QUASI-AMICABLE NUMBERS ARE RARE

PAUL POLLACK

Abstract. Define a quasi-amicable pair as a pair of distinct natural numbers each of which is the sum of the nontrivial divisors of the other, e.g., \{48, 75\}. Here nontrivial excludes both 1 and the number itself. Quasi-amicable pairs have been studied (primarily empirically) by Garcia, Beck and Najar, Lal and Forbes, and Hagis and Lord. We prove that the set of \( n \) belonging to a quasi-amicable pair has asymptotic density zero.

1. Introduction

Let \( s(n) := \sum_{d|n, d<n} d \) be the sum of the proper divisors of \( n \). Given a natural number \( n \), what can one say about the aliquot sequence at \( n \) defined as \( n, s(n), s(s(n)), \ldots \) ? From ancient times, there has been considerable interest in the case when this sequence is purely periodic. (In this case, \( n \) is called a sociable number; see Kobayashi et al. [11] for some recent results on such numbers.) An \( n \) for which the period is 1 is called perfect (see sequence A000396), and an \( n \) for which the period is 2 is called amicable (see sequence A063990). In the latter case, we call \( \{n, s(n)\} \) an amicable pair.

Let \( s^-(n) := \sum_{d|n, 1<d<n} d \) be the sum of the nontrivial divisors of the natural number \( n \), where nontrivial excludes both 1 and \( n \). According to Lal and Forbes [12], it was Chowla who suggested studying quasi-aliquot sequences of the form \( n, s^-(n), s^-(s^-(n)), \ldots \). Call \( n \) quasi-amicable if the quasi-aliquot sequence starting from \( n \) is purely periodic of period 2 (see sequence A005276). Thus, a quasi-amicable pair is a pair of distinct natural numbers \( n \) and \( m \) with \( s^-(n) = m \) and \( s^-(m) = n \) (e.g., \( n = 48 \) and \( m = 75 \)). The numerical data, reproduced in Table 1 from sequence A126160, suggests that the number of such pairs with a member \( \leq N \) tends to infinity with \( N \), albeit very slowly.

While quasi-amicable pairs have been studied empirically (see [8, 12, 1, 10, 2], and cf. [14, 13], [9, section B5]), it appears that very little theoretical work has been done. In this paper, we prove the following modest theorem, which is a quasi-amicable analogue of Erdős’s 1955 result [4] concerning amicable pairs:

**Theorem 1.1.** The set of quasi-amicable numbers has asymptotic density zero. In fact, as \( \epsilon \downarrow 0 \), the upper density of the set of \( n \) satisfying

\[
1 - \epsilon < \frac{s^-(s^-(n))}{n} < 1 + \epsilon
\]

(1.1)

tends to zero.

**Remark.** With \( s \) replacing \( s^- \), Theorem 1.1 follows from work of Erdős [4] and Erdős et al. [7, Theorem 5.1].

2010 Mathematics Subject Classification. Primary: 11A25, Secondary: 11N37.

Key words and phrases. aliquot sequence, quasi-aliquot sequence, quasi-amicable pair, augmented amicable pair.
<table>
<thead>
<tr>
<th>$N$</th>
<th># of quasi-amicable pairs with least member $\leq N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^5$</td>
<td>9</td>
</tr>
<tr>
<td>$10^6$</td>
<td>17</td>
</tr>
<tr>
<td>$10^7$</td>
<td>46</td>
</tr>
<tr>
<td>$10^8$</td>
<td>79</td>
</tr>
<tr>
<td>$10^9$</td>
<td>180</td>
</tr>
<tr>
<td>$10^{10}$</td>
<td>404</td>
</tr>
<tr>
<td>$10^{11}$</td>
<td>882</td>
</tr>
<tr>
<td>$10^{12}$</td>
<td>1946</td>
</tr>
</tbody>
</table>

Table 1

**Notation.** Throughout, $p$ and $q$ always denote prime numbers. We use $\sigma(n) := \sum_{d|n} d$ for the sum of all positive divisors of $n$, and we let $\omega(n) := \sum_{p|n} 1$ stand for the number of distinct prime factors of $n$. We write $P(n)$ for the largest prime divisor of $n$, with the understanding that $P(1) = 1$. We say that $n$ is $y$-smooth if $P(n) \leq y$. For each $n$, its $y$-smooth part is defined as the largest $y$-smooth divisor of $n$.

The Landau–Bachmann $o$ and $O$-symbols, as well as Vinogradov’s $\ll$ notation, are employed with their usual meanings. *Implied constants are absolute unless otherwise specified.*

## 2. Proof of Theorem 1.1

In this section we give the proof of Theorem 1.1, assuming two preliminary results whose proofs are deferred to §3 and §4.

**Proposition 2.1.** For each $\epsilon > 0$, the set of natural numbers $n$ with

$$\frac{\sigma(n+1)}{n+1} - \epsilon < \frac{\sigma(s^-(n))}{s^-(n)} < \frac{\sigma(n+1)}{n+1} + \epsilon. \quad (2.1)$$

has asymptotic density 1.

**Remark.** If $n$ is prime, then $s^-(n) = 0$, and the expression $\sigma(s^-(n))/s^-(n)$ is undefined. This does not contradict Proposition 2.1, since the set of primes has asymptotic density zero.

**Proposition 2.2.** As $\epsilon \downarrow 0$, the upper density of the set of natural numbers $n$ for which

$$1 - \epsilon < \left(\frac{\sigma(n)}{n} - 1\right) \left(\frac{\sigma(n+1)}{n+1} - 1\right) < 1 + \epsilon \quad (2.2)$$

tends to zero.

**Proof of Theorem 1.1.** It suffices to prove the upper density assertion of the theorem. Let $\delta > 0$. We will show that if $\epsilon > 0$ is sufficiently small, then the upper density of the set of $n$ for which (1.1) holds is at most $2\delta$. We start by assuming that both $\sigma(n)/n \leq B$ and $\sigma(n+1)/(n+1) \leq B$, where $B > 0$ is chosen so that these conditions exclude a set of $n$ of upper density at most $\delta$. To see that such a choice is possible, we can use a first moment argument; indeed, since

$$\sum_{n \leq x} \frac{\sigma(n)}{n} = \sum_{n \leq x} \sum_{d|n} \frac{1}{d} \leq x \sum_{d \leq x} \frac{1}{d^2} < 2x,$$
we can take $B = 4/\delta$. Moreover, Proposition 2.1 shows that by excluding an additional set of density 0, we can assume that

$$\left| \frac{\sigma(s^{-}(n))}{s^{-}(n)} - \frac{\sigma(n+1)}{n+1} \right| < \frac{\epsilon}{2B}.$$  

Now write

$$s^{-}(s^{-}(n)) = \frac{s^{-}(n) s^{-}(s^{-}(n))}{n} = \left( \frac{\sigma(n)}{n} - 1 - \frac{1}{n} \right) \left( \frac{\sigma(s^{-}(n))}{s^{-}(n)} - 1 - \frac{1}{s^{-}(n)} \right).$$

If $n$ is a large natural number satisfying (1.1) and our above conditions, then a short computation shows $s^{-}(s^{-}(n))$ is within $\epsilon$ of the product $(\sigma(n)/(n-1))(\sigma(n+1)/(n+1)-1)$. (Keep in mind that since $n$ is composite, we have $s^{-}(n) \geq \sqrt{n}$.) Thus,

$$1 - 2\epsilon < \left( \frac{\sigma(n)}{n} - 1 \right) \left( \frac{\sigma(n+1)}{n+1} - 1 \right) < 1 + 2\epsilon.$$

Finally, Proposition (2.2) shows that if $\epsilon$ is chosen sufficiently small, then these remaining $n$ make up a set of upper density $< \delta$. \hfill $\square$

3. The proof of Proposition 2.1

3.1. Preparation. The proof of the proposition is very similar to the proof, due to Erdős, Granville, Pomerance, and Spiro, that $s(s(n))/s(n) = s(n)/n + o(1)$, as $n \to \infty$ along a sequence of density 1 (see Erdős et al. [7, p. 195]). We follow their argument, as well as the author’s adaptation [15], very closely.

We begin by recalling some auxiliary estimates. The first of these is due to Pomerance [16, Theorem 2].

Lemma 3.1. Let $D$ be a natural number, and let $x \geq 2$. The number of $n \leq x$ for which $D \nmid \sigma(n)$ is $\ll x/(\log x)^{1/\varphi(D)}$.

For a given $\alpha$, we call the natural number $n$ an $\alpha$-primitive number if $\sigma(n)/n \geq 1 + \alpha$ while $\sigma(d)/d < 1 + \alpha$ for every proper divisor $d$ of $n$. The following estimate is due to Erdős [5, p. 6]:

Lemma 3.2. Fix a positive rational number $\alpha$. There is a constant $c = c(\alpha) > 0$ and an $x_0 = x_0(\alpha)$ so that for $x > x_0$, the number of $\alpha$-primitive $n \leq x$ is at most

$$\frac{x}{\exp(c\sqrt{\log x} \log \log x)}.$$  

As a consequence of Lemma 3.2, we obtain the following convergence result, which we will need to conclude the proof of Proposition 2.1.

Lemma 3.3. Fix a positive rational number $\alpha$. Then

$$\sum_{a \text{ \alpha-primitive}} \frac{2\omega(a)}{a} < \infty.$$
Proof. We split the values of $a$ appearing in the sum into two classes, putting those $a$ for which $\omega(a) \leq 20 \log \log a$ in the first class and all other $a$ in the second. If $a$ belongs to the first class, then $2^{\omega(a)} \leq (\log a)^{20 \log 2}$, and Lemma 3.2 shows that the sum over these $a$ converges (by partial summation). To handle the $a$ in the second class, we ignore the $\alpha$-primitivity condition altogether and invoke a lemma of Pollack [15, Lemma 2.4], according to which $\sum_{a: \omega(a) > 20 \log \log a} 2^{\omega(a)} a < \infty$. \hfill \Box

3.2. Proof proper. We proceed to prove Proposition 2.1 in two stages; first we prove that the lower-bound holds almost always, and then we do the same for the upper bound. The following lemma is needed for both parts.

Lemma 3.4. Fix a natural number $T$. For each composite value of $n$ with $1 \leq n \leq x$, write $n + 1 = m_1 m_2$ and $s^-(n) = M_1 M_2$, where $P(m_1 M_1) \leq T$ and every prime dividing $m_2 M_2$ exceeds $T$. Then, except for $o(x)$ (as $x \to \infty$) choices of $n$, we have $m_1 = M_1$.

Proof. At the cost of excluding $o(x)$ values of $n \leq x$, we may assume that

$$m_1 \leq (\log \log x)^{1/2} \left( \prod_{p \leq T} p \right)^{-1} =: R.$$ 

Indeed, in the opposite case, $n + 1$ has a $T$-smooth divisor exceeding $R$, and the number of such $n \leq x$ is

$$\ll x \sum_{e \ T\text{-smooth}} \frac{1}{e} = o(x),$$

as $x \to \infty$. Here we use that the sum of the reciprocals of the $T$-smooth numbers is $\prod_{p \leq T} (1 - 1/p)^{-1} < \infty$. Hence, $m_1 \prod_{p \leq T} p \leq (\log \log x)^{1/2}$, and so Lemma 3.1 shows that excluding $o(x)$ values of $n \leq x$, we can assume that $m_1 \prod_{p \leq T} p$ divides $\sigma(n)$. Since

$$s^-(n) = \sigma(n) - (n + 1),$$

it follows that $m_1$ is the $T$-smooth part of $s^-(n)$. That is, $m_1 = M_1$. \hfill \Box

Proof of the lower bound half of Proposition 2.1. Fix $\delta > 0$. We will show that the number of $n \leq x$ for which the left-hand inequality in (2.1) fails is smaller than $3\delta x$, once $x$ is large.

Fix $B$ large enough that $\sigma(n + 1)/(n + 1) \leq B$ except for at most $\delta x$ exceptional $n \leq x$. That this is possible follows from the first moment argument used in the proof of Theorem 1.1 (e.g., we may take $B = 4/\delta$ again). Next, fix $T$ large enough so that with $m_2$ defined as in Lemma 3.4, we have

$$\frac{\sigma(m_2)}{m_2} \leq \exp(\epsilon/B).$$
QUASI-AMICABLE NUMBERS ARE RARE

except for at most $\delta x$ exceptional $n \leq x$. To see that a suitable choice of $T$ exists, observe that

$$\sum_{n \leq x} \log \frac{\sigma(m_2)}{m_2} \leq \sum_{n \leq x} \sum_{p|n+1 \atop p>T} \log \left(1 + \frac{1}{p} + \frac{1}{p^2} + \ldots \right)$$

$$\leq \sum_{n \leq x} \sum_{p|n+1 \atop p>T} \frac{1}{p-1} \leq 2x \sum_{p>T} \frac{1}{p(p-1)} < \frac{2x}{T}.$$

Hence, we may take $T = [2B/(\delta \epsilon)].$

For large $x$, we have that $n$ is composite (so that $M_1$ is defined) and that $m_1 = M_1$, except for at most $\delta x$ values of $n \leq x$. This follows from Lemma 3.4 and the fact that the primes have density 0.

If $n$ is not in any of the exceptional classes defined above, then

$$\frac{\sigma(s^-(n))}{s^-(n)} = \frac{\sigma(M_1 M_2)}{M_1 M_2} \geq \frac{\sigma(M_1)}{M_1} = \frac{\sigma(m_1)}{m_1} = \frac{\sigma(n+1)/(n+1)}{\sigma(m_2)/m_2}$$

$$\geq \frac{\sigma(n+1)}{n+1} \exp \left(- \frac{\epsilon}{B} \right) > \frac{\sigma(n+1)}{n+1} \left(1 - \frac{\epsilon}{B} \right) \geq \frac{\sigma(n+1)}{n+1} - \epsilon,$$

which is the desired lower bound. Note that at most $3\delta x$ values of $n \leq x$ are exceptional, as claimed. \hfill \Box

Proof of the upper bound half of Proposition 2.1. We may suppose that $0 < \epsilon < 1$. Let $\delta > 0$ be given. Fix $\eta \in (0,1)$ so small that the number of $n \leq x$ which are either prime or which fail to satisfy

$$P(n) > x^\eta \quad \text{and} \quad P(n)^2 \nmid n \quad \text{is smaller than } \delta x,$$

once $x$ is large. The existence of such an $\eta$ follows either from Brun’s sieve or well-known work of Dickman on smooth numbers. Next, using the first moment argument from the proof of Theorem 1.1, choose a fixed number $B \geq 1$ so that all but at most $\delta x$ of the numbers $n \leq x$ satisfy

$$\frac{\sigma(n+1)}{n+1} \leq B. \tag{3.2}$$

We fix rational numbers $\alpha_1$ and $\alpha_2$ satisfying

$$0 < \alpha_1 \leq \frac{\epsilon}{4B}, \quad 0 < \alpha_2 \leq \frac{\alpha_1 \eta}{12}.$$

Finally, we fix a natural number $T$ which is sufficiently large, depending only on the $\alpha_i$, $\delta$, $\eta$, and $B$. The precise meaning of “sufficiently large” will be specified in the course of the proof.

Suppose that the right-hand inequality (2.1) fails for $n$, where we assume that $n$ is composite and satisfies both (3.1) and (3.2). Write

$$n + 1 = m_1 m_2 \quad \text{and} \quad s^-(n) = M_1 M_2,$$

where $P(m_1 M_1) \leq T$ and every prime dividing $m_2 M_2$ exceeds $T$. By Lemma 3.4, we can assume $m_1 = M_1$, excluding at most $\delta x$ values of $n \leq x$. Thus,

$$\frac{\sigma(M_2)/M_2}{\sigma(m_2)/m_2} = \frac{\sigma(s^-(n))/s^-(n)}{\sigma(n+1)/(n+1)} \geq 1 + \frac{\epsilon}{\sigma(n+1)/(n+1)} \geq 1 + \frac{\epsilon}{B} \geq 1 + 4\alpha_1.$$
In particular,
\[
\frac{\sigma(M_2)}{M_2} \geq 1 + 4\alpha_1. \tag{3.3}
\]

We can assume our choice of \(T\) was such that, apart from at most \(\delta x\) exceptional \(n \leq x\), we have
\[
\frac{\sigma(m_2)}{m_2} \leq 1 + \alpha_1. \tag{3.4}
\]

Indeed, the argument for the analogous claim in the proof of the lower-bound shows it is sufficient that \(T > 2(\delta \log(1 + \alpha_1))^{-1}\). Henceforth, we assume (3.4). Now write \(M_2 = M_3M_4\), where every prime dividing \(M_3\) divides \(n + 1\), while \(M_4\) is coprime to \(n + 1\). Note that every prime dividing \(M_3\) divides \(m_2\). Hence,
\[
\frac{\sigma(M_3)}{M_3} \leq \prod_{p|M_3} \left(1 + \frac{1}{p - 1}\right) = \left(\prod_{p|M_3} \frac{p^2}{p^2 - 1}\right) \prod_{q|M_3} \frac{q + 1}{q}
\]
\[
\leq \left(\prod_{p>T} \frac{p^2}{p^2 - 1}\right) \frac{\sigma(m_2)}{m_2} \leq 1 + 2\alpha_1,
\]
using (3.4) and assuming an initial appropriate choice of \(T\). So from (3.3),
\[
\frac{\sigma(M_4)}{M_4} = \frac{\sigma(M_2)/M_2}{\sigma(M_3)/M_3} \geq \frac{1 + 4\alpha_1}{1 + 2\alpha_1} \geq 1 + \alpha_1.
\]

It follows that there is an \(\alpha_1\)-primitive number \(a_1\) dividing \(M_4\), where each prime dividing \(a_1\) exceeds \(T\).

We claim next that there is a squarefree, \(\alpha_2\)-primitive number \(a_2\) dividing \(a_1\) with \(a_2 \leq a_1^{\eta/2}\).

List the distinct prime factors of \(a_1\) in increasing order, say \(T < q_1 < q_2 < \cdots < q_t\), and put \(a_0 := q_1q_2\cdots q_{[\eta t/2]}\), so that
\[
a_0 \leq (q_1\cdots q_{[\eta t/2]})^{[\eta t/2]/t} \leq a_1^{\eta/2}.
\]

We will show that \(\sigma(a_0)/a_0 \geq 1 + \alpha_2\); then we can take \(a_2\) as any \(\alpha_2\)-primitive divisor of \(a_0\).

First, observe that \(\eta t/2 \geq \eta t/3\). Otherwise, \(t < 6/\eta\) and
\[
1 + \alpha_1 \leq \frac{\sigma(a_1)}{a_1} \leq \prod_{1 \leq i \leq t} \left(1 + \frac{1}{q_i - 1}\right) \leq \left(1 + \frac{1}{T}\right)^{6/\eta} \leq \exp\left(\frac{6}{\eta T}\right),
\]
which is false, assuming a suitable initial choice of \(T\). It follows that
\[
\frac{\sigma(a_0)}{a_0} = \prod_{1 \leq i \leq [\eta t/2]} \frac{q_i + 1}{q_i} \geq \left(\prod_{p>T} \frac{p^2 - 1}{p^2}\right) \prod_{1 \leq i \leq [\eta t/2]} \frac{q_i}{q_i - 1},
\]
while
\[
\prod_{1 \leq i \leq [\eta t/2]} \frac{q_i}{q_i - 1} \geq \left(\prod_{1 \leq i \leq t} \frac{q_i}{q_i - 1}\right)^{[\eta t/2]/t}
\]
\[
\geq \left(\frac{\sigma(a_1)/a_1}{a_1}\right)^{\eta/3} \geq (1 + \alpha_1)^{\eta/3} \geq 1 + \frac{\alpha_1 \eta}{6}.
\]
Thus,
\[
\frac{\sigma(a_0)}{a_0} \geq \left( \prod_{p > T} \frac{p^2 - 1}{p^2} \right) \left( 1 + \frac{\alpha_1 \eta}{6} \right) \geq 1 + \frac{\alpha_1 \eta}{12} \geq 1 + \alpha_2,
\]
again assuming a suitable choice of \( T \) to justify the middle inequality.

Observe that \( a_2 \) satisfies
\[
a_2 \leq a_1^{\eta/2} \leq (s^-(n))^{\eta/2} < x^{2\eta/3},
\]
for large \( x \). Write \( n = Pr \), where \( P = P(n) \). Then \( r > 1 \) (since \( n \) is composite) and also, by (3.1),
\[
r \leq x/P \leq x^{1-\eta}.
\]
Moreover, \( a_2 \) divides
\[
s^-(Pr) = P(\sigma(r) - r) + \sigma(r) - 1,
\]
and so
\[
P(\sigma(r) - r) \equiv 1 - \sigma(r) \pmod{a_2}.
\]
We view this as a linear congruence condition on \( P \) modulo \( a_2 \). If there are any solutions, then \( D := \gcd(\sigma(r) - r, a_2) \mid 1 - \sigma(r) \), and in this case there are exactly \( D \) solutions modulo \( a_2 \). Note that if there are any solutions, then \( D \mid r - 1 \). Also note that \( D \) is squarefree, since \( a_2 \) is squarefree.

We now sum over pairs \( a_2 \) and \( r \), for each pair counting the number of possible values of \( P \leq x/r \). By the Brun–Titchmarsh inequality and the preceding remarks about \( D \), we have that the number of possible values of \( n = Pr \) is
\[
\ll \sum_{a_2 \text{ \( \alpha_2 \)-primitive}} \sum_{1 < r \leq x^{1-\eta}} D \frac{x/r}{\varphi(a_2) \log \left( x/(a_2r) \right)} \sum_{D \mid (a_2, r-1)} \frac{1}{\varphi(a_2)} \sum_{D \mid a_2} D \sum_{1 < r \leq x^{1-\eta}} \frac{1}{r}.
\]
The sum on \( r \) is \( \ll \frac{1}{D} \log x \). Moreover, since \( a_2 \) is \( \alpha_2 \)-primitive, we have
\[
\frac{a_2}{\varphi(a_2)} \ll \frac{\sigma(a_2)}{a_2} \leq \frac{3}{2} (1 + \alpha_2) \ll 1,
\]
and so \( \varphi(a_2) \gg a_2 \). Thus, the remaining sum is
\[
\ll \frac{x}{\eta} \sum_{a_2 \text{ \( \alpha_2 \)-primitive}} \frac{1}{a_2} \sum_{D \mid a_2} \frac{1}{D} \sum_{a_2 \geq T} \frac{\varphi(a_2)}{a_2}.
\]
But if \( T \) was chosen sufficiently large, then this last sum is bounded by \( \eta \delta x \) (by Lemma 3.3), leading to an upper bound of \( \ll \delta x \). Since the number of exceptional \( n \) appearing earlier in the argument is also \( \ll \delta x \), and \( \delta > 0 \) was arbitrary, the proof is complete. \( \Box \)
4. Proof of Proposition 2.2

We start by quoting two lemmas. The first was developed by Erdős [3] to estimate the decay of the distribution function of \(\sigma(n)/n\) near infinity. We state the lemma in a slightly stronger form which is supported by his proof.

**Lemma 4.1.** For \(x > 0\), the number of positive integers \(n \leq x\) with \(\sigma(n)/n > y\) is
\[
\leq x/\exp(\exp((e^{-\gamma} + o(1))y)), \quad \text{as } y \to \infty,
\]
uniformly in \(x\), where \(\gamma\) is the Euler–Mascheroni constant.

The next lemma, also due to Erdős [6], supplies an estimate for how often \(\sigma(n)/n\) lands in a short interval; note the uniformity in the parameter \(a\).

**Lemma 4.2.** Let \(x > t \geq 2\) and let \(a \in \mathbb{R}\). The number of \(n \leq x\) with \(a < \sigma(n)/n < a + 1/t\) is \(\ll x/\log t\).

The next two lemmas develop the philosophy that the rough size of \(\sigma(n)/n\) is usually determined by the small prime factors of \(n\). Put \(h(n) := \sum_{d|n} \frac{1}{d}\), so that \(h(n) = \sigma(n)/n\). For each natural number \(T\), set \(h_T(n) := \sum_{d|n,P(d) \leq T} \frac{1}{d}\). The next lemma says that \(h\) and \(h_T\) are usually close once \(T\) is large.

**Lemma 4.3.** Let \(\epsilon > 0\) and \(x \geq 1\). The number of \(n \leq x\) with \(h(n) - h_T(n) > \epsilon\) is \(\ll x/(T\epsilon)\).

**Proof.** Again, we use a first moment argument. We have
\[
\sum_{n \leq x} (h(n) - h_T(n)) \leq \sum_{n \leq x} \sum_{d|n, d>T} \frac{1}{d} \leq x \sum_{d>T} \frac{1}{d^2} \ll x/T,
\]
from which the result is immediate. \(\square\)

**Lemma 4.4.** Let \(T\) be a natural number. Let \(S\) be any set of real numbers, and define \(\mathcal{E}(S)\) as the set of \(T\)-smooth numbers \(e\) for which \(h_T(e) - 1 \in S\). Then for \(n \in \mathbb{N}\), we have \(h_T(n) - 1 \in S\) precisely when \(n\) has \(T\)-smooth part \(e\) for some \(e \in \mathcal{E}(S)\). Moreover, the density of such \(n\) exists and is given by
\[
\sum_{e \in \mathcal{E}(S)} \frac{1}{e} \prod_{p \leq T} (1 - 1/p). \tag{4.1}
\]

**Proof.** It is clear that \(h_T(n)\) depends only on the \(T\)-smooth part of \(n\). So it suffices to prove that the density of \(n\) with \(T\)-smooth part in \(\mathcal{E}(S)\) is given by (4.1).

For each set of \(T\)-smooth numbers \(\mathcal{E}\), let \(\overline{d}_\mathcal{E}\) and \(\underline{d}_\mathcal{E}\) denote the upper and lower densities of the set of \(n\) whose \(T\)-smooth part belongs to \(\mathcal{E}\). If \(\overline{d}_\mathcal{E} = \underline{d}_\mathcal{E}\), then the density of this set exists; denote it by \(d_\mathcal{E}\).

For each \(T\)-smooth number \(e\), a natural number \(n\) has \(T\)-smooth part \(e\) precisely when \(e\) divides \(n\) and \(n/e\) is coprime to \(\prod_{p \leq T} p\), so that the set of such \(n\) has density \(\frac{1}{e} \prod_{p \leq T} (1 - 1/p)\). Since density is finitely additive, it follows that for any finite subset \(\mathcal{E} \subset \mathcal{E}(S)\),
\[
d_\mathcal{E} = \sum_{e \in \mathcal{E}} \frac{1}{e} \prod_{p \leq T} (1 - 1/p).
\]
Now let \(x > 0\), and put \(\mathcal{E}(S) = \mathcal{E}_1 \cup \mathcal{E}_2\), where \(\mathcal{E}_1 = \mathcal{E}(S) \cap [1, x]\) and \(\mathcal{E}_2 = \mathcal{E}(S) \setminus \mathcal{E}_1\). Then \(d_{\mathcal{E}(S)} \geq d_{\mathcal{E}_1}\) for all \(x\), and so letting \(x \to \infty\), we find that \(d_{\mathcal{E}(S)}\) is bounded below by
(4.1). On the other hand, \( \overline{d}_{\mathcal{E}(S)} \leq \overline{d}_{\mathcal{E}_1} + \overline{d}_{\mathcal{E}_2} \). But \( \overline{d}_{\mathcal{E}_1} \) is bounded above by (4.1) for all \( x \), while \( \overline{d}_{\mathcal{E}_2} \leq \sum e \text{ smooth } e^{-1} = o(1) \), as \( x \to \infty \). Thus, letting \( x \to \infty \), we obtain that \( \overline{d}_{\mathcal{E}(S)} \) is bounded above by (4.1).

\( \square \)

Proof of Proposition 2.2. Let \( \delta > 0 \) be sufficiently small. We will show that for

\[ \epsilon < \exp(-4/\delta), \quad (4.2) \]

the number of \( n \leq x \) satisfying (2.2) is \( \ll \delta (\log \log \frac{1}{\delta}) x \), for large \( x \). Note that since \( \delta \log \log \frac{1}{\delta} \to 0 \) as \( \delta \downarrow 0 \), this proves the proposition. In what follows, we fix \( \delta \) and \( \epsilon \), always assuming that \( \delta \) is small and that \( \epsilon > 0 \) satisfies (4.2).

Put \( T := \epsilon^{-1} \delta^{-1} \). We can assume that both \( n \) and \( n + 1 \) have \( T \)-smooth part \( \leq \log x \). Indeed, for large \( x \), this excludes a set of \( n \) size \( \prec \delta x \), since

\[ \sum_{e \text{ smooth} \atop e > \log x} \frac{1}{e} = o(1), \]

as \( x \to \infty \).

Let \( I \) be the closed interval defined by \( I := [\exp(-1/\delta), 2 \log \log \frac{1}{\delta}] \). For large \( x \), Lemmas 4.1 and 4.2 imply that all but \( \ll \delta x \) values of \( n \leq x \) are such that \( h(n) - 1 \in I \) and \( h(n + 1) - 1 \in I \). By Lemma 4.3, excluding \( \ll \delta x \) additional values of \( n \leq x \), we can assume that \( h_T(n) \geq h(n) - \epsilon \) and \( h_T(n + 1) \geq h(n + 1) - \epsilon \). Recalling the upper bound (4.2) on \( \epsilon \), we see that both \( h_T(n) - 1 \) and \( h_T(n + 1) - 1 \) belong to the interval \( J \), where

\[ J := \left[ \frac{1}{2} \exp(-1/\delta), 2 \log \log \frac{1}{\delta} \right]. \]

Moreover (always assuming \( \delta \) sufficiently small),

\[(h_T(n) - 1)(h_T(n + 1) - 1) \geq ((h(n) - 1) - \epsilon)((h(n + 1) - 1) - \epsilon) \geq 1 - 5\epsilon \log \log \frac{1}{\delta}, \quad (4.3)\]

and

\[(h_T(n) - 1)(h_T(n + 1) - 1) \leq (h(n) - 1)(h(n + 1) - 1) \leq 1 + \epsilon. \quad (4.4)\]

Write \( J \) as the disjoint union of \( N := [1/\epsilon] \) consecutive intervals \( J_0, J_1, \ldots, J_{N-1} \), each of length \( 1/N \). We estimate, for each \( 0 \leq i < N \), the number of \( n \) for which \( h_T(n) - 1 \) belongs to \( J_i \). Fix \( 0 \leq i < N \). Since \( h_T(n) - 1 \) belongs to \( J_i \), (4.3) and (4.4) show that

\[ h_T(n + 1) - 1 \in \left[ \frac{1 - 5\epsilon \log \log \frac{1}{\delta}}{x_{i+1}}, \frac{1 + \epsilon}{x_i} \right] := J'_i, \quad (4.5)\]

where \( x_i \) and \( x_{i+1} \) are the left and right endpoints of \( J_i \), respectively. So in the notation of Lemma 4.4, \( n \) has \( T \)-smooth part \( e \in \mathcal{E}(J_i) \) and \( n + 1 \) has \( T \)-smooth part \( e' \in \mathcal{E}(J'_i) \). Clearly, \( \gcd(e, e') = 1 \). That \( n \) and \( n + 1 \) have \( T \)-smooth parts \( e \) and \( e' \), respectively, amounts to a congruence condition on \( n \) modulo \( M := ee' \prod_{p \leq T} p \), where the number of allowable residue classes is \( \prod_{p \mid ee'} (p - 1) \prod_{p \nmid ee', p \leq T} (p - 2) \). For large \( x \),

\[ M \leq (\log x)^2 \prod_{p \leq T} p < (\log x)^3 \leq x. \]
(Recall that $e, e' \leq \log x$.) Thus, the Chinese remainder theorem shows that the number of such $n \leq x$ is

$$\ll \frac{x}{ee'} \prod_{p \mid ee'} (1 - 1/p) \prod_{p \mid ee'} (1 - 2/p)$$

$$\leq x \left( \frac{1}{e} \prod_{p \leq T} (1 - 1/p) \right) \left( \frac{1}{e'} \prod_{p \leq T} (1 - 1/p) \right) \prod_{p \mid ee'} (1 - 1/p)^{-1}.$$ 

But

$$\prod_{p \mid ee'} (1 - 1/p)^{-1} = \frac{e}{\varphi(e)} \frac{e'}{\varphi(e')} \ll \frac{\sigma(e) \sigma(e')}{ee'} \ll \left( \log \log \frac{1}{\delta} \right)^2,$$

since $h(e) - 1, h(e') - 1 \leq 2 \log \log \frac{1}{\delta}$. Summing over $e \in \mathcal{E}(J_i)$ and $e' \in \mathcal{E}(J'_i)$, we find that the number of $n$ under consideration is

$$\ll x \left( \log \log \frac{1}{\delta} \right)^2 \left( \sum_{e \in \mathcal{E}(J_i)} \frac{1}{e} \prod_{p \leq T} (1 - 1/p) \right) \left( \sum_{e' \in \mathcal{E}(J'_i)} \frac{1}{e'} \prod_{p \leq T} (1 - 1/p) \right).$$

Now sum over $0 \leq i < N$. We obtain that the number of remaining $n$ satisfying (2.2) is

$$\ll Lx \left( \log \log \frac{1}{\delta} \right)^2,$$

where

$$L := \sup_{0 \leq i < N} \left\{ \sum_{e' \in \mathcal{E}(J'_i)} \frac{1}{e'} \prod_{p \leq T} (1 - 1/p) \right\} \left( \sum_{e \in \mathcal{E}(J_i)} \left\{ \sum_{e' \in \mathcal{E}(J'_i)} \frac{1}{e'} \prod_{p \leq T} (1 - 1/p) \right\} \right).$$

we use here that the $J_i$ are disjoint, so that

$$\sum_{0 \leq i < N} \sum_{e \in \mathcal{E}(J_i)} \frac{1}{e} \leq \sum_{e \text{ T-smooth}} \frac{1}{e} = \prod_{p \leq T} (1 - 1/p)^{-1}.$$ 

The proof will be completed by showing that $L \ll \delta$. It is enough to argue that each sum

$$\sum_{e' \in \mathcal{E}(J'_i)} \frac{1}{e'} \prod_{p \leq T} (1 - 1/p)$$

is $\ll \delta$, uniformly for $0 \leq i < N$. By Lemma 4.4, this sum describes the density of those natural numbers $m$ for which $h_T(m) - 1 \in J'_i$. We split these $m$ into two classes, according to whether $h(m) - h_T(m) > \epsilon$ or not. The set of $m$ in the former class has upper density $\ll \delta$, by Lemma 4.3. Suppose now that $h(m)$ belongs to the second class. From the expression (4.5) defining $J'_i$ and a short computation, we see that $h_T(m)$ is trapped within a specific interval of length

$$\ll \exp(2/\delta) \left( \log \log \frac{1}{\delta} \right)^2 \epsilon \ll \exp(3/\delta) \epsilon.$$
Since \( m \) belongs to the second class, \( h(m) \) is also trapped within a specific interval of length \( \ll \exp(3/\delta)\epsilon \). By (4.2), \( \exp(3/\delta)\epsilon \leq \exp(-1/\delta) \), and so by Lemma 4.2, the upper density of the set of those \( m \) in the second class is
\[
\ll \frac{1}{\delta^{-1} + O(1)} \ll \delta,
\]
assuming again that \( \delta \) is sufficiently small. \( \square \)

**Remark.** Our argument also shows that the set of *augmented amicable numbers* has density zero (see sequences A007992, A015630). Here an augmented amicable number is an integer which generates a 2-cycle under iteration of the function \( s^+(n) := 1 + \sum_{d|n, d < n} d \), e.g., \( n = 6160 \).

**Sequences discussed.** A000396, A005276, A007992, A015630, A063990, A126160

### 5. Acknowledgements

The author thanks the National Science Foundation for their generous support (under award DMS-0802970).

**References**


*University of Illinois at Urbana-Champaign, Department of Mathematics, 1409 West Green Street, Urbana, Illinois 61801, USA
E-mail address: pppollac@illinois.edu*