Binary Structure and Density Dynamics in Collatz Sequences

Umar Bin Ayaz
Independent Researcher
blackcoffee2.dev@gmail.com

July 26, 2025

Abstract

We investigate the Collatz conjecture through the lens of binary representations, focusing on Mersenne numbers of the form $2^n - 1$ and their behavior under the Collatz map. We establish several rigorous results: (1) a one-step non-return property showing that Mersenne numbers cannot immediately produce another Mersenne number; (2) exact formulas for bit density evolution from maximum density states; (3) a cycle product constraint linking any hypothetical cycle to the irrationality of $\log_2(3)$; (4) complete characterization of numbers achieving maximum density correction through alternating binary patterns; and (5) a unified classification of all odd numbers into four trajectory categories. Our results reveal that Mersenne numbers and alternating patterns represent extreme strategies for managing the fundamental tension between multiplicative growth and binary divisibility in Collatz dynamics. These structural insights, while not proving convergence, significantly constrain the possible behaviors and suggest new avenues for approaching the conjecture.

1 Introduction

The Collatz conjecture, one of the most notorious unsolved problems in mathematics, posits that iterating the map

$$C(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ 3n+1 & \text{if } n \text{ is odd} \end{cases} \tag{1}$$

eventually reaches 1 from any positive integer [6]. Despite its simple formulation, the problem has resisted solution for over 80 years, leading Erdős to famously state that "mathematics is not yet ready for such problems" [3].

In this paper, we analyze the Collatz problem through binary representations, building on a rich tradition of computational and theoretical approaches. Our focus on bit density and special number classes reveals new structural constraints on Collatz dynamics.

1.1 Related Work

Binary analysis of the Collatz conjecture has a substantial history. Terras [11, 12] pioneered density-theoretic approaches in the 1970s, introducing parity vectors and proving that almost every positive integer has finite stopping time. More recently, Tao [10] achieved a major breakthrough by proving that almost all Collatz orbits attain almost bounded values (in the sense of logarithmic density), meaning the minimal elements of the orbits grow at most logarithmically for almost all starting values, though this falls short of proving the conjecture itself.

The behavior of Mersenne numbers ($2^n - 1$) under Collatz iteration has received limited attention. Ohira and Watanabe [8] studied the path lengths of Mersenne primes computationally, while the OEIS sequence A277109 [7] documents empirical patterns, but rigorous structural results have been lacking.

Several researchers have explored binary patterns in Collatz sequences. Colussi [2] developed concatenation theory for binary representations, while Hew [5] reformulated the problem using dyadic rationals. Recent work by Stérin [9] proved that the Collatz process embeds a base-3 to base-2 conversion algorithm, providing new insights into binary dynamics.

Despite extensive computational verification reaching 2^{71} [1] and various theoretical advances, the conjecture remains open. Lagarias's comprehensive survey [6] notes that the problem appears "completely out of reach of present day mathematics."

1.2 Our Contributions

We establish several new results on binary structure in Collatz sequences:

- A *one-step non-return property*: Mersenne numbers cannot produce another Mersenne number in a single Collatz step (Theorem 2.1)
- A *cycle product constraint* connecting hypothetical cycles to the irrationality of $\log_2(3)$ (Theorem 2.4)
- Exact formulas for bit density evolution from maximum density states (Theorem 3.1)
- Complete characterization of which numbers can achieve maximum density and through which pathways (Theorem 3.3)
- A unified classification of all odd numbers into four trajectory categories (Theorem 6.4)
- Identification of Mersenne numbers and alternating patterns as extreme strategies for managing multiplicative growth versus binary divisibility

1.3 Definitions

Definition 1.1. For a positive integer n with binary representation, the *bit density* $d(n)$ is the ratio of 1-bits to total bits.

Definition 1.2. For odd n, the *Collatz odd-successor* $C_{\text{odd}}(n)$ is the odd number obtained by applying $3n + 1$ and dividing by the largest power of 2.

Definition 1.3. The *Collatz function* is defined as:

$$C(n) = \begin{cases} n/2 & \text{if } n \text{ is even} \\ 3n + 1 & \text{if } n \text{ is odd} \end{cases}$$

Definition 1.4. A *Mersenne number* is a number of the form $2^k - 1$ for $k \geq 1$. When $k \geq 2$, these have all 1s in binary representation.

Definition 1.5. An *alternating pattern number* is an odd integer whose binary representation consists of alternating 1s and 0s, starting and ending with 1.

Definition 1.6. For a positive integer m, $Z(m)$ denotes the number of trailing zeros in the binary representation of m.

2 The Descent Property for Mersenne Numbers

We begin by analyzing the behavior of Mersenne numbers under the Collatz map.

Theorem 2.1 (One-Step Non-Return Property)**.** *If $n = 2^k - 1$ (a Mersenne number) for $k \geq 2$, then $C_{odd}(n)$ is not a Mersenne number.*

Proof. From Lemma 2.2 and Lemma 2.3 below. □

Lemma 2.2. *An odd number p produces a Mersenne number $2^j - 1$ via one Collatz step if and only if $3p + 1 = 2^t(2^j - 1)$ for some $t \geq 1$.*

Proof. This gives $p = \frac{2^t(2^j-1)-1}{3}$. For p to be an odd integer:

- We need $2^t(2^j - 1) \equiv 1 \pmod 3$
- Since $2^j \equiv 1$ or $2 \pmod 3$, we have $2^j - 1 \equiv 0$ or $1 \pmod 3$
- If $2^j - 1 \equiv 0 \pmod 3$, then $2^t \cdot 0 \equiv 0 \not\equiv 1 \pmod 3$
- Thus $2^j - 1 \equiv 1 \pmod 3$, requiring j odd and t even

□

Lemma 2.3. *If $p = 2^k - 1$ for some $k \geq 2$, then p cannot produce another Mersenne number in one Collatz step.*

Proof. If $2^k - 1$ produces $2^j - 1$, then:

$$3(2^k - 1) + 1 = 2^t(2^j - 1) \tag{2}$$

The left side equals $3 \cdot 2^k - 2 = 2(3 \cdot 2^{k-1} - 1)$, which can be written as $2^s \times m$ where $s = 1$ and $m = 3 \cdot 2^{k-1} - 1$ is odd.

From our analysis, $m = 2^k + 2^{k-1} - 1$, which in binary is "10" followed by $(k - 1)$ ones. This is not of the form $2^j - 1$ (all ones), proving the lemma. □

2.1 The Cycle Product Constraint

Theorem 2.4 (Cycle Product Constraint)**.** *For any hypothetical k-cycle of odd numbers $n_1 \to n_2 \to \ldots \to n_k \to n_1$ under C_{odd}: $\prod_{i=1}^{k} \left(3 + \frac{1}{n_i}\right) = 2^{\sum_{i=1}^{k} Z(3n_i+1)}$*

Proof. Each transition satisfies $n_{i+1} = \frac{3n_i+1}{2^{Z(3n_i+1)}}$. For a cycle, the product $\prod_{i=1}^{k} \frac{n_{i+1}}{n_i} = 1$ gives: $\prod_{i=1}^{k} \frac{3n_i+1}{n_i \cdot 2^{Z(3n_i+1)}} = \prod_{i=1}^{k} \frac{3+1/n_i}{2^{Z(3n_i+1)}} = 1$ Therefore: $\prod_{i=1}^{k} \left(3 + \frac{1}{n_i}\right) = 2^{\sum_{i=1}^{k} Z(3n_i+1)}$. □

Remark 2.5. Taking logarithms: $\sum_{i=1}^{k} \log(3 + 1/n_i) = \left(\sum_{i=1}^{k} Z(3n_i + 1)\right) \cdot \log(2)$. Since $\log(3 + 1/n_i) > \log(3)$ and $Z(3n_i + 1)$ are positive integers, this equation requires a specific relationship between sums involving $\log(3)$ and integer multiples of $\log(2)$. The irrationality of $\log_2(3)$ suggests achieving this exact balance may be difficult, though this observation does not constitute a proof that cycles are impossible.

3 Bit Density Evolution

We now analyze how bit density changes under Collatz iteration, extending ideas from [11] and recent binary analysis work [9].

Theorem 3.1 (Density Formula for Mersenne Numbers)**.** *For* $n = 2^k - 1$ *where* $k \geq 2$*,* $d(C_{odd}(n)) = \frac{k}{k+1}$*.*

Proof. From $n = 2^k - 1$:

$$3n + 1 = 3(2^k - 1) + 1 = 3 \cdot 2^k - 2 = 2(3 \cdot 2^{k-1} - 1) \tag{3}$$

$$\text{Thus } C_{\text{odd}}(n) = 3 \cdot 2^{k-1} - 1 \tag{4}$$

We can rewrite: $3 \cdot 2^{k-1} - 1 = 2^k + 2^{k-1} - 1$. In binary, this is "10" followed by $(k-1)$ ones, giving:

- Total bits: $k + 1$
- Total ones: k
- Density: $k/(k + 1)$

□

Theorem 3.2 (No Maximum Density Maintenance)**.** *If* $d(n) = 1$ *for odd* n*, then* $d(C_{odd}(n)) < 1$*.*

Proof. If $d(n) = 1$, then $n = 2^k - 1$. By Theorem 3.1, $d(C_{\text{odd}}(n)) = k/(k + 1) < 1$. □

Theorem 3.3 (Maximum Density Achievement)**.** *For any odd* n*,* $C_{odd}(n)$ *satisfies* $d(C_{odd}(n)) = 1$ *if and only if:*

1. $C_{odd}(n) = 1$*, OR*
2. $C_{odd}(n) = 2^j - 1$ *for some* $j \geq 2$

Proof. If $d(C_{\text{odd}}(n)) = 1$, then $C_{\text{odd}}(n)$ must have all 1s in binary, which means $C_{\text{odd}}(n) = 2^m - 1$ for some $m \geq 1$.

- If $m = 1$, then $C_{\text{odd}}(n) = 1$.
- If $m \geq 2$, then $C_{\text{odd}}(n)$ is a Mersenne number.

By Lemma 2.2, $C_{\text{odd}}(n) = 2^j - 1$ (with $j \geq 2$) occurs when $n = \frac{2^t(2^j-1)-1}{3}$ with j odd and t even.

Conversely, both forms clearly have density 1. □

Remark 3.4. The inability to maintain maximum density and the specific conditions required to achieve it suggest fundamental constraints on Collatz dynamics. As we will see, these constraints are intimately connected to the balance between multiplicative growth and binary divisibility.

4 Maximum Density Correction Mechanisms

Building on binary string analysis approaches [5, 4], we characterize the special cases where density can be maximally restored.

Theorem 4.1 (Maximum Density Correction)**.** *The odd numbers achieving maximum density gain are precisely those where* $3n+1$ *is a power of 2.*

Proof. If $3n+1 = 2^s$, then $n = \frac{2^s-1}{3}$ and $C_{\text{odd}}(n) = 1$, giving $d(C_{\text{odd}}(n)) = 1$. For n to be an odd integer, $2^s \equiv 1 \pmod 3$, which occurs when s is even. Thus $n = \frac{4^j-1}{3}$ for $j \geq 1$. □

Theorem 4.2 (Alternating Pattern Characterization)**.** *The numbers* $n = \frac{4^j-1}{3}$ *have binary representations consisting of alternating 1s and 0s.*

Proof. We have $\frac{4^j-1}{3} = (1 + 4 + 4^2 + \ldots + 4^{j-1})$. In binary, $4^i = 1$ followed by $2i$ zeros. The sum gives the pattern $10101\ldots01$ with length $2j-1$. □

This result connects to Colussi's concatenation theory [2] and provides a complete characterization of the alternating patterns observed empirically.

5 Algebraic Connections

Theorem 5.1 (Direct Path to Alternating Pattern)**.** *The Mersenne number* $3 = 2^2 - 1$ *reaches the alternating pattern* $5 = \frac{4^2-1}{3}$ *in one Collatz odd-step:* $C_{odd}(3) = 5$.

Proof.

- $3 \times 3 + 1 = 10 = 2 \times 5$
- Therefore $C_{\text{odd}}(3) = 5$
- We verify: $5 = \frac{16-1}{3} = \frac{4^2-1}{3}$ ✓
- Furthermore, $C_{\text{odd}}(5) = 1$ since $3 \times 5 + 1 = 16 = 2^4$

This demonstrates the pathway: maximum density → alternating pattern → trivial maximum density. □

6 Structural Insights

Theorem 6.1 (Binary Classification)**.** *The odd positive integers partition into three classes under Collatz dynamics:*

1. *Numbers of form* $2^n - 1$ *(Mersenne numbers - all ones in binary)*
2. *Numbers of form* $\frac{4^j-1}{3}$ *(alternating patterns)*
3. *All other odd numbers*

With the properties:

- *Class 1 numbers have* $d(n) = 1$ *and must decrease density*
- *Class 2 numbers satisfy* $3n+1 = 4^j$*, achieving maximum density correction*
- *Class 3 numbers have varied behavior*

Theorem 6.2 (Structural Constraints). *1. For Mersenne numbers $n = 2^k - 1$ with $k \geq 2$, $C_{odd}(n)$ is not a Mersenne number (follows from Theorem 2.1)*

2. *An odd number p produces a Mersenne number $2^j - 1$ ($j \geq 2$) via C_{odd} if and only if $p = \frac{2^t(2^j-1)-1}{3}$ with j odd and t even*

3. *An odd number n achieves $d(C_{odd}(n)) = 1$ if and only if one of the following holds:*
   - *n is an alternating pattern number (Class 2), satisfying $3n + 1 = 2^s$, thus $C_{odd}(n) = 1$*
   - *$n = \frac{2^t(2^j-1)-1}{3}$ for some odd $j \geq 3$ and even $t \geq 2$, thus $C_{odd}(n) = 2^j - 1$*

Remark 6.3. Numbers like 9, 41, 169 are in Class 3 but have the special property of producing Class 1 numbers (Mersenne numbers) in one step. This reveals that Class 3 is not homogeneous but contains special subsets with remarkable properties.

6.1 Complete Trajectory Classification

Theorem 6.4 (Trajectory Classification). *Every odd positive integer belongs to exactly one category:*

1. ***Alternating pattern generators**: Numbers n where $3n + 1 = 2^s$ for some s (producing $C_{odd}(n) = 1$)*

2. ***Mersenne generators**: Numbers $n = \frac{2^t(2^j-1)-1}{3}$ with j odd, t even (producing $C_{odd}(n) = 2^j - 1$)*

3. ***Other numbers with** $Z(3n + 1) \geq 2$: Not in categories 1 or 2*

4. ***Numbers with** $Z(3n + 1) = 1$: The remaining odd numbers*

Proof. The categories are mutually exclusive by construction. Completeness follows from the fact that $Z(3n+1) \geq 1$ for all odd n, and categories 1 and 2 are the only ways to achieve $d(C_{\text{odd}}(n)) = 1$ (by Theorem 3.3 and Theorem 6.2). □

Remark 6.5. This classification reveals that categories 1 and 2 represent extreme solutions to a fundamental constraint in Collatz dynamics. They are the only trajectories that can achieve maximum bit density in one step, representing boundary cases in the space of possible behaviors.

7 Discussion

7.1 Summary of Results

We have established rigorous constraints on Collatz sequences viewed through binary representations:

- **One-Step Non-Return**: Mersenne numbers cannot produce another Mersenne number in a single Collatz step (proven).
- **Density Dynamics**: Maximum density states (all ones) necessarily lose density. We have completely characterized which numbers can achieve maximum density: either by reaching 1 via alternating patterns, or by producing Mersenne numbers via special formulas.

- **Cycle Constraints**: Any hypothetical cycle must satisfy a product constraint linking the cycle members to their trailing zero counts, with connections to the irrationality of $\log_2(3)$.
- **Complete Classification**: Every odd number falls into one of four trajectory categories based on their ability to achieve maximum density and their trailing zero behavior.

7.2 A Unified Perspective

Our results reveal that the special number classes we identify—Mersenne numbers and alternating patterns—are not arbitrary curiosities but represent extreme strategies for a fundamental problem in Collatz dynamics. Consider that:

1. **Mersenne numbers** ($2^k - 1$) start with maximum density but immediately lose it, dropping to density $k/(k+1)$. They represent the most dramatic density loss possible from a maximum state.
2. **Alternating patterns** ($\frac{4^j-1}{3}$) achieve the opposite extreme: they produce numbers with maximum density (either 1 or Mersenne numbers) while guaranteeing strong divisibility by powers of 2.
3. **The cycle product constraint** shows that any cycle must balance multiplicative growth (factors of approximately 3) with binary divisibility in a precise way that appears incompatible with the irrationality of $\log_2(3)$.

These patterns suggest that Collatz dynamics are governed by a tension between multiplicative expansion and binary contraction, with our special number classes representing boundary cases where this tension is resolved in extreme ways.

7.3 Open Questions

- Can the cycle product constraint and the irrationality of $\log_2(3)$ be leveraged to prove no cycles exist?
- Are there other special number classes with unique roles in Collatz dynamics?
- How does the trajectory classification relate to convergence rates and stopping times?

Our theorems constrain the possible dynamics significantly. While we cannot characterize general density increase conditions due to the complex nature of carry propagation in binary multiplication, we have identified the extreme cases where density can be restored to maximum. This suggests that binary analysis, as advocated by researchers like Stérin [9] and others, continues to offer insights into the structure of Collatz sequences.

References

[1] David Barina. Convergence verification of the Collatz problem. *The Journal of Supercomputing*, 77(3):2681–2688, 2021.

[2] Livio Colussi. The convergence classes of Collatz function. *Theoretical Computer Science*, 412(39):5409–5419, 2011.

[3] Paul Erdős. Personal communication. In Richard K. Guy, editor, *Unsolved Problems in Number Theory*. Springer-Verlag, 1980. As cited in R.K. Guy.

[4] Jishe Feng. Using binary string to prove the Collatz conjecture, 2024.

[5] Patrick Chisan Hew. Collatz on the dyadic rationals with fractals: How bit strings change their length under $3x + 1$. *Experimental Mathematics*, 30(4):490–510, 2019.

[6] Jeffrey C. Lagarias. The $3x + 1$ problem: An annotated bibliography. In *The Ultimate Challenge: The* $3x + 1$ *Problem*, pages 267–341. American Mathematical Society, 2010.

[7] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, A277109. `http://oeis.org/A277109`, 2016.

[8] Yusuke Ohira and Naoki Watanabe. A conjecture on the Collatz-Kakutani path length for the Mersenne primes. Technical report, Research Institute for Mathematical Sciences, Kyoto University, 2011.

[9] Tristan Stérin. The Collatz process embeds a base conversion algorithm. In *Reachability Problems*, Lecture Notes in Computer Science, pages 131–147. Springer, 2020.

[10] Terence Tao. Almost all orbits of the Collatz map attain almost bounded values. *Forum of Mathematics, Pi*, 10:e12, 2022. Originally posted to arXiv in 2019.

[11] Riho Terras. A stopping time problem on the positive integers. *Acta Arithmetica*, 30(3):241–252, 1976.

[12] Riho Terras. On the existence of a density. *Acta Arithmetica*, 35(1):101–102, 1979.