On the distribution of amicable numbers. II

By *Carl Pomerance**) at Athens, Georgia

§ 1. Introduction

Let $\sigma(n)$ denote the sum of the divisors of n and let $s(n)=\sigma(n)-n$. Two natural numbers n, m are called an amicable pair if $s(n)=m$ and $s(m)=n$. The least such pair with $n \neq m$ is $n=220$, $m=284$. We say a natural number n is an amicable number if it is a member of an amicable pair. An equivalent definition is $s(s(n))=n$ and still another is $\sigma(n)=\sigma(s(n))$. Amicable numbers have a very long history; they were mentioned (in fact, defined) by Pythagoras, they were investigated by the Arabs during the European Dark Ages, and they were studied by Fermat, Descartes, and Euler.

While it is intuitively clear that the property of being an amicable number is very special and so they should be rare among the natural numbers, this fact is not easy to prove. It was not until 1955 when Erdös [4] showed the amicable numbers have density 0. The best result to date is in [5] where it is shown that $A(x)$, the number of amicable numbers not exceeding x, satisfies

$$A(x) \leqq x \cdot \exp\{-c(\log\log\log x \log\log\log\log x)^{1/2}\} \tag{1}$$

for all large x, where c is a certain positive constant. The reader may refer to [5] for the intermediate history between Erdös's result and (1). Although better than previous results, (1) is still very weak; it does not even give $A(x)=O(x/\log\log x)$.

In this paper we take a different approach to the problem and prove the dramatically stronger result

$$A(x) \leqq x \cdot \exp\{-(\log x)^{1/3}\} \tag{2}$$

for all large x. Note that (2) immediately implies that the sum of the reciprocals of the amicable numbers is finite, a fact not previously known. Also (2) settles a conjecture of Erdös [4] that $A(x)=O(x/(\log x)^k)$ for every k. In [4], Erdös conjectures that $A(x) \gg x^{1-\varepsilon}$ for every $\varepsilon>0$. This conjecture conflicts with that of Bratley, Lunnon, and McKay [1] that $A(x)=o(\sqrt{x})$. If Erdös is right, then (2) might be near to best possible. However, we cannot even prove there are infinitely many amicable numbers.

In our discussion p, q, r will always denote primes and n, m, k, a natural numbers. We will denote by $P(n)$ the largest prime factor of n if $n>1$; also $P(1)=1$.

*) Research supported in part by a grant from the National Science Foundation.

0075-4102/81/0325-0009$02.00

§ 2. The proof of (2)

In this section we prove the

Theorem. *For all large* x, $A(x) \leqq x/e^{(\log x)^{1/3}}$.

Proof. If $n \leqq x$ is an amicable number, then $s(n) = \sigma(n) - n$ is also amicable. Moreover, if x is large enough and $n \leqq x$, we have $s(n) \leqq 2x \log\log x$, since

$$\limsup \sigma(n)/(n \log\log n) = e^{\gamma} < 2.$$

For simplicity of notation, we shall let

$$l = e^{(\log x)^{1/3}}, \quad L = e^{\frac{1}{8}(\log x)^{2/3} \log\log x}.$$

We first show that we may assume

(i) $$P(n) \geqq L^2, \quad P(s(n)) \geqq L^2.$$

Indeed, from de Bruijn [2], the number of $n \leqq 2x \log\log x$ composed only of primes smaller than L^2 is $o(x/l)$.

We next show we may assume that

(ii) if k^a divides n or $s(n)$ where $a \geqq 2$, then $k^a < l^3$.

Indeed, the number of $n \leqq 2x \log\log x$ divisible by some $k^a \geqq l^3$ with $a \geqq 2$ is at most

$$2x \log\log x \sum_{k^a \geqq l^3} k^{-a} \ll x(\log\log x)/l^{3/2} = o(x/l).$$

Now we show that we may assume

(iii) if $p \mid (n, \sigma(n))$, then $p < l^4$.

If $p \mid \sigma(n)$, then there is a prime power $q^a \| n$ with $p \mid \sigma(q^a)$. If $p \geqq l^4$, then

$$q^a > \tfrac{1}{2}\sigma(q^a) \geqq \tfrac{1}{2} p \geqq l^3,$$

so that by (ii) we may assume $a = 1$. Thus n is divisible by p and a prime $q \equiv -1 \pmod{p}$. The number of such n is at most

$$\frac{x}{p} \sum_{\substack{q \equiv -1 (p) \\ q \leqq x/p}} \frac{1}{q} \leqq \frac{cx \log x}{p^2},$$

where c is an absolute constant. Now summing on $p \geqq l^4$, we have the number of n for which (iii) fails is $o(x/l)$.

Next we show we may assume

(iv) $$n/P(n) \geqq L, \quad s(n)/P(s(n)) \geqq L.$$

To see this, say we write $n = mp$ where $p = P(n)$ and $s(n) = m'p'$ where $p' = P(s(n))$. From (i) and (ii) we have

$$p \nmid m, \; p' \nmid m', \; m \leqq x/L^2, \; m' \leqq 2x(\log\log x)/L^2.$$

Thus

$$(3) \qquad m'p' = s(n) = \sigma(mp) - mp = p\big(\sigma(m) - m\big) + \sigma(m),$$

$$(4) \qquad \sigma(m')p' + \sigma(m') = \sigma(m'p') = \sigma(n) = \sigma(m)p + \sigma(m).$$

Multiplying (3) by $\sigma(m)$, (4) by $\sigma(m) - m$, and subtracting, we find that

$$(5) \quad p'\big(\sigma(m)m' - \sigma(m)\,\sigma(m') + m\sigma(m')\big) = m\sigma(m) + \sigma(m)\,\sigma(m') - m\sigma(m').$$

Since the right side of (5) is positive and hence non-zero, we see that m and m' determine p'. Furthermore, if m' and p' are known, then so is n since $n = s(m'p')$. Thus the number of amicable numbers $n \leqq x$ for which (iv) fails is at most the number of pairs m, m' where either

$$m < L, \quad m' \leqq 2x\,(\log\log x)/L^2$$

or

$$m \leqq x/L^2, \quad m' < L.$$

The number of such pairs is less than

$$4L \cdot x\,(\log\log x)/L^2 = o(x/l).$$

Now we show we may assume that

(v) if $m = n/P(n)$, $m' = s(n)/P\big(s(n)\big)$, then $P\big(\sigma(m)\big) \geqq l^4$, $P\big(\sigma(m')\big) \geqq l^4$.

Say $P\big(\sigma(m)\big) < l^4$. Say q is a prime and $q^a \| m$. If $a \geqq 2$, then by (ii) we have $q < l^{3/2}$. If $a = 1$, then $q + 1 | \sigma(m)$, so that $P(q+1) < l^4$. Let

$$\psi(m) = \prod_{q^a \| m} (q+1)\,q^{a-1},$$

so that $P\big(\psi(m)\big) < l^4$. Then not only have we defined an integer $\psi(m)$, but we have given a factorization of this integer. (By "factorization" we mean a representation as a product of integers larger than 1 where the order of the factors does not count.) We ask how many choices for m can map onto not only the same integer $k = \psi(m)$, but also give the same factorization of k. We show there are at most 2 such m. Indeed if $k = \psi(m_1) = \psi(m_2)$ and the factorization of k so described is

$$k = a_1 a_2 \cdots a_t$$

where the $a_i > 1$ are integers, we first note that if some $a_i = q$, a prime, with $q \neq 3$, then $q^2 | m_1$, $q^2 | m_2$. We may thus replace m_1, m_2, and k with m_1/q, m_2/q, and $k = a_1 \cdots a_{i-1} a_{i+1} \cdots a_t$. Continuing in this fashion we may assume that if $q | m_1$, $q \neq 3$, then $q \| m_1$, and similarly for m_2. Next, we argue that if some $a_i \neq 4$ is composite, then $a_i = q + 1$ for some prime q with $q \| m_1$, $q \| m_2$. Replacing m_1, m_2, and k with m_1/q, m_2/q, k/a_i and continuing in this fashion, we finally reach the stage where

$$m_1, m_2 \in \{2^a \cdot 3^b : a = 0 \text{ or } 1,\ b = 0, 1, 2, \ldots\}.$$

Then $\psi(m_1) = \psi(m_2)$ implies $m_1, m_2 \in \{3^{b+1}, 2 \cdot 3^b\}$ for some $b \geqq 1$. Thus there cannot be 3 distinct m which give the same factorization of $k = \psi(m)$.

Let $f(k)$ denote the number of factorizations of k. Thus if z is any quantity, then the number $N(z)$ of m with $\psi(m) \leqq z$ and $P(\psi(m)) < l^4$ satisfies

$$N(z) \leqq 2 \sum_{\substack{k \leqq z \\ P(k) < l^4}} f(k).$$

Following the methods of [3] (Theorem 5. 1) where we use a result of McMahon and a method of Rankin, we have for $c > 3/4$,

$$\begin{aligned} N(z) &\leqq 2z^c \sum_{P(k) < l^4} f(k)/k^c = 2z^c \prod_{\substack{P(k) < l^4 \\ k > 1}} (1 - k^{-c})^{-1} \ll z^c \cdot \exp\Big(\sum_{P(k) < l^4} k^{-c}\Big) \\ &= z^c \cdot \exp\Big(\prod_{p < l^4} (1 - p^{-c})^{-1}\Big) = z^c \cdot \exp\{\exp\big(\sum_{p < l^4} p^{-c} + O(1)\big)\} \end{aligned}$$

where the implied constants are absolute. We shall take

$$c = 1 - (\log x)^{-1/3}.$$

From the prime number theorem and partial summation we find

$$\begin{aligned} \sum_{p < l^4} p^{-c} &= l^{-4c} \pi(l^4) + \int_2^{l^4} c s^{-c-1} li(s)\, ds + \int_2^{l^4} c s^{-c-1} (\pi(s) - li(s))\, ds \\ &= \int_2^{l^4} \frac{s^{-c}}{\log s}\, ds + l^{-4c}(\pi(l^4) - li(l^4)) + 2^{-c} li(2) + \int_2^{l^4} c s^{-2} s^{1-c} (\pi(s) - li(s))\, ds \\ &= \int_2^{l^4} \frac{s^{-c}}{\log s}\, ds + O(1), \end{aligned}$$

since $s^{1-c} \leqq l^{4(1-c)} = e^4$. Also

$$\int_2^{l^4} \frac{s^{-c}}{\log s}\, ds = \int_{2^{1-c}}^{l^{4(1-c)}} \frac{du}{\log u} = \int_{2^{1-c}}^{e^4} \frac{du}{u \log u} + \int_{2^{1-c}}^{e^4} \frac{u-1}{u \log u}\, du = -\log(1-c) + O(1),$$

so that

$$\sum_{p < l^4} p^{-c} = \tfrac{1}{3} \log\log x + O(1).$$

If $z \geqq L$, we thus have

$$\begin{aligned} \frac{N(z)}{z} &\ll z^{c-1} \exp\exp\big(\tfrac{1}{3} \log\log x + O(1)\big) \\ &= \exp\{(c-1)\log z + \exp\big(\tfrac{1}{3}\log\log x + O(1)\big)\} \\ &\leqq \exp\{-\tfrac{1}{8}(\log x)^{1/3} \log\log x + O\big((\log x)^{1/3}\big)\}. \end{aligned}$$

Thus for all large x and $z \geqq L$ we have

$$N(z) \leqq z/l^2. \qquad (6)$$

Note that $\limsup \psi(m)/(m \log\log m) = 6e^{\gamma}/\pi^2 < 2$. By (iv), $m \geqq L$ so we may assume x is large enough so that $\psi(m) < 2m \log\log m$. If $n \leqq x$ is an amicable number and $p = P(n)$, then again from (iv) we have $x/p \geqq L$. Thus the number of such n with the first inequality in (v) failing is by (6) at most

$$\sum_{p \leqq x/L} N\big(2x(\log\log x)/p\big) \leqq \frac{2x \log\log x}{l^2} \sum_{p \leqq x/L} \frac{1}{p} = o(x/l).$$

Similarly the number of n for which the second inequality in (v) fails is at most

$$\sum_{p \leqq 2x(\log\log x)/L} N\big(4x(\log\log x)^2/p\big) = o(x/l).$$

We thus may assume (v).

We now consider amicable numbers $n \leqq x$ that satisfy (i)—(v). In addition, if we double the count and let n range up to $2x \log\log x$, we may also assume

(vi) $P(n) \geqq P\big(s(n)\big)$.

Let $n \leqq 2x \log\log x$ be such an amicable number. Write $n = mp$ where $p = P(n)$. From (v) there is a prime $r \geqq l^4$ with $r | \sigma(m)$. Since $p \nmid m$ (from (i) and (ii)), we have $r | \sigma(n) = \sigma\big(s(n)\big)$. Thus from (ii) there are primes q, q' with $q \equiv q' \equiv -1 \pmod r$ and $q \| m$, $q' \| s(n)$. Thus

$$(7) \qquad \big(\sigma(m) - m\big)p + \sigma(m) \equiv 0 \pmod{q'}.$$

Since $q' > l^4$, we have from (iii) that

$$\sigma(m) - m \not\equiv 0 \pmod{q'}.$$

Thus given m and q', (7) places p in a certain residue class $\bmod q'$, call this class $a(m, q')$. From all of these considerations we have the number of such n at most

$$\sum_{r \geqq l^4} \sum_{\substack{q \equiv -1\,(r) \\ q \leqq x}} \sum_{\substack{m \equiv 0\,(q) \\ m \leqq x}} \sum_{\substack{q' \equiv -1\,(r) \\ q' \leqq 2x \log\log x}} \sum_{\substack{p \equiv a(m,q')\,(q') \\ p \leqq 2x(\log\log x)/m \\ p \geqq q'}} 1 \leqq \sum_r \sum_q \sum_m \sum_{q'} \frac{2x \log\log x}{q'm}$$

$$\ll \sum_r \sum_q \sum_m \frac{x \log x \log\log x}{rm}$$

$$\ll \sum_r \sum_q \frac{x (\log x)^2 \log\log x}{rq}$$

$$\ll \sum_r \frac{x (\log x)^3 \log\log x}{r^2} = o(x/l).$$

This calculation completes the proof of our theorem.

Remark. Some small alterations in the above proof allow us to establish the slightly stronger result that there is some $c > 0$ with

$$A(x) \ll x \cdot \exp\{-c(\log x \log\log x)^{1/3}\}.$$

References

[1] *P. Bratley, F. Lunnon* and *J. McKay,* Amicable numbers and their distribution, Math. Comp. **24** (1970), 431—432.

[2] *N. G. de Bruijn,* On the number of integers $\leqq x$ and free of prime factors $> y$, Nederl. Akad. Wetensch. Proc. Ser. A **54** (1951), 50—60.

[3] *E. R. Canfield, P. Erdös* and *C. Pomerance,* On a problem of Oppenheim concerning "Factorisatio Numerorum", to appear.

[4] *P. Erdös,* On amicable numbers, Publ. Math. Debrecen **4** (1955), 108—111.

[5] *C. Pomerance,* On the distribution of amicable numbers, J. reine angew. Math. **293/294** (1977), 217—222.

Department of Mathematics, University of Georgia, Athens, Georgia 30602, U.S.A.

Eingegangen 13. Oktober 1980