Proof of the Collatz Conjecture

Henok Tadesse
email: entkidmt@yahoo.com

21May 2023

Abstract

Take any positive integer N. If it is odd, multiply it by three and add one. If it is even, divide it by two. Repeatedly do the same operations to the results, forming a sequence. It is found that, whatever the initial number we choose, the sequence will eventually descend and reach number 1, where it enters a closed loop of 1- 4 - 2 - 1. This is known as the Collatz conjecture which states that the sequence always converges to 1. So far no proof has ever been found that this holds for every positive integer. In this paper we prove the Collatz conjecture by studying what happens as the sequence becomes infinitely long, thereby avoiding the chaos of finite length sequences. We have noted that the ratio between the number of odd operations and even operations continues to decrease as the sequence length increases, approaching zero for infinite sequence length. This leads to the only possibility that the sequence must eventually decouple from its staring value and enter a cycle, with the only possible cycle being the 1-4-2-1 cycle. We have obtained an equation for the final sequence of infinite length, which is the 1-4-2-1 closed loop:

$$C_n = \frac{\frac{3}{2^{Kn-1}} + 1}{2^{Kn}}$$

Introduction

The Collatz conjecture, originally proposed by Lothar Collatz in 1937[1], is a famous unsolved problem in number theory. It concerns the behavior of the iteration defined on positive integers by:

$$C(N) = \begin{cases} 3N + 1\,, & \text{if N is odd} \\ \dfrac{N}{2}, & \text{if N is even} \end{cases}$$

The conjecture asserts that for every positive integer N, repeated iterations of C eventually reaches the number 1. Despite extensive computational evidence and numerous partial results, a full proof has remained illusive[2,3].

Several surveys and studies have explored the problem’s rich structure and its generalizations (see, e.g. Lagarias[2]). This paper presents a novel approach that studies the properties of an infinite Collatz sequence, particularly the ratio between the number of even operations and odd operations.

Collatz Sequence of Odd Terms

Definition

Odd Collatz Sequence is an alternative representation containing only the odd terms of the sequence.

A term in a Collatz sequence can be even or odd. We start from the fact that even terms always lead to odd terms because of successive divisions by 2. Therefore, for any initial odd number N, an alternative representation of Collatz sequence can be defined as:

$$N \longrightarrow \frac{3N+1}{2^{K1}} \longrightarrow \frac{3\left(\frac{3N+1}{2^{K1}}\right)+1}{2^{K2}} \longrightarrow \frac{3\left(\frac{3\left(\frac{3N+1}{2^{K1}}\right)+1}{2^{K2}}\right)+1}{2^{K3}}$$

$$\longrightarrow \frac{3\left(\frac{3\left(\frac{3\left(\frac{3N+1}{2^{K1}}\right)+1}{2^{K2}}\right)+1}{2^{K3}}\right)+1}{2^{K4}} \longrightarrow \frac{3\left(\frac{3\left(\frac{3\left(\frac{3\left(\frac{3N+1}{2^{K1}}\right)+1}{2^{K2}}\right)+1}{2^{K3}}\right)+1}{2^{K4}}\right)+1}{2^{K5}}$$

$$\longrightarrow \frac{3\left(\frac{3\left(\frac{3\left(\frac{3\left(\frac{3\left(\frac{3N+1}{2^{K1}}\right)+1}{2^{K2}}\right)+1}{2^{K3}}\right)+1}{2^{K4}}\right)+1}{2^{K5}}\right)+1}{2^{K6}} \longrightarrow \cdot\ \cdot\ \cdot$$

All terms of the above sequence are odd numbers because every time an even term occurs in the sequence it is successively divided by 2 until an odd term occurs. That is, when an even term occurs, we don't include it into the sequence but divide it by an integral power of 2 until we get an odd term, which is included into the sequence. Thus, with this we have created an alternative Collatz sequence as shown above with all terms odd.

By expanding each term we get the following:

$$\frac{3(\frac{3N+1}{2^{K1}})+1}{2^{K2}} = \frac{3^2N+3+2^{K1}}{2^{K1+K2}}$$

$$\frac{3(\frac{3(\frac{3N+1}{2^{K1}})+1}{2^{K2}})+1}{2^{K3}} = \frac{3(\frac{3^2N+3+2^{K1}}{2^{K1+K2}})+1}{2^{K3}} = \frac{3^3N+3^2+3*2^{K1}+2^{K1+K2}}{2^{K1+K2+K3}}$$

$$\frac{3\left(\frac{3\left(\frac{3(\frac{3N+1}{2^{K1}})+1}{2^{K2}}\right)+1}{2^{K3}}\right)+1}{2^{K4}} = \frac{3\left(\frac{3^3N+3^2+3*2^{K1}+2^{K1+K2}}{2^{K1+K2+K3}}\right)+1}{2^{K4}}$$

$$= \frac{3^4N+3^3+3^2*2^{K1}+3*2^{K1+K2}+2^{K1+K2+K3}}{2^{K1+K2+K3+K4}}$$

$$\frac{3\left(\frac{3\left(\frac{3\left(\frac{3N+1}{2^{K1}}\right)+1}{2^{K2}}\right)+1}{2^{K3}}\right)+1}{2^{K4}}\right)+1}{2^{K5}} = \frac{3\left(\frac{3^4N+3^3+3^2*2^{K1}+3*2^{K1+K2}+2^{K1+K2+K3}}{2^{K1+K2+K3+K4}}\right)+1}{2^{K5}}$$

$$= \frac{3^5N + 3^4 + 3^3 * 2^{K1} + 3^2 * 2^{K1+K2} + 3 * 2^{K1+K2+K3} + 2^{K1+K2+K3+K4}}{2^{K1+K2+K3+K4+K5}}$$

and so on.

From the above, if we take the $(3N+1)/2$ term to be the first (n=1) term, then the n^{th} odd Collatz term (C_n) will be:

$$C_n = \frac{3^nN + 3^{n-1} + 3^{n-2} * 2^{K1} + ... +3^2 * 2^{K1+K2+...Kn-3} +3 * 2^{K1+K2+...Kn-2} + 2^{K1+K2+K3+...+Kn-1}}{2^{K1+K2+K3+...+Kn}} \quad ... (1)$$

C_n can be written as:

$$C_n = \frac{3^nN + 3^{n-1}}{2^{K1+K2+K3+...+Kn}} + \frac{3^{n-2} * 2^{K1} + ... +3^2 * 2^{K1+K2+...Kn-3} +3 * 2^{K1+K2+...Kn-2} + 2^{K1+K2+K3+...+Kn-1}}{2^{K1+K2+K3+...+Kn}}$$

The number of even Collatz operations compared to the number of odd operations

Intuitively, the number of even operations is significantly greater than the number of odd operations, and increasingly so at very large or infinitely large numbers. This follows from the fact that Collatz operation on an odd number always results in an even number, whereas an even operation may result in several successive even operations before an odd number occurs again, and this effect is even more pronounced at very large numbers. After every odd operation, the next term is always an even term and the probability that this even term has large powers of two as a factor (e.g. 2^{1000}) indefinitely increases as the length of the sequence approaches infinity. Although the rigor of such proof could be questioned (claiming that it is based on probability), one can arbitrarily increase this probability, making it approach 1 in the limit, that is by assuming infinite length of the sequence.

Assumption 1

The ratio of the number of odd operations to the number of even operations continuously decreases with increase in length of the sequence and approaches zero for infinite length of the sequence. That is:

$$\frac{n}{K1 + K2 + K3 + \ldots + Kn} \longrightarrow 0 \quad \text{at infinity}$$

Where n is the total number of Collatz odd operations and K1+K2+K3+ . . .+Kn is the total number of Collatz even operations in a sequence.

Assumption 2

This assumption concerns the terms in equation (1):

$$\frac{3^n}{2^{K1+K2+K3+\,.\,.\,.\,+Kn}} \longrightarrow 0 \;\; as \;\; n \;\; and \;\; K1 + K2 + K3 + \cdots + Kn \longrightarrow \infty$$

We postpone the proof of these assumptions. Later on we will show that Assumption 2 follows from Assumption 1. We can see that the first term is the coefficient of N. The convergence of these terms to zero at infinity (infinite length of the sequence) would reflect the fact that the number of even Collatz operations is significantly greater than the number of odd operations in a sequence, which intuitively means that the sequence always follows a general descent, eventually reaching the number 1.

Proof of the Collatz Conjecture

Theorem 1

All Collatz sequences converge to the trivial loop 1-4-2-1.

Proof

Using Assumption 2, we can see that the first terms of equation (1) diminish to zero for a sequence of infinite length.

$$C_n = \cancelto{0}{\frac{3^n N + 3^{n-1}}{2^{K1+K2+K3+\ ...\ +Kn}}} + \frac{3^{n-2} * 2^{K1} + \ ...\ +3^2 * 2^{K1+K2+\ ...\ Kn-3} + 3 * 2^{K1+K2+\ ...\ Kn-2} + 2^{K1+K2+K3+\ ...\ +Kn-1}}{2^{K1+K2+K3+\ ...\ +Kn}}$$

$$C_n \to \frac{3^{n-2} * 2^{K1} + 3^{n-3} * 2^{K1+K2} + \ ...\ +3^2 * 2^{K1+K2+\ ...\ Kn-3} + 3 * 2^{K1+K2+\ ...\ Kn-2} + 2^{K1+K2+K3+\ ...\ +Kn-1}}{2^{K1+K2+K3+\ ...\ +Kn}}$$

We now factor out the term 2^{K1} ,

$$C_n \to \frac{2^{K1}\,(3^{n-2} + 3^{n-3} * 2^{K2} + 3^{n-4} * 2^{K2+K3} + \ ...\ +3^2 * 2^{K2+\ ...\ Kn-3} + 3 * 2^{K2+\ ...\ Kn-2} + 2^{K2+K3+\ ...\ +Kn-1})}{2^{K1+K2+K3+\ ...\ +Kn}}$$

Cancelling the 2^{K1} terms from both the numerator and the denominator,

$$C_n \to \frac{(3^{n-2} + 3^{n-3} * 2^{K2} + 3^{n-4} * 2^{K2+K3} + \ ...\ +3^2 * 2^{K2+\ ...\ Kn-3} + 3 * 2^{K2+\ ...\ Kn-2} + 2^{K2+K3+\ ...\ +Kn-1})}{2^{K2+K3+\ ...\ +Kn}}$$

This can be re-written as:

$$C_n \;\longrightarrow\; \frac{(3^{n-2}\,)}{2^{K2+K3+\,...\,+Kn}} + \frac{(3^{n-3} * 2^{K2} \;+ 3^{n-4} * 2^{K2+K3} + ... \;+3^2 * 2^{K2+\,...\,Kn-3} \;+3 * 2^{K2+\,...\,Kn-2} + 2^{K2+K3+\,...\,+Kn-1}\,)}{2^{K2+K3+\,...\,+Kn}}$$

Again, as *n* and K2+K3+ … +Kn approach infinity, the first term diminishes to zero:

$$\frac{(3^{n-2})}{2^{K2+K3+\,...\,+Kn}} \;\longrightarrow\; 0 \qquad (Note\ that\ this\ follows\ from\ \frac{3^n}{2^{K1+K2+K3+\,...\,+Kn}} \;\longrightarrow\; 0 \qquad)$$

Therefore:

$$C_n \;\longrightarrow\; \cancelto{0}{\frac{(3^{n-2})}{2^{K2+K3+\,...\,+Kn}}} + \frac{(3^{n-3} * 2^{K2} \;+ 3^{n-4} * 2^{K2+K3} + ... \;+3^2 * 2^{K2+\,...\,Kn-3} \;+3 * 2^{K2+\,...\,Kn-2} + 2^{K2+K3+\,...\,+Kn-1}\,)}{2^{K2+K3+\,...\,+Kn}}$$

$$C_n \;\longrightarrow\; \frac{(3^{n-3} * 2^{K2} \;+ 3^{n-4} * 2^{K2+K3} + ... \;+3^2 * 2^{K2+\,...\,Kn-3} \;+3 * 2^{K2+\,...\,Kn-2} + 2^{K2+K3+\,...\,+Kn-1}\,)}{2^{K2+K3+\,...\,+Kn}}$$

Again, factoring out 2^{K2} ,

$$C_n \;\longrightarrow\; \frac{2^{K2}(3^{n-3} \;+ 3^{n-4} * 2^{K3} + 3^{n-5} * 2^{K3+K4} \;+ ... \;+3^2 * 2^{K3+\,...\,Kn-3} \;+3 * 2^{K3+\,...\,Kn-2} + 2^{K3+K4+\,...\,+Kn-1}\,)}{2^{K2+K3+\,...\,+Kn}}$$

and cancelling the 2^{K2} terms from both the numerator and the denominator,

$$C_n \;\longrightarrow\; \frac{(3^{n-3} \;+ 3^{n-4} * 2^{K3} + 3^{n-5} * 2^{K3+K4} \;+ ... \;+3^2 * 2^{K3+\,...\,Kn-3} \;+3 * 2^{K3+\,...\,Kn-2} + 2^{K3+K4+\,...\,+Kn-1}\,)}{2^{K3+\,...\,+Kn}}$$

Again, this can be re-written as:

$$C_n \longrightarrow \frac{3^{n-3}}{2^{K3+\,...\,+Kn}} + \frac{(3^{n-4} * 2^{K3} + 3^{n-5} * 2^{K3+K4} +... +3^2 * 2^{K3+\,...\,Kn-3} +3 * 2^{K3+\,...\,Kn-2} + 2^{K3+K4+\,...\,+Kn-1})}{2^{K3+\,...\,+Kn}}$$

Again applying our assumption about convergence to zero at infinity, that is as *n* and K2+K3+ … +Kn approach infinity, the first term diminishes to zero.

$$C_n \longrightarrow \cancelto{0}{\frac{3^{n-3}}{2^{K3+\,...\,+Kn}}} + \frac{(3^{n-4} * 2^{K3} + 3^{n-5} * 2^{K3+K4} +... +3^2 * 2^{K3+\,...\,Kn-3} +3 * 2^{K3+\,...\,Kn-2} + 2^{K3+K4+\,...\,+Kn-1})}{2^{K3+\,...\,+Kn}}$$

$$C_n \longrightarrow \frac{(3^{n-4} * 2^{K3} + 3^{n-5} * 2^{K3+K4} +... +3^2 * 2^{K3+\,...\,Kn-3} +3 * 2^{K3+\,...\,Kn-2} + 2^{K3+K4+\,...\,+Kn-1})}{2^{K3+\,...\,+Kn}}$$

Next we factor out 2^{K3} and repeat the above procedure, then factor out 2^{K4} , and so on.

It can be shown that eventually we get:

$$C_n \longrightarrow \frac{3}{2^{Kn-1+Kn}} + \frac{2^{Kn-1}}{2^{Kn-1+Kn}}$$

$$C_n \longrightarrow \frac{3}{2^{Kn-1+Kn}} + \frac{1}{2^{Kn}}$$

$$C_n \longrightarrow \frac{1}{2^{Kn}}\left(\frac{3}{2^{Kn-1}} + 1\right) = \frac{\frac{3}{2^{Kn-1}} + 1}{2^{Kn}}$$

Now, since C_n can only be a whole number, then the term:

$$\left(\frac{3}{2^{Kn-1}} + 1\right)$$

must be some factor of 2. The only possible value of 2^{Kn-1} is 1. That is:

$$2^{Kn-1} = 1$$

Therefore,

$$\left(\frac{3}{2^{Kn-1}} + 1 \right) = \frac{3}{1} + 1 = 4$$

Therefore:

$$C_n \rightarrow \frac{\frac{3}{2^{Kn-1}} + 1}{2^{Kn}} = \frac{4}{2^{Kn}}$$

Again, since C_n can only be a whole number, only three values are possible for C_n .

$$2^{Kn} = 1 \implies C_n = 4$$

$$2^{Kn} = 2 \implies C_n = 2$$

$$2^{Kn} = 4 \implies C_n = 1$$

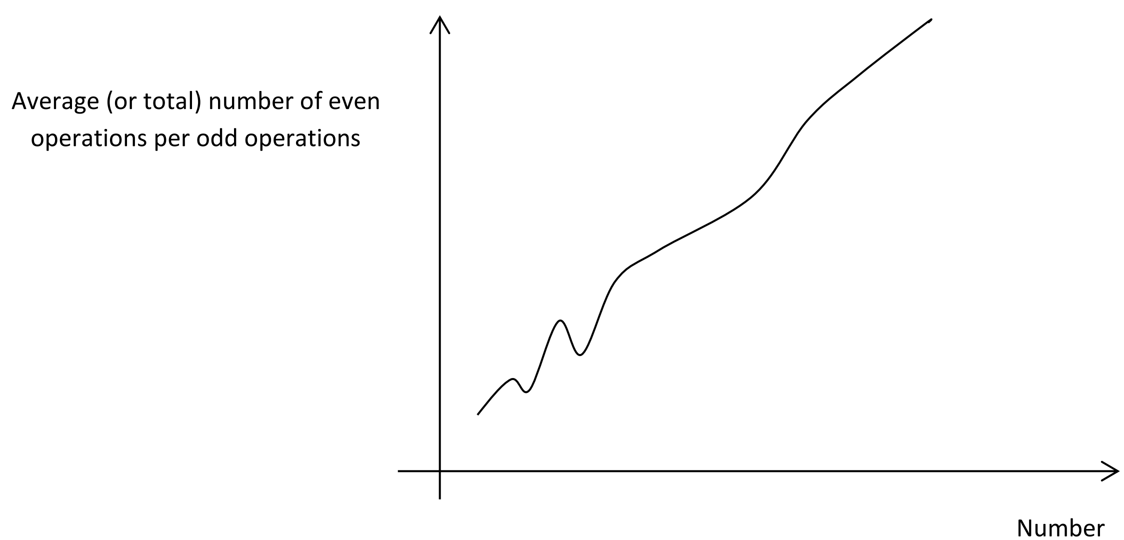
Thus we have proved that (almost?) all Collatz sequences eventually converge to the 1-4-2-1 loop.

Proof of convergence to zero at infinity

So far, we have reduced the problem of the Collatz conjecture to a problem of proving that those terms assumed as converging to zero at infinity indeed converge to zero. Next, I present the proof of convergence to zero of the terms we have assumed as converging to zero at infinity, and perhaps the rigor of the proof could be the only possible one with regard to the Collatz conjecture.

We start from the fact that the total number of even Collatz operations in a given Collatz sequence is significantly greater than the number of total odd Collatz operations, and this effect indefinitely increases for infinitely large numbers and infinitely long sequences. That is, if one starts with a very, very large number, the total number of even operations will be very large compared to the total number of odd operations, resulting in steep descent of the sequence. In fact, the steepness of the general descent indefinitely increases as the starting value approaches infinity. The steepness of the general descent will gradually decrease as the sequence descents towards finitely large numbers and the descent of the sequence becomes less and less pronounced at relatively small numbers, and more chaotic at smaller numbers, but eventually reaching 1 in all cases.

Why does the ratio of the total number of even operations to the number of odd operations, or the average number of even operations following each odd operation, would increase vastly if the sequence diverged to infinity? The reason for this is that at infinitely large numbers, unlike at smaller numbers, even integers can have very large powers of two as a factor, say 2^{10} , 2^{100} , 2^{1000} etc. Thus, since every Collatz operation on an odd number results in an even number, once the sequence lands on an even number it may go through tens or hundreds or thousands of successive even operations before ending in an odd number, that is before another odd operation occurs again, vastly dropping the sequence. There can never be successive odd operations, but there can always be successive even operations. Therefore, although every odd operation results in a term greater than the term before it, that operation always results in even operation(s), which may result in vast descent of the sequence, much more than its ascent by the single odd operation. In other words, at relatively small numbers, an even number can have 2, 2^2 , 2^3 , 2^4 , etc. as a factor and at very large numbers this could be all the way from 2, 2^2 , . . . to 2^{100} and still at immensely large numbers 2, 2^2 , . . . to 2^{10000} and so on , following the same trend indefinitely. Therefore, as the starting odd number increases, the average and total ratio of even to odd operations also increases accordingly. Thus, after just one odd operation, there could be thousands or billions of successive even operations before the next odd operation occurs, always resulting in a steep average descent of the sequence. As the length of the sequence increases towards infinity, the ratio between the total number of odd operations and the total number of even operations continuously tends towards zero. This results from the fact that as the starting value increases more and more even operations occur *per every odd operation*.



Note that the above graph is qualitative, meant only to explain the concept.

Proof of Assumption 2

From Assumption 1 :

$$\frac{n}{K1 + K2 + K3 + \ldots + Kn} \longrightarrow 0 \quad \text{at infinity}$$

Let

$$\frac{n}{K1 + K2 + K3 + \ldots + Kn} = r$$

$$\Longrightarrow \; n = r \; (K1 + K2 + K3 + \cdots + Kn)$$

Therefore,

$$\frac{3^n}{2^{K1+K2+K3+\ .\ .\ .\ +Kn}} = \frac{3^{r\,(K1+K2+K3+\cdots+Kn)}}{2^{K1+K2+K3+\ .\ .\ .\ +Kn}} = (\frac{3^r}{2})^{(K1+K2+K3+\cdots+Kn)}$$

Now since *r* approaches 0 at infinity (Assumption 1), *r* = 0 at infinity:

$$(\frac{3^r}{2})^{(K1+K2+K3+\cdots+Kn)} = (\frac{3^0}{2})^{(K1+K2+K3+\cdots+Kn)} = (\frac{1}{2})^{(K1+K2+K3+\cdots+Kn)} = 0 \;, \quad at\ infinity\,, that\ is\ for\ K1 + K2 + K3 + \ .\ .\ . + Kn \rightarrow \infty$$

Proof of Assumption 1

Lemma 1

Let n_1 and n_2 be any two even positive integers, with $n_2 > n_1$. Let $n_1 = m_1 \times 2^{k1}$ and $n_2 = m_2 \times 2^{k2}$, where m_1 and m_2 are odd numbers. The probability that $k_2 > k_1$ is always greater than the probability that $k_1 > k_2$. In other words, the probability that a larger even number has a given power (k) of two is greater than the probability that a smaller even number has the same power of two.

Lemma 2 follows from the fact that the probability that an even number $n = m * 2^k$ having greater k increases as n increases. This follows from the fact that larger numbers can have higher factors of 2 than smaller numbers. For example, even numbers up to 100 can have 2^1 , 2^2 , 2^3 , 2^4 , 2^5 , 2^6 as factors, whereas even numbers up to 1000 can have 2^1 , 2^2 , 2^3 , 2^4 , 2^5 , 2^6 , 2^7 , 2^8 , 2^9 , as factors, and even numbers up to 1,000,000 can have 2^1 , 2^2 , 2^3 , 2^4 , 2^5 , 2^6 , 2^7 , 2^8 , 2^9 , 2^{10} , 2^{11} , 2^{12} , 2^{13} , 2^{14} , 2^{15} , 2^{16} , 2^{17} , 2^{18} , 2^{19} and even numbers up to 1,000,000,000 can have 2^1 , 2^2 , 2^3 , 2^4 , 2^5 , 2^6 , 2^7 , 2^8 , 2^9 , 2^{10} , 2^{11} , 2^{12} , 2^{13} , 2^{14} , 2^{15} , 2^{16} , 2^{17} , 2^{18} , 2^{19} , 2^{20} , 2^{21} , 2^{22} , 2^{23} , 2^{24} , 2^{25} , 2^{26} , 2^{27} , 2^{28} , 2^{29} and so on.

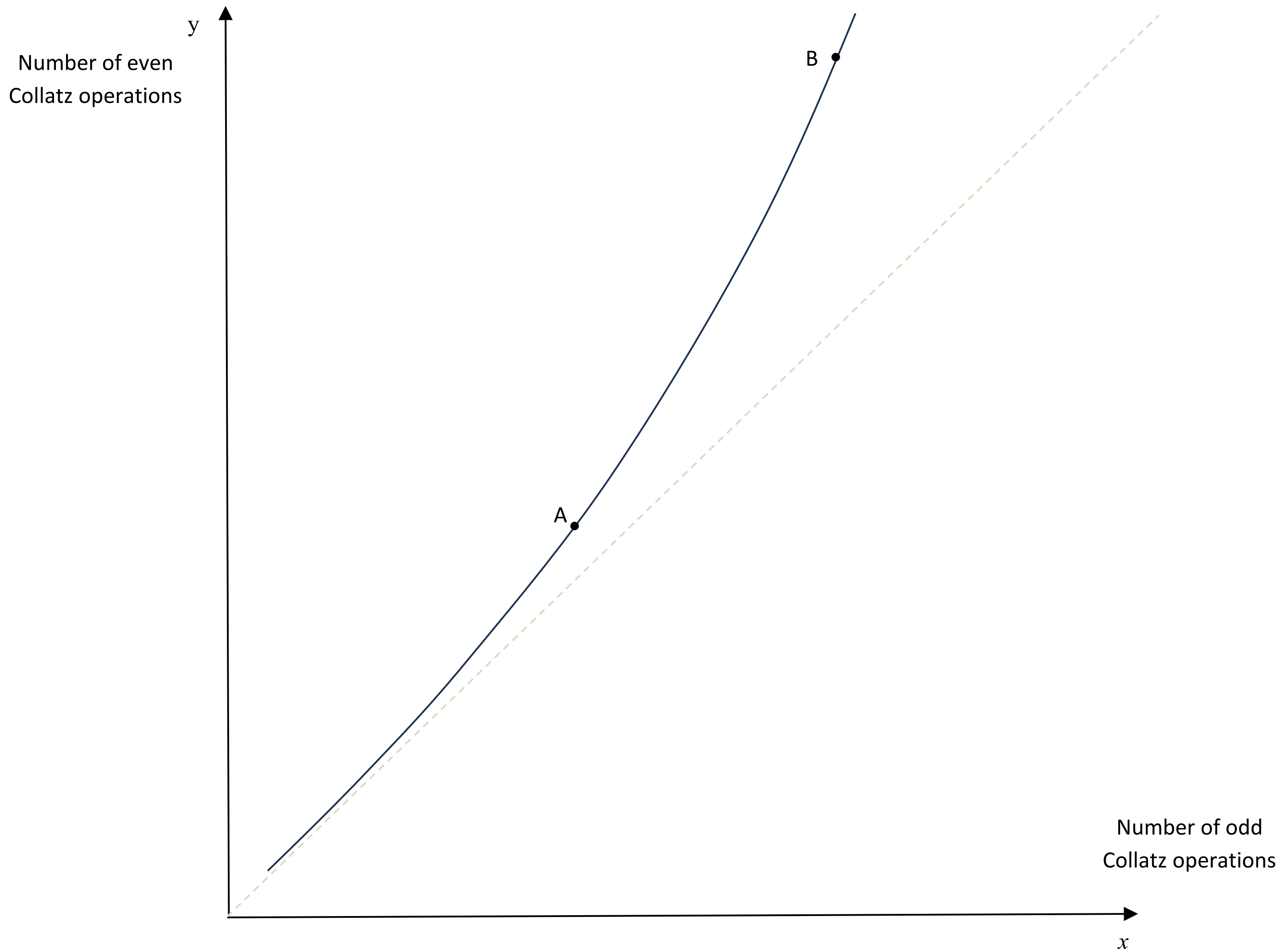
This clearly proves that the probability that a bigger number has a given powers of 2 is greater than the probability that a smaller number has the same power of 2. It follows that the number of even operations after every odd operation increases with increase of the starting value, and increases towards infinity as the starting number goes to infinity.

The next graph illustrates this effect. We can see that the slope of the graph continues to increase as the number of Collatz operations increases. We can see that not only the local slope *dy/dx*, but also the global slope *y/x* , increases with the number of operations.

We can see that the ratio of the total number of even operations to the total number of odd operations is greater at point B than at point A on the curve, indicating a continuous increase of this ratio with increase in the number of Collatz operations (or sequence length), the ratio approaching infinity for infinite sequence length and, conversely, the ratio of odd to even operations approaching zero.

$$\frac{Number\ of\ even\ Collatz\ operations}{Number\ of\ odd\ Collatz\ operations} \simeq \infty \quad (for\ infinite\ length\ of\ sequence)$$

$$\Longrightarrow \frac{Number\ of\ odd\ Collatz\ operations}{Number\ of\ even\ Collatz\ operations} \simeq 0 \quad (for\ infinite\ length\ of\ sequence)$$


y
Number of even
Collatz operations
B
A
Number of odd
Collatz operations
x

Now, there are two possibilities for the Collatz conjecture to be disproved.

1. The sequence indefinitely diverges to infinity.
2. The sequence may enter some closed loop (there might be more than one loop.

Since we have already proved that the Collatz sequence cannot diverge to infinity, next we consider the possibility of another closed loop other than the trivial loop.

Proof that no closed loops exist other than the 1-4-2-1 loop

Corollary 1

There are no closed loops of the Collatz sequence other than the trivial loop of 1-4-2-1.

We propose a property of a closed loop as follows: *any closed loop is decoupled from the sequences and the numbers in the sequences leading to it*. Since multiple Collatz sequences lead to the closed loop, the closed loop must be independent from any sequence (and its terms) leading to that closed loop. It necessarily follows that the coefficient of N vanishes at infinity, in the following equation:

$$C_n = \frac{3^n N + 3^{n-1}}{2^{K1+K2+K3+\,...\,+Kn}} + \frac{3^{n-2} * 2^{K1} + \,...\, +3^2 * 2^{K1+K2+\,...\,Kn-3} + 3 * 2^{K1+K2+\,...\,Kn-2} + 2^{K1+K2+K3+\,...\,+Kn-1}}{2^{K1+K2+K3+\,...\,+Kn}}$$

Note that the coefficient of N here is:

$$\frac{3^n}{2^{K1+K2+K3+\,...\,+Kn}}$$

which in turn brings us back to our previous assumption about the convergence of terms to zero at infinity, which in turn leads to the 1-4-2-1 loop! So we began with an assumption that another closed loop could exist somewhere in the darkness of infinity, which demanded the vanishing of the coefficient of N, which led back to the only known closed loop of 1-4-2-1. This should be a complete, rigorous mathematical proof (by induction) that no other closed loop exists.

In other words, assume that a closed loop exists at some immensely large number. Any closed loop requires that the coefficient of N must diminish to zero in the limit. Assume that the sequence was not decoupled from the starting number even at infinity, that is for infinite length of the sequence. By contradiction, there cannot be such a closed loop because C_n cannot be shown to converge to constant numbers (possible terms in the supposed closed loop), in the same way that we have proved that every Collatz sequence converges to the number 4, 2 or 1. In summary, any closed loop requires the coefficient of N to diminish to zero at infinity, but then this leads to the 1-4-2-1 loop. This proves that no other closed loop can exist.

Arbitrarily Close to One Probability Vs Absolute Certainty

Our analysis and arguments have shown that the probability of a Collatz sequence converging to the number 1 can be shown to be arbitrarily close to *one*. Conversely, the probability of a Collatz sequence diverging to infinity can be shown to be arbitrarily close to *zero*. But one can argue that, irrespective of all the arguments and analysis made so far, these arguments and analysis do not rule out at least one sequence diverging to infinity, out of infinitely many possible sequences.

The question is: if all the purpose of a paper on a proof of the Collatz conjecture is to prove that *all* (not almost all) Collatz sequences converge to the number 1, then what have we achieved in this paper? Is it only a partial result? In this paper, we have proved that (at least) *almost all* Collatz sequences converge to the number 1. We raise the following fundamental questions:

1. If the analysis presented proves that *almost* all Collatz sequences converge, then what is the proof that *all* sequences converge?
2. Do we need a different, unrelated proof for the latter, unrelated to the number theories and Collatz sequence behaviors discussed in this paper?
3. Or, does such a proof even exist?
4. Or, could the ideas and analysis introduced in this paper lead to a complete proof of the conjecture?

We strongly feel that the analysis presented in this paper should also be *part of* the complete proof. If this is really the case, then perhaps there might be some missing link.

Principle of Discrete Probability – Probability 'Quanta'

We hereby introduce/propose a new principle of probability as follows.

If the probability of an event is arbitrarily close to zero, then the probability of that event is (exactly) zero. That is, that event will never happen, with absolute certainty. Likewise, if the probability of an event is arbitrarily close to one, then the probability of that event is (exactly) one. That is, that event will happen, with absolute certainty. Probability only changes in discrete steps: 'quanta'. The probability of any event is an integral multiple of the probability 'quanta', which is the smallest possible non-zero probability.

From our analysis, we have seen that the probability of a Collatz conjecture diverging to infinity can be shown to be arbitrarily close to zero, by assuming infinite sequence length. We have given the explanation for this: the ratio between the total number of odd Collatz operations to the total number of even operations approaches zero, for an infinite sequence length. Therefore, if the Principle of Discrete Probability is true, then we can say with absolute certainty that a Collatz sequence diverging to infinity does not exist.

Unlike conventional probability that is continuous and has infinite different values between 0 and 1, discrete probability has finite number of values between 0 and 1. Unlike conventional probability in which probabilities can arbitrarily approach zero, discrete probability cannot arbitrarily approach zero without eventually becoming (exactly) *zero*, and does not arbitrarily approach one without eventually becoming (exactly) *one*. In discrete probability, *zero* probability of an event means absolute certainty that the event will not occur and probability of *one* means absolute certainty of the event occurring. This is unlike conventional probability in which an event can have a probability arbitrarily close to zero and yet the occurrence of that event cannot be completely ruled out with absolute certainty, which creates conceptual problems.

From our discussions so far, we have shown that the (conventional) probability of an infinite length Collatz sequence is arbitrarily close to zero and yet this has not led to a complete proof of the Collatz conjecture because one encounters a conceptual problem regarding whether arbitrarily close to *zero* means exactly zero or not . Discrete probability principle states that since the probability of a diverging Collatz sequence continuously decreases with increase in the length of the sequence, this probability eventually becomes lower than the minimum possible probability ('probability quanta') and eventually become (exactly) *zero*. Then we can say that a Collatz sequence diverging to infinity does not exist, and we say this with absolute certainty. The only way such a divergent sequence could exist was if our analysis indicated that the probability would finally settle at some finite, non-zero value.

This idea of probability quanta presented here is not entirely pure speculation. In one of my papers, I have proposed that every physical quantity in the universe is discrete and changes in discrete steps (quanta). This theory has been developed based on Zeno's paradoxes. Probabilities decrease (or increase) in discrete steps and cannot decrease (or increase) indefinitely without eventually becoming *zero* (or *one*).

Conclusion

In this paper, we have been able to prove the Collatz sequence, by studying the property of the sequence of infinite length, thereby avoiding the chaos of the finite length sequences. We noted that at infinite lengths, the ratio of odd operations to even operations diminishes, approaching zero. This leads to the eventual decoupling of the sequence from the starting value, ending in the only possible path of all sequences: the 1-4-2-1 loop. We have also introduced a new Principle of Discrete Probability to close the conceptual gap that exists in conventional probability, namely the distinction between a probability arbitrarily approaching *zero* (or *one*) and a probability of exactly *zero* (or *one*).

Glory Be To Almighty God Jesus Christ and His Mother, Our Lady Saint Virgin Mary

References

1. L.Collatz, *On the behavior of sequences defined by simple arithmetic operations*, 1937

2. J.C, Lagarias, *The 3x+1 problem and its generalizations*, American Mathematical Monthly, vol.92, no.1,pp.3-23,1985

3. Jeffrey C.Lagarias (Ed.), *The ultimate challenge : The 3x+1 Problem*, AMS,2010