谈谈素数

TANTAN SUSHU

王　元

上海教育出版社

谈 谈 素 数

王 元

上海教育出版社

内 容 提 要

素数论这一古老的数学分支，包含着许多诸如哥德巴赫问题那样的有趣而又艰深的难题。为了解决这些问题，素数论既借助也带动了其他数学分支的发展，因而素数论迄今仍是一个活跃的领域。

本书旨在介绍素数论的主要内容，书中谈到了许多著名的数论问题和猜想，简介了解决这些问题的方法和近代成果，介绍了我国数学家在这个领域里的重要贡献。本书的前一半只用到了中学的数学知识，而具备一些数学分析的知识后就可以读完后一半。全书写法简洁，深入浅出，可供中学生和广大数学爱好者阅读。

谈 谈 素 数

王 元

上海教育出版社出版

（上海永福路 123 号）

新华书店上海发行所发行　上海崇明印刷厂印刷

开本 787×1092 1/32　印张 2.5　字数 51,000

1978 年 11 月第 1 版　1978 年 11 月第 1 次印刷

印数 1—50,000

统一书号：7150·1947　　定价：0.20 元

序　言

在数学中，数论是研究数的性质，特别是研究整数性质的分支，它和几何学一样，是最古老的数学分支.

素数就是除1与其自身外，没有其他因数的大于1的自然数. 在自然数列中，最初的几个素数是

$$2, 3, 5, 7, 11, 13, 17, \cdots.$$

素数的性质是数论最早的研究课题之一，现在则已发展成为数论的一个独立分支——素数论. 素数论是数论中十分有味与引人入胜的一个分支，它里面有着许多没有解决的奇妙的猜测.

这本小册子将介绍素数论方面的一些结果，前面一部分(§1～§11)是算术部分. 在中学的数学课中，平面几何学是训练逻辑推导最好的课程. 此外，初等数论也能起到这个作用，它有助于培养分析问题和解决问题的能力. 这一部分并不涉及更多的定义与知识，所以只要耐心阅读，高中的同学是可以看得懂的. 但素数论方面的重要与深刻的结果，常常是用精深的数学方法，特别是精深的分析方法得到的. 如果不讲这一部分，就会给人以错觉，好象近代的素数论研究，只要从整数与素数的定义出发，作一些算术推导就行了. 事实当然不是这么回事，所以在§12～§23中，我们将假定读者学过微积分并了解实数的极限概念. 这一部分着重介绍近代素数论的一些问题与结果，而将证明省略了. 讲这一部分的目的是给读者增加一点数学常识. 属于近代数学的那些结论中，

能让非专业人员了解的，也许除数论以外就不多了．从这里也不难看到，虽然素数论中的许多问题表面上提法都很简单，但是近代素数论的重要成就，却往往是在近代数学成就的基础上，通过十分迂回的道路而得到的．反过来，为了解决素数论中的问题，也曾多次刺激并带动了其他不少数学分支的重要发展．因此素数论在数学中并不是孤立的，而是与很多数学分支密切相关的．由上所述，我们认为企图从整数与素数的定义出发，用简单的算术方法来处理这一类问题是不易收效的．不少事例表明这样做往往劳而无功，我们应该从中总结经验教训．总之，我们认为有兴趣于这类经典问题（例如哥德巴赫问题）的人，应该具备相当的数学知识与修养，而且应该先熟习素数论中已有的成果与方法，再作进一步的探讨，才可能会是有益的．

这本小册子取材于华罗庚老师的著作《数论导引》（科学出版社，1957 年），《指数和的估计及其在数论中的应用》（科学出版社，1963 年）及夕尔宾斯基著《关于素数——我们已知和未知的》（波兰，华沙，1961 年）．笔者仅仅作了一些整理与归纳，使读者更便于了解素数论的概貌．另外，由于上面几本著作都已出版十多年了，所以本书也引征了一些新的文献，供作参考．

在撰写的过程中，承蒙陈景润同志的热情支持与帮助，又承蒙于坤瑞、徐广善等同志帮助准备手稿，他们提出了不少宝贵的意见，我谨在此向他们致以最衷心的感谢．限于笔者的水平，错误与不妥之处，还希望读者不吝指教．

王　元

1978 年 5 月于北京

目　录

§1. 素数与复合数

自然数是指

$$1, 2, 3, \cdots$$

中的数. 整数是指

$$\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots$$

中的数. 所以自然数就是正整数.

任意给出二整数 a 与 b, 其中 $b>0$. 如果有一个整数 c 使

$$a=bc,$$

就称 b 可以整除 a, a 称做 b 的倍数, b 称做 a 的因数. 记为 $b|a$. 假若 b 不能整除 a, 就记做 $b\nmid a$. 注意, 这里因数都是正的. 记

$$|a|=\begin{cases} a, & \text{当 } a\geqslant 0, \\ -a, & \text{当 } a<0. \end{cases}$$

我们称 $|a|$ 为 a 的绝对值. 如果 $b|a$, 而且 $1<b<|a|$, 我们就称 b 是 a 的真因数.

显然, 对于任何正整数 a 都有

$$1|a,\ a|0,\ a|a,$$

这说明 a 至少有因数 1 和 a.

自然数可以分成三类:

1) 1, 只有自然数 1 为它的因数.

2) p, 正好有而且只有自然数 1 及 p 为它的因数. 换句话说, p 是大于 1 而又没有真因数的自然数.

3) n, 有两个以上大于1的因数. 换句话说, n 是有真因数的自然数.

第2)类数叫素数. 例如

$$2, 3, 5, 7, 11, 13, 17, 19, 23, \cdots.$$

我们常常用 $p, q, r, p_1, p_2, \cdots$ 等等来表示素数.

第3)类数叫复合数. 例如

$$4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20, 21, 22, \cdots.$$

我们常常用 $n, l, m, a, b, \cdots$ 等等来表示复合数.

2能整除的自然数叫做偶数, 如2, 4, 6, 8, …. 而2不能整除的自然数叫做奇数, 如1, 3, 5, 7, …. 显然大于2的偶数都是复合数. 所以只有一个偶素数2, 其余的素数都是奇素数.

§2. 唯一分解定理

引理1. 大于1的自然数 n 都可以分解成为素数的乘积.

证 如果 n 本身就是一个素数, 那末定理就已经成立了. 现在假定 n 是复合数, 那末 n 总有一个最小的真因数 q_1. 我们先证明 q_1 一定是素数. 如果 q_1 是复合数, 那末 q_1 还有真因数 r_1, 当然 $r_1<q_1$, 而且 r_1 也是 n 的真因数. 这与 q_1 是 n 的最小真因数相矛盾, 所以 q_1 是素数. 记

$$n=q_1n_1, \qquad 1<n_1<n.$$

如果 n_1 已经是素数, 那末定理即成立. 如果 n_1 不是素数, 假定 q_2 是 n_1 的最小素因数, 即得

$$n=q_1q_2n_2, \qquad 1<n_2<n_1<n.$$

我们继续实行上面这种手续，得 $n>n_1>n_2>\cdots>1$. 所以这种手续不能超过 n 次. 最后得

$$n=q_1q_2\cdots q_k,$$

其中 q_1, q_2, $\cdots$, q_k 都是素数（注意：q_1、q_2、$\cdots$、q_k 不一定是互不相同的）. 这个式子叫做 n 的素因数分解式. 引理证完.

例如： $10,725=3^1\cdot5^2\cdot11^1\cdot13^1$.

我们可以把大于 1 的自然数 n 的素因数分解式写成

$$n=p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k},$$

其中 $p_1<p_2<\cdots<p_k$ 都是素数，而 a_1, a_2, $\cdots$, a_k 都是自然数. 这个式子叫做 n 的标准分解式.

引理 2. 如果 p 是素数而且 $p|ab$，那末必定 $p|a$ 或 $p|b$.

证 不妨假定 a, b 都是自然数. 假定引理不成立，那末一定有一个最小的素数 p 使引理不成立. 对于这个素数 p，又有最小的 ab 使引理不成立，即 $p|ab$ 而 $p\nmid a$, $p\nmid b$.

我们先来证明 $a<p$, $b<p$. 假如不然，例如假定 $a>p$. 由于 $p\nmid a$，所以用 p 除 a，所得的余数 a_1 必在 0 与 p 之间，即

$$a=kp+a_1,\quad 0<a_1<p.$$

因此

$$ab=(kp+a_1)b=kbp+a_1b.$$

由 $p|ab$ 及 $p|kbp$ 得 $p|(ab-kbp)$，即 $p|a_1b$. 然而 $p\nmid a_1$, $p\nmid b$，从而有 $a_1b<ab$ 使引理不成立. 这与 ab 是使引理不成立的最小数的定义相矛盾. 所以 $a<p$，同理可知 $b<p$，因此 $ab<p^2$.

现在来证明 $p|ab$ 而 $p\nmid a$, $p\nmid b$ 将引出矛盾. 因 $p|ab$，所以 $ab=lp$. 若 $l=1$，那末 p 有真因数 a 与 b. 这与素数的定义相矛盾. 因此 $l>1$. 另一方面，上面已证 $ab<p^2$，所以 $l<p$.

由引理1的证明可知，假定q是l的最小非1的因数，那末q为素数．由于$l|ab$，所以$q|ab$．因为$q\leqslant l<p$，所以由p是最小的使引理不成立的素数这一假定，可知$q|a$或$q|b$．我们不妨假定$q|a$．记$a=a'q$．由于前设$q|l$．记$l=tq$，代入$ab=lp$得

$$a'qb=tqp.$$

因而$a'b=tp$，即$p|a'b$．但这样$a'b<ab$，$p\nmid a'$，$p\nmid b$．这与关于p与ab的假定相矛盾．引理证完．

定理1(唯一分解定理)．大于1的自然数n的标准分解式是唯一的．换句话说，如果不计次序，那末n只有唯一的方法表示成素数的乘积．

证　由引理2显然可知，如果p是素数，

$$p|ab\cdots c,$$

那末p一定能整除a，b，$\cdots$，c中的一个．又如果a，b，$\cdots$，c都是素数，那末p一定是a，b，$\cdots$，c中的一个．

假定n有两种标准分解式

$$n=p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}=q_1^{b_1}q_2^{b_2}\cdots q_l^{b_l}.$$

那末任何p_i必定为q_1，q_2，$\cdots$，q_l中的一个，任何q_j也必定为p_1，p_2，$\cdots$，p_k中的一个．所以$k=l$．由于

$$p_1<p_2<\cdots<p_k \quad 及 \quad q_1<q_2<\cdots<q_k,$$

所以

$$p_i=q_i, \qquad 1\leqslant i\leqslant k.$$

最后，如果有$a_i>b_i$，这儿$1\leqslant i\leqslant k$，那末以$p_i^{b_i}$除n的标准分解式得

$$p_1^{a_1}\cdots p_i^{a_i-b_i}\cdots p_k^{a_k}=p_1^{b_1}\cdots p_{i-1}^{b_{i-1}}p_{i+1}^{b_{i+1}}\cdots p_k^{b_k}.$$

上式的左边是p_i的倍数，但右边不是，这不可能．同样$a_i<b_i$也是不可能的，所以

$$a_i=b_i, \qquad 1\leqslant i\leqslant k.$$

定理证完.

顺便说一句，我们不把 1 看成素数，是因为如果把 1 看成素数，那末在 n 的标准分解式前面，可以乘上 1 的任何次幂，这就破坏了标准分解式的唯一性了.

虽然在理论上，任何自然数 n 都是可以写成标准分解式的. 但当 n 很大时，具体写出 n 的标准分解式来却是很不容易的事. 有时甚至连 n 的一个素因数也找不出来. 例如人们已经证明 $M_{101}=2^{101}-1$(共 31 位)是两个不同素数的乘积，其中较小的一个至少有 11 位，但我们至今还不知道这两个素因数是什么[1]. 又例如在 1958 年，人们就知道 $F_{1945}=2^{2^{1945}}+1$ 的最小素因数 $p=5\times 2^{1947}+1$. 但至今我们并不知道 F_{1945} 的其他素因数[2]. 由于

$$2^{1945}=32\times 2^{1940}=32\times(2^{10})^{194}>30\times(10^3)^{194}$$
$$=3\times 10^{588},$$

所以

$$F_{1945}>2^{3\times 10^{588}}=(2^{10})^{3\times 10^{588}}>10^{9\times 10^{588}},$$

即 F_{1945} 是一个超过 10^{532} 位的自然数，而 p 则是一个有 587 位的素数.

假定 a 与 b 是两个整数，但不都是 0. 如果 $c|a$, $c|b$, 我们就称 c 是 a 与 b 的公因数. 如果 $a\neq 0$, 那末由 $c|a$ 可得 $a=cd$, 其中 $d\neq 0$ 是整数，即 $|d|\geqslant 1$. 所以，$|a|=|cd|=c|d|\geqslant c$, 即 a 与 b 的公因数 c 不大于 a 的绝对值 $|a|$. 因此，a 与

---

[1] J. Brillhart and G. D. Johnson, On the factors of certain Mersenne numbers, *Math. Comp.*; **14** (1960), 553～555.

[2] R. M. Robinson, A report on primes and on factors of Fermat numbers, PAMS; **9** (1958), 673～681.

b 的公因数中一定有一个最大的，称为 a 与 b 的最大公因数，记为 (a, b). 例如

$$(5, 3)=1, \qquad (20, 45)=5,$$
$$(11, -242)=11, \qquad (0, -377)=377$$

等等. 如果 $(a, b)=1$，就称 a 与 b 互素.

我们用 $r=\min(m, n)$ 表示 r 等于 m 与 n 中较小的一个. 例如 $5=\min(5, 13)$. 我们又用 $s=\max(m, n)$ 表示 s 等于 m 与 n 中较大的一个. 例如 $13=\max(5, 13)$.

定理 2. 假定 a 与 b 是二正整数，把它们写做

$$a=p_1^{a_1}\cdots p_s^{a_s}, \qquad a_1\geqslant 0, \cdots, a_s\geqslant 0,$$
$$b=p_1^{b_1}\cdots p_s^{b_s}, \qquad b_1\geqslant 0, \cdots, b_s\geqslant 0,$$

其中 $p_1<\cdots<p_s$ 都是素数. 那末

$$(a, b)=p_1^{c_1}\cdots p_s^{c_s},$$

其中 $c_i=\min(a_i, b_i)\ (1\leqslant i\leqslant s)$.

证　如果 $c\mid a$, $c\mid b$，那末由引理 2 可知 c 的素因数只能是 $p_1, \cdots, p_s$，即

$$c=p_1^{d_1}\cdots p_s^{d_s}.$$

显然 $d_1\leqslant a_1$, $d_1\leqslant b_1$，所以，$d_1\leqslant\min(a_1, b_1)=c_1$. 同理 $d_i\leqslant c_i\ (2\leqslant i\leqslant s)$. 因此

$$c\leqslant p_1^{c_1}\cdots p_s^{c_s}.$$

即 a、b 的任何公因数 c 不大于 $p_1^{c_1}\cdots p_s^{c_s}$. 另一方面，$p_1^{c_1}\cdots p_s^{c_s}\mid a$, $p_1^{c_1}\cdots p_s^{c_s}\mid b$，即 $p_1^{c_1}\cdots p_s^{c_s}$ 是 a、b 的公因数. 所以 $p_1^{c_1}\cdots p_s^{c_s}$ 是 a 与 b 的最大公因数. 定理证完.

§3. 素数有无穷多

现在发生一个问题，素数究竟只有有限多个呢？还是有

无穷多？这件事早在欧几里德 (Euclid) 就已经知道了：素数有无穷多.

定理 1. 素数有无穷多.

证　如果素数的个数有限，那末我们就可以将全体素数列举如下：

$$p_1, p_2, \cdots, p_k.$$

命

$$q=p_1p_2\cdots p_k-1.$$

q 总是有素因数的. 但我们可证明任何一个 $p_i(1\leqslant i\leqslant k)$ 都除不尽 q：假若不然，由 $p_i|q$ 及 $p_i|p_1p_2\cdots p_k$ 就得到 $p_i|(p_1p_2\cdots p_k-q)$，即 $p_i|1$，这是不可能的. 故任何一个 p_i 都除不尽 q，这说明 q 有不同于 $p_1, p_2, \cdots, p_k$ 的素因数. 这与 $p_1, p_2, \cdots, p_k$ 是全体素数的假定相矛盾，所以素数有无穷多. 定理证完.

由定理 1 的证明立刻可以推出：

定理 2. 假定 $n>2$，那末在 n 与 $n!$ ($n!$表示不超过 n 的自然数的连乘积，即 $n!=1\cdot 2\cdot \cdots \cdot n$) 之间一定有一个素数.

证　假定不超过 n 的素数为 $p_1, p_2, \cdots, p_k$. 又假定 $q=p_1p_2\cdots p_k-1$. 由于 $n>2$，所以 $q>4$. 由定理 1 的证明可知 q 有一个不同于 $p_1, p_2, \cdots, p_k$ 的素因数 p，所以 $p>n$. 另一方面，$p\leqslant q\leqslant n!-1<n!$. 定理证完.

定理 1 的证明方法还可以用来证明更广泛的结果. 例如：

定理 3. 形如 $4n+3$ 的素数有无穷多.

证　如果形如 $4n+3$ 的素数有限，则可假定它们的全体是

$$p_1, p_2, \cdots, p_k.$$

命

$$q=4p_1p_2\cdots p_k-1=4(p_1p_2\cdots p_k-1)+3.$$

从而 q 是形如 $4n+3$ 的，而且任何 $p_i\ (1\leqslant i\leqslant k)$ 都除不尽 q. 由于除掉 2 以外，素数都是奇数，因此奇素数用 4 除以后，所得的余数必定是 1 或 3. 又由于两个 4 除余 1 的数 $4l+1$ 与 $4m+1$ 相乘得

$$(4l+1)(4m+1)=4(4lm+l+m)+1,$$

仍然是一个 $4n+1$ 型的数. 因 q 是 $4n+3$ 型的数，所以 q 的素因数不可能都是形如 $4n+1$ 的数，即 q 还有形如 $4n+3$ 的素因数，但又不能是 $p_1, p_2, \cdots, p_k$ 中的一个. 这与对于 $p_1, p_2, \cdots, p_k$ 的假定相矛盾. 所以形如 $4n+3$ 的素数有无穷多. 定理证完.

读者可以仿照以上证法，证明形如 $6n+5$ 的素数有无穷多. 形如 $4n+1$ 的素数也有无穷多，这将在 §8 中证明.

虽然素数的个数有无穷多，但我们并不能写出任意大的素数来. 目前所知道的最大素数都是通过特殊的方法，而且借助于电子计算机才得到的. 现在我们知道的最大素数是

$$M_{19937}=2^{19937}-1,$$

共 6002 位[1].

§4. 素 数 表

所谓素数表，就是造一张表，其中包括不超过已知自然数 N 的所有素数. 先讲一条引理.

[1] B. Tuckerman, The 24-th Mersenne Prime, PNAS USA, 1971, 2319～2320.

引理 1. 每一个复合数 n 至少有一个素因数 $\leqslant\sqrt{n}$.

证　假定 p 是 n 的最小真因数，那末由引理 2.1(即 §2，引理 1)的证明可知 p 是素数. 现在来证明 $p\leqslant\sqrt{n}$. 由于 n 是复合数，所以可以将 n 写作 $n=pn_1$. 因 p 是 n 的最小因数，所以 $n_1\geqslant p$. 如果 $p>\sqrt{n}$，就有 $n=pn_1>\sqrt{n}\cdot\sqrt{n}=n$. 矛盾. 所以 $p\leqslant\sqrt{n}$. 引理证完.

我们先找出不超过 $\sqrt{N}$ 的全部素数，依次排列如下：

$$2=p_1<p_2<\cdots<p_r\leqslant\sqrt{N}.$$

然后把大于 1，而又不超过 N 的自然数，按大小次序排列如下：

$$2,\ 3,\ \cdots,\ N.$$

在其中留下 $p_1=2$，而把 p_1 的倍数全部划掉，再留下 p_2，而把 p_2 的倍数都划掉，继续这一手续，最后，留下 p_r，而把 p_r 的倍数都划掉. 留下的就是不超过 N 的全体素数了. 这是因为由引理 1 可知，如果 $n\leqslant N$ 而又是复合数，那末 n 必定有一个素因数 $\leqslant\sqrt{N}$，所以被划掉了. 如果 n 是 $\leqslant\sqrt{N}$ 的素数，那末规定 n 留下. 如果 n 是满足 $\sqrt{N}<n\leqslant N$ 的素数，那末 n 不会是任何 $p_i(1\leqslant i\leqslant r)$ 的倍数，所以 n 也留下来了. 因此留下来的是不超过 N 的全体素数.

例如要求出不超过 50 的全体素数，因为不超过 $\sqrt{50}<8$ 的素数是 2, 3, 5, 7，所以在 2, 3, ⋯, 50 中，留下 2, 3, 5, 7，依次划去 2, 3, 5, 7 的倍数

2, 3, 4, 5, 6, 7, 8, 9, 10,
11, 12, 13, 14, 15, 16, 17, 18, 19, 20,
21, 22, 23, 24, 25, 26, 27, 28, 29, 30,
31, 32, 33, 34, 35, 36, 37, 38, 39, 40,
41, 42, 43, 44, 45, 46, 47, 48, 49, 50.

留下的数

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,$$
$$23,\ 29,\ 31,\ 37,\ 41,\ 43,\ 47,$$

就是不超过 50 的全体素数.

上面讲的就是著名的埃拉多斯染尼氏 (Eratosthenés) 筛法. 早在公元前三百年左右，埃氏就提出这一方法. 素数表都是根据这一方法略加变化而造出来的. 埃氏筛法的改进与发展，是近代解析数论的重要工具之一.

1909 年，莱茉[1] 发表了不超过 10^7 的素数表. 在表中凡 $\leqslant 10,170,600$，而又不能被 2, 3, 5, 7 整除的自然数，它的最小素因数都被列了出来. 还有居立刻 (J. F. Kulik, 1793～1863)，他曾造出不超过 10^8 的素数表，他的手稿存放于维也纳科学院内. 1951 年，居立刻(J. P. Kulik)，波来梯与波尔特[2] 曾发表了不超过 1.1×10^7 的素数表，即在莱茉氏表的基础上增加了由 10,006,741 至 10,999,997 之间的所有素数. 他们在造表过程中，用了居立刻(J. F, Kulik)的手稿.

自从有了电子计算机后，更大得多的素数表被制作出来了. 1959 年，贝克尔与格伦贝尔格[3] 制成含有不超过 $p_{6,000,000}=104,395,301$ 的全体素数 (共 6×10^6 个素数) 的微型卡片. 六十年代初，美国学者就曾宣称，他们将在电子计算机的存储系统中存放前 5×10^8 个素数.

---

[1] D. N. Lehmer, Factor table for the first ten millions, Washington, Carnegie Institute, 1909.

[2] J. P. Kulik, L, Poletti and R. J. Porter, Liste des nombres premiers du onzième million (plus précisément de 10, 006, 741 à 10, 999, 997), Amsterdam, 1951.

[3] C. L. Backer and F. L. Gruenberger, The first six million prime numbers. The RAND Corp. Santa Monica, Pub. Microcard Foun; Madison, Wisconsin, 1959.