前排从左至右：林亚南，刘建亚，张智民，付晓青；后排从左至右：张英伯，蔡天新，邓明立，顾沛，罗懋康，汤涛，励建书，游志平（刘青阳 摄）

《数学文化》旨在发表高质量的传播数学文化的文章；主要面向广大的数学爱好者

《数学文化》欢迎投稿，来稿请寄：
Math.Cult@gmail.com

本刊网站：http://www.global-sci.org/mc/
本刊淘宝网：http://mysanco.taobao.com/
本期出版时间：2013年5月

本刊鸣谢国家自然科学基金数学天元基金的支持

Contents | 目录

数学人物

张益唐和北大数学78级　汤　涛　3

数学趣谈

有错必究
——汉明码（Hamming Code）的原理及其应用　万精油　21

善科网 —— 数学趣题专栏　24

枪打出头鸟 ——三人决斗问题趣谈　万精油　26

数学烟云

通过计算实现正义　易延友　29

谷歌数学涂鸦赏析（中）　欧阳顺湘　34

正态分布的前世今生（下）　靳志辉　54

数学教育

给年轻数学家的忠告 Atiyah, Bollobás, Connes, Mcduff, Sarnak　63

高等教育的危机　Nicholas Carr　79

数学经纬

解释离婚情感动态的数学模型　Jose-Manuel Rey　85

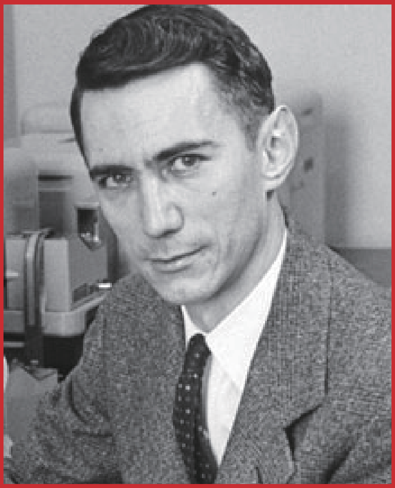
微博上的数学漫游（五）　歌之忆　93

数学家随笔

数学并非你想象的那么战无不胜　曹广福　103

数学人书评

卢昌海《黎曼猜想漫谈》书评　扶　磊　105



张益唐和北大数学78级

谨以此文纪念陈景润诞辰80周年；他影响了那个难忘的时代

汤 涛

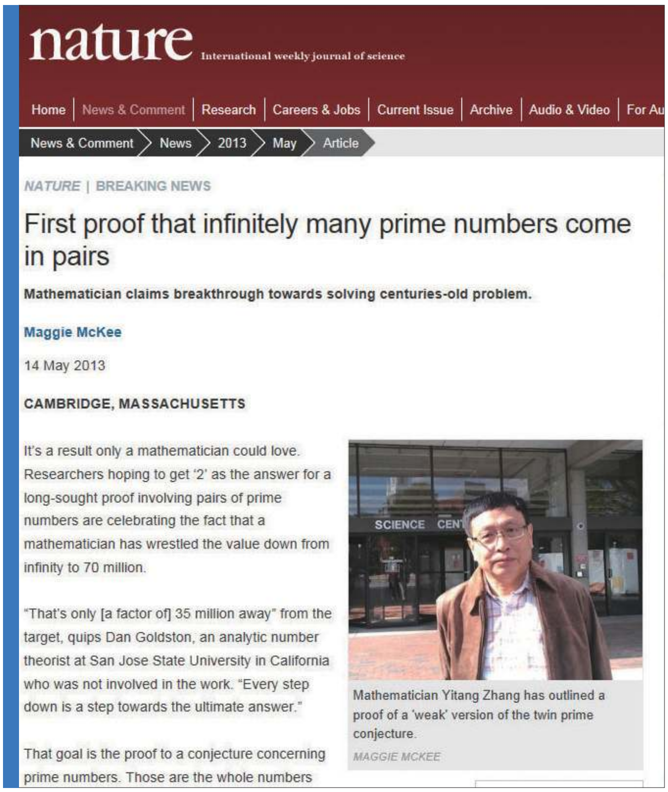
2013年是北京大学数学系成立100周年。百年大庆，作为校友，总觉得应该写点什么作为纪念。

1904年，清政府颁布的《钦定学堂章程》规定"高等算学"隶属格致科（现在称理科），并且规定了算学门的课程。辛亥革命后，京师大学堂于1912年5月1日改名为国立北京大学。同年公布的"民国元年所订之大学学制及其学科"中格致科改名为理科，其中包括数学门。1913年秋，北京大学数学门招收新生，标志着中国现代第一个大学数学系正式开始教学活动。北大数学的早期学生张国焘因为后期的政治生涯为数学系增添了些许另类传奇。

北大数学的一百年培养了大量的人才，先后培养出6000多名本科生，1000多名硕士、博士毕业生，一大批优秀的数学家和其他方面的专家。他们分布在各行业，许多人成绩斐然，得到了社会各界的高度评价。其中有30余位毕业生被选为中国科学院院士。著名数学家江泽涵、许宝騄、段学复、程民德等都是数学系重量级的前辈，获得国家最高科技奖的吴文俊院士、王选院士也任教或毕业于北大数学系。特别值得一提的是，由哈佛大学和麻省理工学院联合举办的2013年度"数学发展前沿"（Current Developments in Mathematics）研讨会上，六位主旨发言的数学家中北大数学毕业生就占了三位：2000级的恽之玮和张伟、1978级的张益唐。另外三位包括菲尔兹奖得主、著名数学家爱德华•威滕（Edward Witten）。

北大数学的辉煌历史，不是我这样一个晚辈敢写的，也不可能写得好。但是为了表达对院庆的祝贺，我就写一个跟我稍稍沾点边的文革后北大数学第一届学生的故事吧。

张益唐与孪生素数



nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Au

News & Comment > News > 2013 > May > Article

NATURE | BREAKING NEWS

First proof that infinitely many prime numbers come in pairs

Mathematician claims breakthrough towards solving centuries-old problem.

Maggie McKee

14 May 2013

CAMBRIDGE, MASSACHUSETTS

It's a result only a mathematician could love. Researchers hoping to get '2' as the answer for a long-sought proof involving pairs of prime numbers are celebrating the fact that a mathematician has wrestled the value down from infinity to 70 million.

"That's only [a factor of] 35 million away" from the target, quips Dan Goldston, an analytic number theorist at San Jose State University in California who was not involved in the work. "Every step down is a step towards the ultimate answer."

That goal is the proof to a conjecture concerning prime numbers. Those are the whole numbers

Mathematician Yitang Zhang has outlined a proof of a 'weak' version of the twin prime conjecture.

MAGGIE MCKEE

《自然》5 月 14 日的报道

北大数学系张益唐沉默 20 多年，几天前突然横空出世，为世纪难题孪生素数猜想的解决做出了突破性的工作。为此，顶级科学杂志《自然》在“突破性新闻”栏目里，报道了张益唐证明了存在无穷多个差值小于 7 千万的素数对，从而在解决孪生素数猜想这一百年数论难题的道路上前进了一大步。

为了认识孪生素数猜想，让我们先做一些简单铺垫。先谈谈素数。素数是数学中美妙的音乐，美丽的女神，有着很多让人捉摸不透的秘密。传说大数学家欧拉说过：“一直以来，数学家总是在孜孜不倦地寻找素数规律，但是很难成功。我们可以把素数看作人类思维无法渗透的奥秘。”

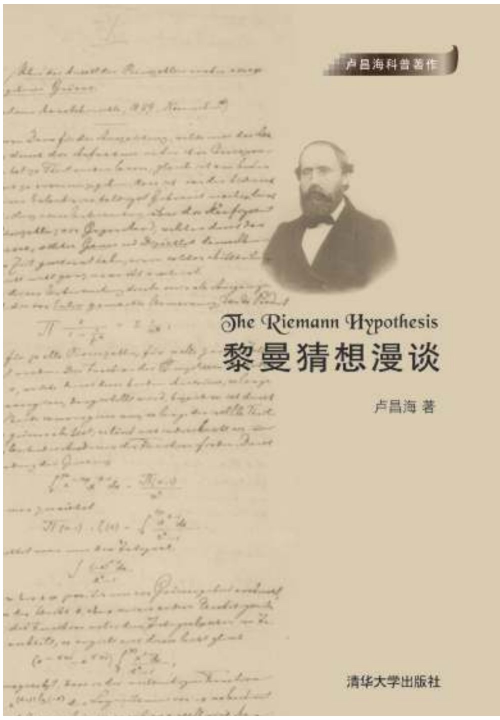
远在中古时代，就产生了自然数的概念，印度人对数学最大的贡献之一就是引进了符号“1, 2, 3, 4, 5, 6, 7, 8, 9, 0”来计数。正是由于有了数的合理记法，对于数的研究才能够代代相传。比零大的正整数统称为自然数，它们共有三类：即 1（民间常说的大佬，无人敢与其争锋）；素数（素数也叫质数，只能被 1 和自身整除的数，比如 5, 7, 11, 19 等）；合数（可以被 1 和自身以外的某个自然数整除的数，如 9, 16, 20）。

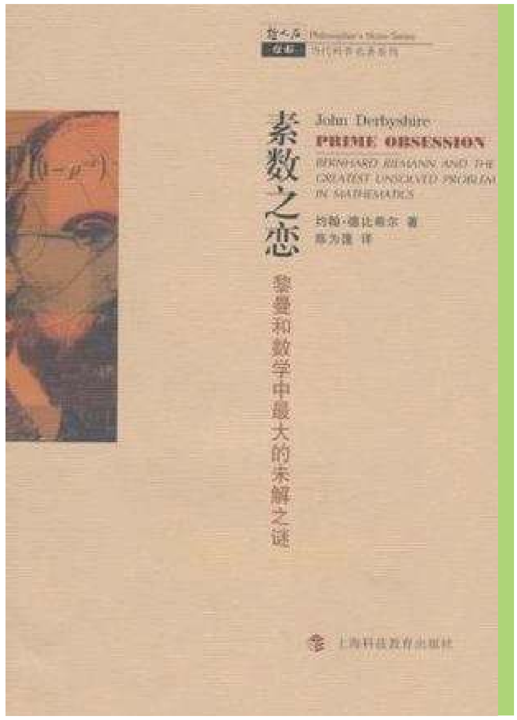
有了自然数的概念以后，很快就有了一个古老的数学分支：数论。它是纯粹数学的分支之一，主要研究自然数的性质。数论在很长一段时间里被称为算术，直到 20 世纪初才开始使用数论的名称。数论的早期铺垫有三大内容：欧几里得证明素数有无穷多个；寻找素数的过筛方法以及欧几里得求最大公约数的辗转相除法；公元 420 至 589 年（中国南北朝时期）的孙子定理。

欧几里得用漂亮的反证法证明了素数的个数有无穷多个。记得我在 80 年代初选修丁石孙的初等数论课时，被他完美的讲课风格和欧几里得伟大的证明所折服，在当时的老二教的上课时光现在还历历在目。

素数在自然数中的分布很奇妙；从公元前三世纪开始至今，吸引了众多数学家的不懈努力。公元前三世纪古希腊数学家、哲学家埃拉托色尼（Eratosthenes, 公元前 275 - 前 193）为了研究这个问题，提出了一个叫“过筛”的方法（the Sieve of Eratosthenes）简称埃氏筛法，造出了世界上第一张素数表，就是按照素数大小排成的表。比方说把一张超大的纸放在沙滩上，然后把自然数按其大小一个一个写上去；然后按下列法则把合数挖掉：

（1）先把 1 删除（因为 1 不是质数）
（2）把 2 留下（最小的偶数质数），然后把 2 的倍数删去
（3）把 3 留下，然后把 3 的倍数删去
（4）把 5 留下，然后把 5 的倍数删去
（5）同理继续进行下去，直到把所有数要么留下，要么删除





这样如果纸上最大的数是 N，则上述方法可以产生 N 以内素数的分布表。

从这个古典的方法中人们可以观察到，素数的分布随着 N 的变大，变得越来越稀疏。比如 1 到 10 之间有 2, 3, 5, 7 四个素数；100 之内有 25 个素数，1000 之内有 168 个素数，100 万之内有 78498 个素数。大量数值试验显示，当 N 变得很大时，在 1 到 N 之间素数的个数和 N 的比值变得很小。那么严格的数学刻画是什么呢？

用 $\pi(N)$ 表示不大于自然数 N 的素数的个数，如 $\pi(2)=1$，$\pi(3)=2$，$\pi(10)=4$。法国大数学家勒让德（Legendre, 1752-1833）于 1808 年建议当 N 非常大时：

$$\pi(N) \sim \frac{N}{\ln N + B},$$

其中 $B=-1.08366$ 被称为勒让德常数。可惜这个公式在相应的级数展开式中仅第一项正确。1792 年，当数学王子高斯刚满 15 岁时，就猜测当 N 非常大时，$\pi(N)$ 和 $N/\ln(N)$ 差不多大；更确切地说，当 N 充分大时，$\pi(N)$ 和 $N/\ln(N)$ 之比接近于 1。用极限的语言来说就是

$$\lim_{N\to\infty} \frac{\pi(N)}{\frac{N}{\ln(N)}} = 1.$$

这个猜想被叫做素数定理。1850 年，俄罗斯数学先驱切比雪夫证明：存在两个正数 a 和 b，使不等式

$$a \le \frac{\pi(N)}{\frac{N}{\ln(N)}} \le b$$

成立，其中 $N \ge 2$。这为证明高斯的素数定理迈进了一大步；并且切比雪夫在证明中用到了微积分。

革命性的变化发生在 1859 年。1859 年 8 月，时年 32 岁的数学家黎曼 (G. F. B. Riemann) 向柏林科学院提交了一篇 8 页纸的论文，题为“论不超过一个给定值的素数的个数”。在这篇论文中，他把素数个数和所谓的 ζ 函数建立了联系，这一联系推动了解析数论的发展；文章中提出的黎曼猜想给数学家们带来了比素数分布更大的挑战。时至

今日，在经历了150多年的认真研究和极力探索后，这个猜想仍然悬而未决。关于黎曼猜想的最权威的科普文章，可以见科普高手卢昌海的《黎曼猜想漫谈》(见本期的书评)。

卢昌海的大作从2010年底在《数学文化》上连载。科学院的王元院士看后非常欣赏。有幸和元老在海淀区知春路上的一个西餐馆吃过一顿饭；其间他对卢昌海的文笔以及卢对黎曼猜想的深刻理解赞不绝口："一个学物理的能把一个这么艰深的数学问题写得如此透彻，真是太不容易了。"之后卢昌海要把文章集结成书时，叫我代请元老写序。一周后，我收到老院士亲笔写的10页纸的手稿。

美国作家约翰•德比希尔（John Derbyshire）根据黎曼猜想的提出和可能的应用，写出了畅销书《素数之恋——伯恩哈德•黎曼和数学中最大的未解之谜》。在这本《素数之恋》中，作者用明晰的笔法，对一个史诗般的数学之谜作了迷人而流畅的叙述，再次展示了素数的魅力。

虽然黎曼没有给出关于 $\pi(N)$ 的具体结果，但他为在黑暗里奋斗的素数分布研究指明了方向。正是沿着这个方向，1896年，法国数学家阿达玛（J. S. Hadamard）和比利时数学家普桑（Charles Jean de la Vallée Poussin）几乎同时独立地证明了素数定理。差不多半个多世纪后的1949年，塞尔伯格（Atle Selberg）和爱多士（Paul Erdös）给出了素数定理的初等证明。前者因此工作以及对筛法的改进获得了1950年的菲尔兹奖。

现在回到孪生素数。孪生素数指差为2的素数对。前几对孪生素数分别是（3, 5），（5, 7），（11, 13），（17, 19），（29, 31），（41, 43），（59, 61）等。一般来说，如果 p 和 $p+2$ 都是素数，则（$p, p+2$）就叫做孪生素数。100以内有8对孪生素数；501到600间只有（521, 523）和（569, 571）两对。更大的孪生素数还有，如（5971847, 5971849）。不过，可以观察到孪生素数的分布也是极不均匀的，并且也是越来越稀疏，与素数的分布相比，还要稀疏得多。

这样问题就来了：比如孪生素数的分布规律是什么？共有多少对孪生素数？或者说有没有一对最大的孪生素数？

于是人们又开始猜想了：有无数对孪生素数。但没有人确切地知道究竟有多少对。到2009年8月6日，已知最大的孪生素数为 $2003663613\cdot2^{195000}\pm1$，这两个数都有10万多位。

上世纪最伟大的数学家大卫•希尔伯特（David Hilbert）在1900年国际数学家大会上提出了著名的23个重要数学难题和猜想，其中孪生素数问题是希尔伯特问题的第8个的一部分，可以这样描述：存在无穷多个素数 p，使得 p 与 $p+2$ 同为素数；而素数对（$p, p+2$）称为孪生素数。数学家们相信这个猜想是成立的。1849年，法国数学家波利尼亚克（Alphonse de Polignac, 1817-1890）提出了更一般的猜想：对所有自然数 k，存在无穷多个素数对（$p, p+2k$）；其中 $k=1$ 的情况就是孪生素数猜想。

2013年4月，张益唐向《数学年刊》（*Annals of Mathematics*）杂志提交了题为"素数间的有界距离"（*Bounded gaps between primes*）的文章。《数学年刊》是研

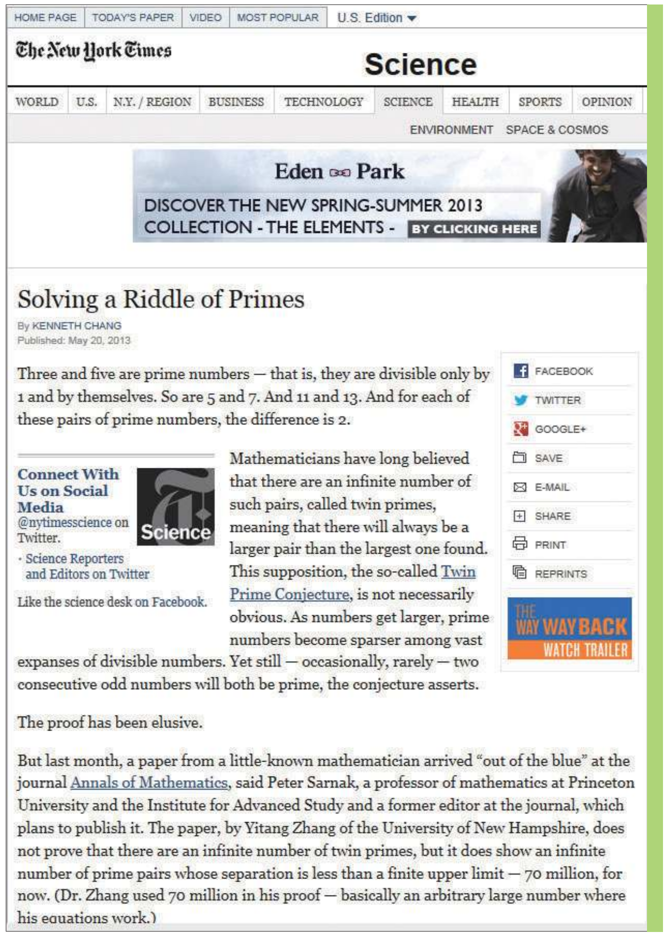
究纯粹数学的学者们最敬仰的期刊，很多人一辈子能在上面发表几篇文章就可以奠定崇高的学术地位。要在这个刊物发表文章必须解决很难的问题，文章一般都很长，还要接受审稿人苛刻的甚至漫长的审稿过程。比如翻开此刊 2013 年的 177 卷，从提交文章到文章被接受，第一期的第一篇文章花了四年半，第二期的第一篇花了两年半，第三期的第一篇花了五年半。根据美国数学会 2012 年 11 月公布的统计资料，2011 年在《数学年刊》所发表的文章，从投稿到接受的平均时间为 24 个月。而张益唐的文章，2013 年 4 月 17 日交稿，5 月 21 日接受，这可能也是这一顶级期刊的一个纪录了。

另外，比较罕见的是这篇文章的审稿人也自报身份：此文的证明被著名的数论专家伊万尼克（Henryk Iwaniec）严格检查。波兰裔的美国数学家伊万尼克被公认为当今最顶级的解析数论专家。他对张益唐的工作赞不绝口：“结果太美了。”（His result is beautiful.）

张益唐的文章到底做了什么？他给出了和上面介绍的波利尼亚克猜想紧密相关的一个命题的证明。他证明了存在无数个素数对 (p, q)，其中每一对中的两个素数之差，即 p 和 q 的距离，不超过七千万。由此推出，存在无穷多个素数对 (p, q)，以及一个不超过七千万的正偶数 h，使得 $p - q = h$。

张益唐的文章基于加州圣荷西州立大学（San Jose State University ）的戈德斯通（Daniel Goldston）小组于 2005 年发表的文章。一般来说，随着数的增大，素数间隙也越来越大；也就是前面说到的越来越稀疏。但戈德斯通的研究小组证明了，即使在很大的数中，仍然存在紧邻的素数。要直接把戈德斯通的方法应用于孪生素数问题却有很本质的困难。这个困难被张益唐巧妙地克服了，他说他是去年夏天的 7 月 3 日在科罗拉多州朋友家的后院里聚会时突然开窍的。

张益唐 2013 年 5 月 13 日在哈佛展示了研究成果。他的证明看起来运用了一些常用的数学技巧，以至于有些人质疑他是否真的正确。但是《数学年刊》审稿人高度评价说：“这项研究是第一流的，作者成功证明了一个关于素数分布的里程碑式的定理。”（The



HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | U.S. Edition

The New York Times

Science

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

ENVIRONMENT | SPACE & COSMOS

Solving a Riddle of Primes

By KENNETH CHANG
Published: May 20, 2013

Three and five are prime numbers — that is, they are divisible only by 1 and by themselves. So are 5 and 7. And 11 and 13. And for each of these pairs of prime numbers, the difference is 2.

Mathematicians have long believed that there are an infinite number of such pairs, called twin primes, meaning that there will always be a larger pair than the largest one found. This supposition, the so-called Twin Prime Conjecture, is not necessarily obvious. As numbers get larger, prime numbers become sparser among vast expanses of divisible numbers. Yet still — occasionally, rarely — two consecutive odd numbers will both be prime, the conjecture asserts.

The proof has been elusive.

But last month, a paper from a little-known mathematician arrived "out of the blue" at the journal Annals of Mathematics, said Peter Sarnak, a professor of mathematics at Princeton University and the Institute for Advanced Study and a former editor at the journal, which plans to publish it. The paper, by Yitang Zhang of the University of New Hampshire, does not prove that there are an infinite number of twin primes, but it does show an infinite number of prime pairs whose separation is less than a finite upper limit — 70 million, for now. (Dr. Zhang used 70 million in his proof — basically an arbitrary large number where his equations work.)

《纽约时报》5 月 20 日的报道

main results are of the first rank；the author has succeeded to prove a landmark theorem in the distribution of prime numbers.）

戈德斯通说发现一个有限大的差距已经是巨大突破了：“我还以为我有生之年看不到这个结果呢。”

刚刚卸任《数学年刊》主编职务的普林斯顿大学教授彼得•萨纳克（Peter Sarnak）对《纽约时报》说，张益唐的观察很深邃（It's a deep insight）、结论很深刻（It's a deep result）。

英国《卫报》文章的正标题是：一个不知名教授的素数大突破，副标题是：鲜为人知的教授在折磨了数世纪数学精英的大问题上迈进了一大步。

数学界又掀起了一股素数热潮。《数学


The Telegraph

HOME | NEWS | WORLD | SPORT | FINANCE | COMMENT | BLOGS | CULTURE | TRAVEL | LIFE | FA

Politics | Obits | Education | Earth | Science | Defence | Health | Scotland | Royal | Celebrities

Science News | Space | Night Sky | Roger Highfield | Dinosaurs | Evolution | Steve Jones | Scien

HOME » SCIENCE » SCIENCE NEWS

Prime number breakthrough by unknown professor

A virtually unknown professor has taken a major step towards solving a numerical problem which has baffled the finest minds in mathematics for centuries.

Dr Yitang Zhang is a researcher at the University of New Hampshire Photo: UNH

By Nick Collins, Science Correspondent

12:02PM BST 22 May 2013

Follow 1,960 followers

168 Comments

Dr Yitang Zhang, who once resorted to working at Subway when he could not find an academic job, has received glowing reviews for his "astounding" new paper on number theory.

Print this article

Share 673

Facebook 565

Twitter 104

Email

LinkedIn 4

英国《卫报》5 月 22 日的报道

《质数的孤独》

文化》特约作者欧阳顺湘在网上留言："听说过小说《素数之恋》、《质数的孤独》什么的，不过我从没看过。不明白作者想用素数谈恋爱还是想说研究者喜欢质数？不过，人们发现素数已经变得更加不孤独了！"

质数是素数的另一个称呼。上面提到的《质数的孤独》是一部关于童年经验、爱与孤独的小说。小说的男女主人公就像两个孪生质数，彼此相近却永远无法靠近。《质数的孤独》出版以来，在意大利掀起了迷恋狂潮，书迷们甚至在城市各角落涂鸦书中佳句。该书表现人性的孤独极其有力，对于揭露造成这种孤独的原因也非常深刻。在销售方面，《质数的孤独》创造了意大利出版史上的处女作销售记录，小说在意大利销售超过 100 万册，在全欧洲更是售出 500 万册，在全球售出 36 个国家的版权。"恐怖"的是全书共有 1907 页——页码只能是质数——造成了这本 300 页左右的书有一个超大的总页数。

北大数学 78 级

人物之一：王鲁燕

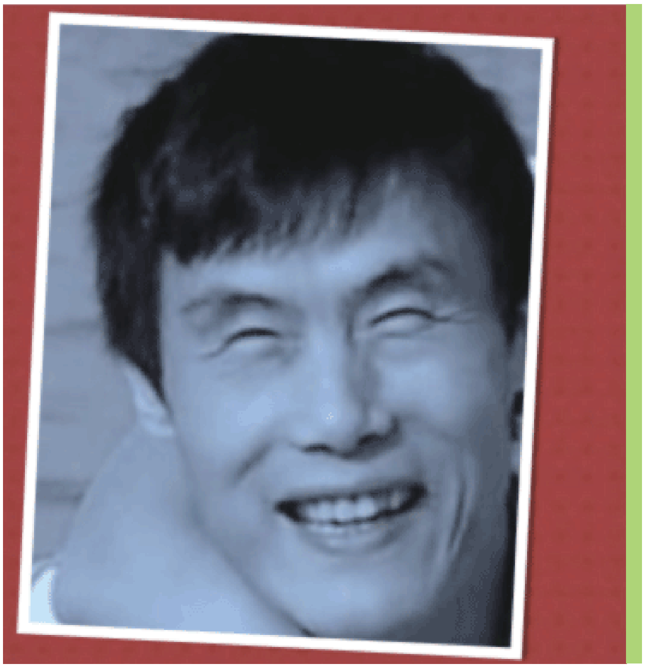
我是 1980 年秋天入读北大数学系。恢复高考后，北大 1977 年没有招生 (这一点很可惜，十年积攒的人才让复旦、中国科大等学校占了很大便宜)。为什么没有招生呢？按当时的副系主任丁石孙在他 2007 年的《自述年谱》所述："(文革后的) 1977 年北大数学系还比较混乱，没有教材，由哪些老师上课也没有安排好，所以七七年就没有招生。"

因此当时 1978 级就是北大数学文革后恢复高考的第一批学生。他们那一届是典型的多年龄段学生混合体，班上最大学生应该是三十三、四岁，最小的十六、七岁吧，还有穿军装的，这些都是现在的大学生难以想象的。我们的班主任刘森老师就是 1978 级的学生，当时 30 几岁吧，是大学三年级的学生。由于他比较成熟严肃，干事情非常认真负责，我们这些比他晚两届的学弟很多见到他都非常害怕。2004 年我们大学毕业 20 年后重聚未名湖畔，很多人都小有成就了，见到刘老师还是毕恭毕敬，“心有余悸”。

78 级不只有成熟的学长，还有些学习上的高手，像考试成绩超好的周青、数学竞赛优胜者严勇、王鲁燕、陈刚。1978 年，在“科学的春天”到来之时，沉寂了十余年的数学竞赛得以恢复，各地的数学竞赛相继开始。数学竞赛优胜者是 1978 级北大数学的一个亮点。这年竞赛里有个小孩，通县一中的王鲁燕，他当时只是一个初二的学生，却在高中水平的数学竞赛中获得了全国第十五名的成绩；并且作为数学竞赛优胜者进了北大。1978 年 6 月 20 日的《光明日报》有下面的励志型的报道，今天读起来颇有“亲切感”：

数学竞赛第十五名王鲁燕是北京市通县一中初二学生，今年只有十四岁。在一年多以前，他是一个上课不听讲，下课爱打架的学生。调皮起哄，全校有名。粉碎“四人帮”之后，整个社会的风气变了，学校的风气变了，王鲁燕的“学不学都上学，会不会都毕业”的思想从根本上动摇了。恰在这个时候，王鲁燕转学到通县一中，模范教师刘纯朴是他的班主任。刘老师因势利导地对王鲁燕进行政治思想教育，又根据他爱动脑筋的长处，加强对他进行智育教育。仅仅一年多的时间，不仅把他培养成为“三好”学生，而且学完了高中数学课程，在数学竞赛中，取得了好成绩，引起了人们的极大注意。

王鲁燕在北大 78 级确是一大名人。第一岁数小，第二长相可爱，细眯的眼睛绝对盖过孙红雷；第三严重偏科。数学尤其是纯数学的课学得极好，据说抽象代数课很多人都在及格的边缘徘徊，他可以得 100 分。那



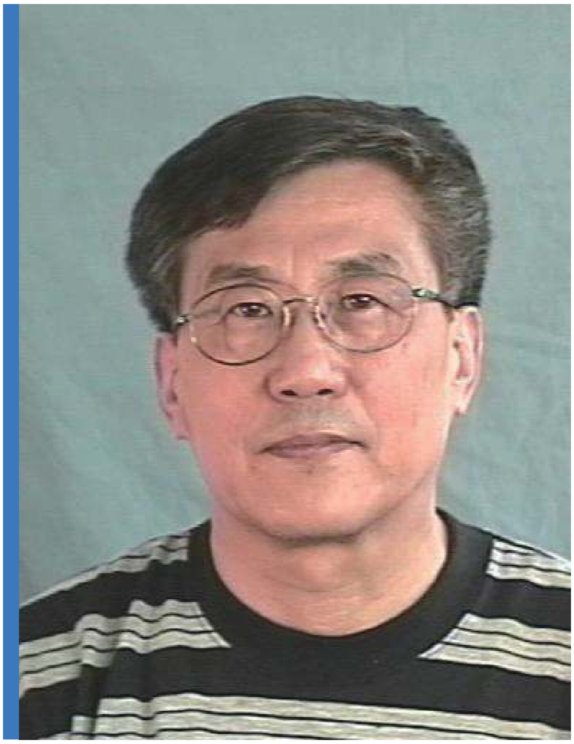
王鲁燕

时给 78 级开课的老师可都是北大老师中的人精，著名数学家丁石孙、聂灵沼、张恭庆、姜伯驹、姜礼尚等都给 78 级上基础课。要在这些人手上拿高分可不是开玩笑的。但是王鲁燕的软肋是英语。一则逸事是他和另一个同学补考英语，他的成绩是 4 分，另一位是 15 分。老师根据成绩誊分，以为是 5 分制的 4 分，本来想给王鲁燕一个良；看到下面这个 15 分才觉得搞错了，两人都不及格。

由于王鲁燕英语长期过不了关，考研究生时考两次都过不了英语最低线，所以一直没有上得了北大研究生。但时任系主任的丁石孙非常喜欢王鲁燕（还有后面要谈的张益唐）的数学才气，破格把他留校当老师了！！

1989 年我在英国博士毕业，第二年到加拿大温哥华当助理教授。两年后北大徐明耀教授到我们系访问。当时我是单身一人，经常周末给不会做饭的徐老师做饭吃，虽然我的厨艺非常一般，但我们饭后的聊天可以忘掉饭菜水平的不足。徐老师很能聊，给我印象深刻。

徐老师也是丁石孙喜欢的人，研究代数的。文革前的大学生、1960 年代北大数



北大徐明耀教授

学系应该算是比较左的，据他告诉我，他们毕业前的最后半年赶上北大的“四清”，他们被进行了半年的毕业教育。近半年的时间，什么也不干，就是政治学习和批判。每个人要彻底地清理思想，人人过关。这次过关的负产品是有两个学生被戴上了“反动学生”的帽子，其中一个就是徐明耀。后来，他就被分配到唐山的一个中学，戴着“帽子”被教育了十年，身体受到极大损害。他文革后又回到北大成为了第一批硕士研究生之一；之后成为北大教授。他给我的印象之一是非常关心政治，他可以大段背诵马克思或列宁的书。这使我认识到他文革倒霉是有自身原因的。第二是对代数充满热爱，我印象很深的是他那时写了本代数书，由于印刷失误，书上多印了一个积分符号。对微积分颇不以为然的他非常生气，“我这么纯洁的书里面居然有积分这个豆芽菜？”想必这本高书就是著名的赵春来、徐明耀《抽象代数 I》或《抽象代数 II》吧。

徐明耀老师的主要研究领域是有限群论、代数数论、计算群论等。他当时到加拿大访问的是我们系的教授阿什巴赫（Brian Alspach），作图论的，也用一些群论；所以徐老师的群论专长也可以发挥得很好。徐老师对王鲁燕一直很欣赏，那个时候王鲁燕在国内读研究生的希望已经不大，所以他就想把王鲁燕推荐到阿什巴赫那里。我们这位阿什巴赫教授有一个爱好，就是每个星期要到赌场去打牌；用数学知识实战赌场。他听徐明耀把王鲁燕的数学才能、智商描绘得很高，很感兴趣，立刻问：“他会打牌吗？”

要知道王鲁燕虽然英语考不及格，桥牌英语绝对一百分。他在大学期间和北大 78 级的袁勇搭档，另一对搭档张林波和沈捷，是北大很有名气的桥牌队。他大学毕业留校后，在数学竞赛界和桥牌界都享有盛名。徐明耀当然毫不犹豫地告诉阿什巴赫王鲁燕的牌技是一流的，说得阿什巴赫眉飞色舞，当时就说只要王鲁燕英语能过最低的托福线，就可以录取他。

一个月后我正好要访问北大；因为我当时和北大的滕振寰老师合作一项研究。徐明耀就叫我去找王鲁燕。见到他后，王鲁燕说他真不想出国，但现在国内没有研究生学位可能不好办，最主要的是结婚后没房子，晚上只能在办公室里拉个帘子，老婆很有意见。所以他对阿什巴赫的好意很有兴趣。但他对最低的英语成绩也毫无把握，所以最终他也没有去加拿大。

后来我听说他还是考过了托福，居然分数还不错。他去了普渡（Purdue）大学，这里也是张益唐读博的地方，只是那时张益唐已经毕业离开了。王鲁燕在普渡的时光，花了不少时间在网络上打桥牌。普渡那时的朋友还记得他的一个桥牌轶事。那时几乎每天中午，有些教授午饭时间在教师俱乐部打桥牌；王鲁燕常坐在某位牌手旁边看牌。经常叫完牌开打前，他就可以猜出其他三位手中的牌，并且是八九不离十。这点让很多老师都佩服得五体投地。

可能用在打牌上的时间太多，他在普渡的学习受到影响，资格考试受挫。于是，他转学来到匹兹堡大学学习；2001 年数学博士毕业，现在在一家计算机软件公司工作。最新在网上看到，王鲁燕在匹兹堡华人教会上

作见证，演讲的第一段是：“我是一个内心非常骄傲的人。尽管我没什么太多骄傲的资本，却看不起任何人。神能让我这样骄傲的人得救是他行的一件大神迹。”

我的文章写在善科问答网站上，因为我觉得这个网站给想写数学文章的作者提供了很多方便，可以很容易地放链接、传相片和打数学公式。第一天我写了文章的三分之一，就被“好事者”传到北大的 BBS 未名空间站。有一个跟帖写得很有意思，传神地描述了王鲁燕出国前在北大的教书生涯：“他在曾经的大一新鲜人中名声如雷贯耳啊。他讲线性代数，从有限群讲起。期中考试，好像定规矩延时可以，但要扣分，然后给了几个人负分。呵呵。老王是个性情中人啊。”也只有北大才可以出这样的“名师”啊。

人物之二：张益唐

张益唐的故事这几天在很多媒体上开始流传，但在 2013 年 5 月 10 日以前，知道他名字的可能屈指可数。知道他行踪的也没有几个。为了潜心研究数学，他几乎与世隔绝，在美国的偏远省份潜伏下来。2000 年初，他的妹妹曾在网上发寻人启事，哥哥张益唐失去联系了。当时在宾州州立大学当教授的张的老同学给他妹妹回了电邮，表示他哥哥健康地活着，在钻研数学呢。

1990 年中，我和沈捷讨论写一本微分方程谱方法的书（这本近 500 页的书终于在 15 年后的 2011 年完稿；由施普林格出版社出版）；为此沈捷到温哥华访问了我一周。研究之余，我们也闲聊了北大时候的一些老师、同学，他就说了张益唐的故事。虽然我大学就听过张益唐的大名，知道他是数学系的高材生，但张益唐的故事很多是从和他接触较多的校友那里知道的。

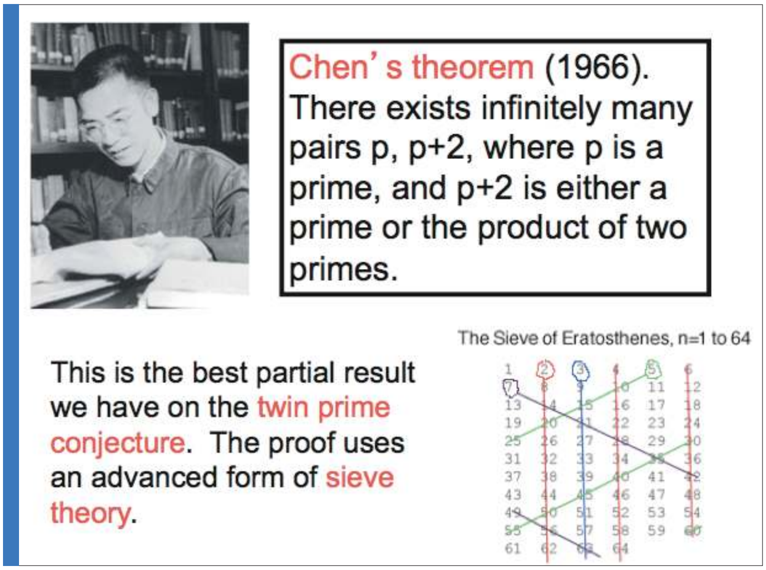
张益唐是北京人，1955 年出生，1978 年进北大数学系。

《人民文学》杂志 1978 年第 1 期上，作家徐迟发表了一篇报告文学《哥德巴赫猜想》，讲述了数学家陈景润刻苦钻研，终于在哥德巴赫猜想研究上取得重大突破的真实故事。《人民日报》和《光明日报》随即转载，一时间陈景润和哥德巴赫猜想变得家喻户晓。关于哥德巴赫猜想的故事，中科院数学院贾朝华在《数学文化》2010 年第 4 期上的文章《从哥德巴赫说开去》是一篇难得的高水平的科普文章；此文文笔优美，从历史的画卷上写数学，让人有品尝美味咖啡的感觉。

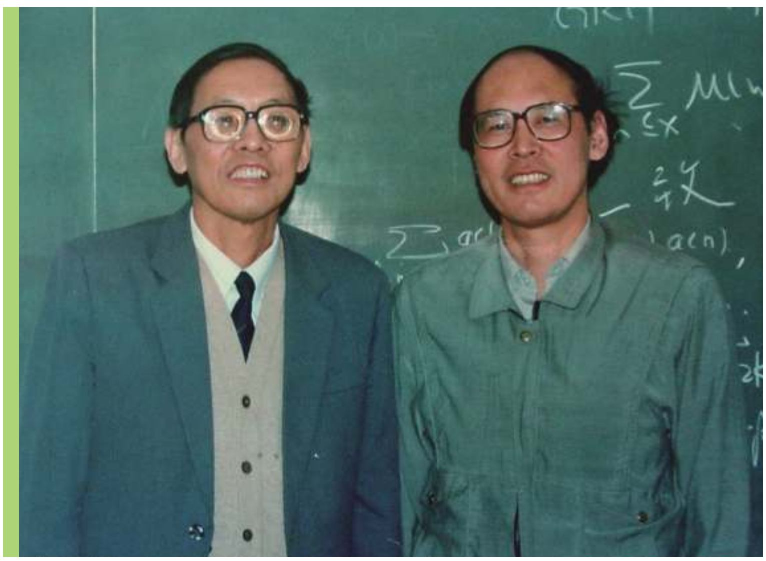
张益唐

像那个时代很多有志青年一样，张益唐也是被徐迟的文章、被陈景润的故事、被哥德巴赫猜想引导到数学系，以致终身投入到数学中去。

北大的 7 年为张益唐打下了坚实的数学基础；那时的北大教书育人之风极强，最顶尖的教师都在讲台上耕耘。代数高手段学复、聂灵沼、丁石孙、潘承彪，都给张益唐这些代数爱好者打下了基础。北大也有很多眼界很高的老师，学富五车，但不轻易落手写小文章，像江泽涵、廖山涛、段学复等，但谈



著名数学家、菲尔兹奖得主陶哲轩对陈景润工作的介绍



1995 年潘承洞与潘承彪于山东大学（展涛 摄）

起大问题颇为津津乐道，这让年轻的王鲁燕、张益唐们“中毒”匪浅。也只有北大，才能够让王鲁燕、张益唐们活得很舒服。很难想象，英语一塌糊涂、在其他学校很可能无法生存的王鲁燕，在考研连续失利的情况下，还可以留在北大任教，系主任人前背后还把他夸成一朵花。要不是结婚后生活压力太大，王鲁燕也许会在数学上留下一篇神话。

张益唐 1982 年毕业后跟随著名数论专家潘承彪读了三年的硕士。潘承彪教授的哥哥就是大名鼎鼎的前山东大学校长、因在哥德巴赫猜想方面的工作而闻名的潘承洞院士；而我们《数学文化》的主编刘建亚、编委蔡天新都是潘院士的高足。潘氏兄弟也是北大数学系校友，哥哥是 1952 年入学，弟弟是 1955 年入学。丁石孙在他的《自述年谱》中对他们有这样的描述：“潘承洞认为潘承彪基本功比较扎实，他有什么想法，就让他弟弟帮着算。”大学毕业后，哥哥去了山东，弟弟留在北京。

北大数学虽然 77 级没有招生，但于 1982 年初还是招了 77 级本科毕业生作为硕士研究生；那是 1966 年后上高中的人第一次成为研究生。那时的北大，大学生戴白底的校徽，研究生戴桔黄底的校徽，教师戴红底的校徽。80 年代初在北大校园最难看到的是那些带橘黄色校徽的天之骄子；大概全校也就 100 多位吧（现在很多学校一个系的研究生就可以顶上那时全北大的研究生数了）。因为历史的原因，1977 级的学生是 78 年的春季入学的，而 78 级的学生是 78 年秋季入学的；所以这两届研究生的入学时间也就相差半年。1982 年初来北大读数学研究生的有南京来的田刚（现北大数学院院长）、巫孝南（现香港浸会大学教授）、山东来的张继平（前任北大数学院院长）、上海来的张来武（现中国科技部副部长）、湖南来的许进超（现美国宾州州立大学讲座教授）、武汉来的吴志坚（现美国阿拉巴马大学数学系主任）等，其中北邮考进来的贾朝华（现中科院数学院的教授、1998 年国家杰出青年基金获得者）师从潘承彪教授，研究解析数论，并且他在北大读完了硕士和博士。张益唐是 1982 年秋天本科毕业后跟随潘承彪的，也是研究数论。

硕士研究生毕业后，张益唐来到了位于美国中西部印第安纳州西拉法叶（West Lafayette）的名校普渡大学读博士。普渡大学最著名的是其工学院；我前年去台湾的成功大学访问，在其校史馆里第一次知道是普渡大学把一个小小的台南学院变为亚洲名校。其中多名普渡教授在 1950-60 年代的台南乡下一住就是十几年，把一个穷孩子硬给拉扯成了一个高富帅的小伙子。美国的普渡

张益唐的博士导师莫宗坚

1993 年北大 78 级同学在沈捷（中）、陈敏（右 2）夫妇家中聚会。左为张益唐（沈捷提供）

大学所在地属于典型的乡下，最大的乐趣也许就是钓鱼，摘摘老玉米了。这应该是研究数学，特别是纯数学的天堂了。

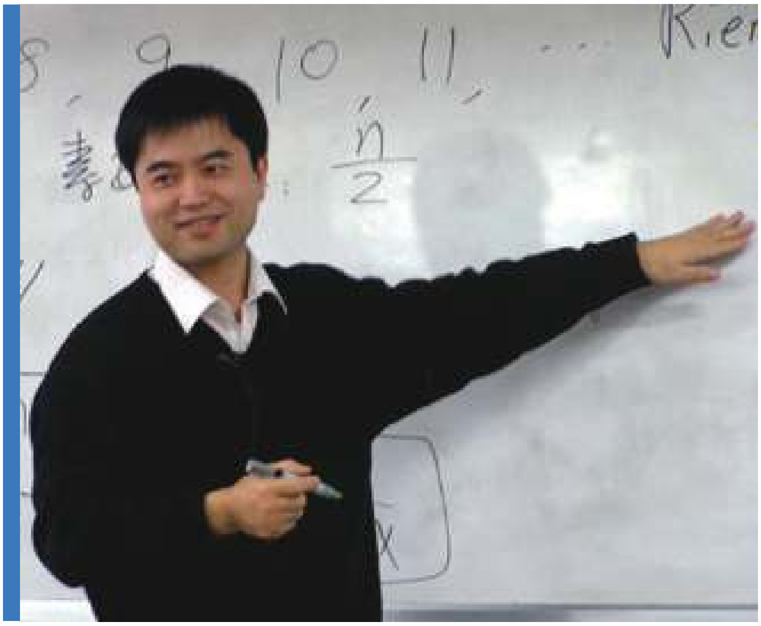
张益唐在美国的导师是从台大走出去的代数专家莫宗坚（Tzuong-Tsieng Moh），他 1969 年从普渡大学开始学术生涯直到今天。北京大学数学丛书里面的莫宗坚的《代数学》是很不错的代数参考书。莫老师也是性情中人，网上流传的他的《少年游——我的大学时代》很精彩，看看小标题就能体会到："读书记；生活记；恋爱记；跷课记…。"文章开头一段："民国四十七年（一九五八年）我从中部的台中一中考上了台大化学系。当我只身远赴台大报到入学时，看到堂堂学府，几排高大的椰子树，眼睛为之一亮。又有梳着清汤挂面头发的女生们奔走在校内路上，真是一个新奇的世界！"就颇为雷人。中间追女生段的开头"经过了一、二年的苦读，我才发现女同学们脸若春花，眸若点漆。有些小和尚们，经也不念，钟也不打：'不拜佛祖，只拜观音'。王俊明坦言：'不论追不追女生，总要花同样的时间想女生'"也算实话实说的经典。

张益唐的博士题目是雅可比猜想 (Jacobian conjecture)；这一猜想由德国数学家科勒（Ott-Heinrich Keller, 1906-1990）于 1939 年提出来，是关于多个变量的多项式里面的一个猜想。这个猜想本身很有名，但更有名的是经常会产生错误的证明。莫教授在普渡大学自己的网页上就挂了两篇短文章，一个是说 Kuo-Parusinski-Paunescu 的雅可比猜想证明是错的，另一篇是说中国科大的苏育才的证明也是错的；这两篇纠错的文章七、八年前就上网了。

张益唐也曾牺牲在雅可比猜想上，并且比较惨重。他 90 年代博士毕业前夕，宣称解决了雅可比猜想，并且有几个专家对他的证明很感兴趣。但是，不幸的是他的证明里的一个引理是其导师莫宗坚的一篇发表的成果，本以为是对的，但再排查时，查出莫教授之前的结果是错的。你应该知道后果很严重吧？张益唐几年的主要心血付之东流了！！

如果雅可比猜想那时被张益唐攻克了，他应该可以毫不费劲地拿下数学界最高的菲尔兹奖，因为这个问题太有名，并且他那时候才三十几岁，低于菲尔兹奖得奖的上线岁数四十。可是这只能是虚拟语气了。想当年天才约翰•纳什作出了著名的纳什嵌入定理（Nash embedding theorem）也在翘望菲尔兹奖，结果未能如愿，导致精神失常。好在三、四十年后，诺贝尔奖委员会根据他研究生博士论文建立的纳什均衡理论给了他一个经济奖，才横扫了失去菲尔兹的阴霾。

虽然未能全部解决雅可比猜想，张益唐对这个问题的部分解决还是有些贡献的，所以他还是于 1991 年拿到了普渡大学的博士学位。可是，因为博士论文的结果没有正式发表，



中科院数学院教授、新罕布什尔大学教授葛力明

加上和导师的关系极不融洽，张益唐变成博士到手失业开始，连个博士后都没有找到。

不过张益唐毕业的时候是历史上数学最难找工作的时候。当时苏联刚刚解体，那里培养的大量数学人才大部分投奔美国、欧洲，很多训练有素、富有成果的俄国数学家向一个个美国的数学系投简历，这严重冲击了美国的数学就业市场。没有发表过论文、没有导师推荐信的张益唐要找到大学教书的工作显然是非常困难的。

对张益唐在普渡的八年，他的博士导师莫宗坚在张益唐出名后的 10 天内，写了一篇文章："张益唐 1985 年 1 月至 1991 年 12 月在普渡的岁月"，描述了张益唐那几年的学习和工作。文中说：有时我很后悔没有帮他找工作（Sometimes I regret not fixing him a job）；他从此再也没有找我写过推荐信（He never came back to me requesting recommendation letters）。在这篇文章中，莫教授还说张益唐是一个有生活情趣的人，曾经被选为普渡大学中国留学生协会主席，并且干得不错；他也认为张益唐的古典文学修养很好。

一面要继续做数学，一面还要糊口。毕业后的前六、七年他干过很多杂活，包括临时会计、餐馆帮手、送外卖。你能想象一代北大数学才子、数学博士好几年送披萨饼、在 KFC 卖炸鸡的情形吗？看到这里，你是否对孟亚圣的"天将降大任于斯人也，必先苦其心志，劳其筋骨"有更深刻的理解呢？

张益唐毕业后基本隐居起来，很少和人来往，他和同学们的联系方式之一就是同学生日时电邮一个生日问候。他居然能记住班上所有同学的生日，这可能是我们这些凡人做不到的。那时他的一些同学已经在美国大学里拿到助理教授、副教授甚至正教授的位置，比如同班同学沈捷、陈敏夫妇已经在宾州州立大学有正式教职。张前几年每年会找这两口子吃顿饭，基本是悄悄地来，悄悄地走，一年之间基本杳无音讯。

从 1999 年到 2005 年，张益唐又回到了学校，到美国的新罕布什尔大学（University of New Hampshire）做一个非编制的助教。新罕布什尔州是位于美国新英格兰地区的一个州，绰号叫"花岗岩州"，因为本州产花岗岩，另外也是因为这个州比较坚守传统观念，政府非常节俭。而新罕布什尔是成立于 1866 年的一所综合性公立大学，学校位置不错，坐落在美丽的达勒姆（Durham）镇上，步行即可到达市中心；曾被《纽约时报》选为全美最好的大学之一。虽然教学量比较大，比起有正式教职的工资性价比低很多，但能回到学校，做自己驾轻就熟的事情，还能利用图书馆、办公室做研究，对一个胸有大志的数学人来说，应该是比较满足的。

2005 年以后，在大学里代课了六年的张益唐获得转正，变成了有比较稳定位置的讲师，主要任务还是上课，据说也只是一年四门课的教学任务，和大学里面的其他教授的工作量相仿，没有行政工作和申请经费的压力。对这个位置，科研是不算工作量的，研究经费是面向那些助理教授、副教授、正教授这些所谓的 faculty 序列的。但是，对于一个胸中有数学的人来说，没有经费的支持可以换来更多平静的思考时间，何尝不是一件好事呢？

张益唐为什么会落脚新罕布什尔大学？这里面有两个主要人物，都是我们北大 80 级的同学；一个是唐朴祁，一个是葛力明。

1999 年初，已经在美国 Intel 实验室工

作了几年的唐朴祁去纽约参加IEEE年会；他在会上发表了一篇关于网络交通（Network traffic）的文章。这个研究解决了数字网络服务质量（QoS）的一类实时量化的问题。唐朴祁和一个同事把多种可能情形的效率进行了研究，结果它们可以把各种情况提升到最优的 $O(n)$ 复杂性，这种线性复杂性已经是不可以做任何改进的了。但有一种情形只有复杂度为 $O(n\ln(n))$ 的结果，这当然不是最优的，会成为此算法在高速网络实时应用时的障碍。开会期间，唐朴祁找到在纽约的张益唐，因为张益唐没有固定电话，找到他还是不容易的。他们见面后，唐朴祁发现老友有些疲惫沉默，于是就和他聊他感兴趣的数学。特别聊到他们在计算复杂性方面遇到的这个困难。大约三周以后，张益唐告诉唐朴祁他看懂了这篇文章并对余下的这个数学问题有主意了。在电话上来回交流几次以后，他们终于找到了最优的线性解。这就导致了唐、张的一个合作的专利。我谷歌了他们二人的英文名字，立刻找到了他们的专利。据说这个专利已经在计算机网络基础设施领域中有广泛应用。三个星期啃下一个有广泛实际用途的计算机算法难题，让张益唐顿觉宝刀不老，信心大增。唐朴祁也对老友的数学实战功夫印象深刻。

同年晚些时候唐朴祁到哈佛参加ACM会议时，特意和在波士顿附近工作的葛力明提起张益唐的强大分析实力和他当时的艰难处境。作为学长的张老师不仅做过他们的习题课老师（应该是抽象代数吧？），也是他们自己组织的大学生讨论班上的常客，他们从张老师那里学到过很多数学。给他们印象最深的是，无论在任何环境下，他们的这位好友都在思考数学、心里总是装着那几个数学大问题。

这次和葛力明见面时，唐朴祁已经不知道张益唐的准确工作地点。此时从宾州大学博士毕业并在大学有固定位置的葛力明似乎更有条件帮一下他们的朋友和老师。经过一番周折，葛力明在美国南方的一个赛百味快餐店（Subway）联系上了张益唐。两三天后张益唐就来到新罕布什尔大学做访问学者了。每过几天，张老师都会说，有进展，应该很快就出来了。但时间过得很快，两个月、

张益唐夫妇（左）新婚蜜月期和唐朴祁全家的合影

三个月，两年、三年……，后来的故事大家都知道了，十四年后，张益唐的文章被权威期刊《数学年刊》接受了。当然，在美国大学里要留一个没有多少资历的人十四年肯定不是一件简单的事，中间也有酸甜苦辣的故事。但有张益唐这样传世的好结果，任何付出都值得。这其中帮他最多的是今年四月十九日刚去世的新罕布什尔大学数学系前系主任凯尼斯•阿佩尔（Kenneth Appel）。阿佩尔也是世界级的数学家，他和沃夫冈•哈肯（Wolfgang Haken）借助电脑在1976年首次证明了四色定理，这也是载入史册的工作。值得欣慰的是，在这位伯乐去世的前两天，他已得知张益唐的结果，这无疑是对他多年来付出的最好报答。

我们80级的帅哥葛力明过去的十年一半时间在中科院数学院工作，教书育人，深得国内同行的好评；同时他也是新罕布什尔大学数学系的大教授了，他在国际数学家大会上应邀做过45分钟报告，这在他们学校应该是凤毛麟角的了。更重要的是，作为北大数学系的80级学弟，对于数学好、又执迷于数学的学长，他真正做到了出手相助！普通数学人崇拜名气大权力大的，但真正数学人

唐朴祁（左）与张益唐

是敬仰那些为了数学而痴迷而献身的！

更重要的是，在新罕布什尔大学数学系，张益唐又可以在数学讨论班上把自己新的想法、新的工作与大家分享。这样，他平生的第二篇文章，关于黎曼猜想的文章（*On the zeros of ζ′(s) near the critical line*）发表在重量级的数学期刊《杜克数学》（*Duke Mathematical Journal*）上。之后，他还有一篇 50 多页的关于朗道 - 西格尔（Landau-Siegel）零点猜想的文章（*On the Landau-Siegel Zeros Conjecture*）放在数学公开平台网上。因为当时张默默无闻，文章被当成民科作品忽略了。现在发现他有真才，文章又被翻出来，有评论说这可能是解析数论的另一个重大突破。而现在已经被承认的孪生素数猜想文章也是在葛力明的讨论班上讲过的，从这里走向世界的。

很多他的大学同学都不知道张益唐成家了。张益唐多年前结婚了，妻子是山东人。在引起公众注意后，张益唐在加州工作的妻子提醒他的第一件事是记得把头发梳理好。

在此节收尾之前我要聊两个题外话。张益唐在北大本科的专业是计算数学专业，最后做了纯数学的大事。无独有偶，著名数学家、北京大学和普林斯顿大学的数学讲席教授田刚，毕业于南京大学计算数学专业；29 岁就成为芝加哥大学数学系正教授、现在是纽约大学柯朗研究所讲席教授的林芳华，毕业于浙江大学的计算数学专业。也许要想做好纯数学，先学计算数学有好处？当然了，这只是小概率事件了。另外我也侃一下我为什么选了计算数学专业。北大数学系选专业是大二的下学期。前两年所有的学生不分专业，都在一起上课。我们大二下学期，每个专业选一个老师来讲自己的专业，属于“忽悠”吧。先上台讲的是概率统计专业的陈家鼎教授，讲了统计的重要性，特别提到概率统计的殿堂级人物柯尔莫戈洛夫的理论如何了得，如何优美。等一会计算数学的黄敦教授开讲，他的第一句就是：“刚才有人大谈特谈柯尔莫戈洛夫。请问在座的各位，谁见过柯尔莫戈洛夫？”冷场半分钟后，黄老师提高嗓门说：“我，黄敦，我见过柯尔莫戈洛夫。”全场鼓掌；那一届计算数学专业学生数远超其他任何专业。

再谈哥德巴赫猜想和孪生素数

前面已经谈了素数分布定理、孪生素数的定义。作为《数学文化》期刊，既然要做一手传播，还需要加点有数学味的部分，供有些数学基础的读者欣赏。因此，我在本文初稿完成后，专门请教了数论专家、我的《数学文化》的搭档刘建亚教授，听了他一个半小时的课，长了很多学问。

1900 年，希尔伯特在第二届国际数学家大会上提出了著名的二十三个希尔伯特问题，其中第八个问题就包括了哥德巴赫猜想和与它类似的孪生素数猜想。很多人认为这两个猜想是紧密相关的，其证明难度也相仿。所以攻克孪生素数猜想对数学发展来说是非常有意义的大事。

所谓的哥德巴赫猜想，其实是 17 世纪一个叫哥德巴赫的业余数学家在给大数学家欧

拉的一封信中提出的一个问题。哥德巴赫算不上数学家，但是他偶然发现的问题却让他扬名了。他发现：任何一个大于等于 6 的偶数，都可以写成两个素数的和。例如 $6 = 3 + 3$，$20 = 7 + 13$，$100 = 3 + 97$ 等等。不过，他只是验证了很多数满足这一规律，却无法给出证明。

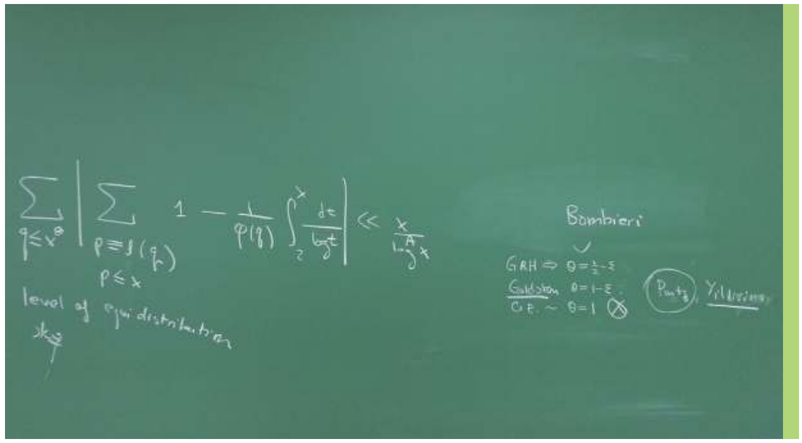
后来，辗转几个世纪，一直没有人证明它。1920 年挪威数学家布朗建议将证明分成若干步骤，就是先证明任何充分大的偶数都可以表示成两个正整数之和，其中一个正整数的素因子个数都不超过 a, 另一个素因子个数都不超过 b。先对比较大的 a 和 b 证明，然后逐步缩小，如果最终缩小到 (1, 1)，那么就证明了哥德巴赫猜想。不过很可惜，到目前这个定理还没有得到证明。最接近的 (1 + 2) 是中国数学家陈景润的成果。1956 年，中国数学家王元证明了命题 (3 + 4)，由此开启了中国数学家在哥德巴赫命题 $(a + b)$ 研究上的先河。之后，王元和另一位中国数学家潘承洞又得到了若干重要的结果，使得我国在哥德巴赫猜想方面的研究达到了国际先进水平。1966 年，陈景润宣布证明了命题 (1 + 2)；1973 年，他发表了 (1 + 2) 的全部证明。这些信息我们可以从徐迟的报告文学或贾朝华的文章里得到。

无论哥德巴赫猜想还是孪生素数猜想，都涉及两个素数。因此，要想证明这两个猜想，必须先了解素数在形如 $qm + a$ 的算术级数中的分布，此处 q 与 a 互素，而 m 跑遍所有正整数。

素数在形如 $qm + a$ 的算术级数中的分布规律如下：当 x 趋于无穷大时，不超过 x 且满足 $p \equiv a \pmod q$ 的素数的总数满足渐近公式：

$$\sum_{p\le x,\, p\equiv a \,(\mathrm{mod}\, q)} 1 \sim \frac{1}{\varphi(q)}\int_2^x \frac{dt}{\log t}.$$

其中左边的那个求和号底下的式子表示考虑所有的素数 p, 这些素数 p 不超过 x 并满足和 a 的差能被 q 整除；而左边的求和就表示满足这两个条件的素数的总个数。



刘建亚给本文作者讲解的公式

减号后面分母下的 $\varphi(q)$ 是欧拉函数。这个公式对每个固定的 q 成立，但是这是不够的；我们需要了解当 q 可以与 x 一起增大的时候，这个公式是否正确。这是一个非常困难的问题。在平均意义下，我们有如下结果

$$\sum_{q\le x^\theta} \max_{(q,a)=1} \left| \sum_{p\le x, p\equiv a \,(\mathrm{mod}\, q)} 1 - \frac{1}{\varphi(q)}\int_2^x \frac{dt}{\log t} \right| \ll \frac{x}{\log^A x}, \qquad (*)$$

其中 A 是任意正数。最左边的求和表示对所有不超过 x 的 θ 次方的正整数求和，这里的 θ 称为素数在算术级数中平均分布的“水平”，θ 越大即水平越高，也就越难以证明。

1948 年，匈牙利数学家任义（Alfred Renyi）得到了 θ 的存在性，但是没有给出其具体数值。任义由此推出哥德巴赫猜想方向上的命题 $(1 + b)$，这里 b 是一个依赖于 θ 的正整数，随着 θ 的增大而减小。因此任义的 θ 是个未知的天文数字。

1962 年潘承洞证明了 θ 可以取 1/3，进而推出命题 (1 + 5)。利用这个 θ，王元证明了 (1 + 4) 。同年，潘承洞进一步证明了 θ 可以取到 3/8, 进而用简单的筛法也证明了 (1 + 4)。1965 年，苏联数学家布什塔波（A. Buchstab ）用这个 θ 加上复杂的筛法证明了 (1 + 3)。

1965 年意大利数学家邦别里（E. Bombieri）以及苏联数学家维诺格拉多夫（A. I. Vinogradov）独立证明了 θ 可以取任何小于

1/2 的正实数。这相当于证明了，广义黎曼猜想的一个重要推论在平均意义下成立，而且这个结果可以在很多场合替代广义黎曼猜想。邦别里因此获得 1974 年的菲尔兹奖。1966 年陈景润利用邦别里 - 维诺格拉多夫的 θ 证明了命题 (1 ＋ 2)，并发表了论文摘要。由于文化大革命的原因，陈景润论文的全文直至 1973 年才得以发表。

上述关于平均分布水平 θ 以及命题 (1 ＋ b) 的历史，同样适用于孪生素数猜想，并给出 (1 − b) 的结论。例如，陈景润的 (1 ＋ 2) 的证明也推出关于孪生素数猜想的命题 (1 − 2)，也即方程 $p - P_2 = 2$ 有无穷多组解，其中 p 是素数，而 P_2 至多有 2 个素因子。

然而孪生素数猜想有其特殊性，也即方程 $p_1 - p_2 = 2$ 的右端是个固定的常数 2，这与哥德巴赫猜想不同。注意到这点，公式 (*) 中的 a 可以简单地取为 2；此外可以要求 q 只取某些满足好的性质的正整数，而不是取遍所有的正整数。有了这些变化，可以把 (*) 当中的水平 θ 做得更大，甚至超过 1/2。注意到这样强的结果，已经不再是广义黎曼猜想的推论了。1980 年代，邦别里、加拿大数学家弗里德兰得（John Friedlander）以及前面提到的张益唐文章的审稿人伊万尼克等致力于证明这样的平均分布结果，他们得到的水平 θ 可以大到 4/7，这个数是大于 1/2 的。

公式 (*) 中的水平 θ 最大可以取多大

北京大学数学 78 级毕业照。2 排左 17 至 21：丁石孙、段学复、江泽涵、庄圻泰、程民德。3 排左 18：王鲁燕，4 排左 21：张益唐。

呢？当 $\theta=1$ 的时候，大家已经知道 (*) 是错的。因此最高的期望是，θ 可以取任何小于 1 的正实数；这个期望被称为艾略特 - 哈伯斯坦猜想（Elliott-Halberstam conjecture）。假设这个大胆的猜想正确，前面提到的加州圣荷西大学的戈德斯通、匈牙利数学家约翰 • 宾兹（János Pintz）、土耳其数学家谢姆 • 伊尔泽姆（Cem Yildirim）在 2010 年合作证明了：存在一个正偶数 $h\le 16$，使得方程 $p_1-p_2=h$ 有无穷多组解，其中 p_1, p_2 都是素数。注意，这是一个条件结果；并且假设的条件艾略特 - 哈伯斯坦猜想未必正确。

张益唐在他的《数学年刊》论文中，首先证明了 (*) 对某个大于 1/2 的水平 θ 成立，进而由此推出：存在一个正偶数 $h\le 70000000$，使得方程 $p_1-p_2=h$ 有无穷多组解，其中 p_1, p_2 都是素数。请注意，这个结果是无条件的！张益唐成功地避开了艾略特 - 哈伯斯坦猜想的假设。

我们可以把张益唐的这个定理简称为 $(1-1=h)$。这个 h 虽然未必是 2，但是上述结果的质量与经典的孪生素数猜想所差不远。

刘建亚告诉我，张益唐的文章不但可以载入数论史，还可以载入数学史，将有深远的影响。另外，受过严格解析数论训练的人都可以看懂这篇文章；山大的数论研究生就可以读懂。他们下次的数论讨论班就会讨论张益唐的文章。

北大数学系 80 级毕业 25 周年相聚于燕园。1 排右 7：刘森；2 排左 9：葛力明；3 排左 2：张占海；3 排左 4：唐朴祁；4 排右 2：孟肖敏；4 排右 5：汤涛

结尾的几句话

过去的二十年是数学史上重要的一个时期；我们在有生之年见证了费马大定理的证明、庞加莱猜想的证明，还看到了孪生素数猜想的重大突破。作为华人，我们为张益唐的成功感到由衷的高兴。

文章结尾前先谈谈张益唐故事的启迪。基础数学的一些极其重大猜想的突破需要极聪明的天才，这些人包括国际数学竞赛的金牌得主（如陶哲轩、佩雷尔曼、吴宝珠），有受过正规数学训练、且被老师同学公认的数学高手（如陈景润、张益唐）。首先这些人要受过正规的数学训练。由徐迟先生力作带动的全民作猜想的做法是不可取的；希望这也不是张益唐的成功带来的副产品。第二，这些数学天才还需要对数学有无限的热爱和执着，同时没有对物质的过高追求。后一点尤其重要。一颗平静的心胜过五颜六色的光环、五花八门的资源。在清贫中做出重大贡献的佩雷尔曼、张益唐的故事再次证明了这一点。

再回到北大数学 100 年，和此文开头呼应。北大数学 78 级的精英分子很多，很难一篇文章写完。本文仅仅通过王鲁燕、张益唐写出那个时代的一些往事。让我自己也过了一把回忆往事的瘾。实际上，78 级还有一些最终没有从事数学的同学，比如我前面提到的刘森，他毕业后调去了中纪委，曾多年在中纪委的华北室工作。在 80 级同学聚会发言中，他告诉我们，在办案中用数学的逻辑，往往可以达到事半功倍的效果。十年前河北、黑龙江那两个大腐败案都是他们室主抓的。为此，居然有人对他的办公室窗户打黑枪。还有些同学离开数学的原因是兴趣发生了变化。比如号称民族主义领军人物的王小东，他自称 78 年入大学头两年成绩上乘，后不读数学，只读英语小说。毕业后学习经济管理并极力宣传民族主义，其著作《全球化阴影下的中国之路》颇有影响。

那个时代的北大数学人确实有多方面的精彩；像我的大学同学孟肖敏毕业后居然当了律师，开了个上海信天诚律师事务所；另一位同学张占海曾三次担任中国南极考察的领队和首席科学家，我们毕业 25 周年给母校的礼物就是张占海带回来的南极石。还有很多这样的故事真让我感到行行出状元，聪明人在哪里都可以出彩。

百年的北大数学培养了很多优秀的大学生和研究生，为中国数学、世界数学输送了众多精英；我相信再过一百年，北大数学将会更上一层楼，将会把更多的世界数学精英吸引到燕园，将会对人类的数学发展作出更大的贡献。

后记：感谢沈捷、唐朴祁、葛力明提供的宝贵信息。感谢刘建亚、卢昌海、蔡天新、游志平、欧阳顺湘、金石、励建书等朋友的支持。本文作者是香港浸会大学数学讲座教授、《数学文化》联合主编。

有错必究 *汉明码（Hamming Code）的原理及其应用*

万精油

上期的题目是帽子的颜色问题。为方便解答，我们把上期题目再列一遍。

帽子的颜色问题：三个人头上都被戴上一顶帽子。帽子的颜色是蓝色或红色，完全独立随机。每个人可以看见别人的帽子，但看不见自己的帽子。每个人可以有两种选择：猜自己帽子的颜色，或者放弃（就是不猜）。每个人把自己的决定写在一张纸上。如果最后的结果是至少一人猜对而且没人猜错，那么他们可以得到一笔巨额奖金。我们的问题是，他们用什么策略才能最大地提高得奖的概率。

这个问题二十年前曾经在美国数学界、计算机界轰动一时。不光因为它是一道趣味题目，而且因为这题目背后蕴藏着计算机编码理论中的一个重要思想。

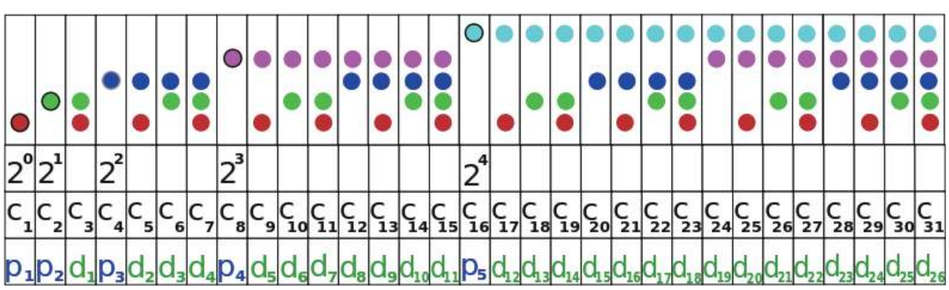
与别的问题不同，这个问题最困难的地方是只要有一个人错则全错。所以不能像别的题那样用数量来搞概率。

如果每个人都随机猜，那么三个人都猜对的可能性是八分之一。除此之外，好像没有什么别的出路。因为帽子都是随机选的，你头上的帽子颜色与别人的帽子颜色独立，似乎没有任何根据让你决定选什么颜色或放弃。其实不然，正因为帽子是随机选的（每个帽子都有二分之一的机会是红色，二分之一的机会蓝色），所以总体帽子的颜色满足一种分布。有些情况多一些，有些情况少一些。我们可以在这上面做文章。

先看三人的情况：三个人的帽子颜色一共有八种情况，红红红，红红蓝，红蓝红，红蓝蓝，蓝红红，蓝红蓝，蓝蓝红，蓝蓝蓝。如果大家商定，当某人看见两个同色的帽子时，他就猜另一种颜色，否则放弃。那么，根据上面的八种分布，我们很容易看出，有六种情况他们都能通过。只有两种情况他们会失败，即全红或全蓝的时候。再仔细数一数，他们答错和答对的时候一样多，都是六次。唯一的区别是，答错的时候大家都一起答错。而答对的时候都只有一人答对，别的人都放弃。

这个题目可以推广到更多人的情况。人数多的时候就不能靠一个情况一个情况地数，必须要有系统方法。这就需要介绍一种叫做汉明码的东西。

现在我们的生活都离不开网络，随时随地都在浏览从网上传来的东西。但是，网上的传递不能保证 100% 都对，经常会出现错误。计算机怎么发现传递有错误？发现了错误以



后又怎样纠正？汉明码就是用来干这个的。

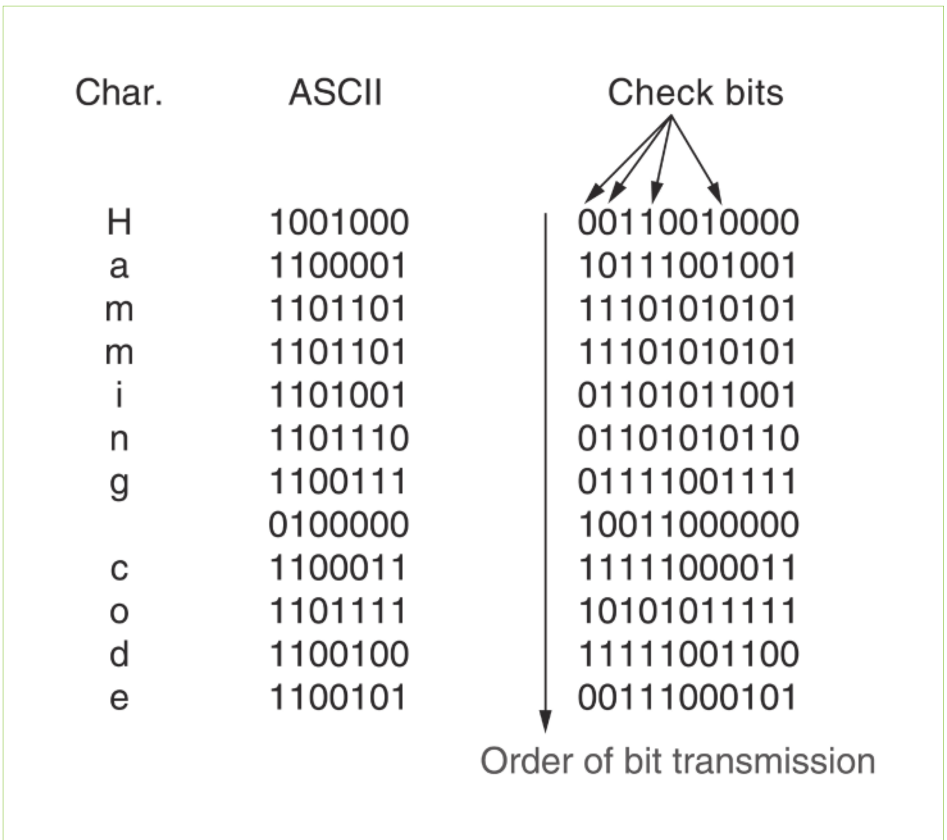
在介绍汉明码以前，先简单介绍一下如何用奇偶性来检查传递的信息是否出错。

如果我们有 8 个比特可以用。那么我们可以用其中的 7 个比特来传递信息，用一个比特来作验证码。如果那 7 个比特传递的信息有奇数个 1，验证码就是 1，否则就是 0。 这样一来，如果信息传递中有一个码出现错误，该是 1 的地方变成了 0，或者该是 0 的地方变成了 1，与这个验证码不符，我们就知道传递有错。这个方法的缺点是它虽然能发现错误，但不能知道错误出在哪里，也不能纠正，只能要求重新传递。汉明码是在用奇偶性来检查传递的信息是否出错的基础上发展出来的更高级的方法。它不但能发现错误，而且能知道错误出现在哪里，从而进行自我纠正。

要知道错误出现在哪里，一个验证码是不够的，汉明码需要用到多个验证码，具体个数根据能够传递的信息长度。假设有 2^N-1 个比特可以传递。那么我们用其中的 N 个比特来做验证码，剩下的 2^N-1-N 个比特来传递信息。要用这 N 个验证码来发现错误并确定其位置，这 N 个码的设置就很有讲究。具体的方法我们用 $N=4$ 的情况作一个说明。

$N=4$ 时，我们有 15 个比特可以用。用其中 11 个来传递信息，4 个来做奇偶性验证码。我们假设这 4 个比特的位置是 1, 2, 4, 8。其余的 11 个比特就是真正要传递的信息。如果我们把这 11 个位置都用二进制表示，每个位置就有 4 个比特，我们把它们叫作位置比特，如“3”的位置比特为“0011”。

第一个验证码验证的是所有位置比特第一比特（自右数）是 1 的位置（其实就是所有



单数位置）的奇偶性。第二个验证码验证的是所有位置比特第二比特是 1 的位置（2, 3, 6, 7, 10, 11, 14, 15）的奇偶性。第三个验证码验证的是所有位置比特第三比特是 1 的位置（4, 5, 6, 7, 12, 13, 14, 15）的奇偶性。第四个验证码验证的是所有位置比特第四比特是 1 的位置（8 到 15）的奇偶性。比如，我们要传递的信息是 01100101110。把这个信息放到 15 个比特里，我们就有 _ _0_110_0101110。注意到在 1, 2, 4, 8 这四个位置上没有数，所以下面看位置比特时也不包括这些位置。所有单数位置的比特是 0100010，有偶数个 1，所以第一验证码是 0。位置比特第二比特是 1 的位置的比特是 0101010，有奇数个 1，所以第二验证码是 1。位置比特第三比特是 1 的位置的比特是 1101110，有奇数个 1，所以第三验证码是 1。位置比特第四比特是 1 的位置的比特是 0101110，有偶数个 1，所以第四验证码是 0。把这些验证码放进原信息，我们就得到全部 15 个传递比特，010111000101110。当然，我们的这种安排法主要是为了便于讲解。真正传递的信息没有必要把原信息打散重新组合。而是把原信息放在前 11 位，验证码放在后 4 位。

这样传递过去的码如果有验证码不符，比如传过去变成了 010111000111110，那么 1, 2, 4 验证码不符，我们就知道第 11 个码出了问题，可以把它纠正过来。

仔细想一想，我们意识到，这个方法只能发现有一个错码的时候。在极少数情况下，如果有两个比特同时出错，这个方法就没有办法发现了。于是人们又设计出再加一个验证码验证总体奇偶性，就可以发现有两个错码的时候。

现在再回到我们帽子颜色的题目中来。当总位数是 2^N-1 时，汉明码的一个特性是所有对错码覆盖了所有 2^N-1 位数。戴帽子的人可以利用这个特性来设计出一套猜帽子颜色的策略。还是拿 $N=4$ 来举例。他们自己从 1 到 15 排一个号。当看见别人的帽子颜色以后，把这些颜色放在相应的比特位置上（红色为 1，蓝色为 0）。再把自己的两个颜色带进去得到两个码。如果两个码都是错码，则放弃。如果一对一错，则猜错的那个颜色。如果实际上所有的帽子颜色形成一个对码（1/16 的可能性），则所有人都猜错。如果实际上所有的帽子颜色形成一个错码（15/16 的可能性），则只有错码位置上那个人会猜，其余的人都放弃。所以，这个方法给猜帽子颜色的人 15/16 通过的概率。

这个问题可以推广到任意 N。N 越大，通过的概率越大，$(2^N-1)/2^N$。对于帽子数不是 2^N-1 的时候，不同的数有不同的解法，还没有通解。

本文只是对汉明码做个简单介绍，对汉明码有兴趣的读者可以找相关的书读一读。

下期题目：监狱里有 $2k$ 个犯人。监狱长把犯人找来，说："你们的名字完全随机地放在这 $2k$ 个盒子里，每盒一个。明天你们轮流到这里来，每个人打开一个盒子，看看是不是自己，不是再开下一个，最多可以开 k 个，看到自己的名字就算通过。如果所有人都通过，就释放你们。现在你们可以讨论一个策略，完了之后不准再有任何形式的交流。"

每个人看到自己名字的概率是 1/2。如果没有策略，释放的概率是 $(1/2)^{2k}$，当 $k=5$ 时，成功率已经低于千分之一，k 更大时就几乎成为不可能事件。能不能设计一个策略，使得全体被释放的概率有一个与 k 无关的正下界？

注：$2k$ 个盒子从左到右一字排开。每次被打开后马上关上，不能挪动。不能做任何记号。事先商量好对策后犯人之间不能有任何交流。

编者按：善科网（http://www.mysanco.com) 是一个高质量的数学网络学习和交流平台，融创意与资源于一体，面向所有层次的数学爱好者。其每日一道的趣味数学题已经出现在其微信 shanketiku 上。感谢善科网授权本刊摘录其中的一些趣题。答案请见下期。

使用只包含数字 4 的算式（譬如 4^4+44），你能计算出 2012 吗？

规则：

① 你可以使用加、减、乘、除、乘方、开根和阶乘运算。

② 你可以使用任何由 4 组成的数，譬如 44.44。

③ 你可以使用括号。

另类的时钟问题：一只劣质钟，分针和时针长短一样。问在正午到凌晨的十二小时之内，有多少个时间无法通过这只钟来判断？（即两个不同的时间，对应指针形状完全一致，则这两个时间均无法辨识。）

时钟的时、分、秒三针，每天重合多少次？（注：三根指针均非作连续性旋转，即只能转动到有刻度的位置）

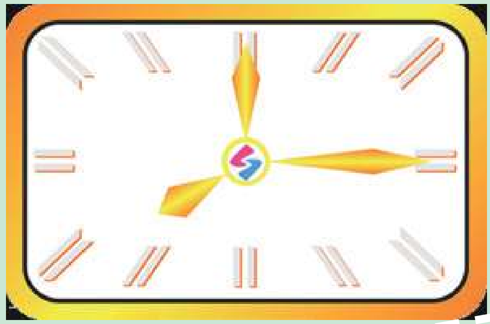
A) 2

B) 12

C) 22

D) 23

E) 24



一位店主收到以下的账单：

22 盒磁带：■ 29.3 ■ 元

其中首尾两个数字弄脏了无法辨认。店主知道，每一盒磁带价格在 25 元以上。 问每盒磁带的单价（以元计）是在以下哪两者之间？

A) 25 和 28

B) 28 和 32

C) 32 和 35

D) 35 和 40

E) 40 和 50

甲、乙、丙、丁和戊玩一种游戏，其中每个人充当狼或羊。狼说的总是假的而羊说的总是真的。

甲说乙是羊；

丙说丁是狼；

戊说甲不是狼；

乙说丙不是羊；

丁说戊和甲是不同的动物。

请问有几只狼?

枪打出头鸟

三人决斗问题趣谈

万精油

三人决斗问题在网上流传很久了，甚至有人已经把它写进书里。这个大家熟悉的题目我本来没有想把它放到我的微博上。可是，上周在《数学文化》的微博上看见推荐一个两人决斗问题，我觉得过于简单，于是把这个三人决斗问题拿出来作比较。题目出来一个星期了，想写一个答案算交差，没想到越写越长，140字的微博不够，于是干脆把它加长成一篇博客文章。

先说那个两人决斗问题。说是两个人搞“俄罗斯轮盘赌”，一个可以装六颗子弹的左轮手枪里装了一颗子弹。随机转盘以后两个人轮流用枪对准对方额头射击。每次打枪后重新转盘。问是先开枪划算还是后开枪划算，并算先开枪和后开枪的存活率。因为每次打枪后重新转盘，所以想都不用想肯定是先开枪的划算。至于先后的存活率，后开枪的人要在第一枪没有被打死的情况下（概率是5/6）才能达到与先开枪的人相同的状态。所以，后开枪的人的存活率是先开枪的人的存活率的5/6 。再加上两人的存活率之和是1，可以得

出先开枪与后开枪的存活率分别为 6/11 和 5/11 。所以我说这个问题过于简单。

其实，上面那个题窜改了“俄罗斯轮盘赌”。真正的“俄罗斯轮盘赌”是随机转盘后对准自己额头打，而且每次打完不再转盘，自动转进下一个子弹位。在这种情况下问先开枪划算还是后开枪划算就是一个很好的条件概率题。第一枪被打死的概率是 1/6 。第二枪被打死的概率是 5/6×1/5，还是 1/6 ，以此类推。当然如果对题目理解的很清楚，根本就不需要算。第 K 枪死的概率就是子弹在第 K 个弹腔的概率，因为是随机的，每个位置的概率都是 1/6，所以先打后打都一样。

三人的情况就要有意思得多。从两人到三人有点像从二体运动到三体运动。因为二体运动必须是平面运动，简单解一解 $F = Ma$ 就可以有结果。三体问题要复杂得多，根本没有解析解。牛顿、庞加莱这些大家都没有办法。当然，这个三人决斗问题只是比两人决斗问题麻烦一点，比三体问题那是要简单多了。 先叙述一下三人决斗问题。

> A，B，C 三人决斗。已知 A 的枪法奇准，百发百中。B 次之，三枪命中两枪。C 最差，三枪只能打中一枪。决斗的方式是三人轮流开枪，每次只能开一枪，可以随便选向谁开枪。为公平起见，他们决定让 C 先开枪。然后是 B(如果还活着)，最后是 A（如果还活着）。如果一轮结束后还有超过一人活着，再按 CBA 循环。问：在上面给出的条件下，每人的最佳策略是什么？如果大家都采用最佳策略，每人的存活率是多少？

首先，在三人都在的情况下，开枪的人应该打另外两人中命中率高的，因为如果他打中就轮到剩下的那个人打他，当然希望命中率不高的人剩下。所以 A, B 肯定互射，而最差的 C 被当作老弱病残保护起来。那么 C 是不是该打 A 呢？如果他打中 A，那么该 B 来打他。他知道有三人存在时 A, B 都不会来打他，打掉一人反倒对他不利。所以他的最佳策略是放空枪。等 A, B 相互之间干掉一人后轮他先打，不管命中率如何差，两人中先开枪总是划算的。这就是所谓鹬蚌相争，渔翁得利。

有了这个策略以后，算存活率就是很直接的概率题了。在 A 的命中率是 1 (100%) 的情况下，B 和 C 的命中率对每人的存活率的影响很不一样。为了求一个通式，我们假设 B 的命中率是 b，C 的命中率是 c。按题目假设，我们有 $1 > b > c > 0$ 。通过一些推导，我们可以得出 A，B，C 的存活率分别为：

A: $(1 - c)\times(1 - b)$

B: $b - b\times c/(b + c - b\times c)$

C: $c + b\times c\times(1/(b + c - b\times c) - 1)$

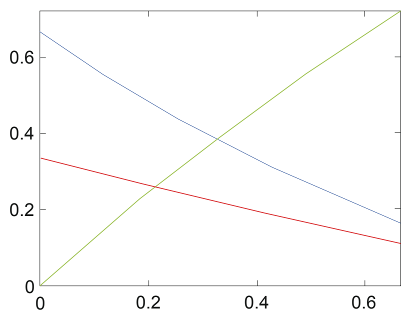
为了不把这篇文章变成数学论文，这个解的具体推导就留成作业好了。

我们最初叙述的这道题就是当 $b = 2/3$ 和 $c = 1/3$ 时的特例。在此情形下，有 A, B, C 的存活率分别是：2/9，8/21 ，25/63 。

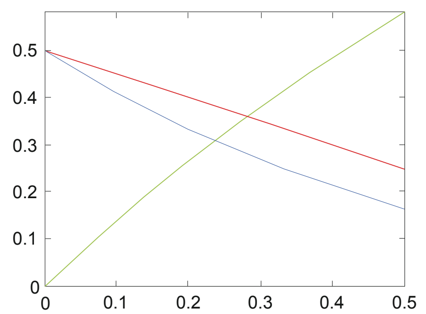
当然，这道题有趣的是在 b, c 取各种值所得的各种结果。我做了三个 A, B, C 存活率的图如下。下图中，b 分别为 2/3，1/2，1/3，横坐标是 C 的命中率，从 0 到相应的 b。Y 坐标是存活率。

可以看到在 $b = 2/3$ 时，虽然 A 的命中率最高，但他的存活率（红色）一直在 B 的存活率（蓝色）下面。甚至当 c 比 0.2 多一点以后，C 的存活率（绿色）也比 A 高。这个图告诉我们在制度不好的时候，优秀人物并不一定混得更好。所谓“枪打出头鸟”，“出头的椽子先烂”，“木秀于林，风必摧之”都是同一个机制。坏制度不能保护他们这些出头鸟。

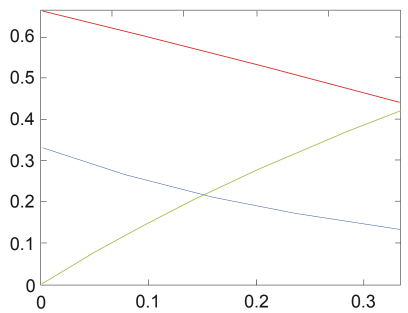
不过，要想比出头鸟混得更好，自己的本事也不能太差。当 $b = 1/2$（或以下）时，蓝线一直在红线之下。也就是说即使有制度保护，B 也永远不会比 A 混得更好。这就是通



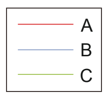
B 的命中率= 0.66667



B 的命中率= 0.5



B 的命中率= 0.33333



A 的存活率（红色）
B 的存活率（蓝色）
C 的存活率（绿色）

常所说的烂泥糊不上墙。阿斗当不好皇帝，虽然有刘皇叔托孤，诸葛亮撑腰。

三个图都有一个共性，那就是当 C 的命中率接近 B 的命中率一半以后，C 的存活率就比 B 还好。这也是一个常见现象，中等水平的人常吃亏。因为他们本事不够，自己上不去，又没有坏到需要制度照顾，最后的结果就是吃亏。美国这边现实的例子就是孩子上大学的学费问题。真正的富人是不在乎这点钱的，而收入不够的人可以申请资助，只有中产阶级，学费压力很大，却不能申请资助。

C 的存活率甚至有时候比 A 还高。不过，当 b 更小的时候（比如 1/3），红线就一直在蓝、绿之上了。这就是为什么许多统治阶层要搞愚民政策。下面的人水平太差以后，无论怎么钻空子（比如开空枪），上面的人都总是有优势。

受过数学训练的人读到这里，想要问的一个很自然的问题就是，什么时候 A, B, C 的存活率相等(都等于 1/3)。有了前面的公式，我们不难算出，当 $c=(5-\sqrt{7})/9$, $b=(\sqrt{7}-1)/3$ 时，A, B, C 的存活率都等于 1/3。（顺便说一下，如果找不到正确方法，要求出这个平衡点需要解一个四次方程。但如果找到正确方法，只需解一个二次方程就可以了，还是留成习题吧。）

这个平衡点表面看起来有点像三权分立，但这种表面上的相等其实很不公平。比 C 优秀差不多 4 倍的 A 在这个规则下得到的结果只不过与 C 相同而已。有点像计划经济体制下的平均主义。从前的大学生毕业，不管好坏一律都是 56 块半的工资。这种制度不能鼓励优秀人士，对社会的整体进步没有好处。

学佛的人常说一滴水珠看世界，所谓“滴水藏海”。我用这个三人决斗的趣味题目来看社会现象，搞笑之作，希望有人能欣赏。

通过计算实现正义

Justice by Computer

易延友

一、检察官突发奇想：传召数学家在刑事案件中作证

在美国证据法上，著名的人民诉科林斯（Peoples v. Collins）一案，国内研究证据法的大体上都耳熟能详。该案发生于 1964 年 6 月 18 日上午大约 11：30。一个老太太在买完东西后沿着一个小巷道回家。她左手提着篮子，装满了从商店购买的东西；右手拄着拐杖。突然从后面过来一名女子，把她推倒在地，抢走了篮子里的钱包，钱包里大约有 35-40 美元。她既没有看到这名女子，也没有听到她走近的声音。她感到一阵疼痛，但挣扎着往上看了一眼，看到一个年轻女子从她身边跑过。根据她事后描述，这名女子体重大约 145 磅，穿着黑色衣服，头发在深棕色和淡棕色之间。

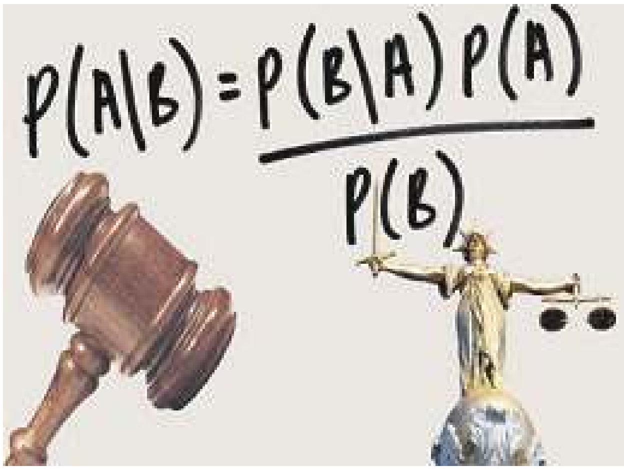
在事件发生的同时，住在该巷道附近的贝斯先生，正在自己的房子前浇草坪。因为听到尖叫声，他往事发方向看了看。他看到一名女子跑出巷道进入一辆停在巷道出口的黄色轿车。轿车马上发动并很快离开。他看到开车的是一个脸上有着络腮胡、嘴上还有小胡须的黑人男子；从巷道跑出的女子则应当是一名高加索裔女子，大约 5 英尺高，黑棕色头发，扎马尾辫。

警察根据证人提供的线索逮捕了珍妮和科林斯夫妇。珍妮的老板作证证明珍妮受雇于一个家政公司担任女佣，案发当天 11：30 左右他丈夫科林斯先生驾驶一辆黄色轿车把她接走，因此她有作案时间。但是珍妮自己作证说她下班后去了一位朋友家，在她家里呆了好几个小时，因此没有作案时间。珍妮的朋友也被传唤作证，证明她确实有一天在她家呆了好几个小时，但不确定是在哪一天。

法庭上，被害人并没有认出珍妮。证人贝斯的辨认也很勉强，因为贝斯在侦查阶段就曾经作过一次辨认，在那次辨认中贝斯没能认出科林斯先生。但贝斯的解释是那次科林斯先生没有蓄胡子。另外被告人提供的证人还证明珍妮在案发当天穿的是浅色衣服，而被害人却证明抢劫犯穿的是黑色衣服。因此控方的证据很不扎实。

不过，担任该案检控的检察官塞内塔是一位颇有创造性的青年才俊。案件发生时塞内塔刚从加州大学洛杉矶分校法学院毕业 2 年，担任长滩地区副检察长才 9 个月。按照他自己的说法，以概率论来证明自己的诉讼主张并不是一个深思熟虑的结果，而是基于一些偶然的因素。他接手该案件以后发现该案有很多瑕疵。尽管被害人很容易获得陪审团的同情，但是她却不能辨认出被告人；证人贝斯的证言也有很多漏洞；警察找到这对夫妇也是有很大的偶然性，因为实际上符合证人描述的夫妇很容易就能找到很多对。检察官左思右想，也想不出很好的对策来赢得他的诉讼。

不过，检察官家族中有一位数学天才，该数学天才在 1962 年以一本概率论方面的著作荣登畅销书榜首。想到这本畅销书，检察官来了灵感：他决定邀请长滩大学数学系的一位年轻教授—— 一个刚刚入职 2 个月的助理教授（Assistant Professor）出庭作证，证明该案中该被告夫妇被冤枉的可能性。该数学教授受到邀请，受宠若惊，觉得他有义务为社会服务，为公众效力，所以欣然接受了出庭作证的邀请。当然，等他来到法庭上时，看到黑压压的人群，庄严肃穆的法庭，他还是有些后悔接这份差事。不过他还是顺利地为检察官完成了这份工作。



检察官向数学家发问时说，假定在案发城市：

(1) 驾驶黄色汽车的人为当地驾车人口的 1/10，(2) 留着小胡须的人为当地男子的 1/4，(3) 梳着马尾辫的占当地妇女的 1/10，(4) 有着金色头发的女子占当地妇女的 1/3，(5) 是黑人且有络腮胡的占当地人口的 1/10，(6) 不同人种之间混婚成为夫妇的占全部婚姻关系的 1/1000；那么，本案被告人被冤枉的可能性是多少？数学家回答说，假定上述比例成立，则任何一对夫妇全部符合上述六个特征的可能性为一千二百万分之一（1/12,000,000），也就是说科林斯夫妇不是本案作案人从而被冤枉的可能性是一千二百万分之一。

不过，检察官在让数学家作证时，并没有提供任何证据证明上述比例在当地是成立的。但检察官坚持说，这些比例数字只是基于说明性的目的，并不表明这些比例就是真的；检察官又说，陪审员可以根据他们自己的经验，把上述比例替换成他们认为真实的比例，并根据自己相信的比例来计算任何一对夫妇完全符合上述特征的概率。

尽管被告人的辩护律师及时对上述数学家的专家证言提出了反对，但是法官却容许了上述证据。数学家作证之后，检察官总结陈词说：根据数学家的证言，科林斯夫妇被冤枉的可能性仅为一千二百万分之一；一千二百万分之一呀！——如果你们连这么一点风险都不愿意冒，那岂不是活的太累了？——于是年轻的检察官赢得了他的诉讼：陪审团做出了被告人有罪的裁决。

第二天，一家当地报纸就以头版头条报道了该案，标题就是：通过计算实现正义（*Justice by Computer*）。5 天之后，该案登上《洛杉矶时报》(*Los Angeles Times*)，标题是“正义召唤科学：概率论促夫妻获刑”（*Justice Invokes Science: Law of Probability Helps Convict Couple*）。一个月后该案就轰动全国（那时候还没有网络，更没有微博，一个月轰动全国已经不容易了）。美国著名的《时代》杂志声称：科林斯夫妇被定罪的原因就是检察官无比聪明地运用了统计学中的概率论；从此以后，我们进入了一个以全新手段检验间接证据的时代。

二、数学家担任法官助理：法院排除数学家的证言

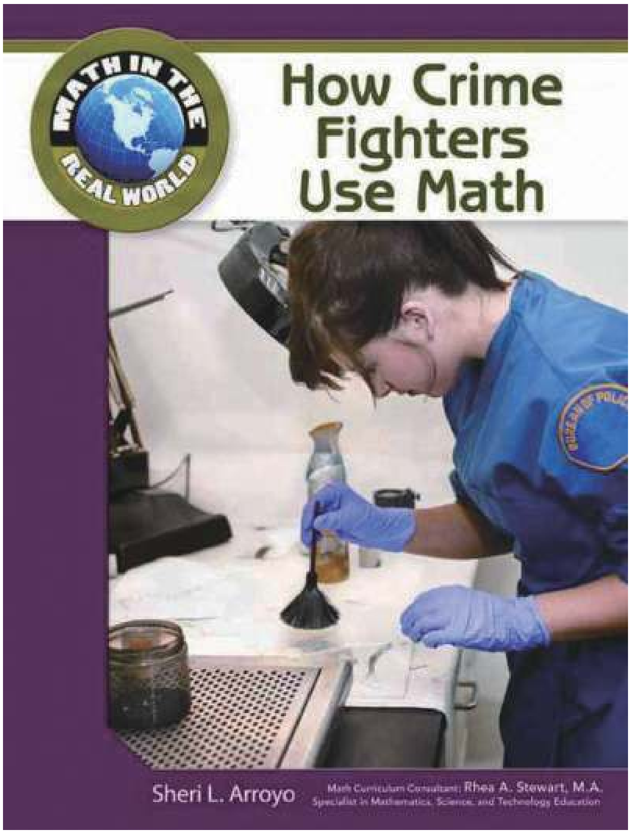
然而案件到此并未结束，检察官塞内塔遇到了自己的克星——一个真正的数学天才：劳伦斯•却伯。却伯于 15 岁入读哈佛大学数学系，并获得博士入读资格；但是却伯不愿意在数学这个领域挥洒自己的智慧和人生，最终放弃了数学博士候选人资格，改读哈佛法学院，并于 1966 年毕业。在科林斯案件于 1967 年上诉至加州最高法院时，却伯正在那里实习，担任马修•拖布雷纳法官的助理。此时，才华横溢的却伯年仅 25 岁。该上诉案主审法官沙利文刚好希望拖布雷纳法官帮助撰写该案的判词，而拖布雷纳则不失时机地将却伯纳入该案的处理程序当中。判词中的主体部分，实际上就是却伯的杰作。

加州最高法院根据却伯的理论，推翻了初审裁决，将该案发回重审。案件发回后，检察官坦率承认，他已经无法再召集原审的证人出庭作证。在此情况下，法官将该案予以撤销。被告人获得释放。

加州最高法院判决的理由主要是：

(1) 尽管数学家将其证言建立在 6 个假定事实的基础上（简要重复一下：黄色汽车 1/10；小胡须 1/4；马尾辫 1/10；金发女 1/3；黑人且络腮胡 1/10；混婚 1/1000），但是检察官没有举出任何证据证明上述比例真实存在。换句话说，没有任何证据证明当地驾驶黄色汽车的人有且只有 1/10；留小胡须的男人有且只有 1/4；梳马尾辫的女人有且只有 1/10……。这些比例完全是检察官自己的臆测。检察官自己也说，这些比例数据纯粹是说明性质的；陪审团成员可以根据自己的经验来替换这些数据。既然这些比例并不存在，或者没有证据证明其真实存在，以之为基础进行的任何进一步的论证都是徒劳的。

(2) 就算上述比例是真实存在的，数学家的计算也仍然存在问题：因为，没有证据表明上述假定事实均属于独立事件。小胡须和络腮胡是独立事件吗？马尾辫和金头发是独立事件吗？如果小胡须和络腮胡是独立事件，那黑人和络腮胡不也是相互独立的事件吗？为什么小胡须和络腮胡相互



独立，黑人和络腮胡就不相互独立了呢？如果没有证据表明上述比例均构成独立事件，拿这些比例来相乘并得到盖然性的结论就会夸大事实。

（3）即使假定上述比例均真实存在，并且均属于独立事件，也不能排除犯罪分子经过伪装的可能性。如果犯罪分子乔装打扮，把自己装扮成络腮胡、小胡须；甚至男扮女装，染上金色头发，梳上马尾辫，那么，检察官的论证将如空中楼阁，轰然坍塌。1200 万分之一的低概率匹配可能性，主要建立在 1/1000 的混合婚姻这一低概率事件的基础上。如果存在男扮女装的事实，或者存在其他伪装的情形，则即便科林斯夫妇完全符合检察官描述的低概率条件，也不能证明他们就是该案中的抢劫犯罪分子。因此，上述计算即便精密，看上去很有道理，但实际上是让人误入歧途的。

（4）哪怕完全不考虑以上三个因素，控诉方也仍然还有一个致命的错误，那就是将随机抽出的夫妇拥有这些特征的可能性与任何一对特定夫妇具有这些特征的可能性等同起来。如果每 1200 万人中有一对夫妇可能具备全部这些特征，那么，我们完全有理由希望在受怀疑的人群中找到两对这样的夫妇；如果这样的话，那么，对其中一对夫妇定罪，其被冤枉的可能性就是 50%，而不是 1200 万分之一。

以上是加州最高法院在科林斯一案中推翻其下级法院判决、不容许概率论证据在本案中运用的基本理由。值得说明一下的是，在本案中，匹配概率和被告人被冤枉的可能性实际上是等同的：在检察官的逻辑中，每 1200 万人中只有一对夫妇有可能完全符合六个条件，因此符合这些条件的这对科林斯夫妇被冤枉的可能性就是 1200 万分之一。但是检察官这一逻辑成立的前提是该城市只有 1200 万人，且只能找出一对完全符合上述特征的夫妇。但是，如果检察官夸大了事实，例如马尾辫的可能性不是 1/10 而是 1/5, 则每 600 万人中就有一对夫妇符合条件，那么在该城市就可以找出 2 对这样的夫妇；根据加州最高法院的逻辑，每一对夫妇被冤枉的可能性均为 50%。

三、法院判决引发学术风暴：数学家之间的较量

加州最高法院的判决引发了一场学术讨论的风暴。美国数学界群起反击，一批数学家和统计学家奋起捍卫他们学科的尊严。他们坚持认为所有的证明都是一种可能性 / 盖然性（概率）；通过概率方法来指导陪审团的思考既是必要的，也是可行的；这个社会需要运用科学方法来指导我们的思维。1970 年，一位统计学博士同时也是哈佛大学肯尼迪政府学院的助理教授和一位律师联合署名在《哈佛法律评论》发表论文，主张数学概率论可以通过贝叶斯定理在司法场域中获得转换适用；贝叶斯定理可以而且应当用来帮助陪审团评价证据。该文实际上是对加州最高法院在科林斯案件采取的立场的批评。

此时，却伯已经从实务界转向理论界，并在哈佛大学法学院担任助理教授。面对来自数学界发向自己曾经的得意之作的挑战，却伯毫不示弱，针锋相对，在 1971 年的《哈佛法律评论》发表论文，专门就加州最高法院对该案的判决做了详细地解释、辩护和发挥。却伯的主要观点是，利用数学计算的方式来运作司法审判以实现正义的方法其实并不是什么新鲜事，可以追溯到古代法定证据制度；根据该制度，每一种证据的证明力大小都由法律预先规定；但是这一制度却导致了邪恶的纠问式诉讼。却伯认为，尽管一个社会所信奉并用以使其审判制度理性化的这些精确的或假冒精确的装置有助于加强这个社会的风俗、文化的可理解性，但是，这些装置在其运行中却无论如何都会导致对法律审判行为所追求和表达的社会意义具有重大价值的东西予以歪曲，并且在某些场合下会毁灭这些价值。所以却伯不赞成以数学计

算的方法来判决案件，尤其是反对以数学的精确性来解释美国刑事司法程序中“排除合理怀疑”这一定罪标准的表述。

从那时起，美国学术界发表了一系列探讨概率论、数学计算方法、贝叶斯理论运用的论文，其中大部分是赞成数理逻辑尤其是概率论在司法案件中加以运用的，只是应当将其限定于适当的领域和不能超出适当的限度。在所有这些论文中，却伯教授是最早对概率论运用于司法提出反思和质疑的学者。

却伯的观点也有不少追随者。1985 年，查尔斯 • 尼桑在《哈佛法律评论》发表论文，强调司法裁决中对过去事实的认定并不是追求数学概率上的精确性，而是追求一种符合常识的可接受性。尼桑举了两个例子以说明其观点。第一个例子是在一个民事侵权行为案件中，甲乙二人均为原告 A 作证。甲作证说只有被告一人 B 特立独行地穿着蓝色西服出现在现场（暗示其他人均穿其他颜色衣服）；乙作证说实施侵权的是一个穿着蓝色西服的男子（暗示不可能是穿其他颜色衣服的人实施侵权行为）。如果甲乙两人的证言均真实可靠，则该案被告人 B 实施了侵权行为当无疑义。依照概率论，被告人实施了侵权行为的可能性取决于证人证言的可靠性（这种可靠性并不单纯建立在证人的道德水准和法制观念基础上，而且建立在证人的观察能力、记忆能力、交流能力等基础上）。假设陪审团认为每个证人的可信度均为 70%，那么，根据概率论，该被告人实施了侵权行为的可能性就低于 50%（两者相乘为 49%）。但是陪审团决不会单独计算各个证人的可信度，然后根据概率论的原理将两者相乘得出 B 实施侵权行为的可能性。相反，陪审团会将两名证人证言的可信度结合起来计算被告人作为现场侵权人身份成立的可能性。

第二个例子：假定一个事件的组成部分为 A 和 B；假定 A 发生的可能性为 70%，不发生的可能性为 30%；B 发生的可能性为 60%，B 不发生的可能性为 40%。根据概率论，A 和 B 都发生（A&B）（同时发生）的可能性为 42%，这一可能性远低于 A 和 B 不同时发生的可能性（100% − A&B = 58%）。但是，A&B 这一组合的可能性大于其他所有组合的可能性（A & NotB；NotA & B；NotA & NotB）。因此在上述概率情况下，陪审团实际上会认定 A&B 同时发生。

通过上述例证，尼桑想要说明，司法实践中关于证据所证明的事实可能性的计算方法并不是一个“更有可能或更无可能”的标准。也就是说，在司法这个场域，不能完全用概率论来解释。

当然，概率论究竟是否应当应用于司法领域，在什么条件下应用于司法领域，这恐怕是一个“猜想级”的问题，相信不断会有既精通数学又心仪法学的青年才俊用他们青春的火花点燃这个智慧殿堂的明灯。本文想说的却是，无论学术界关于概率论在司法场域的应用观点如何，均已经不能影响到当年才华横溢的却伯。如前所述，1968 年，却伯成为哈佛大学助理教授。1971 年，却伯在《哈佛法律评论》发表《通过数学来审判：法律程序中的精度与仪式》（*Trial by Mathematics: Precision and Ritual in the Legal Process*）一文，成为该领域中的先锋之作——当然，也是该领域引证率最高的作品。1972 年，年仅 31 岁的却伯成为哈佛大学教授。

就这样，一个检察官的突发奇想成就了一件名案，一个数学天才在法学领域的杰作成就了一代名家。

参考文献

1. People v. Collins, Supreme Court of California, 68 Cal. 2d 319 (1968).
2. George Fisher, People v. Collins: Historical Postscript, in George Fisher, EVIDENCE, Second Edition, Foundation Press, 2008, pp. 70-75.
3. Finkelstein & Fairley, A Bayesian Approach to Identification Evidence, 83 HARV. L. REV. 489 (1970).
4. Michael O. Finkelstein，William B. Fairley, The Continuing Debate over Mathematics In The Law of Evidence: A Comment on "Triai By Mathematics", 84 Harv. L. Rev. 1329 (1971).
5. Laurence Tribe, Trial by Mathematics: Precision and Ritual in the Legal Process, 84 Harv. L. Rev. 1810 (1971).
6. Laurence Tribe, A Further Critique of Mathematical Proof, 84 Harv. L. Rev. 1810 (1971).
7. Charles Nesson, The Evidence or The Event? On Judicial Proof and The Acceptability of Verdicts, Harvard Law Review 1357 (1985).

作者简介：易延友，清华大学教授，博士生导师，清华大学法学院证据法中心主任。

谷歌数学涂鸦赏析（中）[1]

欧阳顺湘

16 费马诞辰 410 周年

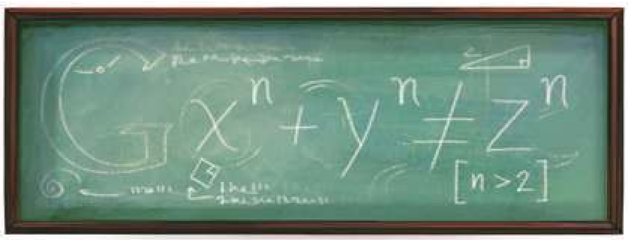


图 66 费马诞辰 410 周年纪念（2011 年 8 月 17 日，全球）

法国数学家皮埃尔 • 德 • 费马（Pierre de Fermat, 1601 年 8 月 17 日-1665 年 1 月 12 日）主职为律师，业余从事数学研究，被称为“业余数学家之王”。他在微积分的早期发展过程中起过重要的作用——研究过求切线和极大极小问题，是概率论和解析几何的开创者之一，还被誉为“近代数论之父”。

纪念费马诞辰 410 周年的谷歌涂鸦很有特色，为很多媒体和读者所留意。涂鸦以黑框灰绿色底的黑板为背景，上书费马大定理的内容：

$$x^n + y^n \neq z^n \ (n > 2).$$

即方程 $x^n + y^n = z^n$ 在 $n > 2$ 时无正整数解。

涂鸦右上角画了一个直角三角形。按勾股定理（即毕达哥拉斯定理），直角三角形两直角边的平方和等于斜边的平方。这提示我们，费马所考虑方程是当 $n = 2$ 时的方程 $x^2 + y^2 = z^2$ 的推广。与费马大定理对 $n > 2$ 情形的论断相反，这时方程 $x^2 + y^2 = z^2$ 有无穷多个整数解[2]。这等价于说有无穷多个整数组（常被称为勾股数组）可以构成一个直角三角形的三条边。

方程 $x^2 + y^2 = z^2$ 的特殊解的首次发现一般被当作各民族和国家数学水平发展的标志：

1. 公元前 19-前 15 世纪的巴比伦泥板中，就记载了 15 组勾股数，其中有 (119, 120, 169)，(3367, 3456, 4825)，(12709, 13500, 18541) 等。
2. 公元前 11 世纪我国《周髀算经》中记载有商高说的：“若勾三，股四，则弦五。”此说明商高已经知道边

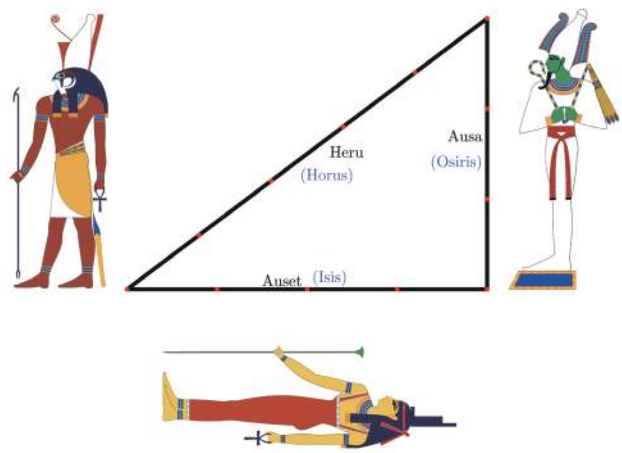


图 67 埃及三角形的象征

[1] 本文续同作者《谷歌数学涂鸦赏析（上）》，《数学文化》第 4 卷第 1 期，页 16-35，2013 年。

[2] 解的通用表达式、丢番图的《算术》以及费马大定理的介绍可参作者《亚历山大城的希帕蒂娅》（《数学文化》第 3 卷第 1 期，页 3–28，2012 年）一文。

图 68 费马

长为 3, 4, 5 的直角三角形。

3. 比例为 3:4:5 的直角三角形在古埃及有着神秘的色彩，常被称为埃及三角形。希腊作家普鲁塔克（Plutarchus，约 46 年-125 年）在他的《道德论集》（*Ethica*，亦作 *Moralia*）第 5 卷中首先描述了这样的三角形：短直角边被称为 Ausa，对应于作为起源的父亲欧西里斯（Osiris）；长直角边被称为 Auset，对应于接受者母亲伊西斯（Isis）；斜边被称为 Heru，对应于前述两者的完满结果：儿子荷鲁斯（Horus）。

费马大定理虽然冠以定理之名，但在 1995 年英国数学家安德鲁•怀尔斯（Andrew Wiles, 1953-）给出这个定理的完整证明之前，并无定理之实。在较早期的文献中，

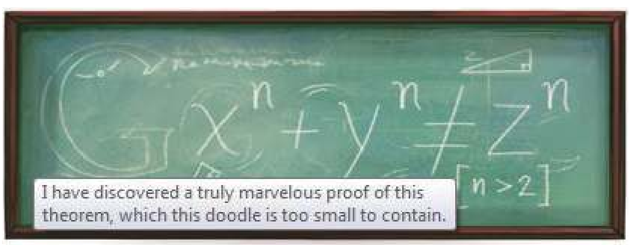


图 69 “费马的名言”

[3] 费马小定理：设 a 为整数，p 为素数，则 p 整除 $a^p - a$。读者可以考虑 $a = 2$，p 分别为 3, 5, 7 的例子：$2^3 - 2 = 2\times3$，$2^5 - 2 = 5\times6$，$2^7\times2 - 2 = 126 = 7\times18$。

它被直接称为费马猜想，后来因为费马的其它猜想或断言均被核实，且只有少数被否定，所以称之为费马最后的定理（Fermat's Last Theorem）。在中文里此定理被称为大定理，是相对于“费马小定理（Fermat's Little Theorem）[3]”而言的。

费马大定理原是费马在 1637 年左右阅读古希腊数学家丢番图的名著《算术》一书的译本时在页边空白的一个断言，当时他还在页边注释“我发现了一个美妙的证明，可惜由于空白太小而不能写下来”。如果你将鼠标移到这幅谷歌涂鸦上，可以读到这句名言的模仿：“我已发现了这个定理一个美妙的证明，可惜这个涂鸦太小而没法写下来。（I have discovered a truly marvelous proof of this theorem, which this doodle is too small to contain.）”怀尔斯的论文长达 100 多页，这篇论文以及他与其学生理查德•泰勒（Richard Taylor）合写的 19 页补充论文，发表在《数学年刊》（*Annals of Mathematics*）1995 年第 3 期整个一期，页边空白和涂鸦空白明显不够！

费马的许多数论结果都像这个论断一样没有给出过证明。不过，费马证明了 $x^4 + y^4 = z^4$ 没有正整数解。这也是现存费马在数论领域给出的仅有的两个证明之一。经过检验，人们发现费马提出的猜想和论断几乎都是对的。但他也“马”失前蹄过。费马知道形如 $2^m + 1$（m 为正整数）的数要为素数，m 必须等于 2 的非负整数幂。接着他检验了 $F_n := 2^{2^n} + 1$（n 为非负整数）的前 5 个数：$F_0 = 2^1 + 1 = 3$，$F_1 = 2^2 + 1 = 5$, $F_2 = 2^4 + 1 = 17$, $F_3 = 2^8 + 1 = 257$, $F_4 = 2^{16} + 1 = 65537$。发现他们都是素数。他猜测所有的 F_n，这些被后人称之为费马数的数，都是素数。但在约 100 年后的 1732 年，计算能力超强的欧拉发现：

$$F_5 = 2^{32} + 1 = 4294967297 = 641\times6700417.$$

这表明第 6 个费马数不是素数，否定了费马的猜测。至今还只知道前面 5 个费马数是素数。

在怀尔斯之前，数学家们已证明了费马猜想的许多特殊情形。在这个过程中催生了许多重要结果。相传希尔伯特不愿意去研究费马大定理的一个原因是不希望杀死这只会生金蛋的鹅。

怀尔斯的证明用到了代数数论中的椭圆曲线理论。陈省身曾在庆祝自然科学基金制设立 15 周年和国家自然科学基金委员会成立 10 周年的讲演《中国的数学——几件数学新闻和对于中国数学的一些看法》中说：“从这个定理我们应认识到：高深的数学是必要的。Fermat 定理的结论虽然简单，但它蕴藏着许多数学的关系，远远超出结论中的数学观念。这些关系日新月异，十分神妙，学问之奥，令人拜赏。”陈省身演讲中的另一段话也是很有借鉴意义的：“我相信，Fermat 定理不能用初等方法证明，这种努力是徒劳的。

数学是一个整体，一定要吸取几千年所有的进步。”

与费马大定理等许多丢番图方程密切相关的是约瑟夫 • 欧斯特勒（Joseph Oesterlé）及大卫 • 梅瑟（David Masser）在 1985 年提出的 *abc* 猜想，这个猜想一旦证明，将一举重新证明或解决包括费马大定理在内的许多问题。

abc 猜想的内容如下：设 a, b 为互质的正整数，$c = a + b$，则对任意 $\varepsilon > 0$，存在常数 $C_\varepsilon > 0$，使得

$$C_\varepsilon < \frac{(\mathrm{rad}(abc))^{1+\varepsilon}}{c}.$$

这里 $\mathrm{rad}(n)$ 表示 n 的所有不同素因子的乘积。例如，对 $n = 2^3 \times 5 \times 11^2$，有 $\mathrm{rad}(n) = 2 \times 5 \times 11 = 110$。

2012 年 8 月，日本京都大学数学家望月新一（Shinichi Mochizuki, 1969-）发布总约五百页的 4 篇文章试图提供 *abc* 猜想的证明。这立即引起了很大的关注，《自然》和《科学》杂志都做了报道。望月新一也用到了椭圆曲线理论，但他接着使用了他自创的全新数学理论。

望月新一的证明目前正由其他数学家检查。他曾经证明过艰难的定理。但数学史上不乏严肃的数学家公开宣称解决了大问题，而后发现有错的例子，如 2007 年，法国数学家吕西安 • 施皮罗（Lucien Szpiro）就曾宣布证明了 *abc* 猜想。

17 哈雷诞辰 355 周年

图 70 哈雷诞辰 355 周年纪念（2011 年 11 月 8 日，英国）

埃德蒙 • 哈雷（Edmond Halley, 1656 年 11 月 8 日-1742 年 1 月 14 日）是英国历史上一位在天文学、数学、物理学、金融和精算等方面做出过重要贡献的人物。他最出名的成就是我们熟知的：成功预言了现在被命名为哈雷彗星的周期（76 年）和椭圆轨道。谷歌在纪念哈雷诞辰 355 周年的涂鸦中就是用星空中的彗星来纪念哈雷的这一成就。1706 年，在他做出对哈雷彗星的预测之后一年，他学会了阿拉伯语，将阿波罗尼奥斯（Apollonius of Perga, 约前 262- 约前 190）的《圆锥曲线论》的第 5 至第 7 卷从阿拉伯文翻译为拉丁文，而且重建了佚失的第 8 卷。圆锥曲线与天体运行轨道密切相关，从此也可以看出哈雷的努力。

吸引公众的天文现象的出现往往是进行科学教育的重

图 71 哈雷

要契机。古代因为对天文现象的不解而常将彗星视为会带来霉运的“扫把星”。现在因为哈雷等人的贡献，我们可以正确认识彗星了。也因为这样一个相关的原因，哈雷彗星出现的时间成了激励美国进行科学教育的信号。在上一次哈雷彗星访问地球的 1985 年，美国提出 2061 计划，期望通过 76 年的努力，在哈雷彗星再次回来的 2061 年，所有能看见这颗彗星的美国人，都能够适应科学技术和社会生活的发展变化。为此目的，在 1985 年美国就需要着手进行科学、数学和技术教育等方面的改革。美国的 2061 计划也促使了中国在科学教育方面的反思，提出“2049 计划”，即全民科学素质行动计划，希望到 2049 年新中国建国 100 周年时，国民的科学素养有大的提高。

哈雷很早就做出了天文学重要贡献。1673 年，年仅 17 岁的他进入牛津大学王后学院学习。两年后他即开始协助皇家天文学家约翰 • 佛兰斯蒂德（John Flamsteed, 1646-1719）进行天文观测。一年后他更是放弃学业去南大西洋上的圣赫勒拿岛（St. Helena）工作了两年，编制了第一个南天星表，补充了当时天文学界只有的北天星表。因此工作，哈雷被选入皇家学会，当时他才 22 岁。哈雷在 1720 年还继任佛兰斯蒂德成为第二位英国皇家天文学家

哈雷对引力的兴趣以及天体运行轨道的疑问，使他在 1684 年 8 月到剑桥去访问牛顿，发现牛顿已得到了万有引力定律，哈雷于是说服牛顿发表该结果，并资助牛顿出版了《自然哲学的数学原理》。类似的故事在英国历史上也重现过：达尔文在赖尔的催促下才在 1859 年出版《物种起源》。实际上他的进化论思想早在 1844 年就已基本定型，并写好了初稿。

哈雷与另一位大数学家拉格朗日的“交往”却是时空相隔的。17 岁以前的拉格朗日喜欢文学，对数学并无感觉，很可能会在父亲的安排下成为一名律师。使得拉格朗日迷上数学的原因是他读到了哈雷所写的一篇介绍牛顿微积分思想的文章。

哈雷在人口统计学与保险方面也有影响深远的工作。他参与到这项工作也有一定的偶然性。

1662 年英国人格兰特（John Graunt, 1620-1674）已经发表了著名的人口统计工作《关于死亡表的自然的和政治的观察》。但他所使用的数据中伦敦人口死亡时的年龄没有记录。大约 30 年后，布雷斯劳（Breslau，即波兰的弗罗茨瓦夫 [Wrocław]）的牧师卡斯帕 • 诺依曼（Caspar Neumann, 1648-1715）发表了一份有关布雷斯劳人口的出生、死亡以及死亡时年龄的详细报告 *Reflexionen über Leben und Tod bey denen in Breslau Geborenen und Gestorbenen*。他将此报告寄给了莱布尼兹，而莱布尼兹转告了英国伦敦皇家学会。1691 年，诺依曼按要求将报告寄给了伦敦皇家学会的秘书。哈雷分析了这份报告，在 1693 年发表了一篇关于人寿保险的文章，改进了格兰特的死亡表并引进了死亡率的定义。

1762 年，世界上第一家保险公司——伦敦公平保险社（The Equitable）在英国成立。这家保险公司应用的技术就来源于哈雷的计算方法。现在很受追捧的职业“精算师”即源于这家公司。

18 华罗庚诞辰 101 周年

图 72 华罗庚诞辰 101 周年纪念（2011 年 11 月 12 日，中国大陆、香港）

2011 年 11 月 12 日是我国著名数学家华罗庚（1910 年 11 月 12 日-1985 年 6 月 12 日）诞辰 101 周年，谷歌以涂鸦进行纪念。涂鸦采用卡通形象的华罗庚一边泡茶一边思考“1 ＋ 1 ＝ ?”的漫画方式，突出了他的两项重要的贡献：一是以哥德巴赫猜想“1 ＋ 1 ＝ ？”所指示的数论研究（或说所指导学生的成就），二是用华罗庚提出的“泡壶茶喝”这个经典的时间安排例子来标志他的统筹学研究。更广一些，两者分别表示他在纯数学研究与应用数学推广方面的贡献。

泡茶喝的例子来自于华罗庚 1965 年 6 月 6 日在《人

图 73 中国科学院数学与系统科学研究院思源楼内的华罗庚雕像（作者摄于 2012 年 11 月 7 日）

民日报》上发表的《统筹方法平话》。读者对此应该不陌生，因为他的文章经节录、编写后入选了中学《语文》教材。华罗庚在文章的引子中先问：“比如，想泡壶茶喝。当时的情况是：开水没有。开水壶要洗，茶壶茶杯要洗；火已升了，茶叶也有了。怎么办？”然后比较甲、乙、丙三种办法：

办法甲：洗好开水壶，灌上凉水，放在火上；在等待水开的时候，洗茶壶、洗茶杯、拿茶叶；等水开了，泡茶喝。

办法乙：先做好一些准备工作，洗开水壶，洗壶杯，拿茶叶；一切就绪，灌水烧水；坐待水开了，泡茶喝。

办法丙：洗净开水壶，灌上凉水，放在火上；坐待水开，开了之后急急忙忙找茶叶，洗壶杯，泡茶喝。

“哪一种办法省时间？”就变得很明显：“谁都能一眼看出，第一种办法好，因为后二种办法都‘窝了工’。”

华罗庚就是从人人都熟悉的泡茶的例子开始，以通俗易懂的语言，深入浅出地说明了“统筹方法，是一种为生产建设服务的数学方法。它的应用范围极为广泛，在国防、在工业的生产管理中和复杂的科研项目的组织和管理中，皆可以应用”。

此后，华罗庚还撰写和出版了另一本科普读物《优选法平话及其补充》。“优选法”可用以改进生产工艺和提高质量。华罗庚组织的“双法”（“优选法”和“统筹法”）推广小分队，共去过 26 个省、自治区和直辖市，有很大的经济效益和社会效益。

华罗庚写作《平话》以及推广“双法”得到过毛泽东

图 74 毛泽东给华罗庚的第一封信

两次亲笔信的鼓励。1964 年初，华罗庚曾给毛泽东致信，表达了自己希望走出书斋，将数学应用于实际工作的愿望。同年 3 月 18 日毛泽东回信：“华罗庚先生：诗和信已经收读，壮志凌云，可喜可贺。肃此。 敬颂教祺！” 1965 年，华罗庚再次致信毛泽东，写了赴三线参加生产实践的感受，并将自己关于统筹方法平话的书寄了一本给毛泽东。1965 年 7 月 21 日，毛泽东再次复信：“华罗庚同志：来信及《平话》，早在外地收到。你现在奋发有为，不为个人，而为人民服务，十分欢迎。听说你到西南视察，并讲学，大有收获，极为庆幸。专此奉复。”

纪念华罗庚涂鸦上的“1 ＋ 1 ＝ ?”表示哥德巴赫猜想“1 ＋ 1 ＝ 2”，含义为任意一个大于或等于 6 的偶数都可以表示为两个奇素数之和，就如我们有 6 ＝ 3 ＋ 3，8 ＝ 3 ＋ 5 以及 100 ＝ 41 ＋ 59 等。这里的“1”和“2”分别用来象征奇素数和偶数。涂鸦中的“1 ＋ 1 ＝ ?”与前文中纪念陈景润的谷歌涂鸦上的“1 ＋ 2”相映成趣。“1 ＋ 2”指陈景润的结果，因此这里的“1”还是表示奇素数，“2”表示不超过两个奇素数的乘积。“1 ＋ 2”是目前对于“1 ＋ 1 ＝ ?”的最好的答案。华罗庚的学生群体对“1 ＋ 1 ＝ ?”的研究可以看做华罗庚对数学的一大贡献。

因陈景润刻苦钻研，勇摘科学皇冠上的明珠的故事而使得形象地描述哥德巴赫猜想的“1 ＋ 1 ＝ 2”以及陈氏定理的“1 ＋ 2”广为人知。然而有不少人对此望文生义地来理解，以为数学家闲得没事做或数学是如此高深莫测，简单的算术 1 ＋ 1 ＝ 2，1 ＋ 2 ＝ 3 都要研究。如此疑虑和牵强附会在公众乃至某些名人的言论中并不鲜见。这反映了数学普及的迫切性。

华罗庚自己做了很多数学普及工作，除了上述两本《平话》，还撰写了《从杨辉三角谈起》、《从祖冲之的圆周率谈起》等深受读者喜爱的普及图书。华罗庚作为一位在数学研究领域有国际地位的数学家在数学普及方面做出如此的努力，必定受到人们的尊敬。

关于华罗庚的传记，感兴趣的读者可以阅读王元著《华罗庚》等作品。

19 赫兹诞辰 155 周年

图 75 赫兹诞辰 155 周年纪念（2012 年 2 月 22 日，全球）

2012 年 2 月 22 日的谷歌涂鸦是波形曲线动画，曲线分段使用蓝、红、黄、绿四种谷歌颜色。这是纪念十九世纪最重要的物理学家之一，德国物理学家海因里希 • 赫兹（Heinrich Hertz, 1857 年 2 月 22 日-1894 年 1 月 1 日）诞辰 155 周年。

涂鸦采用波形曲线，是因为赫兹在 1887 年首先用实验证实了电磁波的存在，验证了麦克斯韦的电磁波理论。电磁波的发现使得我们现在使用电视、广播、无线网络和手机等成为可能。为纪念他，频率的国际单位制单位赫兹就是以他的名字命名的。

在慕尼黑大学求学期间，赫兹听从冯 • 约里（P. G. von Jolly）的建议，研读了拉格朗日、拉普拉斯以及泊松的著作，为日后的研究奠定了坚实的数学基础。他关于数学公式有名言：“人无法摆脱这样的感觉：这些数学公式是独立存在的，有它们自己的智能，它们比我们更聪明，甚至比它们的发现者聪明，因为我们从它们中得到的多于我们赋予它们的 [4]。”跟随

图 76 1994 年德国发行纪念赫兹逝世 100 周年的邮票

[4] 参考贝尔 (Eric T. Bell) 著《数学精英》（*Men of Mathematics, New York, 1937*）：One cannot escape the feeling that these mathematical formulas have an independent existence and an intelligence of their own, that they are wiser that we are, wiser even than their discoverers, that we get more out of them than was originally put into them。

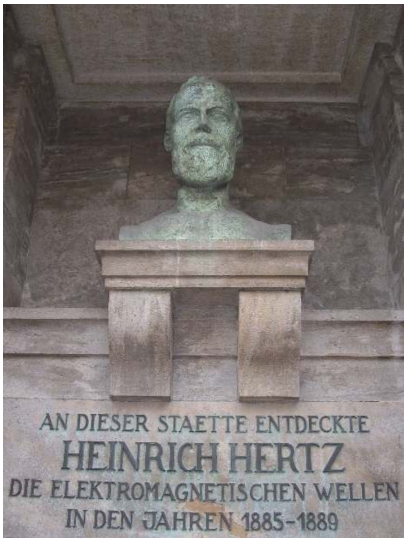


图 77 现德国卡尔斯鲁尔理工学院赫兹纪念雕像，铭文翻译：1885-1889 年，赫兹在这里发现了电磁波。(An dieser Stätte entdeckte Heinrich Hertz die elektromagnetischen Wellen in den Jahren 1885-1889)

基尔霍夫（Gustav Robert Kirchhoff, 1824-1887）以及亥姆赫兹（Hermann Ludwig Ferdinand von Helmholtz, 1821-1894）等大家的学习也使得赫兹的能力得到极大的提高。

虽然因败血症，赫兹年仅 36 岁就去世了。但他的成就是多方面的，除了上述电磁波实验，他还发现了光电效应，改写了麦克斯韦方程组等。我们下面介绍他在力学体系建立方面的贡献。

赫兹是既能够从事实验物理学，也能做理论物理的物理学家，富于哲学思想。他坚信物理学应该把自然现象归结为简单的力学定律，因此开始了基础力学研究。在他生命的最后三年，与疾病斗争，与时间赛跑，撰写了《力学原理》[5]。这本书在赫兹去世后不久出版，很受重视，读者可以在《西方数学中的里程碑式作品：1640-1940》[6] 中读到更多介绍。

传统的经典力学和分析力学体系分别以力和能量为核心概念。赫兹只用到时间、空间和质量这三个量，用类似于《几何原本》中的写作体系和严谨的推理，构建了一种“无力的力学”。亥姆霍兹在为《力学原理》写的序言中说：“未来将证明，本书对揭示自然力的新的、普遍特性有着重大启发价值。”事实上，赫兹的观念后来为爱因斯坦等所继承和发扬。在广义相对论中，万有引力这一自然界最基本的力就是用空间的弯曲来表示。

20 吉泽章诞辰 101 周年

图 78 吉泽章诞辰 101 周年纪念（2012 年 3 月 14 日，全球）

2012 年 3 月 14 日的谷歌涂鸦是用折纸作品表现的，这是为了纪念日本折纸艺术大师、现代折纸之父吉泽章（1911 年 3 月 14 日-2005 年 3 月 14 日）诞辰 101 周年（也恰好为逝世 7 周年）。

纸张是中国四大发明之一，折纸也源自中国，且是中国人的传统之一。作为娱乐，不少朋友少时可能都有过用纸张折飞机和小船经历。用折纸作品祭奠故人也是中国一个重要的传统。但折纸在日本得到了更加广泛的发展。因为吉泽章等的努力，折纸已然成为日本艺术，并被日本视为其国粹之一。折纸的英文词汇“Origami”也是由日文的“折り紙”音译而来。从后文可见，有着深厚折纸“群众基础”的日本，对折纸数学的研究也处于领先地位。

吉泽章出生于日本的栃木县，从 1938 年起开始从事折纸的研究与创作，发表了一系列折纸作品。特别是 1955 年，他在荷兰阿姆斯特丹举办了个人折纸展，在西方引起了很大的震动。吉泽章一生创造了 5 万多件富有创

图 79 吉泽章和他的折纸作品

[5] 第一版：*Die Prinzipien der Mechanik in neuem Zusammenhange dargestellt* (ed. P. Lenard), Leipzig: Barth, 1894。英译：*The principles of mechanics presented in a new form* (trans. D. E. Jones and J. J. Walleye), London: Macmillan, 1899。

[6] *Landmark Writings in Western Mathematics 1640-1940*, I. Grattan-Guinness 著，Elsevier B. V.，2005。

意与趣味的作品。他不但发明了新的折纸法——湿折法（将纸张变潮后再折），还与美国人山姆•兰德列特（Sam Randlett）发明了图解折纸的术语——吉泽章-兰德列特系统，这使得折纸有了自己统一的语言，因而令折纸艺术能够更好地传播。

现代折纸也已经变为一门科学，并且获得很广泛的应用：从日常生活可用到的可折叠不锈钢袋到太空中人造卫星大型太阳能电池板阵列。如同许多学问一样，数学是使折纸从一门游戏成为科学而获得应用的原理。我们介绍这个涂鸦即是因其背后有趣的折纸数学。

人们很早就已经注意到了折纸与几何的密切联系。最明显的是，通过不断将纸张对折，可以将纸张的一边 2 等分，4 等分，也可以从普通的长方形纸张得到正方形等。进一步利用几何原理，人们可以将正方形的一边等分为 3、5、7、9 等等份；还可以得到正三角形、正五边形、正六边形，甚至黄金矩形和白银矩形（长宽比分别为 $(1+\sqrt{5})/2$ 和 $1+\sqrt{2}$）等。

现已退休的日本生物学家芳贺和夫（Kazuo Haga）就是一位折纸数学方面著名的专家。他曾常在做实验的间隙，琢磨折纸。他得到了被人称为芳贺第一、第二和第三定理的折纸定理：用简单的折叠，给出了各种比例。芳贺和夫还与人合作，于 2008 年出版了一本很适合中学数学教学辅助的书《折纸学——通过折纸的数学探索》（*Origamics: Mathematical Explorations Through Paper Folding*），介绍了这些定理及其他相关结果。这本书没有一个折纸实物图，集中关注于几何对象。

作为例子，我们用图 80 来解释芳贺第二定理：设有正方形纸张 $ABCD$，先将纸张对折，分别得到 AD、BC 的中点 E、F。将三角形$\triangle EDC$ 沿 CE 将 D 点折到正方形内部的 D' 点。再沿 ED' 折，得到 AB 与 ED' 之延长线的交点 H。则 $HB=\frac{1}{3}AB$。这个折纸方法还产生了其它很多有趣的结论，如设 G 为 CD' 与 EF 的交点，则三角形$\triangle ED'G$ 有边长比为 3:4:5。

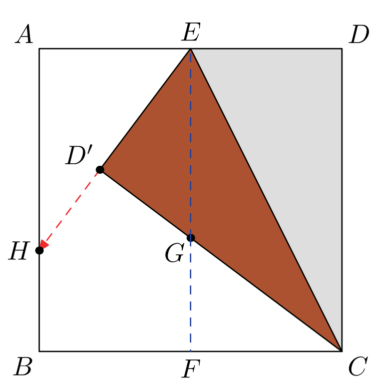


图 80 芳贺第二定理

类似于几何作图规则，折纸有自己的公理系统：

1. 给定两点 p_1 和 p_2，有且仅有一条折痕同时过这两点；
2. 给定两点 p_1 和 p_2，有且仅有一种方法把 p_1 折到 p_2 上；
3. 给定两直线 l_1 和 l_2，可以把 l_1 折到 l_2 上；
4. 给定一点 p 和一条直线 l，有且仅有一种方法过 p 折出 l 的垂线；
5. 给定两点 p_1, p_2 和一条直线 l，可以沿过 p_2 的直线将 p_1 折到 l 上；
6. 给定两点 p_1 和 p_2 以及两直线 l_1 和 l_2，可以一次将 p_1、p_2 分别折到 l_1、l_2 上；
7. 给定一点 p 和两直线 l_1 和 l_2，可以沿着 l_2 的垂线将 p 折到 l_1 上。

这些公理最早由雅克•贾斯汀（Jacques Justin）在 1989 年得到，后来被重新发现：前六条公理由藤田文章（Humiaki Huzita）在 1991 年提出，第七条公理由羽鸟公士郎（Koshiro Hatori）在 2001 年发现。罗伯特•朗（Robert Lang）也发现了第七条公理，并证明这七个公理是完备的。因此这个公理系统被称为：藤田－羽鸟公理或藤田－贾斯汀公理。

折纸公理中的操作很容易直观理解。比如，取一张正方形的纸片（如设为 $ABCD$），沿对角线（BD）将它对折，折成等腰直角三角形的操作就可以用来理解公理 1（过 B 与 D 只有一条折痕 BD）、公理 2（只有一种方法将 A 点折到 C 点）和公理 3（可以将直线 AB 折到 BC 上）等。

折纸公理也可以用数学的语言来解释。比如，公理 2 相当于找由两点 p_1 与 p_2 确定的线段的中垂线。也很自然地可将折纸与尺规作图做比较。前 5 条公理的操作一般最多有两种做法，在数学上相当于最多求解二次方程，与尺规作图的能力相当。例如，在公理 3 中，当两条直线平行时，自然只有一种作法；当两条直线相交，则相当于求角平分线，分别通过两对对顶角做角平分线可以得到两种作法。又如，公理 5 中的做法相当于求一条直线与圆的交点，这需要求解二次方程，一般有两个解。

但公理 6 中的折纸一般有三种方法，比尺规作图要强。它在数学上等价于求一条直线与两抛物线相切，涉及三次方程的求解。这样诸如三等分角、倍立方等本质为三次方程求解的经典尺规作图不能问题就可以用折纸来解决。图 81 所示为如何用折纸解决倍立方问题的例子，从中可以看到公理 6 的应用：设有正方形纸张 $ABCD$，（如通过芳贺第二定理折纸法）先得到 AD、BC 的三等分点 E、G 和 F、H。根据折纸公理 6，可以将点 B、F 分别折到 AD、GH 上成为 B' 和 F' 点。可以证明 $DB'/B'A=\sqrt[3]{2}$。

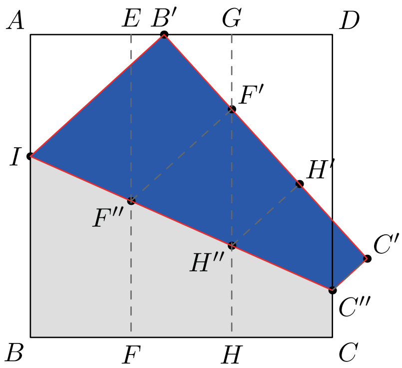


图 81 折纸求解倍立方问题

关于折纸数学，还有其它如折痕的研究等诸多方面，我们就不涉及了。有兴趣的读者可以参考许多著作以及专家罗伯特 • 朗的演讲 http://www.ted.com/talks/robert_lang_folds_way_new_origami.html 及其折纸主题网站 http://www.langorigami.com。实际上，纪念吉泽章的涂鸦折纸就是朗制作的，在涂鸦徽标的说明文档里（http://www.google.com/doodles/akira-yoshizawas-101st-birthday），朗除了介绍吉泽章，还提供了得到相关折纸作品的方法。

21 海亚姆诞辰 964 周年

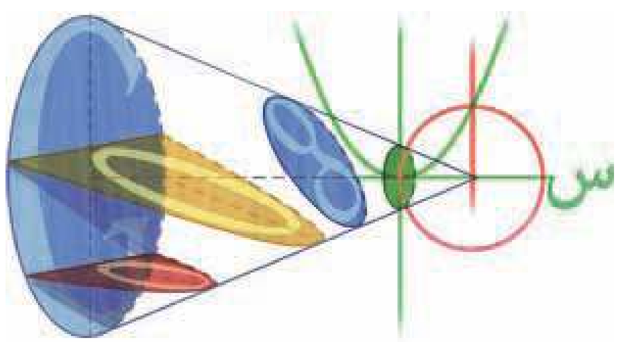


图 82 海亚姆诞辰 964 周年纪念（2012 年 5 月 18 日，阿联酋、巴林、阿尔及利亚、埃及、伊拉克、约旦、科威特、黎巴嫩、利比亚、摩洛哥、阿曼、巴勒斯坦、突尼斯、卡塔尔和沙特阿拉伯）

2012 年 5 月 18 日，谷歌在许多中东和北非国家的谷歌首页纪念古波斯诗人、数学家、天文学家和哲学家欧玛尔 • 海亚姆（Omar Khayyám，又译莪默 • 伽亚谟 [郭沫若用译名]

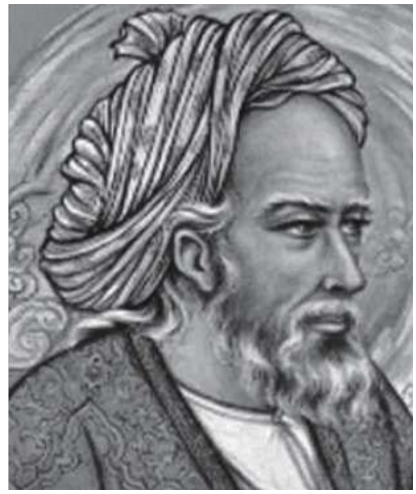


图 83 海亚姆

或峨默 [金庸用译名]，1048 年 5 月 18 日-1131 年 12 月 4 日）诞辰 964 周年。

海亚姆出生于今伊朗东北部霍拉桑（Khorrasan）省的内沙布尔（Neyshabur 或 Nishapur）。当时的伊朗被突厥人建立的塞尔柱帝国统治。海亚姆生活的时代，内沙布尔是重要的文化中心，世界上最大的城市之一。它也是古丝绸之路上的一个城市，至今仍有古代旅馆留存下来。历史上，内沙布尔也是一座充满悲伤的城市。1221 年，成吉思汗的蒙古军队西侵时，曾对内沙布尔进行屠城杀戮，相传死亡人数约 170 万人。

纪念海亚姆的谷歌涂鸦最右边的阿拉伯字符（或说波斯 - 伊朗字符）س[7] 用意就是揭示被纪念者所在地。伊朗人

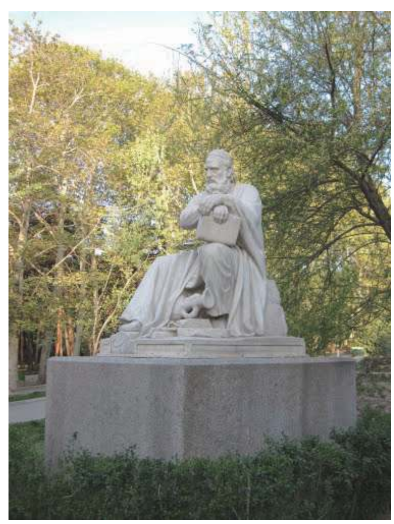


图 84 西班牙马德里康普斯顿大学里的海亚姆雕像

[7] 该字符按理应该用来显示“Google”商标中的字母“e”，但 س 对应的英文字母为“s”。有可能是谷歌失误，因为表示“e”的阿拉伯字符 ی 与 س 相似。也有可能谷歌设计者认为字母“e”已经得到表达，此字符仅是增加“波斯风味”。

图 85　电影《回到未来》(1985）中的情景，女主角所抱的书为 1937 年美国纽约出版的一本费兹哲罗译《鲁拜集》(装在书匣中，书匣表面图案如图中右下部所叠加的图片）

图 86 杜拉克为《鲁拜集》第 12 首（实为费兹哲罗英译第 1 版第 11 首）配的插画

对欧玛尔很推崇，他们把 5 月 18 日定为海亚姆日，每一年的这一天都会在内沙布尔举行纪念活动。

2011 年的 5 月 18 日，远在西班牙马德里康普斯顿大学（Complutense University of Madrid，西班牙文 Universidad Complutense de Madrid）在校园里树立了一座海亚姆的雕像来纪念他。另外，在伊朗的首都德黑兰和罗马尼亚的首都布加勒斯特也都有他的雕像。

海亚姆主要以其诗歌倾倒众人。他善写一种称为鲁拜（Rubai）的四行诗。鲁拜类似于中国的绝句，一首四行，第一、第二和第四行押韵，第三行大抵不押韵。

海亚姆生活在诗的时代。此前在他的家乡霍拉桑省有史称波斯文学之父的诗人鲁达基（Rudaki, Abu Abdollah Ja1far, 850-941）以及著名诗人菲尔多西（Ferdowsi, 约 940-1020）。菲尔多西的代表作《列王纪》是一部波斯民族的英雄史诗巨著。

海亚姆的诗歌在几百年里几乎被人遗忘。直到 1859 年英国的爱德华•费兹哲罗（Edward FitzGerald, 1809-1883）将他的诗从波斯文翻译成英文时才引起极大的关注，现在这个英译已经成为了英国文学的经典。董桥在《画 < 鲁拜集 > 的人》中赞道："这部书饮誉西方文坛一百五十多年，靠的真是菲茨杰罗[8]的英文译本了。都说译文是借句发挥不是依句翻译，海亚姆笔端飘下一片落叶，菲茨杰罗的稿纸上瞬间是满山的秋色。"

此后便有几乎数不清的译本出现。单说中译就有几十种之多。郭沫若、胡适、闻一多、徐志摩和朱湘等名家都翻译过，张鸿年还从波斯文直接翻译。特别是译者黄克孙（1928-），他是一位颇有成就的美籍华人物理学家，麻省理工学院的教授，但在中文世界，他最有名的或许还是他用七言诗的形式将费兹哲罗的英译《鲁拜集》译成中文。

图 87 1967 年迪拜发行的纪念海亚姆的邮票

张承志在《波斯的礼物》里写到："在他们（老读书人）看来，这一西域怪杰，完全可以与整个的波斯文明相匹敌。确实，这位风流诗人的绝妙'鲁拜'，引得中国人译者如蜂，兴而不衰。……这些胡姬当炉的妙歌，它挑逗了中国文人的渴望和趣味，教导了他们个性解放的极致。文人们出于惊喜，争相一译，寄托自由的悲愿。……——于是译笔缤纷，华章比美。"

绿叶衬牡丹，好诗少不了名家的配图。著名的有法国画家杜拉克（Edmund Dulac, 1882-1953）所作的彩色插画。

[8] 即费兹哲罗。

董桥曾赞他的插图“随时翻阅都翻得出一缕古东方的神秘情调”。1967 年，迪拜还以海亚姆的诗歌配图为内容，发行了一套纪念海亚姆的小全张。

爱屋及乌，对诗的喜爱也吸引了很多爱书之人对书的形式的喜爱。董桥不但写了《画 < 鲁拜集 > 的人》，还写了《我集藏的 < 鲁拜集 >》、《彩翎之恋》，谈《鲁拜集》的收藏，以及其插图、装帧、版本乃至书商和轶事。尤其是三篇文章中董桥都念念不忘 1908 年英国书籍装帧名家桑格斯基（Francis Sangorski）所做的一部“封皮上镶了各种宝石缀成的图案，一派东方贵胄气息”的孔雀装《鲁拜集》。这部书几经波折，却不幸在 1912 年随泰坦尼克号邮轮运去美国的途中海葬了。

郭沫若说，在海亚姆的诗里，可以找出李太白的面目来。我们来读读费兹哲罗《鲁拜集》（*Rubaiyat of Omar Khayyam*，Rubaiyat 是 Rubai 的复数）中的第 12 首[9]。这首诗可能是《鲁拜集》中引用最广泛的。

郭沫若译[10]：

树荫下放着一卷诗章，
一瓶葡萄美酒，一点干粮，
有你在这荒原中傍我欢歌——
荒原呀，啊，便是天堂！

接下来我们通过金庸在他的《倚天屠龙记》中所引用

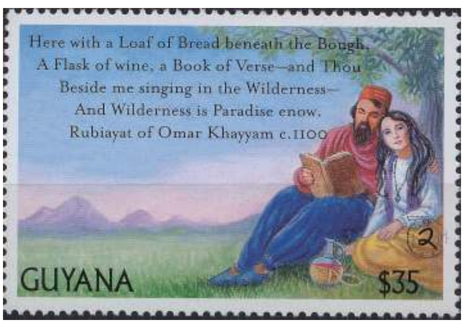


图 88 圭亚那发行纪念海亚姆的邮票（邮票上的英文诗为费兹哲罗英译《鲁拜集》第 1 版第 11 首，即常说的第 12 首的英译）

海亚姆的诗歌与传说来进一步认识《鲁拜集》诗歌中所反映的海亚姆对生命无常的感叹以及传说所折射的他淡泊名利专心学问的态度。我们知道金庸的小说除了恢弘的历史背景，荡气回肠的侠骨柔情，跌宕起伏的情节，也常了无痕迹地寓历史典故、诗词歌赋、琴棋书画、儒释道医、奇门术数于刀光剑影之中。后文在介绍纪念拉马努金的涂鸦中将提到金庸小说里的术数（数学）。

《倚天屠龙记》中小昭、殷离常哼唱令听者倍感悲凉的中土曲子：“来如流水兮逝如风，不知何处来兮何所终！”就是化用了海亚姆的诗的语言和意境。请看郭沫若翻译《鲁拜集》中的第 28 首

我也学播了智慧之种，
亲手培植它渐渐葱茏；
而今我所收获的收成——
只是“来如流水，逝如风”。

而小说中直接使用的“飘飘入世”、“飘飘出世”等词与郭译第 29 首有直接的联系：

飘飘入世，如水之不得不流，
不知何故来，也不知来自何处；
飘飘出世，如风之不得不吹，
风过漠地，又不知吹向何许。

《倚天屠龙记》第三十回中还借用了波斯大哲野芒门下三个杰出的弟子“山中老人”霍山、诗人峨默（即海亚姆）与政治家尼若牟（即尼扎姆 • 穆勒克）的故事。传说三人同窗求学时，曾相约未来若发达，当共享富贵。尼若牟做了首相后，给了霍山官职，但霍山不满足，成立了恐怖组织依斯美良派，专事行刺。后来更遣人刺杀波斯首相尼若牟。而峨默不愿居官，只求一笔年金，以便静居研习天文历数，饮酒吟诗。类似的故事也可以在郭沫若译《鲁拜集》“小引”中找到。这个传说其实不会真实发生，因为三人生活的时代有差异，但故事流传甚广，反映了人们对三人品格的看法。

以海亚姆传奇为主题的 1957 年美国好莱坞电影《欧玛尔 • 海亚姆》（*Omar Khayyam*）和 2005 年美国电影《继承者：欧玛尔 • 海亚姆传奇》（*The Keeper: The Legend of Omar Khayyam*）[11] 都有上述相关情节。金庸小说中更是加以演绎——明教中人的武功出自山中老人。依斯美良派在历史上还与另一位穆斯林大数学家纳西尔丁 • 图西有关，我们将另文讲述他的故事。

海亚姆在数学方面也有杰出的贡献。与其诗歌成就离不开诗的传统一样，他的数学成就的背后是穆斯林数学的整体辉煌。

数学是文明之镜。穆斯林创造了璀璨的文明，在数学

[9] 这里的第 12 首是按通常习惯理解，实际上费兹哲罗的英译有五个版本，第 12 首在第一版中为第 11 首。

[10] 黄克孙则将同一首诗译成七言诗，别有味道：一箪蔬食一壶浆，一卷诗书树下凉。卿为阿侬歌瀚海，茫茫瀚海即天堂。

[11] 故事以历史与现实双线交叉展开：一方面是叙述美国休斯顿 12 岁男孩卡尔曼作为海亚姆的后代，在其家族传承海亚姆故事的“继承者”哥哥去世之后所进行的寻根之旅，另一方面是对海亚姆的传奇经历的回顾。片中还有卡尔曼以孔雀装《鲁拜集》的封面为引子，找到桑格斯基的继承人，见到一部《鲁拜集》的情节。

方面也有着很重要的贡献。穆斯林数学最重要的贡献不仅是翻译继承了希腊数学的传统，使得度过漫长黑暗时代的欧洲人通过十字军东征发现了阿拉伯学者的著作，获得了欧几里得等的著作。作者在《亚历山大城的希帕蒂娅》一文中介绍了穆斯林教徒在六世纪对亚历山大图书馆、希腊数学的毁灭性打击。稳定后的穆斯林世界也开始重视学术研究，在巴格达就建有堪与亚历山大图书馆媲美的智慧宫，专事翻译与学术研究。希腊数学的重要著作如欧几里得的《几何原本》以及托勒密的《天文学大成》，以及印度人的著作等也陆续被翻译成了阿拉伯文。穆斯林数学家学习希腊、印度数学和天文学，并加以评注和改进。他们在代数学、三角学、平面几何与球面几何等方面在希腊数学家的基础上有系统的发展，而且开始考虑几何与代数的关系。我们使用的阿拉伯数字和十进制，就是阿拉伯人从印度引入并传播开来的，而代数学这一名称则直接来源于阿拉伯文。

海亚姆曾研究过在数学史上有重要意义的欧氏几何第五公设问题，但他最主要的著作是《代数问题的论证》（*Treatise on Demonstration of Problems of Algebra*）。在此书中海亚姆研究了开方术，如得到了 44 240 899 506 176 的 5 次方根 536！他也得到了二项式展开公式，即我们熟知的杨辉三角（或贾宪三角）；因此在伊朗也常称该三角为海亚姆 - 帕斯卡三角。其中最重要的成果是第三章中关于三次方程的求解。他的前辈花剌子密系统地解决了一次方程及一元二次方程求解问题。而三次方程的一般解法直到 1545 年才由意大利数学家卡尔达诺在其《大术》中写下。海亚姆将三次方程进行分类并对一些特殊的三次方程给出几何解法，即用两条圆锥曲线的交点来确定方程的根。

谷歌纪念海亚姆的涂鸦显示的就是海亚姆在求解三次方程方面的数学成就。这个涂鸦主要有两部分组成：涂鸦左边是圆锥和各种圆锥截面，并显示了椭圆和抛物线等圆锥曲线；右边是一个抛物线和圆相交的示意图。

在中学学过解析几何的读者知道，解析几何的特点是

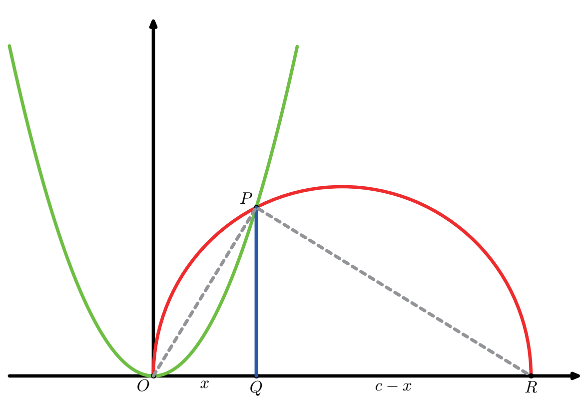


图 89 海亚姆求解三次方程的几何解释

用代数的方法求解几何问题，比如求解两条曲线的交点，只需要求解两曲线方程的联立方程就可以了。但对阿拉伯人来说，代数问题可以且必须用几何来解释。

海亚姆曾求解如下形式的三次方程

$$x^3 + b^2x = b^2c,$$

其中 $b, c > 0$。为方便解释，我们给出了笛卡尔发明的坐标系，其实利用圆锥曲线的性质，这是不必要的。海亚姆的方法如下：作抛物线 $by = x^2$ 以及中心在 $(c/2, 0)$、直径为 c 的圆（参考图 89 并比较涂鸦上的示意图）。设两曲线交点为 P，且 PQ 垂直横坐标，则 OQ 的长就是方程的解。推理不难。设 $OQ = x, PQ = y$。因为三角形△OPR 为直角三角形，易知两直角三角形△OPQ 与△PQR 相似。所以 $y^2 = x(c - x)$。同时又有 $by = x^2$。消去未知量 y 可得 $x^3 = b^2(c - x)$。因此 $x = OQ$ 为方程的解。

三次方程的圆锥曲线解法也可以追溯至古希腊时代。门奈赫莫斯（Menaechmus，约公元前 360 年）在研究倍立方问题时发现了圆锥曲线。倍立方问题是指用尺规作图构造一个两倍于原立方体体积的立方体。从代数学的角度看，这需要构造立方根 $\sqrt[3]{2}$。后人证明这是一个超越数，从而否定了尺规作图的可能性。阿基米德在《论球与圆柱》中的平面截球问题也可归结于三次方程问题。

方程之解的几何构造在十七世纪时吸引了许多数学家的注意。特别是笛卡尔找到了用圆与抛物线相交来求解所有三阶和四阶方程的唯一构造。

1070 年海亚姆向北到古丝绸之路上的名城达撒马尔罕（现为乌兹别克斯坦第二大城市，撒马尔罕州首府）从事数学研究，后赴伊斯法罕进行天文历法改革[12]。海亚姆还在诗歌中记述过他在历法方面的经历：“啊，人们说我的推算高明，我曾把旧历的岁时改正——，谁知道那只是从历书之中，消去未生的明日和已死的昨晨。[13]”

海亚姆的这一姓氏（英文翻译为 Tent-maker）有各种解释。一般认为其意思是帐篷制造者或支架帐篷的人，可能表示其父辈的职业；郭沫若则认为这可能是诗人的雅号。海亚姆在家乡求学后游历四方，终身未婚，晚年回到家乡直至去世。为了纪念海亚姆，1934 年，由多国集资，在他的故乡内沙布尔修建了一座高大的陵墓——一座用八块菱形板围成的圆锥状中空建筑，棱形内部镶嵌着伊斯兰的美丽花纹。这大约是受到海亚姆生前预言的启发：“我的坟墓所在的地方，北风会吹玫瑰花来覆盖。”

[12] 伊朗现行伊朗历比世界通用公历要精确。
[13] 郭沫若译《鲁拜集》第 57 首。

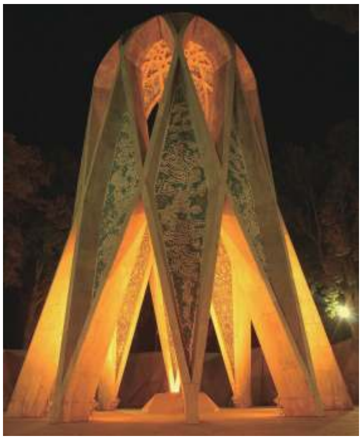


图 90 海亚姆陵墓

(a) (b) (c)

图 91 纪念海亚姆的邮票：(a) 密克罗尼西亚联邦发行；(b) 和 (c) 为阿尔巴尼亚发行

历史上，在文学方面有声誉的数学家（或天文学家）不乏其人，如中国的张衡，法国的帕斯卡、柯西，俄罗斯的柯瓦列夫斯卡娅等。英国的数学家罗素还在 1950 年以《婚姻与道德》获得诺贝尔文学奖。但同时在诗歌、天文学、数学这几个领域如海亚姆一样有杰出成就，而且影响深远者，很难找到另外的人。

关于诗歌与数学的联系，德国数学家魏尔斯特拉斯曾说："一个数学家，如果没有一点诗人的气质，不会是一个完满的数学家 [14]。" 英国 BBC 制作的四集电视节目《数学的故事》之《东方的天才》在介绍海亚姆后评论："学科之间有很大的相似性。诗歌，研究结构和节奏韵律，与构建数学逻辑证明有强烈的共鸣。" 这大约是为什么他能在这些方面取得非凡成就的一个内在原因吧。

1997 年，阿尔巴尼亚发行了两张纪念海亚姆的邮票（邮票上写着为纪念海亚姆诞辰 850 周年，可能有误）。

[14] 德语为 "Ein Mathematiker, der nicht auch etwas Poet ist, wird nie ein vollkommener Mathematiker sein."

[15] 参考谷歌徽标档案网页上该游戏网页：http://www.google.com/doodles/alan-turings-100th-birthday。

可以参考的资料有英国 BBC 制作的纪录片《天才的欧玛尔 • 海亚姆》(*The Genius Of Omar Khayyam*)。也推荐阅读当代诗人与数学家蔡天新教授所著《难以企及的人物——数学天空的群星闪耀》中的《欧玛尔 • 海亚姆的世界》。

22 图灵诞辰 100 周年

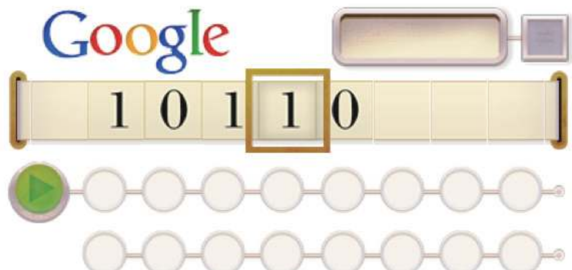


图 92 图灵诞辰 100 周年纪念（2012 年 6 月 23 日，全球；图为图灵机游戏起始页面的计数程序（当前显示的二进制数为 10110，即十进制表示的 22）

2012 年的 6 月 23 日是英国数学家、逻辑学家与计算机科学之父阿兰 • 图灵（Alan Mathison Turing, 1912 年 6 月 23 日 -1954 年 6 月 7 日）诞辰 100 周年的纪念日。谷歌涂鸦以一个交互游戏 [15] 的方式展示了图灵在计算机实现之前提出的抽象计算模型——图灵机（Turing Machine）。

谷歌涂鸦设计者想利用图灵诞辰 100 周年的机会，设计一个好的涂鸦向他们心中的英雄致敬。但图灵在很多方面都有杰出的贡献，而且图灵的许多工作抽象而难以用涂鸦的方式表现使得公众理解。这对谷歌涂鸦设计者是一个挑战。

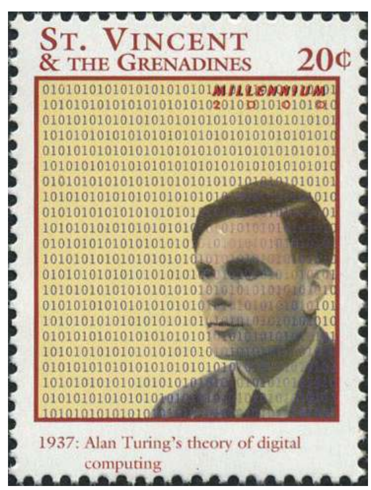
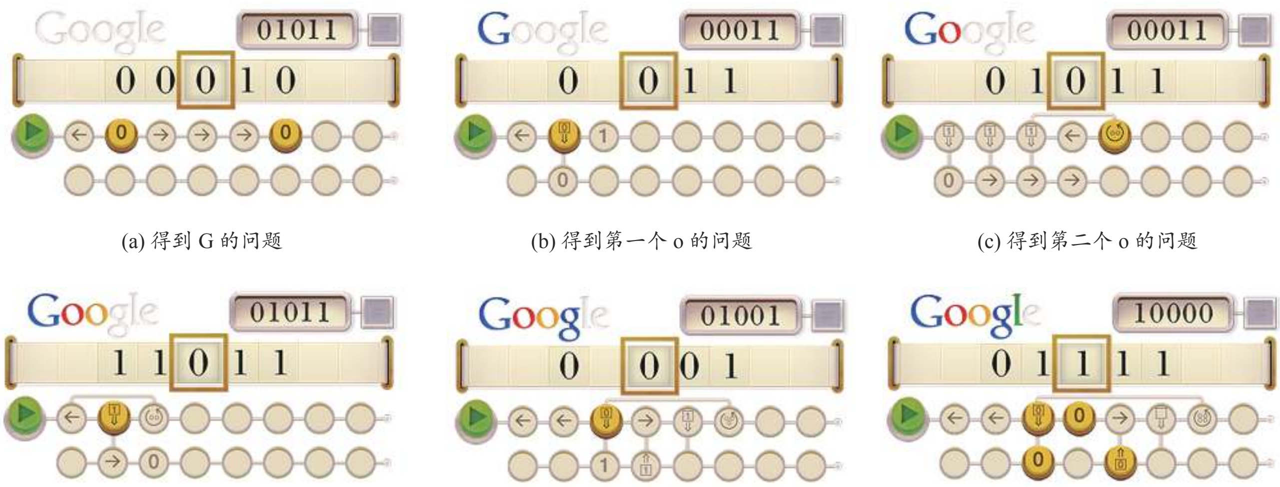


图 93 圣文森特和格林纳丁斯 2000 年发行的千禧年纪念（1900-1950 部分）邮票系列中的一张纪念图灵的邮票，背景为构筑计算机世界基础的 0 和 1



(a) 得到 G 的问题　(b) 得到第一个 o 的问题　(c) 得到第二个 o 的问题

(d) 得到 g 的问题　(e) 得到 l 的问题　(f) 得到 e 的问题

图 94　图灵机游戏第一轮的问题

他们在决定使用这个图灵机游戏方案前，深入研究了图灵的工作，做了许多尝试。

图灵机是图灵在 1936 年，年仅 24 岁时写的一篇论文《论可计算数在判定问题中的应用》（*On computer number with an application to the Entscheidungsproblem*）中提出来的。“Entscheidungsproblem（判定问题的德文表达）”是希尔伯特在 1928 年提出来的三个问题之一。留着德语烙印的问题是当时的热点研究问题，然而，图灵解决问题所描述的图灵机以及通用图灵机却更加要紧，它们奠定了现代计算机发展的理论基础。冯 • 诺伊曼曾说：“如果不考虑巴贝奇等人早先提出的有关思想，现代计算机的概念当属阿兰 • 图灵。”

谷歌纪念图灵的涂鸦可以让我们很好地了解图灵机的原理。本文下部分附录中介绍巴贝奇的相关思想时我们再简单提及图灵机更加形式的描述，这里仅介绍游戏。游戏还使用了循环和条件语句等程序设计思想，是一个很好的智力游戏，充分体现了谷歌涂鸦设计者的水平。

游戏的起始页面是一个用图灵机进行计数的程序：用二进制码表示，从 1 开始，逐次加 1，直至用户点击涂鸦开始游戏第一关。有兴趣的读者可以考虑，在参考下文介绍的游戏指令后，作为练习写出该程序的指令。

游戏的目的是通过正确算法依次得到以 0、1 编码表示的“Google”这六个字母。在游戏中表现为依次点亮左上角的这六个字母，每个字母一个小游戏。实际游戏从简单到复杂，共两轮 12 个小游戏。为方便读者，我们用图 94 给出了第一轮游戏的 6 个问题 [16]。字母编码采用了博多码（Baudot code），共 5 位 0、1 编码。字母 g, o, l, e 对应的博多码（不分大小写）分别为 01011, 00011, 01001, 10000。涂鸦中相应字母的正确编码放在右上方。

图中央是一条被分成多个小方格的纸带，纸带上有待用指令进行修改的二进制码或待填充编码的空位。纸带上有一个方框显示当前指令作用的位置。纸带下方为指令表，用来移动方框（在游戏中表现为纸带的左右移动）或修改其中的值。游戏中已给出一组指令，深色的指令可以通过点击来改选为合适的指令，而灰色的指令不能被修改。当将指令设好后，点击绿色按钮就开始运行指令得到一个修改后的编码。然后游戏逐一检查所得纸带上的各位编码是否与目标值相符（在右上角目标编码后的方框显示等号或不等号），若得到正确答案，即可进行下一关。

指令的作用可分如下几类：

1. 纸带移动：（纸带左移一格，相当于方框右移一格），（纸带右移一格，相当于方框左移一格）。
2. 修改纸带上的值：0 和 1 分别表示在纸带当前空格位置写入 0 和 1。
3. 指令一般从左至右执行：若遇 ，指令继续向右执行，但若遇下列条件，则分别转向或跳转：
   (a) 条件判断指令：这样的指令带有 0 或 1 或空格三种指示值，有上、下箭头两种方向。若方框所

[16] 六个问题的参考解答分别为：1. 将两个 0 均改为 1；2. 将指示值为 1 的条件指令改为指示值为空格的条件指令；3. 将跳转指令连到最左的条件指令，即连线跨越四个指令；4. 同 2；5. 同 2；6. 修改第二行的两个指令：将指示值为 0 的条件指令改为指示值 1 的条件指令，同时将 0 改为 1。

在纸带上位置的值与指令的指示值相符，则执行箭头所指向的指令。例如，表示如果方框所在纸带上位置的值为 1，则执行箭头下方的命令。

(b) 跳转指令：、和等表示指令跳转至连线所指向的指令。弯箭头所包含的小圈数目表示指令最多可以跨越的指令数。

作为例子，我们解释如何用指令“← 1 →→→ 0（这里表示的简写）”将初始位置位于第三个 0 的编码“00$\boxed{0}$10”修改为表示 G 的正确编码“01011”。这里用方框将某位置围住表示当前指令位于该处。对“00$\boxed{0}$10”，依次执行“← 1 →→→ 0”这六个指令：

1. 执行“←”，指令位置左移一格，得到：0$\boxed{0}$010；
2. 执行“1”，当前位置值修改为 1，得到：0$\boxed{1}$010；
3. 执行“→”，指令位置右移一格，得到：01$\boxed{0}$10；
4. 执行“→”，指令位置右移一格，得到：010$\boxed{1}$0；
5. 执行“→”，指令位置右移一格，得到：0101$\boxed{0}$;
6. 执行“1”，当前位置值修改为 1，得到：0101$\boxed{1}$。

每一轮的六个游戏结束后，谷歌涂鸦会重新自动演示一遍全部六个小游戏是如何通过指令逐步改变纸带上的值的。在重演过程中，涂鸦右上角的小方框中会闪烁一只小兔子的图像。如果点击进去，与起始的计数程序类似，涂鸦会自动演示一个更加复杂的程序：用图灵机计算一个用 0、1 编码表示的数列。涂鸦还提供了三行指令，每行 13 个指令，一共 39 个指令。这个程序所计算的序列 1, 110,11010, 110101, 11010110, 110101101,…… 与用 0、1 编码表示的斐波那契兔子序列紧密相关。

图灵的另一重要的贡献是协助军方破译了德国著名密码系统恩尼格玛（ENIGMA），加快了二战的结束。

恩尼格玛是德国人谢尔比乌斯（Arthur Scherbius）于 1918 年创造发明的一种机械密码，使用易操作的恩尼格玛机进行加密和解密。恩尼格玛机所用的一个重要加密方式是用到了多个相连的转子。在加密前，每个转子可设置初始位置，一个转子有 26 种初始设置。一台有三个转子的恩尼格玛密码机就会拥有 26×25×26 ＝ 16 900 个组合。德国海军在战争后期还采用过 10 个转子的恩尼格玛机。法国曾经获得过该种机器，10 余年来无法解密，认为是不可破解的。英国在战后也迟迟不宣告该密码已被破解，希望自己可以应用这套加密系统。

二战中德国海军的 U 型潜艇在大西洋采用“结群战法”：当某艘潜艇发现英国舰艇后，即用无线电将经过恩尼格玛机加密的信息传递给其它潜艇，围歼敌舰。物资和武器补给受到严重威胁的英国，在 1938 年设立了英国政府代码及加密学校（Government Code and Cipher School, GCCS），总部坐落在白金汉郡的布莱切利庄园。在这里，图灵设计了被称为“图灵炸弹（Bombes）”的破译机（仿波兰人 1938 年造的“Bomba”）。这台机器主要由继电器构成，使用了 80 个电子管，由光电阅读器直接读入密码。后为提高速度，布莱切利庄园的工程师汤米 • 弗拉沃尔（Tommy Flowers）在图灵的帮助下设计了 Colossus，这实际上可以看作是最早的电子计算机之一。

图灵对密码的破译也丰富了密码学、数学和计算机科学。例如图灵曾创造了一些新的统计理论。但图灵的这些理论因没有发表而在后来被他人重新发现。统计学中重要的“序贯分析”就是图灵最早使用，后来由瓦尔德（A. Wald）重新提出。实际上，直到 1974 年温特伯坦姆写的《超级机密》(*The*

图 95 图灵机游戏第七关：计算第二轮 G 的问题

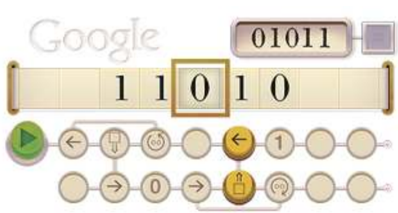


图 96 图灵机游戏第七关：计算第二轮 G 的解答

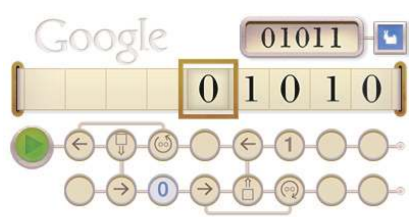


图 97 图灵机游戏第二轮计算的重演（注意右上角的小兔子）

图 98 点击“兔子”后呈现的图灵机演示程序

图 99 位于布莱切利园（Bletchley Park）的图灵雕像

Ultra Secret) 一书出版，图灵在解密方面的贡献才广为人知。

图灵对于人工智能的发展也有重要贡献。1950 年，他提出关于机器思维的问题，设计了一种用于判定机器是否具有智能的试验方法，这种方法现在称作“图灵试验（Turing Test）”。

人机对弈是人工智能领域一个重要的方向。本文的“下”部分在介绍西班牙工程师和数学家托里斯时，会提到他在 1910 年到 1914 年间设计并制造了第一台可以自动下棋的机器 EI Ajedrecista。

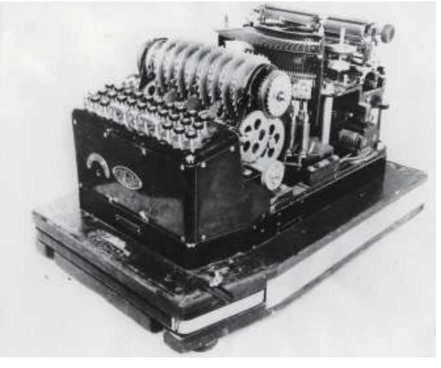


图 100 带 8 个转子的恩尼格玛机

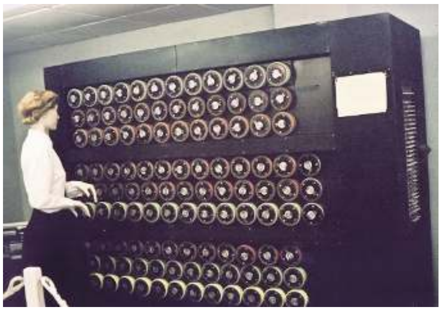


图 101 “图灵炸弹”

图 102 图灵与恩尼格玛机（2005 年圣赫勒拿岛发行，纪念二战结束 60 周年）

历史上，最轰动的人机对弈是 1996 年和 1997 年进行的两场对弈，对阵双方分别为国际象棋世界冠军卡斯帕罗夫（Gary Kasparov）与 IBM 公司的超级电脑“深蓝”。1996 年，深蓝 2–4 失利；1997 年，运算能力翻倍的“深蓝”以 3.5–2.5 的成绩获胜。

图 103 卡斯帕罗夫与图灵的国际象棋程序对弈

图 104 长跑比赛中的图灵

2012 年 6 月 25 日，在曼彻斯特大学举行的图灵诞辰 100 周年纪念大会上，卡斯帕罗夫在发表关于图灵及其“纸机器”的演讲中，表演了一场“人机对弈”。这一次卡斯帕罗夫的对手是图灵在 1952 年设计的程序。他只用 16 步就轻松击败图灵的程序。图灵在写出他的程序时并没有真正的计算机能够来执行他的程序。他只好自己模仿计算机，以每半小时一步的速度，与一位同事下了一盘，结果程序输了。

图灵在生活中也独具魅力。他热爱体育运动，是世界级马拉松长跑运动员，最好成绩 2:46:03 仅比 1948 年奥运会金牌获得者的成绩多 11 分钟。 但遗憾的是，图灵的同性恋倾向不被当时英国的法律认同，被迫接受“治疗”。不甘受辱的图灵于 1954 年 6 月 7 日在家中服用毒苹果自杀，不满 42 岁。2009 年 9 月 15 日，英国首相布朗代表英国政府正式向图灵道歉：我们非常对不起您，您应该得到比这好得多的对待。

人们没有忘记图灵。美国计算机协会（ACM）在 1966 年设立用以表彰在计算机科学方面贡献卓越者的“图灵奖”。该奖有“计算机界诺贝尔奖”之称。2012 年被计算机界定为“图灵年”，全球各地举行了一系列活动来纪念他，特别是在他生前工作和学习过的地方，如美国普林斯顿大学和英国曼彻斯特大学。作为发生在我身边的例子，2012 年一整年中，德国帕德博恩市的海恩茨 • 尼克斯多夫博物馆论坛（简称 HNF，以计算机科学为主题，包括一座计算机博物馆）就分 10 个主题，全面展出了图灵的各个方面。

23 拉齐诞辰 1147 周年

图 105 拉齐诞辰 1147 周年纪念（2012 年 8 月 26 日，巴林、阿富汗、阿尔及利亚、利比亚、阿拉伯联合酋长国、埃及、巴勒斯坦、黎巴嫩、卡塔尔、摩洛哥、约旦、突尼斯、阿曼、伊拉克、科威特、沙特阿拉伯）

拉齐（ Muhammad ibn Zakariyā Rāzī，865 年 8 月 26 日-925 年），或称拉齐斯（Rhazes）是波斯医学家、化学家、哲学家和博物学家。拉齐早年学习伊斯兰文学和数学，在数学上并无大的贡献。

医学上他著有《曼苏尔医书》和《医学集成》等经典，

图 106 拉齐诊治一位麻疹病人

被誉为“阿拉伯的盖伦”、“穆斯林医学之父”。例如他发现了天花与麻疹是两种不同的疾病，最早阐明了过敏和免疫的原理，撰写了最早的儿科护理书籍。

24 比鲁尼诞辰 1039 周年

图 107 比鲁尼诞辰 1039 周年纪念（2012 年 9 月 4 日，巴林、阿富汗、利比亚、阿尔及利亚、沙特阿拉伯、阿拉伯联合酋长国、埃及、巴勒斯坦、黎巴嫩、卡塔尔、摩洛哥、约旦、突尼斯、阿曼、伊拉克和科威特）

阿布 • 拉伊汗 • 穆罕默德 • 本 • 艾哈迈德 • 比鲁尼（Abū al-Rayhān Muhammad ibn Ahmad al-Bīrūnī，英文中常被称作 Al-biruni，973 年 9 月 5 日-1048 年 12 月 13 日）是阿拉伯著名博学家。2012 年 9 月 4 日谷歌以比鲁尼仰望星空的涂鸦，纪念比鲁尼诞辰 1039 周年。

比鲁尼生于花剌子模（现乌兹别克斯坦南部的一个省份）。他毕生从事科学研究与写作，有记载的著作就有 146 部，涉及天文学、历史学、地理学、数学、力学和医学等众多科学领域。科学史学家乔治 • 萨顿（George Sarton, 1884-1956）称他为“伊斯兰最伟大的科学家之一，也是所有时代最伟大

图 108 伊朗德黑兰 Laleh 公园比鲁尼雕像

图 109 比鲁尼诞辰一千年纪念邮票（阿富汗 1973 年发行）

的科学家之一[17]”。

比鲁尼主要在伽色尼王朝极盛时期生活，受到穆斯林第一位有“苏丹”称号的国王马哈茂德及其儿子的庇护。伽色尼王朝幅员辽阔，马哈茂德征服的地盘包括北印度、阿富汗、花剌子模和伊朗大部分地区。马哈茂德统治时，重视学术，吸引了许多学者。

贝尔在其《数学精英》的导言中说：“就整体而言，大数学家是一些多才多艺、精力充沛、机智敏捷、对数学以外的许多事情有着浓厚兴趣的人。”比鲁尼也如前面介绍的海亚姆等许多穆斯林数学家一样，博学多才。阿拉伯文中的“哈基姆（Hakīm）”即是专指这些“智者”。

我们先举些例子来说明比鲁尼在数学之外的成就：

1. 比鲁尼的《编年史》（*Chronology*）讲述了古代各民族的历史和纪元。
2. 发现光的传播速度快于声音。
3. 他精确地测定了 18 种矿石的密度。
4. 他是伊斯兰宗教史专家。马尔代夫盛行伊斯兰教，最早就是由比鲁尼引入的。
5. 他在《药物之书》（*Kitab As Saydalah*）一书中介绍了药物的分类及其在医疗中的应用，包括 720 种药物的医疗作用以及它们在多种语言中的名称。
6. 比鲁尼利用跟随马哈茂德出征印度的机会，在印度居留三十年，写作《印度志》，记载了有关印度自然、社会以及地理学等方面的知识，为印度与阿拉伯之间的文化交流做了许多工作。他的书是后世研究印度历史的重要参考。他关于印度的著作也使得他被认为是最早的人类学家。
7. 比鲁尼最重要的著作之一是《天文典》（*Al Qanun Al Masudi*，英文译为 *The Canon of Al Masudi*，作为献给伽色尼苏丹 Masud 的礼物）。这是一本近 1500 页百科全书式的著作。其中他清晰地解释了月食，提出地球可能绕太阳旋转等观点。

比鲁尼的大多数著作是关于天文学和数学的，特别是他在应用数学方面很有影响。

为宗教服务是穆世林数学家进行研究的一个驱动力。一方面统治者鼓励穆斯林学者研究天地以寻求他们的信仰存在的证明，另一方面穆斯林宗教仪式中也有很多实际活动需要他们研究与数学密切相关的地理学、天文学等。

按照《古兰经》的要求，穆斯林每日要做五次礼拜，分别在日出前、正午后、下午时、日落前和进入夜晚。这要求穆斯林数学家和天文学家研制精准的日晷。清真寺有

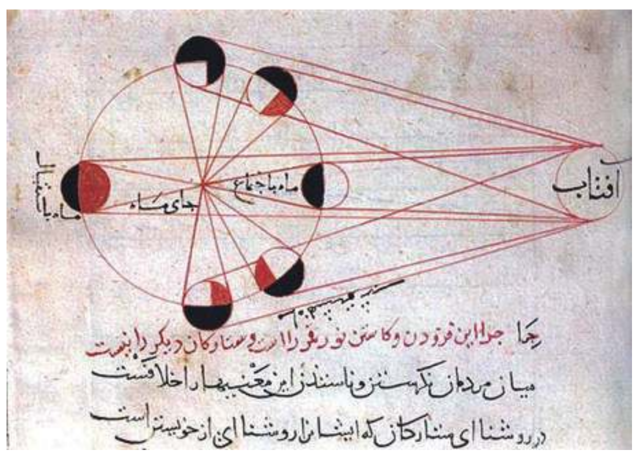


图 110 比鲁尼解释月食的手迹

[17] 原文见 George Sarton 著 *Introduction to the History of Science* 第一卷 707 页：“One of the very greatest scientists of Islam, and, all considered, one of the greatest of all times”。

(a) 突尼西亚开罗安大清真寺广场上的日晷台（Great Mosque of Kairouan, Tunisia）

(b) 开罗安大清真寺日晷台上的日晷

(c) 叙利亚大马士革倭马亚（Umayyid）大清真寺广场上的日晷柱

(d) 倭马亚大清真寺里由著名天文学家 al-Shatir（1304–1375）设计的日晷（图为 18 世纪复制品）

图 111 清真寺中的日晷

专门的司时人员，清真寺外还建有宣礼塔。每到礼拜时间，唤礼者在塔上大声呼唤。此外，穆斯林要确定斋月等圣日的起止日期，这也需要研究数学和天文学。

穆斯林朝拜要面向圣地麦加克尔白。现在每一所清真寺里都有一个凹壁（米哈拉布）指示朝拜方向（Qibla）。在早期伊斯兰世界，如何确定给定点与麦加的方位与距离，是一个至关重要的问题，历代王朝调动了许多学者来参与计算。这里需要分别计算麦加与给定点的纬度和经度，再利用球面三角知识来得出朝拜的方向。现代穆斯林信徒已经可以很方便地掌握礼拜时间和朝拜方位了，智能手机上一个小小的应用程序就可以解决问题。但在古代，这并不简单。即使是现代人，也会对朝拜方位理解错误。

1953 年 4 月 15 日，美国名报《华盛顿每日新闻》（*The Washington Daily News*）登了一篇题为《不能凭借罗盘来建清真寺》（*You Cann't Build That Mosque With a Compass*）的文章，讲述了当地居民对美国首都华盛顿一座在建清真寺的困惑[18]。这座清真寺位于华盛顿哥伦比亚特区麻州大道（Massachusetts Avenue in Washington D. C.），1954 年建立，1957 年开放。当时是西半球最大的清真寺。

在清真寺的建造过程中，一些路人（大部分为穆斯林）

[18] 可参考网页：http://www.surveyhistory.org/can't_build_a_mosque_with_a_compass.htm。

图 112 土耳其伊斯坦布尔圣索菲亚清真寺的米哈拉布

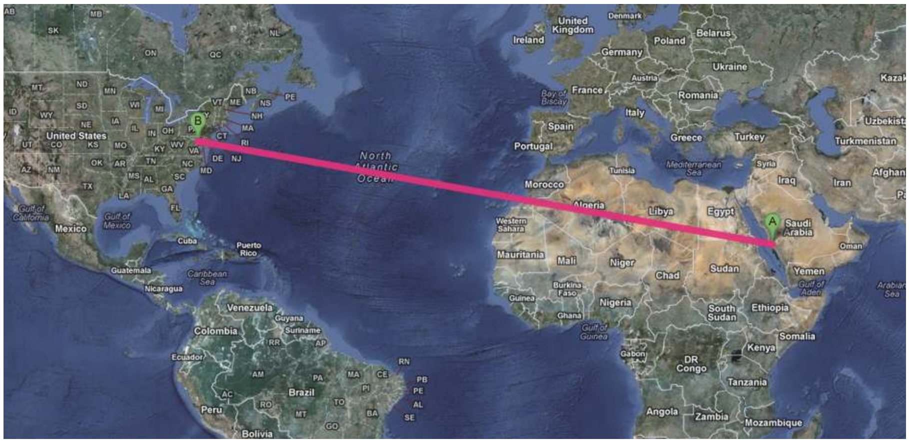


图 113 谷歌地图上从华盛顿到麦加（从左到右）

不时前来参观，其中的细心者注意到该清真寺竟然是偏北的（东偏北 56° 33' 15″），不禁开始担心起来。如同所有清真寺都要朝向麦加圣地一样，华盛顿的这座清真寺也应指向麦加圣地。然而，按照纬度，华盛顿在麦加的南方（华盛顿和麦加的纬度分别为 38° 53'，21° 25'）。在地图上，从麦加往华盛顿方向作直线，也很明显，直线应该稍朝南。

事实上，民众产生疑惑是因为他们在通常意义（墨卡托投影［Mercator projection］）下的平面地图上考虑问题。这上面两点间的直线并不是最短路径。假设地球为理想球体，则地球上两点间的最短路径为经过这两点的最短大圆弧（即以这两点及球心的三点所确定的大圆上的圆弧）。数学上，这样的路径称为测地线。找来一个地球仪，用拇指和食指分别指向这两个城市，在地球仪上比划下就可以形象地理解这一点：最短的路线，从华盛顿出发，以偏北的方向经过纽约附近，穿过北大西洋的冰山。这个道理也用在飞机和海轮的航路设计中。飞机在飞行过程以及海轮在航海过程中都需要不断变换方向，并不是一条直线。

比鲁尼的《城市方位坐标的确定》(*Tahdid, the Demarcation of the Coordinates of Cities* 或 *The Book of the Demarkation of the Limits of the Areas*）就是为了确定穆斯林礼拜朝向而作的。比鲁尼总共撰写了 15 部大地测量学著作。他将数学的精确与严谨引入地理学中，使得他成了科学地理学之父。

确定经度和纬度需要知道地球半径。虽然前人如亚历山大的埃拉托塞尼就已经知道如何利用太阳来测量地球半径的方法[19]，但比鲁尼的方法不像埃拉托塞尼那样依赖地理位置，而且很精密，所得出的地球半径值 6339.6 千米与我们现在所知数值非常接近[20]。

比鲁尼的方法很巧妙。他首先选定某坐山，测出其高 h。然后登到山顶，测出远望地平线的俯角 θ。由三角学知识就可以算出地球半径

$$r=\frac{h\cos\theta}{1-\cos\theta}.$$

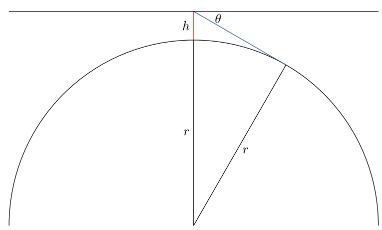


图 114 比鲁尼测量地球半径方法示意图

比鲁尼用一种称为星盘（Astrolabe）的工具来测量角度。他测量山高的方法也是通过该工具测量两次仰角，应用三角学知识得到的：如图 115 所示，先定已知距离为 d 的两点，分别测量其仰望山顶的仰角 α,β，则山高为

$$h=\frac{d\tan\alpha\tan\beta}{\tan\beta-\tan\alpha}.$$

关于山高的测量，我国三国时代刘徽所著的《海岛

[19] 参考作者的《亚历山大城的希帕蒂娅》，《数学文化》第 3 卷第 1 期，2012。

[20] 赤道半径 6378.140 千米，极半径 6356.755 千米，平均半径 6371.004 千米。

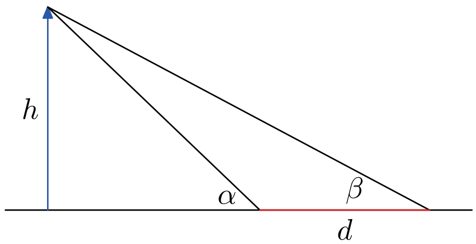


图 115 比鲁尼测量山高方法示意图

算经》中就有望海岛一问。刘徽所用方法和比鲁尼的有些类似，只是《海岛算经》中利用了所立的两个“表”（柱子）进行观测，利用表高、表距等已知量用相似比来计算。

虽然测量方法出奇地简单，但要精确测量也非易事。首先，比鲁尼要选择一个高地以有一定的角度俯视地平线。他选取的山是现今巴基斯坦旁遮普省杰赫勒姆（Jhelum）地区的 Nandana 要塞的一座高山。其次，比鲁尼的方法有赖于角度的精确测量。一般而言，θ 很小，若用弧度表示 θ，则 $\cos\theta$ 近似为 $1-1/2\theta^2$，因此地球半径 R 近似为 $2h/\theta$。由此可见，θ 的微小误差会引起所得结果的精确性。

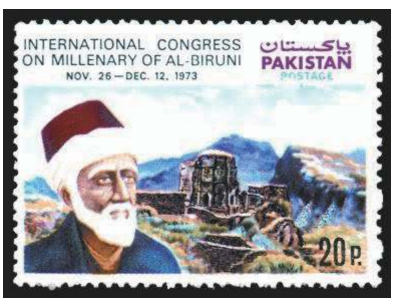
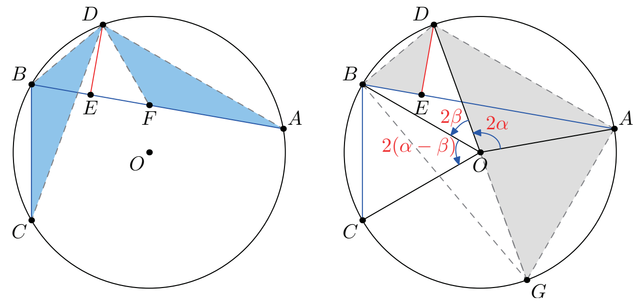


图 116 巴基斯坦 1973 年发行的邮票（背景为比鲁尼测量地球半径所登高山）

我们从比鲁尼测量山高和地球半径的方法可以看到，穆斯林数学家已经能非常熟练地应用三角函数了。确实，阿拉伯人在三角学方面贡献甚大。例如，比鲁尼在他所作《天文典》中研究了圆内接多边形的长、倍角正弦公式、球面三角学的正弦定理以及三角函数的插值法等。

作为一个例子，我们介绍比鲁尼的一本论述圆中弦



(a) 折弦定理及一个证明示意　(b) 折弦定理蕴含两角差的正弦公式

图 117 折弦定理及其推论

的著作里所记载的一个简单而又有趣的定理。这个定理被称为折弦定理（Theorem of Broken Chord）。虽然比鲁尼将它归于阿基米德名下，但该定理实际上等价于两角差的正弦公式，进而可以用之于三角函数表的编制等，因此可以从中了解比鲁尼的兴趣。

折弦定理 [21]：如图 117(a) 所示，设有圆中弦 $AB > BC$，D 为 $\overset{\frown}{AC}$ 弧的中点，且 DE 垂直于 BC，则 $BC + BE = AE$。

折弦定理与两角之差的正弦公式的联系也不难。如图 116（b）所示，设圆半径为 1，圆心为 O，$\angle AOD = 2\alpha$，$\angle BOD = 2\beta$。我们有 $\angle BOC = 2(\alpha-\beta)$。由折弦定理，我们知道 $BC = AE - BE$。另一方面我们知道弦长与三角函数密切相关，有 $BC = 2\sin(\alpha-\beta)$；再注意到 AE、BE 可分别用弦长从而用三角函数表示便可以得到折弦定理的等价表示 [22]：

$$\sin(\alpha-\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta.$$

此即两角差的正弦公式。

有两座以比鲁尼的名字命名的大学：一是乌兹别克斯坦的国立塔什干理工大学（英文名称为 Tashkent State Technical University Named after A.R.Beruni），二是阿富汗卡比萨省于 2000 年成立的比鲁尼大学（Al-Beroni University）。我们将另文简介世界上以数学家的名字命名的大学。

德国比勒费尔德和知屋
2013 年 4 月 10 日

[21] 一个供参考的证明提示：先在 AB 上取点 F，使得 $BC = FA$，再证明△BCD 与△FAD 全等。因此 $BD = FD$。由此立即可得 $BE = EF$。

[22] 设 G 为 D 关于圆心的对称点，易知△DEB、△AED 分别与△DAG、△GBD 相似，这样我们有：$AE = AD\cdot BG/DG$，$BE = AD\cdot BG/DG$。最后将 $DG = 2$，$AD = 2\sin\alpha$，$BD = 2\sin\beta$，$BG = 2\cos\beta$，$AG = 2\cos\alpha$ 代入即可得证。

正态分布的前世今生(下)

靳志辉

六、开疆拓土，正态分布的进一步发展

19 世纪初，随着拉普拉斯中心极限定理的建立与高斯正态误差理论的问世，正态分布开始崭露头角，逐步在近代概率论和数理统计学中大放异彩。在概率论中，由于拉普拉斯的推动，中心极限定理发展成为现代概率论的一块基石；而在数理统计学中，在高斯的大力提倡之下，正态分布开始逐步畅行于天下。

1 论剑中心极限定理

先来说说正态分布在概率论中的地位，这个主要是由于中心极限定理的影响。1776 年，拉普拉斯开始考虑一个天文学中彗星轨道的倾角的计算问题，最终的问题涉及独立随机变量求和的概率计算，也就是计算如下的概率值

$$S_n = X_1 + X_2 + \cdots + X_n$$
$$P(a < S_n < b) = ?$$

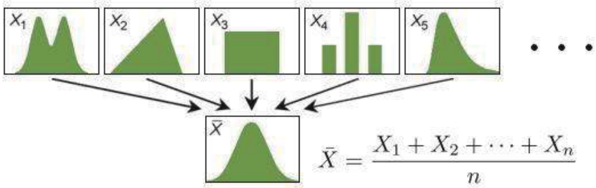
在这个问题的处理上，拉普拉斯充分展示了其深厚的数学分析功底和高超的概率计算技巧，他首次引入了用特征函数（也就是对概率密度函数做傅立叶变换）来处理概率分布的神妙方法，而这一方法经过几代概率学家的发展，在现代概率论里面占有极其重要的位置。基于这一分析方法，拉普拉斯通过近似计算，在他 1812 年发表的名著《分析概率论》中给出了中心极限定理的一般描述：

定理 0.6.1 (拉普拉斯，1812) $e_i\,(i = 1, \cdots n)$ 为独立同分布的测量误差，具有均值 μ 和方差 σ^2，如果 $\lambda_1, \cdots, \lambda_2$ 为常数，$\alpha > 0$, 则有

$$P\left(\left|\sum_{i=1}^{n}\lambda_i(e_i-\mu)\right| \le a\sqrt{\sum_{i=1}^{n}\lambda_i^2}\right) \approx \frac{2}{\sqrt{2\pi}\sigma}\int_0^a e^{-\frac{x^2}{2\sigma^2}}\,dx.$$

这已经是比棣莫弗 - 拉普拉斯中心极限定理更加深刻的一个结论了，理科专业的本科生学习《概率论与数理统计》这门课程的时候，通常学习的中心极限定理的一般形式如下：

定理 0.6.2 (林德伯格 - 列维中心极限定理) 设 $X_1, \cdots, X_n$ 独立同分布，且具有有限的均值 μ 和方差 σ^2 ，则在 $n \to \infty$ 时，有



中心极限定理

$$\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \to N(0,1).$$

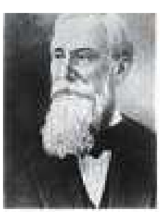
多么奇妙的性质，随意的一个概率分布中生成的随机变量，在序列和(或者等价的求算术平均)的操作之下，表现出如此一致的行为，统一地规约到正态分布。

概率学家们进一步的研究结果更加令人惊讶，序列求和最终要导出正态分布的条件并不需要这么苛刻，即便 $X_1, \cdots, X_n$ 并不独立，也不具有相同的概率分布形式，很多时候他们求和的最终归宿仍然是正态分布。一切的纷繁芜杂都在神秘的正态曲线下被消解，这不禁令人浮想联翩。中心极限定理恐怕是概率论中最具有宗教神秘色彩的定理，如果有一位牧师拿着一本圣经向我证明上帝的存在，我是丝毫不会买账；可是如果他向我展示中心极限定理并且声称那是神迹，我可能会有点犹豫，从而乐意倾听他的布道。如果我能坐着时光机穿越到一个原始部落中，我也一定带上中心极限定理，并劝说部落的酋长把正态分布作为他们的图腾。

中心极限定理虽然表述形式简洁，但是严格证明它却非常困难。中心极限定理就像一张大蜘蛛网，棣莫弗和拉普拉斯编织了它的雏形，可是这张网上漏洞太多，一个多世纪来，数学家们就像蜘蛛一样前赴后继，努力想把所有的漏洞都补上。在19世纪，泊松（Siméon Denis Poisson, 1781-1840）、狄利克雷（Gustav Lejeune Dirichlet, 1805-1859）、柯西（Augustin-Louis Cauchy, 1789-1857）、贝塞尔（Friedrich Bessel, 1784-1846）这些大蜘蛛都曾经试图把这张网的漏洞补上。从现代概率论的角度来看，整个19世纪的经典概率理论并没有能输出一个一般意义下的严格证明。而最终把漏洞补上的是来自俄罗斯的几位蜘蛛侠：切比雪夫（Pafnuty Chebyshev, 1821-1894）、马尔可夫（Andrey Andreyevich Markov, 1856-1922）和李雅普诺夫（Aleksandr Mikhailovich Lyapunov, 1857-1918）。俄罗斯是一个具有优秀数学传统的民族，产生过几位顶尖的数学家，在现代概率论的发展中，俄罗斯的圣彼得堡学派可以算是顶了大半边天，而切比雪夫正是圣彼得堡数学学派的奠基人和领袖。给中心极限定理补漏的方案雏形是从切比雪夫1887年的工作开始的，切比雪夫提出了一个基于矩法的证明，矩法是概率分析中比较传统的方法，使用的数学工具比较基础，不过切比雪夫这个证明也还存在一些漏洞。马尔可夫和李雅普诺夫都是切比雪夫的学生，两人在中心极限定理的严格证明上展开了竞赛。马尔可夫在概率论里面可算是大名鼎鼎，马尔可夫链是应用最为广泛的概率模型之一。马尔可夫和他的老师切比雪夫一样，他们在数学中的研究风格都偏向于使用初等、简单易懂的数学工具来证明复杂艰深的定理。马尔可夫沿着老师的基于矩法的思路在蜘蛛网上辛勤编织，在证明上做了很多的修补，但洞还是补得不够严实。李雅普诺夫不像马尔可夫那样深受老师的影响，他沿着拉普拉斯当年提出的基于特征函数的思路，于1901年给出了一个补洞的方法，切比雪夫对这个方法大加赞赏，李雅普诺夫的证明被认为是第一个在一般条件下的严格证明；而马尔可夫也不甘示弱，在1913年基于矩法也把洞给补严实了。

20世纪初期到中期，中心极限定理的研究几乎吸引了所有的概率学家，这个定理俨然成为了概率论的明珠，成为了各大概率论武林高手华山论剑的场所。不知道大家对中心极限定理中的“中心”一词如何理解，许多人都认为“中心”这个词描述的是这个定理的行为：以正态分布为中心。这个解释看起来确实合情合理，不过并不符合该定理被冠名的历史。事实上，20世纪初概率学家大都称呼该定理为极限定理（Limit Theorem），由于该定理在概率论中处于如此重要的中心位置，如此之多的概率学武林高手为它魂牵梦绕，于是数学家波利亚于1920年在该定理前面冠以“中心”一词，由此后续人们都称之为中心极限定理。

数学家们总是极其严谨苛刻的，给定的一个条件下严格证明了中心极限定理，数学家们就开始探寻中心极限定理成


切比雪夫（1821-1894）


马尔可夫（1856-1922）


李雅普诺夫（1857-1918）


费勒（1906-1970）


列维（1886-1971）

华山论剑

立的各种条件，询问这个条件是否为充分必要条件，并且进一步追问序列和在该条件下以什么样的速度收敛到正态分布。1922 年林德伯格（Jarl Waldemar Lindeberg, 1876-1932）基于一个比较宽泛且容易满足的条件，为中心极限定理提出了一个很容易理解的初等证明，这个条件我们现在称之为林德伯格条件。然后概率学家费勒（William Feller, 1906-1970）和列维（Paul Pierre Lévy, 1886-1971）就开始追问：林德伯格条件是充分必要的吗？基于林德伯格的工作，费勒和列维都于 1935 年独立地得到了中心极限定理成立的充分必要条件，这个条件可以用直观的非数学语言描述如下：

定理 0.6.3（中心极限定理充要条件）假设独立随机变量序列 X_i 的中值为 0，要使序列和 $S=\sum_{i=1}^{n} X_i$ 的分布密度函数逼近正态分布，以下条件是充分必要的：

* 如果 X_i 相对于序列和 S 的散布（可以理解为标准差）是不可忽略的，则 X_i 的分布必须接近正态分布；
* 对于所有可忽略的 X_i，取绝对值最大的那一项，这个绝对值相对于序列和也是可忽略的。

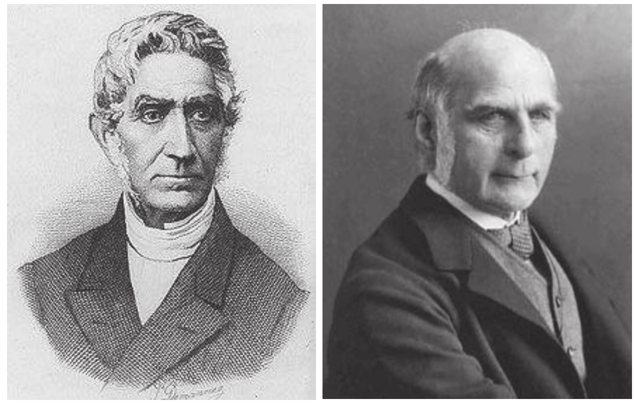
事实上这个充分必要条件发现的优先权，费勒和列维之间还着实出现了一些争论，当然他们俩都是独立的几乎在同一时间解决了这个问题。在列维证明这个充分必要条件的过程中，列维发现了正态分布的一个有趣的性质。我们在数理统计中都学过，如果两个独立随机变量 X, Y 具有正态分布，则 $S=X+Y$ 也具有正态分布；奇妙的是这个定理的逆定理也成立：

定理 0.6.4（正态分布的血统）如果 X, Y 是独立的随机变量，且 $S=X+Y$ 是正态分布，那么 X, Y 也是正态分布。

正态分布真是很奇妙，就像蚯蚓一样具有再生的性质，你把它一刀切两断，它生成两个正态分布；或者说正态分布具有极其纯正的优良血统，正态分布的组成成分中只能包含正态分布，而不可能含有其它杂质。

一流的数学家都是接近上帝的人，善于猜测上帝的意图。1928 年列维就猜到了这个定理，并在 1935 年使用这个定理对中心极限定理的充分必要条件作了证明。有意思的是列维却无法证明正态分布的这个看上去极其简单的再生性质，所以他的证明多少让人觉得有些瑕疵。不过列维的救星很快就降临了，1936 年概率学家克拉美（Harald Cramér, 1893-1985）证明列维的猜想完全正确。

中心极限定理成为了现代概率论中首屈一指的定理，事实上中心极限定理在现代概率论里面已经不是指一个定理，而是指一系列相关的定理。统计学家们也基于该定理不断地完善拉普拉斯提出的元误差理论，并据此解释为何世界上正态分布如此常见。而中心极限定理同时成为了现代统计学中大样本理论的基础。



凯特勒（1796-1874）和高尔顿（1822-1911）

2 进军近代统计学

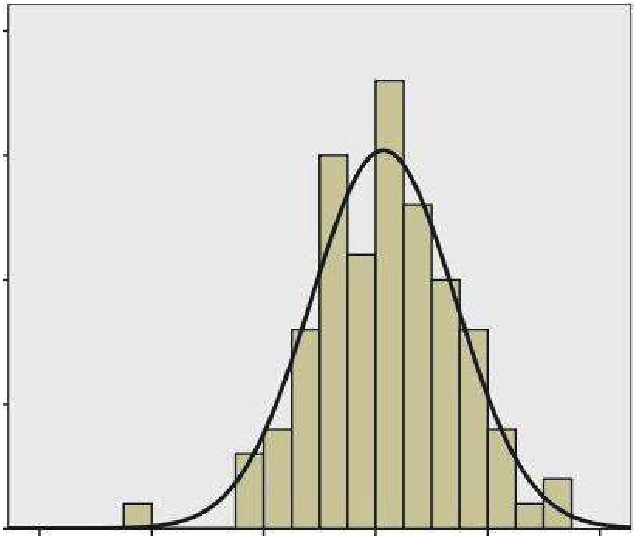
花开两朵，各表一枝。上面说了正态分布在概率论中的发展，现在来看看正态分布在数理统计学中发展的故事。这个故事的领衔主演是凯特勒（Adolphe Quetelet, 1796-1874）和高尔顿（Francis Galton, 1822-1911）。

由于高斯的工作，正态分布在误差分析中迅速确定了自己的地位。有了这么好的工具，我们可能拍脑袋就认为，正态分布很快就被人们用来分析其它的数据，然而事实却出乎我们的意料，正态分布进入社会领域和自然科学领域，可是经过一番周折的。

首先我要告诉大家一个事实：误差分析和统计学是风马牛不相及的两个学科，当然，这个事实存在的时限截止到 19 世纪初。统计学的产生最初是与“编制国情报告”有关，主要服务于政府部门。统计学面对的是统计数据，是对多个不同对象的测量；而误差分析研究的是观测数据，是对同一个对象的多次测量。观测数据和统计数据在当时被认为是两种不同行为获取得到的数据，适用于观测数据的规律未必适用于统计数据。19 世纪的数据统计分析处于一个很落后的状态，和概率论没有多少结合。概率论的产生主要和赌博相关，发展过程中与误差分析紧密联系，而与当时的统计学交集非常小。将统计学与概率论真正结合起来推动数理统计学发展的便是我们的统计学巨星凯特勒。

凯特勒这名字或许不如其它数学家那么响亮，估计很多人不熟悉，所以有必要介绍一下。凯特勒是比利时人，数学博士毕业，年轻的时候曾追随拉普拉斯学习过概率论。此人学识渊博，涉猎广泛，脑袋上的桂冠包括统计学家、数学家、天文学家、社会学家、国际统计会议之父、近代统计学之父、数理统计学派创始人。凯特勒的最大贡献就是将法国的古典概率理论引入统计学，用纯数学的方法对社会现象进行研究。

1831 年，凯特勒参与主持新建比利时统计总局的工作。



用正态分布拟合数据

他开始从事有关人口问题的统计学研究。在这种研究中，凯特勒发现，以往被人们认为杂乱无章的、偶然性占统治地位的社会现象，如同自然现象一样也具有一定的规律性。凯特勒搜集了大量关于人体生理测量的数据，如体重、身高与胸围等，并使用概率统计方法来对数据进行数据分析。但是当时的统计分析方法遭到了社会学家的质疑，社会学家们的反对意见主要在于：社会问题与科学实验不同，其数据一般由观察得到，无法控制且经常不了解其异质因素，这样数据的同质性连带其分析结果往往就有了问题，于是社会统计工作者就面临一个如何判断数据同质性的问题。凯特勒大胆地提出：**把一批数据是否能很好地拟合正态分布，作为判断该批数据是否同质的标准**。

凯特勒提出了一个使用正态曲线拟合数据的方法，并广泛地使用正态分布去拟合各种类型的数据。由此，凯特勒为正态分布的应用拓展了广阔的舞台。正态分布如同一把屠龙刀，在凯特勒的带领下，学者们挥舞着这把宝刀在各个领域披荆斩棘，攻陷了人口、领土、政治、农业、工业、商业、道德等社会领域，并进一步攻占天文学、数学、物理学、生物学、社会统计学及气象学等自然科学领域。

正态分布的下一个推动力来自生物学家高尔顿，当正态分布与生物学联姻时，近代统计学迎来了一次大发展。高尔顿是生物统计学派的奠基人，他的表哥达尔文的巨著《物种起源》问世以后，触动他用统计方法研究遗传进化问题。受凯特勒的启发，他对正态分布怀有浓厚的兴趣，开始使用正态分布去拟合人的身高、胸围以至考试成绩等各类数据，发现正态分布拟合得非常好。他因此相信正态曲线是适用于无数情况的一般法则。

然而，对高尔顿而言，这个无处不在的正态性给他带来一些困惑。他考察了亲子两代的身高数据，发现遵从同一的正态分布，遗传作为一个显著因素是如何发挥作用的？1877年，高尔顿设计了一个叫高尔顿钉板（quincunx, 或者 Galton board）的装置，模拟正态分布的性质，用于解释遗传现象。

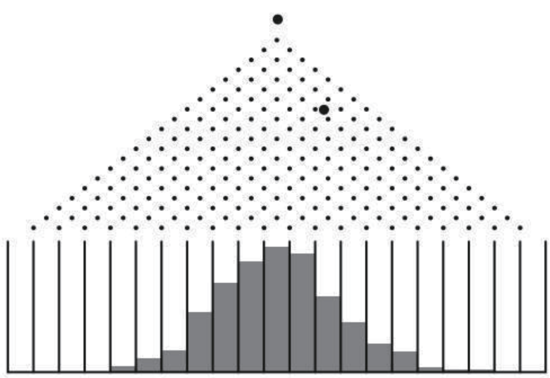
如下图中每一点表示钉在板上的一颗钉子，它们彼此的距离均相等。当小圆球向下降落过程中，碰到钉子后皆以1/2 的概率向左或向右滚下。如果有 n 排钉子，则各槽内最终球的个数服从二项分布 $B(n, 1/2)$，当 n 较大的时候，接近正态分布。

设想在此装置的中间某个地方 AB 设一个挡板把小球截住，小球将在 AB 处聚成正态曲线形状，如果挡板上有许多阀门，打开一些阀门，则在底部形成多个大小不一的正态分布，而最终的大正态分布正是这些小正态分布的混合。

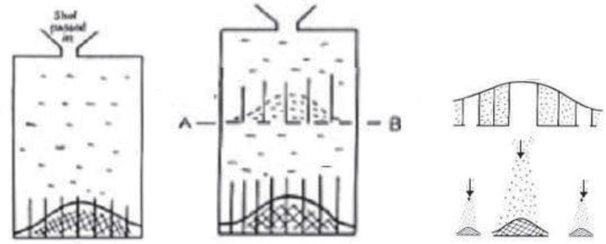
高尔顿利用这个装置创造性地把正态分布的性质用于解释遗传现象。他解释说身高受到显著因素和其它较小因素的影响，每个因素的影响可以表达为一个正态分布。遗传作为一个显著因素，类似图中底部大小不一的正态分布中的比较大的正态分布，而多个大小不一正态分布累加之后其结果仍然得到一个正态分布。

网络上有人开发了一些好玩的程序用于模拟高尔顿钉板，我们可以动态地观察这些小球自上而下滚动的时候，是如何形成正态分布的。一个使用 Java 开发的很漂亮的动态模拟可以在如下网页中观察到：http://www.math.psu.edu/dlittle/java/probability/plinko/index.html。

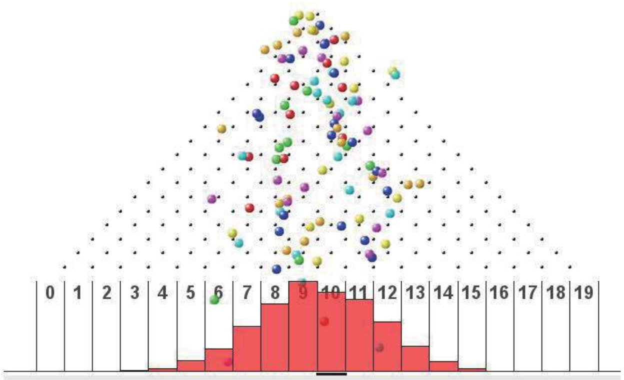
高尔顿在研究身高的遗传效应的时候，同时发现一个奇特的现象：高个子父母的子女，其身高有低于其父母身高的趋势，而矮个子父母的子女，其身高有高于其父母的趋势，即有“回归”到普通人平均身高去的趋势，这也是“回归”一词最早的含义。高尔顿用二维正态分布去拟合父代和子代身高的数据，同时引进了回归直线、相关系数的概念，从而开创了回归分析这门技术。



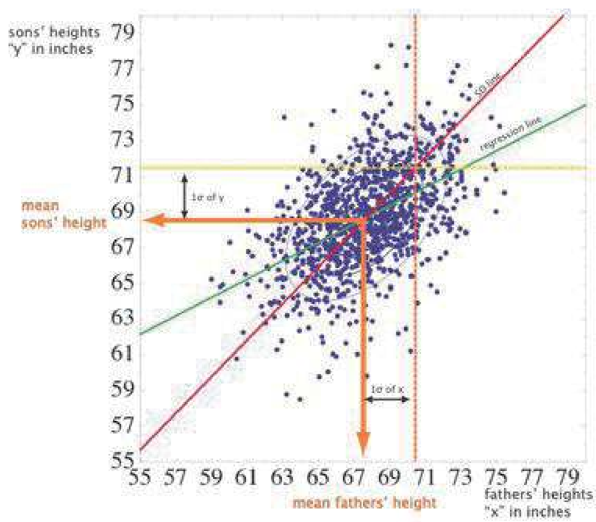
高尔顿钉板



高尔顿钉板解释遗传现象



高尔顿钉板动态模拟程序



儿子与父亲的身高回归线

可以说，高尔顿是用统计方法研究生物学的第一人，他用实际行动开拓了凯特勒的思想，为数理统计学的产生奠定了基础。无论是凯特勒还是高尔顿，他们的统计分析工作都是以正态分布为中心的，在他们的影响下，正态分布获得了普遍认可和广泛应用，甚至是被滥用，以至有些学者认为19世纪是正态分布在统计学中占统治地位的时代。

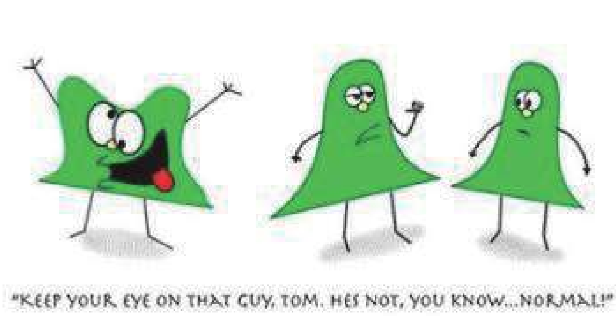
3 数理统计三剑客

最后，我们来到了20世纪，正态分布的命运如何呢？如果说19世纪是正态分布在统计学中独领风骚的话，20世纪则是数理统计学蓬勃发展、百花齐放的时代。1901年，高尔顿和他的学生卡尔•皮尔逊（Karl Pearson,1857-1936）、韦尔登（Walter Frank Raphael Weldon, 1860-1906）创办《生物计量》（*Biometrika*）杂志，成为生物统计学派的一面旗帜，引导了现代数理统计学的大发展。统计学的重心逐渐由欧洲大陆向英国转移，使英国在以后几十年数理统计学发展的黄金时代充当了领头羊。

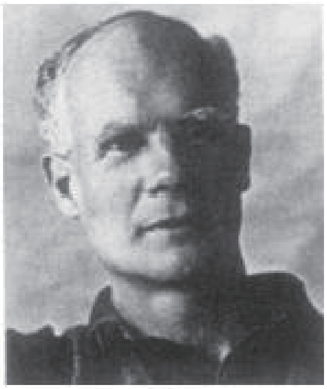
在20世纪以前，统计学所处理的数据一般都是大量的、自然采集的，所用的方法以拉普拉斯中心极限定理为依据，总是归结到正态。到了19世纪末期，数据与正态拟合不好的情况也日渐为人们所注意；进入20世纪之后，人工试验条件下所得数据的统计分析问题，逐渐被人们所重视。由于试验数据量有限，那种依赖于近似正态分布的传统方法开始招受质疑，这促使人们研究这种情况下如何能找到更加准确的统计方法。

在这个背景之下，统计学三大分布 χ^2 分布、t 分布、F 分布逐步登上历史舞台。这三大分布现在的理科本科生都很熟悉。在历史上，这三个分布和来自英国的现代数理统计学的三大剑客有着密切的关系。

第一位剑客就是卡尔•皮尔逊，手中的宝剑就是 χ^2 分布。χ^2 分布这把宝剑最早的锻造者其实是物理学家麦克斯韦，他在推导空气分子的运动速度的分布的时候，发现分子速度在三个坐标轴上的分量是正态分布，而分子运动速度的平方 v^2 符合自由度为3的 χ^2 分布。麦克斯韦虽然造出了这把宝剑，但是真正把它挥舞得得心应手、游刃有余的是皮尔逊。在分布曲线和数据的拟合优度检验中，χ^2 分布可是一个利器，而皮尔逊的这个工作被认为是假设检验的开山之作。皮尔逊继承了高尔顿的衣钵，统计功力深厚，



非正态分布



卡尔·皮尔逊（1857-1936）

戈塞特（1876-1937）

费希尔（1890-1962）

数理统计三剑客

在 19 世纪末 20 世纪初很长的一段时间里，一直被数理统计武林人士尊为德高望重的第一大剑客。

第二位剑客是戈塞特（William Sealy Gosset, 1876-1937），笔名是大家都熟悉的学生氏（Student），而他手中的宝剑是 t 分布。戈塞特是化学、数学双学位，依靠自己的化学知识进酿酒厂工作，工作期间考虑酿酒配方实验中的统计学问题，追随卡尔·皮尔逊学习了一年的统计学，最终依靠自己的数学知识打造出了 t 分布这把利剑而青史留名。1908 年，戈塞特提出了正态样本中样本均值和标准差的比值的分布，并给出了应用上极其重要的第一个分布表。戈塞特在 t 分布的工作开创了小样本统计学的先河。

第三位剑客是费希尔（Ronald Aylmer Fisher, 1890-1962），手持 F 分布这把宝剑，在一片荒芜中开拓出方差分析的肥沃土地。F 分布就是为了纪念费希尔而用他的名字首字母命名的。费希尔剑法飘逸，在三位剑客中当属费希尔的天赋最高，各种兵器的使用都得心应手。费希尔统计造诣极高，受高斯的启发，系统地创立了极大似然估计剑法，这套剑法现在被尊为统计学参数估计中的第一剑法。

费希尔还未出道，皮尔逊已经是统计学的武林盟主了，两人岁数相差了 33 岁，而戈塞特介于他们中间。三人在统计学擂台上难免切磋剑术。费希尔天赋极高，年少气盛；而皮尔逊为人强势，占着自己武林盟主的地位，难免固执己见，以大欺小，费希尔着实受了皮尔逊不少气。而戈塞特性格温和，经常在两位大侠之间调和。毕竟是长江后浪推前浪，一代新人换旧人，在众多擂台比试中，费希尔都技高一筹，而最终取代了皮尔逊成为数理统计学第一大剑客。

由于这三大剑客和统计三大分布的出现，正态分布在数理统计学中不再是一枝独秀，数理统计的领地基本上是被这三大分布抢走了半壁江山。不过这对正态分布而言并非坏事，我们细看这三大分布的数学细节：假设独立随机变量 $X_i \sim N(0,1)$, $Y_j \sim N(0,1)(i=1\cdots n,\ j=1\cdots m)$，则满足三大分布的随机变量可以如下构造出来

1. $\chi_n^2 = X_1^2+\cdots+X_n^2$

2. $t=\dfrac{Y_1}{\sqrt{\frac{X_1^2+\cdots+X_n^2}{n}}}$

3. $F=\dfrac{\frac{X_1^2+\cdots+X_n^2}{n}}{\frac{Y_1^2+\cdots+Y_m^2}{m}}$

你看这三大分布哪一个不是正态分布的嫡系血脉，χ^2、t、F 这三大分布最初都是从正态分布切入进行研究的。所以正态分布在 19 世纪是武则天，进入 20 世纪就学了慈禧太后，垂帘听政了。或者，换个角度说，一个好汉三个帮，正态分布如果是孤家寡人恐怕也难以雄霸天下，有了统计学三大分布作为开国先锋为它开疆拓土，正态分布真正成为了傲视群雄的君王。

20 世纪初，统计学这三大剑客成为了现代数理统计学的奠基人。以戈塞特为先驱，费希尔为主将，掀起了小样本理论的革命，事实上提升了正态分布在统计学中的地位。在数理统计学中，除了以正态分布为基础的小样本理论获得了空前的胜利，其它分布上都没有成功的案例，这不能不让人对正态分布刮目相看。在随后的发展中，相关回归分析、多元分析、方差分析、因子分析、布朗运动、高斯过程等等诸多概率统计分析方法陆续登上了历史舞台，而这些和正态分布密切相关的方法，成为推动现代统计学飞速发展的一个强大动力。

七、正态魅影

> *Everyone believes in it: experimentalists believing that it is a mathematical theorem, mathematicians believing that it is an empirical fact.*
>
> —— *Henri Poincaré*

如果说，充斥着偶然性的世界是一个纷乱的世界，那么正态分布为这个纷乱的世界建立了一定的秩序，使得偶然性现象在数量上被计算和预测成为可能。杰恩斯在《概率论沉思录》中提出了两个问题：

1. 为什么正态分布被如此广泛的使用？
2. 为什么正态分布在实践使用中非常的成功？

杰恩斯指出，正态分布在实践中成功地被广泛应用，主要是因为正态分布在数学方面具有多种稳定性质，这些性质包括：

1. 两个正态分布密度的乘积还是正态分布
2. 两个正态分布密度的卷积还是正态分布，也就是两个独立正态分布随机变量的和还是服从正态分布
3. 正态分布 $N(0, \sigma^2)$ 的傅立叶变换正规化为密度分布后还是正态分布
4. 中心极限定理保证了多个随机变量的求和效应将导致正态分布
5. 正态分布和其它具有相同均值、方差的概率分布相比，具有最大熵

前三个性质说明了正态分布一旦形成，就容易保持该形态的稳定，兰登对于正态分布的推导也表明了，正态分布可以吞噬较小的干扰而继续保持形态稳定。后两个性质则说明，其它的概率分布在各种的操作之下容易越来越靠近正态分布。正态分布具有最大熵的性质，所以任何一个对指定概率分布的操作，如果该操作保持方差的大小，却减少已知的知识，则该操作不可避免地增加概率分布的信息熵，这将导致概率分布向正态分布靠近。

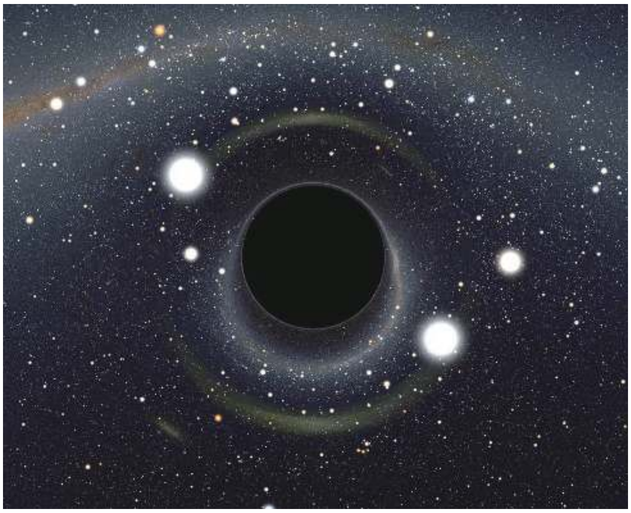
正由于正态分布的多种稳定性质，使得它像一个黑洞一样处于一个中心位置，其它的概率分布形式在各种操作之下都逐渐向正态分布靠拢，杰恩斯把它描述为概率分布中的重力现象（gravitating phenomenon）。

我们在实践中为何总是选择使用正态分布呢，正态分布在自然界中的频繁出现只是原因之一，杰恩斯认为还有一个重要的原因是正态分布的最大熵性质。在很多时候我们其实没有任何的知识知道数据的真实分布是什么，但是一个分布的均值和方差往往是相对稳定的。因此我们能从数据中获取到的比较好的知识就是均值和方差，除此之外没有其它更加有用的信息量。因此按照最大熵的原理，我们应该在给定知识的限制下，选择熵最大的概率分布，而这恰好就是正态分布。即便数据的真实分布不是正态分布，由于我们对真实分布一无所知，如果数据不能有效提供除了均值和方差之外的更多的知识，按照最大熵的原理，正态分布就是这时候的最佳选择。

无处不在的正态分布

当然正态分布还有更多令人着迷的数学性质，我们可以欣赏一下：

* 二项分布 $B(n, p)$ 在 n 很大时逼近正态分布 $N(np, np(1-p))$
* 泊松分布 $Poisson(\lambda)$ 在 λ 较大时逼近正态分布 $N(\lambda, \lambda)$
* $\chi^2_{(n)}$ 在 n 很大的时候逼近正态分布 $N(n, 2n)$
* t 分布在 n 很大时逼近标准正态分布 $N(0, 1)$
* 正态分布的共轭分布还是正态分布
* 几乎所有的极大似然估计在样本量 n 增大的时候都趋近于正态分布
* 克拉美分解定理（之前介绍过）：如果 X, Y 是独立的随机变量，且 $S = X + Y$ 是正态分布，那么 X, Y 也是正态分布
* 如果 X, Y 独立且满足正态分布 $N(\mu, \sigma^2)$，那么 $X + Y$，$X - Y$ 独立且同分布，而正态分布是唯一满足这一性质的概率分布
* 对于两个正态分布 X, Y，如果 X, Y 不相关则意味着 X, Y 独立，而正态分布是唯一满足这一性质的概率分布



正态黑洞

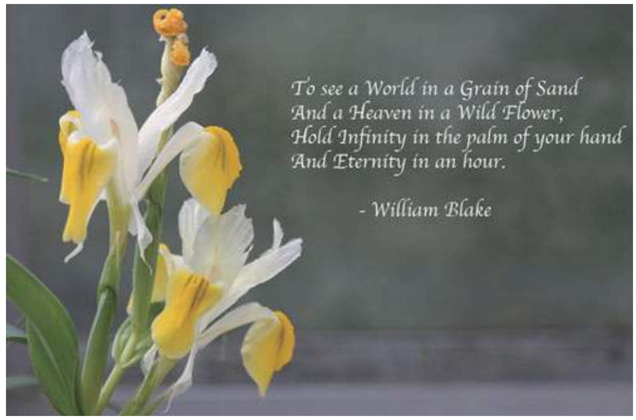
八、大道至简，大美天成

$$\overline{X}=\frac{X_1+X_2+\cdots+X_n}{n}$$

算术平均，极其简单而朴素的一个式子，被人们使用了千百年，在其身后隐藏着一个美丽的世界，而正态分布正是掌管这个美丽世界的女神。正态分布的发现与应用的最初历史，就是数学家们孜孜不倦地从概率论和统计学角度对算术平均不断深入研究的历史。中心极限定理在 1773 年棣莫弗的偶然邂逅的时候，它只是一粒普通的沙子，两百多年来吸引了众多的数学家，这个浑金璞玉的定理不断地被概率学家们精雕细琢，逐渐地发展成为现代概率论的璀璨明珠。而在统计学的误差分析之中，高斯窥视了造物主对算术平均的厚爱，也发现了正态分布的美丽身影。殊途同归，那是偶然中的必然。一沙一世界，一花一天国，算术平均或许只是一粒沙子，正态分布或许只是一朵花，它们却包含了一个广阔而美丽的世界，几百年来以无穷的魅力吸引着科学家和数学家们。

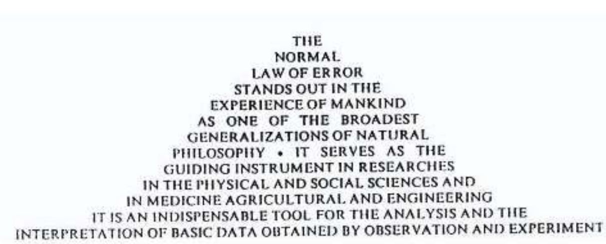
高尔顿对正态分布非常推崇与赞美，1886 年他在人类学研究所的就职演讲中说过一段著名的话："我几乎不曾见过像误差呈正态分布这么美妙而激发人们无穷想象的宇宙秩序。如果古希腊人知道这条曲线，想必会给予人格化乃至神格化。它以一种宁静无形的方式在最野性的混乱中实施严厉的统治。暴民越多，无政府状态越显现，它就统治得越完美。他是无理性世界中的最高法律。当我们从混沌中抽取大量的样本，并按大小加以排列整理时，那么总是有一个始料不及的美妙规律潜伏在其中。"

概率学家卡茨（Mark Kac, 1914-1984）在他的自述传



记《机遇之谜》（*Enigmas of chance: An autobiography*）中描述他与正态分布的渊源："我接触到正态分布之后马上被他深深地吸引，我感到难以相信，这个来自经验直方图和赌博游戏的规律，居然会成为我们日常生活数学的一部分。"另一位概率学家米歇尔•洛伊（Michel Loève, 1907-1979）说："如果我们要抽取列维的概率中心思想，那我们可以这样说，自从 1919 年以后，列维研究的主题曲就是正态分布，他一而再再而三地以她为出发点，并且坚决地又回到她……他是带着随机时钟沿着随机过程的样本路径作旅行的人。"美国国家标准局的顾问约登（W. J. Youden ）用如下一段排列为正态曲线形状的文字给予正态分布极高的评价，意思是：误差的正态分布规律在人类的经验中具有"鹤立鸡群"的地位，它在物理、社会科学、医学、农业、工程等诸多领域都充当了研究的指南，在实验和观测数据的解读中是不可或缺的工具。

几乎所有的人都或多或少地接触数学，虽然各自的目的不同，对数学的感觉也不同。工程师、科学家们使用数学是因为它简洁而实用，数学家们研究数学是因为它的美丽动人。像正态分布这样，既吸引着无数的工程师、科学家，在实践中被如此广泛地应用，又令众多的数学家为之魂牵梦绕的数学存在，在数学的世界里也并不多见。我在读研究生的时候，经常逛北大未名 BBS 的数学板，有一个叫 ukim 的著名 ID 在精华区里面留下了一个介绍数学家八卦的系列


THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY • IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE AGRICULTURAL AND ENGINEERING
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

正态误差态分布律

“Heroes in My Heart”，写得非常的精彩，这些故事在喜欢数学的人群中也流传广泛。最后一个八卦是关于菲尔兹奖得主法国数学家托姆（René Thom）的，它曾经令无数人感动，我也借用来作为我对正态分布的八卦的结语：

在一次采访当中，作为数学家的托姆同两位古人类学家讨论问题。谈到远古的人们为什么要保存火种时，一个人类学家说，因为保存火种可以取暖御寒；另外一个人类学家说，因为保存火种可以烧出鲜美的肉食。而托姆说，因为夜幕来临之际，火光摇曳妩媚，灿烂多姿，是最美最美的……

九、推荐阅读

All knowledge is, in the final analysis, history.
All sciences are, in the abstract, mathematics.
All methods of acquiring knowledge are, essentially, through statistics.

在终极的分析中，一切知识都是历史；
在抽象的意义下，一切科学都是数学；
在理性的基础上，所有的判断都是统计学。

—— *C. R. Rao*

本人并非统计学专业人士，只是凭个人兴趣做一点知识的传播。对统计学历史知识的介绍，专业性和系统性都不是我的目的，我更在乎的是趣味性，因为没有趣味就不会有传播。如果读完这段历史会让你觉得正态分布更加亲切，不再那么遥不可及，那我的目的达到了。如果正态分布是一滴水，我愿大家都能看到它折射出的七彩虹。

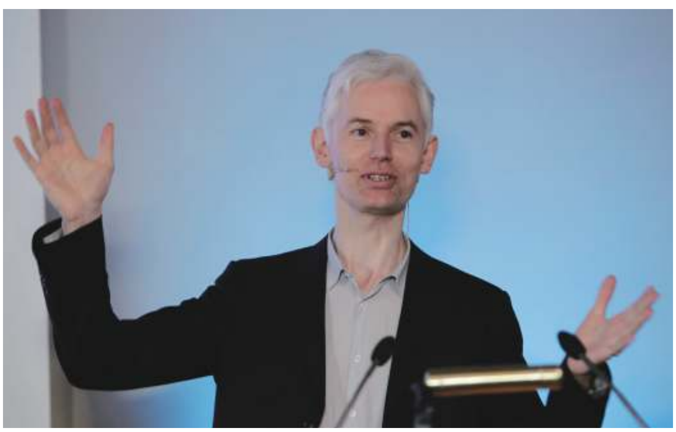
本文所使用的大多是二手资料，有些历史细节并没有经过严格地考证，对于历史资料一定程度上按照个人喜好做了取舍，本文主要基于如下的资料写成，若对历史细节感兴趣，推荐阅读。

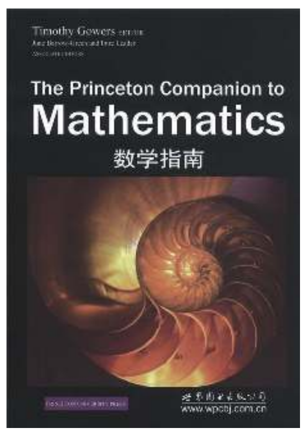
* 陈希孺，数理统计学简史，湖南教育出版社，2000
* 蔡聰明，誤差論與最小平方法，数学传播 21(3):3-13，1994
* 吴江霞，正态分布进入统计学的历史演化，2008
* E.T. Jaynes, Probability Theory: The Logic of Science，Cambridge University Press，2003
* Saul Stahl, The Evolution of the Normal Distribution, Mathematics Magazine, 1996
* Kiseon Kim, Georgy Shevlyakov, Why Gaussianity, IEEE Signal Processing Magazine, 2008
* Stephen M. Stigler, The History of Statistics: The Measurement of Uncertainty before, Belknap Press of Harvard University Press, 1990
* L. Le Cam, The Central Limit Theorem Around 1935, Statistical Science 1(1):78-91, 1986
* Hans Fischer, A History of the Central Limit Theorem: From Classical to Modern Probability Theory, Springer, 2010

作者简介：靳志辉，北京大学计算机系计算语言所硕士，日本东京大学情报理工学院统计自然语言处理方向博士，目前在腾讯科技（北京）有限公司担任研究员，主要参与计算广告学相关的业务，工作内容涉及统计自然语言处理和大规模机器学习方面的工程研究工作。

给年轻数学家的忠告

Atiyah, Bollobás, Connes, Mcduff, Sarnak ／著　　陈跃／译





《普林斯顿数学指南》主编蒂莫西•高尔斯教授

编者按：这篇文章 *Advice to a young mathematician* 出自剑桥大学数学教授、菲尔兹奖得主蒂莫西•高尔斯（Timothy Gowers）主编的《普林斯顿数学指南》，这是一部独具特色的高水平著作，介绍了很多 20 世纪的数学，且都出自数学大家之笔。该书 1000-1010 页有五位顶级数学家对年轻数学人的忠告，对欲进数学研究大门的学子们非常有意义。我们特邀上海师范大学数学系陈跃副教授翻译这一部分，希望对广大数学爱好者有帮助。

年轻的数学家们所需要学习的最重要的东西当然是数学。然而，学习从其他数学家那里得来的经验也是非常有价值的。这篇文章的五位作者被要求按照他们自己研究数学的经验，对新手给出忠告，就好像他们正在开始自己的学术生涯时所愿意收到的忠告一样。由此写出的经验不仅如我们所期待的那样很有意思，而且更使我们感到惊奇的是它们很少有重复。以下就是这五位数学家写的宝贵经验，虽然它们是针对年轻的数学家们而说的，但是肯定会被所有年纪的数学家们所欣赏和阅读。

1 迈克尔·阿蒂亚的忠告

迈克尔·阿蒂亚爵士（Sir Michael Atiyah, 1929- ），英国数学家，主要研究领域为几何学，被誉为当代最伟大的数学家之一。1966 年获菲尔兹奖；2004 年获阿贝尔奖。曾任英国皇家学会会长。

声明

以下仅仅是我个人的看法，主要依据我自己的经验，反映了我的个性、我所研究的数学类型，以及我的工作风格。实际上，数学家们的经验、个性、工作类型和风格可以说是千差万别，你应当遵从你自己的天生爱好。你可以从别人那里学到东西，但应该用你自己的方式来解释你所学到的东西。原创性来自于打破常规，在某种程度上，还来自于过去的实践。

动机

一个数学家做研究，就像一个充满创造力的艺术家一样，必须对所研究的对象极其感兴趣，全神贯注。如果没有强烈的内在动机，你就不可能成功。即使你只是一名数学爱好者，你从解决困难问题中得到的满足感也是巨大的。

研究的头一两年是最为困难的。有那么多的东西要学习，甚至有一些小问题你都无法解决，

这样你就会非常怀疑自己证明新定理的能力。在我从事研究的第二年，我顺利度过了这一艰难的时期。塞尔 (Jean-Pierre Serre) 也许是我们这一代数学家中最杰出的一位，就是他也曾经跟我讲过，他在一段时间里认真地想过是否要放弃数学。

只有凡夫俗子才最相信自己的能力。你越是出色，你为自己定的标准就越高——你可以预见那些不在你目前力所能及范围内的更远目标。

许多有可能成为数学家的人也具有从事其他行业的能力与兴趣，他们可能都会面临着非常艰难的选择：是准备成为一名数学家还是做其他的什么职业。据说伟大的高斯就曾在数学和语言学之间来回摇摆，帕斯卡早年为了研究神学曾经放弃数学，而笛卡尔和莱布尼茨同样也是著名的哲学家。一些数学家后来成了物理学家（例如戴森 (Freeman Dyson)），而另一些人正好相反（例如钱德拉 (Harish Chandra)、博特 (Raoul Bott)），他们从物理学家变成了数学家。你不能将数学看成一个封闭的系统，数学与其他学科之间的相互作用不论对个人还是对社会来说都是健康的。

心理方面

由于在数学中需要精神高度集中，由此产生的心理压力是相当大的，即使是在研究比较顺利的时候也是如此。这个问题是大是小主要看你的性格，不过可以采取措施来降低紧张的情绪。与同学的交流——听讲座、参加讨论班和会议等——都有利于开拓视野和获得很重要的群体支持。过分的孤独与深思可能是比较危险的，有时候表面上看来是散漫的闲谈其实并不是在浪费时间。

一开始的时候，与同学或者导师进行合作研究有许多好处，并且与别人的长时段合作会使人感到特别有信心，无论是在数学方面还是在个人交往方面。当然，个人独自安静的思考总是需要的，不过同朋友们的思想交流与讨论会更有助于这种思考，所以也是不可缺少的。

解决问题还是创建理论

数学家们有时可以被分为“问题解决者”或者“理论创建者”。虽然确实有比较极端的例子显示了这种差别（例如爱尔迪希（Paul Erdös）与格罗滕迪克（Alexande Grothendieck）就是一对），但是绝大多数的数学家都处于他们中间的某个位置，他们同时在解决问题和发展某个理论。实际上，如果一个理论没有导致具体的有趣问题的解决，那么就不值得去建立它。反过来，对于任何真正意义上的深刻问题，在解决它们的过程中总能刺激相关理论的发展（费马大定理就是一个经典的例子）。

这对一个初学者来说有什么启示？虽然人们不得不去读那些书本和论文，以吸收通常的概念与理论方法，但是实际上初学者必须学会去关注一个或更多个具体的问题。这些问题可以让人深思，可以磨砺人们的勇气。一个经过人们仔细研究和理解透彻的特定问题也是检验一个理论是否有效的非常有价值的试金石。

根据研究过程的不同，最后形成的博士论文可能会抛开绝大多数理论的外衣而聚焦于一些本质上的具体问题，也可能是建立一个较为宽广的理论框架使得具体问题纳入其中。

好奇心的作用

驱使人们进行研究的原始动力就是好奇心。一个特定的结论什么时候成立？那是一个最好的证明抑或还有更自然、更简洁的证明？使得结论成立的最一般的情形是什么？

如果你在阅读论文或在听讲座时，总是问自己这样的问题，那么或早或迟答案会隐约浮现——包括一些可能的探索路径。每当这种情形出现时，我就会抽出时间努力追踪这种想法，看它会引到哪里，或者是否经得起仔细琢磨。尽管通常来说十有八九会进入死胡同，但偶尔一次会发现金子。困难在于我们不知道什么时候该停止，有些起初看起来是有效的想法实际上根本没用。这时就应该果断脱身，回到主要的道路上来。人们常常会很犹豫作出这样的决定，事

实上我就是经常回到先前已经丢弃了的想法上来，尝试用另外一种方法来解决问题。

令人想象不到的是，好的想法也会产生于一个不好的讲座或讨论班。在听报告的时候，我经常发现，结果很漂亮，但是证明却很复杂和烦琐。此时我就不会再跟着黑板上的证明，而是在接下来的时间里去构思一个更简洁的证明。虽然这通常来说不太成功，但至少我更好地度过了我的时间，因为我已经用我自己的方式努力地想过这个问题。这远胜过被动地跟随别人的思考。

例子

如果你像我一样，喜欢宏大的和强有力的理论（我虽然受格罗滕迪克的影响，但我不是他的信徒），那么你就必须学会将这些理论运用到简单的例子上，以检验理论的一般性结论。多年以来，我已经构造了一大批这样的例子，它们来自各个分支领域。通过这些例子，我们可以进行具体的计算，有时还能得到详尽的公式，从而帮助我们更好地理解一般性的理论。它们可以让你脚踏实地。非常有意思的是，虽然格罗滕迪克排斥例子，但是很幸运地是他和塞尔有着非常紧密的合作关系，而后者能够弥补他在例子方面的不足。当然在例子与理论之间也没有一条明确的分界线。我喜欢的许多例子都是来自于我早年在经典射影几何中所受到的训练：三次扭曲线、二次曲面或者三维空间中直线的克莱因（Klein）表示等。再没有比这些例子更具体和更经典了，它们不仅都可以同时用代数的方式和几何的方式来进行研究，而且它们每一个都是一大类例子中开头的一个（例子一多慢慢就变成了理论），它们中的每一个都很好地解释了以下这些理论：有理曲线的理论、齐性空间的理论或者格拉斯曼流形 (Grassmannians) 的理论。

例子的另一个作用是它们可以指向不同的研究方向。一个例子可以用几种不同的方式加以推广，或用来说明几种不同的原理。例如一条经典的二次曲线不仅是一条有理曲线，同时又是一个二次超曲面（quadric），或者是一个格拉斯曼流形等。

当然最重要的是，一个好例子就是一件美丽珍宝。它光彩照人，令人信服。它让人洞察和理解。它是（我们对数学理论）信仰的基石。

证明

我们所受到的教育告诉我们，“证明”是数学中最重要的事情，用公理和命题小心编织起来的欧几里得几何体系提供了自文艺复兴以来现代思想的基本框架。相比于其他自然科学家们做实验的检验方法，数学家们为他们绝对准确无误的定理而感到自豪，更不要说在其他领域里那些模模糊糊的思维方式了。

但是自从哥德尔（Kurt Gödel）（发现不完全性定理）以来，数学的绝对真理地位确实发生了动摇，此外繁复冗长的计算机证明的出现也使数学家们的态度变得更谦卑一些。但是不管怎样，证明还是保持着它在数学中的主要作用，如果在你的论文中，你的证明有一个比较严重的漏洞，那么将直接导致退稿。

然而，如果将数学中的全部研究工作仅仅等同于不断作出各种证明的过程，那么你就错了。实际上人们可以说，数学研究中真正带有创造性的那部分工作在写证明的阶段之前就已经完成了。对于后面这个“证明阶段”，我们可以打一个比方：就好比你是在写剧本，必须要从事先的构想出发，发展情节，写出对话，包括给出舞台指导等。最后形成的剧本就可以看成是“证明”：它是事先构想的具体实现。

在数学中，一般是先有思想和概念，然后再提出问题。接下来就开始对问题解答的探寻，人们寻找某种方法或者策略。一旦你自己相信这是一个恰当的问题，并且你又有对此问题合适的工具，那么你接着就会开始努力思索证明的具体技术细节。

但不久你会意识到（也许是通过反例发现）问题提出的方式不对。有时候，在初始的想法与最终的结论之间有较大的反差。你没有注意到一些隐含的假设，或忽略了某个技术细节，或者

你考虑的情形太一般。然后你不得不回过头来，重新修改提出你的问题。如果有人说数学家们总是控制他们提出的问题，以便他们得到答案，这是不公平的夸大其词，但也不是完全没有道理。能够提出一些既有趣又可以被解决的问题，是数学中一种高超的艺术，数学本身其实就是一种艺术。

证明实际上是创造性想象和不断反思推理之间长期相互作用的最终结果。如果没有证明，数学的研究是不完整的，反之，如果没有想象，则研究无从谈起。在这里人们可以看到和其他领域中创造性艺术家（例如作家、画家、作曲家或建筑师）工作的一个相似情形。先有一个幻象，然后发展成一个思路，再不断试验展开，最后便是漫长的艺术品总装完成的技术性过程。技术与幻象之间必须保持接触，各自按照自己的方式不断地修正另外一方。

策略

在上一节中，我讨论了对于证明的看法，以及它在整个创造性过程中的作用。现在让我们转向一个对于年轻的数学家们来说最实际的问题。人们应采取什么策略？你怎样做，才能够找到一个证明？

这个问题如果泛泛而谈的话，没有多大意义。就像我在前面说过的那样，每一个好问题都有它的起源：它来自于某个背景，它有自己的根。为了使问题能够得到进展，你必须要透彻地理解这些根源。这就是为什么发现你自己的问题、提出你自己的想法总是比从你导师那里得到问题要好的缘故。如果你知道一个问题是从哪里来的，为什么要问这个问题，那么你就已经成功了一半。实际上，问一个正确的问题常常和解决这个问题一样困难。找到正确的问题背景是首要的一步。

因此，简要地说，你需要对这个问题的历史有一个很好的了解。你应当知道解决类似的问题是采用什么方法，以及这些方法的局限性又在哪里。

当你被一个问题完全吸引时，应该立即全力以赴地思考这个问题。为了得到解答，除了全力投入外别无他法。你应当考察特殊的情形，以便确定主要困难出现在什么地方。你对问题的背景和先前的解决方法了解得越多，你能够尝试的技巧与方法也就越多。另一方面，有时候对问题与方法的无知也是一件好事情。曾有报道说李特尔伍德 (J. E. Littlewood) 让他的每一个研究生都分别做一个将黎曼猜想装扮起来的问题，直到他们在六个月之后，才知道了真相。他的理由是学生们不会有自信去直接攻克这么有名的难题，但是如果他们不知道他们的对手是大名鼎鼎的黎曼的话，也许他们会获得进展！尽管这种策略不大可能产生一个黎曼猜想的证明，却能够产生一批生气勃勃、敢于攻坚克难的学生。

我自己的方法是尽量避免直接攻击，努力寻找间接的途径。这是因为，将你的问题与各个不同领域中的思想与方法联系起来，可能会带来令人意想不到的结果。如果这个策略成功的话，它将会导致一个非常漂亮的简单证明，同时也“解释”了为什么事情能够成立的原因。实际上我相信：努力地去寻找这样一种解释和理解，是我们真正应该达到的目标。证明可以看成是这个过程的一部分，有时也是这个过程的结果。

拓展你的视野也是你寻找新方法任务中的一部分。与人交谈会提升你的数学素养水平，并且有时会给你带来新思想和新方法。你很有可能由此而获得关于你自己研究的一个有价值的想法，甚至是一个新的方向。

如果你需要学习一个新的课题，除了学习文献之外，最好是能找到这方面的一个比较友善的专家，“从他的嘴里”获取教益——口头的讲解更简洁明快。

在向前看并经常注意新发展的同时，你也不应该忘记过去。在过去的年代中，有许多非常有价值的数学成果被尘封和遗忘了，它们只有被重新发现的时候才显露出光芒。这些结果不容易被发现，部分原因是因为数学的术语和风格改变了，但是它们确确实实是金矿。如果你遇到这样的金矿，你应该要感到非常幸运，你必须报答那些开拓者。

独立或合作

在开始你的研究之前，你与你导师之间的关系是至关重要的，因此要小心地选择你的导师，包括他所研究的方向、人品以及以往的研究工作等都要考虑。当然很少有导师在这三个方面都令人满意。接下来，如果事情在头一两年进行得并不顺利，或者你的兴趣发生了明显转移，则应该毫不犹豫地调换你的导师，甚至是你的大学。这不会冒犯你的导师，或许也是他的解脱！

有的时候，你可能是比较大的研究小组中的一个成员，并且与其他成员也有交流的机会，所以实际上你有不止一个的导师。这可以提供其他的思想来源，以及另外不同的工作方式，这些都是有帮助的。在这样一些大的群体中，你也可以从你的同学那里学到许多，这就是为什么选择一个包含有大的研究生院的数学系是一个好主意的缘故。

当你一旦完成了你的博士论文后，你的研究就进入了一个新的阶段。尽管你可以继续与你的导师进行合作，并待在原来的研究群体中，但是为了你以后进一步的发展，比较健康的做法是用一年或更多的时间去另外的一个地方。这可以让你接受新思想的影响，并获得更多的机会。现在是这样一个时代：你可以有机会在大千数学世界中为自己找到一个位置。一般来讲，在一个相当长的时间里，继续太紧密地停留在你博士论文的课题上不是一个好的主意。你必须要“另立门派”，以显示你的独立性。这不必在研究的方向上作剧烈的改变，只是应该要有确确实实新颖的地方，而不是你博士论文的简单的常规延续。

风格

在你写论文的时候，你的导师通常会指导你如何安排文章的结构和呈现的方式。然而在你的数学研究中也非常需要你自己的个人风格。虽然对于各种类型的数学来说，这方面的要求有所不同，但还是有许多方面的要求适用于所有的数学分支学科。以下便是对怎样写出一篇好论文的几点提示。

(1) 在你开始写作之前，先通盘考虑好整个论文的逻辑结构。

(2) 将很长的复杂证明分成比较短的中间步骤（如引理、命题等），这会帮助读者阅读。

(3) 写通顺简明的英语（或者你选择的语言）。请记住数学也是文学的一种表现形式。

(4) 尽可能地简明扼要，同时又要叙述清楚。要保持这样的平衡是很困难的。

(5) 尽量将论文写成和你所喜欢阅读的论文一样，并模仿它们的风格。

(6) 当你已经完成了你论文的主要部分后，回过头来认真地写一篇引言，在其中要清楚地解释论文的结构和主要结果，以及一般的来源背景。要避免不必要的含糊深奥，要面向一般的数学读者，而不只是少数的专家。

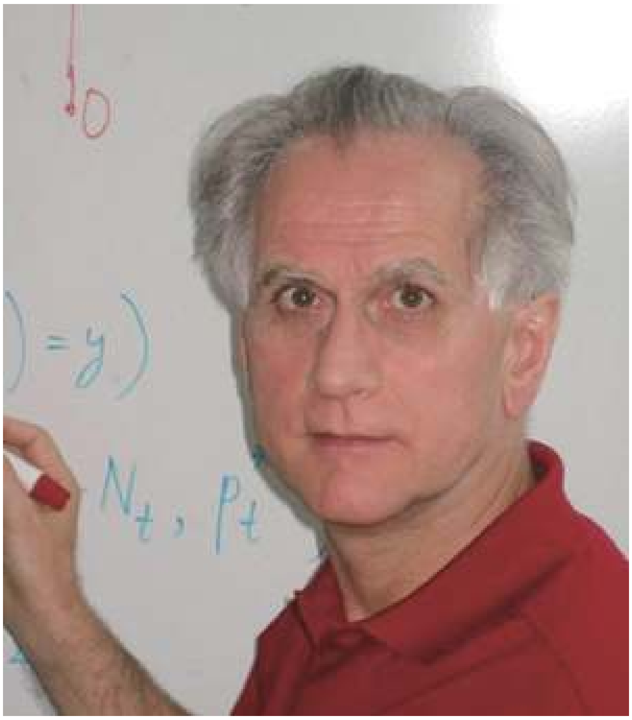
(7) 试着将你的论文初稿让一个同事阅读，并留意任何的建议或评论。如果你最亲近的朋友与合作者都无法理解你的论文，那么你就已经失败了，你需要加倍地努力。

(8) 如果不是非常急着出版，那么将你的论文丢在一边几个星期，做其他的事情。然后再以一种新鲜的视角重新来阅读你的论文，会有一种全然不同的感觉，你将知道怎样去修改它。

(9) 如果你相信重写论文会更加清楚、更容易阅读，那么你就不要吝啬将论文重新写一遍，也许站在一个全新的角度看得更清楚。写得好的论文将成为“经典”，被将来的数学家们广泛阅读。

2 比勒•鲍隆巴斯的忠告

比勒•鲍隆巴斯（Béla Bollobás, 1943- ），匈牙利籍的英国数学家，英国皇家学会院士，在泛函分析、组合数学、图论方面做出过杰出贡献。



“在这个世界上，那些丑陋的数学不可能长久流传。”哈代（G. H. Hardy）曾经这样写过，同样我相信在这个世界上，那些阴郁的缺乏激情的数学家也是没有什么地位的。仅仅当你真正热衷于数学时，你才有可能做好它，即使当你从事另外一种工作，在忙了一整天之后你也会挤出时间来研究数学。就像诗歌和音乐一样，数学不是机械的劳作，而是美好的假日。

品味（或鉴赏力）高于一切。对于我们这门学科来说，对什么是好的数学似乎有比较一致的看法，这简直是一个奇迹。你应当在重要的分支领域里工作，这种领域在相当长的时间里不大可能枯竭，你应该做那些既美丽又重要的问题：在一个好的领域里，这样的问题非常多，而不是仅有那么几个著名的难题。确实来说，如果把目标定得太高，将导致在长时间里出不了成果：虽然在你研究生涯的某个阶段里可以忍受，但是对于新手来说最好要避免这样。

在你的数学活动中要努力保持一种**平衡**：对于真正的数学家来说，研究是放在首要地位的，但是在研究之外，还要进行大量的阅读，以及搞好教学。要充满乐趣地做好各个层面上的数学工作，即使它（几乎）跟你的研究没有任何关系。教学不应该成为负担，而应该是灵感的来源之一。

研究绝不应该变成一种零星的杂务（不像写作那样）：你应该选择你很难**不**去想它的问题。这就是为什么你专注于吸引自己的问题比你去解决别人交给你的问题要更好的缘故。在你研究生涯的早期，当你还是一个研究生的时候，你应该听取你的有经验的导师的意见，让他来帮助你判断你自己喜欢的问题是否合适，这比做他给你的问题要好，因为后者可能不符合你的品味。毕竟，你的导师对某个问题是否值得你去研究，应该会有一个比较好的想法，哪怕是他对你的实力与品味可能还不了解。在你以后的研究生涯中，当你不再依赖你的导师时，与一些比较谈得来的同事的交流也常常会受到启发。

我会建议你们：在任何时候，你所做的数学问题应该包含以下两种类型：

(1)“梦想”类的问题：一个你非常想要解决、但基本上你不可能**期望**解决的大问题。

(2) 非常值得的问题：如果花费足够的时间和努力、并且足够幸运的话，你觉得你很有可能解决的一些问题。

此外，还有两种你应该考虑的问题，虽然它们不如前面的那两种问题重要。

(1) 不时地解决这样一类问题，它们在你的能力之下，你完全有自信会很快地解决它们，使得花在它们上面的时间，不会妨碍你去解决更合适你的问题。

(2) 在更低的层次上，去做那些已经不是真正值得去研究的问题（虽然它们在几年前曾经是这样的问

题），总是一件非常愉快的事情，由于这些问题太好了，所以值得花时间：解决他们将给你带来快乐，并锻炼和提升你的创造能力。

要耐心，要坚持。当你考虑一个问题时，也许你能够采取的最有用的措施是，在所有的时间里都把这个问题放在心上：牛顿就是用这样的方法，其他的许多先人也都是这样。要付出你的时间，尤其是在攻克主要问题的时候，要保证自己在一个大问题上，花费相当数量的时间，但不应期望过多，做完之后，不妨总结一下，然后决定接下来该做什么。在让你的研究不放过一个机会的同时，也应注意不要在一种方法上陷得太深，否则就可能遗漏其他的解决方法。就像爱尔迪希曾经说过的那样，要使头脑敏捷，要让你的大脑保持开放的状态。

不要怕犯错误。一个错误对于一名象棋手来说，可能是致命的，但对一个数学家来说，它相当于是常规的程序。你应该感到恐惧的是面对一张空白稿纸的时候，此时你对问题的思考还不多。等过了一段时间，你的废纸篓里将装满你尝试失败的稿纸，此时也许你还在勤奋地工作。要尽量避免平庸的尝试，而永远是兴致勃勃地投入工作。特别是，研究一个问题的最简单情形，通常来说不会浪费时间，而且可能是非常有效的方法。

当你在一个问题上花费了相当多的时间之后，很容易低估你所取得的进展，并且还同样会低估你将它们全部回忆起来的能力。最好将它们写下来，哪怕是一小部分的结果：你的笔记会节约你以后的大量时间，会给你带来更多的机会。

如果你足够幸运，并且取得了突破的话，你很自然地会对整个项目感到有些厌倦，并且还想靠此荣誉止步不前。要抵制这种想法，并努力寻找是否还有其他的突破在等待着你。

作为一个年轻的数学家，你的主要优势是你有充足的时间做研究。对此你可能没有意识到，但是以后很可能不大会再有像你开始研究生涯时那样充足的时间了。每个人都感觉到没有足够的时间来研究数学，并且随着时间的推移，这种感觉会越来越强烈，时间肯定会越来越少。

现在转向**阅读**。在谈到所阅读过的数量时，年轻人通常都处于不利的地位，因此作为弥补，应尽可能地多阅读，既包括你的一般领域，也包括作为一个整体的其他数学。在你的研究领域，要确信你读的许多论文是由最好的人写的。虽然这些文章常常写得不太仔细，但是它们所包含的高质量的想法会对你的辛苦阅读给出丰厚的回报。不论你读什么，都要保持一种积极参与和质疑的态度：不断设法理解作者的意图，不断努力思考是否还有更好的处理方法。如果作者走的是你已经知道的思路，那么你应该感到高兴，如果他走的是另一条不同的道路，那你就应该进一步思考其中的缘由。对于各种定理与证明，反复问你自己这样的问题，即使它们看起来非常简单：这些问题将极大地帮助你更好地理解数学。

另一方面，对于一个你正在打算做的未解决问题，通常来说**不必**熟读与此相关的所有东西：但在你很深入地思考过之后，可以（而且应该）去读那些其他人写的不成功的文章。

要保持你对新奇事物的感知与惊讶的能力，要能够欣赏你所读过的数学研究成果和思想，不要对这些杰出的思想和成果感觉到没有什么了不起。事后你当然知道事情一定是这样的，这是很容易的：毕竟你已经刚读完了证明。聪明的人会舍得花大量的时间来汲取新思想。对他们来说，只知道一些定理和看懂它们的证明是远远不够的：他们要把定理及其证明融化在他们的血液里。

在你的数学职业生涯进展过程中，永远让你的心智保持对新思想和新方向的开放状态：数学的疆域总是在变化，如果你不想落在后面，那你就必须也要跟着变化。永远要磨砺你的工具，并且不断学习掌握新的工具。

最重要的是，**喜欢数学和热爱数学**。喜欢你自己的研究，到处寻找和阅读相关的新结果，将你对数学的热爱传递给别人，即使你的创造仅仅是一些美丽的小问题，只要是你思考过的，或者是从你的同事那里听说的，都会带来乐趣。

如果我要总结以上这些我们应该遵循的忠告，以便在科学和艺术中获得成功，那么最好是回忆维特鲁威（Marcus Vitruvius Pollio）在两千年前曾经写下的一句话：

从来不学习的天才，以及没有天才的学习者，都不能成为一个完美的艺术家。

3 阿兰·孔涅的忠告

阿兰·孔涅（Alain Connes, 1947- ），法国数学家，在非交换几何学方面的工作对理论和数学物理有重要影响。他是 1982 年菲尔兹奖获得者。

数学是现代科学的支柱，它是许多新概念与新工具的相当有效的源泉，借助于这些新概念和新工具，我们才得以理解置身其中的“现实世界”。这些新概念本身就是人类思维这个蒸馏器经历长期“蒸馏”过程的结果。

我被要求写一些对于年轻数学家的忠告。首先我感到每一位数学家都是一个个特殊的案例，总体来讲数学家们都倾向于成为（喜欢独立的）“费米子”，即他们尽量避免在太大众化的领域里做研究，而物理学家们的表现则更像（群居的）“玻色子”，他们组合成很大的团队，并经常“过分夸大”他们取得的结果——这种态度是会被数学家们鄙夷的。

人们一般总是首先将数学划分成一些相互独立的分支学科，例如几何学、代数学、分析学以及数论等等，其中，几何学主要是试图理解“空间”的概念，代数学主要研究字母符号的操作艺术，而分析学则主要关注涉及“无穷”与“连续”的对象等等。

但是，这种看法与数学这门学科的一个最重要的本性相违背，即把上述那些分支学科中的任何一个在不剥离自身本质的前提下从其余的分支中独立出来是根本不可能的。实际上，整个数学就像一个完整的生命体，只有在结为一体的情况下才能生存，如果它被分割成若干不相连的部分，那么它就会消亡。

数学家们的研究生涯可以被描述成是在“数学的现实世界”王国里的一次探险旅行，他们用自己的知识架构逐步揭开它的神秘面纱。

这个过程往往开始于对现存书本上关于数学王国的教条描述的不满与反叛。想要成为数学家的年轻人开始意识到，他们自己关于数学世界的看法已经抓住了数学的某些特征，而这又与已有的教条不相符合。在大多数的情况下，这种初期的反叛来源于无知，但却不无益处，因为它可以帮助人们从对权威的敬畏中解放出来，使得他们可以依靠他们的直觉，并且用实际的证明来支撑他们的直觉。一旦一个数学家以一种原始且“个人化”的方式真正开始进入数学世界的某个小领域，并且其研究无论初看起来是多么怪异[1]，那么这个探险旅行实际上就已经开始了。当然，很重要的是不要去打破“阿莉阿尼线团”[2]：在始终保持用一种新鲜的眼光看待旅途中遇到的各种问题的同时，还能够在一次次感到迷路的时候回到出发点。

同样重要的是，一直保持对各种数学的兴趣。否则，我们就会冒着一种风险，将自己完全局限于一个已经被高度技术化了的非常狭小的领域里，从而限制了我们对于巨大的变幻莫测的数学

世界洞察力的发挥。

在这方面最基本的要点是：尽管有许多数学家毕其一生在探索数学世界中各不相同的领域，而且看问题的角度是那样的不同，可是他们都同意他们其实是在研究同一对象的各个不同部分。不管我们各自旅行的出发点在哪里，总有一天，当我们走了足够长的距离后，会发现大家都不约而同地走到了数学王国的同一座著名城堡：例如椭圆函数、模形式或者 zeta 函数等。“条条道路通罗马”，数学世界也是相互连通的。当然这并不是说数学的所有各个部分都是相似的，在这里很值得引用格罗滕迪克（在《收获与播种》一书中）对分析学与代数几何学所作的比较，前者是他最初涉足的研究领域，而对后者的研究则耗尽了他之后数学生涯的全部心血：

> 我仍然记得这种强烈的印象（当然完全是主观的），就好像我自己从贫瘠的荒野转瞬间突然来到了“神所授予的”富饶土地，它们无边无际，你可以尽情地在其中探究，施展自己的身手与才华。

大多数的数学家们都很务实地将自己看成是这个“数学世界”的探索者，他们并不是很关心它是否真的存在，他们只是用自己的直觉以及大量的理性思维来揭示这个数学世界的结构。这种直觉离所谓的“理想化的愿望”并不太远（就像法国诗人保罗 • 瓦莱利（Paul Valéry）所强调的那样），而大量的理性思维则需要高度集中的思考时间。

每一代数学家都构建了反映他们自己对这个数学王国理解的智力图景。他们建造了越来越敏锐的智力工具，这样就能够来开发先前未被发现的各种研究领域。

真正有趣的事情是：在数学王国的各个不同的领域之间找到了意想不到的联系桥梁，这种联系在前辈数学家们的智力图景中还是显得非常模糊和遥远的。而当这种情形产生时，就像突然之间一阵清风吹散了笼罩在我们美丽大地上的迷雾。在我自己的工作中，这种类型的巨大惊喜常来自于与物理学密切相关的数学研究。数学概念很自然地来源于物理学，这已经成为了一个基本共识，就像阿达玛（Jacques Hadamard）所曾经指出的那样。对他来说，他们所展示的

> 不仅仅是短暂的能让数学家们陶醉的新奇小技巧，而是那种从事物本身涌现出来的真正富饶多产的新颖。

下面我将用一些比较“实用”的忠告来结束这篇短文。只是要注意每一位数学家都是一个“特殊的案例”，不用太在意这些忠告。

散步

当你正在与一个非常复杂的问题搏斗时（常常涉及计算），一个非常明智的练习是出去走一段距离很长的路（不带纸和笔），一边走一边在脑子里做计算，不用担心这初看起来是否“太复杂了，不可能这样做”。即使不成功，也能够训练超人的记忆力并不断完善自己的方法。

躺下

数学家们一般很难向他们的同事解释，他们研究工作最辛苦紧张的时刻竟然是他们在黑暗中躺在沙发上的时候。很不幸的是，随着 e-mail 和电脑屏幕侵入所有的数学研究机构和场所，将自己完全孤立起来、从而集中心思的弥足珍贵的机会就变得十分稀少了。

再勇敢些

在通向新的数学发现的过程中，一般有好几个阶段。尽管处在后面的只需要我们理性与专心的（证明与计算）核查阶段的工作量大得惊人，但相比较而言，位于前面的更加富有创造性的（探索与构思）阶段则完全不同。在某种程度上，这个探索阶段还需要一种你对自己无知的保护意识，因为总是有千万个理由叫我们不要盯在那些其他许多数学家们都没有解决的问题上。

挫折

在数学家们的研究生涯中，包括很早的时期，他们会不断收到来自竞争者们的论文预印本，这时他们会因为落后而感到倍受打击。在这里我所能给出的建议是，应当努力将这种挫败感转化为鼓励你更加勤奋工作的前进的动力。然而这做起来并不容易。

怀有妒忌的认可

我的一个同事曾经说："我们（数学家）的工作，说到底就是为了得到几个朋友的怀有妒忌的认可。"确实，由于我们的研究工作本质上讲还是相当冷僻孤立的，所以不幸地是我们总是以各种方式极度渴求这种认同感，可是坦率讲我们不应该期望过高。实际上，真正的判断来自于自己。没有人能比自己更明白自己所做的工作究竟是什么，过分地在乎别人的看法是在浪费时间：因为迄今为止还没有一个定理是通过选举的方式获得证明的。就像费曼（Richard Feynman）曾经说过的那样，"你为什么要在乎别人怎么想？"

（本节脚注）

1 我自己最初的出发点是研究多项式根的分布问题。幸运地是我在很小年纪就被邀请参加在西雅图举行的一次会议，在那里我受到引导，后来我所有的在因子理论方面的工作都起源于此。

2 阿莉阿尼是古希腊传说中的克里特国王弥诺斯的长女。国王养了一头怪物，每年要吃七对童男童女。雅典王子忒修斯决心到岛上的迷宫里除掉怪物。在进迷宫前，他偶遇阿莉阿尼公主，公主爱上了他，交给他一个线团，让他将一端放在迷宫外，一端拿在手中，以免迷路。忒修斯杀死怪物后顺线走出迷宫。

4 杜萨•麦克杜夫的忠告

杜萨•麦克杜夫（Dusa McDuff, 1945-），英国数学家，英国皇家学会院士，在辛几何方面做出了杰出贡献。图右是其丈夫菲尔兹奖和沃尔夫奖得主美国数学家约翰•米尔诺（John Milnor, 1931-）。

我是以一种与我的同龄人很不相同的方式开始我的成年生活的。我总是被教导说要有一份独立的职业，我的家庭与学校也非常鼓励我成为一名数学家。与通常学校不同，我所上的女子学校有一位非常杰出的数学教师，他让我领略了欧几里得几何与微积分的美好。相比之下，我对教科学的老师就没有好感，并且由于大学里的科学教师也不是太好，以至于我从来就没有学到任何真正的物理学。

就是在这种很有限的条件下，我变得非常想成为一名专门进行研究的数学家。在某些方面我非常自信的同时，在另外一些方面我又感到十分的欠缺。我所面临的一个基本问题是，有时会莫名地感受到这样一种说法的影响：女性在学术职业上属于二等公民，因此是可以被忽略的。我从来没有女性的朋友，从来没有真正地评估过我自己的智力状况，可能还是和其他女性一样属于比较琐碎和实际的那种，并且不像男人那样真正具有创造力。社会上对于这种情况有许多的说法：妇女们待在家里烧火做饭，而男人们则外出闯世界，妇女可能会冥想，但不会成为诗人，妇女不具有成为一名数学家所需要的真正的灵魂等等。对此还有许多其他的说法。最近在我的一些女性朋友中流传着一封有趣的信件，它列出了各种不同科学领域中一些常见的互相矛盾的偏见，所得出的结论是，女性被认为不论做什么最有价值的事情，都是做不好的。

我不久又遇到的另一个明显问题是，我在只学了非常少的数学的情况下，就要着手完成我的博士论文。我的题目是有关冯•诺依曼 (von Neumann) 代数，这是一个狭小的题目，与其他我真正了解的事情没有任何联系。在那个小领域里我看不到前进的方向，并且我几乎对其他的领域也一无所知。当我在研究生的最后一年到达莫斯科时，盖尔范德（Israel Gelfand）给了我一篇论文，是关于流形上向量场的李代数的上同调的，我不知道什么是上同调，什么是流形，什么是向量场，以及什么是李代数。

尽管这种无知部分地是由于一种过分专门化的教育体系所造成的，但这也是我与更广泛数学世界缺少联系的结果。我曾经用基本上将自己的女性角色与数学家的角色分开来担当的方式，解决了将两者调和起来的问题。我从莫斯科回来后，我的这种数学上的孤独感愈发强烈。在我从泛函分析转向拓扑学的过程中，几乎没有任何指导，并且因为害怕暴露自己的无知而极少提问。此外，当我做博士后时还有了孩子，因此整天忙于应付各种琐碎事务。在这一阶段，由于对研究数学的过程还缺乏理解，我主要是靠阅读来进行学习，没有意识到学会提出问题的基本作用，哪怕是提出自己的也许是比较幼稚的想法。我对自己今后的职业规划也同样没有想法。天上不会掉馅饼：你必须去申请研究职位和工作，并且不断地寻找有兴趣的会议。如果能得到一位良师益友的帮助，那就更好了，他能够对你如何克服所有这些困难给出更好的建议。

对我来说，可能最重要的是学会怎样去提出比较好的数学问题。作为一个学生，你的任务就是：不仅要学习足够多的东西，以便能回答别人提出的问题，而且还要学习怎样去构造一个问题，以便能走向有趣的研究方向。在研究新的东西时，我经常习惯于从中间出发，运用一些别人已经发展起来的复杂理论。不过经常从最简单的问题和例子出发也能够获得进展，这是因为它们能使你更容易地理解基本问题，从而可能发现一条新的途径。例如在辛几何中，我总是很喜欢运用格罗莫夫（Mikhail Gromov）的非挤压（nonsqueezing）定理，它对如何以辛方式控制一个球体给出了限制条件。这个非常基本的几何结果在某种程度上与我产生了共鸣，并由此形成了开始探索的一个牢固的基础。

最近以来，人们越来越感觉到数学是一种依靠群体努力进行建设的学科：即使是最杰出的思想，如果它脱离了与数学整体的联系，也会失去意义。通常来说，一旦你理解了这种与整体的联系，再进行自我独立的研究工作就会变得非常重要和富有成果了。当然，在学习的同时，与别人的联系互动也是必不可少的。

有许多的方法可以成功地建立起这样的交流渠道，例如改变建筑物的结构、会议和会场的安排、数学系里教学研究计划的调整以及改进相对来讲不那么正式的讨论班与讲座的气氛等。有时候在讨论班上，一个新手数学家不是昏昏欲睡和看上去比较厌烦，而是积极地提问，使问

题得到澄清，并且使参加讨论的每一个人都获益，此时我们会惊喜地发现讨论班上的气氛改变了。人们（不管是年轻的还是年长的）通常都因为害怕暴露自己的无知、缺乏想象力和其他固有的缺点而导致沉默。但是在像数学这样看上去既困难又美丽的学科中，其实每一个人都需要从别人那里学到一些东西。现在有许多很好的小型会议和研究班，它们是这样组织的：既容易让人们学习讨论某些已有的理论的细节，又能够对新的方向和新的问题展开讨论和进一步的研究。

怎样将女性角色与数学家角色调和统一起来的问题依然困扰着我，虽然数学是一门本质上不适合女性的学科的观点已经不那么普遍流行了。我不认为我们女性应该尽可能多地出现在数学世界中，只要有足够数量的女数学家，并且不再作为例外而被忽视就可以了。我发现主要由女数学家参加的会议并不理想；满屋子全部都是女性，她们正在谈论着数学，这种气氛比较怪异。同样，人们也越来越能够理解，真正的问题是在于**任何**一个年轻人将怎样既能够过好一个令人满意的个人生活，同时又成为一个有创造力的数学家。只要人们开始以一种严肃的方式认真地对待这个问题，并且努力地工作，我们将真正地拥有一个长远的数学人生。

5 彼得·萨纳克的忠告

彼得·萨纳克(Peter Sarnak, 1953-), 美国数学家，英国皇家学会院士，美国科学院院士，专长数论，是最顶级数学期刊 *Annals of Mathematics* 的前任主编。

多年来我已经指导了不少博士研究生，这也许使我有资格以一个有经验导师的身份来写一些忠告。每当我遇到一个出色的学生（我非常幸运我能够有这样的一些学生），我所能给出的指导仅仅是告诉他，比如可以在某个区域内挖出黄金，或者给出一些含糊不清的建议。一旦他们行动起来，发挥其智慧与才能，结果他们没有发现黄金，但却找到了钻石（当然事后我忍不住会说“我告诉过你会这样”）。在这种情况下，和大多数的情形一样，一个讨论班导师的作用就更像一个教练：

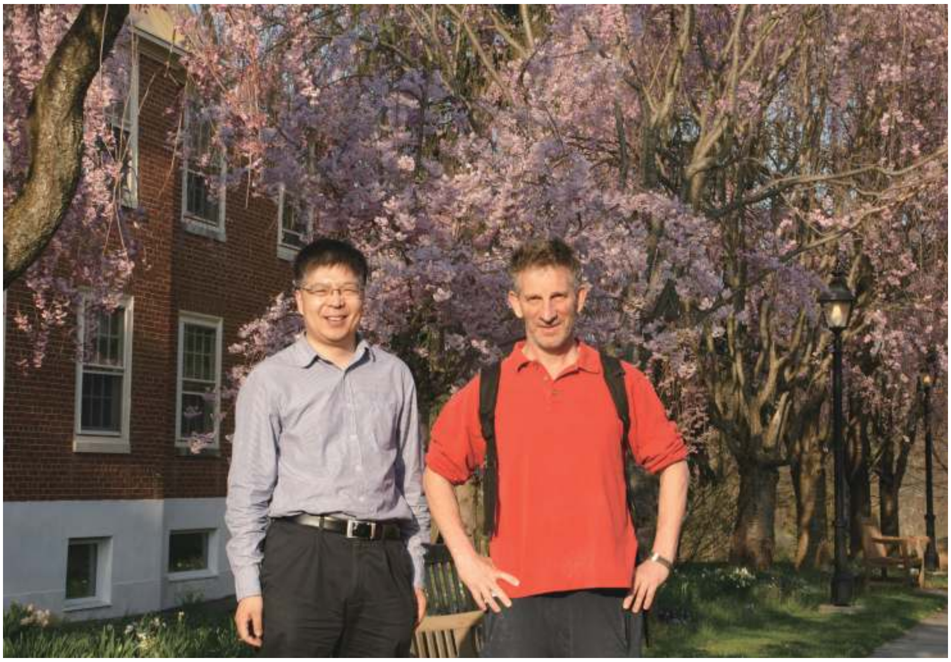
只是不断加以鼓励，并确信所指导的学生正在做的是一些有意思的问题，且清楚其可以采用的工具是什么。这么多年下来，我发现自己经常重复说的某些评论与建议对于学生来说还是很有用的。以下所列是其中的一部分。

（1）当我们学习一个新领域的知识时，我们应该将阅读现代论述与钻研原始论文结合起来，尤其是该领域开创大师的论文。很多学科的现代叙述所产生的主要麻烦是它们太完美了。随着每一个（数学专著与教科书的）新作者都不断地发现和加入更巧妙的证明处理方法，最后形成的理论体系总是倾向于采用“最简短的证明”。很不幸的是，这种形式化的表述经常引起新一代学生们的极大困惑：“人们是怎样想出来的？”通过回到原始的出发点，学生通常就能够看到概念与理论的演变十分自然，并且理解它们是怎样一步步变成现代形式的理论的。（当然面对着那些天才的数学家们所具有的令人意想不到的杰出思想，我们只能感到惊叹不已，但是这种情况要比你想象的少得多。）我想举一个例子，紧李群表示论有许多现代的表述，我在讲解其中的一种时，通常会推荐学生去阅读外尔（Hermann Weyl）的原始论文，看他是如何推导出他的特征标公式的。类似地，我会向已经了解复分析并且想要进一步学习黎曼面现代理论的学生，推荐他写的书《黎曼面的概念》，而黎曼面对于现代数学的许多领域来说具有最基本的重要性。研究和阅读像外尔这样的大数学家的论文选集同样也是十分富有教益的。在学习他们的定理的同时，我们可以发现他们的心智是如何运作的。从一篇论文到下一篇之间，基本上总是有一条线在自然地引导，而且容易看出某些后来的研究发展也是不可避免的。这一切都非常具有启发性。

（2）另一方面，你应当对一些教条和“标准猜想”敢于质疑，即使它们来自于某些大人物。许多标准猜想都是基于一些人们所能够理解的特殊情形作出的。除此之外，剩下的基本上只有人们多少有些一厢情愿的想法：人们很期望一般的图景不会和特殊情形所建议的图景相差太远。我知道几个这样的研究事例，在其中一开始的时候，人们都是先着手证明一个被认为普遍成立的结论，但是没有取得任何进展，直到最后人们才认真地反思它是否真的成立。在说了以上这些后，我也感到，如果不是出于特别好的理由，随便地去怀疑某个特定的猜想（例如黎曼猜想）及其可证性，的确有些不妥。虽然作为科学家的我们，肯定是应该采取一种批判性的态度（特别是针对一些我们数学家所发明的人造对象），但在心理上同样重要的是，我们要相信我们的“数学世界”的存在性，对什么能成立以及什么能被证明抱有信心。

（3）不要将“初等”混同于“容易”：一个证明可以确定是初等的，但却不是容易的。实际上，存在着许多这样的定理：只要用一点点（现代数学的）高端方法就可以使定理的证明变得非常容易理解，并显示蕴涵于其中的思想，相反如果避免使用高端的概念与方法，而只是用初等的方法来证明，则会掩盖定理背后的思想内涵。另一方面，也要注意不要将高端等同于高质量，或者等同于“高级证明”（这是一个我很喜欢用的字眼，它会引起许多我以前学生们的哄笑）。在年轻的数学家们中间确实有一种（盲目）使用新奇的高端数学语言的倾向，以显示他们正在做的工作比较深刻。然而，只有真正理解了现代工具，并且与新的思想相结合，现代的工具才能发挥作用。那些在某些领域（例如数论）工作的人，如果不花时间和实质性的努力去学习掌握这些工具，就会使他们处于非常不利的境地。拒绝学习和掌握这些新工具，就好像只用凿子来拆一座建筑物。即使你使用凿子非常熟练，别人用推土机也比你有巨大的优越性，并且用不着掌握像你那样（使用凿子）的技巧。

（4）在数学中做研究会让人感到受挫，如果你还不习惯于遭受挫折，那么数学就不是你的理想选择。在绝大多数的时间里，你是没有任何进展的，如果不是这样，则要么你是一个天才，要么就是你所遇到的问题属于在开始研究之前你已经知道怎样解决的那种。尽管一些后续的研究



2010年萨纳克与本刊主编之一在普林斯顿高等研究院（任秀敏 摄）

工作也会有相当的（发展）空间，并且达到较高的水准，但是一般来说绝大多数的重大突破都是用艰苦的工作换来的，伴随着许多错误的步骤，长时间里只有微小的进展，甚至还有倒退。有一些方法可以减轻这种痛苦。如今许多人采用合作研究的方式，这种方法除了有让不同专长的人一起攻关的明显优点外，还能让人们来共同承受失败。这对绝大多数的人们来说肯定是很有好处的（至少迄今为止，在数学中分享重大突破所带来的喜悦和荣誉还没有像某些学科那样导致许多严重的名次争议）。我经常劝我的学生在任何可用的时间里，手上要同时有一连串待研究的问题。其中，就是挑战最小的问题也应该有足够的难度，难到解决它以后会给你带来相当的满足感（不然又有什么意思呢？），并且幸运的话可能带动其他问题的解决。这样，你应该考虑一连串更具有挑战性的问题，其中最难的问题就是（该领域）最关注的未解决问题。你应该不时地考虑去攻克它们，从各种不同的视角来审视它们。很重要的是你要敢于让自己去解决非常困难的问题，不然就没有成功的可能性，也许幸运的话，你会从中获取许多。

（5）每周听系里的各种学术报告，并且希望报告的组织者能够挑选好的报告人。在数学中有比较广博的知识是很重要的。在学习了解其他分支领域里的人们解决有趣问题的进展时，或当你听到演讲者在谈论相当不同的研究时，你的心灵会经常受到某些思想的触动。同样，你也可能学到一种方法或理论，或许可以用到你正在做的其中一个问题上。在最近的一段时期里，有好几个长期未能解决的重大问题获得了最令人惊讶的突破，解决它们的方法都是来自于一些不同的数学分支领域思想的意想不到的组合。

译者简介：陈跃，复旦大学数学系77级本科毕业，上海师范大学应用数学硕士，现任上海师范大学数学系副教授，主要研究代数几何的历史，写有《扎里斯基与代数几何的抽象化》等文章。

《中国数学会通讯》是该刊编辑委员会编辑出版的内部刊物，季刊。主要刊登国内外数学界的重要信息，报道中国数学会与各省市自治区数学会、学科分会等的活动情况。主要栏目包括：学术活动信息，数学教育与普及，数学精品文章（数学的历史、进展、价值、趣事等），人物专栏，学科介绍，书讯与书评等。

2013 年《中国数学会通讯》全年的总订费为 50 元（含邮费）。欢迎各省市自治区数学会、学科分会和有关单位以及广大数学工作者、数学爱好者订阅本刊并踊跃投稿。

订阅办法：请将订费邮汇至北京中关村东路 55 号（邮政编码：100190），中国数学会办公室；或行汇至中国数学会，同时请给中国数学会办公室来信告知（或在汇款单附言中注明）订购份数、收刊单位（或个人）详细地址及邮政编码，以便我们及时准确地投寄本刊。

开 户 行：北京工商行海淀西区支行
帐　　号：0200004509089161419
电　　邮：cms@math.ac.cn
电　　话：0086-10-62551022

2013 年第 1 期要目：

- 2013 年数学界迎新春茶话会在京召开
- 中国数学会正副理事长、秘书长会议纪要
- 中国数学会十一届五次常务理事会会议纪要

华罗庚数学奖奖励条例
陈省身数学奖奖励条例
钟家庆数学奖奖励条例
2013 年中国数学奥林匹克
第四届全国大学生数学竞赛决赛在成都举行
创新乃科研之魂
中国多复变学科的创建
访学哥廷根
数学文化的传播和数学文化课程的推广
开拓创新图与超图谱理论

《中国数学会通讯》编辑部供稿

高等教育的危机

Nicholas Carr/ 文　施劲松 / 译

在线版本的大学课程吸引了无数的学生、千百万美元的资金和来自大学管理层的赞美。这是一种时尚，抑或是高等教育必须的一场彻底改造？

100 年前，高等教育似乎处在了技术革命的边缘。现代邮政系统这种强有力的新型通讯网络的广泛使用，使得大学有可能突破校园围墙的限制来传播它的课程。任何拥有一个邮箱的人都可以注册到一个班级里。威斯康星大学著名的历史学家弗雷德里克•杰克逊•特纳（ Frederick Jackson Turner ）写道，远程教育的“机器”，将要在全国范围内把“教育的溪流灌溉进干旱的地区”。很多学校意识到了这一可以招收新学生和获得新收益的历史机遇，于是都匆忙成立了函授部。到了 20 年代，函授课程变得繁荣昌盛，加入它们的人数是所有本国的学院和大学加在一起人数的 4 倍。

人们对远程学习早期形式的希望远远超越了原来的设想。因为可以为每位学生定制一些特别的作业和评估，所以许多教育家相信函授课程比传统的校园教导要好。国家最大的函授系之一，芝加哥大学函授系，告诉潜在的学生，他们将会获得针对其个人的关注。任何一个地方，只要有邮政服务，都可以提供针对于个人时间的服务。系主任声称提供给学生一种亲密的“重视学习中个人差异”的“辅导关系”，他说，函授教育将会证明它优越于“通常美国大学拥挤的教室”所能提供的东西。

今天，我们一直能听到类似的吸引人的声音。另一个强有力的通讯网络——Internet，再一次燃起了高等教育革命的热情。今年（2012 年——译者注）秋天，许多一流的大学，包括麻省理工学院、哈佛大学、斯坦福大学以及普林斯顿大学，都在网上提供了免费的课程，而且全世界范围内有超过一百万的人注册、学习这些课程。这些“巨大的公开在线课程”（massive open online courses），或者叫“MOOCs”，给众多的原本可能无法进入大学的人，包括那些在边远地区的人以及在职人员带来了优秀的大学教育，因而赢得了声誉。通常来说，无论对在校学生还是校外学生，在线教育这种方式正被人们用来作为提高教学质量的一种方法。美国前教育部长

Udacity 创始人之一塞巴斯蒂安•特隆

Coursera 创始人之一达芙妮•科勒

edX 创始人之一阿南特•阿加瓦尔

威廉•班尼特（William Bennett）写道，他意识到“一场类似雅典的文艺复兴”正在形成。斯坦福大学的校长约翰•亨尼斯（John Hennessy）告诉《纽约客》，他看到了“一场正在来袭的海啸”。

正在人们对州立大学教育越来越不满的时候，MOOCs 乘势而来。学士学位的平均学费蹿升至超过 100,000 美元。在大学校园里度过 4 年通常令年轻人或者他们的家庭被高额债务拖累，这不仅是个人财政也是整个经济的负担。甚至许多人担心，即使高等教育的费用增加了，但它的质量却会下降。而且辍学率居高不下，特别是在一些公立大学。对许多毕业生来说，很少有证据表明，大学改善了他们的批判性思维的能力。根据 2011 年由皮尤研究中心（Pew Research Center）所作的调查，接近 60% 的美国人认为本国的学院和大学未能提供给学生“与他们及其家庭所付学费相对应的教育”。MOOCs 的支持者们说在线指导的有效性和灵活性将会提供及时的补救。

但并不是每个人都热衷于此事。一些教育者们担心，在线教育将会给学院的管理层带来困扰。更有甚者，它们最终将会降低校园教育的质量。批评家们指出，早些时候函授教育的狂热就是一个警示。即使在大学都在匆忙扩展它们的函授教育规模的 20 年代，调查也揭示了指导的质量低于承诺的水平，而且仅有极小一部分学生完成了课程。1928 年，杰出的美国教育家亚伯拉罕•弗莱克斯纳（Abraham Flexner）在牛津大学的一次演讲中，提出了一个对函授教育极具毁灭性的控诉，声称它使教育的严格性付出了代价。到了 30 年代，曾经一度热衷于函授教育的员工和管理者，失去了通过邮件来进行教学的兴趣。狂热烟消云散。

这次有什么不同么？先进技术能否使得远程教育革命性的承诺最终被实现？我们仍然不得而知。围绕 MOOCs 的激情，很容易使人忘记它们仍然处于婴儿期。因此在初期，这种全新教育方式的优势和劣势正在成为焦点。

MOOCs 的兴起

“我也不知道那时我正在做什么。”在回忆去年当他决定免费在线提供斯坦福大学人工智能导论的课程时，塞巴斯蒂安•特隆（Sebastian Thrun）轻声笑着说。这名 45 岁的机器人专家有个预感，那就是通常招收数百名大学生的班级，在网络上报名的人也不会少。毕竟，他和他的合作教授彼得•诺尔维格（Peter Norvig）都是硅谷的明星，并且在谷歌拥有顶级研究岗位，外加在斯坦福大学教书。但是当特隆预想招生人数可能达到 10,000 名时，实际数字竟然整整超过了一个更高的数量级，当 2011 年 10 月开班的时候，居然有 160,000 人注册。

经历改变了特隆的生活。1 月，声称“我再也不会在斯坦福任教”的他，宣布联合其他两个机器人专家开展一项雄心勃勃的教育项目，称为“Udacity”。这个自称是“21 世纪大学”的企业，支付薪酬给来自诸如罗格斯学院和弗吉尼亚大学利用源发技术针对 AI 班级提供网上公开课的教授们。

Udacity 提供的 14 个班级中的大多数，被分成计算机领域和数学领域，特隆说，目前，Udacity 将主要着力于这些领域。但是他的理想几乎没有缩小：他认为传统的大学学位是一种过时的人工制品，而且他相信 Udacity 将要提供一种新型的更适合现代劳动市场的终生教育模式。

Udacity 是对 MOOCs 迅速增长的热情正在加以利用的几个公司之一。4 月，来自斯坦福大学计算机系的特隆的两个同事，达芙妮 • 科勒（Daphne Koller）和吴恩达（Andrew Ng），推出了一个叫做 Coursera 的类似项目。和 Udacity 一样，Coursera 是一个得到数百万美元风险投资的营利性商业机构。但和 Udacity 不一样的是，Coursera 正在和许多大的大学合作。特隆想发展一个传统大学的替代品，而科勒和吴恩达正在寻求构建一个系统，以使学院在网络上可以推送他们自己的课程。Coursera 新颖的模式不仅包括了斯坦福，还包括了普林斯顿大学、宾夕法尼亚大学和密西根大学，而且 2012 年暑假，公司宣布还有超过 29 所学校加盟。它已经提供了大约 200 个课程，从统计学领域一直到社会学领域。

5 月，在美国的另外一边，麻省理工学院和哈佛大学合作成立了 edX，这也是一家提供免费在线课程给所有申请学生的非营利性机构。edX 受到来自每所大学 30,000,000 美元的资助，它正在利用一个由麻省理工学院开发的开源教学平台。edX 不但包括了类似于那些盈利机构竞争对手所提供的视频课程和一些论坛，而且还整合了一些学生能够在那儿执行仿真实验的虚拟实验室。在 2012 年夏天，加州大学伯克利分校也加盟了 edX。9 月，项目推出了先期的 7 个课程，它们主要是在数学和工程领域。监督 edX 市场投放的是麻省理工学院计算机科学和人工智能实验室的前主任阿南特 • 阿加瓦尔（Anant Agarwal）。

Udacity，Coursera 和 edX 的领导者并没有限制他们自己增强远程学习的愿望。他们相信，在线指导，将会成为和在校学生一样的大学经历的奠基石。他们说，虚拟教室和现实教室的结合，将会推进学术界的进步。“我们正在彻底改造教育”，阿加瓦尔声称，“这将改变世界。”

机器人教授

在线课程并不是新生事物，大的商业机构，比如凤凰城大学和德锐大学就提供了上千门课程，而且许多公立大学允许学生通过网上修课获取学分。那么 MOOCs 到底有什么不同呢？在特隆看来，秘密存在于“学生参与”。到目前为止，大部分网络课程很大程度上都是由视频录像讲座构成，在特隆眼中，这种模式存在着极大的缺陷。他说，教室讲座通常是“令人乏味的”，而录播讲座更是缺乏吸引力，“你得到的只是最糟糕的部分，而没有丝毫最好的部分。”然而 MOOCs 包括了教授解释概念以及在白板上书写的视频；讲座也被分成一段段的简短的小节，而且不时地在屏幕上穿插练习和测验。特隆认为，事实证明一些强化的教学方法可以加强学生的理解和记忆，因此用问题来刺激学生可以使得他们参与到课程的学习中来。

2012 年早些时候，在 Udacity 班级讲授计算机程序设计的诺尔维格指出 MOOCs 和它们的前任之间的另一个不同。他说，在线教育的经济状况获得了令人瞩目的进步。云计算设备的使用，使得大容量数据的存储以及传送的代价变得很小。课程和测验在 YouTube 和其他流行的提供服务的媒体上免费地流动。而且像 Facebook 这样的社交网络提供了数字校园，在里面学生们能够形成学习群，互相回答问题。在刚刚过去的几年里，随着在线传送交互式多媒体课程的费用陡降，大量的学生可能无需交学费就能接受教育。

Udacity，Coursera 和 edX 都是由计算机科学家来领导的，这不足为奇。为了实现他们更大的承诺——使大学立刻变得价廉物美，MOOCs 需要在大容量数据处理和机器学习的开发上有最新的突破，以使电脑能适应手头的工作。同时传送复杂的课程给成千上万的人，对自动化有着很高的要求。许多被教授和助教按传统方式执行的劳动密集型任务，比如批改考卷、辅导、主持讨论等必须要让计算机来完成。还需要有先进的分析软件来解析课程期间收集到的关于学生行为的大量的信息。通过使用算法来识别数据的模式，程序员希望获得关于学习风格和教学策略的深刻理解，以使它们能够被用于技术的进一步提炼。MOOCs 的先锋们相信，那样的人工智能技术将要把高等教育带出工业时代，带进数字时代。

由于抱负远大，特隆、科勒和阿加瓦尔都强调他们新生的组织仅仅开始从课程中积累和分析信息。“我们还没有系统地利用这些数据，”特隆说。公司要将这些他们收集的信息转换成对教授和学生有价值的新的特征还需要一些时间。为了看到今天计算

机化教学的前沿，你不得不四处看看——特别地，要看一看一小群正在努力地将教学理论翻译到软件代码的学术测试和辅导机构。

这个领域最重要的的思想家之一是说话温和的纽约人大卫•昆茨（David Kuntz）。1994 年，在获得他的哲学硕士学位以后，作为一个在法学院管理委员会（管理 LSAT 考试的一家组织机构）任职的知识学家或者知识理论家，昆茨加入了美国教育测试服务中心（ETS），该中心负责进行大学入学的学术能力评估测试（SAT）。ETS 非常想利用计算机迅速增长的能力来设计出更加精确的考试和更加有效的打分。这使昆茨和其他哲学家面临一个非常大的问题：你怎样使用软件来测量意识、提升学习和评估理解？在万维网向普通大众打开 Internet 时，这个问题变得愈加紧迫。“电子学习”兴趣的激增，开发精致的教学与测试软件和设计出能激发学习兴趣的教育网站这两个方向的努力结合在了一起。

三年前，昆茨加入了曼哈顿一个名为 Knewton 的小项目并出任研究主管。公司专门研究自适应学习的初期训练。像其他教学软件的开拓者（包括隶属于加州大学欧文分校的 ALEKS，卡内基梅隆大学的开放学习项目（Open Learning Initiative ），以及非常著名的汗学院）一样，它致力于发展那种能够适应需求和贯穿他们整个课程指导的、针对个别学生学习风格的在线辅导系统。昆茨说，这个项目，“数据收集得越多就会变得越好。”比如说，针对代数教学的软件能够被写入，从而反映出其它可供选择的用于学习的理论，于是，当许多学生开始运行程序时，这些理论就能够被测试和改进，而软件也得到了提升。数据库越大，系统就会越加熟练地在正确的时间以正确的方式提供给每个学生正确的信息。

Knewton 介绍了一种针对即将入学的学生的数学补习课程，它的技术被纳入到由教科书巨头培生教育集团(Pearson)提供的辅导项目。但是昆茨相信，我们仅仅是刚开始看到教育软件的潜力。通过密集使用数据分析和机器学习技术，他预言这些项目将会通过好几个“层次的自适应性”得到改进，通过更先进的自动化，每个层次可以提供更多的个性化服务。初始层次在很大程度上已经到位，在这一层次中，一个学生通过一门课程所采取的一系列步骤，取决于学生的选择和响应。举例来说，回答一套问题，可以引发对其所掌握概念的进一步的指导，或者通过一个新主题的介绍来推进学生的进一步学习。昆茨解释道：“每个学生都可以选择不同的途径。”按照 Knewton 的计划，在即将达成的下一层次中，材料以能够自动适应每个学生的模式出现。尽管对媒介和学习之间的连接仍有争议，许多教育家还是相信，不同的学生可以通过不同的方式来学习。一些人通过阅读文本可以学得最好，另一些人是通过收看演示，而有些人通过玩游戏来学习，还有些人则是通过对话来完成学习。一个学生的理想学习模式，在课程的每个阶段是可以改变的，甚至是在每天的不同时间里都可以有变化。一个视频讲座对某门课程或许是最好的；但是对下一个课程，或许写作练习才是最好的。通过监控学生是如何与系统自身互动——比如什么时候快一点、什么时候慢一点、他们点击了哪里，一台计算机能够学习预测学生的需求，然后在媒介中传送材料以保证他们的理解和记忆能得到最大程度地巩固。

展望未来，昆茨说计算机最终将能度身定制一份适应每位学生的完整的“学习环境”。举例来说，程序界面的基本元素将会随着计算机感觉到的适宜学生的最佳学习方式而改变。

校园的大数据

辅导项目的先进性在于许诺可以帮助许多大学生、高中生，甚至小学生掌握基本概念。一对一的指导在很长一段时间内被认为可以提供实质性的教育效果，但是高昂的价格却限制了它的使用，特别是在公立学校。如果计算机能够代替教师，那么很可能就有许多学生可以享有辅导项目的好处。根据最近的一项针对本科生在公立大学选择统计课的研究，最新的在线辅导系统似乎和面对面的指导可以产生大致相同的效果。

当 MOOCs 正在将自适应学习程序融合进他们的软件的时候，他们关于数据挖掘的野心远远超出了辅导系统。特隆说我们仅仅看到了“冰山一角”。由于免费在线课程达到了一个史无前例的规模，使得他与其他计算机专家感到特别兴奋，如此规模足以产生有效的机器学习所需的海量数据。科勒说 Coursera 已经建立起了自己的密集型数据收集和记忆分析的系统。课程中的每一次变动都被追踪，当一个学生停下了视频或者增加了回放速度，这个选择就会被 Coursera 的数据库所捕获。同样，当一个学生回答了一个测验问题、修订了一个作业，或者在一个论坛中给出了评论的时候，信息也会被数据库所收集。每一个动作，那怕看上去或许无关紧要，

都会给信息数据的统计提供材料。

科勒说，对学生行为细节的信息汇总达到了分钟级别，这“开辟了理解学习的新的途径”。学生们探索和掌握复杂题材的早期隐藏模式能够被揭示出来。

这种交互方式也承诺将直接给教师和学生带来益处，她补充道。教授们将会收到关于什么在他的班级里起作用，或者不起作用的定期报告。而且，通过确定“最有预见性的成功因素”，MOOCs 软件最终将能够指导每一名学生走上“正确的轨道”。科勒说她希望乌比刚湖——一个“所有学生的智力水平都要比平均值高”的虚构的小镇，将“成为现实”。

麻省理工学院和哈佛大学正在将 edX 设计成为像数字教学平台一样的一个教育研究工具，阿加瓦尔说道。学者们已经开始使用来自于这个系统的数据，以测试关于人们是怎样学习的猜想。而且随着课程系列的增长，研究机会将会激增。除了产生教学的见解，阿加瓦尔还预言了许多针对 edX 数据仓库的实际应用。举例来说，机器学习可以为在线课程中的自动化欺诈检测系统铺平道路，因为这些欺诈行为已成为大学给那些完成了 MOOCs 的学生授予学位甚至给予学分而面临的愈加紧迫的挑战。

看似即将来临的数据爆炸，很容易地就抓住了 MOOCs 建设者们的热情。虽然他们的工作主要集中在计算机上，但他们的目标还是有着深刻的人本精神的。他们指望利用机器学习来培养学生学习，在对人的智力服务中部署人工智能。但是这个热情应当被一种质疑精神所调和。教育中机器学习的益处很大程度上还停留在理论上。而且即使 AI 技术在教学法中产生了真正的进步，那些突破还是只有有限的应用。对程序员来说，当知识主体可以被明确地定义，而且学生的进步可以被明确测量时，实现自动化课程指导是一件事；而尝试在计算机屏幕上复制错综复杂的、有时难以形容的、发生在大学校园里关于教与学的经历，则是另一件完全不同的事了。

MOOCs 的推广者有着一种“关于对大型数据集分析所允许的相当天真的看法”，斯沃思莫尔学院（Swarthmore）的历史学教授蒂莫西 • 伯克（Timothy Burke）说道。他主张，远程教育低于历史的预期不是因为技术原因，而是由这种模式所具有的“深刻的哲学问题”导致的。他认为，在线教育可以在计算机编程以及可以编纂在软件里的固定程序所描述的其他领域提供有效的训练。但是他主张大学教育的本质存在于学生和教师之间微妙的相互作用，无论有多么精心的程序设计，这种相互作用都不能被机器所模仿。

阿兰 • 雅各布斯（Alan Jacobs），伊利诺伊州威顿学院的一位英语教授，提出了类似的观点。在一封他发给我的电子邮件中，他关注到大学生的工作，“会因为他们在教室里所面临的夸张情形以及在实时同步里邂逅的其他人的反应，而戏剧性地受到影响”。这种对话的丰富性是不可能在 Internet 论坛上被复制的，他争论道，“除非在线写作的人拥有技巧高超的小说家的能力，以散文的形式来表达复杂的思维方式和经历”。计算机屏幕绝不会比好大学的一间教室的影子更好。和伯克一样，雅各布斯担心，MOOCs 所反映出的教育理念已经被正在发展平台的计算机学家们曲解了。

翻转课堂

MOOCs 的设计者和推广者们，并不认为计算机将会使课堂变得过时。但是他们认为在线指导将改变校园教学的性质，会使校园教学变得迷人和高效。那种学生到班级里听讲座，然后自己低头完成作业的传统指导模式将要被倒转。学生们将要在他们的计算机上独立地完成听讲座、复习其他解释性的资料（就如同一些初中和高中已经在利用汗学院的视频那样），然后他们集中到教室里，通过和教授们讨论，或者通过实验，来进行较为深入地题材探索。理论上，这种“翻转课堂”将使教学时间的分配更加合理，丰富了教授和学生两者的经历。

当然，这里面也有困惑。值得关注的一个原因就是早期困扰 MOOCs 的高辍学率。在招收了 160,000 名学生的诺尔维格和特隆的 AI 班级里，最终只有 14% 的学生完成了学业。2012 年早些时候，在 155,000 名报名参加麻省理工学院电子电路课程的学生中，仅有 23,000 名学生艰难地完成了第一套问题。大约 7,000 名或者 5% 的同学通过了该课程。不论以哪一种标准，指导成千上万的学生通过一门学院课程都是一种非凡的成就，因为通常每年仅有 175 名麻省理工学院的学生能完成电路课程。但是高辍学率，还是使如何保持在线学生的专注性和积极性这一问题变得突出。诺尔维格承认 MOOCs 最初的入学者是个有着独特品质的自我激励的群体。在大量而且较为典型的一群人选了课程以后，对在校学生的在线指导，将要迎来真正的考验。当

学生们坐在电脑面前进行数周时间学习的时候，MOOCs 将不得不激发各种各样的学生并保持他们的学习兴趣。

对 MOOCs 的批评中，最大的担心是学院在还没有仔细评估可能的缺点之时就急于将在线指导结合进传统课堂。2011 年秋天，在共同创办 Coursera 之前的较短时间内，吴恩达改编了他的斯坦福大学的关于机器学习的课程，以使在线学生能够参与，于是成千上万的人报名了。但是至少有一名学生发现了班级的不足。在博客上，主修计算机科学的本•鲁道夫（Ben Rudolph）抱怨地写道，“学术严谨性”达不到斯坦福的标准。他感觉这种提供自动且即时的提示和指导的计算机化的作业，在激励“批判性思考”上是失败的。他还提到了一种隔离的感觉，就是他“几乎碰不到（这个）班级里的任何一个人”，因为“每一件事情都是在我的房间里单独完成的。”吴恩达坚定地辩护着班级的形式，但是实际上，没有人能真正地知道，逐步增加的计算机化的指导，到底是怎样改变动态的学院生活的。

MOOC 运动的领导者们承认他们面临挑战。阿加瓦尔说，为使模式更好，必须在许多领域具有“精致的发明”，包括从论文分级到证书授予。当在线课程进一步延伸至人文学科的开放探索领域时，这会变得更加困难，因为其中的知识很少有容易编纂的，而且一个班级的成功取决于一个教授拥有的引导学生获得出人意料见解的能力。MOOCs 耕耘的成果，应该更多地告诉我们关于班级的价值，以及在教育系统里他们最终将要扮演的角色。

至少，在线教育对大学提出的基本问题，就像技术挑战一样令人生畏。无论大规模的开放课程是否能履行他们的大肆宣传，它们都要迫使学院管理层和教授们在教学形式和意义等方面重新考虑他们的设想。无论好还是坏，网络破坏性的压力已经来到了学术界的大门口。

致谢：在本文翻译过程中，得到了原文作者 Nicholas Carr 先生的授权以及美国韦恩州立大学数学系郭海龙、王泗奎两位博士生的协助。

后记：

原文发表于《MIT Technology Review》

原文链接：

http://www.technologyreview.com/featuredstory/429376/the-crisis-in-higher-education/

版权问题：

https://s100.copyright.com/AppDispatchServlet?publisherName=mittechreview&publication=1099-274X&orderBeanReset=true&publicationDate=2012-09-27&author=Nicholas%20Carr&title=The%20Crisis%20in%20Higher%20Education&contentid=429376&cc=Computing&ap=Featured%20Story&sp=November/December%202012&end

作者简介：Nicholas Carr（1959-）美国作家，出版了关于技术、商业和文化等方面的多部图书。其著作 *The Shallows: What the Internet Is Doing to Our Brains* 曾入围 2011 年一般非小说类普利策奖。

译者简介：施劲松，中国华东理工大学数学系副教授，美国韦恩州立大学数学系访问学者。一直从事基础数学教育、网络教育及其研究。

解释离婚情感动态的数学模型

Jose-Manuel Rey ／文　　李玉田／译

摘　要

背景

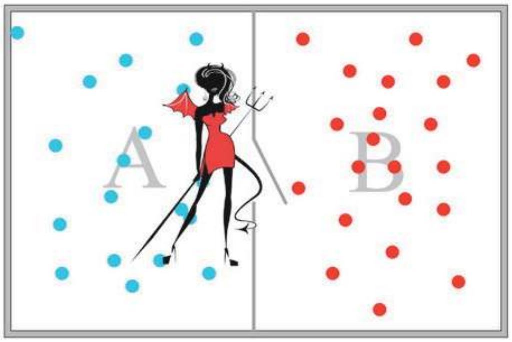
西方社会的离婚是普遍存在的。这无论在理论上还是解决方式上都引发了很严重的科学与社会学问题。学者和问题处理专家认为存在一种**情感关系的热力学第二定律**。仅有爱是不够的，还需努力来维持。

方法 / 主要发现

我们在一个简化的第二定律的基础上用最优控制理论作为建立情感动态模型的一个新处理。我们的分析结果与社会学数据相一致。我们揭示出，如果双方有相似的情感属性，存在一种最优的努力策略使得婚姻长久幸福。这一策略因两方面因素的组合而受到结构性不稳定的制约：一是努力差距，因为最优策略通常带来不愉快；一是由于动力系统的不稳定性使得努力降低到不可持续水平的趋势。

结论 / 重要性

这个模型的数学事实揭示了真实情境中配偶关系破裂的相关机制。在这个框架下，第二定律可以解释一个显见的悖论——一贯想天长地久的婚姻可能会破裂。



引 论

有着浪漫特性的情感关系通常被认为是西方社会中安定的幸福生活的基本组成部分。当人们被问及什么是幸福生活的必备要素时，大家通常将“爱”和“亲密关系”摆在首位。很难想象人类生活的其它方面会涉及如此多的文化、社会、心理及经济学问题。理论研究显示浪漫关系的最初阶段似乎是受化学作用控制，而维护一段情感关系则在较大程度上属于理性决策的范畴。人们通常在仔细考虑之后才投入长期的关系，最典型的就是婚姻。即使在一直是一夫一妻制的西方社会，配偶通常声称他们的意图是维持长久的关系并和睦相处。但欧美大量报道的高离婚率表明了他们在方案执行上的彻底失败。配偶关系破裂这一现象在美国被视为流行病，“每两对夫妇中就有一对以离婚收场”这一统计结果被媒体和学术报告频繁引用。欧盟 27 国的离婚率并不比美国好到哪里去，而且一些欧洲国家表现出了更高的离婚率。

诸多领域的学者们普遍认为，婚姻不稳定性日益加剧的主要原因应归结为二十世纪性别分工的变化所释放的经济力量。然而这一原因并不能解释过去几十年里不断增长的、无处不在的婚姻破裂现象。事实上，人们并不理解为什么这么多配偶最终离婚而有些人却没有离婚。这一理解尤为重要，因为婚姻破裂所诱发的社会变化深深地影响着当代西方社会的社会结构以及福祉。

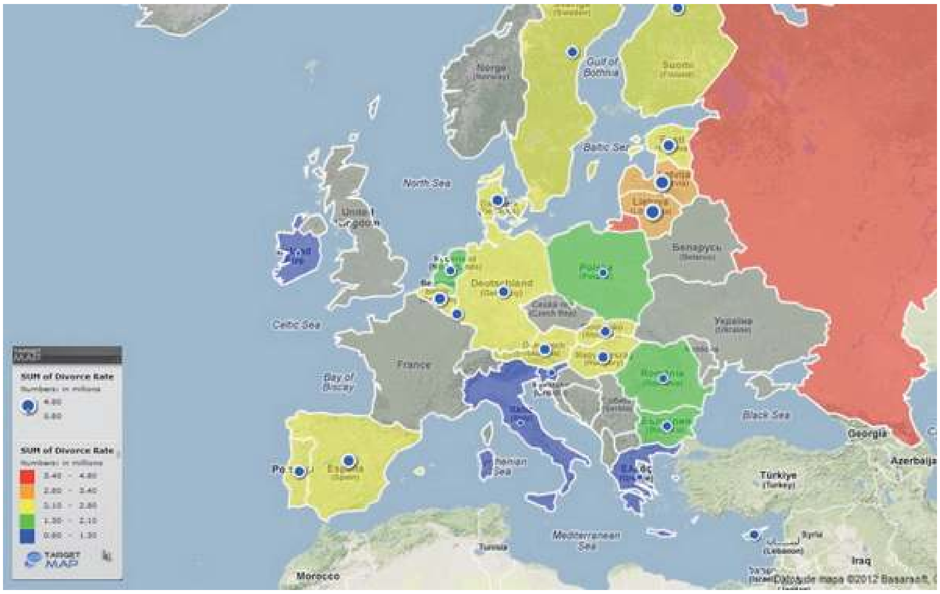
对大多数配偶来说，双方都打算维持长久的关系并为这一目标而努力，这个事实却与报道出来的高离婚率相矛盾。本文称这种矛盾为**失败悖论**。根据戈特曼（J. M. Gottman）等人的研究，婚姻研究领域亟需一个理论，尤其是数学理论。本文旨在满足这一需求。尤其是本文将为这种失败悖论提供一个合理的解释。

戈特曼等人的研究（收录于 *The Mathematics of Marriage-Dynamic Nonlinear Models* 一书）似乎是迄今为止对配偶关系所进行的唯一的一项数学研究。他们用一对非线性差分方程评估了双方在实验观察时的短期互动。斯特罗伽茨（S. H. Strogatz ）率先提出了一个配偶互动的简单动态系统模型。我们这里采用不同的动力系统方法：配偶被视作一个单元（不考虑内部互动），他们的情感动态由其长期和睦共处的意愿来合理地描述。

鉴于配偶分手现象的普遍性，相对于情感关系中的具体瑕疵，寻找能解释分手现象的基本的确定性机制似乎更为合情合理。在社会学数据的基础上，我们提出了一个基于最优控制理论的数学模型来解释同类型配偶如何建立长期关系的理性计划。如果配偶双方有相似的特性，我们称之为同类型配偶。同类型婚配是西方社会情感关系中最普遍的的类型。我们的模型实际上需要的是一种较弱的同类性。我们通过基于情感互动的**热力学第二定律**（以下简称**第二定律**）的动力学方程来描述情感关系的演化。第二定律指出，除非有“能量”供给，否则一段关系将会破裂。这一被广为接受的事实允许我们将情感关系建模成一个控制问题，以努力这一形式存在的能量充当着控制变量。最优控制理论已经被广泛地应用于应用科学，如工程学和经济学。我们的最优控制模型为分析婚姻和亲密关系提供了一个全新的数学方法。

给定一些可行性的条件，我们对于模型的分析表明维持长期成功的关系是可能的，并且可对应于动力系统的平衡路径。然而很明显的是，长期关系中不可或缺的是努力，这一模型最引人注目的发现是保持幸福关系的努力水平**总是**比先验的最优选择（假如只考虑目前情形）的努力水平要高。当这两种水平的**努力差距**在可容忍的范围内，维持一段关系是切实可行的。这一数学分析的最主要结果是：服从于第二定律的情感动力系统本质上是不稳定的。这说明，当努力放松时，情感的逐步恶化会轻易地发生。这一分析还确立了一个合理的机制用来解释逐渐退化将导致感情破裂抑或使人不满的情感生活。

本文的研究结果有助于解答关系失败悖论：根据第二定律，长久幸福关系的优化设计与动态不稳定性并由此导致的可能的分手相容。这一显著的发现解决了这一失败悖论，



欧洲各国和地区 2010 年的离婚率

因为真实的关系被认为有更多不稳定性和不确定性的源头。而且这一结果也许揭示了如何才能保持健康有活力的长久关系。

本文的第 2 部分将展示由社会学数据所支持的关键证据，它们将构成一个检验模型结果一致性的框架。失败悖论这一问题也由社会学证据来导出。随着对于基本假设的深入讨论我们将逐步提出这一模型的构成要素。对于该模型分析得出的主要预测将汇集在本文的第 3 部分，其中一些预测与第 2 部分中提到的事实保持一致。为了便于讨论，数学上的一些技术细节被放置到附录中。

方　法

公认的事实

马丁和巴姆帕斯（Martin and Bumpass ）使用 1985 年的数据表明，在 40 年之内，美国每三个婚姻中就有两个将以分居或离婚而结束。目前可能还没有达到这一比例，但是 2002 年的数据显示实际数字并不低多少。50% 的人在 40 多岁的时候已经离婚不止一次。2005 年的数据表明，报道中经常提到的 50% 这一数字只比欧盟 27 国的平均离婚率（44%）稍高，而在一些欧盟国家，这一比例甚至高达 71%。

如果包括未婚同居，这一数据将进一步上升，尽管基于同居状态的数据很难获得。一项最近的研究证实，总体上非婚姻同居比婚姻更不稳定：49% 的婚前同居在 5 年内分手（10 年后分手率是 62%），然而只有 20% 的婚姻在 5 年内以分居或离婚告终（10 年后这一数字是 33%）。我们正在观察的这一现象的第一个公认的事实可以阐述如下：

断言 1　在爱情关系中存在普遍的失败。

这种臭名昭著的情感关系不稳定性与婚姻或同居关系中的信任缺失并不相关，虽然信任是幸福的主要因素。相反地，人们普遍表示满意的情感关系是建立幸福生活的第一要素。并且人们还声称他们希望自己的伴侣伴随一生：

断言 2　配偶通常相信关系的持久性是追求幸福的重要部分。而且大部分人认为，他们自己的关系不会破裂。

现有的数据支持断言 2。当被问到使得他们幸福的要素时，78% 的美国大学生选择了“与自己理想的伴侣相爱并维持下去”。在美国的一次国家调查中，93.9% 的受访已婚夫妻认为他们离婚或分居的可能性很低 (19.9%) 或相当低 (74%)，而 81.1% 未婚受访者也有同样的回答（32.4% 的受访者认为离婚可能性低，47.7% 的受访者则认为非常低）。

有趣的是，尽管公认分手的可能性非常高，绝大多数的人仍然认为他们自己的情感关系不会破裂。事实上，断言 1 与断言 2 提出了一个明显的悖论。根据上述的数据，一个新组成的伴侣声称有九成的把握走到最后。然而 5 年之后分手的同居者是 50%；10 年之后他们则更有可能不在一起。这一事实可以叙述如下：

失败悖论：打算长久持续的情感关系却怎么极可能破裂？

下面提出的模型表明，在合理的假设之下，断言 1 和断言 2 是相容的。为了进一步测试该模型的一致性，我们还将考虑其它两个公认的事实。

断言 3　配偶分手是一种逐渐恶化过程的结果。

已有的数据支持这一事实。根据加州离婚调解计划中 80% 受访男女的情况，他们离婚的主要原因是“逐渐产生的距离感和渐渐失去的亲密感，或许呆在一起但是情感上却逐步疏离，以致再也不能忍受孤独感”。

断言 4　配偶双方的主观幸福感在婚后逐渐降低。

尽管都认为婚姻比起独身有着较高的幸福感，但已婚人士自己感受到的满足感却在慢慢降低。研究人员发现，结婚后人们对于生活的满足感逐步下降。

模　型

依据以上情景，我们将建立一个简单的动力系统模型。

这一模型基于两个关键的假设，也就是下面假设 2 中要讨论的第二定律和上述断言 2 所说的配偶关系的长期规划。这些假设连同弱同类性（参见下述假设 1）以及情感状态的一个自然的成本 - 收益评估（下面的假设 3），允许我们将配偶的情感关系视为一个最优控制问题。

建模始于（时间 $t=0$）浪漫的热恋期刚刚结束，对彼此的好感达到顶点之时。在这一起始时间，伴侣们有强烈的意愿希望能最终走到一起，并试图去做任何必需的事情以保证两人的长远未来。我们假设：

假设 1 （弱同类性）伴侣双方拥有下述模型所指的相同特质。或者说，一对伴侣作为规划问题的一个决策单元。

这一假设预示着模型中所有的参数、变量和效用结构都是定义在一对由两个类似的个体组成的伴侣上。人们倾向于被具有相同特质的人所吸引，这一事实已经在现有文献中获得广泛认同。西方社会中有充分的证据支持这一事实。所以假设 1 被视作一个法则，而不仅是一个期望。严格来说，我们的理论仅仅要求情感而不是性格上的相似性（参见下面的假设 3）会促成配偶关系，不论是一起约会或者是已婚夫妻。

根据前面的叙述，下面的假设对我们的模型至关重要。

假设 2 （情感关系的热力学第二定律）对于另外一方的感觉有一种消退的趋势。这种惯性必须由有意识的实践来抵消。

在诸多文献中对这一事实有着广泛的共识。没有精心维护的爱会随着时间消退似乎是一个自然法则，雅各布森（N. S. Jacobson）和马戈林（G. Margolin）认为这是婚姻不稳定的主要诱因。他们写道：“婚姻以幸福开始，但是时间的侵蚀是婚姻问题的源头。”流行的格言“爱还不够”反映了这一事实并含蓄地表明侵蚀是可以通过某种办法防止的。假设 2 这种定律的形式取自戈特曼等人的书里，在那里，情感被建议性地解释为“类似于婚姻关系的热力学第二定律的某种东西：一种缺了保持关系鲜活良好的能量就会分崩离析的东西”。

为了将假设 2 转化为数学语言，我们用一个非负变量 $x(t)$ 来表示在 t 时刻 $(t\geq 0)$ 的关系状态。它是一个**感情变量**，可以理解为伴侣对对方的情感。$x(t)$ 是一个有序变量，它用来探测情感关系的质量水平。$x(t)$ 的一个特定值并不包含任何信息，但是在不同时刻 t_1, t_2 的情感水平可以进行比较，比如 $x(t_1)\geq x(t_2)$ 还是 $x(t_1)\leq x(t_2)$。在 $t=0$ 时刻共同的感情水平 $x(0)=x_0$ 被设定为一个非常大的数值。我们假定 $x(t)$ 降低到低于某个阈值 $x_{min}>0$ 时，关系开始变得令人不满意，不同的伴侣的阈值不尽相同。

根据假设 2，对关系的细心处理可以抵消这种退化的惯性。这种处理用一个非负的有序变量 $c(t)$ 来表示，这个变量被称为**努力变量**，并假定这个变量是分片连续的（请看附录 1 中的讨论）。$c(t)$ 包含所有日常生活中加强情感关系的实践，例如，治疗师提出的建设性行动（发问、积极地聆听、一起制定计划），以及宽容的态度（接受配偶的缺点、给她 / 他隐私、尊重不同的口味和习惯），这些只是其中一少部分建议的行动。不论是主动还是被动，努力或者说牺牲的重要性以及对情感关系的益处已经被人们广泛的注意到。

一个简化的第二定律可以用感情变量和努力变量组成的一个微分方程来表示

$$\frac{dx(t)}{dt}=-rx(t)+ac(t),\quad for\ t\geq 0, \tag{1}$$

这里 $r>0$, $a>0$。在没有介入的情况下（也就是 $c(t)=0$），方程 (1) 意味着 $x(t)$ 以 r 的速度递减，这个速度 r 与每对配偶有关，用来表示这对配偶情感退化的力量。这个为人所熟知的简单的线性法则描述了很多自然界和社会现象。事实上，它的一个离散化版本在文献中被用来描述未受影响的配偶在短时间内的婚姻互动中的最初演化过程。对于任何速率，方程 (1) 中 $c(t)=0$ 是对感情退化一个明显的假设。在方程 (1) 中“努力”被用来抵消这种感情退化。很明显，其中的参数 a 表示的是**努力的效率**。选定一个努力策略 $c(t)$，方程 (1) 就给出了情感的演化过程。方程 (1) 隐含地表明 $x(t)$ 是光滑变化的，除非在努力是不连续的情况下。

付出的努力 $c(t)$ 的强度由配偶对来决定，这与（非理性）变量 $x(t)$ 的水平不同，$x(t)$ 不能由配偶决定。努力变量 $c(t)$ 的理性特质允许我们将它解释为最优控制理论里的一个**控制变量**。在这一假设下，受控变量，也就是**状态变量**，是 $x(t)$，并且方程 (1) 是联系两个变量的**状态方程**。

我们接下来也是最后一个假设指的是对于努力和情感水平的成本 - 收益评估。我们考虑的是一个标准的效用性的途径。对于情感演化的感觉的数学表达是较为直接的（参见下面的假设 3）。然而，对于努力评估的表述却需要一番考量。最典型的努力是牺牲，为了亲密关系而忘掉自身利益，这种潜在的收益和成本已反复地在文献中被考虑过。对于牺牲和相关实践的实证研究表明，付出的努力可能同时蕴含着情感成本及其收益。这一明显的矛盾通过对牺牲行为的动机分析得以消弭，而这一动机分析则基于积极处理与刻意回避的不同态度。选择试图取悦伴侣的态度可能会产生积极的情绪影响，而刻意回避则可能引发紧张与痛苦。我们对于付出努力造成的情感差距的解释与努力的强度有关，因为我们认为在一定水平上的努力会获得感情回报，但是超出这一水平就会付出代价（痛苦）。这些形成了下面的假设。

假设 3 （效用结构）有两个不同的效用来源。一个来自于情感水平，另一个是努力强度的后果。

i) 来自感情的效用可以用一个可微函数 $U(x)$ 来表示，$U(x)$ 满足 $U'(x) > 0$, $U''(x) < 0$，当 $x \to \infty$ 时，$U'(x) \to 0$。换句话说，对于任何感情水平，它的边际效用是正的，但是随着感情越强边际效用递减到 0。

ii) 努力 $c \geq 0$ 的负效用用一个可微函数 $D(c)$ 来表示，它满足 $D''(c) > 0$，存在一个 $c^* \geq 0$ 使得 $D'(c^*) = 0$，并且当 $c \to \infty$ 时，$D'(c) \to \infty$。也就是说，努力引发的不满在 c^* 达到一个最低水平，并且随着努力水平的不断增加，对婚姻的不满足会升高到无限大。

注意我们并不需要 U 和 D 的数学表达式。这个理论对于满足以上性质的所有函数都是有效的。

效用这一术语可以与幸福感或者对生活的满足感互换。i) 的假设是消耗一些美好的事物所达致的效用的一个标准。定义在感情上的效用并非一定是一种上层结构：当 x（一个人的感情）与一段关系的情感有关系的时候，$U(x)$ 就产生一个关于这个感情水平 x 的评估，这个评估依据的是个体的判断并或许依赖于以往的经验以及人格特质。比如说，两对配偶对于类似的感情水平或许给予很不相同的评价，所以他们的评估会基于不同的效用函数。感情效用函数的存在性这一假设可以像基于消费的效用的论证方式一样来进行明确合理的探讨。

函数 D 表达的是负效用，它基于的是额外付出的努力会导致对效用的一种消耗。因此它的负数（$-D$）可以理解为一种效用。图 1 给出了这两个函数的一个典型的形状。

在这个模型的动力学假设中，U 和 D 是用来测量瞬间（也就是此时此刻）感情和努力的效用和负效用。D 是非单调的假设反映的是努力在一定的水平之下会自我觉得是值

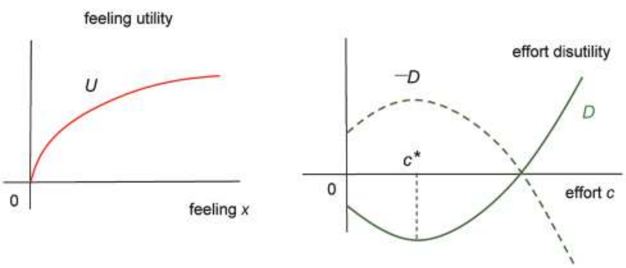


图 1 效用结构：效用与负效用函数的典型形状

得的这一事实。举例来说，计划与配偶的一次消遣活动，这不需要太大的努力并将产生一定的愉悦心情。尽管此刻的努力所达致的将来的益处隐含在感情效用中，因为此刻的努力为的是通过方程 (1) 来提升将来的感觉，如果 D 一直是非递减的，那么此刻努力所达致的即时的益处就没有被考虑到。

当努力水平很低的时候，做出小努力看起来是愉快的，然而当努力水平较高的时候，努力则会付出较大的情感代价。所以在假设 3 的 ii)，当努力在达到水平 c^* 之前所造成的效用是随着努力水平增加的，而在这个水平 c^* 之后随着努力水平的增加，效用却是减少的。这个参数 c^* 代表的是一对配偶的一个先验的最佳努力水平，它在分析中扮演着关键角色。这个理论也允许 D 是一直单调的特殊情况，也就是 $c^* = 0$ 的时候。这对应的是，即使努力水平最低，(此刻的）努力也会导致（此刻的）烦恼这一情形。所建议的 D 的结构似乎允许更加合理的情形。

一对配偶面临的问题就是如何制定一个努力策略，使得他们的关系持久并给双方最大限度的满足感。因此努力水平的演化是由追求最大幸福感的一个理想标准来决定的。这是一个优化问题，可以作如下地描述

问题（P）情感动力学的努力控制问题：假设感情的演化过程由方程 (1) 来描述，效用结构由假设 3 来描述，最初的感觉水平 $x(0) = x_0 \gg 1$，用 $\rho > 0$ 表示不耐烦的因子。在这些条件下，寻找一个努力计划，$c^{\heartsuit}(t) \geq 0$ 对于 $t \geq 0$，使得净效用最大并且双方的感觉和努力的演化是长期可持续的。

总的满意度是 $t \geq 0$ 所有瞬间净效用的总和，它可以用一种标准的方式来表达

$$W = \int_0^\infty e^{-\rho t}(U(x(t)) - D(c(t)))dt.$$

（指数项对应于对将来效用的一个评估。）问题 (P) 是一个标准的无限水平最优控制问题。由于断言 2，这一问题的计划时期设定为无限长。持久性的问题，作为配偶问题的一个关键要求，关注点有两个：可容许性与可行性。不仅长期的感觉和努力水平一定要在可容许的范围内（也就是说感情必须在水平 x_{min} 之上），而且向这些渐进水平的过渡一定要是可行的。

结果和讨论

这个模型所蕴含的主要结论在以下导出并加以评论。引人注目的是，断言 3 和断言 4 中所述的一些实证证据将由这个模型的分析理论地推导出来。并且，在这一模型框架下，断言 1 和断言 2 是相容的，这在一定程度上解决了失败悖论。分析的数学细节在附录 1 中。时刻 t 的最优努力必须满足

$$\frac{dc}{dt}(t) = \frac{1}{D''(c)}((r+\rho)D'(c) - aU'(x)), \quad t \geq 0. \qquad (2)$$

方程 (2) 给出了最优努力的变化法则。方程 (1) 和方程 (2) 组成了一个描述最优的感情和努力水平演化轨迹的方程组。这个最优轨迹记为 $(x^{\heartsuit}(t), c^{\heartsuit}(t))$。

情感平衡

如果方程（1）和（2）的稳态解是可行的，那么基于一套固定的努力程序，则可保证达成长久幸福的感情生活。享受永久的情感回报，却无需担心付出努力所引起的情感波折，显然是情感动力学一个吸引人的特点。这使得平衡态成为一个维系长久关系的理想结构。

存在性和可行性 平衡态由方程 (1) 和方程 (2) 中关于时间的导数为 0 来描述。基于对该模型的详述，在附录 1 中证明了存在唯一的明确定义的情感平衡状态 $E = (x_s^{\heartsuit}, c_s^{\heartsuit})$，该状态描绘在图 2 中。

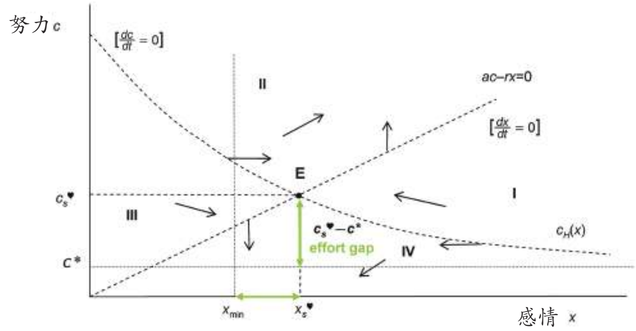


图 2　情感平衡态

$x_s^{\heartsuit}$ 位于 x_{min} 上方，因此这个平衡态是一个可接受的解。模型分析的一个关键发现是平衡的努力水平 $c_s^{\heartsuit}$ 位于 c^* 上方（参看附录 1），如图 2 所示。这是一个很重要的推论，也就是额外的努力 $c_s^{\heartsuit} - c^* > 0$ 对于维持长久的情感动力系统平衡是必要的。如果这对配偶认为这个努力差距 $c_s^{\heartsuit} - c^* > 0$ 并不需付出太大代价，那么这个平衡解就是可行的。当 $x_s^{\heartsuit} - x_{min} > 0$ 是可接受的并且努力差距 $c_s^{\heartsuit} - c^* > 0$ 是舒适的时候，一段关系就处于平衡状态。只要执行固定的努力计划 $c(t) = c_s^{\heartsuit}$，关系就会长期处于平衡状态。因为这一水平是问题（P）以 $E = (x_s^{\heartsuit}, c_s^{\heartsuit})$ 为初始值的唯一解，所以就达到了最大的幸福感。然而努力差距的存在或许会导致平衡解的不可行性。

不稳定性 一个基本的问题是扰动会随着时间消失还是放大。如果扰动会被放大，那么系统就是不稳定的。稳定性会促成一段长久稳固的关系，而不稳定性则是一个严重的缺陷。在不稳定的情况下，小的冲击，通常归咎于较低的努力水平，将会使得感情 - 努力状态远离平衡态。如果没有介入，扰动造成的最终的命运是情感关系的破裂。这将会在整体情感动力系统的分析中予以明确。方程 (1) 和 (2) 的平衡解被证明是不稳定（参见附录 1）。此外，平衡态的局部动力系统属于鞍点类型。这对于整体动力系统有很重要的影响。一个可行的但是不稳定的情感平衡状态在理论上是可持续的，这需要配偶留心由于低于平衡努力水平所造成的扰动，并及时的增加额外努力使得系统恢复到平衡状态。

情感动力学和分手机制

通常来说，情感关系的最初状态并不处于平衡点，这是由于最初对对方的感情一般的来说远远地高于平衡水平 $x_s^{\heartsuit}$。所以我们的讨论必须基于在初始感情 $x_0 \gg x_s$ 的情况下考察动力系统 (1) 和 (2)。我们需要考察相空间的整体构造来解释向平衡点的过渡动态。

整体情感动力系统 图片 3 是感情 - 努力的相空间的一幅定性的图片，使用标准技术获取（参看附录 1）。这个图

片对于所有满足假设 3 的效用和负效用函数都是近似有效的。这个动力系统的构造是一个非线性的鞍点。定向曲线代表的是最优路径（每条路径都使得满足一定初始和最终条件的总的净效用达到最大）。

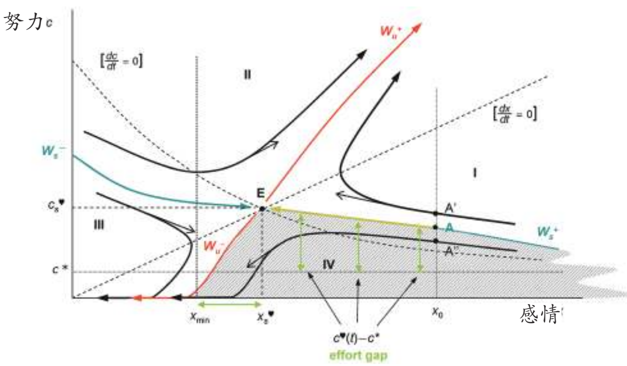


图 3　长久关系

稳定和不稳定流形，由所有接近和远离平衡点的点 (x, c) 组成，在图 3 中由曲线 W_s 和 W_u 表示。一旦轨迹到达 x- 轴，在这和之后的努力必须设定在 $c^{\heartsuit}(t)=0$（参见附录 1）。因此，到达线 $c=0$ 的一条轨迹按照

$$\frac{d}{dt}x(t)=-rx(t)$$

沿着 x- 轴衰退到 $x=0$。

长久关系的瞬时动力学 关键问题在于，给定一个初始的感觉 x_0，是否存在一个努力策略可以达到平衡态，如果存在，这个努力策略的特点是什么。稳定流形就是使得轨迹达到平衡点的唯一的曲线。任何其它的轨迹要么是不可接受的，要么对应于非长久的关系。事实上，位于 W_s 上方区域 I 和 II 中（见图 3）的轨迹是不可接受的，因为它们会导致越来越高的努力水平并最终导致难以忍受的负效用水平（参见附录 1）。另一方面，在有限时间内到达 x- 轴的轨迹会导致此后放弃努力并最终达到的感情水平只能是以结束关系收场。给定 x_0，存在一个合适的水平 $c_0^{\heartsuit}$，使得点 A $=(x_0, c_0^{\heartsuit})$ 位于 W_s^+ 上并演变到平衡点 E（见图片 3）。只要 $x_s^{\heartsuit}$ 大于 x_{min}，这个**目标轨迹** AE 就是成功维系永恒情感关系的（唯一）秘诀。因为目标轨迹 AE 嵌入在 W_s^+ 上，而整个处于 $c=c^*$ 之上，所以在通向平衡点的过程中必须付出一些额外的努力（大过 c^*）。因此，最优轨迹 AE 能够成功需要两个必要的条件，即这种感情盈余 $x_s^{\heartsuit}-x_{min}>0$ 必须是有益的并且努力差距 $c^{\heartsuit}(t)-c^*$ 在所有的 $t\geq 0$ 时刻都必须是可以忍受的。沿此目标路径存在的这一努力差距或许是配偶关系破裂的源头，因为在许多情况下它或许是难以忍受的。

幸福感降低 因为目标路径位于区域 I 中，在达到平衡值之前的路径上 $c^{\heartsuit}(t)$ 递增而 $x^{\heartsuit}(t)$ 递减。由假设 3 和链式法则可以推导出

$$\frac{d}{dt}(U(x)-D(c))=U'(x)\frac{dx}{dt}-D'(c)\frac{dc}{dt}<0。$$

这意味着幸福感沿着最优路径 AE 一直递减直到达到 E 点。这一模型的理论预测与第 2 节中陈述的断言 4 一致。

分手机制 如上解释，典型的动态出现在图片 3 中的阴影区域，因为 x_0 很大并且轨迹导致的努力水平的增加是貌似不合理的。沿着阴影区域的轨迹，努力会最终降低直到撞到 x- 轴并最终设定为 $c^{\heartsuit}(t)=0$。这使得情感关系的维持只能是短暂的。因为不稳定性——由努力的减少引起的——原本沿着 W_s^+ 一条轨迹的偏离会使得系统状态落在阴影区域，而那里的最优路径偏离了目标曲线。这一关键性的特征是情感不稳定的主要源头。

由于努力的疏忽，导致情感关系日渐恶化的一个可能的

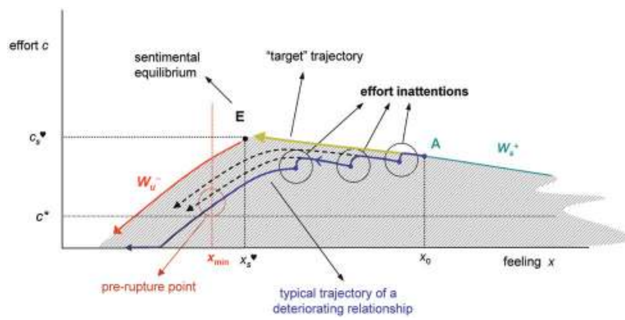


图 4　分手机制

机制在图片 4 中展示出来，并且可以进行如下地描述。假设关系起始于状态 A 点并且沿着位于稳定分支 W_s^+ 上的目标轨迹 AE 进行发展。如果在某点处发生了努力上的疏忽，也就是说如果努力水平降低了，状态就偏离了 W_s^+。如果努力不回到正确的水平并且系统遵循最优的动力系统方程 (1) 和 (2)，新的偏离状态就会作为一个新的初始条件，它导致的轨迹会逐渐偏离目标轨迹。这个新的轨迹可能会持续一段时间，直到新的努力疏忽出现，驱使状态处于更低努力水平的一个新位置，并沿着一个新的、更加远离 W_s^+ 的退化轨迹运行。经过一系列的努力的疏忽，不稳定性导致退化轨迹穿过阈值水平 x_{min}（参看图 4）。这是一个预破裂点，因为感情低于满意水平并且放弃努力只是时间问题。关系或许会持续一段时间但最终会突破承受的极限。在这一最后阶段，情感依恋逐渐消失，而这与专业文献中对离婚现象的描述基本相符。

如果系统沿着一条退化轨迹，通过提高努力水平仍然可以回到最优路径上。然而，做出反应及纠正偏差的耗时越长，状态就离目标轨迹越远，回到目标轨迹的困难就越大。如果

努力被长期忽视，这或许变得不可逆转。相当数量的被报道的不幸婚姻似乎符合这一判断。上述的恶化过程与第 2 节所述的断言 3 一致。

总 结

本文的数学理论给出了一个机制，用来解释在计划长久维持的情感关系中大量出现的恶化和破裂。两种力量一起导致了恶化过程的出现。首先，因为在成功路径上为了持久的关系必须付出额外的努力，假使这个差距是令人不舒服的，伴侣或许会放松并降低努力水平。接着，不稳定性出场，使得感情 - 努力状态偏离持久成功的动态。

一个更重要的发现是这一事实——伴侣建立关系并视其为明确目标与证据显示的他们的关系有可能破裂其实是相容的——这在模型中是很典型的动态过程。这消除了失败悖论，配偶分手的可能性是由于最佳选择下第二定律起作用的后果。

模型的分析或许可以给伴侣提供如何保持长期关系的建议。持久的关系只有在努力差距是可以接受的情况下才是可能的，最佳的努力策略是持续的关注使得其位于目标动态上，或许可以描述为在稳定分支附近的一条轨迹，并徘徊于平衡状态附近，时刻留心让努力保持在正确的水平上。这些类型的关系是常见的，虽然还有一些例外。这与模型中持久成功关系的例外是一致的。

两个明显的事实用来验证本文提出的理论：(i) 模型建立在被接纳的证据之上（亦即，第二定律和配偶打算长久相处的意向）以及 (ii) 模型的机制显示的与离婚及分手的实证事实相一致，这些事实也就是衰退关系中典型的逐渐恶化（第 2 部分中的断言 3）和婚后幸福感的降低（第 2 部分中的断言 4）。对于验证模型的进一步研究要——通过实验室试验或者实地调查——测试这个理论的两个主要发现，亦即努力差距的存在性和感情 - 努力动力系统的不稳定性。

对于配偶持久性的悲观结论应该也适用于并非理想的情景，只要第二定律的提法仍被认为是有效的。更一般的假设，如（弱）不同型配偶、外部冲击的出现或者次优的行为，或许可以作为加强不稳定性的因素来加以考虑。努力差距和揭示出来的不稳定性构成可能的情感失败的基本内在机制。

附录 I：本文理论分析的相关数学推导。详见 http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info:doi/10.1371/journal.pone.0009881.s001

作者简介：何塞 - 曼纽尔 • 雷（José-Manuel Rey），从马德里获得数学博士学位，在圣安德鲁斯大学和伦敦大学学院做过博士后研究，现为西班牙马德里孔普卢顿大学（Universidad Complutense in Madrid）教授。

译者简介：李玉田，吉林大学数学系毕业，香港城市大学博士，香港浸会大学助理教授。

微博上的数学漫游（连载五）

歌之忆 http://weibo.com/wildmath

二十世纪，人类从电气时代迈入了信息时代。伟大的时代需要伟大的数学。而这个伟大时代的主要缔造者，正是一群才华横溢的数学家——诺伯特·维纳、冯·诺依曼、阿兰·图灵、克劳德·香农，还有他们杰出的同行。这一群为第二次世界大战立下卓越功勋的先驱者，以移山填海之势，彻底打破了种种樊篱之见，冲开了旧有的学术屏障，勾画出前所未有的技术帝国的蓝图。在历史的定格中，他们不仅是超一流的数学家，更是从控制论到通信理论、从计算机科学到人工智能、从博弈论到行为科学等一大批新兴学科的开山祖师。正是他们以深邃的洞察力和强悍的冲击力，用无坚不摧的数学，奠定了信息化社会的技术基础，创造出了史诗般的丰功伟绩。

香农喜爱 Dixieland 爵士

每一个稀世天才，都有一颗孤独的灵魂。论才具，香农在麻省理工的硕士论文，把开关电路与逻辑联系在一起，被誉为二十世纪最重要的硕士论文；其攻读数学博士学位的论文，研究的却是遗传学。孤绝的才华，让年轻的香农常常陷入精神的迷乱，只好拿起他的单簧管，沉迷于迪克西兰（Dixieland）爵士乐中。

南方的迪克西兰热爵士，是自由的，是放浪的，也是焦虑的。它征服了芝加哥、征服了纽约，征服了美国，也征服了香农。博士毕业，24 岁的香农迈入普林斯顿高等研究院，追随大数学家外尔做博士后。信息论尚在孕育，置身在遥不可及的科学巨星爱因斯坦、哥德尔的身边，香农满是憧憬和焦灼。

论智力，尚在麻省理工读研究生时的香农已是才华毕现。他甚至还参加了飞行训练。但教练在一众学员中发现香农超凡绝伦，于是向校长建言：请中止香农危险的飞行训练，让科学天才去做更伟大的事业！康普顿校长断然拒绝：你不能

德国数学家外尔（1885-1955）

因为香农有才华就禁止他去冒险，这对学生的身心健康没有好处！

只有真正的教育家而非政客，才有资格去做大学校长。而康普顿不愧是麻省理工历史上最伟大的校长。正是他一手将麻省理工学院从普通的技工学校变成执世界工程之牛耳的超一流大学。鼓励学生大胆冒险，容忍学生的恶作剧，则几乎是历任麻省理工校长的美德。剥夺冒险精神，就是摧残创造力。

循规蹈矩只能造就平庸，唯人性的张扬与精神的自由才能成就奔放无羁的创造力。麻省理工学院的学生们曾经伪造证件，从西海岸竞争对手加州理工学院偷运出对手的镇院之宝加农炮；他们趁黑夜把消防车弄上学校主楼楼顶。精确的计算与大胆的创意，才有了如此之多符合名校气质的高级恶作剧。

在麻省理工这样杰出的大学，满是醉心于探索的天才，而非唯唯诺诺的书呆子。闪耀着天马行空创造力的家伙，往往在孩提时代就已经开始了无拘无束的探索。香农十岁时，就把发报机搭到自家牛栏的铁丝网上，和邻居小朋友玩起了莫尔斯电码。一个伟大的理论，萌芽于一个孩子为之着迷的游戏。

从十岁就开始玩通信的香农，长大后在密西根大学继续攻读电子工程与数学，在麻省理工念研究生时期展现出惊人的创造力，经过贝尔实验室的一段熏陶，终于在普林斯顿高等研究院正式开始了酝酿信息论的深刻思考。他身后不仅是伟大的物理学家爱因斯坦和数学家外尔，还有一代雄才冯•诺依曼。

香农前妻诺玛

香农

香农以一个工程师的敏锐，抛弃了对信息内容的追究，单纯从传输的视角定义了信息的测量方式。在香农眼里，“我爱你爱你爱爱爱你你爱你”，和“我恨你恨你恨恨恨你你恨你”，这两串符号的信息量完全一样。“爱”换成了“恨”，符号变了，概率结构却依旧，信息量只与符号出现的概率有关。

信息量不关乎“爱”与“恨”，但攀登学术最高峰，却多有爱恨交加的精神史。在普林斯顿，约翰•纳什如此，香农也几近崩溃。在最残酷的奋斗中，他的新婚妻子诺玛离他而去。香农终生几乎未在公开记录中谈及前妻。事实上，两人曾经在分手20多年后一度相逢，香农问：你当初为何离我而去？！

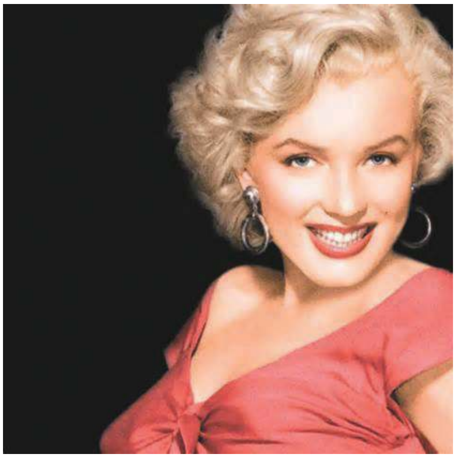
爱情故事的开局大多雷同。麻省理工的香农与拉德克里夫女子学院的诺玛，舞会相识后旋即闪婚。香农踏入普林斯顿高等研究院之后，沉醉于自身的研究，几乎不与外界沟通。诺玛反复告诫香农——你这样太病态了！你得去看心理医生。但香农却不以为然。于是在一顿大吵之后，诺玛毅然离家出走。

19岁嫁给香农的诺玛转眼离开了香农，从东海岸跑到好莱坞当上了电影编剧，又嫁给了同行巴尔兹曼。夫妻两人果然志同道合，一同加入了共产党。于是他们成了冷战时期标准的赤色分子。一天，正在家门口的诺玛被一位大美女叫住：我也叫诺玛（Norma）。警察正在山下记录你家访客的车牌。你们快逃！

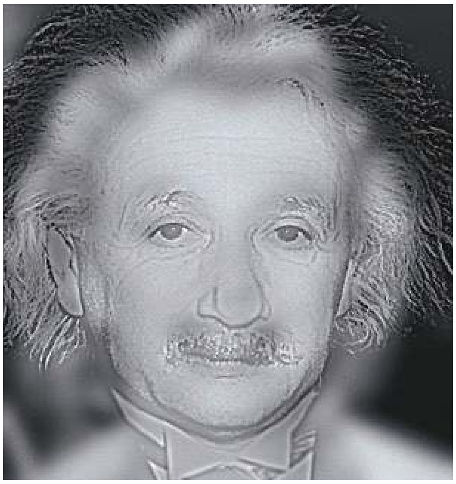
从此，诺玛跟随巴尔兹曼，开始了几十年颠沛流离的欧洲流亡生涯。不过在诺玛看来，与其在普林斯顿给香农尊贵的同事爱因斯坦沏茶，远不及在欧洲同毕加索聊天那般惬意。当然，诺玛更是无法忘记最初提醒她逃走的另一个诺玛——她就是诺玛•珍妮•摩登森（Norma J. Mortenson），日后红极一时的性感明星玛丽莲•梦露。

童年时代遭受过极度屈辱的玛丽莲•梦露，有着一颗水晶般的心，为人厚道善良。尚未成名的她出手搭救香农的前妻诺玛，并不令人意外。性感女郎让无数男子遐想连篇，这当中物理学家表现尤其不够镇定。霍金曾说：假如能坐上时间机器，他一定要去拜会两个人，一位是伽利略，另一位则是梦露。

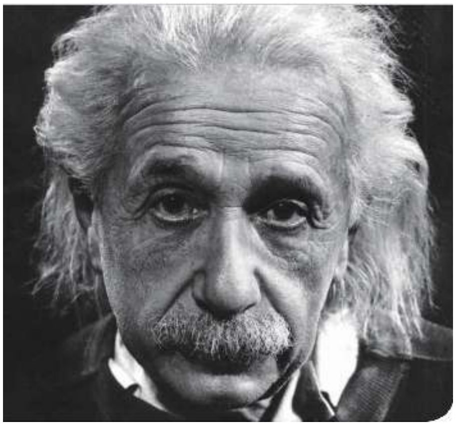
霍金何止想寻访梦露，他在卧室挂了梦露的玉照，外加一幅他与梦露的合成照片。第一流的物理学之所以美，因为第一流的物理学家青春永驻。据说爱因斯坦对梦露也是情有独钟，这启发了麻省理工的奥德•欧丽瓦（Aude Oliva）。她在《新科学家》发表了一幅窥探心灵的画：远观是迷人的梦露，近看是深邃的爱因斯坦。



梦露



梦露与爱因斯坦混合像



爱因斯坦

霍金

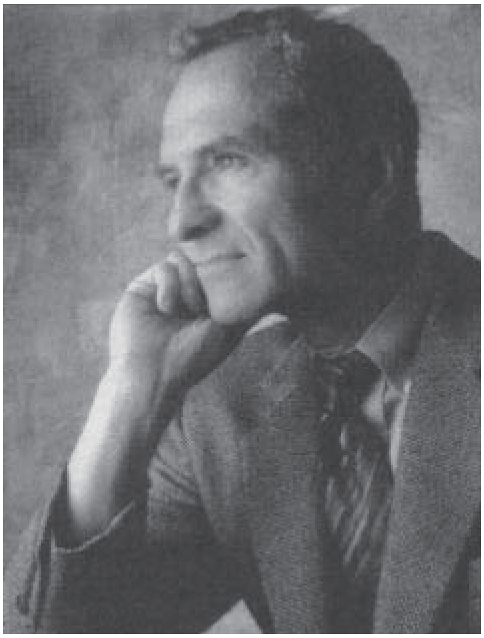
梦露与爱因斯坦多少也算门当户对——2012 年的“福布斯已故名人收入排行版”上，两位名流并列第七名，他们照片的版税年收入都达到 1000 万美元。当年被梦露搭救的诺玛原本是犹太富家千金，而香农只是一介书生。两人重逢后香农告白：我过得很好，妻子贤惠、孩子出色，我已经有了 23 辆汽车！

香农告诉诺玛，现在的他生活美满富足，与 22 年前不可同日而语。当年，他那犹太富豪丈母娘根本不信一个数学博士能挣什么钱，越俎代庖将小两口的寓所装修成一副现代气派，令香农耿耿于怀。时过境迁，功成名就的数学博士 1956 年到麻省理工做教授时，年薪是惊人的 1 万 7 千美元，远远高于同行。

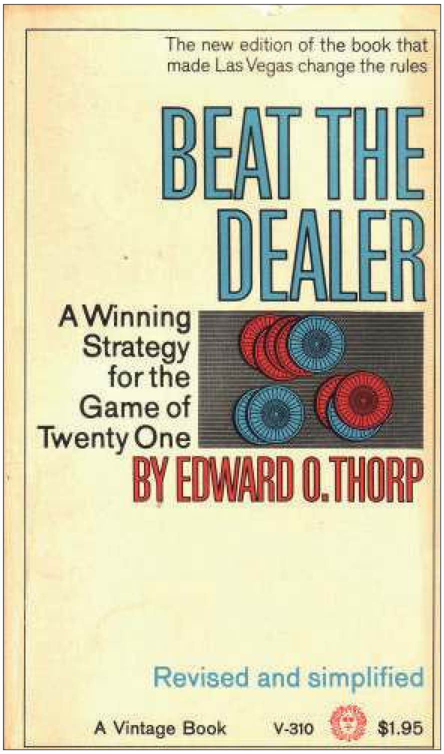
其实香农并不是个痴迷金钱的数学家。有人曾借用香农在麻省理工的办公室，吃惊地发现这里竟然还有过期未兑现的大额支票！香农的第二任妻子贝蒂，发现他居然把生活积蓄全都存在了支票账户，分毫利息皆无。香农，似乎仅仅是个创立了信息论的数学家，一个专注于学问而对金钱毫无感觉的学者。

但有人却天生痴迷金钱。有位出生于大萧条时期的数学博士，从一门心思抠钱发展到钻研赌博术，靠精深的概率计算，琢磨出玩 21 点的取胜之道。他将心得写成论文，祈求当时麻省理工唯一的数学院士香农来推荐发表。香农问了一堆刁钻的问题，得到满意的回答，于是断定，此人的确干出了名堂。

向来心事缜密的香农，断定此人还有猛料。毕竟 21



美国数学家索普（1932- ）



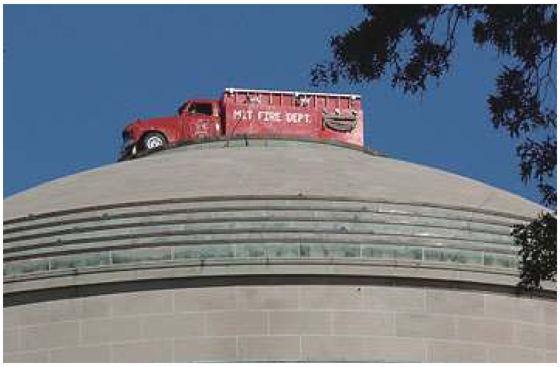
索普的名作《打败庄家》

21 点

点只是赌场里的陪衬。一番旁敲侧击之后，来人招认，他还想进一步钻研轮盘赌。如此正中香农下怀。于是两人合谋发明一个设备，带到赌场大干一场。现代金融史上杀伤力无穷的“量化投资”，由此掀开帷幕。这位访客正是先驱者爱德华•索普（Edward O. Thorp）。

在洛杉矶加州大学取得数学博士学位的索普，在信息论创始人香农的鼎力相助下，发表了《财富密码：21 点取胜之道》；他弄到 1 万美元杀入拉斯维加斯赌场，短短数日便翻了一番，由此一举成名。1962 年出版《战胜庄家》，更是风靡世界。由赌博催生出的概率论，成为赌场高手所向无敌的利器。

中国历来不乏清高的读书人，却鲜有玩得活色生香的学者。1956 年香农刚刚 40 周岁，8 年前他创造的信息论已让世界为之瞠目。某天晚上香农兴起，拉上同事凯利（John L. Kelly）脱下鞋去爬麻省理工滑溜溜的大礼堂顶。警察过来了，大



麻省理工学院：左图：楼顶上的消防车；右图：香农和凯利攀爬的 Kresge 礼堂顶

美国科学家凯利（1923 - 1965）

名鼎鼎的罗伯特 • 法诺（Robert Fano）教授赶紧跑来道歉：这两人可是贝尔实验室顶尖的科学家。

50 年代的学术界已有人把香农同爱因斯坦相提并论。但香农的同事们却说这不公平：就人类的日常生活而言，香农远比爱因斯坦重要！他们更愿意去这么比——在贝尔实验室，智力最接近于香农的人是凯利——这是一位二战退役的飞行员，一位大大咧咧、雪茄不离口的物理学博士，桥牌的绝顶高手。

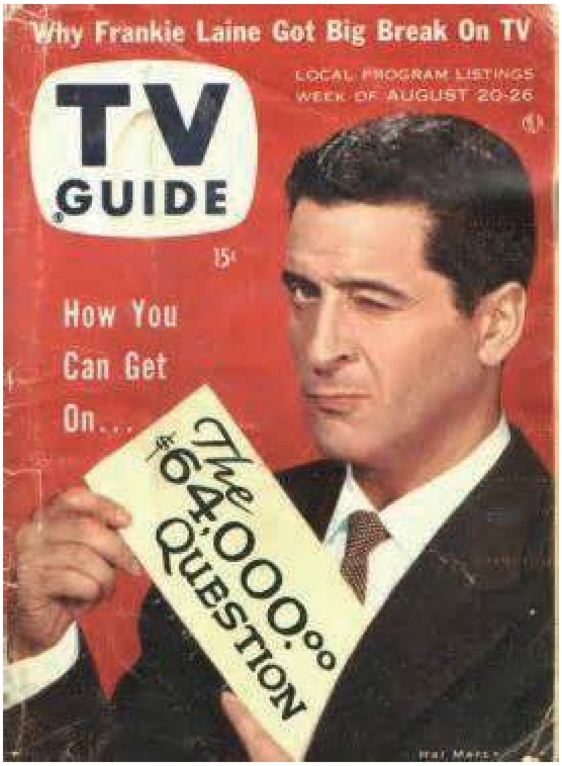
训练有素的数学家往往是好厨师，他们善于把身边的食材烹饪成一桌美味佳肴；而训练有素的物理学家却更像是好猎手，他们善于抓住稍纵即逝的机会去捕获一切“学术猎物”。当卓越的猎手遇上卓越的厨师，学术便被注入了鲜活的灵魂，凡夫俗子眼中杂乱无章的世界，才层层显现出它真正的意蕴。

香农声名鹊起时，无数东施效颦的论文都冠以“信息论”的时髦标签，但物理学博士凯利没有去凑这份热闹，贪玩的他琢磨起一档“64000 美元问答”的电视闯关游戏：闯过一关奖金增加一倍，闯关失败则分文不取。凯利独出心裁地观赏了这个节目。令人吃惊的是：他从中发现了投资策略与信息论之间的绝妙联系。

50 年代风靡美国的“64000 美元问答”电视娱乐节目，火爆堪比“中国好声音”。银屏上的选手们在竞猜答案，银屏下的观众却在玩更大的竞猜游戏——他们踊跃下注去赌谁将胜出。不过，东海岸的节目要迟三个小时才在西海岸播出。于是，西海岸的投机赌客会通过电话打听好谁是赢家，再去下注。

身在东海岸贝尔实验室的凯利，用与众不同的眼光去分析这种电视竞猜游戏。西海岸的观众，事先能搞到谁是赢家的谜底，这样下注等于是利用内幕消息炒股。但炒股时若你真有了内线信息，料你不敢投下全部身家，以免倾家荡产；可你若过度小心却也会错失良机。左右为难之境，究竟该如何下注？

现实中的人经常处在不确定的境地。抛一次



64000 美元电视竞答

香农和妻子贝蒂

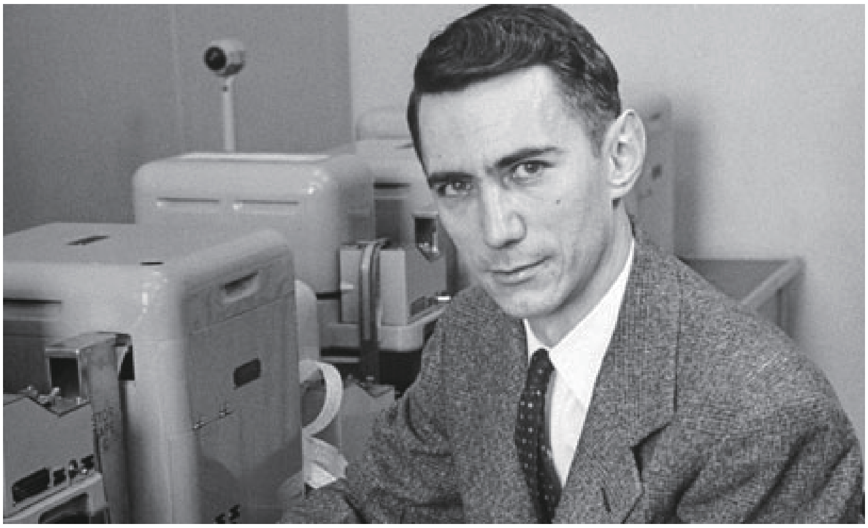
硬币，你无法预测究竟是正面朝上还是背面朝上——数学不回答这个问题，数学只宣称：若重复这样抛一万次硬币，大约有五千次左右是正面朝上。而究竟是哪些轮次正面朝上，它依旧无法告诉你。但这样的数学真理已足以让智者找到战胜不确定性的法宝。

一个电视娱乐节目，让贝尔实验室的高手凯利把投资策略与通信技术这风马牛不相及的两个领域扯到了一起：最高的资本增长指数率等价于通信中的传输速率。换种说法：资本的增长指数率与内幕消息的不确定性之和为常数。物理学博士凯利，在投资理论与信息论之间，架设了一座令人眩目的金桥。

不确定性并非万劫不复的深渊，相反，它是智者通往胜境的桥梁。噪声给通信带来了不确定性，但香农战胜了它。香农能以几乎无错的方式去传输消息；不可靠的内幕消息给投资带来了不确定性，但凯利战胜了它。凯利只需恰当地调整资金投入比例。极致的凯利与极致的香农，相会于无限风光之巅。

贪婪的人，总是妄图打破机会的天平;谨慎的人，常常不敢去接受机会的馈赠;只有真正的智者，才懂得平衡地享受机会。凯利的美妙论文，象香农一样发表在贝尔实验室的期刊上。其通篇只引用了两篇不朽之作，一篇是香农的《通信的数学理论》，另一篇则是冯•诺伊曼的《博弈论和经济行为》。

香农在普林斯顿时就曾纠结于如何命名他研究的“不确定性”。冯•诺伊曼建议——你就命名它为“熵”吧！本来物理学上就有“熵”，但没几个人能整明白。



香农

匈牙利数学家瑞尼（1921- 1970）

将来若有人找你论战，你就不战而胜了！凯利当然明白什么是熵：不确定性每降低一个比特，就等价于一万个基点的资产收益——财富倍增。

那个曾经忘记去兑现大额支票的香农、那个居然把积蓄全部存进支票账户的香农，最终应验了其导师的断言——此人做任何事都将是天才。50 年代后期，这位数学家开始投资股市，到 1986 年他的基金已位于巴伦（Barron）共同基金业绩排行榜首。长达三十多年，基金年收益率高于 28%，业绩超过了巴菲特。

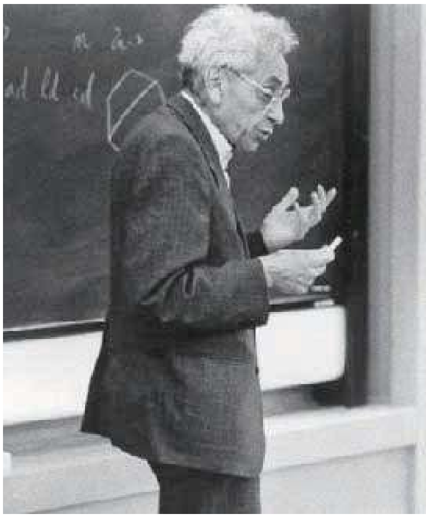
从赌场到股市，无数的传奇来自对“机会”的数学研究。去赌场狂碰运气的卡尔达诺输得一塌糊涂；而从伽利略计算骰子的组合开始，费马、帕斯卡、惠更斯、伯努利、棣莫弗及至香农，纷纷加入了对机会的讨论。数学家们在机会面前不仅仅展示了冷竣的理性，也展示了狂野的血性。

二战期间，匈牙利的反犹恶浪剥夺了青年瑞尼 (Alfréd Rényi) 上大学的机会，但他在数学竞赛中脱颖而出，侥幸上了大学，不料刚毕业就被关进集中营。瑞尼勇敢地逃了出来。半年之后，他父母也被抓进集中营。于是，瑞尼伪装成纳粹分子，拿着假证件闯进集中营，假借“逮捕”之名，救出了自己的父母。

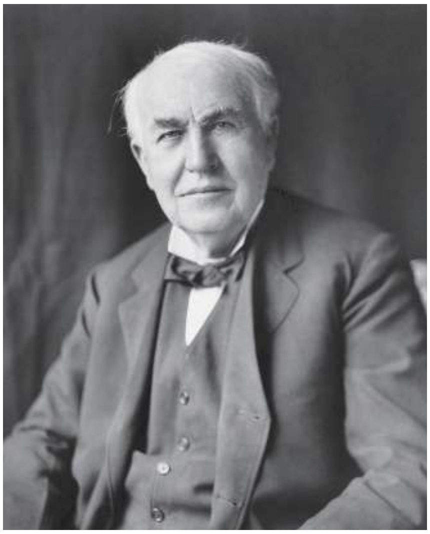
从集中营逃出生天的瑞尼，后来投身到泛函分析大师黎兹（F. Riesz）门下。他的同门师兄拉多（T. Radó）参加过第一次世界大战，在西伯利亚战俘营开始学数学，然后千里逃亡回来。英勇无畏的瑞尼，后来成为匈牙利科学院数学研究所所长，在概率论、信息论和哥德巴赫猜想上作出了出色成就。

研究“不确定性”的高手，也会遭遇不确定性的尴尬。瑞尼有句名言——“数学家就是把咖啡变成定理的机器”，如今却被张冠李戴成了爱多士的名言。不过，瑞尼还有一句值得记取的名言：“不开心的时候，只要去做做数学，立马就会开心起来；开心的时候，依然去做数学，就会一直开心下去”。

瑞尼可以一直开心地去研究数学，但香农却不能。香农 21 岁以一篇硕士论文打开数字逻辑之门、开创了现代集成电路工业；24 岁以一篇博士论文揭示了遗传之规律；32 岁更以一篇旷世之作开辟了现代通信技术产业；40 岁后转战股市取得惊人业绩。但是，改变了世界的天才，却无法摆脱命中注定的孤独。



匈牙利数学家爱多士（1913-1996）



美国发明家爱迪生（1847 -1931）

40 岁后的香农，纯然是独孤求败的境界。向他讨教的人，发现每个问题都被他搞透了。沉思者无心出门，偶尔一次去英国开会，到会场时未被认出。窃窃私语中，大家知道香农来了，顿时簇拥围了上来——有人笑称：这般轰动，无异于开物理学大会，发现牛顿大驾光临！而那一刻的香农，多么孤独！

19 世纪后期，香农的先辈远亲爱迪生，用电灯照亮了整个世界；之后，爱迪生和特斯拉、威斯汀豪斯竞相角逐电气化，成就了光电帝国。但美利坚的富强，绝不仅仅是因为两次世界大战占了便宜。二十世纪强、弱电行业开始分离，美国本土培养的卓越人才撼动了世界。而香农，正是划时代的标志。

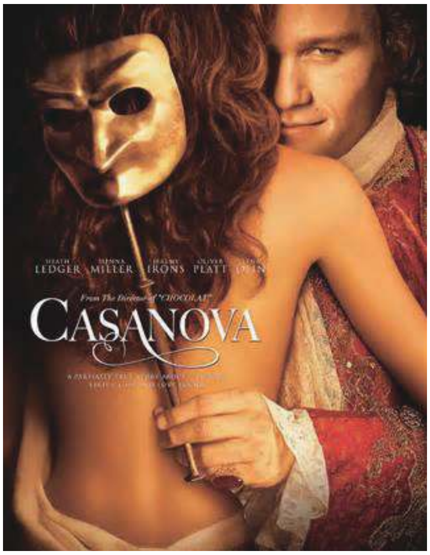
匈牙利数学家福雷叶斯•黎兹（1880-1956）

香农、图灵和冯•诺依曼，共同开启了信息时代的远航。香农定义了测量信息的基本单位：比特（bit）。你抛出一枚硬币，正面朝上还是背面朝上，仅仅不过 1 个比特；某个嫌犯究竟犯了罪还是无罪，最多也不过 1 个比特。信息量不解读内容、不解读人心，但它却承载了我们心灵最深处的柔情与悲愤。

美国数学家杜布（1910-2004）

1956 年，香农回到 MIT。两年前，42 岁的图灵自杀身亡；而一年之后，54 岁的冯•诺依曼因癌症病逝。二战前在普林斯顿，冯•诺依曼帮助香农确立了信息熵的概念。在二战绝密的 X 计划期间，香农与图灵探讨过他研究的信息单位——比特。到 1957 年，开创信息时代的三大巨头，只剩下时年 41 岁的香农。

天才离不开足以比肩的智慧去磨砺，否则锋芒不再。40 多岁的香农，几近退休之态，研究生上门找他讨教问题，见他总是沉迷于单簧管，他已不再乐意教学。瑞尼的导师黎兹更是夸张。这位牛人上课，要带上一位副教授和一位助教。副教授去念他的讲义，助教上黑板去板书，而黎兹自己却闭目养神。

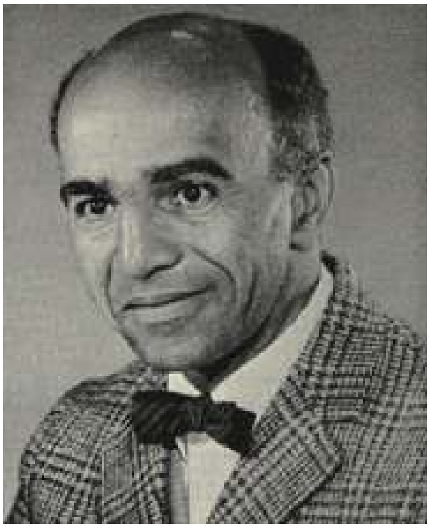


意大利冒险家卡萨诺瓦（1725-1798）

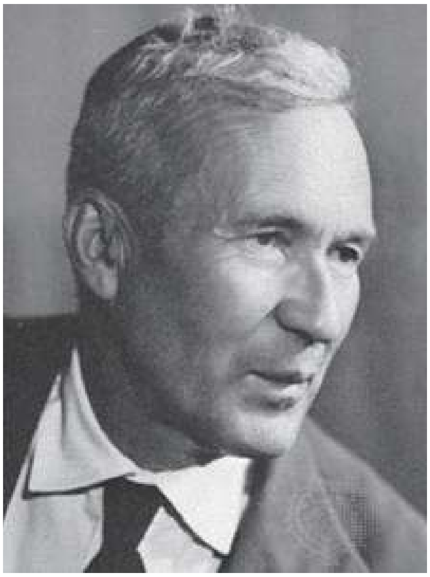
对随机现象极富洞察力的思考，成就了香农的理论。但他的论文发表后，概率论名家杜布（Joseph L. Doob）却公开抨击香农的数学靠不住。香农的思想脱胎于概率论的大数定律，而哈佛出身的杜布也是玩大数定律起家的。香农毫不理睬概率权威的恶评。半个世纪后有了盖棺定论：全部23个定理，完全正确。

从掷骰子到打扑克，人类对随机性的许多认识往往来自赌博。若没有赌博，很难想象概率论和信息学是何种形式。历史上赫赫有名的花花公子、冒险家卡萨诺瓦，赌博成瘾，输钱时必定加倍下注——只要输赢机会均等，大多会把赌本加倍赢回来。如此高风险的玩法，后来就促成了数学上的“鞅”论。

杜布抨击了香农理论，但他自己的弟子、统计学家布莱克维尔（David H. Blackwell）却迷上了信息论，甚至培养出通信理论的杰出人才。据说杜布后来多次声称，他一生犯了两大错误：第一是低估了香农；第二是定义错了“上鞅”——赌徒离开赌场时输了钱，叫“下鞅”比叫“上鞅”更为自然。



美国统计学家布莱克维尔（1919-2010）

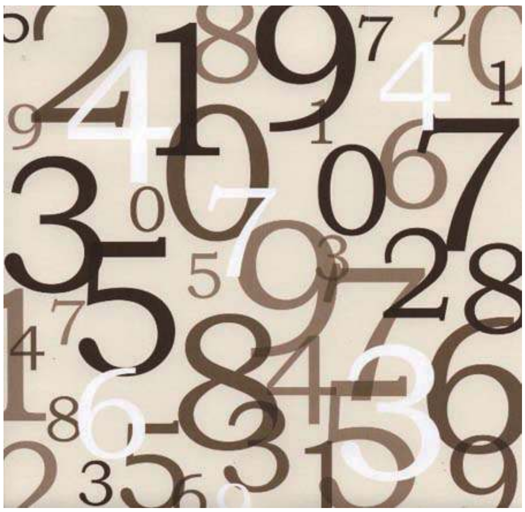


前苏联数学家柯尔莫哥洛夫（1903- 1987）

与杜布相比，50年代初的苏联甚至将香农理论称作蒙昧主义的糟粕。此时，一代数学大师柯尔莫哥洛夫强势出手，不仅支持了香农的理论，甚至揭示了更为深刻的本质——这位对概率论进行了公理化的巨匠，提出了极为不凡的见解：表面上看概率论是信息论的基础，实际上信息论比概率论更为基本！

数学并非你想象的那么战无不胜

曹广福



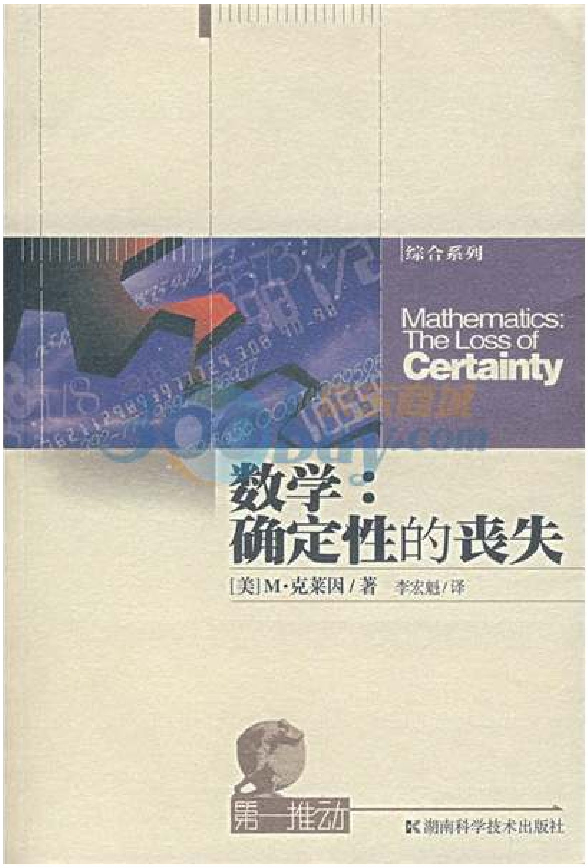
人们习惯于“王婆卖瓜”，数学工作者自然也难免俗，倘若其他专业的人士夸奖自己从事的领域，无疑更是一件无尚荣光的事。任何一门学科都有值得欣赏的地方，很难想象，一门科学如果不能在某个方面改变世界或者推动人类文明的发展，它还能称之为科学。数学之所以得到很多人的推崇，一个很重要的原因是它与自然科学、社会科学的几乎每一个领域都休戚相关。毋庸置疑，没有数学就不可能有今天的人类文明。不过凡事不可以绝对化，任何人的话都不是绝对的真理，伟人的话也是如此。

人们往往强调了数学在科学研究中的作用，却忽略了数学软弱无力甚至逻辑错乱的一面。事实上，虽然数学的发展日新月异，学科分支五花八门，但真正可以解决问题的数学为数不多。很多人高谈阔论数学之有用，其实他们仍然沉浸在数百年前的数学中，对近现代数学所知有限。法国布尔巴基学派的领袖人物丢东尼曾经大肆抨击近代数学研究的某些倾向：“许多数学家在数学王国的一角占据了一席之地，并且不愿意离开。他们不仅差不多完全忽略了与他们的专业领域无关的东西，而且不能理解他们的同事在远离他们的另一个角落使用的语言和术语。即使是受过最广博的训练的人在浩瀚的数学王国的某些领域中也感到迷茫，像庞加莱和希尔伯特这样的人，几乎在每个领域都留下他们天才的印迹，甚至在最伟大的成功者中也是少而又少的极其伟大的例外。”

“数学是什么”是一个纯粹的哲学问题，对他的解答众说纷纭，迄今并没有什么标准答案，也许永远也不可能有标准答案，我倒是倾向于柯朗的一段名言：“‘数学是什么？’这个问题，不能通过哲学概括、语意学定义或者新闻工作者所特有的迂回说法来做出令人满意的回答。”

有一种有趣的现象，尽力鼓吹数学万能的人往往是那些对数学了解不多的人，或许真的像人们说的那样：“学得越多，越知道自己无知。”我知道的数学不算多，也知道数学对很多学科有用，不过我最关心的是它到底有什么用？它在自然科学研究中到底充当了什么角色？人们对数学的认识伴随着数学的发展在不断变化。曾几何时，人们判断数学好坏的重要标准是：“它有什么用？”即使是一个响当当的大定理，我们也需要问一问这个问题。可按照这个标准，很多数学大有被扫进垃圾堆之虞。更重要的是，今天看起来没用的数学也许明天就能挽救世界。在数论、代数几何应用于密码领域之前，谁能知道它对现实社会可以产生什么样的影响？于是人们开始改变好数学与坏数学的评价标准：“有用的或者能给人带来美感的数学就是好的数学。”标准的改变既挽救了一些学科，也让很多人保住了“饭碗”。不过“数学能干什么”依然是大家最关心的问题。即使你从事的数学研究与自然科学没有直接的联系，至少你要找出它与数学其它分支的内在关系，否则难免授人以“闭门造车”之口实。

要列举数学的丰功伟绩实在太容易了，任何人都可以如数家珍般列出一长串，远的不说，仅最近几届诺贝尔经济学奖的工作就足以说明数学之伟大，更不用说物理学、化学等理学分支与数学的深刻关系了。但我们在为数学高唱赞歌之时更应该清楚另一个问题：“数学，它不仅是很不完善的，有些还是相互矛盾的。”数学也不像人们想象的那样无所不能，我相信差不多在每个领域都可以找



出数学无能为力的问题来。有人可能会质疑这一观点:“现在解决不了不等于将来解决不了，也许数学还没发展到那个程度。”不排除这种可能，然而我们不要忽视一个至关重要的问题，自然界有很多东西本身是不确定的，而我们的数学大多是研究确定性问题的。举个例子，很多研究中都需要解决一个重要问题：“风险”，如何度量风险？这是个大难题，没有任何有效的数学工具可以解决这个问题。

克莱因写过一本书《数学：确定性的丧失》，这本书颠覆了数学乃绝对真理、是物质世界不可动摇的知识体系的神话，值得一读。从这本书中我们可以看到，数学的发展常常是不符合逻辑的，然而正是这种不合逻辑才使得数学得以顺利发展。以微积分为例，如果牛顿研究天体物理时早就知道连续函数未必可导，微积分也许就不会产生。读完这本书可以让我们认识到：“数学的确是有用的，而且它在自然科学研究中的有效性会随着研究不断扩大，但数学也只是相对真理，甚至它的基础还是不确定的。”

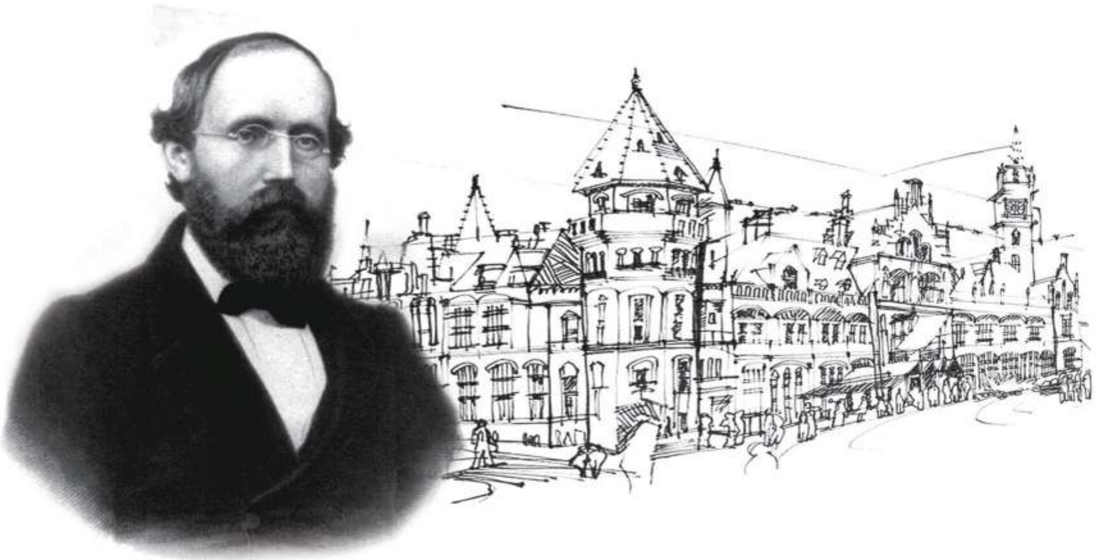
数学在社会科学、自然科学中的作用已被无数的事实证明，但我们同时应该了解，数学是很不完美的，它的发展过程并不像教科书中描述的那样逻辑井然，它也不可能像人们期望的那样无所不能。毕竟世界未必真的是上帝按数学方法创造出来的。数学的尴尬之处在于：1、它自身并不完善，有些甚至相互矛盾；2、它永远是物理世界的近似而非精确描述；3、很多物理世界中的现象很难甚至不能用数学来描述。

要了解数学的本质不是一件简单的事，至少目前很难有一个恰如其分的答案，如果我们无条件地把数学知识奉为解决一切问题的法宝，甘当数学知识的奴隶，恰恰可能囿于传统的数学思维，最终失去了原本拥有的创造力。

如此说来，数学不重要了？我们不需要学好数学了？非也，数学当然重要，因为它是解决问题的钥匙，根本问题在于：“什么叫学好数学？”如果我们拘泥于已有的数学知识，也许永远一事无成。从这个意义上说，所谓学好数学不在于你掌握了多少数学知识，而在于你是否掌握了数学的思维方法。事实上，很多问题的解决虽然运用了数学，但绝不是简单的数学知识的应用，而是在已有数学知识基础之上的再创造。

人有优缺点，知识也有正反两方面，只有以批判的眼光对其有了全面的认识才能真正理解并很好地运用它，否则很容易成为知识的奴隶。

作者简介：曹广福，吉林大学数学博士，广州大学数学院教授、院长，2003 年获首届全国高校教学名师。



卢昌海《黎曼猜想漫谈》书评

扶 磊

勿容置疑，黎曼猜想是当今数学中最重要的未解决的问题。不同于数论中其它猜想（例如费马大定理、哥德巴赫猜想等），将黎曼猜想和黎曼的思想叙述清楚至少需要用到比较高等的复变函数论的知识，因此已有的关于黎曼猜想的科普读物，要么太肤浅而不能将问题说清楚，要么用了太多的艰深的数学工具而晦涩难懂。但卢昌海的《黎曼猜想漫谈》以轻松愉快的语言清晰地介绍了黎曼猜想、黎曼在数论方面的工作和后人在黎曼的影响下所做的后继工作。一个学过复变函数论的大学生应该能看懂这本书，从而能够欣赏黎曼工作的深度和创造性。

黎曼发表的文章不多，但几乎每篇文章都是划时代的甚至是超越时代的绝世佳作，这些文章开辟的新的数学领域包括：复变函数论、黎曼几何、代数几何、解析数论等等。黎曼在 1859 年发表了他唯一的一篇数论方面的论文《小于给定值的素数的个数》，用复变函数 $\zeta(s)$ 研究了素数的分布。黎曼 ζ- 函数是用无穷级数

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

定义的，这个函数欧拉也研究过，并注意到

$$\zeta(s)=\prod_{p}\frac{1}{1-\frac{1}{p^s}}, \qquad (2)$$

这里乘积是对所有素数 p 取的。欧拉只考虑当 s 是实数时 $\zeta(s)$ 的性质，而黎曼是历史上第一个将 $\zeta(s)$ 当作复变函数，并将素数的分布与 $\zeta(s)$ 的零点分布联系起来。上面 $\zeta(s)$ 的表达式的无穷级数和无穷乘积只是当复数 s 的实部 $\mathrm{Re}(s)>1$ 时才收敛。黎曼证明了由这个无穷级数 (1) 和无穷乘积 (2) 定义的函数可以解析延拓得到定义在整个复平面上的亚纯函数。这个亚纯函数就是黎曼 ζ- 函数。黎曼证明了 $\zeta(s)$ 在 $s=1$ 处有个单极点，在复平面其它地方解析，并满足函数方程

$$\zeta(s)=2\Gamma(1-s)(2\pi)^{s-1}\sin\left(\frac{\pi s}{2}\right)\zeta(1-s).$$

当 s 是负的偶数时，我们有 $\sin(\pi s/2)=0$，从函数方程可以推出 $\zeta(s)=0$，这些零点称为 $\zeta(s)$ 的平凡零点。$\zeta(s)$ 的其它零点称为非平凡零点。黎曼考虑辅助函数

$$\xi(s)=\Gamma\left(\frac{s}{2}+1\right)(s-1)\pi^{-\frac{s}{2}}\zeta(s).$$

函数 $\xi(s)$ 在复平面上处处解析，即是个整函数，满足函数方程

$$\xi(s)=\xi(1-s),$$

最重要的：$\xi(s)$ 的零点正好是 $\zeta(s)$ 的非平凡零点。黎曼指出

$$\xi(s)=\xi(0)\prod_{\rho}\left(1-\frac{s}{\rho}\right), \tag{3}$$

这里乘积是对 $\xi(s)$ 的所有零点 ρ 取的。如果 $\xi(s)$ 是个多项式，这个乘积公式是明显的，但事实上 $\xi(s)$ 是个有无穷多个零点的整函数。黎曼给出了这个公式成立的一些理由，严格的证明是 1893 年阿达玛（Jacques Hadamard）给出的。比较两个乘积公式 (2) 和 (3) 并巧妙运用傅里叶分析中的反演公式，黎曼将素数 p 的分布和 $\zeta(s)$ 非平凡零点 ρ 的分布联系起来。更确切地，对任何 $x>1$，黎曼得到等式

$$\begin{aligned}J(x)=\mathrm{Li}(x)-&\sum_{\mathrm{Im}(\rho)>0}(\mathrm{Li}(x^{\rho})+\mathrm{Li}(x^{1-\rho}))\\&-\log 2+\int_x^{\infty}\frac{dt}{t(t^2-1)\log t},\end{aligned} \tag{4}$$

这里 $\mathrm{Li}(x)=\int_0^x\frac{dt}{\log t}$，函数 $J(x)$ 在 $x=0$ 处取值 0，每过一个素数 p 处跳跃 1，每过一个素数平方 p^2 处跳跃 1/2，每过一个素数立方 p^3 处跳跃 1/3，……。记不超过 x 的素数的个数为 $\pi(x)$，则有

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

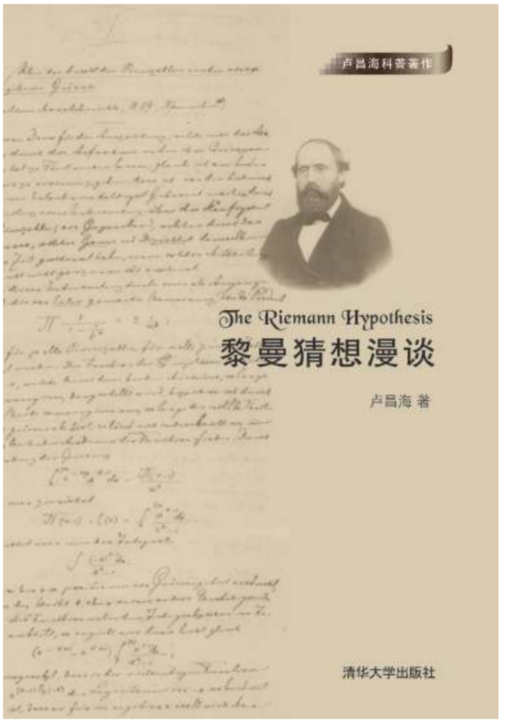
可以证明

$$\pi(x)=J(x)-\frac{1}{2}J(x^{\frac{1}{2}})-\frac{1}{3}J(x^{\frac{1}{3}})+\cdots+\frac{\mu(n)}{n}J(x^{\frac{1}{n}})+\cdots, \tag{5}$$

这里的 $\mu(n)$ 是 Möbius 函数

$$\mu(n)=\begin{cases}0 & \text{如果 } n \text{ 被素数平方整除，}\\ 1 & \text{如果 } n \text{ 是偶数个不同素数乘积，}\\ -1 & \text{如果 } n \text{ 是奇数个不同素数乘积}.\end{cases}$$

这些公式是黎曼文章的主要结果，严格证明是 1895 年冯•曼戈尔特（Hans von Mangoldt）给出。从上面的公式 (4) 和 (5)，我们可以用 $\zeta(s)$ 的非平凡零点 ρ 去表达 $\pi(x)$:



$$\begin{aligned}&\pi(x)-\sum_{n=1}^{N}\frac{\mu(n)}{n}\mathrm{Li}(x^{\frac{1}{n}})\\&=-\sum_{n=1}^{N}\sum_{\rho}\frac{\mu(n)}{n}\mathrm{Li}(x^{\frac{\rho}{n}})+\text{（更小的项）}.\end{aligned} \tag{6}$$

但如果对 $\zeta(s)$ 的非平凡零点的分布没有足够多的信息，还是不能证明

$$\pi(x)\sim\sum_{n=1}^{N}\frac{\mu(n)}{n}\mathrm{Li}(x^{\frac{1}{n}}).$$

特别地，上面的公式 (6) 还不足以说明素数定理成立：

$$\pi(x)\sim\mathrm{Li}(x).$$

1896 年阿达玛和查尔斯•瓦利•普桑（Charles de La Vallée Poussin）得到了证明素数定理所需要的 $\zeta(s)$ 的非平凡零点分布的信息，他们证明了 $\zeta(s)$ 在 $\mathrm{Re}(s)=1$ 上没有任何零点，并在此基础上结合冯•曼戈尔特的公式证明了素数定理。黎曼的著名猜想说 $\zeta(s)$ 的所有非平凡零点都落在复平面的直线 $\mathrm{Re}(s)=1/2$ 上。如果这个猜想成立，则有

$$\pi(x)-\mathrm{Li}(x)=O(x^{\frac{1}{2}}\log x).$$

在黎曼之后的很长一段时间里，人们怀疑黎曼猜想

只是黎曼的一个凭空猜测，没有任何计算或理论方面的依据。但在1932年，卡尔•路德维希•西格尔（Carl Ludwig Siegel）仔细研究了黎曼未发表的手稿，发现了黎曼计算过 $\zeta(s)$ 的很多非平凡零点，可以说黎曼猜想是黎曼经过大量的数值计算和理论推导后提出的，黎曼所发表只是他研究 $\zeta(s)$ 的一小部分结果。黎曼计算 $\zeta(s)$ 的零点时运用了 1932 年之前大家都不知道的公式，为了纪念西格尔在整理黎曼手稿方面的贡献，这个公式现在称为黎曼 - 西格尔（Riemann-Siegel）公式，这个公式计算 $\zeta(s)$ 零点的能力远远超过当时数学界所使用的方法。目前人们用计算机计算了八千多亿个 $\zeta(s)$ 的非平凡零点，它们全都在直线 $\mathrm{Re}(s)=1/2$ 上。

在理论研究方面，1914 年玻尔 - 兰道（Bohr-Laudan）证明了对任何 $\delta>0$, $\zeta(s)$ 在区域 $\{\mathrm{Re}(s)\geq 1/2+\delta, 0\leq \mathrm{Im}(s)\leq T\}$ 中的 $\zeta(s)$ 零点个数是 $O(T)$。可以证明在区域 $\{0\leq \mathrm{Re}(s)\leq 1, 0\leq \mathrm{Im}(s)\leq T\}$ 中 $\zeta(s)$ 的零点约为

$$\frac{T}{2\pi}\log\frac{T}{2\pi}-\frac{T}{2\pi}$$

（这是黎曼文章中的一个结论，严格证明是 1905 年冯•曼戈尔特给出的）。所以玻尔 - 兰道定理说明 $\zeta(s)$ 在区域 $\mathrm{Re}(s)\geq 1/2+\delta$ 中的零点在所有非平凡零点中所占比例为无穷小。也是在 1914 年哈代（G. H. Hardy）证明了 $\zeta(s)$ 有无穷多个零点落在直线 $\mathrm{Re}(s)=1/2$ 上。1921 年哈代 - 李特尔伍德（Littlewood）证明存在常数 K，使得对任何充分大的 T，$\zeta(s)$ 在线段 $\{\mathrm{Re}(s)=1/2, 0\leq \mathrm{Im}(s)\leq T\}$ 上零点个数$\geq KT$。1942 年阿特勒•塞尔贝格（Atle Selberg）将这个这个结果改进为 $\zeta(s)$ 在 $\{\mathrm{Re}(s)=1/2, 0\leq \mathrm{Im}(s)\leq T\}$ 上零点个数大于或等于 $KT\log T$。记线段 $\{\mathrm{Re}(s)=1/2, 0\leq \mathrm{Im}(s)\leq T\}$ 上 $\zeta(s)$ 零点个数为 $N_0(T)$，记区域 $\{0\leq \mathrm{Re}(s)\leq 1, 0\leq \mathrm{Im}(s)\leq T\}$ 中的 $\zeta(s)$ 零点个数为 $N(T)$，1974 年诺曼•莱文森（Norman Levinson）证明了 $N_0(T)\geq 1/3\mathrm{N}_0(T)$。后来人们对莱文森的方法作了改进，目前这方面的最好结果是 1989 年布莱恩•卡恩瑞（Brian Conrey）得到的

$$N_0(T)\geq 2/5N(T).$$

卢昌海对上面这些结果都作了深入浅出的介绍。很难得的是卢昌海对很多结果给出了富有启发性的数学推理，这些启发性的推理虽然不是严格的数学证明，但可以让读者相信推理和结论的合理性，从而去接受书中的数学结果。这些结果的严格证明和精确的数学描述往往很技术化，很容易让读者迷失方向。卢昌海还介绍了如何用黎曼 - 西格尔公式计算 $\zeta(s)$ 的一个零点，让读者亲身体会黎曼当年是如何对 $\zeta(s)$ 进行计算的。整本书给人的感觉一直是轻松愉快、引人入胜，本来我是打算利用闲暇时间翻翻这本书，事实是我拿到这本书后，放下了当天所有其它工作，一口气把这本书看完，看完后还把一些精彩片段复习了一遍。这部科普读物的数学描述部分是准确的，我自己的研究方向是平展上同调论，这是当年格罗登迪克（Alexandre Grothendieck）和德利涅（Pierre Deligne）为了证明韦伊猜想（Weil conjectures）而发展起来的，韦伊猜想在卢昌海的书中被称为“山寨版”黎曼猜想，应该说这是个很专门的话题，**但我发现卢昌海对韦伊猜想方面的介绍是很准确、专业的，让我难以置信的是卢昌海是个物理学博士。**

我发现的书中的一个漏洞是第 34 节最后一段话“有关自守 L- 函数的许多简单很多的性质，比如它的解析延拓及函数方程，也都还是未被普遍证明的东西”。事实上自守 L- 函数的解析延拓及函数方程已经证明了。算术代数几何中所谓的 motivic L- 函数的这些性质还没有证明。事实上罗伯特•朗兰兹（Robert Langlands）引进自守 L- 函数的目的之一就是为了研究 motivic L- 函数，如果能够建立朗兰兹对应，即证明 motivic L- 函数都是自守 L- 函数，则可以得到 motivic L- 函数的解析延拓及函数方程。例如怀尔斯（Andrew Wiles）证明费马大定理时证明了椭圆曲线的 L- 函数 (这是一类 motivic L- 函数) 是某个模形式的 L- 函数，这样就可以得到椭圆曲线的 L- 函数的解析延拓及函数方程。著名的阿廷猜想说数域的有限伽罗瓦扩张的非平凡不可约伽罗瓦群表示的 L- 函数 (这也是一类 motivic L- 函数) 是解析函数，如果能对这些 L- 函数建立朗兰兹对应，则可以证明阿廷猜想，目前阿廷猜想已解决的情形都是这样证明的。

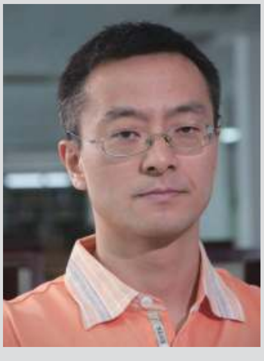
但上面的漏洞相对书中精彩片段实在是微不足道的。我会将该书推荐给每一位想对黎曼猜想有所了解的人。我建议读者在读完卢昌海的书后，看看爱德华兹（H. M. Edwards）的 *Riemann's Zeta Function*。这本书按照历史发展的顺序详细、严格地介绍了黎曼 ζ-函数的性质及其在素数分布的研究中的应用，同时也不乏富有启发性的推理，特别地，卢昌海书中大部分结果的严格数学证明都可以在这本书中找到。

在作者写这个书评前，作者、南开大学张伟平院士和主编有些邮件来往，附在本文后作为注记。

张伟平（2012年10月10日）： 汤涛，也许你想看看一位数论学者对你们期刊上卢昌海的连载文章的读后感。（Dear Tang Tao, I presume you'd like to see a remark from a number theorist on the articles of Lu Changhai published in your journal.）

扶磊（2012年10月9日）： 我读完了《黎曼猜想漫谈》一书。我所付出的代价是读此书期间中午睡不着觉，需要服用安眠药。明天我要把书中最精彩的部分再重读一遍，之后再把书还回图书馆。（I finish reading the book on Riemann hypothesis. One price I paid was that I couldn't fall asleep at noon and had to use sleeping pills. Tomorrow I will re-read the most wonderful part, and then give it to the library.）我在书中发现一个小错误……不过这个小错误完全可以忽略不计。我从书中学到了很多重要事实；这些知识是我很想知道但又找不到足够时间去啃书本得到的。（I spot one error……But this error is absolutely neglectable. I learned many important fact that I wanted to know but couldn't find time to go through those advanced books.）

扶磊（2012年10月8日）： 伟平，你昨天送给图书馆的"黎曼猜想"那本书写得真好，很奇怪的是作者卢昌海是个物理博士，但整本书写得很专业，书中有部分写的是我最熟悉的*Weil*猜想，这是个很专业的内容，但他写得没有任何漏洞，一本科普书能写得这么专业，或一本专业书能写得这么平易近人，真不容易，这本书我可能会推迟很长时间给图书馆，希望你别介意。作者还写了"寻找太阳系的疆界"，"太阳的故事"，这些物理方面的科普他应该写得更好了，你不妨在书店留意他写的其他的书。如果他写本量子物理就更好了。



作者简介：扶磊，1970 年生，武汉大学数学系毕业，美国 Rice 大学数学博士。南开大学陈省身数学研究所所长，长江学者特聘教授。