Journal of Number Theory ••• (••••) •••–•••

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

On the distribution of sociable numbers ☆

Mitsuo Kobayashi [a], Paul Pollack [b,*], Carl Pomerance [a]

[a] Mathematics Department, Dartmouth College, Hanover, NH 03755, United States
[b] Mathematics Department, University of Illinois at Urbana–Champaign, 1409 W. Green St., Urbana, IL 61801, United States

ARTICLE INFO

Article history:
Received 5 September 2008
Available online xxxx
Communicated by R.C. Vaughan

MSC:
11A25

Keywords:
Aliquot cycles
Sociable chains
Aliquot sequences
Sociable numbers
Perfect numbers
Amicable pairs

ABSTRACT

For a positive integer n, define $s(n)$ as the sum of the proper divisors of n. If $s(n) > 0$, define $s_2(n) = s(s(n))$, and so on for higher iterates. *Sociable* numbers are those n with $s_k(n) = n$ for some k, the least such k being the *order* of n. Such numbers have been of interest since antiquity, when order-1 sociables (perfect numbers) and order-2 sociables (amicable numbers) were studied. In this paper we make progress towards the conjecture that the sociable numbers have asymptotic density 0. We show that the number of sociable numbers in $[1, x]$, whose cycle contains at most k numbers greater than x, is $o(x)$ for each fixed k. In particular, the number of sociable numbers whose cycle is contained entirely in $[1, x]$ is $o(x)$, as is the number of sociable numbers in $[1, x]$ with order at most k. We also prove that but for a set of sociable numbers of asymptotic density 0, all sociable numbers are contained within the set of odd abundant numbers, which has asymptotic density about 1/500.

1. Introduction and statement of results

1.1. History

Let $s(n)$ denote the sum of the proper divisors of the positive integer n, so that $s(n) = \sigma(n) - n$. The study of the behavior of this arithmetic function has a long and rich history, going back to the ancient Greeks, who classified the positive integers as *deficient*, *perfect*, or *abundant*, according as $s(n)$ is less than, equal to, or greater than n, respectively. For example, 5 is deficient, 6 is perfect, and 12 is abundant.

☆ The authors were supported in part by NSF grants DMS-0703850 and DMS-0802970.
* Corresponding author.
E-mail addresses: mitskobay@dartmouth.edu (M. Kobayashi), pppollac@illinois.edu (P. Pollack), carl.pomerance@dartmouth.edu (C. Pomerance).

0022-314X/$ – see front matter
doi:10.1016/j.jnt.2008.10.011

Another concept steeped in history (some scholars have traced it to the Old Testament) is that of an amicable pair. Two distinct numbers m, n are called *amicable* (and are said to form an *amicable pair*) if $s(m) = n$ and $s(n) = m$. The first example consists of the pair 220 and 284.

It has seemed interesting to iterate the function s starting at an arbitrary natural number n. For example, if $n = 15$, the sequence is $15, 9, 4, 3, 1, 0$, and so the sequence terminates, while if $n = 6$ or $n = 220$, the sequence is purely periodic. Let $s_0(n) = n$, and inductively define $s_k(n) = s(s_{k-1}(n))$ when $k \geqslant 1$ and $s_{k-1}(n) > 0$. The *aliquot sequence* at n is $n, s_1(n), s_2(n), \ldots,$ which either terminates at 0 or is an infinite sequence. In 1888, Catalan [4] proposed the "empirical theorem" that all infinite aliquot sequences reach a perfect number. Perrott [27] promptly pointed out that this fails for the sequence starting at 220, and Dickson [8] later amended Catalan's conjecture to the claim that all aliquot sequences are bounded. This latter claim is now known as the Catalan–Dickson conjecture. Though we know no counterexamples, Guy and Selfridge [19] have made the counter-conjecture that in fact unbounded sequences are fairly common. The least n for which we do not know whether the aliquot sequence starting at n is unbounded is 276.

It is common to call a natural number n whose aliquot sequence is purely periodic a *sociable* number, with the least $k > 0$ with $s_k(n) = n$ called the *order* of n. Thus, perfect numbers are the sociable numbers of order 1 and amicable numbers are those of order 2. If n is sociable of order k, the set $\{n, s(n), \ldots, s_{k-1}(n)\}$ is called a *sociable k-cycle*.

There are many results on these topics from a statistical point of view. Building on work of Schoenberg dealing with Euler's function φ, Davenport [6] showed that the arithmetic function $\sigma(n)/n$ has a continuous distribution. That is, for each real number u, the set

$$\{n \in \mathbb{N}:\ \sigma(n)/n \leqslant u\}$$

has an asymptotic density, call it $D(u)$, and this function of u is continuous and strictly increasing for $u \geqslant 1$, with $D(1) = 0$ and $\lim_{u\to+\infty} D(u) = 1$. Thus, using this result, we see that the deficient numbers and the abundant numbers each have a positive asymptotic density, while the continuity of $D(u)$ implies that the perfect numbers have asymptotic density 0. It is now known after work of Behrend [2], Wall et al. [35,36], and Deléglise [7] that the density of the abundant numbers is slightly less than a quarter, between 0.2475 and 0.2480 (see also [18, p. 75]).

In 1954, Kanold [22] gave a "direct" proof that the set of perfect numbers has asymptotic density zero; that is, a proof that does not invoke the distribution function $D(u)$. After many intermediate results we now know after Hornfeck and Wirsing [21] and Wirsing [37] that the number of perfect numbers up to x is $O(x^{c/\log\log x})$ for some constant $c > 0$. This result is presumably still far from the truth.

Less is known about the distribution of amicable pairs, and unlike for perfect numbers, there are contrary opinions about what one should expect to hold. On the basis of numerical evidence up to 10^8, Bratley, Lunnon, and McKay [3] conjecture that $A(x)$, the number of amicable numbers up to x, is $o(x^{1/2})$. (These computations are extended in [32] and [17].) Interestingly, Erdős held the contrary belief that for each $\epsilon > 0$, one should have $A(x) > x^{1-\epsilon}$ for all sufficiently large x. The first proved result on the distribution of amicable numbers is that of Kanold [22], who showed that the amicable numbers have upper density < 0.204. Shortly thereafter, Erdős [12] established that the density is zero. In Pomerance [29] it is shown that

$$A(x) \leqslant x \exp\bigl(-(\log x)^{1/3}\bigr)$$

for all sufficiently large x. In particular, the sum of the reciprocals of the amicable numbers is finite. Lower bounds for their distribution appear much more difficult to obtain; though almost twelve million amicable pairs are known ([26]; see also [17]), we have no proof that there are infinitely many.

The question of whether there exist sociable numbers of order > 2 was raised by Meissner in 1907 [24]. The first examples, of length 5 and 28, were given by Poulet in 1918 [30]. Today there are 175 known aliquot cycles of length > 2, all but ten of which have length 4. See [25] for a complete list of known cycles together with the relevant references.

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

The goal of this paper is to study the asymptotic distribution of sociable numbers, and in particular to make progress towards the following conjecture.

Conjecture 1. *The set of sociable numbers has asymptotic density zero. That is, almost all numbers are not sociable.*

1.2. Results

In the same paper [12] where Erdős established that the amicable numbers have density zero, he remarked that a similar argument would establish the same result for the sociable numbers of order k, for any fixed k. These ideas were developed in [14]. In that paper Erdős introduces the following conjecture.

Conjecture A. *For each $\delta > 0$ and positive integer J, we have both*

$$\frac{s_{j+1}(n)}{s_j(n)} > \frac{s(n)}{n} - \delta \tag{1.1}$$

and

$$\frac{s_{j+1}(n)}{s_j(n)} < \frac{s(n)}{n} + \delta \tag{1.2}$$

for all $1 \leqslant j \leqslant J$, except for a set of numbers n of density zero.

Erdős [14] contains a proof that (1.1) holds for all $j \leqslant J$ except on a set of density zero. (Also see Lenstra [23].) For (1.2), a proof is claimed, but not given. This claim was later retracted in [16], where a proof of (1.2) was given for the case $J = 1$.

To see how Conjecture A relates to the distribution of sociable numbers, we recall the distribution function $D(u)$ discussed above. Fix $k \geqslant 1$, and suppose $\epsilon > 0$. Suppose $n \leqslant x$ is the smallest member of a sociable k-cycle. Notice that n is nondeficient, i.e., that $s(n)/n \geqslant 1$. Choosing $\delta > 0$ small enough, we can assume, at the cost of excluding at most ϵx values of $n \leqslant x$, that $s(n)/n \geqslant 1 + \delta$. (It is enough to choose δ so that $D(2+\delta) - D(2) < \epsilon$, which is possible by the continuity of $D(u)$.) But for each of these values of n, the lower inequality (1.1) fails for $j = k-1$. Indeed,

$$s\big(s_{k-1}(n)\big) = s_k(n) = n \leqslant s_{k-1}(n),$$

so that $s_{k-1}(n)$ is not abundant, contradicting (1.1) (with the same value of δ). It follows (from the *proved* half of Conjecture A) that the upper density of these n is at most ϵ, and so the upper density of the sociable numbers of order k is at most $k\epsilon$.

While this argument suffices to show that the sociable numbers of order k have density zero for each fixed k, it yields only very poor explicit upper bounds, and also fails to give a result which is uniform in any reasonable range of k. Our first theorem partially addresses these issues. Let $\log_1 x = \max\{\log x, 1\}$ and inductively define $\log_k x = \max\{\log(\log_{k-1} x), 1\}$.

Theorem 1.

(a) *The number of sociable cycles all of whose terms are contained in $[1, x]$ is at most $x/L(x)^{1+o(1)}$, where*

$$L(x) = \exp(\sqrt{\log_3 x \log_4 x}).$$

(b) *The number of sociable $n \leqslant x$ of order at most k is bounded by*

$$k(2\log_4 x)^k \frac{x}{L(x)^{1+o(1)}}.$$

Here the $o(1)$-term tends to zero as $x \to \infty$, uniformly in $k \geqslant 1$.

Part (a) of this theorem follows from the method used in [28] to study amicable pairs. Part (b) is obtained by inserting into that argument an estimate obtained by Erdős in his investigation into the behavior of the distribution function $D(u)$ for large u.

We can obtain results sharper than those of Theorem 1 if we restrict ourselves to a special class of sociable numbers. Let $\{n_1, \dots, n_k\}$ be a sociable cycle of order k, and let a be the greatest common divisor of the n_i. Following Cohen [5], we call $\{n_1, \dots, n_k\}$ a *regular* sociable cycle if a is a unitary divisor of each of the n_i, i.e., if $\gcd(a, n_i/a) = 1$ for all $1 \leqslant i \leqslant k$. A sociable number is called *regular* if it belongs to a regular cycle. This classification is of interest in light of a theorem of Dickson [8], which asserts the *ir*regularity of all sociable numbers of odd order $k > 1$. Since [8] is no longer easily accessible, we include a variant of Dickson's proof in Section 3.

Theorem 2.

(a) *The number of irregular sociable cycles all of whose terms are contained in* $[1, x]$ *is*

$$\ll \frac{x}{\sqrt{\log_2 x \log_3 x}}.$$

In particular, this estimate holds for the number of sociable cycles of odd order contained in $[1, x]$.

(b) *The number of irregular sociable numbers* $n \leqslant x$ *of order at most* k *is*

$$\ll k(2\log_4 x)^k \frac{x}{\sqrt{\log_2 x \log_3 x}},$$

where the implied constant is uniform in $k \geqslant 1$.

Theorem 1 provides a nontrivial upper bound on the number of sociable numbers of order at most k, as long as k does not exceed $(1-\epsilon)\sqrt{\log_3 x \log_4 x}/\log_5 x$. The bound of Theorem 2(b) for irregular sociable numbers is nontrivial for k up to $(\frac{1}{2} - \epsilon)\log_3 x/\log_5 x$.

We have been able to prove some additional results that lend support for Conjecture 1.

Theorem 3. *The set of deficient sociable numbers has density zero. In fact, the number of deficient sociable numbers up to x is* $O(x/L(x)^{1/12})$, *where* $L(x)$ *is as defined in Theorem* 1.

As we will see later, a plausible hypothesis on the behavior of $D(u)$ around $u = 2$ would permit one to improve the upper bound in Theorem 3 to $O(x(\log_3 x)^3/(\log_2 x)^{1/8})$.

Theorem 4. *All but* $O(x/\log_3 x)$ *sociable numbers* $n \leqslant x$ *belong to a cycle with more than* $K(x)$ *consecutive terms exceeding x, where* $K(x) = \frac{1}{9}\log_3 x/\log_6 x$.

Note that as a consequence of Theorem 4, almost no numbers up to x are sociable of order at most $K(x)$. As we discuss below, a version of Theorem 4 with a tighter bound on the exceptional set can be proved for *odd* sociable numbers, if one assumes that odd perfect numbers do not exist.

Theorem 5. *The set of even abundant sociable numbers has density zero. In fact, the number of even abundant sociable numbers not exceeding x is* $O(x/\log_3 x)$.

Taken together, Theorems 3 and 5 imply that all sociable numbers, except for a set of asymptotic density zero, are contained within the set of odd abundants. Call an odd abundant sociable number n a *special sociable* if the number preceding n in its cycle exceeds $nL(n)^{1/2} = n\exp(\frac{1}{2}\sqrt{\log_3 n \log_4 n})$. Our next theorem asserts that these special sociables form the only obstruction to establishing Conjecture 1.

Theorem 6.

(a) *The set of sociable numbers whose cycle contains no special sociable number has density zero. In fact, there are $O(x/\log_3 x)$ such numbers up to x.*
(b) *If the set of special sociable numbers has density zero, then Conjecture* 1 *holds.*

In the final section we make some further observations about the set of odd abundant sociable numbers. In addition we show that the odd abundant numbers have density about 1/500, while an upper bound for the upper density of the special sociable numbers is about 1/6000.

Notation. In addition to the usual notation of analytic number theory and our notation $\log_k$ for iterated logarithms (explained above), we find it convenient to give meaning to iterates of s indexed by negative integers. We let s_{-1} denote the inverse of the restriction of s to the sociable numbers, and for $j > 1$, we let s_{-j} be the jth iterate of s_{-1}. Thus, if n is sociable, $s_{-j}(n)$ is the unique sociable number m with $s_j(m) = n$. In general, we reserve the letters p, q, r for primes.

2. Preliminaries

2.1. Results from the literature

Lemma 1. *Let m be a positive integer. The number of $n \leqslant x$ for which $m \nmid \sigma(n)$ is*

$$\ll x/(\log x)^{1/\varphi(m)},$$

where the implied constant is absolute.

Proof. From [28, Theorem 2] we have that for all but $O(x/(\log x)^{1/\varphi(m)})$ numbers $n \leqslant x$, there is a prime $p \equiv -1 \pmod m$ for which $p \parallel n$. Then $m \mid p+1 \mid \sigma(n)$. □

Theorem A. *The number of positive integers $n \leqslant x$ for which n is abundant but $s(n)$ is deficient is at most*

$$x/L(x)^{1+o(1)}, \quad \text{as } x \to \infty, \tag{2.1}$$

where $L(x)$ is defined in Theorem 1. *The same upper bound holds on the number of sociable $n \leqslant x$ for which n is deficient but $s(n)$ is abundant.*

Sketch of the proof. The first half of Theorem A is established (implicitly) in the argument of [28] on amicable pairs. In that paper, the result appears with an unspecified constant in place of $(1 + o(1))$; however, as remarked there (see p. 221), this result is valid with $c + o(1)$ in the exponent, provided the number of primitive abundant numbers up to x is

$$\leqslant x/\exp\bigl((c+o(1))\sqrt{\log x \log\log x}\bigr). \tag{2.2}$$

Work of Avidon [1] shows that we may take $c = 1$ in (2.2). (We say n is *primitive abundant*—more technically, primitive nondeficient—if $s(n) \geqslant n$ and $s(d) < d$ for all $d \mid n$ with $d < n$.)

The second half of Theorem A is proved in a similar manner. Notice that since n is sociable, $m := s(n)$ determines n, so that it is enough to bound the number of possible values of m. Also, $m < n \leqslant x$. If n is deficient but m is abundant, then for some primitive abundant number a we have $a \mid m$ but $a \nmid n$; so $a \nmid \sigma(n)$. As in [28], we may assume that $a \leqslant (\log\log x)^{1-1/\log_4 x}$, since from (2.2) the number of $m \leqslant x$ with a primitive abundant divisor exceeding $(\log\log x)^{1-1/\log_4 x}$ is bounded by (2.1) (cf. the proof of [28, Theorem 3]). From Lemma 1, we have that the number of $n \leqslant x$ for which $a \nmid \sigma(n)$

is $\ll x/(\log x)^{1/\varphi(a)}$, with an absolute implied constant. Replacing $\varphi(a)$ by a and summing over the possibilities for a, we obtain a (crude) upper estimate of

$$\ll \frac{x \log\log x}{(\log x)^{1/(\log\log x)^{1-1/\log_4 x}}} = \frac{x \log\log x}{\exp(\exp(\log_3 x/\log_4 x))},$$

which is negligible compared to our target upper bound (2.1). □

The next few results concern the distribution of $\sigma(n)/n$ on the positive integers $n \leqslant x$.

Theorem B. *For every $x > 0$, the number of positive integers $n \leqslant x$ with $\sigma(n)/n > y$ is*

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

uniformly in x, with γ the Euler–Mascheroni constant.

Theorem B, which appears implicitly in [11], refines the well-known upper estimate

$$\sigma(n) \leqslant \bigl(e^{\gamma} + o(1)\bigr)n \log\log n.$$

If we take the estimate of Theorem B, divide by x, and let x tend to infinity, we obtain that

$$1 - D(u) \leqslant \exp\bigl(-\exp\bigl(\bigl(e^{-\gamma} + o(1)\bigr)u\bigr)\bigr) \quad (\text{as } u \to \infty),$$

and this "tail-estimate" is essentially Erdős's Theorem 1 in [11]. However, a careful reading of the proof of that theorem reveals that his argument actually proves the more uniform estimate of Theorem B.

The next result is the main theorem of Erdős's paper [13]. Its applicability to the study of sociable numbers was noted already by Erdős and Rieger [15] (see also [31]), who used it to prove that there are $O(x/\log_3 x)$ amicable numbers up to x.

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

If we fix a number ρ in the range of $\sigma(n)/n$, then, as discussed in [13], Theorem C is in some sense best possible. However, if ρ is the right-endpoint of our interval rather than the left, then we can do better. The following estimate is due to Toulmonde (cf. [33, Théorème 1] and the discussion in §10 of that paper).

Theorem D. *Fix a number ρ in the range of $\sigma(n)/n$. Then for t sufficiently large (which may depend on ρ) and $x > 0$, the number of $n \leqslant x$ for which $\sigma(n)/n \in [\rho - 1/t, \rho)$ is*

$$\ll \frac{x}{\exp(\frac{1}{5}\sqrt{\log t \log\log t})}.$$

Here the implied constant is absolute.

2.2. Conjecture A revisited

As mentioned above, Erdős's Conjecture A is a theorem in the special case when $J = 1$. That is, on a set of n of density 1, we have $s_2(n)/s(n) = s(n)/n + o(1)$. For our purposes it is important to have an explicit version of this result for sociable numbers:

Theorem 7. *For all but $O(x/(\log_2 x)^{1/4})$ sociable numbers $n \leqslant x$, we have*

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

Proof. The argument borrows heavily from the proof of [16, Theorem 5.2]. We may suppose that $\sigma(n)$ is divisible by all primes and prime powers up to $A := \frac{2}{3}\log_2 x/\log_3 x$, since by Lemma 1 the exceptions up to x make up a set of size

$$\ll \sum_{\substack{p^a \leqslant A \\ p \text{ prime},\ a \geqslant 1}} \frac{x}{(\log x)^{1/\varphi(p^a)}} \leqslant \frac{x}{(\log x)^{1/A}} \sum_{\substack{p^a \leqslant A \\ p \text{ prime},\ a \geqslant 1}} 1, \tag{2.3}$$

which is

$$\ll \frac{x}{(\log_2 x)^{1/2}(\log_3 x)^2},$$

and so is negligible. We may further assume that the prime factorizations of n and $s(n)$ agree on all primes up to $B := A^{1/2}$. (That is, $v_p(n) = v_p(s(n))$ for all $p \leqslant B$, where the p-adic valuation v_p is defined so that $p^{v_p(m)} \parallel m$.) Indeed, if n and $s(n)$ fail to agree at the prime p, then we must have $v_p(n) \geqslant v_p(\sigma(n))$. But for the numbers n under consideration, $v_p(\sigma(n)) \geqslant e_p$, where e_p is an integer such that

$$p^{e_p} \leqslant A = \frac{2}{3}\log_2 x/\log_3 x < p^{e_p+1}.$$

The numbers n which fail to agree with $s(n)$ at some prime $p \leqslant B$ thus make up a set of size at most $\sum_{p \leqslant B} x/p^{e_p}$. We estimate this sum by considering those terms with $e_p = 2$ and $e_p > 2$ separately. (Notice that $e_p \geqslant 2$ for all primes p by our choice of B.) The terms with $e_p = 2$ contribute

$$\ll \sum_{A^{1/3} < p \leqslant B} \frac{x}{p^2} \ll \frac{x}{A^{1/3}\log A},$$

while those with $e_p > 2$ contribute

$$\ll x \sum_{p \leqslant A^{1/3}} \frac{1}{p^{e_p}} \leqslant \frac{x}{A} \sum_{p \leqslant A^{1/3}} p \ll \frac{x}{A^{1/3}\log A}.$$

Once again this is negligible.

Let $n \leqslant x$ be a sociable number. By Theorem B, we have $\sigma(n)/n > 2\log B$ for at most

$$x/\exp\bigl(\exp\bigl(\bigl(2e^{-\gamma} + o(1)\bigr)\log B\bigr)\bigr) < x/\exp(B) = o(x/\log_2 x)$$

values of $n \leqslant x$. So we can assume $\sigma(n)/n \leqslant 2\log B$. Write $n = m_0 n_0$ and $s(n) = m_1 n_1$, where m_0 and m_1 are the B-smooth parts of n and $s(n)$, respectively. (By the 'B-smooth part' we mean the largest divisor supported on the primes up to B.)

We can also assume that

$$\sigma(n_0)/n_0 < \exp(1/\sqrt{B}).$$

To see this we use an averaging argument: Observe that

$$\frac{\sigma(n_0)}{n_0} \leqslant \prod_{p|n,\ p>B} (1-1/p)^{-1} \leqslant \exp\Bigg(\sum_{p|n,\ p>B} \frac{1}{p-1}\Bigg), \tag{2.4}$$

so that any n for which $\sigma(n_0)/n_0 \geqslant \exp(1/\sqrt{B})$ satisfies

$$\sum_{p|n,\ p>B} \frac{1}{p-1} \geqslant \frac{1}{\sqrt{B}}.$$

But

$$\sum_{n\leqslant x} \sum_{p|n,\ p>B} \frac{1}{p-1} \leqslant \sum_{B<p\leqslant x} \frac{x}{p(p-1)} \ll \frac{x}{B\log B},$$

and so there can be

$$\ll \frac{x}{\sqrt{B}\log B} \ll \frac{x}{(\log_2 x)^{1/4}(\log_3 x)^{3/4}}$$

such values of $n \leqslant x$, which fits within our final bound for the exceptional set.

Now assuming all the above conditions are satisfied, we have

$$\begin{aligned} \frac{s(n)}{n} - \frac{s_2(n)}{s(n)} &= \frac{\sigma(m_0)}{m_0}\left(\frac{\sigma(n_0)}{n_0} - \frac{\sigma(n_1)}{n_1}\right) \\ &\leqslant (2\log B)\big(\exp(1/\sqrt{B}) - 1\big) < 3\frac{\log B}{\sqrt{B}} < \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}} \end{aligned}$$

for large x, which proves one of the inequalities implicit in the theorem statement. Notice that in proving this half of the theorem we have not needed the hypothesis that n is sociable.

We can prove the other half in the same way, beginning with the identity

$$\frac{s_2(n)}{s(n)} - \frac{s(n)}{n} = \frac{\sigma(m_0)}{m_0}\left(\frac{\sigma(n_1)}{n_1} - \frac{\sigma(n_0)}{n_0}\right),$$

if we show that $\sigma(n_1)/n_1 < \exp(1/\sqrt{B})$ for all but $O(x/(\log_2 x)^{1/4})$ values of $n \leqslant x$. To prove this, notice that we can assume $s(n) \leqslant 2x\log_4 x$, since by Theorem B this fails for only $O(x/\log_2 x)$ values of $n \leqslant x$. But if $\sigma(n_1)/n_1 \geqslant \exp(1/\sqrt{B})$, then the same averaging argument employed above shows that the number of possibilities for $s(n) \leqslant 2x\log_4 x$ is $\ll x\log_4 x/(\sqrt{B}\log B)$, which is again $\ll x/(\log_2 x)^{1/4}$. Since n is sociable, $s(n)$ determines n, so that this is also an upper bound on the number of possibilities for n. $\square$

We shall frequently use Theorem 7 for a string of consecutive terms of a sociable cycle to show that usually these terms behave approximately like a geometric progression. In particular, we have the following result.

Corollary 1. *For each positive integer J, all but $O(Jx/(\log x)^{1/4})$ sociable numbers n with $\{n, s(n), \dots, s_{J-1}(n)\} \subset [1, x]$ have*

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

for each integer $1 \leqslant j \leqslant J$.

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

3. Proofs of Theorems 1 and 2

Proof of Theorem 1. As mentioned in the introduction, there are much stronger estimates known for the number of perfect numbers up to x, so that we may restrict attention to sociable numbers of order > 1. Observe that every sociable cycle of order > 1 contains an abundant number n for which $s(n)$ is deficient (e.g., the element preceding the largest term in the cycle). Theorem 1(a) now follows immediately from Theorem A, since a cycle is determined by any one of its elements.

For the proof of (b), it is enough to show that the number of cycles of length at most k which contain a term not exceeding x is

$$\ll (2\log_4 x)^k x/L(x)^{1+o(1)}. \tag{3.1}$$

Suppose we are given such a cycle. Consider first the case when this cycle contains a term n for which $s(n)/n > 2\log_4 x$. We may suppose n is the first term of the cycle for which this happens, where we view the cycle as starting with its smallest term. Then $n \leqslant x(2\log_4 x)^k$, and so by Theorem B, the number of possibilities for n (and hence the cycle containing n) is at most

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

which satisfies the bound (3.1). (Notice that $2e^{-\gamma} > 1$.) If, on the other hand, $s(n)/n \leqslant 2\log_4 x$ for all terms n of the cycle, then the cycle is entirely contained in $[1, x(2\log_4 x)^k]$ and the result follows from (a). □

Proof of Theorem 2. It is enough to prove part (a), as then the proof of (b) follows exactly the same pattern as the proof of Theorem 1(b).

Let $\{n_1, \ldots, n_k\}$ be an irregular cycle of length k entirely contained in $[1, x]$, and put $a := \gcd(n_1, \ldots, n_k)$. Note that necessarily $k > 1$. As established in the proof of Theorem 7, we have that $\sigma(n)$ is divisible by all primes and prime powers $p^j \leqslant \frac{2}{3}\log_2 x/\log_3 x$ for all but $O(x/((\log_2 x)^{1/2}(\log_3 x)^2))$ values of $n \leqslant x$. So we can assume each $\sigma(n_i)$ is divisible by all these small prime powers.

We take two cases. First suppose $a = 1$. Then no n_i can have a prime factor up to $\frac{2}{3}\log_2 x/\log_3 x$, since otherwise a simple induction shows that p divides n_j for every j and so also divides a. Clearly n_i is abundant for some i, and so for this i we have (by (2.4))

$$\sum_{p \mid n_i,\ p > \frac{2}{3}\frac{\log_2 x}{\log_3 x}} \frac{1}{p-1} > \log 2.$$

But

$$\sum_{n \leqslant x} \sum_{p \mid n,\ p > \frac{2}{3}\frac{\log_2 x}{\log_3 x}} \frac{1}{p-1} \leqslant x \sum_{p > \frac{2}{3}\frac{\log_2 x}{\log_3 x}} \frac{1}{p(p-1)} \ll \frac{x}{\log_2 x},$$

so that the number of possibilities for n_i (and so for its cycle) is $\ll x/\log_2 x$.

Now suppose $a > 1$. Write $n_i = am_i$ for $1 \leqslant i \leqslant k$. We can suppose that none of the n_i have a prime power divisor p^j, $j \geqslant 2$, with $p^j \geqslant \frac{2}{3}\log_2 x/\log_3 x$, as otherwise n_i belongs to a set of size

$$\ll \frac{x}{\log_3 x\sqrt{\log_2 x/\log_3 x}} = \frac{x}{\sqrt{\log_2 x \log_3 x}}.$$

Cycling the n_i around if necessary, we can assume that a is not a unitary divisor of n_1. So there is a prime p dividing both a and m_1. Let p^e be the exact power of p dividing a; then p^{e+1} divides n_1, so that $p^{e+1} \leqslant \frac{2}{3} \log_2 x/\log_3 x$. This implies that p^{e+1} divides each $\sigma(n_i)$, which in turn shows that p^{e+1} divides each n_i. Thus p^{e+1} divides a, contrary to the choice of e. □

As promised, we present the proof of Dickson's result from the introduction:

Theorem E. *Let $\{n_1, \ldots, n_k\}$ be a sociable cycle of odd length $k > 1$. Suppose $a := \gcd(n_1, \ldots, n_k) > 1$. Then a cannot be a unitary divisor of all of $n_1, \ldots, n_k$.*

Proof. We assume the n_i are numbered so that $n_{i+1} = s(n_i)$, where the indices are taken modulo k. Write $\sigma(a)/a = b/c$ in lowest terms, and write $n_i = am_i$ for $1 \leqslant i \leqslant k$. Suppose that a is a unitary divisor of each n_i. Then $\sigma(n_i) = (ab/c)\sigma(m_i)$ for each i, and since $n_{i+1} = \sigma(n_i) - n_i$, we find that

$$c(m_i + m_{i+1}) = \frac{c}{a}(n_i + n_{i+1}) = \frac{c}{a}\sigma(n_i) = b\sigma(m_i).$$

Since b and c are relatively prime, it follows that b divides $m_i + m_{i+1}$, i.e., $m_i \equiv -m_{i+1} \pmod{b}$. Since k is odd, iterating this observation shows that modulo b,

$$m_i \equiv -m_{i+1} \equiv m_{i+2} \equiv \cdots \equiv -m_{i+k} = -m_i,$$

and so b divides $2m_i$. Since this holds for each i and $\gcd(m_1, \ldots, m_k) = 1$, we must have either $b = 1$ or $b = 2$. If $b = 1$, then $\sigma(a)/a = 1$, and this contradicts $a > 1$. So $b = 2$, and since $c < b$, we have $c = 1$. Hence $\sigma(a)/a = 2$ and a is perfect. Since a is a common divisor of the n_i, each n_i is nondeficient. But any sociable chain composed entirely of nondeficient numbers consists of a single perfect number, and this contradicts that $k > 1$. □

4. Proofs of Theorems 3–6

Proof of Theorem 3. Put $\delta := (\log_3 x)^3/(\log_2 x)^{1/8}$. We can assume $s(n)/n \leqslant 1 - \delta$, since by Theorem D, the number of $n \leqslant x$ for which $s(n)/n \in (1 - \delta, 1)$ is

$$\leqslant x \exp\left(-\left(\frac{1}{5\sqrt{8}} + o(1)\right)\sqrt{\log_3 x \log_4 x}\right).$$

Let $J := \lceil (\log_2 x)^{1/8} \rceil$. By Corollary 1, for all but

$$\ll Jx/(\log_2 x)^{1/4} \ll x/(\log_2 x)^{1/8}$$

values of $n \leqslant x$, we have

$$s_j(n)/s_{j-1}(n) \leqslant 1 - \delta + (j-1)(\log_3 x)^2/(\log_2 x)^{1/4}$$

for all $1 \leqslant j \leqslant J$. So we can assume these J conditions all hold. The right-hand side of this inequality is always at most $1 - \delta/2$, which implies that $s_J(n) \leqslant x(1 - \delta/2)^J$. But s_J is injective on the set of sociable numbers, so that the number of sociable n remaining is at most

$$x(1 - \delta/2)^J \leqslant x \exp(-J\delta/2) \leqslant x \exp\left(-\frac{1}{2}(\log_3 x)^3\right),$$

which is negligible.

This proves the slight weakening of Theorem 3 where 12 is replaced by $5\sqrt{8} + o(1)$. To see that the theorem is correct as stated, we note that according to Toulmonde [33, Théorème 1, Remarque 1], the constant $1/5$ in Theorem D can be replaced by any constant smaller than $1/4$. Then the above analysis shows that we may replace $5\sqrt{8}$ by $4\sqrt{8} < 12$. □

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

Remarks. If ρ belongs to the range of $\sigma(n)/n$, then it is not hard to prove that the right-hand derivative of D at ρ is infinite. By contrast, very little is known about the left-hand derivative (see [13, pp. 59–60]). Toulmonde has conjectured (see [34, Eq. (4)]) that the left-hand derivative always vanishes for these ρ. This would fit well with Erdős's result [10] that $D(u)$ is singular, i.e., that $D' = 0$ almost everywhere.

Suppose that the difference quotient of $D(u)$ remains bounded as one approaches the particular value $\rho = 2$ from the left. (This is quite a bit weaker than Toulmonde's conjecture.) By a theorem of Elliott [9, Theorem 5.6],

$$\frac{1}{x}\sum_{\substack{n\leqslant x\\ \sigma(n)/n\leqslant u}} 1 = D(u) + O\left(\frac{\log_2 x}{\log x \log_3 x}\right)$$

for all u. Applying this estimate with $u = 2$ and $u = 2 - \delta$, with δ as in the proof of Theorem 3, shows that the number of $n \leqslant x$ for which $s(n)/n \in (1-\delta, 1)$ is

$$\ll x(\log_3 x)^3/(\log_2 x)^{1/8}. \tag{4.1}$$

The rest of the proof of Theorem 3 goes through unchanged and shows that (4.1) serves as an upper bound for the number of deficient sociable numbers $n \leqslant x$. Moreover, since every sociable cycle of order > 1 contains a deficient term, we deduce that this is also an upper bound for the number of sociable cycles contained entirely in $[1, x]$; i.e., we have a conditional improvement of Theorem 1(a).

Proof of Theorem 4. Define δ and J as in the proof of Theorem 3. By Theorem 3 and Theorems B and C, at the cost of excluding $O(x/\log_3 x)$ values of n, we may restrict our attention to those $n \leqslant x$ for which

$$1 + \delta \leqslant s(n)/n \leqslant 2\log_5 x - 1.$$

Put $y := x(2\log_5 x)^K$ where $K = K(x) = \frac{1}{9}\log_3 x/\log_6 x$, so that

$$y = x(\log_2 x)^{1/9+o(1)}. \tag{4.2}$$

We may assume that for all $1 \leqslant j < J$ for which $s_{j-1}(n) \leqslant y$, we have

$$\left|\frac{s(s_j(n))}{s_j(n)} - \frac{s_j(n)}{s_{j-1}(n)}\right| < \frac{(\log_3 y)^2}{(\log_2 y)^{1/4}} < \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}}, \tag{4.3}$$

since by Theorem 7, the number of sociable n for which this fails is

$$\ll J\frac{y}{(\log_2 y)^{1/4}} \ll \frac{y}{(\log_2 x)^{1/8}} = \frac{x}{(\log_2 x)^{1/72+o(1)}},$$

which is negligible.

We can further assume that $\{n, s(n), \ldots, s_J(n)\}$ is not entirely contained in the interval $[1, y]$. Indeed, if it is, then by (4.3), we have

$$s_j(n)/s_{j-1}(n) \geqslant 1 + \delta - (j-1)(\log_3 x)^2/(\log_2 x)^{1/4} \geqslant 1 + \delta/2$$

for all $1 \leqslant j \leqslant J$, so that

$$n \leqslant s_J(n)(1+\delta/2)^{-J} \leqslant y\exp\left(\left(-\frac{1}{2} + o(1)\right)(\log_3 x)^3\right).$$

This upper bound is $o(x/\log_3 x)$ by (4.2), and so these n may be ignored.

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

Say $1 \leqslant j \leqslant J$ is such that $s_j(n) > y = x(2\log_5 x)^K$. Since the cycle containing n has at most K consecutive terms exceeding x, there is some $1 \leqslant k \leqslant j$ with $s_k(n)/s_{k-1}(n) > 2\log_5 x$. Moreover, if k is the smallest such index, we must have $\{n, s(n), s_2(n), \ldots, s_{k-1}(n)\} \subset [1, y]$. But this is impossible, as (4.3) implies that

$$\begin{aligned} s_k(n)/s_{k-1}(n) &\leqslant s(n)/n + (k-1)(\log_3 x)^2/(\log_2 x)^{1/4} \\ &\leqslant 2\log_5 x - 1 + o(1) < 2\log_5 x. \end{aligned}$$

This completes the proof. □

Remark. The proof of Theorem C (cf. [33, Lemme 7]) shows that for t sufficiently large, if $n \leqslant x$ satisfies $\sigma(n)/n \in [2, 2+1/t)$, then either n belongs to an exceptional set of size

$$\ll x/\exp\left(\frac{1}{5}\sqrt{\log t \log\log t}\right)$$

or n has a small perfect number divisor. If, as is conjectured, there do not exist odd perfect numbers, then every odd n automatically has no perfect divisors. This gives a conditional improvement of Theorem C for $\rho = 2$ and odd n, which in turn allows one to establish the following variant of Theorem 4: *If no odd perfect numbers exist, then all but $O(x/L(x)^{1/5\sqrt{8}+o(1)})$ odd sociable numbers $n \leqslant x$ belong to a cycle with more than $\frac{1}{9}\log_3 x/\log_5 x$ consecutive terms exceeding x.* For the same reasons as in the proof of Theorem 3, one can replace the exponent on $L(x)$ by $1/12$. We leave the details to the reader.

Proof of Theorem 5. The idea of the proof is the same as that of Theorem 3, but now we trace backwards through our cycle instead of forwards. Choose δ and J as in the proof of Theorem 3. We can restrict attention to even abundant sociable n which satisfy $s(n)/n \geqslant 1+\delta$, since by Theorem C the exceptional n make up a set of size $\ll x/\log_3 x$.

In analogy with the proof of Theorem 3, we would like to prove that for all but $O(x/\log_3 x)$ of the remaining n, we have $s_{-(j-1)}(n)/s_{-j}(n) \geqslant 1+\delta/2$ for all $1 \leqslant j \leqslant J$. We prove something a little stronger: all but $O(x/\log_3 x)$ of the even abundant sociable $n \leqslant x$ satisfying $s(n)/n \geqslant 1+\delta$ also satisfy

$$s_{-j}(n) \quad \text{is even,} \quad \text{and} \quad \frac{s(s_{-j}(n))}{s_{-j}(n)} \geqslant \frac{s(s_{-(j-1)}(n))}{s_{-(j-1)}(n)} - \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}} \tag{4.4}$$

for every $1 \leqslant j \leqslant J$. To see this, suppose we have an n for which (4.4) fails for some $1 \leqslant j \leqslant J$, and let $j(n)$ be the least integer where either of the conditions of (4.4) is violated. Set $m = s_{-j(n)}(n)$. Since both conditions hold up to $j(n)-1$, we have

$$\frac{s_{-(i-1)}(n)}{s_{-i}(n)} \geqslant 1+\delta - i\frac{(\log_3 x)^2}{(\log_2 x)^{1/4}} \geqslant 1+\frac{\delta}{2}$$

for all $1 \leqslant i \leqslant j-1$. So

$$s(m) = s_{-(j-1)}(n) = n\prod_{i=1}^{j-1}\frac{s_{-i}(n)}{s_{-(i-1)}(n)} \leqslant x(1+\delta/2)^{-(j-1)},$$

and in particular $s(m) \leqslant x$. Suppose that at step j, it is the first condition of (4.4) that is violated, so that m is odd. Since $s(m) = s_{-(j-1)}(n)$ is even, it must be that $\sigma(m)$ is odd, forcing $m = l^2$ for some odd l. But the number of values of l for which $s(l^2) \leqslant x$ is $\ll x/\log x$. (Indeed, $s(l^2) \geqslant l^{3/2}$ unless l is prime, in which case $s(l^2) = l + 1$.) So m can assume at most $O(x/\log x)$ possible values, and the same is therefore true for $n = s_j(m)$. If the first condition of (4.4) holds but the second is violated, then

$$\frac{m}{2} \leqslant s(m) \leqslant x,$$

so that $m \leqslant 2x$ and

$$\frac{s(m)}{m} < \frac{s(s(m))}{s(m)} - \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}} < \frac{s(s(m))}{s(m)} - \frac{(\log_3 (2x))^2}{(\log_2 (2x))^{1/4}}.$$

By Theorem 7, there are only $O(x/(\log_2 x)^{1/4})$ possibilities for m and so also for $n = s_j(m)$. Summing over $j \leqslant J$, we see that the number of possibilities for n violating one of the conditions (4.4) is

$$\ll J\frac{x}{\log x} + J\frac{x}{(\log_2 x)^{1/4}} \ll \frac{x}{(\log_2 x)^{1/8}},$$

which is $o(x/\log_3 x)$.

So we can assume n is such that (4.4) holds for all $j \leqslant J$. But then

$$s_{-J}(n) \leqslant n(1+\delta/2)^{-J} \leqslant x\exp\left(-\left(\frac{1}{2} + o(1)\right)(\log_3 x)^3\right),$$

which is $o(x/\log_3 x)$. Since s_{-J} is injective on the set of sociable numbers, the number of these n is also $o(x/\log_3 x)$. $\square$

Proof of Theorem 6(a). By Theorem 3 we may restrict attention to abundant n. Choose δ and J as in the proof of Theorem 5. Then all but $O(x/\log_3 x)$ of the abundant $n \leqslant x$ satisfy $s(n)/n \geqslant 1 + \delta$, so that we can restrict attention to these n. We claim that all but $o(x/\log_3 x)$ of the $n \leqslant x$ under consideration in this theorem satisfy the second condition of (4.4), i.e.,

$$\frac{s(s_{-j}(n))}{s_{-j}(n)} \geqslant \frac{s(s_{-(j-1)}(n))}{s_{-(j-1)}(n)} - \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}} \tag{4.5}$$

for all $1 \leqslant j \leqslant J$. Once this is established, the proof is completed just as for Theorem 5.

We partition the n which do not satisfy the entire sequence of inequalities (4.5) into sets $\mathcal{S}_j$ according to the first index $j \leqslant J$ for which (4.5) is violated, and we estimate the size of each $\mathcal{S}_j$. Let n be an element of $\mathcal{S}_j$ and put $m := s_{-j}(n)$. As in the proof of Theorem 3, we have that $s(m)$ is abundant and bounded by x.

Let us show that for all but $x/L(x)^{1/2+o(1)}$ elements $n \in \mathcal{S}_j$, the integer m is abundant. Since the cycle corresponding to n does not contain any special sociable numbers, the integer $s(m)$ is not special, and so (using $s(m) \leqslant x$)

$$m = s_{-1}\bigl(s(m)\bigr) \leqslant s(m)L\bigl(s(m)\bigr)^{1/2} \leqslant xL(x)^{1/2} =: y,$$

say. But by Theorem A, the number of sociable $m \leqslant y$ for which m is deficient but $s(m)$ is abundant is

$$\leqslant y/L(y)^{1+o(1)} = x/L(x)^{1/2+o(1)}.$$

Thus m, and hence also $n = s_j(m)$, can assume at most $x/L(x)^{1/2+o(1)}$ values.

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

Suppose now that $n \in S_j$ is such that m is abundant. Then $m < s(m) \leqslant x$. Since (4.5) fails to hold, we have

$$\frac{s(m)}{m} \leqslant \frac{s(s(m))}{s(m)} - \frac{(\log_3 x)^2}{(\log_2 x)^{1/4}}.$$

But then the number of possibilities for m, and hence also for $n = s_j(m)$, is $O(x/(\log_2 x)^{1/4})$ by Theorem 7.

Summing over $1 \leqslant j \leqslant J$, we find that the total number of possibilities for n for which (4.5) fails to hold for some $j \leqslant J$ is

$$\ll Jx/L(x)^{1/2+o(1)} + Jx/(\log_2 x)^{1/4} \ll x/(\log_2 x)^{1/8},$$

which is $o(x/\log_3 x)$. □

Proof of Theorem 6(b). We are to show that if the set of special sociables has density zero, then the set of sociable numbers also has density zero. By Theorem 3 we can restrict our attention to abundant n.

Let $\epsilon > 0$. By the continuity of the distribution function $D(u)$, we may fix $\delta > 0$ so that $D(2+\delta) - D(2) < \epsilon$. Then for large x, we have $s(n)/n \geqslant 1 + \delta$ for all but ϵx abundant $n \leqslant x$. Now let J be a large, fixed positive integer, to be specified more precisely momentarily. We claim that for all but $o(x)$ of the abundant sociable $n \leqslant x$ satisfying $s(n)/n \geqslant 1 + \delta$, we have

$$\frac{s(s_{-j}(n))}{s_{-j}(n)} \geqslant \frac{s(s_{-(j-1)}(n))}{s_{-(j-1)}(n)} - \frac{\delta}{2J} \tag{4.6}$$

for all $1 \leqslant j \leqslant J$. Notice that for such n, we have $s_{-(j-1)}(n)/s_{-j}(n) \geqslant 1 + \delta/2$ for all $1 \leqslant j \leqslant J$, which implies that

$$s_{-J}(n) \leqslant x(1+\delta/2)^{-J}.$$

Since s_{-J} is injective on sociables, it will follow that the upper density of the sociable numbers is at most

$$\epsilon + (1+\delta/2)^{-J},$$

which is less than 2ϵ if J was fixed sufficiently large. So it is enough to prove the claim.

As in the proof of (a), we partition those n which fail to satisfy (4.6) for all $1 \leqslant j \leqslant J$ into sets $\mathcal{S}_j$ according to the first index for which (4.6) fails to hold. Since J is fixed, to prove our claim above it is enough to show that each $\mathcal{S}_j$ has size $o(x)$. For each j, let $\mathcal{S}'_j$ be that subset of $\mathcal{S}_j$ consisting of those n for which $s(s_{-j}(n))$ is not special; essentially the same argument as given in (a) shows that $\mathcal{S}'_j$ contains only $o(x)$ elements. But if $n \in \mathcal{S}_j \setminus \mathcal{S}'_j$, then $s_{-(j-1)}(n)$ is a special sociable number l (say), not exceeding x. But, by assumption, there are only $o(x)$ possibilities for l, and so also for $n = s_{j-1}(l)$. □

5. Remarks on odd abundant numbers

5.1. Nonsociable odd abundants

Though we cannot prove that a positive proportion of odd abundant numbers are not sociable, we can prove that infinitely many of them are not sociable. In fact, we can prove that the sum of the reciprocals of the nonsociable odd abundants is infinite.

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

Theorem 8. *The number of nonsociable odd abundant numbers in* $[1, x]$ *is at least* $x/(\log x)^{1/2+o(1)}$ *as* $x \to \infty$.

Proof. Let m be an arbitrary but fixed odd abundant number with the property that $\gcd(\sigma(m), m) = 1$. (Note that if m_0 is any squarefree odd abundant number and $p > m_0$ is any prime, then $m = m_0^{p-1}$ is odd abundant and $\gcd(\sigma(m), m) = 1$.) Let $\mathcal{A}$ denote the set of integers a such that $\gcd(a, 2m\sigma(m)) = 1$ and $\gcd(\sigma(a), am) = 1$. Let

$$\rho(m) = \prod_{\substack{q|m \\ q \text{ prime}}} \left(1 - \frac{1}{q-1}\right).$$

We first show that

$$\sum_{\substack{a \leqslant x \\ a \in \mathcal{A}}} \frac{1}{a} \gg \frac{(\log x)^{\rho(m)}}{(\log_2 x)^{5/2}}. \tag{5.1}$$

Since we are looking for a lower bound for $\sum 1/a$ it suffices to consider just the squarefree members of $\mathcal{A}$. Such numbers a have the properties that for each prime $r \mid a$ we have $r \nmid 2m\sigma(m)$ and $\gcd(r + 1, m) = 1$. Let $\mathcal{R}$ be the set of these primes r, so that the density of $\mathcal{R}$ within the set of primes is $\rho(m)$. Let

$$u = \lfloor \rho(m) \log_2 x - 2 \log_3 x \rfloor.$$

Then, for x sufficiently large, we have

$$R_1 := \sum_{\substack{r \in \mathcal{R} \\ (\log_2 x)^2 < r \leqslant x^{1/u}}} \frac{1}{r} > u. \tag{5.2}$$

If we choose u distinct primes from $\mathcal{R}$ that are in $((\log_2 x)^2, x^{1/u}]$, then their product a will be at most x, and $a \in \mathcal{A}$ provided $\gcd(\sigma(a), a) = 1$. To ensure this last condition, we insist that when we choose the primes to form a, we never choose a pair r_1, r_2 with $r_1 \mid r_2 + 1$. Let

$$R_2 = \sum_{\substack{r \in \mathcal{R} \\ (\log_2 x)^2 < r \leqslant x^{1/u}}} \frac{1}{r^2}, \quad R_3 = \sum_{\substack{r_1, r_2 \in \mathcal{R} \\ (\log_2 x)^2 < r_1, r_2 \leqslant x^{1/u} \\ r_1 | r_2 + 1}} \frac{1}{r_1 r_2}.$$

Thus,

$$\sum_{\substack{a \in \mathcal{A} \\ a \leqslant x}} \frac{1}{a} \geqslant \frac{1}{u!} R_1^u - \frac{1}{(u-2)!} (R_2 + R_3) R_1^{u-2}. \tag{5.3}$$

We have

$$R_2 \ll (\log_2 x)^{-2}$$

and from the Brun–Titchmarsh inequality, we have

$$R_3 \ll (\log_2 x)^{-1}.$$

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

Thus, (5.2) and (5.3) imply that

$$\sum_{\substack{a\in\mathcal{A}\\ a\leqslant x}} \frac{1}{a} \geqslant \frac{1}{u!} R_1^u\bigl(1-u^2(R_2+R_3)R_1^{-2}\bigr) = \frac{1}{u!} R_1^u\bigl(1+O\bigl((\log_2 x)^{-1}\bigr)\bigr).$$

Using the inequality $u! \ll u^{u+1/2}/e^u$, we thus have

$$\sum_{\substack{a\in\mathcal{A}\\ a\leqslant x}} \frac{1}{a} \gg \left(\frac{eR_1}{u}\right)^u \frac{1}{u^{1/2}} > \frac{e^u}{u^{1/2}} \gg \frac{(\log x)^{\rho(m)}}{(\log_2 x)^{5/2}},$$

which establishes the estimate (5.1).

Next, consider integers $map \leqslant x$ where $a \in \mathcal{A}$ and $p \nmid ma$ is prime. We suppose that $a \leqslant x^{1/3}$. We have $\gcd(s(ma), \sigma(ma)) = 1$, and so by the lower bound estimate in the sieve (see, e.g., [20, Theorem 2.5′]), the number of primes $p \in (ma, x/ma]$ for which $s(map) = s(ma)p + \sigma(ma)$ has no prime factors below $(\log x)^2 + 1$ is

$$\gg x/(ma \log x \log_2 x).$$

The numbers $n = map$ constructed this way are easily seen to be distinct and bounded by x, and, by (5.1), their number is

$$\gg \frac{x}{m \log x \log_2 x} \sum_{\substack{a\in\mathcal{A}\\ a\leqslant x^{1/3}}} \frac{1}{a} \gg \frac{x}{(\log x)^{1-\rho(m)}(\log_2 x)^{7/2}}, \tag{5.4}$$

since m is assumed fixed.

But note that for a number n so constructed we have that n is odd abundant and

$$\begin{aligned} s\bigl(s(n)\bigr) = s(n)\left(\frac{\sigma(s(n))}{s(n)} - 1\right) &< s(n)\left(\prod_{\substack{q\mid s(n)\\ q \text{ prime}}} \left(1+\frac{1}{q-1}\right) - 1\right) \\ &\leqslant s(n)\left(\left(1+\frac{1}{(\log x)^2}\right)^{\omega(s(n))} - 1\right), \end{aligned}$$

where we write $\omega(k)$ for the number of prime divisors of k. Now for $n \leqslant x$ we have $s(n) \ll x \log_2 x$ and $\omega(s(n)) \ll \log x/\log_2 x$. Thus,

$$s_2(n) \ll \frac{x}{\log x}.$$

Since the function s_2 is injective on the sociable numbers, it follows that the number of n counted above that are sociable is at most $O(x/\log x)$. Thus, using (5.4) we have that

$$\sum_{\substack{n\leqslant x,\, m\mid n\\ n \text{ odd abundant}\\ n \text{ not sociable}}} 1 \gg \frac{x}{(\log x)^{1-\rho(m)}(\log_2 x)^{7/2}}. \tag{5.5}$$

We now make a judicious choice for m. Let B be a large number, let m_0 be the smallest squarefree abundant number composed of primes all greater than B, and let $m = m_0^{p-1}$, where $p > m_0$ is prime. It is easy to see that as $B \to \infty$, this construction gives an odd abundant number m with $\rho(m) \to 1/2$. But for each such m we have (5.5). This completes the proof of the theorem. $\square$

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011

5.2. The distribution function for odd numbers

Let $\mathscr{B}(u)$ denote the set of odd numbers n with $\sigma(n)/n \geqslant u$ and let $B(u)$ denote the asymptotic density of $\mathscr{B}(u)$. It follows from the Erdős–Wintner theorem that $B(u)$ exists, is continuous, and is strictly decreasing for $u \geqslant 1$.

The number $B(2)$, as the density of the odd abundant numbers, is of interest to us since it stands as our upper bound for the upper density of the sociable numbers. We can also prove an upper bound for the upper density of the special sociable numbers which is considerably smaller than $B(2)$. In fact, we will show in Theorem 10 that its value is about $1/6000$, while $B(2) \approx 1/500$.

Theorem 9. *The set of special sociable numbers has upper density at most α, where*

$$\alpha = \int_1^\infty \frac{B(1+u)}{u^2}\,du.$$

Here α is the asymptotic proportion of odd abundant numbers $n \leqslant x$ with $s(n) > x$.

Proof. Let $\mathcal{D}_0(x)$ denote the set of odd abundant numbers $n \leqslant x$, and let $\mathcal{D}_1(x)$ denote the set of $n \in \mathcal{D}_0(x)$ with $s(n) \leqslant x$. Let $\mathcal{T}_i(x) = \mathcal{S} \cap \mathcal{D}_i(x)$ for $i = 0, 1$, where $\mathcal{S}$ is the set of sociable numbers. Let $\mathcal{T}_2(x)$ denote the set of integers $s(n)$ with $n \in \mathcal{T}_1(x)$. Since s is injective on $\mathcal{S}$ it follows that $\#\mathcal{T}_2(x) = \#\mathcal{T}_1(x)$. Also, let $\mathcal{T}_3(x)$ denote the set of odd abundant sociable numbers in $[1, x]$ that are not special. Clearly, each member of $\mathcal{T}_2(x)$ is not special, so we have $\mathcal{T}_2'(x) \subset \mathcal{T}_3(x)$, where $\mathcal{T}_2'(x)$ is the set of odd abundant members of $\mathcal{T}_2(x)$. As we have seen, the set of odd abundant numbers n for which $s(n)$ is not odd abundant has density 0, so that $\#\mathcal{T}_2(x) = \#\mathcal{T}_2'(x) + o(x)$. Thus, if $\mathcal{T}_4(x)$ denotes the set of special sociable numbers in $[1, x]$, then

$$\mathcal{T}_4(x) = \mathcal{T}_0(x) \setminus \mathcal{T}_3(x) \subset \mathcal{T}_0(x) \setminus \mathcal{T}_2'(x),$$

so that

$$\begin{aligned}\#\mathcal{T}_4(x) &\leqslant \#\mathcal{T}_0(x) - \#\mathcal{T}_2'(x) = \#\mathcal{T}_0(x) - \#\mathcal{T}_1(x) + o(x)\\ &= \#\big(\mathcal{T}_0(x) \setminus \mathcal{T}_1(x)\big) + o(x) \leqslant \#\big(\mathcal{D}_0(x) \setminus \mathcal{D}_1(x)\big) + o(x)\\ &= \#\mathcal{D}_0(x) - \#\mathcal{D}_1(x) + o(x).\end{aligned}$$

Now,

$$\frac{\#\mathcal{D}_0(x)}{x} = B(2) + o(1)$$

as $x \to \infty$, and by the continuity of $B(u)$, we have

$$\frac{\#\mathcal{D}_1(x)}{x} = B(2) - \int_1^\infty \frac{B(1+u)}{u^2}\,du + o(1)$$

as $x \to \infty$. We complete the proof by subtracting the second asymptotic relation from the first. □

We remark that if we have strict inequality in Theorem 9, i.e., if the upper density of the special sociable numbers is $< \alpha$, then there is a positive proportion of odd abundant numbers that are not sociable. This follows from the argument given for Theorem 6(a): Say a sociable number n is a *climber* if there is some integer $j \geqslant 0$ such that $s_{-j}(n)$ is a special sociable number, and each of $s_{-i}(n)$ for $i = 0, 1, \ldots, j$ is odd abundant. From the proof of Theorem 6(a) we have that the set of sociable numbers that are not climbers has density 0. But the number of odd abundant numbers $n \leqslant x$ with $s(n) > x$ is $(\alpha + o(1))x$. Clearly, no two of these numbers can be climbers corresponding to the same special sociable. Thus, if the upper density of the special sociables is $\beta < \alpha$, then at least $(\alpha - \beta + o(1))x$ odd abundant numbers up to x are not climbers, and so at least $(\alpha - \beta + o(1))x$ odd abundant numbers up to x are not sociable.

5.3. Numerical estimates

Using methods found in Deléglise [7], we compute bounds for $B(u)$, for various values of u. We prove the following.

Theorem 10. *The density $B(2)$ of the set of odd abundant numbers satisfies*

$$0.002042 < B(2) < 0.002071,$$

and the constant α in Theorem 9 satisfies

$$0.0001600 < \alpha < 0.0001772.$$

To prove this theorem, we introduce certain subsets of $\mathscr{B}(u)$, namely

$$\mathscr{B}_y(u) = \bigl\{n\colon\ \sigma(n)/n \geqslant u,\ \gcd\bigl(n, \Pi(y)\bigr) = 1\bigr\},$$

where $\Pi(y)$ denotes the product of the primes $p \leqslant y$ and $y \geqslant 2$. Denote the density of $\mathscr{B}_y(u)$ by $B_y(u)$. (This density is denoted $A_{\pi(y)}(u)$ in [7].) Let $P(n)$ be the largest prime factor of a number n if $n > 1$ and $P(1) = 1$, and define $\mathscr{N}_y$ to be the set of odd numbers n such that $P(n) \leqslant y$. As with Proposition 1.1 in [7], we have the following result.

Proposition 1. *For each $u \geqslant 1$, we have*

$$B(u) = \sum_{n \in \mathscr{N}_y} \frac{1}{n} B_y\left(\frac{u}{\sigma(n)/n}\right).$$

The summation here has infinitely many terms, so to get numerical estimates it is convenient to truncate and estimate the error. Let

$$\mathscr{N}_y(z) = \mathscr{N}_y \cap [1, z].$$

We have, imitating the argument in [7], the following result.

Proposition 2. *Let $y \geqslant 2$. Then the density $B(u)$ has bounds*

$$F_y \sum_{\substack{n \in \mathscr{N}_y(z) \\ \sigma(n)/n \geqslant u}} \frac{1}{n} \leqslant B(u) \leqslant \sum_{n \in \mathscr{N}_y(z)} \frac{1}{n} B_y\left(\frac{u}{\sigma(n)/n}\right) + \frac{1}{2} - F_y \sum_{n \in \mathscr{N}_y(z)} \frac{1}{n},$$

where $F_y = \prod_{p \leqslant y}(1 - 1/p)$, the product taken over primes $p \leqslant y$.

In order to estimate $B_y(u)$ from above, let $h_y(n)$ be the multiplicative function with $h_y(p^a) = 1$ for $p \leqslant y$ and $h_y(p^a) = \sigma(p^a)/p^a$ for $p > y$. The idea is to compute estimates for high moments of h_y. Towards this goal, let

$$H_{y,j}(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) h_y(d)^j,$$

so that $H_{y,j}(n) \geqslant 0$ for all n. Then,

$$\sum_{n \leqslant x} h_y(n)^j = \sum_{d \leqslant x} H_{y,j}(d) \left\lfloor \frac{x}{d} \right\rfloor \leqslant x \sum_{d=1}^{\infty} \frac{H_{y,j}(d)}{d} = x M_{y,j},$$

say. Since $h_y(n) \geqslant 1$ for all n, we have

$$\begin{aligned} B_y(u) &= F_y \cdot \lim_{x \to \infty} \frac{1}{x} \sum_{\substack{n \leqslant x \\ h_y(n) \geqslant u}} 1 \\ &\leqslant F_y \cdot \lim_{x \to \infty} \frac{1}{x} \sum_{\substack{n \leqslant x \\ h_y(n) \geqslant u}} \frac{h_y(n)^j - 1}{u^j - 1} \leqslant F_y \frac{M_{y,j} - 1}{u^j - 1}. \end{aligned}$$

In [7], the mean value $M_{y,j}$ is denoted by $\Lambda_{\pi(y)}(j)$; upper bounds are computed there for $y = 500$ and $j = 2^i$ for $i = 0, 1, \ldots, 12$. We have recalculated these values, using the program PARI/GP, and present them in the table below. Note that our values are slightly smaller than those in [7] when $j = 2^i$, $0 \leqslant i \leqslant 10$, but are larger for $j = 2^{11}$ and 2^{12}.

j	$M_{500,j} - 1 \leqslant$	j	$M_{500,j} - 1 \leqslant$
1	0.0002732982	128	0.03814735
2	0.0005469571	256	0.08377620
4	0.001095360	512	0.20651323
8	0.002196521	1024	0.7076500
16	0.004416412	2048	12.96156
32	0.008927653	4096	1.661395×10^{17}
64	0.01824571		

With these bounds, we calculated upper and lower bounds for $B(2)$ using $y = 500$ and $z = 10^{16}$, as well as bounds for $B(u)$ using $y = 500$ and $z = 10^{14}$, where $u = 2.0000, \ldots, 2.0500$ in increments of 0.0005, $u = 2.051, \ldots, 2.400$ in increments of 0.001, and $u = 2.405, \ldots, 2.700$ in increments of 0.005. These estimates for $B(u)$ were used to calculate our upper and lower bounds for α. We warmly thank Professor Deléglise for kindly providing us with his program for calculating upper and lower bounds for these values of $B(u)$.

References

[1] M.R. Avidon, On the distribution of primitive abundant numbers, Acta Arith. 77 (1996) 195–205.
[2] F. Behrend, Über numeri abundantes. I, Sitzungsberichte Akad. Berlin (1932) 322–328;
F. Behrend, Über numeri abundantes. II, Sitzungsberichte Akad. Berlin (1933) 280–293.
[3] P. Bratley, F. Lunnon, J. McKay, Amicable numbers and their distribution, Math. Comp. 24 (1970) 431–432.
[4] E. Catalan, Propositions et questions diverses, Bull. Soc. Math. France 16 (1888) 128–129.
[5] H. Cohen, On amicable and sociable numbers, Math. Comp. 24 (1970) 423–429.
[6] H. Davenport, Über numeri abundantes, Sitzungsberichte Akad. Berlin (1933) 830–837.
[7] M. Deléglise, Bounds for the density of abundant integers, Experiment. Math. 7 (1998) 137–143.
[8] L.E. Dickson, Theorems and tables on the sum of the divisors of a number, Q. J. Math. 44 (1913) 264–296.

[9] P.D.T.A. Elliott, Probabilistic Number Theory. I, Grundlehren Math. Wiss., vol. 239, Springer-Verlag, New York, 1979.
[10] P. Erdős, On the smoothness of the asymptotic distribution of additive arithmetical functions, Amer. J. Math. 61 (1939) 722–725.
[11] P. Erdős, Some remarks about additive and multiplicative functions, Bull. Amer. Math. Soc. 52 (1946) 527–537.
[12] P. Erdős, On amicable numbers, Publ. Math. Debrecen 4 (1955) 108–111.
[13] P. Erdős, On the distribution of numbers of the form $\sigma(n)/n$ and on some related questions, Pacific J. Math. 52 (1974) 59–65.
[14] P. Erdős, On asymptotic properties of aliquot sequences, Math. Comp. 30 (1976) 641–645.
[15] P. Erdős, G.J. Rieger, Ein Nachtrag über befreundete Zahlen, J. Reine Angew. Math. 273 (1975) 220.
[16] P. Erdős, A. Granville, C. Pomerance, C. Spiro, On the normal behavior of the iterates of some arithmetic functions, in: Analytic Number Theory, Allerton Park, IL, 1989, in: Progr. Math., vol. 85, Birkhäuser Boston, Boston, MA, 1990, pp. 165–204.
[17] M. García, J.M. Pedersen, H.J.J. te Riele, Amicable pairs, a survey, in: High Primes and Misdemeanours: Lectures in Honour of the 60th Birthday of Hugh Cowie Williams, in: Fields Inst. Commun., vol. 41, Amer. Math. Soc., Providence, RI, 2004, pp. 179–196.
[18] R.K. Guy, Unsolved Problems in Number Theory, Problem Books in Math., Springer-Verlag, New York, 2004.
[19] R.K. Guy, J.L. Selfridge, What drives an aliquot sequence? Math. Comp. 29 (1975) 101–107; Math. Comp. 34 (1980) 319–321, Erratum, collection of articles dedicated to Derrick Henry Lehmer on the occasion of his seventieth birthday.
[20] H. Halberstam, H.-E. Richert, Sieve Methods, London Math. Soc. Monogr., vol. 4, Academic Press, London–New York, 1974.
[21] B. Hornfeck, E. Wirsing, Über die Häufigkeit vollkommener Zahlen, Math. Ann. 133 (1957) 431–438.
[22] H.-J. Kanold, Über die Dichten der Mengen der vollkommenen und der befreundeten Zahlen, Math. Z. 61 (1954) 180–185.
[23] H.W. Lenstra Jr., Problem 6064, Amer. Math. Monthly 82 (1975) 1016, solution by the proposer in Amer. Math. Monthly 84 (1977) 580.
[24] O. Meissner, Über einige zahlentheoretische Funktionen, Arch. Math. Phys. 12 (1907) 199–202.
[25] D. Moews, A list of aliquot cycles of length greater than 2, available from the author's website: http://djm.cc/sociable.txt.
[26] J.M. Pedersen, Tables of aliquot cycles, electronic resource available at http://amicable.homepage.dk/.
[27] J. Perrott, Sur une proposition empirique énoncée au Bulletin, Bull. Soc. Math. France 17 (1889) 155–156.
[28] C. Pomerance, On the distribution of amicable numbers, J. Reine Angew. Math. 293/294 (1977) 217–222.
[29] C. Pomerance, On the distribution of amicable numbers. II, J. Reine Angew. Math. 325 (1981) 183–188.
[30] P. Poulet, Question 4865, L'interméd. des Math. 25 (1918) 100–101.
[31] G.J. Rieger, Bemerkung zu einem Ergebnis von Erdős über befreundete Zahlen, J. Reine Angew. Math. 261 (1973) 157–163.
[32] H.J.J. te Riele, Computation of all the amicable pairs below 10^{10}, Math. Comp. 47 (1986) 361–368, S9–S40, with a supplement.
[33] V. Toulmonde, Module de continuité de la fonction de répartition de $\varphi(n)/n$, Acta Arith. 121 (2006) 367–402.
[34] V. Toulmonde, Sur les variations de la fonction de répartition de $\varphi(n)/n$, J. Number Theory 120 (2006) 1–12.
[35] C.R. Wall, Density bounds for the sum of divisors function, in: The Theory of Arithmetic Functions, Proc. Conf., Western Michigan Univ., Kalamazoo, Mich., 1971, in: Lecture Notes in Math., vol. 251, Springer, Berlin, 1972, pp. 283–287.
[36] C.R. Wall, P.L. Crews, D.B. Johnson, Density bounds for the sum of divisors function, Math. Comp. 26 (1972) 773–777; Math. Comp. 31 (1977) 616, Erratum.
[37] E. Wirsing, Bemerkung zu der Arbeit über vollkommene Zahlen, Math. Ann. 137 (1959) 316–318.

Please cite this article in press as: M. Kobayashi et al., On the distribution of sociable numbers, J. Number Theory (2009), doi:10.1016/j.jnt.2008.10.011