

Looping and Divergence in the Collatz Conjecture

Jai Sharma^{*}, Akshat Jha[†], Sambhabi Bose[‡]
Garrett Heller, Dr. Nick Castro[§]

July 2020

^{*}Monta Vista High School, Cupertino, CA, United States

[†]AK Children Academy, Ghaziabad, Uttar Pradesh, India

[‡]McNair Academic High School, Jersey City, NJ, United States

[§]University of Arkansas, AR, United States

Abstract

In this paper, we investigate the possible scenarios in which a number does not satisfy the Collatz Conjecture. Specifically, we examine numbers which may have a looping Collatz reduction sequence as well as numbers which may lead to a diverging Collatz reduction sequence. In order to investigate these, we look at the parity of the numbers in a general Collatz reduction sequence. Further, we examine cases in which these parity cycles repeat themselves infinitely in the reduction sequence. Through the research conducted in the paper, we formulate a necessary condition for looping in the Collatz Conjecture. We also prove that if a number has a diverging reduction sequence, then it must generate an infinite non-repeating parity cycle.

Introduction to the Collatz Conjecture

In this paper, we will discuss the ideas of looping and divergence in the Collatz Conjecture. As background, the following is the Collatz Conjecture problem statement:

Conjecture (The Collatz Conjecture) Define a piecewise function f over the domain of positive integers as the following:

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

The Collatz Conjecture states that, for all x , $f^n(x) = 1$ for some non-negative integer n , where $f^i(x)$ represents the Collatz Conjecture function being applied to x exactly i times.

Summary of Findings

Consider applying the Collatz Conjecture function repeatedly to number m :

Definition (Collatz Reduction Sequence) Let $f(x)$ be the result of the Collatz Conjecture function on the positive integer x .

If the number m satisfies the Collatz Conjecture, then the **Collatz Reduction Sequence** is the following:

$$m, f(m), f^2(m), f^3(m), \dots, 1.$$

If the number m does not satisfy the Collatz Conjecture, then the **Collatz Reduction Sequence** is the following:

$$m, f(m), f^2(m), f^3(m), \dots, f^a(m)$$

for arbitrarily large a .

Suppose that the number m does not satisfy the Collatz Conjecture. Then, there are two ways this can occur:

1. $f^i(m) = f^j(m)$ for some positive integers i and j such that $j < i$. We call this **Looping**. Looping has two cases:
 - **Case 1:** $f^i(m) = m$ for some positive integer i . We call this **Direct Looping**.
 - **Case 2:** $f^k(m) \neq m$ for any positive integer k . We call this **Indirect Looping**. In this case, there exists some positive integer j such that $f^j(m)$ leads to Direct Looping in its Collatz Reduction Sequence. Here, we only focus on numbers that lead to Direct Looping; if we find the positive odd integers m that lead to Direct Looping, then we can compute all of the numbers that lead to Indirect Looping as $f^{-1}(m)$, $f^{-2}(m)$, $f^{-3}(m)$, etc. for all m (where $f^{-1}(m)$ is the inverse function of m). Therefore, any restrictions we put on Direct Looping also apply for Indirect Looping. Henceforth, all references to "Looping" will be references to **Direct Looping**.
2. $f^i(m)$ increases without bound as i increases to ∞ and $f^i(m) \neq 1$ for any positive integer i . We call this **Divergence**.

We address both of these cases separately after a critical definition:

Definition (Parity Sequence) A **Parity Sequence** is the sequence formed by the parities of the elements of the Collatz Reduction Sequence for a number.

For example, take the number 3. The following is the Collatz Reduction Sequence for 3:

$$3, 10, 5, 16, 8, 4, 2, 1.$$

Then, the corresponding Parity Sequence is the following:

$$1 \text{ odd}, 1 \text{ even}, 1 \text{ odd}, 1 \text{ even}, 1 \text{ even}, 1 \text{ even}, 1 \text{ even}, 1 \text{ odd},$$

or more concisely:

$$1 \text{ odd}, 1 \text{ even}, 1 \text{ odd}, 4 \text{ evens}, 1 \text{ odd}.$$

Suppose that for some number m , the Collatz Reduction Sequence, and therefore the Parity Sequence, is infinite in length and periodic with some finite period. Call one period of the infinitely long Parity Sequence of m a **Parity Cycle**.

In the case of Looping, the Parity Sequence must be infinite and periodic, as the elements in the Collatz Reduction Sequence themselves repeat; therefore, an infinitely repeating Parity Cycle of finite length must exist. Additionally, in the case of Divergence, the Parity Sequence must be infinite, and there is a chance that this infinite Parity Sequence is periodic; therefore, an infinitely repeating Parity Cycle of finite length can exist in this case as well.

In this paper, we focus on these infinitely repeating Parity Cycles of finite length that occur in the said cases. *Specifically, we use the idea of these Parity Cycles to set restrictions on numbers which can lead to Looping or Divergence in the Collatz Conjecture.*

Note that the only case not covered here by the concept of Parity Cycles which can still result in a number not satisfying the Collatz Conjecture is when the number generates an infinite aperiodic Parity Sequence through its Collatz Reduction Sequence, which can only happen in the case of Divergence.

Consider an odd positive integer m which generates an infinitely repeating Parity Cycle of finite length through its Collatz Reduction Sequence. Let the following be the general Parity Cycle of m :

$$1 \text{ odd}, p_1 \text{ evens}, 1 \text{ odd}, p_2 \text{ evens}, 1 \text{ odd}, p_3 \text{ evens}, \dots, 1 \text{ odd}, p_b \text{ evens}.$$

Notice that we start with an odd number and end with an odd number after the final series of even numbers. This Parity Cycle can then be thought of as a function that takes in an odd integer, performs the Collatz Conjecture function according to the Parity Cycle, and outputs an odd integer which is also in the domain of the same function (meaning that we can put the resulting number into the Parity Cycle again by definition). For future reference, let this Parity Cycle function be $g(x)$, where $x = m$ in the above example.

From this parity cycle, we define the following:

$$\begin{aligned} r &= 3^b, \\ q &= 3^{b-1} + 3^{b-2}2^{p_1} + 3^{b-3}2^{p_1+p_2} + \dots + 2^{p_1+p_2+p_3+\dots+p_{b-1}}, \\ s &= p_1 + p_2 + p_3 + \dots + p_b. \end{aligned}$$

We can then establish our findings on Looping in the Collatz Conjecture:

Theorem (Looping in the Collatz Conjecture) An odd integer m generates a Looping Collatz Reduction Sequence if and only if it is of the form:

$$m = \frac{q}{2^s - r},$$

where r , q , and s are defined by the Parity Cycle of m and $2^s > r$. We can also establish our findings on Divergence in the Collatz Conjecture:

Theorem (Divergence in the Collatz Conjecture) An odd integer m generates an infinite Diverging Collatz Reduction Sequence only if the infinitely long Parity Sequence corresponding to this Collatz Reduction Sequence is aperiodic.

Proof of Findings

Note that all of the definitions used in this proof (that have not been defined here) are defined in the Summary of Findings.

Consider $g(x)$, a Parity Cycle function that takes an odd integer input, performs the Collatz Conjecture according to

some Parity Cycle, and returns an odd integer that is also in the domain of g . Let the Parity Cycle corresponding to $g(x)$ be:

$$1 \text{ odd}, p_1 \text{ evens}, 1 \text{ odd}, p_2 \text{ evens}, 1 \text{ odd}, p_3 \text{ evens}, \dots, 1 \text{ odd}, p_b \text{ evens}.$$

By the Collatz Conjecture function, for each "odd," we can multiply by 3 and add 1 to our number. Similarly, for each "even," we can divide our number by 2.

Then, $g(x)$ can be computed in terms of the p_i 's:

$$\begin{aligned} g(x) &= \frac{3 \left(\frac{3 \left(\frac{3 \left(\frac{3x+1}{2^{p_1}} \right) + 1 \right) + 1}{2^{p_2}} \right) + 1}{\vdots} \right) + 1}{2^{p_b}} \\ &= \frac{3^b x + 3^{b-1} + 3^{b-2} 2^{p_1} + 3^{b-3} 2^{p_1+p_2} + \dots + 2^{p_1+p_2+p_3+\dots+p_{b-1}}}{2^{p_1+p_2+p_3+p_4+\dots+p_b}}. \end{aligned}$$

Due to the bulkiness of this representation, we make some substitutions to make this form more readable. We define the following:

$$\begin{aligned} r &= 3^b, \\ q &= 3^{b-1} + 3^{b-2} 2^{p_1} + 3^{b-3} 2^{p_1+p_2} + \dots + 2^{p_1+p_2+p_3+\dots+p_{b-1}}, \\ s &= p_1 + p_2 + p_3 + \dots + p_b. \end{aligned}$$

From the definition of q , we deduce that $q \equiv 1 \pmod{2}$. Now, we substitute these values back into the representation for $g(x)$:

$$g(x) = \frac{rx + q}{2^s}.$$

By the definition of a Parity Cycle, $g(x)$ may be applied repeatedly to an x in the domain of g . Let $g^a(x)$ be the composition of $g(x)$ on the number x exactly a times. Therefore, this would mean that $g^a(x)$ is an odd integer which is also in the domain of g for any positive integer a .

Similar to how we expanded $g(x)$, we can also expand $g^a(x)$ using the concise form we just constructed:

$$\begin{aligned} g^a(x) &= \frac{r \left(\frac{r \left(\frac{r \left(\frac{rx+q}{2^s} \right) + q \right) + q}{\vdots} \right) + q}{2^s} \\ &= \frac{r^a x + qr^{a-1} + qr^{a-2} 2^s + qr^{a-3} 2^{2s} + \dots + q 2^{(a-1)s}}{2^{as}} \\ &= \frac{r^a x + q (r^{a-1} + r^{a-2} 2^s + r^{a-3} 2^{2s} + \dots + 2^{(a-1)s})}{2^{as}}. \end{aligned}$$

The sum in the numerator encased by parentheses is a geometric sequence with common ratio $2^s/r$, so this form can be simplified by taking the sum of a geometric sequence:

$$\begin{aligned} & r^{a-1} + r^{a-2}2^s + r^{a-3}2^{2s} + \dots + 2^{(a-1)s} \\ &= \frac{2^{as} - r^a}{2^s - r}. \end{aligned}$$

Substituting this for the geometric series in the numerator of $g^a(x)$, we get:

$$\begin{aligned} g^a(x) &= \frac{r^a x + q \left(\frac{2^{as} - r^a}{2^s - r} \right)}{2^{as}} \\ g^a(x) &= \left(\frac{r}{2^s} \right)^a \left(x - \frac{q}{2^s - r} \right) + \frac{q}{2^s - r}. \end{aligned}$$

This final form represents the *odd integer* given after a applications of the Parity Cycle corresponding to g on a number x in the domain of g . Note that once x is determined, r , q , and s are uniquely defined *constants*.

Now, we enforce the restriction that $g^a(x) \equiv 1 \pmod{2}$, which should be true by the definition of g :

$$g^a(x) = \left(\frac{r}{2^s} \right)^a \left(x - \frac{q}{2^s - r} \right) + \frac{q}{2^s - r} \equiv 1 \pmod{2}.$$

Multiplying both sides of this congruence statement by the odd number $2^s - r$, we get

$$\left(\frac{r}{2^s} \right)^a (x(2^s - r) - q) + q \equiv 2^s - r \equiv 1 \pmod{2}.$$

We know $q \equiv 1 \pmod{2}$ by definition, so when we subtract q on both sides, we get:

$$\left(\frac{r}{2^s} \right)^a (x(2^s - r) - q) \equiv 1 - q \equiv 0 \pmod{2}.$$

Thus, we see that

$$\left(\frac{r}{2^s} \right)^a (x(2^s - r) - q)$$

is an even integer. By the definition of g , a can be arbitrarily large and every other variable is constant, including x . From here, we have a few cases.

Case 1: Our first case is when

$$x(2^s - r) - q \neq 0.$$

In this case, since r and 2^s are relatively prime by definition, we must have that:

$$2^{as} \mid x(2^s - r) - q$$

for all a . However, $x(2^s - r) - q$ is constant, and 2^{as} is unbounded since a is unbounded. Therefore, the quotient of the two cannot be an integer and this case is impossible.

Case 2: Our second case is when

$$x(2^s - r) - q = 0.$$

Rearranging this, we find:

$$x = \frac{q}{2^s - r}.$$

Therefore,

$$\begin{aligned} g^a(x) &= \left(\frac{r}{2^s}\right)^a \left(x - \frac{q}{2^s - r}\right) + \frac{q}{2^s - r} \\ &= \left(\frac{r}{2^s}\right)^a (0) + \frac{q}{2^s - r} \\ &= 0 + \frac{q}{2^s - r} \\ &= x \end{aligned}$$

for all a . Thus, this case corresponds to Looping, not Divergence. Specifically, not only do the *parities* of the numbers repeat after every iteration of the Parity Cycle corresponding to g , but the *numbers* themselves repeat.

Before examining this any further, we see that it is not possible for a number to generate an infinite Diverging Collatz Reduction Sequence for which the corresponding infinitely long Parity Sequence is periodic with finite period. Therefore, a number can only Diverge if its infinitely long Parity Sequence is aperiodic, proving our finding on Divergence.

Now, we return to **Case 2**. We see that a number can lead to Looping if and only if

$$x = \frac{q}{2^s - r}.$$

Note that both the “if” part and the “only if” part are proven by the computation above for $g^a(x)$.

Further, since $x > 0$ and $q > 0$, we must have $2^s - r > 0$, or

$$2^s > r$$

if x leads to Looping.

Hence, we have also proved our finding on Looping, and the proof is complete. \square

Recommendations For Future Research

A Parity Cycle corresponding to a number x can lead to Looping if and only if:

$$x = \frac{q}{2^s - r}.$$

For future research, we can try to verify if an expression of this form can be an integer.

That is, we can try to find a sequence of numbers say p_1, p_2, \dots, p_b (which would correspond to the Parity Cycle) such that

$$2^s - r = (2^{p_1+p_2+\dots+p_b} - 3^b) \mid (3^{b-1} + 3^{b-2}2^{p_1} + 3^{b-3}2^{p_1+p_2} + \dots + 2^{p_1+p_2+\dots+p_b-1}).$$

The only known solution to this form is when $b = 1$ and $p_1 = 2$ (there are no p_2, p_3 , etc. because $b = 1$). In this case, we

simply arrive at the trivial loop

1, 2, 4, 1, 2, 4, ...

A word of caution: Finding a sequence p_1, p_2, \dots, p_b which satisfies the form presented above would prove that there exists a Parity Sequence for which Looping can occur, but not necessarily whether any number satisfies this discovered Parity Cycle.

References

- [1] "Collatz Conjecture." Wikipedia, Wikimedia, 19 Sept. 2020, https://en.wikipedia.org/wiki/Collatz_conjecture.
- [2] Tao, Terence. "Almost All Orbits of the Collatz Map Attain Almost Bounded Values." ArXiv.org, 7 June 2020, <https://arxiv.org/abs/1909.03562>.