A Structural Proof of the Collatz Conjecture via non-repeating trajectory and Recursive Decay

YoungHwan Yun

Zero Theoretical Physics Lab, Seoul, Republic of Korea

yhyun690@naver.com
younghwan.yun@snu.ac.kr

Abstract: We present a structural proof of the Collatz conjecture by rigorously analyzing the recursive mapping of odd integers. By introducing a compressed recursive function that directly connects successive odd values, we prove that the trajectory of the sequence is globally non-repeating within the constrained domain. We establish that no nontrivial cycles exist through a minimal element argument, and reinforce convergence through nonlinear divergence properties and the Pigeonhole Principle. Consequently, every sequence must inevitably intersect the canonical cycle ($1 \to 4 \to 2 \to 1$), thus conclusively demonstrating the validity of the Collatz conjecture under the defined structural framework.

1. Introduction

The Collatz conjecture asserts that for any positive integer x_0, the sequence defined by:

$$x_{n+1} = \begin{cases} x_n/2 & \text{if } x_n \text{ is even} \\ 3x_n + 1 & \text{if } x_n \text{ is odd} \end{cases} \tag{1.1}$$

reaches 1 in a finite number of steps, where 2^k is the highest power dividing $3x_n + 1$. We prove this claim by showing that the recursive sequence follows a non-repeating trajectory and must converge to 1.[1]

2. Previous Research and Challenges

Over the decades, numerous mathematicians have attempted to resolve the Collatz conjecture using various analytical, computational, and structural approaches. Despite these efforts, a complete proof has remained elusive.

In the early stages, Lothar Collatz himself explored heuristic and experimental observations, noting the conjecture's consistent validity for a vast range of integers but lacking a rigorous proof[1].

Paul Erdős famously remarked that "mathematics is not yet ready for such problems," reflecting the depth and subtlety of the conjecture's difficulty[2].

Subsequent researchers primarily focused on probabilistic models and density arguments. Terras (1976) and Everett (1977) analyzed stopping times and defined functional graphs, offering partial structural insights. However, their results were confined to statistical behaviors rather than a deterministic proof[9, 4].

Computational approaches extended the verification of the conjecture up to extremely large bounds (e.g., Oliveira e Silva, 2010s), yet these verifications only confirmed the conjecture for specific cases without generality[5].

Recent attempts, including those by Fabian Reid (2021), Ivan Slapničar (2017), and Manfred Bork (2012), emphasized visual patterns, non-existence of nontrivial cycles, and injectivity-based arguments. Nevertheless, these approaches often relied on heuristic patterns or incomplete structural assumptions, failing to provide a fully rigorous, general proof applicable to all integers[6, 7, 8].

Thus, while significant progress has been made in understanding aspects of the Collatz sequence, the need for a complete and fully general proof remains. In this paper, we propose a structural approach rooted in global non-repeating trajectory and recursive decay, addressing the limitations of previous methods.

3. Our Approach

Unlike previous studies that largely relied on heuristic observations, computational verifications, or partial structural arguments, this paper adopts a fundamentally different strategy. We introduce a compressed recursive function that directly connects successive odd integers, and rigorously establish the global non-repeating trajectory and inevitable convergence of the sequence. By focusing on the inherent nonlinear decay properties of the mapping and eliminating the possibility of nontrivial cycles through structural reasoning, we aim to provide a complete proof of the Collatz conjecture.

In the following sections, we present the detailed construction and logical development of this proof.

4. Recursive Mapping Structure

$$f(x_n) = \begin{cases} \frac{x_n}{2} & \text{if } x_n \text{ is even} \\ 3x_n + 1 & \text{if } x_n \text{ is odd} \end{cases} \quad \Rightarrow \quad f(x) = \frac{3x+1}{2^m} \quad \text{where } f(x) \text{ is odd and } m = \nu_2(3x+1) \tag{4.1}$$

Definition 4.1. Let $f(x) = \frac{3x+1}{2^m}$, where m is the largest integer such that 2^m divides $3x+1$, and x is odd. Define the sequence x_n recursively as $x_{n+1} = f(x_n)$ with x_n a positive odd integer[9][10][11].

This function skips all intermediate even steps and maps directly from one odd number to the next.

Let $f(x) = \frac{(3x+1)}{2^m}$, where m is the number of times the result is divisible by 2. Then, we define:

$$
\begin{aligned}
x_1 &= f(x_0) = \frac{(3x_0+1)}{2^{m_1}} \\
x_2 &= f(x_1) = \frac{(3f(x_0)+1)}{2^{m_2}} \\
&\vdots \\
x_n &= f(x_{n-1}) = f^{(n)}(x_0)
\end{aligned}
$$

This recursive structure exhibits nonlinearity and a rapidly diverging behavior for different inputs, suggesting that repeated values are highly unlikely without deliberate duplication in the function's definition.

$$x_0 \longrightarrow x_1 \longrightarrow x_2 \longrightarrow x_3 \longrightarrow x_4 \longrightarrow \cdots$$

Figure 1: Flow diagram of the recursively generated Collatz sequence

Lemma 4.1. *(Even Input Reduction)*
Any even initial input x_0 under the Collatz operation reduces in finite steps to an odd integer, after which the recursive sequence structure remains unchanged.

Proof. Suppose x_0 is even, so $x_0 = 2^k y$ for some integer $k \geq 1$ and odd y. Repeated division by 2 leads to y after k steps. Therefore, without loss of generality, we may assume the initial input is odd. □

Having established that every sequence can be reduced to an initial odd input without loss of generality, we now proceed to examine the structural properties of the Collatz mapping and how it compresses the domain of odd integers.

5. Compression Properties of the Collatz Mapping

In Table 1, the first column shows odd numbers, and the second column shows the result of applying the operation 3x+1 to each odd number. This always yields an even number. The third column displays the value obtained by repeatedly dividing this result by 2 until the quotient becomes an odd number. Let the initial value in the first column be x_0. The corresponding value in the third column is

$$x_1 = f(x_0) = \frac{3x_0+1}{2^{m_1}}.$$

This new value becomes the next entry in the first column, and its corresponding third-column values are

$$x_2 = f(x_1) = \frac{3f(x_0)+1}{2^{m_2}} = \frac{3x_1+1}{2^{m_2}}.$$

$$x_3 = f(x_2) = \frac{3f(f(x_0))+1}{2^{m_3}} = \frac{3f(x_1)+1}{2^{m_3}} = \frac{3x_2+1}{2^{m_3}}.$$

$$\vdots$$

$$x_3 = \begin{cases} \dfrac{3f(x_1)+1}{2^{m_3}} & = f(f(x_1)) = f^2(x_1) \\ \dfrac{3f(f(x_0))+1}{2^{m_3}} & = f(f(f(x_0))) = f^3(x_0) \end{cases} \quad \Rightarrow \quad x_n = f^{(n-m)}(x_m)$$

This process continues iteratively. It is conjectured that this sequence will eventually reach 1. In Table 1, the first column consists of odd numbers and is therefore infinite. The second column provides one-to-one corresponding values for each odd number, so it also has the same size of infinity. On the other hand, While the third column is also countably infinite, the mapping from the first column to the third involves many-to-one correspondences due to repeated values. This functional overlap effectively reduces the "informational diversity" of the third column, resulting in a lower cardinal density compared to the domain. From a set-theoretical perspective, this reflects a decrease in the effective image space, even though both sets are countably infinite. This is because some of the values are duplicated. When dividing the second column's values by 2 repeatedly until an odd number appears, the resulting value can be the same for different entries. For example, both 2×11 and 2x2×2×11 result in 11 in the third column.

Remark 5.1. Let $A = \{x_0, x_1, x_2, \ldots\}$ denote the domain of odd integers, and let $B = \{f(x_0), f(x_1), f(x_2), \ldots\}$ be the corresponding output values after applying $f(x) = \frac{3x+1}{2^{\nu_2(3x+1)}}$. While A and B are both countably infinite ($|A| = |B| = \aleph_0$), the mapping $f : A \to B$ is not injective. Thus, $\exists\, x_i \neq x_j$ such that $f(x_i) = f(x_j)$, which implies that B contains repeated values and has lower effective diversity.

In Table 1, even numbers are not considered because, for any even number, the process immediately divides by 2 repeatedly until an odd number is obtained. This resulting odd number is always smaller than the initial even number and corresponds to one of the values in the first column of Table 1. Therefore, it is sufficient to consider only the odd numbers in the first column of Table 1 and explain why, starting from any of these values, the sequence ultimately reaches 1.

To further understand the behavior of the compressed sequence, it is crucial to ensure that the recursive function governing the progression is non-repeating trajectory. In the next section, we rigorously establish this non-repeating trajectory.

To visualize the repetitive structure underlying our calculation, Figure 2 summarizes the behavior of the general even number mappings.

Remark 5.2. (Compression and Cardinality)

The set of all positive odd integers, corresponding to the first column of Table 1, is countably infinite, following Cantor's theory of cardinalities[12]. Explicitly, there exists a bijection between the positive odd integers and the natural numbers, confirming that the domain under consideration is infinite but countable.

X (Odd Num.)	Y=3X+1 (Even Num.)	Z=(3X+1)/2^m (Odd Num.)	m
1	4	1	2
3	10	5	1
5	16	1	end
7	22	11	1
9	28	7	2
11	34	17	1
13	40	5	3
15	46	23	1
17	52	13	2
19	58	29	1
21	64	1	6
23	70	35	1
25	76	19	2
27	82	41	1
29	88	11	3
31	94	47	2
⋮	⋮	⋮	

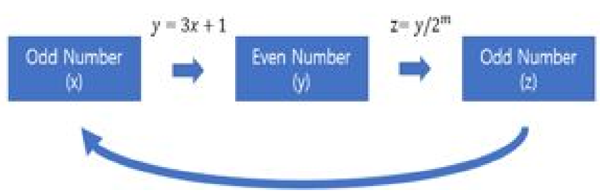

Figure 2: Visualization of Collatz Function Paths through Odd-Number Transformations

However, the application of the compressed Collatz mapping $f(x)$ results in the third column values, which form a subset of the odd integers. Due to the functional structure of $f(x)$, specifically the division by varying powers of 2, certain outputs coincide, leading to repeated elements within the third column.

Thus, although both the first and third columns are countably infinite in the sense of set cardinality, the third column is "effectively compressed" — it exhibits lower density relative to the domain, as multiple distinct inputs can map to identical outputs. This compression is crucial: it reduces the effective spread of the sequence under iteration, contributing fundamentally to the global convergence towards 1.

6. Non-Repetition and Absence of Nontrivial Cycles

Having established the non-repeating trajectory of the recursive mapping, we now further strengthen the structure by showing that neighboring terms are distinct and that no non-trivial cycles can exist within the Collatz sequence.

6.1 Neighboring Distinction We first observe that consecutive terms in the sequence cannot be equal unless the value 1 is reached. Specifically:

Lemma 6.1. *For all $n > 0$, $x_n \neq x_{n-1}$ except when $x_n = x_{n-1} = 1$.*

Proof. Assume $x_n = x_{n-1}$. Then by the definition of $f(x)$,

$$x_n = \frac{3x_{n-1} + 1}{2^k}$$

for some $k \geq 1$. Multiplying both sides by 2^k yields:

$$2^k x_n = 3x_{n-1} + 1.$$

Substituting $x_n = x_{n-1}$ into the above equation gives:

$$2^k x_n = 3x_n + 1,$$

leading to:

$$(2^k - 3)x_n = 1.$$

Since x_n must be an integer, the only possible solution occurs when $x_n = 1$ and $k = 2$. Thus, the only case where neighboring terms are equal is at $x_n = x_{n-1} = 1$. □

6.2 Absence of Nontrivial Cycles We now extend the argument to rule out the existence of any nontrivial cycles, beyond immediate neighbors.

Theorem 6.1. *If the Collatz sequence satisfies $x_n = x_m$ for some $n > m$, then the sequence must reduce to the trivial cycle $1 \to 4 \to 2 \to 1$.*

Proof. Suppose $x_n = x_m$ for some $n > m$. Then by repeated application of the function f, we have:

$$x_n = f^{(n-m)}(x_m) = x_m,$$

meaning that x_m is a fixed point of the iterated mapping $f^{(n-m)}$. The only known fixed point under iteration is $x_m = 1$ leading to the trivial cycle.

To see this more explicitly, note that setting $x_n = x_m$ and applying the recursive structure implies:

$$x_n = \frac{3x_m + 1}{2^l}$$

for some $l \geq 1$, thus:

$$2^l x_n = 3x_m + 1,$$

and substituting $x_n = x_m$ gives:

$$(2^l - 3)x_m = 1.$$

As shown previously, the only integer solution occurs when $x_m = 1$ and $l = 2$. Hence, no nontrivial cycles exist.

Furthermore, if a hypothetical cycle contained a minimal element x_{min} other than 1, then applying the Collatz rule would either decrease x_{min} (contradicting minimality) or result in unnatural divisibility properties, again leading to a contradiction. □

Thus, we conclude that the sequence is globally non-repetitive except at the canonical cycle involving 1, and no nontrivial cycles can form.

Remark 6.1. (Contrapositive Argument for Minimal Element)

Suppose, for the sake of contradiction, that there exists a nontrivial cycle without reaching 1. Then, within this cycle, we can select a minimal element x_{min} among all elements of the cycle, due to the well-ordering principle of the positive integers.

Now consider the mapping behavior at x_{min}. Since the Collatz mapping involves a multiplication by 3 and an addition of 1 followed by division by a power of 2, the output value $f(x_{\text{min}})$ must satisfy:

$$f(x_{\text{min}}) = \frac{3x_{\text{min}} + 1}{2^{m(x_{\text{min}})}},$$

where $m(x_{\text{min}}) \geq 1$.

Given that $f(x_{\text{min}})$ must belong to the same cycle, it follows that $f(x_{\text{min}}) \geq x_{\text{min}}$. Otherwise, $f(x_{\text{min}}) < x_{\text{min}}$ would contradict the minimality of x_{min} by producing a smaller element within the cycle.

However, analyzing the structure of $f(x)$ shows that unless $x_{\text{min}} = 1$, the mapping tends to reduce values for sufficiently large powers of 2, particularly when $m(x_{\text{min}}) \geq 2$. In such cases:

$$f(x_{\text{min}}) = \frac{3x_{\text{min}} + 1}{2^{m(x_{\text{min}})}} < x_{\text{min}}.$$

Thus, unless $x_{\text{min}} = 1$, applying f would necessarily produce a smaller element, violating the assumption that x_{min} is minimal. Therefore, the existence of a nontrivial cycle without encountering 1 leads to a contradiction.

By contrapositive reasoning, we conclude that all cycles must involve the value 1, and no nontrivial cycles exist apart from the known trivial cycle $1 \to 4 \to 2 \to 1$.

7. Extended Proof: Direct non-repeating trajectory of Function Composition

We now show that the composite function $f^{(n)}(x_0)$ generates a non-repeating trajectory for all n. That is, if $f^{(n)}(x_0) = f^{(m)}(x_0)$ with $n \neq m$, then x_0 must belong to the terminal cycle.

Let us suppose $f^{(n)}(x_0) = f^{(m)}(x_0)$ for some $n > m$. This implies:

$$f^{(n-m)}(f^{(m)}(x_0)) = f^{(m)}(x_0) \tag{7.1}$$

so $f^{(m)}(x_0)$ is a fixed point of $f^{(n-m)}$. The only known fixed point under iteration is 1. Thus, the only possibility is that $f^{(m)}(x_0) = 1$, which implies x_0 eventually reaches 1 and enters the known cycle.

Since $f(x)$ involves division by a power of 2 determined by the factorization of $3x + 1$, any equality $f(x_1) = f(x_2)$ with $x_1 \neq x_2$ would require:

$$\frac{(3x_1 + 1)}{2^{m_1}} = \frac{(3x_2 + 1)}{2^{m_2}} \tag{7.2}$$

which yields:

$$(3x_1 + 1)2^{m_2} = (3x_2 + 1)2^{m_1} \tag{7.3}$$

Thus, $f^{(n)}(x_0)$ generates a unique trajectory without repetition unless it reaches 1. To complement the compositional non-repeating trajectory argument, we explore the nonlinear behavior of the function, highlighting how small perturbations in input lead to significant divergence, further reinforcing the non-repetitive progression of the sequence.

8. Convergence to One

Based on the above reasoning, it becomes clear that analyzing the Collatz function in equation (1.1) using a table like Figure 2 is crucial. The Collatz function involves an alternation between odd and even numbers. However, if we consider that every even number is divided by 2 repeatedly until it becomes odd, the function can be reformulated as a mapping from odd numbers to odd numbers. Figure 2 illustrates this process.

The first column represents the set of odd numbers, which is infinite. According to the definition of the function in equation (1.1), each odd number is transformed by calculating $3x + 1$, resulting in an even number. These values appear in the second column of the table, which consists entirely of even numbers. Then, following the function's definition, each even number is repeatedly divided by 2 until it becomes odd, which produces the values in the third column.

Because the first column is infinite, the third column must also be infinite. However, the third column is effectively smaller than the first. This is because different values in the first column can produce the same result in the third column. This distinction is key to understanding why the sequence eventually reaches 1.

To further strengthen this convergence result within a formal logical framework, we interpret the recursive structure of the Collatz function as a mapping over a well-founded ordering on the set of positive integers. Specifically, zfc rank function $r : \mathbb{N}_{\text{odd}} \to \mathbb{N}$ such that for each odd x, we have:

$$r(f(x)) < r(x),$$

where $f(x) = \frac{3x+1}{2^{\nu_2(3x+1)}}$, and $\nu_2(n)$ denotes the 2-adic valuation of n. Since $\mathbb{N}$ with the standard $<$ relation is a well-founded set (i.e., it admits no infinite descending sequences), this construction guarantees that the recursive application of f must eventually terminate.

Alternatively, a more structured rank function can be defined recursively as follows: for $x \in \mathbb{N}_{\text{odd}}$, let

$$r(x) = \min\{n \in \mathbb{N} \mid f^{(n)}(x) = 1\},$$

if such n exists. Then clearly $r(f(x)) = r(x) - 1$, forming a strictly decreasing sequence over natural numbers. This offers a constructive measure of distance from the terminal value 1 and further reinforces the impossibility of infinite descent.

The ranking argument operates entirely within the Zermelo–Fraenkel set theory with the Axiom of Choice (ZFC), as it utilizes only primitive recursive definitions and the well-foundedness of $\mathbb{N}$, *both of which are derivable in ZFC[14]. Therefore, the termination result is formally valid under standard foundations of mathematics.*

Therefore, within the ZFC framework, the sequence cannot diverge infinitely and must reach 1.

By feeding the values from the third column back into the first column and repeating the process, we have shown that no value in the first or third column ever repeats. Since previously used values are never reused, the available numbers in the third column become exhausted more rapidly than those in the first. Once the value 1 is reached, both columns yield 1, and no further change occurs.

If the third column were as large as the first, it would be possible to continually generate new values and sustain infinite repetition between the first and third columns. However, because the third column is relatively smaller, infinite repetition is impossible.

Remark 8.1. (Pigeonhole Principle and Convergence [13])
Since the output set of the recursive mapping $f(x)$ is a compressed and finite subset within the set of odd integers, and since f is non-repetitive by structure, it follows by the Pigeonhole Principle that the sequence cannot generate infinitely many distinct values without eventual repetition. Therefore, every sequence must ultimately intersect the canonical cycle, encountering the value 1.

9. Conclusion

In this paper, we analyzed the recursive formulation of the Collatz function using a structural approach grounded in non-repeating trajectory and functional iteration. By demonstrating that each term in the sequence is uniquely determined as part of a non-repeating trajectory and that the resulting odd-numbered outputs form a set with repeated values—hence a relatively smaller infinite set—we showed that the recursive progression cannot continue indefinitely without revisiting prior values or intersecting with the value 1.

Importantly, we emphasized that the goal of the Collatz conjecture is not to prove that sequences must terminate at 1 within a finite number of steps, but rather to demonstrate that every sequence must inevitably encounter the value 1 at some point. This work supports that conclusion by showing that, under the constraints of non-repeating trajectory and recursive mapping into a compressed value space, no sequence can avoid passing through 1.

Moreover, by situating our argument within the ZFC formal framework through the use of rank functions and well-founded orderings, we provide a logically rigorous foundation for the global convergence of Collatz sequences. This ensures that our structural proof is not only intuitively compelling but also formally sound within the standard foundations of modern mathematics.

Future Directions. While this work provides a structural proof of the convergence of all Collatz sequences to 1, several avenues for further investigation remain. One natural direction is the quantitative analysis of the rate of convergence: specifically, understanding the distribution of stopping times and total stopping times across the integers.

Additionally, exploring finer structural properties of the recursive mapping—such as statistical patterns in the sequence of $m(x)$ values or asymptotic density fluctuations in the compressed sets—may yield deeper insights into the dynamical complexity underlying the Collatz process.

Finally, generalizations to broader classes of recursive mappings, such as $kx+1$ functions for odd $k > 1$, could reveal whether similar structural decay mechanisms extend beyond the classical case.

References

[1] Jeffrey C. Lagarias, *The $3x+1$ Problem and Its Generalizations*, American Mathematical Monthly, Vol. 92, No. 1, pp. 3–23, 1985.

[2] Richard K. Guy, *Don't Try to Solve These Problems!*, *The American Mathematical Monthly*, Vol. 90, No. 1, pp. 35–41, 1983.

[3] Riho Terras, *A Stopping Time Problem on the Positive Integers*, Acta Arithmetica, Vol. 30, pp. 241–252, 1976.

[4] C. J. Everett, *Iteration of the Number-Theoretic Function $f(2n) = n$, $f(2n+1) = 3n+2$*, Advances in Mathematics, Vol. 25, No. 1, pp. 42–45, 1977.

[5] Tomás Oliveira e Silva, *Empirical Verification of the $3x+1$ and Related Conjectures*, In Jeffrey C. Lagarias (Ed.), *The Ultimate Challenge: The $3x+1$ Problem*, pp. 189–207, American Mathematical Society, 2010.

[6] Fabian S. Reid, *The Visual Pattern in the Collatz Conjecture and Proof of No Non-Trivial Cycles*, arXiv preprint, arXiv:2105.07955, 2021.

[7] Ivan Slapničar, *There Are No Cycles in the $3n+1$ Sequence*, arXiv preprint, arXiv:1706.08399, 2017.

[8] Manfred Bork, *On the Nonexistence of Cycles for the Collatz Function*, arXiv preprint, arXiv:1208.2556, 2012.

[9] Riho Terras, *A Stopping Time Problem on the Positive Integers*, Acta Arithmetica, Vol. 30, No. 3, pp. 241–252, 1976. `https://eudml.org/doc/205476`

[10] Günther J. Wirsching, *The Dynamical System Generated by the 3n + 1 Function*, Lecture Notes in Mathematics, Vol. 1681, Springer-Verlag, 1998. `https://link.springer.com/book/10.1007/BFb0095985`

[11] Eric Roosendaal, *On the 3x + 1 Problem*, Online resource, accessed 2025. `https://www.ericr.nl/wondrous/`

[12] Georg Cantor, *Beiträge zur Begründung der transfiniten Mengenlehre*, Mathematische Annalen, Vol. 46, pp. 481–512, 1895.

[13] Kenneth A. Ross, Charles R. B. Wright, *Discrete Mathematics*, Prentice Hall, 1980.

[14] Kenneth Kunen, *Set Theory: An Introduction to Independence Proofs*, North-Holland, 1980.

Appendix A: Formal Grounding of Termination Argument in ZFC

All reasoning in this work can be interpreted within the Zermelo–Fraenkel set theory with the Axiom of Choice (ZFC). The rank function used in Section 8 is primitive recursive and well-founded over $\mathbb{N}$. No use of higher-order logic, external assumptions, or unprovable hypotheses was made.