ON SOME FRIENDS OF THE SOCIABLE NUMBERS

PAUL POLLACK

Abstract. Let $s(n)$ denote the sum of the proper divisors of n. Set $s_0(n) = n$, and for $k > 0$, put $s_k(n) := s(s_{k-1}(n))$ if $s_{k-1}(n) > 0$. Thus, *perfect* numbers are those n with $s(n) = n$ and *amicable numbers* are those n with $s(n) \neq n$ but $s_2(n) = n$. We prove that for each fixed $k \geq 1$, the set of n which divide $s_k(n)$ has density zero, and similarly for the set of n for which $s_k(n)$ divides n. These results generalize the theorem of Erdős that for each fixed k, the set of n for which $s_k(n) = n$ has density zero.

1. Introduction

For each natural number n, let $s(n) := \sum_{d|n, d<n} d$ be the sum of the proper divisors of n. If $s(n) = n$, then n is said to be a *perfect number*, while if $s(n) \neq n$ but $s(s(n)) = n$, then n is said to be *amicable*. More generally, let $s_0(n) = n$ and for $k > 0$, define $s_k(n) = s(s_{k-1}(n))$ if $s_{k-1}(n) > 0$. We call n *sociable* if $s_k(n) = n$ for some k.

The distribution of sociable numbers remains rather mysterious. For example, it is not known whether there are infinitely many sociable numbers, nor is it known if the set of sociable numbers has density zero. Some recent progress in the direction of the latter statement is reported in [7].

If we focus attention on solutions to $s_k(n) = n$ for a fixed value of k, then more can be said. Indeed, it is implicit in the work of Erdős [4] that such numbers have density zero. (This is also a special case of our Lemma 2 below.) The aim of this note is to prove the following two generalizations of this result:

Theorem 1. *For each fixed $k \geq 1$, the set of n for which $s_k(n)$ exists and is a multiple of n has density zero.*

Theorem 2. *For each fixed $k \geq 1$, the set of n for which $s_k(n)$ exists and is a divisor of n has density zero.*

2010 *Mathematics Subject Classification.* 11A25.

Key words and phrases. sociable number, amicable number, perfect number, multiamicable number, aliquot parts, sociable cycle.

The author is supported by NSF award DMS-0802970.

When $k = 1$, these theorems reduce to results of Kanold: Indeed, Theorem 1 asserts in this case that the set of *multiply perfect numbers* has density zero, which was established by Kanold [6] in 1955. As for Theorem 2, a moment's reflection on the definition of $s(n)$ reveals that $s(n)$ can only be a divisor of n if $s(n) = 1$ or if $s(n) = n$, i.e., if n is prime or if n is perfect. So Theorem 2 follows from elementary results in prime number theory and Kanold's earlier theorem [5] that the perfect numbers form a set of density zero.

2. Proof of Theorem 1

The following theorem of Erdős is the main result of [4]:

Theorem A. *Fix an integer $k \geq 1$ and an $\epsilon > 0$. For all n outside of a set of density zero the following holds: All of n, $s(n)$, $\ldots$, $s_k(n)$ are defined, and moreover*

$$\frac{s_{j+1}(n)}{s_j(n)} > \frac{s(n)}{n} - \epsilon \quad \textit{for all} \quad 1 \leq j < k.$$

It is immediate from Theorem A that, for each k, the set of n with $s_k(n) = 0$ has density zero. Indeed, if $s_k(n) = 0$, then $s_{k+1}(n)$ is undefined. Moreover, Theorem A easily implies that for each fixed k, the set of n with $s_k(n) = n$ has density zero (this argument is spelled out explicitly in [7]). Thus, to complete the proof of Theorem 1, it is enough to establish the following proposition:

Proposition 1. *For each fixed $k \geq 1$, the number of $n \leq x$ for which n divides $s_k(n)$ and $s_k(n)/n \geq 2$ is*

$$\ll_k \frac{x \log_3 x}{\log_2 x} (\log_4 x)^{2k}. \tag{1}$$

Here $\log_1 x := \max\{\log x, 1\}$, and $\log_{k+1} x = \max\{1, \log_k x\}$ for $k > 0$.

Remark. In [1], Cohen et al. study pairs m, n for which $s(m)$ divides n while $s(n)$ divides m; they call these *multiamicable pairs*. Their [1, Proposition 3] asserts that the set of numbers belonging to a multiamicable pair has density zero. Now if n divides $s_2(n)$, then n and $s(n)$ form a multiamicable pair, and so their Proposition 3 implies the case $k = 2$ of our Theorem 1. Their proof, which they attribute to the referee, was the starting point of our argument for Proposition 1.

The following lemma, due to Erdős, is proved in [2] but stated in a slightly weaker form there (see [2, Theorem 1]). Below, we write $\sigma(n) := \sum_{d|n} d$ for the sum of the positive divisors of n.

Lemma 1. *For each $x > 0$, the number of $n \le x$ with $\sigma(n)/n > y$ is*

$$\le x/\exp(\exp((e^{-\gamma} + o(1))y)), \quad \text{as } y \to \infty,$$

uniformly in x. Here γ is the Euler-Mascheroni constant.

We also need a lemma of Pomerance (see [8, Theorem 2]):

Lemma 2. *Let $x \ge 3$ and let m be any positive integer. The number of $n \le x$ for which $m \nmid \sigma(n)$ is $\ll x/(\log x)^{1/\phi(m)}$, where the implied constant is absolute.*

Proof of Proposition 1. Let Z denote expression appearing on the right of (1). Clearly we may assume $n > Z$. We may also assume that for all $0 \le j < k$, we have

$$\frac{s_{j+1}(n)}{s_j(n)} \le 2\log_4 x.$$

Indeed, if this fails, let j be the first index for which it fails. Then $m := s_j(n) \le x(2\log_4 x)^j$ and $s(m)/m > 2\log_4 x$. Let $\mathcal{A}$ be the set of $m \le x(2\log_4 x)^j$ for which $s(m)/m > 2\log_4 x$. By Lemma 1,

$$\#\mathcal{A} \le \frac{x(2\log_4 x)^k}{\exp(\exp((e^{-\gamma} + o(1))(2\log_4 x)))} \ll_k \frac{x}{(\log_2 x)^{k+2}}. \tag{2}$$

(Here we use that $2e^{-\gamma} > 1$.) Hence $s_k(n) = s_{k-j}(s_j(n)) \in \mathcal{B}$, where

$$\mathcal{B} := \{s_{k-j}(m) : m \in \mathcal{A} \text{ and } s_{k-j}(m) \text{ exists}\}.$$

From the classical determination of the maximal order of the sum-of-divisors function, we also have $s_k(n) \le X$, where

$$X := x(2\log_2 x)^k.$$

Since n divides $s_k(n) > 0$, the number of possibilities for n is at most $\sum_{r\in\mathcal{B}\cap[1,X]} d'(r)$, where $d'(r) := \sum_{d|r, r>Z} 1$ denotes the number of divisors of the natural number r that are at least Z. Trivially $d'(r) \le r/Z$, so that by (1) and (2),

$$\sum_{r\in\mathcal{B}\cap[1,X]} d'(r) \le (X/Z)\#\mathcal{B} \le (X/Z)\#\mathcal{A} \ll_k \frac{x}{\log_2 x \log_3 x (\log_4 x)^{2k}}.$$

Summing over $0 \le j < k$, we see the number of n of this type is $o(Z)$ and may be neglected.

So we may assume that if $s_k(n)$ divides n, then $\alpha := s_k(n)/n$ is an integer with $2 \le \alpha \le (2\log_4 x)^k$. We now fix an integer

$$2 \le \alpha \le (2\log_4 x)^k$$

and estimate the number of $n \le x$ which satisfy $s_k(n) = \alpha n$. Let L be a power of α chosen to satisfy

$$\frac{\log_2 x}{\log_3 x}(2\log_4 x)^{-k} < L \le \frac{\log_2 x}{\log_3 x}.$$

We can assume that L divides $\sigma(s_j(n))$ for all $0 \le j < k$. Indeed, if this is false for a certain value of $0 \le j < k$, then by Lemma 2 there are

$$\ll \frac{x(2\log_4 x)^k}{(\log x)^{1/L}} \le \frac{x(2\log_4 x)^k}{\log_2 x}$$

possibilities for $s_j(n)$. But $s_j(n)$ determines αn through the relation $\alpha n = s_k(n) = s_{k-j}(s_j(n))$. Summing over the k possibilities for j and the at most $(2\log_4 x)^k$ possibilities for α, we find that the number of n that can arise in this way is

$$\ll_k \frac{x(\log_4 x)^{2k}}{\log_2 x}, \tag{3}$$

which is again $o(Z)$ and so is negligible.

Assuming now that L divides each $\sigma(s_j(n))$, it follows that

$$s_{j+1}(n) = \sigma(s_j(n)) - s_j(n) \equiv -s_j(n) \pmod{L}$$

for all $0 \le j < k$. Hence

$$\alpha n = s_k(n) \equiv (-1)^k s_0(n) = (-1)^k n \pmod{L},$$

so that L divides $(\alpha + (-1)^{k+1})n$. Since L is a power of α, we have that L is coprime to $\alpha + (-1)^{k+1}$, so it must be that L divides n. Hence the number of possibilities for n is

$$\ll x/L \ll \frac{x\log_3 x}{\log_2 x}(2\log_4 x)^k. \tag{4}$$

Summing over $2 \le \alpha \le (2\log_4 x)^k$, we obtain a total of

$$\ll_k \frac{x\log_3 x}{\log_2 x}(\log_4 x)^{2k}$$

possible values of n. □

3. Proof of Theorem 2

Our proof of Theorem 2 depends on the following result:

Proposition 2. *Fix an integer $k \ge 1$ and a real number $\alpha > 0$. For $x \ge 3$, the number of solutions $n \le x$ to*

$$\frac{s_k(n)}{n} = \alpha$$

is $O_k(x/\log_3 x)$. The implied constant here depends only on k, and in particular is independent of α.

A weaker version of Proposition 2, without uniformity in α, would suffice for our application. But it seems that Proposition 2 as stated is of some interest in itself.

The following result is proved in [7]; see the proof of [7, Theorem 7].

Lemma 3. *For all sufficiently large x, there are sets $\mathcal{E}_1(x)$ and $\mathcal{E}_2(x)$ with*

$$\max\{\#\mathcal{E}_1(x), \#\mathcal{E}_2(x)\} \ll \frac{x}{(\log_2 x)^{1/4}}$$

and for which the following holds: If $n \le x$, then

$$\left|\frac{s(s(n))}{s(n)} - \frac{s(n)}{n}\right| \le \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}}$$

or $n \in \mathcal{E}_1(x)$ or $s(n) \in \mathcal{E}_2(x)$.

The next lemma is proved by Erdős in [3]:

Lemma 4. *Let ρ be any real number and $t > 1$. If $x > t$, then the number of $n \le x$ for which $\sigma(n)/n \in [\rho, \rho + 1/t)$ is $O(x/\log t)$. Here the implied constant is absolute.*

Proof of Proposition 2. When $k = 1$, one has the much stronger result, due to Wirsing [9], that the number of solutions $n \le x$ to $s(n)/n = \alpha$ is $O_\epsilon(x^\epsilon)$, and so we may assume that $k \ge 2$.

We may also assume that for all $0 \le j < k$, we have $s_{j+1}(n)/s_j(n) \le 2\log_4 x$. To see this, suppose this inequality fails, and let j be the minimal index for which it fails. Then as in the proof of Proposition 1, Lemma 1 forces $m := s_j(n)$ to belong to a set of size at most

$$\frac{x(2\log_4 x)^k}{\exp(\exp((e^{-\gamma} + o(1))(2\log_4 x)))} \le \frac{x(2\log_4 x)^k}{\log_2 x},$$

once x is large. But m determines $s_{k-j}(m) = s_k(n) = \alpha n$, which in turn determines n (since $\alpha \ne 0$), and so the same bound holds on the number of possibilities for n. Summing over j we see that a negligible number of n can arise this way, and so our assumption is validated.

In particular, we may assume that $\{n, s(n), \dots, s_k(n)\} \subset [1, X]$, for $X := x(2\log_4 x)^k$. We now consider two cases, according to whether $s(n)/n$ is particularly close to $\alpha^{1/k}$ or not. Suppose first that

$$\left|s(n)/n - \alpha^{1/k}\right| < k\frac{(\log_3 x)^2}{(\log_2 x)^{1/4}}; \tag{5}$$

then by Lemma 4, n belongs to a set of size $\ll_k x/\log_3 x$, and we are done. So we may suppose that (5) does not hold. By repeated application of Lemma 3, either one of $n, s(n), \dots, s_{k-2}(n)$ belongs to $\mathcal{E}_1(X)$, one of $s(n), s_2(n) \dots, s_{k-1}(n)$ belongs to $\mathcal{E}_2(X)$, or

$$\left|\frac{s_{j+1}(n)}{s_j(n)} - \frac{s_j(n)}{s_{j-1}(n)}\right| \le \frac{(\log_3 X)^2}{(\log_2 X)^{1/4}} < \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}} \tag{6}$$

for all $1 \le j < k$. Since each of $n, s(n), \dots, s_{k-1}(n)$ determines n, the exceptional possibilities give rise to

$$\ll_k \#\mathcal{E}_1(X) + \#\mathcal{E}_2(X) \ll \frac{X}{(\log_2 X)^{1/4}} \ll_k x\frac{(\log_4 x)^k}{(\log_2 x)^{1/4}}$$

values of n, which is negligible. If (6) holds for all $1 \le j < k$, then

$$\left|\frac{s_{j+1}(n)}{s_j(n)} - \frac{s(n)}{n}\right| \le j\frac{(\log_3 x)^2}{(\log_2 x)^{1/4}}$$

for all $0 \le j < k$. Since we are supposing that (5) fails, it follows that all the ratios $s_{j+1}(n)/s_j(n)$ lie strictly on the same side of $\alpha^{1/k}$. But then $s_k(n)/n = \prod_{0 \le j < k}(s_{j+1}(n)/s_j(n))$ cannot equal α. □

Proof of Theorem 2. Fix $\epsilon > 0$. Choose $u > 0$ so that the n for which $s(n)/n < u$ form a set of upper density at most ϵ. (To see that such a choice is possible, note that if $s(n)/n < u$, then n has no prime factors up to u^{-1}; the result now follows from an elementary sieve argument. Alternatively, the needed result follows directly from Lemma 4.) We claim that if $s(n)/n \ge u$ and $s_k(n)$ divides n, then n belongs to a set of density zero. It follows that the n for which $s_k(n)$ divides n comprise a set of upper density at most ϵ.

To prove the claim, suppose that $s(n)/n \ge u$. By Theorem A, after throwing away a set of density zero, we may assume that

$$s_{j+1}(n)/s_j(n) > u/2 \quad \text{for all } 1 \le j < k,$$

so that

$$\frac{n}{s_k(n)} = \prod_{j=0}^{k-1} \frac{s_j(n)}{s_{j+1}(n)} \le (2/u)^k.$$

Thus if $s_k(n)$ divides n, then $s_k(n)/n \in \{1/1, 1/2, \dots, 1/B\}$, where $B := \lfloor (2/u)^k \rfloor$. But Theorem 2 implies that the set of n with this property has density zero. □

Acknowledgements

I would like to thank Carl Pomerance and Kevin Ford for useful conversations during the preparation of this paper.

References

[1] G. L. Cohen, S. F. Gretton, and P. Hagis, Jr., *Multiamicable numbers*, Math. Comp. **64** (1995), 1743–1753.

[2] P. Erdős, *Some remarks about additive and multiplicative functions*, Bull. Amer. Math. Soc. **52** (1946), 527–537.

[3] ———, *On the distribution of numbers of the form $\sigma(n)/n$ and on some related questions*, Pacific J. Math. **52** (1974), 59–65.

[4] ———, *On asymptotic properties of aliquot sequences*, Math. Comp. **30** (1976), 641–645.

[5] H.-J. Kanold, *Über die Dichten der Mengen der vollkommenen und der befreundeten Zahlen*, Math. Z. **61** (1954), 180–185.

[6] ———, *Über mehrfach vollkommene Zahlen*, J. Reine Angew. Math. **194** (1955), 218–220.

[7] M. Kobayashi, P. Pollack, and C. Pomerance, *On the distribution of sociable numbers*, J. Number Theory **129** (2009), 1990–2009.

[8] C. Pomerance, *On the distribution of amicable numbers*, J. Reine Angew. Math. **293/294** (1977), 217–222.

[9] E. Wirsing, *Bemerkung zu der Arbeit über vollkommene Zahlen*, Math. Ann. **137** (1959), 316–318.

Mathematics Department, University of Illinois, 1409 West Green Street, Urbana, IL 61801

E-mail address: `pppollac@illinois.edu`