# THE ALIQUOT CONSTANT, AFTER BOSMA AND KANE

CARL POMERANCE

Abstract. Let $s(n)$ be the sum of those positive divisors of the natural number $n$ other than $n$ itself. A conjecture of Catalan–Dickson is that the "aliquot" sequence of iterating $s$ starting at any $n$ terminates at 0 or enters a cycle. There is a "counter" conjecture of Guy–Selfridge that while Catalan–Dickson may be correct for most odd numbers $n$, for most even seeds, the aliquot sequence is unbounded. Lending some support for Catalan–Dickson, Bosma and Kane recently showed that the geometric mean of the numbers $s(n)/n$ for $n$ even tends to a constant smaller than 1. In this paper we reprove their result with a stronger error term and with a finer calculation of the asymptotic geometric mean. In addition we solve the analogous problems for certain subsets of the even numbers, such as the even squarefrees and the multiples of 4.

## 1. Introduction

Let $s(n) = \sigma(n) - n$ denote the sum of the "proper" divisors of the natural number $n$; that is, the sum of the positive divisors of $n$ that are smaller than $n$. The Catalan–Dickson conjecture asserts that every "aliquot" sequence $n, s(n), s(s(n)), \dots$ either terminates at 0 or enters into a cycle, so it is always bounded. The first $n$ in doubt is 276. The Guy–Selfridge counter conjecture, see [7], is that Catalan–Dickson is correct for asymptotically all odd numbers and false for asymptotically all even numbers.

On average the ratios $s(n)/n$ are $\zeta(2) - 1 = 0.6449\cdots < 1$, perhaps lending some credence to the Catalan–Dickson conjecture. However, restricted to odd numbers, the average ratio $s(n)/n$ is $\frac{3}{4}\zeta(2) - 1 = 0.2337\dots$ and restricted to even numbers it is $\frac{5}{4}\zeta(2) - 1 = 1.0561 > 1$, perhaps lending some credence to the Guy–Selfridge counter conjecture. (Since the function $s(n)$ usually, but not always, satisfies $s(n) \equiv n \pmod 2$, it seems appropriate to separate the problem by parity.)

In a recent paper [1], Bosma and Kane take the view that if one is to look at averages, the geometric mean is more appropriate than the arithmetic mean. Further they found that restricted to odd numbers (larger than 1) the geometric mean of the numbers $s(n)/n$ is $o(1)$ and restricted to even numbers it is $< 1$. Specifically, they prove the following result.

**Theorem 1.1.** *We have*

$$\frac{2}{x} \sum_{\substack{1<n\le x \\ n \text{ odd}}} \log(s(n)/n) = -2e^{-\gamma} \log\log x + O(\log\log\log x),$$

*where $\gamma$ is Euler's constant, and there is a constant $\lambda < -0.03$ such that*

$$\frac{2}{x} \sum_{\substack{n\le x \\ n \text{ even}}} \log(s(n)/n) = \lambda + O(1/\log x). \tag{1.1}$$

They call the negative number $\lambda$ the "aliquot constant" and they suggest that their theorem lends some support to the Catalan–Dickson conjecture.

In this note we strengthen Theorem 1.1 by computing $\lambda$ to higher precision and we provide a power-saving error estimate in (1.1).

**Theorem 1.2.** *To* 13 *decimal places we have* $\lambda = -0.03325\,94808\,010\ldots$, *and*

$$\frac{2}{x}\sum_{\substack{n\le x\\ n\text{ even}}}\log(s(n)/n) = \lambda + O(x^{-0.08317}).$$

In addition, we discuss the average geometric mean over certain subsets of the even numbers, such as the even squarefrees and those that are 0 (mod 4). The function $s(n)$ not only usually preserves parity, it also usually preserves the property of being squarefree, the property of being *not* squarefree, the property of being 0 (mod 4) and the property of being 2 (mod 4). So, investigating the average geometric mean over these subsets has some relevance. Finally we remark that our high order precision calculation of $\lambda$ is not accomplished through heroic computation, but rather through some standard ideas for accelerating the convergence of certain series.

It appears after numerical computations in [1] and [2] of the sum in (1.1) that the convergence to $\lambda$ is fairly rapid, perhaps of order $x^{-1+o(1)}$ as $x\to\infty$. It is clear that it cannot be $O(x^{-1})$ as individual terms can be of a slightly larger order. In [2] some numerical experiments are done for higher iterates of $s$. Let $s_k(n)$ denotes the $k$th iterate, when it exists. It is conjectured in [11] that averaged over those even numbers for which $s_k(n)>0$, we have $\log(s_k(n)/s_{k-1}(n))$ the same for every fixed $k$, namely the aliquot constant $\lambda$. This is proved for $k=2$ in [11]. In the numerical experiments in [2] convergence to $\lambda$ seems possible for higher values of $k$, but if so, it is quite slow.

## 2. Average value of a multiplicative function

The following result corrects an oversight and broadens [1, Lemma 3.12].

**Proposition 2.1.** *Let $f$ be a multiplicative function and suppose that $\kappa$ is a positive integer with $|f(p^m)-1|\le \kappa/p$ for every prime $p$ and positive integer $m$. For $x\ge \max\{e^\kappa, 20\}$,*

$$\sum_{n\le x} f(n) = x\prod_p\left(\left(1-\frac{1}{p}\right)\sum_{m\ge 0}\frac{f(p^m)}{p^m}\right) + O\Big(\frac{1}{(2\kappa)!}(\log x+2\kappa)^{2\kappa}x^{1.4/\log\log x}\Big),$$

*where the $O$-constant is absolute.*

(In [1, Lemma 3.12] it is assumed that each $f(p^m)$ is in $[0,1]$, and the error term is asserted to be $O((\log x)^C)$. Even if $C$ is allowed to depend on $\kappa$, the proof there does not seem to support such a small error estimate.)

*Proof.* Let $g$ be the multiplicative function which satisfies $g(p^m)=f(p^m)-1$ for each prime power $p^m$.

$$\sum_{n\le x} f(n) = \sum_{n\le x}\prod_{p^m\|n}(1+g(p^m)) = \sum_{n\le x}\sum_{d\|n} g(d) = \sum_{d\le x} g(d)\sum_{\substack{k\le x/d\\ \gcd(k,d)=1}} 1.$$

The inner sum here, via an inclusion-exclusion over the squarefree divisors of $d$, is $x\varphi(d)/d^2 + O(2^{\omega(d)})$, so that

$$\sum_{n\le x} f(n) = x\sum_{d\le x} g(d)\frac{\varphi(d)}{d^2} + O\left(\sum_{d\le x}|g(d)|2^{\omega(d)}\right), \tag{2.1}$$

where $\omega(d)$ is the number of different primes that divide $d$.

Letting $\mathrm{rad}(d)$ denote the largest squarefree divisor of $d$, note that

$$\sum_{d\le x}|g(d)|2^{\omega(d)} \le \sum_{d\le x}\frac{(2\kappa)^{\omega(d)}}{\mathrm{rad}(d)} = \sum_{n\le x}\frac{(2\kappa)^{\omega(n)}}{n}\sum_{\substack{d\le x\\ \mathrm{rad}(d)=n}} 1. \tag{2.2}$$

The inner sum here is $O(x^{1.4/\log\log x})$ for $x > e$. (A stronger result is implicit in the proof of [6, Theorem 11] and explicit in [10, Lemma 4.2].) Further, as is easy to show by an induction argument,

$$\sum_{n\le x}\frac{k^{\omega(n)}}{n} \le \frac{1}{k!}(\log x + k)^k \tag{2.3}$$

for all positive integers $k$ and for all $x \ge 1$. Thus, (2.2) implies that

$$\sum_{d\le x}|g(d)|2^{\omega(d)} = O\Big(\frac{1}{(2\kappa)!}(\log x + 2\kappa)^{2\kappa}x^{1.4/\log\log x}\Big). \tag{2.4}$$

It remains to consider the main term in (2.1). We will show that

$$\sum_{d>x}|g(d)|\frac{\varphi(d)}{d^2} = O\Big(\frac{1}{\kappa!}(\log x + \kappa)^{\kappa}x^{-1+1.4/\log\log x}\Big), \tag{2.5}$$

which implies that $\sum_d |g(d)|\varphi(d)/d^2$ converges. Assume this for now. Note that

$$1 + \Big(1-\frac{1}{p}\Big)\sum_{m\ge1}\frac{g(p^m)}{p^m} = 1 + \Big(1-\frac{1}{p}\Big)\sum_{m\ge1}\frac{f(p^m)-1}{p^m} = \Big(1-\frac{1}{p}\Big)\sum_{m\ge0}\frac{f(p^m)}{p^m},$$

so that

$$\sum_d g(d)\frac{\varphi(d)}{d^2} = \prod_p\Big(1 + \Big(1-\frac{1}{p}\Big)\sum_{m\ge1}\frac{g(p^m)}{p^m}\Big) = \prod_p\Big(\Big(1-\frac{1}{p}\Big)\sum_{m\ge0}\frac{f(p^m)}{p^m}\Big).$$

With (2.5) we thus have

$$x\sum_{d\le x} g(d)\frac{\varphi(d)}{d^2} = x\prod_p\Big(1-\frac{1}{p}\Big)\sum_{m\ge0}\frac{f(p^m)}{p^m} + O\Big(\frac{1}{\kappa!}(\log x+\kappa)^{\kappa}x^{1.4/\log\log x}\Big). \tag{2.6}$$

To show (2.5) first note that as in (2.2), for $t > e$,

$$\sum_{d\le t}|g(d)| \le \sum_{n\le t}\frac{\kappa^{\omega(n)}}{n}\sum_{\substack{d\le t\\ \mathrm{rad}(d)=n}} 1 \le \frac{1}{\kappa!}(\log t+\kappa)^{\kappa}t^{1.4/\log\log t},$$

using (2.3). Since the derivative of $-t^{-1+1.4/\log\log t}$ is $\gg t^{-2+1.4/\log\log t}$ for $t \ge 20$,

$$\begin{aligned}\sum_{d>x}\frac{|g(d)|}{d} &\le \int_x^\infty \frac{1}{\kappa!}(\log t+\kappa)^\kappa t^{-2+1.4/\log\log t}\,dt\\ &\ll \int_x^\infty \frac{1}{\kappa!}(\log t+\kappa)^\kappa\, d(-t^{-1+1.4/\log\log t})\\ &= \frac{1}{\kappa!}(\log x+\kappa)^\kappa x^{-1+1.4/\log\log x} + \int_x^\infty \frac{1}{(\kappa-1)!}(\log t+\kappa)^{\kappa-1}t^{-2+1.4/\log\log t}dt.\end{aligned}$$

Under the assumption that $t \ge x \ge \max\{e^\kappa, 20\}$,

$$\frac{1}{(\kappa-1)!}(\log t+\kappa)^{\kappa-1} \le \frac{1}{2\kappa!}(\log t+\kappa)^\kappa,$$

so that the prior calculation implies that

$$\sum_{d>x}\frac{|g(d)|}{d} \ll \frac{1}{\kappa!}(\log x+\kappa)^\kappa x^{-1+1.4/\log\log x},$$

which implies (2.5). Using (2.6) with (2.4) and (2.1) completes the proof. □

**Corollary 2.2.** *Suppose that $f, \kappa$ are as in Proposition 2.1 and denote by $E(x)$ the expression in the $O$-term. For any positive integer $A$ and for all $x$ with $x/A \ge \max\{e^\kappa, 20\}$ we have*

$$\sum_{\substack{n\le x\\ \gcd(n,A)=1}} f(n) = \frac{\varphi(A)}{A}x\prod_{p\nmid A}\left(\Big(1-\frac{1}{p}\Big)\sum_{m\ge 0}\frac{f(p^m)}{p^m}\right) + O\big(E(x)\big),$$

$$\sum_{\substack{n\le x\\ A\parallel n}} f(n) = \frac{\varphi(A)f(A)}{A^2}x\prod_{p\nmid A}\left(\Big(1-\frac{1}{p}\Big)\sum_{m\ge 0}\frac{f(p^m)}{p^m}\right) + O\big(|f(A)|E(x/A)\big),$$

*Proof.* Let $f_A(n)$ be the same as $f(n)$ except that it is 0 when $\gcd(n, A) > 1$. The first assertion follows directly from Proposition 2.1 applied to $f_A$. The second assertion follows directly from the first assertion. □

**Remark**. The estimate (2.3) in the proof of Proposition 2.1 is proved, as mentioned, by induction. However, the proof in mind is for the larger function $\tau_k(n)$ which is the number of ordered factorizations of $n$ into $k$ positive integral factors. Some difficulties are entailed working directly with $k^{\omega(n)}$, but if one is willing to push through them, as in Exercises 55 and 56 in [12], one can do better. This in turn would lead to a better error estimate in Theorem 1.2.

## 3. A geometric mean

In this section we compute the geometric mean of the numbers $\sigma(n)/n$. Let

$$\alpha = \sum_{p^m}\frac{\log(1+1/(\sigma(p^m)-1))}{p^m}, \tag{3.1}$$

where the summation is over all prime powers $p^m$ with $m \ge 1$.

**Proposition 3.1.** *For $x \geq 3$ we have*

$$\sum_{n \leq x} \log(\sigma(n)/n) = \alpha x + O(\log\log x).$$

*Proof.* Let $h(n) = \sigma(n)/n$. For prime powers $p^m$ with $m \geq 1$ let $\Lambda_\sigma(p^m) = \log(h(p^m)/h(p^{m-1}))$, and for all other positive integers, let $\Lambda_\sigma(n) = 0$. Note that

$$\Lambda_\sigma(p^m) = \log\left(1 + \frac{1}{\sigma(p^m) - 1}\right) \in \left(0, \frac{1}{p^m}\right) \quad \text{and} \quad \sum_{d|n} \Lambda_\sigma(d) = \log(h(n)).$$

We thus have

$$\sum_{n \leq x} \log(\sigma(n)/n) = \sum_{n \leq x} \sum_{d|n} \Lambda_\sigma(d) = \sum_{d \leq x} \Lambda_\sigma(d) \left\lfloor \frac{x}{d} \right\rfloor.$$

Using this identity we immediately have the upper bound

$$\sum_{n \leq x} \log(\sigma(n)/n) \leq x \sum_{d} \frac{\Lambda_\sigma(d)}{d} = \alpha x. \tag{3.2}$$

For a lower bound we have

$$\sum_{d \leq x} \Lambda_\sigma(d) \left\lfloor \frac{x}{d} \right\rfloor \geq \left(\alpha - \sum_{d > x} \frac{\Lambda_\sigma(d)}{d}\right) x - \sum_{d \leq x} \Lambda_\sigma(d).$$

The proposition follows upon noting that $\sum_{d>x} \Lambda_\sigma(d)/d \ll 1/x$ and $\sum_{d \leq x} \Lambda_\sigma(d) \ll \log\log x$. □

As with Corollary 2.2 we now have the following result.

**Corollary 3.2.** *Let $A$ be a positive integer. For $x \geq 3$ we have*

$$\sum_{\substack{n \leq x \\ \gcd(n,A)=1}} \log(\sigma(n)/n) = \alpha_1(A)x + O\big(2^{\omega(A)} \log\log x\big),$$

$$\sum_{\substack{n \leq x \\ A \parallel n}} \log(\sigma(n)/n) = \alpha_2(A)x + O\big(2^{\omega(A)}(\log\log x + \log(\sigma(A)/A))\big),$$

*where*

$$\alpha_1(A) = \frac{\varphi(A)}{A}\left(\alpha - \sum_{\gcd(d,A)>1} \frac{\Lambda_\sigma(d)}{d}\right), \quad \alpha_2(A) = \frac{\alpha_1(A)}{A} + \frac{\varphi(A)}{A^2} \log\left(\frac{\sigma(A)}{A}\right).$$

## 4. The geometric mean of $s(n)/n$

To compute the geometric mean of $s(n)/n$ we use the identity

$$\log(s(n)/n) = \log(\sigma(n)/n - 1) = \log(\sigma(n)/n) - \sum_{j \geq 1} \frac{1}{j}(n/\sigma(n))^j, \tag{4.1}$$

which holds for all $n > 1$. Let $f_j(n) = (n/\sigma(n))^j$. We have

$$1 > f_j(p^m) > \left(1 - \frac{1}{p}\right)^j \geq 1 - \frac{j}{p}$$

for all primes $p$ and positive integers $m, j$. We thus may apply Proposition 2.1 and Corollary 2.2 to $f_j$ with $\kappa = j$. Let

$$\beta_{p,j} = \Big(1 - \frac{1}{p}\Big) \sum_{m \ge 0} \frac{f_j(p^m)}{p^m}, \quad \beta'_{p,j} = \Big(1 - \frac{1}{p}\Big) \sum_{m \ge 1} \frac{f_j(p^m)}{p^m}, \tag{4.2}$$

and let

$$M_j = \prod_p \beta_{p,j} - \frac{1}{2} \prod_{p>2} \beta_{p,j} = \beta'_{2,j} \prod_{p>2} \beta_{p,j}. \tag{4.3}$$

It follows from Proposition 2.1 and Corollary 2.2 that

$$\sum_{\substack{n \le x \\ 2 \mid n}} f_j(n) = M_j x + O\Big(\frac{1}{(2j)!}(\log x + 2j)^{2j} x^{1.4/\log\log x}\Big) \tag{4.4}$$

for $x \ge \max\{40, 2e^j\}$.

In addition to (4.4), we have the trivial estimate

$$\sum_{\substack{n \le x \\ 2 \mid n}} f_j(n) \le \frac{1}{2}\Big(\frac{2}{3}\Big)^j x. \tag{4.5}$$

Indeed, if $n$ is even then $n/\sigma(n) \le \frac{2}{3}$ and so $f_j(n) \le (\frac{2}{3})^j$. Thus, (4.5) follows since the sum there has at most $\frac{1}{2}x$ terms. Dividing (4.4) by $x$ and letting $x \to \infty$, the estimate (4.5) implies that

$$M_j \le \frac{1}{2}\Big(\frac{2}{3}\Big)^j, \tag{4.6}$$

so that

$$\beta := \sum_j \frac{1}{j} M_j \tag{4.7}$$

converges. Let

$$\lambda = 2\alpha - 2\alpha_1(2) - 2\beta,$$

where $\alpha$ is defined in (3.1) and $\alpha_1(2)$ is defined in Corollary 3.2.

**Theorem 4.1.** *We have*

$$\frac{2}{x} \sum_{\substack{n \le x \\ 2 \mid n}} \log(s(n)/n) = \lambda + O(x^{-0.08317}).$$

*Proof.* Using (4.1), we have

$$\sum_{\substack{n \le x \\ 2 \mid n}} \log(s(n)/n) = A - B,$$

where

$$A = \sum_{\substack{n \le x \\ 2 \mid n}} \log(\sigma(n)/n) \quad \text{and} \quad B = \sum_{j \ge 1} \frac{1}{j} \sum_{\substack{n \le x \\ 2 \mid n}} f_j(n).$$

By Corollary 3.2, we have

$$A = (\alpha - \alpha_1(2))x + O(\log\log x). \tag{4.8}$$

Let $J$ be a large number to be determined shortly. By (4.4) we have

$$\sum_{j\le J}\frac{1}{j}\sum_{\substack{n\le x\\ 2\mid n}} f_j(n) = x\sum_{j\le J}\frac{M_j}{j} + O\left(\sum_{j\le J}\frac{1}{(2j)!}(\log x+2j)^{2j}x^{1.4/\log\log x}\right). \tag{4.9}$$

Further, from (4.5) and (4.6) we have

$$\sum_{j>J}\frac{1}{j}\sum_{\substack{n\le x\\ 2\mid n}} f_j(n) \ll \left(\frac{2}{3}\right)^J x, \quad \sum_{j>J}\frac{M_j}{j} \ll \left(\frac{2}{3}\right)^J.$$

Using this with (4.9) we have

$$B = \beta x + O\Bigg(\left(\frac{2}{3}\right)^J x + \sum_{j\le J}\frac{1}{(2j)!}(\log x+2j)^{2j}x^{1.4/\log\log x}\Bigg). \tag{4.10}$$

It remains to choose the optimal value of $J$, which is close to $0.205131\log x$. With this choice a calculation shows the error term in (4.10) is $O(x^{0.91683})$, which together with (4.8), proves the theorem. □

With a few simple changes, we have the following results.

**Corollary 4.2.** *With $\delta = 0.08317$, we have*

$$\frac{4}{x}\sum_{\substack{n\le x\\ 2\,\|\,n}}\log\left(\frac{s(n)}{n}\right) = \lambda_2 + O\big(x^{-\delta}\big) \quad \textit{and} \quad \frac{4}{x}\sum_{\substack{n\le x\\ 4\mid n}}\log\left(\frac{s(n)}{n}\right) = \lambda_4 + O\big(x^{-\delta}\big),$$

*where*

$$\lambda_2 = 4\alpha_2(2) - \sum_{j\ge 1}\frac{1}{j}\left(\frac{2}{3}\right)^j\prod_{p>2}\beta_{p,j}, \qquad \lambda_4 = 2\lambda - \lambda_2.$$

4.1. **The even squarefree case.** We can also obtain an estimation for

$$\sum_{\substack{n\le x\\ 2\mid n\\ \mu(n)^2=1}}\log\left(\frac{s(n)}{n}\right).$$

To this end it is straightforward by following the proof of Proposition 3.1 to obtain the following result:

$$\sum_{\substack{n\le x\\ \mu(n)^2=1}}\log\left(\frac{\sigma(n)}{n}\right) = \frac{6\alpha_0}{\pi^2}x + O\big(x^{1/2}\big),$$

where

$$\alpha_0 = \sum_p \frac{\log(1+1/p)}{p+1}.$$

Further, as in Corollary 3.2,

$$\sum_{\substack{n\le x\\ \mu(n)^2=1\\ 2\nmid n}}\log\left(\frac{\sigma(n)}{n}\right) = \frac{4}{\pi^2}\Big(\alpha_0 - \frac{\log(3/2)}{3}\Big)x + O\big(x^{1/2}\big),$$

so that

$$\sum_{\substack{n\le x\\ \mu(n)^2=1\\ 2\mid n}} \log\Big(\frac{\sigma(n)}{n}\Big) = \Big(\frac{2\alpha_0}{\pi^2}+\frac{4\log(3/2)}{3\pi^2}\Big)x + O\big(x^{1/2}\big). \tag{4.11}$$

For a multiplicative function $f$ satisfying the hypothesis of Proposition 2.1 it is possible to prove an analog of that result for $\mu(n)^2 f(n)$ (with an extra factor of $x^{1/2}$ in the error term), but we leave this for another time. For now we content ourselves with a non-uniform result which follows directly from Tenenbaum [12, Theorem I.3.12]:

$$\sum_{\substack{n\le x\\ \mu(n)^2=1\\ 2\mid n}} f_j(n) = \frac{x}{4} f_j(2)\prod_{p>2}\Big(\Big(1-\frac{1}{p}\Big)\Big(1+\frac{f_j(p)}{p}\Big)\Big) + o(x), \quad (x\to\infty).$$

We now use (4.1). Since the number of even squarefree numbers to $x$ is $\frac{2}{\pi^2}x + O(x^{1/2})$, we have the average value of $\log(s(n)/n)$ for even squarefree numbers to $x$ is $\sim \lambda_0 x$ as $x\to\infty$, where

$$\lambda_0 = \alpha_0 + \frac{2}{3}\log(3/2) - \frac{\pi^2}{8}\sum_{j\ge1}\frac{1}{j}\Big(\frac{2}{3}\Big)^j \prod_{p>2}\Big(\Big(1-\frac{1}{p}\Big)\Big(1+\frac{f_j(p)}{p}\Big)\Big).$$

## 5. Evaluation of constants

In this section we discuss the numerical evaluation of the various constants introduced, and in particular $\alpha$, defined in (3.1), and $\beta$, defined in (4.7). Calculations were done with Mathematica.

5.1. **The calculation of $\alpha$.** Let $N_0 = 15\,485\,863$ denote the one-millionth prime. Let

$$A_p = \frac{1}{p}\log\Big(1+\frac{1}{p}\Big) + \log\Big(1-\frac{1}{p^2}\Big).$$

A minor calculation shows that $-1/p^3 < A_p < -1/(2p^3)$ for all $p$. Note, as well, that

$$p^{-m}\log\Big(1+\frac{1}{\sigma(p^m)-1}\Big) < p^{-2m}.$$

Let

$$m_p = \lfloor \log N_0/\log p\rfloor.$$

We have

$$\begin{aligned}\alpha - \log\zeta(2) &= \sum_p A_p + \sum_{p,m\ge2}\frac{1}{p^m}\log\Big(1+\frac{1}{\sigma(p^m)-1}\Big)\\ &= \sum_{p\le N_0} A_p + \sum_{p\le N_0}\sum_{2\le m\le 2m_p}\frac{1}{p^m}\log\Big(1+\frac{1}{\sigma(p^m)-1}\Big) + E,\end{aligned} \tag{5.1}$$

where

$$-\sum_{p>N_0}\frac{1}{p^3} < E < -\sum_{p>N_0}\frac{1}{2p^3} + \sum_{p>N_0}\frac{1}{p^2(p^2-1)} + \sum_{p\le N_0}\frac{1}{p^{2m_p}(p^2-1)}.$$

It isn't difficult to get a reasonable estimate for the sum on the left. One can use what is known about the evaluation of the "prime zeta-function" (for example, see Glaisher [8]), or merely compute $\log\zeta(3)+\sum_{p\le N_0}\log(1-p^{-3})$. We find that

$$1.22\times 10^{-16} < \sum_{p>N_0} p^{-3} < 1.23\times 10^{-16}. \tag{5.2}$$

We have

$$\sum_{p>N_0}\frac{1}{p^2(p^2-1)} < \frac{1}{N_0^2}\sum_{p>N_0}\frac{1}{p^2-1} < \frac{1}{N_0^3} < 2.7\times 10^{-22},$$
$$\sum_{p\le N_0}\frac{1}{p^{2m_p}(p^2-1)} < 2.4\times 10^{-20},$$

so that $-1.23\times 10^{-16} < E < -6.0\times 10^{-17}$. With (5.1), we have

$$\alpha = 0.44570\,89175\,47339\,15 \pm 4\times 10^{-17}\,. \tag{5.3}$$

5.2. **The calculation of $\beta$.**

**Lemma 5.1.** *Using the notation of* (4.2)*, we have for each prime $p$ and positive integer $j$,*

$$1-\frac{j}{p^2} < \beta_{p,j} < 1-\frac{j}{p^2}+\frac{j^2}{(p-1)^3}.$$

*Proof.* For the lower bound note that

$$\beta_{p,j} > \Bigl(1-\frac{1}{p}\Bigr)\Bigl(1+\frac{1}{p-1}\Bigl(1-\frac{1}{p}\Bigr)^j\Bigr) > \Bigl(1-\frac{1}{p}\Bigr)\Bigl(1+\frac{1}{p-1}\Bigl(1-\frac{j}{p}\Bigr)\Bigr) = 1-\frac{j}{p^2}.$$

The upper bound trivially holds for $p\le j$ since $\beta_{p,j}<1$. So assume that $p\ge j+1$. Then

$$\begin{aligned}\beta_{p,j} &< \Bigl(1-\frac{1}{p}\Bigr)\Bigl(1+\frac{1}{p-1}\Bigl(1-\frac{1}{p+1}\Bigr)^j\Bigr)\\ &< \Bigl(1-\frac{1}{p}\Bigr)\Bigl(1+\frac{1}{p-1}\Bigl(1-\frac{j}{p+1}+\binom{j}{2}\frac{1}{(p+1)^2}\Bigr)\Bigr)\\ &= 1-\frac{j}{p(p+1)}+\binom{j}{2}\frac{1}{p(p+1)^2} < 1-\frac{j}{p^2}+\frac{j^2}{(p-1)^3},\end{aligned}$$

completing the proof. □

A simple calculation now verifies the following consequence.

**Corollary 5.2.** *For $p>j$ we have*

$$\Bigl|\log\Bigl(\Bigl(1-\frac{1}{p^2}\Bigr)^{-j}\beta_{p,j}\Bigr)\Bigr| < \frac{2j^2}{p^3}.$$

We use this for $p$ large. For $p$ somewhat smaller we use the following result. Recall that $f_j(n)=(n/\sigma(n))^j$.

**Lemma 5.3.** *For every prime $p$ and positive integers $j, m_1$, we have*

$$\frac{(1-1/p)^j}{p^{m_1}(p-1)}+\sum_{m=0}^{m_1}\frac{f_j(p^m)}{p^m} < \frac{\beta_{p,j}}{1-1/p} < \frac{f_j(p^{m_1+1})}{p^{m_1}(p-1)}+\sum_{m=0}^{m_1}\frac{f_j(p^m)}{p^m}.$$

Recalling the definition of $M_j$ in (4.3), we also have the following simple result.

**Lemma 5.4.** *For each positive integer $j$, we have $M_{j+1}/M_j < 2/3$.*

*Proof.* For every $p, j, m$ we have

$$\frac{f_{j+1}(p^m)}{f_j(p^m)} = \frac{p^m}{\sigma(p^m)} \le 1,$$

with the inequality strict for $m \ge 1$. Thus $\beta_{p,j+1} < \beta_{p,j}$. In the case $p = 2$ and $m \ge 1$, the ratio above is maximal when $m = 1$ and is $\frac{2}{3}$ there. Thus, $\beta'_{2,j+1} \le \frac{2}{3}\beta'_{2,j}$. This completes the proof. □

As before, $N_0$ is the one-millionth prime and $m_p = \lfloor \log N_0/\log p \rfloor$. For each prime $p \le N_0$, let

$$\beta^-_{p,j} = \Big(1 - \frac{1}{p}\Big)\Big(\frac{(1-1/p)^j}{p^{m_p}(p-1)} + \sum_{m=0}^{m_p} \frac{f_j(p^m)}{p^m}\Big),$$

$$\beta^+_{p,j} = \Big(1 - \frac{1}{p}\Big)\Big(\frac{f_j(p^{m_p+1})}{p^{m_p}(p-1)} + \sum_{m=0}^{m_p} \frac{f_j(p^m)}{p^m}\Big),$$

and similarly let $\beta'^{\pm}_{p,j}$ be the corresponding quantities where the sums start at $m = 1$. Let

$$M^{\pm}_{j,N_0} = \Big(\frac{3}{4}\zeta(2)\Big)^{-j} \beta'^{\pm}_{2,j} \prod_{2<p\le N_0} \beta^{\pm}_{p,j}\big(1 - p^{-2}\big)^{-j}.$$

We have from Lemma 5.3 that

$$M^-_{j,N_0} \prod_{p>N_0} \beta_{p,j}\big(1-p^{-2}\big)^{-j} < M_j < M^+_{j,N_0} \prod_{p>N_0} \beta_{p,j}\big(1-p^{-2}\big)^{-j}.$$

With Corollary 5.2, we thus have for $j < N_0$ that

$$M^-_{j,N_0} \exp\Big(-2j^2 \sum_{p>N_0} p^{-3}\Big) < M_j < M^+_{j,N_0} \exp\Big(2j^2 \sum_{p>N_0} p^{-3}\Big). \tag{5.4}$$

Using (5.2),

$$M^-_{j,N_0} \exp\big(-2.46 \times 10^{-16}\, j^2\big) < M_j < M^+_{j,N_0} \exp\big(2.46 \times 10^{-16}\, j^2\big). \tag{5.5}$$

We have computed that

$$2.14 \times 10^{-13} < M_{65} < 2.141 \times 10^{-13},$$

so by Lemma 5.4, we have that

$$3.29 \times 10^{-15} < \sum_{j\ge 65} \frac{1}{j} M_j < 9.65 \times 10^{-15}. \tag{5.6}$$

We have also computed that

$$\sum_{j=1}^{64} \frac{1}{j} M^-_{j,N_0} \exp\big(-2.46 \times 10^{-16}\, j^2\big) > 0.36578\,82599\,73748\,99\,,$$

$$\sum_{j=1}^{64} \frac{1}{j} M^+_{j,N_0} \exp\big(2.46 \times 10^{-16}\, j^2\big) < 0.36578\,82599\,73751\,53\,.$$

Thus, with (5.5) and (5.6), we have

$$\beta = \sum_j \frac{1}{j} M_j = 0.36578\,82599\,73757\,23 \pm 5 \times 10^{-15}\,. \tag{5.7}$$

5.3. **The calculation of $\lambda$.** We have

$$2\alpha - 2\alpha_1(2) = \alpha + \sum_{m\geq 1} \frac{\log(1 + 1/(2^{m+1} - 2))}{2^m}.$$

The infinite sum is easily computed to be $0.25260\,81215\,99091\,81384\ldots$, so that with (5.3) and (5.7),

$$\lambda = 2\alpha - 2\alpha_1(2) - 2\beta = -0.03325\,94808\,01084 \pm 1.1 \times 10^{-14}. \tag{5.8}$$

We have done similar calculations for $\lambda_0, \lambda_2, \lambda_4$. In particular, to 13 decimal places, we have

$$\begin{aligned} \lambda_0 &= -0.33843\,54384\,114\ldots, \\ \lambda_2 &= -0.24129\,50555\,350\ldots, \\ \lambda_4 &= +0.17477\,60939\,329\ldots. \end{aligned}$$

## 6. Discussion

According to Bosma and Kane [1], the fact that the average value of $\log(s(n)/n)$ for $n$ even is negative lends support to the Catalan–Dickson conjecture. The reasoning is as follows. One should think of the sequence $n, s(n), s(s(n)), \ldots$ as usually approximately geometric. Let $s_j(n)$ denote the $j$th iterate of $s$ at $n$. In fact, it is shown in Erdős [4] that for each fixed $k$ and $\epsilon > 0$, the ratio $s_{k+1}(n)/s_k(n)$ is at least $s(n)/n - \epsilon$ on a set of $n$ of asymptotic density 1. He also claims the analogous result that $s_{k+1}(n)/s_k(n)$ is at most $s(n)/n + \epsilon$ almost always, but this was later retracted in [5]. It is still thought though that this is the case, just the claim of proof is retracted. In [5] the assertion for $k = 1$ is proved.

But why do we restrict to even numbers? It is easy to see that $s(n) \equiv n \pmod 2$ except if $n$ is of the form $2^a b^2$, and of course such numbers are very sparsely distributed. So, as a simplifying assumption, perhaps it is reasonable to assume that at high levels, an aliquot sequence tends to maintain its parity. It is less obvious that it maintains other properties. For example, if $n$ is even, then $s(n) \equiv n \pmod 4$ except if the odd part of $n$ is of the form $pb^2$, where $p$ is a prime with $p \equiv 1 \pmod 4$. So, $s(n)$ maintains parity with "probability" about $1/\sqrt{n}$, but for even $n$ it maintains the residue mod 4 with "probability" about $1/\log n$.

Persistence of divisibility or lack thereof by other small primes is even weaker. For example, for an odd prime $p$, the "probability" that $p \nmid s(n)$ given that $p \mid n$ is about $1/(\log n)^{1/(p-1)}$, and the same holds for $p \mid s(n)$ given that $p \nmid n$. However, as noted in Guy–Selfridge [7], these events are not independent. For example, if $n \equiv 2 \pmod 4$, then $3 \mid \sigma(n)$, so $3 \mid n$ if and only if $3 \mid s(n)$. This gives rise to the idea of a "driver". If it should turn out for example that $n \equiv 6 \pmod{12}$ (and $n > 6$), then not only do we have $s(n) > n$, but it is highly likely that $s(n) \equiv 6 \pmod{12}$, and so on. What it would take to break this behavior is to arrive at a point where we pick up an extra factor of 2, which may occur with a chance about $1/\log n$, as discussed. So, the persistence of 3 is not a $1/(\log n)^{1/2}$ chance of breaking, but rather $1/\log n$. Note that if $n \equiv \pm 6 \pmod{36}$, then it is certain that $s(n) \equiv 2 \mod 4$ and so it is certain that $s(n) \equiv 6 \pmod{12}$, and so it is

certain that $6 \mid s_2(n)$. However, the chance of breaking being $\pm 6 \pmod{36}$ is about $1/(\log n)^{1/2}$.

One complication of this kind of thinking is that we don't know much about what is "normal" for $s_k(n)$, even for a fixed number $k \geq 2$. We do know that a positive proportion of even integers are not values of $s(n)$ (see Erdős [3]) and that a positive proportion of even integers are values of $s(n)$ (see Luca–Pomerance [9]), but we know very little about even numbers that are of the form $s_2(n)$.

While proofs can be, and often are, complicated, if they are to be useful and believable, heuristics should be simple. So perhaps the Bosma–Kane point of view is helpful. But consider the situation for those $n$ divisible by 4. If the chance for this to break is $1/\log n$ and the aliquot sequence is about geometric with ratio $e^{\lambda_4} = 1.1909\ldots$, by the time we move $\log n$ steps further, the numbers we are dealing with are larger by a factor of about $n^{\lambda_4}$, and so now the chance of breaking divisibility by 4 is somewhat less. It is perhaps not so convincing, but maybe most aliquot sequences that start at a multiple of 4 will diverge to infinity.

## Acknowledgments

I thank Anton Mosunov for his interest in this paper. In addition, he suggested that it was of interest to compute the average of $\log(s(n)/n)$ over even squarefree numbers. I am grateful to Mits Kobayashi for helping me with a Mathematica issue. I also thank the referee for a careful reading.

## References

1. W. Bosma and B. Kane, *The aliquot constant*, Quart. J. Math. **63** (2012), 309–323.
2. K. Chum, R. K. Guy, M. J. Jacobson, Jr., and A. S. Mosunov, *Numerical and statistical analysis of aliquot sequences*, preprint, 2017.
3. P. Erdős, *, Über die Zahlen der Form $\sigma(n) - n$ und $n - \varphi(n)$*, Elem. Math. **28** (1973), 83–86.
4. ———, *On asymptotic properties of aliquot sequences*, Math. Comp. **30** (1976), 641–645.
5. P. Erdős, A. Granville, C. Pomerance, and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions*, Analytic number theory (Allerton Park, IL, 1989), Progr. Math., vol. 85, Birkhäuser Boston, Boston, MA, 1990, pp. 165–204.
6. P. Erdős, F. Luca, and C. Pomerance, *On the proportion of integers coprime to a given integer*, Proceedings of the Anatomy of Integers Conference, Montreal, March 2006, J.-M. De Koninck, A. Granville, F. Luca, eds., CRM Proceedings and Lecture Notes, vol. 46 (2008), 47–64.
7. R. K. Guy and J. L. Selfridge, *What drives an aliquot sequence?*, Math. Comp. **29** (1975), 101–107.
8. J. W. L. Glaisher, *On the sums of inverse powers of the prime numbers*, Quart. J. Pure and Appl. Math. **25** (1891), 347–362.
9. F. Luca and C. Pomerance, *The range of the sum-of-proper-divisors function*, Acta Arith. **168** (2015), 187–199.
10. P. Pollack, *The greatest common divisor of a number and its sum of divisors*, Michigan Math. J. **60** (2011), 199–214.
11. C. Pomerance, *The first function and its iterates*, Connections in Discrete Mathematics: A Celebration of the Work of Ron Graham, S. Butler, J. Cooper, and G. Hurlbert, eds., Cambridge U. Press, to appear.
12. G. Tenenbaum, *Introduction to analytic and probabilistic number theory, third edition*, Graduate Studies in Math. **163**, American Math. Soc., Providence, 2015.

Department of Mathematics, Dartmouth College, Hanover, NH, 03755, USA
*E-mail address*: carl.pomerance@dartmouth.edu