On amicable numbers

Carl Pomerance

To Professor Helmut Maier on his sixtieth birthday

Abstract Let $\mathscr{A}(x)$ denote the set of integers $n \le x$ that belong to an amicable pair. We show that $\#\mathscr{A}(x) \le x/e^{\sqrt{\log x}}$ for all sufficiently large x.

1 Introduction

Let σ denote the sum-of-divisors function and let $s(n) = \sigma(n) - n$. Two different positive integers a, b with $s(a) = b$ and $s(b) = a$ are said to form an amicable pair. This concept is attributed to Pythagoras and has been studied over the millennia since both by numerologists and number theorists. The first example of an amicable pair is 220 and 284. About 12 million pairs are now known, but we don't have a proof of their infinitude.

Say a positive integer is amicable if it belongs to an amicable pair and let $\mathscr{A}$ denote the set of amicable numbers. Kanold [9] was the first to consider $\mathscr{A}$ from a statistical viewpoint, announcing in 1954 that $\mathscr{A}$ has upper density smaller than 0.204. Soon after, Erdős [5] showed they have asymptotic density 0. In the period 1973 to 1981 there were several papers getting successively better upper bounds for $\#\mathscr{A}(x)$, where $\mathscr{A}(x) = \mathscr{A} \cap [1, x]$. Somewhat simplifying the expressions, these upper bounds have progressed as follows:

Carl Pomerance
Dartmouth College, Hanover, NH 03755, USA
carl.pomerance@dartmouth.edu

Mathematics Subject Classification: 11A25, 11N25
Key Words: amicable number

$$\frac{x}{(\log\log\log\log x)^{1/2}}, \quad \frac{x}{\log\log\log x}, \quad \frac{x}{\exp\left((\log\log\log x)^{1/2}\right)}, \quad \frac{x}{\exp\left((\log x)^{1/3}\right)},$$

see [16], [7], [14], and [15], respectively. In this paper we are able to replace the exponent $1/3$ in the last estimate with $1/2$. In particular, we prove the following theorem.

Theorem 1.1 *As $x \to \infty$, we have*

$$\#\mathscr{A}(x) \le x/\exp\left(\left(\tfrac{1}{2} + o(1)\right)\sqrt{\log x \log\log\log x}\right).$$

The proof largely follows the plan in [15], but with some new elements. In particular, a separate argument now handles the case when an amicable number n is divisible by a very large prime p. Thus, assuming the largest prime factor p of n is not so large, it can be shown that, usually, a fairly large prime divides $\sigma(n/p)$. The argument to handle the case of p large is reminiscent of the paper [13] which deals with Lehmer's problem on Euler's function φ, also see the newer paper [1]. In addition, we use a result in [2] to streamline the argument that $\sigma(n/p)$ has a large prime factor.

The argument in [15] was subsequently used to estimate the distribution of numbers n with $\varphi(n) = \varphi(n+1)$, and some similar equations; see [6], [8]. These results were since improved in [17]. However, it is not clear if the method of [17] can be used for the distribution of amicable numbers.

We record some of the notation used. Let $P(n)$ denote the largest prime factor of $n > 1$ and let $P(1) = 1$. We say an integer n is squarefull if for each prime $p \mid n$ we have $p^2 \mid n$. We use the notation (a,b) for the greatest common divisor of the positive integers a,b. We write $d \parallel n$ if $d \mid n$ and $(d, n/d) = 1$.

Note that if n is large we have $s(n) < 2n\log\log n$.

2 A lemma

Let $\Phi(x,y)$ denote the number of integers $n \in [1,x]$ with $P(\varphi(n)) \le y$. In [2, Theorem 3.1] it is shown that for any fixed $\varepsilon > 0$, we have $\Phi(x,y) \le x\exp(-(1 + o(1))u\log\log u)$ as $u \to \infty$, where $u = \log x/\log y$ and $(\log\log x)^{1+\varepsilon} \le y \le x$. At first glance one might think there is a typographical error here, since the corresponding result of de Bruijn [3] counts $n \in [1,x]$ with $P(n) \le y$, and the upper bound is $x\exp(-(1+o(1))u\log u)$, which is supported by a corresponding lower bound, see [4]. However, a heuristic argument indicates that it is likely that for $\Phi(x,y)$, replacing $\log u$ with $\log\log u$ is correct (see [2, Section 8] and [10]). Until we have a corresponding lower bound, we will not know for sure.

In this paper we will require an estimate for the number of $n \in [1,x]$ with $P(\sigma(n)) \le y$. The function $\sigma(n)$ closely resembles the function $\varphi(n)$; we see "$p+1$" in $\sigma(n)$, where we would see "$p-1$" in $\varphi(n)$. However, it is not as simple as this, since σ treats higher powers of primes differently than φ. It is possible to overcome this difference since most numbers n are not divisible by a large squarefull number. But to keep things simple, we restrict our numbers n in the following result to squarefree numbers. Let $\Sigma(x,y)$ denote the number of squarefree numbers $n \in [1,x]$ with $P(\sigma(n)) \le y$.

Lemma 2.1 *For each fixed $\varepsilon > 0$, we have $\Sigma(x,y) \le x\exp(-(1+o(1))u\log\log u)$ as $u \to \infty$, where $u = \log x/\log y$ and $(\log\log x)^{1+\varepsilon} \le y \le x$.*

As indicated above, the proof of this result follows from small cosmetic changes to the proof of the corresponding result on $\Phi(x,y)$ in [2].

3 Proof of Theorem 1.1

Let x be large and let

$$L = L(x) = \exp\left(\tfrac{1}{2}\sqrt{\log x \log\log\log x}\right).$$

(i) *For $n \in \mathscr{A}(x)$, we may assume that $n > x/L$ and $s(n) > x/L$.*

The first assertion is obvious and the second follows from the fact that each $n \in \mathscr{A}(x)$ is determined by $s(n)$.

(ii) *For $n \in \mathscr{A}(x)$ we may assume that the largest divisor d of n with $P(d) \le L^2$ has $d \le x^{1/3}$, and similarly for $s(n)$. In particular, we may assume that $P(n) > L^2$ and $P(s(n)) > L^2$.*

Indeed, by [3], for $z \ge x^{1/3}$, the number of integers $d \le z$ with $P(d) \le L^2$ is $O(z/L)$, so by partial summation, the number of $n \in \mathscr{A}(x)$ divisible by such a number $d > x^{1/3}$ is $O(x(\log x)/L)$. A parallel argument holds for $s(n)$. The assertions about $P(n), P(s(n))$ now follow from (i).

(iii) *For $n \in \mathscr{A}(x)$ we may assume the largest squarefull divisor of n is at most L^2, and the same for $s(n)$.*

Since the number of squarefull numbers $d \le z$ is $O(\sqrt{z})$ for all $z \ge 1$, partial summation implies that the number of integers $n \le x$ divisible by a squarefull number $d > L^2$ is $O(x/L)$. A similar estimate holds for $s(n)$.

Note that if $n \in \mathscr{A}(x)$, then (ii) and (iii) imply that $P(n) \,\|\, n$ and $P(s(n)) \,\|\, s(n)$. For the remainder of the proof, for $n \in \mathscr{A}(x)$ we write:

$$n = pm,\ p = P(n) \nmid m, \quad s(n) = n' = p'm',\ p' = P(n') \nmid m'.$$

Note too that for x large we have $n' < 2x\log\log x$.

(iv) *For $n \in \mathscr{A}(x)$, we may assume that $P((n,s(n))) \le L$.*

Suppose $n \in \mathscr{A}(x)$, $r = P((n,s(n)))$, and that $r > L$. Then $r \mid \sigma(n)$, so there is a prime power $q^j \mid n$ with $r \mid \sigma(q^j)$. We use an elementary inequality found in the proofs of [11, Lemma 3.6] and [12, Lemma 3.3]: for any positive integer d,

$$\sum_{\substack{q^j \le x \\ d \mid \sigma(q^j)}} \frac{1}{q^j} \ll \frac{(\log x)^2}{d}. \tag{1}$$

We Apply (1) with $d = r$, getting that the number of $n \le x$ with $r \mid n$ and $r \mid \sigma(n)$ is $O(x(\log x)^2/r^2)$, and so the number of n which violate (iv) is at most a constant times

$$x(\log x)^2 \sum_{r > L} \frac{1}{r^2} \ll \frac{x(\log x)^2}{L}.$$

This estimate shows that we may assume (iv).

(v) *For $n \in \mathscr{A}(x)$, we may assume that $mm' > x/L$.*

Suppose $n \in \mathscr{A}(x)$. We have

$$p'm' = s(n) = \sigma(pm) - pm = ps(m) + \sigma(m), \tag{2}$$

$$p'\sigma(m') + \sigma(m') = \sigma(p'm') = \sigma(n) = p\sigma(m) + \sigma(m). \tag{3}$$

Multiplying (2) by $\sigma(m)$, (3) by $s(m)$ and subtracting to eliminate p, we have

$$p'm'\sigma(m) - p'\sigma(m')s(m) - \sigma(m')s(m) = \sigma(m)^2 - \sigma(m)s(m) = m\sigma(m).$$

Thus,

$$p'(m'\sigma(m) - \sigma(m')s(m)) = \sigma(m')s(m) + m\sigma(m). \tag{4}$$

Since the right side of (4) is positive, we see that m, m' determine p', and by symmetry, they also determine p. So the number of cases for which (v) fails is at most

$$\sum_{mm' \le x/L} 1 = O(x(\log x)/L).$$

(vi) *For $n \in \mathscr{A}(x)$, we may assume that $p, p' \le x^{3/4}L$.*

Suppose $n \in \mathscr{A}(x)$ and $p > x^{3/4}L$. Then $m < x^{1/4}/L$, so by (v), $m' > x^{3/4}$. Since $n' < 2x \log\log x$, we have $p' < 2x^{1/4} \log\log x$. Write the prime factorization of n' as $p_1 p_2 \dots p_t$, where $p_1 = p'$ and $p_1 \ge p_2 \ge \cdots \ge p_t$, and for $j = 1, 2, \dots, t$, let $D_j = p_1 p_2 \dots p_j$. By (i), some $D_j > x^{1/2}$; let D be the least such divisor of n'. Since $p_1 < 2x^{1/4} \log\log x$, we have

$$x^{1/2} < D \le 2x^{3/4} \log\log x. \tag{5}$$

Further, by (ii) and (iii), D is squarefree, $D \parallel n'$, and every prime dividing D is larger than L^2. From the identity

$$\begin{aligned} s(m)\sigma(n') = s(m)\sigma(n) &= \sigma(m)\sigma(n) - \sigma(n)m = \sigma(m)\sigma(n) - \sigma(m)(p+1)m \\ &= \sigma(m)\sigma(n) - \sigma(m)(n+m) = \sigma(m)n' - \sigma(m)m, \end{aligned}$$

we have

$$s(m)\sigma(D)\sigma(M) = \sigma(m)DM - m\sigma(m).$$

Reading this equation as a congruence modulo $\sigma(D)$, we have

$$\sigma(m)DM \equiv m\sigma(m) \pmod{\sigma(D)}.$$

The number of choices for $M < 2x(\log\log x)/D$ which satisfy this congruence is at most

$$1 + \frac{2x\log\log x}{D\sigma(D)/(\sigma(m)D,\sigma(D))} \le 1 + \frac{2x\sigma(m)(D,\sigma(D))\log\log x}{D^2}.$$

Since $(D,\sigma(D)) \mid (n,n')$ and every prime dividing D exceeds $L^2 > L$, (iv) implies that $(D,\sigma(D)) = 1$. So, for a given choice of m, D, the number of choices for M is at most

$$1 + \frac{4xm(\log\log x)^2}{D^2}.$$

We sum this expression for D satisfying (5) and $m < x^{1/4}/L$ and so get that the number of choices for $n \in \mathscr{A}(x)$ with $p > x^{3/4}L$ is $O(x(\log\log x)/L)$. A similar argument holds if $p' > x^{3/4}L$. We conclude that the number of cases where (vi) fails is negligible.

For $n = pm \in \mathscr{A}(x)$ we write $m = m_0m_1$ where m_1 is the largest squarefree number with $m_1 \| m$, and we similarly write $m' = m_0'm_1'$.

(vii) *For $n \in \mathscr{A}(x)$, we may assume that $P(\sigma(m_1)) > L$ and $P(\sigma(m_1')) > L$.*

Assume for $n \in \mathscr{A}(x)$ that the first condition in (vii) fails. (The argument for the second condition will follow similarly.) By (iii) and (vi), $pm_0 \le x^{3/4}L^3$. For given choices of p and m_0 we count the number choices of squarefree integers $m_1 \le x/pm_0$ with $P(\sigma(m_1)) \le L$. For this, we use Lemma 2.1. Let $u = \log(x/pm_0)/\log L$. Since $pm_0 \le x^{3/4}L^3$, we have

$$u \ge \frac{\log\left(x^{1/4}/L^3\right)}{\log L} = \left(\tfrac{1}{2} + o(1)\right)\sqrt{\frac{\log x}{\log\log\log x}},$$

so that

$$u\log\log u = \left(\tfrac{1}{2} + o(1)\right)\sqrt{\log x \log\log\log x} = (1 + o(1))\log L$$

as $x \to \infty$. Hence, we uniformly have that the number of choices for m_1 is at most

$$\frac{x}{pm_0L^{1+o(1)}}$$

as $x \to \infty$. We now sum on p, m_0 getting that the number of $n \in \mathscr{A}(x)$ for which $P(\sigma(m_1)) \le L$ is at most $x/L^{1+o(1)}$ as $x \to \infty$.

We now turn to the conclusion of the argument. We suppose that $n \in \mathscr{A}(x)$ and that (i)–(vii) hold. At the cost of doubling our count and letting n run up to

$2x\log\log x$, we may assume that $p > p'$. (That $p \neq p'$ can be seen from (ii), (iv).) By (vii), there is a prime $r \mid \sigma(m_1)$ with $r > L$. Thus, there is a prime $q \mid m$ with $q \equiv -1 \pmod{r}$. This implies that $q > L$, so that by (iv), $q \nmid n'$. But $\sigma(n) = \sigma(n')$, so there is a prime power $\ell^j \mid n'$ with $\ell \neq q$ and $r \mid \sigma(\ell^j)$. Note that, by (1), $\sum 1/\ell^j \ll (\log x)^2/r$. We have

$$n' = s(n) = ps(m) + \sigma(m) \equiv 0 \pmod{\ell^j}. \tag{6}$$

Say $\ell^i = (\ell^j, s(m))$, so that p is in a residue class $a(m) \pmod{\ell^{j-i}}$. Also, (6) implies that $\ell^i \mid \sigma(m)$, so $\ell^i \mid m$. Further, using (ii), (iii), and $p > \ell$, we may assume that $p > \ell^j$. The number of such numbers $n \leq 2x\log\log x$ is at most

$$\begin{aligned}
&\sum_{r>L} \sum_{\substack{q<x \\ q\equiv -1 \pmod{r}}} \sum_{\substack{\ell^j<x \\ r\mid\sigma(\ell^j)}} \sum_{i\leq j} \sum_{\substack{m<x \\ q\ell^i\mid m}} \sum_{\substack{p\leq 2x\log\log x/m \\ p\equiv a(m) \pmod{\ell^{j-i}} \\ p>\ell^j}} 1 \\
&\leq \sum_r \sum_q \sum_{\ell^j} \sum_i \sum_m \frac{2x\log\log x}{\ell^{j-i}m} \ll \sum_r \sum_q \sum_{\ell^j} \sum_i \frac{x\log x \log\log x}{q\ell^j} \\
&\ll \sum_r \sum_q \sum_{\ell^j} \frac{x(\log x)^2 \log\log x}{q\ell^j} \ll \sum_r \sum_q \frac{x(\log x)^4 \log\log x}{rq} \\
&\ll \sum_r \frac{x(\log x)^5 \log\log x}{r^2} \ll \frac{x(\log x)^5 \log\log x}{L},
\end{aligned}$$

where we treated r, q, ℓ, p as integer variables. This calculation along with the previous cases finishes the proof.

Acknowledgments The author would like to thank Hanh Nguyen, Paul Pollack, and Lola Thompson for their interest in this work.

References

1. A. Anavi, P. Pollack, and C. Pomerance, On congruences of the form $\sigma(n) \equiv a \pmod{n}$, IJNT **9** (2012), 115–124.
2. W. D. Banks, J. B. Friedlander, C. Pomerance, and I. E. Shparlinski, Multiplicative structure of values of the Euler function, in High primes and misdemeanours: Lectures in honour of the sixtieth birthday of Hugh Cowie Williams (A. van der Poorten, ed.), Fields Inst. Comm. **41** (2004), 29–47.
3. N. G. de Bruijn, On the number of integers $\leq x$ and free of prime factors $> y$, Nederl. Akad. Wetensch. Proc. Ser. A **54** (1951), 50–60.
4. E. R. Canfield, P. Erdős, and C. Pomerance, On a problem of Oppenheim concerning 'Factorisatio Numerorum', J. Number Theory **17** (1983), 1–28.
5. P. Erdős, On amicable numbers, Publ. Math. Debrecen **4** (1955), 108–111.
6. P. Erdős, C. Pomerance, and A. Sárközy, On locally repeated values of certain arithmetic functions, II, Acta Math. Hungar. **49** (1987), 251–259.
7. P. Erdős and G. J. Rieger, Ein Nachtrag über befreundete Zahlen, J. reine angew. Math. **273** (1975), 220.

8. S. W. Graham, J. J. Holt, and C. Pomerance, On the solutions to $\varphi(n) = \varphi(n+k)$, in Number theory in progress, K. Gyory, H. Iwaniec, and J. Urbanowicz, eds., vol. 2, de Gruyter, Berlin and New York, 1999, pp. 867–882.
9. H. J. Kanold, Über die Dichte der Mengen der vollkommenen und der befreundeten Zahlen, Math. Z. **61** (1954), 180–185.
10. Y. Lamzouri, Smooth values of the iterates of the Euler phi-function, Canad. J. Math. **59** (2007), 127–147.
11. P. Pollack and L. Thompson, Practical pretenders, Publ. Math. Debrecen **82** (2013), 651–667.
12. P. Pollack and J. Vandehey, Some normal numbers given by arithmetic functions, Canad. Math. Bull. **58**, 160–173.
13. C. Pomerance, On composite n for which $\varphi(n) \mid n-1$, I, Acta Arith. **28** (1976), 387–389.
14. C. Pomerance, On the distribution of amicable numbers, J. reine angew. Math. **293/294** (1977), 217–222.
15. C. Pomerance, On the distribution of amicable numbers, II, J. reine angew. Math. **325** (1981), 183–188.
16. G. J. Rieger, Bemerkung zu einem Ergebnis von Erdős über befreundete Zahlen, J. reine angew. Math. **261** (1973), 157–163.
17. T. Yamada, On equations $\sigma(n) = \sigma(n+k)$ and $\varphi(n) = \varphi(n+k)$, arXiv:1001.2511.