型，通过纷繁复杂的排列方式，呈现出绚烂多姿的样貌。选手需要通过观察，在脑海中把这些模型拆解成不同的单元，然后计算出模型所用的巴克球数量。

在这场挑战中，80个巴克球模型让所有嘉宾都忍不住感叹造物的神奇，有的嘉宾表示像是观看珠宝展。这些绝美的模型中暗藏玄机，模型中共点、共边、共面的情况会增加计算的难度，不少弧面结构还需要选手在脑海中把它展开成平面结构以方便计算——这对选手的空间想象力和计算能力，都是极大的挑战。

本文并非针对某一具体的巴克球模型介绍其计算方法，而是把巴克球作为道具，介绍两个著名的数学难题。

正四棱锥方程

法国数学家卢卡斯早在1875年便向《新数学年鉴》提出如下问题。

用球（例如巴克球）堆成正九棱锥（如图9-1所示，各层球数分别为1,4,9,16,…），整个锥体所堆球数能不能是完全平方数？换言之，方程$1^2+2^2+3^2+4^2+\ldots+n^2=m^2$（$m$，$n$均是整数）有无解？

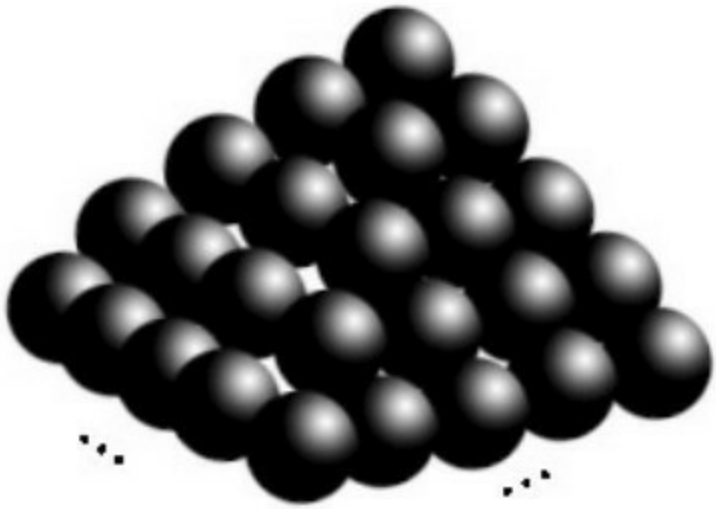


图9-1

卢卡斯此时已有相当大的名气，他在提出该问题时，竟武断地认为此方程无解。但次年，勃兰克便给出一组解：

$n=24$，$m=70$

解的唯一性问题直到1918年才由瓦特松利用椭圆函数给出了一个严格的证明，从而否定了当初卢卡斯的断言。老虎也有打盹的时候，在数学中大师也会犯错误，这也正常。但提醒大家：千万不能凭借不完全的归纳或直觉，妄下断言。

由$1^2+2^2+3^2+\ldots+23^2+24^2=70^2$，我们不由想到完全正方形问题，

从上式可以看出边长分别为1～24的正方形的面积的和恰好与一个边长为70的正方形的面积相等。我们也许会想这样一个问题：边长为70的正方形，可否恰好分割成边长为1,2,…,23,24的小正方形？答案是否定的。迄今为止，最好的答案是大正方形分割为24个小正方形（其中边长为1的正方形有2个，边长为7的正方形却没有，如图9-2所示），仅剩下总面积为48的6个小条，人们称其为“拟完美正方形”。

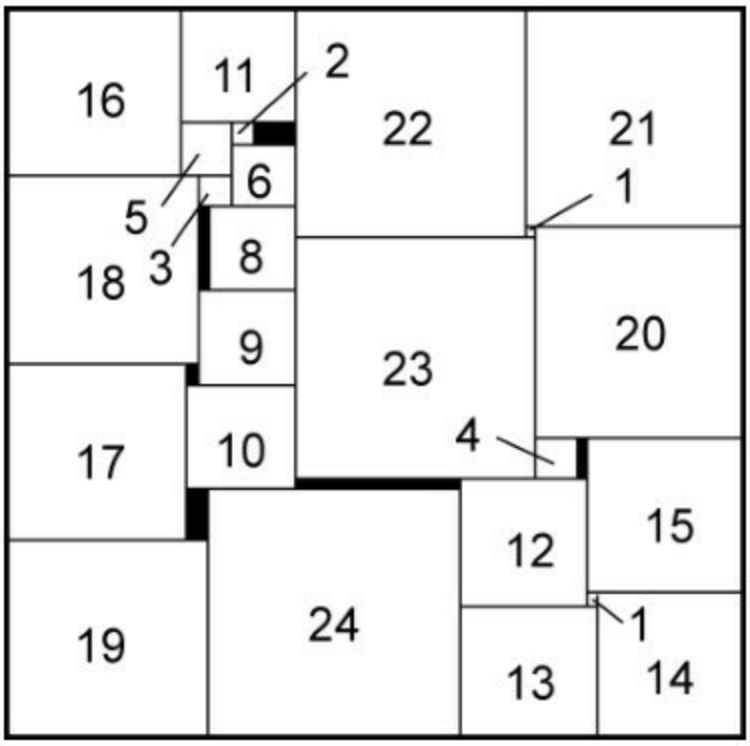


图9-2

十三球问题

1694年，英国牛津的天文学家格雷戈里与他的朋友——大名鼎鼎的牛顿讨论一个问题：一个单位球能否与13个（互不相交的）单位球相切？就像一个巴克球能否在其表面吸附13个巴克球，牛顿认为不可能，而格雷戈里则猜测一个单位球能够与13个单位球相切。

如图9-3所示，当单位球A与单位球O相切时，点O与球A的切线形成一个圆锥。这个圆锥含有球面A的一个球冠，切点A_1就是球冠的极（顶点）。点A_1与该球冠上任一点的球面距离$d\leqslant\frac{\pi}{6}$，$\frac{\pi}{6}$即为这个球冠的半径。同样，对另一个与圆O相切的单位球B，也有一个以切点B_1为极、$\frac{\pi}{6}$为半径的球冠。于是，格雷戈里的猜测等价于下面的问题：球面上能否有13个半径为$\frac{\pi}{6}$的球冠，且它们互不重叠？

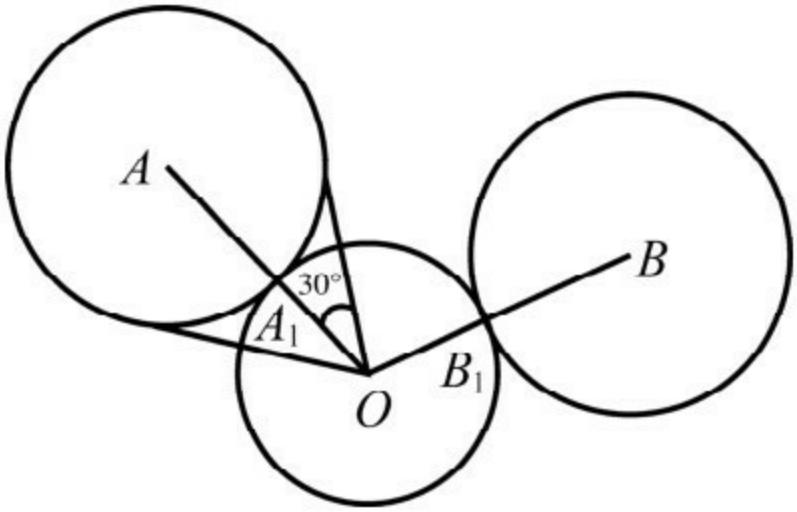


图9-3

由于$\overset{\frown}{A_1B_1} \geqslant 2\times\frac{\pi}{6}=\frac{\pi}{3}$，所以$A_1B_1 \geq OA_1=1$，即每两个切点之间的（直线）距离$d \geq 1$。格雷戈里猜测还可以再换一个等价的说法：单位球面上能否有13个点满足每两个点之间的直线距离$d \geq 1$？

在平面上，问题就简单多了，二维的“球”就是圆。设⊙(O ,1)是圆心为O 、半径为1的圆，至多有多少个互不重叠的单位圆与⊙(O ,1)相切？图9-4表明可以有6个互不重叠的单位圆与⊙(O ,1)相切，它们的圆心组成一个边长为2的正六边形［正好是⊙(O ,2)的内接正六边形］。这6个圆“挤”得很紧，第7个圆显然没有立锥之地。但是直观上的显然，不能代替严谨的证明，怎样证明与⊙(O ，1)相切的圆至多有6个呢？

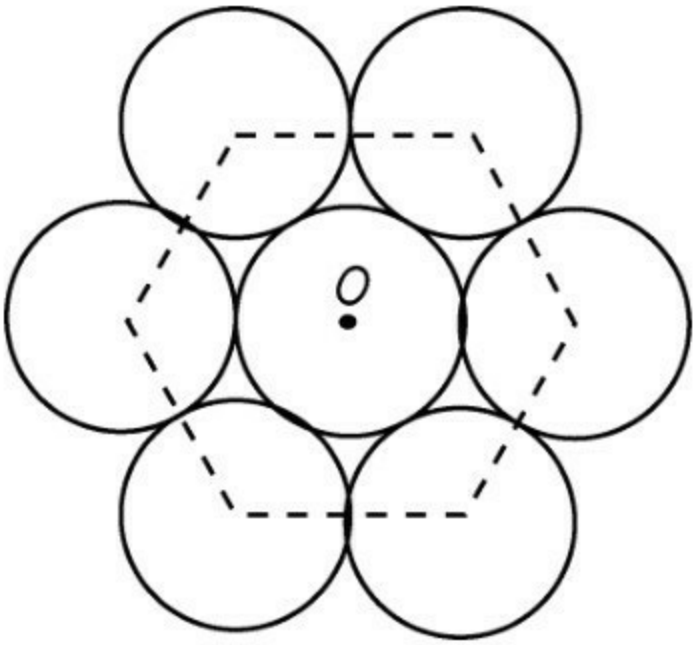


图9-4

证明并不困难。如图9-5所示，在⊙(A ,1)和⊙(B ,1)互不重叠而且均与⊙(O ,1)相切时，$AB \geq 1+1=2$，$OA=OB=1+1=2$。所以在△AOB 中，边AB 最大，从而$\angle AOB \geq 60°$。因为点O 处的周角是360°，所以像⊙(A ,1),⊙(B ,1),…这样互不重叠且均与⊙(O ,1)相切的圆至多有6(=360°÷60°)个。

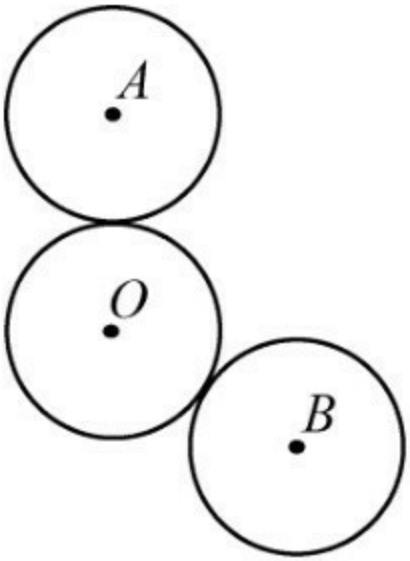


图9-5

空间的情形复杂得多。不难举例说明，可以有12个单位球与同一个单位球相切。最简单的例子是将单位球一层一层地堆起来，最上面1个，第2层4个，第3层9个，第4层16个。那么，在第3层核心的那个球既与上一层的4个球相切，又与下一层的4个球相切，还与同一层的4个球相切。

另一个很自然的例子就是球的内接正二十面体。它有12个顶点，20个面，每个面均为正三角形。不难算出这个正二十面体的每条棱长为

$$2\sin\left(\arccos\frac{1}{2\sin\frac{\pi}{5}}\right)=\frac{1}{5}\sqrt{50-2\sqrt{125}}=1.0514\cdots>1$$

即12个顶点两两间的距离大于1，这12个顶点可以作为12个单位球与球O 的切点。

另一方面，我们可以证明与单位球O 相切又互不重叠的单位球不超

过14个，也就是在球面O 上两两距离不小于1的点的个数不超过14个。为此，考虑半径为$\frac{\pi}{6}$ 的球冠（如图9-6所示），球冠的高

$$h = A_1H = 1 - OH = 1 - \frac{\sqrt{3}}{2}$$

所以，球冠的面积$2\pi h = \pi(2-\sqrt{3})$ 。因为球面积为4π，所以互不重叠、半径为$\frac{\pi}{6}$ 的球冠的个数

$$\leqslant \frac{4\pi}{2\pi h} = \frac{4}{2-\sqrt{3}} = 4(2+\sqrt{3}) = 14.928\cdots$$

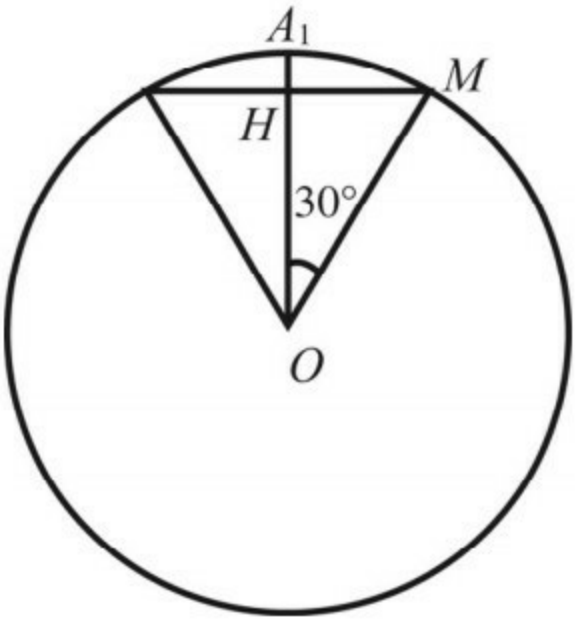


图9-6

因此，15个半径为$\frac{\pi}{6}$ 的球冠必有重叠。换句话说，互不重叠且半径为$\frac{\pi}{6}$ 的球冠个数不超过14。

虽然大多数人倾向于牛顿的观点，认为至多有12个单位球同时与一个单位球相切，但严格的证明姗姗来迟，直至1953年（将近260年之后）才由许特与著名代数学家范德瓦尔登给出证明。1956年，加拿大的利奇给出了一个较为简单的证明，这一证明需要掌握一些球面几何知识，有兴趣的读者可参阅文献【7】。

第十章 质数和密码

《最强大脑》第二季晋级赛第二场中，有一个项目叫“国家宝藏”。屏幕上出现一个六面体，各个面上都填满了五位数。把六面体展开来，屏幕上满满一片数字，大约有2000个。在这些数字中，有7个是质数（也叫素数）。由这7个质数连成的直线仅有两条互相相交，其交点即宝藏之藏匿点，要求求出该点的坐标。显然，挑战成功的核心就是找出这些质数，也就是要找出这7个不能作因数分解的五位数来。

早在2000多年前，古希腊人就对质数作了定义。此外，欧几里得做了一个质数的个数是无限的经典证明，埃拉托塞尼创造了“筛法”，以找出整数集中的质数。现在，高年级小学生都学会了因数分解。

自古以来，质数都是纯数学的研究对象，常年躺在象牙塔中，除数学家外无人问津。但时光荏苒，万物反转。1978年，竟有一种“不可破译”的密码出现。据说，它令不少数论学家身价倍增，成了保密部门争相抢夺的人才。一边是简单易懂的因数分解，另一边是神秘的密码体系，如此意外的联姻，其黏合剂竟是大整数的因数分解，实在令人吃惊。

在此，笔者不想谈这些五位数是如何找出的，只是与读者一起进行这次“质数·密码”的穿越时空的奇幻旅行。这自然要从传统密码说起。

传统密码示例

甲把一条信息发送给乙，要求除甲乙二人之外，不能让第三人得知该信息的含义，这种通信称为保密通信。用自然语言或其他能为常人所识的符号传递信息的信号称为明文，把明文作某种变换与伪装后得到的信号称为密文。把明文变换成密文的过程称为加密，加密规则叫作密钥；把密文还原成明文的过程称为解密。

保密通信自古有之，不但至今不衰，而且已经发展为一个受到各国政府与军方十分重视的研究课题。这一节就与大家共享传统密码的轶事与其方法。古希腊有斯巴达天书，传说斯巴达派一奴隶到前线给莱桑德将军送来一条腰带，如图10-1所示。莱桑德拿到腰带后，把它缠在一根木棍上（如图10-2所示）得明文

Kill King（杀死暴君）

在这个例子中，图10-1中的符号是密文，图10-2中的符号是明文，密钥是“腰带缠木棍”。

图10-1

图10-2

用数学语言来表示，上述密钥等价于把$i \equiv 1 \pmod 4$ 的符号抄出即得明文。密文中其余的字母是随机的，目的是为了加密，使外人不易破译。可惜当时数学还没有发展到这个程度，莱桑德将军还得试用直

径合适的木棍，否则就可能会出现乱码，不明其意。

还有一种密码称为代换密码。据说是古罗马时期的恺撒大帝首先发明的，因此又称恺撒密码。首先，分别用数字0至25代表英文字母A至Z。密钥是加一整数k 后以26为模，1至25的整数皆可以当作密钥的特定值（若k 为0即为不加密）。例如，若k =7，要传递的信息为CAT，则密码为JHA。当收到JHA后，应将每个字母的序号减7，得到的数字对应的信息即为CAT。

法国密码专家维吉尼亚利用0到25对应26个拉丁字母编号，再取一个英语单词作密钥，发明了一种加强版的代换密码，写成Radio=（17，0， 3，8，14）。如果明文是I am Wang Shuhe，写成8，0，12，22，0，13，6， 18，7，20，7，4。加密过程如下：

	I	a	m	W	a	n	g	S	h	u	h	e
+)	R	a	d	i	o	R	a	d	i	o	R	a
	8	0	12	22	0	13	6	18	7	20	7	4
+)	17	0	3	8	14	17	0	3	8	14	17	0
	25	0	15	30	14	30	6	21	15	34	24	4

对上述求得的数列进行处理，即求被26除的余数（mod 26），得数列

25 0 15 4 14 4 6 21 15 8 24 4

这个数列在自然顺序的字母表（以0到25编号）中对应的字母为

Z A P E O E G V P I Y E

此即密文！解密过程如下：

	Z	A	P	E	O	E	G	V	P	I	Y	E
−)	R	A	D	I	O	R	A	D	I	O	R	A
	(25-17)	(0-0)	(15-3)	(30-8)	(14-14)	(4-17)	(6-0)	(21-3)	(15-8)	(8-14)	(24-17)	(4-0)
	8	0	12	22	0	−13	6	18	7	−6	7	4
mod 26	8	0	12	22	0	13	6	18	7	20	7	7
	I	A	M	W	A	N	G	S	H	U	H	E

此外，还有一种要用到乘法规则的密码。例如，密文是
00110100101010110111100000111001111000101010110000110111011010(
00111011100001111001111110101010111110011100100111111000101010
1011011，密钥规则是（α_1 ,α_2 ,α_3 ,α_4 ,α_5 ）·（1,2,4,8,16），解密过程如下。先把上述密文以每5个码为一组进行点乘，得出

(0, 0, 1, 1, 0) · (1, 2, 4, 8, 16) =4+8=12=L

(1, 0, 0, 1, 0) · (1, 2, 4, 8, 16) =1+8=19=I

(1, 0, 1, 0, 1) · (1, 2, 4, 8, 16) =1+4+16=21=U

(1, 0, 1, 1, 1) · (1, 2, 4, 8, 16) =1+4+8+16=29=

(1, 0, 0, 0, 0) · (1, 2, 4, 8, 16) =1=A

(0, 1, 1, 1, 0) · (1, 2, 4, 8, 16) =2+4+8=14=N

(0, 1, 1, 1, 1) · (1, 2, 4, 8, 16) =2+4+8+16=30=

(0, 0, 0, 1, 0) · (1, 2, 4, 8, 16) =8=H

(1, 0, 1, 0, 1) · (1, 2, 4, 8, 16) =1+4+16=21=U

(1, 0, 0, 0, 0) · (1, 2, 4, 8, 16) =1=A

(1, 0, 0, 0, 0) · (1, 2, 4, 8, 16) =1+2+8+16=27=

(1, 0, 1, 1, 0) · (1, 2, 4, 8, 16) =1+4+8=13=M

(1, 0, 0, 1, 0) · (1, 2, 4, 8, 16) =1+8=9=I

(0, 1, 1, 1, 0) · (1, 2, 4, 8, 16) =2+4+8=14=N

(1, 1, 1, 0, 0) · (1, 2, 4, 8, 16) =1+2+4=7=G

(0, 0, 1, 1, 1) · (1, 2, 4, 8, 16) =4+8+16=28=

(1, 0, 0, 1, 1) · (1, 2, 4, 8, 16) =1+8+16=25=Y

（1，1，1，1，0）·（1，2，4，8，16）=1+2+4+8=15=O

（1，0，1，0，1）·（1，2，4，8，16）=1+4+16=21=U

（0，1，1，1，1）·（1，2，4，8，16）=2+4+8+16=30=

（1，0，0，1，1）·（1，2，4，8，16）=1+8+16=25=Y

（1，0，0，1，0）·（1，2，4，8，16）=1+8=9=I

（1，1，0，1，1）·（1，2，4，8，16）=1+2+8+16=27=

（1，1，0，0，0）·（1，2，4，8，16）=1+2=3=C

（1，0，1，0，1）·（1，2，4，8，16）=1+4+16=21=U

（0，1，1，1，0）·（1，2，4，8，16）=2+4+8=14=N

（1，1，0，1，1）·（1，2，4，8，16）=1+2+8+16=27=

可将密文译成

Liǔ àn huā míng yòu yīcūn

汉语译文就是：柳暗花明又一村。

当然，密文中不只有ABCD或0和1，还有一些奇特的符号。下面介绍一个有趣的例子。欧洲中世纪有不少秘密团体。其中，阿尔卑斯山区的石工人数众多，势力不小，他们发明了一种颇为别致的代换密码。这种密码用两种九宫格与两个交叉十字来代表英语中的26个字母（如图10-3所示）。

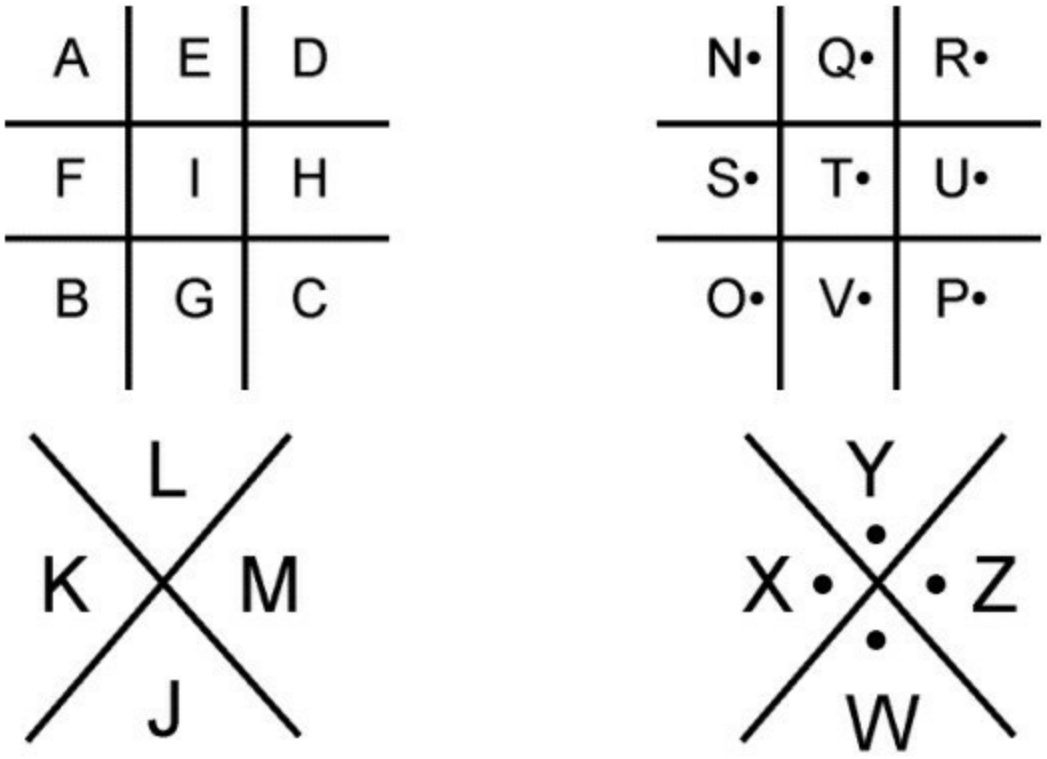


图10-3

如下图形的密码

译成英语就是The King is deed（国王死了）。原来当时在位的英国国王死后秘不发丧，因此石工们传出了这一秘密消息。

这些例子举不胜举，方法多种多样，目的只有一个，就是不让第三者知道消息的内容。但解密者自然不甘落后，用尽浑身解数来一窥机密。接下来，不妨来看一看解密者的智慧。

福尔摩斯巧破密码案

在英国，最有名的侦探小说是亚瑟·柯南道尔所写的《福尔摩斯探案全集》，其中《归来记》里的《跳舞的人》就是一篇以密码为主题的短篇小说。现在，我们来看一下福尔摩斯在此密码案中所展现的智慧与才华，领略解密的逻辑与技巧。当然，此处仅提及与密码术有关的情节。

一年前才与艾尔西结为连理的丘比特写了一封信给福尔摩斯，信内附了一张纸条，纸条上横着画了一些正在跳舞的奇形怪状的小人，如图10-4所示。

图10-4

因为艾尔西一看到这张纸条便昏倒了，所以丘比特写了一封信给福尔摩斯并于次日到伦敦拜访。跟福尔摩斯会面后的隔天早上，丘比特发现另一系列跳舞的小人被用粉笔画在了工具房门上，如图10-5所示。

图10-5

过了两个早上，又出现了新的小人，如图10-6所示。

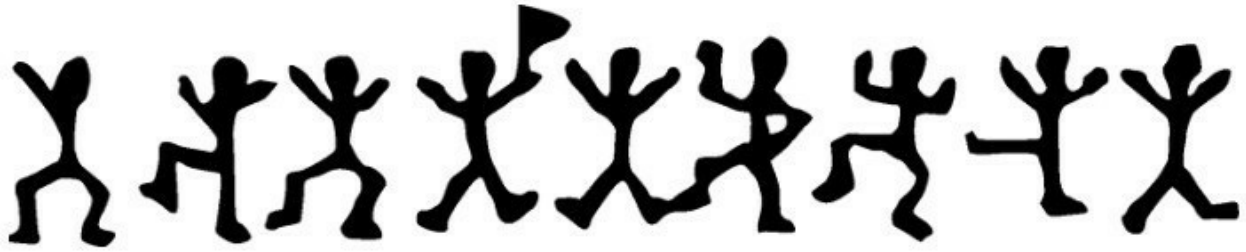


图10-6

三天后，在日晷仪上发现一张纸条，很潦草地画了一行小人，跟上次的完全一样。那夜在工具房门上又有人画了一行跳舞的小人，排列方式跟前两次的完全相同。隔天早上那扇门上除了已经有的那行小人外，又添了几个新画的，如图10-7所示。

图10-7

图10-5到图10-7是丘比特第二次拜访福尔摩斯时提供的。当时福尔摩斯十分兴奋，丘比特的背影一消失，他就急忙把所有纸条都摆在面前，开始分析。最后，他满意地叫了一声，显然已有重大突破。后来，他发了一封很长的电报给某人，过后就等着回电。第二天晚上来了一封信，丘比特说他家里平静无事，只是那天清早日晷仪上又出现了一长行跳舞的小人，他临摹了一张，附在信里寄了过来，如图10-8所示。

图10-8

福尔摩斯对着这些怪异的图案看了几分钟，发出一声沮丧的喊叫，接着对华生表示他们应尽快赶到马场村庄园。过后没多久，他所盼望的电报来了，看完后表示急需让丘比特知道目前的情况。隔了一天，他与华生抵达马场村庄园时，丘比特先生已中弹身亡，而他的太太艾尔西也中弹且情况相当危急。福尔摩斯问了一些问题后，就叫人送了一则短信息给艾尔里奇斯农场的阿奥·斯兰尼先生。处理完这一切，福尔摩斯向华生和警长解释了他是如何破解那几张画有小人的纸条的。他说只要看出这些符号（即跳舞的小人）代表字母，再应用密文的规则来分析，就不难找出答案。

“第一张纸条上的那句话很短，我只能稍有把握确定图10-9的小人代表字母E，因为在英文句子中E最常见，它出现的次数多到即便在一个较短的句子中也会重复。第一张纸条上的15个符号中有4个完全一样，因此把它看成E是合理的。这些图形中有的带一面小旗，有的没有，从小旗的分布来看，带旗的图形可能是用来把句子分成一个个单词的。我把这看作一个可以接受的假设，同时记下E是用图10-9表示的。

图10-9

“最难的问题来了，因为除了E以外，英文字母出现次数的顺序并不明确，大致来说出现次数多少的顺序为

T，A，O，I，N，S，H，R，D，L。

其中T，A，O，I出现的次数不相上下（英文字母中最常出现的前9个的频率是e：0.127；t：0.091；a：0.082；o：0.075；i：0.070；n：0.067；s：0.063；n：0.061；r：0.060），要是把每一种组合都试一遍，那会是一项没完没了的工作，所以只好等新材料来了再说。丘比特第二次来访的时候给了我另外两句短语和似乎只有一个单词的话，就是这几个不带小旗的符号。在这个由5个字母组成的单词中，第二个和第四个都是E，这个单词可能是sever（切断），也可能是lever（杠杆）或者nerer（决不）。无疑，使用最后这个词来回答一项请求的可能性极大，而且很可能是丘比特太太写的答复。假如这个判断正确，现在就多了3个符号分别代表N，V和R。但困难仍很大，然后一个很妙的想法使我知道了另外几个字母：我想这些恳求是来自一个在艾尔西年轻时就跟她亲近的人，那么一个两头是E，当中有3个别的字母的组合很可能就是ELSIE（艾尔西）这个名字。这一来我就找出了L， S和I。可是究竟恳求什么呢？在ELSIE前面的一个词只有4个字母，末尾是E，这个词必定是COME（来），我试过其他多种以E结尾的4字母单词，都不符合情况。这样就找出了C，O和M，现在可以回头再分析第一句话，把它分成单词，还不知道的字母就用点代替。经过这样的处理，这句话就变成

·M·ERE··ESL·NE·

“现在，第一个字母只能是A，这是最有帮助的发现，因为在这个短句中出现了3次。第二个词开头是H也是显而易见的，这句话就变成了

AMHEREA·ESLANEY

再把名字中所缺的字母添上

AMHEREABESLANEY（我来了，阿贝·斯兰尼）

有了这么多字母，就有把握解释第二句话了。这一句是这样的

A·ELRI·ES

我一看缺字母的地方只有加上T和G才有意义（意为住在艾尔里奇斯），并确定这个名字是写信人住的旅店或地方。‘后来怎么样，先生？’警长问。福尔摩斯说：‘我有充分的理由猜想阿贝·斯兰尼是美国人，因为阿贝是美式语言，而这些麻烦的起因又是来自从美国寄来的一封信，我也有充分的理由认为这件事带有犯罪的内情。女主人说的那些暗示她过去的话和她拒绝把实情告诉她丈夫的行为都使我往这方面去想。所以，我才给纽约警方的一个朋友发了一封电报，他回电说：阿贝·斯兰尼正是芝加哥最危险的骗子。就在我接到回电的那天晚上，丘比特寄来了最后一行小人，用已知的字母译出来就是

ELSIE·RE·ARETOMEETTHYGO

再加上两个P，这句话就完整了（艾尔西准备见上帝）。这说明这个流氓已从劝诱改为恐吓，所以我和华生立即赶去，但不幸的是仍晚了一步。”

福尔摩斯一说完，警长就急着去艾尔里奇斯逮捕斯兰尼。福尔摩斯说不必了，斯兰尼很快就会送上门来。不一会儿，斯兰尼来了，一进门就被警长戴上了手铐。斯兰尼承认罪行，并说这种秘密文字是艾尔西的父亲发明给他们在芝加哥的帮派所使用的，斯兰尼和艾尔西订过婚，但艾尔西无法容忍他们帮派的行径，就独自逃到伦敦。斯兰尼找到艾尔西的住处后，就发出了秘密信息。

奇怪的是，为何斯兰尼会自投罗网呢？福尔摩斯写的信息如图10-10所示。

图10-10

你会发现它的意思是“马上到这里来”（COMEHEREATONCE）。斯兰尼当然确信这必定是艾尔西发来的，因为除了她别人不会知道这种秘密文字的书写规则。因此，他走进了这个圈套。

公钥密码的孕育

自古以来，加密与解密就如同魔道斗法，“魔高一尺，道高一丈”，此消彼长，循环不止。特别是在第二次世界大战期间，战事紧急，密码纷飞，隐藏战线的斗争丝毫不亚于冲锋肉搏的惨烈。待第二次世界大战结束后，硝烟散尽，世界平静，有智者终于抚平伤痛，潜心反思：密码的根本属性是什么？

我们分析一下上两节中关于传统密码的几个示例，可以发现一个共性，即密码的安全性完全依赖于密钥的秘密性。如何打破这个困境呢？当然得从密钥这边来动脑筋。传统密码最大的缺陷在于其加密钥匙和解密钥匙是对称的，也可以说解密钥匙很容易从加密钥匙推导出来，甚至有时候简单到解密钥匙就是加密钥匙。所以，突破之法就在于：打破加密钥匙和解密钥匙之间的对称性。　这样的话即使给你加密钥匙，你也没有办法计算出或猜到解密钥匙。

在思考这个问题时，我们会从钥匙联想到门。许多公共建筑的大门从门内到门外时只要轻轻一推即可，毫无困难；但反过来则不行，必须有钥匙才能从门外进入建筑物内。从门内推门，表面上好像不需要钥匙，实际上是因为推的动作每个人都知道，可以看成公开的钥匙。门里门外是全然不同的两个世界，应如何打破加密钥匙与解密钥匙之间的对称性呢？乍看起来似乎不可能，然而门的比喻给了我们一些启发与暗示：出去简单、容易、快速；进来复杂、困难、缓慢。　这样的东西到底是什么呢？是一种演算法，还是一个函数？这引起了数学家们的关注。

美国斯坦福大学的两位教授惠特菲尔德·迪菲及马丁·赫尔曼在1976

年发表了一篇对现代编码理论有重要影响的论文。在论文中，他们提出了一个所谓单一方向进行的密码函数，这种函数须具备下述特性。

①对每一个自然数，有唯一自然数$y=f$（x）与之对应；

②反函数f^{-1}存在，使$f^{-1}[f(x)]=x$；

③对这种函数f及其反函数f^{-1}存在有效的演算法；

④即使密码函数f及其运算被知道了，反函数也无法得到。

在上述4个条件中，条件④当然是最关键的。

在探索的历史过程中，质数判定和大数分解之间的不同步现象，引起了数学家的高度重视。我们可以判定一个一二百位的数是否是质数，也可以找出这么多位的质数，但无法有效地分解一个一二百位的大合数。这种不同步现象称为“大数分解困难”，这种不对称、不同步性正好符合一个方向简单、容易、快速，另一个方向复杂、困难、缓慢的编码理论条件。在此指导思想下，一种公开密钥的密码体系产生了。

RSA密码体系

1978年，数学家罗纳德·李维斯特、阿迪·萨莫尔和伦纳德·阿德曼创造了一种公开密钥体系，称为RSA密码。详细介绍RSA密码在这本书中是不可能也无必要的，下面简要介绍其过程，并以示例说明。

（1）RSA的加密

①把字母序列*ABC …XYZ* 编成平凡码：*A* =01，*B* =02，*C* =03，…，*Z* =26。

②宣布密钥 (*e* ,*n*)，*e* 、*n* 是自然数。

③对用平凡码编写的明文*M* 进行切分，这里*M* 是一个自然数。

$M=M_1M_2\ldots M_k$，使$M_i<n$，$i=1,2,\ldots,k$。

④计算$M_i^e=M_i'\pmod n$即得密文$M'=M_1'M_2'\cdots M_k'$。

（2）密钥 (*e* ,*n*)的制作

①选取两个百位级大奇数*p* ，*q* （这是绝密的）。

②公布*n* =*p* ×*q* 的*n* 值。

③记*r* =(*p* -1)(*q* -1)，*r* 不公开。

④ *e* 为一个自然数，且*e* 、*r* 的最大公约数是1（*e* 、*r* 互质），公布*e* 。

（3）解密方法

①选取自然数*d* ，使1≤*d* ≤*r* ，$ed\equiv 1\pmod r$，*d* 保密。

②$(M_i')^d\equiv M_i\pmod n$，得明文$M=M_1M_2\ldots M_k$。

上面即RSA密码的制作及解密方法。下面取两个小的质数举例说明：取*p* =73，*q* =97。于是有

① *n* =73×97=7081，*r* =(73-1)×(97-1) =6912。

②取e，使其满足$(e,r)=(e,6912)=1$，求得若干解，并取$e=101$，$d=2669$，其中$ed\equiv1(\mathrm{mod}\ 6912$，这里$r=6912)$。

③ 若取明文 M =Pounds，则明文的平凡码为$M=161521140419$ ($p=16,o=15,u=21,n=14,d=04,s=19$)。

④切分M 使$M=M_1M_2M_3$，$M_1=1615$，$M_2=2114$，$M_3=0419$。

⑤制作密文：$M'=M'_1M'_2M'_3$ ($e=101,d=2699$)。

$1615^e\equiv1615^{101}\equiv4226(\mathrm{mod}\ 7081)$，$M'_1=4226$。

$2114^e\equiv2114^{101}\equiv1582(\mathrm{mod}\ 7081)$，$M'_2=1582$。

$0419^e\equiv419^{101}\equiv0765(\mathrm{mod}\ 7081)$，$M'_3=0765$。

于是密码$M'=422615820765$。

⑥接收者解密：$d=2669$。

$4226^d\equiv4226^{2669}\equiv1615(\mathrm{mod}\ 7081)$

$1582^d\equiv1582^{2669}\equiv2114(\mathrm{mod}\ 7081)$

$0765^d\equiv0765^{2669}\equiv0419(\mathrm{mod}\ 7081)$

于是得明文$M=161521140419$（与明文的平凡码完全一致）。

对于RSA密码体系，第三者解密的难度非常大，难度的焦点就在于大质数$n=p\times q$ 的分解。事实上，当今用40秒便可以判定一个一二百位的自然数是否是质数，所以制造RSA者可以很快确定p 与q 这两个大质数，进而定出$n=p\times q$。但第三者想要把已知的n （一个两百位以上的自然数）分解成两个质数之积，大约需进行12×10^{23} 次运算，以每次运算所需时间为10^{-6} 秒计算，也需要3.8×10^9 年。可见，RSA密码体系中的p 与q 的保密程度是极高的。不知p 、q 则不知r ，进而不知d ，从而不能把密文M，译成明文M 。

不过，对RSA体系的保密性也得客观评价。1999年，科学家们在互联网上破译了信用卡及其他绝密信息所用的公开密钥机制，特别是学者

们破译了欧洲信用卡传输与保密电子邮件RSA-155的编码机制。此过程需把一个155位的质数分解为两个质数的乘积。

这个庞大的数字是由荷兰首都阿姆斯特丹的一个由赫尔曼·德·里莱领导的集团用300台个人电脑及一台超级计算机成功分解得到的。现在美国通常用232位数进行加密，而美国政府则用309位数作为行政与军事的加密标准。但随着数学的进步与超级计算机的迅猛发展，难道还要用400位或以上的质数密钥吗？

109417386415705274218097073220403576120037329454492059909138421314763499842889347847179972578912673324976257528997818337970765372440171467435315933543338971026395928297411057720541965739916759007165678080380668033419335217907113077791066034883801684548209272203600128786792079585759892815222706082371930628086431793

第十一章 分形之美

《最强大脑》第四季第五期中的一个挑战项目叫“分形之美”。它以分形几何中的经典集合——朱利亚集（简称J集）为研究对象，以此推演图形中的数字逻辑。J集是由一个迭代函数$Z_{n+1}=Z^2_n+c$ 来定义的，其中c 为复数，根据$c=x+y$ i来确定。当c 确定并固定y_0 的值后，可以得到一系列Z 值，也就得到了一些不断变化的图像。节目组公布了Z_0 和Z_{24} 的图像，选手们将在其中间部分的23张，根据嘉宾指定的3张图像来确定相应的数值。

这个项目考察选手的推理能力和计算能力，有一定难度。但广大观众对什么是J集，什么是分形尚缺乏足够的了解，这不免与节目组“让科学流行起来”的初衷有一定差距，未能达到理想效果。

为此，笔者对什么是分形，什么是J集，以及分形几何的核心——分数维作一简单的介绍。

什么是分形

1967年，在美国《科学》杂志上出现了一篇划时代的论文《英国的海岸线有多长？统计自相似性与分数维数》，论文作者曼德布罗特是当代的一位美籍法国数学家和计算机专家。他的答案颇出人意料，他认为无论你做得多么认真细致，都不可能得到准确答案，因为根本不会有准确答案。

关于“海岸线有多长”的问题看似很简单，其中却存在一些问题。测量数值保留的精确度不同，测出的结果会有较大区别。如结果精确到千米，则短于1千米的迂回曲折都被忽略掉了。若精确到米，则能多测出一些迂回曲折，数值将变大。精确度越高，测得的数值愈大，这些愈来愈大的数值将趋于一个确定值，这个极限值就是海岸线的长度。

但是曼德布罗特发现：当精确度不断提高时，所得的值是无限增大的，他认为海岸线的长度是不确定的，或者说在一定意义上海岸线是无限长的。为什么呢？答案就在于海岸线的极不规则和极不光滑。我们知道，经典几何研究规则图形，平面解析几何研究一次和二次曲线，微分几何研究光滑的曲线和曲面。传统上总是将自然界中存在的大量不规则形体近似规则后再进行处理，就像在测量海岸线时，总会先把海岸线折线化（无论是有意还是无意），然后得出一个有意义的长度值。但曼德布罗特突破了这一点。

海岸线长度问题最初是由曼德布罗特在英国数学家理查逊的遗稿中发现的，这个问题引起了他极大的兴趣，开始潜心研究。当初，理查逊为了了解一些国家锯齿形的海岸线长度，查了西班牙、葡萄牙、比利时和荷兰的百科全书，结果发现书上在估计同一国家的海岸线长度时，竟

然有百分之二十的误差。理查逊指出，这种误差是因为他们使用不同精度的量尺所导致的。他同时发现海岸线长度**L** 与量尺精度**s** 的关系——log（1/**s** ）与log（**L** ）呈线性关系。

理查逊得出了一个关于边界长度的经验公式，曼德布罗特敏锐地发现该公式中的一个参数可以用来描述海岸线的特征，他称之为“量规维数”，这就是分数维数之一。自此，曼德布罗特的分形概念开始萌芽生长。

谈到分形，就不得不提非线性的概念。所谓线性是直线的属性，也意味着系统的简单性；而非线性是指变量之间的数学关系，不是直线，而是曲线、曲面或不确定的属性。分形满足非线性的概念，由于其本身的复杂性，至今没有一个确切的科学定义。英国数学家法尔科内认为，应该将分形看作具有以下特征的集合F。

① F具有精细结构，即在任意小的比例尺度内包含整体。

② F是不规则的，以至于不能用传统的几何语言来描述。

③ F通常具有某种自相似性。

④ F在某种方式下定义的“分数维”通常大于F的拓扑维数。

⑤ F的定义通常是非常简单的，或许是递归的。

简单来说，分形的系统具有自相似性和分数维数。自相似性容易辨认，例如一块磁铁中的每一部分都像整体一样具有南北两极，可以不断分割下去，每一部分都具有和整体相同的磁场。在适当放大或缩小几何尺寸时，这种自相似的特性不变。具有自相似性的形态广泛存在于自然界中，如连绵的山川、飘浮的云朵、岩石的断裂口、粒子的布朗运动，及树冠、叶子、花菜、大脑皮层等。分形的概念也早已从最初所指的形态或结构上具有自相似性的几何对象的狭义分形扩展到了功能、信息、时间、空间上等具有自相似性的广义分形。连《格列佛游记》中也有4句诗，精彩地给出了一个“自相似”的例子。

啊！博物学家看到一只跳蚤

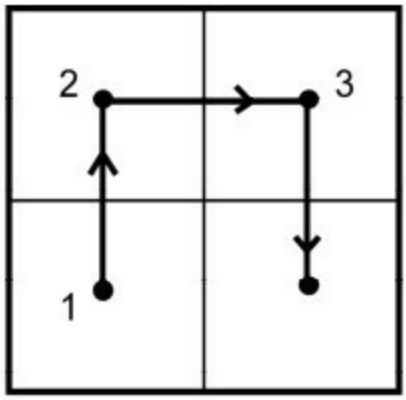
有一只更小的跳蚤在它身上吸血
再又看到更小的跳蚤在小跳蚤身上吸血
如此直至无穷……

分数维数

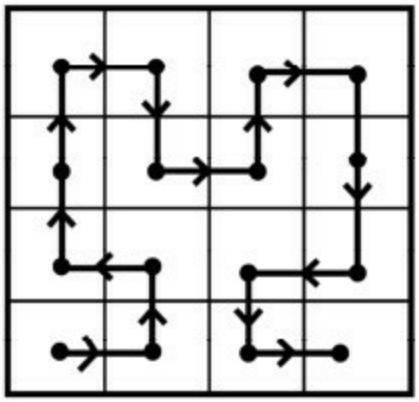
分形结构或分形点集是几维的？在欧氏空间中，欲确定一条直线上点的位置只需一个坐标值，确定平面上一点的位置需要两个坐标值，确定我们生活空间中一点的位置需要3个坐标值。用一组实数（x_1，x_2，…，x_n）来确定一个点的位置所形成的空间，称为n 维空间。可见，在我们的心目中，维数即确定点的位置所需要的坐标数目。这种通常的维数概念受到了分形几何的挑战，必须加以推广。

事情要追溯到1890年，意大利数学家皮亚诺（1858—1932）发明了一种曲线，可以装满一个正方形区域。如果作为点集那么此曲线是一维的，它装满正方形区域后就应该说是二维的。这种自相矛盾的结论是怎么回事呢？我们先来看看皮亚诺是如何构造这种奇怪的曲线的。

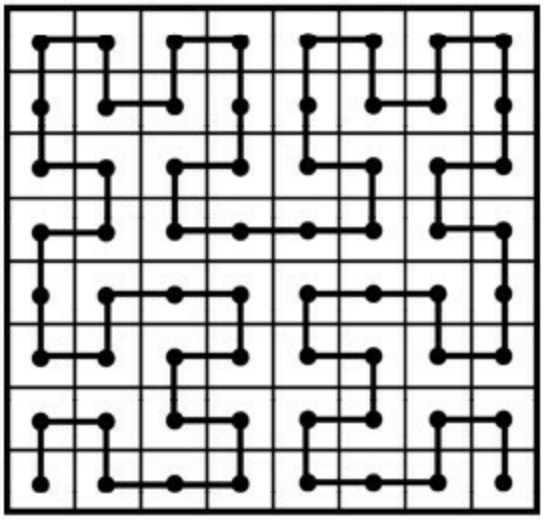
把单位正方形R 划分成4个全等的一级正方形，按顺时针方向依次历经这4个小区间，沿图11-1（a）中折线走遍一级正方形的中心。再把每个一级正方形划分成4个二级正方形，并行遍所有的二级正方形，具体过程如下。先行遍一级1号正方形中的4个二级正方形，再进入一级2号正方形，行遍一级2号正方形内4个二级正方形后，进入一级3号正方形，行遍其4个二级正方形后进入一级4号正方形，且行遍其中的二级正方形［如图11-1（b）所示］。图11-1（c）画的是n =3的情况，令n →+∞，则得一连续曲线L，L装满了单位正方形R 。



(a)



(b)



(c)

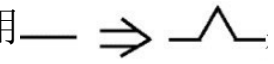
图11-1

1919年，德国著名的数学家豪斯道夫（1868—1942）提出了分数维

数的概念。其实从理论上讲允许各种不同的维数定义存在，但每种维数定义必须满足欧氏空间的传统维数定义。

有一种维数叫“盒子维”，设一个有界点集S 属于m 维欧氏空间，我们把m 维欧氏空间划分成棱长为ε 的m 维小方盒，再数一数含S 中点的小盒子的数目N（ε），则称$D=\lim\limits_{\varepsilon\to 0}\dfrac{\ln N(\varepsilon)}{\ln\frac{1}{\varepsilon}}$ 为S 的盒子维数。

用该定义可以求正方形区域的维数：把单位正方形区域划分成棱长为$\dfrac{1}{n}$ 的n^2 个方盒子，即$\varepsilon=\dfrac{1}{n}$，代入上式，得$D=\lim\limits_{\varepsilon\to 0}\dfrac{\ln n^2}{\ln n}=2$。类似地，可用上述定义求得线段是一维的，立方体是三维的，可见这种维数定义不违反传统维数的结论。用盒子维来计算皮亚诺曲线的维数，得到这条曲线所形成的点集恰为二维的，从而克服了传统观点中皮亚诺曲线与正方形区域之间维数的矛盾。事实上，对于皮亚诺曲线，$\varepsilon=\dfrac{1}{2^n}$，$N(\varepsilon)=4^n\ (n\geqslant 1)$，有 $D=\lim\limits_{\varepsilon\to 0}\dfrac{\ln 4^n}{\ln 2^n}=2$

还有一种所谓的相似形维数，若某图形可由边长缩小为$\dfrac{1}{a}$ 的b 个自相似图形拼成，则定义其维数为$D=\dfrac{\ln b}{\ln a}$。例如，图11-2所示的柯克曲线是由一段单位直线逐次应用  变换而成的。

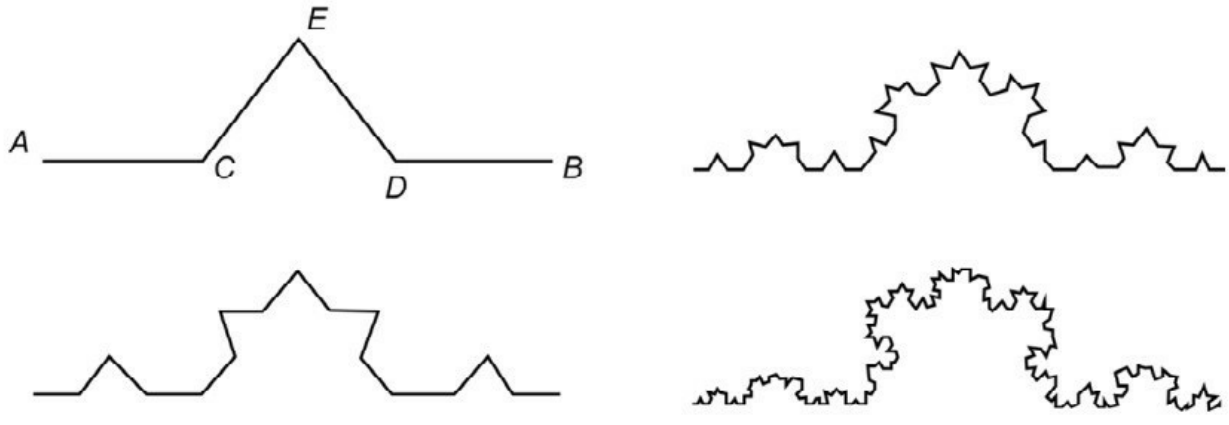


图11-2

柯克曲线的构造过程如下：取长为1的线段AB，AB 的三等分点分别是 C、D，把线段从 C、D 点断开，从此处凸起一个折线 CED，使 $CE=ED=\frac{1}{3}AB$，再以AC、CE、ED、DB 分别扮演AB 的角色，依次类推。在这里，柯克曲线的a =3，b =4，于是$D=\frac{\ln 4}{\ln 3}\approx 1.26$。

这里就出现了非整数维数这样一类新事物，柯克曲线维数的准确值是一个无理数。由此可见，维数可以是非负整数，也可以不是整数，甚至可以是无理数。

下面再介绍一个几何分形图，它是由波兰数学家谢尔宾斯基制作的，如图11-3所示。谢尔宾斯基分形图的构造十分简单，大家见图自明。下面来计算它的维数。从图11-3的左图开始，将三角形的边长每次缩小1/2，然后把3个小图放在一起，如此迭代下去，就得到一个分形的谢尔宾斯基三角形，其维数是$D=\frac{\ln 3}{\ln 2}\approx 1.585$。

图11-3

必须指出，上面介绍的只是最简单的情况，分数维数的计算往往是“一事一议”，没有统一的法则和公式。目前的研究已经算出了许多自然分形和人工分形的分数维数，如下所示。

海岸线的维数：$1<D<1.3$

山地表面维数：$2.1<D<2.9$

河流水系维数：$1.1<D<1.85$

云的维数：D =1.35

金属断裂纹的维数：D =1.27±0.02

人的肺的维数：$D \approx 2.17$

血管直径分布的维数：$D \approx 2.3$

人脑表面的维数：$2.73 < D < 2.79$

人的脑电图的维数：$1.9 < D < 2.4$

J集和M集

本章一开始，我们就说起过J集即朱利亚集，那么朱利亚是何许人也？朱利亚可归于神童一类，出生于阿尔及利亚，8岁时第一次进小学就直接读五年级。后来，朱利亚获得奖学金到巴黎学习数学。在第一次世界大战期间，法国卷入战争，21岁的朱利亚在一次战斗中受了重伤，被炸掉了鼻子。后长期在医院接受治疗，他仍以顽强的毅力研究数学。在医院的几年中他完成了博士论文，25岁那年，他在《纯粹数学和应用数学》杂志上发表了描述函数迭代的长达300页的杰作，因之一举成名。但不幸的是，几年后这个有关迭代函数的工作似乎被人们遗忘了，一直到20世纪七八十年代，由曼德布罗特奠基的分形几何被广泛应用到各个领域后，朱利亚的名字才因此传播开来。

J集研究的是函数$f(z)=z^2+c$（等价于$Z_{n+1}=Z_n^2+c$）迭代的结果。为了说明它的迭代过程，先取c　　为实数，方便大家理解。例如当c = 0.75，并取$z_0=1$，我们得到：

$z_1=f(1.0)=1.0^2-0.75=0.25$

$z_2=f(0.25)=0.25^2-0.75=-0.6875$

$z_3=f(-0.6875)=(-0.6875)^2-0.75=-0.2773$

$z_4=f(-0.2773)=(-0.2773)^2-0.75=-0.6731$

$z_5=f(-0.6731)=(-0.6731)^2-0.75=-0.2970$

…

显然，这些结果都在坐标轴的x 轴上，但当c 取复数$a+ib$ 时，迭代

结果就在一个复平面上。对于有的初始值z_0，z 很快趋向无穷大，有的则不是。设想给复平面上的点着色，凡是趋向无穷大的着彩色，反之，则着黑色，其边界就是J集。也就是说，对于有些z_0，函数值始终约束在某一范围内，满足该情况的所有初始值z_0的集合称为J集。

借助于计算机，可以把它们画出来，每一个c 有它自己的J集。它的形状简直复杂得令人难以置信，有位法国画家说出了内心的惊叹：“它们有的像天上的浮云，有的如多刺的荆棘，有的则像节日之夜，烟火熄灭后仍在空中飘动闪烁的火星……其足以使许多画家掷笔兴叹，顶礼膜拜。”

上面提到的非线性迭代公式$Z_{n+1}=Z_n^2+c$，正是后来被曼德布罗特所研究的。现在，就来说说曼德布罗特其人。本华·曼德布罗特（1924—2010）是一位成衣批发商和牙医的儿子，幼年时喜爱数学，特别是几何。后来，他的研究范围更加广泛，包括棉花价格、股票涨落、语言中的词汇分布等，从物理学、天文学、地理学，到经济学、生理学……都有所涉及。他一直在IBM公司做研究，又曾在哈佛大学教授经济学，在耶鲁大学教授工程，在爱因斯坦医学院教授生理学。也许正是这些风马牛不相及的多个领域的研究经验，使曼德布鞯狒创立了跨学科的分形几何。他被誉为20世纪后半叶少有的影响深远而广泛的科学伟人之一。1993年，身为美国科学院院士的曼德布罗特获得了沃尔夫物理学奖，图11-4即曼德布罗特的肖像。

图11-4

2010年10月14日，曼德布罗特因胰腺癌逝世，享年85岁。他离世之后，法国总统萨科齐称其具有“从不被革命性的、惊世骇俗的猜想所吓退的强大而富有独创性的头脑”。

在用公式$Z_{n+1}=Z_n^2+c$迭代时，开始平面上有两个固定点：c和Z_0。若先取$Z_0=0$，然后就有$Z_1=c$，我们将每次Z的位置用亮点表示，也就是说，开始时平面上的原点是亮点。第一次迭代后亮点移动到c，再后我们可以计算Z_2，它等于c^2+c，亮点移动到Z_2……经过一次次迭代，代表复数Z的亮点在平面上的位置不停地变化，从Z_0开始，$Z_1,Z_2,\ldots,Z$，亮点跳来跳去，很难看出它的跳动有什么规律。但是，我们感兴趣的是当迭代次数K趋于无穷大时，亮点会在哪里？是在有限的范围内转悠呢？还是会跳到无穷远处不见踪影？显然，无限迭代时Z的行为取决于

复数c 的大小。

这样，我们得到了曼德布罗特集的定义：“所有使得无限迭代后的结果保持有限数位的复数C 的集合，构成曼德布罗特集（简称M集）。”前面说过，J集的函数值若约束在某一范围内，则可以称之为是连通的，否则为不连通。故曼德布罗特集严谨的定义是：使J集为连通的参数c 的集合，即

$$M=\{c \in C \mid J(Pc)是连通的\}$$

从而J集和M集成了亲戚。也正因为M集是以Z_0 =0开始迭代的，而J集是以某一不确定的初始值Z_0 开始迭代的。在图像上可以看清这一点。在M集的计算机图像中，任意一点都有对应的J集，如图11-5所示，中间那个大图是被人称为“数学恐龙”的M集图形，而任意一点都有相应的J集图形。

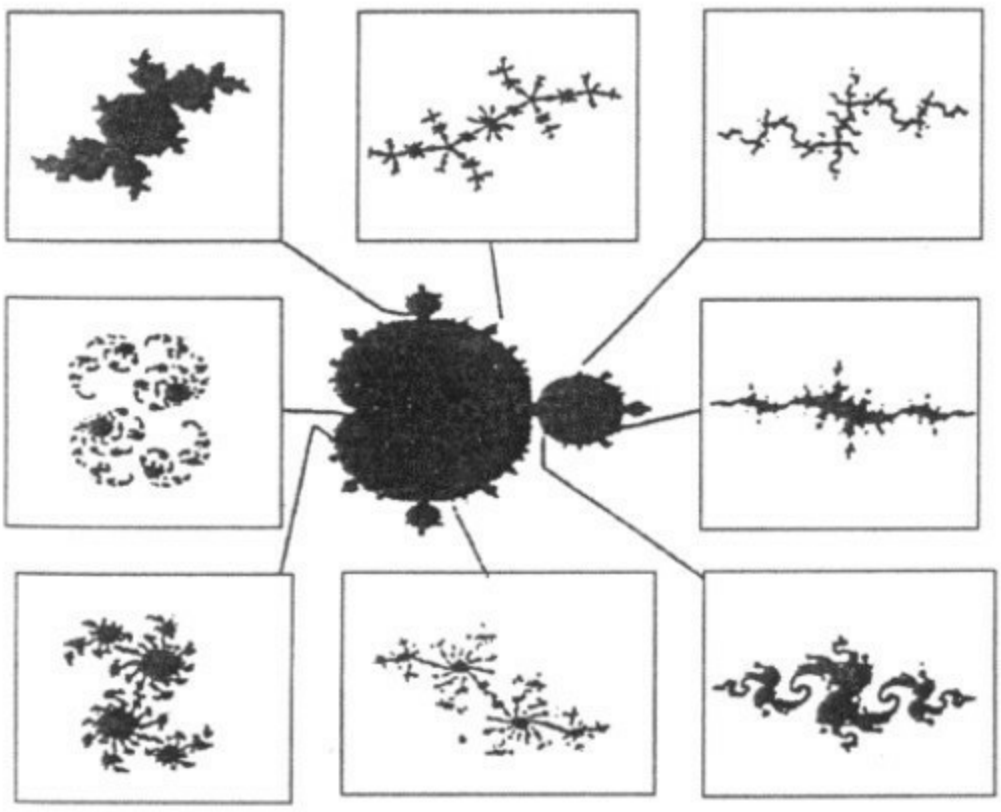


图11-5

最后，引用曼德布罗特在1975年出版的《大自然的分形几何学》一书中精彩的一段话来结束本节。

“云不只是球体，山不只是圆锥，海岸线不是圆形，树皮不是那么光滑，闪电传播的路径更不是直线。它们是什么呢？它们都是简单而又复杂的分形。”

分形数列

上几节中，我们讲的都是有关分形几何的内容。实际上，在分形数学的大花园中，还有一株鲜为人知的小草——分形数列（也称自相似数列）。下面举几个例子，让大家对它有个基本了解。

（1）克拉科斯基数列

克拉科斯基数列是一个仅由1和2组成的无限数列，通过“自描述”进行定义。它在整数数列大全网站上排名第二，足见该数列在组合数学中的重要性。

它的定义很简单，若把数列中相同的数归为一组，令$a(1)=1$，$a(2)=2$，…则$a(n)$等于第n组数的长度。可以根据这个定义来推算第三项以后的数：例如由于$a(2)=2$，因此第2组数的长度是2，所以$a(3)=2$；而$a(3)=2$就表明第三组数的长度为2，即数列接下来要有两个1，因此$a(4)=a(5)=1$，所以第四组数和第五组数的长度都为1，因而$a(6)=2$，$a(7)=1$，以此类推。它的前几项为

1,2,2,1,1,2,1,2,2,1,2,2,1,1,2,1,1,2,2,1,2,1,1,2,1,2,2,1,1,…

如果我们将相同的数字组成一组，那么可以得到 1,2,2 ,1,1 ,2,1,2,2 ,1,2,2 ,1,1 ,2,1,1 ,2,2 ,1,2,1,1 ,2,1,2,2 ,1,1 ,2,1,1 ,2,1,2,2 ,1,2,2 ,1,1 ,2,1,2,2 ,…每一组数的个数为1,2,2,1,1,2,1,2,2,1,2,2,1,1,2,1,1,2,2,1,2,1,1,2,1, 2,2,1,1,2,…也就是说，将数列中相同的数以其个数合并，得到的仍将是数列本身，这就是一个分形数列。

随着n 的增大，数列中1和2的个数是否趋于相等呢？现在还没有证明出来。最近有研究表明，数列中1和2个数比的极限值可能不是1/2。赫瓦塔尔已经证明了数列中1的密度的上界为0.50084。另外，面对如此

简单的数列，目前竟还没有发现可以用解析方式来表达的通项公式。

（2）莫尔斯-瑟厄数列

莫尔斯-瑟厄数列中的每一代都可以由其前代加上它的“补数列”得出。从0开始，第二代补上“1”得到01，第三代看见上一代是01，加上它的“补数列”得到0110，第四代在0110后面加上“补数列”1001，就得到01101001。数列的生长极其迅速，下面是它的第8代。

0110100110010110100101100110100110010110011010010110100110
1010010110011010010110100110010110011010011001011010010110011010
1001011001101001011010011001011001101001100101101001011001101010
1010011001011010010110011010011001011001101001011010011001011010

如果把该数列逐项相间地截取，可以复制出全部数列。与此类似，交替地保留相邻的两项，也能复制出全部数列。有兴趣的读者不妨亲自尝试。

莫尔斯-瑟厄数列有许多出人意料的应用：象棋问题，准晶体理论，合金中的振动态，数学物理，文学组合，微分几何，数论以及连续函数的迭代。E.普洛海特在1851年发表过一篇论文，文末说：“从文献中找寻频频出现的莫尔斯-瑟厄数列，将使你在迷人的数学领域中获得一次愉快的散步。”

（3）黄金数列

这里说的并不是斐波那契数列，而是由下述方法生成的数列。从1、10开始，在后面不断地添加上一行的数字，如第三行加1，成101，第四行在第三行的后面加上第二行的10，成10110，依次类推。

1

10

101

10110

10110101

1011010110110

……

数列中1与0的个数比近似于黄金比例，而且加进去的项数越多，1与0个数的比值越接近 $\phi=\frac{\sqrt{5}+1}{2}$ 的值，这也是称它黄金数列的原因。

在黄金数列中划出任意子数列——例如10 1 10 10 1 10 1 10 …，你会发现前后两个10之间的位数为：2122121…（这里计算位数的规则是后一个10中的“1”与前一个10中的“1”相隔的位数）。如果这里的2用1代替，　1用0代替，则黄金数列又会重现，这就显示了在不同尺度下的自相似性，也表明了它确实是一个分形数列。

（4）签名数列

这是一个十分奇怪的数列，先确定一个无理数，例如 $\sqrt{2}$ 。现在将 i，j 都从正整数1,2，…，n 开始取值，然后计算 $i+j\sqrt{2}$ 的结果并取近似值，从小到大予以排列。

$$1+1\times\sqrt{2}=2.414$$

$$2+1\times\sqrt{2}=3.414$$

$$1+2\times\sqrt{2}=3.828$$

$$3+1\times\sqrt{2}=4.414$$

$$2+2\times\sqrt{2}=4.828$$

$$1+3\times\sqrt{2}=5.243$$

$$4+1\times\sqrt{2}=5.414$$

$$3+2\times\sqrt{2}=5.828$$

$$2+3\times\sqrt{2}=6.243$$

$$5+1\times\sqrt{2}=6.414$$

现在取其 i 值，形成数列 1,2,1,3,2,1,4,3,2,5,1,4,3,6,2,5,1,4,7,3,6,2,5,8,1,4,7,3,6,9,2,5,8,…这就是$\sqrt{2}$ 对应的签名数列。如果删去首次出现的每一个整数，你将看到剩余数列同原来的数列一模一样，所以它真的是分形数列啊！

那么究竟什么是分形数列呢？曼德布罗特在他的名著《大自然的分形几何》中所给出的定义是：“一个无界集合r (s)如果同s 能全等的话，则称它有比值r 的自相似性。”试考虑一个整数数列$x_1, x_2, x_3, x_4, x_5,$.如果$x_2, x_4, x_6,$…与$x_1, x_2, x_3,$…是等同的，这说明该数列具有比值为的自相似性。当然，我们也可以推广：如果存在某个满足关系式$1\le d \le r$的整数d，使得$x_d, x_{(r+d)}, x_{(2r+d)}, x_{(3r+d)}, x_{(4r+d)},$…与

$_4, x_5$,…等同，可将该数列称为关于比值r 的自相似性数列。自然，也可以不拘泥于“比值r ”的限制，一切具有“部分同整体相似”的性质的数列都可以称为分形数列。

从几何到代数，分形世界无处不在。

附录一 同余数的基本概念

设a ,b 均为整数，n 为一正整数。若a -b 可被n 整除，或说a -b 为n 的整数倍（即a 、b 除以m 所得余数相等），则称a ,b 对模n 同余，记作$a \equiv b(\mathrm{mod}\ n)$ 。反之，若a 与b 对模n 不同余，则表示为$a \not\equiv b(\mathrm{mod}\ n)$，例如$32 \equiv 5(\mathrm{mod}\,3)$，$3 \not\equiv -2(\mathrm{mod}4)$ 。

同余的概念在日常生活中经常遇到，如我国传统文化中的十二生肖，在称某人属哪一生肖时，即为以12为模的算法。再如，若今天是星期二，18天后是星期几？将2+18的和除以7得余数6，即知18天后为星期六。

同余关系有下述性质：

①反身性。即对每一整数a 及正整数n，$a \equiv a(\mathrm{mod}\ n)$。

②对称性。即对任意整数a ,b 及正整数n，若$a \equiv b(\mathrm{mod}\ n)$，则$b \equiv a(\mathrm{mod}\ n)$ 。

③递推性。即对任意整数a ,b ,c 及正整数n，若$a \equiv b(\mathrm{mod}\ n)$，且$b \equiv c(\mathrm{mod}\ n)$，则$a \equiv c(\mathrm{mod}\ n)$ 。

④设有任意整数a ,b ,c ,d 及正整数n，若$a \equiv b(\mathrm{mod}\ n)$，且$c \equiv d(\mathrm{mod}\ n)$，则$a + c \equiv b + d(\mathrm{mod}\ n)$，且$ac \equiv bd(\mathrm{mod}\ n)$。

由性质④得，若$a \equiv b(\mathrm{mod}\ n)$，其中$a$ 、b 为整数，n 为正整数，则对每一非负整数k，$ka \equiv kb(\mathrm{mod}\,n)$，且$a^k \equiv b^k(\mathrm{mod}\,n)$ 。但是除法不成立。例如$2 \times 2 \equiv 8 \times 2(\mathrm{mod}\,4)$，且$2 \not\equiv 0(\mathrm{mod}\,4)$，但

$2 \neq 8(\mathrm{mod}\,4)$ 。不过，可证明对任意整数a ,b ,c ,d 及正整数n ，若$ac \equiv bc(\mathrm{mod}\,n)$ ，且（c ，n =1，则$a \equiv b(\mathrm{mod}\,n)$ 。在此对任意整数a 、b ，以（a ，b ）表示其最大公约数，（a ，b ）=1即表示a 、b 互质。

以下举一例，让大家熟悉以上性质并得以应用。求2^{31} 除以11的余数。

解：首先$2^{30}=(2^6)^5=64^5$

而$64 \equiv 9(\mathrm{mod}11)$，故$64^5 \equiv 9^5(\mathrm{mod}11)$

类似可得 $9^5 \equiv 81 \times 81 \times 9 \equiv 4 \times 4 \times 9 \equiv 1(\mathrm{mod}11)$

故知 $2^{30} \equiv 1(\mathrm{mod}11)$

又$2 \equiv 2(\mathrm{mod}11)$，故$2^{31} \equiv 2(\mathrm{mod}11)$

附录二 答案

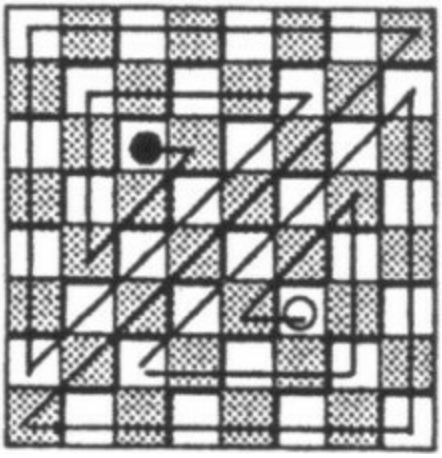


图2-7的答案

图3-10中迷宫的答案

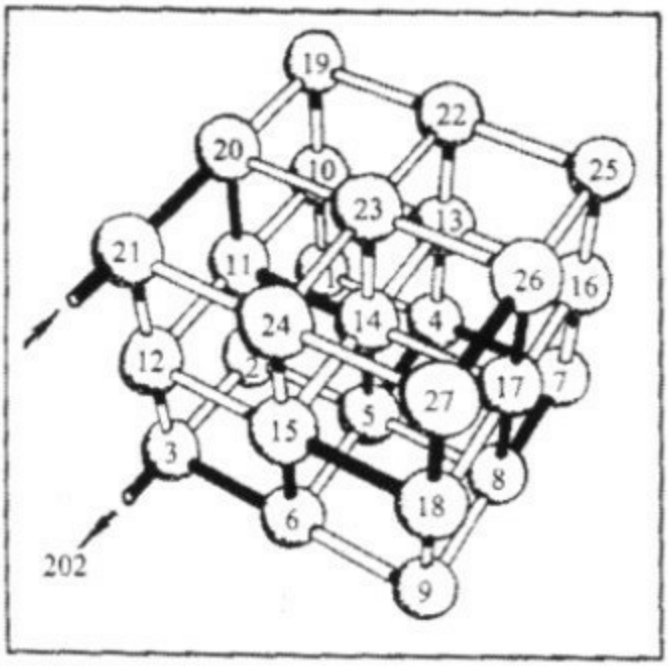


图3-11中迷宫的答案

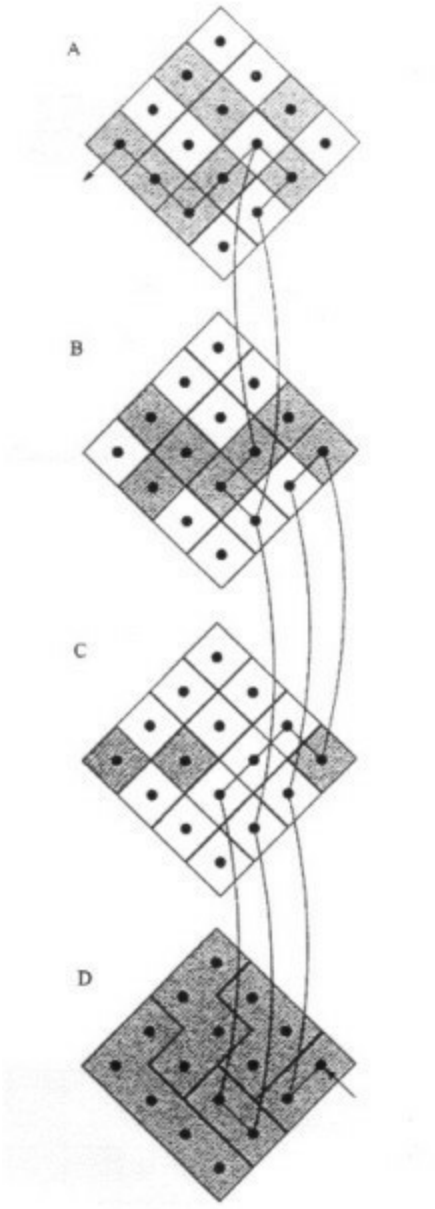


图3-12中迷宫的答案

参考文献

[1] 孙泽瀛.数学方法趣引[M].上海：上海科学技术出版社，1959.

[2] 梁宗巨.一万个世界之谜（数学分册）[M].武汉：湖北少年儿童出版社，1995.

[3] 吴鹤龄，毛晚堆.迷宫趣活[M].北京：人民邮电出版社，2007.

[4] 伊莱·马奥尔.三角之美：边边角角的趣事（第2版）[M].曹雪林，边虹挪，译.北京：人民邮电出版社，2018.

[5] 吴鹤龄.幻立方及其他——娱乐数学经典各题[M].北京：科学出版社，2004.

[6] 李大潜.十万个为什么(数学) [M].北京：少年儿童出版社，2013.

[7] 墫单 .十个有趣的数学问题[M].上海：华东师范大学出版社，2013.

[8] 程钊.欧拉关于七桥问题的解——从数学史与数学教育的角度看[J].数学传播（台湾）, 2012, 36（4）: 42-74.

[9] MCGUIRE G, TUGEMANN B, CIVARIO G.There is no 16-clue Sudoku: solving the Sudoku minimum number of clues problem via hitting set enumeration [J].Experimental Mathematics, 2014, 23(2): 190-217.