Mathematika 64 (2018) 330–342 © 2018 University College London
doi:10.1112/S0025579317000535

DIVISOR-SUM FIBERS

PAUL POLLACK, CARL POMERANCE AND LOLA THOMPSON

Abstract. Let $s(\cdot)$ denote the sum-of-proper-divisors function, that is, $s(n) = \sum_{d|n,\ d<n} d$. Erdős, Granville, Pomerance, and Spiro conjectured that for any set $\mathscr{A}$ of asymptotic density zero, the preimage set $s^{-1}(\mathscr{A})$ also has density zero. We prove a weak form of this conjecture: if $\epsilon(x)$ is any function tending to 0 as $x \to \infty$, and $\mathscr{A}$ is a set of integers of cardinality at most $x^{1/2+\epsilon(x)}$, then the number of integers $n \leqslant x$ with $s(n) \in \mathscr{A}$ is $o(x)$, as $x \to \infty$. In particular, the EGPS conjecture holds for infinite sets with counting function $O(x^{1/2+\epsilon(x)})$. We also disprove a hypothesis from the same paper of EGPS by showing that for any positive numbers α and ϵ, there are integers n with arbitrarily many s-preimages lying between $\alpha(1-\epsilon)n$ and $\alpha(1+\epsilon)n$. Finally, we make some remarks on solutions n to congruences of the form $\sigma(n) \equiv a \pmod n$, proposing a modification of a conjecture appearing in recent work of the first two authors. We also improve a previous upper bound for the number of solutions $n \leqslant x$, making it uniform in a.

§1. *Introduction.* Let $s(n)$ denote the sum-of-proper-divisors function of an integer n, i.e. $s(n) = \sigma(n) - n$. The function $s(\cdot)$ has been a source of fascination since the time of the ancient Greeks, who classified numbers as *perfect*, *abundant*, or *deficient*, according to whether $s(n) = n$, $s(n) > n$, or $s(n) < n$. Two thousand years later, the desire to understand statistical properties of $s(n)$ played a motivating role in the early development of probabilistic number theory by figures such as Schoenberg, Davenport, Erdős, and Wintner.

It is interesting that the function s can map sets of asymptotic density 0 to sets of positive density. Indeed, if $\mathscr{A}$ is the set of numbers pq, where p, q are primes, then $\mathscr{A}$ has zero asymptotic density, yet $s(\mathscr{A})$ has asymptotic density $1/2$, the latter claim coming from the fact that the Goldbach conjecture has at most a zero-density set of exceptions. We also know of sets $\mathscr{A}$ of positive density such that $s^{-1}(\mathscr{A})$ not only has zero density, but, in fact, is empty (see Erdős [6]).

Our focus in this paper is on what can be said about $s^{-1}(\mathscr{A})$ when $\mathscr{A}$ has asymptotic density 0. Part of the landscape here is a conjecture proposed by Erdős, Granville, Pomerance, and Spiro [7] (hereafter "EGPS") in their study of the iterates of $s(n)$.

CONJECTURE 1.1. *Let $\mathscr{A}$ be a set of asymptotic density zero. Then $s^{-1}(\mathscr{A})$ also has asymptotic density zero.*

If true, a consequence (see [7]) would be that for each fixed positive integer k, but for a set of numbers n of asymptotic density zero, if $s(n) < n$, then $s_k(n) <$

Received 12 June 2017, published online 3 April 2018.
MSC (2010): 11N37 (primary), 11N64 (secondary).

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

$s_{k-1}(n) < \cdots < n$, where s_j is the j-fold iteration of s. With the inequality signs reversed, this is an unconditional theorem of Erdős. Another consequence of Conjecture 1.1 is given in [**17**].

Some special cases of the EGPS conjecture can be read out of the literature. For example, it is proved in [**11**] that if $\mathscr{A}$ is the set of primes, then the counting function of $s^{-1}(\mathscr{A})$ is $O(x/\log x)$. In addition, Troupe showed in [**20**] that $s^{-1}(\mathscr{A}_\epsilon)$ has density zero for each of the sets

$$\mathscr{A}_\epsilon = \{m : |\omega(m) - \log\log m| > \epsilon \log\log m\}.$$

In [**12**], it is shown that $s^{-1}(\mathscr{A})$ has density zero when $\mathscr{A}$ is the set of palindromes (in any given base). All of these arguments make critical use of structural features of $\mathscr{A}$; the methods do not carry over to arbitrary sets with similar counting functions.

In the present paper, we make some partial progress on Conjecture 1.1. In contrast with the aforementioned results, the structure of $\mathscr{A}$ is irrelevant in the theorem, but we must assume a strong condition on the count of elements of $\mathscr{A}$.

THEOREM 1.2. *Let $\epsilon = \epsilon(x)$ be a fixed function tending to 0 as $x \to \infty$. Suppose that $\mathscr{A}$ is a set of at most $x^{1/2+\epsilon(x)}$ positive integers. Then, as $x \to \infty$,*

$$\#\{n \leqslant x : s(n) \in \mathscr{A}\} = o_\epsilon(x),$$

uniformly in the choice of $\mathscr{A}$.

As an example, as $s(n) < 2n \log\log n$ for all large n and as there are only $O_b(\sqrt{x \log\log x})$ base-b palindromes up to $2x \log\log x$, Theorem 1.2 implies the theorem of [**12**] alluded to above.

Our proof of Theorem 1.2 is presented in §2. We borrow some ideas from recent work of Booker [**3**]; Booker's arguments by themselves almost immediately give the slightly weaker result with $x^{1/2+\epsilon(x)}$ replaced by $x^{1/2-\epsilon}$ for any fixed (constant) $\epsilon > 0$.

EGPS [**7**, p. 170] point out that their Conjecture 1.1 would be a consequence of the following assertion about the sizes of elements in a fiber.

HYPOTHESIS 1.3. *For each positive number θ there exists a constant C_θ such that for all positive integers m there exist at most C_θ numbers $n \leqslant \theta m$ with $s(n) = m$.*

The authors write in [**7**]: "We are not sure we believe this hypothesis and in fact it may be possible to disprove it". Our second theorem, shown in §3, disproves Hypothesis 1.3 in a strong way.

THEOREM 1.4. *There is a constant $c > 0$ for which the following holds. Let α and ϵ be positive real numbers. There are infinitely many m with at least $\exp(c \log m/\log\log m)$ s-preimages that lie in the interval $(\alpha(1-\epsilon)m, \alpha(1+\epsilon)m)$.*

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

Our proof of Theorem 1.4 shows that $c = 1/7$ is admissible.

In §4, we use ideas presented in the previous sections to study solutions to equations of the form $\sigma(n) = kn + a$, where k and a are integers and $k \geqslant 0$. Such equations have been studied extensively by the second author over the course of his career, dating back to a 1975 paper [**18**], and are connected to many classical problems in number theory. We show that a conjecture in [**2**], that was modified in [**16**], needs to be further modified. Furthermore, we show that an upper bound for the count of $n \leqslant x$ satisfying $\sigma(n) \equiv a \pmod n$, which is given in [**18**], can be made uniform in a. In particular, we prove the following.

THEOREM 1.5. *For any integer a, the number of $n \leqslant x$ with $\sigma(n) \equiv a \pmod n$ is $O(x/\log x)$, uniformly in a.*

We remark that [**18**, Corollary 3] appears to give a uniform upper bound, but the dependence on a is suppressed in the notation.

Notation. Throughout this paper, n and m denote positive integers and ℓ, p, and q are primes. We let $P(n)$ and $P^-(n)$ denote the largest and smallest prime factors of n, respectively, and we let $\mathrm{rad}(n) = \prod_{p|n} p$. We write $\log_k x$ for the kth iterate of the natural logarithm function at x.

§2. *The preimages of very sparse sets: proof of Theorem 1.2.* By replacing $\epsilon(x)$ with $\max\{\epsilon(x), 1/\log\log x\}$, we can assume that $x^{\epsilon(x)} \geqslant x^{1/\log\log x}$. We start by introducing a set $\mathscr{E}$ of "exceptional preimages". We let $\mathscr{E}$ be the set of $n \leqslant x$ such that at least one of the following holds:

(a) n has no prime factor in $(1, \log x]$;
(b) n has a divisor in $(x^{1/2-10\epsilon(x)}, x^{1/2+10\epsilon(x)})$;
(c) n has a squarefull part exceeding $x^{2\epsilon(x)}$;
(d) $n \leqslant \sqrt{x}$.

By a simple sieve (inclusion–exclusion), there are $\ll x/\log\log x$ numbers $n \leqslant x$ satisfying (a). Known results on the distribution of divisors (see, e.g., the sharp result of Ford [**8**, Theorem 1]) imply that the number of $n \leqslant x$ satisfying (b) is $o(x)$. The number of $n \leqslant x$ satisfying (c) is $\ll x^{1-\epsilon(x)}$, which is also $o(x)$. Finally, the number of n satisfying (d) is trivially $O(x^{1/2})$. We conclude that $\#\mathscr{E} = o(x)$. Thus, for the sake of proving the theorem, it suffices to bound the number of non-exceptional n with $s(n) \in \mathscr{A}$.

We will show that for each $a \in \mathscr{A}$, the number of non-exceptional preimages of a is $\ll x^{1/2-9\epsilon(x)}$. As $\#\mathscr{A} \leqslant x^{1/2+\epsilon(x)}$, the theorem follows.

Let $a \in \mathscr{A}$, and let n be a non-exceptional preimage of a. Write $n = de$, where d is the largest divisor of n not exceeding $\sqrt{x}$. As n is non-exceptional, $d \leqslant x^{1/2-10\epsilon(x)}$.

As $n > \sqrt{x}$, we have $e > 1$. Let p be any prime divisor of e. Then $dp \mid n$. By the choice of d, we have $dp > \sqrt{x}$, and, as n is non-exceptional,

$$dp \geqslant x^{1/2+10\epsilon(x)}. \tag{2.1}$$

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

Thus,

$$p \geqslant x^{1/2+10\epsilon(x)}/d \geqslant x^{20\epsilon(x)}. \tag{2.2}$$

If p divides both d and e, then n has squarefull part at least $p^2 \geqslant x^{40\epsilon(x)}$, contradicting that n is non-exceptional. Hence, $\gcd(d, e) = 1$.

Observe that

$$\begin{aligned}\frac{\sigma(e)}{e} &= \prod_{p^k \| e}\left(1 + \frac{1}{p} + \cdots + \frac{1}{p^k}\right) \leqslant \prod_{p|e}\left(1 + \frac{1}{p-1}\right)\\ &\leqslant \exp\left(\sum_{p|e}\frac{1}{p-1}\right) \leqslant \exp\left(\frac{\omega(e)}{P^-(e)-1}\right) \leqslant \exp\left(\frac{2\omega(e)}{P^-(e)}\right),\end{aligned}$$

where $\omega(e)$ is the number of distinct primes dividing e. But $\omega(e) \leqslant \log e/\log 2 \leqslant \log x/\log 2$ whereas, from (2.2), $P^-(e) \geqslant x^{20\epsilon(x)} \gg (\log x)^2$. So, $\exp(2\omega(e)/P^-(e)) = 1 + O(\log x/P^-(e))$, and hence

$$\frac{s(e)}{e} = \frac{\sigma(e)}{e} - 1 \ll \frac{\log x}{P^-(e)}.$$

Keeping in mind that $de = n \leqslant x$, we deduce that

$$s(e) \ll e\frac{\log x}{P^-(e)} \leqslant \frac{x \log x}{d \cdot P^-(e)} \leqslant x^{1/2-10\epsilon(x)} \log x; \tag{2.3}$$

here the last inequality follows from (2.1).

As $\gcd(d, e) = 1$, we also have that

$$a = s(de) = \sigma(d)s(e) + s(d)e. \tag{2.4}$$

So if $g = \gcd(\sigma(d), s(d))$, then $g \mid a$, and $a/g = (\sigma(d)/g)s(e) + (s(d)/g)e$. Hence,

$$s(e)(\sigma(d)/g) \equiv a/g \pmod{s(d)/g}.$$

As $\sigma(d)/g$ and $s(d)/g$ are relatively prime, given d we see that $s(e)$ lies in a uniquely determined residue class modulo $s(d)/g$. Combined with (2.3), we deduce that the number of possibilities for $s(e)$ is $\ll 1+x^{1/2-10\epsilon(x)} \log x \cdot g/s(d)$. Moreover, $s(d) \geqslant d/P^-(d) \geqslant d/\log x$. So the number of choices for $s(e)$, given d, is

$$\ll 1 + x^{1/2-10\epsilon(x)}(\log x)^2 g/d.$$

From (2.4), d and $s(e)$ determine e, and hence this last displayed quantity is also a bound on the number of possibilities for $n = de$, given d.

We now sum the upper bound displayed above on possible values of g and d. Each d under consideration has the form gh, where $h \leqslant x^{1/2-10\epsilon(x)}/g$. Thinking of g as fixed and summing on $d = gh$ gives a bound of $\ll (\log x)^3 x^{1/2-10\epsilon(x)}$. Summing on the $\tau(a)$ divisors g of a bounds the number of possibilities for n by $\ll \tau(a)(\log x)^3 x^{1/2-10\epsilon(x)}$. As $a < \sigma(n) \ll x \log\log x$, by the maximal order of the divisor function, $\tau(a) \leqslant x^{0.7/\log\log x}$ for large x. As $x^{0.7/\log\log x}(\log x)^3 < x^{1/\log\log x} \leqslant x^{\epsilon(x)}$ for large x, we see that the number of n that arise in this way is $\ll x^{1/2-9\epsilon(x)}$, as claimed.

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

§3. *Disproof of Hypothesis 1.3: proof of Theorem 1.4.* The following can be deduced from [**1**, Theorem 2.1].

THEOREM 3.1. *For each $\epsilon > 0$ and number x sufficiently large depending on ϵ, there is a finite set $\{m_1, m_2, \dots, m_t\}$ of integers, where t depends only on ϵ and each $m_i > \log x$, with the following property. If $m \leqslant x^{5/12-\epsilon}$ and m is not divisible by any of $m_1, \dots, m_t$, then for each integer u coprime to m, there are $\gg_\epsilon x/\varphi(m) \log x$ primes $p \leqslant x$ with $p \equiv u \pmod m$.*

We prove the following result, which generalizes ideas of Prachar [**19**] and Erdős [**5**].

THEOREM 3.2. *There is a positive absolute constant c such that, for all pairs of integers a, b with $a \neq 0, b > 0$, there are infinitely many integers k with more than $\exp(c \log k/\log\log k)$ representations as $(bp+a)(bq+a)$ with p, q primes.*

Proof of Theorem 3.2. We fix a choice for a, b and let $v = \max\{|a|, b\}$. We choose x, ϵ in Theorem 3.1 with $x > \mathrm{e}^v$ and $\epsilon = \frac{1}{13}$. Let $\mathscr{E}$ be the smallest set of primes such that each exceptional m_i from Theorem 3.1 has a prime factor in $\mathscr{E}$. Let M be the product of all odd primes $\leqslant \frac{1}{3}\log x$ that are not in $\mathscr{E}$ and do not divide ab. As $x > \mathrm{e}^v$, it follows that $|a|, b < \log x$, so the number of prime factors of ab is $\ll \log\log x$. We have $\omega(M) \sim \frac{1}{3}\log x/\log\log x$ and $M = x^{1/3+o(1)}$. For each $d \mid M$, let r_d be the solution to $br_d + a \equiv 0 \pmod d$. As a, b are coprime to d, we must have r_d coprime to d as well. Consider the primes $p \leqslant x$ with $bp + a \equiv 0 \pmod d$. These are the primes $p \equiv r_d \pmod d$. By Theorem 3.1 there are $\gg x/\varphi(d)\log x$ such p.

Now also look at primes $q \leqslant x$ with $(bq + a, M) = M/d$. In order that $r \bmod M$ be a coprime residue class with $(br + a, M) = M/d$, it is necessary and sufficient that

$$r \equiv -a/b \pmod \ell \quad \text{for all } \ell \,\Big|\, \frac{M}{d}$$

and

$$r \not\equiv 0, -a/b \pmod \ell \quad \text{for all } \ell \mid d.$$

There are precisely $\prod_{\ell\mid d}(\ell-2)$ such residue classes $r \bmod M$, and so the number of primes $q \leqslant x$ belonging to one of these classes is

$$\gg \frac{x}{\log x}\frac{\prod_{\ell\mid d}(\ell-2)}{\varphi(M)}.$$

The number of pairs p, q is therefore

$$\gg \frac{x}{\varphi(d)\log x}\cdot\frac{x\prod_{\ell\mid d}(\ell-2)}{\varphi(M)\log x} = \frac{x^2}{\log^2 x}\cdot\frac{1}{\varphi(M)}\prod_{\ell\mid d}\frac{\ell-2}{\ell-1} \gg \frac{x^2}{M\log^2 x}.$$

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

Because of the condition $(bq + a, M) = M/d$, different values of d correspond to different values of q. Thus, we must have

$$\gg \frac{x^2\tau(M)}{M\log^2 x}$$

pairs p, q running over all values d. Map each pair p, q to $(bp + a)(bq + a) \ll v^2x^2$. The number of integers $\ll v^2x^2$ divisible by M is $\ll v^2x^2/M$. By the Pigeonhole Principle, there exists some $k \ll v^2x^2$ with $\gg (x^2\tau(M)/M\log^2 x)/(v^2x^2/M)$ representations as $(bp + a)(bq + a)$, which is

$$\begin{aligned}\gg \frac{\tau(M)}{v^2\log^2 x} > \frac{\tau(M)}{\log^4 x} &= \frac{2^{\omega(M)}}{\log^4 x}\\ &= \frac{2^{(1/3+o(1))(\log x/\log\log x)}}{\log^4 x} = 2^{(1/3+o(1))\log x/\log\log x}\\ &= 2^{(1/6+o(1))\log(v^2x^2)/\log\log(v^2x^2)} \geqslant 2^{(1/6+o(1))\log k/\log\log k}.\end{aligned}$$

As x can be arbitrarily large, this argument produces infinitely many integers k with at least $\exp(c\log k/\log\log k)$ representations as $(bp+a)(bp+q)$ with p, q primes, where $c = 1/10$. □

Remark. The best we can do using Theorem 3.1 would be to replace $\frac{1}{3}$ with any number smaller than $\frac{5}{12}$, which gives the result for any number $c < (5/24)\log 2$. In particular, $c = 1/7$ works.

Proof of Theorem 1.4. We may assume that $0 < \epsilon < 1$. It is well known that the values $s(n)/n$ are dense in $(0, \infty)$. Thus, we may fix an integer $n_0 > 1$ with

$$s(n_0)/n_0 \in (\alpha^{-1}(1 - \tfrac{1}{2}\epsilon), \alpha^{-1}(1 + \tfrac{1}{2}\epsilon)).$$

If p and q are distinct primes not dividing n_0, then

$$\begin{aligned}s(n_0pq) &= \sigma(n_0)(p + 1)(q + 1) - n_0pq\\ &= s(n_0)pq + \sigma(n_0)(p + q + 1),\end{aligned}$$

so that

$$s(n_0)s(n_0pq) = (s(n_0)p + \sigma(n_0))(s(n_0)q + \sigma(n_0)) + s(n_0)\sigma(n_0) - \sigma(n_0)^2. \tag{3.1}$$

By Theorem 3.2, there are infinitely many integers k having more than $\exp(c\log k/\log\log k)$ representations in the form

$$k = (s(n_0)p + \sigma(n_0))(s(n_0)q + \sigma(n_0)),$$

with p, q distinct and

$$p, q > \max\{n_0, \sigma(n_0)/n_0 \cdot (s(n_0)/n_0 \cdot \epsilon/12)^{-1}\}.$$

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

(These latter conditions on p and q exclude only $O(1)$ representations of k.) Letting k be a large integer of this kind, define m in terms of k by

$$m = \frac{k + s(n_0)\sigma(n_0) - \sigma(n_0)^2}{s(n_0)}.$$

Then $m < k$ and m has at least $\exp(c \log m/\log\log m)$ representations in the form $s(n_0 pq)$. Moreover,

$$m = s(n_0 pq) \geqslant s(n_0)pq \geqslant (n_0\alpha^{-1}(1 - \tfrac{1}{2}\epsilon))pq = n_0 pq \cdot \alpha^{-1}(1 - \tfrac{1}{2}\epsilon).$$

Thus,

$$n_0 pq \leqslant \alpha m(1 - \tfrac{1}{2}\epsilon)^{-1} < (1 + \epsilon)\alpha m.$$

On the other hand, using the bounds on p, q,

$$\begin{aligned}\frac{s(n_0 pq)}{n_0 pq} &= \frac{s(n_0)}{n_0} + \frac{\sigma(n_0)}{n_0}\left(\frac{1}{p} + \frac{1}{q} + \frac{1}{pq}\right)\\ &< \frac{s(n_0)}{n_0} + \frac{s(n_0)}{n_0} \cdot \frac{\epsilon}{4} < \alpha^{-1}\left(1 + \frac{7}{8}\epsilon\right).\end{aligned}$$

Rearranging,

$$n_0 pq > \alpha \cdot s(n_0 pq)(1 + 7\epsilon/8)^{-1} > (1 - \epsilon)\alpha m.$$

Thus, m has at least $\exp(c \log m/\log\log m)$ preimages $n = n_0 pq$ in $((1 - \epsilon)\alpha m, (1 + \epsilon)\alpha m)$. □

Remark. In [17], the second author writes "it seems difficult to show there are infinitely many even n with $\#s^{-1}(n) \geqslant 3$". This can now be shown. Following the above construction with $n_0 = 2$, one obtains infinitely many even m with more than $\exp(c \log m/\log\log m)$ s-preimages of the form $2pq$.

§4. *Solutions to $\sigma(n) = kn + a$.* Solving $s(n) = a$ is of course the same as solving $\sigma(n) = n + a$. In [2, 15], and [18] the authors study the more general equation

$$\sigma(n) = kn + a, \tag{4.1}$$

for given k and a. In this context, an integer n is called a *regular* solution to (4.1) if

$$n = pm \quad \text{where } p \nmid m,\ \sigma(m)/m = k,\ \text{and } \sigma(m) = a.$$

All other solutions are called *sporadic*. If there are any regular solutions at all, then there are $\gg_{a,k} x/\log x$ regular solutions up to x, for large x. In [2, Theorem 1], it is shown that the count of sporadic solutions is much smaller: for any integer $k \geqslant 0$ and any integer a with $|a| \leqslant x^{1/4}$, the count of sporadic solutions up to x is at most $x^{1/2+o(1)}$, as $x \to \infty$, uniformly in k and a.

The authors of [2] claim it is plausible that the upper bound $x^{1/2+o(1)}$ can be replaced with a bounded power of $\log x$, even in the wider range $|a| \leqslant x/2$.

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

However, this is provably false in the case $k = 0$, as shown by the method in Erdős [4]. And it is provably false in the case $k = 1$ as there are infinitely many positive integers a with $\gg a/\log^2 a$ representations as $p + q + 1$ where p, q are unequal primes (so that $\sigma(pq) = pq + a$). The first two authors of the present paper go on to state this as a conjecture in [16, Conjecture 2.4], taking into account that k should not be 0 or 1.

CONJECTURE 4.1. *Let $k \geqslant 2$. Let $x \geqslant 3$ and $a \in \mathbb{Z}$ with $|a| \leqslant x/2$. The number of solutions $n \leqslant x$ to $\sigma(n) = kn + a$ is $\ll (\log x)^C$, where both the implied constant and C are absolute constants.*

This statement of Conjecture 4.1 in [16] inadvertently omits the word "sporadic", which is still needed. However, it turns out that confining attention to sporadic solutions is still not enough. For example, when $k = 3$, there are values of a for which the equation $\sigma(n) - 3n = a$ has at least $x^{1/2+o(1)}$ solutions with $n \leqslant x$ of the form $n = 120pq$. This construction depends on the fact that 120 is 3-perfect, i.e., $\sigma(120) = 3 \cdot 120$. One might attempt to further salvage Conjecture 4.1 by barring constructions like $\sigma(n) - 3n = a$ with $n = 120pq$. However, even with this modification, the upper bound on the number of solutions to $\sigma(n) = kn + a$ given in Conjecture 4.1 is too small.

PROPOSITION 4.2. *Let k be a positive integer, and let $C > 0$. There are infinitely many positive integers a for which the equation $\sigma(n) = kn + a$ has more than $(\log a)^C$ solutions $n \leqslant a$ not of the form mp or mpq for any k-perfect number m.*

By considering values of $x = 2a$, we see that Conjecture 4.1 fails for every k, even after restricting n to sporadic solutions not of the form mpq for a k-perfect number m.

Proof (sketch). Fix a positive integer n_0 with $\sigma(n_0)/n_0 > 2k$. We consider numbers $n = n_0pq$, where p, q are distinct primes not dividing n_0. Then, writing $T = \sigma(n_0) - kn_0$, one finds that

$$T(\sigma(n) - kn) = (Tp + \sigma(n_0))(Tq + \sigma(n_0)) + T\sigma(n_0) - \sigma(n_0)^2. \qquad (4.2)$$

(This is the analogue of (3.1).) Following the proof of Theorem 1.4, we produce infinitely many positive integers a that can be written as $\sigma(n) - kn$ for more than $\exp(c \log a/\log\log a)$ numbers $n = n_0pq$. All of our $n = n_0pq$ satisfy $\sigma(n)/n \geqslant \sigma(n_0)/n_0 > 2k$, so that $n < (\sigma(n) - kn)/k = a/k \leqslant a$. Moreover, n does not have the form mpq for a for a k-perfect number m. Indeed, for any primes p and q, both $\sigma(mp)/mp$ and $\sigma(mpq)/mpq$ are bounded above by $(\sigma(m)/m)(1 + 1/2)(1 + 1/3) = 2k$, whereas $\sigma(n)/n > 2k$. As $\exp(c \log a/\log\log a) > (\log a)^C$ for large a, the proposition follows. □

Thus, we cannot hope to get the bound of $(\log x)^C$ stated in Conjecture 4.1. Instead, one might make the following conjecture.

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

CONJECTURE 4.3. *Let k be a positive integer. Let $x \geqslant 3$ and $a \in \mathbb{Z}$ with $|a| \leqslant x$. The number of sporadic solutions $n \leqslant x$ to $\sigma(n) = kn + a$ is at most $x^{1/2+o(1)}$, as $x \to \infty$, uniformly in k, a.*

We have the following variant.

THEOREM 4.4. *Let k be a positive integer. Let $x \geqslant 3$ and $a \in \mathbb{Z}$. The number of sporadic solutions $n \leqslant x$ to $\sigma(n) = kn + a$ is at most $x^{3/5+o_k(1)}$, as $x \to \infty$, uniformly in a.*

Note that the conclusion here is slightly weaker than that of [2, Theorem 1]: we have $x^{3/5+o(1)}$ in place of $x^{1/2+o(1)}$, and the result is not claimed to be uniform in k. The advantage of Theorem 4.4 is that a is no longer restricted to satisfy $|a| \leqslant x^{1/4}$. However, for all solutions $n \leqslant x$, we have $|a| = |\sigma(n) - kn| \ll_k x \log\log x$.

Proof. We can assume $a \neq 0$; if $a = 0$, we are counting multiply perfect numbers, and Hornfeck and Wirsing [9] have shown that the count of such $n \leqslant x$ is $O(x^\epsilon)$ for any fixed $\epsilon > 0$.

If $n = p$ is a prime solution to $\sigma(n) = kn + a$, then $p(1 - k) = a - 1$. If $k \geqslant 2$, this equation determines p, and so there is at most one solution with n prime. If $k = 1$, then for there to be any solution we must also have $a = 1$; but then $n = p$ is a regular solution to $\sigma(n) = kn + a$, not a sporadic solution.

Thus, we may restrict attention to solutions n with at least two prime factors, counted with multiplicity. We write each such n in the form

$$n = n_0 pq \quad \text{where } q = P(n),\ p = P(n/q).$$

We now show that the cases $n_0 \leqslant x^{3/5}$ contribute at most $x^{3/5+o(1)}$ solutions. To start with, we suppose that $(pq, n_0) = 1$ and $p \neq q$.

Let $T = \sigma(n_0) - kn_0$. If $T = 0$, then

$$a = \sigma(n) - kn = \sigma(n_0)(p + 1)(q + 1) - kn_0 pq = kn_0(p + q + 1).$$

Thus, n_0 and $p + q + 1$ are divisors of a. As $|a| \ll x \log\log x$, there are only $x^{o(1)}$ divisors of a, and so there are only $x^{o(1)}$ possibilities for n_0, as well as only $x^{o(1)}$ possibilities for q, given p. As $p \leqslant x^{1/2}$, the case when $T = 0$ leads to only $x^{1/2+o(1)}$ possibilities for $n = n_0 pq$.

Now suppose that $T \neq 0$. Then (cf. (4.2)),

$$Ta - T\sigma(n_0) + \sigma(n_0)^2 = (Tp + \sigma(n_0))(Tq + \sigma(n_0)).$$

If the left-hand side is non-vanishing, appealing once again to the maximal order of the divisor function we may conclude that there are only $x^{o(1)}$ possibilities for p, q, and so only $x^{3/5+o(1)}$ possibilities for $n = n_0 pq$. If the left-hand side is 0, then $\sigma(n_0) = -Tp$ or $\sigma(n_0) = -Tq$. If $\sigma(n_0) = -Tp$, then

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

$$\begin{aligned}\sigma(n_0p) &= -Tp(p+1) = kn_0p(p+1) - \sigma(n_0)p(p+1)\\ &= kn_0p(p+1) - p\sigma(n_0p);\end{aligned}$$

hence, $\sigma(n_0p) = k \cdot n_0p$. As

$$kn_0pq + a = \sigma(n_0pq) = \sigma(n_0p)(q+1) = kn_0p(q+1) = kn_0pq + kn_0p,$$

we have

$$a = kn_0p = \sigma(n_0p).$$

Therefore $n = (n_0p)q$ is a regular solution to $\sigma(n) = kn + a$ and so should not be counted here. A similar analysis applies when $\sigma(n_0) = -Tq$.

It still remains to deal with the cases when either p or q divides n_0 or $p = q$. If q divides n_0, then $p = q$. As there are only $O(\log x)$ primes dividing n_0, there are only $O(x^{3/5}\log x)$ possibilities for $n = n_0pq$. So we may suppose that $q \nmid n_0$. When $p = q$ (and $q \nmid n_0$),

$$\begin{aligned}a &= \sigma(n_0p^2) - kn_0p^2\\ &= (\sigma(n_0) - kn_0)p^2 + \sigma(n_0)(p+1).\end{aligned}$$

Given n_0, the final right-hand side is a polynomial in p of degree 1 or 2, and so p is determined in at most two ways, yielding $O(x^{3/5})$ possibilities for n. We are left with the case when p, q are distinct primes, $q \nmid n_0$, but $p \mid n_0$. In this case, p is determined in at most $O(\log x)$ ways from n_0. Moreover,

$$a = \sigma(n_0pq) - kn_0pq = \sigma(n_0p)(q+1) - kn_0pq = q(\sigma(n_0p) - kn_0p) + \sigma(n_0p).$$

Thus, as long as $\sigma(n_0p) \neq kn_0p$, we see q is determined by n_0 and p, and hence $n = n_0pq$ is determined in $O(x^{3/5}\log x)$ ways. But if $\sigma(n_0p) = kn_0p$, it is easy to see from the last display that $n = (n_0p)q$ is actually a regular solution rather than a sporadic solution.

Thus, we have shown that those cases where $n_0 \leqslant x^{3/5}$ contribute at most $x^{3/5+o(1)}$ solutions. So we may assume that $n_0 > x^{3/5}$ and thus that

$$pq = n/n_0 \leqslant x/n_0 < x^{2/5}.$$

Hence, $p < x^{1/5}$, and there is a divisor d_0 of n in the interval $(x^{2/5}, x^{3/5}]$. Indeed, we may choose d_0 as a certain product of the largest several prime factors of n.

The rest of the argument follows the proof of the main theorem in [2].

With d_0 chosen as above, we let d be the unitary divisor of n satisfying $\mathrm{rad}(d) = \mathrm{rad}(d_0)$. Write $n = de$. Our strategy is to count the number of values of e corresponding to fixed d, d_0, then to sum on d, d_0. As

$$a = \sigma(d)\sigma(e) - kde, \tag{4.3}$$

we have

$$kde \equiv a \pmod{\sigma(d)},$$

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

and thus e is placed in a determined residue class modulo $\sigma(d)/\gcd(\sigma(d), kd)$. As $e \leqslant x/d$, the number of possibilities for e, given d and d_0, is

$$\ll \frac{x \cdot \gcd(\sigma(d), kd)}{d\sigma(d)} + 1 \ll_k \frac{x \gcd(\sigma(d), d)}{d\sigma(d)} + 1 \ll \frac{x \gcd(\sigma(d), d)}{d^2} + 1.$$

Given d_0, the number of $d \leqslant x$ with $\mathrm{rad}(d_0) = \mathrm{rad}(d)$ is $x^{o(1)}$ (see, e.g., [**10**, Lemma 4.2]), and so when summing on d, d_0 the "+1" terms contribute at most $x^{3/5+o(1)}$. This is acceptable for us. We deal with the remaining terms as follows. Put $g = \gcd(d, \sigma(d))$; from (4.3), we have that $g \mid a$. Thus,

$$\sum_{d,d_0} \frac{x \gcd(\sigma(d), d)}{d^2} \leqslant x \sum_{g \mid a} \sum_{\substack{d > x^{2/5} \\ g \mid d}} \frac{g}{d^2} \sum_{d_0 \mid d} 1.$$

The inner sum is just $\tau(d) \leqslant x^{o(1)}$. Now writing $d = gh$, and summing on $h > x^{2/5}/g$, we see that the above expression is

$$\leqslant x^{1+o(1)} \sum_{g \mid a} \frac{1}{g} \sum_{h > x^{2/5}/g} \frac{1}{h^2} \ll x^{3/5+o(1)} \sum_{g \mid a} 1 \leqslant x^{3/5+o(1)}.$$

Collecting all of our estimates completes the proof of the theorem. □

We close with the following result, which makes the upper bound in [**18**, Corollary 3] completely uniform.

Proof Sketch of Theorem 1.5. Let a be an arbitrary integer. We show there are at most $O(x/\log x)$ values of $n \leqslant x$ with $P(n) \mid \sigma(n) - a$. Via standard estimates, we may assume that:

(a) $x/\log x < n \leqslant x$;
(b) $P(n) > x^{1/\log\log x}$;
(c) n is not divisible by a proper power larger than $\log^2 x$.

Let $p = P(n)$. By (b) and (c), we may assume that $p^2 \nmid n$. Write $n = pm$, so that $\sigma(n) = \sigma(m)(p+1)$. We have

$$\sigma(m) \equiv a \pmod{p}. \tag{4.4}$$

Assume that $p > x^{1/2} \log x$ and x is large. Then $\sigma(m) < m \log x < p$. Thus, for a given p, all of the solutions $m_1, m_2, \ldots, m_t$ to (4.4), have $\sigma(m_1) = \sigma(m_2) = \cdots = \sigma(m_t)$. Now for any integer c, the number of solutions $m \leqslant y$ to $\sigma(m) = c$ is $\leqslant y^{1-(1+o(1))\log_3 y/\log_2 y}$ as $y \to \infty$, uniformly in c. This is the σ-analogue of a result in [**13**] (also see [**14**]) for φ. So, there is an absolute constant y_0 such that if $y \geqslant y_0$, then the number of solutions $m \leqslant y$ to $\sigma(m) = c$ is $\leqslant y^{1-1/\log_2 y}$. But for $O(x/\log x)$ choices for $n \leqslant x$, we may assume that $p \leqslant x/y_0$. Thus, the number of m in (4.4) for a given p is at most $(x/p)^{1-1/\log_2(x/p)}$. We now sum this expression on p with $x^{1/2} \log x < p \leqslant x/y_0$. If such a p is in an interval

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

$(x/\mathrm{e}^{j+1}, x/\mathrm{e}^{j}]$, then the number of m corresponding to p is at most $\mathrm{e}^{j(1-1/\log j)}$. Further, the number of choices for p in this interval is $\ll x/(\mathrm{e}^{j}\log(x/\mathrm{e}^{j})) \ll x/(\mathrm{e}^{j}\log x)$, using $p > x^{1/2}\log x$. Thus, the number of choices for n is

$$\ll \frac{x}{\mathrm{e}^{j/\log j}\log x}.$$

Summing on j gets us $\ll x/\log x$.

We now assume that $p \leqslant x^{1/2}\log x$. This implies that $m > x^{1/2}/\log^2 x$. Let $m = uq$ where $q = P(m)$. By standard estimates on smooth numbers, the number of integers $mp \leqslant x$ with $m > x^{1/2}/\log^2 x$ and $P(m) \leqslant \log x$ is at most $x^{1/2+o(1)}$ as $x \to \infty$, so we may assume that $q > \log x$, and so, by (c), $q \nmid u$. Thus, (4.4) implies that

$$\sigma(u)q \equiv a - \sigma(u) \pmod{p}. \tag{4.5}$$

We may assume that $p \nmid \sigma(n)$. Indeed, otherwise, there is a prime power $r^a \mid n$ with $p \mid \sigma(r^a)$, and as $r^a > \frac{1}{2}\sigma(r^a) \geqslant \frac{1}{2}p$, (b) and (c) imply that $a = 1$, that is, $r \equiv -1 \pmod{p}$. As we may assume that x is large enough that $p > 3$, we have $r > p$, contradicting $p = P(n)$. Thus, we may assume that $p \nmid \sigma(u)$ in (4.5). Hence, given u, p, (4.5) completely determines q, using $q < p$. But the number of choices for u, p with $up < x/\log x$ and $p = P(up)$ is $< x/\log x$, and so this estimate completes the proof. □

Acknowledgements. The first author is supported by the National Science Foundation under Grant No. DMS-1402268. This work was completed while the second and third authors were in residence at the Mathematical Sciences Research Institute, during which time they were supported by the National Science Foundation under Grant No. DMS-1440140. The third author is also supported by an AMS Simons Travel Grant.

References

1. W. R. Alford, A. Granville and C. Pomerance, There are infinitely many Carmichael numbers. *Ann. of Math.* (2) **139** (1994), 703–722.
2. A. Anavi, P. Pollack and C. Pomerance, On congruences of the form $\sigma(n) \equiv a \pmod{n}$. *Int. J. Number Theory* **9** (2012), 115–124.
3. A. R. Booker, Finite connected components of the aliquot graph. *Math. Comp.*, doi:10.1090/mcom/3299.
4. P. Erdős, On the normal number of prime factors of $p-1$ and some related problems concerning Euler's φ-function. *Quart. J. Math. (Oxford Ser.)* **6** (1935), 205–213.
5. P. Erdős, On integers which are the totient of a product of two primes. *Quart. J. Math. (Oxford Ser.)* **7** (1936), 16–19.
6. P. Erdős, Über die Zahlen der Form $\sigma(n) - n$ und $n - \varphi(n)$. *Elem. Math.* **28** (1973), 83–86.
7. P. Erdős, A. Granville, C. Pomerance and C. Spiro, On the normal behavior of the iterates of some arithmetic functions. In *Analytic Number Theory, Proc. Conf. in Honor of Paul T. Bateman* (ed. P. T. Bateman *et al*), Birkhauser (Boston, 1990), 165–204.
8. K. Ford, The distribution of integers with a divisor in a given interval. *Ann. of Math.* (2) **168** (2008), 367–433.
9. B. Hornfeck and E. Wirsing, Über die Häufigkeit vollkommener Zahlen. *Math. Ann.* **133** (1957), 431–438.

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press

10. P. Pollack, On the greatest common divisor of a number and its sum of divisors. *Michigan Math. J.* **60** (2011), 199–214.
11. P. Pollack, Some arithmetic properties of the sum of proper divisors and the sum of prime divisors. *Illinois J. Math.* **58** (2014), 125–147.
12. P. Pollack, Palindromic sums of proper divisors. *Integers* **15A** (2015), Paper No. A13, 12 pp.
13. P. Pollack and C. Pomerance, Popular values of Euler's function. *Mathematika* **27** (1980), 84–89.
14. P. Pollack and C. Pomerance, Two methods in elementary analytic number theory. In *Number Theory and Applications* (ed. R. A. Mollin), Kluwer Academic (Dordrecht, 1989), 135–161.
15. P. Pollack and C. Pomerance, On the distribution of some integers related to perfect and amicable numbers. *Colloq. Math.* **130** (2013), 169–182.
16. P. Pollack and C. Pomerance, Some problems of Erdős on the sum-of-divisors function. *Trans. Amer. Math. Soc. Ser. B* **3** (2016), 1–26.
17. P. Pollack and C. Pomerance, The first function and its iterates. In *Connections in Discrete Mathematics: A Celebration of the Work of Ron Graham* (eds S. Butler, J. Cooper and G. Hurlbert), Cambridge University Press (Cambridge, 2018).
18. C. Pomerance, On the congruences $\sigma(n) \equiv a \pmod{n}$ and $n \equiv a \pmod{\varphi(n)}$. *Acta Arith.* **26** (1975), 265–272.
19. K. Prachar, Über die Anzahl der Teiler einer natürlichen Zahl, welche die Form $p-1$ haben. *Monatsh. Math.* **59** (1955), 91–97.
20. L. Troupe, On the number of prime factors of values of the sum-of-proper-divisors function. *J. Number Theory* **150C** (2015), 120–135.

Paul Pollack,
Department of Mathematics,
University of Georgia,
Athens, GA 30602,
U.S.A.
E-mail: pollack@uga.edu

Carl Pomerance,
Department of Mathematics,
Dartmouth College,
Hanover, NH 03755,
U.S.A.
E-mail: carlp@math.dartmouth.edu

Lola Thompson,
Department of Mathematics,
Oberlin College,
Oberlin, OH 44074,
U.S.A.
E-mail: lola.thompson@oberlin.edu

https://doi.org/10.1112/S0025579317000535 Published online by Cambridge University Press