# The range of the sum-of-proper-divisors function

Florian Luca
Mathematical Institute, UNAM Juriquilla
76230 Santiago de Querétaro, México and
School of Mathematics, University of the Witwatersrand
P. O. Box Wits 2050, South Africa
fluca@matmor.unam.mx

Carl Pomerance
Department of Mathematics
Dartmouth College
Hanover, NH 03755-3551, USA
carl.pomerance@dartmouth.edu

**Abstract**

Answering a question of Erdős, we show that a positive proportion of even numbers are in the form $s(n)$, where $s(n) = \sigma(n) - n$, the sum of proper divisors of $n$.

*2000 Mathematics Subject Classification:* Primary 11A25, Secondary 11N37

## 1 Introduction

For a positive integer $n$, let $s(n) = \sigma(n) - n$, the sum of the proper divisors of $n$. The function $s$ has been studied since antiquity; it may be the first function ever defined by mathematicians. Beginning with Pythagoras, we have looked for cycles in the dynamical system formed when iterating $s$. There are still a number of unsolved problems connected with this dynamical system: Are there infinitely many cycles? Examples of cycles are $6 \to 6$ and $220 \to 284 \to 220$; about 12 million are known. Does the set of numbers involved in some cycle have asymptotic density 0? We know the upper density is bounded above by about 0.002. Is there an unbounded orbit? The least starting value in question is $n = 276$. (For references on these questions, see [9].)

Perhaps a more basic question with the function $s$ is to identify its image: What numbers are of the form $s(n)$? Note that if $p, q$ are different primes then $s(pq) = p+q+1$. Not many even numbers are of this form, but a slightly stronger version of Goldbach's conjecture (every even number starting with 8 is the sum of two different primes) implies that every odd number starting with 9 is in the range of $s$. Since $s(2) = 1$, $s(4) = 3$, and $s(8) = 7$, while $s(n) = 5$ has no solutions, it then follows from this slightly stronger Goldbach conjecture that every odd number except 5 is in the range of $s$. Moreover, this slightly stronger form of Goldbach's conjecture is known to be usually true. There are many papers in this line, a recent survey is [10].

So, almost all odd numbers (in the sense of asymptotic density) are of the form $s(n)$. In a short beautiful paper, Erdős [4] looked at the even values of $s$, showing that a positive proportion of even numbers are missed. He raises the issue of whether the asymptotic density of even values exists, saying that is not even known if the lower density is positive. Similar questions are raised for the function $s_\varphi(n) := n - \varphi(n)$, where $\varphi$ is Euler's function. Again, almost all odd numbers are attained by $s_\varphi$, but even less is known about even values, compared with $s(n)$. In fact, the Erdős argument that shows that $s$ misses a positive proportion of even values fails for $s_\varphi$.

These thoughts were put in a more general context in [5]. There the following conjecture is formulated.

**Conjecture 1.** *If $\mathcal{A}$ is a set of natural numbers of asymptotic density 0, then $s^{-1}(\mathcal{A})$ also has asymptotic density 0.*

If this is true, one consequence would be that the set of even values of $s$ does not have density 0. Indeed, if $\mathcal{A}$ is the set of even numbers in the range of $s$, then

$$s^{-1}(\mathcal{A}) = \{n \text{ even} : n, n/2 \text{ not squares}\} \cup \{n^2 : n \text{ odd}\},$$

so $s^{-1}(\mathcal{A})$ has asymptotic density $\frac{1}{2}$. Thus, if Conjecture 1 is true, then $\mathcal{A}$ does not have asymptotic density 0.

In this paper we prove the following theorem.

**Theorem 1.** *The set of even numbers of the form $s(n)$ for some integer $n$ has positive lower density.*

Essentially the same proof will show that even numbers of the form $n - \varphi(n)$ comprise a set of positive lower density. It is hoped that the methods in this paper can be of help in proving Conjecture 1.

It seems likely that the asymptotic density of even numbers in the range of $s$ exists. In some numerical work in [12] it appears that the even numbers in the range have density about $\frac{1}{3}$ and the density of even numbers missing is about $\frac{1}{6}$. In [1] it is shown that the lower density of the set of even numbers missing from the range is at least 0.06. The proof of Theorem 1 that we present is effective, but we have made no effort towards finding some explicit lower bound for the lower density of even values of $s$.

## 2 Notation and lemmas

We have the letters $p, q, r$, with or without dashes or subscripts representing prime numbers. We let $\tau(n)$ denote the number of positive divisors of $n$. We say a positive integer $n$ is deficient if $s(n) < n$. We let $P(n)$ denote the largest prime factor of $n$ when $n > 1$, and we let $P(1) = 1$. We say a positive integer $n$ is $z$-smooth if $P(n) \le z$. For each prime $p$ and natural number $n$, we let $v_p(n)$ denote the exponent of $p$ in the prime factorization of $n$. For each large number $n$, let

$$y = y(n) = \log\log n / \log\log\log n.$$

**Lemma 1.** *On a set of asymptotic density* 1 *we have*

(1) $p^{2a} \mid \sigma(n)$ *for every prime power* $p^a \le y$,

(2) $P(\gcd(n, \sigma(n)) \le y$,

(3) $\sigma(n)/\gcd(n, \sigma(n))$ *is divisible by every prime* $p \le y$,

(4) *and every prime factor of* $s(n)/\gcd(n, \sigma(n))$ *exceeds* $y$.

*Proof.* (1) Let $x$ be large, let $y = y(x)$, and let $d$ be an integer with $1 < d \le y$. The integers $n \le x$ with $d^2 \mid \sigma(n)$ include all $n \le x$ which are precisely divisible (i.e., divisible to just the first power) by two different primes $p_1, p_2$ in the residue class $-1 \pmod d$. The number of $n \le x$ which do not have this property is, by the sieve,

$$\begin{aligned}
&\ll x\left(1+\sum_{\substack{p\le x\\ p\equiv -1 \pmod d}}\frac{1}{p}\right)\prod_{\substack{p\le x\\ p\equiv -1 \pmod d}}\left(1-\frac{1}{p}+\frac{1}{p^2}\right)\\
&\ll \frac{x\log\log x}{\varphi(d)}\exp\left(-\frac{\log\log x}{\varphi(d)}\right)\\
&\le \frac{x\log\log x}{\varphi(d)}\exp\left(-\frac{\log\log x}{d}\right)\le \begin{cases}\frac{x}{\varphi(d)}, & \text{if } \frac{1}{2}y<d\le y,\\ \frac{x}{\varphi(d)\log\log x}, & \text{if } d\le \frac{1}{2}y.\end{cases}
\end{aligned}$$

Letting $d$ run over primes and powers of primes, we see that the number of integers $n\le x$ which do not have the property in (1) is $\ll x/\log y = o(x)$ as $x\to\infty$.

(2) In [6, Theorem 8], it is shown that on a set of asymptotic density 1, $\gcd(n,\varphi(n))$ is the largest divisor of $n$ supported on the primes at most $\log\log n$. Virtually the same proof establishes the analogous result for $\gcd(n,\sigma(n))$, so that for almost all $n$, $\gcd(n,\varphi(n))=\gcd(n,\sigma(n))$. (Also see [3, 5, 7, 11].) That the assertion (2) usually holds, it suffices to note that the number of $n\le x$ divisible by a prime in $(y,\log\log x]$ is $o(x)$ as $x\to\infty$.

(3) Let $y=y(x)$, where $x$ is large. This assertion will follow from (1) for $n\le x$ if for each prime power $p^a$ with $p^a\le y<p^{a+1}$, we have $p^{2a}\nmid n$. But, the number of $n\le x$ which fail to have this condition is at most

$$x\sum_{p\le y}\frac{1}{y}=\frac{x}{y}\pi(y)=o(x),\quad x\to\infty.$$

(4) For this part, we have seen that we may assume that for each prime $p\le y$, we have $v_p(\sigma(n))>v_p(n)$. Thus, $v_p(s(n))=v_p(n)=v_p(\gcd(n,\sigma(n))$ for such primes $p$. $\square$

**Lemma 2.** *The set of deficient numbers $n$ for which $s(n)$ is non-deficient has asymptotic density* 0*.*

This result follows from [5, Theorem 5.1] and the continuity of the distribution function for $\sigma(n)/n$.

**Lemma 3.** *On a set of integers $n$ of asymptotic density* 1 *we have $\tau(s(n))=(\log n)^{\log 2+o(1)}$ as $n\to\infty$.*

This result follows from the estimates in [13]. We remark that our proof does not depend on this lemma, we could have used the weaker inequality $\tau(s(n)) \le n^{o(1)}$ which holds for all $n$ as $n \to \infty$, but we thought it good to highlight some other recent research concerning the statistical study of $s(n)$.

**Lemma 4.** *On a set of integers $n$ of asymptotic density 1 we have*

$$\sum_{\substack{r|\sigma(n) \\ r>(\log\log n)^2}} \frac{1}{r} \le 1.$$

This follows by the method of proof of [2, Lemma 5].

# 3 Proof of the theorem

In this section we prove Theorem 1.

*Proof.* We identify a set of integers $\mathcal{A}$ such that every member of $s(\mathcal{A})$ is even and $s(\mathcal{A})$ has positive lower density. We shall pile on a number of conditions for $\mathcal{A}$ to satisfy. For our initial choice for $\mathcal{A}$, we take the set of even deficient numbers. This set has a positive density, see [8]. Let $x$ be large; we study $\mathcal{A}(x) := \mathcal{A} \cap [1, x]$. We assume that each member $n$ of $\mathcal{A}(x)$ is of the form

$$\begin{aligned} &n = pm, \quad p \in \left(\frac{x}{2m}, \frac{x}{m}\right], \quad m = q\ell = qrk, \\ &k \le x^{1/60}, \quad r \in (x^{1/15}, x^{1/12}], \quad q \in (x^{7/20}, x^{11/30}]. \end{aligned}$$

So $n = pm = pq\ell = pqrk$. Note that $n, m, \ell, k$ are all even deficient numbers, each running through a positive proportion of numbers to their respective bounds: $n \le x$, $m \le x^{7/15}$, $\ell \le x^{1/10}$, and $k \le x^{1/60}$. We assume that each of these 4 variables satisfy the properties in the lemmas. We also assume that $k$ has no prime factors in $(y(k), y(x)]$.

Let $y = y(x)$. For each $y$-smooth integer $d$, let $\mathcal{A}_d(x)$ denote the subset of $\mathcal{A}(x)$ consisting of those members $n$ with largest $y$-smooth divisor equal to $d$. There is some number $c$ and a set $\mathcal{D} \subseteq [1, y^c]$ of $y$-smooth numbers $d$ such that

$$\sum_{d \in \mathcal{D}} \frac{1}{d} \gg \log y, \qquad \#\mathcal{A}_d(x) \gg \frac{x}{d \log y},$$

where the latter inequality holds uniformly for $d \in \mathcal{D}$.

For $d \in \mathcal{D}$ and a positive integer $s$, let $r_d(s)$ denote the number of representations of $s$ in the form $s(n)$ for $n \in \mathcal{A}_d(x)$. Clearly,

$$\sum_s r_d(s) = \#\mathcal{A}_d(x) \gg \frac{x}{d \log y}$$

uniformly for all $d \in \mathcal{D}$. Note too that if $d \neq d'$, then we cannot have both $r_d(s), r_{d'}(s) > 0$. Indeed, by Lemma 1, if $r_d(s) > 0$, then $d$ is the largest $y$-smooth divisor of $s$.

We will show that

$$\sum_s r_d(s)^2 \ll \frac{x}{d \log y} \tag{1}$$

uniformly for each $d \in \mathcal{D}$, so that from Cauchy's inequality, it will follow that

$$\#s(\mathcal{A}(x)) = \sum_{d \in \mathcal{D}} \#s(\mathcal{A}_d(x)) \geq \sum_{d \in \mathcal{D}} \frac{\left(\sum_s r_d(s)\right)^2}{\sum_s r_d(s)^2} \gg \sum_{d \in \mathcal{D}} \frac{x}{d \log y} \gg x.$$

The sum $\sum_s r_d(s)^2$ counts solutions to $s(n) = s(n')$ for $n, n' \in \mathcal{A}_d(x)$, with $n = pm, n' = p'm'$. Suppose that $m = m'$. From the equation

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

and $m > 1$ (which implies that $s(m) > 0$), we deduce that $p = p'$. This situation contributes $\sum_s r_d(s)$ to $\sum_s r_d(s)^2$, which is easily seen to be $\ll x/(d \log y)$. Thus, we may assume that $m \neq m'$.

By Lemma 1, we have $\gcd(m, \sigma(m)) = \gcd(m', \sigma(m')) = d$, so that $d \mid (s(m), s(m'))$. Write $\gcd(s(m), s(m')) = dh$. By Lemma 1, every prime factor of $h$ exceeds $y$.

We have from (2),

$$p\frac{s(m)}{dh} - p'\frac{s(m')}{dh} = \frac{\sigma(m') - \sigma(m)}{dh}. \tag{3}$$

For fixed $m, m'$, we count the number of pairs of primes $p,\ p'$ that satisfy this equation. Note that $\sigma(m) \neq \sigma(m')$, since if they would be equal, we would then get from (2) that $ps(m) = p's(m')$, and since $\min\{p, p'\} > \max\{m, m'\} > \max\{s(m), s(m')\}$, we would get that $s(m) = s(m')$, so $m = m'$, which is false. Let $u, u'$ be the integral solution of the linear equation (3) in $p,\ p'$ with $u > 0$ and minimal. Then

$$p = u + \frac{s(m')}{dh}t \qquad \text{and} \qquad p' = u' + \frac{s(m)}{dh}t$$

are both primes and $0 \le t \le (x/m)/(s(m')/dh) = xdh/(ms(m'))$. Let

$$A = \frac{s(m)}{dh} \times \frac{s(m')}{dh} \times \frac{|\sigma(m) - \sigma(m')|}{dh} =: A_1 A_2 A_3, \text{ say.}$$

By the sieve, the number of such $p \le x/m$ is

$$\ll \frac{xdh}{ms(m')(\log(xdh/ms(m')))^2} \frac{A}{\varphi(A)} \ll \frac{xdh}{mm'(\log x)^2} \frac{A_1}{\varphi(A_1)} \frac{A_2}{\varphi(A_2)} \frac{A_3}{\varphi(A_3)}, \tag{4}$$

where the above inequality follows because $ms(m') \le mm' \le x^{14/15}$ and $s(m') \gg m'$. Since $s(m)/(dh)$ and $s(m')/(dh)$ are deficient, it follows that

$$\frac{A_1}{\varphi(A_1)} \ll 1, \qquad \frac{A_2}{\varphi(A_2)} \ll 1$$

However, $A_3/\varphi(A_3)$ is not small. In fact, by Lemma 1, we may assume that $A_3$ is divisible by all primes $\le y = y(x)$, so $\log y \ll A_3/\varphi(A_3) \ll \log\log x$. Write $A_3 = A_{3,1}A_{3,2}A_{3,3}$, where $A_{3,1}$ is the largest divisor with $P(A_{3,1}) \le (\log\log x)^2$ and $A_{3,2}$ is the largest divisor of what remains with $P(A_{3,2}) \le \log x$. Since $A_3$ has $O(\log x/\log\log x)$ distinct prime factors, it follows that $A_{3,3}/\varphi(A_{3,3}) \sim 1$ as $x \to \infty$ and so

$$\frac{A_1 A_2 A_3}{\varphi(A_1)\varphi(A_2)\varphi(A_3)} \ll \frac{A_3}{\varphi(A_3)} \ll \frac{A_{3,2}}{\varphi(A_{3,2})} \log y. \tag{5}$$

Let $A'_{3,2}$ be the largest divisor of $A_{3,2}$ which is coprime to $\sigma(m)$. By Lemma 4, we may assume that $A_{3,2}/\varphi(A_{3,2}) \ll A'_{3,2}/\varphi(A'_{3,2})$. From (4), we now have the problem of showing that for $d \in \mathcal{D}$,

$$\frac{x \log y}{(\log x)^2} \sum_{m,m'} \frac{dhA'_{3,2}}{mm'\varphi(A'_{3,2})} \ll \frac{x}{d \log y}, \tag{6}$$

where $dh = \gcd(s(m), s(m'))$.

We first sum over $m, m'$ with $h > x^{1/3}$, showing that the contribution to (6) is small. With $m = q\ell$ and $h \mid s(m)$, we have

$$s(m) = qs(\ell) + \sigma(\ell) \equiv 0 \pmod{h}. \tag{7}$$

In addition, $h$ and $\sigma(\ell)$ are coprime. Indeed, if some prime $\pi \mid \gcd(h, \sigma(\ell))$, then $\pi = q$ or $\pi \mid s(\ell)$. In the latter case, $\pi \mid \ell$, so $\pi \mid n$. But $\pi \mid \sigma(\ell)$ implies that $\pi \mid \sigma(n)$, so we have a contradiction to our assumption that the properties in Lemma 1 hold. If $\pi = q$, since $\pi \mid \sigma(\ell)$, we again get

$\pi \mid \gcd(n, \sigma(n))$, a contradiction. So, given $h, \ell$ we have from (7) that $q$ is in a fixed coprime residue class modulo $h$; say

$$q \equiv a_{h,\ell} \pmod h.$$

Similarly, we have $m' = q'\ell'$ and $q' \equiv a_{h,\ell'} \pmod h$.

Since $h \mid \gcd(s(m), s(m'))$, (2) implies that $h \mid \sigma(m) - \sigma(m'))$, so that $m \equiv m' \pmod h$. With (7) we get that

$$\frac{\ell\sigma(\ell)}{s(\ell)} \equiv -q\ell = -m \equiv -m' = -q'\ell' \equiv \frac{\ell'\sigma(\ell')}{s(\ell')} \pmod h,$$

which implies

$$s(\ell')\ell\sigma(\ell) - s(\ell)\ell'\sigma(\ell') \equiv 0 \pmod h. \tag{8}$$

The absolute value of the left-hand side is $< 2\max\{\ell^3, \ell'^3\} < 2x^{3/10}$. Thus, for $h > x^{1/3}$, then it must be the case that the integer in the left–hand side of the above congruence must be the zero integer. We thus get that

$$\frac{\ell\sigma(\ell)}{s(\ell)} = \frac{\ell'\sigma(\ell')}{s(\ell')}, \quad \text{or equivalently,} \quad \frac{\ell^2}{s(\ell)} + \ell = \frac{\ell'^2}{s(\ell')} + \ell'. \tag{9}$$

For us, $\gcd(\ell, s(\ell)) = \gcd(\ell', s(\ell')) = d$. Further, by property (3) in Lemma 1, $d\,\mathrm{rad}(d) \mid \gcd(\sigma(\ell), \sigma(\ell'))$, where $\mathrm{rad}(d)$ is the largest squarefree divisor of $d$. Hence, $\gcd(\ell^2, s(\ell)) = d$ and the same is true for $\gcd(\ell'^2, s(\ell'))$. Putting $\ell = d\lambda$, $\ell' = d\lambda'$, we get that

$$\frac{d\lambda^2}{s(\ell)/d} - \frac{d\lambda'^2}{s(\ell')/d} = \ell - \ell',$$

and the two fractions appearing in the left-hand side above are reduced. So, their denominators must be equal, that is, $s(\ell)/d = s(\ell')/d$, therefore $s(\ell) = s(\ell')$. Now equation (9) gives

$$\ell^2 + \ell s(\ell) = \ell'^2 + \ell' s(\ell),$$

and since the function $t^2 + ts(\ell)$ is increasing in $t$, this gives $\ell = \ell'$. Thus, in the case $h > x^{1/3}$, we must have $\ell = \ell'$ and the congruence classes $a_{h,\ell}, a_{h,\ell'}$ of $q$ and $q'$ modulo $h$ are the same.

Summing the expression in (6) over $m, m'$ where $h \mid \gcd(s(m), s(m'))$, $h > x^{1/3}$, and using the maximal order of $A'_{3,2}/\varphi(A'_{3,2})$ , we have

$$\frac{dx \log\log x}{(\log x)^2} \sum_{m,m',h} \frac{h}{mm'} = \frac{dx \log\log x}{(\log x)^2} \sum_{q,q',\ell,h} \frac{h}{qq'\ell^2}.$$

Since $\ell = \ell'$ and $m \neq m'$, we have $q \neq q'$; assume that $q > q'$. Since $q \equiv q' \equiv a_{h,\ell} \pmod h$, the sum of $1/q$ above is $O((\log x)/h)$ and the sum of $1/q'$ is $O(1)$, even forgetting that $q, q'$ are prime. Thus, the above sum reduces to

$$\frac{dx \log\log x}{(\log x)^2} \sum_{\ell,h} \frac{\log x}{\ell^2} \leq \frac{dx \log\log x}{\log x} \sum_{\ell} \frac{\tau(s(m'))}{\ell^2}.$$

The sum of $1/\ell^2$ is $O(x^{-1/10})$, so by Lemma 3, we have the estimate

$$x^{9/10}(\log x)^{O(1)} = O\left(\frac{x}{d\log y}\right),$$

which is consistent with (6).

We now turn to values of $h$ with $h \leq x^{1/3}$. Since $s(m')$ is deficient, $s(m')/\varphi(s(m')) \ll 1$, so that $A'_{3,2}/\varphi(A'_{3,2}) \ll A''_{3,2}/\varphi(A''_{3,2})$, where $A''_{3,2}$ is the largest divisor of $A'_{3,2}$ coprime to $s(m')$. Fix $m'$, $h$ with $h \mid s(m')$ and consider numbers $m$ that can arise. As noted before,

$$m \ \equiv\ \sigma(m) \ \equiv\ \sigma(m') \ \equiv\ m' \pmod h.$$

Since $h \mid s(m)$ and $\gcd(m, \sigma(m)) = d$, we have $\gcd(m, h) = \gcd(\sigma(m), h) = 1$. Thus, the above congruences, rewritten as

$$qrk \equiv (q+1)(r+1)\sigma(k) \equiv m' \pmod h,$$

determine $qr \pmod h$ and $q + r \pmod h$. Hence, there are at most $\tau(h)$ pairs $a, b$ such that $q \equiv a \pmod h$ and $r \equiv b \pmod h$. Fix one of these pairs $a, b$.

Define

$$f(m) = \sum_{\substack{\pi \mid \sigma(m)-\sigma(m') \\ (\log\log x)^2 < \pi \leq \log x \\ \pi \nmid h\sigma(m)}} \frac{1}{\pi},$$

where $\pi$ runs over primes. Note that if $f(m) \leq 1$, then $A''_{3,2}/\varphi(A''_{3,2}) \ll 1$. Say $m = q\ell$ and $\pi$ have $\pi \mid \sigma(m) - \sigma(m')$ and $\pi \nmid \sigma(m)$. Since

$$q\sigma(\ell) = -\sigma(\ell) + \sigma(m) \equiv -\sigma(\ell) + \sigma(m') \pmod \pi,$$

if $\ell, \pi$ are fixed, then $q$ is in a residue class modulo $\pi$, say $c_{\pi,\ell} \pmod \pi$. To summarize, with $m', h, \ell, a, b$ fixed, if $m = q\ell = qrk$ has $\pi \mid A''_{3,2}$, we have $q \equiv c_{\pi,\ell} \pmod \pi$, $q \equiv a \pmod h$, $r \equiv b \pmod h$. Since $\pi \nmid h$, the two congruences for $q$ may be combined to put $q$ in a single residue class modulo $\pi h$. Thus, using $h \le x^{1/3}$ and $\pi \le \log x$,

$$\sum_m \frac{f(m)}{m} \ll \sum_\pi \frac{1}{\pi} \sum_k \frac{1}{k} \sum_r \frac{1}{r} \sum_q \frac{1}{q} \ll \sum_\pi \frac{1}{\pi^2 h} \sum_k \frac{1}{k} \sum_r \frac{1}{r}.$$

To estimate $\sum_r \frac{1}{r}$ we consider two ranges for $h$. Since $r \equiv b \pmod{h}$, we have

$$\sum_r \frac{1}{r} \ll \begin{cases} \frac{\log\log x}{h} + \frac{\log x}{x^{1/15}}, & \text{if } h > x^{1/20}, \\ \frac{1}{h}, & \text{if } h \le x^{1/20}. \end{cases}$$

Thus, in the case that $h > x^{1/20}$, we have

$$\begin{aligned} \sum_m \frac{f(m)}{m} &\ll \sum_\pi \frac{1}{\pi^2 h} \sum_k \frac{1}{k} \left( \frac{\log\log x}{h} + \frac{\log x}{x^{1/15}} \right) \\ &\ll \sum_\pi \frac{\log x \log\log x}{\pi^2 h^2 d \log y} + \sum_\pi \frac{(\log x)^2}{\pi^2 h d x^{1/15}} \\ &\ll \frac{\log x}{h^2 d \log y \log\log x} + \frac{(\log x)^2}{h d x^{1/15}}, \end{aligned} \tag{10}$$

while in the case $h \le x^{1/20}$, a similar calculation shows that

$$\sum_m \frac{f(m)}{m} \ll \sum_\pi \frac{1}{\pi^2 h} \sum_k \frac{1}{kh} \ll \sum_\pi \frac{\log x}{\pi^2 h^2 d \log y} \ll \frac{\log x}{h^2 d \log y (\log\log x)^2}. \tag{11}$$

The expression in (6) for $h \le x^{1/3}$ can be dealt with as follows. Fix $m', h, a, b$. Since $A'_{3,2}/\varphi(A'_{3,2}) \ll 1$ or $\log\log x/\log y$ depending on whether $f(m) \le 1$ or $f(m) > 1$,

$$\begin{aligned} \frac{x \log y}{(\log x)^2} \sum_m \frac{dhA'_{3,2}}{mm'\varphi(A'_{3,2})} &\ll \frac{x \log y}{(\log x)^2} \frac{dh}{m'} \left( \sum_{f(m)\le 1} \frac{1}{m} + \sum_{f(m)>1} \frac{\log\log x}{m \log y} \right) \\ &\le \frac{x \log y}{(\log x)^2} \frac{dh}{m'} \sum_m \frac{1}{m} + \frac{x \log\log x}{(\log x)^2} \frac{dh}{m'} \sum_m \frac{f(m)}{m} \\ &= S_1 + S_2, \quad \text{say.} \end{aligned}$$

First assume that $x^{1/20} < h \le x^{1/3}$. Writing $m = qrk$ and with $m', h, a, b$

fixed, we have

$$\sum_m \frac{1}{m} \ll \sum_k \frac{1}{k} \sum_r \frac{1}{r} \sum_q \frac{1}{q} \ll \frac{1}{h} \sum_k \frac{1}{k} \sum_r \frac{1}{r}$$
$$\ll \frac{1}{h} \sum_k \frac{1}{k} \left( \frac{\log\log x}{h} + \frac{\log x}{x^{1/15}} \right) \ll \frac{\log x \log\log x}{h^2 d \log y} + \frac{(\log x)^2}{h d x^{1/15} \log y}.$$

Thus, summing over choices for $m', h, a, b$,

$$\sum_{m',h,a,b} S_1 \ll \frac{x \log y}{(\log x)^2} \sum_{m',h,a,b} \frac{dh}{m'} \left( \frac{\log x \log\log x}{h^2 d \log y} + \frac{(\log x)^2}{h d x^{1/15} \log y} \right)$$
$$\leq \frac{x}{\log x} \sum_{m',h} \left( \frac{\tau(h) \log\log x}{h m'} + \frac{\tau(h) \log x}{m' x^{1/15}} \right)$$
$$\ll \frac{x \log\log x}{\log x} \sum_{m',h} \frac{\tau(h)}{x^{1/20} m'}.$$

Now

$$\sum_{h | s(m')} \tau(h) \leq \tau(s(m'))^2 \leq (\log x)^{1.4},$$

using Lemma 3. Thus,

$$\sum_{m',h,a,b} S_1 \ll x^{19/20} (\log x)^{O(1)} \sum_{m'} \frac{1}{m'} \leq x^{19/20} (\log x)^{O(1)},$$

which is consistent with our goal in (6).

For $S_2$ we use our estimate (10) for the sum of $f(m)/m$ and summing over choices for $m', h, a, b$, we get that

$$\sum_{m',h,a,b} S_2 \ll \frac{x \log\log x}{(\log x)^2} \sum_{m',h} \frac{dh\tau(h)}{m'} \left( \frac{\log x}{h^2 d \log y \log\log x} + \frac{(\log x)^2}{h d x^{1/15}} \right)$$
$$= \frac{x}{\log x \log y} \sum_{m',h} \frac{\tau(h)}{h m'} + x^{14/15} \log\log x \sum_{m',h} \frac{\tau(h)}{m'}$$
$$\ll \frac{x}{\log x \log y} \frac{(\log x)^{1.4}}{x^{1/20}} \sum_{m'} \frac{1}{m'} + x^{14/15} \log\log x (\log x)^{1.4} \sum_{m'} \frac{1}{m'}$$
$$\leq x^{19/20} (\log x)^{O(1)},$$

which is also consistent with (6).

It remains to consider the case $h \le x^{1/20}$. For a given choice of $m', h, a, b$, we have

$$\sum_m \frac{1}{m} \ll \sum_k \frac{1}{k} \sum_q \frac{1}{q} \sum_r \frac{1}{r} \ll \frac{1}{h^2} \sum_k \frac{1}{k} \ll \frac{\log x}{dh^2 \log y}.$$

Since $\sum_h \tau(h)/h \le (\sum_h 1/h)^2 < 4$, we have

$$\begin{aligned} \sum_{m',h,a,b} S_1 &\ll \frac{x \log y}{(\log x)^2} \sum_{m',h,a,b} \frac{dh}{m'} \frac{\log x}{dh^2 \log y} \\ &\le \frac{x}{\log x} \sum_{m',h} \frac{\tau(h)}{hm'} \ll \frac{x}{\log x} \sum_{m'} \frac{1}{m'} \ll \frac{x}{d \log y}. \end{aligned}$$

This is again consistent with the goal in (6).

Finally, we use (11) to see that

$$\begin{aligned} \sum_{m',h,a,b} S_2 &\ll \frac{x \log\log x}{(\log x)^2} \sum_{m',h} \frac{dh\tau(h)}{m'} \frac{\log x}{h^2 d \log y (\log\log x)^2} \\ &\ll \frac{x}{\log x \log y \log\log x} \sum_{m',h} \frac{\tau(h)}{hm'} \\ &\ll \frac{x}{\log x \log y \log\log x} \sum_{m'} \frac{1}{m'} \\ &\ll \frac{x}{d (\log y)^2 \log\log x}, \end{aligned}$$

which also is in line with (6)

This calculation allows us to conclude that (6) holds, which we have seen then implies our theorem. This completes the proof. $\square$

## 4 Acknowledgements

Research of F. L. on this project was carried on while he visited the Mathematics Department of Dartmouth College in Spring 2014. F. L. thanks the Mathematics Department of Dartmouth College for their hospitality. We thank Paul Pollack for his interest in this paper.

## References


[1] Y.-G. Chen and Q.-Q. Zhao, *Nonaliquot numbers*, Publ. Math. Debrecen **78** (2011), no. 2, 439–442.

[2] J.-M. de Koninck and F. Luca, *On the composition of the Euler function and the sum of divisors function*, Colloq. Math. **108** (2007), no. 1, 31–51.

[3] P. Erdős, *On perfect and multiply perfect numbers*, Ann. Mat. Pura Appl. **42** (1956), 253–258.

[4] P. Erdős, *Über die Zahlen der Form $\sigma(n) - n$ und $n - \varphi(n)$*, Elem. Math. **11** (1973), 83–86.

[5] P. Erdős, A. Granville, C. Pomerance, and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions* Analytic Number Theory, Proc. Conf. in honor of Paul T. Bateman, B. C. Berndt, et al. eds., Birkhauser, Boston, 1990, pp. 165–204.

[6] P. Erdős, F. Luca, and C. Pomerance, *On the proportion of numbers coprime to a given integer*, Anatomy of integers, 47–64, CRM Proc. Lecture Notes, 46, Amer. Math. Soc., Providence, RI, 2008.

[7] I. Kátai and M. V. Subbarao, *Some further remarks on the iterates of the $\varphi$ and the $\sigma$-functions*, Annales Univ. Sci. Budapest, Sect. Comp. **26** (2006), 51–63.

[8] M. Kobayashi, *On the density of abundant numbers*, PhD thesis, Dartmouth College, 2010.

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

[10] J. Pintz, *Recent results on the Goldbach conjecture*, Elementare und analytische Zahlentheorie, 220–254, Schr. Wiss. Ges. Johann Wolfgang Goethe Univ. Frankfurt am Main, 20, Franz Steiner Verlag Stuttgart, Stuttgart, 2006.

[11] P. Pollack, *On the greatest common divisor of a number and its sum of divisors*, Michigan Math. J. **60** (2011), 199–214.

[12] C. Pomerance and H.-S. Yang, *Variant of a theorem of Erdős on the sum-of-proper-divisors function*, Math. Comp. **83** (2014), 1903–1913.

[13] L. Troupe, *On the number of prime factors of values of the sum-of-proper-divisors function.* arXiv 1405.3587.