Abundant Numbers and the Riemann Hypothesis

Keith Briggs

CONTENTS

2000 AMS Subject Classification: 11M26, 11N64, 11Y55

Keywords: Riemann hypothesis, abundant numbers

In this note I describe a computational study of the successive maxima of the relative sum-of-divisors function $\rho(n) := \sigma(n)/n$. These maxima occur at superabundant and colossally abundant numbers, and I also study the density of these numbers. The values are compared with the known maximal order $e^\gamma \log\log(n)$; theorems of Robin and Lagarias relate these data to a condition equivalent to the Riemann hypothesis. It is thus interesting to see how close these conditions come to being violated.

1. INTRODUCTION

Many conjectures equivalent to the Riemann hypothesis (RH) are known [Conrey 05]. Let us refer to an attempt to disprove the Riemann hypothesis by a purely numerical computation as an *attack*. Some may regard the failure of an attack as evidence for the truth of the Riemann hypothesis. Two of the most extensive attacks attempted involve:

1. explicit computations of zeros of the Riemann zeta function: the zetagrid project [Wedeniwski 05] has verified that the first 10^{12} zeros have real part $\frac{1}{2}$, and Gourdon [Gourdon 04] has verified 10^{13} zeros;

2. the Turán inequalities, on which Varga and colleagues have based an attack [Norfolk et al. 92].

It is interesting to speculate which of these attacks is the stronger, in the sense of having the largest "evidence-to-computation-time" ratio. In general, one would like to minimize the amount of floating-point computation (which always entails difficult round-off error analysis) in favor of as much exact calculation with integers as possible. Thus, it is worthwhile to look for other approaches.

Robin [Robin 84] has shown that

$$\text{RH} \iff \frac{\sigma(n)}{n} < e^\gamma \log\log(n) \quad \text{for} \quad n > 5040, \qquad \text{(R)}$$

where $\sigma(n)$ is the sum of divisors of the positive integer n. Building on this, Lagarias [Lagarias 02] showed

1058-6458/2006 $0.50 per page

the equivalence of the Riemann hypothesis to a condition on harmonic sums $H_n := \sum_{i=1}^{n} 1/i$, namely

$$\text{RH} \Longleftrightarrow \sigma(n) \leqslant H_n + \exp(H_n)\log(H_n) \quad \text{for all } n. \quad \text{(L)}$$

These inequalities suggest an attack: to disprove the RH, we need to find n with large *relative abundance* $\rho(n) := \sigma(n)/n$ and that violate one or other of these inequalities. How difficult might this be? We can get a clue from another theorem of Robin [Robin 84]:

Theorem 1.1. *Independently of the RH, except for $n = 1, 2, 12$,*

$$\rho(n) < e^{\gamma} \log\log(n) + \frac{\left(\frac{7}{3} - e^{\gamma} \log\log(12)\right) \log\log(12)}{\log\log(n)}. \quad (1\text{–}1)$$

The numerator in the last term is about 0.6482. Note also that it is known that $\limsup_{n\to\infty} \rho(n)/\log\log(n) = e^{\gamma}$. Thus, possible violations of inequality (R) must have $\rho(n)$ exceeding $e^{\gamma}\log\log(n)$ in the small gap allowed by Theorem 1.1.

2. ABUNDANT NUMBERS

The numbers called *perfect* in classical number theory have $\sigma(n) = 2n$, so $\rho(n) = 2$. *Abundant* numbers have $\rho(n) > 2$, and t-abundant numbers have $\rho(n) > t$. Some properties of these have been studied in [Davenport 33, Deléglise 98].

The study of numbers with $\rho(n)$ large in various other senses was initiated by Ramanujan [Ramanujan 15, Ramanujan 97] and further developed by Alaoglu, Erdős, and Nicolas [Alaoglu and Erdős 44, Erdős and Nicolas 75]. A positive integer n is called *superabundant* (SA) if $\rho(k) < \rho(n)$ for all $k < n$.

Here is a summary of some known facts about SA numbers, upon which I will later base an algorithm:

1. The exponent sequence does not increase: if $n = 2^{a_2}\, 3^{a_3} \cdots m^{a_m}$ is SA, where m is the maximal prime factor, then $a_2 \geqslant a_3 \geqslant \cdots \geqslant a_m$.
2. All exponents are determined within a range ± 1 by the first exponent: if $1 < j < i \leqslant m$, then $|a_i - \lfloor a_j \log_i j \rfloor| \leqslant 1$.
3. The last exponent is unity except in two cases: $a_m = 1$ unless $n = 4$ or 36.
4. The first exponent provides an upper bound for all exponents: for $i > 2$ we have $i^{a_i} < 2^{a_2+2}$.

A certain subset of the SA numbers allows an even more precise characterization: n is *colossally abundant* (CA) if there exists $\epsilon > 0$ such that

$$\frac{\sigma(n)}{n^{1+\epsilon}} \geqslant \frac{\sigma(k)}{k^{1+\epsilon}} \quad \text{for all } k > 1. \quad (2\text{–}1)$$

CA numbers may be viewed as maximizers of $\log(\sigma(n)/n) - \epsilon\log(n)$ for fixed ϵ; thus, we penalize n that are too large. Robin [Robin 84] has shown that the structure of the set of CA numbers is determined by the following properties: We first form the set E of *critical* ϵ *values*,

$$E_p := \bigcup_{\alpha=1,2,3,\ldots} \left\{ \log_p\left(1 + \frac{1}{\sum_{i=1}^{\alpha} p^i}\right) \right\}, \quad (2\text{–}2)$$

$$E := \bigcup_p E_p,$$

and then label the elements of E in decreasing order: $\epsilon_1 = \log_2\left(\frac{3}{2}\right) > \epsilon_2 = \log_3\left(\frac{4}{3}\right) > \epsilon_3 = \log_2\left(\frac{7}{6}\right) > \cdots$. We then have the following theorem (paraphrased from [Robin 84, p. 190]:

Theorem 2.1.

(i) *If $\epsilon \notin E$, then $\sigma(n)/n^{1+\epsilon}$ has a unique maximum attained at the number n_ϵ with prime exponents given by*

$$a_p(\epsilon) = \left\lfloor \log_p\left(\frac{p^{1+\epsilon} - 1}{p^{\epsilon} - 1}\right)\right\rfloor - 1. \quad (2\text{–}3)$$

(ii) *If ϵ satisfies $\epsilon_{i+1} < \epsilon < \epsilon_i$ for $i = 1, 2, 3, \ldots$, then n_ϵ is constant and we call it N_i. We have $N_1 = 2$, $N_2 = 6$, $\ldots$.*

(iii) *If the sets E_p are disjoint (which is likely, but not certainly known), then the set of CA numbers is equal to the set of N_i, $i = 1, 2, 3, \ldots$. If this is the case, then $\sigma(n)/n^{1+\epsilon_i}$ attains its maximum at the two points N_i and N_{i+1}.*

(iv) *If the sets E_p are not disjoint, then for each $\epsilon_i \in E_q \cap E_p$, $\sigma(n)/n^{1+\epsilon_i}$ attains its maximum at the four points N_i, qN_i, rN_i and $N_{i+1} = qrN_i$.*

Finally, Robin has proven that if the Riemann hypothesis is false, then there will be a counterexample n (violating the inequality (R)) that is a CA number. There might also be SA counterexamples. We could also try to find violations of Lagarias's inequality (L), but this would involve more difficult estimation (using asymptotic expansions) of the harmonic sum.

Thus, my procedure will simply be as follows: I compute successively larger CA numbers and check Robin's inequality for each. As a further check, I compute some SA numbers, but these cause extra difficulties to be mentioned below.

3. ALGORITHMS

The computation of CA and SA numbers allows a very compact representation: since the prime exponents form a slowly decreasing sequence, with a very long tail of ones, we may just store for each exponent the number of primes with that exponent. In this way I was able to reach CA numbers as large as $10^{10^{10}}$. The required quantities for inequality testing, $\log(n)$ and $\rho(n)$, are computed directly from this representation using high-precision floating-point arithmetic. For this I used the `mpfr` library [Hanrot et al. 04].

3.1 Colossally Abundant Numbers

My algorithm to compute colossally abundant numbers is as follows: I keep a list z of records, each containing a prime p, $\log p$, its exponent a, and a critical ϵ_c, which is the value of ϵ at which this exponent will next change (as ϵ is decreased). I also maintain a variable ι, which counts the number of exponents equal to unity.

We first initialize:

- Fix $0 < \epsilon \leqslant 1$. Then, for each prime p, compute $a = \left\lfloor \log_p\left(\frac{p^{1+\epsilon}-1}{p^\epsilon-1}\right)\right\rfloor - 1$, and if $a \geqslant 2$, store it in the z list. If $a = 1$, just increment the variable ι. Stop when $a = 0$. During this p loop, also update $\log(n)$ and $\rho(n)$, using $\rho(n) = \prod_i \frac{p_i - p_i^{-a_i}}{p_i - 1}$.

Now each step of the main loop consists in determining which of the possible events A, B, or C occurs:

- A: A new prime (with exponent 1) is added, so we increment ι. This happens when $\epsilon_{\text{ext}} := \log_p(1+p)$ is maximal, where p is the new prime.

- B: The first prime with exponent 1 has its exponent raised to 2. This happens when

  $$\epsilon_{\text{inc}} := \log_p\left(\frac{p+1+\frac{1}{p}}{p+1}\right)$$

  is maximal, where p is the prime in question.

- C: A prime with exponent greater than or equal to 2 has its exponent incremented. This happens when

  $$\epsilon_{\max} := \log_p\left(\frac{1-p^{a+1}}{p-p^{a+1}}\right)$$

  is maximal, where p is the prime in question and a its exponent.

This algorithm will correctly compute all CA numbers in sequence, as long as the floating-point arithmetic is accurate enough to ensure that all tests are decided correctly. In all tests I computed primes using Dan Bernstein's version of Atkin's sieve [Bernstein 99].

3.2 Superabundant Numbers

In contrast, to my knowledge, no algorithm for computing all SA numbers up to a given maximum is known, although it is possible that a method of Robin for a related problem might be adapted [Robin 82]. I have therefore used the following method, which generates a list an initial portion of which contains only SA numbers (and all SA numbers up to the maximum of the initial portion), but it is not possible to determine when the first incorrect entry in the list occurs.

The algorithm is as follows: For each initial term 2^{a_2}, $a_2 = 1, 2, 3, \ldots$, we recursively extend the list of prime factors with every possible exponent a_p, subject to the conditions $a_2 \geqslant a_3 \geqslant \cdots \geqslant a_m$; if $1 < j < i \leqslant m$, then $|a_i - \lfloor a_j \log_i j \rfloor| \leqslant 1$; and $i^{a_i} < 2^{a_2+2}$, $i \geqslant 2$. Each extension satisfying all these conditions is a candidate SA number. When no more extensions are possible, we move to the next power of two. After reaching the largest desired power of two, we check the list to remove non-SA numbers by sorting it and keeping only n corresponding to successive maxima of $\rho(n)$. I was able to reach 2^{14} this way, at which n is approximately 10^{10^5}. It is easy to determine that the smallest SA candidate divisible by 2^{14} has $\log(n) > 154$, so my SA data up to at least this n will be correct.

4. RESULTS

Figure 1 shows that $\log\log(n)$ at CA numbers n appears to be asymptotically an affine function of $\log\epsilon$.

This may be verified by some asymptotic estimates: for fixed small positive ϵ, we have that the maximal prime $m(\epsilon)$ in the colossally abundant number associated with ϵ satisfies $m(\epsilon) \sim \frac{-1}{\epsilon\log(\epsilon)}$. This is obtained from equation (2–3) by solving

$$\log_p\left(\frac{p^{1+\epsilon}-1}{p^\epsilon-1}\right) = 2.$$

Similarly, the number of distinct primes $k(\epsilon)$ satisfies $k(\epsilon) \sim 1/(\epsilon\log^2(\epsilon))$. From bounds given in [Robin 84],

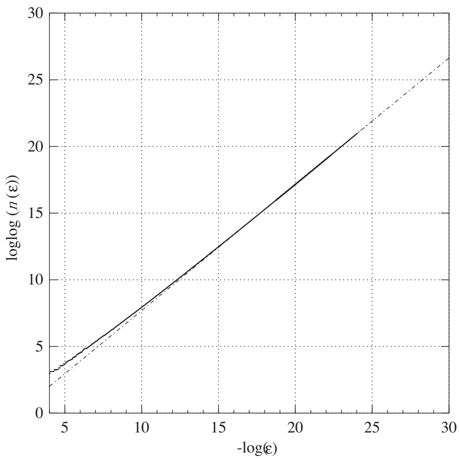

FIGURE 1. The dependence of the colossally abundant number $n(\epsilon)$ on ϵ. The straight line confirms the asymptotic affine dependence of $\log\log(n)$ on $\log\epsilon$.

we have

$$\sum_{p \leqslant m(\epsilon)} \log p \leqslant \log n(\epsilon) \leqslant \sum_{p \leqslant m(\epsilon)} \log p + \sum_{p \leqslant x_2} a_2 \log p, \tag{4–1}$$

where $x_2 < \sqrt{2x}$ and $a_2 < \log_2 \frac{2^{1+\epsilon}-1}{2^{\epsilon}-1}$. It follows that $n(\epsilon) \sim -1/(\epsilon \log(\epsilon))$.

Using the computed data on SA and CA numbers, we may estimate their density. Defining $Q(x)$ to be the number of SA numbers less than or equal to x, we see the graph of Figure 2(top). It is known from [Erdős and Nicolas 75] that

$$\liminf_{n\to\infty} \frac{\log(Q(x)}{\log\log(x)} \geqslant 1 + \frac{5}{48} \approx 1.1042;$$

my data gives about 1.2 for the left-hand side ratio at $\log\log(n) = 5$. For CA numbers no comparable results are known; my data is shown in Figure 2(bottom).

Defining $\delta(n)$ as the difference between the right- and left-hand sides of Robin's inequality,

$$\delta(n) := e^{\gamma} \log\log(n) - \frac{\sigma(n)}{n} \tag{4–2}$$

(so that $\delta < 0$ implies that the Riemann hypothesis is false), we see the behavior in Figure 3 (top). This suggests that $\log\delta$ is asymptotically an affine function of $x = \log\log(n)$, with slope close to $-\frac{1}{2}$. We may subtract a line of this slope in order to study more closely the oscillations. There are also fast oscillations as shown

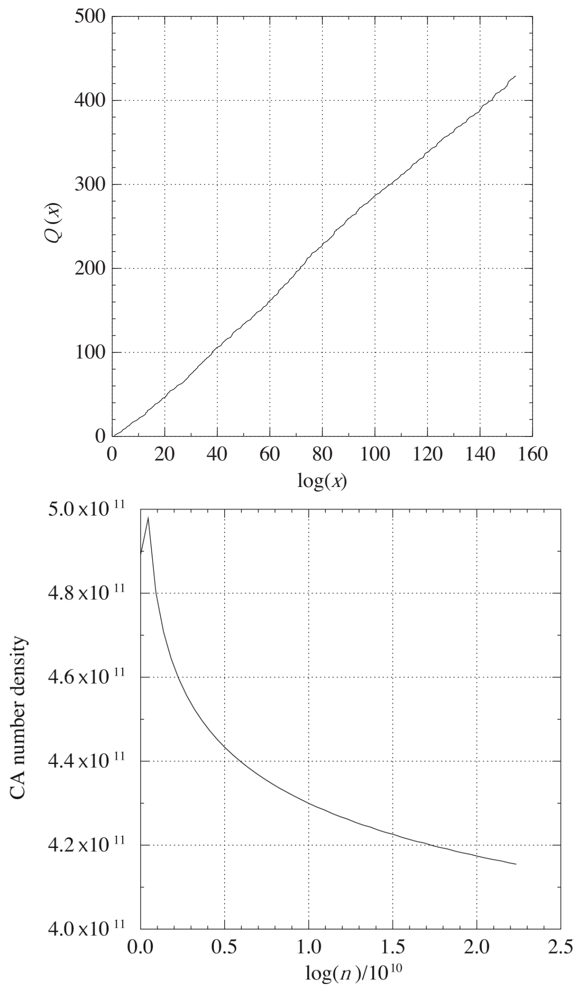

FIGURE 2. Cumulative number of SA numbers (top). Density of CA numbers (bottom).

in Figure 4 (top). This shows the difference between $\log\delta(n)$ and a conjectured best-fit line $a - x/2$, where $x := \log\log(n)$.

Considering now superabundant numbers, I observe the behavior shown in Figure 4 (bottom). It appears that SA numbers come almost as close to minimizing δ as do CA numbers.

On the basis of these data, we are led to a final conjecture:

Conjecture 4.1. *Assuming the Riemann hypothesis, then for colossally abundant numbers we have*

$$\log\delta(n) \sim -\frac{1}{2}\log\log(n) - o(\log\log(n)) \tag{4–3}$$

as $n \to \infty$.

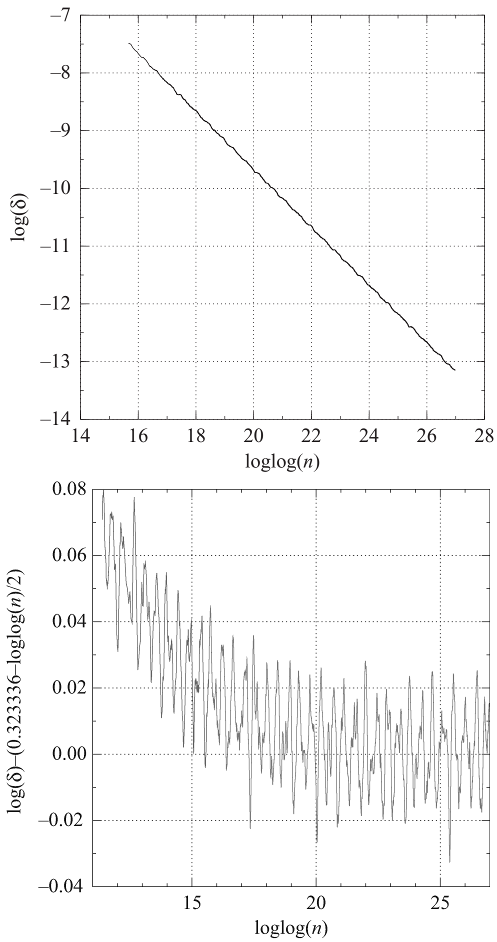

FIGURE 3. Dependence of $\log \delta(n)$ on n at CA numbers (top). Deviation of $\log \delta$ from a best-fit line at CA numbers (bottom).

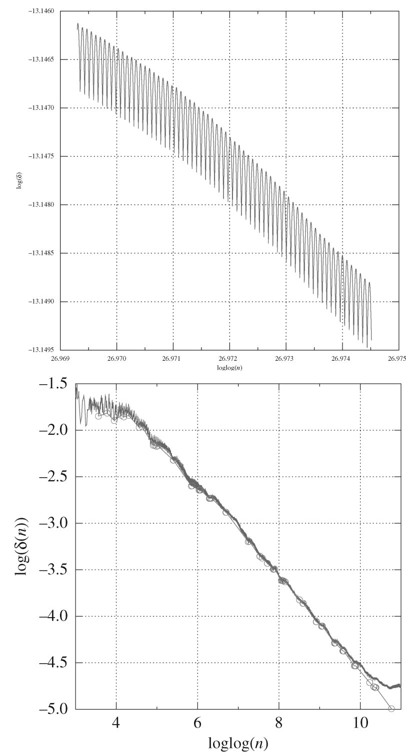

FIGURE 4. Dependence of δ on n at CA numbers (top); $\log \delta$ at CA numbers (lower line and circles) and at SA numbers (upper line). The SA number data are probably incorrect past about $\log\log(n) = 10$ (bottom).

5. CONCLUSION

The most surprising observation is that the oscillations in δ are so small. There is no sign of the near-violations of the Riemann hypothesis that are seen in ζ-zero calculations. The present calculations were done on a Pentium 4 and took several days using `mpfr` set at 100-bit precision. Some calculations were verified with interval arithmetic. The need to use floating-point software creates difficulties: not only is there a speed penalty, but also there is really a need for some kind of dynamic precision control. It is not easy to see how to implement this (perhaps some form of exact real arithmetic [Briggs 06]), but if it could be achieved, a much longer run would be possible and should provide worthwhile new data. The limiting factor in the current implementation is the storage needed for internal data structures.

ACKNOWLEDGMENTS

I wish to thank Jeff Lagarias, Jean-Louis Nicolas, Pieter Moree, and an anonymous referee for useful suggestions and comments.

REFERENCES

[Alaoglu and Erdős 44] L. Alaoglu and P. Erdős. "On Highly Composite and Similar Numbers." *Trans. Amer. Math. Soc.* 56 (1944), 448–469.

[Bernstein 99] D. Bernstein. "Primegen-0.97." Available online (http://cr.yp.to/primegen.html), 1999.

[Briggs 06] K. M. Briggs. "Implementing Exact Real Arithmetic in Python, C++, and C." *J. Theor. Comp. Sci.* 351:1 (2006), 74–81.

[Conrey 05] J. B. Conrey. "The Riemann Hypothesis." *Notices of the AMS* 50 (2005), 341–353.

[Davenport 33] H. Davenport. "Über Numeri abundantes." *Sitzungsber. Preuss. Akad. Wiss.* (1933), 830–837.

[Deléglise 98] "M. Deléglise. "Bounds for the Density of Abundant Integers." *Exp. Math.* 7 (1998), 137–143.

[Erdős and Nicolas 75] P. Erdős and J.-L. Nicolas. "Répartition des nombres superabondants." *Bull. Math. Soc. Fr.* 103 (1975), 65–90.

[Gourdon 04] X. Gourdon. "The 10^{13} First Zeros of the Riemann Zeta Function, and Zeros Computation at Very Large Height." Available online: (http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeroscompute.html), 2004

[Hanrot et al. 04] G. Hanrot, V. Lefèvre, P. Pélissier, P. Zimmermann, et al. "The MPFR Library." Available online: (http://www.mpfr.org), 2004.

[Lagarias 02] J. C. Lagarias. "An Elementary Problem Equivalent to the Riemann Hypothesis." *Am. Math. Monthly* 109 (2002), 534–543.

[Norfolk et al. 92] T. S. Norfolk, A. Ruttan, and R. S. Varga. "A Lower Bound for the de Bruijn–Newman Constant II." In *Progress in Approximation Theory*, edited by A. A. Gonchar and E. B. Saff, pp. 403–418. New York: Springer-Verlag, 1992.

[Ramanujan 15] S. Ramanujan. "Highly Composite Numbers." *Proc. London Math. Soc.* 14 (1915), 347–409.

[Ramanujan 97] S. Ramanujan. "Highly Composite Numbers" (annotated and with a foreword by J.-L. Nicolas and G. Robin). *Ramanujan J.* 1 (1997), 119–153.

[Robin 82] G. Robin. "Méthodes d'optimisation pour un problème de théorie des nombres." *RAIRO Informatique théorique* 17 (1982), 239–247.

[Robin 84] G. Robin. "Grandes valeurs de la fonction somme des diviseurs et hypothèse de Riemann." *J. Math. Pures Appl.* 63 (1984), 187–213.

[Wedeniwski 05] S. Wedeniwski. "Zetagrid Project." Available online: (http://zetagrid.net), 2005.

Keith Briggs, Mobility Research Center, BT, Polaris 134, Adastral Park, Martlesham, Suffolk IP5 3RE, UK (Keith.Briggs@bt.com)

Received July 8, 2005; accepted November 4, 2005.