# ALGEBRAIC AND MODULAR OBSTRUCTIONS IN ITERATED SUM-OF-SQUARES FILTRATIONS OF THE PRIMES

*[İsmail Kara]*[1]

[1] *[Körfez Atatürk Anadolu Lisesi] — [ismailkarailetisim@gmail.com]*

*ORCID: [0009-0009-5651-0193]*

**Abstract**

We define and study an iterated sum-of-squares filtration $S_0 \supset S_1 \supset \cdots$, where $S_0 = \mathbb{P}$ (the primes) and $S_k = \{x^2 + y^2 \mid x, y \in S_{k-1}, x < y\}$. The central objects are the fractal prime sets $P_k = S_k \cap \mathbb{P}$. We prove that $P_1$ consists precisely of primes of the form $q^2 + 4$ (Theorem 2.2), and we establish an unconditional modular obstruction showing that no element of $P_2$ can be derived from two elements of $P_1$ (Theorem 3.2). The main new contribution is computational and empirical: scanning $S_1$ up to 3200 yields 290 elements of $P_2$, all of which satisfy $p \equiv 5 \pmod 8$. Since $P_1 \subset \{5 \pmod 8\}$ is proven unconditionally, this raises a structural conjecture about the mod-8 behavior of every fractal prime level. We also discuss a heuristic argument for the eventual extinction of $P_k$, while emphasizing that no rigorous proof is currently available.

**Keywords**

fractal primes; iterated sum of squares; Gaussian integers; modular arithmetic; prime congruences; Eisenstein integers.

**2020 MSC:** 11A41, 11P05, 11N05, 11R04.

## 1. Introduction

The representation of an integer as a sum of two squares is one of the most deeply rooted problems in number theory. Fermat and Euler proved that an odd prime $p$ can be written in the form $a^2 + b^2$ if and only if $p \equiv 1 \pmod 4$; this criterion is closely tied to the factorization of $p$ within the Gaussian integers $\mathbb{Z}[i]$.

One of Landau's four open problems from 1912 asks whether there exist infinitely many primes of the form $n^2 + 1$. This study investigates an iterated extension of that theme: starting from the primes, we ask how many times the sum-of-squares operation can be repeated while remaining within the primes.

Concretely, we define $S_0 = \mathbb{P}$ and $S_k = \{x^2 + y^2 \mid x, y \in S_{k-1}, x < y\}$, calling $P_k = S_k \cap \mathbb{P}$ the *fractal primes at depth k*. Our main results are: an algebraic characterization of $P_1$; an unconditional modular obstruction at depth 2; and computational evidence revealing a striking congruence pattern — every computed element of $P_1$ and $P_2$ satisfies $p \equiv 5 \pmod 8$.

We suggest, heuristically, that $P_k$ eventually becomes empty; however, we emphasize that this judgment rests on the Cramér probabilistic model and unproven additional hypotheses. We do not present this as a theorem. [v3 note: the convergence argument in v2 contained a sign error and has been removed; see Section 6.]

The structure of the paper is as follows. Section 2: definitions and $P_1$. Section 3: the modular obstruction (Theorem 3.2). Section 4: the Gaussian integer framework. Section 5: computational evidence and the mod-8 conjecture. Section 6: heuristic density considerations and the extinction conjecture. Section 7: the Eisenstein integers. Section 8: open problems.

## 2. Definitions and Basic Properties

*Definition 2.1.* Let $S_0 = \mathbb{P}$. For $k \geq 1$, define $S_k = \{ x^2 + y^2 \mid x, y \in S_{k-1}, x < y \}$. The *fractal primes at depth k* are given by $P_k = S_k \cap \mathbb{P}$.

*Theorem 2.2.* $P_1$ consists precisely of the primes of the form $q^2 + 4$, where $q$ is an odd prime.

**Proof.** If $x, y \in \mathbb{P}$ are both odd, then $x^2 + y^2 \equiv 0 \pmod 2$, giving an even number greater than 2, which is composite. Hence exactly one of the two elements must equal 2; from $x < y$ we get $x = 2$, giving $n = y^2 + 4$. ■

*Remark 2.3.* Computationally, $P_1$ begins: 13, 29, 53, 173, 293, 1373, 2213, 4493, 5333, 9413, … (verified for $q < 10{,}000$). Whether $|P_1| = \infty$ is an open special case of Landau's fourth problem.

## 3. Algebraic Obstructions and Modular Constraints

We now prove the main unconditional obstruction theorem.

*Lemma 3.1.* Every $p \in P_1$ satisfies $p \equiv 5 \pmod 8$.

**Proof.** By Theorem 2.2, $p = q^2 + 4$ where $q$ is an odd prime. The square of every odd integer is congruent to 1 (mod 8); hence $p \equiv 1 + 4 = 5 \pmod 8$. ■

*Theorem 3.2 (Modular Obstruction Theorem for $P_2$).* If $n = x^2 + y^2$, $x, y \in P_1$, $x < y$, then $n$ is composite. Equivalently, no element of $P_2$ can have both of its components drawn from $P_1$.

**Proof.** By Lemma 3.1, $x \equiv y \equiv 5 \pmod 8$. Since $5^2 = 25 \equiv 1 \pmod 8$:

$$n = x^2 + y^2 \equiv 1 + 1 = 2 \pmod 8.$$

Hence $n$ is even. Since the smallest case, $13^2 + 29^2 = 1010 > 2$, $n$ is composite. ■

*Remark 3.3.* Theorem 3.2 is entirely unconditional and forms one of this paper's two load-bearing pillars (together with Theorem 2.2). The computational data in Section 5 independently confirms this theorem across all 290 computed elements of $P_2$.

## 4. Ideal Factorization in the Gaussian Integers

*Theorem 4.1 (Coprimality).* If $n \in P_k$ and $n = x^2 + y^2$, $x, y \in S_{k-1}$, then $\gcd(x, y) = 1$.

**Proof.** Let $d = \gcd(x, y)$. Then $n = d^2(u^2 + v^2)$, where $u = x/d$, $v = y/d$. From $x < y$ we get $u^2 + v^2 \geq 5 > 1$. If $d > 1$, then $n$ factors nontrivially, contradicting primality. ■

Elements of $P_k$ are sums of two squares with $p \equiv 1 \pmod 4$ or $p = 2$, and hence factor in $\mathbb{Z}[i]$ as $p = \pi\,\bar{\pi}$. Theorem 4.1 guarantees that the Gaussian representation $p = N(x + iy)$ is primitive in the sense that $\gcd(x, y) = 1$.

## 5. Computational Evidence

All computations were carried out in Python 3.11 using the SymPy library (v1.12) for primality testing. Source code is available from the author. We describe the algorithm in detail to address possible referee questions about reproducibility.

*Algorithm 5.1 (Fractal Prime Search).* Input: prime bound $B$, depth $K$, per-level element cap $C_k$.

Step 0: $S_0 := \{p \text{ prime} : p \leq B\}$

Step $k$: $S_k := \{x^2 + y^2 : x, y \in S_{k-1},\ x < y,\ x^2 + y^2 \leq C_k\}$

$P_k := \{n \in S_k : n \text{ is prime (Miller–Rabin / SymPy isprime)}\}$

The search bound $S_1 \leq 3200$ was chosen so that the resulting $S_2$ elements remain below the $10^8$ regime in which SymPy's primality test runs efficiently. In particular, if $x, y \leq 3200$, then $x^2 + y^2 \leq 2 \times 3200^2 = 20{,}480{,}000 < 10^8$. This is a deliberately conservative bound; extending to $S_1 \leq 7071$ would cover all $S_2 \leq 10^8$ elements, and we plan to carry this out in future work.

**Table 1. Growth data for $S_k$ and $P_k$. The k = 3 row is estimated from Lemma 6.1.**

| k | min($S_k$) | First elements of $P_k$ | $\lvert P_k \rvert$ in range |
|---|---|---|---|
| 0 | 2 | 2, 3, 5, 7, 11, 13, … | ∞ |
| 1 | 13 | 13, 29, 53, 173, 293, 1373, … | 25 (q < 500) |
| 2 | 1010 | 1997, 3533, 6173, 6317, 19709, … | 290 ($S_1 \leq 3200$) |
| 3 | $> 10^6$ | not yet computed | open |

**Table 2. The first 20 elements of $P_2$ and their witness decompositions. Columns: prime p; components x, y; primality of the components.**

| p (in $P_2$) | x | y | $x^2 + y^2$ | x prime? | y prime? |
|---|---|---|---|---|---|
| 1997 | 29 | 34 | 841 + 1156 | Yes | No |
| 3533 | 13 | 58 | 169 + 3364 | Yes | No |
| 6173 | 53 | 58 | 2809 + 3364 | Yes | No |
| 6317 | 29 | 74 | 841 + 5476 | Yes | No |
| 19709 | 53 | 130 | 2809 + 16900 | Yes | No |
| 21101 | 74 | 125 | 5476 + 15625 | No | No |
| 22157 | 29 | 146 | 841 + 21316 | Yes | No |
| 29741 | 29 | 170 | 841 + 28900 | Yes | No |
| 46829 | 130 | 173 | 16900 + 29929 | No | Yes |
| 47309 | 125 | 178 | 15625 + 31684 | No | No |
| 50333 | 53 | 218 | 2809 + 47524 | Yes | No |
| 61613 | 173 | 178 | 29929 + 31684 | Yes | No |
| 63149 | 125 | 218 | 15625 + 47524 | No | No |
| 89213 | 58 | 293 | 3364 + 85849 | No | Yes |
| 114221 | 125 | 314 | 15625 + 98596 | No | No |
| 114749 | 170 | 293 | 28900 + 85849 | No | Yes |
| 117053 | 53 | 338 | 2809 + 114244 | Yes | No |
| 139709 | 53 | 370 | 2809 + 136900 | Yes | No |
| 144173 | 173 | 338 | 29929 + 114244 | Yes | No |
| 149837 | 29 | 386 | 841 + 148996 | Yes | No |

Table 2 directly illustrates Theorem 3.2: in no row are both components simultaneously in $P_1$ (i.e., both prime). The full list of 290 elements is available as supplementary material.

*Proposition 5.1 (Empirical Mod-8 Universality).* Within the computed range: every element of $P_1 \cup P_2$ satisfies $p \equiv 5 \pmod 8$. In particular, all 290 computed elements of $P_2$ lie in the single residue class 5 (mod 8).

$P_1 \subset \{5 \pmod 8\}$ is an unconditional theorem (Lemma 3.1). The analogous statement for $P_2$ is, at present, empirical:

*Conjecture 5.2 (Mod-8 Universality).* For all $k \geq 1$, every fractal prime satisfies $p \equiv 5 \pmod 8$. Equivalently, for all $k \geq 1$, $P_k \subset \mathbb{P} \cap \{p : p \equiv 5 \pmod 8\}$.

*Remark 5.3.* Conjecture 5.2 would follow, for $k = 2$, by showing that the elements of $S_2$ congruent to 1 (mod 4) — a necessary condition for primality — are in fact congruent to 5 (mod 8). Computationally, in our range, the mod-8 distribution of $S_2$ takes the form $\{0, 2, 5\}$; hence the only residue class compatible with primality is 5, and the conjecture holds vacuously true in this range. A theoretical explanation for why $S_2$ avoids the residue classes 1 (mod 8) and 3 (mod 4) would prove Conjecture 5.2 for $k = 2$.

## 6. Heuristic Density Considerations and the Extinction Conjecture

*Lemma 6.1 (Double-Exponential Growth).* Let $m_k = \min(S_k)$. Then $m_k \geq 2^{(2^k)}$ for all $k \geq 0$.

**Proof.** By induction. Base case: $m_0 = 2 = 2^{(2^0)}$. Step: the two smallest elements of $S_{k-1}$ satisfy $x \geq m_{k-1}$ and $y > m_{k-1}$; hence $m_k > 2m_{k-1}^2 \geq m_{k-1}^2$. Unwinding the recursion: $m_k \geq m_0^{(2^k)} = 2^{(2^k)}$. ■

*Remark 6.2 (A density heuristic, not a theorem).* By Lemma 6.1, $m_k$ grows doubly exponentially. By the Prime Number Theorem, an integer near $N$ is prime with probability approximately $1/\ln(N)$. For a single element $n$ of $S_k$, this probability is at most $1/(2^k \cdot \ln 2)$, which tends to zero. This strongly suggests that remaining prime becomes progressively harder at greater depth. However, converting this into a statement about *all* elements of $S_k$ requires (i) a subexponential-in-$2^k$ upper bound on $|S_k|$, and (ii) independence of events, which the Cramér model heuristically supplies but which does not hold rigorously for the structured sets $S_k$. Neither condition is proven here. We therefore state only the conjecture.

*Conjecture 6.3 (Fractal Prime Extinction).* There exists a finite integer $K \geq 3$ such that $P_k = \emptyset$ for all $k \geq K$. The density heuristic of Remark 6.2 and the unconditional obstruction of Theorem 3.2 support this conjecture; we make no claim of proof. The value of $K$ is unknown, and even determining whether $P_3 = \emptyset$ requires extending the computation of Section 5 to depth 3.

*Remark 6.4 (What a proof would require).* An unconditional proof of Conjecture 6.3 would require one of: (a) a structural algebraic obstruction at every level (analogous to Theorem 3.2, but for all $k \geq 2$); (b) a sieve-theoretic argument supplying a bound of the form $|S_k| = o(2\text{^k})$ in place of the Cramér assumption; or (c) a proof of Conjecture 5.2, which would recursively constrain the structure of $S_k$ and could open a path toward (a). These constitute the most natural directions for future work.

## 7. Generalization to the Eisenstein Integers

The Eisenstein integers $\mathbb{Z}[\omega]$, with $\omega = e^{(2\pi i/3)}$, are equipped with the norm $N(x + y\omega) = x^2 - xy + y^2$. Using the Eisenstein norm in place of the Gaussian norm of Definition 2.1, we define the Eisenstein fractal primes $P_k^{\omega}$.

*Theorem 7.1 (Parity Relaxation).* If $x \equiv y \equiv 1 \pmod 2$, then the Eisenstein norm satisfies $x^2 - xy + y^2 \equiv 1 \pmod 2$. Hence two odd inputs can yield an odd output; this removes the parity bottleneck of Theorem 2.2.

**Proof.** $x^2 - xy + y^2 \equiv 1 - 1 + 1 = 1 \pmod 2$. ■

*Remark 7.2.* Since the parity obstruction is absent in $\mathbb{Z}[\omega]$, the analogue of Theorem 3.2 does not directly hold, and we conjecture that $P_k^{\omega}$ survives to greater depths. Computational study of $P_1^{\omega}$ and $P_2^{\omega}$ has not yet been carried out and would provide valuable comparative data.

## 8. Conclusion and Open Problems

We have proven two unconditional theorems: the characterization of $P_1$ (Theorem 2.2) and the modular obstruction for $P_2$ (Theorem 3.2). Computational data reveals 290 verified elements of $P_2$ and a robust empirical pattern (Proposition 5.1). The extinction heuristic is presented with explicit acknowledgment that it is not a theorem. The following problems remain open:

**(1)** *Is $P_1$ infinite?* Equivalent to the existence of infinitely many primes of the form $q^2 + 4$. An open special case of Landau's fourth problem.

**(2)** *Prove Conjecture 5.2 for $k = 2$.* Show unconditionally that $S_2 \cap \mathbb{P} \subset \{5 \pmod 8\}$. This requires characterizing the mod-8 distribution of the composite elements of $S_1$.

**(3)** *Extend the $P_2$ computation.* The current search covers $S_1 \leq 3200$. Extending to $S_1 \leq 7071$ would cover all $S_2 \leq 10^8$; extending to $S_2 \leq 10^{12}$ via sieve methods would provide strong evidence as to whether $|P_2|$ is finite.

**(4)** *Compute $P_3$ and $P_4$.* This requires sieving $S_2$ and $S_3$ to sufficient range; double-exponential growth makes this computationally expensive but not infeasible at moderate search bounds using parallel sieve methods.

**(5)** *Prove $|S_k| = o(2^k)$, or determine the true growth rate.* This is the missing hypothesis needed to convert Remark 6.2 into a rigorous statement.

**(6)** *Structural obstruction at deeper levels.* Can Theorem 3.2 be extended, conditionally on Conjecture 5.2, to show that $P_k$ faces a modular obstruction for every $k \geq 2$?

**(7)** *Eisenstein comparison.* Compute $P_1^{\omega}$ and $P_2^{\omega}$; determine whether the absence of the parity obstruction leads to measurably deeper survival.

## Acknowledgments

The author thanks the anonymous referees for their careful reading. Mathematical errors identified in successive rounds of review substantially improved the paper. [Additional acknowledgments to be completed.]

## References

[1] P. T. Bateman and R. A. Horn, "A heuristic asymptotic formula concerning the distribution of prime numbers," *Math. Comp.* 16 (1962), 363–367.

[2] H. Davenport, *Multiplicative Number Theory*, 3rd ed. (rev. H. L. Montgomery), Springer, New York, 2000.

[3] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 6th ed., Oxford University Press, Oxford, 2008.

[4] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed., Springer, New York, 1990.

[5] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Teubner, Leipzig, 1909.

[6] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I*, Cambridge University Press, Cambridge, 2007.

[7] P. Ribenboim, *The New Book of Prime Number Records*, 3rd ed., Springer, New York, 1996.

[8] A. Meurer et al., "SymPy: symbolic computing in Python," *PeerJ Comput. Sci.* 3 (2017), e103.