Preprints.org

Article

Not peer-reviewed version

Numerical Approach on Collatz Conjecture

Youngik Lee *

Posted Date: 3 May 2024

doi: 10.20944/preprints202404.1570.v2

Keywords: collatz conjecture; 3x+1 problem; Terras Theorem

Article

Numerical Approach on Collatz Conjecture

Youngik Lee

Brown University, Department of Physics, Box 1843, 182 Hope Street, Barus & Holley 545, Providence, RI 02912, USA; s6ygleee@gmail.com

Abstract: We can analyze the properties of the Collatz conjecture by using Terras' Theorem [1,2] and computational graphical calculations. To do this, we introduce the concept of "tree" which shows the relationship between the input and output positive integers of the Collatz process. This approach offers a novel way for determining the non-existence of loops and diverge trees through computer simulations.

Keywords: Collatz conjecture; 3x+1 problem; Terras Theorem

1. Introduction

Collatz conjecture is if one considers the following process on an arbitrary positive integer: [3,4]

$$f(n) = \begin{cases} n/2 & \text{if } n \equiv 0 \pmod{2}, \\ (3n+1) & \text{if } n \equiv 1 \pmod{2}. \end{cases} \tag{1}$$

In the end, the final value will eventually reach the number 1, regardless of the initial positive integer chosen. For convenience, let us refer to applying Equation (1) on any arbitrary positive integer as applying **Collatz process**.

2. Theory

For the first step, let us parametrize the Equation (1) with $x = \ln(n)$, $y = \ln(f(n))$

$$y = \ln(f(n)) = \begin{cases} \ln(e^x/2) & \text{if } n \equiv 0 \pmod{2}, \\ \ln(3e^x + 1) & \text{if } n \equiv 1 \pmod{2}. \end{cases} \tag{2}$$

$$y = \begin{cases} x - \ln(2) & \text{if } n \equiv 0 \pmod{2}, \\ x + \ln(3) + \ln\left(1 + \frac{1}{3e^x}\right) & \text{if } n \equiv 1 \pmod{2}. \end{cases} \tag{3}$$

There is a chain of a Collatz process if we start from an odd number,

$$odd\ (= a) \to even\ (= b) \to \begin{cases} even\ (= c) \\ odd\ (= c) \end{cases} \tag{4}$$

If we using Equation (3), we can rewrite Equation (4) as,

$$\ln(c) = \ln(a) + \ln(3/2) + \ln\left(1 + \frac{1}{3a}\right) \tag{5}$$

So if we call $\ln(a)$ as x_i as $\ln(c)$ as x_{i+1} then,

$$x_{i+1} - x_i = \ln(3/2) + \ln\left(1 + \frac{1}{3e^{x_i}}\right) \tag{6}$$

We can apply the exponential function to both sides of Equation (6),

$$\frac{e^{x_{i+1}}}{e^{x_i}} = \left(\frac{3}{2} + \frac{1}{2e^{x_i}}\right) \tag{7}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

$$e^{x_{i+1}} = \frac{3}{2}e^{x_i} + \frac{1}{2} \tag{8}$$

If there is a chain of Collatz processes for a list of odd numbers from n_i to n_j, then by using $x_i = \ln(n_i)$,

$$n_j = \left(\frac{3}{2}\right)^{j-i-1}(n_i+1) - 1 \tag{9}$$

Let us define the function $\nu_2(n)$ as exponents of 2 after factorization of n.

$$n = 2^a \cdot 3^b \cdot ..., \quad \nu_2(n) = a \tag{10}$$

We can define $m = \nu_2(n_i+1) = j-i-1$, and $k = \nu_2((\frac{3}{2})^m(n_i+1)-1)$. Then, by applying multiple Collatz processes to Equation (9), we can obtain,

$$Col_1(n_i) = n_{j+k} = \frac{(\frac{3}{2})^m(n_i+1)-1}{2^k} \tag{11}$$

Lemma 1. *If we prove the Collatz conjecture for all odd positive integers, it is equivalent to proving it for all positive integers.*

Proof. Let us assume that all odd positive integers converge to 1 after a finite number of steps in the Collatz process. Then, if we have an even number $x_n = 2n \in \mathbf{N}$ for some n, let $k = \nu_2(x_n)$. The number $y_n = x_n/2^k$ is odd. Therefore, x_n also converges to 1 after k more steps than y_n. □

According to Lemma 1, to investigate the Collatz conjecture, we only need to examine the convergence properties of odd positive integers.

Definition 1. *The set of odd positive integers G_k is defined as follows:*

$$G_k = \{x_n | {Col_1}^{k-1}(x_n) = 1, x_n = 2n-1\}$$

where n and k are integers greater than 1, and $G_1 = \{1\}$. The properties of G_k is:

$$^\forall x \in G_{k+1},\ ^\exists y \in G_k, \quad satisfies \quad Col_1(x) = y$$

$$G_k = \{Col_1(x_n) | x_n \in G_{k+1}\}$$

for all $k \geq 1$

According to Definition 1, we can categorize the set of odd numbers as follows:
Now, let us make a function that maps each odd number to its order number.

$$ord(q) = \frac{q+1}{2}, \quad q \equiv 1(mod2) \tag{12}$$

Here q is the odd number, and $ord(q)$ represents its order number. For example, $ord(5) = 3$, indicating that 5 is the third odd number (1, 3, 5, 7...). Then for $n \in \mathbf{N}$ we can define the Equation (13) as an update of Equation (11) in odd number order space:

$$Col_2(n) = ord(Col_1(2n-1)) = \frac{1}{2}\left(\frac{3(\frac{3}{2})^m n - 1}{2^k} + 1\right) \tag{13}$$

Here m and k are defined as by using the Equation (10) as:

$$m = \nu_2(n), \quad k = \nu_2\left(3\left(\frac{3}{2}\right)^m n - 1\right) \tag{14}$$

If we denote n as X, then we can rewrite Equation (13) as:

$$Col_2(X) = \frac{2^k - 1}{2^{k+1}} + \frac{3}{2^{k+1}}\left(\frac{3}{2}\right)^m X \tag{15}$$

Lemma 2. *If we prove the order number, $ord(q)$, of all odd positive integers converges to 1 after apply finite step of applying $Col_2 : \mathbf{N} \to \mathbf{N}$, it is equivalent to proving Collatz conjecture for all odd positive integers.*

Proof. Let us assume that for all odd positive integers, their order numbers converge to 1 after a finite number of steps of Col_2. Now, if we have an odd number $x_{n_1} = 2n_1 - 1 \in \mathbf{N}$ for some n_1, then $ord(x_{n_1}) = n_1$. This is because $Col_2(n_1) = ord(Col_1(x_{n_1}))$, and we obtain $Col_2(n_1) = (Col_1(x_{n_1}) + 1)/2$. Then let us assume that n_1 converge to 1 after k steps of Col_2:

$$Col_2(n_1) \to Col_2(n_2) \to Col_2(n_3) \to ... \to 1$$

Here $n_{k+1} = Col_2(n_k) = (x_{n_k} + 1)/2$. Then we can map this chain as:

$$\frac{Col_1(x_{n_1}) + 1}{2} \to \frac{Col_1(x_{n_2}) + 1}{2} \to \frac{Col_1(x_{n_3}) + 1}{2} \to ... \to \frac{Col_1(x_{n_k}) + 1}{2} = 1$$

Multiply each term by 2 and then subtract 1 for all terms.

$$Col_1(x_{n_1}) \to Col_1(x_{n_2}) \to Col_1(x_{n_3}) \to ... \to Col_1(x_{n_k}) = 1$$

Here, $x_{n_{k+1}} = Col_1(x_{n_k})$. This implies that $Col_1(x_{n_k}) = 1$, which implies the convergence of x_{n_1} in a finite number of steps of the Collatz process. □

Definition 2. *The set of positive integers $\mathcal{G}_k$ is defined as follows:*

$$\mathcal{G}_k = \{x_n | {Col_2}^{k-1}(x_n) = 1, x_n \in \mathbf{N}\}$$

where n and k are integers greater than 1, and $\mathcal{G}_1 = \{1\}$. The properties of $\mathcal{G}_k$ is:

$$^{\forall}x \in \mathcal{G}_{k+1},\ ^{\exists}y \in \mathcal{G}_k, \quad satisfies \quad Col_2(x) = y$$

$$\mathcal{G}_k = \{Col_2(x_n) | x_n \in \mathcal{G}_{k+1}\}$$

for all $k \geq 1$

When we associate Table 1 with Equation (12), we can get Table 2

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

Table 1. The categorization of odd numbers, G_k (e.g., G_1 to G_{23}).

ine G_1	1
ine G_2	3, 5, 21, 85, 151, 227, 341...
ine G_3	13, 15, 23, 35, 53, 75, 113...
ine G_4	7, 11, 17, 61, 69, 93, 141...
ine G_5	9, 19, 29, 37, 45, 81, 117...
ine G_6	25, 49, 51, 67, 77, 87, 101...
ine G_7	33, 39, 43, 59, 65, 79, 89...
ine G_8	57, 105, 115, 157, 173, 177, 211...
ine G_9	127, 135, 139, 153, 159, 187, 191...
ine G_{10}	123, 169, 185, 247, 249, 283, 339...
ine G_{11}	111, 167, 219, 223, 225, 251, 295, 329...
ine G_{12}	297, 351, 393, 445, 527, 585, 595...
ine G_{13}	175, 263, 395, 519, 593, 615, 699...
ine G_{14}	103, 155, 207, 233, 311, 467, 701...
ine G_{15}	91, 137, 183, 275, 413, 549, 621...
ine G_{16}	63, 95, 121, 143, 215, 243, 323...
ine G_{17}	31, 47, 71, 107, 161, 253, 381...
ine G_{18}	27, 41, 55, 83, 125, 165, 189...
ine G_{19}	73, 109, 147, 195, 199, 221, 291...
ine G_{20}	97, 145, 259, 265, 387, 389, 471...
ine G_{21}	129, 193, 235, 240, 343, 345, 353...
ine G_{22}	171, 240, 257, 313, 415, 427, 457...
ine G_{23}	379, 417, 427, 543, 553, 569, 609...
ine	

Table 2. Part of the reparameterizing of Table 1 based on Equation (12) (e.g., from G_1 to G_{10}.)

ine $\mathcal{G}_1$	1
ine $\mathcal{G}_2$	2, 3, 11, 43, 76, 114, 171...
ine $\mathcal{G}_3$	7, 8, 12, 18, 27, 38, 57...
ine $\mathcal{G}_4$	4, 6, 9, 31, 35, 47, 72...
ine $\mathcal{G}_5$	5, 10, 15, 19, 23, 41, 59...
ine $\mathcal{G}_6$	13, 25, 26, 34, 39, 44, 51...
ine $\mathcal{G}_7$	17, 20, 22, 30, 33, 40, 45...
ine $\mathcal{G}_8$	29, 53, 58, 79, 87, 89, 106...
ine $\mathcal{G}_9$	64, 68, 70, 77, 80, 94, 96...
ine $\mathcal{G}_{10}$	62, 85, 93, 124, 125, 142, 170...
ine	

If we represent Table 2 as a diagram, we can obtain Figure 1.

 doi:10.20944/preprints202404.1570.v2

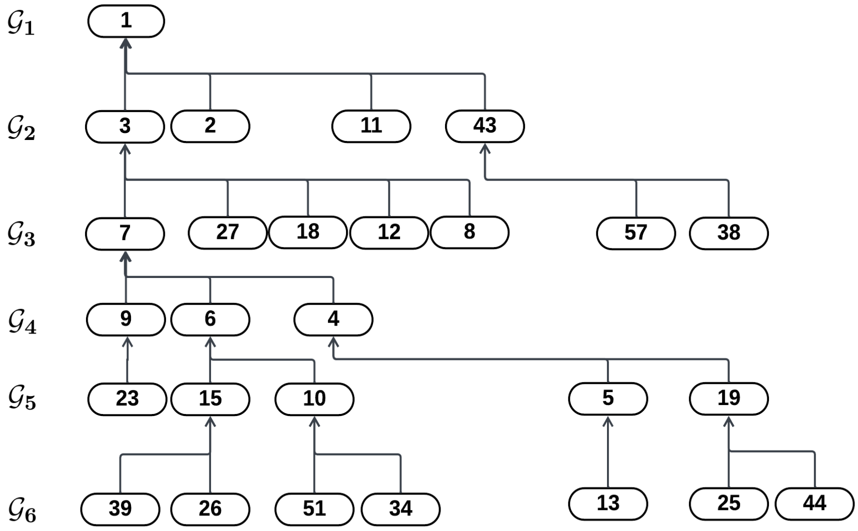

Figure 1. A diagram expression of Table 2 with directed arrows.

Let us define this tree-like diagram as a tree with a converging value x.

Definition 3. *If we link two integers with an arrow from the input to the output value using Equation (15) for all integers, it generates the diagram, which we call a "tree", denoted as $(T, <_T)$. Here, $<_T$ represents the partial ordering based on Equation (15),*

$$^{\exists}k \in \mathbf{N},\ Col_2{}^k(a) = b \to b <_T a$$

Then tree $(T, <_T)$ has a several properties [5]:

(1) $<_T$ is a well-founded partial ordering on T.
(2) $^{\forall}t \in T$, the set $\{u \in T | u <_T t\}$ is linearly ordered by $<_T$.
(3) $^{\forall}t, u \in T$, there exists $v \in T$ such that $v \leq_T t$ and $v \leq_T u$. Here, $t \leq_T t$ essentially means $t <_T t$ or $t = t$.

Definition 4. *For arbitrary tree $(T, <_T)$, $^{\forall}x \in T$ if $^{\exists}A \leq_T x$, then A is called the "***converge value of the tree***", and the tree is called "***converge tree***". However, if there is no such A then the tree is called "***diverge tree***".*

3. Lemmas and Theorems

3.1. Proof of Uniqueness of the Converge Tree

According to Definition 3, when we apply the function $Col_2(x)$ to all positive integers, after k iterations, some values converge to 1. We can categorize these values into groups, denoted as $\mathcal{G}_k$. If we recursively apply $Col_2(x)$ multiple times and there exist some values of x that satisfy $Col_2{}^k(x) - x = 0$, then those values remain fixed throughout this recursive process.

$$^{\forall}x \in \mathcal{G}_k,\ Col_2(x) = x \to Col_2{}^k(x) = x \tag{16}$$

Lemma 3. *For some positive integer X, to satisfy $Col_2(X) = X$, the corresponding k values in Equation (15) lie within a length-1 interval,*

$$\frac{\ln(3/2)}{\ln 2}(m+1) < k < \frac{\ln(3/2)}{\ln 2}(m+1) + 1 \tag{17}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

Proof. When $Col_2(X) = X$, then X should satisfy Equation (18) according to Equation (15).

$$X = \frac{2^k - 1}{2^{k+1}} + \frac{3}{2^{k+1}} \left(\frac{3}{2}\right)^m X \tag{18}$$

$$X = \frac{2^m(2^k - 1)}{2^{k+m+1} - 3^{m+1}} = \frac{(2^k - 1)}{2(2^k - (3/2)^{m+1})} \tag{19}$$

Therefore, the condition to make X as a positive integer is:

$$2^k > (3/2)^{m+1} \rightarrow k > \frac{\ln(3/2)}{\ln 2}(m+1) \tag{20}$$

If we want to make X as an integer, then the numerator of Equation (19) should be divisible by the denominator of Equation (19). Therefore, the numerator must be greater than the denominator.

$$2^k - 1 \geq 2(2^k - (3/2)^{m+1}) \tag{21}$$

$$k \leq \frac{\ln(2(3/2)^{m+1} - 1)}{\ln 2} < \frac{\ln(2(3/2)^{m+1})}{\ln 2} \tag{22}$$

$$k < \frac{\ln(3/2)}{\ln 2}(m+1) + 1 \tag{23}$$

And since we know that the integer multiplication of $\ln(3/2)/\ln 2$ can not be an integer, k should exist within the range of two transcendental numbers Q and $Q + 1$.

$$Q = \frac{\ln(3/2)}{\ln 2}(m+1) \rightarrow Q < k < Q + 1 \tag{24}$$

So, this proves the lemma. $\square$

Next Equation (19) can be expressed as,

$$X = \frac{even \cdot odd}{odd} \tag{25}$$

If we want to make X as an integer, the odd number in the numerator should be divisible by the odd number in the denominator. Therefore the odd number of the numerator must be greater than that of the denominator.

$$2^k - 1 \geq 2^{m+1}(2^k - (3/2)^{m+1}) \tag{26}$$

Therefore, the available values of k should lie within the region:

$$k \leq \frac{1}{\ln 2} \ln\left(\frac{3^{m+1} - 1}{2^{m+1} - 1}\right) = \mathcal{K} \tag{27}$$

Then, we want to show that $\mathcal{K} < R$ in Equations (27) and (28).

$$R = \lfloor Q + 1 \rfloor = \left\lfloor \frac{\ln(3/2)}{\ln 2}(m+1) \right\rfloor + 1 \tag{28}$$

To achieve this, first let us utilizing the lemmas.

Lemma 4.

$$\int_1^\infty \frac{\sin(kA)}{k} dk \leq \sum_{k=1}^\infty \frac{\sin(kA)}{k} \tag{29}$$

Here, $A = 2\pi \frac{\ln(3/2)}{\ln 2} L$

Proof. First, let us define:

$$\int_k^{k+1} \frac{\sin(At)}{t} dt = \mathrm{Si}(A(k+1)) - \mathrm{Si}(Ak) \tag{30}$$

$$\sum_{k=1}^{\infty} \int_k^{k+1} \frac{\sin(At)}{t} dt = \mathrm{Si}(\infty) - \mathrm{Si}(A) \tag{31}$$

By using Taylor Series expansion, [2]

$$-\ln(1-z) = \sum_{k=1}^{\infty} \frac{z^k}{k} \tag{32}$$

Hence, we can rewrite the right-hand side of Equation (29) as:

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{\sin(Ak)}{k} &= \frac{1}{2i} \sum_{k=1}^{\infty} \frac{e^{iAk} - e^{-iAk}}{k} \\ &= \frac{1}{2i} \ln(e^{i(\pi - A)}) \\ &= \frac{1}{2}\left((\pi - A) - 2\pi \left\lfloor \frac{\pi - A}{2\pi} \right\rfloor\right)\end{aligned} \tag{33}$$

Therefore we want to show:

$$\frac{1}{2}\left((\pi - A) - 2\pi \left\lfloor \frac{\pi - A}{2\pi} \right\rfloor\right) \geq \frac{\pi}{2} - \mathrm{Si}(A) \tag{34}$$

$$\frac{\mathrm{Si}(A)}{\pi} \geq \frac{A}{2\pi} + \left\lfloor \frac{1}{2} - \frac{A}{2\pi} \right\rfloor \tag{35}$$

Here, $A = 2\pi \frac{\ln(3/2)}{\ln 2} L = 2\pi u$, $L \in \mathbf{N}$

$$l(u) = \frac{\mathrm{Si}(2\pi u)}{\pi} \geq u + \left\lfloor \frac{1}{2} - u \right\rfloor = r(u) \tag{36}$$

Then, for every u, there exists an x such that:

$$\lfloor x \rfloor = \lfloor u \rfloor, \quad \frac{\mathrm{Si}(2\pi x)}{\pi} = \frac{1}{2} \tag{37}$$

Then, for some $N \in \mathbf{N}$, we have $x = N + \epsilon$, where $0 < \epsilon < 1$. Considering the periodicity properties of $\sin(2\pi x)/x$, there are two available values of x for each u. Let us define that when $0 < \epsilon \leq 1/2$, then $\epsilon = \alpha$, and if $1/2 < \epsilon < 1$, then $\epsilon = \beta$.

We can categorize the region of u into three regions.

(1) $N \leq u \leq N + \alpha$
(2) $N + \alpha \leq u \leq N + \frac{1}{2} \rightarrow l(u) \geq 1/2 > r(u)$
(3) $N + \frac{1}{2} \leq u \leq N + 1 \rightarrow l(u) > 0 \geq r(u)$

In regions (2) and (3), Equation (36) is valid. In the case of region (1), let us calculate the tangential line of $l(u)$ at $N + \alpha$.

$$y(u) = \frac{\sin(2\pi\alpha)}{\pi(N+\alpha)}(u - N - \alpha) + \frac{1}{2} \tag{38}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

Therefore, when $u = N$, the line passes through the point,

$$y(N) = \frac{1}{2} - \frac{\alpha \sin(2\pi\alpha)}{\pi(N+\alpha)} \geq \frac{1}{2} - \frac{\alpha}{\pi(N+\alpha)} > \frac{1}{2} - \frac{1}{2\pi} > 0$$

because $0 < \alpha < 1/2$. Furthermore, Figure 3 shows that based on computer simulation, when $N \geq 1$, $y(N) > \alpha$.

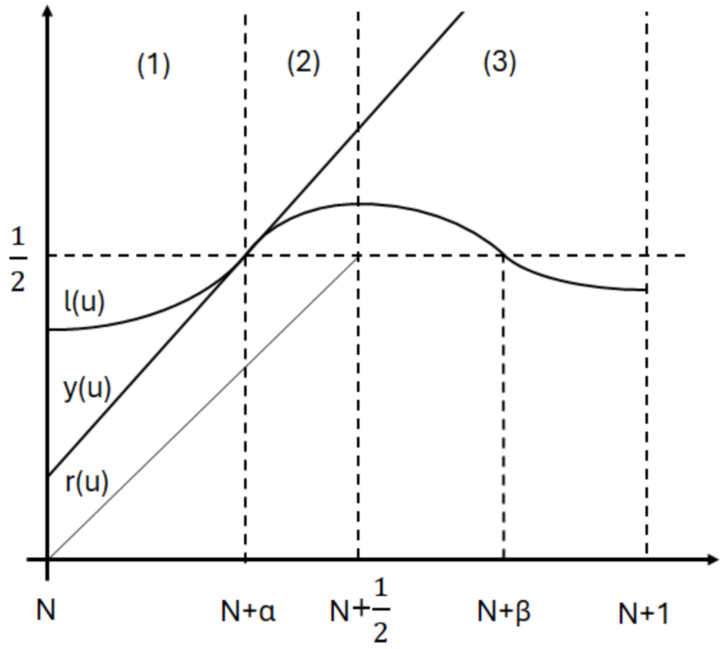

Figure 2. Graphical representation of regions (1) to (3) and functions $l(u)$, $y(u)$, $r(u)$. Here, the x-axis represents u when $\lfloor u \rfloor = N$.

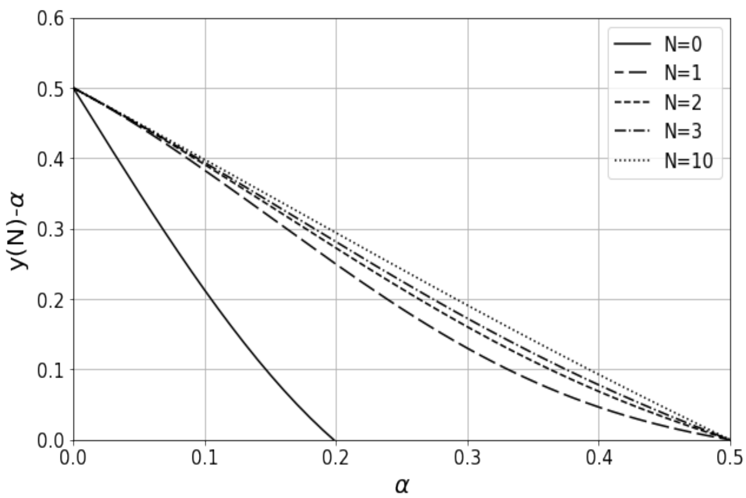

Figure 3. The results of $y(N) - \alpha$ in the region where $0 < \alpha < 1/2$. Except for the case $N = 0$, all other graphs yield positive results across all α regions.

Therefore, in region (1), $l(u) > y(u) > r(u)$. Hence, Equation (36) holds for all u, and Equation (34) also holds. If we define:

$$s(N, L) = \sum_{k=1}^{N} \left(\frac{\sin(kA)}{k} - \text{Si}(A(k+1)) + \text{Si}(Ak) \right) \tag{39}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

then we can conclude that for any $L \in \mathbf{N}$,

$$\lim_{N \to \infty} s(N, L) \geq 0 \tag{40}$$

We can use computer simulation to double-check that $s(N, L)$ converges as we increase the values of N.

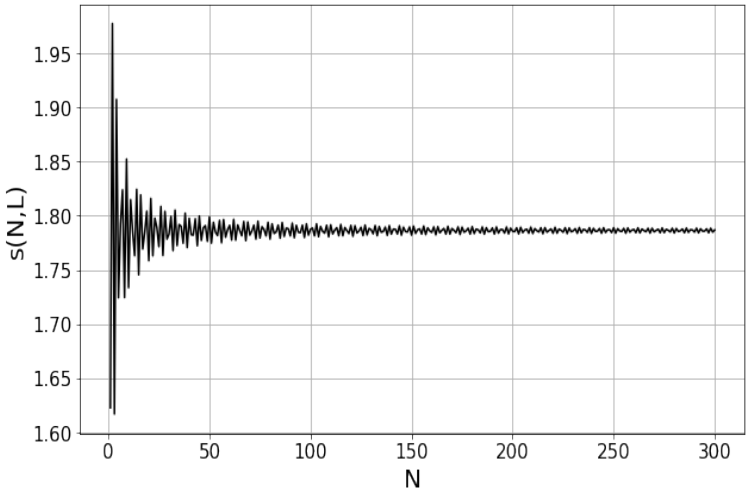

Figure 4. The summation of the right-hand side of Equation (39), which converges to positive real values, is calculated for N ranging from 1 to 300, and $L = 1$.

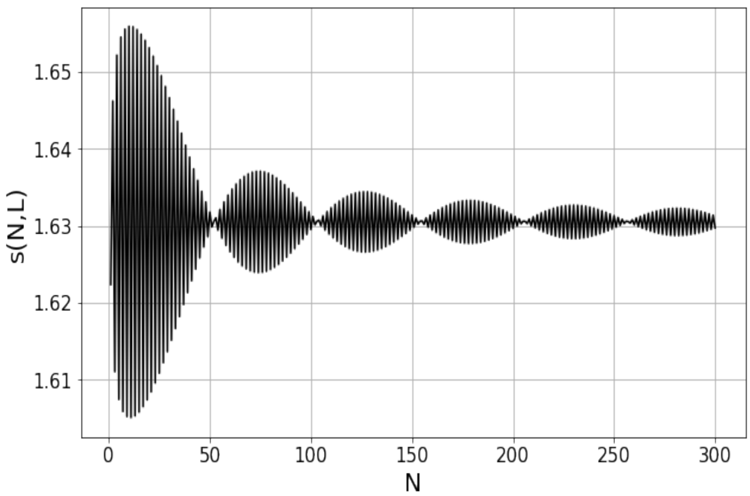

Figure 5. The summation of the right-hand side of Equation (39), which converges to positive real values, is calculated for N ranging from 1 to 300, and $L = 6$.

Based on computer simulation, we can double-check that $s(N, L)$ converges as N increases, as shown in Figure 7. Therefore we can get,

$$\sum_{k=1}^{\infty} \left(\frac{\sin(kA)}{k} - \text{Si}(A(k+1)) + \text{Si}(Ak) \right) \geq 0 \tag{41}$$

$$\sum_{k=1}^{\infty} (\text{Si}(A(k+1)) - \text{Si}(Ak)) \leq \sum_{k=1}^{\infty} \frac{\sin(kA)}{k} \tag{42}$$

So if we utilize Equation (30), (31), (35) and (42),

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

$$\sum_{k=1}^{\infty}\int_{k}^{k+1}\frac{\sin(At)}{t}dt=\int_{1}^{\infty}\frac{\sin(At)}{t}dt\leq\sum_{k=1}^{\infty}\frac{\sin(kA)}{k} \tag{43}$$

which concludes the proof. □

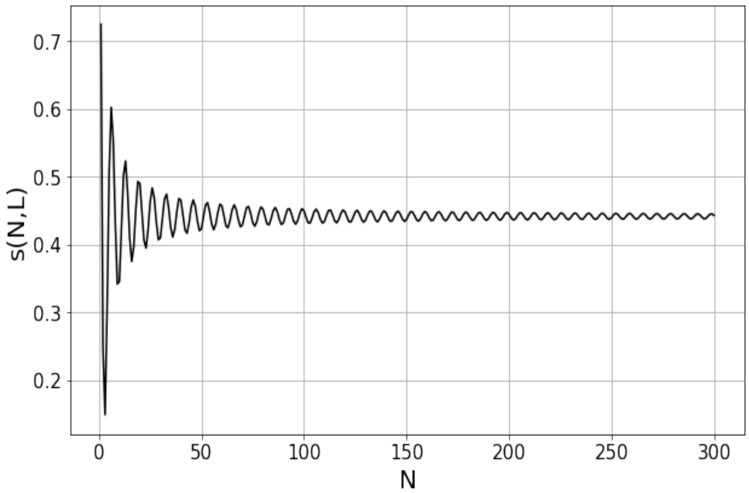

Figure 6. The summation of the right-hand side of Equation (39), which converges to positive real values, is calculated for N ranging from 1 to 300, and $L = 10$.

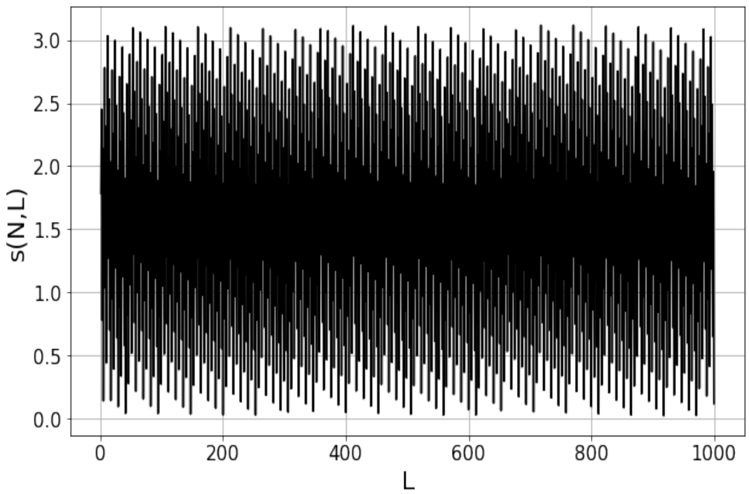

Figure 7. The summation of the right-hand side of Equation (39), which converges to positive real values, is calculated for L ranging from 1 to 1000, and $N = 2000$.

Lemma 5.

$$\frac{1}{\ln 2}\ln\left(\frac{3^{m+1}-1}{2^{m+1}-1}\right)<\left\lfloor\frac{\ln(3/2)}{\ln 2}(m+1)\right\rfloor+1 \tag{44}$$

Proof. Let us set $L = m + 1$, then we can simplify Equation (44) as

$$\frac{1}{\ln 2}\ln\left(\frac{3^{L}-1}{2^{L}-1}\right)<\left\lfloor\frac{\ln(3/2)}{\ln 2}L\right\rfloor+1 \tag{45}$$

and if we use the floor function and fractional part function notation $\lfloor x\rfloor = x - \{x\}$ [6],

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 May 2024 doi:10.20944/preprints202404.1570.v2

$$\left\{ \frac{\ln(3/2)}{\ln 2} L \right\} < 1 - \frac{1}{\ln 2} \ln\left(\frac{1 - 1/3^L}{1 - 1/2^L} \right) \tag{46}$$

Here, we can utilize the Fourier series expansion of the fractional part function [7]:

$$\{x\} = \frac{1}{2} - \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{\sin(2\pi k x)}{k} \tag{47}$$

Therefore, Equation (46) is:

$$\ln\left(\frac{1}{\sqrt{2}} \frac{1 - 1/3^L}{1 - 1/2^L} \right) < \frac{\ln 2}{\pi} \sum_{k=1}^{\infty} \frac{\sin\left(2\pi k \frac{\ln(3/2)}{\ln 2} L\right)}{k} \tag{48}$$

By using Lemma 4, Equation (48) holds when Equation (49) is valid (with $A = 2\pi \frac{\ln(3/2)}{\ln 2} L$).

$$\ln\left(\frac{1}{\sqrt{2}} \frac{1 - 1/3^L}{1 - 1/2^L} \right) < \frac{\ln 2}{\pi} \left(\int_1^{\infty} \frac{\sin(kA)}{k} dk \right) \tag{49}$$

$$\ln\left(\frac{1}{\sqrt{2}} \frac{1 - 1/3^L}{1 - 1/2^L} \right) < \frac{\ln 2}{\pi} \left(\frac{\pi}{2} - \mathrm{Si}(A) \right) \tag{50}$$

$$\ln\left(\frac{1 - 1/3^L}{1 - 1/2^L} \right) < \ln 2 - \frac{\ln 2}{\pi} \mathrm{Si}\left(2\pi \frac{\ln(3/2)}{\ln 2} L \right) \tag{51}$$

And we can get the range:

$$\ln\left(\frac{1 - 1/3^L}{1 - 1/2^L} \right) \le \ln(4/3) \tag{52}$$

Let us find the region that satisfies:

$$\ln(4/3) < \ln 2 - \frac{\ln 2}{\pi} \mathrm{Si}\left(2\pi \frac{\ln(3/2)}{\ln 2} L \right) \tag{53}$$

$$\mathrm{Si}\left(2\pi \frac{\ln(3/2)}{\ln 2} L \right) < \frac{\ln(3/2)\pi}{\ln 2} \tag{54}$$

The available range for L is:

$$0 < L < 0.7757, \quad 0.939 < L$$

And $L = m + 1 \ge 1$, which satisfying the condition. □

Theorem 1. *There is no positive integer point x that satisfies $Col_2(x) - x = 0$, except for $x = 1$.*

Proof. According to Lemma 3, 4, and 5, we can conclude that the available region for k is given by:

$$Q < k < \mathcal{K} < R < Q + 1$$

which implies that X and k in Equation (18) cannot both be integers at the same time. Consequently, no integer value of k exists to satisfy Equation (18). □

Theorem 2. *There is only one converge tree, and all elements in the converge tree converge to 1.*

Proof. Let us assume there is another converge tree with converging values $a \ne 1$. Then, since a is a converging value, we have $Col_2(a) = a$. However, according to Theorem 1 only 1 goes to itself, so

$Col_2(a)$ is not an integer value other than a. This contradicts the assumption, thus confirming that there is only one converge tree in an integer system.

Therefore, there is only one converge tree, and if any integer is present in the converge tree, it will eventually converge to 1. □

3.2. Proof of the Non-Existence of a Diverge Tree

Theorem 3. *There is no diverge tree.*

Proof. The part of the proof refers to [1] (1976) by R. Terras. Let us call A_1 as the set of all available converge trees. A_2 is the set of all available sets of diverge trees.

Let us assume $A_2 \neq \emptyset$. Then we can select the minimum element $N \in A_2$. For some k, if N is not a power of 2, then there always exists $2^k < N < 2^{k+1}$. Thus, we can express $N = 2^k + b$ $(b < N)$. Since N is the minimum element of A_2, we know that ${}^\forall x < N, x \in A_1$. According to Theorem 2, ${}^\forall x \in A_1$ converges to 1, and based on Definition 1.5 in [1], this implies that all x have a **stop time** except when $x = 1$. Here, stop time refers to the point at which, after a finite number of steps of the Collatz process, the value goes smaller than its initial value. For example, $3 \to 8 \to 4 \to 2 \to 1$, the stop time is 3.

According to periodicity Theorem 1.2 in [1], if $x \equiv b \pmod{2^k}$ then x has the same **encoding vector** as b.

(1) When $b \neq 1$, N also has the same stop time as b.
(2) When $b = 1$, $N = 2^k + 1$. If we apply the Collatz process, $N \to N' = 3 \cdot 2^{k-2} + 1$. If $k \geq 2$, $N' = 3 \cdot 2^{k-2} + 1 < 2^k + 1 = N$. When $k = 1$, $N = 3$ which also converges to $1 < N$.
(3) When $N = 2^k$, applying the Collatz process results in convergence to 1, which is $1 < N$.

Therefore, in all cases, N also has a stop time. However, $N \in A_2$, so it should not have a stop time. This leads to a contradiction, so $A_2 = \emptyset$. □

4. Conclusions and Outlook

Theorem 4. *All positive integers converge to 1 after a finite number of steps of the Collatz process.*

Proof. According to Theorem 2 and 3, in a positive integer system, there is no diverge tree and there is only one converge tree, with the converging value 1. □

In this research, we prove that all positive integers converge to 1 after a finite number of steps of the Collatz process.

As an outlook, we can investigate the continuous generative function of Equation (15) and the periodicity of the encoding vector of the numbers in Table 2. Additionally, there is another way of proving Lemma 5 from Equation (46), which is equivalent to show:

$$\{L \cdot \log_2(3)\} < 1 - \log_2\left(\frac{1 - 1/3^L}{1 - 1/2^L}\right) \tag{55}$$

We know that $\log_2(3)$ is a transcendental number with the value, [8,9]

$$\log_2(3) = 1.584962500721156181... \tag{56}$$

Then, with particular values of L, we can obtain:

$$\begin{aligned} \{79335 \cdot \log_2(3)\} &= 0.999994712... \\ \{190537 \cdot \log_2(3)\} &= 0.999999906... \\ \{64497107 \cdot \log_2(3)\} &= 0.99999998... \end{aligned} \tag{57}$$

Thus, to show the validity of Equation (55) is the same as to show how quickly its right-hand side converges to 1 compared to $\{L \cdot \log_2(3)\}$. Moreover, besides this approach, various other methods

 doi:10.20944/preprints202404.1570.v2

such as cycle analysis [10] or bounding functions [11,12] are employed in the investigation of the Collatz conjecture. Exploring potential relations between these methods and the current result could be a relevant next step.

References

1. Riho Terras, "A stopping time problem on the positive integers", ACTA ARITIMETICA (1976), http://matwbn.icm.edu.pl/ksiazki/aa/aa30/aa3034.pdf
2. Eric Roosendaal, "The Terras Theorem", http://www.ericr.nl/wondrous/terras.html
3. Jeffrey C. Lagarias, "The 3x + 1 problem and its generalizations". The American Mathematical Monthly. 92 (1): 3–23, (1985), https://doi.org/10.1080/00029890.1985.11971528
4. Jeffrey C. Lagarias, "The 3x+1 Problem: An Overview", (2021), https://arxiv.org/abs/2111.02635
5. Spencer Unger, "Trees in set Theory", https://www.math.toronto.edu/sunger/cmu/TreesTalk.pdf
6. R. L. Graham, D. E. Knuth, O. Patashnik, "Concrete Mathematics", Addison-Wesley (1994), p. 70.
7. E. C. Titchmarsh, D. R. Heath-Brown, "The Theory of the Riemann Zeta-function" (2nd ed.), Oxford U. P. (1986), p. 15, Equation 2.1.7
8. OEIS, "Decimal expansion of $log\_2(3)$", https://oeis.org/A020857
9. Simon Plouffe, "log(3)/log(2) to 10000 digits", https://www.plouffe.fr/simon/constants/log3log2.txt
10. Shalom Eliahou, "Le problème 3n+1: y a-t-il des cycles non triviaux? (III)", Images des Maths, CNRS, (2011)
11. Terence Tao, "Almost all orbits of the Collatz map attain almost bounded values.", Forum of Mathematics, (2022) https://arxiv.org/abs/1909.03562
12. Ilia Krasikov, Jeffrey C. Lagarias, "Bounds for the 3x+1 problem using difference inequalities". Acta Arith., 109, 237–258, (2003)