Finding Prime Numbers
as Fixed Points of Sequences

Enrique Navarrete[1]
Daniel Orellana[2]

In this note we describe a method for finding prime numbers as fixed points of particularly simple sequences. Some basic calculations show that success rates for identifying primes this way are over 99.9%. In particular, it seems that the set of odd primes can be obtained as fixed points of the sequence which we call $A(1)$, the sequence of smallest divisors of triangular numbers, where the divisors are positive numbers that have not yet appeared in the sequence.

Keywords: Prime-generating function, formula for primes, sequences, sieve methods, fixed points, triangular numbers.

1. Introduction

Throughout history, there have been many attempts to find either formulas or polynomials that produce prime numbers. The most notable recent results of this type are due to Matiyasevich (1971) and Jones, Sato, Wada and Wiens (1976). The first result, Matiyasevich's Theorem, is an existence result which states that there exists a finite set of Diophantine equations from which the set of prime numbers can be obtained. The second result constructs an explicit polynomial of degree 25 in 26 variables whose positive values is the set of prime numbers. Since there is plenty of bibliography on these topics, the interested reader is encouraged to look at the references.

A different recent approach to produce primes has been developed by Rowland (2008), who found a recurrence that produces primes and the value 1 (both with repetitions). In this note we also detour from closed-form prime-generating approaches (*ie.* formulas, polynomials), to find a class of sequences where odd primes are produced as fixed points of these sequences. We describe the method below and some computational results in the next sections.

---

[1]Universidad Pontificia Bolivariana; enrique.navarrete@upb.edu.co
[2]Interijento Consulting, daniel.orellana@interijento.net

2. Key sequences and Fixed-Point Algorithm

The method we use to generate odd primes in the correct positions in the sequence of positive numbers relies on the construction of a class of sequences which we will denote by $\mathcal{A}(p)$, where sequences $A(p)$ in the class are themselves indexed by the odd primes p. A sequence $A(p)$ can be described as the sequence of smallest divisors of p-multiples of triangular numbers, where the divisors are positive numbers that have not yet appeared in the sequence.

We describe the method as follows:

1. Choose an odd prime p, produce its multiples kp and take the partial sums

$$q(n) = \sum_{k=0}^{n-1} (kp), \ n \geq 1; \ \text{these are the } p\text{-multiples of triangular numbers,}$$

   OEIS A000217.

2. Define a sequence with offset 1 as follows: let $a(1) = 1$ and let $a(n)$ be the smallest positive number not yet in the sequence such that $a(n)$ divides $q(n)$. Denote the sequence constructed this way by $A(p)$. Note that this sequence is well defined, that $a(1) = 1$ divides $q(1) = 0$, and that $a(2) = q(2) = p$.

3. Under some qualifications mentioned below, we get odd prime numbers as fixed points of the sequence $A(p)$, *ie.* as the terms $a(n)$ such that $a(n) = n$.

To illustrate, Table 1 shows the first few terms of A(7).

As we can see from the table, the fixed points of the sequence are the first values of the set of odd primes (with the exception of 7, which does not appear as a fixed point since $a(2) = p$ in all sequences $A(p)$, as mentioned above).[3]

Even though this method is indeed very precise for finding primes, as well as highly efficient, it does have a minimal amount of error, *ie.* the fixed points of the sequences $A(p)$ miss some primes. Fortunately, as will be discussed below, the primes missed by $A(p)$ can be detected by some other sequence $A(p')$, $p' \neq p$. Besides this minimal amount of error, the only pitfall encountered in the method is that it also produces nonprimes as fixed points (which we call "false negatives" from now on, using statistical terminology). To separate both kind of errors (miss primes *vs.* detect nonprimes), in Section 3 we describe results leaving out false negatives, which we will return to in Section 4. Figure 1 summarizes

---

[3]Note that the value 1 also appears as a fixed point but is not a prime; however, it will not be counted as a "false negative" or Type II error in Section 4 since, by definition, $a(1) = 1$ in sequences $A(p)$.

Fixed Points	n	Mult of 7	Sum $q(n)$	$a(n)$
fixed point	1	0	0	1
	2	7	7	7
fixed point	3	14	21	3
	4	21	42	2
fixed point	5	28	70	5
	6	35	105	15
	7	42	147	21
	8	49	196	4
	9	56	252	6
	10	63	315	9
fixed point	11	70	385	11
	12	77	462	14
fixed point	13	84	546	13
	14	91	637	49
	15	98	735	35
	16	105	840	8
fixed point	17	112	952	17
	18	119	1071	51
fixed point	19	126	1197	19
	20	133	1330	10
	21	140	1470	30
	22	147	1617	33
fixed point	23	154	1771	23
	24	161	1932	12
	25	168	2100	20

Table 1: The first terms of $A(7)$.

the two types of mistakes using terms from statistics or binary classification in general.[4]

	PRIME	NONPRIME
DETECT	Correct	False negative
DON'T DETECT	False positive	Correct

Figure 1: Classification Matrix.

---

[4]Assuming the null hypothesis "H_0: the number n is prime", a Type I error occurs when H_0 is rejected when it is true, or in our method, when a prime number n is not detected as a fixed point of the sequence $A(p)$. A Type II error occurs when H_0 is false and is not rejected; that is, when a number is nonprime and is detected as a fixed point. In hypothesis testing, a positive result corresponds to rejecting the null hypothesis, while a negative result corresponds to failing to reject it; the term "false" means the conclusion inferred is incorrect. Hence a Type I error is also lnown as a "false positive" and a Type II error as a "false negative".

3. Additional Observations and Results for Correctly Identified Primes

To check success rates of sequences $A(p)$ to detect primes, the first 10,000 terms of such sequences were computed for $p = 3, 5, 7, 11, 13, 17, 19, 23, 41, 97$ and 199. Table 2 below shows that for p as low as $p = 97$, the rate of success for detecting primes is as large as 99.92%. Before describing the results, we highlight an additional observation that appeared in all cases computed, which we state as a conjecture.

Conjecture 3.1. *The fixed points of sequences $A(p)$, p an odd prime, are odd.*[5]

Another observation stated as a conjecture is the following:

Conjecture 3.2. *A prime n in the sequence $A(p)$ appears either as the fixed point $n = a(n)$ or as $n = a(n+1)$.*

For example, the earliest failure in all cases computed occurs in $A(3)$ for the prime 17, which appears at $a(18)$.

Note that if Conjecture 3.2 is true, it implies that if we allow a prime n in the sequence $A(p)$ to appear either as the term $a(n)$ or as the term $a(n+1)$ (not too restrictive as compared to other methods), then the success rate for detecting primes is 100%.[6]

If we stick to the original requirement that prime n must appear in the term $a(n)$ of the sequence $A(p)$,[7] then the natural question to ask is how well the sequences $A(p)$ perform in detecting prime numbers depending upon the selection of p. As Table 2 and Figure 2 show, the success rates are very favorable and even though they are initially high (94.5% in $A(3)$), they seem to keep increasing with the selection of p in $A(p)$.

Success Rates of Sequences $A(p)$	$p = 3$	$p = 5$	$p = 7$	$p = 11$	$p = 41$	$p = 97$	$p = 199$
A) Number of Matches $n = a(n)$	1160	1145	1166	1176	1220	1226	1226
B) Number of Matches $n = a(n+1)$	67	82	61	51	7	1	1
C) Total Primes in 10,000 terms*	1227	1227	1227	1227	1227	1227	1227
Success Rates (A/C)	94.54%	93.32%	95.03%	95.84%	99.43%	99.92%	99.92%

* Excluding 2 and p

Table 2: Succes rates of sequences $A(p)$.

---

[5]Actually, this also seems to be the case for $A(1)$, the sequence of smallest divisors of triangular numbers not yet in the sequence, as will be discussed in Section 5.

[6]Keeping again in mind the issue of false negatives, which will be discussed in Section 4.

[7]That is, the prime of value n in its correct place in the sequence of positive numbers (not the nth prime).
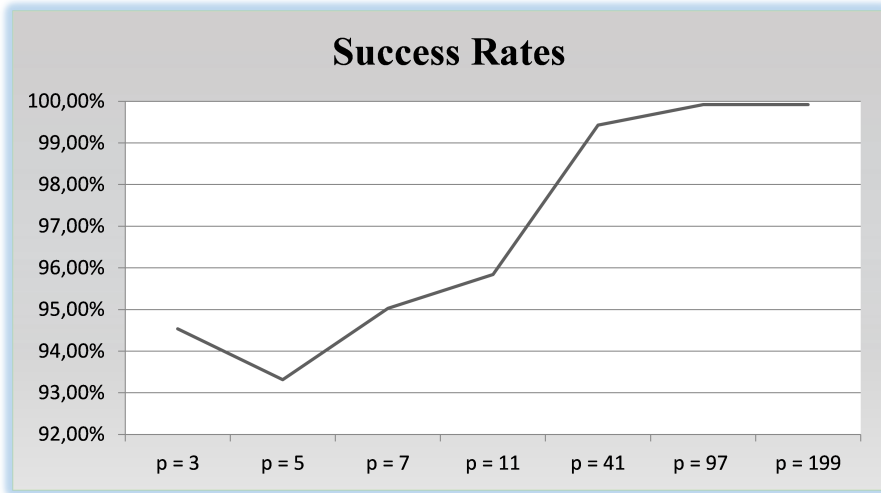
Figure 2: Success rates in sequences $A(p)$.

Note that there are 1229 primes in the first 10,000 terms of the positive numbers; however, in computing success rates for detecting odd primes in the sequences $A(p)$ we only consider 1227 primes since we exclude 2 as it is even and we exclude p itself since it appears as the second term of every sequence $A(p)$, as discussed in Section 1 (hence it is never a fixed point).

As Table 2 shows, success rates for identifying odd primes by the first 10,000 terms of the sequences $A(p)$ start at 95.5% for $p = 3$ and reach up to 99.9% for $p = 199$. The table distinguishes between primes identified by the correct term of the sequence, *ie.* "matches $n = a(n)$" and primes detected by the following term, $a(n + 1)$. Naturally, success rates in the table are computed using "true" matches $n = a(n)$, *ie.* the fixed points, since as mentioned above, using both types of matches would give a success rate of 100% for all sequences $A(p)$, as can be checked from the table.

Note from Table 2 that, aside from the issue of false negatives, the sequences $A(97)$ and $A(199)$ make only one mistake in identifying correctly the first 1227 primes (excluding 2 and p). It turns out that the prime they both miss is the 844th prime, which is 6529. The natural question that arises here is whether the same primes are missed by the sequences $A(p)$, as well as what kind of primes are missed. *ie.* primes of the form $4k+1$, $4k-1$, etc. These and other questions will be pursued further on.

4. Results taking into account False Negatives

We obtain the results below when false negatives are considered, *ie.* nonprimes detected as fixed points of sequences $A(p)$.

False Negatives in Sequences $A(p)$	$p = 3$	$p = 5$	$p = 7$	$p = 11$	$p = 41$	$p = 97$	$p = 199$
A) Number of false negatives in 10,000 terms	1179	1233	1248	1415	1478	1518	1526
B) Number of nonprimes in first 10,000 terms	8771	8771	8771	8771	8771	8771	8771
% (A/B)	13.44%	14.06%	14.23%	16.13%	16.85%	17.31%	17.40%

Table 3: Number of False Negatives in sequences $A(p)$.

As we see from Table 3, the percentage of false negatives in the first 10,000 terms of $A(p)$ increases with p, with percentages ranging from 13.4% to 17.4%. These percentages are the ratio of the number of nonprimes detected as fixed points in the first 10,000 terms of $A(p)$ over the total number of nonprimes in the first 10,000 positive numbers.[8]

Another way of summarizing the performance of the sequences $A(p)$ in detecting primes is by means of the classification matrix, an example of which is shown below for the first 10,000 terms of $A(3)$. Naturally, we would like the errors (off-diagonals) to be as small as possible.

	PRIME	NONPRIME
DETECT	94.54%	13.44%
DON'T DETECT	5.46%	86.56%

Figure 3: Classification Matrix for $A(3)$.

From Tables 2 and 3 we see there is a tradeoff between success rates and percentage of false negatives since both increase with p.

Even though the percentages of false negatives seem to be high, we see from the actual sequences computed that many of these wrongly identified nonprimes are multiples of 3 or 5, so they are easily detected. It would be interesting to compute the percentage of false negatives after removing these nonprimes; it is likely that these percentages will decrease considerably.

5. Results for nonprime values of of p in $A(p)$

What kind of sequences $A(p)$ do we get when we select p nonprime? Table 4 below provides a brief summary.

---

[8]Recall that the value 1 appears as a fixed point in all sequences $A(p)$ but is not counted as a false negative since it is by definition that $a(1) = 1$ in $A(p)$ (see footnote 3).

p	$q(n)$	OEIS	$A(p)$
1	Triangular Numbers	A000217	A111273*
2	Oblong numbers	A002378	Positive Integers
4	4 times the triangular numbers	A046092	Positive Integers
6	6 times the triangular numbers	A028896	Positive Integers
9	9 times the triangular numbers	A027468	Same as for $p = 3$

*$A(1)$ is A111273 if 0, the first triangular number, is omitted.

Table 4: Some sequences $A(p)$ for p nonprime.

Table 4 shows sequences $q(n)$, their classification number in OEIS, and the corresponding sequences $A(p)$ for some values of p nonprime. We see that we don't get any new sequences $A(p)$ except for $p = 1$.[9] In this case the sequence $q(n)$ is given by the triangular numbers $n(n + 1)/2 = \{0, 1, 3, 6, 10, 15, 21, \ldots\}$ (A000217 in OEIS). To get the sequence $A(1)$, if we drop 0, the first triangular number, then $A(1)$ is A111273 in OEIS, which starts as $\{1, 3, 2, 5, 15, 7, 4, 6, 9, 11, 22, 13, \ldots\}$. The fixed points of this sequence start with $\{1, 9, 25, 49, 57, 65, \ldots\}$. This sequence of fixed points, which is A113659 in OEIS, seems to produce some odd square numbers. However, if we compute $A(1)$ including the first triangular number 0, then by selecting 1 as the divisor of 0 (or $a(1) = a(2) = 1$), we get $A(1) = \{1, 1, 3, 2, 5, 15, 7, \ldots\}$, which is A111273 with an initial value of 1 appended. This shifts A111273 to the right by 1 and now the fixed points of the sequence are the odd prime numbers as in our previous results. Interestingly, the success rate of the shifted sequence up to the first 10,000 terms is 100% since in this case all 1228 odd primes in the first 10,000 positive numbers are detected as fixed points. On the other hand, the number of false negatives is 1532, for a rate of 17.47%, the largest one in our results.[10]

After checking the first 10,000 terms of the sequence $A(1)$, we state the following conjecture:

Conjecture 5.1. *Let $A(1)$ be the sequence of smallest divisors of triangular numbers not yet in the sequence with $a(1) = a(2) = 1$. Then the set of odd primes can be obtained as fixed points of this sequence.*[11]

6. Conclusion and Further Lines of Research

The method described for finding odd prime numbers is simple and efficient since, unlike other very intricate algorithms or formulas, it only requires sums

---

[9]Actually, for some values of p which are product of primes we get sequences $A(p)$ for some p prime but with the first terms permuted.

[10]Note that considering the first triangular number 0 when computing $A(1)$ is consistent with our method described in Section 2, which computes the partial sum $q(n)$ of multiples kp starting from $k = 0$.

[11]We don't use the term "the fixed points" due to the false negatives.

of multiples of primes, divisions to obtain the terms in the sequences $A(p)$, and an algorithm to check that the smallest divisor $a(n)$ of $q(n)$ is a positive number not yet in the sequence.

With respect to precision, Section 3 shows that success rates for finding primes are very high for the first 10,000 terms of the sequences $A(p)$. Furthermore, we have seen that these rates improve as the selection of p in $A(p)$ increases, to the extent of missing just one prime in the first 10,000 terms of the sequences $A(97)$ and $A(199)$. Hence, we would expect that no primes will be missed after a sufficiently high selection of p. However, until further results are obtained, we state these expectations as the following conjecture:

Conjecture 6.1. *Let p be a fixed odd prime and $a(n) \in A(p)$ the smallest positive number not yet in the sequence such that $a(n)$ divides $q(n) = \sum_{k=0}^{n-1}(kp)$, $a(1) = 1$. Then for large p, the set of odd primes (excluding p) can be obtained as fixed points of the sequence $A(p)$.*[12,13]

Regarding the errors present in the method, the "true errors", or primes not detected, seem to be manageable, since the computations show that if some prime is not picked up by a particular sequence $A(p)$, it is detected by another (and there are plenty of sequences $A(p)$ to work with). More important seems to be the "noise" introduced by false negatives, or nonprimes detected as fixed points, and even though they may be a drawback, many of them are multiples of 3 or 5, as mentioned previously, so they can be easily detected and discarded.[14]

It seems that further work should concentrate on determining what kind of primes are missed by the sequences $A(p)$, as well as what are the "best" individual sequences or combinations of sequences to minimize the percentage of false negatives.

---

[12]We exclude the value p since recall that $p = a(2)$ in all sequences $A(p)$. Also, we don't use the term "the fixed points" due to the false negatives.

[13]Actually, at the time of submission, all 1227 primes (excluding 2 and p) were correctly identified by the first 10,000 terms of $A(p)$ for $p = 541$, the 100th prime, for a success rate of 100%.

[14]Note that if Conjecture 3.1 is true, false negatives are odd.

References

[1] R. L. Goodstein and C. P. Wormell, Formulae for primes, *The Mathematical Gazette* **51** (1967) 35 - 38.

[2] James Jones, Formula for the nth prime number, *Canadian Mathematical Bulletin* **18** (1975) 433 - 434.

[3] James Jones, Daihachiro Sato, Hideo Wada, and Douglas Wiens, Diophantine representation of the set of prime numbers, *The American Mathematical Monthly* **83** (1976) 449 - 464.

[4] Yuri Matiyasevich, Diophantine representation of the set of prime numbers (in Russian), *Doklady Akademii Nauk SSSR* **196** (1971).

[5] Rowland, Eric S. (2008), "A Natural Prime-Generating Recurrence", *Journal of Integer Sequences*, **11**: 08.2.8, arXiv:0710.3217.

[6] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, published electronically at http://oeis.org.

[7] C. P. Willans, On formulae for the nth prime number, *The Mathematical Gazette* **48** (1964) 413 - 415.