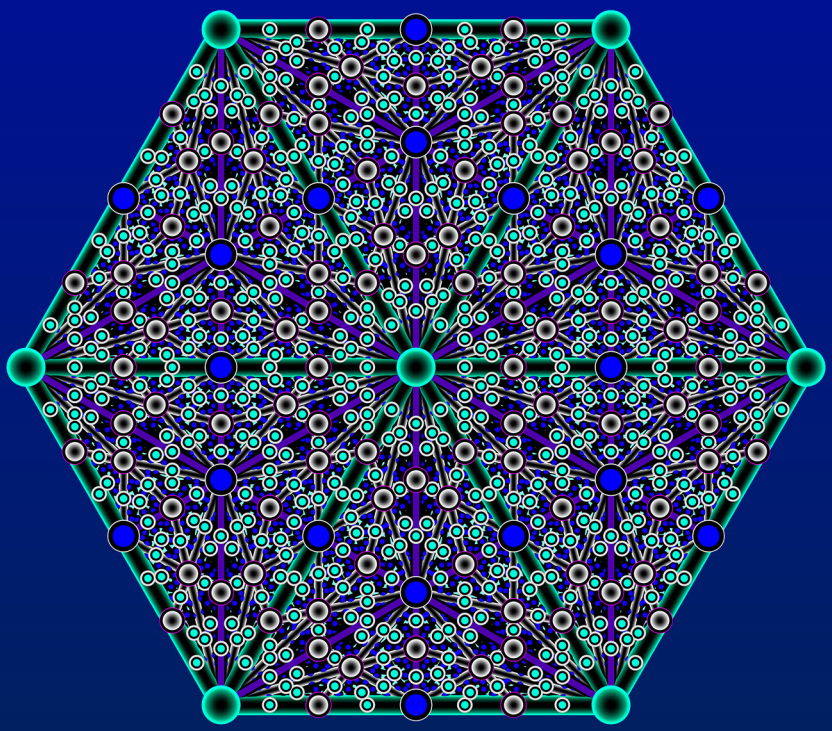
JOURNAL OF EXPERIMENTAL MATHEMATICS



ASSOCIATION FOR MATHEMATICAL RESEARCH

VOLUME 2 ISSUE 1

2026

JOURNAL OF EXPERIMENTAL MATHEMATICS

Editor-in-Chief
Jacob Tsimerman, University of Toronto (Canada)

Managing Editor
Djordje Milićević, Bryn Mawr College (USA)

Editorial Board
David Bailey, Lawrence Berkeley National Laboratory (USA)
Kathrin Bringmann, University of Cologne (Germany)
Joe Buhler, Reed College (USA)
Rafael de la Llave, Georgia Institute of Technology (USA)
Jan Draisma, University of Bern (Switzerland)
Sinan Güntürk, Courant Institute of Mathematical Sciences, New York University (USA)
Rob Kusner, University of Massachusetts at Amherst (USA)
Mark Levi, Pennsylvania State University (USA)
Alex Lubotzky, Hebrew University of Jerusalem & Weizmann Institute of Science (Israel)
Al Marden, University of Minnesota (USA)
Kaisa Matomäki, University of Turku (Finland)
Igor Rivin, Temple University (USA)
Peter Sarnak, Princeton University & Institute for Advanced Study (USA)
Rich Schwartz, Brown University (USA)
Sergei Tabachnikov, Pennsylvania State University (USA)
Yuri Tschinkel, Courant Institute of Mathematical Sciences, New York University (USA)
Akshay Venkatesh, Institute for Advanced Study (USA)

ISSN 3069-1125

PUBLISHED BY

Association for Mathematical Research, amathr.org,
Davis, CA; Jenkintown, PA.

JOURNAL OF EXPERIMENTAL MATHEMATICS, 2(1):194–204, 2026
jexpmath.org

Prime Numbers and Dynamics of the Polynomial $x^2 - 1$

Ivan Penkov * Michael Stoll

Received 03 Apr 2025; Revised 09 Sep 2025; Accepted 09 Nov 2025

Abstract:

Let $n \in \mathbb{Z}_{\geqslant 2}$. By $P(n)$ we denote the set of all prime divisors of the integers in the sequence $n, n^2-1, (n^2-1)^2-1, \ldots$. We ask whether the set $P(n)$ determines n uniquely under the assumption that $n \neq m^2-1$ for $m \in \mathbb{Z}_{\geqslant 2}$. This problem originates in the structure theory of infinite-dimensional Lie algebras. We show that the sets $P(n)$ generate infinitely many equivalence classes of positive integers under the equivalence relation $n_1 \sim n_2 \iff P(n_1) = P(n_2)$. We also prove that the sets $P(n)$ separate all positive integers up to 2^{29}, and we provide some heuristics on why the answer to our question should be positive.

Key words and phrases:
Arithmetic dynamics; Quadratic polynomial; Prime divisors; Lie algebra

1 Statement of the problem

We consider the following question.

Question 1.1. Let $n > 1$ be an integer, not of the form $n = m^2 - 1$. Consider the sequence $(a_k)_{k\geqslant 0} = (a_k(n))_{k\geqslant 0}$ defined recursively by $a_0 = n$, $a_{k+1} = a_k^2 - 1$. Let $P(n)$ be the set of all prime divisors of all a_k. Is n determined uniquely by $P(n)$?

Note that if p divides a_k, then p will divide infinitely many terms of the sequence, as the sequence considered modulo p will be periodic $0 \mapsto -1 \mapsto 0$ from that point on. In particular, $P(n^2-1) = P(n)$, and so $P(a_k(n)) = P(n)$ for all k.

*IP was supported in part by the DFG grant PE 980/9-1.

© 2026 Ivan Penkov and Michael Stoll
Licensed under a Creative Commons "Attribution-NonCommercial 4.0 International" license (CC BY-NC 4.0)
DOI: 10.56994/JXM.002.001.008

The above question arose in the structure theory of certain infinite-dimensional Lie algebras, see [PH22]. More precisely, for any $n \in \mathbb{Z}_{\geqslant 2}$ one has an interesting chain of inclusions of Lie algebras

$$\mathfrak{sl}(n) \subset \mathfrak{sl}(n^2-1) \subset \mathfrak{sl}((n^2-1)^2-1) \subset \dots,$$

where the natural representation of each Lie algebra restricts to the adjoint representation of the preceding Lie algebra. It is natural to ask when the direct limits of such chains are isomorphic as Lie algebras. As explained in [PH22], the notion of Dynkin index allows to infer that a positive solution to Question 1.1 implies that the direct limit Lie algebras arising from sequences as above starting respectively with $\mathfrak{sl}(n_1)$ and $\mathfrak{sl}(n_2)$ for $n_1 < n_2$ are isomorphic if and only if n_2 is a member of the sequence $n_1, n_1^2-1, (n_1^2-1)^2-1, \dots$.

An equivalent formulation of Question 1.1 is as follows.

Question 1.2. For a prime p, let $\bar{S}(p)$ be the subset of the finite field $\mathbb{F}_p$ consisting of all a such that iterating $x \mapsto x^2-1$ on a eventually produces 0. Denote by $S(p)$ the preimage of $\bar{S}(p)$ in $\mathbb{Z}$. Is it true that for each subset T of the set of all primes the inequality

$$\#\left(\{n \in \mathbb{Z}_{>0} : n \neq m^2-1\} \cap \bigcap_{p \in T} S(p) \cap \bigcap_{p \notin T} (\mathbb{Z} \setminus S(p))\right) \leqslant 1$$

holds?

On heuristic grounds (which we will detail below in Section 5), this appears to be very likely, but we also expect it to be very hard to prove.

The purpose of this paper is to gather some experimental evidence and to propose some heuristics that support a positive answer to the questions above.

We record here that some primes occur in all sets $P(n)$ and so do not provide useful information toward the answer of Question 1.1.

Lemma 1.3. *For all n, $\{2,3,7,23,19207\} \subset P(n)$, and these are the only primes below 10^5 with that property.*

Proof. One checks by a computation that $\bar{S}(p) = \mathbb{F}_p$ for the five primes occurring in the statement and for no other primes $p < 10^5$. □

This begs another question:

Question 1.4. Is $\bigcap_n P(n)$ finite or infinite?

See Section 5 for some heuristics regarding the likely answer.

Remark 1.5. In principle, the same questions can be asked for any polynomial with integral coefficients. We are not aware of any other specific polynomial for which it is possible to actually provide an answer to Question 1.1 with proof (other than trivial special cases like monomials).

We do want to point out, however, that the polynomial x^2-1 is somewhat special in that its critical point 0 is periodic. This has the effect that a prime dividing one term of the iteration sequence will divide infinitely many of them, which gives the problem a distinctly different flavor compared to the generic case.

This paper is structured as follows. In Section 2 we show that all sets $P(n)$ are infinite. Then in Section 3 we provide experimental evidence in favor of a positive answer to Questions 1.1 and 1.2. In Section 4 we show that there are infinitely many distinct sets $P(n)$. In Section 5 we present some heuristic considerations pertaining to the questions asked above. Finally, Section 6 has some rather speculative thoughts on how one might try to apply a result due to Huang to reduce Question 1.1 to Vojta's Conjecture.

2 The sets $P(n)$ are infinite (but likely sparse)

Before we state and prove the claim in the title of this section, we need some auxiliary results.

Lemma 2.1. *Let $n \in \mathbb{Z}_{\geqslant 1}$ and let $a_k = a_k(n)$ for $k \geqslant 0$ be defined as above. Then for all $k \geqslant 0$ we have the divisibility $a_k \mid a_{k+2}$.*

Proof. Note that $a_{k+2} = a_{k+1}^2 - 1 = (a_k^2 - 1)^2 - 1 = a_k^2(a_k^2 - 2)$. □

We set $f := x^2 - 1 \in \mathbb{Z}[x]$ and write f^m for its m-th iterate (i.e., $f^0 = x$ and $f^{m+1} = (f^m)^2 - 1 = f^m(x^2 - 1)$).

Lemma 2.2. *For every $m \geqslant 0$, the polynomial $f^m - 1$ is irreducible (in $\mathbb{Z}[x]$). Furthermore, 2 is the only ramified prime in the splitting field of $f^m - 1$ (for $m \geqslant 1$).*

Proof. The constant term of $f^m - 1$ is -1 when m is even and -2 when m is odd, and the constant term of $f^m(x-1) - 1$ is -2 when m is even and -1 when m is odd (we use that $f^{m+2}(0) - 1 = f^{m+1}(-1) - 1 = f^m(0) - 1$). The image of f^m in $\mathbb{F}_2[x]$ is x^{2^m} when m is even and $(x+1)^{2^m}$ when m is odd. Both observations together imply that for each m either $f^m - 1$ or $f^m(x-1) - 1$ is irreducible by the Eisenstein criterion at the prime 2.

It is not hard to show that $\operatorname{disc} g(x^2 - 1) = (-4)^{\deg g} g(-1)(\operatorname{disc} g)^2$, where $g \in \mathbb{Z}[x]$ and $\operatorname{disc} g$ is the discriminant of g; see, e.g., [Jon08, Lemma 2.6] for a more general statement. This implies that the discriminant of $f^m - 1$ is a power of 2 up to sign, so 2 is the only prime that can possibly ramify in the splitting field of $f^m - 1$. There are no unramified nontrivial extensions of $\mathbb{Q}$, so 2 indeed has to ramify when $m \geqslant 1$. (See also [BJ09, page 222], where the result regarding the ramification is stated without proof.) □

Now we show that $P(n)$ is infinite.

Proposition 2.3. *Let $n \in \mathbb{Z}_{\geqslant 2}$. Then $P(n)$ is an infinite set of prime numbers.*

Proof. From Lemma 2.1, we can deduce that the set of primes dividing one of $a_0, \dots, a_{k+1}$ is the same as the set of primes dividing $a_k a_{k+1}$. Now $a_k^2 - 2$ is coprime to the odd part of $a_k a_{k+1} = a_k(a_k^2 - 1)$ and is divisible by 2 at most once. This implies that unless $a_k = 2$, $a_{k+2} = a_k^2(a_k^2 - 2)$ has a prime divisor not dividing $a_k a_{k+1}$. So, with one possible exception (which occurs only when $n = 2$), each a_k contributes at least one new prime to $P(n)$. In particular, $P(n)$ must be infinite. □

We remark that this is a special case of the much more general Thm. 6.1 in [Jon08], which builds on results by Silverman [Sil07]. A similar result for rational functions such that 0 is preperiodic is given in [IS09].

On the other hand, the sets $P(n)$ are likely sparse in the following sense.

Conjecture 2.4. *Let $n \in \mathbb{Z}_{\geqslant 2}$. Then $P(n)$ is a set of prime numbers of density zero.*

Note that we are in the exceptional case $k=-1$ of [Jon08, Thm. 1.2(iii)].

The main result of [BGH+13] (in the simple form as given in the abstract there) at least implies that $P(n)$ does not have full density (by taking $\phi(t)=t^2-1$ and $\alpha=0$).

Let G_m denote the Galois group of $f^m(x)-1$ over $\mathbb{Q}$. Then Conjecture 2.4 would follow from the following statement.

Conjecture 2.5. *Let δ_m be the proportion of elements $\sigma \in G_m$ such that σ fixes at least one root of $f^m(x)-1$. Then $\lim_{m\to\infty}\delta_m = 0$.*

Note that by the Main Theorem in [ABC+22], the corresponding group $G_m(a)$ for $f^m(x)-a$ is the level-m quotient M_m of the 'arithmetic basilica group' M_∞ for all m when a is outside a 'thin set'. The statement of Conjecture 2.5 is expected to hold in these cases (Rafe Jones, private communication). When $a=1$, the limit group $G_\infty = \varprojlim G_m$ is of infinite index in M_∞, so this case requires additional work and is still open (with the expectation being that Conjecture 2.5 above should hold).

3 Separation of numbers up to a bound

The following definition is useful.

Definition 3.1. Let X be a positive integer and let P be a set of prime numbers. We say that P *separates the numbers up to X*, if the sets $P(n)\cap P$ are pairwise distinct for all $n \leqslant X$ not of the form $n=m^2-1$.

In other words, Questions 1.1 and 1.2 have a positive answer when restricted to $n \leqslant X$, and this can be verified by only considering divisibility by primes in P.

With this notion, we have the following experimental data.

Theorem 3.2. *Write $P_{\leqslant m}$ to denote the set of prime numbers $p \leqslant m$. Then:*

(1) *The numbers up to 10^1 are separated by $P_{\leqslant 47}$.*

(2) *The numbers up to 10^2 are separated by $P_{\leqslant 223}$.*

(3) *The numbers up to 10^3 are separated by $P_{\leqslant 379}$.*

(4) *The numbers up to 10^4 are separated by $P_{\leqslant 919}$.*

(5) *The numbers up to 10^5 are separated by $P_{\leqslant 2137}$.*

(6) *The numbers up to 10^6 are separated by $P_{\leqslant 3001}$.*

(7) *The numbers up to 10^7 are separated by $P_{\leqslant 4793}$.*

(8) *The numbers up to 10^8 are separated by $P_{\leqslant 5791}$.*

Proof. We run the following algorithm. The input is $X=10^k$ with $k=1,2,\ldots,8$.

1. Initialize $\mathcal{N} := \big\{\{n \in \mathbb{Z}_{>1} : n \leqslant X, \neg\exists m\colon n = m^2 - 1\}\big\}$, a set of finite sets of positive integers.
2. Set $p := 3$.
3. Repeat the following steps until $\mathcal{N}$ is empty.
   a. Replace p by the next larger prime number.
   b. Compute $\bar{S}(p)$.
   c. Replace each set N in $\mathcal{N}$ by the sets in the list $N \cap S(p)$, $N \setminus S(p)$ that have at least two elements.
4. Return p.

Taking into account that the primes 2 and 3 do not give information by Lemma 1.3, it is clear that this algorithm will return the minimal p such that $P_{\leqslant p}$ separates the numbers up to X when it terminates. The algorithm does in fact terminate for all $X = 10^k$ with $k \leqslant 8$ and returns the bounds given in the statement. □

We note that the growth of the bound on the primes that are necessary to separate the numbers up to X is numerically consistent with a growth of order $(\log X)^2$. The considerations in Section 5 below would predict $(\log X)^2 \log\log X$, which is also consistent with our numbers above ($\log\log X$ grows too slowly to allow distinguishing the two possibilities by experimental data).

We clearly get the "best" effect from using the information at a prime p when the set $\bar{S}(p)$ comprises close to half the elements of $\mathbb{F}_p$. So, to get a more efficient algorithm than the one used in the proof above, we do the following.

1. Pre-compute the sets $\bar{S}(p)$ for all p up to a suitable bound.
2. Sort the list of pairs $(p, \bar{S}(p))$ by increasing value of $|\#\bar{S}(p)/p - 1/2|$.
3. Use the primes in the order that is given by the sorted list of pairs.

The effect is that we can get similar results with less computation, because we need fewer primes to get separation.

For example, considering the primes up to 10000, the first ten primes in the sorted list, together with the value of $|\#\bar{S}(p)/p - 1/2|$, are

$$(2713, 0.00350),\quad (2137, 0.00726),\quad (1399, 0.0232),\quad (5927, 0.0534),\quad (8681, 0.0637),$$
$$(4799, 0.0741),\quad (3079, 0.0746),\quad (71, 0.0775),\quad (919, 0.0833)\quad (7951, 0.0875)\,.$$

The actual splitting of the sets in $\mathcal{N}$ can be done in a breadth-first (like in the algorithm in the proof above) or in a depth-first way. The latter is more space-efficient, but there is no significant difference in run times (as long as there is sufficient memory available; see below). The time complexity should be $\asymp X \log X$ (this is corroborated by the running times), with a memory requirement of $\asymp X$.

Using this improved algorithm, we can show:

Theorem 3.3. *The primes up to* 10000 *separate the numbers up to* $2^{29} \approx 5.37 \cdot 10^8$.

For comparison, our Magma implementation [Sto25] of the algorithm described in the proof of Theorem 3.2 takes a bit over two hours to verify the result for $X = 10^8$ on the second author's current laptop, while both the breadth-first and the depth-first versions of the second algorithm take about 35 minutes with the same bound (but only prove the slightly weaker result that the numbers up to 10^8 are separated by the primes below 10000). The computation verifying the statement of Theorem 3.3 takes a bit less than four hours (using the depth-first version and with some other tasks being executed in parallel; the breadth-first version requires too much memory to run in reasonable time, probably caused by the overhead incurred when working with a very large number of sets at the same time in Magma). We expect that a low-level implementation in C could be made sufficiently (time and space) efficient to be able to extend the bounds further, but it is perhaps not so clear that the additional effort spent for writing, testing and debugging such an implementation justifies the somewhat marginal improvement in experimental evidence.

4 Infinitely many classes

We can define an equivalence relation on $\mathbb{Z}_{>0}$ by declaring

$$n \sim m \;:\Longleftrightarrow\; P(n) = P(m)\,.$$

Then Theorem 3.3 shows that there are at least $536\,847\,742 = 2^{29} - \lfloor\sqrt{2^{29}+1}\rfloor$ distinct equivalence classes. We can in fact show more. But first we need a lemma.

Lemma 4.1. *Let $p \geqslant 3$ be a prime number. Then $\{-1,0,1\} \subset \bar{S}(p)$ (as subsets of $\mathbb{F}_p$), with equality if and only if $p \equiv \pm 3 \bmod 8$.*

Proof. Since 1 and -1 both map to 0 under $x \mapsto x^2-1$, the inclusion $\{-1,0,1\} \subset \bar{S}(p)$ holds for all p. The only preimage of -1 is 0. So $\bar{S}(p)$ is larger if and only if there are preimages of 1 in $\mathbb{F}_p$. But this is equivalent to 2 being a quadratic residue mod p, which is well known to be equivalent to $p \equiv \pm 1 \bmod 8$. □

Theorem 4.2. *There are infinitely many equivalence classes under the relation "$\sim$" defined above.*

Proof. We show that the sets $P(p)$, where $p \equiv \pm 3 \bmod 8$ is a prime, are pairwise distinct. This follows from the observation that $p \in P(p)$ and that for $q < p$ with $q \equiv \pm 3 \bmod 8$ we have $1 < q < p-1$, so $q \not\equiv -1,0,1 \bmod p$, which means that $p \notin S(q)$ by Lemma 4.1, hence $q \notin P(p)$. (A similar argument works for $P(p-1)$ or $P(p+1)$ with the same primes.) □

One can ask whether primes $p \equiv \pm 3 \bmod 8$ are sufficient to separate all positive integers not of the form m^2-1. However, the answer is "no". Indeed, note that

$$P'(n) := P(n) \cap \{p : p \equiv \pm 3 \bmod 8\} = \big\{p : p \equiv \pm 3 \bmod 8,\, p \mid (n-1)n(n+1)\big\}\,,$$

in particular,

$$\begin{aligned} P'(2) = P'(7) = P'(17) &= \{3\} \\ P'(4) = P'(5) = P'(6) = P'(9) = P'(16) &= \{3,5\} \\ P'(10) = P'(11) = P'(21) &= \{3,5,11\}\,. \end{aligned}$$

Moreover, looking at the sets $P'(n)$ for n up to 10000 seems to suggest that any given finite set of primes $p \equiv \pm 3 \bmod 8$ that contains 3 occurs infinitely often.

5 Heuristics

From the discussion in Section 3, it is reasonable to consider asymptotic properties of the distribution of the relative sizes $\#\bar{S}(p)/p$ as p gets large.

Here is a heuristic model for the size of $\bar{S}(p)$: We start with $1 \to 0 \leftrightarrow -1$ and follow 1 backwards. For a given $x \neq -1$, the chances that $x+1$ is or is not a square in $\mathbb{F}_p$ are equal, so we add two preimages with a probability of $1/2$ and recurse. This is related to some of the arguments in [HB17], where it is shown that the forward orbit of 0 under x^2+c on a finite field of odd size q has size $\ll q/\log\log q$.

We write

$$F(z) = \sum_{n\geq 0} \mathbb{P}(\#\bar{S}(p) = n) z^n = z^2 G(z) \in \mathbb{Q}[\![z]\!],$$

where

$$G(z) = \sum_{n\geq 0} g_n z^n$$

with g_n the probability that a binary tree that is randomly generated by the following procedure has n nodes in total: start with a root node and for each node in the tree, add two children with probability $1/2$. This tree is obtained by removing -1 and 0 from $\bar{S}(p)$; x is a child of y when $y = x^2 - 1$.

Alternatively, we obtain the same distribution starting from the root node by leaving it alone with probability $1/2$ and otherwise adding two random trees of the same kind to it as the left and right sub-trees. This leads to the equation

$$G(z) = \frac{z}{2}\bigl(1 + G(z)^2\bigr) \quad \Longrightarrow \quad G(z) = \frac{1-\sqrt{1-z^2}}{z},$$

so

$$F(z) = z\bigl(1-\sqrt{1-z^2}\bigr) = \sum_{n\geqslant 1} (-1)^{n+1} \binom{1/2}{n} z^{2n+1}.$$

The coefficient of z^{2n+1} is

$$\left|\binom{1/2}{n}\right| = \frac{1}{2n-1}\left|\binom{-1/2}{n}\right| = \frac{4^{-n}}{2n-1}\binom{2n}{n} \sim \frac{1}{2n\sqrt{\pi n}}.$$

For the regime of relative sizes $< 1-\varepsilon$, the heuristic model above should be fairly accurate. Indeed, counting the primes $p < 10^5$ such that $\#\bar{S}(p) = 2n+1$ shows a reasonably good agreement with the prediction from the model for, say, $n \leqslant 20$ (for larger n the numbers of primes are too small for a meaningful comparison).

The model would predict that for $0 < a < b < 1$, we should expect there to be about

$$\text{const} \int_3^X \left(\int_{ax}^{bx} \frac{dt}{t^{3/2}}\right) d\pi(x) \approx \text{const}\, \frac{\sqrt{X}}{\log X}\left(\frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}\right)$$

primes $p \leqslant X$ such that $ap \leqslant \#\bar{S}(p) \leqslant bp$.

For instance, we expect there to be infinitely many primes p such that

$$\left|\frac{\#\bar{S}(p)}{p} - \frac{1}{2}\right| \leqslant \frac{1}{10}.$$

The first few primes satisfying this inequality are

$$5,\ 71,\ 919,\ 1399,\ 2137,\ 2713,\ 3079,\ 4799,\ 5927,\ 7951,\ 8681,\ 10271,\ 10711,\ 11369,$$
$$12487,\ 12577,\ 22409,\ 22871,\ 24623,\ 24631,\ 27647,\ 29641,\ 46457,\ 54751,\ 84559,$$
$$87583,\ 99929,\ 103703,\ 105449,\ 106753,\ 120199,\ 120607,\ 123289,\ 131111,\ 147703.$$

This fits reasonably well with the heuristic growth.

Each such prime will lead to a nearly 50-50 split of the numbers n, and so the very likely fact that there are infinitely many of them gives another strong indication that the sets $P(n)$ do actually separate all positive integers n not of the form m^2-1.

For the "probability" that $\#\bar{S}(p) = p$ (i.e., $\bar{S}(p) = \mathbb{F}_p$), the model predicts something like $O(p^{-3/2})$, which would suggest that the answer to Question 1.4 is "finite".

However, the model does not take into account that $\#\bar{S}(p) \leqslant p$. If we include the cases where the size would be larger than p according to the model, the "probability" grows to $p^{-1/2}$, which would indicate that the answer is "infinite"! A refined model will be necessary to obtain reasonable predictions in the regime of $\bar{S}(p)$ close to maximal. So we now develop another heuristic regarding the existence of infinitely many primes with $\bar{S}(p) = \mathbb{F}_p$.

The directed graph on the vertex set $\mathbb{F}_p$ with edges $x \to y$ when $y = x^2-1$ consists of a number $N(p)$ of connected components, each of which contains precisely one cycle. Define the polynomials Φ_n so that

$$\prod_{d|n} \Phi_d = f^n - x,$$

where $f = x^2-1$ and f^n denotes the n-th iterate of f. So, e.g.,

$$\begin{aligned}
\Phi_1 &= x^2 - x - 1\\
\Phi_2 &= x^2 + x = x(x+1)\\
\Phi_3 &= x^6 + x^5 - 2x^4 - x^3 + x^2 + 1\\
\Phi_4 &= x^{12} - 6x^{10} + x^9 + 12x^8 - 4x^7 - 7x^6 + 4x^5 - 4x^4 + x^3 + 4x^2 - 2x + 1\\
&\text{etc.}
\end{aligned}$$

At least up to $m = 6$, we have

$$\mathrm{Gal}\Big(\prod_{n\leqslant m} \Phi_n\Big) = \prod_{n\leqslant m} \mathrm{Gal}(\Phi_n)$$

(where $\mathrm{Gal}(g)$ denotes the Galois group of a polynomial g over $\mathbb{Q}$) and (for $n \neq 2$)

$$\mathrm{Gal}(\Phi_n) = G_n := C_n \wr S_{k_n} \quad \text{with} \quad k_n = \frac{\deg \Phi_n}{n} = \frac{1}{n}\sum_{d|n} \mu\Big(\frac{n}{d}\Big) 2^d.$$

Here, $C_n \wr S_{k_n}$ denotes the wreath product, i.e., the semidirect product of $C_n^{\#S_{k_n}}$ with S_{k_n} acting via permutation of the factors, and C_n is the cyclic group of order n.

If we assume that this remains true for larger m (this is the case for generic polynomials f), then the density of primes p such that f has no cycle of length n ($\neq 2$) in $\mathbb{F}_p$ is (by the Chebotarev Density Theorem) the fraction of elements in G_n without a fixed point, which is

$$\sum_{k=0}^{k_n} \left(\sum_{j=0}^{k_n-k} \frac{(-1)^j}{j!} \right) \frac{(1-1/n)^k}{k!} .$$

This is quite close to

$$e^{-1} e^{1-1/n} = e^{-1/n} .$$

So under this model and assuming that the events "f has a cycle of length m on $\mathbb{F}_p$" are independent for $3 \leqslant m \leqslant p$, the expected "probability" for a prime p to have $N(p) = 1$ is close to

$$\frac{1}{2} \exp\left(- \sum_{3 \leqslant n \leqslant p} \frac{1}{n} \right) \sim \frac{1}{2} \exp\bigl(-(\log p + \gamma - 3/2)\bigr) = \frac{c}{p}$$

for a constant c. This would lead us to expect roughly $c \log\log X$ such primes up to X. In particular, this indicates that the answer to Question 1.4 is "infinite".

A bit more realistically, since (except for -1) every element in a cycle has *two* preimages, we can stop at $m = (p-3)/2$ (subtract 3 for $1, 0, -1$), which basically doubles c. We'd then actually expect a prime ≈ 163 and one ≈ 2130 and then one around 100 000 in the intersection of all sets $P(n)$. Of course, these are not precise predictions, and they serve only as an indication of the expected growth of the numbers in $\bigcap_n P(n)$.

See also [JKMT16] for some general results on the proportion of preperiodic points mod p under iteration of polynomials or rational functions.

Remark 5.1. The use of heuristics based on the assumption that polynomial maps on finite fields behave randomly is fairly standard in the study of arithmetic dynamical systems; see, e.g., [BGH+13, Section 5].

6 A speculation

Assuming Vojta's Conjecture, Huang [Hua20], inspired by Silverman's re-interpretation in [Sil05] of a result by Bugeaud, Corvaja and Zannier [BCZ03], shows a general result that says the following. Let $f_1, f_2 \in \mathbb{Z}[x]$ be polynomials of degree d satisfying some mild condition, let ε be a fixed positive real number, and let a and b be integers. If the sequence $\bigl((f_1^n(a), f_2^n(b)\bigr)_{n \geq 0}$ is Zariski-dense in the affine plane, then

$$\gcd\bigl(f_1^n(a), f_2^n(b)\bigr) \ll e^{\varepsilon d^n} .$$

(We again write f_j^n for the nth iterate of f_j.)

We can try to apply this with $f_1 = f_2 = f^2$ (where $f^2(x) = x^2(x^2-2)$ is the second iterate of $f = x^2 - 1$); it is not hard to see that a positive answer to the analogue of Question 1.1 for f^2 implies a positive answer to the original question.

Write $P^2(n)$ for the set of primes dividing $a_{2k}(n)$ for some $k \geq 0$. Assume that $P^2(n) = P^2(n')$ for $n, n' \geq 2$. *If* we could show that this implies that there is some $\delta > 0$ such that

$$\gcd\bigl(f^{2k}(n), f^{2k}(n')\bigr) > e^{\delta 4^k}$$

for infinitely many k, then Huang's result would show (if we also assume Vojta's Conjecture) that the sequence $\bigl((f^{2k}(n), f^{2k}(n'))\bigr)_{k\geq 0}$ is not Zariski dense, so there is some polynomial $F \in \mathbb{Z}[u,v]$ such that

$$F\bigl(f^{2k}(n), f^{2k}(n')\bigr) = 0$$

for all $k \geq 0$. It seems rather plausible that this would imply that n' is in the f^2-orbit of n or vice versa.

Note that when $p \in P^2(n) \cap P^2(n')$, then there is some $k_0(p)$ such that

$$v_p\bigl(\gcd\bigl(f^{2k}(n), f^{2k}(n')\bigr)\bigr) \geq 2^{k-k_0(p)} \qquad \text{for all } k \geq k_0(p).$$

Here $v_p(m)$ denotes the exponent of p in the prime factorization of m. The desired lower bound would follow if we could prove something like

$$\sum_{p:\, k_0(p)\leq k} 2^{-k_0(p)} \log p \gg 2^k \qquad \text{as } k \to \infty,$$

where the sum is over all $p \in P^2(n) = P^2(n')$ such that $k_0(p) \leq k$. This probably comes down to showing that the first index k such that a prime p divides $a_{2k}(n)$ is not too far away (in most cases) from the corresponding index for n'. Unfortunately, this still seems to be rather difficult.

References

[ABC+22] Faseeh Ahmad, Robert L. Benedetto, Jennifer Cain, Gregory Carroll, and Lily Fang. The arithmetic basilica: a quadratic PCF arboreal Galois group. *J. Number Theory*, 238:842–868, 2022. 197

[BCZ03] Yann Bugeaud, Pietro Corvaja, and Umberto Zannier. An upper bound for the G.C.D. of $a^n - 1$ and $b^n - 1$. *Math. Z.*, 243(1):79–84, 2003. 202

[BGH+13] Robert L. Benedetto, Dragos Ghioca, Benjamin Hutz, Pär Kurlberg, Thomas Scanlon, and Thomas J. Tucker. Periods of rational maps modulo primes. *Math. Ann.*, 355(2):637–660, 2013. 197, 202

[BJ09] Nigel Boston and Rafe Jones. The image of an arboreal Galois representation. *Pure Appl. Math. Q.*, 5(1):213–225, 2009. 196

[HB17] D. R. Heath-Brown. Iteration of quadratic polynomials over finite fields. *Mathematika*, 63(3):1041–1059, 2017. 200

[Hua20] Keping Huang. Generalized greatest common divisors for orbits under rational functions. *Monatsh. Math.*, 191(1):103–123, 2020. 202

[IS09] Patrick Ingram and Joseph H. Silverman. Primitive divisors in arithmetic dynamics. *Math. Proc. Cambridge Philos. Soc.*, 146(2):289–302, 2009. 196

[JKMT16] Jamie Juul, Pär Kurlberg, Kalyani Madhu, and Tom J. Tucker. Wreath products and proportions of periodic points. *Int. Math. Res. Not. IMRN*, 2016(13):3944–3969, 2016. 202

[Jon08] Rafe Jones. The density of prime divisors in the arithmetic dynamics of quadratic polynomials. *J. Lond. Math. Soc. (2)*, 78(2):523–544, 2008. 196, 197

[PH22] Ivan Penkov and Crystal Hoyt. *Classical Lie algebras at infinity*. Springer Monographs in Mathematics. Springer, Cham, 2022. 195

[Sil05] Joseph H. Silverman. Generalized greatest common divisors, divisibility sequences, and Vojta's conjecture for blowups. *Monatsh. Math.*, 145(4):333–350, 2005. 202

[Sil07] Joseph H. Silverman. *The arithmetic of dynamical systems*, volume 241 of *Graduate Texts in Mathematics*. Springer, New York, 2007. 196

[Sto25] M. Stoll. Magma code verifying the results of Section 3, 2025. https://www.mathe2.uni-bayreuth.de/stoll/magma/index.html. 199

AUTHORS

Ivan Penkov

Constructor University Bremen
Campus Ring 1
28759 Bremen
Germany

ivanpenkov@yahoo.com
http://math.constructor.university/penkov/

Michael Stoll

Mathematisches Institut
Universität Bayreuth
Universitätsstraße 30
95440 Bayreuth
Germany

Michael.Stoll@uni-bayreuth.de
http://www.mathe2.uni-bayreuth.de/stoll/