A statistical investigation of a divisor-sum function

Ivan Aidun

April 6, 2019

Abstract The sum of proper divisors function $s(n)$ has been studied for more than 2000 years. In this thesis we study statistical properties of the related function $S_s(n) := \sum_{d|n} s(d)$. This function arises from a generalization of the practical numbers. We prove that $S_s(n)/n$ has an asymptotic distribution function, and that its values are dense in the interval $[0, \infty)$. We also establish mean value computations for $S_s(n)$ and $S_s(n)/n$, and provide uniform bounds for the higher order moments of $S_s(n)/n$.

Contents

1 Background: arithmetic functions

1.1 Arithmetic functions

We define an *arithmetic function* as a function $f : \mathbb{N} \to \mathbb{C}$. Although this definition is quite general, as Hardy and Wright [7, Chapter XVI, p. 302] note, in practice we want arithmetic functions to meaningfully encode some arithmetic property of their input. In this section we recall several natural examples of arithmetic functions, as well as basic operations on them.

Definition 1.1. We define the function φ by setting $\varphi(n)$ equal to the number of integers k in the interval $[1, n]$ such that $\gcd(k, n) = 1$. This function is known as *Euler's φ function.*

Definition 1.2. We define the function τ by setting $\tau(n)$ equal to the number of positive divisors of n. This function is known as the *number of divisors function*. (Sometimes also simply the *divisor function*, a name we avoid here to maintain clarity.)

Definition 1.3. We define the function σ by setting $\sigma(n)$ equal to the sum of all the positive divisors of n. This function is known as the *sum of divisors function*.

All three of these functions are multiplicative, i.e., they each satisfy $f(ab) = f(a)f(b)$ whenever $\gcd(a, b) = 1$. The functions τ and σ both represent examples of functions with representations as sums over the divisors of n. In those cases we can write

$$\tau(n) = \sum_{d|n} 1$$

and

$$\sigma(n) = \sum_{d|n} d.$$

The subscripts on these sums should be read as "sum over all (positive) d which divide n".

Sums of this form play an important role in the study of multiplicative functions for the following reason: whenever f and g are multiplicative, the function

$$h(n) := \sum_{d|n} f(d)g(n/d)$$

is also multiplicative. In this case, h is called the *Dirichlet convolution* of f and g, expressed symbolically as $h = f * g$. Using this notation we can rewrite $\tau = \mathbb{1} * \mathbb{1}$ and $\sigma = \mathrm{id} * \mathbb{1}$, where $\mathrm{id}(n) = n$ is the identity function and $\mathbb{1}(n) = 1$ is the function which is identically one.

The operation $*$ is commutative and associative, and distributes over pointwise addition. Moreover, the function

$$I(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise} \end{cases}$$

satisfies $I * f = f * I = f$ for all arithmetic functions f. Thus, the arithmetic functions form a commutative ring under pointwise addition and $*$.

A function that is often discussed in the context of the $*$ operator is the *Möbius μ function*, defined as

$$\mu(n) = \begin{cases} 1 & \text{for } n \text{ squarefree with an even \# of prime factors} \\ -1 & \text{for } n \text{ squarefree with an odd \# of prime factors} \\ 0 & \text{if } n \text{ has a square factor.} \end{cases}$$

This function satisfies $\mu * \mathbb{1} = I$. Thus, if the arithmetic functions f and g are related by $f = g * \mathbb{1}$ then $g = f * \mu$. This is known as the *Möbius inversion formula*.

1.2 The function $s(n)$

A very natural arithmetic function to consider is the *sum of proper divisors function* $s(n)$. That is, $s(n)$ is the sum over all positive divisors of n which are strictly less than

n. Using the sum notation from the previous section, we can write

$$s(n) = \sum_{\substack{d|n \\ d<n}} d.$$

Notice that this is the same as the definition of $\sigma(n)$, except that we do not add the value n itself. Thus, we may write $s(n) = \sigma(n) - n$. Notice that $s(n)$ is neither additive nor multiplicative.

The function $s(n)$ has an ancient history, having been considered by the Pythagoreans. Pomerance [15] goes so far as to call $s(n)$ "the first function". Some properties considered by the Pythagoreans include classifying integers by whether they satisfy $s(n) < n$, $s(n) > n$, or $s(n) = n$. Such integers are called *deficient*, *abundant*, or *perfect* numbers, respectively. It is natural to wonder how many of each of these numbers there are. There are known to be infinitely many abundant numbers; indeed, every multiple of 6 greater than 6 itself is abundant. It is not currently known if there are infinitely many perfect numbers. Euclid first showed that a number of the form $2^{p-1}(2^p - 1)$ is perfect if $2^p - 1$ is prime, and Euler showed that all even perfect numbers must have this form. No odd perfect numbers are known.

Because of the very restrictive form perfect numbers can take, it is not surprising that they are rare: there are only 3 perfect numbers less than 1000, and only 4 less than 100,000. A tool used to measure this disparity in how frequent these different kinds of numbers appear in the sequence of integers is the *asymptotic density*, defined in §2.1. Intuitively, if an integer n is chosen at random from a large interval $[1, N]$, the asymptotic density of a set $A \subset \mathbb{N}$ measures how likely it is that $n \in A$. Since every multiple of 6 is abundant, this chance is at least $1/6$ for the abundant numbers. However, for the perfect numbers it appears that this chance is next to zero. Indeed, it was first shown by Davenport [13] that the deficient, abundant, and perfect numbers all have asymptotic densities, and that the density of the perfect numbers is 0.

The Pythagoreans also considered when an integer n satisfies $s(s(n)) = n$. An example of this behavior is the number 220, which satisfies $s(s(220)) = s(284) = 220$. Notice that also $s(s(284)) = s(220) = 284$. In cases such as this, the pair $(n, s(n))$ is known as an *amicable pair*. Again, it is not known if there are infinitely many amicable pairs, but it is known that the amicable numbers have density 0. There are unresolved questions about further iterates of $s(n)$: the Catalan-Dickson conjecture asserts that the orbit of every integer under $s(n)$ remains bounded, while the competing Guy-Selfridge conjecture asserts that for "most" even n, the $s(n)$-iterates of n are unbounded.

Another open question involving the function $s(n)$ is a question of preimages: if A is a set of integers, then how large can $s^{-1}(A)$ be? A conjecture due to Erdős, Granville, Pomerance, and Spiro [6] asserts that if A has density 0 then $s^{-1}(A)$ also has density zero. Though the theorem is not proven in general, it has been proven in special cases (e.g., [12, 14, 29]).

1.3 The f-practical numbers

The function $s(n)$ is closely related to other questions about sums of divisors. The *practical numbers*, introduced by Srinivasan in [24], are positive integers n such that every number between 1 and n can be represented as a sum of distinct divisors of n.

For example, $n = 12$ is practical, since the divisors of 12 are 1, 2, 3, 4, and 6, and we can write $5 = 1 + 4$, $7 = 1 + 6$, $8 = 2 + 6$, $9 = 3 + 6$, $10 = 4 + 6$, and $11 = 1 + 4 + 6$.

Erdős [5] claimed to have a proof that the practical numbers have density 0, but gave no details. Complete criteria for a number to be practical were given by Stewart [25] and Sierpiński [23]. However, specific information about the distribution of practical numbers is more difficult to come by. Let $P(x)$ denote the number of practical numbers less than or equal to x. The first bound on $P(x)$ was given by Hausman and Shapiro [8] who asserted that $P(x) \leq x/(\log x)^{\beta+o(1)}$ (though their original proof was flawed). Tenenbaum [26] established the sharper result $P(x) \leq \frac{x}{\log x}(\log\log x)^{O(1)}$. Improving on this, Saias [19] showed that there exist absolute constants c_1, c_2 such that

$$c_1 \frac{x}{\log x} \leq P(x) \leq c_2 \frac{x}{\log x}.$$

The most recent progress in this direction was made by Weingartner [30], who showed that there exists a positive constant c such that $P(x) \sim cx/\log(x)$.

An analog of the practical numbers arises in relation to divisors of polynomials of the form $x^n - 1$. Recall that the cyclotomic polynomial $\Phi_k(x)$ is the monic degree $\varphi(k)$ polynomial with integer coefficients whose roots are exactly the primitive kth roots of unity. Since the roots of $x^n - 1$ are all the nth roots of unity, we have the factorization

$$x^n - 1 = \prod_{d|n} \Phi_d(x).$$

Notice that the degree of the right side is $\sum_{d|n} \varphi(d)$, which is equal to n. Thus, $x^n - 1$ has a divisor of every degree less than or equal to n if and only if every number between 1 and n can be written as a sum of $\varphi(d)$ for distinct divisors d of n. Such integers n are now known as φ-practical.

Let $P_\varphi(x)$ denote the number of φ-practical numbers less than or equal to x. There are no Stewart-like criteria for determining whether a number n is φ-practical; however, in [28], Thompson showed that there exist positive constants A, B such that

$$\frac{Ax}{\log x} \leq P_\varphi(x) \leq \frac{Bx}{\log x}.$$

Later, Pomerance, Thompson, and Weingartner [16] showed that there exists a constant C such that $P_\varphi(x) \sim Cx/\log x$.

Following this, Schwab and Thompson [22] generalized this construction to f-practical numbers for positive-integer-valued arithmetic functions f: a number n is f practical if every integer between 1 and $S_f(n) = \sum_{d|n} f(d)$ can be written as a sum of $f(d)$, for distinct divisors d of n. Notice that $S_f(n)$ is the largest number that could be written as a sum of $f(d)$ where the d are distinct, so it is the natural upper bound for the interval where we can expect this property to hold. The original practical numbers and the φ-practical numbers correspond to the f-practical numbers for $f = \mathrm{id}$ and $f = \varphi$, respectively.

1.4 Goals of this paper

In this paper we will prove several results about the function

$$S_s(n) := \sum_{d|n} s(n).$$

In the spirit of the classical work of Davenport [3] on $n/\sigma(n)$ and Schoenberg [20, 21] on $\varphi(n)/n$, it is natural to consider whether the function $S_s(n)/n$ possesses a distribution function. We prove the following result in §3.2 and 3.3.

Theorem 3.2. The function $S_s(n)/n$ has a continuous asymptotic distribution function.

Schoenberg also proved that the function $\varphi(n)/n$ has image dense in the interval $[0,1]$. Analogous to this result, we prove the following.

Theorem 3.1. The values $S_s(n)/n$ are dense in the interval $[0,\infty)$.

We also establish mean value computations for $S_s(n)$ and $S_s(n)/n$, and provide uniform bounds for the higher order moments of $S_s(n)/n$.

2 Background: probabilistic number theory

2.1 Asymptotic density

If we are given some natural number n, what are the chances that n is even? That n is divisible by 3? Prime? Perfect? These are examples of "probabilistic" questions we can ask about numbers, and some of them seem like they should have straightforward answers. It seems, for instance, that an arbitrary natural number should have a 1/2 chance of being even. Indeed, if we know that n is chosen uniformly at random in the interval $1 \leq n \leq N$, then we know exactly that

$$P(n \text{ is even}) = \begin{cases} 1/2 & \text{if } N \text{ is even} \\ 1/2 - 1/(2N) & \text{if } N \text{ is odd.} \end{cases}$$

If N is quite large, then regardless of its parity, the chance of the chosen n being even is quite close to 1/2. Similar reasoning should show that, as long as N is large, the chance that n is divisible by k should be $1/k$.

However, if we want to extrapolate from this a statement like "a natural number chosen uniformly at random has a $1/k$ chance of being divisible by k", we will encounter a problem. Although a uniform distribution on a finite number of natural numbers is well-defined, there does not seem to be such a notion for the entire set of natural numbers. Intuitively, it would seem that any such probability distribution would have to assign probability 0 to most if not all numbers. We can formalize this intuition, but first we need the following result, known as Mertens' Second Theorem [11, Theorem 3.15].

Theorem 2.1 (Mertens' Second Theorem)**.** There exists an absolute constant $C > 0$ for which

$$\prod_{p \leq x} (1 - 1/p) = e^{-C}/\log x + O(1/(\log x)^2).$$

With this tool under our belt, we can prove that there is no "nice" probability distribution on the natural numbers. Our proof here is modeled after [27, III.1 Theorem 1]. In what follows, let ξ be a random variable taking values in $\mathbb{N}$. We abbreviate the probability $P(\xi = n)$ by $P(n)$, and for a subset $S \subset \mathbb{N}$ we abbreviate $P(\xi \in S)$ by $P(S)$. Let $k\mathbb{N}$ represent the set of positive integer multiples of k.

Theorem 2.2. There does not exist a probability distribution on $\mathbb{N}$ satisfying $P(k\mathbb{N}) = 1/k$ for every $k \in \mathbb{N}$.

Proof. We suppose such a distribution exists, seeking a contradiction. First, we note that if p and q are distinct primes, then we know from elementary number theory that $pq \mid n$ if and only if $p \mid n$ and $q \mid n$. That is, $p\mathbb{N} \cap q\mathbb{N} = pq\mathbb{N}$. Thus, the events $\xi \in p\mathbb{N}$ and $\xi \in q\mathbb{N}$ are independent, since $P(p\mathbb{N} \cap q\mathbb{N}) = P(pq\mathbb{N}) = 1/pq = P(p\mathbb{N})P(q\mathbb{N})$. From this it follows that $\xi \in (\mathbb{N} \setminus p\mathbb{N})$ and $\xi \in (\mathbb{N} \setminus q\mathbb{N})$ are also independent events. For any fixed $n \in \mathbb{N}$, we know that n is certainly not in any set $p\mathbb{N}$ for prime $p > n$. So, for any $x > n$, n is an element of the intersection $S_x \coloneqq \bigcap_{n<p\leq x} \mathbb{N} \setminus p\mathbb{N}$. Hence, $P(n) \leq P(S_x)$. But, since each of the events $\xi \in \mathbb{N} \setminus p\mathbb{N}$ are independent, and since $P(\mathbb{N} \setminus p\mathbb{N}) = 1 - 1/p$ by hypothesis, we have

$$\begin{aligned} P(S_x) &= \prod_{n<p\leq x} P(\mathbb{N} \setminus p\mathbb{N}) \\ &= \prod_{n<p\leq x} (1 - 1/p). \qquad (1) \end{aligned}$$

Since there are only finitely many primes less than n, the product (1) is a constant multiple of the product appearing in Mertens' Second Theorem, so (1) is $O(1/\log x)$. In particular, as $x \to \infty$, the product (1) goes to 0, so $P(n) = 0$. However, since n was arbitrary, this shows that $P(n) = 0$ for all $n \in \mathbb{N}$, which is impossible. Thus, no such probability distribution exists. □

This theorem says that any truly "probabilistic" study of the integers cannot have this natural theorem that $P(k\mathbb{N}) = 1/k$. However, if we circumvent some of the requirements of a probability space, we can arrive at a notion which is similar enough to be useful. We noted already that in any interval $I = \{n \colon 1 \leq n \leq N\}$ we have a well-defined uniform probability distribution. With this distribution, if $A \subset \mathbb{N}$ then we can say precisely $P(\xi \in A) = \#(A \cap I)/N$. Instead of hoping for a probability distribution on all of $\mathbb{N}$, we can come close by instead taking the limit of these "partial probabilities," if it exists. This motivates the following definition.

Definition 2.3. We define the *asymptotic density* (also called the *natural density* or simply *density*) of a subset $A \subset \mathbb{N}$ to be

$$\mathbf{d}A = \lim_{N\to\infty} \frac{\#\{a \in A \colon a \leq N\}}{N},$$

when the limit exists.

We can verify quickly that under this definition, $\mathbf{d}\, k\mathbb{N} = 1/k$. We note that the number of multiples of k less than N is $\lfloor N/k \rfloor$, which satisfies $N/k - 1 \leq \lfloor N/k \rfloor \leq N/k$. Since

$$\lim_{N\to\infty} \frac{N/k - 1}{N} = \lim_{N\to\infty} \frac{N/k}{N} = 1/k,$$

we conclude that the density $\mathbf{d}\, k\mathbb{N} = 1/k$. In fact, it can be shown that an increasing sequence $a_1 < a_2 < a_3 < \dots$ has density α if and only if $\lim_{n\to\infty} n/a_n = \alpha$ [27, p. 271]. Below we list some other nice properties of the asymptotic density.

- All finite sets have asymptotic density 0.

- If $A, B \subset \mathbb{N}$ both have asymptotic densities, and $A \cap B = \emptyset$, then $\mathbf{d}(A \cup B) = \mathbf{d}A + \mathbf{d}B$.
- Let $A \Delta B = (A \cup B) \setminus (A \cap B)$ be the set symmetric difference of A and B. If A and B have asymptotic densities, and $\mathbf{d}A\Delta B = 0$, then $\mathbf{d}A = \mathbf{d}B$.

Not all sets have asymptotic densities. For example, let S be the set of all integers with an odd number of digits (when written base 10). Letting $S_N = \{s \in S \colon s \le N\}$, if $N = 10^{2m} - 1$ we have $\#S_N/N = 1/11 + O(1/100^m)$, while if $N = 10^{2m+1} - 1$ we have $\#S_N/N = 10/11 + O(1/100^m)$. Since these subsequences converge to different values, the density $\mathbf{d}S$ cannot exist. We can define the upper density of a set A as

$$\overline{\mathbf{d}}A = \limsup_{N \to \infty} \frac{\#\{a \in A \colon a \le N\}}{N}$$

and the lower density $\underline{\mathbf{d}}A$ similarly with the lim sup replaced by lim inf. Then for the set S above we have $\overline{\mathbf{d}}S = 10/11$ and $\underline{\mathbf{d}}S = 1/11$.

2.2 Mean values

Figure 2.2 shows a table of values of some arithmetic functions encountered in a first course in number theory, in the interval $2510 \le n \le 2530$. The values have been computed with Mathematica, and the interval has been somewhat arbitrarily chosen to be the ten numbers on either side of 2520—the least common multiple of the numbers 1 through 10.

We can see from this table that values of these arithmetic functions are quite sporadic. For example, τ takes on its largest value at 2520, and φ takes on its smallest value there—but at the prime 2521 τ is at a local minimum and φ a local maximum. It seems the values of $\tau(n)$ are perhaps hovering around 6, and the values of $\varphi(n)$ around 1500 (with considerable variation). Meanwhile, the values of μ seem to follow no discernible pattern.

Each of these functions is multiplicative, so we know about their values at particular integers n only insofar as we know about the prime factorization of n, which can vary wildly. It would be nice to be able to make some predictions knowing only the size of n, and not having to do the extra work of factoring it.

Here again, we are aided by treating our inputs as uniformly distributed random variables in some interval. If we do, then these arithmetic functions, in addition to looking like they take "random" values, now will take on values that are truly random. With this view, we can begin to ask certain statistical questions about our functions, and get to know them the way a statistician gets to know a data set. Perhaps the most natural first statistic to examine is the mean value, or average.

Definition 2.4. For an arithmetic function f, we define the *mean value of f over $n \le x$*, for x some positive real number, to be

$$M_x(f) = \frac{1}{x} \sum_{n \le x} f(n).$$

Furthermore, we define the *mean value of f* to be $M(f) = \lim_{x \to \infty} M_x(f)$ when the limit exists.

n	$\mu(n)$	$\tau(n)$	$\varphi(n)$
2510	-1	8	1000
2511	0	10	1620
2512	0	10	1248
2513	1	4	2148
2514	-1	8	836
2515	1	4	2008
2516	0	12	1152
2517	1	4	1676
2518	1	4	1258
2519	1	4	2280
2520	0	48	576
2521	-1	2	2520
2522	-1	8	1152
2523	0	6	1624
2524	0	6	1260
2525	0	6	2000
2526	-1	8	840
2527	0	6	2052
2528	0	12	1248
2529	0	6	1680
2530	1	16	880

Figure 1: Table of values for common arithmetic functions

We can apply this definition to compute the mean value of some of the arithmetic functions we have discussed. We present only the computation of $M_x(\sigma)$, as the computations for τ, φ, and S_σ are very similar. To perform this computation, we will need the following estimates [1, Theorem 3.2].

Lemma 2.5. For $x \geq 1$ we have:

1. $\displaystyle\sum_{n\leq x} \frac{1}{n} = \log x + C + O(1/x).$

2. $\displaystyle\sum_{n\leq x} \frac{1}{n^s} = \zeta(s) + \frac{x^{1-s}}{1-s} + O(x^{-s})$ for $s > 0, s \neq 1.$

Theorem 2.6. For $x \geq 1$ we have

$$\sum_{n\leq x} \sigma(n) = \frac{\zeta(2)}{2} x^2 + O(x \log x).$$

It follows that $M_x(\sigma) = \zeta(2)x/2 + O(\log x)$.

Proof. We write

$$\sum_{n\leq x} \sigma(n) = \sum_{n\leq x} \sum_{d|n} d.$$

If we have $d \mid n$, then by definition there is some q such that $n = dq$. Thus, we can rewrite the above sum as a sum over all pairs (d, q) satisfying $dq \leq x$:

$$\begin{aligned}\sum_{n\leq x}\sum_{d\mid n} d &= \sum_{n\leq x}\sum_{\substack{d\colon \exists q,\\ dq=n}} d \\ &= \sum_{\substack{d,q\\ dq\leq x}} d.\end{aligned}$$

We now interpret this sum by first choosing $q \leq x$, then summing over all the d satisfying $dq \leq x$. The d satisfying this are exactly the $d \leq x/q$. Thus, we write

$$\sum_{\substack{d,q\\ dq\leq x}} d = \sum_{q\leq x}\sum_{d\leq x/q} d.$$

Notice now that $\sum_{d\leq x/q} d = \lfloor x/q\rfloor(\lfloor x/q\rfloor + 1)/2 = x^2/2q^2 + O(x/q)$. Therefore, we write

$$\begin{aligned}\sum_{q\leq x}\sum_{d\leq x/q} d &= \sum_{q\leq x}\left(\frac{1}{2}\left(\frac{x}{q}\right)^2 + O(x/q)\right) \\ &= \frac{x^2}{2}\sum_{q\leq x}\frac{1}{q^2} + O\left(x\sum_{q\leq x}\frac{1}{q}\right) \\ &= \frac{x^2}{2}\left(\zeta(2) - \frac{1}{x} + O(1/x^2)\right) + O(x\log x),\end{aligned}$$

where the last equality follows from the estimates in Lemma 2.5. From here the result follows from a brief calculation. □

We can interpret Theorem 2.6 as saying if ξ is chosen randomly from $\{1, 2, \ldots, N\}$, then $\sigma(\xi)$ is on average $\zeta(2)N/2$.

Methods similar to those used in Theorem 2.6 can show that $M_x(\tau) = \log x + O(1)$ and $M_x(\varphi) = x/(2\zeta(2)) + O(\log x)$. The methods used above, coupled with the above result, can show that $M_x(S_\sigma) = \zeta(2)^2 x/2 + O((\log x)^2)$.

Mean values of functions are not merely tools to fuel intuition—the computation of mean values can imply deep facts about the arithmetic behavior of integers. For example, the prime number theorem is equivalent to the assertion that $M(\mu) = 0$ [1, Theorems 4.14, 4.15].

2.3 Distribution functions

In the previous two sections we have seen that we can recover definitions similar to those of probability theory by considering the interval $[1, n] \cap \mathbb{Z}$ equipped with a uniform probability distribution. We replace our notion of probability with the notion of density, and we replace our notion of expected value with the mean value introduced in the previous section. Now, we introduce an analogue to the distribution function.

In probability, given a real random variable X following some distribution, the distribution function F associated to the distribution is $F(x) = P(X \leq x)$. Any function arising this way will be non-decreasing and right-continuous (i.e., $\lim_{x\to x_0^+} F(x) =$

$F(x_0)$). Moreover, such a function will satisfy $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$. We take these properties to define a general distribution function.

Definition 2.7. A non-decreasing function F is a distribution function (d.f.) if F is right-continuous and satisfies $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

For our purposes, our "random variable" will be an arithmetic function f, and we will replace the probability $P(X \le x)$ with the density $\mathbf{d}\{f(n) \le x\}$. If the function f is well-behaved, then the function which appears will be a true distribution function according to the above definition.

Definition 2.8. Given an arithmetic function f, we define the sequence of functions

$$F_N(x) = \frac{\#\{n \le N : f(n) \le x\}}{N}.$$

We say f has asymptotic distribution function (a.d.f.) F if the functions F_N converge pointwise to a function F, and if F is a distribution function.

Example 2.9. Let A be a set of natural numbers with $\mathbf{d}A = \alpha$. Recall that the characteristic function χ_A is defined as

$$\chi_A(n) = \begin{cases} 1 & \text{if } n \in A \\ 0 & \text{otherwise.} \end{cases}$$

Let us examine the "partial" distribution functions F_N associated with χ_A. The functions $F_N(x)$ will all equal 0 for $x < 0$, since $\chi_A(n) \ge 0$. For $0 \le x < 1$, $F_N(x)$ is equal to the number of $n \le N$ for which $\chi_A(n) = 0$, i.e., the number of $n \le N$ satisfying $n \notin A$. This value converges to $1 - \alpha$ by the assumption that A has density α. Then, for $x \ge 1$, $F_N(x) = 1$ since $\chi_A(n) \le 1$. Thus, χ_A has an a.d.f. F given by

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1-\alpha & \text{if } 0 \le x < 1 \\ 1 & \text{if } x \ge 1. \end{cases}$$

The preceding example shows a simple case in which a question of interest (the asymptotic density of A) is translated into a question about the limiting distribution of an arithmetic function (in this case, χ_A). If an arithmetic function f has an a.d.f. F, then F contains much of the information about f. Thus, sufficiently precise knowledge of F can be used to answer questions about f. We give one more example of a way this translation can occur, again involving asymptotic density.

Example 2.10. Suppose an arithmetic function f has a continuous a.d.f. F. Then for any $x \in \mathbb{R}$, $\mathbf{d}\{n : f(n) = x\} = 0$. To see this, note that for any $\varepsilon > 0$ we have

$$\mathbf{d}\{n : f(n) = x\} \le \mathbf{d}\{n : x - \varepsilon < f(n) \le x + \varepsilon\} = F(x+\varepsilon) - F(x-\varepsilon).$$

Since F is continuous, as $\varepsilon \to 0$ the right-hand side goes to 0. Thus, $\mathbf{d}\{n : f(n) = x\} = 0$.

It will become important to us in Section 3 to know something of the kinds of analytic behavior a d.f. can display. First of all, because of the assumption of right-continuity, if a d.f. is discontinuous at a point x, then it must increase by a jump at x. A simple example of such a jumping function is the distribution

$$H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0. \end{cases}$$

If a d.f. only increases at its discontinuities—i.e., it is constant in any interval not containing a discontinuity—then it is said to be a *discrete* d.f.

For any d.f. F, since F is monotone, a theorem of real analysis says that F is differentiable except perhaps on a set of measure 0. Suppose F is continuous, with derivative h almost-everywhere. If h is Lebesgue-integrable and F satisfies

$$F(x) = \int_{-\infty}^{x} h(t)\, dt,$$

then F is said to be *absolutely continuous.* On the other hand, if $h(x) = 0$ almost-everywhere then F is said to be *purely singular.* The usefulness of these three classifications comes in the following theorem [27, III.2 Theorem 1].

Theorem 2.11 (Lebesgue decomposition theorem)**.** Any d.f. F may be written as $F = a_1F_1 + a_2F_2 + a_3F_3$, where $a_1 + a_2 + a_3 = 1$, and F_1, F_2, F_3 are d.f.s such that F_1 is absolutely continuous, F_2 is purely singular, and F_3 is discrete.

This theorem says that we can break every d.f. up into a sum of three distribution functions, each of whose behavior is more easily understandable.

2.4 Theorems of Erdős-Wintner and Schoenberg

Because of the utility of a.d.f.s, a rich theory has been established on the subject of when certain arithmetic functions have an a.d.f. A powerful theorem in this vein is the Erdős-Wintner Theorem, which completely answers the question of the existence of an a.d.f. in the case of additive arithmetic functions.

Theorem 2.12 (Erdős-Wintner, 1939)**.** Fix any real number $R > 0$. A real additive function $f(n)$ has a limiting distribution if and only if the following three series converge simultaneously:

$$\text{(i)} \sum_{|f(p)|>R} \frac{1}{p}; \qquad \text{(ii)} \sum_{|f(p)|\leq R} \frac{f(p)^2}{p}; \qquad \text{(iii)} \sum_{|f(p)|\leq R} \frac{f(p)}{p}.$$

In this case, all three sums converge for all $R > 0$. The limiting d.f. is either absolutely continuous, purely singular, or discrete. It is continuous if and only if

$$\sum_{f(p)\neq 0} \frac{1}{p} = \infty.$$

The Erdős-Wintner theorem gives insight not only into additive functions, but also multiplicative functions. If g is a strictly positive multiplicative function satisfying certain reasonable conditions [1], then g possesses an a.d.f. ψ if and only if the additive function $\log g$ possesses an a.d.f. ω. In this case, $\omega(x) = \psi(e^x)$.

Perhaps most surprising is that the Erdős-Wintner Theorem does not require considering $f(p^\alpha)$ for any $\alpha > 1$. In some sense this tells us that if an additive function f has an a.d.f., then for almost all n, the value of $f(n)$ is almost determined by its value on the squarefree part of n.

As an application of the Erdős-Wintner theorem, we show how it can be used to prove the classic theorem of Davenport [3] that $n/\sigma(n)$ has a continuous distribution function. The same kind of argument can be applied to the functions $\varphi(n)/n$ and $n/S_\sigma(n)$ to show that they, too, have a.d.f.s.

Corollary 2.13. The function $f(n) = n/\sigma(n)$ has a continuous asymptotic distribution function.

Proof. We will show that the additive function $\log f$ has a continuous a.d.f. by considering the sums of the Erdő-Wintner Theorem at $R = 1$. This will immediately show the existence of a continuous a.d.f. for f. From the Taylor expansion of $\log(1 + x)$ and of $(1 + x)\log(1 + x)$, we have $x/(1 + x) \leq \log(1 + x) \leq x$ for $x > -1$. Notice first that $\log f(p) = \log(1 - 1/(p + 1))$, and so $-1/p \leq \log f(p) \leq -1/(p + 1)$. Thus, $|\log f(p)| \leq 1/p < 1$. Therefore, the sum (i) is empty, while the sums (ii) and (iii) are over all positive primes. For these sums we have

$$\left|\frac{(\log f(p))^2}{p}\right| \leq \left|\frac{\log f(p)}{p}\right| \leq \frac{1}{p^2}.$$

Hence, both of the sums (ii) and (iii) converge absolutely, and so by the first part of the Erdős-Wintner Theorem, $\log f$ has an a.d.f. F.

Notice also that

$$\sum_{f(p)\neq 0} \frac{1}{p} = \sum_{p} \frac{1}{p},$$

which diverges. Therefore, by the second part of the Erdős-Wintner Theorem, F is continuous. □

This theorem has implications for the study of deficient, abundant, and practical numbers. In light of Example 2.10, the continuity of the distribution function F implies that $\mathbf{d}\{n\colon n/\sigma(n) = 1/2\} = 0$. The numbers for which $n/\sigma(n) = 1/2$ are exactly the numbers for which $n = s(n)$, i.e., the practical numbers. Similarly, the value of $F(1/2)$ is the density of the abundant numbers, and $1 - F(1/2)$ is the density of the deficient numbers.

Another, earlier, theorem guaranteeing the existence of a distribution function is the following, due to Schoenberg [21, Theorem 1].

[1] g cannot be almost everywhere almost zero, i.e., it cannot be the case that for all $\varepsilon > 0$, $\mathbf{d}\{n\colon g(n) > \varepsilon\} = 0$. An example of a function failing this condition is $f(n) = 1/n$. See [2, Theorem 4].

Theorem 2.14 (Schoenberg, 1928)**.** Let $f(n)$ be a multiplicative arithmetic function satisfying

i. $f(p^\alpha) > 0$

ii. the series

(1)
$$\sum_p \frac{1}{p} \|\log f(p)\|$$

converges, where $\|x\| := \min(1, |x|)$.

Then $f(n)$ has an asymptotic distribution function, $F(x)$. The function F is increasing on the closure of $\operatorname{im} f(n)$. Moreover, if there exists an increasing sequence of primes $q_1, q_2, \ldots$ such that $f(q_i) \neq f(q_j)$ whenever $i \neq j$ and such that

$$\sum_{i=1}^{\infty} \frac{1}{q_i}$$

diverges then $F(x)$ is everywhere continuous.

This theorem can also be applied to $\varphi(n)/n$, $n/\sigma(n)$, and $n/S_\sigma(n)$. However, in this paper we will be most concerned with the result about the points on which the d.f. F is increasing. Schoenberg showed that the values of $\varphi(n)/n$ are dense in $[0, 1]$ [20], and the same argument extends to $n/\sigma(n)$ and $n/S_\sigma(n)$. The arguments are very similar, so we will present only the one for $n/\sigma(n)$. We rely on the following lemma.

Lemma 2.15. Let $(a_i)_i$ be a decreasing sequence of positive real numbers satisfying $\lim_{i\to\infty} a_i = 0$ and $\sum_{i=1}^{\infty} a_i = \infty$. Then for any $L > 0$, there is a subsequence $(a_{i(k)})_k$ satisfying $\sum_{k=1}^{\infty} a_{i(k)} = L$.

Proof. Let $i(1)$ be the smallest index i such that $a_{i(1)} \leq L$; such an index exists since $a_i \to 0$. If we have equality, we are done. Recursively define

$$s_n = \sum_{k=1}^{n} a_{i(k)}$$

and let $i(k+1)$ be the smallest index $i > i(k)$ such that $a_{i(k+1)} \leq L - s_k$. If at any point we achieve equality, we are done. Now, the sequence of partial sums s_n is increasing and bounded above by L, and so approaches some limit. Moreover, since the sum of the a_i diverges, for every $N > 0$, we cannot have $i(k+1) = i(k) + 1$ for all $k \geq N$, or else $s_n = \sum_{k=1}^{n} a_{i(k)}$ would also diverge. Thus, infinitely often we have $i(k+1) > i(k) + 1$, which by the way we chose $i(k+1)$ means that $L - s_k < a_{i(k)+1}$. Since $i(k) + 1$ goes to ∞ with k, $a_{i(k)+1}$ goes to 0 with k, and so s_n converges to L. □

Theorem 2.16 (Schoenberg)**.** The values $n/\sigma(n)$ are dense in $[0, 1]$.

Proof. It suffices to show that the values $\sigma(n)/n$ for n squarefree are dense in $[1, \infty)$. Note that if n is squarefree, then we can write

$$\sigma(n)/n = \prod_{i=1}^{r} \frac{p_i + 1}{p_i},$$

where $p_1, \ldots, p_r$ are the distinct prime factors of n. Taking logarithms we find

$$\log \sigma(n)/n = \sum_{i=1}^{r} \log \frac{p_i + 1}{p_i}.$$

Thus, it suffices to show that sums of this form are dense in $[0, \infty)$. Now, by the same logarithm inequality as before we have $\log((p_i + 1)/p_i) = \log(1 + 1/p_i) \geq 1/(p_i + 1) > 1/2p_i$. Thus, the series

$$\sum_{p} \log \frac{p + 1}{p}$$

diverges. Note that as p grows, $\log((p + 1)/p)$ goes to 0, so this sequence satisfies the conditions of the preceding lemma. Thus, for every $L > 0$ there is a sequence of sums $\sum_{i=1}^{r} \log(p_i + 1)/p$ converging to L, i.e., such sums are dense in $[0, \infty)$, which is what we had to show. □

Since $s(n)/n = \sigma(n)/n - 1$, it follows from the above that the values of $s(n)/n$ are dense in $[0, \infty)$. The above result also provides us with the following corollary.

Corollary 2.17. Let F be the a.d.f. associated with $n/\sigma(n)$. Then F is increasing on exactly the set $[0, 1]$.

Proof. By Theorem 2.14, F is increasing on all the points $\{n/\sigma(n)\}$, as well as on all limit points of this set. Note that for all n, $n/\sigma(n) \in [0, 1]$. By Theorem 2.16, every point of $[0, 1]$ is a limit point of $\{n/\sigma(n)\}$, and so F is increasing on $[0, 1]$. □

Since $S_s(n)/n$ is not multiplicative, we cannot apply either the Erdős-Wintner Theorem nor Theorem 2.14 to yield an a.d.f. the way we can for the related functions $\sigma(n)/n$ and $S_\sigma(n)/n$. Moreover, for the function $f(n) = \log(S_s(n)/n)$, $f(p)$ is negative and unbounded, so there exists a prime p_0 so that $|f(p)| > R$ for all $p \geq p_0$. Thus, the sum (i) in the Erdős-Wintner theorem will diverge for this function. Similarly, for $g(n) = S_s(n)/n$, the sum (1) of Theorem 2.14 diverges. However, we will use the continuous distribution functions for $\sigma(n)/n$ and $S_\sigma(n)/n$ furnished by these theorems in the next section to show $S_s(n)/n$ has a continuous distribution function.

3 Results

3.1 $S_s(n)/n$ is dense in $\mathbb{R}^+$

In this section we will show that the values $S_s(n)/n$ are dense in $[0, \infty)$. Since the function $S_s(n)/n$ is not multiplicative, the same sort of argument as in Theorem 2.16 will not work. We are able to extract the fact that $s(n)/n$ is dense in $[0, \infty)$ by writing it in terms of the function $\sigma(n)/n$, so we might hope that there is a similar representation for $S_s(n)/n$. For example, we can write

$$\begin{aligned} S_s(n) &= \sum_{d|n} s(d) \\ &= \sum_{d|n} (\sigma(d) - d) \\ &= \sum_{d|n} \sigma(d) - \sum_{d|n} d \\ &= S_\sigma(n) - \sigma(n). \end{aligned}$$

Then $S_s(n)/n = S_\sigma(n)/n - \sigma(n)/n$. However, determining whether the values of $S_s(n)/n$ are dense in $[0, \infty)$ from this seems to require we are able to simultaneously control the growth of $S_\sigma(n)/n$ and $\sigma(n)/n$, which seems difficult.

To circumvent these problems, we introduce another relationship involving S_s: if a and b are relatively prime integers, then $S_s(ab) = S_s(a)S_s(b) + \sigma(a)S_s(b) + \sigma(b)S_s(a)$. To see this, we write

$$\begin{aligned} S_s(ab) &= S_\sigma(ab) - \sigma(ab) \\ &= S_\sigma(a)S_\sigma(b) - \sigma(a)\sigma(b) \\ &= (S_s(a) + \sigma(a))(S_s(b) + \sigma(b)) - \sigma(a)\sigma(b) \\ &= S_s(a)S_s(b) + \sigma(a)S_s(b) + \sigma(b)S_s(a). \end{aligned}$$

We will also make use of the following result, known as Bertrand's Postulate.

Theorem 3.1 (Bertrand's Postulate)**.** If $x > 1$ then there is a prime p in the interval $[x, 2x]$.

We now proceed with the result.

Theorem 3.2. The values $S_s(n)/n$ are dense in $[0, \infty)$.

Proof. Let $x \in [0, \infty)$, and index the primes in order $p_1, p_2, \ldots$. If $x = 0$, then the sequence $S_s(p_i)/p_i = 1/p_i$ converges to x. Otherwise, $x > 0$, and let p_k be the least prime so that $S_s(p_k)/p_k < x$. Notice that if q is a prime not dividing n, then

$$\begin{aligned} \frac{S_s(nq)}{nq} &= \frac{S_s(n)S_s(q)}{nq} + \frac{\sigma(n)S_s(q)}{nq} + \frac{\sigma(q)S_s(n)}{nq} \\ &= \frac{1}{q}\left(\frac{S_s(n)}{n} + \frac{\sigma(n)}{n}\right) + \frac{q+1}{q}\frac{S_s(n)}{n} \\ &= \frac{1}{q}\frac{S_\sigma(n)}{n} + \frac{q+1}{q}\frac{S_s(n)}{n} \\ &\geq \frac{q+1}{q}\frac{S_s(n)}{n}. \end{aligned}$$

For $n \geq k$ let $P_m = \prod_{i=k}^{m} p_i$. Since the product $\prod_i \frac{p_i+1}{p_i}$ diverges, there exists an $m \geq k$ so that $S_s(P_m)/P_m < x \leq S_s(P_{m+1})/P_{m+1}$.

Let $n_1 = P_m$. Beginning with the value $S_s(n_1)/n_1$ as an approximation for x, we will recursively construct better approximations for x. Let q_1 be any prime which does not divide n_1 and which satisfies

$$q_1 > \frac{S_\sigma(n_1)/n_1 + S_s(n_1)/n_1}{x - S_s(n_1)/n_1} \geq p_{m+1}.$$

Then we have

$$\begin{aligned}\frac{S_s(n_1q_1)}{n_1q_1} &= \frac{1}{q_1}\frac{S_\sigma(n_1)}{n_1} + \frac{q_1+1}{q_1}\frac{S_s(n_1)}{n_1} \\ &= \frac{S_s(n_1)}{n_1} + \frac{1}{q_1}\left(\frac{S_\sigma(n_1)}{n_1} + \frac{S_s(n_1)}{n_1}\right) \\ &< \frac{S_s(n_1)}{n_1} + \left(x - \frac{S_s(n_1)}{n_1}\right) = x.\end{aligned}$$

On the other hand, since $p_{m+1} > 1$, by Bertrand's Postulate we can choose q_1 so that

$$q_1 \leq 2\frac{S_\sigma(n_1)/n_1 + S_s(n_1)/n_1}{x - S_s(n_1)/n_1}.$$

Hence,

$$\frac{S_s(n_1q_1)}{n_1q_1} \geq \frac{x + S_s(n_1)/n_1}{2}.$$

We may now take $n_2 = n_1q_1$ and repeat the process, finding a q_2 so that

$$x > \frac{S_s(n_2q_2)}{n_2q_2} \geq \frac{x + S_s(n_1)/n_1}{2}.$$

Since $x - S_s(n_j)/n_j \leq (x - S_s(n_{j-1})/n_{j-1})/2$, the values $S_s(n_j)/n_j$ converge to x. Since x was arbitrary, this shows $\{S_s(n)/n\}$ is dense in $[0, \infty)$. □

3.2 Continuous distribution function: an analytic approach

Note: there is a mistake in this section: the distribution functions of $\sigma(n)/n$ and $S_\sigma(n)/n$ are both purely singular. The main result still holds, as discussed in the next section.

In this and the following section, we present two proofs that the function $S_s(n)/n$ has a continuous a.d.f. Both proofs will make use of the existence of distribution functions for the functions $\sigma(n)/n$ and $S_\sigma(n)/n$; however, they will vary in their approach. The proof presented in this section will make use of the analytic properties of the distribution functions for $\sigma(n)/n$ and $S_\sigma(n)/n$, as well as techniques for the manipulation of random variables found in probability theory. The proof presented in the next section, by contrast, will make use of results by Lebowitz-Lockard and Pollack from [9], which are established through explicitly arithmetic techniques.

We present both proofs as they represent a common difference of approach in probabilistic number theory. One approach uses the translation from arithmetic problems into probabilistic problems to bring to bear the tools of analysis, and then the desired arithmetic result is translated back from the probabilistic theorem. Another approach is to derive the result by purely arithmetic methods from the probabilistic formulation.

We now proceed with the main result of this section.

Theorem 3.3. The function $S_s(n)/n$ has a continuous a.d.f.

Proof. Recall from the preceding section that $S_s(n) = S_\sigma(n) - \sigma(n)$. Our method for showing that $S_s(n)/n$ has a continuous distribution function will then be to show that the distribution functions for $S_\sigma(n)/n$ and $\sigma(n)/n$ are absolutely continuous, and then we can use a standard result from probability theory to express the asymptotic distribution of $S_s(n)/n$.

By arguments similar to Corollary 2.13, both the functions $\sigma(n)/n$ and $S_\sigma(n)/n$ have continuous asymptotic distribution functions. Let these distribution functions be F and G, respectively, and note that both F and G are increasing on $[1, \infty)$. Now, both the additive functions $\log \sigma(n)/n$ and $\log S_\sigma(n)/n$ also have continuous distribution functions, given respectively by $F(e^x)$ and $G(e^x)$. By Theorem 2.12, these distribution functions are either absolutely continuous or purely singular. However, since these functions are increasing for all x, they cannot be singular, and so must be absolutely continuous.

Since these functions are absolutely continuous, they admit derivatives f, g almost-everywhere, and satisfy

$$F(x) = \int_1^x f(y)\ dy \qquad\qquad G(x) = \int_1^x g(y)\ dy.$$

We recall a standard theorem of probability theory (see for example [18, §6.3]) that if $Z = X + Y$ then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t) f_Y(z-t)\ dx.$$

Thus, there exists a continuous a.d.f. $H(x)$ for $S_s(n)/n$, and it satisfies

$$H(x) = \int_0^x \int_{-\infty}^{-1} g(t-y) f(-y)\ dy\ dt.$$

□

3.3 A discrete method

We now present another proof of Theorem 3.3. This proof will make use of concepts and theorems from [9]. If f is a real-valued arithmetic function, we say f *clusters around the real number* x if there exists a real number $d > 0$ such that for all $\varepsilon > 0$,

$$\overline{\mathbf{d}}\{n \colon x - \varepsilon < f(n) < x + \varepsilon\} \geq d.$$

If f does not cluster around any x, we say f is *nonclustering*. Recall that in Example 2.10 we showed that if f has a continuous a.d.f. F, then f is nonclustering. In fact, the converse holds.

Lemma 3.4. If the arithmetic function f has an a.d.f. F, and if f is nonclustering, then F is continuous.

Proof. Recall that F is the pointwise limit of the "partial" distribution functions F_N defined as

$$F_N(x) = \frac{\#\{n \leq N \colon f(n) \leq x\}}{N}.$$

Thus, we have

$$\begin{aligned}F(x+\varepsilon)-F(x-\varepsilon) &= \lim_{N\to\infty} F_N(x+\varepsilon)-F_N(x-\varepsilon)\\ &= \mathbf{d}\{n\colon x-\varepsilon<f(n)\le x+\varepsilon\}\\ &\le \overline{\mathbf{d}}\{n\colon x-\varepsilon<f(n)<x+\varepsilon\}.\end{aligned}$$

Thus, by the assumption that f is nonclustering, as $\varepsilon \to 0$, we have $F(x+\varepsilon)-F(x-\varepsilon) \to 0$. Therefore, F is continuous. □

We will use the following two theorems, which appear as Theorem 1 and Proposition 5 in [9], respectively.

Theorem 3.5 (Lebowitz-Lockard and Pollack)**.** Let $f_1, ..., f_k$ be multiplicative arithmetic functions taking values in the nonzero real numbers and satisfying the following conditions:

1. f_k does not cluster around 0
2. for all $i < j$ with $i, j \in \{1, 2, ..., k\}$, the function f_i/f_j is nonclustering.
3. for each i, whenever p and p' are distinct primes, we have $f_i(p) \neq f_i(p')$.

Then for all nonzero $c_1, ..., c_k \in \mathbb{R}$, the arithmetic function $F := c_1 f_1 + \cdots c_k f_k$ is nonclustering.

Theorem 3.6 (Lebowitz-Lockard and Pollack)**.** Let $f_1, \ldots, f_k$ be positive-valued multiplicative functions each possessing a distribution function. Then for any $c_1, \ldots, c_k \in \mathbb{R}$, the function $c_1 f_1 + \cdots + c_k f_k$ also has a distribution function.

Both of these theorems are proven by explicit estimation of upper densities by using the arithmetic properties of the functions f_i. This is in contrast to the previous section, in which the analytic properties of the distribution functions were used. We now proceed with another proof of Theorem 3.3.

Proof of Theorem 3.3. Recall that we can write

$$S_s(n) = \sum_{d|n} (\sigma(d) - d) = S_\sigma(n) - \sigma(n).$$

Thus,

$$\frac{S_s(n)}{n} = \frac{S_\sigma(n)}{n} - \frac{\sigma(n)}{n}$$

is a difference of two multiplicative functions.

Let $f_1 = S_\sigma(n)/n$, $f_2 = \sigma(n)/n$, and $F = f_1 + (-1)f_2$. We have previously stated that f_1 and f_2 have distribution functions, so by Theorem 3.6 above, F has an a.d.f. To show that the distribution function for F is continuous, by Lemma 3.4 it suffices to show that it satisfies the hypotheses of Theorem 3.5. We may apply Theorem 2.12 to the additive functions $\log f_1$, $\log f_2$, and $\log(f_1/f_2)$ shows that f_1, f_2 and f_1/f_2 have continuous a.d.f.s. Thus, conditions (1)-(3) of Theorem 3.5 are satisfied. Therefore, F is non-clustering. Since a distribution function for an arithmetic function F is continuous precisely when F is non-clustering, it follows that F is continuous. □

3.4 Mean values of $S_s(n)$ and $S_s(n)/n$

In this section we will compute the mean values $M_x(S_s(n))$ and $M(S_s(n)/n)$. These results will ground our discussion in the following section of uniform estimates for the moments of $S_s(n)/n$. To begin with, notice that the mean value defined in §2.2 is linear: if f and g are arithmetic functions and α, β are real numbers then

$$\begin{aligned} M_x(\alpha f + \beta g) &= \frac{1}{x}\sum_{n\le x} \alpha f(n) + \beta g(n) \\ &= \frac{\alpha}{x}\sum_{n\le x} f(n) + \frac{\beta}{x}\sum_{n\le x} g(n) \\ &= \alpha M_x(f) + \beta M_x(g). \end{aligned}$$

Thus, from our results from §2.2 we can immediately derive the following.

Theorem 3.7. We have

$$M_x(S_s(n)) = \frac{\zeta(2)(\zeta(2)-1)}{2}x + O((\log x)^2).$$

Proof. By linearity of M_x, we compute

$$\begin{aligned} M_x(S_s(n)) &= M_x(S_\sigma(n) - \sigma(n)) \\ &= M_x(S_\sigma(n)) - M_x(\sigma(n)) \\ &= \frac{\zeta(2)(\zeta(2)-1)}{2}x + O((\log x)^2). \end{aligned}$$

□

We will derive our computation of $M(S_s(n)/n)$ by applying *partial summation.* Partial summation is a discrete analogue of integration by parts. Given a sequence of integers (a_n) and a differentiable function f, let $A(x) = \sum_{n\le x} a_n$. Then partial summation states that

$$\sum_{n\le x} a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)\ dt.$$

Theorem 3.8. We have

$$M(S_s(n)/n) = \zeta(2)(\zeta(2)-1).$$

Proof. Consider the sum $\sum_{n\le x} S_s(n)/n$. Applying partial summation to this sum with $a_n = S_s(n)$ and $f(n) = 1/n$ we find

$$\begin{aligned} \sum_{n\le x}\frac{S_s(n)}{n} &= \frac{1}{x}\sum_{n\le x} S_s(n) + \int_1^x \frac{\sum_{n\le t} S_s(n)}{t^2}\ dt \\ &= M_x(S_s(n)) + \int_1^x \frac{M_t(S_s(n))}{t}\ dt \\ &= \frac{\zeta(2)(\zeta(2)-1)}{2}x + O((\log x)^2) + \int_1^x \left(\frac{\zeta(2)(\zeta(2)-1)}{2} + O((\log t)^2/t)\right) dt \\ &= \frac{\zeta(2)(\zeta(2)-1)}{2}x + O((\log x)^2) + \left(\frac{\zeta(2)(\zeta(2)-1)}{2}t\right)\Bigg|_1^x + O((\log t)^3\big|_1^x) \\ &= \zeta(2)(\zeta(2)-1)x + O((\log x)^3). \end{aligned}$$

Thus, $M(S_s(n)/n) = \lim_{x\to\infty}\frac{1}{x}\sum_{n\le x} S_s(n)/n = \zeta(2)(\zeta(2)-1)$. □

3.5 Estimates of the moments of $S_s(n)/n$

Analogous to the mean value, the *kth moment* of an arithmetic function f is defined to be

$$\lim_{x\to\infty} \frac{\sum_{n\le x} f(n)^k}{x},$$

when the limit exists. The computation of the moments of an arithmetic function can give a more complete set of statistics describing that function's behavior, and moment estimates have been computed for several arithmetic functions of interest.

In this section, we aim to estimate the moments of $S_s(n)/n$, i.e., the quantities

$$\mu_k = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (S_s(i)/i)^k.$$

We will make use of a powerful tool known as Wintner's Mean Value Theorem for multiplicative functions [17, Theorem 1, p. 138].

Theorem 3.9 (Wintner's Mean Value Theorem)**.** If g is a multiplicative function satisfying

$$\text{i. } \sum_p \frac{|g(p)-1|}{p} < \infty$$

$$\text{ii. } \sum_p \sum_{\nu=2}^{\infty} \frac{\left|g(p^\nu) - g(p^{\nu-1})\right|}{p^\nu} < \infty$$

then the mean value of g exists and is finite.

There are a few other facts we will make use of to establish our estimates for μ_k. We will use the following expressions for the functions σ and S_σ, derivations for which can be found in the Appendix:

$$\sigma(p^\nu) = p^\nu\left(1 + \frac{1}{p-1}\right) - \frac{1}{p-1}, \tag{1}$$

$$S_\sigma(p^\nu) = p^\nu\left(1 + \frac{1}{p-1}\right)^2 - \frac{\nu+1}{p-1} - \frac{p}{(p-1)^2}. \tag{2}$$

Additionally, let μ'_k be the kth moment of the function $n/\varphi(n)$. We will use the estimates for μ'_k appearing in the proof of [10, Proposition 4.3], in particular,

$$\log \mu'_k \ll k \log\log k.$$

We may now proceed with the result.

Theorem 3.10. The moments μ_k exist and are finite. Moreover, they satisfy

$$\log \mu_k \ll k \log\log k.$$

Proof. First, by the binomial formula, we get

$$\begin{aligned}(S_s(i)/i)^k &= \frac{(S_\sigma(i)-\sigma(i))^k}{i^k} \\ &= \frac{1}{i^k}\sum_{j=0}^{k}\binom{k}{j}(-1)^j(\sigma(i))^j(S_\sigma(i)^{k-j}).\end{aligned}$$

Each of the functions $h_{k,j}(i) = (\sigma(i))^j(S_\sigma(i))^{k-j}/i^k$ is multiplicative, and below we will use Wintner's Mean Value Theorem to show that each has finite mean. From the existence of mean values for the $h_{k,j}$, we conclude that the moments μ_k exist and are finite.

For sum i. in Wintner's Mean Value Theorem, note that using expression (2)

$$\begin{aligned}h_{k,j}(p)-1 &\le \left(\frac{S_\sigma(p)}{p}\right)^k - 1 \\ &< \left(1+\frac{1}{p-1}\right)^{2k} - 1 \\ &= \frac{p^{2k}-(p-1)^{2k}}{(p-1)^{2k}} \\ &= \frac{p^{2k}-(p^{2k}-2kp^{2k-1}+\text{terms of lower degree})}{(p-1)^{2k}} \\ &\ll \frac{p^{2k-1}}{(p-1)^{2k}} \\ &\ll \frac{1}{p}.\end{aligned}$$

Thus, for each $h_{k,j}$, the terms of sum i. are $O(1/p^2)$, and so the sum converges. For the double sum ii., we begin by using expressions (1) and (2) to estimate

$$\begin{aligned}h_{k,j}(p^\nu) &= \left(\frac{\sigma(p^\nu)}{p^\nu}\right)^j\left(\frac{S_\sigma(p^\nu)}{p^\nu}\right)^{k-j} \\ &= \left(\left(1+\frac{1}{p-1}\right)+O\left(\frac{1}{p^{\nu+1}}\right)\right)^j\left(\left(1+\frac{1}{p-1}\right)^2+O\left(\frac{\nu}{p^{\nu+1}}\right)\right)^{k-j} \\ &= \left(1+\frac{1}{p-1}\right)^{2k-j}+O\left(\frac{\nu}{p^{\nu+1}}\right).\end{aligned}$$

Thus, the numerator of the inner sum ii. is $O(\nu p^{-(\nu+1)})$. Therefore, the terms of the inner sum are $O(\nu p^{-(2\nu+1)})$. By taking the derivative of the geometric series $\sum_{\nu=2}^{\infty} p^{-(2\nu+2)}$ it can be verified that

$$\sum_{\nu=2}^{\infty}\frac{\nu}{p^{2\nu+1}} = \frac{2p^2-1}{(p^2-1)^2p^3}.$$

So, we conclude that the inner sum converges to a value that is $O(p^{-5})$. Therefore, the double sum converges. Having checked that the hypotheses of Wintner's Mean Value Theorem hold, we conclude that each $h_{k,j}$ has a finite mean value.

By our computations above, $S_s(n)/n \le S_\sigma(n)/n \le (n/\varphi(n))^2$, so we can use the estimates for $n/\varphi(n)$ to deduce that

$$\begin{aligned}\log \mu_k &\le \log \mu'_{2k} \\ &\ll 2k \log\log 2k \\ &\ll k \log\log k,\end{aligned}$$

as desired. □

A consequence of Theorem 3.10 is yet another method of showing that $S_s(n)/n$ has a distribution function. By our computations above, we also have

$$\log \mu_{2k} \ll k \log\log k,$$

so there exists some index k_0 and constant A so that $\log \mu_{2k} \le Ak \log\log k$ for all $k \ge k_0$. Hence, for all $k \ge k_0$ we have

$$\begin{aligned}\mu_k &\le \exp(Ak \log\log k) \\ &= (\log k)^{Ak}.\end{aligned}$$

Therefore, for $k \ge k_0$,

$$\frac{\mu_{2k}^{1/2k}}{k} \le \frac{(\log k)^{A/2}}{k}.$$

Thus, the condition $\limsup_{k\to\infty} \mu_{2k}^{1/2k}/k < \infty$ needed to apply Theorem 3.3.12 from [4] is satisfied, and therefore $S_s(n)/n$ has an a.d.f. As in Section 3.3, the results of Lebowitz-Lockard and Pollack suffice to show this a.d.f. is continuous.

Appendix

In this section we will derive the expressions for $\sigma(n)/n$ and $S_\sigma(n)/n$ used in §3.5:

$$\sigma(p^\nu) = p^\nu\left(1 + \frac{1}{p-1}\right) - \frac{1}{p-1}, \tag{1}$$

$$S_\sigma(p^\nu) = p^\nu\left(1 + \frac{1}{p-1}\right)^2 - \frac{\nu+1}{p-1} - \frac{p}{(p-1)^2}. \tag{2}$$

We have

$$\begin{aligned}\sigma(p^\nu) &= \sum_{i=0}^{\nu} p^i \\ &= \frac{p^{\nu+1}-1}{p-1} \\ &= p^\nu \frac{p}{p-1} - \frac{1}{p-1} \\ &= p^\nu\left(1 + \frac{1}{p-1}\right) - \frac{1}{p-1}.\end{aligned}$$

For the second derivation we have

$$
\begin{aligned}
S_\sigma(p^\nu) &= \sum_{i=0}^{\nu} \sigma(p^i) \\
&= \sum_{i=0} \left(p^i \left(\frac{p}{p-1} \right) - \frac{1}{p-1} \right) \\
&= \frac{1}{p-1} \left(\sum_{i=0}^{\nu} p^{i+1} - \sum_{i=0}^{\nu} 1 \right) \\
&= \frac{1}{p-1} \left(\frac{p^{\nu+2} - p}{p-1} - (\nu+1) \right) \\
&= p^\nu \left(1 + \frac{1}{p-1} \right)^2 - \frac{\nu+1}{p-1} - \frac{p}{(p-1)^2}.
\end{aligned}
$$

Bibliography

[1] T. M. Apostol. *Introduction to Analytic Number Theory.* Springer, 1976.

[2] G. J. Babu. On the distributions of multiplicative functions. *Acta Arithmetica*, 36:331–340, 1980.

[3] H. Davenport. Über numeri abundantes. *Sitzungberichte Akad. Berlin*, pages 830–837, 1933.

[4] R. Durret. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, 4th edition, 2010.

[5] P. Erdős. On a diophantine equation. *Mat. Lapok*, 1:192–210, 1950.

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

[7] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, 6 edition, 2008.

[8] M. Hausman and H.N. Shapiro. On practical numbers. *Comm. Pure Appl. Math.*, 37(5):705–713, 1984.

[9] N. Lebowitz-Lockard and P. Pollack. Clustering of linear combinations of multiplicative functions. *Journal of Number Theory*, 180:660–672, 2017.

[10] G. Martin, P. Pollack, and E. Smith. Averages of the number of points on elliptic curves. *Algebra Number Theory*, 8:813–836, 2014.

[11] P. Pollack. *Not Always Buried Deep: A Second Course in Elementary Number Theory.* American Mathematical Society, 2009.

[12] P. Pollack. Some arithmetic properties of the sum of proper divisors and the sum of prime divisors. *J. Math.*, 58:125–147, 2014.

[13] P. Pollack and C. Pomerance. Paul Erdős and the rise of statistical thinking in elementary number theory. In L. Lovász, I. Ruzsa, and V. Sós, editors, *Erdős Centennial*, pages 512–523. János Bolyai Math. Soc. and Springer-Verlag, 2013.

[14] P. Pollack, C. Pomerance, and L. Thompson. Divisor-sum fibers. *Mathematika*, 64:330–342, 2018.

[15] C. Pomerance. The first function and its iterates. In *Connections in Discrete Mathematics: A Celebration of the Work of Ron Graham*, pages 125–138. Cambridge University Press, 2018.

[16] C. Pomerance, L. Thompson, and A. Weingartner. On integers n for which $x^n - 1$ has a divisor of every degree. *Acta Arithmetica*, 175(3):225–243, 2016.

[17] A. G. Postnikov. *Introduction to analytic number theory*, volume 68 of *Translations of Mathematical Monographs.* American Mathematical Society, 1988.

[18] S. Ross. *A First Course in Probability.* Pearson Prentice Hall, 8 edition, 2010.

[19] E. Saias. Entiers á diviseurs denses. *J. Number Theory*, 62(1):163–191, 1997.

[20] I. J. Schoenberg. Über die asymptotische verteilung reeller zahlen *mod 1. Mathematische Zeitschrift*, 28:171–200, 1928.

[21] I. J. Schoenberg. On asymptotic distributions of arithmetical functions. *Transactions of the American Mathematical Society*, 39(2):315–330, 1936.

[22] N. Schwab and L. Thompson. A generalization of the practical numbers. *International Journal of Number Theory*, 14(5):1487–1503, 2018.

[23] W. Sierpiński. Sur une propriété des nombres naturels. *Annali di Matematica Pura ed Applicata*, 39(1):69–74, 1955.

[24] A.K. Srinivasan. Current Science, 17, 179-180, 1948.

[25] B.M. Stewart. Sums of distinct divisors. *American Journal of Mathematics*, 76(4):779–785, 1954.

[26] G. Tenenbaum. Sur un probléme de crible et ses applications. *Ann. Sci.École Norm. Sup.*, 19(1):1–30, 1986.

[27] G. Tenenbaum. *Introduction to Analytic and Probabilistic Number Theory*, volume 163 of *Graduate Studies in Mathematics.* Cambridge University Press, 3rd edition, 2015.

[28] L. Thompson. Polynomials with divisors of every degree. *Journal of Number Theory*, 132:1038–1053, 2012.

[29] L. Troupe. On the number of prime factors of values of the sum-of-proper-divisors function. *J. Number Theory*, 150C:120–135, 2015.

[30] A. Weingartner. Practical numbers and the distribution of divisors. *The Quarterly Journal of Mathematics*, 66:743–758, 2015.