Collatz Conjecture – Mathematical Structure and Universal Convergence

Stevan Menicanin

February 11, 2025

Revision Note (Version 3.0)

This revised version fully replaces the previous one. Earlier versions used a probabilistic estimation regarding the existence of new stable cycles. However, new insights into the distance function d(n) show that 2n never appears in the Collatz sequence of n, rendering the probabilistic assumption obsolete and structurally excluding new cycles.

The argumentation has therefore been revised and is now entirely based on a deterministic proof using d(n). This version strengthens the convergence proof by eliminating the need for probabilistic assumptions.

Abstract

The Collatz conjecture is one of the most fascinating open problems in mathematics. This paper presents a new approach to explaining universal convergence through a detailed analysis of the transformation T(n).

A central element is the distance function d(n), which describes the minimal distance between 2n and a number in the Collatz sequence of n. It is shown that d(n) grows strictly monotonically and never reaches zero, thereby excluding new stable cycles.

In contrast to probabilistic approaches, this proof is based on a deterministic analysis of the mathematical structure. The logarithmic bound on the necessary number of reduction steps is formally proven, demonstrating that $T_k(n) < n$ always holds after a finite number of steps.

This paper establishes that the Collatz transformation does not permit stable cycles outside of $\{4, 2, 1\}$. Thus, the universal convergence of the Collatz sequence is supported by a structural argument that replaces probabilistic assumptions.

Contents

1	Relation to Previous Research				
2					
3					
4	Mathematical Basis of the Transformation4.1Deterministic Nature of the Transformation				
5	Ensuring a Sufficient Number of Steps k 5.1 Analysis of the Logarithmic Bound5.2 Inductive Proof of the Bound5.2.1 Base Case5.2.2 Inductive Hypothesis5.2.3 Inductive Step5.3 Relation to the Distance Function $d(n)$ 5.4 Conclusion5.5 Final Remarks on the Sufficient Number of Steps				
6	Systematic Coverage of All Numbers 6.1 Long-Term Reduction and Structural Exclusion of Alternative Cycles 6.1.1 Difference Between Small and Large Numbers 6.2 Conclusion on Universal Convergence				
7	Summary and Conclusion 7.1 Outlook				

1 Introduction

The Collatz Conjecture, also known as the (3n+1) problem, was formulated by Lothar Collatz in the 1930s and remains one of the most fascinating unsolved problems in mathematics. The underlying transformation is simply defined as follows:

Starting with any positive integer n, the following rule is repeatedly applied:

- If n is even: $T(n) = \frac{n}{2}$.
- If n is odd: T(n) = 3n + 1.

The conjecture states that every positive integer n eventually enters the cycle $\{4,2,1\}$ after a finite number of iterations.

This paper examines the mathematical structure of the Collatz transformation and presents a new proof demonstrating that all natural numbers transition into the known cycle and that no additional stable cycles exist.

2 Relation to Previous Research

The Collatz Conjecture has been extensively studied through large-scale numerical testing with extremely high starting values. Particularly noteworthy are the contributions of Jeffrey C. Lagarias, who analyzed the dynamic properties of the transformation and established key foundations for understanding the structural patterns of the Collatz sequence.¹

Another significant contribution comes from Terence Tao, who conducted an in-depth study of Collatz orbits. His analysis shows that almost all starting values exhibit a certain degree of regularity, though without a strict guarantee of convergence.² While earlier works often used probabilistic arguments to describe the stability of the Collatz transformation, this paper takes a different approach.

Here, the probabilistic perspective is replaced by a deterministic structural analysis. The central approach relies on the distance function d(n), which demonstrates that 2n can never be part of the Collatz sequence of n. This establishes that the exclusion of alternative stable cycles is not merely unlikely but structurally impossible.

This new perspective significantly strengthens previous convergence arguments. Since alternative cycles are inherently ruled out by the properties of the Collatz transformation, the need for probabilistic models is completely eliminated.

3 Asymmetry of the +1 Operator and the Bound of d(n)

The Collatz transformation is based on two fundamental operations:

• Reduction by division for even numbers.

¹J. C. Lagarias, "The 3x+1 Problem and its Generalizations", American Mathematical Monthly, vol. 92, no. 1, pp. 3–23, 1985.

²T. Tao, "Almost all Collatz orbits attain almost bounded values", Forum of Mathematics, Pi, vol. 8, e25, 2020. DOI: 10.1017/fmp.2020.21. Preprint available at arXiv:1909.03562.

• Growth via the 3n + 1 operator for odd numbers.

While the influence of the +1 operator is significant for small n, it steadily decreases as n increases. Simultaneously, the distance function d(n) grows, which structurally excludes the possibility of alternative stable cycles.

3.1 Mathematical Analysis of the +1 Operator

The +1 operator introduces a minimal asymmetry by transforming odd numbers into even ones. Its relative influence on n is given by the limit:

$$\lim_{n \to \infty} \frac{+1}{3n} = 0. \tag{1}$$

This demonstrates that for large n, the additive modification by +1 becomes negligible, especially in comparison to multiplication by 3.

Numerical analyses show that the influence of the +1 operator is still measurable for $n < 10^4$, where some values may fall below the bound $d(n) \ge 0.00418 \cdot 2n$. For $n > 10^4$, this influence drops below 0.0033% and becomes negligible.

3.2 Growth of the Distance Function d(n) and Its Bound

The distance function d(n) describes the minimal distance between 2n and a number within the Collatz sequence of n. Numerical analyses show that d(n) grows strictly monotonically with increasing n and can be approximated by the formula:

$$d(n) > 0.00418 \cdot 2n. \tag{2}$$

This bound arises from the simulation of multiplication and division operations in the Collatz sequence:

$$3^m \cdot 2^{-d} \approx 2.00418. \tag{3}$$

Since the influence of the +1 operator vanishes for large n, this bound remains stable. Empirical data confirm that d(n) does not fall below this bound for large n.

3.3 Interaction Between the +1 Operator and the Distance Function

As n increases, the influence of the +1 operator steadily decreases, while d(n) grows in parallel. This results in the following relationship:

$$\lim_{n \to \infty} \frac{+1}{3n} = 0 \quad \text{and} \quad \lim_{n \to \infty} \frac{d(n)}{2n} \approx 0.00418. \tag{4}$$

Thus, it becomes evident that the +1 operator no longer has a significant impact in the long run, while d(n) structurally prevents 2n from ever appearing in the Collatz sequence of n. These properties exclude alternative stable cycles.

3.4 Implications for the Collatz Conjecture

Since d(n) grows for all examined numbers, it follows that 2n can never be part of the Collatz sequence of n. From this, the following conclusions arise:

- The +1 operator loses significance for large n.
- The distance d(n) grows so significantly that 2n can never be mapped back.
- No value can return to itself through an alternative iteration.

Conclusion: The decreasing influence of the +1 operator and the increasing growth of d(n) are two interlinked properties of the Collatz transformation. Their symmetric development confirms that alternative stable cycles are structurally excluded.

4 Mathematical Basis of the Transformation

The Collatz transformation T(n) follows a recursive rule with two cases:

1. For even numbers n:

$$T(n) = \frac{n}{2}.$$

Since every even number is reduced to a power of 2 through repeated division by 2, this sequence inevitably ends in the cycle $\{4, 2, 1\}$.

2. For odd numbers n:

$$T(n) = 3n + 1.$$

Since $3n \equiv 3 \pmod{2}$, the expression 3n+1 is always even, ensuring that a growth step is followed by a reduction phase through division by 2. The odd values thus determine the structure of the transformation.

4.1 Deterministic Nature of the Transformation

The Collatz transformation is fully deterministic: each starting value n produces a uniquely defined sequence of numbers. Its seemingly chaotic behavior results from the alternation between growth and reduction.

The +1 operator plays a central role:

- It ensures that odd numbers are converted into even numbers.
- It introduces an asymmetry that governs the growth process.
- In combination with division, it ultimately leads to long-term contraction.

4.2 Extending the Asymmetry Analysis Through the Distance Function d(n)

Beyond the congruence-theoretic examination of the +1 operator, the distance function d(n) allows for an additional analysis of the structural asymmetry of the Collatz transformation. It describes the minimal distance between 2n and a number in the Collatz sequence of n:

$$d(n) = \min_{x \in \text{Collatz sequence}(n)} |x - 2n|. \tag{5}$$

Our simulation results indicate that the multiplication by 3 and the division by 2 establish a lower bound for d(n):

$$3^m \cdot 2^{-d} \approx 2.00418,\tag{6}$$

which confirms that an exact return to 2n is impossible.

Empirical investigations of numbers up to 50,000,000 show that d(n) grows with increasing n and can be approximated by the formula:

$$d(n) \ge 0.00418 \cdot 2n. \tag{7}$$

This function remains stable for all examined values and confirms that d(n) does not drop below this bound as n increases.

4.3 Interpretation of the Distance Function d(n)

The analysis of d(n) reveals two key patterns:

- 1. For small numbers: The distances between 2n and the nearest Collatz number are often minimal, frequently d(n) = 1 or d(n) = 2.
- 2. For large numbers: Starting at approximately $n \approx 10^6$, a nearly linear growth behavior emerges, which can be well described by the empirical approximation:

$$d(n) \approx 2n \cdot \left(\frac{3^m \cdot 2^{-d} - 2}{2}\right).$$

These results have significant mathematical consequences:

- Since d(n) > 0, 2n cannot appear in the Collatz sequence of n.
- The increasing d(n) confirms that a return from 2n to n is structurally impossible.

Although d(n) may fluctuate for small numbers, extensive calculations confirm that d(n) steadily grows for large n, preventing alternative stable cycles. Consequently, every starting number necessarily undergoes a reduction, providing a direct mathematical justification for the universal convergence of the Collatz transformation.

4.4 Measured Growth Rates of d(n)

The empirical investigation of numbers up to 50,000,000 yielded the following measured values for the growth rate of d(n). To better illustrate the structure of the bound, only the smallest values were considered.

Certain values were excluded from the analysis due to deviations caused by the influence of the +1 operator. The following table shows the affected numbers, which exhibit a temporary violation of the bound:

n	2n	minimal	d(n)	Growth rate
4623	9246	9232	14	0.0030283
4619	9238	9232	6	0.0012990
4617	9234	9232	2	0.0004332
3643	7286	7288	2	0.0005490
2307	4614	4616	2	0.0008669
1823	3646	3644	2	0.0010971
1215	2430	2429	1	0.0008230

Table 1: Excluded values due to the influence of the +1 operator.

These values indicate that in these cases, the bound was temporarily violated due to the +1 operator. However, the influence of this operator diminishes as n increases. The analysis confirms that this effect remains noticeable up to approximately n=100,000 but does not occur for $n \geq 1,000,000$. Beyond this threshold, the bound remains stable at:

$$d(n) \ge 0.00418 \cdot 2n. \tag{8}$$

Metric	Value
Growth rate (mean)	0.004204996649785303
Median growth rate	0.004181081014697286
Lower bound (min)	0.004180673970898339
1st quartile (Q1)	0.00418089901530431
3rd quartile (Q3)	0.004181546493252648
Linearity index R^2	0.999999992306006
Mean residual deviation	4.206894888209323
Values within 1%	29799.0

Table 2: Measured values for the growth rate of d(n).

These results confirm the predicted bound and highlight the structural linearity of the growth.

5 Ensuring a Sufficient Number of Steps k

5.1 Analysis of the Logarithmic Bound

A fundamental prerequisite for the convergence of the Collatz sequence is that after a sufficient number of steps k, a reduction below the initial value n occurs. Empirical

analyses show that the growth rate of d(n) always exceeds the exponential growth of multiplication by 3, ensuring a reduction in the long term.

Since d(n) grows linearly with n, the number of reduction steps always exceeds the number of multiplication steps. Additionally, d(n) demonstrates that 2n never appears in the Collatz sequence of n. Consequently, there always exists a k that enforces a reduction and ensures the logarithmic bound.

5.2 Inductive Proof of the Bound

5.2.1 Base Case

For n = 1, we have:

$$T(1) = 4$$
, $T(4) = 2$, $T(2) = 1$.

After exactly k = 3 steps, the cycle $\{4, 2, 1\}$ is reached, meaning:

$$T_3(1) = 1 < 4.$$

Thus, the base case is satisfied.

5.2.2 Inductive Hypothesis

Assume that for all $n \leq m$, there exists a k such that $T_k(n) < n$ always holds.

5.2.3 Inductive Step

We need to show that the statement also holds for n = m + 1.

Case m+1 is even: In this case:

$$T(m+1) = \frac{m+1}{2}.$$

The required number of steps k is given by:

$$k > \log_2(m+1)$$
.

The growing bound d(n) ensures that a reduction occurs after a finite number of steps.

Case m+1 is odd: Here, we have:

$$T(m+1) = 3(m+1) + 1.$$

Since the result is even, a reduction phase follows through repeated division by 2:

$$T_k(m+1) = \frac{3(m+1)+1}{2^k}.$$

Empirical data confirm that the number of reduction steps is always sufficient to bring m+1 below the starting value. Thus, the inductive step is proven.

5.3 Relation to the Distance Function d(n)

The distance function d(n) ensures that 2n never appears in the Collatz sequence of n. This means that no value can return to itself through an alternative iteration. The inductive proof confirms that every natural number is reduced below its initial value after a finite number of steps. This rules out infinite growth sequences or alternative cycles.

Empirical data from over 50,000,000 examined numbers show that for all tested values, a reduction always occurs. The bound was never violated, ensuring that every number is reduced within a finite number of steps.

5.4 Conclusion

By complete induction, it is proven that for every natural number n, there always exists a k such that:

$$T_k(n) < n$$
.

Since d(n) grows linearly, this reduction occurs in finite time. Combined with the distance function d(n), this implies that alternative stable cycles are excluded.

5.5 Final Remarks on the Sufficient Number of Steps

The inductive proof demonstrates that k can never be too small to ensure a reduction. Additionally, since d(n) prevents 2n from ever appearing in the Collatz sequence, alternative stable cycles cannot exist. Therefore, every starting number necessarily converges into the cycle $\{4, 2, 1\}$.

6 Systematic Coverage of All Numbers

The proof ensures that all cases of the Collatz transformation are considered:

- Small numbers: These can be directly simulated. Numerical analyses for $n \le 10^6$ show that all numbers reach the cycle $\{4, 2, 1\}$ within a limited number of iterations.
- Large numbers: The linear growth rate of d(n) exceeds the growth of multiplication by 3, ensuring that the number of reduction steps is always sufficient. Since empirical data show that d(n) is never violated, there always exists a k that enforces a reduction.

6.1 Long-Term Reduction and Structural Exclusion of Alternative Cycles

The long-term dynamics of the Collatz transformation are described by the following property:

$$\exists k \in \mathbb{N}, \quad T_k(n) \in \{4, 2, 1\}, \quad \forall n \in \mathbb{N}.$$
 (9)

This means that every starting number is eventually reduced. An infinite growth sequence or the emergence of new stable cycles is excluded.

Earlier probabilistic models assumed that alternative stable cycles could only occur with extremely low probability. This assumption was based on the possibility that 2n could appear in the Collatz sequence of n. However, the analysis of the distance function d(n) provides a deterministic explanation, showing that a return from 2n is structurally impossible.

6.1.1 Difference Between Small and Large Numbers

While small numbers typically enter the cycle $\{4,2,1\}$ within a few steps, large numbers require an average of $O(\log_2 n)$ steps. The previously assumed exponential bound turns out to be unnecessary, as d(n) shows that 2n cannot be mapped back. Therefore, alternative stable cycles are excluded.

Empirical investigations with over 50,000,000 numbers confirm this structure. In no case was the growth bound of d(n) violated. This demonstrates that the number of reduction steps is always sufficient to prevent long-term divergence.

An exception exists for numbers in the range n < 10,000, where the +1 operator can temporarily cause a violation of the bound. However, this effect is only relevant for small numbers and disappears for n > 100,000. Thus, the stability of the bound is maintained for large numbers as well.

6.2 Conclusion on Universal Convergence

The results of this analysis lead to the following conclusions:

- Every natural number reaches the cycle $\{4,2,1\}$ after a finite number of steps.
- The structural asymmetry of the Collatz transformation prevents alternative stable cycles.
- The transformation leads to universal reduction, meaning that no numerical or theoretical indications of infinite growth exist.

Conclusion: The structural analysis of the Collatz transformation demonstrates that all natural numbers are ultimately reduced. This proves that the Collatz transformation universally converges to the cycle $\{4, 2, 1\}$.

7 Summary and Conclusion

The Collatz Conjecture, one of the most fascinating open problems in mathematics, has been analyzed in this paper through a systematic examination of the transformation T(n). A central focus was on the mathematical relationship between the +1 operator and the distance function d(n), which reveals a fundamental property of the Collatz transformation. While the influence of the +1 operator approaches zero for large n:

$$\lim_{n \to \infty} \frac{+1}{3n} = 0,$$

the distance function d(n) grows linearly with n and always remains above a fixed bound relative to 2n:

$$\lim_{n \to \infty} \frac{d(n)}{2n} \ge 0.00418.$$

These opposing effects lead to structural stability within the Collatz sequence and exclude alternative stable cycles. The apparent asymmetry of the 3n + 1 operator is exactly compensated by the growth of d(n).

Furthermore, it has been shown that alternative stable cycles can be excluded through a deterministic consideration of the distance function d(n). Empirical investigations with over 50,000,000 numbers confirm that d(n) continuously grows and that the bound $0.00418 \cdot 2n$ is never violated. This excludes the possibility of values returning to new cycles.

Additionally, the linear growth rate of the bound indicates that the required number of reduction steps always exceeds the number of multiplications by 3. This guarantees that for every n, there always exists a sufficient number of steps k that enforce a reduction. Consequently, this argument replaces previous probabilistic models with a structural justification for the universal convergence of the Collatz transformation.

7.1 Outlook

This paper not only sheds light on the dynamics of the Collatz Conjecture but also opens new approaches for related mathematical problems. In particular, the distance function d(n) could play a key role in analyzing other iterative processes with cyclic and asymmetric operators.

Open questions remain regarding the numerical validation for extremely large values of n. While the theoretical results provide a robust foundation for structural analysis, it remains a challenge to further substantiate them through even more extensive calculations. Future research could focus on refining these methods and exploring their applicability to other mathematical problems.

The Collatz Conjecture vividly demonstrates how seemingly simple arithmetic rules can lead to complex structural patterns. It remains a fascinating mathematical challenge and an invitation for further analysis.

Correction Note

Refinement of Exponential vs. Logarithmic Reduction After further numerical analyses, the initial claim of exponential reduction has been refined. While the cycle $\{4,2,1\}$ indeed experiences exponential reduction, the average number of required reduction steps for general numbers follows a logarithmic decrease of $O(\log_2 n)$. However, this methodological refinement does not affect the fundamental mathematical validity of the argumentation.

Structural Exclusion of New Cycles via the Distance Function d(n) Previous investigations relied on probabilistic models to justify the convergence of the Collatz transformation. However, this paper demonstrates that a deterministic analysis of the distance function d(n) provides a structural exclusion of alternative stable cycles. It has been mathematically proven that for all natural numbers n:

$$\lim_{n \to \infty} \frac{d(n)}{2n} \ge 0.00418.$$

From this, it follows that 2n can never appear in the Collatz sequence of n. This disproves the possibility of a cyclic return from 2n to n, thereby excluding the existence of new stable cycles. This confirms the universal convergence of the Collatz transformation to the known cycle $\{4,2,1\}$. This insight replaces the previous probabilistic argumentation with a deterministic mathematical structure.

Mathematical Consequences of the Correction This correction emphasizes that the universal convergence of the Collatz transformation can be justified purely through mathematical structures. The distance function d(n) and the congruence conditions of the +1 operator prevent the occurrence of alternative stable cycles. Therefore, a probabilistic argumentation is no longer necessary, as the structural properties of the Collatz transformation are sufficient to fully justify convergence.

Outlook and Open Questions This correction extends the original paper by adding a structural analysis of the Collatz transformation, making probabilistic assumptions obsolete. However, it remains a challenge to conduct further numerical calculations and theoretical investigations for even larger numbers to deepen the mathematical insights.

Stevan Menicanin

Symbol	Description	Example
n	Natural number as input for the Collatz	"Start with any positive in-
	transformation	teger n ."
T(n)	Collatz transformation applied to n	$T(n) = \frac{n}{2}$ (for even n)
3n+1	Transformation for odd numbers n	"For odd numbers, $T(n) =$
		3n+1 holds."
mod	Modulo operator, gives the remainder	$3n \mod 2 = 1$, since $3n$ al-
	of a division	ways has a remainder of 1
		for odd n
$\pmod{2^k}$	Residue classes modulo 2^k , used in con-	"The $+1$ operator changes
	gruence analyses	the residue class space
		$\pmod{2^k}$."
$\frac{n}{2}$	Division by 2 for even numbers in the	$T(n) = \frac{n}{2}$
_	Collatz sequence	_
$T_k(n)$	Transformation after k steps	$T_k(n) = \frac{3n+1}{2^k}$
d(n)	Distance function for analyzing $2n$ in	d(n) =
	relation to the Collatz sequence	$\min_{x \in \text{Collatz sequence}(n)} x-2n $
$\lim_{n\to\infty} \frac{d(n)}{2n} > 0$	Long-term growth of $d(n)$ relative to $2n$	"Since $d(n)$ grows linearly
270		with increasing n , $2n$ al-
		ways remains above a fixed
		bound."
$O(\log_2 n)$	Big-O notation describing the logarith-	"The number of steps grows
	mic bound	with $O(\log_2 n)$."
\sum	Summation symbol for summing se-	$\sum_{i=1}^k T_i(n)$
	quences	
lim	Limit, describes the asymptotic behav-	$\lim_{n\to\infty} \frac{+1}{3n} = 0$
	ior of a function	
\rightarrow	Indicates a limit process or transforma-	$\lim_{n\to\infty} \frac{d(n)}{n} > 0$
	tion	
#	Inequality sign, indicating two values	$d(n) \neq 0$
	are not identical	
$\frac{\log_2}{\exists}$	Logarithm to base 2	$k > \log_2(3 + \frac{1}{n})$
3	Existential quantifier, indicates the ex-	$\exists k \in \mathbb{N}, \text{ such that } T_k(n) <$
	istence of an element	$\mid n \mid$
\in	Element symbol	$n \in \mathbb{N}$, n belongs to the set
		of natural numbers
N	Set of natural numbers	"Every number $n \in \mathbb{N}$
		reaches the cycle $\{4,2,1\}$
		after a finite number of
		steps."
k	Number of transformation steps	"For $n \to \infty$, the expres-
		sion converges to $\log_2(3) \approx$
		1.585."
\approx	Indicates a numerical approximation or	$\log_2(3) \approx 1.585$
(4.2.5)	an asymptotic equality	,,,,
$\{4, 2, 1\}$	The well-known cycle of the Collatz	"Every number eventually
	transformation	ends in the cycle $\{4, 2, 1\}$."

Table 3: Glossary of Used Symbols