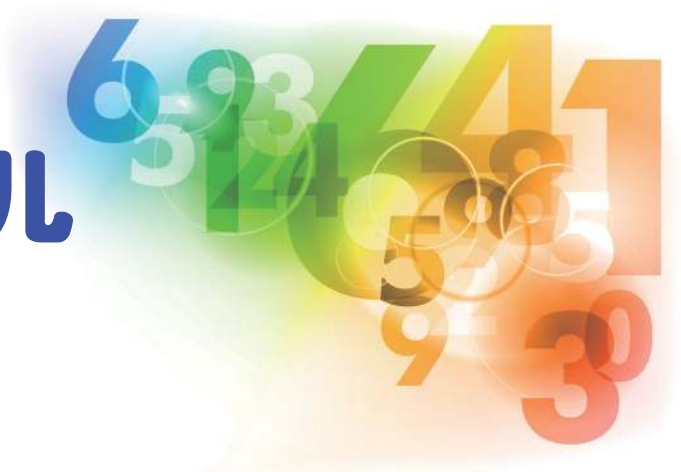
素数的那些事儿



陆俊

1 引子

近几年开设《初等数论》课程，我总是要抽出一定的时间专门给学生科普一下关于素数的故事。这篇科普文章正是基于这些讲稿整理出来的。

素数是整个数论的灵魂。然而多数学生对素数的了解非常少。很多人不明白：为什么我们要研究素数？素数如何与众不同？素数到底有趣在哪里？素数对数学很重要吗？……如果学生在上完一个学期的数论课后，却仍然对素数茫然无知，那无疑是一种讽刺——这就好比你看完一场戏，不知道主角做了些什么。

写这篇文章的另一目的也是为了给那些依然执着于证明哥德巴赫猜想的民科们做一次扫盲的尝试——尽管他们中的大多数会继续执着下去。然而我们不得不承认这样一个现实：民科们对素数的热情与执着确实远远超过很多数学系的本科生——这多少会让我们这些老师感到沮丧。

2 素数有多少？

我们说一个整数 a 能被另一个整数 b 整除，就是指 $\frac{a}{b}$ 是整数。有时我们也把 b 称作 a 的因子。

一个素数（Prime Number）是指这样一种正整数：除了1和它本身之外，其它任何正整数都不可能整除它。我们也可以这么定义素数：它不能写成两个大于1的正整数的乘积。有时我们也将素数称作质数。通常我们不承认1是素数。这样做的好处下面会介绍。除了1和素数之外，其他的正整数统称为合数。

最初的几个素数是2, 3, 5, 7, 11,…。显然6不是素数，因为 $6=2\times3$，所有的素数中只有2是偶数！这件看似平凡的事，其实很重要。在许多数学研究中，2和其他素数会对我们所考虑的问题产生不同的影响。你可能会问：为什么我们把这样的数命名为“素数”呢？这实际上来自于素数最基本的结论——**算术基本定理**：

> 任何大于1的正整数 n 都可以唯一地分解成一些素数的乘积 $n=p_1\cdots p_s$，这里 $p_1\le p_2\le\cdots\le p_s$ 都是素数（允许相同）。

无论如何，素数本身不能再进一步分解成一些更小的正整数乘积，因此它在此意义下是最基本或最本质的数——类似于朴素的原子论——从而命名它们为“素数”或者“质数”。容易看到，假如我们承认1是素数，那么算术基本定理就不能保证分解是唯一的了。因为1可以写为任意多个自身的乘积。

接下去，一个最自然不过的问题当然是：究竟有多少素数？无限多个还是仅有有限个？这个问题的答案早由欧几里德在两千多年前解决了。他用初等方法巧妙地证明：存在无限多个素数！具体言之，我们假设所有正整数中只有有限个素数 $p_1,\cdots,p_n$，那么可以构造一个正整数

$$N=p_1p_2\cdots p_n+1\ .$$

很容易发现，左边的 N 分解成素数乘积的话，不可能包含任何素数 p_i，因此它的分解式中必定含有这些 p_i 之外的新素数。这就和我们的假设矛盾！

这个证明包含了富有启发性的思想。事实上，证明本身并没有提供构造出所有素数的具体方法。但是它却能告诉我们素数有无限个！这就是数学中所谓的“存在性证明”：它告诉你某些对象存在，但是却没有具体构造出来。“存在性”是数学哲学中的一个深刻话题，涉及到数学大厦的根基。数学史上曾经关于这类问题有过广泛而激烈的争论，有人反对这种类型的证明，有人却支持它们。这场争论涉及了许多重要的数学家，产生了许多和数学、逻辑、哲学相关的理论。有兴趣的读者可以参考相关书籍，此处不再赘述。

类似欧几里德的证明，你也可以轻松断言：所有被4除余数为3的素数有无限个！换言之，就是等差数列3, 7, 11,

15, …中包含无穷多个素数。这就产生了一个有趣的问题：

一个等差数列 $a, a+b, a+2b, \cdots, a+nb, \cdots$ 中是否包含无限多个素数？

数学家狄利克雷回答了这一问题（**狄利克雷定理**）：

> 假如 a 和 b 是互素的（就是说它们不能同时被一个大于 1 的正整数整除），那么答案是肯定的！

不要以为欧几里德的方法可以轻松解决这一问题哦。事实上，除了少数情形之外，这个问题是不可能用它来简单解决的。

如果我们把等差数列换成其他数列，结论会怎样呢？比如考虑以下的数列：

$$2,5,10,17,26,\cdots,n^2+1,\cdots$$

其中是否有无限多个素数呢？让人颇为失望的是，这至今仍是一个未解决的难题。

3 素数是怎么分布的？

知道“素数有无限多个”仅仅是个开始。我们还想知道更多！比如，素数在所有自然数中所占的比率多大？当然，我们首先要说明“比率”在这里意味着什么。对任何正实数 x，我们用 $\pi(x)$ 表示不超过 x 的素数的个数。比如 $\pi(1)=0$, $\pi(4)=2$, $\pi(2.5)=1$ 等等。我们用 $\frac{\pi(x)}{x}$ 来反映所有不超过 x 的正整数中，素数所占的比率——也称作平均分布密度。

一个简单的结论告诉我们：当 x 非常非常大时，$\frac{\pi(x)}{x}$ 几乎就等于 0。换句话说，素数在所有正整数中极为罕见，可以说少得几乎没有——尽管我们知道它们有无穷多个！这就好比宇宙中有生命的星球也许有无限多个，但是它们相隔得太远，相对整个宇宙来说实在是十分稀疏罕见的。

对一般人来说，这个结论似乎已经让我们走到了问题的尽头。但是天才数学家高斯却不这么认为。在那个没有计算机的年代（1792-1793 年间），他通过大量的手工计算，单凭超人的直觉，竟然得到了一个让人吃惊的猜测（但其本人并未证明）：

> 当 x 非常大时，素数出现的比率 $\frac{\pi(x)}{x}$ 约等于 $\frac{1}{\log x}$。换言之，$\frac{\pi(x)}{x/\log x}$ 约等于 1，这里 $\log x$ 是 x 的对数函数。

高斯原始的猜测要比上面的表述式更为精确。在高斯之后，数学家勒让德实际上也通过数值计算得到过类似的猜测公式（1800 年左右），但没有高斯的精确。证明这一结论是极其困难的工作。直到 19 世纪中叶，俄国数学家切比雪夫才有了突破性进展，他证明了：

$$C_1 \le \frac{\pi(x)}{x/\log x} \le C_2 ,$$

这里 C_1 和 C_2 是确定的常数。此猜想大约到 19 世纪末，才由法国数学家阿达玛和 Paussin 几乎同时独立证明。人们将它称作**素数定理**。阿达玛等人的证明是建立在天才数学家黎曼的研究基础上的，用到了极为高深的函数理论。到了 1949 年前后，才由数学家爱尔特希和塞尔伯格给出了初等证明。请注意，这里所谓的“初等”只是说没有用到太多高深的数学理论，但是证明本身是很复杂的，也较为难懂。数学中有很多这样的问题（比如哥德巴赫猜想），它们表面上很简单，但实际上要证明它们往往是极其困难的。

素数定理只是在大样本范围内描述了一种统计规律。素数本身的分布位置极不规则。当你确定一个素数之后，很难预测在它之后的下一个素数是多少。尽管如此，我们仍有一些猜测和结论来描绘素数在整数集中分布性态。有趣的是，猜想要比结论多得多。

首先是著名的**伯特兰猜想**（后被切比雪夫证明），它断言：

对任何大于 1 的正整数 n，必定有素数落在 n 和 $2n$ 之间。

比如 n 取 4，那么在 4 到 8 之间我们可以找到素数 5 和 7。当 n 非常大时，这一结论显然是素数定理的直接推论。

你可以随手举出很多类似伯特兰 - 切比雪夫定理的猜想，比如在 n^2 和 $(n+1)^2$ 之间是否必有素数存在？这一看似简单的问题实际上至今仍未解决！

其次是著名的**孪生素数猜想**：

是否存在无限多个素数 p，使得 $p+2$ 也是素数？

我们将这样的一对素数 $(p, p+2)$ 称为孪生素数对。比如 (3, 5), (5, 7), (11, 13) 等等都是孪生素数对。类似地，你也可以定义三生素数对（$p, p+2, p+6$），亦即要求这三个数同时为素数。三生素数猜想就是问：是否存在无限个三生素数对？回答仍是“不知道”。我们也可以定义 n 生素数对，并提出类似的猜测。有趣的是，有人证明：n 生素数猜想和以下的三角不等式猜测互为矛盾——也就是说不可能同时正确：

$$\pi(x+y) \le \pi(x)+\pi(y) ,$$

这里 $\pi(x)$ 定义同前。

另一个著名的猜想就是在国内广为人知的**哥德巴赫猜想**：

> 任何大于等于 6 的偶数必定能写成两个奇素数之和；任何大于等于 9 的奇数都是三个奇素数之和。

这个猜想和陈景润的名字联系在一起，带有很多现代历史的色彩。许多民科投身于哥德巴赫猜想的证明也与此有关。哥德巴赫只是一个普通的数学家，除了提出这个猜想之

外没有什么数学贡献。他将这一猜测告诉了天才数学家欧拉。遗憾的是，后者未能证明它，但是该猜想却得以被很多人知道。容易看到，哥德巴赫猜想第二部分只不过是第一部分的简单推论。但有趣的是，第二部分反而先被证明了（称作**三素数定理**），第一部分却迟迟得不到解决。目前最好的结果是陈景润的“1+2”定理，即充分大的偶数都可以写成一个素数和一个不超过两个素数乘积的数之和。哥德巴赫猜想的研究是十分艰难的，它本质上涉及到十分深刻的函数论知识，不可能如那些民科所妄想的那样，拍拍脑袋就能用初等方法做出来。

虽然我们无法彻底证实这个猜想，但是却可以退而求其次，用所谓的*密率方法*得到以下的有趣结论：

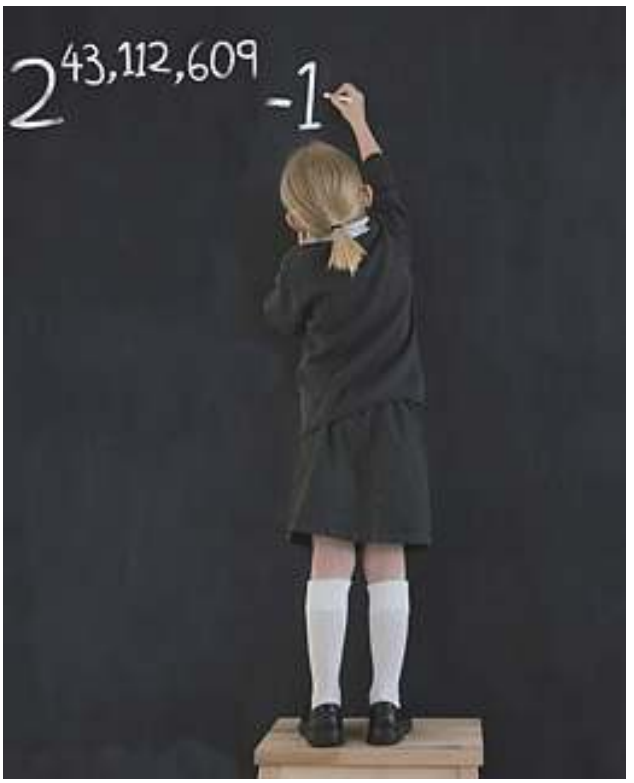
> 任何大于1的正整数必可写成不超过26个素数之和。

4 如何构造素数？

上面的讨论只是介绍了素数在整数中的分布情况，但是我们至今还没有具体构造出这些素数来。一个基本的问题就是：如何构造素数？最原始的办法就是古典筛法。比如我们要找出所有不超过 100 的素数，那么首先将所有从 $4=2^2$ 开始的偶数全部从这 100 个数中去除掉；接着将所有从 $9=3^2$ 开始的 3 的倍数全部去除掉；再将所有从 $25=5^2$ 开始的 5 的倍数全部去除掉……以此类推，最终通过筛选剩下的数恰好就是所有不超过 100 的素数。

上面的筛法虽然可以逐一列出不超过某个上限 N 的全部素数，但是当 N 很大时，其工作量也是巨大的。因此人们开始寻找其他方法来构造素数。通常的思路是构造一个有规律的数列 $\{a_n\}_{n=1}^{\infty}$，使得数列中每一项都是素数。这样的数列称作*素数公式*。比如费马构造了以下数列（*费马数*），并

数学家群星图；排在最上面的是数论结果及高斯和欧拉等数论先驱



美国一家网络安全基金会悬赏超长素数；1 千万位素数的十万美元被加州大学的 Edson Smith 领走。他找到了第 46 个梅森数（见上图）

猜测它们都是素数：

$$F_n = 2^{2^n}+1, \qquad n=0,1,2,\cdots$$

他计算了前五项，即 3, 5, 17, 257, 65537，确实都是素数。然而欧拉以其高超的计算能力手算验证了 $F_5 = 641\times 6700417$ 不是素数。事实上，由目前的计算机验算可知，从 F_5 到 F_{11} 都不是素数。是否存在无限多个费马素数？这是一个未解之谜。

尽管欧拉的计算粉碎了利用费马数构造素数公式的企图，但是这并不表示研究费马数没有意义。高斯在年少时期，证明了一个让人无比惊叹的奇妙结论，让费马素数声名大噪。这个结论（**正 *m* 边形尺规作图**）是说：

> 一个正 m 边形能用尺规作图得到的充分必要条件是：$m=2^e p_1 p_2\cdots p_s$，这里诸 p_i 是费马素数，且两两不同。

比如 $m = 17$ 是费马素数，因此可以用尺规作图得到正十七边形！要知道，在高斯之前的几百年，有那么多人研究尺规作图问题，但谁也没有想到正十七边形居然可以尺规作图得到。这个结论的重要性在于，它将几何（代数）问题和数论问题这两个看似无关的领域奇妙地结合起来。

与费马数对应的是著名的梅森数列：

$$M_p = 2^p-1, \qquad p=2,3,5\cdots$$

这里 p 依次取遍所有的素数。人们也曾猜测梅森数都是素数。比如前几项分别为 3, 7, 31, 127 都是素数。但是 $M_{11} = 23\times 89$ 不是素数。一个有趣的结论断言：

> 如果素数 p 被 4 除余数为 3，并且 $2p+1$ 也是素数，那么 M_p 必定不是素数。

梅森数的遭遇与费马数类似，尽管没有能够达到原始的构造素数公式的目的，但是它却和另一个著名的定理联系起来。为了叙述该定理，我们做些准备工作。给一个正整数 n，我们把它的所有可能的因子（就是能整除 n 的那些正整数）加起来得到的总和记作 $\sigma(n)$。如果 n 满足 $\sigma(n) = 2n$，那么我们就称 n 是完全数。比如 $n = 6$ 就是完全数，因为 n 的因子只有 1, 2, 3, 6, 加起来正好是 12。同样地，$n = 28$ 也是完全数。我们有如下的**偶完全数定理：**

> 所有的偶完全数都可以写成 $\frac{1}{2}p(p+1)$，这里 p 是梅森素数。反之，这样的表达式得到的数也必定是偶完全数。

上面说的 $n = 6$ 恰好写成 $\frac{1}{2}3(3+1)$，其中 3 是梅森素数；28 可以写为 $\frac{1}{2}7(7+1)$，其中 7 是梅森素数。

由此产生另一个有趣的问题：奇完全数存在吗？这又是数论中一个至今悬而未决的著名猜想。人们借助计算机检验了 10^{300} 以内所有数，竟然都没能找到奇完全数！另外一个同样让人沮丧的事实是，至今我们还不知道是否有无穷多个梅森素数。

欧拉提出了另一类构造素数的方式。比如考虑多项式 $n^2 - n + 41$。当 n 从 0 取到 40 时，多项式的值皆素数。我们同样可以考虑多项式

$$n^2-n+p\ ,$$

这里 p 是素数，使得当 n 从 0 开始直至某个数 N 为止逐一代入时，上述多项式取值始终为素数。对任意 N，我们是否总能找到这样的 p 满足上面的要求呢？这个有趣的问题也是未解决的难题之一。让我们在上述多项式中分别取 $n = 1, 2, 3$, 那么得到的三个值恰好为 $p, p+2, p+6$。如果上面的问题答案是肯定的话，这就立刻证实了前文所述的孪生素数猜想和三生素数猜想！因此很显然上面的问题要远远难于孪生或三生素数猜想。

有趣的是，假如我们要求上述 $N=p-1$ 的话，那么这个历史上著名的难题有一个极为漂亮的解答（**Rabinovtch 定理**）：

> 设 $m\geq 2$，二次多项式 $f_m(x)=x^2-x+m$ 当 $x=0, 1, \cdots m-1$ 时总取素数值的充分必要条件是 m 为以下正整数之一：2，3，5，11，17，41.

（注记：上面的 m 的六种取值是有深刻背景的，它们恰好对应所有类数为 1 的虚二次域。）

看了这样几个例子之后，你也许会问：是否真有素数公式呢？有是有的，但这类公式通常意义不大。我们这里举一个例子来说明。为此需要一些准备。对任何实数 x，我们用 $[x]$ 表示不超过 x 的最大整数。比如 $[1.2]=1, [0]=0, [-1.2]=-2$。这个记号是高斯引进的，称作 x 的取整函数或高斯函数。此外我们用 $n!$ 表示乘积 $1\times 2\times 3\cdots n$，它称作 n 的阶乘。

现在我们可以构造素数公式

$$a_n=n+(n-2)\times\left(\left[\frac{(n-1)!+1}{n}\right]+\left[-\frac{(n-1)!+1}{n}\right]\right).$$

对任何大于 1 的正整数 n，项 a_n 都是素数，比如最前面的几项分别为 2, 3, 2, 5, 2, 7, … 这个素数公式表面上看似乎很神奇，其实并没有太多的新东西。它只是利用了和素数有关的**威尔逊定理**：

> 一个大于 1 的正整数 n 是素数的充分必要条件为：$\frac{(n-1)!+1}{n}$ 是整数。

比如 $n=5$ 是素数，$4!+1=25$ 显然是 5 的倍数。另一方面，对任何非整数的实数 x，总有 $[x]+[-x]=-1$；而对整数 x，则有 $[x]+[-x]=0$。现在我们能够看到，a_n 实际上在 n 是素数时就等于 n，在其他情况下都取值 2。

5 素数和方程

素数的很多有趣的性质都是从方程开始的。在经典数论（即研究整数性质的理论）中，人们关心的一个问题就是某些多项式方程是不是有整数解。比如著名的勾股方程 $x^2+y^2=z^2$ 或者佩尔方程 $x^2-dy^2=1$ 等。

当我们只关心这类方程的整数解（或有理数解等等）时，这样的方程通常就称为不定方程或者丢番图方程。我们这里介绍一些和素数有关的方程。它们中的一些对数学的发展起到了很大的作用。以下设 p 是素数。

(1) 我们考虑不定方程

$$x(x-1)\cdots(x-(p-1))=py.$$

对任一整数 n，假设它被 p 除的余数是 r。那么 $n-r$ 显然被 p 整除。因此对任何整数 n，p 总是能整除连乘积 $n(n-1)\cdots(n-(p-1))$。这样，当我们把 $x=n$ 代入上述方程的左边，则可求出整数解 y。利用这个简单的不定方程和下面的费马小定理，人们可以得到前面提及的威尔逊定理。

(2) **费马小定理**断言：不定方程

$$x^p-x=py,$$

对任何整数 $x=n$，都有整数解 y 存在。换句话说，对任何整数 n，素数 p 总是能整除 n^p-n。费马小定理的“小”字是相对费马大定理（也称费马最后的定理或者费马猜想）而言的，他声称方程

$$X^n+Y^n=Z^n,\qquad n>2,$$

没有非零的整数解 (X, Y, Z)。但费马只证明了 $n=4$ 的情形。其余情形经过许许多多数学家的艰苦努力，终于在 1994 年前后由英国数学家怀尔斯彻底解决。如果我们将素数 p 换成一般的整数 m，那么费马小定理可以推广到一般情形——即欧拉定理。我们这里不打算讨论它。

(3) 假设 n 是一个整数，并且不为 p 整除。我们来求解不定方程

$$nx-py=1$$

的整数解 (x, y)。由费马小定理，我们取 $x=n^{p-2}$，即可求得整数 y。从几何的角度看，上面的方程在平面上描绘了一条直线。因此我们相当于要找出该直线上的整数点 (x, y)。

(4)（**平方剩余**）求解以下方程的整数解是初等数论的核心问题之一：

$$x^2-py=q.$$

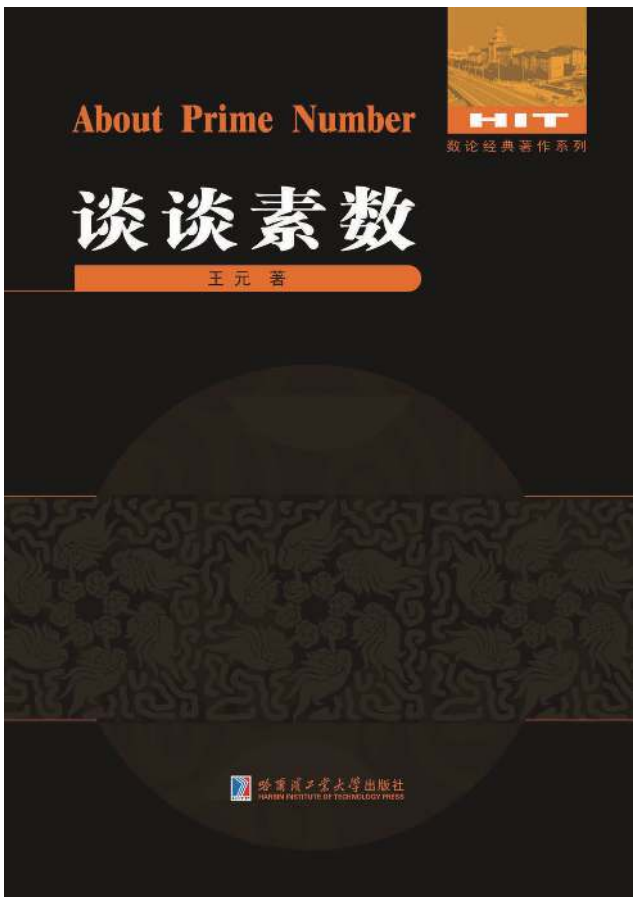
这里 q 是给定的正整数。高斯在其名著《算术研究》中对此作了深入研究，给出了一系列漂亮的研究成果。与前面几个方程不同的是，该方程有可能无整数解，比如

$$x^2-3y=2.$$

如果 $x^2-py=q$ 有解，我们就说 q 在模 p 下是平方剩余的——这里不打算解释该术语的意思。为了表示该方程是否有解，勒让德引进了一个方便的符号

$$\left(\frac{q}{p}\right)=\begin{cases}1, & \text{若有解},\\ -1, & \text{若无解}.\end{cases}$$

若 $p=2$ 的话，方程当然有解，因此我们只关心 p 是奇素数的情形。此时高斯用巧妙的初等方法（但并不显而易见）首



王元院士 2011 年的科普新作《谈谈素数》

先得到以下结果

$$\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}.$$

这就是说，如果 p 被 8 除余数为 1 或 7，那么方程 $x^2-py=2$ 必定有解；否则就无解。类似地还有

$$\left(\frac{-1}{p}\right)=(-1)^{\frac{p-1}{2}}.$$

当 q 也是奇素数时，高斯发现了数论史上具有里程碑意义的重要结果，即著名的**二次互反律**：

$$\left(\frac{q}{p}\right)\cdot\left(\frac{p}{q}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

高斯将此定理比喻为“数论之酵母”，意思是说这个深刻定理对数论的研究极其重要。近代数学的研究也验证了这一观点。事实上这一定理经过许多数学家——比如希尔伯特、高木贞治等——的努力，推广到代数数论研究中，发展成了深刻的理论——类域论。二次互反律也可以这样解释：

当 p, q 中有一个被 4 除余数是 1 时，以下两个方程

$$x^2-py=q,\qquad z^2-qw=p$$

要么同时有解要么同时无解；当 p, q 被 4 除余数都是 3 时，那么其中一个方程有解而另一方程无解。

对一般的整数 q，由算术基本定理，我们可以将它写成素因子乘积 $q=\pm p_1\cdots p_s$，从而可以通过以下关系求出勒让德符号的值以判断方程是否有解：

$$\left(\frac{q}{p}\right)=\left(\frac{\pm1}{p}\right)\left(\frac{p_1}{p}\right)\cdots\left(\frac{p_s}{p}\right).$$

从几何角度看，平方剩余问题对应的方程就是平面上的一条抛物线。因此我们等价于寻找抛物线上的整数点 (x, y)。

(5) **两平方和问题**：一个素数 p 什么时候可以写成两个平方数之和？比如

$$5=1^2+2^2,\quad 13=2^2+3^2,\quad 17=1^2+4^2,\cdots$$

容易检验，7 不可能写成两平方数之和。这个问题等价于求不定方程

$$x^2+y^2=p$$

的整数解。从几何上看，相当于寻找一个圆周上的整数点。费马早在 1640 年前后就得到了以下的结论，但是并未正式发表；欧拉最早给出了其正式证明。

> 如果 p 被 4 除余数是 3，那么上述方程无解，即不能分解为两个平方数之和；如果 p 被 4 除余数是 1，则上述方程有唯一解，即 p 能唯一表示成两个平方数之和。

这个结论的第一部分可以从平方剩余问题 (4) 简单地得到。事实上，由方程 (3)，我们可以找到一个整数 z，使得 $yz-1$ 能被 p 整除。由于

$$pz^2=(xz)^2+(yz)^2=(xz)^2+(yz-1+1)^2$$
$$=(xz)^2+1+(yz-1)^2+2(yz-1),$$

故得

$$(xz)^2-p\left(z^2-2\left(\frac{yz-1}{p}\right)-p\left(\frac{yz-1}{p}\right)^2\right)=-1.$$

因此上面方程有解也就意味着

$$1=\left(\frac{-1}{p}\right)=(-1)^{\frac{p-1}{2}},$$

这就相当于 p 被 4 除余数为 1。

无论如何，要具体求出 p 的平方数之和的精确表达式并非易事。这就需要用到更加高深的工具了。

上面这些不定方程的研究为古典数论提供了许多重要的原动力，发展出了很多重要的数学工具。如何求解一般的不定方程实际上是个很困难的问题。希尔伯特在他著名的 23 个数学问题中也作了探讨。近代的一种重要思想是：先退而求其次，求方程的有理数解。对每个素数 p，将方程放到 p-adic 数域情形去求解（p-adic 数是和素数 p 有关的一类很特殊的“数”）。我们将相应的 p-adic 解综合起来，根据这些解的信息，试图构造出真正的有理数解。这一思想曾被成功地应用到二次不定方程上。它可以归结成著名的 **Hasse-Minkowski 定理**：

> 假设 $f(x, y)$ 是二次多项式（比如 x^2+y^2-1 等等），方程 $f=0$ 有有理数解的充分必要条件是：它有实数解，并且对每个素数 p，它也有 p-adic 解。

本节最后，我们再介绍一个与素数和方程有关的重要猜想—— *abc* **猜想**：

> 任意给定正实数 $\varepsilon>0$，则最多只有有限个三元组 (a, b, c)，满足以下条件：
> (1) a, b, c 皆整数，并且两两之间互素，
> (2) $a+b=c$，
> (3) $c>d^{1+\varepsilon}$，此处 d 是乘积 abc 中所有互不相同的素因子的乘积。

如果 *abc* 猜想正确，那么立刻可以推出费马猜想对充分大的方幂 n 都成立。这还不止，其实有许许多多重要的猜想和定理竟然都能从 *abc* 猜想推出来，而后者本身看上去却如此简单！

6 素数和代数

素数在整数中具有如此特殊的地位，它除了包含许多有趣深刻的性质之外，也留下诸多未解之谜。人们自然会想到将素数的概念推广到更一般的数域上来研究，比如实数域、复数域等等。高斯首先研究了这样一些复数（称为高斯整数）：

$$a+b\sqrt{-1}\ ,$$

这里 a, b 都取整数。所有这些复数构成的集合记为 $\mathbb{Z}[\sqrt{-1}]$。我们可以像在整数情形一样定义高斯整数的加减乘运算，甚至我们还可以定义两个高斯整数的带余数除法等。完全类似地，我们自然也可以定义 $\mathbb{Z}[\sqrt{-1}]$ 中的“素数”概念。

整数集合里的素数在 $\mathbb{Z}[\sqrt{-1}]$ 中不一定是素数哦。比如当 p 被 4 除余数为 1 时，根据上一节的结论，它可以写成两平方数之和

$$p=a^2+b^2.$$

因此 p 在 $\mathbb{Z}[\sqrt{-1}]$ 中可以写为两个高斯整数的乘积

$$p=(a+b\sqrt{-1})(a-b\sqrt{-1})\ .$$

高斯的一项有趣的工作，就是找出了 $\mathbb{Z}[\sqrt{-1}]$ 中的所有素数 P：

> (1) $P=1+\sqrt{-1}$；
> (2) $P=a+b\sqrt{-1}$，使得 $p=a^2+b^2$ 是整数集合里被 4 除余数为 1 的素数；
> (3) P 就是整数集合里被 4 除余数为 3 的素数。

在整数中我们不把 ±1 这样的数作为素数；同样地，在 $\mathbb{Z}[\sqrt{-1}]$ 中我们也不考虑以下诸如 ±1, $\pm\sqrt{-1}$ 这样的数——单位数。如果两个高斯整数只相差一个单位数，它们的数论性质一般没什么差别，所以我们有时只挑选其中的一个代表来讨论。接着，我们同样可以得到高斯整数的算术基本定理等等一系列的数论性质。

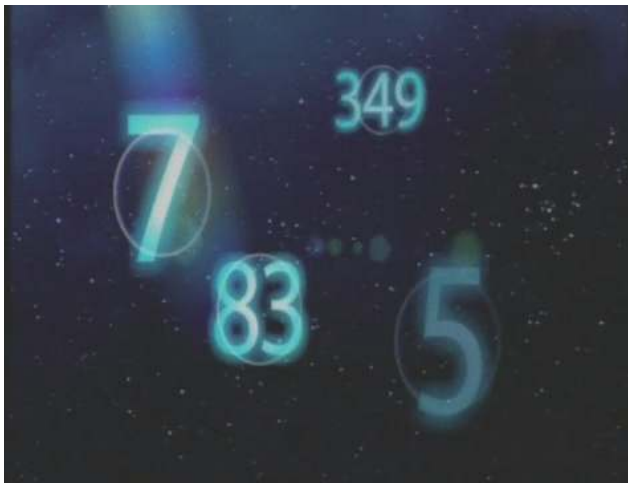
高斯对于 $\mathbb{Z}[\sqrt{-1}]$ 的研究可以说是代数数论的重要起源之一。为什么高斯要研究高斯整数呢？原来高斯一开始在研究四次剩余问题（即不定方程 $x^4-py=q$ 求解）时，没有能够找到类似二次互反律那样的有效算法来判断方程是否有解。他在研究中逐渐意识到，这一问题不能只局限于整数范围内考虑，而应该扩展到 $\mathbb{Z}[\sqrt{-1}]$ 上研究其算术性质，用高观点来探讨这一问题。这是一种富有启发性的数学思想，当人们把视野扩大后，很多问题的答案也许就会变得清晰起来。事实正是如此，很快高斯就在 $\mathbb{Z}[\sqrt{-1}]$ 中找到了四次互反律。不过他并未给出证明，而是由雅可比和艾森斯坦后来分别独立给出了证明。

进一步，如果我们考虑除法，整数集合 $\mathbb{Z}$ 可以扩张到有理数域 $\mathbb{Q}$ 上。因此类似地，高斯整数集合也可以扩张到高斯有理数域 $\mathbb{Q}[\sqrt{-1}]$ 上，里面的数无非是两个高斯整数的比值而已。很自然地，我们也可以考虑更一般的数域——二次域：

$$\mathbb{Q}[\sqrt{m}]=\left\{a+b\sqrt{m}\,\middle|\,a,b\in\mathbb{Q}\right\}\ ,$$

这里 m 是任意整数，并且我们可以假设 m 不含平方因子。$m=-1$ 就是高斯有理数域。$\mathbb{Q}[\sqrt{m}]$ 中的“整数”是什么样的呢？答案与我们想象的稍微不同：

情形一：如果 m 被 4 除余数是 2 或者 3，那么二次域中的“整数”（称作代数整数）都可以写为，



$$a+b\sqrt{m}\ ,$$

这里 a, b 是整数。

情形二：如果 m 被 4 除余数是 1，那么二次域中的“整数”可以写为

$$a+b\left(\frac{1+\sqrt{m}}{2}\right),$$

这里 a, b 是整数。

同样地，人们可以考虑这样的“整数”什么时候能称作“素数”等等基本问题，这里我们不再详细展开。那么著名的算术基本定理在这时是否一定成立呢？答案是否定的！这里我们举一个简单的例子。在 $\mathbb{Q}[\sqrt{-5}]$ 中，21 竟然有两种完全不同的素因子分解：

$$21=3\times 7=(1+2\sqrt{-5})(1-2\sqrt{-5}).$$

在历史上，人们一开始并未意识到这一问题。最初，人们之所以引进这样的数域，是为了研究著名的费马猜想，就是证明不定方程

$$X^n+Y^n=Z^n$$

当 n 是大于 2 的正整数时没有非零整数解。比如我们可以在高斯整数的意义下证明 $n=4$ 的情形没有非零解，因此在通常整数意义下就更不可能有非零解了。类似地，高斯在数域 $\mathbb{Q}[\sqrt{-3}]$ 中也巧妙地证明了 $n=3$ 的情形也没有非零解——这一证法远比欧拉的证明更简洁且更具启发性。一般情形下，人们将有理数域扩展到由 n 次单位根生成的数域上（所谓单位根就是方程 $x^n=1$ 的根）。这样，费马方程的左边在该数域下就可以分解为一次因式的乘积。如果我们事先知道这样的数域中也有算术基本定理，那就很容易推出矛盾，从而证明费马猜想。

当这一想法第一次被正式提出时，遭到了很多数学家的质疑和反对。事实上，数学家库莫此前早已意识到这一问题，即复数情形下“素因子”分解不一定唯一！为了弥补这一缺陷，库莫引入了理想数的概念，证明理想数有类似于算术基本定理那样的唯一分解性质，从而成功证明费马猜想在 $n<100$ 时成立。

其实理想数并不是真正的数，而是一组数的集合。但有趣的是，我们也可以定义这种集合之间的乘法运算，并且定义出类似素数的东西——素理想，最终证明理想数唯一分解定理——算术基本定理的推广。理想数的引入可以说是极为关键的。它使数论的研究观点和方法产生了质的飞跃，促使了代数数论的发展。继库莫的工作之后，戴德金将理想数推广到了更一般情形，从而发展成了系统的理想理论。这一理论是交换代数等学科中的核心内容之一。它不但对数论发展极为重要，而且还深入到其他各个数学领域中，特别是对代数几何等等学科有着重要的影响。

素数和函数

上一节我们从一个侧面看到素数及算术基本定理对于数学的重要影响，它们的推广促进了代数数论等领域的发展。同样地，素数对于函数的研究也有极为深刻的影响。如前所述，除了算术基本定理，素数另一重要的基本结论就是“素数个数无限”。我们曾经介绍了欧几里德关于这一结论的存在性证明。实际上，欧拉还给了另一个巧妙的证明，这个证明极富启发性，常被人们视为解析数论之发端。下面我们用不太严格的方式来介绍一下。欧拉的证明利用了下面几个简单的事实

(1)：对任何介于 0 和 1 之间的实数 x，都有无限求和公式

$$\frac{1}{1-x}=1+x+x^2+\cdots+x^n+\cdots;$$

(2)：将下式左边展开并利用算术基本定理得到

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

左边的大写希腊字母在这里表示求乘积符号，就是把每个素数 p 对应的项 $(1+\frac{1}{p}+\cdots)$ 都相乘起来。

(3)：结合上面两个式子，则有

$$\prod_p\left(1-\frac{1}{p}\right)^{-1}=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}+\cdots.$$

如果素数只有有限个，那么上式左边是个有限的数。但无论如何，上式右边的值是无穷大（数学上叫做发散），这就推出矛盾！因此素数个数必定无限。

从上面的讨论，我们可以定义一个重要的函数——黎曼 ζ 函数

$$\zeta(s)=1+\frac{1}{2^s}+\frac{1}{3^s}+\cdots+\frac{1}{n^s}+\cdots .$$

稍微推广一下前面的恒等式，就得到有趣的恒等式

$$\prod_p\left(1-\frac{1}{p^s}\right)^{-1}=\zeta(s) .$$

黎曼 ζ 函数是数学中极其重要的函数。黎曼首先研究了这种函数的诸多深刻性质，并且第一次发现了黎曼 ζ 函数居然和素数之间存在着极为深刻的内在联系！按照上述方式定义的黎曼 ζ 函数的定义域还比较小。通过一定的数学技巧，我们可以把它的定义域扩大到除了 $s=1$ 以外的整个复平面上（即对所有不等于 1 的复数 s 都有定义）。人们感兴趣方程 $\zeta(s)=0$ 的根——通常称作零点。黎曼 ζ 函数有许许多多零点，其中一部分很容易求出来，我们把它们叫做平凡零点。剩下那部分非平凡零点落在哪里呢？黎曼做出了一个重要的猜测：

> $\zeta(s)$ 的非平凡零点一定可以写成形如 $s=\frac{1}{2}+b\sqrt{-1}$ 的复数（b 是实数）。换言之，所有的非平凡零点都落在直线 $\mathrm{Re}(s)=\frac{1}{2}$ 上，这里 Re 表示取复数的实部。

这个猜想被称为**黎曼猜想**，它被列为千禧年七大数学猜想之一，也被希尔伯特收入到 23 个著名数学问题中。这是个极其困难的问题，它远比费马猜想艰深得多。我们注意到，《数学文化》刊登的卢昌海博士的系列文章对黎曼猜想作了非常生动的介绍。

为什么黎曼猜想如此重要呢？黎曼以其深刻的洞察力，发现素数的许多性质和黎曼函数的解析性质密切相关。黎曼函数也可以稍稍变化，改成更一般的狄利克雷级数

$$D(s)=f(1)+\frac{f(2)}{2^s}+\frac{f(3)}{3^s}+\cdots+\frac{f(n)}{n^s}+\cdots .$$

此外还有许多更复杂的函数。研究表明，许许多多重要的数论猜想或定理本质上都和研究这类函数的零点位置有关。比如前面说的素数定理等等。因此从这样的深刻背景来看，我们有理由预见，那些表面上看似简单的未解决之难题，即使有证明也必定是极其艰深的，绝不可能用简单的初等方法获得。

黎曼的这一杰出工作，可以说是开了解析数论之先河，给数论研究提供了强有力的研究工具和技巧。尽管我们至今无法证实黎曼猜想，但是可以退而求其次，想办法证明所有这些零点落在一条狭窄的区域里面——这个区域当然要包含整条直线 $\mathrm{Re}(s)=\frac{1}{2}$。我们要做的是将这个区域不断地缩小。如果最终能压缩成直线 $\mathrm{Re}(s)=\frac{1}{2}$，那就等于证明了黎曼猜想。一个十分有趣的现象是：区域缩得越小，那么就能得到越多的关于素数的深刻定理。

另一种迂回的方式，则是企图在函数域情形探讨黎曼猜想的一个模拟。事实上，Weil 在有限域代数曲线上建立了这样的类似猜想——Weil 猜想。Weil 本人于 1948 年证明了该情形的猜想。对有限域上高维代数簇情形，Weil 也提出了类似猜想。数学家 Deligne 利用代数几何等理论工具于 1973 年证明了它。尽管这一猜想离原始的黎曼猜想还相差很远，但这一杰出的工作已经影响深远，极大地推动了数学各领域的发展，特别是代数几何理论。

8 一些题外话

我最早是通过汤涛教授的新浪微博了解《数学文化》的，并立刻被深深吸引住了，成为其忠实粉丝。在这里，我要感谢《数学文化》各位老师给我这次难得的锻炼机会，也要感谢他们在科普传播方面的辛勤劳作，让我们能够看到数学有如此生动有趣的一面。

作者介绍：

陆俊，华东师范大学数学系讲师，代数几何方向，师从谈胜利及陈志杰教授。