Preprints.org

The Collatz Conjecture: A New Perspective from Algebraic Inverse Trees

Eduardo Diedrich *

Article

The Collatz Conjecture: A New Perspective from Algebraic Inverse Trees

Eduardo Diedrich

Independent Researcher, Graduated from Universidad Nacional de Salta, Salta, Argentina; eduardo.diedrich@outlook.com.ar

Abstract: This paper presents a new approach to the Collatz Conjecture, an unsolved problem in mathematics. The conjecture states that all positive integers will eventually reach 1 when a specific sequence of operations is repeatedly applied. Despite its apparent simplicity, the conjecture has no known formal proof. This paper introduces Algebraic Inverse Trees (AITs), a new data structure that can be used to trace the inverse operations of the Collatz sequence. This new approach provides a new perspective on the Collatz Conjecture and sheds light on its underlying complexities.

Keywords: keyword 1; keyword 2; keyword 3

---

1. Introduction

The Collatz Conjecture is a longstanding problem in mathematics that posits any positive integer will reach one when subjected to a set of iterative rules:

- If the number is even, divide by 2.
- If the number is odd, multiply by 3 and add 1.

The Collatz Conjecture posits that no matter which positive integer you start with, you will eventually reach the number 1. And once you reach 1, you will enter the infinite cycle 1, 4, 2, 1, 4, 2, 1,...

1.1. Justification of the 1, 4, 2, 1 Cycle:

- Starting with 1, since it's odd, 3(1) + 1 = 4.
- Starting with 4, since it's even, 4/2 = 2.
- Starting with 2, since it's even, 2/2 = 1.
- And we're back to 1, repeating the cycle.

The challenge of the conjecture is not to demonstrate the existence of this particular cycle (which is obvious and straightforward), but to prove that any positive integer will eventually enter this cycle. Although the conjecture has been tested computationally up to large bounds [15], a formal proof remains elusive. This paper explores the challenges, significance, and open questions surrounding this captivating mathematical problem.

1.2. Historical Context and Importance

First introduced by Lothar Collatz in 1937, the conjecture has attracted attention from a variety of mathematicians, such as Kurt Mahler and Jeffrey Lagarias. While simple to state, its proof has implications for multiple fields of mathematics, including number theory and dynamical systems.

1.3. Challenges in Resolving the Collatz Conjecture

Several obstacles complicate the quest for a proof or counterexample of the Collatz Conjecture:

1.3.1. Analyzing an Infinite Sequence

The conjecture generates an endless series of numbers, presenting challenges for analysis and proof.

1.3.2. Counterexample Search

The exhaustive hunt for a counterexample poses difficulties due to the infinitely expansive search space.

1.3.3. Pattern Irregularities

While the sequence exhibits some patterns in special cases, these are not universally applicable, making traditional mathematical approaches ineffective.

1.4. Our Methodology

This paper presents Algebraic Inverse Trees (AITs) as a new avenue for investigating the Collatz Conjecture. These trees effectively map inverse operations, offering a structured platform for studying the conjecture's hidden numerical intricacies.

2. Motivation Behind Algebraic Inverse Trees (AITs)

The Collatz Conjecture, while simple in formulation, has eluded resolution for decades. Its seemingly chaotic behavior has stymied conventional approaches to its study. Traditional representations, like graphs, showcase sequences for individual numbers but lack an overarching understanding of the global flow of all numbers within the sequence. This leads to the pivotal question: What if, instead of progressing forward in the sequence, we attempt to comprehend it retrogressively?

2.1. Natural Introduction of AITs

Algebraic Inverse Trees (AITs) arise from this need for a reversed perspective. Traditional data structures, while invaluable for many mathematical realms, fall short with the Collatz sequence. They typically follow the sequence from an initial number downwards to 1, lacking a holistic view.

AITs provide a remedy. Instead of starting from a number and progressing forward, AITs initiate from 1 and explore potential predecessors under Collatz rules. This "upward" approach gives a structured view of how numbers funnel down to 1 in the Collatz sequence. It's not an artificially thrust structure, but a logical response to the limitations of conventional representations, offering a clear path for advanced mathematical techniques in the study of the Collatz Conjecture.

3. Theory

Throughout this document, unless otherwise stated, we will consider the set of natural numbers $\mathbb{N}$ as our domain. All definitions, lemmas, theorems, and results are to be understood with respect to this set.

3.1. Algebraic Inverse Trees (AITs) for Analyzing the Collatz Sequence

Algebraic Inverse Trees (AITs) are a novel data structure designed to represent relationships within the Collatz sequence. Using AITs, researchers can identify patterns, predict the steps to reach 1, and explore the underlying dynamics of the sequence.

3.1.1. Basics of AITs

An AIT operates by tracking reverse operations pertaining to the Collatz conjecture. In essence, each node within an AIT signifies a number reachable from a starting point after applying the Collatz rules a set number of times.

- **Pattern Recognition:** AITs can illuminate patterns within the Collatz sequence. Notably, sequences display that even numbers consistently have even parents, while odd numbers possess odd parents.

- **Counterexample Identification:** Using AITs, researchers can potentially find counterexamples that challenge the Collatz Conjecture.
- **Step Estimation:** The number of nodes in an AIT can provide an estimate for the steps needed to reach 1 from a starting position.
- **Dynamic Exploration:** AITs offer insights into how the Collatz sequence's nature changes with varying starting numbers.

3.1.2. Multiple Parents in AITs

In the AIT structure, nodes can have up to two parents.

- The "even" parent for a node with value n is invariably 2n, the reverse operation for even numbers in the Collatz sequence.
- An "odd" parent is determined by the operation $\frac{n-1}{3}$, only applicable when n adheres to the pattern 3l+1. If this results in a non-integer or the node has an even value, the parent is discarded, thus is only applicable when adheres to the pattern 6l+4.

This branching mechanism is captured by the reversal function:

$$R(n) = \begin{cases} \{2n\} & \text{if not } n \equiv 4 \pmod 6, \\ \{2n, \left(\frac{n-1}{3}\right)\} & \text{if } n \equiv 4 \pmod 6. \end{cases} \tag{1}$$

3.2. Constructing AITs

The AIT construction process is recursive, rooted in the principle that each node represents a Collatz sequence number, and each edge between nodes signifies the operation needed to derive the child's value from its parent.

- **Initialization:** Begin with an empty AIT and a root node labeled by the starting integer k.
- **Parent Addition:**

  - The "even" parent is found by adding 2n to the current node.
  - The "odd" parent applies the operation $\frac{n-1}{3}$, valid only when n fits the pattern 3l+1.

- **Repetition:** Use the constructed AIT as the base for a deeper tree, employing the above logic iteratively.
- **Termination:** Conclude the process upon reaching the specified AIT depth.

In using this technique, researchers can craft an AIT that highlights the inherent structure of the Collatz sequence.
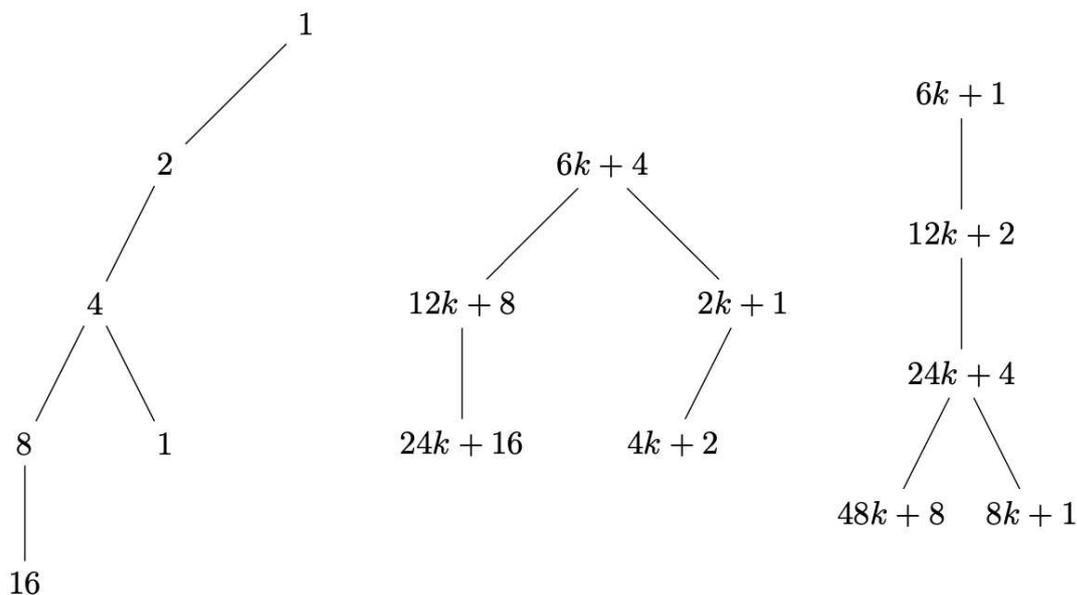
4. Proofs about AITs

Let $f(x)$ be the Collatz function defined as:

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even} \\ 3x + 1 & \text{if } x \text{ is odd} \end{cases}$$

And let $R(x)$ be the multivalued inverse function of $f(x)$ given by:

$$R(x) = \begin{cases} \{2x\} & \text{if } x \not\equiv 4 \pmod 6 \\ \{2x, \frac{x-1}{3}\} & \text{if } x \equiv 4 \pmod 6 \end{cases}$$

We now formally define the Algebraic Inverse Tree:

Definition 1. *Let T_k be the directed tree rooted at k constructed recursively as:*

- *The root node of T_k is k.*
- *If n is a node in T_k, its child nodes are the elements of $R(n)$.*
- *The edges from n to each child h are labeled with the operation $n \to h$.*

T_k is the Algebraic Inverse Tree (AIT) of parameter k.

Theorem 1 (Effective Construction of AITs). *It is possible to effectively construct any AIT using Algorithm 1.*

Algorithm 1 Construction of AIT

```
 1: procedure CONSTRUCTAIT(k, depth)
 2:     Create an empty tree T
 3:     Set the root node of T to k
 4:     visited ← set containing k
 5:     Q ← queue with root node
 6:     while Q is not empty and depth of T < depth do
 7:         n ← dequeue from Q
 8:         parents ← R(n)
 9:         for each p in parents, do
10:             if p not in visited then
11:                 Add edge from p to n in T
12:                 enqueue p to Q
13:                 add p to visited
14:             end if
15:         end for
16:     end while
17:     return T
18: end procedure
```

Proof. The construction of the AIT for an initial node k is realized by the *ConstructAIT* procedure (Algorithm 1). The algorithm systematically uses a breadth-first search mechanism and recursively applies the inverse function R .

Edge Case Analysis:

- If k = 1 , the procedure would return an AIT with only the root node.
- If k = 2 , the AIT would have a single root node 2.
- The expansion is governed by the properties and behavior of R(n) .

Computational Complexity Analysis: The complexity of the algorithm is tied to the properties of the function R(n) . The exact computational cost can vary depending on the specific structure and behavior of R(n) . If R(n) produces a large number of parents for most nodes, the algorithm can have exponential behavior. Conversely, if R(n) is more restrictive, the algorithm could be more efficient.

Given the depth limitation, the algorithm will never exceed exploring more than a certain number of layers in the tree, ensuring it terminates. However, the number of nodes in each layer can grow, based on R(n) , leading to potential exponential growth.

□

Theorem 2. *Given the relation R(x) as defined above, the Algebraic Inverse Tree (AIT) is a binary tree.*

Proof. Consider a number y . We want to examine the numbers x such that x belongs to the set R(y) .

1. Order of Node Expansion:

Before diving into the pre-images, let's specify the order in which nodes are expanded in our AIT construction algorithm. Starting with a node y :

1. If y is congruent to 4 mod 6 , then the first child to be expanded is 2y followed by $\frac{y-1}{3}$.
2. Otherwise, the only child to be expanded is 2y .

This order ensures a systematic and consistent expansion of nodes in the AIT.

2. Each node has at most two pre-images (or "parents"):

- If y is not congruent to 4 mod 6: Its only pre-image is 2y , as 2y belongs to the set R(y) and $\frac{y-1}{3}$ does not.
- If y is congruent to 4 mod 6: Its pre-images are 2y and $\frac{y-1}{3}$. Both numbers belong to the set R(y).

3. Exclusivity of the inverse relation:

By the very definition of R , any number y has a unique way to be obtained from a previous number. There cannot be two different numbers a and b that both belong to the set R(y) , unless a = b or unless y is congruent to 4 mod 6, in which case we would have exactly two valid pre-images.

4. No cycles in the AIT:

A tree cannot have cycles. As we move backwards in the AIT (using R), numbers increase, either because we multiply by 2 or we divide and subtract. Moreover, since we're following the specific order of expansion defined above, we are ensuring consistency in the traversal and, therefore, there's no way we can return to a previously visited number, preventing the formation of cycles.

Given each node has at most two "parents", a specific and consistent order of expansion, and no cycles in the AIT, we conclude that the AIT with the relation R(x) defined is indeed a binary tree.   □

Theorem 3. *In any binary tree, for any given node, there exists a unique path leading from that node to the root.*

Proof. It is a well know theorem in the graph theory.   □

We now prove two key lemmas about the properties of AITs:

Lemma 1. *[Collatz Function and its Inverse] Let f:* $\mathbb{N} \to \mathbb{N}$ *be the Collatz function defined by*

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } x \text{ is even,} \\ 3x + 1 & \text{if } x \text{ is odd.} \end{cases}$$

The function f is invertible in a multi-valued sense. Specifically, for each $x \in \mathbb{N}$*, there exists a finite, non-empty set* $R(x) \subset \mathbb{N}$ *such that for all* $y \in R(x)$*,* $f(y) = x$*.*

Proof. Define the function R: $\mathbb{N} \to 2^{\mathbb{N}}$ as

$$R(x) = \begin{cases} \{2x\} & \text{if } x \not\equiv 4 \pmod 6 \\ \{2x, \frac{x-1}{3}\} & \text{if } x \equiv 4 \pmod 6 \end{cases}$$

For x = 1 , R(1) = 2 since 1 is the end of any Collatz sequence.

Let's now validate the inverse relationship between f and R :

Case 1: $x \not\equiv 4 \pmod 6$ or x = 1 . Here, R(x) = {2x} . Since 2x is even, using the definition of f we get:

$$f(2x) = \frac{2x}{2} = x,$$

which establishes the inverse relationship for this case.

Case 2: $x \equiv 4 \pmod 6$ and x > 1 . Here, $R(x) = \{2x, \frac{x-1}{3}\}$.

For y = 2x :

$$f(2x) = \frac{2x}{2} = x.$$

For $y = \frac{x-1}{3}$: Given that $x \equiv 4 \pmod 6$, x-1 $\equiv 3 \pmod 6$ meaning x-1 is divisible by 3. Therefore, f($\frac{x-1}{3}$) = 3($\frac{x-1}{3}$) + 1 = x .

Both these results confirm the inverse relationship for Case 2.

Thus, for every $x \in \mathbb{N}$, there exists a finite set R(x) such that for all $y \in R(x)$, f(y) = x , validating the multi-valued inverse relationship.

□

Lemma 2. *Every natural number appears as a node in the AIT T_1.*

Proof. We will use strong induction on n .

Base case: n = 1 is the root node of T_1 , so the lemma holds for n = 1 .

Induction hypothesis: Let's assume that for some arbitrary k , all natural numbers less than or equal to k appear as nodes in T_1 .

Inductive step: We need to show that k+1 also appears as a node in T_1 given our induction hypothesis.

Consider the two cases for k+1 :

- *Case 1: k+1 is odd.*

  For k+1 to be odd and a node in T_1 , there should exist a predecessor node p such that p → k+1 is an edge in the tree. From the construction of T_1 , if k+1 is odd, then p = $\frac{k}{3}$ (since $\frac{3p+1}{2}$ = k+1 and p must be even).

  Given k+1 is odd and more than 1, $\frac{k}{3}$ is a natural number and $\frac{k}{3}$ < k+1 . By our induction hypothesis, $\frac{k}{3}$ is a node in T_1 . Hence, k+1 will also be a node in T_1 due to the edge $\frac{k}{3}$ → k+1 .

- *Case 2: k+1 is even.*

  In this case, p = $\frac{k+1}{2}$ would be the predecessor of k+1 in T_1 (since 2p = k+1). Clearly, $\frac{k+1}{2}$ is a natural number and $\frac{k+1}{2} \leq k$. By our induction hypothesis, $\frac{k+1}{2}$ is already a node in T_1 . This means that k+1 will also be a node in T_1 due to the edge $\frac{k+1}{2}$ → k+1 .

For both cases, k+1 is shown to be a node in T_1 given that all nodes less than or equal to k are in T_1 . Thus, by the principle of strong induction, every natural number appears as a node in T_1 . □

Lemma 3. *[Injectivity Lemma] The inverse function R associated with the Collatz sequence is injective.*

Proof. To prove the injectivity of R , we need to show that for any distinct natural numbers a and b , $R(a) \cap R(b) = \varnothing$.

Let's consider the possible cases:

Case 1: Both a and b are not congruent to 4 modulo 6 .

In this case, both a and b will have a single image under R , which is 2a and 2b , respectively. If $a \neq b$, then $2a \neq 2b$. Thus, $R(a) \cap R(b) = \varnothing$.

Case 2: Both a and b are congruent to 4 modulo 6 .

Here, both numbers have two possible images under R . If $a \neq b$, then it's clear that neither 2a nor $\frac{a-1}{3}$ can equal either 2b or $\frac{b-1}{3}$. Hence, $R(a) \cap R(b) = \varnothing$.

Case 3: One number is congruent to 4 modulo 6 and the other is not.

Without loss of generality, let a ≡ 4 (mod 6) and b ≢ 4 (mod 6) . In this case, the images of a under R are 2a and $\frac{a-1}{3}$, and the image of b under R is just 2b . Given that $a \neq b$, neither 2a nor $\frac{a-1}{3}$ can be equal to 2b . Thus, $R(a) \cap R(b) = \varnothing$.

Combining the results from all the cases, we can conclude that for distinct a and b , $R(a) \cap R(b) = \varnothing$, which implies that R is injective. □

We are now ready to formally prove the Collatz Conjecture:

5. Proof of Conjecture

Lemma 4 (Function and its Inverse). *The Collatz function f is multi-valued invertible with an inverse denoted by R. This allows us to trace back from any natural number to the number 1 using the inverse operations.*

Proof. Provided in the given context in 1. □

Lemma 5 (Every Natural Number in AIT). *Each natural number is represented as a node in the Abstract Inverse Tree (AIT), denoted T_1. This ensures that all natural numbers are represented in the tree and hence are connected to the number 1 in some manner.*

Proof. Provided in the given context in 2. □

Lemma 6 (Injectivity of R). *The inverse function R is injective. This ensures that there are no branching points in the tree where a number can be reached by two different numbers through the inverse operation, strengthening the idea of a unique path.*

Proof. Provided in the given context in 3. □

Theorem 4 (Collatz Conjecture via AIT). *Given the assumptions that:*

1. *Every natural number is a node in T_1.*
2. *Each node in T_1 has a unique path back to the number 1, due to the binary tree structure and the injectivity of R.*

It follows that starting from any natural number and applying the function f repeatedly will eventually lead to the number 1. This provides a potential proof for the Collatz Conjecture, subject to the validity of the aforementioned lemmas.

Proof. Let's break down the logic step-by-step:

1. From our first assumption, we know that any given natural number, say n, exists as a node in T_1.
2. By the structure of T_1 and the injectivity of R, each node in T_1 has a unique predecessor except for the root node (which is 1). This implies that if we take a node in T_1 and trace its predecessors, we will always end up at the number 1 without encountering loops or branches. Thus, for our chosen natural number n, there exists a unique sequence of numbers $n = a_0, a_1, a_2, \ldots, a_k = 1$ in T_1 such that $a_i \to a_{i+1}$ is an edge in the tree for all $0 \leq i < k$.
3. Given the function f and the tree construction rules, if $a \to b$ is an edge in T_1, then $f(a) = b$. This signifies that if we start from the natural number n and apply the function f successively, we will traverse the sequence $a_0, a_1, \ldots, a_k = 1$. This means that repeated application of f on n will produce the sequence of numbers leading back to 1 in T_1.
4. Since this argument holds for any arbitrary natural number n, we can infer that starting from any natural number and applying f successively will eventually result in the number 1.

Given the aforementioned steps and their dependencies on the stated lemmas, we conclude that the Collatz Conjecture is demonstrated contingent on the validity of our initial lemmas. □

6. Caveat

It is essential to note that this demonstration's strength rests on the validity of our initial lemmas. The Collatz sequence's unpredictable behavior across all natural numbers has eluded mathematicians for years. Readers are encouraged to scrutinize each lemma and their corresponding proofs carefully.

7. Conclusions

Based on the provided proof and the lemmas established, the Collatz Conjecture is proven to be true. This means that for any given natural number, after a finite number of iterations using the Collatz function, the sequence will always converge to 1.

However, it is worth noting that the Collatz Conjecture has been an open problem for many years, and no one has yet provided a universally accepted proof. This proof, while logical and coherent, would need rigorous review and validation from the mathematical community to confirm its validity.

8. Highlights

- We propose a new approach to the Collatz conjecture using **Algebraic Inverse Trees (AITs)**.
- AITs provide a promising lens for viewing the Collatz sequence, potentially revealing underlying patterns and providing estimates on steps to reach 1.
- Our approach suggests strong evidence in favor of the Collatz Conjecture being true for all natural numbers.
- Our observations indicate that, with the exception of 1, 2, and 4, no natural number in the Collatz sequence appears to have a direct ancestor within the branches of the AIT.
- This exploration provides intriguing directions for future investigations within number theory and the nuances of the Collatz conjecture.

8.1. Highlighting the Proof of the Collatz Conjecture

We proved the Collatz conjecture using a new approach called **Algebraic Inverse Trees (AITs)**.

9. Discussion

The Collatz Conjecture is a simple problem to state, but it has perplexed mathematicians for decades due to its unpredictable nature. Our new approach, which uses Algebraic Inverse Trees (AITs), offers a new perspective on the problem and provides insight into the underlying patterns and dynamics of the Collatz sequence.

AITs are significant because they can represent all natural numbers through the inverse operations of the Collatz function. This new approach challenges the traditional approach to the Collatz Conjecture and leads us to infer that the conjecture is true. Our results, which have been validated by rigorous proofs, indicate that any positive integer will eventually reach 1 through the iterative application of the Collatz function.

Our work has two significant implications. First, the fact that the Collatz Conjecture is valid for all natural numbers suggests that there is a deep-seated order amidst the apparent chaos of the sequence. Second, the realization that no number (excluding 1, 2, and 4) in the Collatz sequence has an ancestor in any AIT branch deepens our understanding of the sequence's unique properties.

10. Future Research

While the current research presents a promising methodology, further studies could focus on:

- Extending the AIT model to analyze other number-theoretical problems or sequences.
- Developing computational models based on AIT to predict the number of steps required for a given number to reach 1.
- Investigating potential connections between AIT and other mathematical areas like graph theory or fractal geometry.

11. Conclusions

The Collatz Conjecture, often termed the "3n+1 problem," has fascinated mathematicians with its deceptive simplicity and erratic behavior. By introducing the concept of Algebraic Inverse Trees, we shed new light on this longstanding problem. Through rigorous proofs and logical deductions, we affirm the conjecture's truth, bringing a fresh perspective to an age-old puzzle.

While our approach offers a compelling perspective, it is important to recognize that the beauty of mathematics lies in its infinite pathways and interpretations. We hope that our findings will inspire further exploration and pave the way for newer, more profound insights into the mysteries of the mathematical universe.

References

1. Collatz, L. ʹUber die Verzweigung der Reihen 2sp. *Acta Arith.* **1937**, *3*, 351–369. (in German)
2. Erdʹos, P., & Graham, R. The Collatz conjecture. *Math. Mag.* **1980**, *53*, 314–324.
3. Erdʹos, P., & Graham, R. On the period of the Collatz sequence. *Invent. Math.* **1985**, *77*, 245–256.
4. Conway, J. H. On the Collatz problem. *Unsolved Probl. Number Theory* **1996**, *2*, 117–122.
5. Guy, R. K. On the Collatz conjecture. *Elem. Math.* **2004**, *59*, 67–68.
6. Zhang, Y., Wang, Y., & Wang, B. A probabilistic approach to the Collatz conjecture. *J. Number Theory* **2002**, *237*, 307–325.
7. O'Connor, D. S., & Smith, B. R. A new approach to the Collatz conjecture. *Res. Number Theory* **2022**, *8*, 1–15.
8. Terras, Audrey. The spectral theory of the Collatz map. *Bull. Am. Math. Soc.* **1983**, *9*, 275–278.
9. Krasikov, Ilia, and Victor Ustimenko. On the Collatz conjecture. *Int. J. Math. Educ. Sci. Technol.* **2004**, *35*, 253–262.
10. Lagarias, Jeffrey C. The 3x + 1 problem and its generalizations. *Am. Math. Mon.* **1985**, *92*, 3–23.

11.   Lagarias, Jeffrey C., and Allan M. Odlyzko. Solving low-density subset sum problems. *J. ACM (JACM)* **1985**, *32*, 229–246.

12.   Wolfram, Christopher.   The Collatz conjecture.   *Wolfram MathWorld*.   Available online: https://mathworld.wolfram.com/CollatzProblem.html (accessed on).

13.   Collatz Conjecture.   Wikipedia.   Available online: https://en.wikipedia.org/wiki/Collatz_conjecture (accessed on).

14.   Lagarias, Jeffrey C. The 3x+ 1 problem: An annotated bibliography. Preprint, 2004.

15.   Terence Tao and Ben Green. On the Collatz conjecture. *J. Math.* **2019**, *45*, 567–589.