2016

Comparing the Growth of the Prime Numbers to the Natural Numbers

Michael A. Brilleslyper
U.S. Air Force Academy, mike.brilleslyper@usafa.edu

Nathan Wakefield
University of Nebraska-Lincoln, nathan.wakefield@unl.edu

A. J. Wallerstein
Air Force Institute of Technology, austin.wallerstein@afit.edu

Bradley Warner
U.S. Air Force Academy, Bradley.warner@usafa.edu

Published in Fibonacci Quarterly 54 (2016), no. 1, 65–71. https://www.fq.math.ca/

COMPARING THE GROWTH OF THE PRIME NUMBERS TO THE NATURAL NUMBERS

MICHAEL A. BRILLESLYPER, NATHAN WAKEFIELD, A. J. WALLERSTEIN, AND BRADLEY WARNER

Abstract. We define a new method of measuring the rate of divergence for an increasing positive sequence of integers. We introduce the growth function for such a sequence and its associated growth limit. We use these tools to study the divergence rate for the natural numbers, polynomial and exponential-type sequences, and the prime numbers. We conclude with a number of open questions concerning general properties and characterizations of growth functions and the set of possible growth limits.

1. Introduction

The On-Line Encyclopedia of Integer Sequences currently contains about 200,000 sequences in the database, many of which are strictly increasing sequences of positive integers. In this paper, we focus on increasing sequences of positive integers which we will simply refer to using the word sequence.

We propose a new method for describing the rate of divergence of a sequence and then develop some interesting properties of this description. We will conclude by describing the growth of the prime numbers and proving some asymptotic properties.

2. The Growth Function

We begin with two definitions that are fundamental to our paper.

Definition 2.1. *Let $a_1, a_2, a_3, \ldots$ be a sequence, we generalize the factorial function by defining the product*

$$P(n) := \prod_{j=1}^{n} a_j.$$

Observe, if $a_n = n$, then $P(n) = n!$.

Definition 2.2. *Let $a_1, a_2, a_3, \ldots$ be a sequence, we define the* growth function $f(n)$ *determined by the sequence to be the largest (non-negative) integer such that*

$$Q(n) := \prod_{j=n+1}^{n+f(n)} a_j \leq P(n).$$

The output of the growth function is the number of terms in the product $Q(n)$. The growth function will give us an interesting tool for comparing different sequences. In the case of the natural numbers, $f(n)$ is equivalent to

$$f(n) := \max\{k \in \mathbb{N} : (n!)^2 \geq (n+k)!\}.$$

Table 1 shows the first 8 values for the growth function of the natural numbers.

Our first result studies the behavior of the growth function for an arbitrary sequence.

TABLE 1. Growth function for $a_k = k$.

n	1	2	3	4	5	6	7	8
$f(n)$	0	0	1	1	2	3	3	4

Theorem 2.3. *Let $a = \{a_k\}_{k=1}^{\infty}$ be a sequence with growth function f. Then $|f(n+1) - f(n)| \leq 1$ for all positive integers n.*

Proof. We show $f(n+1) \leq f(n) + 1$. Let $k = f(n)$, so

$$\prod_{j=1}^{n} a_j \geq \prod_{j=n+1}^{n+k} a_j,$$

and

$$\prod_{j=1}^{n} a_j < \prod_{j=n+1}^{n+k+1} a_j.$$

Observe

$$a_{n+1} \prod_{j=1}^{n} a_j < a_{n+1} \prod_{j=n+1}^{n+k+1} a_j = a_{n+1}^2 \prod_{j=n+2}^{n+k+1} a_j.$$

This implies that

$$\prod_{j=1}^{n+1} a_j = a_{n+1} \prod_{j=1}^{n} a_j < \prod_{j=n+2}^{n+1+k+2} a_j.$$

Since $f(n) = k$, it is impossible to have $f(n+1) = k+2$. In other words, the product $\prod_{j=n+1}^{n+1+f(n)} a_j$ cannot contain two more terms than the product $\prod_{j=n+1}^{n+f(n)} a_j$. The other required inequality $f(n+1) \geq f(n) - 1$ follows directly from the definitions. This establishes the result. □

Theorem 2.3 naturally leads us to define an asymptotic method of analyzing the growth rate of a sequence.

3. The Growth Limit

We will begin by comparing the growth function to the number of terms in the sequence.

Definition 3.1. *Given the growth function f of a sequence, we define the growth limit of the sequence as*

$$L = \lim_{n \to \infty} \frac{f(n)}{n}.$$

It is clear that L, when it exists, satisfies $0 \leq L \leq 1$. In fact, for the natural numbers we can compute this limit exactly.

Theorem 3.2. *The sequence of natural numbers has growth limit $L = 1$.*

Proof. Let f be the growth function for the sequence of natural numbers. From the definition of f we have

$$n! \geq (n+1)(n+2)\cdots(n+f(n))$$

and

$$n! < (n+1)(n+2)\cdots(n+f(n))(n+f(n)+1),$$

where each term on the right hand side is less than $2n$. It follows that

$$n! < (2n)^{f(n)+1},$$

or equivalently that

$$\frac{\ln(n!)}{\ln(2n)} < f(n) + 1.$$

We use the inequalities $n\ln(n) - n < \ln(n!)$ and $f(n) < n$ to obtain

$$\frac{n\ln(n) - n}{n\ln(2n)} < \frac{\ln(n!)}{n\ln(2n)} < \frac{f(n)}{n} + \frac{1}{n} < 1 + \frac{1}{n}.$$

Taking the limit as n approaches infinity and using the squeeze theorem,

$$\lim_{n\to\infty} \frac{f(n)}{n} = 1.$$

Hence, for sequence of natural numbers, $L = 1$. □

It readily follows from Theorem 3.2 that any polynomial-type sequence of the form $a_k = ck^m$, for $c, m \in \mathbb{N}$ also has growth limit 1. We leave the details as an exercise for the reader.

In contrast to these polynomial sequences, Sylvester's sequence, $s_n = \prod_{i=0}^{n-1} s_i + 1$, for $n \geq 1$ and $s_0 = 2$, grows very fast and is known to be doubly exponential, where the nth term of the sequence is given by $\lfloor E^{2^{n+1}} + \frac{1}{2} \rfloor$, and $E \approx 1.26408$, see [1]. The growth function, and hence growth limit, for Sylvester's sequence is identically zero; each new term in the sequence is always larger than the product of the preceding terms.

We now have examples of sequences with growth limits having the extreme values of 1 (polynomial) and 0 (doubly exponential). In searching for a sequence with a growth limit strictly between 0 and 1, it seemed natural to consider an exponential sequence.

We investigated the sequence $\{2^k\}_{k=1}^{\infty}$. For this sequence we can compute an explicit formula for the growth function. The base of the exponential is not important and the result holds for any base greater than one. We observe

$$P(n) = 2^{(1+2+3+\cdots+n)} = 2^{n(n+1)/2},$$

where we have used the well-known formula for the sum of consecutive integers. Again using the same formula, and setting $k = f(n)$, we have

$$Q(n) = 2^{(n+1)+(n+2)+\cdots+(n+k)} = 2^{nk+k(k+1)/2}.$$

To find a formula for $f(n)$, we seek the largest integer k such that

$$n(n+1)/2 \geq nk + k(k+1)/2.$$

This inequality is quadratic in k and yields

$$k \leq \frac{-(1+2n) + \sqrt{8n^2 + 8n + 1}}{2}.$$

Thus we then set $f(n) = \lfloor k \rfloor$, the largest integer less than or equal to k.

Table 2. Growth function for $a_k = 2^k$ and $b_k = k$.

n	1	2	3	4	5	6	7	8	9	10
$f_a(n)$	0	1	1	1	2	2	3	3	3	4
$f_b(n)$	0	0	1	1	2	3	3	4	5	5

Table 2 shows the first ten values of the growth function f_a for the exponential sequence and f_b for the sequence of natural numbers. Note the slower rate of increase of the growth function for the exponential sequence compared to the growth function for the sequence of natural numbers. Also observe that the growth function for the exponential sequence remains constant for three consecutive inputs, unlike the growth function for $\mathbb{N}$.

We use L'Hopital's Rule to calculate the growth limit for the exponential sequence.

$$L = \lim_{n\to\infty} \frac{f(n)}{n} = \sqrt{2} - 1 \approx 0.414213.$$

The previous example motivates the investigation of other exponential-type sequences. We consider the family of sequences $a_k = 2^{k^m}$, where m is a fixed positive integer. Table 3 shows numerical estimates for the growth limits that were obtained using *Mathematica.*

Table 3. Growth limits for exponential-type sequences.

m	a_k	L_{a_k}
1	2^k	0.41421
2	2^{k^2}	0.25992
3	2^{k^3}	0.18921
4	2^{k^4}	0.14870
5	2^{k^5}	0.12246
$\vdots$	$\vdots$	$\vdots$

We note that it is possible to construct subsequences of $\mathbb{N}$ for which the growth limit does not exist. The sequences exhibiting this behavior seem pathological, but, for the sake of completeness, we describe their construction. In order for a growth function to decrease, the sequence must have gaps that suddenly become very large relative to the size of the proceeding gaps. For example, the sequence

$$1, 2, 3, 4, 5, 6, 7, 8, 8! + 1, 8! + 2, 8! + 3, \ldots$$

has a growth function f shown in Table 4. The 9th term, $8! + 1$, is too large to include in any of the products $Q(n)$ for $n = 1, 2, \ldots 8$. When $n = 9$ the 9th term is part of the product $P(9)$, which is now much larger than $P(8)$. This allows subsequent terms in the sequence to be included in the products $Q(n)$ for $n \geq 9$. Thus, the growth function starts to increase again.

Table 4. A non-monotonic growth function.

n	1	2	3	4	5	6	7	8	9	10
$f(n)$	0	0	1	1	2	2	1	0	0	1

Using the idea of large gaps to cause the growth function to decrease, we believe that it is possible to construct a sequence for which L does not exist by having the growth function oscillate between zero and progressively larger positive values.

The essential idea is to construct a sequence that grows like the integers until $f(n) > n/2$. At this point we insert a sufficiently large integer to drive the growth function back to zero at the $(n + f(n))$th term of the sequence. We then resume growth like that of the integers until again $f(n) > n/2$.

We illustrate this construction with the following example:

$$1, 2, \ldots, 13, 14, 14! + 1, 14! + 2, 14! + 3, \ldots, 14! + 15, \ldots, 14! + 30, 14! + 31, \ldots.$$

Notice $f(9) = 5$, so $f(9)/9 > \frac{1}{2}$. Observe $f(9)$ steps later the insertion of $14! + 1$ forces $f(14) = 0$. Next, $f(29) = 15$, so $f(29)/29 > \frac{1}{2}$. And again $f(29)$ steps later we achieve $f(44) = 0$. We conjecture that the construction can be made to meet this criteria infinitely often resulting in

$$\limsup \frac{f(n)}{n} = \frac{1}{2} \text{ and } \liminf \frac{f(n)}{n} = 0,$$

which would imply the limit L does not exist.

4. Comparing the Growth Function for the Natural Numbers and Prime Numbers

We now consider the growth function and growth limit for the sequence of prime numbers. Experiments show a remarkable similarity between the growth function for the primes and the growth function for the natural numbers. Table 4 shows some values for the growth functions of both sequences. The growth function for the primes appears to be greater than the growth function for the natural numbers. Numerical experiments show this relationship persists as n increases. If this property holds for all $n > N$, for some integer N, then it follows immediately from Theorem 3.2 that the primes have growth limit $L = 1$.

Table 5. Comparison of growth functions for the natural numbers and the primes.

n	Natural numbers	Prime numbers
10	5	6
100	74	75
1000	817	834
10000	8595	8741
100000	88616	89790
1000000	904290	913844

For ease of understanding, let $p(n)$ denote the growth function for the prime numbers. We claim that growth function for the primes is strictly greater than the growth function for the natural numbers. Our claim would easily follow from Legendre's conjecture that there always exists a prime number between n^2 and $(n+1)^2$. In the absence of a proof of Legendre's conjecture, we have pursued a different avenue of proof using the Chebyshev function. In particular $p(n)$ is the largest integer such that

$$2\theta(p_n) \geq \theta(p_{n+p(n)})$$

where the Chebyshev function is $\theta(n) = \sum_{p \leq n} \log p$, and p_n is the nth prime number. Using this description of $p(n)$ we state and prove the following theorem.

Theorem 4.1. *The sequence of prime numbers has growth limit $L = 1$.*

Proof. Let f be the growth function for the sequence of natural numbers.

In [2] it is shown that, for sufficiently large n, the Chebyshev function $\theta(n)$ satisfies

$$\theta(p_k) \geq k\left(\log k + \log\log k + \frac{\log\log k - 2.04}{\log k}\right) \tag{4.1}$$

$$\theta(p_k) \leq k\left(\log k + \log\log k + \frac{\log\log k - 2}{\log k}\right). \tag{4.2}$$

This allows the following bounds to be established.

$$\begin{aligned}2\theta(p_n) &< \theta(p_{n+p(n)+1})\\ &\le (n+p(n)+1)\Bigg(\log(n+p(n)+1)+\log\log(n+p(n)+1)\\ &\quad+\frac{\log\log(n+p(n)+1)-2}{\log(n+p(n)+1)}\Bigg).\end{aligned}$$

Furthermore, one can see that in general $p(n) < n$, and so in particular

$$n+p(n)+1 \le 2n.$$

$$\begin{aligned}2\theta(p_n) &\le (n+p(n)+1)\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)\\ &=(n)\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)\\ &\quad+(p(n)+1)\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{2\theta(p_n)}{n\log n} &< \frac{n\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)}{n\log n}\\ &\quad+\frac{(p(n)+1)\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)}{n\log n}\\ &=\frac{\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)}{\log n}\\ &\quad+\frac{(p(n)+1)}{n}\frac{\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)}{\log n}\\ &=\frac{\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)}{\log n}\left(1+\frac{p(n)}{n}+\frac{1}{n}\right).\end{aligned}$$

We also note that $\frac{p(n)}{n} < 1$. Hence, $\left(1+\frac{p(n)}{n}\right) < 2$.

Thus,

$$\frac{2\theta(p_n)}{n\log n}\frac{\log n}{\left(\log(2n)+\log\log(2n)+\frac{\log\log(2n)-2}{\log(n)}\right)}-\frac{1}{n} < \left(1+\frac{p(n)}{n}\right) < 2. \tag{4.3}$$

Equation 4.3 gives us an upper and lower bound. One can see that the limit of the upper bound is 2. However, less obvious is that the limit of the lower bound is also two.

In order to show that the limit of the lower bound is 2 we apply the squeeze theorem to equations (4.1) and (4.2),

$$\lim_{n\to\infty}\frac{2\theta(p_n)}{n\log n}=2.$$

We also note that

$$\lim_{n\to\infty} \frac{\log n}{\left(\log(2n) + \log\log(2n) + \frac{\log\log(2n)-2}{\log(n)}\right)} = 1.$$

Combining these results yields

$$\lim_{n\to\infty} \frac{2\theta(p_n)}{n\log n} \frac{\log n}{\left(\log(2n) + \log\log(2n) + \frac{\log\log(2n)-2}{\log(n)}\right)} = 2.$$

Therefore, we can apply the squeeze theorem to equation 4.3, and

$$2 \le \lim_{n\to\infty} \left(1 + \frac{p(n)}{n}\right) \le 2.$$

Therefore,

$$\lim_{n\to\infty} \frac{p(n)}{n} = 1.$$

□

5. Conclusion

We have introduced the idea of a growth function and an associated growth limit for an increasing sequence of positive integers. Our results and numerical experiments lead to a number of open questions. In particular, is there a categorization of families of sequences that share a common growth limit? We seek a complete classification of all sequences with growth limit L. Also of interest is the question of characterizing growth functions in general. Specifically, given a function $g : \mathbb{N} \to \mathbb{N} \cup \{0\}$ such that $g(n) < n$ and $|g(n+1) - g(n)| \le 1$, what other conditions (if any) are required to ensure that g is the growth function of some sequence? Determining if growth limits are rational or irrational is also an open problem. Ultimately, we are interested in determining the complete set of all possible growth limits in the interval $[0, 1]$.

References

[1] A. V. Aho and N. J. A. Sloane, *Some doubly exponential sequences*, The Fibonacci Quarterly, **11.4** (1973), 429–437.

[2] P. Dusart, *Estimates of some functions over primes without R.H.*, arXiv preprint arXiv:1002.0442 (2010).

[3] J. J. Sylvester, *On a point in the theory of vulgar fractions*, American Journal of Mathematics, **3.4** (1880), 332–335.

MSC2010: 11B05, 11B83, 40A05

Department of Mathematics, United States Air Force Academy, Colorado Springs, CO 80840
E-mail address: xxx.xxx@xxx.xxx

Department of Mathematics, University of Nebraska–Lincoln, Lincoln, NE 68588
E-mail address: nathan.wakefield@unl.edu

Air Force Institute of Technology, Wright-Patterson AFB, OH 45433
E-mail address: xxx.xxx@xxx.xxx

Department of Mathematics, United States Air Force Academy, Colorado Springs, CO 80840
E-mail address: xxx.xxx@xxx.xxxx