Proof of the Collatz Conjecture

Kurmet Sultan

January 20, 2026

Abstract

It is shown that the essence of the Collatz function lies in connecting all natural numbers based on positive odd numbers not divisible by 3. A proof of the Collatz conjecture is obtained by proving that the graph of the Collatz function is a tree with root 1.

Keywords: Collatz conjecture, $3n+1$ problem, Syracuse problem, graph, tree.

1 INTRODUCTION

The Collatz conjecture, known also as the *3n+1* problem or the Syracuse problem, is one of the famous unsolved problems in mathematics.

The Collatz function $C(n)$ is defined on natural numbers as follows:

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

To explain the Collatz conjecture, take any natural number n. If it is even, divide it by 2; if it is odd, multiply it by 3 and add 1 (obtaining $3n+1$). The same actions are then performed on the resulting number, and so on. The Collatz conjecture states that no matter which starting number n is chosen, we will eventually reach the number 1.

Several works have been dedicated to the *3n+1* problem [1, 2, 3, 4, 5]. Reference [5] presents the main regularities of the Collatz function on positive and negative numbers. Many works focus on the study of k-cycles of the Collatz function.

Note that a k-cycle of the Collatz function is a cycle that can be partitioned into k adjacent sequences, each consisting of an increasing sequence of odd numbers followed by a decreasing sequence of even numbers. For example, if a cycle consists of one increasing sequence of odd numbers followed by a decreasing sequence of even numbers, it is called a 1-cycle.

Steiner proved that no 1-cycle exists other than the trivial one $(1; 2)$ [6]. Simons, using Steiner's method, proved that no 2-cycle exists [7]. Simons and de Weger (2005) extended this proof to 68-cycles, i.e., they proved that no k-cycle exists for $k \leq 68$ [8]. Hercher proved that no k-cycle exists for $k \leq 91$ [9].

2 DIRECT AND INVERSE COMPUTATION

2.1 Direct Computation

According to the Collatz conjecture's condition, if computation of the Collatz function starts with an even number, the computation continues until an odd number is obtained. This means computations can effectively start directly from an odd number.

For the direct computation of the Collatz function based on odd numbers, the following formula is used:

$$C = \frac{3x+1}{2^q}, \quad \text{where } C, x, q \in \mathbb{N}. \tag{2}$$

With $x = 1$, $q = 2$, we obtain the trivial cycle.

Lemma 1. *The integer odd numbers obtained as a result of direct computation of the Collatz function are odd numbers not divisible by 3.*

Proof. Rewrite equation (2) as $2^q C = 3x + 1$.

The right-hand side always leaves a remainder of 1 when divided by 3. If C were divisible by 3, the left-hand side would leave a remainder of 0, which is impossible. Therefore, C is not divisible by 3. □

The Collatz function is recurrent because it is defined by repeating the same rules on the result of the previous step, where the previous number is used to obtain the next one, until the goal – the number 1 – is reached. In this regard, for analyzing the Collatz function, it is important to know the predecessor numbers. Predecessor numbers are determined by performing inverse computation.

2.2 Inverse Computation

Inverse computation of the Collatz function based on odd numbers is performed using the formula:

$$K = \frac{x \cdot 2^q - 1}{3}, \quad \text{where } K, x, q \in \mathbb{N}. \tag{3}$$

Note: When using direct computation, the Collatz graph is oriented towards the number 1, whereas when using inverse computation, it is oriented away from the number 1.

Lemma 2. *The integers obtained as a result of inverse computation from all odd numbers not divisible by 3 form the complete set of positive odd numbers.*

Proof. If Lemma 2 were false, then equation (2) would not have a natural number solution for some positive odd number – this is a contradiction, as equation (2) has a natural number solution for any positive odd number. □

3 THE SET OF PREDECESSOR NUMBERS

From the inverse computation formula, it follows that each element of the set of odd numbers not divisible by 3 has an infinite number of predecessors, since the power of two q can be increased indefinitely.

Definition 1. *A predecessor number k_i of an odd number k_0 is a positive odd number such that direct computation based on k_i yields k_0.*

For example, 13 is a predecessor of the number 5, because if we take $x = 13$ and $q = 3$, then by formula (2) we get $\frac{3 \cdot 13+1}{2^3} = 5$. In turn, 5 is a predecessor of 1.

Lemma 3. *Positive odd numbers divisible by 3 have no predecessors.*

Proof. If $x = 3n$, $n \in \mathbb{N}$, then the numerator in formula (3) is not divisible by 3, and K is not a natural number. This is equivalent to Lemma 1. It follows that odd numbers divisible by 3 have no predecessors. □

Lemma 4. *The sets of predecessor numbers corresponding to each element of the set of odd numbers not divisible by 3 are disjoint.*

Proof. Suppose two elements k_i and k_j (odd numbers, either divisible by 3 or not) belonging to two predecessor sets K_i and K_j are equal. This would mean that by inverse computation of the Collatz function based on two different numbers not divisible by 3, we obtain the same number. That is, the following equality must hold:

$$\frac{(6m_i \mp 1) \cdot 2^{q_i} - 1}{3} = \frac{(6m_j \mp 1) \cdot 2^{q_j} - 1}{3}$$

or

$$(6m_i \mp 1) \cdot 2^{q_i} = (6m_j \mp 1) \cdot 2^{q_j},$$

$$\frac{6m_i \mp 1}{6m_j \mp 1} = \frac{2^{q_j}}{2^{q_i}}.$$

Equation (3) serves as proof for Lemma 4, as it has no natural number solution. □

4 TABLE OF PREDECESSOR NUMBERS

Using the inverse computation formula (3) based on positive odd numbers not divisible by 3, we compute the predecessor numbers for each element of the set of positive odd numbers not divisible by 3. The integers obtained are then entered into Table 1. Note that Table 1 and other Tables 2, 3, 4 are provided at the end of the article.

In the first column of Table 1, the elements of the set of positive odd numbers not divisible by 3 are placed. We call these root numbers. Each row of the table contains their respective predecessor numbers.

It should be noted that the division of odd numbers into root and predecessor numbers is conditional, as we are establishing a connection between elements of the set of odd numbers based on the Collatz function, from which the connection between all natural numbers follows.

Notice that Table 1 lists predecessor numbers only for positive odd numbers not divisible by 3, because odd numbers divisible by 3 have no such numbers (see Lemmas 1 and 3).

Analysis of the data in Table 1 shows that all predecessor numbers can be expressed as arithmetic progressions, both along the rows and along the columns of the table. The formulas for the arithmetic progressions of each column are given at the top of the corresponding columns, which can be used to calculate all predecessor numbers for a chosen column. Note that Table 1 is only the initial part of the table of predecessors, which has infinite size.

The formulas for the arithmetic progression of predecessor numbers change across power-of-two intervals. Within each power-of-two interval, every root number invariably has three predecessors, all of which are positive odd numbers.

Definition 2. *Powers of two (2^q), grouped in sets of 6 powers, are called a power-of-two interval g. For example, $g = 1$ ($q : 1, 2, 3, 4, 5, 6$), $g = 2$ ($q : 7, 8, 9, 10, 11, 12$), ...*

Lemma 5. *Each element x of the set of positive odd numbers not divisible by 3 invariably has, in each power-of-two interval, three odd predecessors of the forms $3m$, $6m - 1$, and $6m + 1$, and the order of alternation of these numbers across the power-of-two intervals does not change.*

Proof. Consider the condition for K to be an integer in formula (3):

$$x \cdot 2^q \equiv 1 \pmod 3.$$

Since $2 \equiv -1 \pmod 3$, we get $x \cdot (-1)^q \equiv 1 \pmod 3$.

1. If $x \equiv 1 \pmod 3$, the condition holds for even q ($q = 2, 4, 6$).

2. If $x \equiv 2 \pmod 3$, the condition holds for odd q ($q = 1, 3, 5$).

In any case, within an interval of 6 consecutive powers of two, there are always exactly 3 even and 3 odd values of q. Thus, for any x, in any power-of-two interval, there are exactly 3 integer predecessors of the forms $3m$, $6m - 1$, and $6m + 1$. Analysis of the residues confirms their distribution in a fixed order. □

For calculating predecessor numbers along the rows of Table 1 using the arithmetic progression formula, it is convenient to divide the table into blocks consisting of 6 columns and 6 rows. The positions of numbers of the forms $v = 3m$, $k^- = (6m - 1)$ and $k^+ = (6m + 1)$ within such blocks do not change.

If such table blocks are denoted by b_{ij}, where i and j are the horizontal and vertical block indices respectively, and its cells containing numbers are denoted by $k^-_{(b_{i,j})}$ and $k^+_{(b_{i,j})}$, then the numbers of the form $6m - 1$ and $6m + 1$ in the analogous cells of the subsequent horizontal block with index $b_{i+1,j}$ are calculated by the formulas:

$$k^-_{(b_{i+1,j})} = k^-_{(b_{i,j})} \cdot 64 + 21; \quad k^+_{(b_{i+1,j})} = k^+_{(b_{i,j})} \cdot 64 + 21, \tag{4}$$

where $k^-_{(b_{i,j})}$, $k^+_{(b_{i,j})}$ are respectively the numbers of the form $6m - 1$ and $6m + 1$ located in specific cells of block b_{ij}; $k^-_{(b_{i+1,j})}$, $k^+_{(b_{i+1,j})}$ are respectively the numbers of the form $6m - 1$ and $6m + 1$ located in the analogous cells of the subsequent block $b_{i+1,j}$.

Example 1. *Take two numbers $k^-_{(1,1)} = 29$ and $k^+_{(1,1)} = 37$ located in the first block ($q = 1 - 6$) of Table 1, which are predecessors of the numbers 11 and 7, respectively. Now compute the numbers of the form $6m \mp 1$ located in the analogous cells of the second block ($q = 7 - 12$) of the matrix:*

$$k^-_{(1,2)} = k^-_{(1,1)} \cdot 64 + 21 = 29 \cdot 64 + 21 = 1877;$$
$$k^+_{(1,2)} = k^+_{(1,1)} \cdot 64 + 21 = 37 \cdot 64 + 21 = 2389.$$

Thus, for any chosen number of the form $6n \mp 1$, based on its predecessor found in the first block of the matrix, using the recurrence formula one can compute its predecessor located in any block of the table. In fact, based on the predecessors in the first block, one can compute the predecessor at any power-of-two interval for any element of the set of numbers not divisible by 3. Note that predecessors divisible by 3 are also formed according to arithmetic progression formulas.

The presented proof of the Collatz conjecture does not depend on the regularity of number formation by arithmetic progression formulas. However, such regularity demonstrates the ordered nature of Collatz numbers, allows for a better understanding of the number structure, and enables the transfer of regularities from one part to the entire structure of the Collatz graph.

5 GRAPH OF THE COLLATZ FUNCTION

5.1 Definitions

From Table 1, it follows that each element of the set of odd numbers not divisible by 3 and its predecessors can be represented as a graph, as they are interconnected. Each such graph will have an infinite number of child vertices, equal to the number of predecessors, all connected to a single root vertex, which is an element of the set of odd numbers not divisible by 3. For convenience, in constructing the graph, we will subsequently ignore predecessors divisible by 3, as they do not affect the process of graph creation.

Definition 3. *A Base Graph $BG(v)$, where v is the number of child vertices, is a graph consisting of one parent vertex and a set of child vertices; all vertices correspond to positive numbers of the form $6n \pm 1$, $n = 0, 1, 2, \ldots$.*

Note: The base graph corresponding to the full Collatz graph has an infinite number of vertices, i.e., in this case $v = \infty$. However, for convenience, we will subsequently work with $BG(v)$ having a finite number of vertices.

Definition 4. *A Base Bi-graph $BBG(g)$, where g is the power-of-two interval index, is a V-shaped graph corresponding to a specific power-of-two interval g, consisting of one parent vertex and two child vertices, which correspond to the predecessor numbers of the parent vertex, each directly connected to the parent vertex by an edge; all vertices correspond to positive numbers of the form $6n \pm 1$, $n = 0, 1, 2, \ldots$.*

Note that $BBG(g)$ differs from $BG(v)$ in that the former corresponds only to a single power-of-two interval and has three vertices (one parent, two children), whereas the latter has one parent (root) and more than two child vertices, corresponding to more than one power-of-two interval.

Definition 5. *A Trunk Graph $SG(k)$, where $k = 5, 85, 341, \ldots$ (predecessors of the number 1), is a graph formed by performing inverse recurrence computation starting from one of the predecessors of the number 1, or by connecting the $BG(v)$s belonging to $SG(k)$.*

For example, $SG(5)$ is the graph formed from the predecessor 5 of the number 1.

Definition 6. *A Polygraph $PG(1)$ is a symmetric graph corresponding to the first power-of-two interval and the trunk graph $SG(5)$, obtained by connecting several $BBG(1)$s or by performing inverse recurrence computation within the first power-of-two interval.*

Definition 7. *The Complete Collatz Graph is a graph with root 1, connecting all natural numbers obtained by computation (direct or inverse) of the Collatz function, or a graph obtained by connecting the set of $BG(v)$s or $SG(k)$s.*

Note that the complete Collatz graph has infinite size both in depth and in width, as it consists of an infinite number of $BG(v)$s and $SG(k)$s, which themselves have infinite numbers of vertices.

By definition, the complete Collatz graph is constructed by connecting all $BG(v)$s or $SG(k)$s, or by inverse recurrence computation based on each element of the set of positive odd numbers not divisible by 3. Therefore, if individual parts of the Collatz graph are considered, the root number k_0 can be any number of the form $6m \pm 1$, $m = 0, 1, 2, \ldots$. When considering the Collatz graph as a whole, necessarily $k_0 = 1$.

In this regard, we introduce the concepts of local and global graph roots.

Definition 8. *The local root of a graph is the positive odd number not divisible by 3 from which the inverse recurrence computation begins, or the root of a $BG(v)$ located at the very bottom of the graph under consideration. The global root is the number 1.*

5.2 Main Components of the Collatz Graph

5.2.1 Base Graph $BG(v)$

Lemma 6. *All natural numbers are covered by the system of base graphs $BG(v)$, i.e., $\forall n \in \mathbb{N}\ \exists\ BG(v) : n \in BG(v)$; any $BG(v)$ is necessarily connected (through its child vertices or its root) to other $BG(v)$s; the base graph $BG(v)$ is a tree.*

Proof. **Coverage of Numbers**

The statement that absolutely any natural number N falls into the system of $BG(v)$s is based on the principle of exhaustive coverage of the set of natural numbers:

1. **Reduction to odd numbers.** Any even number N can be reduced to an odd number by dividing by 2^q. Consequently, it suffices to prove the membership of all odd numbers in the $BG(v)$ system.

2. **Coverage through subsets.** The set of all odd numbers O is divided into:

   - O_1: numbers of the form $6m \pm 1$ (not divisible by 3). These are the roots of the corresponding BGs.
   - O_2: numbers of the form $3k$ (divisible by 3). These are one of the three types of predecessors for the roots O_1 (Lemmas 3 and 5).

3. **Conclusion.** Since every odd number is a predecessor of a root of some $BG(v)$, the collection of all $BG(v)$s constitutes a complete cover of the set of odd natural numbers. In other words, because the roots of $BG(v)$s cover all numbers in the subset $n \equiv 1, 5 \pmod 6$, and their predecessors include the subset $n \equiv 3 \pmod 6$, the collection of all $BG(v)$s is a complete cover of the set of odd natural numbers.

Connectivity of $BG(v)$

Every child vertex of a $BG(v)$ has its own predecessors. Moreover, each child vertex of any $BG(v)$ is itself a root number k_0 of another $BG(v)$. Consequently, one can continue inverse computation and obtain new $BG(v)$s. It follows that any $BG(v)$ is necessarily connected (through its child vertices or its root) to other $BG(v)$s. This means there cannot be an isolated component consisting of a single $BG(v)$.

$BG(v)$ **is a Tree**

By definition, a tree is a connected acyclic graph. Connectivity implies the existence of a path between any pair of vertices; acyclicity implies the absence of cycles. In $BG(v)$, by definition, a path exists between any pair of vertices: (1) between the root and any vertex, there is a direct path; (2) between any pair of child vertices, a path exists via the root.

Furthermore, if a cycle existed in $BG(v)$, it could be a 1-cycle or a 2-cycle. However, the absence of non-trivial 1-cycles and 2-cycles in the Collatz sequence has been proven. Additionally, the absence of 1 or 2 cycles is not difficult to prove using formulas (2) and (3). Therefore, it can be asserted that any $BG(v)$ is acyclic. □

5.2.2 Trunk Graph $SG(k)$

By Definition 5, trunk graphs $SG(k)$, where $k = 5, 85, 341, \ldots$ (predecessors of the number 1), are formed from one of the predecessors of the number 1, by performing inverse recurrence computation starting from that predecessor, or by connecting the $BG(v)$s belonging to $SG(k)$.

Algorithm for Connecting $BG(v)$s. If the parent vertex of one $BG(v)$ coincides with one of the child vertices of another $BG(v)$, then the corresponding vertices are identified; consequently, such $BG(v)$s are connected vertically. Connecting $BG(v)$s in this manner yields $SG(k)$. Note that connecting $BG(v)$s is equivalent to performing inverse recurrence computation starting from one of the predecessors of 1.

Since the roots of $BG(v)$s are positive odd numbers not divisible by 3, the vertices of $BG(v)$s corresponding to odd numbers divisible by 3 do not participate in graph formation.

This raises the following question: Could there exist a set of mutually connected $BG(v)$s that are not connected to any $SG(k)$, i.e., to the graph having root 1?

Lemma 7. *For any base graph $BG(v)$, there exists a trunk graph $SG(k)$ such that*

$$BG(v) \subset SG(k),$$

where k is a predecessor of 1.

Proof. Assume there exists a $BG(v)$ not connected to any $SG(k)$. Then, since all its child vertices have their own predecessors, the $BG(v)$ under consideration must, through its child vertices, be connected to other BGs, and those in turn to other $BG(v)$s, and this would continue indefinitely. Now consider the root of this $BG(v)$. Two cases are possible: (1) The root of $BG(v)$ has no connection to another odd number not divisible by 3; (2) The root of $BG(v)$ has a connection to another odd number not divisible by 3.

The first case is only possible if the root number forms a cycle with itself; such a cycle exists only for the number 1. Therefore, the first case is impossible.

The second case means that the considered $BG(v)$, through its root vertex, is connected to another $BG(v)$, and that one in turn to a third $BG(v)$, and so on. Since all $BG(v)$s whose roots are predecessors of the number 1 form trunk graphs, it follows that any $BG(v)$ belongs to some $SG(k)$. The lemma is proven. □

From the above, it follows that the construction rules for $SG(k)$ and their branches are identical. The identity of construction rules implies: all $SG(k)$ are isomorphic as abstract graphs; they have the same local structure; the formulas for predecessors are the same.

Lemma 2 covers all odd numbers as "children" (in inverse computation), i.e., each element of the set of positive odd numbers not divisible by 3 (set M) has infinitely many predecessors. This might raise the question: Is every element of M, except 1, also a "parent" (has children in direct computation)? In this regard, we formulate the following lemma.

Lemma 8. *Let* $M = \{n \in \mathbb{N} \mid n \text{ is odd}, \ n \not\equiv 0 \pmod 3\}$. *Then*

$$\forall k \in M \setminus \{1\} \ \exists \ q \geq 1 : k = \frac{3x+1}{2^q},$$

where the result $k \in M$.

Proof. Assume there exists $k \in M$, $k \neq 1$, such that k is not a predecessor of any other $x \in M$ (i.e., it is not a parent, hence has no downward connection).

Since the Collatz function is defined for all natural numbers and continues indefinitely (in inverse computation) or reaches 1 (in direct computation), the trajectory from any vertex of a component must reach k in a finite number of steps. However, after reaching k in the forward direction, the computation must continue, meaning there must exist another $x \in M$ such that k is its predecessor. Therefore, any $k \in M$, $k \neq 1$, is a parent and thus has a downward connection. The lemma is proved.

A proof of Lemma 8 based on modular arithmetic is provided in Appendix 1. □

Lemma 9. *The trunk graph $SG(k)$ is a tree (connected acyclic graph).*

Absence of cycles in $SG(k)$. **Analytical method**

A cycle in a graph occurs if the end of a path (trajectory) coincides with its beginning. Therefore, if the root vertex coincides with one of the child vertices of $SG(k)$, a cycle is formed. Since every vertex in $SG(k)$ is both a child vertex (of a higher-level $BG(v)$) and a parent vertex (of a lower-level $BG(v)$ within the same trunk), the coincidence of the root and a child vertex would mean that two different numbers not divisible by 3 have equal predecessor numbers. According to Lemma 4, this cannot happen, as predecessor sets are disjoint, except for the number 1. Since the starting point of any trajectory in the considered $SG(k)$ is a predecessor of the number 1, this implies the absence of cycles in $SG(k)$.

Graphical method

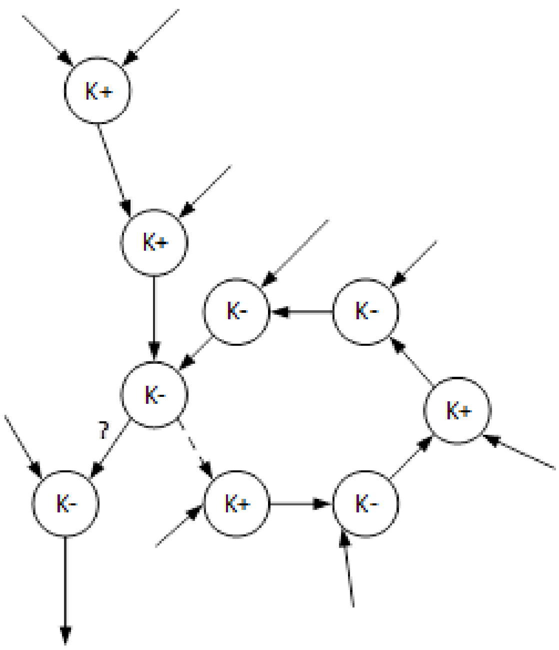
Figure 1 shows a scheme illustrating a hypothetical looped path formed by bending a branch of $SG(k)$. In this scenario, the edge connecting the branch to the tree would cease to exist, replaced by another edge (shown as a dotted line) connecting to the cycle, as any vertex at the bottom can have only one incoming edge (in the directed sense towards 1). This scheme demonstrates that if a looped path existed, it would exist separately from the main graph rooted at 1.

□

Thus, we have proven that $SG(k)$, obtained by connecting $BG(v)$s, is a tree with root 1 and contains no cycles.

Consequently, the complete Collatz graph has a single root, 1, from which infinitely many trunk graphs $SG(k)$ are formed. The roots of these trunk graphs correspond to the predecessors of the number 1. Each trunk graph consists of a set of $BG(v)$s and has branches at each node.

Figuratively speaking, $BG(v)$ (with a finite number of vertices) is like a dandelion head (pappus), and its multitude of "parachutes" are the vertices of $BG(v)$. If we connect other



:

Figure 1: Scheme illustrating cycle formation

dandelion heads to each ”parachute” of the first head, then connect more heads to each ”parachute” of the connected heads, and so on, continuing this process to a given height, we obtain an image of the Collatz graph. However, this would be an image of a Collatz graph formed from a single trunk. The complete Collatz graph, on the other hand, consists of infinitely many trunk graphs, all sharing the same root. Note that the first dandelion head is $BG(v)$ with root 1, and its ”parachutes” are the predecessors of the number 1.

Theorem 1. *The graph induced by the Collatz function on the set of natural numbers is a connected, directed, acyclic graph of a tree-like type with root at vertex 1. No other connected components disjoint from vertex 1 exist.*

Proof. Consider the directed Collatz graph, where vertices are natural numbers, and edges are defined by the action of the Collatz function. By Lemma 6, each base graph $BG(v)$ is a tree, and all natural numbers are covered by a system of such interconnected $BG(v)$s. By Lemma 7, each $BG(v)$ belongs to some trunk graph $SG(k)$, where k is a predecessor of 1.

Assume, for contradiction, that there exists a connected component not containing vertex 1. Let this component have a minimal root $k \in M$, $k \neq 1$, such that k is not a predecessor of any other $l \in M$ (i.e., no downward connection from k in the forward direction, meaning k is a ”leaf” in the component when considering edges oriented towards 1).

Since the Collatz function is defined for all natural numbers and the process continues indefinitely, the trajectory from any vertex in this component, when followed forward (according to $C(n)$), must eventually either:

- enter a cycle,
- diverge to infinity,

- or reach 1.

By construction of the component as a set of inverse trajectories (backwards from some set of numbers), every vertex in the component eventually reaches k after a finite number of forward steps (as k is assumed to be a leaf/minimal element). Now consider the forward trajectory from k itself:

- If the trajectory from k cycles, this contradicts Lemma 4 (disjoint predecessor sets implying uniqueness) and the established results on the non-existence of non-trivial k-cycles.
- If the trajectory from k diverges, this contradicts the fact that k belongs to a component generated by finite inverse steps from its own predecessors (the component exists, so there are numbers that eventually lead to k, implying k itself leads somewhere within the component, not to infinity).
- If the trajectory from k eventually reaches 1, then the component contains a path from k to 1, meaning 1 is in the component, contradicting the assumption that the component does not contain 1.
- The only remaining possibility consistent with the Collatz function's deterministic nature is that the trajectory from k leads to some number within the component other than k itself. This implies k has a child in the forward direction within the component, contradicting the definition of k as a leaf with no downward connection.

Therefore, a separate component without vertex 1 is impossible. Consequently, the Collatz graph is a single connected tree with root 1. All trajectories converge to 1.

Thus, every trajectory of the Collatz function on natural numbers converges to 1, which is equivalent to the truth of the Collatz conjecture. □

5.3 Base Bi-graph $BBG(g)$ and Polygraph $PG(q)$

To demonstrate the regularities in the formation of $BG(v)$ and $SG(k)$ with actual numbers, we consider their individual parts, namely the Base Bi-graph $BBG(g)$ and the Polygraph $PG(g)$. We recall that the regularities of $BG(v)$ apply to $BBG(g)$ as well, since each $BG(v)$ and its corresponding $BBG(g)$ share a common root (k_0).

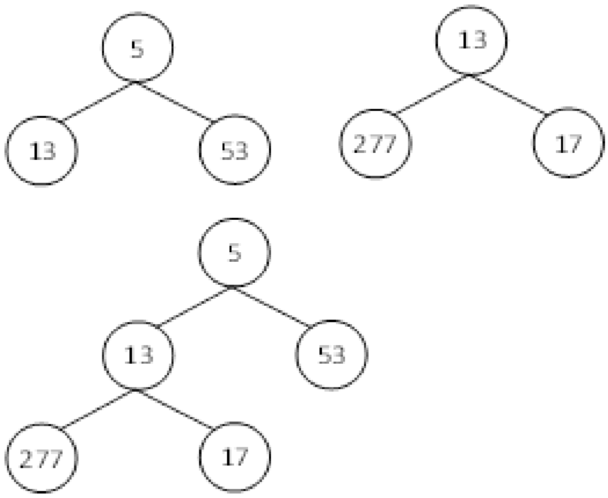
Algorithm for Connecting $BBG(g)$s. If the parent vertex of one $BBG(g)$ coincides with one of the child vertices of another $BBG(g)$, then the corresponding vertices are identified.

According to Lemma 4, the predecessor numbers for each element of the set of numbers not divisible by 3 are unique; i.e., they do not repeat for other elements. For convenience in constructing the graph, we will ignore numbers divisible by 3, as they do not affect the graph's structure.

Figure 2 shows an example of connecting two $BBG(1)$s, corresponding to the first power-of-two interval.

Next, we will examine in more detail $PG(1)$ corresponding to $SG(5)$, which is essentially a vertical "slice" of the complete Collatz graph.

The bottom vertex of any $BBG(1)$ corresponds to one element of the set of positive odd numbers not divisible by 3; i.e., the parent vertices of $BBG(1)$s are distinct. Each of the two child vertices of any $BBG(1)$ also corresponds to one element of the set of positive



:

Figure 2: Scheme illustrating cycle formation

odd numbers not divisible by 3; i.e., the child vertices of $BBG(1)$s are also distinct. It follows that the pair of child vertices of any $BBG(1)$ are the roots of two other $BBG(1)$s. Therefore, all $BBG(1)$s are interconnected and form $PG(1)$.

Thus, all $BBG(1)$s corresponding to $SG(5)$ are connected and form the polygraph $PG(1)$. Figure 2 shows the beginning of the directed $PG(1)$ corresponding to the first power-of-two interval.

$PG(1)$, shown in Figure 3, can also be obtained by performing inverse recurrence computation of the Collatz function based on the number 5 within the first power-of-two interval ($q = 1..6$). The computation proceeds as follows:

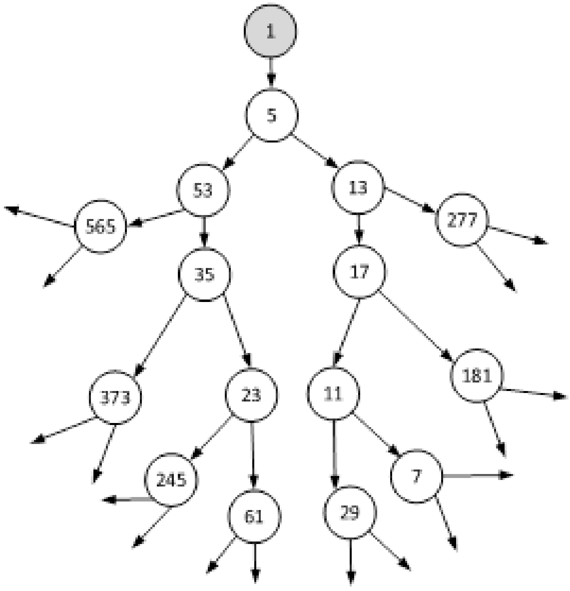
1. In the first iteration of inverse computation within the first power-of-two interval based on the number 5, we obtain the numbers 13 and 53, which are predecessors of 5.
2. Perform the same operation on the number 13, obtaining the numbers 17 and 277.
3. Perform the same operation on the number 53, obtaining the numbers 35 and 565.

Continuing this process yields $PG(1)$ up to any desired boundary number.

If all natural numbers are shown in $PG(1)$, it would resemble the graph shown in Figure 4. The figure shows a directed $PG(1)$ (corresponding to direct computation), containing even numbers ($2n$), numbers divisible by 3 ($3n$), and odd numbers not divisible by 3 (k^- and k^+).

5.4 Scheme of Connection Formation Between Vertices

Next, we show the formation of connections between vertices of $BBG(1)$s corresponding to $SG(5)$. For this, based on Table 1, we compile Table 2, where root numbers k_0 and their predecessors are presented according to their residues modulo 18. Table 2 shows that each root number k_0 has predecessors corresponding to 6 classes modulo 18 of odd numbers not divisible by 3, and these patterns repeat across the columns of the table. The residues also repeat every 18 powers of two; extending Table 3 would reveal this repetition of residues modulo 18.



:

Figure 3: Scheme illustrating cycle formation

Next, based on Table 2, we compile Table 3, which is divided into 6 sub-tables for each type of root number. The six sub-tables show the two predecessors k_1 and k_2 of the root number k_0, corresponding to the first power-of-two interval, along with their residues modulo 18. These six sub-tables also show the multipliers t_0, t_1, t_2 for the root and predecessor numbers k_0, k_1, k_2 respectively.

For convenience, we denote these numbers with two indices k_{ij}, where $i = 0, 1, 2$ indicates the status of the number (root or predecessor); $j = 1, 2, 3, 4, 5, 6$ indicates the type of number, i.e., the residue class modulo 18: 1(1); 2(5); 3(7); 4(11); 5(13); 6(17); $t_{ij} = 0, 1, 2, \dots$ is the multiplier.

Consequently, all odd numbers not divisible by 3, presented in the six sub-tables, have the following representations:

$$k_{i1} = 1 + 18t_{i1}; \quad k_{i2} = 5 + 18t_{i2}; \quad k_{i3} = 7 + 18t_{i3};$$
$$k_{i4} = 11 + 18t_{i4}; \quad k_{i5} = 13 + 18t_{i5}; \quad k_{i6} = 17 + 18t_{i6}.$$

Table 4 shows the formulas for the multipliers of the predecessor numbers. Note that all root numbers have multipliers spanning the entire set of non-negative integers. However, the multipliers of the predecessor numbers, obtained using the formulas in Table 4, do not cover the entire set of non-negative integers. This is because Table 4 is specifically for predecessors corresponding to the first power-of-two interval; formulas for multipliers corresponding to other power-of-two intervals are absent from Table 4. For example, none of the multiplier formulas in Table 4 yields a multiplier $t_{ij} = 90$, because the predecessor number $1621 = 1+18 \cdot 90$ corresponds to the root number 19 with a power of two exponent 8, i.e., it belongs to the second power-of-two interval.

As follows from the 6 sub-tables of Table 3, all root numbers and their predecessors are interconnected, as each of the 6 classes of odd numbers, depending on their multiplier, transforms into the same 6 classes of odd numbers. Since the multiplier of each root number corresponds to the entire set of non-negative integers, any of them is connected to others through their predecessors. This means that all $BBG(1)$s are interconnected.

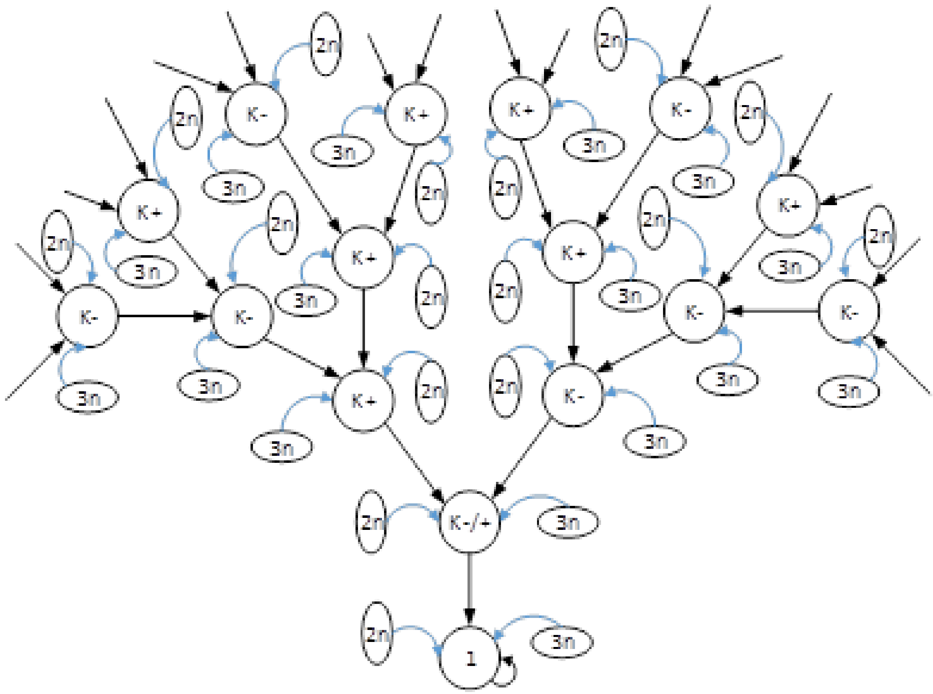


Figure 4: Scheme illustrating cycle formation

Example 2. *Take from Table 3.3 the number* $k_2 = 5$ *with multiplier* $t_2 = 72$*; this number is a predecessor of the root number* $k_0 = 7$ *with multiplier* $t = 3$ *(these numbers are on the same row of the table). Next, we look for a predecessor of the same type with the same multiplier; in Table 3.2 we find* $k_1 = 7$ *with multiplier* $t_1 = 3$*, which is a predecessor of the number* $k_0 = 5$ *with multiplier* $t = 1$*. Continuing this process, we compile the following chain:*

Step	*Source*	k *(type)*	t *(multiplier)*	k_0 *(root)*	t_0 *(multiplier)*
1	*Table 3.3*	$k_2 = 5$	$t_2 = 72$	$k_0 = 7$	$t = 3$
2	*Table 3.2*	$k_1 = 7$	$t_1 = 3$	$k_0 = 5$	$t = 1$
3	*Table 3.6*	$k_1 = 5$	$t_1 = 1$	$k_0 = 17$	$t = 1$
4	*Table 3.6*	$k_1 = 17$	$t_1 = 1$	$k_0 = 17$	$t = 2$
5	*Table 3.2*	$k_2 = 17$	$t_2 = 2$	$k_0 = 5$	$t = 0$
6	*Table 3.1*	$k_2 = 5$	$t_2 = 0$	$k_0 = 1$	$t = 0$

Using actual numbers, the above example corresponds to the following sequence: $1301 \to 61 \to 23 \to 35 \to 53 \to 5 \to 1$.

Thus, we have demonstrated that $PG(1)$, corresponding to $SG(5)$, is a tree rooted at 1, and all $BBG(1)$s forming $PG(1)$ are interconnected with root 1.

We have also shown that by performing inverse computation within the first power-of-two interval, we can construct $PG(1)$ to any depth, which consists of a set of $BBG(1)$s, each of which is entirely contained within $PG(1)$. Consequently, if $BBG(1)$s are interconnected, they are all entirely contained within $PG(1)$. It should be noted that $BBG(g)$s belonging to other power-of-two intervals corresponding to $SG(5)$ are directly connected to $PG(1)$, as they are branches of $PG(1)$.

Conclusion

This work has shown that computing the Collatz function on the set of positive odd numbers not divisible by 3 is equivalent to computing the Collatz function on the entire set of natural numbers. It was further proven that by performing inverse iterative computation using the formula $((6n \pm 1) \cdot 2^q - 1)/3$, with the exponent of two q increasing sequentially, each element of the set of positive odd numbers not divisible by 3 corresponds to an infinite set of positive odd predecessor numbers, forming arithmetic progressions. Each element of the set of positive odd numbers not divisible by 3, together with its predecessors, forms a base graph $BG(v)$.

For any base graph $BG(v)$, there exists a trunk graph $SG(k)$, which is formed by connecting the $BG(v)$s belonging to $SG(k)$ or by performing inverse recurrence computation starting from one of the predecessors of the number 1. It follows that the graph of the Collatz function consists of interconnected base graphs $BG(v)$, which are united into a single tree-like structure with root at vertex 1. Since this structure covers all natural numbers and excludes the existence of isolated cycles or disconnected components, all trajectories of the Collatz function converge to the number 1, which is equivalent to the truth of the Collatz conjecture.

Acknowledgments

I am grateful to my longtime friend Mars Gabbasov for valuable comments, as well as to all purchasers of my book "The Hardest Problem of Mathematics: 3n+1. Ingenious Solution."

References

[1] J. C. Lagarias, *The Ultimate Challenge: The 3x+1 Problem.* American Mathematical Society, 2010.

[2] R. K. Guy, *Unsolved Problems in Number Theory* (3rd ed.). Springer-Verlag, 2004.

[3] L. Collatz, On the motivation and origin of the $(3n+1)$ Problem, *J. Qufu Normal University, Natural Science Edition*, 12(3) (1986) 9–11.

[4] T. Tao, Almost all orbits of the Collatz map attain almost bounded values, *Forum of Mathematics, Pi*, 10, E12. doi:10.1017/fmp.2022.8, 2022.

[5] K. Sultan, *The Hardest Problem of Mathematics: 3n+1. Ingenious Solution*, ISBN-13:979-8840633250, 2022.

[6] R. P. Steiner, A theorem on the Syracuse problem, *Proceedings of the 7th Manitoba Conference on Numerical Mathematics*, pp. 553–559, 1977.

[7] J. L. Simons, On the nonexistence of 2-cycles for the $3x+1$ problem, *Math. Comp.*, 74:1565–1572, 2005.

[8] J. Simons, B. de Weger, Theoretical and computational bounds for m-cycles of the $3n+1$ problem, *Acta Arithmetica*, 117(1):51–70, 2005.

[9] C. Hercher, There are no Collatz m-cycles with $m \leq 91$, *Journal of Integer Sequences*, 26(3):Article 23.3.5, 2023.

Table 1. Predecessor Numbers

	$3+12t$	$1+24t$	$13+48t$	$5+96t$	$53+192t$	$21+384t$	$213+768t$	$85+1536t$	$853+3072t$	$341+6144t$	$3413+12288t$	$1365+24576t$
	$7+12t$	$9+24t$	$29+48t$	$37+96t$	$117+192t$	$149+384t$	$469+768t$	$597+1536t$	$1877+3072t$	$2389+6144t$	$7509+12288t$	$9557+24576t$
	$11+12t$	$17+24t$	$45+48t$	$69+96t$	$181+192t$	$277+384t$	$725+768t$	$1109+1536t$	$2901+3072t$	$4437+6144t$	$11605+12288t$	$17749+24576t$
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1		1		5		21		85		341		1365
5	3		13		53		213		853		3413	
7		9		37		149		597		2389		9557
11	7		29		117		469		1877		7509	
13		17		69		277		1109		4437		17749
17	11		45		181		725		2901		11605	
19		25		101		405		1621		6485		25941
23	15		61		245		981		3925		15701	
25		33		133		533		2133		8533		34133
29	19		77		309		1237		4949		19797	
31		41		165		661		2645		10581		42325
35	23		93		373		1493		5973		23893	
37		49		197		789		3157		12629		50517
41	27		109		437		1749		6997		27989	
43		57		229		917		3669		14677		58709
47	31		125		501		2005		8021		32085	
49		65		261		1045		4181		16725		66901
53	35		141		565		2261		9045		36181	

Note: This is only the initial part of a table of infinite size.

Table 2. Formulas for calculating the predecessor numbers of non-multiples of 3 on the first interval of a power of two

k_0	q = 1	$q = 2$	$q = 3$	$q = 4$	$q = 5$	$q = 6$
$1 + 18t$		$1 + 24t$		$5 + 96t$		
$5 + 18t$			$13 + 48t$		$53 + 192t$	
$7 + 18t$				$37 + 96t$		$149 + 384t$
$11 + 18t$	$7 + 12t$		$29 + 48t$			
$13 + 18t$		$17 + 24t$				$277 + 384t$
$17 + 18t$	$11 + 12t$				$181 + 192t$	

Table 3.Root Numbers and Their Predecessors Represented by Residues Modulo 18

#	r	k_0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1	1	**1**		1		5		3		13		17		15		7		11		9
2	5	**5**	3		13		17		15		7		11		9		1		5	
3	7	**7**		9		1		5		3		13		17		15		7		11
4	11	**11**	7		11		9		1		5		3		13		17		15	
5	13	**13**		17		15		7		11		9		1		5		3		13
6	17	**17**	11		9		1		5		3		13		17		15		7	
7	1	**19**		7		11		9		1		5		3		13		17		15
8	5	**23**	15		7		11		9		1		5		3		13		17	
9	7	**25**		15		7		11		9		1		5		3		13		17
10	11	**29**	1		5		3		13		17		15		7		11		9	
11	13	**31**		5		3		13		17		15		7		11		9		1
12	17	**35**	5		3		13		17		15		7		11		9		1	
13	1	**37**		13		17		15		7		11		9		1		5		3
14	5	**41**	9		1		5		3		13		17		15		7		11	
15	7	**43**		3		13		17		15		7		11		9		1		5
16	11	**47**	13		17		15		7		11		9		1		5		3	
17	13	**49**		11		9		1		5		3		13		17		15		7
18	17	**53**	17		15		7		11		9		1		5		3		13	
19	1	**55**		1		5		3		13		17		15		7		11		9
20	5	**59**	3		13		17		15		7		11		9		1		5	
21	7	**61**		9		1		5		3		13		17		15		7		11
22	11	**65**	7		11		9		1		5		3		13		17		15	
23	13	**67**		17		15		7		11		9		1		5		3		13
24	17	**71**	11		9		1		5		3		13		17		15		7	
25	1	**73**		7		11		9		1		5		3		13		17		15
26	5	**77**	15		7		11		9		1		5		3		13		17	
27	7	**79**		15		7		11		9		1		5		3		13		17
28	11	**83**	1		5		3		13		17		15		7		11		9	
29	13	**85**		5		3		13		17		15		7		11		9		1
30	17	**89**	5		3		13		17		15		7		11		9		1	
31	1	**91**		13		17		15		7		11		9		1		5		3
32	5	**95**	9		1		5		3		13		17		15		7		11	
33	7	**97**		3		13		17		15		7		11		9		1		5
34	11	**101**	13		17		15		7		11		9		1		5		3	
35	13	**103**		11		9		1		5		3		13		17		15		7
36	17	**107**	17		15		7		11		9		1		5		3		13	

Table 4.1: $k_{01} = 1 + 18t$

k_{01}	t	k_{11}	t_{11}	k_{21}	t_{21}
1	0	1	0	5	0
1	1	7	1	11	5
1	2	13	2	17	10
1	3	1	4	5	16
1	4	7	5	11	21
1	5	13	6	17	26
1	6	1	8	5	32
1	7	7	9	11	37
1	8	13	10	17	42

Table 4.2: $k_{02} = 5 + 18t$

k_{02}	t	k_{12}	t_{12}	k_{22}	t_{22}
5	0	13	0	17	2
5	1	7	3	11	13
5	2	1	6	5	24
5	3	13	8	17	34
5	4	7	11	11	45
5	5	1	14	5	56
5	6	13	16	17	66
5	7	7	19	11	77
5	8	1	22	5	88

Table 4.3: $k_{03} = 7 + 18t$

k_{03}	t	k_{13}	t_{13}	k_{23}	t_{23}
7	0	1	2	5	8
7	1	7	7	11	29
7	2	13	12	17	50
7	3	1	18	5	72
7	4	7	23	11	93
7	5	13	28	17	114
7	6	1	34	5	136
7	7	7	39	11	157
7	8	13	44	17	178

Table 4.4: $k_{04} = 11 + 18t$

k_{04}	t	k_{14}	t_{14}	k_{24}	t_{24}
11	0	7	0	11	1
11	1	1	1	5	4
11	2	13	1	17	6
11	3	7	2	11	9
11	4	1	3	5	12
11	5	13	3	17	14
11	6	7	4	11	17
11	7	1	5	5	20
11	8	13	5	17	22

Table 4.5: $k_{05} = 13 + 18t$

k_{05}	t	k_{15}	t_{15}	k_{25}	t_{25}
13	0	17	0	7	15
13	1	5	2	13	36
13	2	11	3	1	58
13	3	17	4	7	79
13	4	5	6	13	100
13	5	11	7	1	122
13	6	17	8	7	143
13	7	5	10	13	164
13	8	11	11	1	186

Table 4.6: $k_{06} = 17 + 18t$

k_{06}	t	k_{16}	t_{16}	k_{26}	t_{26}
17	0	11	0	1	10
17	1	5	1	13	20
17	2	17	1	7	31
17	3	11	2	1	42
17	4	5	3	13	52
17	5	17	3	7	63
17	6	11	4	1	74
17	7	5	5	13	84
17	8	17	5	7	95

Table 5. Formulas for multipliers of predecessor numbers

$\mathbf{k_{01} = 1 + 18t}$			
$k_1 = 1 + 18t_1$	$t_1 = 4x$	$k_2 = 5 + 18t_2$	$t_2 = 16x$
$k_1 = 7 + 18t_1$	$t_1 = 1 + 4x$	$k_2 = 11 + 18t_2$	$t_2 = 5 + 16x$
$k_1 = 13 + 18t_1$	$t_1 = 2 + 4x$	$k_2 = 17 + 18t_2$	$t_2 = 10 + 16x$
$\mathbf{k_{02} = 5 + 18t}$			
$k_1 = 13 + 18t_1$	$t_1 = 8x$	$k_2 = 17 + 18t_2$	$t_2 = 2 + 32x$
$k_1 = 7 + 18t_1$	$t_1 = 3 + 8x$	$k_2 = 11 + 18t_2$	$t_2 = 13 + 32x$
$k_1 = 1 + 18t_1$	$t_1 = 6 + 8x$	$k_2 = 5 + 18t_2$	$t_2 = 24 + 32x$
$\mathbf{k_{03} = 7 + 18t}$			
$k_1 = 1 + 18t_1$	$t_1 = 2 + 16x$	$k_2 = 5 + 18t_2$	$t_2 = 8 + 64x$
$k_1 = 7 + 18t_1$	$t_1 = 7 + 16x$	$k_2 = 11 + 18t_2$	$t_2 = 29 + 64x$
$k_1 = 13 + 18t_1$	$t_1 = 12 + 16x$	$k_2 = 17 + 18t_2$	$t_2 = 50 + 64x$
$\mathbf{k_{04} = 11 + 18t}$			
$k_1 = 7 + 18t_1$	$t_1 = 2x$	$k_2 = 11 + 18t_2$	$t_2 = 1 + 8x$
$k_1 = 1 + 18t_1$	$t_1 = 1 + 2x$	$k_2 = 5 + 18t_2$	$t_2 = 4 + 8x$
$k_1 = 13 + 18t_1$	$t_1 = 1 + 2x$	$k_2 = 17 + 18t_2$	$t_2 = 6 + 8x$
$\mathbf{k_{05} = 13 + 18t}$			
$k_1 = 17 + 18t_1$	$t_1 = 4x$	$k_2 = 7 + 18t_2$	$t_2 = 15 + 64x$
$k_1 = 5 + 18t_1$	$t_1 = 2 + 4x$	$k_2 = 13 + 18t_2$	$t_2 = 36 + 64x$
$k_1 = 11 + 18t_1$	$t_1 = 3 + 4x$	$k_2 = 1 + 18t_2$	$t_2 = 58 + 64x$
$\mathbf{k_{06} = 17 + 18t}$			
$k_1 = 11 + 18t_1$	$t_1 = 2x$	$k_2 = 1 + 18t_2$	$t_2 = 10 + 32x$
$k_1 = 5 + 18t_1$	$t_1 = 1 + 2x$	$k_2 = 13 + 18t_2$	$t_2 = 20 + 32x$
$k_1 = 17 + 18t_1$	$t_1 = 1 + 2x$	$k_2 = 7 + 18t_2$	$t_2 = 31 + 32x$

Notes:

1. $t, x = 0, 1, 2, \ldots$
2. The six sub-tables in Table 3 show only initial root numbers and their predecessors. In reality, there are infinitely many such numbers; nevertheless, the above multiplier formulas apply to all numbers.
3. Tables 3 and 4 illustrate the mechanism for forming Collatz function trajectories and the reason for the zigzag pattern of the graph.

A Proof of Lemma 8 Based on Modular Arithmetic

For k to have a "downward connection" in the inverse Collatz graph, we need to find an $x \in M$ such that a direct application of the Collatz function (the operation $3x + 1$ followed by division by 2^q) yields k.

In terms of an inverse step, this means:

$$x = \frac{k \cdot 2^q - 1}{3},$$

where x must be an integer, odd, and not divisible by 3 (i.e., $x \in M$).

The number x is divisible by 3 (i.e., $x \notin M$) if and only if

$$k \cdot 2^q - 1 \equiv 0 \pmod 9 \Rightarrow k \cdot 2^q \equiv 1 \pmod 9.$$

The powers of 2 modulo 9 have a cycle of length 6: $2, 4, 8, 7, 5, 1$. Since $k \in M$, $k \bmod 9 \in \{1, 2, 4, 5, 7, 8\}$. Consider all possible cases for $k \bmod 9$ and the parity of q (even or odd) required for x to be an integer (from the condition $x \cdot 2^q \equiv 1 \pmod 3$). For each combination, examine the values of $k \cdot 2^q \bmod 9$ over the relevant half-cycles of q (the 3 allowed q values per parity).

In each category and for each possible $k \bmod 9$:

- In each half-cycle of three allowed q values, exactly one value yields $k \cdot 2^q \equiv 1 \pmod 9$ (meaning x would be divisible by 3).
- The other two yield $k \cdot 2^q \not\equiv 1 \pmod 9$ (meaning x is not divisible by 3, and thus $x \in M$).

This occurs because multiplication by a fixed k (where $\gcd(k, 9) = 1$) is a bijection in the multiplicative group $(\mathbb{Z}/9\mathbb{Z})^\times$, which permutes the elements of the cycle of powers of 2 but does not fix them all in one class. Since the cycle period is 6 and the half-cycles (even/odd q) have length 3, the distribution holds: 2/3 "good" q values in each half-cycle. Consequently, for any $k \in M$.