An Elementary Problem Equivalent to the Riemann Hypothesis

Jeffrey C. Lagarias

(May 5, 2001 version)

ABSTRACT. The problem is: Let $H_n = \sum_{j=1}^{n} \frac{1}{j}$ be the n-th harmonic number. Show, for each $n \geq 1$, that

$$\sum_{d|n} d \leq H_n + \exp(H_n)\log(H_n),$$

with equality only for $n = 1$.

AMS Subject Classification (2000): Primary 11M26, Secondary 11A25

Keywords: Riemann hypothesis, colossally abundant numbers

1. Introduction

We consider the following problem.

Problem E. *Let $H_n = \sum_{j=1}^{n} \frac{1}{j}$. Show that, for each $n \geq 1$,*

$$\sum_{d|n} d \leq H_n + \exp(H_n)\log(H_n), \tag{1.1}$$

with equality only for $n = 1$.

The function $\sigma(n) = \sum_{d|n} d$ is the *sum of divisors function*, so for example $\sigma(6) = 12$. The number H_n is called the *n-th harmonic number* by Knuth, Graham and Patashnik [9, Sect. 6.3], who detail various properties of harmonic numbers.

The 'E' in Problem E might stand for either 'easy' or 'elementary'. Perhaps 'H' for 'hard' would be a better letter to use, since our object is to show the following equivalence.

Theorem 1.1 *Problem E is equivalent to the Riemann hypothesis.*

The Riemann hypothesis, stated by Riemann [17] in 1859, concerns the complex zeros of the Riemann zeta function. The Riemann zeta function $\zeta(s)$ is defined by the Dirichlet series

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s},$$

which converges for $\Re(s) > 1$, and it has an analytic continuation to the complex plane with one singularity, a simple pole with residue 1 at $s = 1$. The Riemann hypothesis states that the nonreal zeros of the Riemann zeta function $\zeta(s)$ all lie on the line $\Re(s) = \frac{1}{2}$. One reason for the great interest in the Riemann hypothesis, regarded by many as the most important unsolved problem in pure mathematics, is its connection with the distribution of prime numbers, described below. More significantly, the Riemann hypothesis is a special case of questions concerning generalizations of the zeta function (L-functions) and their connections with problems in number theory, algebraic geometry, topology, representation theory and perhaps even physics, see Berry and Keating [3], Katz and Sarnak [8] and Murty[11].

The connection of the Riemann hypothesis with prime numbers was the original question studied by Riemann [17]. Let $\pi(x)$ count the number of primes p with $1 < p \leq x$. C. F. Gauss noted empirically that $\pi(x)$ is well approximated by the logarithmic integral

$$Li(x) = \int_2^x \frac{dt}{\log t},$$

which itself satisfies

$$Li(x) = \frac{x}{\log x} + O(\frac{x}{(\log x)^2}).$$

The Riemann hypothesis is equivalent to the assertion that for each $\epsilon > 0$ there is a positive constant C_ϵ such that

$$|\pi(x) - Li(x)| \leq C_\epsilon x^{1/2+\epsilon}$$

for all $x \geq 2$, see Edwards [5, p. 90]. The force of the Riemann hypothesis lies in the small size of the error term. The strongest form known of the *Prime Number Theorem* with error term asserts that

$$|\pi(x) - Li(x)| \leq C_1 x \, \exp(-C_2(\log x)^{3/5-\epsilon}),$$

for any positive ϵ, for certain positive constants C_1 and C_2 depending on ϵ; this result is due to Vinogradov and Korobov in 1958.

Problem E encodes a modification of a criterion of Guy Robin [18] for the Riemann hypothesis. Robin's criterion states that the Riemann hypothesis is true if and only if

$$\sigma(n) < e^\gamma n \log\log n \quad \text{for all} \quad n \geq 5041, \tag{1.2}$$

where $\gamma \approx 0.57721$ is Euler's constant. This criterion is related to the density of primes, as explained in §2. Our aim was to obtain a problem statement as elementary as possible, containing no undefined constants. However the hard work underlying the equivalence resides in the results of Robin stated in §3, where we give a proof of Theorem 1.1.

Before coming to the proof, in the next section we describe how the Riemann hypothesis is related to the sum of divisors function. The connection traces back to Ramanujan's work on highly composite numbers, and involves several results of Erdős with coauthors.

2. Colossally Abundant Numbers

The Riemann hypothesis is encoded in the criterion of Theorem 1.1 in terms of the very thin set of values of $\sigma(n)$ that are "large." The sum of divisors function is given by

$$\sigma(n) = \prod_{p^a||n} (1 + p + p^2 + ... + p^a) = n \prod_{p^a||n} (1 + \frac{1}{p} + ... + \frac{1}{p^a}), \tag{2.1}$$

where the product is taken over all primes p dividing n and the notation $p^a||n$, means p^a divides n but p^{a+1} does not divide n. Most values of $\sigma(n)$ are on the order of Cn. The average size of $\sigma(n)$ was essentially found by Dirichlet in 1849, and is given by the following result of Bachmann, cf. Hardy and Wright [7, Theorem 324].

Theorem 2.1 *(Bachmann) The average order of $\sigma(n)$ is $\frac{\pi^2}{6}n$. More precisely,*

$$\sum_{j=1}^{n} \sigma(j) = \frac{\pi^2}{12}n^2 + O(n \log n)$$

as $n \to \infty$.

The maximal order of $\sigma(n)$ is somewhat larger, and was determined by Gronwall in 1913, see Hardy and Wright [7, Theorem 323, Sect. 18.3 and 22.9].

Theorem 2.2 *(Gronwall) The asymptotic maximal size of $\sigma(n)$ satisfies*

$$\limsup_{n\to\infty} \frac{\sigma(n)}{n \log\log n} = e^{\gamma},$$

where γ is Euler's constant.

This result can be deduced from Mertens' theorem, which asserts that

$$\prod_{p \le x} (1 - \frac{1}{p}) \sim \frac{e^{\gamma}}{\log x}.$$

as $x \to \infty$. A much more refined version of the asymptotic upper bound, due to Robin [18, Theorem 2], asserts (unconditionally) that for all $n \ge 3$,

$$\sigma(n) < e^{\gamma} n \log\log n + 0.6482 \frac{n}{\log\log n}. \tag{2.2}$$

In §3 we will show that

$$H_n + \exp(H_n)\log(H_n) \le e^{\gamma} n \log\log n + \frac{7n}{\log n},$$

the right side of which is only slightly smaller than that of (2.2). Comparing this bound with (2.2) shows that the inequality (1.1), if ever false, cannot be false by very much.

The study of extremal values of functions of the divisors of n is a branch of number theory with a long history. Let $d(n)$ count the number of divisors of n (including 1 and n itself). *Highly composite numbers* are those positive integers n such that

$$d(n) > d(k) \qquad \text{for} \qquad 1 \le k \le n-1.$$

Superior highly composite numbers are those positive integers for which there is a positive exponent ϵ such that

$$\frac{d(n)}{n^{\epsilon}} \ge \frac{d(k)}{k^{\epsilon}} \qquad \text{for all} \qquad k > 1,$$

so that they maximize $\frac{d(n)}{n^{\epsilon}}$ over all n; these form a subset of the highly composite numbers. The study of these numbers was initiated by Ramanujan. One can formulate similar extrema for the sum of divisors function. *Superabundant numbers* are those positive integers n such that

$$\frac{\sigma(n)}{n} > \frac{\sigma(k)}{k} \qquad \text{for} \qquad 1 \le k \le n-1.$$

n	Factorization of n	$\frac{\sigma(n)}{n}$
2	2	1.500
6	$2 \cdot 3$	2.000
12	$2^2 \cdot 3$	2.333
60	$2^2 \cdot 3 \cdot 5$	2.800
120	$2^3 \cdot 3 \cdot 5$	3.000
360	$2^3 \cdot 3^2 \cdot 5$	3.250
2520	$2^3 \cdot 3^2 \cdot 5 \cdot 7$	3.714
5040	$2^4 \cdot 3^2 \cdot 5 \cdot 7$	3.838
55440	$2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11$	4.187
720720	$2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13$	4.509
1441440	$2^5 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 13$	4.581
4324320	$2^5 \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 13$	4.699
21621600	$2^5 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13$	4.855
367567200	$2^5 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17$	5.141
6983776800	$2^5 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 19$	5.412
160626866400	$2^5 \cdot 3^3 \cdot 5^2 \cdot 7 \cdots 23$	5.647
321253732800	$2^6 \cdot 3^3 \cdot 5^2 \cdot 7 \cdots 23$	5.692
9316358251200	$2^6 \cdot 3^3 \cdot 5^2 \cdot 7 \cdots 29$	5.888
288807105787200	$2^6 \cdot 3^3 \cdot 5^2 \cdot 7 \cdots 31$	6.078
2021649740510400	$2^6 \cdot 3^3 \cdot 5^2 \cdot 7^2 \cdot 11 \cdots 31$	6.187
6064949221531200	$2^6 \cdot 3^4 \cdot 5^2 \cdot 7^2 \cdot 11 \cdots 31$	6.238
224403121196654400	$2^6 \cdot 3^4 \cdot 5^2 \cdot 7^2 \cdot 11 \cdots 37$	6.407

Table 1: Colossally abundant numbers up to 10^{18}.

Colossally abundant numbers are those numbers n for which there is a positive exponent ϵ such that

$$\frac{\sigma(n)}{n^{1+\epsilon}} \geq \frac{\sigma(k)}{k^{1+\epsilon}} \qquad \text{for all} \qquad k > 1.$$

so that n attains the maximum value of $\frac{\sigma(k)}{k^{1+\epsilon}}$ over all k. The set of colossally abundant numbers is infinite. They form a subset of the superabundant numbers, a fact which can be deduced from the definition. Table 1 gives the colossally abundant numbers up to 10^{18}, as given in [1]. Robin showed that if the Riemann hypothesis is false, there will necessarily exist a counterexample to the inequality (1.2) which is a colossally abundant number, cf. [18, Proposition 1 of Section 3]; the same property can be established for counterexamples to (1.1). (There could potentially exist other counterexamples as well.)

Superabundant and colossally abundant numbers were studied[1] in detail by Alaoglu and Erdős [1] in 1944. As evidenced in the table, colossally abundant numbers consist of a product of all the small primes up to some bound, with exponents which are nonincreasing as the prime increases. The values of these exponents have a characteristic smooth shape which can be almost completely described (see (2.3) below). In fact these classes of numbers had been studied earlier, by Ramanujan, in his 1915 work on highly composite numbers [14]. The notes in Ramanjuan's Collected Papers report: "The London Mathematical Society was in some financial difficulty at the time, and Ramanujan suppressed part of what he had written in order to save expense." [14, p. 339]. Only the first 52 of 75 sections were printed. The manuscript of the unpublished part was eventually rediscovered among the papers of G. N. Watson, in "Ramanujan's Lost Notebook" (see Andrews [2]

[1] Alaoglu and Erdős [1] use a slightly stronger definition of colossally abundant number; they impose the additional requirement that $\frac{\sigma(n)}{n^{1+\epsilon}} > \frac{\sigma(k)}{k^{1+\epsilon}}$ must hold for $1 \leq k < n$. With their definition the colossally abundant numbers are exactly those given by (2.3) below.

and Ramanujan [15, pp. 280–308]) and finally published in 1997, in [16]. In it superabundant and colossally abundant numbers are considered in Section 59, as special cases of the concepts of generalised highly composite numbers and superior generalised highly composite numbers, for the parameter value $s = 1$. Ramanujan derived upper and lower bounds for the maximal order of generalised highly composite numbers in Section 71, assuming the Riemann hypothesis. (The Riemann hypothesis was assumed from Section 40 onward in his paper.) His bounds imply, assuming the Riemann hypothesis, that (1.2) holds for all sufficiently large n.

The results of Alaoglu and Erdős in their 1944 paper are unconditional, and mainly concerned the exponents of primes occurring in highly composite and superabundant numbers. In considering the exponents of primes appearing in colossally abundant numbers, they raised the following question, which is still unsolved.

Conjecture. (Alaoglu and Erdős) If p and q are both primes, is it true that p^x and q^x are both rational only if x is an integer?

This conjecture would follow as a consequence[2] of the *four exponentials conjecture* in transcendental number theory, which asserts that if a_1, a_2 form a pair of complex numbers, linearly independent over the rationals $\mathbb{Q}$, and b_1, b_2 is another such pair, also linearly independent over the rationals, then at least one of the four exponentials $\{e^{a_i b_j} : \ i, j = 1 \text{ or } 2\}$ is transcendental, see Lang [10, pp. 8-11]. Alaoglu and Erdős [1, Theorem 10] showed that for "generic" values of ϵ there is exactly one maximizing integer n , and the exponent $a_p(\epsilon)$ of each prime p in it is

$$a_p(\epsilon) = \lfloor \frac{\log(p^{1+\epsilon} - 1) - \log(p^{\epsilon} - 1)}{\log p} \rfloor - 1; \tag{2.3}$$

furthermore for all $\epsilon > 0$ this value of n is maximizing. However for a discrete set of ϵ there will be more than one maximizing integer. Erdős and Nicolas [6, p. 70] later showed that for a given value of ϵ there will be exactly one, two or four integers n that attain the maximum value of $\frac{\sigma(k)}{k^{1+\epsilon}}$. If the conjecture of Alaoglu and Erdős above has a positive answer, then no values of ϵ will have four extremal integers. This would imply that the ratio of two consecutive colossally abundant numbers, the larger divided by the smaller, will always be a prime, and that every colossally abundant number has a factorization (2.3) for some nonempty open interval of "generic" values of ϵ.

Now we can explain the relevance of the Riemann hypothesis to the extremal size of $\sigma(n)$. Colossally abundant numbers are products of all the small primes raised to powers that are a smoothly decreasing function of their size. Fluctuations in the distribution of primes will be reflected in fluctuations in the growth rate of $\frac{\sigma(n)}{n}$ taken over the set of colossally abundant numbers. Recall that the Riemann hypothesis is equivalent to the assertion that for each $\epsilon > 0$,

$$|\pi(x) - Li(x)| < x^{1/2+\epsilon}$$

holds for all sufficiently large x. It is also known that if the Riemann hypothesis is false, then there will exist a specific positive constant δ such that the one-sided inequality

$$\pi(x) > Li(x) + x^{1/2+\delta}$$

is true for an infinite set of values with $x \to \infty$. Heuristically, if we choose a value of x where such an excess of primes over $Li(x)$ occurs, and then take a product of all primes up to this bound, choosing

[2]For non-rational x consider $a_1 = \log p$, $a_2 = \log q$, $b_1 = 1$ and $b_2 = x$ in the four exponentials conjecture. For non-integer rational x a direct argument is used.

appropriate exponents, one may hope to construct numbers n with $\sigma(n)$ exceeding $e^{\gamma} n \log\log n$ by a small amount. If the Riemann hypothesis holds, there is a smaller upper bound for excess of the number of primes above $Li(x)$, and from this one can deduce a slighly better upper bound for $\sigma(n)$. The analysis of Robin [18] gives a quantitative version of this, as formulated in Propositions 3.1 and 3.2 below.

One can prove unconditionally that the inequality (1.1) holds for nearly all integers. Even if the Riemann hypothesis is false, the set of exceptions to the inequality (1.1) will form a very sparse set. Furthermore, if there exists any counterexample to (1.1), the value of n will be very large.

3. Proofs

The proof of Theorem 1.1 is based on the following two results of Robin [18].

Proposition 3.1 *(Robin) If the Riemann hypothesis is true, then for each* $n \geq 5041$*,*

$$\sum_{d|n} d \leq e^{\gamma} n \log\log n \ , \tag{3.1}$$

where γ *is Euler's constant.*

Proof. This is Theorem 1 of Robin [18]. ■

Proposition 3.2 *(Robin) If the Riemann hypothesis is false, then there exist constants* $0 < \beta < \frac{1}{2}$ *and* $C > 0$ *such that*

$$\sum_{d|n} d \geq e^{\gamma} n \log\log n + \frac{C n \log\log n}{(\log n)^{\beta}} \tag{3.2}$$

holds for infinitely many n*.*

Proof. This appears in Proposition 1 of Section 4 of Robin [18]. The constant β can be chosen to take any value $1 - b < \beta < \frac{1}{2}$, where $b = \Re(\rho)$ for some zero ρ of $\zeta(s)$ with $\Re(\rho) > \frac{1}{2}$, and $C > 0$ must be chosen sufficiently small, depending on ρ. The proof uses ideas from a result of Nicolas [12], [13, Proposition 3], which itself uses a method of Landau. ■

We prove two preliminary lemmas.

Lemma 3.1 *For* $n \geq 3$*,*

$$\exp(H_n)\log(H_n) \geq e^{\gamma} n \log\log n \ . \tag{3.3}$$

Proof. Letting $\lfloor t \rfloor$ denote the integer part of t and $\{t\}$ the fractional part of t, we have

$$\int_1^{\infty} \frac{\lfloor t \rfloor}{t^2} dt = \sum_{1 \leq r \leq t \leq n} \int_1^{\infty} \frac{dt}{t^2} = \sum_{r=1}^{n} (\frac{1}{r} - \frac{1}{n}) = H_n - 1.$$

Thus

$$H_n = 1 + \int_1^{n} \frac{t - \{t\}}{t^2} dt = \log n + 1 - \int_1^{n} \frac{\{t\}}{t^2} dt. \tag{3.4}$$

From this we obtain

$$H_n = \log n + \gamma + \int_n^\infty \frac{\{t\}}{t^2} dt, \tag{3.5}$$

where we have set

$$\gamma := 1 - \int_1^\infty \frac{\{t\}}{t^2} dt.$$

This is Euler's constant $\gamma = 0.57721\ldots$, since letting $n \to \infty$ yields

$$\gamma = \lim_{n\to\infty} (H_n - \log n),$$

which is its usual definition. Now (3.5) gives

$$H_n > \log n + \gamma,$$

which on exponentiating yields

$$\exp(H_n) \geq e^\gamma n. \tag{3.6}$$

Finally $H_n \geq \log n$ so $\log(H_n) \geq \log\log n > 0$ for $n \geq 3$. Combining this with (3.6) yields (3.3). ∎

Lemma 3.2 *For $n \geq 20$,*

$$H_n + \exp(H_n)\log(H_n) \leq e^\gamma n \log\log n + \frac{7n}{\log n} \ . \tag{3.7}$$

Proof. The formula (3.4) implies, for $n \geq 3$,

$$\log H_n \leq \log(\log n + 1) \leq \log\log n + \frac{1}{\log(n+1)} \ . \tag{3.8}$$

Next, (3.5) yields

$$\begin{aligned} \exp(H_n) &= \exp\left(\log n + \gamma + \int_n^\infty \frac{\{t\}}{t^2} dt\right) \\ &\leq e^\gamma n \exp\left(\int_n^\infty \frac{dt}{t^2}\right) \\ &= e^\gamma n \exp\left(\frac{1}{n}\right). \end{aligned} \tag{3.9}$$

Since

$$\exp(x) \leq 1 + 2x \quad \text{for} \quad 0 \leq x \leq 1 \ ,$$

we obtain from (3.9) that

$$\exp(H_n) \leq e^\gamma n \left(1 + \frac{2}{n}\right) \ . \tag{3.10}$$

Combining this bound with (3.8) yields

$$\begin{aligned} \exp(H_n)\log(H_n) &\leq e^\gamma n \log\log n + \frac{e^\gamma n}{\log(n+1)} + 2e^\gamma(\log\log n + 1) \\ &\leq e^r n \log\log n + \frac{6n}{\log n} \end{aligned}$$

for $n \geq 10$. Using $H_n \leq \log n + 1$ and $\frac{n}{\log n} \geq \log n + 1$ for $n \geq 20$ yields (3.7). ∎

Proof of Theorem 1.1

$\Leftarrow$ Suppose the Riemann hypothesis is true. Then Proposition 3.1 and Lemma 3.1 together give, for $n \geq 5041$,

$$\sum_{d|n} d \leq e^{\gamma} n \log\log n < H_n + \exp(H_n) \log H_n \ .$$

For $1 \leq n \leq 5040$ one verifies (1.1) directly by computer, the only case of equality being $n = 1$.

$\Rightarrow$ Suppose (1.1) holds for all n. We argue by contradiction, and suppose that the Riemann hypothesis is false. Then Proposition 3.2 applies. However its lower bound, valid for infinitely many n, contradicts the upper bound of Lemma 3.2 for sufficiently large n. We conclude that the Riemann hypothesis must be true. ■

Acknowledgements. I thank Michel Balazard for bringing Robin's criterion to my attention, and J.-L. Nicolas for references, particularly concerning Ramanujan, Eric Rains for checking (1.1) for $1 \leq n \leq 5040$ on the computer, and Jim Reeds for corrections. I thank the reviewers for simplifications of the proofs of Lemmas 3.1 and 3.2.

References

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

[2] G. E. Andrews, An Introduction to Ramanujan's Lost Notebook, Amer. Math. Monthly **86** (1979), 89–108.

[3] M. V. Berry and J. P. Keating, The Riemann zeros and eigenvalue asymptotics, SIAM Review **41** (1999), 236–266.

[4] M. Davis, Hilbert's Tenth Problem: Aspects of a negative solution, in: *Mathematical Developments arising from Hilbert Problems*, Proc. Symp. Pure Math. No. XXVIII, American Math. Society: Provindence 1976, pp. 323–378.

[5] H. E. Edwards, *Riemann's Zeta Function*, Academic Press: New York 1974.

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

[7] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers. Fifth Edition.* Oxford Univ. Press: Oxford 1979.

[8] N. Katz and P. Sarnak, Zeros of zeta functions and symmetry, Bull. Amer. Math. Soc. *36* (1999), 1–26.

[9] D. Knuth, R. L. Graham and O. Patashnik, *Concrete Mathematics* (Second Edition), Addison-Wesley: Reading, MA 1994.

[10] S. Lang, *Transcendental Number Theory*, Addison-Wesley: Reading, Mass. 1966.

[11] M. R. Murty, A motivated introduction to the Langlands program, in: *Advances in Number Theory*, Oxford U. Press: New York 1993, pp. 37–66.

[12] J. L. Nicolas, Petites valeurs de la fonction d'Euler et hypothèse de Riemann, *Seminar on number theory, Paris 1981-82 (Parius 1981/1982)*, Birkhäuser, Boston 1983, pp. 207–218.

[13] J. L. Nicolas, Petites valeurs de la fonction d'Euler, J. Number Theory **17** (1983), 375–388.

[14] S. Ramanujan, Highly composite numbers, Proc. London Math. Soc. **14** (1915), 347–407. (Also, S. Ramanujan, *Collected Papers,*, Chelsea 1962, pp. 78–129, notes pp. 338-339.)

[15] S. Ramanujan, *The Lost notebook and other unpublished papers.* With an introduction by George E. Andrews, Narosa: New Delhi and Springer-Verlag: Berlin-New York 1988.

[16] S. Ramanujan, Highly composite numbers. Annotated and with a foreword by J.-L. Nicholas and G. Robin, Ramanujan J. **1** (1997), 119–153.

[17] G. F. B. Riemann, Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse, Monatsber. Akad. Berlin (1859), 671–680. (Also in: G. F. B. Riemann, *Gesammelte Mathematische Werke* (Ed. H. Weber), Dover reprint 1953.) [English translation in: Edwards (1974), pp. 299-305.]

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

AT&T Labs–Research, Florham Park, NJ 07932-0971, USA
email address: `jcl@research.att.com`