ARTICLE

Aran Nayebi
On the distribution of Carmichael numbers
Rejecta Mathematica, Vol. 2, No. 1, pp. 24–43, June 2011
2011 Rejecta Publications
math.rejecta.org

An open letter concerning

On the distribution of Carmichael numbers

Aran Nayebi

In my paper, entitled "On the distribution of Carmichael numbers", I investigate the distribution of Carmichael numbers. The importance of Carmichael numbers is that they test the limits of Fermat's primality test, which ultimately led mathematicians to formulate more effective primality tests in the twentieth century. There have been two important conjectures regarding the distribution of these numbers up to sufficiently large bounds, one made by Paul Erdős in 1956 and a subsequent sharpening of this conjecture by Carl Pomerance in 1981. However, neither of these conjectures are well-supported by the Carmichael number counts famously performed by Richard Pinch up to 10^{21}. The inaccuracies of these two aforementioned conjectures are understandable, since not too much is known about Carmichael numbers. In fact, after a century of investigation regarding these numbers, it was only a decade ago that the infinitude of Carmichael numbers was proven! In this paper, I present two conjectures (which sharpen Erdős' and Pomerance's conjectures) regarding the distribution of Carmichael numbers that fit proven bounds, are roughly supported by Pinch's data (as well as data from other papers and resources), that closely model the true distribution of Carmichael numbers, and are supported by many theorems and conjectures put forth by renowned mathematicians such as Alford, Erdős, Galway, Granville, Harman, Pomerance, Wagstaff, Selfridge, and Szymiczek. The reader may wonder why *two* conjectures are presented. The reason is that due to the lack of information regarding Carmichael numbers and their distribution, both conjectures are viable to their own merit.

Unfortunately, although I feel that the results in this paper are important and would satisfy the interests of the mathematical community, the paper was rejected by three journals.

The first journal the paper was submitted to was *Mathematics of Computation*. The referee stated that "the paper deals with interesting topics and might be generally appropriate for Math. Comp. However, the paper is written very poorly and it needs a lot of work before it can be properly considered." Thus, I humbly took the advice of the referee, and I spent the better part of two months revising the paper rigorously with a colleague of mine. I made the paper more readable, the notation more recognizable, and I added six data tables from various cited sources (some of the data I collected myself), all in support of my conjecture. Similarly, through this revision process, we disproved many of my conjectures and theorems, and we sharpened and strengthened many of my proofs. However, the only conjecture that we were *unable to disprove* was my conjecture regarding Carmichael numbers. Furthermore, I discussed my paper with several mathematicians who are known for their work on Carmichael numbers and pseudoprimes (which are a superset of Carmichael numbers), all of whom agreed with the majority of my ideas. I also requested feedback from a mathematician who had not published any papers in this field, who stated: "I read through your paper on pseudoprimes, and while the subject is not my area of expertise, it is clear that you are familiar with the mathematical literature and are making a serious contribution."

Aran Nayebi
On the distribution of Carmichael numbers
***Rejecta Mathematica*, Vol. 2, No. 1, pp. 24–43, June 2011**
2011 Rejecta Publications
math.rejecta.org

After revising the paper thoroughly, I then submitted the paper to *The American Mathematical Monthly*. Although they were unable to find any mistakes (both mathematical and style-wise) and this time the paper received a good editorial review, the paper was rejected because "the Monthly tries to publish expositions of mathematics that are accessible to a broad mathematical audience. The material in your paper is rather technical, and we feel that many Monthly readers will find it forbidding. We will therefore not be able to accept it for publication. These are difficult decisions. The Monthly receives a large number of submissions each year, and we are able to publish only a small fraction of them."

I could not help but be amused by this rejection notice; however, I was somewhat flustered. Carmichael numbers are important in number theory because of their rarity (there are only 20138200 Carmichael numbers up to 10^{21}), and their existence demonstrates the ineffectiveness of the Fermat primality test. Furthermore, the fact that not too much is known about these numbers after almost a century of research, means that more work about them should be considered for inclusion within mathematical literature. Also, my paper is not forbidding as there are tables which present my assertions in non-verbiage form and these tables are even explained in detail. The notation is also entirely readable and widely-recognized.

As a final straw, I sent the paper to Carl Pomerance, in the hopes of a more extensive and in-depth peer review. At the time, Conjecture 1.0.4 (the second conjecture) had not been included in the manuscript; only Corollary 1.0.3 (the first conjecture) was presented as the main result. Although my correspondence with him was brief (parts of which I include in my paper), his advice was helpful. Pomerance's arguments in support of his conjecture compelled me to propose a second conjecture that was a refinement to his original 1981 one, mainly by utilizing finer estimates for the distribution of smooth numbers (a practice which he stated had not yet been done before). This conjecture, which later became Conjecture 1.0.4, gave extremely accurate counts for $C(x)$, the number of Carmichael numbers up to x, at least for smaller bounds (although asymptotically it is the same result as Pomerance's).

With these adjustments made, I submitted my manuscript to *Experimental Mathematics* as it is "a journal devoted to the experimental aspects of mathematics research." Unfortunately, two months later, they rejected the submission on the grounds that "the two conjectures presented by the author can each be substantially simplified by using known (or easily derived) asymptotics for the constituent parts....The first conjecture is extremely unlikely to be true, if only for the reason that it postulates an asymptotic formula for the number of Carmichael numbers up to x, while no other conjecture makes such a strong statement....Also, in the second conjecture, the author claims to be including more explicit secondary terms, but the $(1+o(1))$ factor just washes them out anyway. In short, the statements would need to be substantially simplified and polished to make this paper worth publishing in a strong journal such as EM." I agree with the referee that the statements would have to be simplified; a task which I had completed prior to submission, even going so far as to provide numerical estimates for the various constants used in the statement of Corollary 1.0.2. However, my points of contention with the referee are that the first conjecture cannot simply be disregarded as untrue due to the strength of its assertions (and in fact the numerical evidence compiled in my paper demonstrates its viability) and that the second conjecture must include secondary terms in it so that the discrepancies pointed out by Pinch will not occur.

If anything, the second conjecture appears to be more plausible than the first; however, both conjectures provide different and intriguing insights into the distribution of Carmichael numbers. The first conjecture asserts that an asymptotic formula for $C(x)$ easily follows based on the com-

putation of numerical constants. The second conjecture indicates to us that if secondary terms exist, then the properties of smooth number counting functions *must* be examined further in order to effectively prove an equality for $C(x)$.

Frankly, submitting the paper to another peer-reviewed journal and waiting a few months to a year for a referee look over a paper which has already been examined by several mathematicians of the same expertise (if not more) is a waste of time. I have submitted my paper to *Rejecta Mathematica* in the hopes of advancing mathematics and the investigation of pseudoprimes and their variants.

On the distribution of Carmichael numbers

Aran Nayebi*

Abstract

Erdős conjectured in 1956 that there are $x^{1-o(1)}$ Carmichael numbers up to x. Pomerance made this conjecture more precise and proposed that there are $x^{1-\frac{\{1+o(1)\}\log\log\log x}{\log\log x}}$ Carmichael numbers up to x. At the time, his data tables up to $25 \cdot 10^9$ appeared to support his conjecture. However, Pinch extended this data and showed that up to 10^{21}, Pomerance's conjecture did not appear well-supported. Thus, the purpose of this paper is two-fold. First, we build upon the work of Pomerance and others to present an alternate conjecture regarding the distribution of Carmichael numbers that fits proven bounds and is better supported by Pinch's new data. Second, we provide another conjecture concerning the distribution of Carmichael numbers that sharpens Pomerance's heuristic arguments. We also extend and update counts pertaining to pseudoprimes and Carmichael numbers, and discuss the distribution of One-Parameter Quadratic-Base Test pseudoprimes.

1 Introduction

Fermat's "little" theorem states that if b is an integer prime to n, and if n is prime, then

$$b^n \equiv b \pmod{n}. \tag{1.0.1}$$

When $\gcd(b, n) = 1$, we can divide by b,

$$b^{n-1} \equiv 1 \pmod{n}. \tag{1.0.2}$$

A composite natural number n for which $b^{n-1} \equiv 1 \pmod{n}$ for any fixed integer $b \geq 2$ is a base b pseudoprime. A positive composite integer n is a Carmichael number if $b^{n-1} \equiv 1 \pmod{n}$ for all integers $b \geq 2$ with $\gcd(b, n) = 1$. The importance of Carmichael numbers is that they test the limits of the Fermat primality test, which ultimately led mathematicians to formulate more effective tests. Furthermore, there is little that is known about them; for instance, the infinitude of Carmichael numbers has only recently been proven by Alford, Granville, and Pomerance [3].

Let $\mathscr{P}_b(x)$ denote the number of base b pseudoprimes $\leq x$ and let $C(x)$ denote the number of Carmichael numbers $\leq x$. In 1899, Korselt [4] provided a method for identifying Carmichael numbers

Theorem 1.0.1. *An odd number n is a Carmichael number iff n is squarefree and $p-1 \mid n-1$ for all $p \mid n$, where p is a prime number.*

As a consequence of Theorem 1.0.1, it is easy to see that Carmichael numbers have at least three prime factors.

In 1910, Carmichael [24] found the smallest Carmichael number to be $561 = 3 \cdot 11 \cdot 17$.

*727 Moreno Avenue, Palo Alto, California, United States of America 94303-3618. Email: aran.nayebi@gmail.com

Table 1: Counts of k-prime Carmichael numbers

Bound	$C_3(x)$	$C_4(x)$	$C_5(x)$	$C_6(x)$	$C_7(x)$	$C_8(x)$	$C_9(x)$	$C_{10}(x)$	$C_{11}(x)$	$C_{12}(x)$	$C(x)$
10^3	1	0	0	0	0	0	0	0	0	0	1
10^4	7	0	0	0	0	0	0	0	0	0	7
10^5	12	4	0	0	0	0	0	0	0	0	16
10^6	23	19	1	0	0	0	0	0	0	0	43
10^7	47	55	3	0	0	0	0	0	0	0	105
10^8	84	144	27	0	0	0	0	0	0	0	255
10^9	172	314	146	14	0	0	0	0	0	0	646
10^{10}	335	619	492	99	2	0	0	0	0	0	1547
10^{11}	590	1179	1336	459	41	0	0	0	0	0	3605
10^{12}	1000	2102	3156	1714	262	7	0	0	0	0	8241
10^{13}	1858	3639	7082	5270	1340	89	1	0	0	0	19279
10^{14}	3284	6042	14938	14401	5359	655	27	0	0	0	44706
10^{15}	6083	9938	29282	36907	19210	3622	170	0	0	0	105212
10^{16}	10816	16202	55012	86696	60150	16348	1436	23	0	0	246683
10^{17}	19539	25758	100707	194306	172234	63635	8835	240	1	0	585355
10^{18}	35586	40685	178063	414660	460553	223997	44993	3058	49	0	1401664
10^{19}	65309	63343	306310	849564	1159167	720406	196391	20738	576	2	3381806
10^{20}	120625	98253	514381	1681744	2774702	2148017	762963	114232	5804	56	8220777
10^{21}	224763	151566	846627	3230120	6363475	6015901	2714473	547528	42764	983	20138200

Based on Korselt's criterion, Erdős [21] formulated a method for constructing Carmichael numbers, which was the basis for the proof of Alford, Granville, and Pomerance [3]. His notion was to replace "$p-1 \mid n-1$ for all $p \mid n$" in Theorem 1.0.1 with $L \mid n-1$ for $L := \operatorname{lcm}_{p|n}(p-1)$. By focusing primarily on L, Erdős found every p for which $p-1 \mid L$ and then tried to find a product of those primes in which $\equiv 1 \pmod{L}$ [2]. His results heuristically suggested that for sufficiently large x,

$$C(x) = x^{1-o(1)}. \tag{1.0.3}$$

More convincingly, Theorem 4 of [3] shows that (1.0.3) holds if one assumes widely-believed assumptions regarding primes in arithmetic progressions. However, drawing upon available data at the time, Shanks [12] was skeptical of (1.0.3) because the counts of Carmichael numbers seemed to have noticeably fewer prime factors than those predicted by Erdős' heuristic.

Granville and Pomerance [2] conjectured that the reason for the difference between the computational evidence and the argument of (1.0.3) stems from a grouping of Carmichael numbers into two distinct classes, namely primitive and imprimitive. If we let $g = g(n) := \gcd(p_1-1, p_2-1, \cdots, p_k-1)$ for a squarefree integer $n = p_1p_2\cdots p_k$ and put $pa_i = p_i - 1$ for some integer a_i, then n is a primitive Carmichael number if $g(n) \leq [a_1, \cdots, a_k]$, and imprimitive if otherwise. Thus, since the observations of Shanks are more applicable to imprimitive Carmichael numbers and those of Erdős are more applicable to primitive Carmichael numbers, and most Carmichael numbers are in fact primitive whereas most Carmichael numbers with a fixed number of prime factors are imprimitive, then the two conjecturers easily reached different conclusions.

Interestingly, Pinch's counts of k-prime Carmichael numbers up to 10^{21} [25] reproduced in Table 1 imply that the number of prime factors of primitive Carmichael numbers tends to increase as x gets larger. As can be implied from Table 1, for the maximum number of distinct prime factors $k(x) \ll \frac{\log x}{\log^{(2)} x}$,

$$C(x) = C_3(x) + C_4(x) + C_5(x) + \cdots + C_{k(x)}(x), \tag{1.0.4}$$

where $\log^{(j)} x$ denotes the j-fold iteration of the natural logarithm for $j \geq 2$ (we shall use this notation from now on). Moreover, if we allow $C_k(x)$ to represent the number of Carmichael numbers $\leq x$ with precisely $k \geq 3$ prime factors, then it is conjectured that

$$C_k(x) = \Omega_k(x^{1/k}/\log^k x). \tag{1.0.5}$$

Table 2: Values of $h(x)$

Bound	$h(x)$
10^3	2.93319
10^4	2.19547
10^5	2.07632
10^6	1.97946
10^7	1.93388
10^8	1.90495
10^9	1.87989
10^{10}	1.86870
10^{11}	1.86421
10^{12}	1.86377
10^{13}	1.86240
10^{14}	1.86293
10^{15}	1.86301
10^{16}	1.86406
10^{17}	1.86472
10^{18}	1.86522
10^{19}	1.86565
10^{20}	1.86598
10^{21}	1.86619

Returning to the Erdős-Shanks controversy, Pomerance [2] sharpened the conjecture in (1.0.3) for all large x in order to be consistent with both Shanks' and Erdős' observations. Define the function $h(x)$ as

$$C(x) = x \cdot \exp\Big\{ - h(x)\frac{\log x \log^{(3)} x}{\log^{(2)} x}\Big\}. \tag{1.0.6}$$

According to Pomerance, distribution of Carmichael numbers is given by

$$C(x) = x^{1-\frac{\{1+o(1)\}\log^{(3)} x}{\log^{(2)} x}}, \tag{1.0.7}$$

for x sufficiently large. Unfortunately, according to Pinch [26], there appears to be no limiting value on h as indicated by the recent counts of Carmichael numbers up to 10^{21}. It is obvious that (1.0.7) holds iff $\lim h = 1$ in (1.0.6). However, Pinch [26] explains that the decrease in h is reversed between 10^{13} and 10^{14}, which is presented in Table 2. In fact, there is no clear evidence that suggests $\lim h = 1$.

As a result, we present an alternate conjecture

Conjecture 1.0.2.

$$C(x) \sim \frac{C_3(x)\mathscr{P}_b(x)}{\mathscr{P}_{b,2}(x)}, \tag{1.0.8}$$

where $\mathscr{P}_{b,2}(x)$ is the number of two-prime base b pseudoprimes $\leq x$ and $C_3(x)$ is the number of three-prime Carmichael numbers $\leq x$.

From Conjecture 1.0.2, we derive a corollary that the number of Carmichael numbers up to x sufficiently large is

Corollary 1.0.3.

$$C(x) \sim \frac{\psi' x^{\frac{5}{6}}}{\log x \cdot L(x)} \sim \frac{\psi_1' x^{\frac{1}{2}} \log^2 x \displaystyle\int_2^{x^{\frac{1}{3}}} \frac{dt}{\log^3 t}}{L(x)}, \tag{1.0.9}$$

where $L(x) = \exp\{\frac{\log x \log^{(3)} x}{\log^{(2)} x}\}$, $\psi' = \frac{\tau_3}{C}$, $\psi_1' = \frac{\tau_3}{27C}$. *If we let* p, q, *and* d *be odd primes, and we define* $\omega_{a,b,c}(p)$ *as the number of distinct residues modulo* p *represented by* a, b, c, *then the constants* C *and* τ_3 *are explicitly given as such,*

$$C = 4T \sum_{s \geq 1} \sum_{\substack{r>s \\ \gcd(r,s)=1}} \frac{\delta(rs)\rho(rs(r-s))}{(rs)^{\frac{3}{2}}} \tag{1.0.10}$$

$$T = 2\prod_{d} \frac{1-2/d}{(1-1/d)^2},$$

$$\rho(m) = \prod_{d|m} \frac{d-1}{d-2},$$

$$\delta(m) = \begin{cases} 2, & \text{if } 4 \mid m; \\ \\ 1, & \text{if otherwise.} \end{cases}$$

$$\tau_3 = \kappa_3 \lambda, \tag{1.0.11}$$

$$\lambda := 121.5 \prod_{p>3} \left(\frac{1-3/p}{(1-1/p)^3} \right),$$

$$\kappa_3 = \sum_{n \geq 1} \frac{\gcd(n,6)}{n^{4/3}} \prod_{\substack{p|n \\ p>3}} \frac{p}{p-3} \sum_{\substack{a<b<c,\, n=abc \\ a,b,c \, pairwise \, coprime}} \delta'(a,b,c) \prod_{\substack{p \nmid n \\ p>3}} \frac{p - \omega_{a,b,c}(p)}{p-3},$$

$$\delta'(a,b,c) = \begin{cases} 2, & \text{if } a \equiv b \equiv c \not\equiv 0 \pmod 3; \\ \\ 1, & \text{if otherwise.} \end{cases}.$$

Based upon the computation of C made by Galway [29] and the evaluation of κ_3 by Chick and Davies [16], we believe that ψ' will approach 69.51 and ψ_1' will approach 2.57; although these values are not yet borne out by the data. We also demonstrate that Corollary 1.0.3 fits the proven upper and lower bounds for $C(x)$, that ψ' and ψ_1' appear to approach constant values based upon Pinch's data, and we support Conjecture 1.0.2 through computational efforts.

In private communication [13], Pomerance suggests to us that the reason for $h(x)$ not approaching its conjectural limit of 1 is that "some secondary terms may be present. So, say in my conjecture, one replaces "$\log^{(3)} x$" with "$\log^{(3)} x + \log^{(4)} x$". It is the same conjecture, since the two are asymptotic...and so the Pinch phenomenon is banished". Hence, if secondary terms do indeed exist, then another conjecture regarding $C(x)$ would be to sharpen the heuristic arguments in [5] which, as a consequence, may better match the *actual* counts of Carmichael numbers. Since these heuristic arguments are dependent upon the number of y-smooth numbers up to x, represented by $\Psi(x,y)$, with y in the vicinity of $\exp\{(\log x)^{\frac{1}{2}}\}$, then it would suffice to utilize improvements concerning the asymptotic distribution of these numbers in the aforementioned region. As a result of these endeavors, we obtain the more precise heuristic:

Conjecture 1.0.4. *Let $\pi(x)$ be the prime counting function, for x sufficiently large $C(x)$ is*

$$x^{1-\frac{\{1+o(1)\}\log^{(3)}x+1}{\log^{(2)}x}}. \tag{1.0.12}$$

In Table 3, we define the function

$$a(x) := \left(\frac{(\log^{(2)}x)^2\pi((\log x)^{\log^{(2)}x})\exp\{-\{1+o(1)\}\log^{(2)}x\log^{(3)}x\}}{\log x}\right)^{\log x/(\log^{(2)}x)^2}.$$

Although Conjecture 1.0.4 states the same result and is a much more simplified version of $a(x)$, $a(x)$ is a slightly more precise version (for $x < 10^{100}$) of the conjecture and is thus used in the table instead of (1.0.12).

The reader may wonder why *two* conjectures are presented. The reason is that due to the lack of information regarding Carmichael numbers and their distribution. Corollary 1.0.3 asserts that if the values of ψ' and ψ_1' can be accurately determined then an asymptotic formula for $C(x)$ easily follows. Conjecture 1.0.4 indicates to us that if secondary terms exist, then the relation between the functions $\Psi(x,y)$ and $\Psi'(x,y)$ *must* be examined further (we explain this concept fully in §3.4) to effectively prove an equality for $C(x)$. We should note that the values of $C(x)$ predicted by Corollary 1.0.3 and Conjecture 1.0.4 appear to be closer to the actual values of $C(x)$ than Pomerance's conjecture in (2.2.2). Moreover, at least up to 10^{21}, it appears that Conjecture 1.0.4 is presenting more accurate values of $C(x)$ than Corollary 1.0.3; although, this may cease to be the case for larger bounds. In fact, (1.0.12) is asymptotically the same as (1.0.7); however, the usage of secondary terms in the former equation provides sharper estimates for smaller bounds than does (1.0.7).

Table 3: Comparisons between the actual and predicted Carmichael number counts

Bound	$C(x)$	$\frac{69.51x^{\frac{5}{6}}}{\log x\cdot L(x)}$	$\frac{2.57x^{\frac{1}{2}}\log^2 x\int_2^{x^{\frac{1}{3}}}\frac{dt}{\log^3 t}}{L(x)}$	$a(x)$	$x^{1-\frac{\{1+o(1)\}\log^{(3)}x}{\log^{(2)}x}}$
10^3	1	301.95	1092.82	3.50	94.89
10^4	7	594.43	2835.17	7.81	365.59
10^5	16	1316.29	6640.29	18.18	1485.33
10^6	43	3131.53	14806.24	43.43	6224.10
10^7	105	7826.17	32411.27	107.50	26636.80
10^8	255	20282.91	71150.56	274.074	115803.60
10^9	646	54070.80	159157.24	724.86	509769.35
10^{10}	1547	147451.71	367012.00	1926.56	2267174.18
10^{11}	3605	409716.38	878601.38	5245.56	10171329.99
10^{12}	8241	1156637.85	2188667.23	14488.22	45977679.09
10^{13}	19279	3309970.24	5664006.88	40424.93	209219668.02
10^{14}	44706	9585268.36	15162465.67	114558.014	957710051.36
10^{15}	105212	28049810.91	41763706.96	329251.92	4407472357.25
10^{16}	246683	82852448.55	117743387.56	955940.22	20382638275.29
10^{17}	585355	246785788.13	338238941.70	2796027.81	94682736406.04
10^{18}	1401644	740679196.52	986503770.93	8260103.95	441642695710.74
10^{19}	3381806	2238429061.23	2913197684.15	24637581.64	2067911761776.64
10^{20}	8220777	6807841639.58	8692508977.60	74026750.39	9717200728399.57
10^{21}	20138200	20826296835.28	26167265004.43	224193470.90	45814162191297.01

2 Preliminaries

Before delving into the main results of this paper, we shall first present results regarding pseudoprimes and Carmichael numbers that we will explicitly use later on in our derivations.

2.1 Pseudoprimes

Currently, the tightest bounds for pseudoprime distribution have been proven by Pomerance [27] [5].

Theorem 2.1.1 (R. A. Mollin 1989, Pomerance 1981)**.** *For the base* 2 *pseudoprime counting function,* $\exp\{(\log x)^{\frac{85}{207}}\} \leq \mathscr{P}_2(x) \leq x \cdot L(x)^{\frac{-1}{2}}$*, where* $L(x) = \exp\{\frac{\log x \log^{(3)} x}{\log^{(2)} x}\}$*. These bounds are applicable to* $\mathscr{P}_b(x)$ *for* $x \geq x_0(b)$*.*

Theorem 2.1.2 (Pomerance 1981)**.** *If we allow* $l_2(n)$ *to denote the exponent with multiplicative order of* 2 *modulo* n*, then* n *is a pseudoprime* (base 2) *iff* $l_2(n) \mid n-1$*.*

Conjecture 2.1.3 (Pomerance 1981)**.** *The number of solutions* w *for all* n *and* x *sufficiently large is,*

$$\#\{w \leq x : l_2(w) = n\} \leq x \cdot L(x)^{-1+\theta(x)}, \lim_{x\to\infty} \theta(x) = 0. \tag{2.1.1}$$

As a result, the number of base b *pseudoprimes for sufficiently large* $x \geq x_0(b)$ *is conjectured to be,*

$$\mathscr{P}_b(x) \sim x \cdot L(x)^{-1}. \tag{2.1.2}$$

Galway [29] has recently conjectured a formula for the distribution of pseudoprimes with two distinct prime factors, p and q, based on a longstanding conjecture of Hardy and Wright concerning the density of prime pairs. He noticed that a majority of these pseudoprimes satisfy the relation $\frac{p-1}{q-1} = \frac{r}{s}$, where r and s are small coprime integers. Thus, we heuristically have

Conjecture 2.1.4 (Galway 2004)**.** *Allow* p*,* q*, and* d *be odd primes, allow* $\mathscr{P}_{b,2}(x)$ *to represent the counting function for odd pseudoprimes with two distinct prime factors, and* $\mathscr{P}_{b,2}(x) := \#\{n \leq x : n = pq, p < q, \mathscr{P}_b(n)\}$*. Hence, as* $x \to \infty$*,*

$$\mathscr{P}_{b,2}(x) \sim \frac{Cx^{\frac{1}{2}}}{\log^2 x}, \tag{2.1.3}$$

where

$$C = 4T\sum_{s\geq 1} \sum_{\substack{r>s \\ \gcd(r,s)=1}} \frac{\delta(rs)\rho(rs(r-s))}{(rs)^{\frac{3}{2}}} \approx 30.03, \tag{2.1.4}$$

$$T = 2\prod_{d} \frac{1-2/d}{(1-1/d)^2} \approx 1.32, \tag{2.1.5}$$

$$\rho(m) = \prod_{d\mid m} \frac{d-1}{d-2}, \tag{2.1.6}$$

$$\delta(m) = \begin{cases} 2, & \text{if } 4 \mid m; \\ \\ 1, & \text{if otherwise.} \end{cases} \tag{2.1.7}$$

Table 4: Values of C

Bound	$\mathscr{P}_{b,2}(x)$	C
10^3	0	0
10^4	11	9.331
10^5	34	14.251
10^6	107	20.423
10^7	311	25.550
10^8	880	29.860
10^9	2455	33.340
10^{10}	6501	34.468
10^{11}	17207	34.908
10^{12}	46080	35.181
10^{13}	123877	35.100
10^{14}	334567	34.767
10^{15}	915443	34.534
10^{16}	2520503	34.210
10^{17}	7002043	33.928

Galway's conjecture is somewhat supported by Table 4 for it appears that C is slowly approaching its predicted constant value of 30.03:

Let $\omega(n)$ represent the number of different prime factors of n. Also, given an integer sequence $\{m_i\}_{i=1}^{\infty}$, note that a prime p is said to be a primitive prime factor of m_i if p divides m_i but does not divide any m_j for $j < i$.

Lemma 2.1.5 (Erdős 1949)**.** *Let n be a base 2 pseudoprime. For every k, there exist infinitely many squarefree base 2 pseudoprimes with $\omega(n) = k$* [20].

Theorem 2.1.6. *There exist infinitely many squarefree base b pseudoprimes n for any $b \geq 2$ with $\omega(n) = k$ distinct prime factors.*

Proof. Let $\{n_j\}_{j=1}^{\infty}$ be an integer sequence of base b pseudoprimes such that each term is greater than its preceding term, and $\omega(n_i) = k-1$, for any n_i in $\{n_j\}_{j=1}^{\infty}$. Let p_i be one of the primitive prime factors of $b^{n_i-1}-1$. Since $b^{n_i-1} \equiv 1 \pmod{p_i \cdot n_i}$ and $b^{p_i-1} \equiv 1 \pmod{p_i}$, $p_i \cdot n_i$ is a pseudoprime to base b. We observe that $b^{p_i-1} \equiv 1 \pmod{n_i}$ because $p_i - 1 \equiv 0 \pmod{(n_i-1)}$. As a result, it follows that $b^{n_i-1} \equiv 1 \pmod{n_i}$. Also, $b^{n_i p_i-1} \equiv 1 \pmod{p_i \cdot n_i}$ since $b^{n_i p_i-1} = b^{(n_i-1)(p_i-1)} \cdot b^{n_i-1} \cdot b^{p_i-1}$. Hence, $p_i \cdot n_i$ is squarefree and $\omega(p_i \cdot n_i) = k$. Moreover, every integer satisfying $p_i \cdot n_i$ is different because n_i is composite, $p_i > n_i$, and $p_i \equiv 1 \pmod{(n_i - 1)}$. ■

Theorem 2.1.7. *For any base b pseudoprime, $b \geq 2$, having $k \geq 2$ distinct prime factors and for x sufficiently large,*

$$\mathscr{P}_{b,k+1}(x) \geq \mathscr{P}_{b,k}(\log_b x). \tag{2.1.8}$$

Proof. Let n be a pseudoprime with $k > 1$ distinct prime factors. Since $n-1$ is the smallest exponent ϵ such that $p \mid b^{\epsilon} - 1$ and ϵ divides an exponent h such that $p \mid b^h - 1$, it follows from Fermat's little theorem that $p \mid b^{p-1} - 1$. Thus, from Zsigmondy's theorem, there exists a prime $p > n$ for which $p \mid b^{n-1} - 1$ and $n - 1 \mid p - 1$ for $b \geq 2$. As a result,

$$np \mid b^{n-1} - 1. \tag{2.1.9}$$

On the other hand, since $np-1 = n(p-1)+n-1$ and $n-1 \mid p-1$, $n-1 \mid np-1$ and $np \mid b^{np-1}-1$. If we let $n, m \in \mathbb{N}^*$, the set of positive natural numbers, such that $n \neq m$ and $p > n$, $q > m$, then $np \neq mq$ for primes p and q. However, suppose we let $np = mq$ and $p > n$, then $m \mid p$. Hence, $m \geq p$ and $m > n$. Unfortunately, the latter statement is contradictory, and as a result $np \neq mq$. If n and m are two different base b pseudoprimes with $k \geq 2$ distinct prime factors, then np and mq are distinct pseudoprimes as well.

From (2.1.9),

$$p \mid (b^{\frac{n-1}{2}} - 1)(b^{\frac{n-1}{2}} + 1), \tag{2.1.10}$$

and

$$p \leq b^{\frac{n-1}{2}} + 1 < b^{\frac{n}{2}}. \tag{2.1.11}$$

If $n \leq \log_b x$, then $pn < x^{\frac{1}{2}} \log_b x < x$. It then follows that for every base b pseudoprime n with k distinct prime factors, $n = p_1 p_2 \cdots p_k \leq \log_b x$, there is at least one base b pseudoprime such that $p_1 p_2 \cdots p_k p < x$. ■

2.2 Carmichael Numbers

Improving upon Erdős' results in [21], Pomerance [5] sharpened the upper bound on $C(x)$.

Theorem 2.2.1 (Pomerance 1981)**.**

$$C(x) \leq x \cdot \exp\Big\{ -\frac{\log x}{\log^{(2)} x} \left(\log^{(3)} x + \log^{(4)} x + \frac{\log^{(4)} x - 1}{\log^{(3)} x} + O\left(\left(\frac{\log^{(4)} x}{\log^{(3)} x} \right)^2 \right) \right) \Big\}. \tag{2.2.1}$$

In the other direction, Alford, Granville, and Pomerance proved a lower bound for $C(x)$ for x sufficiently large [3].

Theorem 2.2.2 (Alford-Granville-Pomerance 1994)**.**

$$C(x) > x^{\frac{2}{7}}, \tag{2.2.2}$$

thus there are infinitely many Carmichael numbers.

Recently, Harman improved this lower bound [15].

Theorem 2.2.3 (Harman 2005)**.**

$$C(x) > x^{0.33336704}, \tag{2.2.3}$$

It is not yet even known if $C(x) > x^{\frac{1}{2}}$.

We provide in Table 5 a computation of the exponent β for which $C(x) = x^{\beta}$ for a sufficient value of x up to 10^{21}.

Conjecture 2.2.4 (Granville-Pomerance 2001)**.** *If we let* $C_3(x)$ *be the counting function for Carmichael numbers with* 3 *distinct prime factors, then*

$$C_3(x) \sim \tau_3 \frac{x^{\frac{1}{3}}}{\log^3 x} \sim \frac{\tau_3}{27} \int_2^{x^{\frac{1}{3}}} \frac{dt}{\log^3 t}, \tag{2.2.4}$$

where

$$\tau_3 = \kappa_3 \lambda \approx 2100, \tag{2.2.5}$$

Table 5: Values of β

Bound	10^3	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}
$C(x)$	1	7	16	43	105	255	646	1547
β	0	0.21127	0.24082	0.27224	0.28874	0.30082	0.31225	0.31895

Bound	10^{11}	10^{12}	10^{13}	10^{14}	10^{15}	10^{16}
$C(x)$	3605	8241	19279	44706	105212	246683
β	0.32336	0.32633	0.32962	0.33217	0.33480	0.33700

Bound	10^{17}	10^{18}	10^{19}	10^{20}	10^{21}
$C(x)$	585355	1401644	3381806	8220777	20138200
β	0.33926	0.34148	0.34364	0.34575	0.34781

$$\lambda := 121.5\prod_{p>3}\left(\frac{1-3/p}{(1-1/p)^3}\right) \approx 77.1727, \tag{2.2.6}$$

$$\kappa_3 = \sum_{n\geq 1}\frac{\gcd(n,6)}{n^{4/3}}\prod_{\substack{p\mid n\\ p>3}}\frac{p}{p-3}\sum_{\substack{a<b<c,\, n=abc\\ a,b,c\, pairwise\, coprime}}\delta^{'}(a,b,c)\prod_{\substack{p\nmid n\\ p>3}}\frac{p-\omega_{a,b,c}(p)}{p-3}, \tag{2.2.7}$$

$$\delta^{'}(a,b,c) = \begin{cases} 2, & \text{if } a\equiv b\equiv c\not\equiv 0 \pmod 3;\\ \\ 1, & \text{if otherwise.}\end{cases}, \tag{2.2.8}$$

and $\omega_{a,b,c}(p)$ is the number of distinct residues modulo p represented by a, b, c.

Recent provisional estimates by Chick and Davies [16] of the slowly converging infinite series κ_3 suggest that $\kappa_3 = 27.05$ which gives $\tau_3 = 2087.5$.

3 On the Distribution of Carmichael Numbers

3.1 Two Conjectures Regarding k-prime Pseudoprimes and k-prime Carmichael numbers

We conjecture the following relations:

Conjecture 3.1.1. *For any fixed $k \geq 2$, let $\mathscr{P}_{b,k}(x)$ denote the counting function for base b pseudoprimes with k distinct prime factors, and let $\mathscr{P}_b(x)$ denote the counting function for base b pseudoprimes. Asymptotically,*

$$\frac{\mathscr{P}_{b,k}(x)}{\mathscr{P}_b(x)} = o(1). \tag{3.1.1}$$

In other terms, for any fixed base $b > 1$, the k-prime base b pseudoprimes, $\mathscr{P}_{b,k}(x)$, form a set of relative density 0 in the set of all *base b pseudoprimes, $\mathscr{P}_b(x)$, for that same value of b.*

We are only able to partially support Conjecture 3.1.1. First, we express the ratio $\frac{\mathscr{P}_{b,k}(x)}{\mathscr{P}_b(x)}$ as,

$$\frac{\mathscr{P}_{b,k}(x)}{\mathscr{P}_b(x)} = \frac{\mathscr{P}_{b,k}(x)}{\sum\limits_{i=2}^{k(x)}\mathscr{P}_{b,i}(x)}, \tag{3.1.2}$$

where the maximum number of distinct prime factors, $k(x)$, of any integer $\leq x$ is $k(x) \ll \frac{\log x}{\log^{(2)} x}$. Let $\log_b^{(j)} x$ denote the the j-fold iteration of the base b logarithm. Thus,

$$\sum_{i=2}^{g(x)} \mathscr{P}_{b,i}(x) = \sum_{i=2}^{k-1} \mathscr{P}_{b,i}(x) + \mathscr{P}_{b,k}(x) + \sum_{i=k+1}^{k(x)} \mathscr{P}_{b,i}(x). \tag{3.1.3}$$

Due to Theorem 2.1.7, for any $h \leq k$ in (3.1.3), $\mathscr{P}_{b,k}(x) \geq \mathscr{P}_{b,h}(\log_b^{(k-h)} x)$, and for any $w \geq k$ in (3.1.3), $\mathscr{P}_{b,w}(x) \geq \mathscr{P}_{b,k}(\log_b^{(w-k)} x)$. Hence,

$$\sum_{i=2}^{k-1} \mathscr{P}_{b,i}(x) \leq \mathscr{P}_{b,2}(\log_b^{(k-2)} x) + \mathscr{P}_{b,3}(\log_b^{(k-3)} x) + \cdots + \mathscr{P}_{b,k-1}(\log_b x) \tag{3.1.4}$$

We cut off the terms from proceeding until $\frac{\log x}{\log^{(2)} x}$ because if such were the case, then no x could be sufficiently large to satisfy (3.1.5),

$$\sum_{i=k+1}^{k(x)} \mathscr{P}_{b,i}(x) \geq \mathscr{P}_{b,k+1}(\log_b x) + \cdots + \mathscr{P}_{b,r(x)}(\log_b^{(r(x)-k)} x), \tag{3.1.5}$$

where $r(x)$ is any function that grows slower than $\log^* x$, the iterated logarithm. We explicitly define $\log^* x$ as

$$\log^* x := \begin{cases} 0 & \text{if } x \leq 1; \\ 1 + \log^*(\log x) & \text{if } x > 1 \end{cases}. \tag{3.1.6}$$

Remark 3.1.2. We should note that the support for Conjecture 3.1.1 is rather weak. This is largely due to the weakness of Szymiczek's construction, $\mathscr{P}_{b,k+1}(x) \geq \mathscr{P}_{b,k}(\log_b x)$, in Theorem 2.1.7. We believe that the latter relation can be strengthened if a polynomial decrease can be proven. In other words, if $\mathscr{P}_{b,k+1}(x) \geq \mathscr{P}_{b,k}(x^c)$ for some $c \in (0,1)$. Similarly, in our support for Conjecture 3.1.1, we defined the function $r(x)$ as any function that grows slower than $\log^* x$, the iterated logarithm. Although it is not hard to see that any function growing faster than $\log^* x$ will fail, it is not obvious whether any function growing at the same rate as $\log^* x$ will succeed. However, we have several reasons to strongly believe that $r(x) = \log^* x$. First, for practical values of $x \leq 2^{65536}$ the iterated logarithm grows much more slowly than the logarithm. Second, the iterated logarithm's relation to the super-logarithm also supports its slow growth. Third, higher bases give smaller iterated logarithms, and $\log^* x$ is well defined for any base greater than $\exp\left\{\frac{1}{e}\right\}$. This implies that for any base $b \geq 2$, the iterated logarithm will grow even more slowly for higher pseudoprime bases.

Conjecture 3.1.3. *For any fixed $k \geq 3$, let $C_k(x)$ denote the number of k-prime Carmichael numbers up to x, and let $C(x)$ denote the Carmichael number counting function. Asymptotically,*

$$\frac{C_k(x)}{C(x)} = o(1). \tag{3.1.7}$$

3.2 Support for Conjecture 3.1.1 and Conjecture 3.1.3

So far, the claim established by Conjecture 3.1.1 is not yet borne out by the data in Table 6. We believe that the ratio $\frac{\mathscr{P}_{b,2}(x)}{\mathscr{P}_2(x)}$ will approach 0, but may do so slowly at first. On the other hand, it appears that the ratio $\frac{C_3(x)}{C(x)}$ in Table 7 rapidly approaches 0, thereby supporting Conjecture 3.1.3.

Table 6: Values of $\frac{\mathscr{P}_{b,2}(x)}{\mathscr{P}_2(x)}$

Bound	$\mathscr{P}_{b,2}(x)$	$\mathscr{P}_2(x)$	$\frac{\mathscr{P}_{b,2}(x)}{\mathscr{P}_2(x)}$
10^3	0	3	0.00
10^4	11	22	0.50
10^5	34	78	0.44
10^6	107	245	0.44
10^7	311	750	0.41
10^8	880	2057	0.43
10^9	2455	5597	0.44
10^{10}	6501	14884	0.44
10^{11}	17207	38975	0.44
10^{12}	46080	101629	0.45
10^{13}	123877	264239	0.47
10^{14}	334567	687007	0.49
10^{15}	915443	1801533	0.51
10^{16}	2520503	4744920	0.53
10^{17}	7002043	12604009	0.56

Furthermore, Pomerance, Selfridge, and Wagstaff's famous results [10] support both conjectures. In Conjecture 1 of their paper, they believe that for each $\epsilon > 0$, there is an $x_0(\epsilon)$ such that for all $x \geq x_0(\epsilon)$,

$$C(x) > x \cdot \exp\Big\{\frac{-\{2+\epsilon\}\log x \cdot \log^{(3)} x}{\log^{(2)} x}\Big\}. \tag{3.2.1}$$

Pomerance, Selfridge, and Wagstaff [10] show that $\mathscr{P}_{b,k}(x) \leq O_k(x^{2k/(2k+1)})$. If (3.2.1) is true, then the pseudoprimes "with exactly k prime factors form a set of relative density 0 in the set of all [pseudoprimes]" [10]. Similarly, in Theorem 7 of Granville and Pomerance [2], it is proven that $C_k(x) \leq x^{2/3+o_k(1)}$, and if (3.2.1) holds, "then for each k, $C_k(x) = o(C(x))$" [10].

Interestingly, we can also support the statements in Conjecture 3.1.1 and Conjecture 3.1.3 by relating them to their composite superset. Let the number of composites $\leq x$ with k distinct prime factors be denoted by $\pi_k(x)$ and let the number of composites $\leq x$ with k prime factors (not necessarily distinct) be represented by $\tau_k(x)$. Hence, we can prove upper and lower bounds for $\pi_k(x)$. In 22.18.2 of Hardy and Wright [14] for $k \geq 1$,

$$k!\pi_k(x) \leq \Pi_k(x) \leq k!\tau_k(x), \tag{3.2.2}$$

where $\Pi_k(x) = \frac{\vartheta_k(x)}{\log x} + O(\frac{x}{\log x})$ in 22.18.5. In 22.18.24, since $\vartheta_k(x) = \Pi_k(x)\log x - \displaystyle\int_2^x \frac{\Pi_k(x)}{t}dt \sim kx(\log^{(2)} x)^{k-1}$ for $k \geq 2$ and $\displaystyle\int_2^x \frac{\Pi_k(x)}{t}dt = O(x)$, $\Pi_k(x) \sim \frac{kx(\log^{(2)} x)^{k-1}}{\log x}$. As a result, it follows that

$$\pi_k(x) \leq (1+o(1))\frac{x(\log^{(2)} x)^{k-1}}{(k-1)!\log x}. \tag{3.2.3}$$

In the same respect, a lower bound for π_k can be formulated. In 22.18.3 it is proven that,

$$\tau_k(x) - \pi_k(x) \leq \sum_{p_1p_2\cdots p_{k-1}^2 \leq x} 1 \leq \sum_{p_1p_2\cdots p_{k-1} \leq x} 1 := \Pi_{k-1}(x). \tag{3.2.4}$$

Table 7: Values of $\frac{C_3(x)}{C(x)}$

Bound	$C_3(x)$	$C(x)$	$\frac{C_3(x)}{C(x)}$
10^3	1	1	1.00
10^4	7	7	1.00
10^5	12	16	0.75
10^6	23	43	0.53
10^7	47	105	0.45
10^8	84	255	0.33
10^9	172	646	0.27
10^{10}	335	1547	0.22
10^{11}	590	3605	0.16
10^{12}	1000	8241	0.12
10^{13}	1858	19279	0.096
10^{14}	3284	44706	0.073
10^{15}	6083	105212	0.058
10^{16}	10816	246683	0.044
10^{17}	19539	585355	0.033
10^{18}	35586	1401644	0.025
10^{19}	65309	3381806	0.019
10^{20}	120625	8220777	0.015
10^{21}	224763	20138200	0.011

Since $\pi_k(x) \geq \tau_k(x) - \Pi_{k-1}(x)$ and $\pi_k(x) \geq \frac{\Pi_k(x)}{k!} - \Pi_{k-1}(x)$,

$$\pi_k(x) \geq O\left(\frac{x(\log^{(2)} x)^{k-1}}{(k-1)!\log x}\right) - \frac{(k-1)x(\log^{(2)} x)^{k-2}}{\log x} + O\left(\frac{x}{\log x}\right).$$

We can improve the upper bound given in (3.2.3) to an equality,

$$\pi_k(x) \sim \frac{x(\log^{(2)} x)^{k-1}}{(k-1)!\log x}. \tag{3.2.5}$$

By the Erdős-Kac Theorem [22], we can formulate the probability that a number near x has k distinct prime factors using the fact that these numbers are distributed with a mean and variance of $\log^{(2)} x$. Hence, setting $\log^{(2)} x$ as the λ of the Poisson distribution $\mathrm{P}(k;\lambda)$ and taking its limit for any fixed k,

$$\lim_{x\to\infty} \mathrm{P}(k;\lambda) = \lim_{x\to\infty} \frac{(\log^{(2)} x)^{k-1}\exp\{-\log^{(2)} x\}}{(k-1)!} = 0, \tag{3.2.6}$$

where the asymptotic error bound is given by $O(\frac{1}{\log^{(2)} x})$ [17]. However, we caution the reader to consider that just because the probability of a general composite near x having k distinct prime factors goes to 0, does not necessarily fully prove that this probability will hold for either $\mathscr{P}_{b,k}(x)$ or $C_k(x)$.

3.3 An Alternate Conjecture

From Conjecture 3.1.1 and Conjecture 3.1.3, it is evident that the k-prime pseudoprimes and the k-prime Carmichael numbers are much more sparsely distributed than the set of all pseudoprimes

and Carmichael numbers, respectively. We hypothesize that if k is minimized for both the k-prime pseudoprimes and the k-prime Carmichael numbers, then the ratios $\frac{\mathscr{P}_{b,2}(x)}{\mathscr{P}_b(x)}$ and $\frac{C_3(x)}{C(x)}$ will roughly achieve the same values for large enough x. We also recommend using the minimum number of distinct prime factors for both the pseudoprimes and the Carmichael numbers because first, there is no overlap between the three-prime Carmichael numbers and two-prime pseudoprimes and second, the distinct prime factors cannot be arbitrarily chosen. This idea leads us to believe that,

$$C(x) \sim \frac{C_3(x)\mathscr{P}_b(x)}{\mathscr{P}_{b,2}(x)}.$$

As a result, assuming Conjecture 2.1.3, Conjecture 2.1.4, Conjecture 2.2.4, and Conjecture 1.0.2, the amount of Carmichael numbers $\leq x$ given by the counting function $C(x)$ is conjectured to be for x sufficiently large,

$$C(x) \sim \frac{\psi' x^{\frac{5}{6}}}{\log x \cdot L(x)} \sim \frac{\psi_1' x^{\frac{1}{2}} \log^2 x \displaystyle\int_2^{x^{\frac{1}{3}}} \frac{dt}{\log^3 t}}{L(x)}, \tag{3.3.1}$$

where

$$\psi' = \frac{\tau_3}{C} \tag{3.3.2}$$

and

$$\psi_1' = \frac{\tau_3}{27C}. \tag{3.3.3}$$

In Table 8, the computed values of ψ' and ψ_1' up to 10^{21} are given. Hence, not only does Corol-

Table 8: Values of ψ' and ψ_1'

Bound	$C(x)$	$\frac{69.51x^{\frac{5}{6}}}{\log x \cdot L(x)}$	$\frac{2.57x^{\frac{1}{2}} \log^2 x \int_2^{x^{\frac{1}{3}}} \frac{dt}{\log^3 t}}{L(x)}$	ψ'	ψ_1'
10^3	1	301.95	1092.82	0.2302	0.0024
10^4	7	594.43	2835.17	0.8185	0.0063
10^5	16	1316.29	6640.29	0.8449	0.0062
10^6	43	3131.53	14806.24	0.9545	0.0075
10^7	105	7826.17	32411.27	0.9326	0.0083
10^8	255	20282.91	71150.56	0.8739	0.0092
10^9	646	54070.80	159157.24	0.8305	0.0104
10^{10}	1547	147451.71	367012.00	0.7293	0.0108
10^{11}	3605	409716.38	878601.38	0.6116	0.0105
10^{12}	8241	1156637.85	2188667.23	0.4953	0.0097
10^{13}	19279	3309970.24	5664006.88	0.4049	0.0087
10^{14}	44706	9585268.36	15162465.67	0.3242	0.0076
10^{15}	105212	28049810.91	41763706.96	0.2607	0.0065
10^{16}	246683	82852448.55	117743387.56	0.2070	0.0054
10^{17}	585355	246785788.13	338238941.70	0.1649	0.0044
10^{18}	1401644	740679196.52	986503770.93	0.1315	0.0037
10^{19}	3381806	2238429061.23	2913197684.15	0.1050	0.0030
10^{20}	8220777	6807841639.58	8692508977.60	0.0839	0.0024
10^{21}	20138200	20826296835.28	26167265004.43	0.0672	0.0020

lary 1.0.3 fit the proven bounds for $C(x)$ given in Theorem 2.2.1 and Theorem 2.2.3, but both ψ' and ψ_1' appear to be approaching constant values. However, there are several reasons as to why Corollary 1.0.3 may not be necessarily borne out by the data in the above table. For instance, the infinite series κ_3 is slowly convergent, and it is not until 10^{24} that κ_3 appears to approach its estimated value of 2087.5. However, the primary source of inaccuracy is due to Conjecture 2.1.3. Since Pomerance's conjecture for the distribution of pseudoprimes is applicable for sufficiently large x and pseudoprime counts have only recently been conducted to 10^{17} by Galway and Feitsma, we are not sure how "sufficiently large" x must be for Conjecture 2.1.3 to be an accurate model for pseudoprime distribution. Lastly, x must also be immensely large in order for $\frac{\mathscr{P}_{b,2}(x)}{\mathscr{P}_b(x)} = o(1)$.

3.4 An Improved Heuristic Argument

As mentioned before, Pomerance's heuristic arguments supporting his conjecture in (2.2.2) involve the distribution of smooth numbers. And, if secondary terms exist, then it would be worthwhile to sharpen these heuristics to produce a conjecture for $C(x)$. Let $\Psi(x,y)$ denote the number of y-smooth numbers $\leq x$ and let $\Psi'(x,y)$ denote the number of primes $p \leq x$ for which $p-1$ is squarefree and its prime factors are $\leq y$ [10]. It is conjectured in [5] that for $\exp\{\frac{1}{2}(\log x)^{\frac{1}{2}}\} \leq y \leq \exp\{(\log x)^{\frac{1}{2}}\}$,

$$\frac{1}{x}\Psi(x,y) \sim \frac{1}{\pi(x)}\Psi'(x,y). \tag{3.4.1}$$

If $0 < \alpha < 1$, it is well-known [1] that

$$\Psi\left(x, \exp\{c(\log x)^{\alpha}(\log^{(2)} x)^{\beta}\}\right) = x\exp\{-\{(1-\alpha)/c + o(1)\}(\log x)^{1-\alpha}(\log^{(2)} x)^{1-\beta}\}. \tag{3.4.2}$$

Concerning Carmichael numbers, we are interested in the case for which $\alpha = \frac{1}{2}$, $\beta = 0$, and $c = 1$. Hence,

$$\Psi\left(x, \exp\{(\log x)^{\frac{1}{2}}\}\right) = x\exp\{-\{1/2 + o(1)\}(\log x)^{\frac{1}{2}}(\log^{(2)} x)\}. \tag{3.4.3}$$

From (3.4.1) and (3.4.3), we make the following

Conjecture 3.4.1. *For* $\exp\{\frac{1}{2}(\log x)^{\frac{1}{2}}\} \leq y \leq \exp\{(\log x)^{\frac{1}{2}}\}$,

$$\Psi'(x,y) = \pi(x)\exp\{-\{1/2 + o(1)\}(\log x)^{\frac{1}{2}}(\log^{(2)} x)\}. \tag{3.4.4}$$

Let $A(x)$ denote the product of the primes $p \leq \log x/(\log^{(2)} x)^2$. Thus, $A(x) < x^{2/\log^{(2)} x}$ as in [10]. If we allow $r_1, \ldots, r_q$ to be the primes in the interval $\left(\log x/(\log^{(2)} x)^2, (\log x)^{\log^{(2)} x}\right)$ with $r_i - 1 \mid A(x)$. By Conjecture 3.4.1 we have for x sufficiently large,

$$q = \pi\left((\log x)^{\log^{(2)} x}\right)\exp\{-\{1 + o(1)\}\log^{(2)} x \log^{(3)} x\}. \tag{3.4.5}$$

Let $m_1, \ldots, m_N$ be the squarefree composite integers $\leq x$ composed of r_i and let

$$l = \left\lfloor \log x/(\log^{(2)} x)^2 \right\rfloor.$$

As discussed in [10], we have

$$N \geq \binom{q}{l} \geq \left(\frac{q}{l}\right)^l. \tag{3.4.6}$$

As a result,

$$N \geq \left(\frac{(\log^{(2)} x)^2 \pi((\log x)^{\log^{(2)} x}) \exp\{-\{1+o(1)\} \log^{(2)} x \log^{(3)} x\}}{\log x} \right)^{\log x/(\log^{(2)} x)^2}. \tag{3.4.7}$$

Since Euler's φ function and Carmichael's λ function are virtually the same, the lower bound in (3.4.7) should be applicable to $C(x)$. In fact, from the values of $a(x)$ in Table 3 and the precision of Conjecture 3.4.1, we have reason to believe that this result is asymptotically close to the actual value of $C(x)$.

4 One-Parameter Quadratic-Base Pseudoprimes: A Sidenote

As mentioned earlier, the discovery of Carmichael numbers demonstrated the fallability of Fermat's primality test therefore lending to the development of efficient probabilistic primality tests. Baillie, Pomerance, Selfridge, and Wagstaff [10] [23] have determined a primality test that is an amalgamation of the Miller-Rabin test and a Lucas test. However, even though Pomerance [7] presented a heuristic argument that the number of counter-examples up to x was $\gg x^{1-\epsilon}$ for $\epsilon > 0$, we have not been able to find any counter-examples up to 10^{17}. In fact, no precise probability of error has been given about this test either [30].

Grantham [18] has also provided a probable prime test known as the RQFT that has a *known* worst-case probability of error of 1/7710 per iteration.

An even stronger test known as the One-Parameter Quadratic-Base Test (OPQBT) has been given by Zhang [30], and is a version of the Baillie-PSW test that not only has a known probability of error but is more efficient than the RQFT except for a thin set of cases. We let $u(\neq \pm 2) \in \mathbb{Z}$, let $T_u = T \pmod{T^2 - uT + 1}$, and define the ring associated with parameter u as

$$R_u = \mathbb{Z}[T]/(T^2 - uT + 1) = \{a + bT_u : a, b \in \mathbb{Z}\}.$$

We then define an odd integer $n > 1$ as an OPQBT pseudoprime for $0 \leq u < n$ with

$$\epsilon = \left(\frac{u^2 - 4}{n} \right) \in \{-1, 1\},$$

where in the ring R_u, n must pass

$$T_u^{n-\epsilon} \equiv 1 \pmod{n}. \tag{4.0.8}$$

Moreover, n is defined as an OPQBT strong pseudoprime if for some $i = 0, 1, \cdots, k-1$, either

$$T_q^u \equiv 1 \pmod{n}, \tag{4.0.9}$$

or

$$T_u^{2^i q} \equiv -1 \pmod{n}, \tag{4.0.10}$$

in which for q odd, $n - \epsilon = 2^k q$ [30].

We have verified that there are no OPQBT pseudoprimes up to 10^{17}. Let the counting function $\mathscr{O}(x)$ denote that number of OPQBT pseudoprimes $\leq x$ and let $\mathscr{SO}(x)$ denote the number of strong OPQBT pseudoprimes $\leq x$. The best upper bound we are able to prove is

$$\mathscr{SO}(x) \leq \mathscr{O}(x) \leq x \cdot L(x)^{\frac{-1}{2}}, \tag{4.0.11}$$

since an upper bound on the pseudoprimes is applicable to an upper bound on the OPQBT pseudoprimes and strong OPQBT pseudoprimes.

Based upon Erdős' construction [21] and Pomerance's heuristics [7], in the interval $\left[H, H^j\right]$, for any fixed $j > 4$ and H sufficiently large, there are most likely $\exp\{H^2(1 - 4/j)\}$ counter-examples to Zhang's primality test, meaning that there are at least $x^{1-4/j}$ counter-examples below $x = \exp\{H^2\}$. Thus, for arbitrary j, the number of counter-examples to the OPQBT becomes generalized to $\gg x^{1-\epsilon}$ for $\epsilon > 0$. In other words, there are infinitely many counter-examples to Zhang's OPQBT.

Acknowledgements

I would like to express my gratitude to Charles R. Greathouse IV who provided invaluable guidance in the direction of this paper; Carl Pomerance who offered helpful comments on this paper in its initial stages; Harvey Dubner who provided guidance on the experimental aspects of this paper; Johan B. Henkens who helped me count the base 2 pseudoprimes and 2-strong pseudoprimes up to 10^{17} using Galway and Feitsma's data; Kazimierz Szymiczek who clarified Theorem 2.1.7; and David H. Low who assisted me with formatting and typesetting errors. I would also like to thank William F. Galway for sharing his viewpoints regarding the pseudoprimes with k distinct prime factors.

References

[1] A. Granville, *Smooth numbers: computational number theory and beyond*, Mathematical Sciences Research Institute Publications, **44** (2008) 1–58. `http://www.math.leidenuniv.nl/~psh/ANTproc/09andrew.pdf`.

[2] A. Granville and C. Pomerance, *Two Contradictory Conjectures Concerning Carmichael Numbers*, Math. Comp. **71** (2001): 883–908.

[3] A. Granville, C. Pomerance, and W. R. Alford, *There are Infinitely Many Carmichael Numbers*, Ann. of Math. **140** (1994): 703–722.

[4] A. Korselt, *Problème chinois*, L'intermèdiaire de mathématiciens **6** (1899): 142–143.

[5] C. Pomerance, *On the Distribution of Pseudoprimes*, Math. Comp. **37** (1981): 587–93.

[6] C. Pomerance, *A New Lower Bound for the Pseudoprime Counting Function*, Illinois J. Math. **26** (1982): 4–9.

[7] C. Pomerance, *Are there counter-examples to the Baillie-PSW primality test?*, Dopo Le Parole aangeboden aan Dr. A. K. Lenstra (H. W. Lenstra, jr., J. K. Lenstraand, P. Van Emde Boas, eds.), Amsterdam, 1984.

[8] C. Pomerance and R. Crandall, *Prime Numbers: A Computational Perspective*, New York: Springer, 2005.

[9] C. Pomerance and D. M. Gordon, *The Distribution of Lucas and Elliptic Pseudoprimes*, Math. Comp. **57** (1991): 825–38.

[10] C. Pomerance, J. L. Wagstaff, and S. S. Wagstaff, jr, *Pseudoprimes to* $25 \cdot 10^9$, Math. Comp. **35** (1980): 1003–026.

[12] D. Shanks, *Solved and unsolved problems in number theory*, 3rd ed., Chelsea, New York, 1985.

[13] Emails exchanged between A. Nayebi and C. Pomerance.

[14] E. M. Wright and G. H. Hardy, *An Introduction to the Theory of Numbers*, Oxford: Clarendon P, Oxford UP, 1998.

[15] G. Harman, *On the number of Carmichael numbers up to x*, Bull. Lond. Math. Soc. **37** (2005): 641–650.

[16] G. H. Davies and J. M. Chick, *The Evaluation of κ_3*, Math. Comp. **77** (2008): 547–550.

[17] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge: Cambridge UP, 1995.

[18] J. Grantham, *A probable prime test with high confidence*, J. Number Theory, **72** (1998) 32–47.

[19] K. Szymiczek, *On Pseudoprimes which are Products of Distinct Primes*, Amer. Math. Monthly **74** (1967): 35–37.

[20] P. Erdős, *On the Converse of Fermat's Theorem*, Amer. Math. Monthly, **56** (1949): 623–24.

[21] P. Erdős, *On pseudoprimes and Carmichael numbers*, Publ. Math. Debrecen **4** (1956): 201–206.

[22] P. Erdős and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math **62** (1940): 738–742.

[23] R. Baillie and S. S. Wagstaff, jr., *Lucas pseudoprimes*, Math. Comp. **35** (1980): 1391–1417.

[24] R. D. Carmichael, *Note on a new number theory function*, Bull. Amer. Math. Soc. **16** (1910): 232–238.

[25] R. G.E. Pinch, *The Carmichael Numbers up to* 10^{21}, Proceedings Conference on Algorithmic Number Theory, Turku, May 2007. Turku Centre for Computer Science General Publications 46, edited by Anne-Maria Ernvall-Hytönen, Matti Jutila, Juhani Karhumäki and Arto Lepistö.

[26] R. G.E. Pinch, *The Carmichael Numbers up to* 10 *to the* 21, Eighth Algorithmic Number Theory Symposium ANTS-VIII May 17–22, 2008 Banff Centre, Banff, Alberta (Canada).

[27] W. F. Galway, *The Pseudoprimes below* 2^{64}, Simon Fraser University, 2002, `http://oldweb.cecm.sfu.ca/pseudoprime/psp-search-slides.pdf`.

[28] W. F. Galway, *Tables of pseudoprimes and related data*, 2002, `http://oldweb.cecm.sfu.ca/pseudoprime/psp1e15.gz`.

[29] W. F. Galway, *Research Statement*, 2004, `http://www.math.uiuc.edu/~galway/research-statement.pdf`.

[30] Z. Zhang, *A one-parameter quadratic-base version of the Baillie-PSW probable prime test*, Math. Comp. **71** (2002): 1699–1734.