

9-18-2021

## Novel Theorems and Algorithms Relating to the Collatz Conjecture

Michael R. Schwob

*The University of Texas at Austin*

Peter Shiue

*University of Nevada, Las Vegas, peter.shiue@unlv.edu*

Rama Venkat

*University of Nevada, Las Vegas, rama.venkat@unlv.edu*

Follow this and additional works at: [https://digitalscholarship.unlv.edu/math\\_fac\\_articles](https://digitalscholarship.unlv.edu/math_fac_articles)



Part of the [Mathematics Commons](#), and the [Theory and Algorithms Commons](#)

---

### Repository Citation

Schwob, M. R., Shiue, P., Venkat, R. (2021). Novel Theorems and Algorithms Relating to the Collatz Conjecture. *International Journal of Mathematics and Mathematical Sciences*, 2021 1-10.

<http://dx.doi.org/10.1155/2021/5754439>

This Article is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Article in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Article has been accepted for inclusion in Mathematical Sciences Faculty Publications by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact [digitalscholarship@unlv.edu](mailto:digitalscholarship@unlv.edu).

## Research Article

# Novel Theorems and Algorithms Relating to the Collatz Conjecture

Michael R. Schwob <sup>1</sup>, Peter Shiue,<sup>2</sup> and Rama Venkat <sup>3</sup>

<sup>1</sup>Department of Statistics and Data Sciences, University of Texas at Austin, Austin 78712, TX, USA

<sup>2</sup>Department of Mathematical Sciences, University of Nevada, Las Vegas 89154, NV, USA

<sup>3</sup>Howard R. Hughes College of Engineering, University of Nevada, Las Vegas 89154, NV, USA

Correspondence should be addressed to Rama Venkat; [rama.venkat@unlv.edu](mailto:rama.venkat@unlv.edu)

Received 24 June 2021; Accepted 26 August 2021; Published 18 September 2021

Academic Editor: Fernando Bobillo

Copyright © 2021 Michael R. Schwob et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Proposed in 1937, the Collatz conjecture has remained in the spotlight for mathematicians and computer scientists alike due to its simple proposal, yet intractable proof. In this paper, we propose several novel theorems, corollaries, and algorithms that explore relationships and properties between the natural numbers, their peak values, and the conjecture. These contributions primarily analyze the number of Collatz iterations it takes for a given integer to reach 1 or a number less than itself, or the relationship between a starting number and its peak value.

## 1. Introduction

In 1937, the German mathematician Lothar Collatz proposed a conjecture that states the following sequence will always reach 1, regardless of the given positive integer  $n$ : if the previous term in the sequence is even, the next term is half of the previous term; if the previous term in the sequence is odd, the next term is 3 times the previous term plus 1 [1]. This conjecture is considered impossible to prove given modern mathematics. In fact, Paul Erdős claimed that “mathematics may not be ready for such problems.” [2] Jeffrey Lagarias echoed this sentiment and provided a thorough summary of numerous results concerning the conjecture in [3].

Despite today’s supposed inability to prove the Collatz conjecture, several papers have outlined discoveries related towards the seemingly impossible proof. Perhaps the most notable recent development was made by Terence Tao, who showed that most orbits of the Collatz map attain almost bounded values [4]. An excellent review of his paper was published in the College Mathematics Journal [5]. This significant development illuminates one potential avenue to proving the Collatz conjecture; however, other paths to the proof remain viable.

Some mathematicians have attempted to break the Collatz conjecture up into several smaller problems or have rephrased the conjecture in terms that may be more tractable. For example, Ren [6] proved that if his reduced Collatz conjecture (RCC) is proven, then that proof can be applied to the Collatz conjecture. Ren also proved that half of the natural numbers follow reduced Collatz dynamics; however, the second half of the proof for the RCC remains unsolved [7]. While exploring reduced Collatz dynamics may be a promising route for the proof of the Collatz conjecture, it is likely beneficial to propose other potential routes to the proof [8, 9].

Several manuscripts have outlined various, unsuccessful paths that mathematicians have taken to solve the conjecture [10–12]. Perhaps the proof of this mesmerizing conjecture is so desired because the conjecture itself finds use in various applications. The Collatz conjecture is used in high-uncertainty audio signal encryption [13], image encryption [14], dynamic software watermarking [15], and information discovery [16]. Although the lack of a proof does not preclude applications of the Collatz conjecture, mathematical phenomenon related to the conjecture may be of interest in other applications, such as cryptography or information theory.

In Section 2, we present novel theorems and corollaries that explore a mapping that follows the Collatz conjecture; this section is broken up into two sections with the first focusing on determining the number of Collatz iterations for a natural number to reach 1 and the second section focusing on an analysis of peak values. In Section 3, an algorithm to obtain a directed graph of the conjecture is provided, as well as an algorithm that analyzes iterations of the conjecture and an algorithm that determines peak values in Collatz sequences. Lastly, we summarize these novel theorems, corollaries, and algorithms in Section 4, and we relate them to the Collatz conjecture.

### 2. The Conjecture and Related Theorems

Let  $\mathbb{N}$  be the set of all positive integers and  $n \in \mathbb{N}$ . Collatz defined the following map:

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

A sequence can be formed by iterating over  $T$ :  $n, T(n), T^2(n) = T(T(n)), \dots, T^k(n) = T(T^{k-1}(n))$  for  $k \in \mathbb{N}$ . The Collatz conjecture asserts that the sequence obtained by iterating  $T$  always reaches integer 1, no matter which positive integer  $n$  begins the sequence. A diagram that depicts the relationship between positive integers is shown in Figure 1. We start our exploration of the Collatz conjecture with the following theorem.

**Theorem 1.** Let  $a_{n+1} = (3b + 1)a_n + b$  and  $n \geq 2$ , where  $a_1$  and  $b$  are given positive, odd integers. Then,

- (i)  $T(a_{n+1}) = (3b + 1)T(a_n)$
- (ii)  $T(a_{n+1}) = (3b + 1)^n T(a_1)$

*Proof.* For (i): by the Collatz conjecture, we have

$$\begin{aligned} T(a_{n+1}) &= 3(a_{n+1}) + 1 = 3[(3b + 1)a_n + b] + 1 \\ &= 9ba_n + 3a_n + 3b + 1 = (3b + 1)(3a_n + 1) \\ &= (3b + 1)T(a_n). \end{aligned} \tag{2}$$

For (ii): we prove through a combination of (i) and mathematical induction. Let  $n = 1$ . Then,

$$T(a_{n+1}) = (3b + 1)^n T(a_1) = (3b + 1)T(a_1), \tag{3}$$

which is true by (i). Now, assume that  $T(a_{n+1}) = (3b + 1)^n T(a_1)$ . If we show that  $T(a_{(n+1)+1}) = (3b + 1)^{(n+1)} T(a_1)$ , then we prove (ii). By (i),

$$T(a_{(n+1)+1}) = T(a_{n+2}) = (3b + 1)T(a_{n+1}), \tag{4}$$

and by our assumption,

$$(3b + 1)T(a_{n+1}) = (3b + 1)(3b + 1)^n T(a_1) = (3b + 1)^{n+1} T(a_1). \tag{5}$$

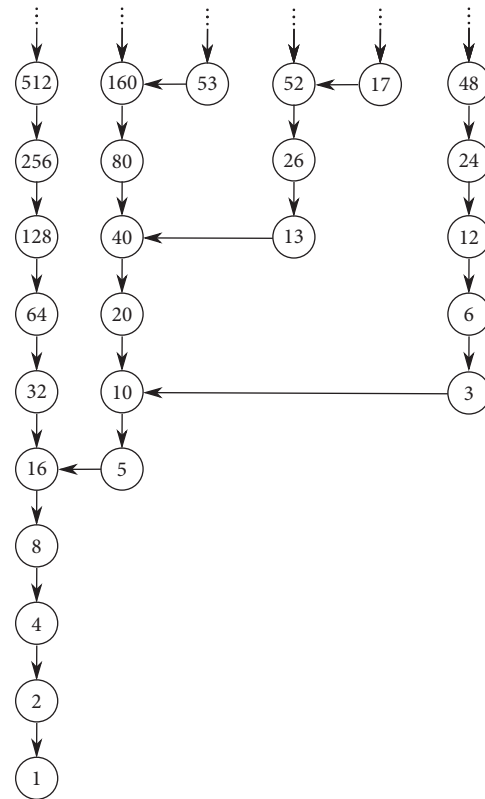


FIGURE 1: Collatz conjecture visualized.

Therefore, we have proved through mathematical induction that  $T(a_{n+1}) = (3b + 1)^n T(a_1)$ .  $\square$

In Theorem 1, we state that  $a_1$  is an odd integer. Due to the nature of Collatz sequences, an even integer  $e$  will eventually reach an odd integer  $o$ , where  $o$  can be used in this theorem. Note that there is one trivial case for  $e$ ; if  $e$  can be written as  $e = 2^k$ , where  $k \in \mathbb{N}$ , then the first odd integer that  $e$  will reach in its Collatz sequence is 1. Thus, if  $a_1$  is even, then  $a_1$  will be an odd integer for certain steps, which can be used in Theorem 1. Therefore, the statement that  $a_1$  is odd in this theorem does not limit its application to Collatz sequences that start with some positive, odd integer.

From this theorem, we can establish a unique relationship for a particular choice in  $b$ . This relationship is delineated in the following corollary.

**Corollary 1.** Let  $a_{n+1} = 4^k a_n + 4^k - 1/3$ . Then,

- (i)  $T(a_{n+1}) = 4^k T(a_n)$ ,
- (ii)  $T(a_{n+1}) = (4^k)^n T(a_1)$ ,

which follow from Theorem 1 when  $3b + 1 = 4^k$ .

We can then use (i) from Corollary 1 to help establish the following.

**Corollary 2.** Let  $k = 1, b = 1$ , and  $a_n = 2j + 1$ , where  $j \in \mathbb{N}$ . Since  $a_{n+1} = 4a_n + 1 = 8j + 5$ , we have  $T(8j + 5) = 4T(2j + 1)$ .

Suppose  $j = 1$  and examine  $T(13)$ . By Corollary 2,  $T(13) = 4T(3) = 2^2T(2j + 1)$ . We can extract a convenient and meaningful relationship from this corollary. Note that the sequence starting with integer 13 takes 9 iterations to reach 1 in the Collatz conjecture, and the sequence that starts at integer 3 takes 7 iterations. Therefore, the sequence starting at 13 takes 2 more iterations than the sequence starting at 3. Now, take  $j = 2$  and examine  $T(21)$ . By Corollary 2,  $T(21) = 2^2T(5)$ . Through the Collatz conjecture, the sequence starting at integer 5 takes 5 iterations to reach 1, while the sequence starting at integer 21 takes 7 iterations to reach 1. Therefore, there is a difference of two iterations between sequences starting at  $8j + 5$  and  $2j + 1$  for any  $j \in \mathbb{N}$ . We see that if  $a_n = 2j + 1$  reaches 1 in the conjecture, then  $a_{n+1}$  reaches 1.

*2.1. Counting Iterations.* Most approaches to proving the Collatz conjecture will inevitably detail the number of iterations it takes for a given number  $n$  to reach 1. Hence, we provide relationships between  $\mathbb{N}$  and the number of Collatz iterations it would take to complete a sequence ending in 1; this will be of direct interest to the academic community attempting to prove the conjecture.

In the theorem that follows, we provide a general relationship that implies if  $a_n$  reaches 1 in finitely many iterations, then  $a_{n+1} = (3b + 1)a_n + b$  also reaches 1 in finitely many iterations.

**Theorem 2.** *Suppose  $a_n$  reaches 1 with  $j$  iterations of the Collatz conjecture. Then,  $a_{n+1} = 4^k a_n + (4^k - 1/3)$  reaches 1 with  $2k + j$  iterations.*

*Proof.* Note that  $T^j(a_n) = 1$ , where  $j$  is the number of iterations required of the Collatz conjecture for  $a_n$  to reach 1. From Corollary 1, we have  $T(a_{n+1}) = 4^k T(a_n) = 2^{2k} T(a_n)$ . Since  $T^{2k}(2^{2k}) = 1$ , we have

$$T^{2k}(2^{2k} T(a_n)) = T(a_n), \tag{6}$$

which gives us

$$T^{2k}(T(a_{n+1})) = T^{2k}(2^{2k} T(a_n)) = T(a_n). \tag{7}$$

For any positive integer  $j$ ,

$$T^{j-1}[T^{2k+1}(a_{n+1})] = T^{j-1}[T(a_n)] = T^j(a_n). \tag{8}$$

Since  $T^j(a_n) = 1$ , we have  $T^{2k+j}(a_{n+1}) = 1$ .  $\square$

*Example 1.* Fix  $k = 1$ , and suppose we let  $a_1 = 3$ . If we take  $a_{n+1} = 4a_n + 1$ , then  $a_2 = 13$ ,  $a_3 = 53$ , and  $a_4 = 213$ . By Theorem 2,  $a_2$ ,  $a_3$ , and  $a_4$  will reach 1 in a finite number of iterations so long as  $a_1$  does. From Figure 1, we know that  $T^7(3) = 1$ . Therefore, 13 reaches 1 with  $7 + 2 = 9$  iterations. Similarly, 53 and 213 reach 1 with 11 and 13 iterations, respectively. This is depicted in Figure 2. Another case, where  $a_1 = 5$ , is depicted in Figure 3.

A special case arises from Theorem 2 when  $a_n = 1$  and, consequently,  $a_k = a_{n+1} = 4^k + (4^k - 1/3) = 4^{k+1} - 1/3$ . By

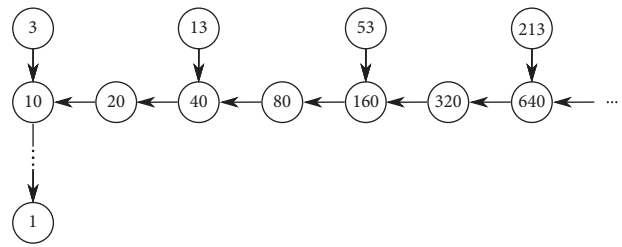


FIGURE 2: Collatz conjecture with  $a_1 = 3$ .

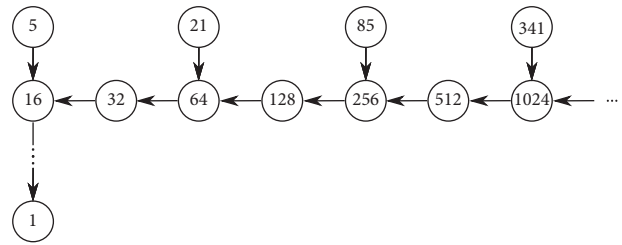


FIGURE 3: Collatz conjecture with  $a_1 = 5$ .

Theorem 2,  $a_k$  reaches 1 with  $2k + 3$  Collatz iterations. Note that  $a_k$  takes on the form of a geometric series with ratio 4:

$$a_k = 1 + 4 + 4^2 + \dots + 4^k. \tag{9}$$

Therefore, this particular geometric series shares a unique relationship with the Collatz conjecture. Perhaps this discovery may inspire new approaches to the Collatz conjecture's proof that utilize series and their properties.

While general relationships would be the most helpful to establish a proper proof of the Collatz conjecture, an analysis of specific relationships may prove nontrivial. In the following theorem, we explore such relationships similar to the special sequences relating to the Collatz conjecture in [17].

**Theorem 3.** *Let  $k \in \mathbb{N}$  and  $t \in \mathbb{N} \cup \{0\}$ . We have the following relationships:*

- (i) *If  $n = 4k + 1$ , then the number of iterations until the Collatz sequence reaches a number less than  $n$  is 3. That is,  $T^3(4k + 1) < 4k + 1$ .*
- (ii) *If  $n = 16t + 3$ , then  $T^6(16t + 3) < 16t + 3$ .*
- (iii) *If  $n = 32t + 11$  or  $n = 32t + 23$ , then  $T^8(n) < n$ .*
- (iv) *If  $n = 128t + 7$ ,  $n = 128t + 15$ , or  $n = 128t + 59$ , then  $T^{11}(n) < n$ .*
- (v) *If  $n = 256k + 39$ ,  $n = 256k + 79$ ,  $n = 256k + 95$ ,  $n = 256k + 123$ ,  $n = 256k + 199$ , or  $n = 256k + 219$ , then  $T^{13}(n) < n$ .*

*Proof.* This proof can be obtained analytically or computationally. Since all points in this proof follow the same structure, we only analytically prove (i) and (ii).

For (i): we have that  $T(4k + 1) = 3(4k + 1) + 1 = 12k + 4$ . Similarly, we can obtain  $T^2(4k + 1) = T(12k + 4) = 12k + 4/2 = 6k + 2$  and  $T^3(4k + 1) = T(6k + 2) = 3k + 1 < 4k + 1$ .

For (ii): we continue to follow the Collatz sequence. We have  $T(16t + 3) = 48t + 10$ ,  $T^2(16t + 3) = T(48t + 10) = 24t + 5$ ,  $T^3(16t + 3) = T(24t + 5) = 72t + 16$ ,  $T^4(16t + 3) = T(72t + 16) = 36t + 8$ ,  $T^5(16t + 3) = T(36t + 8) = 18t + 4$ , and  $T^6(16t + 3) = T(18t + 4) = 9t + 2 < 16t + 3$ .

The proof for the remaining portions of this theorem is obtained through a similar process, which can be supplemented with computation. Examples of computations for the last three portions of this theorem are provided in Tables 1–3. □

The consequence of this theorem is that each of the positive integers  $n$  that can be written as any of the above forms will eventually become a number less than itself. If  $n$  does not abide by the above forms, its pattern is unknown. If a similar, general relationship was to be discovered such that each positive integer  $n$  becomes a number less than itself, one can then prove the Collatz conjecture.

The next three examples reveal the number of iterations it takes for an integer to reach a number less than itself. The first of these examples arises from a special case of Theorem 3 (i), and the last two examples pertain to two particular geometric series.

*Example 2.* Consider  $3^m$ , where  $m$  is even. Then, we have

$$3^m \equiv (-1)^m \equiv 1 \pmod{4}. \tag{10}$$

We can then write  $3^m = 4k + 1$  for some  $k \in \mathbb{N}$ . By Theorem 3 (i),  $T^3(3^m) < 3^m$ . Similarly,  $7^m = 4k_1 + 1$  and  $11^m = 4k_2 + 1$  for some  $k_1, k_2 \in \mathbb{N}$ . In general, we can write  $n = (3 + 4t)^m = 4k + 1$  for some  $k \in \mathbb{N}$  and even  $m$ . Then, by Theorem 3,  $T^3[(3 + 4t)^m] < (3 + 4t)^m$ .

*Example 3.* Let  $s_m = 1 + 4^n + 4^{2n} + \dots + 4^{mn}$ . Then,  $s_m$  will reduce to a number less than itself in 3 Collatz iterations. We can show this by rewriting

$$s_m = 1 + 4(4^{n-1} + \dots + 4^{m-1}) = 1 + 4t, \tag{11}$$

for some  $t \in \mathbb{N}$ . Therefore, by Theorem 3 (i),  $s_m$  reduces to a number less than itself in 3 Collatz iterations. That is,  $T^3(s_m) < s_m$ .

*Example 4.* Let  $s_m = 1 + (an + b) + (an + b)^2 + \dots + (an + b)^m$ , where  $\gcd(a, b) = 4$ . Then,  $s_m$  reduces to a natural number less than itself in 3 Collatz iterations. We can rewrite  $s_m$  as

TABLE 1: Theorem 3 (iii) examples.

Iteration	$t = 1: n = 43$	$t = 7: n = 247$	$t = 14: n = 459$
1	130	742	1378
2	65	371	689
3	196	1114	2068
4	98	557	1034
5	49	1672	517
6	148	836	1552
7	74	418	776
8	37	209	388

TABLE 2: Theorem 3 (iv) examples.

Iteration	$t = 1: n = 135$	$t = 7: n = 911$	$t = 14: n = 1851$
1	406	2734	5554
2	203	1367	2777
3	610	4102	8332
4	305	2051	4166
5	916	6154	2083
6	458	3077	6250
7	229	9232	3125
8	688	4616	9376
9	344	2308	4688
10	172	1154	2344
11	86	577	1172

TABLE 3: Theorem 3 (v) examples.

Iteration	$t = 1: n = 295$	$t = 7: n = 1871$	$t = 14: n = 3707$
1	886	5614	11122
2	443	2807	5561
3	1330	8422	16684
4	665	4211	8342
5	1996	12634	4171
6	998	6317	12514
7	499	18952	6257
8	1498	9476	18772
9	749	4738	9386
10	2248	2369	4693
11	1124	7108	14080
12	562	3554	7040
13	281	1777	3520

$$\begin{aligned} s_m &= 1 + (an + b)[1 + (an + b) + \dots + (an + b)^{m-1}] \\ &= 1 + 4\left(\frac{a}{4}n + \frac{b}{4}\right)[1 + (an + b) + \dots + (an + b)^{m-1}] \\ &= 1 + 4t, \end{aligned} \tag{12}$$

for some  $t \in \mathbb{N}$ . By Theorem 3 (i),  $T^3(s_m) < s_m$ .

It is important to note that these last two examples can be proven by computation without using Theorem 3 (i).

*Proof.* For the first example, let  $s_m = 1 + 4^n + 4^{2n} + \dots + 4^{nm} = (4^n)^{m+1} - 1/4^n - 1 = Q_1$ , which will be an odd integer. Then,  $Q_2 = T(Q_1) = 3(4^n)^{m+1} + 4^n - 4/4^n - 1 = 4[3(4^n)^{m+1} + 4^{n-1} - 1]/4^n - 1$ . Since  $Q_2$  is even,  $Q_4 = T(Q_3) = T(T(Q_2)) = 3(4^n)^{m+1} + 4^{n-1} - 1/4^n - 1$ . Note that we can rewrite  $Q_4$  as

$$\begin{aligned} \frac{3(4^n)^{m+1} + 4^{n-1} - 1}{4^n - 1} &= \frac{4(4^n)^{m+1} - 4^{n(m+1)-1} + 4^{n-1} - 1}{4^n - 1} \\ &= \frac{4^{n(m+1)} - 4^{n(m+1)-1} + 4^{n-1} - 1}{4^n - 1}. \end{aligned} \tag{13}$$

For all  $m, n \geq 1$ , we have  $4^{n-1} < 4^{n(m+1)-1}$ . So,

$$Q_4 < \frac{4^{n(m+1)} - 4^{n(m+1)-1} + 4^{n(m+1)-1} - 1}{4^n - 1} = \frac{4^{n(m+1)} - 1}{4^n - 1} = Q_1. \tag{14}$$

Therefore,  $T^3(s_m) < s_m$ . □

*Proof.* For the second example, let  $s_m = 1 + (12n+4) + (12n+4)^2 + \dots + (12n+4)^m$ , where  $a = 12$  and  $b = 4$ . So,  $s_m = (12n+4)^{m+1} - 1/12n+4 - 1 = (12n+4)^{m+1} - 1/3(4n+1) = Q_1$ . Since  $Q_1$  is odd, we can obtain  $Q_2 = (12n+4)^{m+1} - 1 + 4n + 1/4n + 1 = 4[(3n+1)(12n+4)^m + n]/4n + 1$ . Here,  $Q_2$  is even, so we can obtain  $Q_3 = 2[(3n+1)(12n+4)^m + n]/4n + 1$  and  $Q_4 = (3n+1)(12n+4)^m + n/4n + 1$ . We want to show that

$$\begin{aligned} \frac{(3n+1)(12n+4)^m + n}{4n+1} &< \frac{(12n+4)^{m+1} - 1}{3(4n+1)} \\ \Rightarrow 3(3n+1)(12n+4)^m + 3n + 1 &< (12n+4)^{m+1} \\ \Rightarrow 3(12n+4)^m + 1 &< \frac{(12n+4)^{m+1}}{3n+1} = \frac{(3n+1)^{m+1} 4^{m+1}}{3n+1} \\ \Rightarrow 3(12n+4)^m + 1 &< 4(12n+4)^m, \end{aligned} \tag{15}$$

which is clearly true  $\forall n \in \mathbb{N}$ . Therefore,  $T^3(s_m) < s_m$ . □

**2.2. Peak Values.** Analyzing the number of iterations required for given  $n$  to complete a Collatz sequence may lead to interesting results related to the Collatz conjecture, as can be seen above. Another fascinating set of values that may be of interest are those that are the maximum value within a Collatz sequence; we call these values “peak values.” The following seven theorems relate to peak values and their properties.

**Theorem 4.** Let  $n \in \mathbb{N}$ , and let  $T^*(n) = p$  denote the peak value of the Collatz sequence starting at  $n$ . Then,  $p \equiv 0 \pmod{4}$  (i.e.,  $4 \mid p$ ).

*Proof.* First, we show that  $p$  is even through contradiction. If  $p$  is odd, then  $p = 2m + 1$ , where  $m \in \mathbb{N}$ . Then,

$$T(p) = 3p + 1 = 6m + 4 > 2m + 1 = p, \tag{16}$$

which means that  $p$  is not the peak value in its Collatz sequence. Thus,  $p$  is even (i.e.,  $p/2 \in \mathbb{N}$ ). Now, we claim that  $p/2$  is even. If  $p/2$  was odd, then

$$T\left(\frac{p}{2}\right) = 3\left(\frac{p}{2}\right) + 1 = \frac{3p+2}{2} > p, \tag{17}$$

so  $p$  is not the peak value. By contradiction,  $p/2$  is even. Hence,  $p/4 \in \mathbb{N} \Rightarrow 4 \mid p$ . □

**Theorem 5.** Let  $u$  be the odd number immediately before  $p$ . Then,  $u \equiv 1 \pmod{4}$  (i.e.,  $u = 4l + 1$  for some  $l \in \mathbb{N}$ ).

*Proof.* Since  $u$  is odd, either  $u = 4l + 1$  or  $u = 4l + 3$  for some  $l \in \mathbb{N}$ . Suppose that  $u = 4l + 3$ . Then,  $p = T(u) = 12l + 10 = 2(6l + 5)$ . Since  $2 \nmid 6l + 5$ ,  $4 \nmid p$ . However,  $4 \mid p$  by Theorem 4. So, through contradiction, we know that  $u = 4l + 1$ . □

**Theorem 6.** Let  $v$  be the odd number immediately after  $p$ . Then,  $v < u$ .

*Proof.* By Theorem 2,  $u = 4l + 1 \Rightarrow p = T(u) = 12l + 4$ . So,  $p/4 = 3l + 1$ . Let  $3l + 1 = 2^e v$ , where  $e$  is the largest power such that  $2^e \mid 3l + 1$  (i.e.,  $v = 3l + 1/2^e$ ). Then,  $3l + 1/2^e < 4l + 1 \Rightarrow v < u$ . □

**Theorem 7.** Let  $u = 2 \cdot 4^t k + (4^{t+1} - 1/3)$ , where  $t \geq 1$ . If  $k$  is odd, then  $2 \cdot 4^t \mid p$  and  $v = 3k + 2$ .

*Proof.* We see that  $p = T(u) = 6 \cdot 4^t k + 4^{t+1} = 4^t(6k + 4) = 2 \cdot 4^t(3k + 2)$ . Since  $k$  is odd and  $3k + 2 \equiv 0 \pmod{2}$ , we obtain  $2 \cdot 4^t \mid p$  and  $v = 3k + 2$ . □

In the following, we have three special cases of Theorem 7 and two examples that demonstrate the use of this theorem.

*Special Case 1.* Let  $t = 1$ . Then,  $u = 8k + 5 \Rightarrow 8 \mid p$  and  $v = 3k + 2$  so long as  $k$  is odd.

*Special Case 2.* Let  $t = 2$ . Then,  $u = 32k + 21 \Rightarrow 32 \mid p$  and  $v = 3k + 2$  so long as  $k$  is odd.

*Special Case 3.* Let  $t = 3$ . Then,  $u = 2 \cdot 4^3 k + 85 \Rightarrow 128 \mid p$  and  $v = 3k + 2$  so long as  $k$  is odd.

*Example 5.* Let  $u = 13 = 8 \cdot 1 + 5$  and  $p = T(u) = 40$ . Then,  $8 \mid p$  and  $v = 3 \cdot 1 + 2 = 5$ .

*Example 6.* Let  $u = 53 = 32 \cdot 1 + 21$  and  $p = 160$ . Then,  $32 \mid p$  and  $v = 3 \cdot 1 + 2 = 5$ .

While Theorem 7 explores the case when  $k$  is odd, the following theorem covers the case when  $k$  is even.

**Theorem 8.** Let  $u = 2 \cdot 4^t k + (4^{t+1} - 1/3)$ . If  $k$  is even and  $k = 2l$  for some  $l \in \mathbb{N}$ , then we have the following cases:

- (1) If  $l$  is even, then  $4^{t+1} | p$  and  $v = 3l + 1$ .
- (2) If  $l$  is odd, let  $2^e | (3l + 1)$  and  $2^{e+1} \nmid (3l + 1)$ . Then,  $2^{2t+e+2} | p$  and  $v = 3l + 1/2^e$ .

*Proof.* We let  $p = 2 \cdot 4^t (3k + 2) = 2 \cdot 4^t (6l + 2) = 4 \cdot 4^t (3l + 1) = 4^{t+1} (3l + 1)$ .

- (1) If  $l$  is even, then  $3l + 1 \equiv 0 \pmod{2}$ . Hence,  $4^{t+1} | p$  and  $v = 3l + 1$ .
- (2) If  $l$  is odd, then let  $2^e | (3l + 1)$  and  $2^{e+1} \nmid (3l + 1)$ . We have  $p = T(u) = 4^{t+1} \cdot 2^e \cdot 3l + 1/2^e$ . Thus,  $2^{2t+e+2} | p$  and  $v = 3l + 1/2^e$ . □

Just as we did for Theorem 7, we provide the following special cases and examples for Theorem 8.

*Special Case 4.* Let  $t = 1$ . Then,  $u = 8k + 5$  and  $k = 2l$ .

- (1) If  $l$  is even, then  $16 | p$  and  $v = 3l + 1$ .
- (2) If  $l$  is odd, let  $2^e | (3l + 1)$  and  $2^{e+1} \nmid (3l + 1)$ . Then,  $2^{4+e} | p$  and  $v = 3l + 1/2^e$ .

*Special Case 5.* Let  $t = 2$ . Then,  $u = 32k + 21$  and  $k = 2l$ .

- (1) If  $l$  is even, then  $64 | p$  and  $v = 3l + 1$ .
- (2) If  $l$  is odd, let  $2^e | (3l + 1)$  and  $2^{e+1} \nmid (3l + 1)$ . Then,  $2^{6+e} | p$  and  $v = 3l + 1/2^e$ .

*Special Case 6.* If  $3l + 1/2^e = 1$ , then  $3l + 1 = 2^e \implies l = 2^e - 1/3$ . Hence,  $l = 1$  if  $e = 2$ ,  $l = 5$  if  $e = 4, \dots$ ; thus,

$$u = 2 \cdot 4^t (2l) + \frac{4^{t+1} - 1}{3} = \frac{1}{3} [4^{t+1} \cdot 2^e - 1] \tag{18}$$

$$= \frac{1}{3} [4^{t+1} \cdot 4^m - 1] = \frac{1}{3} [4^a - 1],$$

where  $e = 2m$  is a nonzero even integer and  $a = t + 1 + m \in \mathbb{N}$ . Hence, if  $u$  is an odd number such that  $u = (1/3)[4^a - 1]$ , then  $p = T(u) = 4^a$  and  $v = 1$ . Note that  $u$  will reach 1 in  $2a + 1$  steps.

*Example 7.* Let  $t = 1$ ,  $u = 21 = 8 \cdot 2 + 5$ ,  $k = 2$ , and  $p = 64$ . Then,  $2^e | (3l + 1) \implies 2^e | 4$  and  $2^{e+1} \nmid 4$ . Thus,  $e = 2$  and  $2^{2t+2+e} = 2^6 = 64 | p$ . Lastly,  $v = 3l + 1/2^e = 1$ .

*Example 8.* Let  $t = 1$ ,  $u = 37 = 8 \cdot 4 + 5$ ,  $k = 4$ , and  $p = 112$ . Then,  $2^e | (3l + 1) \implies 2^e | 7$  and  $2^{e+1} \nmid 7$ . Thus,  $e = 0$  and  $2^{2t+2+e} = 2^4 = 16 | p$ . Lastly,  $v = 3l + 1/2^e = 7$ .

In the following, we present a general result with some accompanying examples for a certain form in  $u$  and its respective peak values.

**Theorem 9.** Let  $u = 4^t k + (4^t - 1/3)$ . If  $k$  is even (i.e.,  $k = 2l$  for some  $l \in \mathbb{N}$ ) and  $t \in \mathbb{N}$ , then  $4^t | p$  and  $v = 6l + 1$ .

*Proof.* We have  $p = T(u) = 3k(4^t) + 4^t = 4^t (3k + 1) = 4^t (6l + 1)$ . Since  $6l + 1 \equiv 0 \pmod{2}$  (i.e.,  $2 \nmid (6l + 1)$ ),  $4^t | p$  and  $v = 6l + 1$ . □

*Example 9.* Consider the case when  $t = 1$ . Then,  $u = 4k + 1 = 8l + 1$ ,  $p = 4(6l + 1) \implies 4 | p$ , and  $v = 6l + 1$ . Suppose  $u = 65 = 8(8) + 1$ . Then,  $l = 8$  and  $T(u) = 4(6 \cdot 8 + 1) = 196$ . That is,  $4 | 196$ , and  $v = 6l + 1 = 49$ .

*Example 10.* Consider the case when  $t = 2$ . Then,  $u = 4^2 k + 5 = 32l + 5$ ,  $p = 4^2 (6l + 1) \implies 4^2 | p$ , and  $v = 6l + 1$ . Suppose  $u = 101 = 32(3) + 5$ . Then,  $l = 3$ ,  $p = 4^2 (6l + 1) = 16(6 \cdot 3 + 1) = 16 \cdot 19 = 304$ , and  $v = 6l + 1 = 19$ .

*Example 11.* Consider the case when  $t = 3$ . Then,  $u = 4^3 k + 21 = 128l + 21$ ,  $p = 4^3 (6l + 1) \implies 4^3 | p$ , and  $v = 6l + 1$ . Suppose  $u = 149 = 128(1) + 21$ . Then,  $l = 1$ ,  $64 | p$ , and  $v = 6l + 1 = 7$ .

The following theorem is similar to Theorem 9; however, it explores the cases that arise when  $k$  is odd, specifically when  $l$  is odd and when  $l$  is even. Four accompanying examples are provided.

**Theorem 10.** Let  $u = 4^t k + (4^t - 1/3)$  and  $k$  be odd (i.e.,  $k = 2l + 1$  for some  $l \in \mathbb{N}$ ).

- (1) If  $l$  is odd, then  $2^{2t+1} | p$  and  $v = 3l + 2$ .
- (2) Suppose  $l$  is even, and let  $2^e | (3l + 2)$  and  $2^{e+1} \nmid (3l + 2)$ . Then,  $2^{2t+e+1} | p$  and  $v = 3l + 2/2^e$ .

*Proof.* We use a similar proofing technique as we used for Theorem 8.

- (1) We have that  $p = T(u) = 3 \cdot 4^t k + 4^t = 3 \cdot 4^t (2l + 1) + 4^t = 4^t (6l + 4) = 2 \cdot 4^t (3l + 2) = 2 \cdot 2^{2t} (3l + 2) = 2^{2t+1} (3l + 2)$ . Since  $2 \nmid (3l + 2)$ , we must have that  $2^{2t+1} | p$  and  $v = 3l + 2$ .
- (2) From (1),  $p = T(u) = 2 \cdot 4^t (3l + 2)$ . Since  $l$  is even,  $2^e | (3l + 2)$ , and  $2^{e+1} \nmid (3l + 2)$ , we have  $p = T(u) = 2 \cdot 4^t \cdot 2^e (3l + 2/2^e) = 2^{2t+e+1} (3l + 2/2^e)$ . Thus,  $2^{2t+e+1} | p$  and  $v = 3l + 2/2^e$ . □

*Example 12.* Let  $t = 1$ . Then,  $u = 4k + (4 - 1/3) = 4k + 1$ . Suppose  $k = 3$  and  $l = 1$ . Then,  $u = 13$  and  $p = T(u) = 40$ . Here, we have  $2^{2t+1} = 8 | p$  and  $v = 3l + 2 = 5$ .

*Example 13.* Let  $t = 2$ . Then,  $u = 16k + (4^2 - 1/3) = 16k + 5$ . Suppose that we still have  $k = 3$  and  $l = 1$ . Then,  $u = 53$  and

```

(1) Obtain the maximum number for the algorithm to consider:  $M$ 
(2) for int  $i = 3; i \leq n; i = i + 2$  do
(3)   Set counters  $c_0 = 0$  and  $c_1 = 0$ , and set  $i_0 = i$ 
(4)   while  $i \neq 1$  do
(5)     if  $i \% 2 == 0$  then
(6)       Set  $i = i / 2$ 
(7)     else  $\{i \% 2 \neq 0\}$ 
(8)       Set  $i = 3 * i + 1$ 
(9)     end if
(10)    Increase  $c_0 = c_0 + 1$ 
(11)    if  $i < i_0$  &&  $c_1 == 0$  then
(12)      Set  $c_1 = c_0$ 
(13)    end if
(14)  end while
(15)  Output  $i, c_0, c_1$ 
(16)  Convert  $i$  to bases 2, 3, 4, then output
(17) end for

```

ALGORITHM 1: Iteration analysis.

```

(1) Obtain the maximum number for the algorithm to consider :  $N$ 
(2) for int  $i = 3; i \leq N; i = i + 2$  do
(3)   Set tmp =  $i$  and peak =  $i$ , and set counter  $c = 0$ 
(4)   while  $i \neq 1$  do
(5)     if  $tmp \% 2 == 0$  then
(6)       Set tmp =  $tmp / 2$ 
(7)     else  $\{tmp \% 2 \neq 0\}$ 
(8)       Set tmp =  $3 * tmp + 1$ 
(9)     end if
(10)    if peak < tmp then
(11)      Set peak = tmp
(12)    end if
(13)     $c = c + 1$ 
(14)  end while
(15)  Output  $i, \text{peak}, c$ 
(16) end for

```

ALGORITHM 2: Peak analysis.

$p = T(u) = 160$ . Here, we have  $2^{2t+1} = 32|p$  and  $v = 3l + 2 = 5$ .

*Example 14.* Let  $t = 1$ . Then,  $u = 4k + 1$ . Let  $l = 4$  and  $k = 9$ . Then,  $u = 37$  and  $p = T(u) = 112$ . Now,  $2^e|(3l + 2) \Rightarrow 2^e|14$ , so  $e = 1$ . Hence,  $2^{2t+e+1} = 16$  and  $16|112$ . Lastly,  $v = 3l + 2/2^e = 7$ .

*Example 15.* Let  $t = 2$ . Then,  $u = 16k + 5$ . Let  $l = 4$  and  $k = 9$ . Then,  $u = 149$  and  $p = T(u) = 448$ . Now,  $2^e|(3l + 2) \Rightarrow 2^e|14$ , so once again  $e = 1$ . Hence,  $2^{2t+e+1} = 64$  and  $64|448$ . Lastly,  $v = 3l + 2/2^e = 7$ .

While these theorems and examples are valuable insights into Collatz dynamics, there are certainly more relationships to uncover. Further peak value analysis can be executed using Algorithm 1 in Section 3.

### 3. The Algorithms

The provided theorems and corollaries serve as a great introduction to a family of mathematical relationships that may cover all of  $\mathbb{N}$ . There are undoubtedly more relationships to discover pertaining to the Collatz conjecture. It is our hope that the three algorithms that follow will aid in the discovery of such relationships and, consequently, the proof of the conjecture.

*3.1. Visualizing the Collatz Conjecture.* Most attempts to visualize the Collatz conjecture come in the form of inverse binary trees. For example, in Figure 1, the left parent of a node is the even integer that results in the child node, and the right parent of a node, if it exists, is the odd integer that results in the child node. Some nodes, such as “8,” only have a left parent node: 16. Note that there does not exist an odd integer  $n$  such that  $3n + 1 = 8$ ; therefore, “8” does not have a



TABLE 4: Output from Algorithm 3.

$N$	Iterations until 1	Iterations until $<N$	Base 2	Base 3	Base 4
3	7	6	11	10	3
5	5	3	101	12	11
7	16	11	111	21	13
9	19	3	1001	100	21
11	14	8	1011	102	23
13	9	3	1101	111	31
15	17	11	1111	120	33
17	12	3	10001	122	101
19	20	6	10011	201	103
21	7	3	10101	210	111

```

(1) Obtain integer  $n$  for the current child node
(2) while input does not fail do
(3)   for (int  $i = 0; i < 9; i = i + 1$ ) do
(4)     Output  $n * \text{pow}(2, i + 1)$ 
(5)   end for
(6)   if right parent exists for  $n$  then
(7)     Output  $n - 1/3$ 
(8)     for (int  $i = 0; i < 9; i = i + 1$ ) do
(9)       Output  $n - 1/3 * \text{pow}(2, i + 1)$ 
(10)      if right parent exists for  $m = n - 1/3$  then
(11)        Output  $m - 1/3$ 
(12)      end if
(13)    end for
(14)  end if
(15)  Enter a new value  $n$  for a child node
(16) end while

```

ALGORITHM 3: Inverse binary tree generation.

right parent. However, “16” has both a left and a right parent because  $T(32) = 16$  and  $T(5) = 16$ .

Algorithm 2 provides a framework for building a visualization tool, where one can construct a custom, interactive interface for navigating an inverse binary tree similar to that in Figure 1. Such an algorithm can facilitate efficient exploration of the properties of the Collatz conjecture.

Console output formatting is subject to the reader’s taste. Note the for-loop cycle through 10 values. This number of cycles was arbitrarily chosen and can be changed without affecting the accuracy of the algorithm. The code used to generate Figure 1 is available as discussed in the data availability section at the end of the manuscript.

**3.2. Analyzing Iterations of the Collatz Conjecture.** As stated in the previous section, a rigorous study of relationships regarding the Collatz conjecture will aid in this conjecture’s proof. While much attention has been placed on proving the Collatz conjecture for subsets of  $\mathbb{N}$ , we propose an alternative route: analyzing the number of iterations it takes for some positive integer  $n$  to either (1) reach 1 or (2) reach a value less than itself ( $<n$ ).

Algorithm 3 generates the output containing the number of iterations each value  $n$  requires to reach 1, the number of

iterations until it reaches a value less than  $n$ , and the representation of  $n$  in bases 2–4. An analysis of this output for various subsets of  $\mathbb{N}$  may result in novel discoveries that can ignite new lines of attack to prove this conjecture. For example, some positive integers such as 381,727 and 2,788,008,987 require 282 and 729 iterations, respectively, to reach a value less than itself. Perhaps an analysis of the properties of these numbers in various bases will render fruitful discussion. An example of the output from this algorithm is shared in Table 4.

This algorithm can be easily modified to obtain the maximum number of iterations required to reach 1 for all natural numbers up to  $M$ , which we refer to as the starting integer since it is specified at the beginning of the algorithm. The output can then be analyzed to determine how the number of iterations increases as  $M$  increases. Figure 4 denotes this relationship up to  $M = 3$  billion, where the  $x$ -axis is  $\log(M)$  to improve readability; this provides hope that the conjecture is true since it appears that the number of iterations converges to a finite number as  $M$  increases.

**3.3. Determining Peaks in Collatz Sequences.** It may be of interest to explore the highest value that a given natural number  $n$  will reach before completing its Collatz sequence.

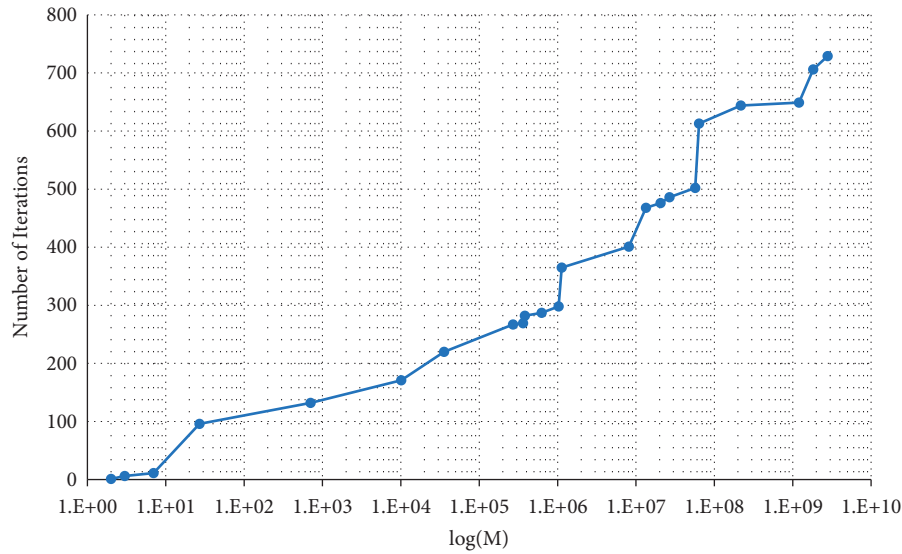


FIGURE 4: Logarithm of starting integer vs. number of iterations required for the Collatz sequence.

For example, the sequence starting with  $n = 15$  has a maximum value (or peak) of 160 before it reaches 1 on its 17th iteration. Perhaps an analysis of these peak numbers will reveal insights related to Collatz sequences and the conjecture. Algorithm 1 determines the peak values of the Collatz sequences starting at each odd natural number up to a given  $N$ , as well as the number of iterations it takes for each number to reach 1. An example of a similar output can be found in tables in [18].

In this algorithm, we restrict ourselves to positive, odd integers up to 100,000. However, one can modify the for-loop in line 2 to include as large a natural number as necessary. One intriguing observation is that, of the first 1539 odd natural numbers, for 407 (26.44%) of them, the peaks occur at 9232. No other peak number has this high of a percentage, with the second most frequent peak in the 1% range. It is not obvious if 9232 has any significance or if the peaks are somehow related.

#### 4. Conclusion

Although the proof of the Collatz conjecture seems intractable at this time, the contributions of various scholars from around the world are forming an excellent foundation for future approaches. Relationships that cover subsets of  $\mathbb{N}$  decrease the amount of natural numbers that need to be checked against the conjecture. If enough relationships are provided or if a generic relationship can be established, the proof of the Collatz conjecture will be within reach.

Traditional approaches to proving the conjecture are rather direct, whether mathematical or computational in nature. These approaches attempt to show that each natural number will eventually reach 1 in its Collatz sequence. In Section 2.1, we contribute to traditional approaches by

presenting a few relationships between subsets of  $\mathbb{N}$  and their Collatz sequences; the delineated work should assist in the discovery of more relationships that decrease the amount of positive integers that need to be checked against the Collatz conjecture. If more relationships were to be discovered, then we can focus on proving that subsets of  $\mathbb{N}$  satisfy the Collatz conjecture rather than all of  $\mathbb{N}$ . Ideally, one would establish enough relationships that cover the entire space of  $\mathbb{N}$ .

Once the underlying patterns in Section 2.1 are identified, more relationships between subsets of  $\mathbb{N}$  and their Collatz sequences can be explored. The presented relationships are certainly not exhaustive of all that can and should be discovered. However, the work in Section 2.1 can be built upon by discovering similar relationships, which will reduce the workload of checking  $\mathbb{N}$  with the Collatz conjecture.

While Section 2.1 is our attempt to aid the proof of the Collatz conjecture in a traditional manner, Section 2.2 is our attempt using a novel approach. In Section 2.2, seven theorems, multiple special cases, and numerous examples are provided that suggest an alternative approach to proving the Collatz conjecture: if every natural number  $n$  eventually becomes a number less than itself while iterating through the Collatz sequence, the conjecture will be proven. This approach is yet to be declared in the academic literature, yet it seems promising. The delineated work provides a foundation for future work to use, and it hopefully acts as a catalyst for an alternative approach to proving the Collatz conjecture.

In Section 3, the three algorithms that were presented in pseudo-code merely augment the mathematical work completed in Section 2. These computational tools will enable other researchers to visualize the Collatz conjecture (Section 3.1), explore various properties between  $\mathbb{N}$  and

Collatz sequences (Section 3.2), and determine peak values in these sequences (Section 3.3). While presented in pseudocode, C++ implementations may be found in the following section.

To prove the Collatz conjecture is to solve a marvelous mathematical puzzle that has stumped countless minds. It will undoubtedly be a collective effort that solves this puzzle—an effort that we invite the reader to join.

## Data Availability

All data, algorithms, and codes used for this article can be found at <https://github.com/michaelschwob/CollatzConjecture>.

## Disclosure

This article's preprint can be found in [19].

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

The authors would like to thank Professor Curtis Cooper for providing [4, 5] and Professor Aruhn Venkat for computational assistance.

## References

- [1] Ş. Andrei and C. Masalagiu, "About the Collatz conjecture," *Acta Informatica*, vol. 35, no. 2, pp. 167–179, 1998.
- [2] R. Guy, *Unsolved Problems in Number Theory*, Springer Science & Business Media, Berlin, Germany, 2004.
- [3] J. C. Lagarias, Ed., *The Ultimate Challenge: The  $3x + 1$  Problem*, American Mathematical Society, Providence, RI, USA, 2010.
- [4] T. Tao, "Almost all orbits of the collatz map attain almost bounded values," 2019, <https://arxiv.org/abs/1909.03562>.
- [5] K. Hartnett, "Mathematician Proves Huge Result on Dangerous Problem," *Quanta Magazine*, 2019.
- [6] W. Ren, "Reduced Collatz dynamics data reveals properties for the future proof of Collatz conjecture," *Data*, vol. 4, no. 2, pp. 89–93, 2019.
- [7] W. Ren, "A new approach on proving Collatz conjecture," *Journal of Mathematics*, vol. 2019, no. 1, 12 pages, Article ID 6129836, 2019.
- [8] A. N. Arslan, "Methods for constructing Collatz numbers," *Notes on Number Theory and Discrete Mathematics*, vol. 24, no. 2, pp. 47–54, 2018.
- [9] R. Zarnowski, "A refinement of the  $3x + 1$  conjecture," *Notes on Number Theory and Discrete Mathematics*, vol. 26, no. 3, pp. 234–244, 2020.
- [10] J. C. Lagarias, "The  $3x + 1$  problem and its generalizations," *The American Mathematical Monthly*, vol. 92, no. 1, pp. 3–23, 1985.
- [11] K. S. LaTourette, *Explorations of the Collatz Conjecture*, Citeseer, 2007.
- [12] J. P. Van Bendegem, "The Collatz conjecture. A case study in mathematical problem solving," *Logic and Logical Philosophy*, vol. 14, pp. 7–23, 2005.
- [13] D. Renza, S. Mendoza, and L. D. M. Ballesteros, "High-uncertainty audio signal encryption based on the Collatz conjecture," *Journal of Information Security and Applications*, vol. 46, pp. 62–69, 2019.
- [14] D. Ballesteros, J. Peña, and D. Renza, "A novel image encryption scheme based on Collatz conjecture," *Entropy*, vol. 20, no. 12, pp. 901–903, 2018.
- [15] H. Ma, C. Jia, S. Li, W. Zheng, and D. Wu, "Xmark: dynamic software watermarking using Collatz conjecture," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 11, pp. 2859–2874, 2019.
- [16] M. A. Idowu, "A novel theoretical framework formulated for information discovery from number system and Collatz conjecture data," *Procedia Computer Science*, vol. 61, pp. 105–111, 2015.
- [17] D. Jenber, "Collatz theorem," 2018, <https://arxiv.org/abs/1811.08500>.
- [18] dCode, "Collatz conjecture," 2021, <https://www.dcode.fr/collatzconjecture>.
- [19] M. Schwob, P. Shiue, and R. Venkat, "Novel theorems and algorithms relating to the collatz conjecture," 2021, <https://arxiv.org/abs/2104.10713>.