Local behavior of the composition of the aliquot and co-totient functions

Florian Luca and Carl Pomerance

For Krishna Alladi on his sixtieth birthday

Abstract We study the local behavior of the composition of the aliquot function $s(n) = \sigma(n) - n$ and the co-totient function $s_\varphi(n) = n - \varphi(n)$, where σ is the sum-of-divisors function and φ is the Euler function. In particular, we show that $s \circ s_\varphi$ and $s_\varphi \circ s$ are independent in the sense of Erdős, Győry, and Papp.

1 Introduction

In [5], two arithmetic functions $f(n)$ and $g(n)$ are called independent if for all $k \geq 2$ and permutations $i_1, \ldots, i_k$ and $j_1, \ldots, j_k$ of $\{1, 2, \ldots, k\}$, there exist infinitely many n such that

$$\begin{aligned} f(n+i_1) &< f(n+i_2) < \cdots < f(n+i_k), \\ g(n+j_1) &< g(n+j_2) < \cdots < g(n+j_k). \end{aligned} \tag{1}$$

In [5], it was shown that the number-of-prime-divisors function, denoted ω, and the number-of-divisors function, denoted τ, are independent. They also showed that σ,

Florian Luca
School of Mathematics, University of the Witwatersrand, Private Bag X3, Wits 2050, South Africa; Centro de Ciencias Matemáticas, UNAM, Morelia, Mexico e-mail: florian.luca@wits.ac.za

Carl Pomerance
Department of Mathematics, HB 6188, Dartmouth College, Hanover, NH 03755, USA e-mail: carl.pomerance@dartmouth.edu

Mathematics Subject Classification: 11N25, 11N36
Key Words: sum of divisors function, Euler function, applications of sieve methods.

the sum-of-divisors function, and φ, Euler's function, are not independent (when $k \geq 5$). In [2], it was shown that φ and the Carmichael function λ are independent. In [8], it was shown that the compositions $\sigma \circ \varphi$ and $\varphi \circ \sigma$ are independent.

Here, we put

$$s(n) = \sigma(n) - n \quad \text{and} \quad s_\varphi(n) = n - \varphi(n).$$

These functions are well-known in the literature, and the first has an ancient history, dating to Pythagoras. It is not known if the sets of values of these functions has an asymptotic density, though recent progress was made in [7]. Due to the result in [5] that σ and φ are not independent, it seems likely that s and s_φ are also not independent. Our principal result is the following theorem.

Theorem 1 *The functions $s \circ s_\varphi$ and $s_\varphi \circ s$ are independent.*

We also show the following result, by somewhat different methods.

Theorem 2 *The closure of the set of rationals*

$$\left\{ \frac{s(n)}{s_\varphi(n)} : n > 1 \right\}$$

is the interval $[1, \infty)$. The closure of the set of rationals

$$\left\{ \frac{(s \circ s_\varphi)(n)}{(s_\varphi \circ s)(n)} : n \text{ composite} \right\}$$

is $[0, \infty)$. The same is true for the rationals $(s \circ s)(n)/(s_\varphi \circ s_\varphi)(n)$ with n composite.

Note that $s(1) = s_\varphi(1) = 0$ and $(s \circ s_\varphi)(p) = (s_\varphi \circ s)(p) = 0$ for p prime, and this is why there are certain values of n excluded in the sets in Theorem 2.

One can also ask about typical behavior; we find it to be markedly different.

Theorem 3 *There is a set $\mathscr{A}$ of asymptotic density 1 such that*

$$\frac{(s \circ s_\varphi)(n)}{n} \sim \frac{s(n)s_\varphi(n)}{n^2} \sim \frac{(s_\varphi \circ s)(n)}{n}$$

as $n \to \infty$, $n \in \mathscr{A}$. In particular,

$$\lim_{n \to \infty, n \in \mathscr{A}} \frac{(s \circ s_\varphi)(n)}{(s_\varphi \circ s)(n)} = 1.$$

In addition, as $n \to \infty$, $n \in \mathscr{A}$,

$$\frac{(s \circ s)(n)}{n} \sim \left(\frac{s(n)}{n} \right)^2, \quad \frac{(s_\varphi \circ s_\varphi)(n)}{n} \sim \left(\frac{s_\varphi(n)}{n} \right)^2.$$

That $s(s(n))/n$ is normally asymptotic to $(s(n)/n)^2$ is essentially [4, Theorem 5.1 and (5.1)]. Probably $s(n)/s_\varphi(n)$ has a continuous and strictly increasing distribution function on $[1,\infty)$, but we haven't been able to show this. The existence of a distribution function may follow from the methods of [13, Section 3] and [11, Section 3].

Throughout this paper, we use the Vinogradov symbols O, o and the Landau symbols $\gg$, $\ll$ with their usual meaning. The constants implied by them might depend on the fixed parameter k. For a set $\mathscr{A}$ of integers and a real number $t \ge 1$, let $\mathscr{A}(t) = \mathscr{A} \cap [1,t]$. The letters p,q run over primes.

2 The proof of Theorem 1

Let $i \ge 1$ be an integer and

$$f_i(t) = \left(\frac{\sigma(i)}{i}t - 1\right)\left(1 - \frac{\varphi(i)}{it}\right) \quad \text{for real} \quad t \ge 1.$$

Clearly,

$$f_i'(t) = \frac{\sigma(i)}{i} - \frac{\varphi(i)}{it^2} > 0 \quad \text{for} \quad t > 1,$$

so $f_i(t)$ is increasing for $t \ge 1$. Let $k \ge 2$, $C_k := \max\{2, f_i(2) : 1 \le i \le k\}$ and consider two permutations $i_1,\ldots,i_k$ and $j_1,\ldots,j_k$ of $\{1,\ldots,k\}$. Choose real numbers

$$C_k < \alpha_{j_1} < \cdots < \alpha_{j_k},$$

and solve the equations

$$f_i(u_i) = \alpha_i \quad \text{for} \quad u_i > 2 \quad \text{and} \quad i \in \{1,\ldots,k\}.$$

This is possible because $\alpha_i > C_k \ge f_i(2)$. Now choose real numbers

$$0 < \beta_{i_1} < \beta_{i_2} < \cdots < \beta_{i_k} \le 1$$

and put

$$v_i = \frac{\varphi(i)/i}{1 - \beta_i/\alpha_i + (\beta_i/\alpha_i)(\varphi(i)/(iu_i))}.$$

Since $\beta_i/\alpha_i \le \frac{1}{2}$ and $0 < \varphi(i)/(iu_i) < 1$, we have

$$0 < \frac{\varphi(i)}{i} < v_i < \frac{\varphi(i)/i}{1 - \beta_i/\alpha_i} \le 2\frac{\varphi(i)}{i} \le 2,$$

so that for $i = 1,\ldots,k$, we have $0 < v_i < u_i$. Further, the way we have chosen v_i and u_i gives

$$\left(\frac{\sigma(i)}{i}u_i - 1\right)\left(1 - \frac{\varphi(i)}{iu_i}\right) = \alpha_i, \tag{2}$$

$$\left(\frac{\sigma(i)}{i}u_i - 1\right)\left(1 - \frac{\varphi(i)}{iv_i}\right) = \beta_i, \tag{3}$$

for all $i \in \{1,\ldots,k\}$.

Let $\mathscr{Q}$ be the set of odd primes $3 = q_1 < q_2 < \cdots$ such that for all $\ell \ge 2$, q_ℓ is the smallest odd prime with $q_\ell \not\equiv 1 \pmod{q_j}$ for any $j = 1,\ldots,\ell-1$. The first elements of $\mathscr{Q}$ are $3, 5, 17, \ldots$. Erdős [3] showed that

$$\#\mathscr{Q}(t) = (1+o(1))\frac{t}{\log t \log\log t} \quad \text{for} \quad t \to \infty. \tag{4}$$

In particular, by Abel summation,

$$\sum_{\substack{a<q<b\\ q\in\mathscr{Q}}} \frac{1}{q} = \log\log\log b - \log\log\log a + o(1) \tag{5}$$

uniformly in $b > a$ and $a \to \infty$. We now let x be large, and put

$$y := (\log\log x)^4, \quad z := e^{(\log\log x)^{1/2}}, \quad \varepsilon := (\log\log x)^{-1}.$$

By (5), it follows easily that

$$\prod_{\substack{y<q<z\\ q\in\mathscr{Q}}} \left(1 + \frac{1}{q}\right) = \left(\frac{1}{2} + o(1)\right)\frac{\log\log\log x}{\log\log\log\log x} \text{ for } x \to \infty. \tag{6}$$

We choose pairwise disjoint sets of primes

$$\mathscr{Q}_i \subset \mathscr{Q} \cap (y,z) \quad \text{for all} \quad i \in \{1,\ldots,k\}, \tag{7}$$

such that

$$\prod_{q\in\mathscr{Q}_i} \left(1 + \frac{1}{q}\right) \in (u_i - \varepsilon, u_i + \varepsilon) \quad \text{for all} \quad i \in \{1,\ldots,k\}. \tag{8}$$

We select subsets $\mathscr{R}_i \subseteq \mathscr{Q}_i$ such that

$$\prod_{q\in\mathscr{R}_i} \left(1 + \frac{1}{q}\right) \in (v_i - \varepsilon, v_i + \varepsilon) \quad \text{for all} \quad i \in \{1,\ldots,k\}. \tag{9}$$

All this is possible because of (6) and because $v_i < u_i$ for all $i \in \{1,\ldots,k\}$. Put

$$Q_i = \prod_{q\in\mathscr{Q}_i} q \quad \text{and} \quad R_i = \prod_{q\in\mathscr{R}_i} q \quad \text{for all} \quad i \in \{1,\ldots,k\}.$$

We now choose for each $i \in \{1, \dots, k\}$, U_i to be the smallest prime such that

$$\begin{aligned} U_i &\equiv -1 + R_i + 2R_i^2 \pmod{R_i^3} \\ U_i &\equiv 2 + 2(Q_i/R_i)^2 \pmod{(Q_i/R_i)^3} \\ U_i &\equiv 2 + 2Q_j^2 \pmod{Q_j^3} \quad \text{for all} \quad j \in \{1, \dots, k\} \backslash \{i\}. \end{aligned} \tag{10}$$

Note that $U_1, \dots, U_k$ are distinct and $\{U_1, \dots, U_k\}$ is disjoint from $\bigcup_{i=1}^k \mathcal{Q}_i$ because any prime among the U's is larger than any member of $\bigcup_{i=1}^k \mathcal{Q}_i$. The congruences (10) put U_i in a certain arithmetic progression modulo $(\prod_{i=1}^k Q_i)^3$. By a result of Xylouris [14], we have

$$U_i \ll \left(\prod_{i=1}^k Q_i^3 \right)^5 < \left(\prod_{p<z} p \right)^{15}. \tag{11}$$

Finally, let

$$P = \prod_{\substack{2k<p<3z^2 \\ p \notin \bigcup_{i=1}^k \mathcal{Q}_i \cup \{U_1, \dots, U_k\}}} p.$$

Consider the Chinese Remainder Theorem system of congruences:

$$\begin{aligned} n &\equiv 0 \pmod{(2k)!P} \\ n &\equiv -i + Q_i^2 U_i \pmod{Q_i^3 U_i^2} \quad \text{for all} \quad i = 1, \dots, k. \end{aligned} \tag{12}$$

Congruences (12) put n into an arithmetic progression of modulus

$$M := (2k)!P \prod_{i=1}^k (Q_i^3 U_i^2). \tag{13}$$

Recalling that implied constants depend on the choice of k, note that from (11),

$$\begin{aligned} M &\ll P \left(\prod_{j=1}^k Q_i \right)^3 \left(\prod_{i=1}^k U_i \right)^2 \ll P \left(\prod_{i=1}^k Q_i \right)^{3+30k} \\ &< \left(\prod_{p<3z^2} p \right) \left(\prod_{p<z} p \right)^{3+30k} \ll e^{4z^2}, \end{aligned} \tag{14}$$

so that $M \le e^{(\log x)^{o(1)}}$ as $x \to \infty$. Write

$$n = M\lambda + N_0,$$

where $0 < N_0 < M$ is the smallest positive integer in the progression. Then

$$\begin{aligned} n+i \ &= \ M\lambda + (i+N_0) = iQ_i^2U_i(M_i\lambda+N_i), \\ \text{where } \ M_i &:= \frac{M}{iQ_i^2U_i} \\ \text{and } \ N_i &:= \frac{i+N_0}{iQ_i^2U_i} \quad \text{for all} \quad i\in\{1,\ldots,k\}. \end{aligned} \tag{15}$$

Fix $i\in\{1,\ldots,k\}$. We start by noting that M_i is divisible by all primes $p\le 3z^2$, while the numbers N_i are coprime to all primes $p\le 3z^2$ for $i\in\{1,\ldots,k\}$. To justify this claim, first of all let us note that for large x, we have $y>2k$, so all primes in $\mathscr{Q}_i$, all primes dividing P, and all the primes $U_1,\ldots,U_k$ are larger than $2k$. Then:

(i) Since $i\le k$, we get $i^2 \mid (k!)^2 \mid (2k)! \mid M$, so $i\mid M_i$.
(ii) If $p\in(k,2k)$, then $p\mid(2k)!$ and $p\nmid i$ for any $\in\{1,\ldots,k\}$, so $p\mid M_i$.
(iii) Since, $Q_i^3\mid M$, we have $Q_i\mid M_i$.
(iv) If $U_i\le 3z^2$ for some $i\in\{1,\ldots,k\}$, then $U_i^2\mid M$, so $U_i\mid M_i$.
(v) If $p\in(2k,3z^2]$ and $p\not\in\bigcup_{i=1}^k\mathscr{Q}_i\cup\{U_1,\ldots,U_k\}$, then $p\mid P\mid M_i$.

From (i)–(v) above, we get that $p\mid M_i$ for all $p\le 3z^2$.

Similar observations show that N_i is not a multiple of any prime $p\le 3z^2$. Indeed, if $p\mid i$, then $p^2\mid i^2\mid(2k)!\mid N_0$, so p does not divide $(i+N_0)/i$. If $p\le k$ does not divide i, then p divides N_0 but not N_0+i, so $p\nmid N_i$. If $p\in(k,2k)$, then $p\mid N_0$ and $p\nmid i$, so p does not divide $i+N_0$, and in particular $p\nmid N_i$. If $p\in\mathscr{Q}_i$, then $p^2\,\|\,i+N_0$, so $p\nmid N_i$. Similar arguments show that p does not divide N_i if p is either in $\{U_1,\ldots,U_k\}$, or if it divides P.

So, N_i is coprime to all primes $p\le 3z^2$ as well as with the primes in $\{U_1,\ldots,U_k\}$. In particular,

$$\gcd(M_i,N_i)=1.$$

Further, $M_i\lambda+N_i$ is neither a multiple of any prime $p\le 3z^2$ nor a multiple of any of the primes in $\{U_1,\ldots,U_k\}$. Then we have for $i=1,\ldots,k$,

$$\begin{aligned} s_\varphi(n+i) &= (n+i)-\varphi(n+i) \\ &= iQ_i^2U_i(M_i\lambda+N_i)-\varphi(i)Q_i(U_i-1)\varphi(M_i\lambda+N_i)\prod_{q\in\mathscr{Q}_i}(q-1) \\ &= iQ_iT_i, \end{aligned} \tag{16}$$

where

$$T_i := Q_iU_i(M_i\lambda+N_i)-\varphi(i)(U_i-1)(\varphi(M_i\lambda+N_i)/i)\prod_{q\in\mathscr{Q}_i}(q-1).$$

Also,

$$\begin{aligned} s(n+i) &= \sigma(n+i)-(n+i) \\ &= \sigma(i)(U_i+1)\sigma(M_i\lambda+N_i)\prod_{q\in\mathscr{Q}_i}(q^2+q+1)-iQ_i^2U_i(M_i\lambda+N_i) \\ &= iR_iS_i, \end{aligned} \tag{17}$$

where

$$S_i = \sigma(i)((U_i+1)/R_i)(\sigma(M_i\lambda+N_i)/i)\prod_{q\in\mathscr{Q}_i}(q^2+q+1) - Q_i(Q_i/R_i)U_i(M_i\lambda+N_i).$$

Now we start sieving. Note that if $\lambda \le x/M$, then $M_i\lambda+N_i < n < 2x$. Let us throw away some values of $\lambda \le x/M$ in such a way that at each step we only throw an amount of λ of order of magnitude

$$o\left(\frac{x}{M}\right) \quad \text{as} \quad x\to\infty.$$

There exists an absolute constant c_0 such that for every $i\in\{1,\dots,k\}$, the set Λ_1 of $\lambda \le x/M$ such that $\varphi(M_i\lambda+N_i)$ or $\sigma(M_i\lambda+N_i)$ is not divisible by all numbers $m < c_0\log\log x/\log\log\log x$ is of cardinality $\ll x/(M\log\log x)$. For φ and without the arithmetic progression, this follows from Lemma 2 in [6]. The proof of that lemma can be adapted in a straightforward way to yield the current result. So, we ignore $\lambda\in\Lambda_1$, and assume from now on that

$$\varphi(M_i\lambda+N_i), \sigma(M_i\lambda+N_i) \ \text{ are multiples of all numbers } \ m \le c_0\frac{\log\log x}{\log\log\log x}.$$

In particular, this implies that T_i, S_i are integers.

We eliminate $\lambda\in\Lambda_2$, where this set is such that for some $i\in\{1,\dots,k\}$ we have that $M_i\lambda+N_i$ is a multiple of a prime $p > x/M$. Assume $\lambda\in\Lambda_2$. Then for some $i\in\{1,\dots,k\}$ we have that $M_i\lambda+N_i = pm$, where $m < 2M$. Fixing m, this puts $\lambda\le x/M$ into a certain progression modulo m, such that $M_i\lambda+N_i\equiv 0 \pmod m$ and $(M_i\lambda+N_i)/m = p$ is prime. Thus, the number of $\lambda\le x/M$ satisfying these conditions is

$$\ll \frac{M_i m}{\varphi(M_i m)}\frac{x/Mm}{\log(x/Mm)} \ll \frac{x\log\log x}{Mm\log x},$$

using the minimal order of φ and (14). Summing on $m<2M$ and on $i=1,\dots,k$ and again using (14), we have

$$\#\Lambda_2 \ll \frac{x\log M\log\log x}{M\log x} \ll \frac{xz^2\log\log x}{M\log x} \ll \frac{x}{M\sqrt{\log x}} = o\left(\frac{x}{M}\right)$$

as $x\to\infty$.

We eliminate $\lambda\in\Lambda_3$ such that $M_i\lambda+N_i$ is not squarefree. So, assume that $i\in\{1,\dots,k\}$ and $M_i\lambda+N_i$ is not squarefree. Thus, some $p^2 \mid M_i\lambda+N_i$. Assume first that $p^2 < x/M$. Then the number of such $\lambda\le x/M$ is $\ll x/(Mp^2)$. Summing over all $i\in\{1,\dots,k\}$ and $p>3z^2$, we get a bound of

$$\ll \sum_{p>3z^2}\frac{x}{Mp^2} \ll \frac{x}{M}\sum_{p>3z^2}\frac{1}{p^2} \ll \frac{x}{Mz^2} = o\left(\frac{x}{M}\right) \quad \text{as} \quad x\to\infty. \tag{18}$$

Assume now that $p^2 > x/M$. Since $M_i\lambda + N_i < 2x$, we have $p < \sqrt{2x}$. Moreover, each such p gives rise to at most one value of $\lambda \le x/M$. Thus, the number of choices of λ in this case is at most $\pi(\sqrt{2x}) < \sqrt{x} = o(x/M)$ as $x \to \infty$, by (14). We deduce from (18) that

$$\#\Lambda_3 = o\left(\frac{x}{M}\right) \quad \text{as} \quad x \to \infty.$$

We eliminate $\lambda \in \Lambda_4$ such that for some $i \in \{1,\dots,k\}$,

$$\omega(M_i\lambda + N_i) \ge 10 \log\log x.$$

For this, write $M_i\lambda + N_i = m'm$ where the least prime factor of m' exceeds the greatest prime factor of m and m is maximal with $m \le x/M$. Since $\sqrt{M_i\lambda + N_i} < \sqrt{2x} < x/M$ and since $M_i\lambda + N_i$ is squarefree (using $\lambda \notin \Lambda_3$), it follows that $\omega(m) \ge \frac{1}{2}\omega(M_i\lambda + N_i)$. Thus, summing over $i = 1,\dots,k$,

$$\#\Lambda_4 \ll \sum_{\substack{m \le x/M \\ m \text{ squarefree} \\ \omega(m) \ge 5 \log\log x}} \frac{x}{Mm} \le \frac{x}{M} \sum_{j \ge 5 \log\log x} \frac{1}{j!} \left(\sum_{p \le x/M} \frac{1}{p} \right)^j \ll \frac{x}{M(\log x)^3},$$

since the inner sum on p is $\log\log x + O(1)$. Thus, $\#\Lambda_4 = o(x/M)$ as $x \to \infty$.

We now eliminate those $\lambda \in \Lambda_5$ where for some $i = 1,\dots,k$, we have a prime $p \mid M_i\lambda + N_i$ with $\omega(p-1) > 5 \log\log x$. Since $\lambda \notin \Lambda_2$, we may assume that $p \le x/M$. For a given p with $\omega(p-1) > 5 \log\log x$ and a given $i \in \{1,\dots,k\}$, the number of $\lambda \le x/M$ with $M_i\lambda + N_i$ divisible by p is $\ll x/Mp < x/M(p-1)$. Writing $p - 1 = m$ and ignoring that p is prime, we have

$$\#\Lambda_5 \ll \frac{x}{M} \sum_{\substack{m \le x/M \\ \omega(m) > 5 \log\log x}} \frac{1}{m} \le \frac{x}{M} \sum_{j > 5 \log\log x} \frac{1}{j!} \left(\sum_{q^a \le x/M} \frac{1}{q^a} \right)^j,$$

where q^a runs over prime powers. Since the inner sum is $\log\log x + O(1)$, we have, as with the calculation for Λ_4, that $\#\Lambda_5 = o(x/M)$ as $x \to \infty$.

We eliminate $\lambda \in \Lambda_6$ such that for some $i \in \{1,\dots,k\}$, we have that

$$\gcd((M_i\lambda + N_i)U_i, \varphi(M_i\lambda + N_i)) > 1.$$

Since $\lambda \notin \Lambda_3$, we have that $M_i\lambda + N_i$ is squarefree. Thus, if for some $i \in \{1,\dots,k\}$, the number $M_i\lambda + N_i$ and its Euler function are not coprime, then there are primes p and q with $pq \mid M_i\lambda + N_i$ and $p \mid q - 1$. There are two cases here to consider. If $pq \le x/M$, then we fix i, p, q and we get that the number of such $\lambda \le x/M$ is $\ll x/Mpq$. Summing up this inequality over all $q \equiv 1 \pmod{p}$ with $q \le x/M$ (using [12, Theorem 1, Remark 1]), then over all $p \in (3z^2, x/M)$, then over all $i \in \{1,\dots,k\}$, we get a bound of

$$k \frac{x \log\log x}{Mz^2} = o\left(\frac{x}{M}\right) \quad \text{for} \quad x \to \infty.$$

The other case to consider is when $pq > x/M$. We then write $M_i\lambda + N_i = pqm$, where $m < 2M$ and fix m. Since $q > p$, we get that $p < 2x^{1/2}$. So, $pm < 4x^{1/2}M < x/M$ for large x. Fixing also p, we get that $\lambda \le x/M$ is in a certain arithmetic progression modulo pm such that $pm \mid M_i\lambda + N_i$ and $(M_i\lambda + N_i)/pm$ is prime. The number of such λ is, using the minimal order of φ,

$$\ll \frac{M_i m}{\varphi(M_i m)} \frac{x}{Mpm\log x} \ll \frac{x\log\log x}{Mpm\log x}.$$

Summing over $p \le 2x^{1/2}$, $m < 2M$, $i \le k$, we get an estimate that is

$$\ll \frac{x(\log\log x)^2 \log M}{M\log x} = o\left(\frac{x}{M}\right),$$

as $x \to \infty$, using (14). Finally consider the case that $U_i \mid \varphi(M_i\lambda + N_i)$. Then there is a prime $q \equiv 1 \pmod{U_i}$ with $q \mid M_i\lambda + N_i$. Again using [12], $\sum 1/q \ll (\log\log x)/U_i$, so the number of such $\lambda \le x/M$ is $o(x/M)$ as $x \to \infty$. Thus, $\#\Lambda_6 = o(x/M)$ as $x \to \infty$.

We eliminate $\lambda \in \Lambda_7$ such that for some $i \in \{1, \dots, k\}$ we have

$$\gcd(M_i\lambda + N_i, U_i - 1) > 1.$$

Assume that $i \in \{1, \dots, k\}$ and that there is a prime $p \mid \gcd(M_i\lambda + N_i, U_i - 1)$. Then $p > 3z^2$. Fixing p, we have $p \le x/M$ because $\lambda \notin \Lambda_2$, therefore the number of such $\lambda \le x/M$ is $\le 1 + x/Mp \le 2x/Mp$. Summing this over all the prime divisors $p > 3z^2$ of $U_i - 1$, we get a bound of

$$\ll \frac{x\omega(U_i - 1)}{Mz^2} \ll \frac{x\log U_i}{Mz^2 \log\log U_i} \ll \frac{x\log M}{Mz^2\log\log M} \ll \frac{x}{M\log\log M}$$

where we used the maximal order of $\omega(m)$ together with (14). Summing this up over $i \in \{1, \dots, k\}$, we get

$$\#\Lambda_7 \ll \frac{x}{M\log\log M} = o\left(\frac{x}{M}\right) \quad \text{as} \quad x \to \infty.$$

Now let us look at

$$T_i = Q_i U_i (M_i\lambda + N_i) - \varphi(i)(U_i - 1)(\varphi(M_i\lambda + N_i)/i)\varphi(Q_i).$$

Note that

$$\gcd(Q_i U_i (M_i\lambda + N_i), \varphi(i)(U_i - 1)\varphi(Q_i)) = 1;$$
$$\gcd(U_i(M_i\lambda + N_i), \varphi(M_i\lambda + N_i)) = 1.$$

Indeed, Q_i is coprime to $\varphi(Q_i)$ due to the definition of the set $\mathcal{Q}$. The other relations follow from the sizes of the primes involved, the definition of U_i, and since $\lambda \notin \Lambda_6 \cup \Lambda_7$. So, it follows that

$$\gcd(Q_iU_i(M_i\lambda+N_i),\varphi(i)(U_i-1)(\varphi(M_i\lambda+N_i)/i)\varphi(Q_i))=W_i,$$

where

$$W_i:=\gcd(Q_i,\varphi(M_i\lambda+N_i)).$$

We may write

$$T_i=W_i'T_i',$$

where W_i' is a multiple of W_i and is the largest divisor of T_i supported on the prime factors of W_i, and where T_i' is coprime to W_i and its least prime factor exceeds $c_0\log\log x/\log\log\log x$ (because $\lambda\not\in\Lambda_1$).

We continue with the sieving. We put

$$Y:=\exp\left(\frac{\log x\log\log\log x}{\log\log x}\right),$$

and eliminate $\lambda\in\Lambda_8$ such that for some $i\in\{1,\ldots,k\}$ we have that the largest prime factor P of $M_i\lambda+N_i$ satisfies $P\le Y$. Since the conditions (1.10) for the main theorem in [1] are fulfilled, we get that

$$\#\Lambda_8\ll\frac{x}{M\exp(u\log u+u\log\log u)}=o\left(\frac{x}{M}\right)\quad\text{as}\quad x\to\infty,$$

where $u=\log(xM_i/M)/\log Y=(1+o(1))\log\log x/\log\log\log x$ as $x\to\infty$.

We eliminate $\lambda\in\Lambda_9$ such that $\omega(T_i')>100\log\log x$. Fix $i\in\{1,\ldots,k\}$ and assume that $\omega(T_i')>100\log\log x$. Write $M_i\lambda+N_i=Pm$, where $Y<P\le x/M$ and $m<2x/Y$. This is possible because $\lambda\not\in\Lambda_2\cup\Lambda_8$. Further, P and m are coprime, and m is squarefree because $\lambda\not\in\Lambda_3$. Substituting $M_i\lambda+N_i=Pm$ into

$$iT_i=iQ_iU_i(M_i\lambda+N_i)-\varphi(i)(U_i-1)\varphi(M_i\lambda+N_i)\varphi(Q_i)=iW_i'T_i',$$

we get

$$A_iP+B_i=iW_i'T_i',$$

where

$$A_i=iQ_iU_im-\varphi(i)(U_i-1)\varphi(m)\varphi(Q_i),\quad B_i=\varphi(i)(U_i-1)\varphi(m)\varphi(Q_i).$$

It follows from the definition of T_i' that T_i' is coprime to $\gcd(A_i,B_i)$, and so Q_i,m,T_i' are pairwise coprime. We consider the prime factors of T_i' in three ranges:

(i) At least $x^{1/(20\log\log x)}$ (the number of such is at most $20\log\log x$);
(ii) At most $x^{1/(\log\log x)^2}$;
(iii) In the interval $I=[x^{1/(\log\log x)^2},x^{1/(20\log\log x)}]$.

Suppose that T_i' has at least $5\log\log x$ prime divisors in the range (ii) above and let τ be the product of $\lceil 5\log\log x\rceil$ of them. Then $\tau<x^{6/\log\log x}<\sqrt{Y}$. The relation $\tau\mid A_iP+B_i$ with τ coprime to $\gcd(A_i,B_i)$ puts P in a certain arithmetic progression modulo τ, which for a fixed value of m, puts λ in a particular arithmetic progression modulo $m\tau$. Ignoring that P is prime and since $m\tau < 2x/\sqrt{Y} < x/M$, the number of such $\lambda \le x/M$ is

$$\ll \frac{x}{Mm\tau}.$$

We now sum over all possible $m < 2x/Y$ and all possible τ, a number which is at most x/M but has $\lceil 5\log\log x\rceil$ distinct prime factors. We get

$$\frac{x}{M}\sum_{m\le x}\frac{1}{m}\sum_{\substack{\tau\le x/M\\ \omega(\tau)\ge 5\log\log x}}\frac{1}{\tau} \ll \frac{x\log x}{M}\sum_{\substack{\tau\le x/M\\ \omega(\tau)\ge 5\log\log x}}\frac{1}{\tau} \ll \frac{x}{M(\log x)^2} = o\left(\frac{x}{M}\right),$$

as $x\to\infty$, where the last bound follows from the argument above for Λ_4 or Λ_5.

Assume now that T_i' has at least $K := \lfloor 5\log\log\log x\rfloor$ prime divisors in I and let τ' be a product of K of them. Then

$$\tau' < x^{5\log\log\log x/20\log\log x} < \sqrt{Y}.$$

Thus, for fixed τ', m, we have the number of $\lambda \le x/M$ with λ in the particular class mod $m\tau'$ and $(M_i\lambda + N_i)/m = p$ prime, is at most the number of primes $p \le 2M_ix/(Mm)$ which are in a fixed congruence class modulo $M_i\tau'$, and this is

$$\ll \frac{M_i\tau'}{\varphi(M_i\tau')}\frac{x}{Mm\tau'\log((2x)/(Mm\tau'))} \ll \frac{x\log\log x}{Mm\tau'\log Y},$$

using $m\tau' \ll x/\sqrt{Y}$ and (14). (We have also used the minimal order of φ.) Thus, by an argument similar to the preceding one we get that the number of such numbers $\lambda \le x/M$ is

$$\ll \frac{x(\log\log x)^2}{M\log x}\sum_{m\le x}\frac{1}{m}\sum_{\substack{\tau'\ \text{squarefree}\\ \omega(\tau')=K\\ p\mid\tau'\Rightarrow p\in I}}\frac{1}{\tau'} \ll \frac{x(\log\log x)^2}{M}\frac{1}{K!}\left(\sum_{p\in I}\frac{1}{p}\right)^K.$$

The inner sum is $\log\log\log x + O(1)$, and since $5(\log 5 - 1) > 3$, the estimate is $o(x/M)$ as $x\to\infty$. Thus, we get that the number of λ for which there are at least K prime factors of T_i' in I is $o(x/M)$ as $x\to\infty$. To summarize, except for a set of λ of cardinality $o(x/M)$, the number T_i' has at most $25\log\log x + 5\log\log\log x$ prime factors. Thus, Λ_9 has cardinality $o(x/M)$ as $x\to\infty$ and we may assume therefore that $\omega(T_i') < 100\log\log x$.

Now we have

$$\begin{aligned}(s\circ s_\varphi)(n+i) &= \sigma(s_\varphi(n+i)) - s_\varphi(n+i)\\ &= \sigma(iQ_iW_i'T_i') - iQ_iW_i'T_i'\\ &= s_\varphi(n+i)\left(\frac{\sigma(i)}{i}\frac{\sigma(Q_iW_i')}{Q_iW_i'}\frac{\sigma(T_i')}{T_i'} - 1\right).\end{aligned}$$

Thus,

$$\begin{aligned}\frac{(s\circ s_\varphi)(n+i)}{n+i} &= \frac{(s\circ s_\varphi)(n+i)}{s_\varphi(n+i)}\,\frac{s_\varphi(n+i)}{n+i}\\ &= \left(\frac{\sigma(i)}{i}\,\frac{\sigma(Q_iW_i')}{Q_iW_i'}\,\frac{\sigma(T_i')}{T_i'}-1\right)\\ &\quad\times\left(1-\frac{\varphi(i)}{i}\left(1-\frac{1}{U_i}\right)\frac{\varphi(Q_i)}{Q_i}\,\frac{\varphi(M_i\lambda+N_i)}{M_i\lambda+N_i}\right).\end{aligned}$$

So, let us see what we have. Since all primes dividing W_i' are in $\mathcal{Q}_i$, it follows that

$$\frac{\sigma(Q_i)}{Q_i}\le\frac{\sigma(Q_iW_i')}{Q_iW_i'}\le\frac{\sigma(Q_i)}{Q_i}\exp\left(O\left(\sum_{q>y}\frac{1}{q^2}\right)\right)=(1+o(1))\frac{\sigma(Q_i)}{Q_i},$$

for $x\to\infty$. Further, since primes dividing T_i' exceed $c_0\log\log x/\log\log\log x$ and the number of them is $<100\log\log x$, we get that for large x,

$$\frac{\sigma(T_i')}{T_i'}=1+o(1)\quad\text{as}\quad x\to\infty.$$

Since $\lambda\notin\Lambda_4$, $M_i\lambda+N_i$ has $O(\log\log x)$ prime factors all larger than $3z^2>y=(\log\log x)^4$, so

$$\frac{\varphi(M_i\lambda+N_i)}{M_i\lambda+N_i}=1+o(1)\quad\text{as}\quad x\to\infty.$$

Finally,

$$\frac{\sigma(Q_i)\varphi(Q_i)}{Q_i^2}=\prod_{q\in\mathcal{Q}_i}\left(1-\frac{1}{q^2}\right)=\exp\left(O\left(\sum_{q>y}\frac{1}{q^2}\right)\right)=\exp\left(O\left(\frac{1}{y}\right)\right),$$

which is $1+o(1)$ as $x\to\infty$. Summarizing all these observations, we get that

$$\begin{aligned}\frac{(s\circ s_\varphi)(n+i)}{n+i} &= (1+o(1))\left(\frac{\sigma(i)}{i}\,\frac{\sigma(Q_i)}{Q_i}-1\right)\left(1-\frac{\varphi(i)}{i}\,\frac{Q_i}{\sigma(Q_i)}\right)\\ &= (1+o(1))f_i(u_i)=(1+o(1))\alpha_i,\end{aligned}$$

as $x\to\infty$.

Now we deal with $s(n+i)$ given by formula (17). Here, much like in the case of $s_\varphi(n+i)$, we have

$$\begin{aligned}&\gcd(Q_iU_i(M_i\lambda+N_i),\sigma(i)((U_i+1)/R_i)\sigma(Q_i^2))=1;\\ &\gcd(U_i(M_i\lambda+N_i),\sigma(M_i\lambda+N_i))=1.\end{aligned}$$

Indeed, because the primes in $Q_iU_i(M_i\lambda+N_i)$ are all large, they do not divide $\sigma(i)$. Similarly, $M_i\lambda+N_i$ is coprime to $\sigma(Q_i^2)$. We have U_i coprime to $\sigma(Q_i^2)$ by (10). The fact that all prime factors in $\mathscr{Q}_i$ are congruent to 2 modulo 3, implies they cannot divide $\sigma(q^2) = q^2 + q + 1$ for any prime q, so Q_i is coprime to $\sigma(Q_i^2)$. Also, the fact that $\gcd(M_i\lambda + N_i, U_i + 1) = 1$ can be achieved by removing a set of values of λ similar to Λ_7 whose cardinality is $o(x/M)$ for $x \to \infty$. As for the second line above, this is certainly true if we exclude a set of λ similar to Λ_6, namely the set of $\lambda \in \Lambda_{10}$ such that for some $i \in \{1, \ldots, k\}$ we have $\gcd((M_i\lambda + N_i)U_i, \sigma(M_i\lambda + N_i)) > 1$, a set which by the arguments used to deal with Λ_6 can be proved to have cardinality $o(x/M)$ as $x \to \infty$. It then follows that

$$\begin{aligned} &\gcd(\sigma(i)((U_i+1)/R_i)(\sigma(M_i\lambda+N_i)/i)\sigma(Q_i^2), Q_iU_i(Q_i/R_i)(M_i\lambda+N_i)) \\ &\quad = \gcd(Q_i^2/R_i, \sigma(M_i\lambda+N_i)) =: Z_i, \end{aligned}$$

say. Writing

$$S_i = Z_i' S_i',$$

where $Z_i \mid Z_i'$, Z_i' is the largest divisor of S_i supported on the primes from Z_i, we have that S_i' is coprime to Z_i' and all its prime factors exceed $c_0 \log\log x / \log\log\log x$. This last condition holds since $\lambda \notin \Lambda_1$. Since $M_i\lambda + N_i$ is squarefree (using $\lambda \notin \Lambda_3$), we have $\omega(Z_i') \le \sum_{p \mid M_i\lambda+N_i} \omega(p+1)$. Eliminating a set of λ's similar to Λ_5, let's call it Λ_{11}, but for which there exists $i \in \{1, \ldots, k\}$ and a prime factor p of $M_i\lambda + N_i$ with $\omega(p+1) > 10\log\log x$, a set whose cardinality is $o(x/M)$ for $x \to \infty$, we get that Z_i' has $O((\log\log x)^2)$ distinct prime factors all of which exceed $y = (\log\log x)^4$, so

$$\frac{\varphi(Z_i')}{Z_i'} = 1 + O\left(\frac{1}{\log\log x}\right). \tag{19}$$

Finally, eliminating a subset of λ denoted Λ_{12} similar to Λ_9 and of cardinality $o(x/M)$ as $x \to \infty$, we can assume that $\omega(S_i') < 100\log\log x$. As in the previous case, this implies that

$$\frac{\varphi(S_i')}{S_i'} = 1 + o(1),$$

as $x \to \infty$. Thus, using (19),

$$\begin{aligned} \frac{(s_\varphi \circ s)(n+i)}{s(n+i)} &= 1 - \frac{\varphi(iR_iS_i)}{iR_iS_i} = 1 - \frac{\varphi(i)}{i}\frac{\varphi(R_iZ_i')}{R_iZ_i'}\frac{\varphi(S_i')}{S_i'} \\ &= (1+o(1))\left(1 - \frac{\varphi(i)}{i}\frac{\varphi(R_i)}{R_i}\right), \\ &= (1+o(1))\left(1 - \frac{\varphi(i)}{i}\frac{R_i}{\sigma(R_i)}\right) \end{aligned}$$

as $x \to \infty$, while

$$\begin{aligned}\frac{s(n+i)}{n+i} &= \frac{\sigma(i)}{i}\left(1+\frac{1}{U_i}\right)\frac{\sigma(M_i\lambda+N_i)}{M_i\lambda+N_i}\prod_{q\in\mathcal{Q}_i}\left(1+\frac{1}{q}+\frac{1}{q^2}\right)-1\\ &= (1+o(1))\left(\frac{\sigma(i)}{i}\frac{\sigma(Q_i)}{Q_i}-1\right).\end{aligned}$$

Thus,

$$\begin{aligned}\frac{(s_\varphi\circ s)(n+i)}{n+i} &= (1+o(1))\left(\frac{\sigma(i)}{i}\frac{\sigma(Q_i)}{Q_i}-1\right)\left(1-\frac{\varphi(i)}{i}\frac{R_i}{\sigma(R_i)}\right)\\ &= (1+o(1))\left(\frac{\sigma(i)}{i}u_i-1\right)\left(1-\frac{\varphi(i)}{iv_i}\right)\\ &= (1+o(1))\beta_i\end{aligned}$$

for $i\in\{1,\dots,k\}$. Since $n+i=(1+o(1))n$, we get that

$$(s\circ s_\varphi)(n+i)=(\alpha_i+o(1))n \quad \text{while} \quad (s_\varphi\circ s)(n)=(\beta_i+o(1))n$$

as $x\to\infty$. This certainly implies that inequalities (1) hold for $(f,g)=(s\circ s_\varphi, s_\varphi\circ s)$ and for our large n as $x\to\infty$, which finishes the proof of the theorem.

3 The proof of Theorem 2

For $n>1$, let $f(n)=s(n)/s_\varphi(n)$. Simple arguments show that for a prime p, we have $f(np)>f(n)$ (one considers the two cases: $p\mid n$, $p\nmid n$). Further $f(p)=1$. Thus, $f(n)\ge 1$ for all $n>1$.

Let $\alpha>1$ be arbitrary. Assume n is squarefree. Then

$$\sigma(n)\varphi(n)=n^2\prod_{p\mid n}\left(1-\frac{1}{p^2}\right),$$

so if n runs over any sequence of squarefree numbers with least prime tending to infinity, we have $\sigma(n)/n\sim n/\varphi(n)$. Since the reciprocal sum of the primes is divergent, it follows that there is a sequence of squarefree integers $n_1<n_2<\dots$ such that the least prime factor of n_i tends to infinity as $i\to\infty$, and at the same time, $\sigma(n_i)/n_i\to\alpha$. Then

$$f(n_i)=\frac{\sigma(n_i)/n_i-1}{1-\varphi(n_i)/n_i}\to\frac{\alpha-1}{1-1/\alpha}=\alpha, \quad \text{as } i\to\infty.$$

This proves the first assertion of the theorem.

We now turn to the second assertion in Theorem 2. If p,q are different primes, note that

$$\frac{(s\circ s_\varphi)(pq)}{(s_\varphi\circ s)(pq)}=\frac{s(p+q-1)}{s_\varphi(p+q+1)}.$$

Thus it suffices to show that the set of limit points of the rationals of the form $s(m-1)/s_\varphi(m+1)$, where m runs over those numbers that are the sum of two distinct primes, is $[0,\infty)$. If we knew the slightly stronger form of Goldbach's conjecture which asserts that all even numbers at least 8 are the sum of two distinct primes, we could assume that m runs over all even numbers at least 8. It turns out this slightly stronger form of Goldbach's conjecture is "almost" true.

Theorem 4 *There is a positive constant c such that if x is sufficiently large, the number of even numbers in $[1,x]$ which are not the sum of two distinct primes is at most x^{1-c}.*

This result is due to Montgomery and Vaughan [9], with later improvements due to Pintz and others, see [10].

Let $\alpha > 0$ be an arbitrary real number. Let x be large, let $m_1 = m_1(x)$ be the product of all of the odd primes to $y := (\log\log x)^{1/2}$, and let $m_2 = m_2(x) < z := e^{e^y}$ be an integer not divisible by any prime $p \le y$ and such that $\sigma(m_2)/m_2 \to \alpha+1$ as $x \to \infty$. Now let $m \le x$ run over even integers with

$$m \equiv -1 \pmod{m_1}, \quad m \equiv 1 \pmod{m_2}. \tag{20}$$

Note that the number of solutions m to x of this system is of magnitude $x/m_1m_2 > x/z^2$, which is huge compared with the exceptional set in Theorem 4. Thus, most of these numbers m are of the form $p+q$ where p,q are distinct primes. We now show that most of these m also satisfy $s(m-1)/s_\varphi(m+1) \to \alpha$ as $x \to \infty$.

It is clear that $\varphi(m+1) = o(m)$ as $x \to \infty$. Write $m-1 = m_2m_3 = m_2'm_3'$, where m_2' is the largest divisor of $m-1$ supported on the primes dividing m_2. Since the primes in $m-1$ all exceed y, it is clear that

$$\frac{\sigma(m_2')}{m_2'} = (1+o(1))\frac{\sigma(m_2)}{m_2}, \quad \text{as } x \to \infty.$$

Further, $1 \le \sigma(m_3')/m_3' \le \sigma(m_3)/m_3$. For m_1, m_2 fixed, $m_3 \le (x-1)/m_2$ runs through an arithmetic progression with modulus m_1. Let $g(m) = \sum_{p|m} 1/p$. Since no integer $\le x$ is divisible by two primes $p > \sqrt{x}$,

$$\sum_{m_3} g(m_3) \le \sum_{m_3} \frac{1}{\sqrt{x}} + \sum_{y<p\le\sqrt{x}} \frac{1}{p} \sum_{\substack{m_3 \le (x-1)/m_2 \\ p|m_3}} 1.$$

The inner sum is $\ll x/pm_1m_2$. Thus,

$$\sum_{m_3} g(m_3) \ll \frac{x}{m_1m_2y}.$$

We conclude that the number of choices for m_3 with $g(m_3) > 1/y^{1/2}$ is $o(x/m_1m_2)$ as $x \to \infty$. Hence there are $\gg x/m_1m_2$ choices of $m \le x$ where $g(m_3) \le 1/y^{1/2}$. Applying Theorem 4, we may also assume that these numbers m are the sum of two distinct primes. But

$$\frac{\sigma(m_3)}{m_3} \ll e^{g(m_3)},$$

so we may assume that $\sigma(m_3)/m_3 \to 1$ as $x \to \infty$. Putting the above observations together, we have for our numbers m that

$$\frac{s(m-1)}{s_\varphi(m+1)} = \frac{(1+o(1))(\alpha+1)m - m}{m - o(1)m} = (1+o(1))\alpha, \quad \text{as } x \to \infty.$$

This completes the proof of the second assertion in Theorem 2.

For the last assertion of Theorem 2, we again assume n is of the form pq where p,q are distinct primes. Then

$$\frac{(s \circ s)(n)}{(s_\varphi \circ s_\varphi)(n)} = \frac{s(m+1)}{s_\varphi(m-1)},$$

where $m = p+q$. By interchanging "-1" and "1" in the system (20), the above argument allows us to complete the proof of the theorem.

4 The proof of Theorem 3

For an integer $n > 20$, let $a(n)$ denote the largest divisor of n supported on the primes to $y(n) := \log\log n/\log\log\log n$. It follows from [7, Lemma 2.1] that on a set of asymptotic density 1 we have $a(n) = a(s(n)) = \gcd(n, s(n))$. Moreover, the same proof shows that on a set of asymptotic density 1, we have $a(n) = a(s_\varphi(n)) = \gcd(n, s_\varphi(n))$. Let

$$h(n) = \sum_{\substack{p \mid ns(n) \\ y(n) < p < (\log n)^2}} \frac{1}{p}.$$

We will show that there is a set $\mathscr{A}$ of asymptotic density 1 such that

$$H_{\mathscr{A}}(x) := \sum_{n \in (x,2x] \cap \mathscr{A}} h(n) = o(x) \quad \text{as } x \to \infty.$$

It will follow that there is a subset $\mathscr{A}'$ of $\mathscr{A}$ of asymptotic density 1 on which $h(n) = o(1)$ as $n \to \infty$.

Let $y = y(x)$, so that

$$H_{\mathscr{A}}(x) \le \sum_{y < p < (\log(2x))^2} \frac{1}{p} \sum_{\substack{n \in (x,2x] \cap \mathscr{A} \\ p \mid ns(n)}} 1.$$

The contribution to $H_{\mathscr{A}}(x)$ from the case $p \mid n$ is

$$\ll \sum_{y<p<(\log(2x))^2} \frac{x}{p^2} \ll \frac{x}{y} = o(x) \quad \text{as } x \to \infty.$$

Thus, we may concentrate on the case $p \mid s(n)$. Write $n = Pm$, where P is the largest prime factor of n. By a well-known result of de Bruijn, the number of $n \in (x, 2x]$ with $P \le z := x^{1/\log\log x}$ is $\ll x/\log x$, so we may assume that $\mathscr{A}$ captures the condition $P > z$. Fixing a value of $m \le 2x/z$ and a prime $p \in (y, (\log(2x))^2)$, we consider those primes $P \le x/m$ with $p \mid s(Pm)$. Discarding the case where $P^2 \mid n$ as negligible, we have

$$s(Pm) = Ps(m) + \sigma(m).$$

Since $a(n) = a(s(n)) = \gcd(n, s(n))$ may be assumed to hold for members of $\mathscr{A}$, we have $P \nmid \sigma(n)$, so in particular $P \nmid \sigma(m)$. Thus, having $p \mid s(Pm)$ puts P in a residue class mod p. So, ignoring the condition that P is prime, the number of choices for $P \le 2x/m$ is $\ll x/mp$. Hence

$$H_{\mathscr{A}}(x) \ll \sum_{y<p<(\log x)^2} \sum_{m \in (x/P, 2x/P]} \frac{x}{mp^2} \ll \sum_{y<p<(\log x)^2} \frac{x}{p^2} \ll \frac{x}{y},$$

which is $o(x)$ as $x \to \infty$.

By an analogous argument, the same holds if we change s to s_φ. Note also that for any $n \in (x, 2x]$, we have

$$\sum_{\substack{p \ge (\log n)^2 \\ p \mid ns(n)s_\varphi(n)}} \frac{1}{p} \ll \frac{1}{\log x} = o(1) \quad \text{as } x \to \infty.$$

Thus, there is a set $\mathscr{A}$ of asymptotic density 1 such that for $n \in \mathscr{A}$, we have $a(n) = a(s(n)) = a(s_\varphi(n))$ and

$$\sum_{\substack{p > y(n) \\ p \mid ns(n)s_\varphi(n)}} \frac{1}{p} = o(1) \quad \text{as } n \to \infty.$$

For each fixed $\varepsilon > 0$, let $\mathscr{A}_\varepsilon$ denote the subset of $\mathscr{A}$ consisting of those numbers n where $\varepsilon < s(n)/n < 1/\varepsilon$. By the continuity of the distribution function for $s(n)/n$, the density of $\mathscr{A} \setminus \mathscr{A}_\varepsilon$ tends to 0 as $\varepsilon \to 0$. On $\mathscr{A}_\varepsilon$ each of $s(n)/n$, $(s \circ s)(n)/s(n)$, and $(s \circ s_\varphi)(n)/s_\varphi(n)$ is asymptotically equal to $s(a(n))/a(n)$. And, each of $s_\varphi(n)/n$, $(s_\varphi \circ s_\varphi)(n)/s_\varphi(n)$, and $(s_\varphi \circ s)(n)/s(n)$ is asymptotically equal to $s_\varphi(a(n))/a(n)$. The various assertions in Theorem 3 now follow.

Acknowledgments

This paper was written when F. L. visited Dartmouth College in Summer 2015. He thanks the Mathematics Department there for its hospitality and support. We thank Paul Pollack for suggesting the relevance of [13]. We also thank the referee for a very careful reading.

References

1. A. Balog and C. Pomerance, "The distribution of smooth numbers in arithmetic progressions", *Proc. Amer. Math. Soc.* **115** (1992), 33–43.
2. N. Doyon and F. Luca, "On the local behavior of the Carmichael λ-function", *Michigan Math. J.* **54** (2006), 283–300.
3. P. Erdős, "On a problem of G. Golomb", *J. Austral. Math. Soc.* **2** 1961/1962, 1–8.
4. P. Erdős, A. Granville, C. Pomerance, and C. Spiro, "On the normal behaviour 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.
5. P. Erdős, K. Győry, and Z. Papp, "On some new properties of functions $\sigma(n)$, $\varphi(n)$, $d(n)$ and $\nu(n)$", *Mat. Lapok* **28** (1980), 125–131.
6. F. Luca and C. Pomerance, "On some problems of Makowski–Schinzel and Erdős concerning the arithmetical functions φ and σ", *Colloq. Math.* **92** (2002), 111–130.
7. F. Luca and C. Pomerance, "The range of the sum-of-proper-divisors function", *Acta Arith.* **168** (2015), 187–199.
8. M.-O. Hernane and F. Luca, "On the independence of $\sigma(\varphi(n))$ and $\varphi(\sigma(n))$", *Acta Arith.* **138** (2009), 337–346.
9. H. L. Montgomery and R. C. Vaughan, "The exceptional set in Goldbach's problem", *Acta Arith.* **27** (1975), 353–370.
10. J. Pintz, "Recent results on the Goldbach conjecture", in: *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 and C. Pomerance, "Some problems of Erdős on the sum-of-divisors function", *Trans. Amer. Math. Soc., Ser. B* **3** (2016), 1–26, http://dx.doi.org/10.1090/btran10.
12. C. Pomerance, "On the distribution of amicable numbers", *J. Reine Angew. Math.* **293/294** (1977), 217–222.
13. H. N. Shapiro, "Addition of functions in probabilistic number theory", *Comm. Pure Appl. Math.* **26** (1973), 55–84.
14. T. Xylouris, "Über die Nullstellen der Dirichletschen *L*-Funktionen und die kleinste Primzahl ineiner arithmetischen Progression", Doctoral dissertation, Universiteit Bonn, 2011.