On Carmichael numbers with 3 factors and the strong pseudoprime test

Warren D. Smith*
wds@research.nj.nec.com

August 18, 1997

Abstract — **The well known "strong pseudoprime test" has its highest probability of error ($\approx 1/4$) when the numbers being tested are certain Carmichael numbers with 3 prime factors. We present a nonrigorous plausibility argument that the count $C_3(x)$ of 3-factor Carmichaels up to x, is asymptotically $C_3(x) \sim K x^{1/3} (\ln x)^{-3}$ where $K \approx 478 \pm 110$ is an absolute constant given by**

$$K = 3^3 \sum_{1 \le a < b < c} \frac{1}{(abc)^{4/3}} \prod_p F_{abc}(p) \quad (1)$$

where the sum is over 3-tuples (a, b, c) of *pairwise relatively prime* sorted distinct positive integers, the product is over *prime* p, and $F_{abc}(p)$ is defined as follows. Let y, $y \in \{0, 1, 2, 3\}$, be the number of distinct values among the *nonzero* $\{a, b, c\}$ mod p. Then

$$F_{abc}(p) = \frac{(p-y)p^2}{(p-1)^3}. \quad (2)$$

The 3-factor Carmichaels which cause error probability $\sim 1/4$ in the strong pseudoprime test also should obey a law of the same form, but with a different constant $K' \approx 383 \pm 81$, defined by the same formula as K, except that in the sum, a, b, and c are now also required to be *odd*.

It is even more plausible (since fewer conjectural assumptions are required) that $C_3(x) x^{-1/3} (\ln x)^3$ is bounded between two constants.

The only other numbers which can cause a probability of error $> 1/8$ in the strong pseudoprime test are certain numbers with only 2 prime factors and certain prime powers. However if these cases are somehow known to apply, then we show how to improve the strong pseudoprime test so that its probability of error on N is $O(1/\sqrt{\ln N})$.

Keywords — Carmichael numbers, Fermat test, pseudoprimes

1 Notation

An odd composite integer N is a "Carmichael number" if $a^N \equiv a \bmod N$ for all integers a. It is a "3-factor Carmichael" if it has exactly 3 prime factors. $a|b$ means "a divides b" and $a \nmid b$ means "a does not divide b." In expressions $\sum_p$ or $\prod_p$, we mean for p to range over the primes $\{2, 3, 5, 7, 11, ...\}$.

2 Introduction

Given an odd number N, is it prime?

The famous "strong pseudoprime test," often ascribed to Gary Miller 1976, or Atkin & Larson 1982, (other claimants include M.Rabin, L.Monier, and J.Selfridge, and really, it was essentially known 100 years ago) is as follows.

Strong pseudoprime test (for odd $N \ge 3$)

1. Choose a random integer a with $1 < a < N$.
2. Suppose $N - 1 = k2^s$ with k odd. Let $b = a^k \bmod N$.
3. If $b = 1$, or if $b^{2^m} \bmod N = -1$ for some m with $0 \le m < s$, then "N is probably prime (evidence: a)."
4. Otherwise, "N is composite and a proves it."

If N is prime, this procedure will always output "N is probably prime." If N is composite, then the probability (and this probability arises purely from the random choice of a, and does not involve any probabilistic assumption about the input N) that the output is "N is composite" is $\ge 3/4$. Although a 1/4 probability of getting a misleading answer isn't acceptable, this probability may be decreased to 4^{-m} by m repeated runs using independently generated random bits each time.

The strong pseudoprime test is an improvement over the simpler "Fermat primality test" that $a^N \equiv a \bmod N$, because unfortunately, the Fermat test is passed by an infinite number of composite numbers N, called Carmichael numbers, for all a. I.e., if N is Carmichael the Fermat test will fail with probability 1. If N has ω prime factors, then the strong pseudodprime test will only fail (for a random a) with probability $\le 2^{1-\omega}$, because there will be 2^ω square roots of 1 (mod N), all arising equiprobably, but only 2 of them are the allowed values ± 1.

*NEC Research Institute, 4 Independence way, Princeton NJ 08544

This focuses our attention on the worst composite numbers N – those which cause the strong pseudoprime test to fail with a high probability.

It was shown by Damgard et al. 1993 that a failure probability $> 1/8$ was possible only if N was certain Carmichael numbers with 3 prime factors, or certain numbers with 2 prime factors, or a prime power. The latter two cases will be dealt with in the final section.

3 Carmichael numbers with 3 prime factors

N is a "Carmichael number with 3 prime factors" if $N = pqr$ where $p < q < r$ are distinct odd primes and

$$\mathrm{LCM}(p-1, q-1, r-1) | N-1. \tag{3}$$

The probability of failure in the strong pseudoprime test is $\sim 1/4$ if $p-1$, $q-1$, and $r-1$ all contain the same number of factors of 2 and p is large.

Let $C_3(x)$ be the number of 3-factor Carmichaels $\leq x$.

It is easy to verify the (well known) fact that

$$N = (6m+1)(12m+1)(18m+1) \tag{4}$$

is Carmichael if all 3 factors are prime. (Verify that $6m$, $12m$, and $18m$ all divide $N-1$.) If we make the usual heuristic assumption that a number $y = 6m+1$ is prime with "probability" $3/\ln y$ and naively assume that all these probabilities are "independent," then we would conclude (nonrigorously) that

$$C_3(x) \geq \frac{81 \cdot 6^{2/3}}{4} \frac{x^{1/3}}{(\ln x)^3}. \tag{5}$$

Note $81 \cdot 6^{2/3}/4 \approx 66.86$. But in fact, it is wrong to assume "independence" of the 3 primality "probabilities," because if $p = 6m+1$ is prime, then that makes it less likely that q is prime, since, e.g. it is more likely that $q = 12m+1 = 2p-1$ is divisible by 5: since $5 \not| 2p$, there is a $1/4$, not a $1/5$, probability that $5|q$, for example. Similarly if p and $q = 2p-1$ are prime, then $r = 18m+1 = 3p-2$ is divisible by 5 with probability $1/3$, not $1/4$ or $1/5$. This thinking (but considering all primes, not just "5") leads (similarly to the famous Hardy-Littlewood 1922 heuristic derivation of the count of twin primes) to the more justified (but still nonrigorous) bound

$$C_3(x) \geq \frac{81 \cdot 6^{2/3}}{4} K_5 \frac{x^{1/3}}{(\ln x)^3} \approx 42.4 \frac{x^{1/3}}{(\ln x)^3} \tag{6}$$

where

$$K_m = \prod_{p \geq m} \frac{(p-3)p^2}{(p-1)^3}, \quad K_5 \approx 0.635. \tag{7}$$

In fact, the number of Carmichaels of the form (EQ 4) below 10^{18} is 783 ($\approx$ 2.2% of the Carmichaels in this range), which is about midway between the two right hand sides of (EQ 6) and (EQ 5) above with $x = 10^{18}$ (respectively 596 and 939, rounded to the nearest integer). I believe that the discrepancy between 596 and 783 is simply due to the fact that 10^{18} is not large enough.

In fact, it is possible to construct larger families of 3-factor Carmichael numbers. For example $N = (2m+1)(4m+1)([4z+2]m+1)$ where all three parenthesized factors are distinct primes and $2z+1|3+4m$. Also $N = (3m+1)(9m+1)([18z-3]m+1)$ is a Carmichael whenever all 3 factors are prime and $6z-1|1+12m$. This latter family has all prime factors p, q, and r such that $p-1$, $q-1$, and $r-1$ all contain the same power of 2, so that it will achieve the maximal failure probability $\sim 1/4$ in the strong pseudoprime test.

These are both special cases of the following family with two free parameters a and m, and one constrained parameter z:

$$N = (am+1)(a^2m+1)([az-1]am+1) \tag{8}$$

is Carmichael if all 3 factors are distinct primes and $az-1|a^2m+a+1$.

To count these, choose z and then a and then m. The number of ways to choose $z \leq x$ is x. The number of ways to choose a so that $a^5z < x$ is $\sim (x/z)^{1/5}$. Finally, the number of ways to choose m so that $za^5m^3 < x$ and so that $az-1|a^2m+a+1$ is is $\sim (xa^{-5}z^{-1})^{1/3}/(az-1)$ if $x/(a^8z^4) \to \infty$. If we apply this estimate in the regime $x > a^5z$ and $x \to \infty$, in which it is less justified, we estimate that the total count of suitable triplets (a, z, m) is thus asymptotically

$$\sum_{z \leq x} \sum_{a < (x/z)^{1/5}} (\frac{x}{a^5 z})^{1/3} \frac{1}{az-1} \tag{9}$$

which is

$$\sim x^{1/3} \sum_{z \geq 1} \sum_{a \geq 1} \frac{1}{a^{5/3} z^{1/3}} \frac{1}{az-1} \sim \theta x^{1/3} \tag{10}$$

for some absolute constant θ. This has not yet taken into account the "probability" that the 3 factors are each prime. It is possible but exceedingly tedious to express this probability conjecturally exactly in terms of sums of products over primes. It is clear that the result is that $C_3(x) \gtrsim Cx^{1/3}/(\ln x)^3$ for some absolute constant C, and the simplest method of estimating C is just to count all the Carmichaels of the form (EQ 8) below 10^{18} and find that there are 8623 of them ($\approx$ 24% of all Carmichaels in this region), which leads to the estimate $C \approx 133$, probably accurate to about ± 30. One may also get a quick and rough estimate of this number by approximating $az-1 \approx az$ in which case the double sum in (EQ 10) is $\zeta(4/3)\zeta(8/3) \approx 4.624$, and then estimating the probability that all three factors are primes as $\approx 3^3/(\ln x)^3$ suggests that $C \approx 3^3\zeta(4/3)\zeta(8/3) \approx 125$. This agrees reasonably with 133.

Notice that the additional free parameter a and the constrained parameter z in (EQ 8) did not cause the presumed count of 3-factor Carmichaels up to x, to increase by more than a constant factor versus just using the 1-parameter family (EQ 4).

Now we'll present an even more general construction, which in fact yields a presumed asymptotic form for $C_3(x)$, not just a lower bound. As was pointed out in Damgard et al. 1993, all 3-factor Carmichael numbers $N = pqr$ may (by letting $g = GCD(p-1, q-1, r-1)$) be written in the form $N = (ga+1)(gb+1)(gc+1)$ where a, b, c are *pairwise* relatively prime[1], and $a < b < c$. If a, b, c are known, then g mod abc is uniquely determined by

$$a|b+c+bcg, \quad b|a+c+acg, \quad c|b+a+bag \tag{11}$$

via the Chinese remainder theorem. (These congruences arise from (EQ 3).)

Now, suppose we make the conjecture that the solutions g of (EQ 11), among pairwise relatively prime 3-tuples a, b, c, (when $1 \le a < b < c < z$, or when $1 \le a < b < c$ with $abc < z$, as $z \to \infty$) are *uniformly distributed* mod abc, or, more precisely, that $G = (g \bmod abc)/(abc)$ is asymptotically uniformly distributed[2] in the real interval $(0, 1)$.

We'll also conjecture that the probability that the three factors $p = ag+1$, $q = bg+1$, and $r = cg+1$ that result are all prime, is of order $[(\ln p)(\ln q)(\ln r)]^{-1}$. More precise estimates for this primality probability will be discussed later; for the moment let us merely point out that for p, q, and r all to be odd primes, it is necessary that g be even (because at least 2 of a, b, c will be odd), which automatically cuts the count by a factor of 2.

Case 1: Suppose $g > abc$. In that case, the number of g with $g^3 abc \le x$, g in a particular residue class mod abc, and g even, and $g \ge abc$, is asymptotic to $x^{1/3}(abc)^{-4/3}/2$ when $x/(abc)^2 \to \infty$. In fact, our uniformity conjecture implies validity (as an expectation value) in the regime defined by the much weaker restrictions $x > abc$ and $x \to \infty$; also, we may even avoid depending on the uniformity conjecture (cf. footnote 2) if we redefine "case 1" (also "case 2" will need redefinition; it is the complement of case 1) to be "$g > abc\kappa$" where κ is some constant which may be made arbitrarily large (instead of the current definition, $\kappa = 1$) so that the fractional error in our counting formula will be at most $1/\kappa$, i.e. arbitrarily small. Then the number of eligible 4-tuples (a, b, c, g), is asymptotic to

$$\frac{x^{1/3}}{2} \sum_{1 \le a < b < c} (abc)^{-4/3} \tag{12}$$

where the sum is over triples a, b, c with $a < b < c$ and pairwise relatively prime. The value of the sum is $\approx 2.55 \pm 0.3$.

Case 2: suppose $1 \le g \le abc$. Then g is uniquely determined by a, b, c. We may suppose $2^{-s}abc < g \le 2^{1-s}abc$ for some s, $1 \le s \le \lg x$. In that case, $abc \le (2^{3s}x)^{1/4} = Z$, so that the number of suitable (a, b, c) is at most the number of ways to write all the numbers $\le Z$ as a product of 3 factors, i.e. $O(Z^{1+\epsilon})$ for any $\epsilon > 0$. Now, by our uniform distribution assumption, this happens with probability 2^{-s}, so the total number of 3-tuples (a, b, c) that result is $O(x^{1/4+\epsilon} \sum_{s \ge 1} 2^{(\epsilon - 1/4)s}) = O(x^{1/4+\epsilon})$. This is asymptotically negligible[3] in comparison to case 1.

Now, in case 1, we have not yet used any estimate of the probability that all three of $ag+1$, $bg+1$, and $cg+1$ are prime, assuming g is even. Again, we could naively assume this probability is $\ge 3^3 2^3 (\ln x)^{-3}$, (because, e.g.: if $fgh = x$ and $1 \le f < g < h$ then $\ln f \ln g \ln h \le (\ln x)^3/3^3$; the "$2^3$" is because we have already assumed g is even) which would lead to the belief that

$$C_3(x) \gtrsim 275 x^{1/3} (\ln x)^{-3}. \tag{13}$$

(Here "275" is an estimate accurate to about ± 30.) However, this estimate is naive in several ways. First, Dirichlet's theorem tells us that the "probability" that a number n, where $n \equiv 1 \bmod m$, is prime, is $m/[\phi(m) \ln n]$, so we presumably should insert the correction factors $a/\phi(a)$, $b/\phi(b)$, and $c/\phi(c)$, in each summand. (This would increase the "292 ± 10" to about 659 ± 200, but this still is not right.) Second, the 3 primality probabilities are not "independent," for the same reason as in our discussion of (EQ 4).

The right estimate of the probability that $ag+1$, $bg+1$, and $cg+1$ are simultaneously prime, assuming g is random (and a, b, c are fixed and pairwise relatively prime) is

$$\frac{1}{\ln(ag+1)\,\ln(bg+1)\,\ln(cg+1)} \prod_p F_{abc}(p) \tag{14}$$

where $F_{abc}(p)$ is defined as follows. Suppose there are y, $y \in \{0, 1, 2, 3\}$, distinct values among the nonzero $\{a, b, c\}$ mod p. Then

$$F_{abc}(p) = [\frac{1}{p} + (1 - \frac{1}{p})(1 - \frac{y}{p-1})](1 - \frac{1}{p})^{-3}, \tag{15}$$

because with probability $1/p$, g is divisible by p and hence p cannot divide any of $\{ag+1, bg+1, cg+1\}$, boosting the mutual primeness probability by a factor of $(1-1/p)^{-3}$ above the naive estimate; and with probability $1-1/p$, g is not divisible by p, in which case there is a probability $1 - y/(p-1)$, not $(1-1/p)^3$, that at least one of the three are divisible by p. Again this is reminiscent

[1] Actually, so far it has only been obvious that $GCD(a, b, c) = 1$. However, after a short consideration of (EQ 11), one realizes that in fact, a, b, c must be pairwise relatively prime.

[2] Note: among 3-factor Carmichaels below x, G when $G < 1$ is definitely *not* uniformly distributed in $(0, 1)$, much preferring to be small. This is because smaller G makes it more likely to result in the Carmichael lying below x. This is a different question which has nothing to do with our uniformity conjecture. Anyhow, we do not need the distribution to be precisely uniform; for example any bounded probability density would suffice.

[3] However, in R.Pinch's list of the 35585 3-factor Carmichaels below 10^{18}, 24396 ($\approx 69\%$) are of type 1 and still 11189 ($\approx 31\%$) are of type 2, which seems non-negligible. This is presumably because $x = 10^{18}$ is not large enough, so that $x^{1/4+\epsilon}$ is not sufficiently smaller than $x^{1/3}$. (Anyhow, even the very weak conjecture that type 2 numbers are of asymptotically smaller order than type 1 numbers, would suffice for us.)

of Hardy & Littlewood 1922. Upon simplification, (EQ 15) becomes (EQ 2).

This leads to the estimate

$$C_3(x) \sim \frac{3^3 x^{1/3}}{(\ln x)^3} \sum_{1 \le a < b < c} \frac{1}{(abc)^{4/3}} \prod_p F_{abc}(p) \qquad (16)$$

where again in the sum a, b, c are pairwise relatively prime. Here we are assuming that only an asymptotically negligible fraction of the summands have $ax^\epsilon < b$ for any fixed $\epsilon > 0$, when $x \to \infty$. This allowed us to conclude that the four quantities $\ln(ag+1)$, $\ln(bg+1)$, $\ln(cg+1)$ and $\ln x^{1/3} = (\ln x)/3$ are all approximately equal (to within a factor of $1+\epsilon$), which led to the term $3^3/(\ln x)^3$ in (EQ 16).

Numerical evaluation of the sum yields 17.7, with a numerical uncertainty of about ± 4 so we finally conclude (nonrigorously) that

$$C_3(x) \sim K \frac{x^{1/3}}{(\ln x)^3} \qquad (17)$$

where $K \approx 478 \pm 110$.

R.Pinch showed that $C_3(10^{18}) = 35585$, but it is more relevant for us to consider only the 3-factor Carmichaels of *type 1* (i.e. with $g > abc$), which we've conjectured to be asymptotically 100% of them, but of which there are only 24396 below 10^{18}. This latter figure would lead to an estimate 1737 for the coefficient in (EQ 17) which we estimated from first principles to be $478 \approx 17.7 \times 3^3$, i.e., $3.6\times$ smaller.

What is the cause of this factor-3.6 discrepancy? I do not know. It could be that 10^{18} is too small, or it could be an error in my reasoning – though I do not see one. 3.6 is not as bad as it sounds, if one considers that the term $(\ln x)^3$ in the formula is perhaps $(\eta + \ln x)^3$ to some closer degree of approximation, and $\eta = -13$ would be enough to explain the discrepancy. There are definitely noise terms η of rough size at least $|\eta| \approx 13$ in this position, as may be seen by considering the fact that $C_3(10^3) = 1$. Another reason 3.6 is not as bad as it sounds is our earlier remark that we could redefine "case 1" to be $g > abc\kappa$ where κ is an arbitrarily large constant, and still get the same asymptotic answer. This device was in fact necessary if we wished to avoid the uniformity conjecture. But with this new definition of "type 1," the number of type 1 3-factor Carmichaels below 10^{18} is smaller than 24396, in fact may be made as small as we like, so the numerical disagreement may be forced to vanish! (Somehow unsatisfying, I admit...)

Even if one does not believe the conjectured exact limiting value of the constant factor, it still has been made *very* plausible that $C_3(x)$ is bounded between two positive constants times $x^{1/3}/(\ln x)^3$.

Finally, the 3-factor Carmichaels with a, b, c all odd are of interest because they yield the maximal error probability ($\sim 1/4$ when $\min(a,b,c) \to \infty$) in the strong pseudoprime test. The same asymptotic counting formula conjecturally applies, but with the sum defining K replaced by the same sum but with the requirement that a, b, c all be odd. This sum has value 14.2 ± 3 instead of 17.7 ± 4, leading to $K' \approx 383 \pm 81$ versus $K \approx 478 \pm 110$. It makes intuitive sense that $2K' \approx K$ because we already knew that at most one of $\{a, b, c\}$ could be even, and the chance of that one happening to have odd parity would seem to be roughly 1/2. However, again, the numerical evidence in Pinch's list of the 3-factor Carmichaels below 10^{18} seems to contradict both our reasoning and our intuition. Of the 35585 3-factor Carmichaels below 10^{18}, only 5630 ($\approx 16\%$) have a, b, c all odd – a discrepancy which is hard to explain! On the other hand, $K' x^{1/3}/(\ln x)^3$ is 5379 when $x = 10^{18}$, which is fairly near to 5630.

4 A note on numerical evaluation of certain slowly convergent infinite sums

These were evaluated by truncating the sum at 2^n instead of ∞, for $n \in \{4, 5, 6, 7, 8, 9, 10, 11\}$. The resulting numbers S_n were "accelerated" by use of the "Wynn Epsilon algorithm" (Wynn 1961). The accuracy of the result cannot be vouched for rigorously. The use of certain number theoretic restrictions on the summands (such as that a, b, c be pairwise relatively prime) unfortunately substantially reduces the effectiveness of Wynn-acceleration. Also, sometimes when it appeared that we knew the asymptotic form of the remainder in a sum, we took advantage of this to estimate the infinite sum more directly. It is unfortunate that the sums defining our constants converge so slowly – this is yet another reason it is hard to confirm or refute our conjectures.

Here is one example. Consider the sum S in (EQ 12). Truncating the triple sum by additionally requiring $c < 2^n$ instead of allowing $c \to \infty$, we get the following values of S_n: $S_4 = 0.607629$, $S_5 = 0.909215$, $S_6 = 1.17898$, $S_7 = 1.42615$, $S_8 = 1.63993$, $S_9 = 1.82310$, and $S_{10} = 1.97611$. From the form of the sum we see that the truncation error should be roughly proportional to $2^{-n/3}$ times a constant, and solving for this constant allows one to extrapolate S_∞. Using S_{10} and S_9 yield $S_\infty = 2.565$, while S_9 and S_8 had yielded $S_\infty = 2.528$. Meanwhile the Wynn epsilon algorithm applied to S_4,... S_{10} yields $S_\infty = 2.548$ as its allegedly best estimate, although earlier main-diagonal Wynn estimates were 2.752 and 3.087. So, all this suggests that the true value of the sum is about $2.55 \pm .3$, which is really largely guesswork since all I know for sure is that it is at least 1.976 and in this case I can prove an upper bound of 3.3.

5 Summary of knowledge about 3-factor Carmichael numbers

Very little is *rigorously* known about $C_3(x)$. On the one hand, it is not even known that there are an infinite number of 3-factor Carmichaels. The best known upper bound, Damgard et al. 1993's

$$C_3(x) \le \frac{1}{4} x^{1/2} (\ln x)^{11/4}, \qquad (18)$$

also seems weak.

By using

1. Conjectures about "primeness probabilities" reminiscent of Hardy & Littlewood 1922's heuristic enumeration of twin primes,
2. The conjecture that we may treat the solutions g mod abc of (EQ 11) as if $G = (g \bmod abc)/(abc)$ were a uniformly distributed random variable in the real interval (0,1)
3. The preceding "uniformity conjecture," which seems more doubtful than the Hardy-Littlewood type conjectures, may be replaced by certain weaker sounding conjectures. Even further weakening is possible if we merely want to conclude that $C_3(x)x^{-1/3}(\ln x)^3$ is bounded between two constants, rather than that it tends in the limit to a particular constant K which we express in closed form.

we concluded that

$$C_3(x) \sim Kx^{1/3}(\ln x)^{-3} \tag{19}$$

with K given by (EQ 1).

Examination of the numerical evidence (Pinch's list of all 3-factor Carmichaels below 10^{18} yielded neither a satisfying confirmation nor a convincing rebuttal.

It would be interesting either to confirm our conjectures by enumerating additional 3-factor Carmichaels (say out to 10^{30}, which might be feasible) or demonstrate some flaw in our reasoning.

6 Improving the strong pseudoprime test for N known to have ≤ 2 prime factors

It is readily tested whether N is a perfect power, with a computational cost of the same order as a single iteration of the strong pseudoprime test.

Now suppose that N is a product of two primes $N = pq$, then wlog we may write p and q in the form $p = gc+1$, $q = gd+1$ where $GCD(c, d) = 1$ and $1 \leq c \leq d$. (That is, $g = GCD(p-1, q-1)$.) In this case, the error probability E of a Fermat test (which is an upper bound on the error probability for a strong pseudoprime test) is $E = 1/(cd)$. Write $\Delta = d - c$. Observe that $4Ncd + \Delta^2$ is a square.

Therefore, whenever we know that N has at most 2 prime factors, we may simply test $4Ncd + (c - d)^2$ for squareness for all relatively prime pairs (c, d) of numbers with $1 \leq c \leq d$ and $cd \leq 1/E$. If any squares are found, we may deduce the values dp and cq from the known values of N, Δ and cd via $\sqrt{4Ncd + \Delta^2} = dp + cq$ and $\Delta = d - c = dp - cq$.

On the other hand, if no squares or factorizations are found, we would (assuming we knew N had at most 2 prime factors) then be able to conclude that each iteration of the Fermat or strong pseudoprime test would have error probability below E, allowing us to get more confidence faster than in the regular strong pseudoprime test.

If we use $E = 1/\sqrt{\ln N}$, the additional computational cost won't exceed the cost of just one strong pseudoprime test.

7 References

We have given the original references to the Hardy-Littlewood heuristic count of the number of twin primes up to x, and to Wynn's "epsilon algorithm" for sequence acceleration. However, these topics are also treated in many books respectively on number theory and numerical analysis.

A.O.L. Atkin & R.G. Larson: On a primality test of Solovay and Strassen, SIAM Journal on Computing, 11,4 (1982) 789-791.

I.Damgard, P.Landrock, C.Pomerance: Average case error estimates for the strong probable prime test, Math. of Computation 61 (1993) 177-194.

G.H.Hardy & J.E.Littlewood: On some problems of partitio numerorum: III: on the expression of a number as a sum of primes, Acta Math. 44 (1922) 1-70. (Reprinted in "Collected papers of G.H.Hardy vol I" p.561-630.)

G.L.Miller: Riemann's hypothesis and tests for primality, J.Comput.System Sci. 13 (1976) 300-317.

R.G.E.Pinch: The Carmichael numbers up to 10^{15}, Math. of Comput. 61 (1993) 381-391. Later results (available at Pinch's web site `http://www.math.nus.sg/ rpinch`) now are available reaching to 10^{16}, and with 3-factor Carmichaels to 10^{18}.

P.Wynn: The rational approximation of functions which are formally defined by a power series expansion, Math. of Computation 14 (1960) 147-192.