Formulas for Primes

Eric Rowland

University of Liège

2015 February 20

The sequence of primes

$$\cancel{1}\ 2\ 3\ \cancel{4}\ 5\ \cancel{6}\ 7\ \cancel{8}\ \cancel{9}\ \cancel{10}\ 11\ \cancel{12}\ 13\ \cancel{14}\ \cancel{15}\ \cancel{16}\ 17\ \cancel{18}\ 19\ \cancel{20}\ \cancel{21}\ \cancel{22}\ 23\ \cancel{24}\ \cancel{25} \cdots$$

The sieve of Eratosthenes generates the sequence of primes:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, \ldots$$

Two questions:

- Is it easy to tell when a number is prime? (1 slide)
- Are there formulas that easily produce primes? (24 slides)

Can primality be determined quickly?

Trial division: Test divisibility by all numbers $2 \leq m \leq \sqrt{n}$.

Wilson's Theorem (Lagrange, 1773)

If $n \geq 2$, then n is prime if and only if n divides $(n-1)! + 1$.

For example, 5 divides $4! + 1 = 25$, but 6 doesn't divide $5! + 1 = 121$.

But determining whether an ℓ-digit number is prime using Wilson's theorem requires multiplying $\approx 10^{\ell}$ numbers.

Is there a polynomial-time algorithm for testing primality? Yes.

Agrawal, Kayal, & Saxena (2002) provided an algorithm that determines whether an ℓ-digit number is prime in $c \cdot \ell^{12}$ steps.

Willans' formula for the *n*th prime (1964)

$$p_n = 1 + \sum_{i=1}^{2^n} \left\lfloor \left(\frac{n}{\sum_{j=1}^{i} \left\lfloor \left(\cos \frac{(j-1)!+1}{j} \pi \right)^2 \right\rfloor} \right)^{1/n} \right\rfloor$$

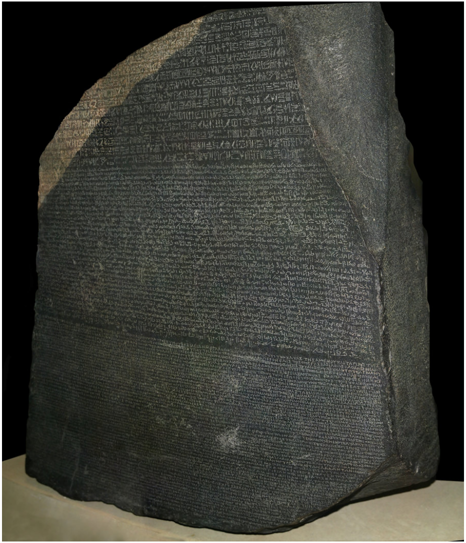
$$\frac{(j-1)!+1}{j} = \begin{cases} \text{an integer} & \text{if } j = 1 \text{ or } j \text{ is prime} \\ \text{not an integer} & \text{if } j \geq 2 \text{ and } j \text{ is not prime} \end{cases}$$

$$\left\lfloor \left(\cos \frac{(j-1)!+1}{j} \pi \right)^2 \right\rfloor = \begin{cases} 1 & \text{if } j = 1 \text{ or } j \text{ is prime} \\ 0 & \text{if } j \geq 2 \text{ and } j \text{ is not prime} \end{cases}$$

$$\sum_{j=1}^{i} \left\lfloor \left(\cos \frac{(j-1)!+1}{j} \pi \right)^2 \right\rfloor = \pi(i) + 1$$

$$\left\lfloor \left(\frac{n}{\pi(i)+1} \right)^{1/n} \right\rfloor = \begin{cases} 1 & \text{if } i < p_n \\ 0 & \text{if } i \geq p_n \end{cases}$$

Translation



Fermat primes

n	0	1	2	3	4	5	6	7	8	9	10
$2^n + 1$	2	3	5	9	17	33	65	129	257	513	1025

If $2^n + 1$ is prime and $n \geq 1$, must n be a power of 2? Yes:

$$a^m - b^m = (a - b) \cdot \left(a^{m-1} + a^{m-2}b + a^{m-3}b^2 + \cdots + b^{m-1}\right)$$

Suppose n is divisible by some odd number m.
Then letting $a = 2^{n/m}$ and $b = -1$ shows that
$a - b = 2^{n/m} + 1$ divides $a^m - b^m = (2^{n/m})^m - (-1)^m = 2^n + 1$.

$2^{16} + 1 = 65537$ is also prime!
Fermat conjectured that $2^{2^n} + 1$ is prime for all $n \geq 0$.

But Euler factored $2^{32} + 1 = 4294967297 = 641 \times 6700417$.
And no Fermat primes have been found since!

Mersenne primes

n	0	1	2	3	4	5	6	7	8	9	10
$2^n - 1$	0	1	3	7	15	31	63	127	255	511	1023

Marin Mersenne (1588–1648)

If $2^n - 1$ is prime, must n be prime? Yes:

$$2^{km} - 1 = (2^k - 1) \cdot \left(2^{(m-1)k} + \cdots + 2^{3k} + 2^{2k} + 2^k + 1\right).$$

However, $2^{11} - 1 = 2047 = 23 \times 89$.

Great Internet Mersenne Prime Search

Testing primality of $2^p - 1$ is (relatively) easy: Lucas–Lehmer test.

GIMPS is a distributed computing project begun in 1996.
`http://mersenne.org`

All 48 known Mersenne primes:

$2^{2}-1, 2^{3}-1, 2^{5}-1, 2^{7}-1, 2^{13}-1, 2^{17}-1, 2^{19}-1, 2^{31}-1, 2^{61}-1, 2^{89}-1, 2^{107}-1, 2^{127}-1, 2^{521}-1,$
$2^{607}-1, 2^{1279}-1, 2^{2203}-1, 2^{2281}-1, 2^{3217}-1, 2^{4253}-1, 2^{4423}-1, 2^{9689}-1, 2^{9941}-1, 2^{11213}-1, 2^{19937}-1,$
$2^{21701}-1, 2^{23209}-1, 2^{44497}-1, 2^{86243}-1, 2^{110503}-1, 2^{132049}-1, 2^{216091}-1, 2^{756839}-1, 2^{859433}-1,$
$2^{1257787}-1, 2^{1398269}-1, 2^{2976221}-1, 2^{3021377}-1, 2^{6972593}-1, 2^{13466917}-1, 2^{20996011}-1, 2^{24036583}-1,$
$2^{25964951}-1, 2^{30402457}-1, 2^{32582657}-1, 2^{37156667}-1, 2^{42643801}-1, 2^{43112609}-1, 2^{57885161}-1$

Largest known prime: $2^{57885161} - 1$.
It was discovered in January 2013 and has 17425170 decimal digits.

A prime-generating double exponential

In 1947, William Mills proved the existence of a real number b such that $\lfloor b^{3^n} \rfloor$ is prime for $n \geq 1$.

If the Riemann hypothesis is true, the smallest such b is

$$b = 1.3063778838630806904686144926026057 \cdots$$

and generates the primes

$$2, 11, 1361, 2521008887, 16022236204009818131831320183, \ldots.$$

But the only known way of computing digits of b
is by working backward from known large primes!

Euler's polynomial (1772)

Euler observed that $n^2 - n + 41$ is prime for $1 \leq n \leq 40$:

$$41, 43, 47, 53, 61, 71, 83, 97, 113, 131, 151, 173, 197, 223, 251,$$
$$281, 313, 347, 383, 421, 461, 503, 547, 593, 641, 691, 743, 797, 853,$$
$$911, 971, 1033, 1097, 1163, 1231, 1301, 1373, 1447, 1523, 1601$$

But for $n = 41$ the value is $1681 = 41^2$.

Does there exist a polynomial $f(n)$ that only takes on prime values?

Yes. The constant polynomial $f(n) = 3$ does!

Prime-generating polynomials

What about a non-constant polynomial?

Suppose $f(n)$ is prime for all $n \geq 1$.
Let $p = f(1)$. Then

$$f(1 + pk) = f(1) + p \times (\text{higher order terms})$$

is divisible by p for each $k \geq 1$.

No.

What about a multivariate polynomial?

Theorem (Jones–Sato–Wada–Wiens, 1976)

The set of positive values taken by the following degree-25 polynomial in 26 variables is equal to the set of prime numbers.

$$
\begin{aligned}
(k+2) \times \bigl(1 &- (wz+h+j-q)^2 \\
&- ((gk+2g+k+1)(h+j)+h-z)^2 \\
&- (2n+p+q+z-e)^2 \\
&- (16(k+1)^3(k+2)(n+1)^2+1-f^2)^2 \\
&- (e^3(e+2)(a+1)^2+1-o^2)^2 \\
&- ((a^2-1)y^2+1-x^2)^2 \\
&- (16r^2y^4(a^2-1)+1-u^2)^2 \\
&- (((a+u^2(u^2-a))^2-1)(n+4dy)^2+1-(x+cu)^2)^2 \\
&- (n+l+v-y)^2 \\
&- ((a^2-1)l^2+1-m^2)^2 \\
&- (ai+k+1-l-i)^2 \\
&- (p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m)^2 \\
&- (q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x)^2 \\
&- (z+pl(a-p)+t(2ap-p^2-1)-pm)^2\bigr)
\end{aligned}
$$

Corollary: If $k+2$ is prime, then there is a proof that $k+2$ is prime consisting of 87 additions and multiplications.

Programming with polynomials

The set of positive values taken by

$$n \cdot \left(1 - (n - 2m)^2\right)$$

for positive integers n, m is the set of positive even numbers:

$$(n - 2m)^2 = 0 \text{ has a solution in integers} \quad \Leftrightarrow \quad n \text{ is even.}$$

Given a system of equations whose integer solutions characterize primes, you can use the same trick.

The JSWW multivariate polynomial is an implementation of a primality test in the “programming language” of polynomials.

Hilbert’s 10th problem

Hilbert’s 10th problem

Is there an algorithm to determine whether a polynomial equation has positive integer solutions?

$$\begin{aligned} x^2 = y^2 + 2 \quad &\rightarrow \quad \text{no solution} \\ x^2 = y^2 + 3 \quad &\rightarrow \quad \text{solution exists } (x = 2, y = 1) \\ x^3 + y^3 = z^3 \quad &\rightarrow \quad \text{no solution} \\ x^3 + xy + 1 = y^4 \quad &\rightarrow \quad ??? \end{aligned}$$

Work of Davis, Matiyasevich, Putnam, and Robinson, 1950–1970:
No. Any set of positive integers output by a computer program (running forever) can be encoded as the set of positive values of a polynomial.

But where are the primes?

In practice, those “formulas for primes” don’t generate primes at all!

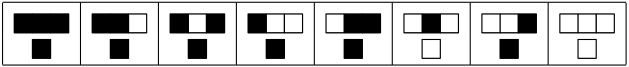
They are engineered to generate primes we already knew.

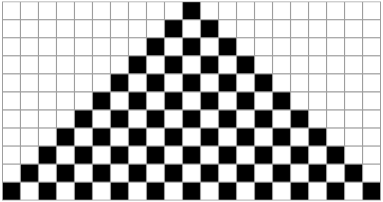
Are there formulas that generate primes we didn’t already know?

— INTERLUDE —

Cellular automata

- alphabet Σ (for example, $\Sigma = \{\square, \blacksquare\}$)
- a bi-infinite sequence on Σ (the initial condition)
- function $f : \Sigma^d \to \Sigma$ (the local update rule)

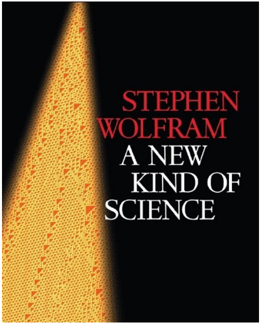




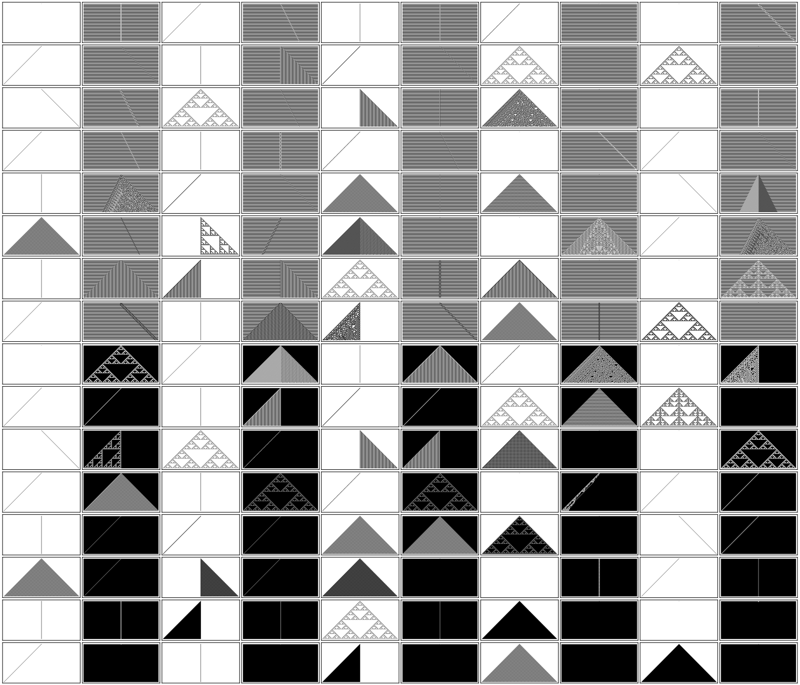
The “game of life” is a famous two-dimensional cellular automaton.

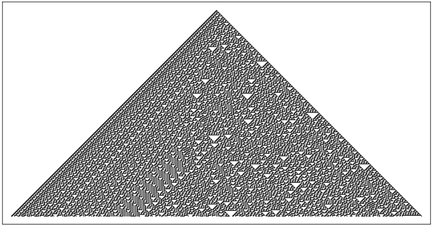
A New Kind of Science

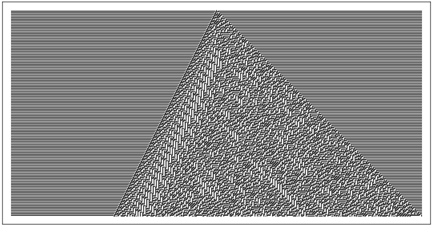
In 2002 Stephen Wolfram published a book on simple programs, including cellular automata.

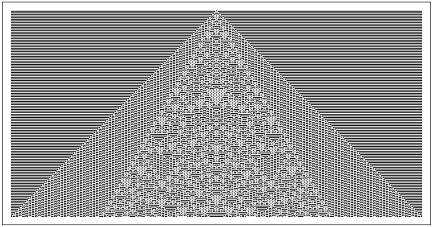


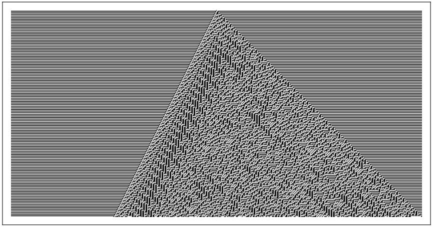
Wolfram's approach: Systematically look at all possible rules.

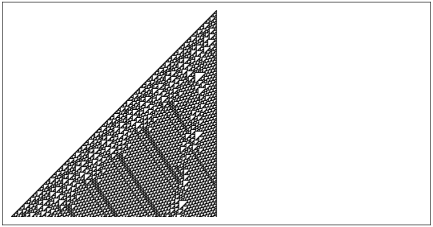


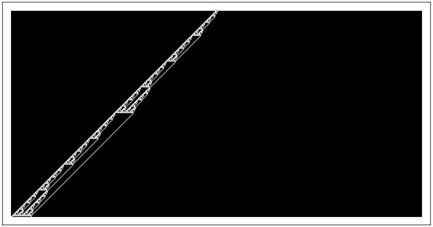






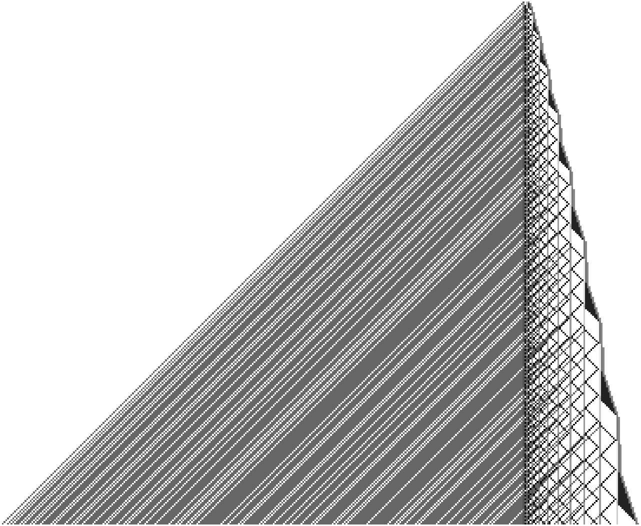






Programming with cellular automata

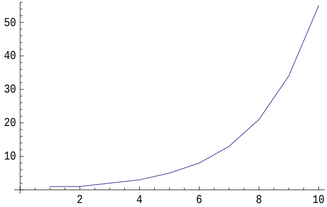
A 16-color rule depending on 3 cells that computes the primes:

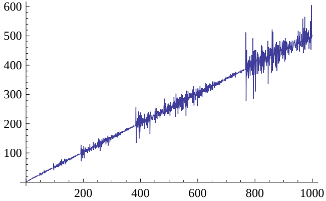


Recurrences

In 2003, Matthew Frank applied the approach to recurrences.

Fibonacci recurrence:
$s(n) = s(n-1) + s(n-2)$



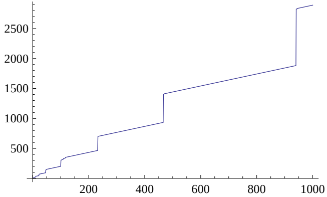


Hofstadter recurrence:
$s(n) = s(n - s(n-1)) + s(n - s(n-2))$

A new recurrence

Frank systematically substituted different functions into a template recurrence and looked at the pictures they generate.

This one caught his eye:
$s(n) = s(n-1) + \gcd(n, s(n-1))$



$$\begin{aligned}
s(1) &= 7\\
s(2) &= 7 + \gcd(2,7) = 7 + 1 = 8\\
s(3) &= 8 + \gcd(3,8) = 8 + 1 = 9\\
s(4) &= 9 + \gcd(4,9) = 9 + 1 = 10\\
s(5) &= 10 + \gcd(5,10) = 10 + 5 = 15\\
s(6) &= 15 + \gcd(6,15) = 15 + 3 = 18\\
s(7) &= 18 + \gcd(7,18) = 18 + 1 = 19\\
s(8) &= 19 + \gcd(8,19) = 19 + 1 = 20\\
s(9) &= 20 + \gcd(9,20) = 20 + 1 = 21\\
s(10) &= 21 + \gcd(10,21) = 21 + 1 = 22\\
s(11) &= 22 + \gcd(11,22) = 22 + 11 = 33\\
s(12) &= 33 + \gcd(12,33) = 33 + 3 = 36\\
s(13) &= 36 + \gcd(13,36) = 36 + 1 = 37
\end{aligned}$$

Difference sequence $s(n) - s(n-1) = \gcd(n, s(n-1))$:

1, 1, 1, 5, 3, 1, 1, 1, 1, 11, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 23, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 47, 3, 1, 5, ...

The difference sequence

1, 1, 1, 5, 3, 1, 1, 1, 1, 11, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 23, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 47, 3, 1, 5, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 101, 3, 1, 1, 7, 1, 1, 1, 1, 11, 3, 1, 1, 1, 1, 1, 13, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 233, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 467, 3, 1, 5, 3, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, …

$s(n) - s(n-1)$ appears to always be 1 or prime!

Prime-generating recurrence

Theorem (Rowland, 2008)

Let $s(1) = 7$, and for $n > 1$ let

$$s(n) = s(n-1) + \gcd(n, s(n-1)).$$

For each $n \geq 2$, $\gcd(n, s(n-1))$ is either 1 or prime.

This recurrence can generate primes we didn't expect to see!

Does it generate primes efficiently? No.

Without a shortcut, $\frac{p-3}{2}$ consecutive 1s precede p.
The shortcut requires computing the smallest prime divisor of $2n-1$ at each step.

Which primes appear?

5, 3, 11, 3, 23, 3, 47, 3, 5, 3, 101, 3, 7, 11, 3, 13, 233, 3, 467, 3, 5, 3, 941, 3, 7, 1889, 3, 3779, 3, 7559, 3, 13, 15131, 3, 53, 3, 7, 30323, 3, 60647, 3, 5, 3, 101, 3, 121403, 3, 242807, 3, 5, 3, 19, 7, 5, 3, 47, 3, 37, 5, 3, 17, 3, 199, 53, 3, 29, 3, 486041, 3, 7, 421, 23, 3, 972533, 3, 577, 7, 1945649, 3, 163, 7, 3891467, 3, 5, 3, 127, 443, 3, 31, 7783541, 3, 7, 15567089, 3, 19, 29, 3, 5323, 7, 5, 3, 31139561, 3, 41, 3, 5, 3, 62279171, 3, 7, 83, 3, 19, 29, 3, 1103, 3, 5, 3, 13, 7, 124559609, 3, 107, 3, 911, 3, 249120239, 3, 11, 3, 7, 61, 37, 179, 3, 31, 19051, 7, 3793, 23, 3, 5, 3, 6257, 3, 43, 11, 3, 13, 5, 3, 739, 37, 5, 3, 498270791, 3, 19, 11, 3, 41, 3, 5, 3, 996541661, 3, 7, 37, 5, 3, 67, 1993083437, 3, 5, 3, 83, 3, 5, 3, 73, 157, 7, 5, 3, 13, 3986167223, 3, 7, 73, 5, 3, 7, 37, 7, 11, 3, 13, 17, 3, 19, 29, 3, 13, 23, 3, 5, 3, 11, 3, 7972334723, 3, 7, 463, 5, 3, 31, 7, 3797, 3, 5, 3, 15944673761, 3, 11, 3, 5, 3, 17, 3, 53, 3, 139, 607, 17, 3, 5, 3, 11, 3, 7, 113, 3, 11, 3, 5, 3, 293, 3, 5, 3, 53, 3, 5, 3, 151, 11, 3, 31889349053, 3, 63778698107, 3, 5, 3, 491, 3, 1063, 5, 3, 11, 3, 7, 13, 29, 3, 6899, 3, 13, 127557404753, 3, 41, 3, 373, 19, 11, 3, 43, 17, 3, 320839, 255115130849, 3, 510230261699, 3, 72047, 3, 53, 3, 17, 3, 67, 5, 3, 79, 157, 5, 3, 110069, 3, 7, 1020460705907, 3, 5, 3, 43, 179, 3, 557, 3, 167, . . .

$p = 2$ cannot occur. But all odd primes below 587 do occur.

Theorem (Chamizo–Raboso–Ruiz-Cabello, 2011)

The difference sequence contains infinitely many distinct primes.

A variant

Benoit Cloitre looked at the recurrence

$$s(n) = s(n-1) + \operatorname{lcm}(n, s(n-1))$$

with $s(1) = 1$.

He observed that $\frac{s(n)}{s(n-1)} - 1$ seems to be 1 or prime for each $n \geq 2$:

2, 1, 2, 5, 1, 1, 1, 1, 5, 11, 1, 13, 1, 5, 1, 17, 1, 19, 1, 1, 11, 23, 1, 5, 13, 1, 1, 29, 1, 31, 1, 11, 17, 1, 1, 37, 1, 13, 1, 41, 1, 43, 1, 1, 23, 47, 1, 1, 1, 17, 13, 53, 1, 1, 1, 1, 29, 59, 1, 61, 1, 1, 1, 13, 1, 67, 1, 23, 1, 71, 1, 73, 1, 1, 1, 1, 13, 79, 1, 1, 41, 83, 1, 1, 43, 29, 1, 89, 1, 13, 23, 1, 47, 1, 1, 97, 1, 1, 1, 101, 1, 103, 1, 1, 53, 107, 1, 109, 1, 1, 1, 113, 1, 23, 29, 1, 59, 1, 1, 1, 61, 41, 1, 1, 1, 127, 1, 43, 1, 131, 1, 1, 67, 1, 1, 137, 1, 139, 1, 47, 71, 1, 1, 29, 73, 1, 1, 149, 1, 151, 1, 1, 1, 1, 1, 157, 1, 53, 1, 1, 1, 163, 1, 1, 83, …

No proof yet!