

Article

Not peer-reviewed version

Symbolic Dynamic Formulation for the Collatz Conjecture: II. Quasi-Global Behavior

[Eric Sakk](#)*

Posted Date: 30 June 2025

doi: 10.20944/preprints202506.2510.v1

Keywords: collatz conjecture; dynamical systems; stability; symbolic dynamics



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Symbolic Dynamic Formulation for the Collatz Conjecture: II. Quasi-global Behavior

Eric Sakk

Department of Computer Science, Morgan State University, Baltimore, MD, 21251; eric.sakk@morgan.edu

Abstract

In this work, we continue the development of a symbolic dynamical formulation for the Collatz mapping. This formulation has been previously demonstrated to induce a ternary alphabet for characterizing the behavior of trajectories generated by this mapping. While the previous installation of this series focused on local and quasi-global behaviors, this work presents more quasi-global results that contribute to further understanding of the Collatz mapping.

Keywords: collatz conjecture; symbolic dynamics; expansive sequences; system stability

1. Introduction

In this work, we review and further characterize a symbolic dynamical formulation for the discrete dynamical system

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

for any $x \in \mathbb{Z}^+$. Much has been written about this mapping [1] and it has been observed (but, not rigorously proved) that all initial conditions tested so far converge to the periodic solution $1 \rightarrow 2 \rightarrow 4 \rightarrow 1 \rightarrow \dots$ [4]. The goal of this paper is to continue and extend previous work that introduced this formulation [5]. We begin this report by reviewing the main features and definitions in Section 2. We then present results that extend various quasi-global characterizations previously presented.

2. Summary of Relevant Definitions

Definition 2.1. A *node* with respect to Equation (1) is any $x \in \mathbb{Z}^+$ such that x is odd.

Definition 2.2. A *hub* is a node that is not of the form $4k + 1$ for any positive odd integer $k = 1, 3, 5, \dots$.

The hub is the starting point for an object we refer to as a *C-ladder*. One side of the C-ladder is formed by starting with a hub and then *pumping the associated odd number k using successive applications of $4k + 1$* ; hence, one side of the ladder consists of a sequence of nodes. In base 2, this sequence is formed by two left shifts then adding a value of one. The 'rungs' of the ladder are generated using the mapping $f(x) = 3x + 1$. The other side of a C-ladder is naturally induced by Equation (1). For example, the odd integer 3 is a hub and generates the C-ladder

$$\begin{array}{ccccccc} 3 & \Rightarrow & 13 & \Rightarrow & 53 & \Rightarrow & 213 & \Rightarrow & \dots \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ 10 & \leftarrow & 40 & \leftarrow & 160 & \leftarrow & 640 & \leftarrow & \dots \end{array}$$

In this diagram, $4k + 1$ pumps are denoted by ' \Rightarrow ' and application of Equation (1) is denoted by ' \rightarrow '. Finally, we have the following

Definition 2.3. A *C-tree* is the connected union of all C-ladders.

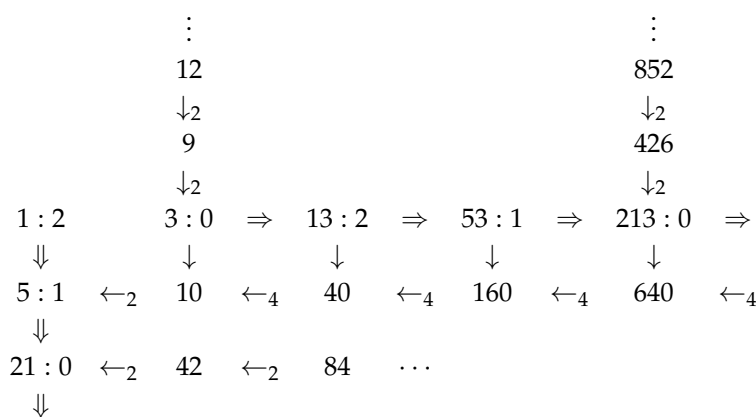
Definition 2.4. A *0-node* is an odd integer of the form $6m + 3$ for any nonnegative integer $m = 0, 1, 2, \dots$.

Since $6m + 3 = 3(2m + 1)$, it should be clear that set of all 0-nodes represents the set of nonnegative odd integers that are divisible by three.

Definition 2.5. A 1-node is an odd integer of the form $6m + 5$ for any nonnegative integer $m = 0, 1, 2, \dots$.

Definition 2.6. A 2-node is an odd integer of the form $6m + 1$ for any nonnegative integer $m = 0, 1, 2, \dots$.

Justification for these node definitions is given in Appendix A. It was thought that the justification was obvious; however, several comments on previous work indicate an explanation is needed. If a trajectory contains a sequence of 1-nodes, asymptotically, values will be multiplied by a factor of $3/2$. If a trajectory contains a sequence of 2-nodes, then values will be multiplied by a factor of $3/4$. Hence, 1-nodes are the result of a local increase (expansion) and 2-nodes are the result of a local decrease (dissipation). According to Equation (1), 1-nodes and 2-nodes must always have both an even preimage and an odd preimage. 0-nodes only possess an even preimage. The node information can be depicted as part of the directed graph introduced above



where ' \leftarrow_2 ' means divide by 2 and ' \leftarrow_4 ' means divide by 4. Using this approach, we can represent a trajectory using the symbols $\{0, 1, 2\}$. For example, the sequence generated by the Collatz mapping $9 \rightarrow 7 \rightarrow 11 \rightarrow 17$ can be represented as 0211. This mod 6 characterization is useful as it makes clear the local stability properties of a node. It tells us what has happened as a result of a transition from a previous hub.

3. Periodic Behavior

Much has been written regarding the periodic behavior of the Collatz mapping [5,6]. In this section, we highlight a specific property useful for relating symbolic sequences. To do this, we briefly review the mod 8 characterization that tells us about the next transition (as opposed to telling us about the previous transition using the mod 6 approach). Given a nonnegative integer p , if the current node is of the form

- $8p + 1$, then the current node will transition to a 2-node
- $8p + 3$ or $8p + 7$, then the current node will transition to a 1-node
- $8p + 5$, then the current node is $4k+1$ pump

where the 'x' in the bottom line means a $4k + 1$ pump, $1^{(3)}$ means $8p + 3$ will transition to a 1-node, $1^{(7)}$ means $8p + 7$ will transition to a 1-node and $2^{(1)}$ means $8p + 1$ will transition to a 2-node. The periodic waltz motif '1-2-1-x-1-2-1-x ...' has been pointed out in [5]. In this work, we delve a bit deeper into its significance. It can be readily seen that simply adding a value of 2 to a node will transform it periodically to a another node form. However, this observation is a special case of a more general rule.

Theorem 3.1. Given a node r and a nonnegative, even integer n , successively adding the term $2 \cdot 3^n$ will result in the periodic transition sequence

$$\dots 1^{(7)} - 2^{(1)} - 1^{(3)} - x - 1^{(7)} - 2^{(1)} - 1^{(3)} - x \dots$$

Given the table above, this statement is clearly true for $n = 0$. Appendix B provides the proof for this statement and the following statement.

Theorem 3.2. Given a node r and a nonnegative, odd integer n , successively adding the term $2 \cdot 3^n$ will result in the periodic transition sequence

$$\dots 1^{(3)} - 2^{(1)} - 1^{(7)} - x - 1^{(3)} - 2^{(1)} - 1^{(7)} - x \dots$$

Qualitatively, these theorems are true because they effectively result in the downsampling of Table 1 by a factor of $2 \cdot 3^n$ for a specific value of n . Given this periodic behavior, it is important to note that downsampling with a value of $4 \cdot 2 \cdot 3^n$ leads to an invariance where the node mapping remains constant.

Table 1. Periodic behavior of the mod 8 formulation.

$8p+1:$	1				9			17	
$8p+3, 8p+7:$		3		7		11		15	19
$8p+5:$			5				13		
	$2^{(1)}$	$1^{(3)}$	x	$1^{(7)}$	$2^{(1)}$	$1^{(3)}$	x	$1^{(7)}$	$2^{(1)}$
									$1^{(3)}$

As an example of applying these theorems, consider a starting value of 3 which is of the form $8p + 3$ and will map to a $1^{(3)}$ -node having a value of 5 upon application of Equation 1. Consider successively adding a value of $2 \cdot 3^1 = 6$, Theorem 3.2 implies that the next value computed should be of the form $8p + 1$ and map to a 2-node. In other words $3 + 6 = 9$ which is of the form $8p + 1$ is a 2-node. Adding a value of 6 again leads to $9 + 6 = 15$ which of the form $8p + 7$ and will map to a $1^{(7)}$ -node having a value of 23. Adding a value of 6 again leads to $15 + 6 = 21$ which of the form $8p + 5$ and, hence, is a $4k + 1$ pump. Finally, the cycle repeats with $21 + 6 = 27$ which is of the form $8p + 3$ and will map to a $1^{(3)}$ -node and so on. Additionally, it should be clear that adding a value of $4 \cdot 2 \cdot 3^1 = 24$ to any starting node will periodically result in a node with similar behavior.

4. Trajectory Catalog

4.1. Catalog of Primitive Itineraries

Recalling some basic terms from [5],

Definition 4.1. A *primitive itinerary* of Equation (1) is a sequence that begins at a 0-node and ends at a pumped $4k + 1$ node where k is odd.

More directly, a primitive itinerary begins at a 0-node and *traverses a sequence of hubs* that are either 1-nodes and 2-nodes until a pumped node is encountered.

Example 4.2. $21 : 0$ is a primitive itinerary of length one as it is both a 0-node and a pumped $4k + 1$ node where $k = 5$.

Example 4.3. $15 : 0 \rightarrow 23 : 1 \rightarrow 35 : 1 \rightarrow 53 : 1$ is a primitive itinerary that generates the sequence $\{0111\}$.

If the value of the pumped node is less than that of the 0-node, then we say the primitive itinerary is *dissipative*. If the value of the pumped node is greater than that of the 0-node, then we say the primitive itinerary is *expansive*. In [5], it was demonstrated how to construct a catalog of all possible primitive itineraries given the set of all 0-nodes. The reasoning behind doing this is so that trajectories having the same symbolic sequence can be grouped into families with similar

dynamics. This formulation goes a bit further than the encoding sequence approach of [6] and, as will be demonstrated shortly, will enable a method for understanding the dynamics beyond a given primitive itinerary. This process is important to understand because another major dissipative effect is falling down through the rungs of a Collatz ladder which results in a decrease of at least $3/8$ from the pumped node.

To see how to apply the theorems of the previous section, consider Table 2.

Table 2. Periodicity of primitive itinerary 01.

Initial condition	Trajectory	Sequence
3:	$3 \rightarrow 5$	01]
15:	$15 \rightarrow 23 \rightarrow 35 \rightarrow 53$	0111]
27:	$27 \rightarrow 41 \rightarrow 31 \rightarrow 47 \rightarrow 71 \dots$	01211\dots
39:	$39 \rightarrow 59 \rightarrow 89 \rightarrow 67 \rightarrow 101$	01121]
51:	$51 \rightarrow 77$	01]
63:	$63 \rightarrow 95 \rightarrow 143 \rightarrow 215 \rightarrow 323 \rightarrow 485$	011111]
75:	$75 \rightarrow 113 \rightarrow 85$	012]
87:	$87 \rightarrow 131 \rightarrow 197$	011]
99:	$99 \rightarrow 149$	01]

In this table, we use the symbol ']' to indicate the last node is the end of a primitive itinerary (i.e. the last node is a $4k + 1$ pump). By successively adding a value of $2^2 \cdot 3 = 12$ to the initial condition we periodically revisit the primitive sequence 01 of length 2. This result easily follows from the results of [6] where it is known that encoding vectors of length k are periodic modulo 2^k . In this work, we are particularly interested in the special case where the period is taken modulo $2^k \cdot 3$. Additionally, this table also demonstrates the additive rules in Appendix B (where the 'x-1-2-1-x' motif is clearly present) as well as the transition rules of Appendix C.

One important consequence of this formulation when combined with results of [6] is that when considering primitive itineraries of length L , when initial conditions are separated by a value of

$$2^{L+n_2} \cdot 3 \quad (2)$$

(where $n_2 =$ the number of 2-nodes in the sequence), the values at the end of the sequence will be separated by a value of

$$2 \cdot 3^L. \quad (3)$$

This is why the theorems of the previous section are central to the theme of this section. It is straightforward to check the above table where $L = 2$ and $n_2 = 0$ that initial conditions separated by a value of $2^2 \cdot 3$ lead to values separated by a value of $2 \cdot 3^2 = 18$. Considering the above example and using the well-known formula

$$f^{p-1}(2^p \cdot h - 1) = 2 \cdot 3^{p-1} \cdot h - 1.$$

for the case where $n_2 = 0$. If we let $x_1 = 2^p \cdot h - 1$, then $f^{p-1}(x_1) = 2 \cdot 3^{p-1} \cdot h - 1$. If $x_2 = 2^p \cdot (h + 3) - 1$, then $f^{p-1}(x_2) = 2 \cdot 3^{p-1} \cdot (h + 3) - 1$. It follows that $x_2 - x_1 = 2^p \cdot 3$ and $f^{p-1}(x_2) - f^{p-1}(x_1) = 2 \cdot 3^p$ as Equations (2) and (3) also indicate. Given the 'x-1-2-1-x' periodicity, Equations (2) and (3) imply that the final value of equivalent primitive itineraries of length L initially separated by a value of $4 \cdot 2^{L+n_2} \cdot 3$ must always be separated by a value of $4 \cdot 2 \cdot 3^L$ regardless of the value of n_2 . We consider the value $4 \cdot 2 \cdot 3^L$ to be an important 'anchor', 'invariant' or 'constant of the motion' for primitive itineraries of length L .

4.2. Traversal of Intersecting Primitive Itineraries

Once the end of a primitive itinerary is encountered, a given trajectory will fall down a Collatz ladder by at least a factor of $3/8$ depending how far that trajectory has been pumped and the type of node encountered at the bottom of the ladder. Furthermore, the trajectory will land at the image of

a hub contained within another primitive itinerary constrained by the addition and transition rules of Appendices B and C. Based upon the addition and transition rules, the node encountered at the bottom of the ladder can either be a 1-node hub, a 2-node hub (either of which can continue along another primitive itinerary) or a $4k + 1$ pumped node (which can be either a 1-node or a 2-node).

As mentioned above, the final value of equivalent primitive itineraries of length L initially separated by a value of $4 \cdot 2^{L+n_2} \cdot 3$ must always be separated by a value of $\hat{\Delta} \equiv 4 \cdot 2 \cdot 3^L$ (which is independent of the value of n_2). Therefore, $4k + 1$ pumps at the end of a primitive itinerary of length L are periodically separated by a value of $\hat{\Delta}$. This implies that there exists at least one pump below the end of the primitive itinerary where these 'pre-pumps' are periodically separated by a value of $\hat{\Delta}/4 = 2 \cdot 3^L$. The addition rules of Appendix B then imply that the pumped node (the 'pre-pump') directly below that of the end of the primitive itinerary will periodically obey the 'x-1-2-1-x' pattern. This result is of importance because it enables us to characterize the least decrease after falling down the Collatz ladder. The pre-pump will either be of the form $8k + 1, 8k + 3, 8k + 5, 8k + 7$. Nodes of the form $8k + 3$ or $8k + 7$ will map to 1-node and this would represent the drop of least decrease (as opposed to $8k + 1$ or $8k + 5$). So, let us consider, for example, the case where a 1-node is encountered after a pre-pump. Initial values separated by a value of $4 \cdot 4 \cdot 2^{L+n_2} \cdot 3$ would lead to the end of primitive itineraries separated by $4\hat{\Delta} = 4 \cdot 4 \cdot 2 \cdot 3^L$. Assuming the drop leads to a $1^{(7)}$ -node from the pre-pump (if not, keep cycling through x-1-2-1-x until a 1-node is encountered), the resulting 1-nodes would be separated by a value of $2 \cdot 2 \cdot 3^{L+1} = 4 \cdot 3 \cdot 3^L$. This implies that the Collatz mapping always has the effect of contracting sets of primitive itineraries of length L (for any value of L) after the terminal is encountered. In other words, in the worst case, when the terminals of primitive itineraries of length L are separated by a value of $4 \cdot 4 \cdot 2 \cdot 3^L = 32 \cdot 3^L$, there exists a set of 1-nodes in the next primitive itinerary separated by a value of $4 \cdot 3 \cdot 3^L = 12 \cdot 3^L$. For the 2-node case, there will exist a set of 2-nodes in the next primitive itinerary separated by a value of $2 \cdot 3 \cdot 3^L = 6 \cdot 3^L$ where an extra factor of $1/2$ must be introduced. If a $4k + 1$ pump is encountered, there will continue to be an even greater contractive effect. So, the 1-node case is the 'least contractive' case where primitive itineraries of length L will be separated by a value of $12 \cdot 3^L$ after dropping down the Collatz ladder to 1-node hub.

5. Conclusions

In this work, we continue a symbolic dynamical formalism for characterizing all trajectories generated by Equation (1). This symbolic approach enables the establishment of dynamical similarities within the presented formalism. Furthermore, this framework is useful for drawing dynamical conclusions when, after dropping down a Collatz ladder, primitive itineraries intersect new primitive itineraries.

Appendix A. Justification of Mod 6 for Node Definitions

1-Node Definition

Looking for odd images of numbers that obey

$$\begin{aligned} x &= \frac{3(2n+1)+1}{2} \\ &= 3n+2 \end{aligned}$$

Only when $n = 2p + 1$, can x be an odd number.

$$\begin{aligned} x &= 3(2p+1)+2 \\ &= 6p+5 \end{aligned}$$

2-node definition

Looking for odd images of numbers that obey

$$\begin{aligned} x &= \frac{3(2n+1)+1}{4} \\ &= \frac{3}{2}n + 1 \end{aligned}$$

Only when $n = 4p$, can x be an odd number.

$$\begin{aligned} x &= \frac{3}{2}(4p) + 1 \\ &= 6p + 1 \end{aligned}$$

Appendix B. Proofs of Addition Rules*Appendix B.1. Transform $2^{(1)} \rightarrow 1$*

Let \mathbb{N}_0 represent the set of non-negative integers. Since transitions to 1-nodes involve both $8p + 3$ and $8p + 7$, we divide $2 \rightarrow 1$ transitions into 2 cases.

Appendix B.1.1. $2^{(1)} \rightarrow 1^{(3)}$:

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 1 + 2 \cdot 3^{2m} = 8q + 3$$

Proof:

Rearranging the above equation leads to

$$3^{2m} = 4t + 1$$

where $t = q - p$. Using induction, we can see that when $m = 0$, there exists $t = 0$ so the base case holds. We then assume that there exists a t such that

$$3^{2m} = 4t + 1 \tag{4}$$

is true. We must demonstrate that there exists an r such that

$$3^{2(m+1)} = 4r + 1$$

Using Equation 4, it follows that such an r must obey

$$3^2 \cdot 3^{2m} = 4r + 1 \Rightarrow 3^2(4t + 1) = 4r + 1$$

This condition implies that the equality

$$36t + 9 = 4r + 1$$

must hold. Therefore, given t we have that r exists according to the condition

$$r = 9t + 2.$$

Appendix B.1.2. $2^{(1)} \rightarrow 1^{(7)}$:

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 1 + 2 \cdot 3^{2m+1} = 8q + 7$$

Proof:

Rearranging the above equation leads to

$$3^{2m+1} = 4t + 3$$

where $t = q - p$. Using induction, we can see that when $m = 0$, there exists $t = 0$ so the base case holds. We then assume that there exists a t such that

$$3^{2m+1} = 4t + 3 \tag{5}$$

is true. We must demonstrate that there exists an r such that

$$3^{2(m+1)+1} = 4r + 3$$

Using Equation 5, it follows that such an r must obey

$$3^2 \cdot 3^{2m+1} = 4r + 3 \Rightarrow 3^2(4t + 3) = 4r + 3$$

This condition implies that the equality

$$36t + 27 = 4r + 3$$

must hold. Therefore, given t we have that r exists according to the condition

$$r = 9t + 6.$$

Appendix B.2. Transform $1 \rightarrow 2^{(1)}$

Since transitions to 1-nodes involve both $8p + 3$ and $8p + 7$, we divide transitions into 2 cases.

Appendix B.2.1. $1^{(3)} \rightarrow 2^{(1)}$

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 3 + 2 \cdot 3^{2m+1} = 8q + 1$$

Proof:

Rearranging the above equation leads to

$$3^{2m+1} = 4t - 1$$

where $t = q - p$. Using induction, we can see that when $m = 0$, there exists $t = 1$ so the base case holds. We then assume that there exists a t such that

$$3^{2m+1} = 4t - 1 \tag{6}$$

is true. We must demonstrate that there exists an r such that

$$3^{2(m+1)+1} = 4r - 1$$

Using Equation 6, it follows that such an r must obey

$$3^2 \cdot 3^{2m+1} = 4r - 1 \Rightarrow 3^2(4t - 1) = 4r - 1$$

This condition implies that the equality

$$36t - 9 = 4r - 1$$

must hold. Therefore, given t we have that r exists according to the condition

$$r = 9t - 2.$$

Appendix B.2.2. $1^{(7)} \rightarrow 2^{(1)}$

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 7 + 2 \cdot 3^{2m} = 8q + 1$$

Proof:

Rearranging the above equation leads to

$$3^{2m} = 4t - 3$$

where $t = q - p$. Using induction, we can see that when $m = 0$, there exists $t = 1$ so the base case holds. We then assume that there exists a t such that

$$3^{2m} = 4t - 3 \tag{7}$$

is true. We must demonstrate that there exists an r such that

$$3^{2(m+1)} = 4r - 3$$

Using Equation 7, it follows that such an r must obey

$$3^2 \cdot 3^{2m} = 4r - 3 \Rightarrow 3^2(4t - 3) = 4r - 3$$

This condition implies that the equality

$$36t - 27 = 4r - 3$$

must hold. Therefore, given t we have that r exists according to the condition

$$r = 9t - 6.$$

Appendix B.3. Transform $1 \rightarrow x$

Appendix B.3.1. $1^{(3)} \rightarrow x$

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 3 + 2 \cdot 3^{2m} = 8q + 5$$

Proof:

Rearranging the above equation leads to

$$3^{2m} = 4t + 1$$

where $t = q - p$. This condition is similar to the $2 \rightarrow 1^{(3)}$ case above implying that, given t , the condition

$$r = 9t + 2$$

must hold.

Appendix B.3.2. $1^{(7)} \rightarrow x$

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 7 + 2 \cdot 3^{2m+1} = 8q + 5$$

Proof:

Rearranging the above equation leads to

$$3^{2m+1} = 4t - 1$$

where $t = q - p$. This condition is similar to the $1^{(3)} \rightarrow 2$ case above implying that, given t ,

$$r = 9t - 2.$$

must hold.

Appendix B.4. Transform $x \rightarrow 1$

Appendix B.4.1. $x \rightarrow 1^{(3)}$

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 5 + 2 \cdot 3^{2m+1} = 8q + 3$$

Proof:

Rearranging the above equation leads to

$$3^{2m+1} = 4t - 1$$

where $t = q - p$. This condition is similar to the $1^{(3)} \rightarrow 2$ and $1^{(7)} \rightarrow x$ cases above implying that, given t ,

$$r = 9t - 2.$$

must hold.

Appendix B.4.2. $x \rightarrow 1^{(7)}$

$\forall m, p \in \mathbb{N}_0, \exists q \in \mathbb{N}_0$ such that

$$8p + 5 + 2 \cdot 3^{2m} = 8q + 7$$

Proof:

Rearranging the above equation leads to

$$3^{2m} = 4t + 1$$

where $t = q - p$. This condition is similar to the $2 \rightarrow 1^{(3)}$ and $1^{(3)} \rightarrow x$ cases above implying that, given t , the condition

$$r = 9t + 2$$

must hold.

Appendix C. Node Transition Rules

The periodic 'x-1-2-1-x-...' motif is pervasive. In this section, we justify the addition rules when applied to node transition rules. This section exhaustively lists the complete set of allowed transitions along a primitive itinerary.

Appendix C.1. $8p + 3$ Node Transitions

We already know that $8p + 3$ nodes will transition to a 1-node. First, we check if the next node is a $4k + 1$ pump.

$$\begin{aligned} f(8p + 3) &= \frac{24p + 10}{2} \\ \Rightarrow 12p + 5 &= 8k + 5 \text{ must be satisfied} \\ \Rightarrow k &= \frac{3}{2}p \end{aligned}$$

meaning that if p is even, we end up at a $4k + 1$ pump which would signal the end of a primitive itinerary.

When p is odd, After two transitions, we see that

$$\begin{aligned} f^2(8p + 3) &= f(12p + 5) \\ &= \frac{36p + 16}{2} \\ &= 18p + 8 \end{aligned}$$

which requires another division by 2

$$\frac{18p + 8}{2} = 9p + 4$$

indicating the $1^{(3)}$ -node transitions to a 2-node when p is odd. Let $q = 2p + 1$ so that

$$\begin{aligned} 9p + 4 &= 9(2q + 1) + 4 \\ &= 18q + 13 \text{ for } q = 0, 1, 2, 3 \dots \end{aligned}$$

Using a mod 4 congruence where $q = 4k + r$ for $k = 0, 1, 2, 3 \dots$ and $r = 0, 1, 2, 3$, this gives

$$\begin{aligned} 18q + 13 &= 18(4k + r) + 13 \\ &= 72k + 18r + 13 \\ &= 8(9k) + 18r + 13 \end{aligned}$$

This leads again to the cyclic motif 'x-1-2-1-x ...'. So, for p odd, we have the transitions

$r = 0 :$	$= 8(9k) + 13$	$8(9k + 1) + 5$	x
$r = 1 :$	$= 8(9k) + 31$	$8(9k + 3) + 7$	1
$r = 2 :$	$= 8(9k) + 49$	$8(9k + 6) + 1$	2
$r = 3 :$	$= 8(9k) + 67$	$8(9k + 8) + 3$	1

$$\begin{aligned}
 &1^{(3)} 2] \\
 &1^{(3)} 2 1 \\
 &1^{(3)} 2 2 \\
 &1^{(3)} 2 1 \\
 &1^{(3)} 2] \\
 &\vdots
 \end{aligned}$$

Here, we are using the ']' symbol to indicate the end of a primitive itinerary.

Combining and interleaving results with even values of p , as p increases, allowed transitions are of the form:

$$\begin{aligned}
 &1^{(3)}] \\
 &1^{(3)} 2 1 \\
 &1^{(3)}] \\
 &1^{(3)} 2 2 \\
 &1^{(3)}] \\
 &1^{(3)} 2 1 \\
 &1^{(3)}] \\
 &1^{(3)} 2] \\
 &1^{(3)}] \\
 &\vdots
 \end{aligned}$$

These and other 'lego blocks' introduced below are how symbols are appended to sequences generated by the Collatz mapping.

Appendix C.2. $8p + 7$ Nodes Cannot Transition to an $8k + 5$ Node

$$\begin{aligned}
 f(8p + 7) &= \frac{24p + 22}{2} \\
 \Rightarrow 12p + 11 &= 8k + 5 \text{ must be satisfied} \\
 \Rightarrow k &= \frac{3}{2}p + \frac{3}{4}
 \end{aligned}$$

Given an integer p , no integer solution exists to yield a value of k . While this information is important for understanding disallowed transitions, the next subsection discusses what is allowed.

Appendix C.3. $8p + 7$ Nodes Must Always Transition to Two Consecutive 1-Nodes

We already know that $8p + 7$ nodes will transition to a 1-node. After two transitions, we see that

$$\begin{aligned}
 f^2(8p + 7) &= f(12p + 11) \\
 &= \frac{36p + 34}{2} = 8p + 17 \\
 &= 6(3p + 2) + 5 \\
 &= 6r + 5
 \end{aligned}$$

also ends up at a 1-node. Using the same cyclic argument as above, allowed transitions are of the form:

$$\begin{aligned} &1^{(7)} 1] \\ &1^{(7)} 1 1 \\ &1^{(7)} 1 2 \\ &1^{(7)} 1 1 \\ &1^{(7)} 1] \\ &\vdots \end{aligned}$$

Appendix C.4. $8p + 1$ Transitions

Analyses similar to the above lead to

$$\begin{aligned} &2] \\ &2 1 \\ &2 2 \\ &2 1 \\ &2] \\ &\vdots \end{aligned}$$

References

1. Lagarias, J.C.. The $3x + 1$ Problem and its Generalizations, The American Mathematical Monthly, 92:3-23, 1985.
2. Lagarias, J.C.. The $3x + 1$ problem: An annotated bibliography (1963 – 1999), <https://arxiv.org/abs/math/0309224>.
3. Lagarias, J.C.. The $3x + 1$ problem: An annotated bibliography (2000 – 2009), <https://arxiv.org/abs/math/0608208>.
4. Barina, D.. Convergence verification of the Collatz problem, The Journal of Supercomputing, 77, 2021.
5. Sakk, E.. Symbolic dynamic formulation for the Collatz Conjecture:
I. Local and quasi-global behavior. Arxiv: 2403.19699, 2024.
6. Terras, R. . Stopping time problem on the positive integers, Acta Arith., 30:241-252, 1976.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.