理论数学前沿
2021 年 12 月第 3 卷第 4 期

素数分两半
——SJTU-125 全球“科学”问第一题的思考

钱　进

中南财经政法大学，武汉

摘　要 | 本文独立发现了“素数分两半”客观存在，并用数学名家数学结论尚值得商榷的案例，印证不容忽视。是否数核内细节的微小差异，使得素数如此特别？

关键词 | 素数；宿集数；数核；甲素数；戊素数

1　EUCLID 素数无穷性证明的遗漏

人们谈素数，总是从它有无穷多个谈起。加拿大 P·里本伯姆（P·Ribenboim）在《更大素数的一本小书》[10] 中是这样，最近，SJTU 和《科学》提出 125 问第一题的应征入选述评《素数——为什么它们如此令人兴奋》[11] 也是如此。

欧几里得在《几何原本》[3]（《EUCLID's ELEMENTS》）第九章 P275 中，以命题 IX.20 证明了论断：**“预先给定几个质数，那么有比它们更多的质数。”**用的是给定数的最小公倍数加 1；后来给定的是 n 之前的所有素数之

作者简介：钱进，武汉市人，中南财经政法大学计算机经济信息管理专业退休教师。E-mail：qjcd@qq.com。科学研究人员国际唯一学术标识码：ORCID 0000-0001-8201-2360。
文章引用：钱进. 素数分两半——SJTU-125 全球“科学”问第一题的思考 [J]. 理论数学前沿，2021，3（4）：55-70.
https://doi.org/10.35534/tms.0304011

积。采纳都伯纳（Dubner）1987 年建议的符号 p# 表示 p 的“素连乘（primerial）”：$p\#=p_1 \cdot p_2 \cdots p_n$，$p_i$ 是不大于 p 的所有素数，p#+1 就有不同于“素连乘”中 n 个素数的**“更多的质数”**做因子。这一改，遗漏出来了。因为 $p_1=2$，$p_2=3$，p#+1 相当于 $6 \cdot k+1$（$k= p\#/p_1 \cdot p_2$），是一个“模六余一数”——笔者发现它，并称之为“甲素数”（或 A_α 素数，以下“戊素数”同样，后文交代），比如：$p_3\#+1=6 \times 5+1=31$，$p_4\#+1=6 \times 35+1=211$，… p#+1 没包括“模六余五数”，即 $6 \cdot k-1$，笔者称之为“戊素数”（或 E_ε 素数）。因为这个证明的每一步“都没有给出新素数的任何信息”［10］，P4，p#+1 可能是素数，也可能是合数。素数就是甲素数，合数可能有戊素数因子，既不明确，也不必然。一直到两千多年之后，库默尔（Kummer）在 1878 年用 p#−1 重新证明了素数有无穷多个，才补足了戊素数无穷性的证明，比如：$p_5\#-1=6 \times 385-1=2309$，$p_6\#-1=6 \times 5005-1=30029$，…。

2　因 ASK 算法研究热尔曼素数

笔者因为工作需要，20 世纪 80 年代初，自学了“均匀设计”[4, 5]，作为计算机模拟的实验设计，“工作报告”被总结在一册专著[7]中。2002 年 3 月 3 日、3 月 18 日、4 月 1 日，《光明日报》发布**“《奇素数和图》准确评定哥德巴赫猜想的结果”等。本人看到，对照手上正将定稿的［7］P41 § 2.4.2“偶数表为宿集数和图”**，已经覆盖了罗先生的工作，即刻去信商榷，对其结论给予质疑，《光明日报》未置可否。原来他们不过是为了赶在 2002 年 8 月“第 24 届国际数学家大会”在北京召开之前，一而再、再而三地以“公告”形式做个广告而已。于是，把“了无音讯”的论文**“奇素数和定理质疑”**转投一个学术年会发表[8]，主要内容后来纳入[7] P192 § 9.1 节中。后来在“面书 Facebook”“腾讯 QQ”等社交媒体上，发布“偶数表为宿集数和图”[7] P226 图 2.4.2，不久更新为彩色“两分划素数和图”，现更名为“戊素甲素和图”，公示近十年，没有异议，只有不间断“加朋友”的邀请。

成文进展梗概如下：1994 年，和“均匀设计”发明人二度联系上，受到鼓励，酝酿把“工作报告”修改完成[7]。1995 年春，请武汉大学李国平老师提了签；次年老师病逝，遂成绝笔，潜心挖掘以不负所望。到 2000 年，获得“全国高协

组织教育发展中心”的出版资助，又修改一年半，到了 2002 年出版。同年 11 月，法国**《科学与生活》**月刊登载**“新算法解决了质数难题”**一文，知晓了 AKS 算法第一稿[1]：“十一步简单的代数运算结果，可验算出一个数字是否是素数”，恰好和［7］同年同月发表。到 2012 年，获得［1］即**《PRIMES is in P》**。在读到共 13 步“算法”之第 7 步，“判断”if（$q \geqslant 4\sqrt{r}\log n$）and（$n^{(r-1/r)} \neq 1 \pmod r$）时，就意识到，笔者在“素数宿集”内用“筛法”划去合数，只相当“加法、存取”，没有这么复杂。此算法第 2 步至 10 步，是一个重复结构：

“2、r=2;　　* 置初值
3、while（r<n）{　　* 用计算机科学术语，就是“遍历”［2，n］
……
9、r ← r+1;　　* 步长是 1
10、}”

而笔者是用从小到大的素数“跨步”，相当于 $r \leftarrow r+p_i$，（i=3，…，k，$p_k \leqslant n$）。2013 年，又看到与［1］同名的［2］，正式发表在美国**《数学年鉴》**上，步骤已经减少到 6 步。虽然达到确定的多项式时间复杂性 $O \sim (\log^{15/2} n)$，但是，引用了两个假设：费马小定理和热尔曼素数密度猜想。比如后者：“Heuristically, our algorithm does better: under a widely believed conjecture on the density of **Sophie Germain primes**（primes p such that 2 · p+1 is also prime），the algorithm takes only $O \sim (\log^6 n)$ steps.”当年看到这里，笔者立即用甲素数 6 · k+1 做“p”代入 2 · p+1 得（3 × 4 · k+2）+1 = 3 ×（4 · k+1），有 3 的因子，不可能成为热尔曼素数。戊素数中：

k=1，6 · k−1=5，2 · p+1=11；

k=2，6 · k−1=11，2 · p+1=23，得到 5 的二次链；

k=3，6 · k−1=17，2 · p+1=35，得到合数，排除 17 非热尔曼素数；从这步可知，p 为素数不是 2 · p+1 为素数（从而 p 为热尔曼素数）的充分条件。

k=4，6 · k−1=23，2 · p+1=47，得到 11 的二次链，得到 5 的三次链。这个链到此为止，因为 47 和上一步 17 一样，个位数是 7，乘 2 加 1 之后，个位变成 5，是 5 的倍数。

https://doi.org/10.35534/tms.0304011

k=5，6·k-1=29，2·p+1=59，得到个位数同样是 9 的数，所以，唯有这样的数才可能存在 p 的多次链。

k=6，6·k-1=35，2·p+1=71；从这步可知，p 为素数不是 2·p+1 为素数的必要条件，合数 35 的 2 倍加 1，也是戊素数 71。

综上所述，个位是 5 的合数、甲素数、个位是 7 的戊素数，被排斥在热尔曼素数之外。p 是素数，既不是 2·p+1 也是素数的充分条件，又不是 2·p+1 也是素数的必要条件。从个位是 3 或者 1 的戊素数可以找得某个热尔曼素数，但是，前者不可能有二次链，后者也只能到二次链。这样，唯有大约占一成的个位是 9 的戊素数，始终以个位是 9 的戊素数“单薄”延伸，本人感觉不能较快被证明，干脆“绕过它”。沿用笔者**“先算后查”**[7] P35 § 2.3.2 的思想——均匀度用“同余逆”筛选存储、二项系数设计“杨辉 – 贾宪表”[7] P51 表 3.2.3——设计两列线性表，自动生成以 a=5 和 7 为表头、d=6 为公差的“算术级数”P（d，a），平行存于两列表中，再施以“Eratosthenes 筛法”，划去合数。然后，直接用“给定数”计算表址，判断表项是否为零？如果非零，则一步到位地检测出了素数；如果为零，遇到合数，则从相应地址找出在此处被“筛选”划去的“因子”，通过“因子链表”，进行因子分解。

现在，因为修改**“N 内嵌套双子集：素宿集、数核”**[9] 的论文，需要介绍所研究问题的背景、现状和笔者的独特工作，又来回顾［1］［2］两文，发现［1］的第二页和［2］的第二页在论文第一节“引言”（Introduction）中，说了同样的话：“our algorithm does much better：under a widely believed conjecture…”。仅仅是［2］把［1］中的“much”删除了，先后都很注重热尔曼素数。查询［10］知道，长为 k 的第一类 Cunningham 链，就是热尔曼素数，最大长度 **k=14**，其链头素数 P21 是：**143 748 292 422 532 838 039**，由 T-Forbs 于 1997 年发现。请注意，个位数字 9！笔者做过计算，接着还查到，一万之内，有 190 个热尔曼素数：1 至 1 千之内 37 个，1 千至 2 千、2 千至 3 千的一千范围内，各有 23 个，此后，3 千到 4 千等等每隔 1 千的一千范围内的热尔曼素数个数，分别是：18、14、17、18、17、11、12；有递减趋势，也出现回调。

比照第一类 Cunningham 链，检验数 p，q 的素性，计算 q=2·p+1，p、q 都

是素数，也就找出了热尔曼素数 p。对于第二类 Cunningham 链，给出 p 计算 q=2 · p-1，检验数 p、q 的素性，p、q 都是素数，也就展示出这种甲素数的来源。诸如：7 × 2-1=13，19 × 2-1=37，31 × 2-1=61，（37，73），（79，157），（97，193）…仿上查询［10］P229，知道：最大长度为 k=16 的第二类 Cunningham 链，其链头素数 P19 是：**3 203 000 719 597 029 781**，同第一类 Cunningham 链一样，也是由 T · Forbs 于 1997 年发现。从这里可看出，Cunningham 和 Forbs 已经发现素数分两类：所谓“第一类链”是戊数链，所谓“第二类链”是甲数链，如果把戊数按 q=2 · p-1 计算，q=2 · （6 · k-1）-1=3 · （4 · k-1），又像找热尔曼素数用甲数按 q=2 · p+1 计算一样，得到“丙数”，有 3 的因子。所以，这两类数是互相封闭的。也许是既有 2 · p ± 1，又类似如笔者用 6 · k ± 1 验算，出现 12 · k ± 1，不能让前辈数学家们发现两类数用（2、2 × 3=6 或 2^2 × 3=12）哪一个做模来界定！？

“AKS 算法”以“热尔曼素数”做基础，遇到忒大数字，在前述两列线性表“表项”空缺、“先算后查”查不到时，可以用它检测兜底！当然，改造 AKS 算法“遍历”步长用“表格”内的素数，以提高效率——在闲暇时，再用本人的筛法补充计算，填满这个忒大数字留出的前导空白，后面再继续延伸到一定预留备用长度。可不断增长的这个“线性”表，实际上，就是笔者发现的“数核”，由大于 4 的全体素数组成，存储在两列；左列存放戊素数，右列存放甲素数。由于素数无穷，实际运用要求有不同数量的表项，形成不同规模的“数核”。具体“联姻”方法，［9］内再陈述。——本文抽取［9］基本完成的相关内容，旨在尽快回应 SJTU125 的问题，以促进、激励全球数学对此问题的探索、讨论。

3 用“Eratosthenes 筛法”求数核

80 年代的工作，为了求得“高维均匀点列”用作试验设计，需要挑选合适的素数。因为笔者是系统软件程序设计员出身，用到的理论，是“离散数学”：用“模六留余”给自然数集做“分划”——也考虑过 3 和 5 的其他倍数，不是太小分不开，就是太大不得要领。结果发现，素数只能存在于**余五**和**余一**的两个等价类（同余类）中。

https://doi.org/10.35534/tms.0304011

图 1 中，“正六边形”表示自然数集合 N，被等分成六个正三角形，笔者用“模六留余”运算分划得到余数 $\gamma=0$，1，…，5，形成 6 个同余类，分别记为 I_0，I_1，…，I_5，则因为 I_0，I_2，I_4 内是 2 的倍数，I_3 内是 3 的倍数，所以，素数只可能包含在 I_1 和 I_5 两个等价类中。参照“甲醛、乙醇、丙酮、丁烷”，用中国传统“天干”计序法，特用“甲乙丙丁…”标记 1，…，5，0 六个同余类（英文用希腊字母做序号，在此不赘）。用 $\in$ 表示元素“包含”于集合，$\notin$ 表示元素“不包含”于集合；用 a|b 表示 a 整除 b，a∤b 表示 a 不能整除 b。于是有：

【定义 1】素数宿集：如果 $i_1 \in I_1$ 甲类，$i_5 \in I_5$ 戊类，则 $i_1 \equiv 1$，$i_5 \equiv 5(\mathrm{mod}\ 6)$，$\hat{S} = I_1 \cup I_5$，即甲类和戊类自然数组成素数宿集 $\hat{S}$①。

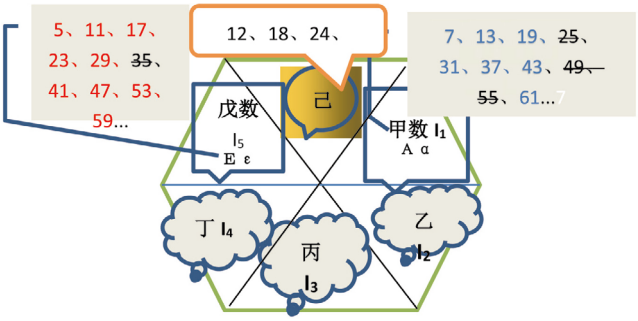

图 1　模六分划 N 示意

Figure 1　N is divided into six ports by module 6 residuals

把“挨拉托斯特尼”筛法（Eratosthenes，早期译为“厄拉多塞筛法”）施与 $\hat{S}$[7] P12（§2.1 节）后，留下大于 4 的素数集合 $\check{N}$。把从 I_1 中筛出的素数记为 $\check{N}_1$，从 I_5 中筛出的素数记为 $\check{N}_5$，可见素数分别有甲类素数和戊类素数，分别称之为：甲素数子集 $\check{N}_1$，戊素数子集 $\check{N}_5$。以下行文，有时也简称为**“甲素”“戊素”**。

【定义 2】数核：自然数集中，和小于自身的所有数互素而大于 4 的全体，组成数核 $\check{N}$。

① 刘启新先生当年，亦独立用算术级数提出素数所在两子集：P（d，a）=a+kd，d=6，a=1 或 5.［7］p41 和 p196 有记载。

$\check{N}_1$ 中素数相互乘积组成甲素数的闭集，就是 I_1。$\check{N}_1$ 比 $\check{\mathbf{N}}$ 小，为何不把它定义为数核呢？因为还有另一半戊素数，也具有“和小于自身的所有数互素”的特性，就是 $\check{N}_5$，虽然其中元素对乘法不封闭，但是，$\check{\mathbf{N}}=\check{N}_1 \cup \check{N}_5$ 可以产生乘法的闭集。这里，素数宿集 $\hat{S}$ 和 “$\check{\mathbf{N}}$ 内元素相互乘积”形成的闭包，仅差“1”这个单元数。存 1，抑或排斥 1，各有利弊：保留 1，除了对于乘法封闭外，还有了单位 1，便于构造“代数系统”的研究；排除 1，有利于从母集 $\hat{S}$ 中，通过“剥皮搡肉”求出数核 $\check{\mathbf{N}}$。通过“定义域筛法”，作“模六留余”剥去 2 和 3 的倍数之“皮”。这与原子核由质子和中子组成类似，虽然电子不在核内，却影响核的中性或电性。数核也有和原子核相似的两种“对立统一”特性，除了“同余类”的区别，还有“孪生”或“独生”区别。是故，“数核元素的相互乘积”组成的闭集，也记为 $\hat{S}$，在本文中和“素数宿集”一致。以下行文“素数宿集”常简称为**“素宿集”**，素宿集内的自然数常简称为**“宿集数”**。

【定理 1】大于 4 的所有素数即数核，包含在素数宿集 $\hat{S}$ 中.

证明： if $\forall 4 < p \in P$，then $2 \nmid p \cap 3 \nmid p$，即 $p \notin I_0 \cup I_2 \cup I_3 \cup I_4$，相当于 $p \in N-I_0 \cup I_2 \cup I_3 \cup I_4=I_1 \cup I_5 = \hat{S}$；故而此素数是宿集数，即数核是素宿集的真子集：$\check{\mathbf{N}} \subset \hat{\mathbf{S}}$，全部数核素数可从素宿集筛得。

［7］P38 指出了这两个等价类中数的运算关系：

性质 1： 如果 $i, j \in I_1$，则 $i \times j \in I_1$，$i+j \in I_2$；

性质 2： 如果 $k, l \in I_5$，则 $k \times l \in I_1$，$k+l \in I_4$；

性质 3： 如果 $i \in I_1$，$k \in I_5$，则 $i \times k \in I_5$，$i+k \in I_0$。

性质 4： 如果把 I_5 和 I_1（I_1 从 7 开始）中的宿集数，依从小到大并行排成两列，则序号 k 或 l 与列表中的数 m 或 n，有如下对应关系：

ⅰ　$m = 6 \cdot k-1$，显然 $m \in I_5$；　$k =（m+1）/ 6$　（1）

ⅱ　$n = 6 \cdot l + 1$，显然 $n \in I_1$；　$l=（n - 1）/ 6$　（2）

$\because k, l=1, 2, 3, \cdots \therefore m, n=5, 7; 11, 13; 17, 19; \cdots, \cdots$

【定理 2】宿集数同类相乘落于甲类 I_1，异类相乘落于戊类 I_5.

此乃［7］P38［引理 1］用现在术语同义表述，由以上**性质 1、2** 和**性质 3** 归纳。

【注】数据结构设计，把 I_5 和 I_1 中的宿集数并行排成两列，中间插一空列，

https://doi.org/10.35534/tms.0304011

作为孪生素数“标志”。

显然，$\check{N}=P-\{2，3\}$，P 表示素数集合。2 和 3 虽然被放到数核外，却完全离不开：$1+2+3=2\times3$，是第一个完全数；是能够把 N 内合数同素数完整分开的唯一“模”数，也是最能体现两半素数子集合不同特性的唯一“模”数。

依笔者“算法设计”对应的数据结构，利用 对应关系（1）或（2）由序号生成列表中所需规模的数，$\check{N}_5$ 分列在左列 I_5 内，$\check{N}_1$ 分列在右列 I_1 内。从小到大逐次取素数，把“挨拉托斯特尼”筛法施与 $\hat{S}$，利用**定理 2**，从 I_5 到 I_1 或从 I_1 到 I_5 相互跳转，将 I_5、I_1 内合数请零，保留素数 $\check{N}_5$ 和 $\check{N}_1$ 分别在 I_5 和 I_1 中的相对位置，以便于利用序号与列表中数的对应关系（1）或（2）定位“检索”，“开销”极小！这就是笔者判素性算法优于 AKS 算法的内在原因。算法形式的表述、正确性证明以及算法复杂性分析，请参看［9］。

宿集数前 100 行，绝大多数合数是 2- 殆素数，少数 3- 殆素数。P_2 被较小因子“筛选”后，留下大因子将产生“数核”数据。先了解概念：

【定义 3】[10] **P212，k- 殆素数：不超过 k 个素数的乘积。所有 k- 殆素数组成的集合，表示成 P_k。**

【例 1】前 100 行左列前 7 次“筛选”示例。

比如，用因子 5 筛选 I_1 中合数，将出现：① $4、5\times5=25$，② $9、5\times11=55$，③ $14、5\times17=85$，④ $19、5\times23=115$，⑤ $24、5\times29=145$，⑥ 29、0（~~$5\times5\times7$~~ $\in P_3$），⑦ $34、5\times41=205$…弃掉 5 剩下的 5、11、17、23、29、0（或者 ~~5×7~~（$\in P_2$））、41…就是 $\check{N}_5$ 前六个数字（实际在“线性表内”第 29 行，冲零的 ~~$35=5\times7$~~（$\in P_2$）将在第 6 行再次遇到，如（1）6 步所示。

另外，用因子 7 筛选 I_5 中合数，将出现：（1）6、~~$5\times7=35$~~，（2）13、$11\times7=77$，（3）20、$17\times7=119$，（4）27、$23\times7=161$，（5）34、$29\times7=203$，（6）41、0（~~$245=5\times7\times7$~~ $\in P_3$），（7）$48、41\times7=287$…弃掉 7 剩下的 5、11、17、23、29、0（~~5×7~~ $\in P_2$）、41…就是 $\check{N}_5$ 前六个数字。

同理，ⅰ 9、$55=5\times11$，ⅱ 20、$121=11\times17$，ⅲ 31、$187=11\times17$，ⅳ 42、$253=11\times23$，ⅴ 53、$319=11\times29$，ⅵ 64、0（$385=$~~$5\times7\times11$~~ $\in P_3$）、ⅶ 75、$451=11\times41$… 弃 11，剩下 $\check{N}_5$ 前六个数字；接下来，㈠ $11、5\times13=65$，

㈡ 24、13 × 11=143， ㈢ 37、13 × 17=221， ㈣ 50、13 × 23=299， ㈤ 63、13 × 29=377，㈥ 76、0（~~455=5 × 7 × 13~~ ∈ P_3），㈦ 89、13 × 41=533…弃 13，也就剩下 $\check{N}_5$ 前六个数字。（6）41、㈥ 64、㈥ 76 这 3 步遇到因子~~35~~，已经在（1）6 步事先清零。

注意 P_k 经过一次筛选后，留下 P_{k-1}，是故，在有限规模内，k-1 次筛选的结果，唯剩“数核”数字。是故，将所发现的数学对象 $\check{N}=\check{N}_1 \cup \check{N}_5$ 称“数核”，乃“实至名归”！

下文用到两个概念，在此特复述［10］§4.3 节 P192、P196 的定义。

【定义 4】[10] **孪生素数：若 p 和 p+2（=q）均为素数，则（p，q）称为孪生素数。**

【定义 5】[10] **k 阶孪生素数束：如果 2k 个相邻素数形成 k 个孪生素数对，则称之。**

由于 2 和 3 是相邻素数，3，5 和 5，7 是相邻孪生素数，容易使人产生错觉。这也表明去掉 2、3 及其倍数之“皮”以凸显数核中素数独立而唯一特性的重要。为了佐证，下面举例。

【例 2】模糊认识，影响作者传达信息的确切性。

［10］引用“N.B.Backhouse 告诉我 1 ~ 7 阶的最小孪生素数束，它们的起始素数分别为 3，5，5，9419，…”［10］P196：“In 1996，N.B. Backhouse communicated to me the smallest clusters of twin primes of order 1 to 7. Their initial primes are 3，5，5，9419，909 287，325 267 931，and 678 771 479，respectively.”

本来，Backhouse 有心将邻接一起的（3，5）（5，7）（11，13）（17，19）前 4 个孪生素数分拆开，不把它们的起始素数写成“分别为 3，3，3，3，…”是为了散开，所以，把（3，5）（5，7）拆开，却又不把（3 后第一个 5）（5，7）（11，13）和（3、5 后第 2 个 5）（5，7）（11，13）（17，19）拆开，角度就不一致了。如果因**数字 5 的重复**，在用过（3，5）之后，既不再用（3，5），也不用有 5 重复的（5，7），角度就一致了。于是，自然将（3，5）作 1 阶，将（11，13）（17，19）作 2 阶。3 阶的最小孪生素数束，是（179，181）（191，193）（197，199），比较**“5”**起始的 3 阶，179 和 191 的数字比 5、11 和 17 拉开

https://doi.org/10.35534/tms.0304011

了，更好表达了“孪生素数”相邻的含义，再观察 4 阶最小：（9419，9421）（9431，9433）（9437，9439）（9461，9463），“它们的起始素数分别为 3，11，179，9419，…”；1 位，2 位，…，4 位数，分布就比较均衡了。

本节简述了数核两种“对立统一”特性，如果将“同余类”的区分、和“孪生”或“独生”区分一起检视，会发现一些素数产生了新性质。本节提及的一些结论，只不过“冰山一角”，希望读者们共同深入探讨。下节，再谈谈最小距离引起的误判…

4　数核内素数的最小距离和频率

例 2 内，比较“5”起始的 3 阶孪生素数束，“179”起始的 3 阶 179 和 191 的数字比 5、11 和 17“为什么说”拉开了，更好表达了“孪生素数”相邻的含义？因为（179，181）（191，193）两对素数距离 d=12，**“5”**起始的三对素数距离都是 d=6，“孪生素数”相邻的含义，不是说“孪生素数”之间，不可以存在“合数”。“相邻”不是紧邻！

［10］§ 4.6 节 P211“哥德巴赫猜想：每个偶数 $2n \geqslant 4$ 均是两个素数之和”。比如 $2 \cdot 57_E=23_P+91_q$，就是“哥德巴赫猜想”的一对解。从猜想命题的“条件” $2n \geqslant 4$，备案“6=3+3”“8=3+5”之后，可以修改为“在数核内”对偶数 $2n \geqslant 10$ 求解。［10］中将用已数涵盖乙数和丁数，即缩小定义域而不影响求解结论的正确。

提到“哥德巴赫猜想”，人们自然联想到**陈景润**，因为“目前最好的结果是他给出的（1966 年宣布，细节 1973、1978 年发表）：**每个充分大的偶整数 2n 都可表成 2n=p+m，其中 p 为素数而 m ∈ P_2.**”同时，他证明了一个“伴随”结果：“存在无穷多个素数 P，使得 p+2 ∈ P_2.”这十分接近“孪生素数无穷”的猜想，打通了两个问题间的联系。恰好，笔者仅考察 I_0 中已数，自然地把求“哥猜”解与“孪生素数”联系起来了。

【例 3】“陈氏定理”默认了 2- **殆素数** m 大于素数 p。

［10］P197：“The best result to date，with sieve methods，is due to Chen（announced in 1966，published in 1973，1978）；he proved that 2 may be written in infifinitely many ways in the form 2 =m−p，with p prime，and m a product of at most two primes（not necessarily distinct）.”这里 m−p 默认了 2- 殆素数 m 大于素数 p，如 55>53，85>83，91>89，

不妥！正好反映了“孪生素数模六次序非“唯一”的错觉。客观上，这里的 p，只能局限于戊素数。所以，严格表述应该是“他证明了 2 可以有无穷多种方法表示成 2= | m−p |，…”这样，p 也就可取甲素数了，如 37>35，79>77，97>95。总之，陈氏证明“1+2”的结论，“1”既可以是“戊素”，也可以是“甲素”。

为了突出数核中“素数间隙”的特点，特用“定理”形式把切合客观的认识固定下来。

【定理 3】数核区域内，所有同类素数最小距离是 6；异类素数最小距离是 4 或 2。甲素数在前时，后面戊素数的最小距离一定为 4.

严格来说，素宿集中，宿集数也有这特点。这是设计算法操纵“数据结构”时的一个结果，不用证明。比如上述甲数在前时，模六余一的结果，后一戊数必须 +4，才能得到模六余五：1+4=5. 下面，把孪生素数（p，q）中的 p 叫“小双”，q=p+2 叫“大双”。

【定理 4】数核区域内，孪生素数的“小双”，必然是戊素数，“大双”必然是甲素数 .

证明：孪生素数是距离为 2 的素数对，根据定理 3，最小距离为 2 只可能发生在异类素数之间，如果孪生素数对（p，q）中，$p \in I_1$，则 $q \in I_5$，q−p=4，和已知“孪生素数”条件相悖，是故，$p \in I_5$，所以“异类”：$q \in I_1$，定理 4 得证。

作为定理 4 的一个令人兴奋的结果，家住加拿大的约翰·霍普金斯大学二年级学生 Felix Yu 读到这里，联系［10］P3 起始的内容，有关欧几里得、库默尔的证明，他从《离散数学》已经学到这两种关于“素数无穷性”证明。这时，他从前两个素数积为 6，联想到戊数和甲数；距离又为 2，“其中隐藏有孪生素数”？ Y 推断有可能，抽取前 n 个素连乘 $p\#=p_1p_2\cdots p_n$，仿［1］P3–4“Euclid’s 和 **Kummer’s proof**，来证明孪生素数的无穷大。

【猜想 1】当 n 足够大时，p# ± 1 会同时出现素数。

遗憾的是，素数并不眷顾“素连乘”前后元，［10］P5“记录”说，在 12 万之内，欧几里得用过的 p# 后元仅 19 个，库默尔用过的 p# 前元仅 18 个，而且，只有当 p=3，5，11 三个值时，前后元才同时为素数。即：$P_3\#=2\times3\times5=30$，$P_3\# \pm 1$ 产生孪生素数（29，31）；$P_5\#=30\times7\times11=2310$，$P_5\# \pm 1$ 有孪生素数（2309，

https://doi.org/10.35534/tms.0304011

2311）；$P_{11}\#=2310\times13\times17\times19\times23\times29\times31=25\ 878\ 772\ 920$，$P_{11}\#\pm1$ 有孪生素数（25 878 772 919，25 878 772 921）. 虽然如此，素数奔向无穷时，往往有意想不到的变化趋势，就像热尔曼素数给笔者的印象一样。鉴于从“素数间隙”证明比较复杂、不直观，从 p# ± 1 同时是素数入手，简易明了，仍然值得期待。

素数间隙仅仅是局部，素数在一定范围内有多少，即素数出现的频率，关系素数分布的全局。下面，图示“**π（X）**”，让人们对千内或百内**素数、戊素、甲素以及孪生几个素数的变化曲线**有个具体印象。图 2 内，已经列出［10］推荐的一种证明简单、表示清晰的素数计算函数，［10］P133 证明了 **π（m）**这个公式，核心是 F（i）：下面，笔者对它设计了 3 个函数因子，一个对原计数函数 π 提高效率，2 个扩大到对戊素数 π_5 和甲素数 π_1 的计数。

$$\pi(m)=\sum_{i=2}^{m}F(i)=\sum_{i=2}^{m}\left[\frac{(i-1)!+1}{i}-\left[\frac{(i-1)!}{i}\right]\right] \tag{3}$$

【定理 5】[10] **P17 Wilson 定理，若 p 为素数，则（p−1）! ≡ −1（mod p）.**

$$F(i)=\left[\frac{(i-1)!+1}{i}-\left[\frac{(i-1)!}{i}\right]\right] \tag{4}$$

π（m）公式（3）用到核心式（4），式中，若 i 为素数 p，根据定理 5 有（i−1）!+1 ≡ 0（mod i），即 i|（（i−1）!+1），假设整除商是 k，可写成（i−1）!+1=i•k，或者 $\frac{(i-1)!+1}{i}=k$. 于是，$\left[\frac{(i-1)!}{i}\right]=\left[\frac{(i-1)!+1}{i}-\frac{1}{i}\right]=\left[k-\frac{1}{i}\right]=k-1$，F（i）$=[k-(k-1)]=1$.

若 i 不是素数 p，有 i=a•b，（i−1）! 中有 a 和 b 的乘子；当 a=b 时，有 a 和 2a 同时包含在（i−1）! 内，所以，$\frac{(i-1)!}{ab}$ 是一个整数，记为 k，于是，$\left[\frac{(i-1)!}{i}\right]=k$，F（i）$=\left[\left(k+\frac{1}{i}\right)-k\right]=\left[\frac{1}{i}\right]=0$. 那么求和时，$\sum_{i=2}^{m}F(i)$，i 是素数就加 1，其它忽略，求和结果，就是 m 内的素数个数了。

继续谈谈改进。图 2 左上方文本框内，提出的第一个函数因子 JL（i），L 表示勒让德符号 $\left(\frac{a}{p}\right)$，$p\xi P$；J 表示雅可比符号 $\left(\frac{a}{i}\right)$，$i\notin P$，都和二次剩余有关。

【定义 6】[10] **P37 § 2.2H 二次剩余：若奇素数 p ł a，如果存在整数 b 使得 $a\equiv b^2$（mod p），a 叫做模 p 的二次剩余，否则叫 a 是模 p 的非二次剩余。**

【定义 7】[10] **勒让德（Legendre）符号** $(\frac{a}{p})=\begin{cases}+1, & a\text{ 为模 p 的二次剩余}\\ 0, & \text{当 p|a 时}\\ -1, & \text{否则}\end{cases}$

【定义 8】[10] **雅可比（Jacobi）符号** $(\frac{a}{b})=\prod_{p|b}(\frac{a}{p})$，p 多则取幂。比如 $(\frac{i}{36})=(\frac{i}{3})^2\cdot(\frac{i}{2})^2$。

选择改进因子：$JL(i)=(\frac{i}{36})=\left((\frac{i}{3})\cdot(\frac{i}{2})\right)^2$ （5）

当 i=1 时，$(\frac{1}{3})$ 和 $(\frac{1}{2})$ 同为 +1，JL（1）=1；当 i=2、3、4 和 6 时，有 $(\frac{i}{p})=0$，JL（i）=0；

当 i=5 时，$(\frac{5}{3})$ 同 $(\frac{2}{3})=-1$、$(\frac{5}{2})$ 同 $(\frac{1}{2})=1$，$(\frac{5}{36})=\left((\frac{5}{3})\cdot(\frac{5}{2})\right)^2=(-1)^2$，JL（5）=1。

勒让德（Legendre）符号有一个重要性质，如果存在整数 a，b 有 $a\equiv b(\mathrm{mod}\ p)$，则 $(\frac{a}{p})=(\frac{b}{p})$. 所以，以上相当于“模 6 取余”一个周期的求值，其它自然数“凡是同余类则 JL（i）值相同”！综上所述，JL（i）函数，仅在”模 6 余一”和”模 6 余五”时取值 1，其它全为 0。

$$\begin{aligned}\pi(m)&=\sum_{i=2}^{m}JL(i)\cdot F(i)\\&=\sum_{i=2}^{m}(\frac{i}{36})\cdot\left[\frac{(i-1)!+1}{i}-\left[\frac{(i-1)!}{i}\right]\right]\end{aligned}\tag{6}$$

所以，公式（6）贯彻了**定理 1** 的结论：在素数宿集中筛选所有素数。i=2 ~ m 遍历时，只有当遇到甲数和戊数时，再计算 $F(i)$ 内的阶乘，如果确实遇到素数，**π（m）**会自动增 1…下面：

$$\begin{aligned}\pi_1(m)&=\sum_{i=2}^{m}M_1(i)\cdot F(i)\\&=\sum_{i=2}^{m}\left[cos\,\pi\left(((i\,mod\,6)-1)/6\right)\right]\cdot\left[\frac{(i-1)!+1}{i}-\left[\frac{(i-1)!}{i}\right]\right]\end{aligned}\tag{7}$$

$$\begin{aligned}\pi_5(m)&=\sum_{i=2}^{m}M_5(i)\cdot F(i)\\&=\sum_{i=2}^{m}\left[cos\,\pi\left((5-(i\,mod\,6))/6\right)\right]\cdot\left[\frac{(i-1)!+1}{i}-\left[\frac{(i-1)!}{i}\right]\right]\end{aligned}\tag{8}$$

https://doi.org/10.35534/tms.0304011

因为余弦函数 cos（x）仅在 x=0 时，取 +1；其它取值，除了 x=−1・π，cos（x）虽亦为 1，但在被 6 除的情况下仍然小于 1，全被“取整函数”化为 0. 显然（7）式仅在 $i \equiv 1 \pmod 6$、（8）式仅在 $i \equiv 5 \pmod 6$ 时，必须用后一函数因子 F（i）判断是否是素数，实际是分别在统计甲素数或戊素数。到此处，也许有读者会质问，既然有 M_1（i）和 M_5（i），把它们的和 M_1（i）+M_5（i）代替 JL（i）做改进素数计算函数的因子，理由不是很简单吗？恰巧正是 JL（i）能够比 M_1（i）和 M_5（i）从根本上揭示其内在联系——体现了**定理 2** 反映的本质：**同类相乘落于甲类**，故而，$\pm 1 \pmod 6$ 同时让雅可比符号（$\frac{i}{36}$）得 1，一致性地把这 2 类和其它分离出来了。

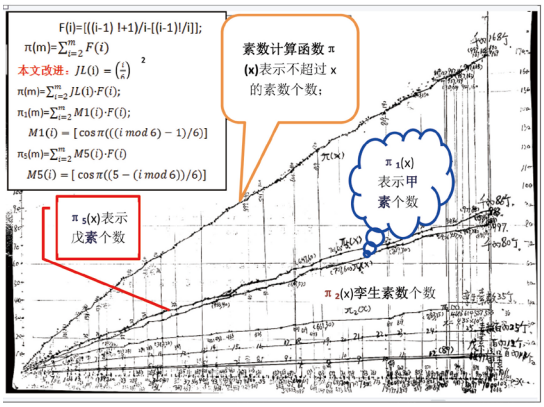

图 2　素数计算函数及其千内变化趋势

Figure 2　Prime number calculation function and its variation trend in thousand points

图 2 中，戊素数（百内 12、千内 86 个），比较甲素数（百内 11、千内 80 个），稍微多一点，甲素数线 π_1（m），基本上在整体素数线 π（m）（百内 25、千内 168 个）二分之一位置、从下方“贴近”戊素数线 π_5（m）移动，时而张开，时而并拢。定理 2 启发笔者用 $\binom{n}{1}$（$\binom{n}{1}$）−2（$\binom{n}{2}$）=n^2−n（n−1）=n 做分子，估计 π_5（m）−π_1（m）的微小差别. 当然，正如［10］在 P21 所说：“但是，Wilson 这种对素数的刻画方式对于判定 N 的素性没有实际价值，因为目前没有好的算法来快速计算 N!”，（6）−（8）式也只有理论上的意义。［10］建议要真求 π（m）的值，不如拿一张“素数表”数数。笔者在［9］中正要建议储存“一定规模的数核”供给备查。

上面，素连乘（primerial）P# ± 1 和阶乘 n! ± 1；再联系［9］归纳的 $m=2^{2k}+1$ 和 $n=2^{2k-1}-1$ 一组，$m=k\cdot 10^n+1$（k mod 3 ≡ 1），或者 $m=k\cdot 10^n-1$（k mod 3 ≡ 2）两种形式，都对称地出现戊数和甲数。追朔到检验数的特性，如果是素数，它是素数内的“细小划分”；如果是合数，它是素数的乘积。总之，是通过两类素数里面一定条件的运算或操作，找出某种特性之数；都归结为数核是此数的来源。更不用重复数域的扩充［6］P16（**I.3.1 主要的数系**）。还有好多可以用数核讨论的数，比如 2 的偶数幂次的**费尔马数 $F_n=2^{2^n}+1$**，就是戊数，**梅森数**也是 2 的幂；还有和梅森数一样源于一元二次方程参数的**斐波那契数**及**卢卡斯数列**等等，鳞次栉比，不一而足。好在来日方长，不必毕其功于一役。且看数核→素数→自然数→整数→有理数→实数→复数，式（5）中，$\frac{(i-1)!+1}{i}-\left[\frac{(i-1)!}{i}\right]$，就有有理数；$\pi=3.1415926535$ 就是实数，取小数点后 5 位扩大十万倍，314159 就是戊数；黄金分割系数 0.618033989，是方程在复数域的根 $a\pm b(\sqrt{5}-1)$ 取实部或虚部得到，扩大千倍除六取整，得到甲数 103…不管有没有用，尽情让思想飞翔…总之，数核是“万数之源”。按照普林斯顿研究院的精神，本文自我发现的抛砖引玉应“适可而止”了。

参考文献

［1］Agrawal M，Kayal N，Saxena N. PRIMES is in P［J/OL］. http://www.cse.iitk.ac.in/news/primality.html, 2002.

［2］Agrawal M，Kayal N，Saxena N. PRIMES is in P［J］. Ann.of Math，2004，160（2）：782-793.

［3］EUCLID，EUCLID’s ELEMENTS vol 13［M］. Translated by Yan Xiaodong，Nanjing：Jiangsu People’s Publishing House，2011.

［4］Fang K T. Uniform design［J］. Journal of Applied Mathematics，1980（4）：363-372.

［5］方开泰，王元. 数论方法在统计中的应用［M］. 北京：科学出版社，1996.

［6］Timothy Gowers. 普林斯顿数学指南［M］. 齐民友，译. 北京：科学出版社，2018.

https://doi.org/10.35534/tms.0304011

[7] Qian J. Uniform design principle and automatic realization [M]. H.K: China Science and Culture Press, 2002: 234.

[8] Qian J. “Sum of odd prime numbers and theorem” questioned [C] // Zhang F•X, National Computer New Technology and Computer Continuing Education Paper Collection, Part 1, 2002. Chengdu: Southwest Jiaotong University Press, 2002 (7): 1-7.

[9] Qian J, “N Nested Double Subsets: Primitive Sets, Nuclei” (to be published) Temporarily search the author's QQ colorful mathematics “1+1” space http://qzone.qq.com/115 3740 959 .

[10] Ribenboim P. The Little Book of Bigger Primes [M]. Second edition, N.Y: Springer Inc, 2004.

[11] Shalit E · d. Prime Numbers-Why are They So Exciting? [J/OL]. Young Reviewers, Israel Arts and Science Academy. https://kids.frontiersin.org/articles/10.3389/frym.2018.00040.

Prime Numbers are Divided Into Two Parts —Thoughts on the First Question of SJTU-125 Global “Science” Question

JIN Chien
ORCID 0000-0001-8201-2360

Zhongnan University of Economics and Law, Wuhan, China

Abstract: This article independently discovered the objective existence of “Prime number set is divided into two halves”, and used famous mathematicians mathematics conclusions worthy of discussion, it is impossible to neglect. Whether the miniatur differences from half in the nucleus of number, make the prime so special ?

Key words: Prime number; Number of hostel; Numerucleus