SHIFTS OF THE PRIME DIVISOR FUNCTION OF ALLADI AND ERDŐS

Snehal M. Shekatkar
Centre for Modeling and Simulation, S.P. Pune University, Pune, Maharashtra, India
snehal.shekatkar@cms.unipune.ac.in

Tian An Wong
University of British Columbia, Vancouver, British Columbia, Canada
wongtianan@math.ubc.ca

Received: 3/8/19, Revised: 10/20/19, Accepted: 1/4/20, Published: 1/8/20

Abstract

We introduce a variation on the prime divisor function $B(n)$ of Alladi and Erdős, a close relative of the sum of proper divisors function $s(n)$. After proving some basic properties regarding these functions, we study the dynamics of their iterates and discover behavior that is reminiscent of the aliquot sequences generated by $s(n)$. In particular, we show that for iterates of $B(n)$, no unbounded sequences occur, analogous to the Catalan-Dickson conjecture. We also discuss possible directions for further investigation, motivated by numerical observations.

1. Introduction

Let n be a positive integer with prime factorization $n = p_1^{r_1} \dots p_k^{r_k}$. Consider the sum of prime divisors function

$$B(n) = \sum_{i=1}^{k} r_i p_i$$

and the sum of distinct prime divisors function

$$\beta(n) = \sum_{i=1}^{k} p_i.$$

They can be viewed as variants of the sum of proper divisors function

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

which has been studied since Pythagoras. On the other hand, the arithmetic properties of $B(n)$ were studied by Alladi and Erdős [1], and those of $\beta(n)$ were studied by Hall [7] in the 1970s. These are both large additive functions in the sense that they have the same average order as the largest prime factor of n, which is expected by the well-known result of Hardy and Ramanujan that $\Omega(n)$ and $\omega(n)$ have the same average order of $\log\log n$, where $\Omega(n)$ and $\omega(n)$ denote the number of total and distinct prime divisors of n, respectively.

In this paper, we introduce variations of $\beta(n)$ and $B(n)$ by shifting the values of these functions at fixed points. Clearly $B(n) = \beta(n) = n$ if n is prime. We then define for a fixed positive integer a the function

$$B_a(n) = \begin{cases} n+a, & \text{if } n \text{ is prime,} \\ B(n), & \text{otherwise.} \end{cases}$$

We also define $\beta_a(n)$ analogously. These new functions are no longer additive, but still bear similarities to the original functions, as we shall see.

The main motivation for introducing these functions lies in their iterates, and the relation to classical conjectures on the behavior of $s(n)$. The aliquot sequence $n, s(n), s(s(n)), \ldots$ is known to stop at either primes, or cycles of length two (amicable pairs) or longer (sociable numbers). Iterating $B_a(n)$ and $\beta_a(n)$, we encounter similar phenomena. We find that for fixed a only a handful of cycles are observed, suggesting a variant of the Catalan-Dickson conjecture, which states that there do not exist unbounded aliquot sequences (or, alternatively, the Guy-Selfridge counter-conjecture [6] which gives certain candidate counterexamples). Let us define the iterates by $B_a^2(n) = B_a(B_a(n))$, and similarly for $\beta_a(n)$. It is obvious that for $a = 0$ the iterates $B_0^k(n)$ and $\beta_0^k(n)$ eventually reach a fixed point, but for general a there may a priori be cases in which the sequences are unbounded. In support of this, we note that the celebrated Green-Tao theorem guarantees that there exists integers n and a such that

$$n < B_a(n) < \cdots < B_a^k(n), \text{ and } n < \beta_a(n) < \cdots < \beta_a^k(n)$$

for any $k > 0$ [5]. This is the analogue of a result of Lenstra, later improved by Erdős [2], that for every $k > 0$ there is an integer n for which

$$n < s(n) < \cdots < s^k(n).$$

Nonetheless, we shall show that in fact no unbounded sequences occur, which is the principal result of this paper.

Theorem 1. *Let $a \in \mathbf{N}$. Then there exist finitely many cycles generated by iterates of $B_a(n)$, and for any $n \in \mathbf{N}$, $B_a(n)$ iterates to one of these cycles.*

Finally, in Section 3, we provide numerical data that lead us to pose questions regarding the behavior of $B_a(n)$.

2. Iterates

We consider the dynamics of iterates of $B_a(n)$ and $\beta_a(n)$, for a nonnegative integer a. While most of the results below concern $B_a(n)$, many of them can be carried over to $\beta_a(n)$ without difficulty, which we leave to the interested reader. Throughout we will also raise several problems regarding the iterates $B_a^k(n)$, which can also be posed for $\beta_a^k(n)$.

2.1. Cycles

We call n a *periodic point* if $B_a^k(n) = B_a(n)$ for some positive integer k, and eventually periodic if $B_a^l(n)$ is periodic for some positive integer l. Define a cycle to be the *orbit* of a periodic point n. We will sometimes refer to the fixed points $B_a(n) = n$ as trivial cycles.

Lemma 1. *For any $a \in \mathbf{N}$, we have that $B_a(n) = n$ if and only if $n = 4$.*

Proof. Primes are not fixed points by definition. So the only possible fixed points are those of $B(n)$ for n composite, which is $n = 4$, as we see that $B(n) < n$ for any composite $n > 4$. □

We then have the following result for $a = 1$.

Proposition 1. *$B_1^k(n)$ is eventually periodic for all $n \in \mathbf{N}$, with cycles $(1), (4)$ and $(5, 6)$.*

Proof. For small n, say $n \leq 6$, this can be checked directly, so we may assume that $n > 6$. For fixed n, define the *stopping time* for B_a to be the integer

$$t_a(n) = \inf\{k : B_a^k(n) < n\}.$$

We first claim that $B_1^k(n)$ has stopping time $t_1(n) = 2$ if n is prime, otherwise $t_1(n) = 1$. It suffices to check this for $n = p$. To do so, set $p + 1 = 2m$. Then we have that

$$B_1^2(p) = B_1(p + 1) = 2 + B_1(m) \leq 3 + m < p.$$

It follows from the claim that $B_1^k(n) < n$ for all $k > 1$, thus $B_1^{2k}(n)$ is strictly decreasing until it reaches the cycle $(5, 6)$. □

We observe that the proof above fails for $a > 1$, since $p + a$ may be prime if a is even, while the final inequality $2 + m < p$ is no longer guaranteed if a is odd. Nonetheless, we can ask the following question: do all primes above 3 occur in some 2-cycle for some a? We can answer this in the affirmative.

Proposition 2. *Every $p > 3$ occurs in a 2-cycle $(p, B_a(p))$ for some $a \in \mathbf{N}$.*

Proof. We have to produce an a such that $B_a(p+a) = p$. This forces $p+a$ to be composite, so it suffices to show that $B_0(p+a) = p$ for some a. Let n be a composite solution to $B_0(n) = p$. Then setting $a = n-p$, we have that $B_a : p \mapsto n \mapsto p$ as desired.

It remains to show that $B_0(n) = p$ always has a composite solution. Let q be the largest prime less than p. If $p-q$ is prime, then choose $n = q(p-q)$. If not, then let a be a prime dividing $p-q$, and write $p-q = ab$. Then choose $n = qa^b$, and the claim follows. □

Remark 2. Observe that if a was chosen to be the largest prime dividing $p-q$, then the construction will in fact produce the smallest composite solution n.

How many kinds of cycles appear for a fixed a? We find that up to $a \leq 200$ there are at most 4 distinct nontrivial cycles. For example, at $a = 39$ we have the following cycles:

$$(43, 82), (13, 52, 17, 56), (7, 46, 25, 10), (5, 44, 15, 8, 6).$$

Note that from this example, we see that different k-cycles can occur for fixed k and a. Also, $a = 21$ has distinct 5-cycles $(5, 26, 15, 8, 6)$ and $(7, 28, 11, 32, 10)$.

It is natural then to consider the following variant of the Catalan-Dickson conjecture for $s(n)$: Are the iterates $B_a^k(n)$ unbounded as k tends to infinity? Heuristically, the average order $B_a(n)$ is $(\pi^2 n)/(6\log n)$. We can express this as follows.

Lemma 2. *We have*

$$\sum_{n\leq x} B_a(n) \sim \frac{\pi^2 x^2}{12\log x},$$

and also with $B_a(n)$ replaced by $\beta_a(n)$.

Proof. This follows simply from the fact that

$$\sum_{n\leq x} B_a(n) = \sum_{n\leq x} B(n) + a\sum_{p\leq x} 1,$$

and similarly for $\beta_a(n)$, then applying the corollaries to [1, Theorem 1.1]. The additional sum over p is $a\pi(x)$, which gives a smaller error $ax/\log x$. □

This suggests that an unbounded sequence of iterates should not exist, in contrast to $s(n)/n$ which is slightly greater than 1 on average. We can now turn to the proof of the main theorem.

Proof of Theorem 1. The case $a = 1$ is covered by Proposition 1, so we may assume that $a > 1$. We shall first show that given n large enough, the smallest integer k such that $B_a^{k-1}(n)$ is composite, whence $B_a^k(n) < B_a^{k-1}(n)$, will in fact be such that $B^k(n) < n$. In other words, iterating the function B_a starting at n, its orbit eventually reaches an integer less than n. Certainly this is true if n is composite, since $B_a(n) < n$, so we may as well assume that n is equal to some prime p.

Now let s be the smallest prime that does not divide a, which is less than $2a$ by Bertrand's postulate. Then it follows that $a \bmod s$ is nonzero, hence relatively prime to it. On the other hand, at least one of $p + a, p + 2a, \ldots, p + sa$ is divisible by s as a consequence of the pigeonhole principle, and therefore composite. Thus, iterating B_a at p reaches a composite number after at most s iterates.

On the other hand, we claim that for composite n, we have

$$B_a(n) \leq 2 + \frac{n}{2}, \tag{2.1}$$

with equality when n is twice a prime. Since n is composite, it suffices to prove (2.1) for $a = 0$. Let $n = p^m q$ where p is the smallest prime factor of n and q coprime to p. If $q = 1$, then $B(p^m) = mp$ and we are done. On the other hand, if $q \geq 1$, then by additivity we have

$$\begin{aligned} B(n) &= mp + B(q) \\ &\leq mp + q \\ &\leq 2 + \Big(\frac{p^m}{2} - 1\Big)q + q \\ &= 2 + \frac{n}{2}, \end{aligned}$$

where the last inequality follows since

$$p^m \leq 2 + \Big(\frac{p^m}{2} - 1\Big)q$$

and $q > 1$. This proves the claim.

Now to prove the theorem, it suffices to show that for any shift a, there exists a constant $C(a)$ such that iterating B_a from any starting point n eventually enters into a cycle contained within the interval $[1, C(a)]$. Let us assume first that such a constant exists. Then given $n > C(a)$, it follows from (2.1) and $s < 2a$ that we can iterate B_a starting at n until $B_a^k(n) \leq C(a)$ for some k. Suppose now that $n \leq C(a)$. Then if n is composite, then $B_a(n) < C(a)$, while on the other hand if n is equal to a prime p, then we see that there exists some k such that

$$B_a(p + ka) \leq 2 + \frac{p + sa}{2} < C(a).$$

From this we see that it suffices to take $C(a) = 2a^2 + 10$ for the bound to hold. It follows then that once the orbit of n under B_a enters into the range $[1, C(a)]$, it will remain bounded within the same interval. □

3. Further Directions

In this section, we discuss some possible directions for further investigation, motivated by numerical observations and analogous properties of $s(n)$. We first note that there is a natural extension of $B_a(n)$ to $\mathbf{Z}_{\geq 0}$ by setting $B(1) = 1$ and $B(0) = 0$, and to negative integers by setting $B(-n) = B(n)$. We may also extend $\beta_a(n)$ in a similar manner. Thus iterating our functions can be viewed as studying dynamics on $\mathbf{Z}$ itself. It is also possible to extend to $\mathbf{Q}$ by defining $B(\frac{x}{y}) := B(x) - B(y)$ for reduced fractions $\frac{x}{y}$, analogous to the logarithm, though upon iterating once we return to $\mathbf{Z}$, and from there $\mathbf{Z}_{\geq 0}$; another possible way of producing more interesting extensions is by setting $B(-n) = -B(n)$, and $B(\frac{x}{y}) = B(x)/B(y)$.

3.1. Cycle Length

What are the lengths of cycles, and how do they depend on a and n? What are the stopping times $t_a(n)$? We have not studied this question in detail, but numerical experiments suggest that both are small relative to $s(n)$. For example, Table 1 below lists the distinct nontrivial cycles found for small a and checking n up to 10^6.

a	cycles	a	cycles
1	$(5, 6)$	11	$(5, 15, 8, 6)$
2	$(5, 7, 9, 6)$	12	$(5, 17, 29, 41, 53, 65, 18, 8, 6)$
3	$(5, 8, 6), (7, 10)$	13	$(5, 16, 8, 6)$
4	$(5, 9, 6)$	14	$(5, 19, 33, 14, 9, 6), (7, 21, 10)$
5	$(7, 12)$	15	$(5, 20, 9, 6), (19, 34)$
6	$(7, 13, 19, 25, 10)$	16	$(7, 23, 39, 16, 8, 6, 5, 21, 10)$
7	$(5, 12, 7, 14, 9, 6)$	17	$(7, 24, 9, 6, 5, 22, 13, 30, 10), (11, 28)$
8	$(5, 13, 21, 10, 7, 15, 8, 6)$	18	$(5, 23, 41, 59, 77, 18, 8, 6), (7, 25, 10)$
9	$(5, 15, 9, 6), (13, 22)$	19	$(5, 24, 9, 6)$
10	$(5, 15, 8, 6)$	20	$(5, 25, 10, 7, 27, 9, 6)$

Table 1: Nontrivial cycles for $n \leq 10^6$.

3.2. Sign Patterns

Any k-cycle $(n, B_a(n), \ldots, B_a^k(n))$ can be ordered so that n is the least term in the sequence, making n prime and $B_a^k(n)$ composite. We adopt the following notation: we will assign either $+$ or $-$ to denote in a cycle whether a number is prime or composite. For example, the cycle $(5, 7, 9, 6)$ in $a = 2$ has sign pattern $(+, +, -, -)$.

We can now pose the following question: what are the possible sign patterns allowed in a k-cycle? All non-trivial cycles of length $k > 2$ must have sign patterns of the form $(+,\dots,-)$. Do all combinations occur in between? For example, with $k=3$ we find that both combinations $(+,+,-)$ and $(+,-,-)$ to occur.

3.3. Prime-divisor Fibres

A question related to cycles is: For fixed a and p, what is the cardinality of the set of solutions to the equation $B_a(n)=p$? The solution sets are the same for every a, except possibly the pre-image $\{p-a\}$, which will be counted if it is prime. Note that the proof of Theorem 2 provides the smallest composite solution.

If a is prime, the solutions to $B_a(n)=m$ for a fixed m are given by the prime partitions $\kappa(m)$ of m, as was already observed in [8, Theoren 2.7], and there is at least one composite solution for all $n\geq 5$. From this fact we can immediately deduce from [4, VIII.26] the asymptotic

$$\log(\#\{n : B_a(n)=m\})\sim 2\pi\sqrt{\frac{m}{3\log m}}$$

as $m\to\infty$.

Indeed, one even has a recursive definition for $\kappa(n)$ in terms of $\beta(n)$,

$$\kappa(n)=\frac{1}{n}\big(\beta(n)+\sum_{i=1}^{n-1}\kappa(n-i)\beta(i)\big),$$

with the initial condition $\kappa(1)=0$. In other words, the number solutions of $B(n)=m$ are determined by the values of $\beta(i)$ for $i\leq m$. More generally, this question can be phrased in terms of prime-divisor sum fibres, with reference to [9]: Let $\mathcal{A}\subset\mathbf{N}$ be a set of asymptotic density zero (for example, the set of prime numbers). Erdős, Granville, Pomerance, and Spiro [3] then conjecture that the fibre $s^{-1}(\mathcal{A})$ has asymptotic density zero. What is the preimage $B_a^{-1}(\mathcal{A})$ for a fixed a?

Acknowledgments. The authors would like to thank Jeff Lagarias for suggesting the proof of Theorem 1, and the reviewer for helpful suggestions and remarks. SMS acknowledges the funding from the DST-INSPIRE Faculty Fellowship (DST/INSPIRE/04/2018/002664) by DST India.

References

[1] K. Alladi and P. Erdős, On an additive arithmetic function, *Pacific J. Math.* **71** (1977), 275–294.

[2] P. Erdős, On asymptotic properties of aliquot sequences, *Math. Comp.* **30** (1976), 641–645.

[3] P. Erdős, A. Granville, C. Pomerance, and C. Spiro, On the normal behavior of the iterates of some arithmetic functions, in *Analytic Number Theory*, Birkhäuser Boston, Boston, MA, 1990.

[4] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.

[5] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. of Math. (2)* **167** (2008), 481–547.

[6] R. K. Guy and J. L. Selfridge, What drives an aliquot sequence? *Math. Comput.* **29** (1975), 101–107.

[7] R. R. Hall. On the probability that n and $f(n)$ are relatively prime, *Acta Arith.* **17** (1970), 169–183.

[8] R. Jakimczuk, Sum of prime factors in the prime factorization of an integer, *Int. Math. Forum* **7** (2012), 2617–2621.

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