第一部分　数　　学

孪生素数猜想[①]

2003年3月28日，在美国数学研究所(American Institute of Mathematics)位于加州帕洛阿尔托(Palo Alto)的总部，一群来自世界各地的数学家怀着极大的兴趣聆听了圣荷西州立大学(San José State University)数学教授戈德斯通(Daniel Goldston)所做的一个学术报告。在这个报告中，戈德斯通介绍了他和土耳其海峡大学(Boğaziçi University)的数学家伊尔迪里姆(Cem Yıldırım)在证明孪生素数猜想(twin prime conjecture)方面所取得的一个进展。这一进展——如果得到确认的话——将把人们在这一领域中的研究大大推进一步。

那么，什么是孪生素数(twin prime)？什么是孪生素数猜想？戈德斯通和伊尔迪里姆所取得的进展又是什么呢？本文将对这些问题做一个简单介绍。

要介绍孪生素数，首先当然要说一说素数(prime number)这一概念。素数

① 本文撰写于2003年4月，是我的第一篇数学科普，填补了作为本人兴趣主要组成部分之一的数学在我网站的空白。自那以后，本文曾以"补注"形式对若干后续进展作了简单提及，并于2014年9月进行了不改变基本结构的轻微修订。

是除了 1 和自身以外没有其他因子的自然数。在数论中，素数可以说是最纯粹、也最令人着迷的概念。关于素数，一个最简单的事实就是：除了 2 以外，所有素数都是奇数(因为否则的话，除了 1 和自身以外还会有一个因子 2，从而不满足定义)。由这一简单事实可以得到一个简单推论，那就是：大于 2 的两个相邻素数之间的最小可能的间隔是 2。所谓孪生素数指的就是这种间隔为 2 的相邻素数，它们之间的距离已经近得不能再近了，就像孪生兄弟一样。不难验证，在孪生素数中，最小的一对是(3，5)，在 100 以内则还有(5，7)、(11，13)、(17，19)、(29，31)、(41，43)、(59，61)和(71，73)等另外 7 对，总计为 8 对。进一步的验证还表明，随着数字的增大，孪生素数的分布大体上会变得越来越稀疏，寻找孪生素数也会变得越来越困难。

那么，会不会在超过某个界限之后就再也不存在孪生素数了呢？

这个问题让我们联想到素数本身的分布。我们知道，素数本身的分布也是随着数字的增大而越来越稀疏的，因此也有一个会不会在超过某个界限之后就再也不存在的问题。不过幸运的是，早在古希腊时代，著名数学家欧几里得(Euclid)就证明了素数有无穷多个(否则的话——即假如素数没有无穷多个的话——今天的许多数论学家恐怕就得另谋生路了)。长期以来数学家们普遍猜测，孪生素数的情形与素数类似，虽然其分布随着数字的增大而越来越稀疏，总数却是无穷的。这就是与哥德巴赫猜想(Goldbach conjecture)齐名、集令人惊异的表述简单性与令人惊异的证明复杂性于一身的著名猜想——孪生素数猜想。

孪生素数猜想：存在无穷多个素数 p，使得 $p+2$ 也是素数。

究竟是谁最早明确地提出这一猜想我没有考证过，但 1849 年法国数学波利尼亚克(Alphonse de Polignac)曾提出过一个猜想：对于任意偶数 $2k$，存在无穷多组以 $2k$ 为间隔的素数。这一猜想被称为波利尼亚克猜想(Polignac's conjecture)。对于 $k=1$，它就是孪生素数猜想。因此人们有时把波利尼亚克作为孪生素数猜想的提出者。值得一提的是，人们对不同的 k 所对应的素数对

的命名是很有趣的：$k=1$(即间隔为2)的素数对我们已经知道叫做孪生素数；$k=2$(即间隔为4)的素数对被称为cousin prime(表兄弟素数)，比“孪生”稍远；而$k=3$(即间隔为6)的素数对竟被称为sexy prime！这回该相信“书中自有颜如玉”了吧？不过别想歪了，之所以称为sexy prime，其实是因为sex正好是拉丁文中的“6”(因此sexy prime的中文译名乃是毫无联想余地的“六素数”)。

孪生素数猜想还有一个更强的形式，是英国数学家哈代(Godfrey Hardy)和李特伍德(John Littlewood)于1923年提出的，有时被称为哈代-李特伍德猜想(Hardy-Littlewood conjecture)或强孪生素数猜想(strong twin prime conjecture)①。这一猜想不仅提出孪生素数有无穷多组，而且还给出其渐近分布为

$$\pi_2(x) \sim 2C_2\int_2^x \frac{\mathrm{d}t}{(\ln t)^2}$$

其中$\pi_2(x)$表示小于x的孪生素数的数目，C_2被称为孪生素数常数(twin prime constant)，其数值为

$$C_2 = \prod_{p\geqslant 3} \frac{p(p-2)}{(p-1)^2} \approx 0.660\,161\,181\,584\,686\,957\,392\,781\,211\,001\,45\cdots$$

强孪生素数猜想对孪生素数分布的拟合程度可以由表1看出。很明显，拟合程度是相当漂亮的。假如可以拿观测科学的例子来作比拟的话，如此漂亮的拟合几乎能跟英国天文学家亚当斯(John Couch Adams)和法国天文学家勒维耶(Urbain Le Verrier)运用天体摄动规律对海王星位置的预言，以及爱因斯坦(Albert Einstein)的广义相对论对光线引力偏转的预言等最精彩的观测科学成就相媲美，可以算同为理性思维的动人篇章。这种拟合对于纯数

① 确切地说，哈代和李特伍德于1923年所提出的猜想共有两个，分别称为第一哈代-李特伍德猜想(first Hardy-Littlewood conjecture)和第二哈代-李特伍德猜想(second Hardy-Littlewood conjecture)。其中第一哈代-李特伍德猜想又称为k-tuple猜想(k-tuple conjecture)，它给出了所有形如$(p, p+2m_1, \cdots, p+2m_k)$(其中$0<m_1<\cdots<m_k$)的素数$k$-tuple的渐进分布。强孪生素数猜想只是$k$-tuple猜想的一个特例。

学的证明来说虽起不到实质帮助，却大大增强了人们对孪生素数猜想的信心。

表 1

x	孪生素数数目	强孪生素数猜想给出的数目
100 000	1 224	1 249
1 000 000	8 169	8 248
10 000 000	58 980	58 754
100 000 000	440 312	440 368
10 000 000 000	27 412 679	27 411 417

在这里还可以顺便提一下，强孪生素数猜想所给出的孪生素数分布规律可以通过一个简单的定性分析来“得到”①：我们知道，素数定理（prime number theorem）表明对于足够大的 x，在 x 附近素数的分布密度大约为 $1/\ln(x)$，因此两个素数位于宽度为 2 的区间之内（即构成孪生素数）的概率大约为 $2/\ln^2(x)$。这几乎正好就是强孪生素数猜想中的被积函数——当然，两者之间还差了一个孪生素数常数 C_2，而这个常数显然正是哈代和李特伍德的功力深厚之处②。

除了强孪生素数猜想与孪生素数实际分布之间的漂亮拟合外，对孪生素数猜想的另一类“实验”支持来自于对越来越大的孪生素数的直接寻找。就像对大素数的寻找一样，这种寻找在很大程度上成为了对计算机运算能力的一种检验。1994 年 10 月 30 日，这种寻找竟然使人们发现了英特尔（Intel）奔腾（Pentium）处理器浮点除法运算的一个瑕疵（bug），在工程界引起了不小的震动。截至 2002 年底，人们发现的最大的孪生素数是：

$$(33\,218\,925\times2^{169\,690}-1,\ 33\,218\,925\times2^{169\,690}+1)$$

这对素数中的每一个都长达 51 090 位。许多年来这种纪录一直被持续而成功地刷新着，它们对于纯数学的证明来说虽也起不到实质帮助，却同样有助于

① 这种定性分析被澳大利亚数学家陶哲轩（Terence Tao）称为“概率启发式理由”（probabilistic heuristic justification），它不是证明，但对于判断命题成立与否有一定的启示性。

② 对孪生素数常数 C_2 也存在“概率启发式理由”，感兴趣的读者可参阅美国数学家查基尔（Don Zagier）的“The First 50 Million Prime Numbers”，Math. Intel. 0，221-224（1977）。

增强人们对孪生素数猜想的信心①。

好了，介绍了这么多关于孪生素数的资料，现在该说说人们在证明孪生素数猜想上所走过的征途了。

迄今为止，在证明孪生素数猜想上的成果大体可以分为两类。第一类是非估算性的，这方面迄今最好的结果是 1966 年由中国数学家陈景润利用筛法(sieve method)所取得的②。陈景润证明了：存在无穷多个素数 p，使得 $p+2$ 要么是素数，要么是两个素数的乘积。这个结果的形式与他关于哥德巴赫猜想的结果很类似③。目前一般认为，由于筛法本身所具有的局限性，这一结果在筛法的范围之内已很难被超越。

证明孪生素数猜想的另一类结果则是估算性的，戈德斯通和伊尔迪里姆所取得的结果就属于这一类。这类结果估算的是相邻素数之间的最小间隔，更确切地说是：

$$\Delta = \liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\ln p_n}$$

翻译成白话文，这个表达式所定义的是两个相邻素数之间的间隔与其中较小的那个素数的对数值之比在整个素数集合中所取的最小值。很明显，孪生素数猜想要想成立，Δ 必须为 0。因为孪生素数猜想表明 $p_{n+1}-p_n=2$ 对无穷多个 n 成立，而 $\ln(p_n)\to\infty$，因此两者之比的最小值对于孪生素数集合——从而对于整个素数集合也——趋于零。不过要注意，$\Delta=0$ 只是孪生素数猜想成立的**必要条件**，而不是充分条件。换句话说，如果能证明 $\Delta\neq 0$，则孪生素数猜想就被推翻了；但证明了 $\Delta=0$，却并不意味着孪生素数猜想一定成立。

---

① 截至 2011 年底，这一纪录已被刷新为了：$(3\,756\,801\,695\,685\times 2^{666\,669}-1,\ 3\,756\,801\,695\,685\times 2^{666\,669}+1)$，这对素数中的每一个都长达 200 700 位。

② 顺便说一下，美国数学研究所在介绍本文开头所提到的戈德斯通和伊尔迪里姆的结果的简报中提到陈景润时所用的称呼是“伟大的中国数学家陈”(the great Chinese mathematician Chen)。

③ 陈景润关于哥德巴赫猜想的结果——被称为陈氏定理(Chen's theorem)——是：任何足够大的偶数都可以表示成两个数的和，其中一个是素数，另一个要么是素数，要么是两个素数的乘积。

对 Δ 最简单的估算来自于素数定理。按照素数定理，对于足够大的 x，在 x 附近素数出现的几率为 $1/\ln(x)$，这表明素数之间的平均间隔为 $\ln(x)$，从而 $(p_{n+1}-p_n)/\ln(p_n)$ 给出的其实是相邻素数之间的间隔与平均间隔的比值，其平均值显然为 1①。平均值为 1，最小值显然是小于等于 1，因此素数定理给出 $\Delta\leqslant 1$。

对 Δ 的进一步估算始于哈代和李特伍德。1926 年，他们运用圆法（circle method）证明了假如广义黎曼猜想（generalized Riemann hypothesis）成立，则 $\Delta\leqslant 2/3$。这一结果后来被苏格兰数学家兰金（Robert Alexander Rankin）改进为 $\Delta\leqslant 3/5$。但这两个结果都有赖于本身尚未得到证明的广义黎曼猜想，因此只能算是有条件的结果。1940 年，匈牙利数学家埃尔德什（Paul Erdös）利用筛法率先给出了一个不带条件的结果：$\Delta<1$（即把素数定理给出的结果中的等号部分去掉了）。此后意大利数学家里奇（Giovanni Ricci）于 1954 年，意大利数学家蓬皮埃利（Enrico Bombieri）、英国数学家达文波特（Harold Davenport）于 1966 年，以及英国数学家赫克斯利（Martin Huxley）于 1977 年，分别将 Δ 的估算值推进到了 $\Delta\leqslant 15/16$，$\Delta\leqslant(2+\sqrt{3})/8\approx 0.4665$，以及 $\Delta\leqslant 0.4425$。戈德斯通和伊尔迪里姆之前最好的结果则是德国数学家梅尔（Helmut Maier）于 1986 年得到的 $\Delta\leqslant 0.2486$。

以上这些结果都是在小数点后面做文章，戈德斯通和伊尔迪里姆的结果将这一系列努力大大推进了一步，并且——如果得到确认的话——将在一定意义上终结对 Δ 进行数值估算的长达几十年的漫漫征途。因为戈德斯通和伊尔迪里姆所证明的结果是 $\Delta=0$。当然，如我们前面所述，$\Delta=0$ 只是孪生素数猜想成立的必要条件，而不是充分条件，因此戈德斯通和伊尔迪里姆的结果即便得到确认，离最终证明孪生素数猜想仍有相当的距离，但它无疑将是近十几年来这一领域中最引人注目的结果。

一旦 $\Delta=0$ 被证明，下一个努力方向会是什么呢？一个很自然的方向将是研究 Δ 趋于 0 的方式。孪生素数猜想要求 $\Delta\sim[\ln(p_n)]^{-1}$（因为 $p_{n+1}-$

① 这个“归一”性也正是在 Δ 的表达式中引进 $\ln(p_n)$ 的原因。

$p_n=2$ 对无穷多个 n 成立）。戈德斯通和伊尔迪里姆的结果所给出的则是 $\Delta\sim[\ln(p_n)]^{-1/9}$，两者之间还有不小的差距①。但是看过戈德斯通和伊尔迪里姆手稿的一些数学家认为，戈德斯通和伊尔迪里姆所用的方法还存在改进空间。也就是说，他们的方法还有可能对 Δ 趋于 0 的方式作出更强的估计。从这个意义上讲，戈德斯通和伊尔迪里姆这一结果的价值不仅仅在于结果本身，更在于它有可能成为一系列未来研究的起点。这种带传承性的系列研究对于数学来说有着双重的重要性，因为一方面，这种研究可能取得的新结果将是对数学的直接贡献；另一方面，这种研究对戈德斯通和伊尔迪里姆的结果会起到反复推敲与核实的作用。现代数学早已超越了一两个评审花一两个小时就可以对一个数学证明做出评判的时代。著名的四色定理（four color theorem）和费马大定理（Fermat's Last Theorem）都曾有过一个证明时隔几年、甚至十几年才被发现错误的例子。因此，一个复杂的数学结果能成为进一步研究的起点，吸引其他数学家的参与，对于最终判定其正确性具有极其正面的意义②。

2003 年 4 月 6 日写于纽约

2014 年 9 月 15 日最新修订

---

① 本文发布之后，关于戈德斯通和伊尔迪里姆的工作又有了一些重要的后续发展，其中包括：2003 年 4 月 23 日，英国数学家格兰维尔（Andrew Granville）和印度数学家桑德拉拉扬（Kannan Soundararajan）发现了戈德斯通和伊尔迪里姆原始证明中的一个错误，并得到了戈德斯通和伊尔迪里姆的承认；2005 年初，戈德斯通和伊尔迪里姆"伙同"匈牙利数学家平兹（János Pintz）"卷土重来"，再次证明了 $\Delta=0$。他们所证明的 Δ 的新的渐进行为是：$\Delta\sim[\ln\ln(p_n)]^2/[\ln(p_n)]^{1/2}$。

② 2013 年 5 月 14 日，《自然》（*Nature*）等科学杂志及大量中外媒体报道了旅美数学家张益唐在孪生素数猜想研究中所取得的一个重要的新进展，即证明了存在无穷多个素数对，其间隔小于 7 000 万。这一进展——如果得到确认的话——相当于证明了波利尼亚克猜想至少对某个小于 3 500 万的 k 成立。用 Δ 来表述的话，则相当于不仅证明了 $\Delta=0$，而且给出了与孪生素数猜想所要求的相同的渐进行为：$\Delta\sim[\ln(p_n)]^{-1}$（不过，这一渐进行为跟 $\Delta=0$ 一样，只是孪生素数猜想成立的必要条件，而不是充分条件）。张益唐的证明用到了戈德斯通、平兹、伊尔迪里姆等人的结果，并于 2013 年 5 月 21 日被《数学年刊》（*Annals of Mathematics*）所接受。张益唐的结果也存在改进空间，截至 2014 年 3 月，陶哲轩等数学家已将其中的 7 000 万这一素数间隔"压缩"到了 246。

魔方与“上帝之数”[①]

2008年7月，来自世界各地的很多优秀的魔方玩家聚集在捷克共和国(Czech Republic)中部城市帕尔杜比采(Pardubice)，参加魔方界的重要赛事：捷克公开赛。在这次比赛上，荷兰玩家阿克斯迪杰克(Erik Akkersdijk)创下了一个惊人的纪录：只用7.08秒就复原了一个颜色被打乱的魔方。无独有偶，在这一年的8月，人们在研究魔方背后的数学问题上也取得了重要进展。在本文中，我们就来介绍一下魔方以及它背后的数学问题。

一、风靡世界的玩具

1974年春天，匈牙利布达佩斯应用艺术学院(Budapest College of Applied Arts)的建筑学教授鲁比克(Ernö Rubik)萌生了一个有趣的念头，那就是设计一个教学工具来帮助学生直观地理解空间几何中的各种转动。经过思考，他决定制作一个由一些小方块组成的，各个面能随意转动的3×3×3的立方体。这样的立方体可以很方便地演示各种空间转动。

① 本文曾发表于《科学画报》2008年第12期(上海科学技术出版社出版)。