◎丁大千 著

LUN SHUANGSHENG DUIOU SANJI SUSHUZU FENBU

论双生、对偶、三奇素数组分布

——关于《哥德巴赫猜想》、《双生素数猜想》的初等证明

合肥工业大学出版社

论双生、对偶、三奇素数组分布

——关于《哥德巴赫猜想》、《双生素数猜想》的初等证明

丁大千 著

主要结论

$m > Q_L \geqslant \sqrt{m}/2$，当 $m \geqslant 25$ 时；

$m > Q_2 \geqslant \sqrt{m}/4$，当 $m \geqslant 50$ 时；

$m^2 > Q_3 \geqslant m/13$，当 $m \geqslant 60$ 时；

当 m 很大、充分大时有　$Q_L \gg \sqrt{m}/2$；　$Q_2 \gg \sqrt{m}/4$；

$Q_3 \gg m/13$；

在极限情形下有　$\lim\limits_{m\to\infty} Q_L = \infty$，　$Q_2 = \infty$，　$Q_3 = \infty$.

内 容 提 要

本书在研究分析素数分布规律基础上，运用一般数学（非高等数学），用埃拉多斯染尼氏（Eratosthenés）筛法，即古典筛法，考虑起筛数，找到了 双生、对偶、三奇素数组分布计算模式，获得了各精确解、简化解及重要结论，由 $R_{\mathrm{Ln}} \geqslant 1$，$R_{\mathrm{pnk}} \geqslant 1$，$R_{\mathrm{Ln1}} \geqslant 1$，有 $Q_{\mathrm{L}} = R_{\mathrm{Ln}} \times \sqrt{m}/2 \geqslant \sqrt{m}/2$，$Q_2 \geqslant Q_{\mathrm{L}}/2 = R_{\mathrm{Ln}} \times \sqrt{m}/4 \geqslant \sqrt{m}/4$，$Q_3 \approx R_{\mathrm{pnk}} \times R_{\mathrm{Ln1}} \times m/13 \geqslant m/13$；故有 $\lim\limits_{m\to\infty} Q_{\mathrm{L}} = \infty$，$Q_2 = \infty$，$Q_3 = \infty$；对于这些素数组的实际分布情形，都用电子计算机作了精心的检查。对于主要计算理论公式也都上机计算，获得了初步的检查验证。可以总结说哥德巴赫猜想、双生素数猜想的结论都是肯定的。书中所列的各素数组分布检查资料，对今后数学界的研究工作，可能有一定的查核参考价值。

图书在版编目(CIP)数据

论双生、对偶、三奇素数组分布：关于《哥德巴赫猜想》、《双生素数猜想》的初等证明 / 丁大千著. — 合肥：合肥工业大学出版社，2005.5

ISBN 7-81093-213-6

Ⅰ.论… Ⅱ.丁… Ⅲ.素数—研究 Ⅳ.O156.2

中国版本图书馆 CIP 数据核字(2005)第 040125 号

论双生、对偶、三奇素数组分布

——关于《哥德巴赫猜想》、《双生素数猜想》的初等证明

丁大千　著　　　　责任编辑　权　怡

出　版	合肥工业大学出版社	版　次	2005 年 5 月第 1 版
地　址	合肥市屯溪路 193 号	印　次	2005 年 5 月第 1 次印刷
邮　编	230009	开　本	787×1092　1/16
电　话	总编室：0551-2903038	印　张	7.625　**插页**　15
	发行部：0551-2903198	字　数	240 千字
网　址	www.hfutpress.com.cn	印　刷	中国科学技术大学印刷厂
E-mail	press@hfutpress.com.cn	发　行	全国新华书店

ISBN 7-81093-213-6/O·20　　　定价：23.00 元

序　一

余友丁大千热爱数学。虽技术上造诣较深，于二十世纪五十年代即因工程设计创新荣获全国先进生产者称号，但业余乐与数字游戏(play with number)。从1979年起，即将几乎全部业余时间投入数论，同时迅速掌握了电子计算技术(这在我国工程技术界，可以说是较早一批的)。藉助电力系统有关单位的计算机，作了许多探查计算。八十年代末退休后，更将全部精力投入其中，并自费购置微机多台，日夜探索。主攻方向主要在素数和素数组分布方面。二十五年多辛劳的成果，大致可按两个阶段划分：

一、二十世纪八十年代至九十年代

从应用埃拉多斯染尼氏筛法，建立起明晰的对偶素数组分布等概念。以此为基础，摸索解析公式、计算模型，并上机计算，于1984年～1986年，先后写出"双生素数猜想、哥德巴赫猜想浅证""素数、双生、对偶、三奇素数组分布的精确解、简化解和检查记录择要"。

二、二十世纪九十年代至今

为检验计算理论，编制了素数、素数组分布的各精确解、检查值等的电算程序，并利用微机，经过十多万个有效运算机时的计算，得出：

分别按精确解、简化解和素数定理进行计算与检查对照；绘制出素数、双生、对偶、三奇素数组的分布简图；素数及各素数组分布值一览表；关于素数分布、双生素数组分布、对偶素数组分布和三奇素数组分布无限的证明。

应当指出的是，十数万个有效运算机时，积累了大量的有益资料。

我认为，本书思虑周密，论证严谨，为数论领域添了新砖，加了新瓦，必能为我国数学界服务，作出应有的贡献。是为序。

郭协万
2005年2月

序　二

二十世纪七十年代后期，国内政治形势发生了重大变化，对科技的重视被提到议事日程上来，特别是1978年4月全国科学大会的召开，号召全民向科学技术进军，舆论媒体紧密配合，报道科技方面的先进人物及先进事迹，对科技的发展起推波助澜的作用，其中的典型代表是对我国科学家陈景润先生哥德巴赫猜想证明的报道。这篇报告文学发表后，激励起成批数学爱好者对被誉为“数学王冠上的明珠”——哥德巴赫猜想证明的兴趣。本文作者丁大千先生就是其中的一员。但是，能够二十几年如一日坚持不辍至今，这种精神难能可贵，而像丁大千先生这样作出成果的，实属凤毛麟角。这说明丁大千先生不是凭一时冲动而为，而是一位在科学上孜孜不倦的探索者。

事实也正是如此，作者凭借多年科学实验经验，从实验的角度对哥德巴赫猜想进行论证，并同时进行检查，花了大量的时间和精力，取得了丰硕的中间成果。这些成果可供正在从事这方面工作的同行们借鉴、参考。

哥德巴赫猜想的证明是一项重大科学课题，本文的结论应由从事这方面工作的专家来定论，愿本文成为专家们讨论的基础，愿“1 + 1”的证明之花开在中国的土地上。

韩立业

2005.2.23 于合肥

序　三

在科学技术高速发展的今天，各个学科领域的发展都离不开计算机网络和信息安全，然而，在网络和信息安全的领域里，密码学又是一个重中之重的学科。在这个学科中，数论是一个非常重要和有用的工具。初等数论又是这一切的基础。与数论相关的一个最令人瞩目的问题是“哥德巴赫猜想”。德国数学家哥德巴赫早在 1774 年提出了这样一个猜想：

1. 任何不小于 6 的偶数，都是两个奇质数之和。

2. 任何不小于 9 的奇数，都是三个奇质数之和。

此后，这个猜想就一直牵动着地球人类的脑神经，无数人为它魂牵梦萦，无数人为它奋斗终身，人们把是否可以求得这个猜想的题解视为是对人类最高智力的最高考验之一。历史上，许多著名的数学家都研究过这个问题，有些人做出了卓著的贡献。有些人一生都致力于证明哥德巴赫猜想却毫无结果。证明哥德巴赫猜想的难度，远远超出了人们的想象。为此人们把哥德巴赫猜想比喻为“数学王冠上的明珠”。

近几十年来，存在着一个与计算机的普及几乎是同步的现象，我们不妨将它称为“哥德巴赫猜想热”现象，那就是，越来越多的数学爱好者和计算机密码学的爱好者们对“哥德巴赫猜想”发生了兴趣，他们以极大的热情和志趣投入到对“哥德巴赫猜想”问题的求解的热潮中。自发的形成了一个民间的理论大团体及其相应的交流体系。在这个大论坛中，不乏有烁烁发光的数学思想，也不乏有鱼目混珠的奇谈怪论。其中的很多人将对“哥德巴赫猜想”的研究作为自己终身的乐趣和志向。加入这支队伍的人员成分非常广泛，他们有工程技术人员、高级知识分子、教师、工人、农民、企业人士、公务员等等。在他们中间所产生的一些有价值的研究成果，往往不能得到学术界的充分肯定和认同。但他们对科学和数学的执著，又使得他们不甘心把大半生的心血和研究成果白白的浪费。他们希望能够通过图书的形式保留他们的那些对后人可能是有用的思想和结论，也希望自己的研究成果在未来的某一个应用领域中可以被利用。总之，他们希望自己的工作对人类的进步能够有所影响。其实这种愿望不应该只是象牙塔中的科学家和数学家们可以有的，而是任何一个有自信的科学爱好者和研究者都可

以有的。正是出于这样纯粹的、无功利的愿望，作者编写了这本《论双生、对偶、三奇素数组分布》。

在这本书里作者用初等数论的方法，修改了通常所用的筛法，根据此筛法建立了有自己特色的筛法精确解，并用初等数论的方法给予了详尽的证明。更难能可贵的是，作者在书中详细的介绍了自己二十几年来进行研究的方法和历程，这给年轻的科学工作者树立了一个好的榜样。通过阅读此书，可以学习到老一辈科学工作者的风范。此书的另一个亮点是，给出了由精确解公式得到的素数、素数组分布计算系数表。就目前而言，在国内外尚无任何人曾经确确实实的给出过像这样可供查证的结果。另外，本书作者还给出了对偶素数组分布普查分析记录整理图表，以及三奇素数组分布普查记录表。这些图表和数据不仅从直观上说明了素数分布状态和趋热，也给信息工程上的查阅需要和运用需求带来许多方便。

在这百花齐放、百家争鸣的大好时代里，此书的另一个重要作用是，通过图书的传播，将学习科学、热爱科学、研究科学的风气在中国大地上普及开来，将严肃的科学精神与广泛的科学讨论联系起来，希望我国在更多领域里都能出现像“哥德巴赫猜想热”这样的学术气氛，使我国在各领域、各方面的科学讨论得到普及。这也是此书作者的一个心愿。

于筑国　2005 年 3 月
合肥工业大学计算机与信息学院

自　序

我于1979年夏萌发了探索素数、素数组分布的思想。首先，作为练习，应用埃拉多斯染尼氏(Eratosthenés)筛法，筛选了 $m=10000$ 以内的素数；随后于1979年末，制成了六支木检查尺[①]，成功地对 $m=12000$ 以内的任意偶数作了对偶素数组组数的检查，逐渐建立起明晰的对偶素数组分布等概念。同时着重探索其计算模型、解析公式，并上机计算各简化解计算结果与检查值比较，于1984年～1985年间写有(中间结果 一)：《素数、素数组分布解析与检查——〈双生素数猜想〉、〈哥德巴赫猜想〉浅证》(一、二、三稿)；于1986年写有(中间结果 二)：《素数、素数组分布问题之一——素数、双生、对偶及三奇素数组分布精确解》，《素数、素数组分布问题之二——素数、双生、对偶及三奇素数组分布简化解和主要结论》,《素数、素数组分布问题之三——素数、双生、对偶及三奇素数组分布检查记录择要》，均为内部资料。

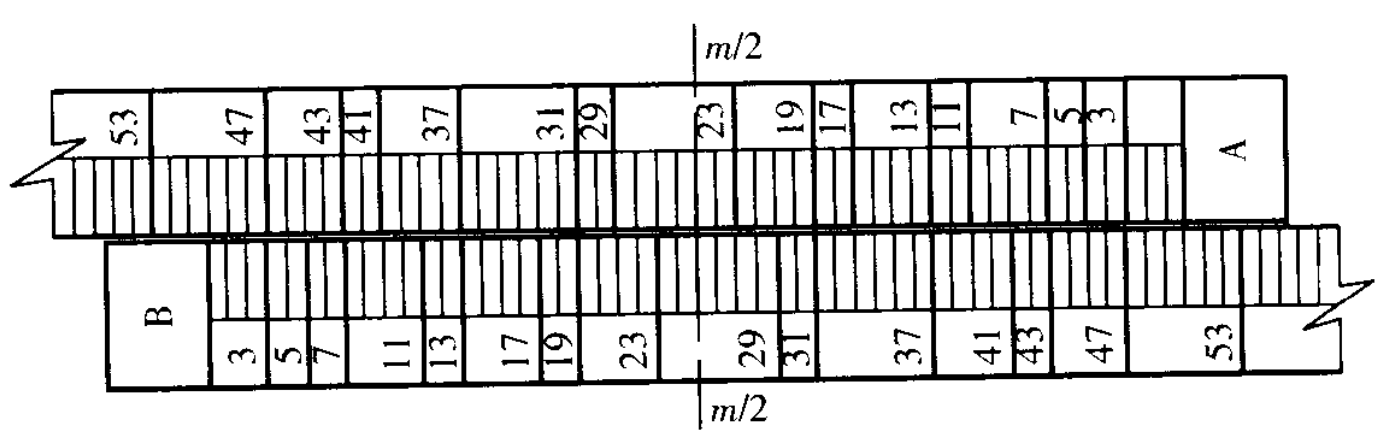
自二十世纪八十年代末本人退休后，着重探索各精确解暨检查值等电算程序的编制，1993年起，购置微机并上机计算。由于这些精确解暨检查值，主要是双生素数组分布精确解暨三奇素数组分布检查值，耗费电算机时过多，再三突破计划，添置机器，直到2001年底才完成双生素数组分布精确解计算，于2004年秋完成了三奇素数组分布检查计划。仅双生素数组分布精确解、计算上限 m

---

①注：关于检查尺及用检查尺检查对偶素数组分布的说明：

(1) 六支木检查尺。每支检查尺为 20mm×20mm×2200mm，各有四个尺面。每个尺面各按序位正确绘有 $dm=1000$ 以内的素数：12个A尺面为由小到大正排列，12个B尺面为由大到小倒排列。

(2) 检查大偶数 m 内有多少组对偶素数组。将A尺的0线对正B尺的 m 线(同时B尺的0线也对正A尺的 m 线)，则两个尺面上位置在同一条线的素数就是要找的对偶素数组了。就A尺而言，从 $0\sim m/2$ 内(含 $m/2$) 所有这样的组数就是我们要找的对偶素数组组数。参见简图0.1，$m=50$，$Q2=4$：3、47；7、43；13、37；19、31。



简图　0.1

= 100 万、抽查 42 个数、实际耗费微机运算有效机时 6.9 万多个，参见附表 2.1。(一年平年只有 8760 小时)。三奇素数组分布检查值则耗费微机机时更多。

之所以坚持要解决精确解和正确值检查的电算问题，我认为此类数学问题也应遵循“实践是检验真理的唯一标准”，电算是对此类计算理论起码的检验。

之所以重视精确解，我以为对于证明“猜想”而言，精确解是水之源、木之本，是建立正确的计算模型的依据；而正确的计算模型是证明“猜想”的前提。

有人认为，(现今) 要解决证明“哥德巴赫猜想”问题，必须应用精深的高等数学。对这话仔细推敲，似尚有商榷余地。

对具体问题作具体分析。分析“哥德巴赫偶数猜想”问题有四点：(1) 素数分布问题；(2) 交错重筛数问题；(3) 优善系数(F) 问题；(4) 常数项(H) 问题。

对于优善系数(F)，已往数学家们都是对少数几个素数因子作具体计算；对于计算系数(H)，有的专家建议采用接近下限值 0.6601…；对于素数分布问题，已有素数定理解决了(素数定理是应用精深的高等数学求到的)；只剩下交错重筛数问题。可以认为交错重筛数问题与素数分布问题相仿，都是舍弃“筛除余数”，但交错重筛数有“起筛数”问题部分补偿，其计算误差应更小些，更可参照素数分布问题、应用素数定理来加以解决。似也可不必再另觅途径。

一般认为素数定理是证明“哥德巴赫猜想”“双生素数猜想”的充分条件。用哲学语言写，可以作如下的充分条件假言判断：

“如果素数定理真，那么‘哥德巴赫猜想’‘双生素数猜想’的结论都是肯定的；

如果素数定理非真，那么‘哥德巴赫猜想’‘双生素数猜想’的结论未必不是肯定的。”

事实上，不借助素数定理，仅用素数分布密度系数及扩展系数的方法也可证明素数、双生、对偶、三奇素数组分布都是无限的，即上述“猜想”的结论都是肯定的。对不很大的 m 可以计算其数组的具体数值；对很大的、充分大的 m，可用不等式表达。

当然，用不等式表达不如用等式表达完美。故就这个意义上说，素数定理只是证明“哥德巴赫猜想”“双生素数猜想”的充分条件(而非必要条件)。

值得一提的是，按本人计算应用素数定理(对数公式) 计算素数分布，其计算误差并不比本人探讨的用素数分布密度系数或扩展系数方法的计算误差更小些。

本人在此项探索研究中，关于对偶、三奇素数组分布，都是采用较低值或最

低值“组合数”，而不采用考虑不同排列的“排列组合数”，可认为这是更落实的取值（一般对偶素数组含两个奇素数 pa、pb，有两种不同排列 pa、pb；pb、pa，参见简图 0.1，以 $m/2$ 为界，前、后两两对应相同；一般三奇素数组有三个奇素数，有六种不同排列）。

本书的论点是个人的观点，书中所列素数、各素数组分布检查数值，都是用电子计算机精心计算的，力求准确和精确，可供研究者检查、参考。

在漫长的此项探索研究工作中，先后得到安徽省电力设计院、安徽电力中心调度所、水电部电力规划设计研究院、水电部各级领导与组织的关心及支持；先后得到韩立业、朱志强、胡新舜、荆彩凤、陈苏宝、吴义应、郭放、王红、赵先进、徐磊等许多同志的热情帮助；先后得到安徽电力中心调度所计算机室、水电部电力规划院计算机室、安徽省电力设计院计算机室领导及几乎全体同志们的大力帮助，本人在此表示由衷的感激和感谢。

因本人数学知识面和专业水平很有限，且晚年精力不足，本书难免有缺点、错误和欠妥之处，诚恳希望读者惠予批评、指正。

丁大千
2004 年 5 月

符号说明

M, m—— 大数

m_2—— 大偶数(用于哥德巴赫偶数猜想问题)

m_3—— 大奇数(用于哥德巴赫奇数猜想问题)

p—— 素数

n—— 计算大数 m 的上限筛素数 p_n 在素数表上的序号,$p_n = \mathrm{INTP}(\sqrt{M})$

na—— 计算大数之半 $m/2$ 的上限筛素数 p_{na} 在素数表上的序号,

$p_{na} = \mathrm{INTP}(\sqrt{M/2})$

$n1$—— 计算大数 $m_3/1.2$ 的上限筛素数 p_{n1} 在素数表上的序号,

$p_{n1} = \mathrm{INTP}(\sqrt{M_3/1.2})$

$na1$—— 计算大数 $m_3/2.4$ 的上限筛素数 p_{na1} 在素数表上的序号,

$p_{na1} = \mathrm{INTP}(\sqrt{M_3/2.4})$

nk—— 大数 $m_3/3$ 内的最大素数 p_{nk} 在素数表上的序号,$p_{nk} = \mathrm{INTP}(M_3/3)$

nkk—— 计算大数 $m_3/3$ 的上限筛素数 p_{nkk} 在素数表上的序号,

$p_{nkk} = \mathrm{INTP}(\sqrt{M_3/3})$

Q_p—— 素数个数

Q_L—— 双生素数组组数

Q_{L3}—— 三生素数组组数

Q_{L4}—— 四生素数组组数

Q_{L5}—— 五生素数组组数

Q_2—— 对偶素数组组合数,简称对偶素数组组数

PQ_2—— 对偶素数组含不同排列的组合数,简称对偶素数组排列组合数

Q_3—— 三奇素数组组合数,简称三奇素数组组数

PQ_3—— 三奇素数组含不同排列的组合数,简称三奇素数组排列组合数

c_{pnj}—— 素数分布密度系数,$c_{pnj} = \prod_{i=1}^{nj}\left(1-\frac{1}{p_i}\right)$

R_{pnj}—— 素数分布扩展系数,$R_{pnj} = \prod_{i=2}^{nj}\frac{p_i - 1}{p_{i-1}}$

c_{Lnj}—— 双生素数组分布密度系数，$c_{Lnj} = 2 \cdot H_{nj} \cdot c_{pnj}^2 = \frac{1}{2} \cdot \prod_{i=2}^{nj} \frac{p_i - 2}{p_i}$

R_{Lnj}—— 双生(及对偶)素数组分布扩展系数，$R_{Lnj} = \prod_{i=3}^{nj} \frac{p_i - 2}{p_{i-1}}$

H_{nj}—— 计算系数，$H_{nj} = \prod_{i=2}^{nj} \left(1 - \frac{1}{(p_i - 1)^2}\right)$

G_{nj}—— 计算系数，$G_{nj} = \prod_{i=2}^{nj} \left(1 + \frac{1}{(p_i - 1)^3}\right)$

F_{na}—— 对偶素数组分布优善系数，$F_{na} = \prod_{i=1}^{x} \frac{p_{xi} - 1}{p_{xi} - 2}$；$p_{xi}$ 为大数 m 所含奇素数因子，即 $p_{xi} \mid m_2$，且有 $2 < p_{xi} \leqslant p_{na}$

F_3—— 三奇素数组分布优善系数，$F_3 = F_{3.1} \cdot F_{3.2} \cdot F_{3.3} \cdot F_{3.4} \geqslant 1$；$F_{3.1}$、$F_{3.2}$、$F_{3.3}$、$F_{3.4}$ 详见第四章

t_3——(三奇素数组分布)重复折减系数，建议取 $t_3 = 0.6$

ZB——(对偶素数组分布)校正系数；

$$ZB = 1 - \frac{\ln 2 \cdot (1 + 4/\sqrt{m_{2.1}}) - \ln m_{2.1} \cdot (4 - \sqrt{8})/\sqrt{m_{2.1}}}{2 \cdot H_{na1} \cdot F_3}$$

$\div$—— 整除符号

/—— 一般除法符号

$p \mid m$——p 能整除 m

$p \nmid m$——p 不能整除 m

\\—— 能整除才做(不能整除则不做)的除法符号

$\sum$—— 连加符号

$\prod$—— 连乘符号

$p_{iM}!$—— 素数阶乘符号；$p_{iM}! = \prod_{i=1}^{iM} p_i$

INT—— 取整符号

INTP—— 取整素数符号

PC_n^i—— 素数组合符号，自 n 个素数中取 i 个素数组合连乘取积

e—— 计算系数， $e = c_{pna}/c_{pn}$

$e1$—— 计算系数， $e1 = c_{pna1}/c_{pn1}$

d_{na}——(对偶素数组的)交筛系数，$d_{na} = \sum_{k=3}^{na} \prod_{i=3}^{k} \frac{p_{i-1} - 2}{p_i - 1}$

L_{na}—— 计算系数，$L_{na} = \dfrac{1 - d_{na}}{2}$ 或 $L_{na} = \prod_{i=2}^{na}\left(1 - \dfrac{1}{p_i - 1}\right)$

B_i—— 比值，$\dfrac{\text{素数个数 对数}}{\text{素数组组数对数}}$ 与大数对数之比，$B_i = \dfrac{\log(Q_i)}{\log(M_i)}$

A_i—— 仰角，过双对数坐标之0点与素数、素数组分布曲线上之某点的连线与 $\log(M_i)$ 轴线之夹角，$A_i = \arctan(B_i)$

目　　录

第一章　引言——关于素数分布的讨论

1.1　素数分布精确解

设有大数m,计算上限筛素数$p_n = \text{INTP}\sqrt{m}$,$n$为$p_n$在素数表上的序号;又小于等于$m$的最大素数阶乘为$p_{iM}! = \prod_{i=1}^{iM} p_i$,则这个大数内有素数个数$Q_p$按下式决定:

$$Q_{pA} \equiv \sum_{i=0}^{iM} (-1)^i m \div \prod C_n^i + (n-1) \qquad (1-1-A)$$

(1－1－A)式适用范围:$m \geqslant 4(p_n \geqslant 2, n \geqslant 1)$,没有上限,即适用于$m = 4$及以上的任何自然数。(1－1－A)式是精确解,故用恒等符号。

对(1－1－A)式的推导说明如次:

一个大数m被筛素数p_1筛除的(有效)筛除数是:

$$m \div p_1;$$

被筛素数p_2筛除的筛除数是$m \div p_2$;其中与p_1重筛的部分是$m \div p_1 \div p_2\{= m \div (p_1 \cdot p_2)$,下同$\}$,应扣除,故被$p_2$筛除的有效筛除数是:

$$m \div p_2 - m \div (p_1 \cdot p_2);$$

被p_3筛除的是$m \div p_3$;其中与p_1、p_2重筛的分别是$m \div (p_1 \cdot p_3)$和$m \div (p_2 \cdot p_3)$,也应扣除;还应考虑与p_1及p_2的双重筛,其数值是$m \div (p_1 \cdot p_2 \cdot p_3)$是多扣的部分,应予以校正,故被$p_3$筛除的有效筛除数是

$$m \div p_3 - [m \div (p_1 \cdot p_3) + m \div (p_2 \cdot p_3)] + m \div (p_1 \cdot p_2 \cdot p_3);$$

同理,被p_4筛除的有效筛除数是:

$$m \div p_4 - [m \div (p_1 \cdot p_4) + m \div (p_2 \cdot p_4) + m \div (p_3 \cdot p_4)] + [m \div (p_1 \cdot p_2 \cdot p_4) + m \div (p_1 \cdot p_3 \cdot p_4) + m \div (p_2 \cdot p_3 \cdot p_4)] - m \div (p_1 \cdot p_2 \cdot p_3 \cdot p_4);$$

余类推,一直筛到小于等于$\sqrt{m}$的最大筛素数p_n为止。从大数m中减去上列各有效筛除数,剩下的就是要找的素数个数。

在这里被p_1、$p_2 \cdots p_n$筛除时,p_1、$p_2 \cdots p_n$本身还要保留着,所以还应加n;又"1"也是筛剩数,却是不计在素数内的,即应减1,故实有素数个数为