White Paper
A Computational Approach to Collatz Conjecture

by

Shailesh Kadakia

August 25, 2025

Abstract

In Mathematics, Collatz Conjecture is concerned with a peculiar shrinking number pattern 4, 2, and 1, end result after specified instructions are performed. The stream of numbers, a sequence is generated as follows. You begin with a positive integer n. Next, you perform set of operations depending on the selected number being even or odd. First, if the number is even, you divide it by 2 repeatedly till you get an odd number. Second, if the number is odd, you multiply it by three and add one. Since three times any odd number plus one is always even, you divide it by 2. You repeat first and second operations indefinitely. The conjecture states that regardless of value of n, the end sequence of numbers generated after several executions of first and second step is always 4, 2, and 1. For instance, starting with odd number n = 13, you get sequence 40, 20, 10, 5, 16, 8, 4, 2, and 1. Also, for starting with an even number n = 36, you get sequence 18, 9, 28, 14, 7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, and 1. In this white paper, author applies pattern analysis technique to prove the conjecture.

A Computational Approach to Collatz Conjecture

Shailesh Kadakia
August 25, 2025

Abstract

In Mathematics, Collatz Conjecture is concerned with a peculiar shrinking number pattern 4, 2, and 1, end result after specified instructions are performed. The stream of numbers, a sequence is generated as follows. You begin with a positive integer n. Next, you perform set of operations depending on the selected number being even or odd. First, if the number is even, you divide it by 2 repeatedly till you get an odd number. Second, if the number is odd, you multiply it by three and add one. Since three times any odd number plus one is always even, you divide it by 2. You repeat first and second operations indefinitely. The conjecture states that regardless of value of n, the end sequence of numbers generated after several executions of first and second step is always 4, 2, and 1. For instance, starting with odd number n = 13, you get sequence 40, 20, 10, 5, 16, 8, 4, 2, and 1. Also, for starting with an even number n = 36, you get sequence 18, 9, 28, 14, 7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, and 1. In this white paper, author applies pattern analysis technique to prove the conjecture.

1. Introduction

For extremely humongous size numbers such as 2^71 = 2305,8430,0921,3693,952 decimal, validation that the number converges under Collatz rules is a challenging and formidable task. Even for the most powerful computer system it takes several days of computation time and investment of manpower/funding the effort. As of 2025, it is verified that all numbers up to 2^71 have converged to 4, 2, 1 sequence at the conclusion of Collatz rules. We will refer to 4, 2, 1 sequence as Collatz sequence in our discussion. It would be of interest to see how the number of steps grow as compared to size of numbers that have the most steps before Collatz sequence is generated. In Table 1, we have summarized number size as a multiple of powers of tens vs. number of steps. It is easy to see that even a Python program to perform execution and validation for Collatz sequence is computationally exhaustive problem. Further, according to information in Table 1, Collatz conjecture is verified for numbers of size a trillion, it is lot lower than verification of Goldbach Conjecture Proof. It is reported that Goldbach Conjecture is proven for number of size 4 x 10^18. Therefore, it is imperative that we develop a generalized proof of the conjecture for unlimited size of starting number odd or even.

2. Preliminaries

The Collatz conjecture specifies iterations of integer odd or even under application of a simple function P(n). The function states: "Starting from any positive integer n, repeated operation sequence specified in P(n) will end in number 1. Thereafter, the number sequence will loop through 1, 4, 2, 1…endlessly. [Lagarias, 2010].

Let us develop a description of Collatz function P(n) notation. To ensure the process, we define basic terms odd and even integers.

A positive integer n is characterized as an even integer E when n mod (2) = 0. Further, an even integer n implies that E = 2M, M is a positive integer greater than 0.
A positive integer n is categorized as an odd integer O when n mod (2) = 1. With the same token as an even integer, for odd values of n, O = 2M + 1.

With this nomenclature for odd and even integers in mind, we shall formally define the Collatz function P(n)

$$P(n) = \begin{cases} n/2 \text{ if n of type E} \\ (3n + 1) \text{ if n is of type O} \end{cases}$$

The Collatz function is named after mathematician Lohar Collatz in 1937. Though, in recent times Collatz conjecture is verified using modern Computers with vast amount memory and the fastest processing speeds, very little progress has been made for proving the conjecture. Several proofs are reported to date but none of them are accepted by mathematics community at large [Marshall 2022]. This inspired the author to develop a convincing proof for Collatz conjecture. First, we will prove the conjecture by heuristic arguments. Then we will describe formal proof in its entirety. For sake of convenience, we will refer P(n) operations: division of even number by 2 as step a and multiplication of an odd number by 3 and subsequent addition of 1 as step b. Further, we will partition our proof by classifying even and odd number types into subcategories: 1). Numbers those are perfect powers of 2, 3, 5, and other primes such as 2^a, 3^b, etc. and 2). Numbers those are multiples of odd and even primes.

First, we will explain why shrinking number sequence 4, 2, and 1 occurs when numbers are decaying. An obvious reason is if at any time during execution of step a or b if we run into a number that is perfect power of two, we will terminate our sequence to 1 by repeated application of step a, divide by two. In neither step a or b, we are dividing by number other than 2. Therefore, even if our starting number is perfect power of other numbers such as 3^b, 5^c, 7^d, and etc., we can not develop sequence such as 27, 9, 3, 1 or 125, 25, 5, and 1 or 343, 49, 7, and 1. Next, we will explain why number sequence is always converges in a more general case of starting number even or odd.

Let us consider a case where starting number is even E = 2N, N is an integer. Repeated execution step a will always produce an odd number O. In the worst case, O = 2M + 1 after first iteration. Next step will be always b, where we multiply the number by three and add one. Notice that after step b, the result will be always an even number. Therefore, we are forced to execute step a. Examination of execution sequence indicates that we will cycle through atleast one even and odd pair of numbers. Our result will be close to a number 2N ★ ¾ because division by 2 occurs twice where as multiplication by 3 occurs once.

Next, we consider a case where the starting number is odd O = (2M + 1). This time, we execution step b first, where we multiply the number by three and add one. However, this will always produce an even number. The reason is when you multiply two odd numbers, 3 and starting number, you get an odd number. After you add 1 in the step you get even number, whose value is close to (2M + 1) ★ 3. Next, we divide the sum by 2 in step a. we get a number value close to (2M + 1) ★ 3/2. This number could be even or odd that has equal probability. If the number is odd in the worst case, the number will grow temporarily but after a cycle or two it will decay because it will be even after step b. Next, for even number, we execute step a. Examination of execution sequence indicates that we will cycle through atleast one odd and even pair of numbers. In that case the result will be close to a number $(2M + 1) \star \frac{3}{4}$.

In summary, on the average, the number will shrink after k iterations of odd and even sequences by a factor of $(3/4)^k < 1, k > 1$. This conclusion does not depend on type of starting number even or odd. In general, starting number E or O is divided by two more frequently while multiplied by three sparingly. Further, when you apply step b, some of the even numbers happens to be multiples of 4, 8, 16, 32, and etc. Therefore, indefinite growth of number is inhibited, leading to collapsing sequence of numbers. An obvious reason is $2^k \geq 3$ for $k > 1$.

To prove 2^{nd} part of conjecture statement we inspect the range of values associated with given number E=(2N) even or given number O = (2M + 1) odd while we execute steps a and b. It is obvious that there are numbers S that are exact multiples of whole powers of number 2 which are all even. Also, there are numbers T that are exact multiples of whole powers of 3 which are odd numbers. However, powers of 3 are sparser than powers of 2, which is a smaller base than base 3. Therefore, we can always find a number S that is larger than number O odd or E even such that S would converge into 16, 8, 4, 2, 1 series for any arbitrary large size number E or O. Our argument is if conjecture holds good for S it should hold for E and O because S > E and O both. Furthermore, S is undergoing decaying sequence of numbers upon repetitive execution of operations stated in steps a.

Next, we will describe formal proof of the conjecture. In order to improve reading, we have partitioned our proof in three Lemmas. First, lemma proves that for any odd or even starting number recursive application of P(n) will not cause a sequence that repeats (loop) the same number again and again. In second lemma, we prove that growing sequence after application of P(n) function is impossible. In the third lemma, we prove that only sequence 4, 2, and 1 is possible. Under no condition other decreasing sequence such as 27, 9, 3, and 1 or any other multiples of prime in decreasing order is possible.

3. Formal Proof

First, we will show that for starting number even E and odd O, Collatz function never enters into a loop indefinitely. Also, it can never produce sequence of numbers such as 2, 2, 2, … or 3, 3, 3…, or n, n, n, … for any value of even number E or odd number O.

Lemma 1. Indefinite repeat number sequences are not possible.

If starting number even, after step a $N = E/2$.
However, $E/2 = E$ will result in value $E = 0$, not possible because E is positive integer $E \geq 1$
If starting number is odd, after step b, $N = 3 \bullet O + 1$
However, $3 \bullet O + 1 = O$ will result in value of $O = -1/2$ not possible because O is positive integer $O \geq 1$

Next, we will explain, why at the end of several passes of Collatz function for an odd or an even starting number always results in a diminishing number sequence 4, 2, and 1.

Lemma 2. Only decaying sequences are possible.

For even number E is reduced to odd number after k successive divisions by number 2.

Step a. E reduced by a factor 2, 4, 8, 16 etc. $k \geq 1$ definition of even number.

$E \div 2^k = (2M + 1) = \text{Odd number } O = E \bullet (1/2)^k,\ k > 0$

Every even number E is transformed to odd number O after step a. Next execute b.

Step b. Multiply O by 3 and add 1 New even = 3 ★ O + 1

New Even = 3 ★ (2M + 1) + 1 = 6M + 4 always even = 2 (3M + 2)

Next, execute step a divide by 2 we get 2(3M + 2)/2 = 3M + 2 and iterate the function P(n).

First time after step b when you execute a, notice that $(3M + 2) \div 3 = M + 2/3$ which is not a whole number. Therefore, 3M + 2 is not divisible by 3. This is the main reason that decaying sequence of numbers 81, 27, 9, 3, and 1 is not possible after step b. Based on results of multiple execution of steps a and b we can conclude that starting even number E is reduced by a factor of $3^l/2^{(l+k)} < 1$ for $k > 0$ Therefore, end result after successive execution of P(n) for an even starting number will be $S = E \bullet (3/2)^l \bullet (1/2)^k = E \bullet 3^l/2^{(l+k)}$

We can perform similar analysis for the case where we start with odd number O. In that case we will perform step b first and after l cycles of growth we will decay for k cycles performing step a. Again, the final number we get is $S = O \bullet (3/2)^l \bullet (1/2)^k = O \bullet 3^l/2^{(l+k)}$. The starting number O is again reduced by factor $3^l/2^{(l+k)} < 1$ for $k > 0$ Thus, we have discovered that for starting number even E or odd O, the number will shrink by the same factor $3^l/2^{(l+k)} < 1$. The fact that both cases, starting number odd or even results in reduction of the number by the same factor, it is possible that sequence will terminate in form 4, 2, and 1. A rigorous proof exist [Li 2022] which shows that upon recursive application of P(n) the number sequence terminates in form 4, 2, and 1 by method of induction. However, the proof does not show why other sequences are not possible such as 27, 9, 3, and 1 or 343, 49, 7, and 1.

For the sake of completeness, further examination of decaying sequence created by specific odd or even starting numbers during execution of P(n) steps is required. Specifically, we will focus, recursive execution of step b in P(n) for the case of an odd starting number.

Lemma 3. Only 4, 2, and 1 sequence is possible upon recursion of P(n).

Let us consider a case starting number is perfect power, two raised to an integer ≥ 4. In that case, we are forced to perform step a repeatedly. This results in shrinking sequence of numbers 8, 4, 2, and 1 naturally. However, this repeated execution of step b is not guaranteed in case starting number is odd whole power of number 3. In fact, decaying sequence 27, 9, 3, and 1 can never occur because we proved that 3M + 2 is not divisible by 3. It is interesting to note that 3M + 2 value could result in an odd number divisible by 5. This could result in number sequence such as 40, 20, 10, 5 in step a from step b. But 3 ● 5 + 1 = 16 from step b cause number sequence to converge 4, 2, and 1 from a.

Next, we will explore various possibilities for values of number 3M + 2 during execution of step b. If M is even, 3M + 2 has even value therefore next step will be a, making starting number small O ● $3^{l}/2^{(l+k)}$. If M is odd 3M + 2 will be odd that has factors greater than 3. In that case the starting number will be amplified by factor of 3/2 few times. Eventually, 3M + 2 will have a value that is multiple of ten. This value will lead to 20, 10, 5 values in step a and then value 16 from step b. Finally, the sequence will end into 8, 4, 2, and 1.

Further, if we examine numbers developed from sequence 3M + 2 by assigning integer values to M we find the following sequence 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35, 38, 41, 44, 47, 50, 53, 56, 59, 62, 65, 68, 71, 74, 77, 80 and so on. The sequence skips multiple of numbers two times for instance 11 is part of sequence but 22 and 33 are not. In fact, next multiple of member in sequence is multiplied by (3p + 1) and displacement of the next term is (3p + 2) , here p = 1, 2, 3, …The number 8 is part of sequence but 16 and 24 are not. This skip feature of 3M + 2 sequence inhibits decaying sequence such as 27, 9, 3, 1 and 125, 25, 5, 1. Further, the displacement term is higher distance away then (3n + 1) amplification value in step b of P(n). Mathematically, the skip in the pattern developed by step b can be observed by testing consecutive values of M.

For even values of M, (3M + 2) attains an even number directs us to step a. For odd values of M if we perform division by 3, 5, 7, 9, 11, 13, 15, 17 we find a sequence that does not loop for reasons stated below.

(3M + 2) ÷ 3 = X is never divisible by 3 for odd or even values of M
(3M + 2) ÷ 5 = X is divisible by 5 that leads to sequence 5 * 3 + 1 = 16 even 8, 4, 2, 1
(3M + 2) ÷ 5 = Y is divisible by 10 that leads to 80, 40, 20, 10, and 5.
(3M + 2) ÷ 7 = Z is never a whole number for two consecutive loops of step b.
● ● ●
● ● ●
● ● ●
● ● ●

(3M + 2) ÷ A = not a whole number for A > 5 for two consecutive loops of step b.

Moreover, a close examination of numbers created by 3M + 2 Arithmetic sequence reveals successive terms are relatively prime to each other. It means that ratio of next term in the sequence to present term is between one and two, not a whole number. i. e. $1 < (3M + 5) / (3M + 2) < 2$. In fact, the next whole multiple of (3M +2) appears in the sequence only after 3M + 2 terms. The skip in multiple of present term leaves gaps in continum of multiples of terms relative to its own position. This is the primary cause the sequence can not terminate with any other pattern except reduction in value either from multiples of two or multiples of five. Therefore, other converging sequence does not occur during repeated execution of step b. Thus, we have proved both parts of the conjecture by heuristic arguments.

4. Computational Considerations

The proof can easily be verified when we implement steps of Collatz Conjecture on a Computer which uses binary arithmetic number system. For instance, in Computer arithmetic division by 2 is achieved, shift right one place binary number and zero fill most significant bit. Also, multiplication by 3 is accomplished by shifting left given number once and adding the same number once to shifted result. By observing bit sequence of results one can conclude that requirement for proof of the conjecture is satisfied. Also, size of the number is extensible to any number of bits because one can store very large size numbers in an array and computer Arithmetic Logic Units (ALU) processes numbers in floating point decimal system according to IEEE Standard 754-1985.

5. Concluding Remarks

This white paper presented a simple and elegant proof of Collatz Conjecture by applying pattern analysis method and heuristic arguments. We will appreciate any constructive comments and suggestion from reviewing community. This will help us improve proofs for this conjecture as well as other problems related to Collatz conjecture.

Acknowledgements

I wish to thank Dr. Derek Muller at Veritasium for producing a video, an excellent review of Collatz conjecture title, “The Simplest Math Problem No One Can Solve”. His YouTube inspired me to work on this and other similar projects. Also, I wish to extend sincere thanks to my friend Nia Stefani, Managing Partner Xnergy Financial and her son West Stefani for supporting this effort. Also, I wish to thank my colleague at Mathnasium Mr. Marcus Stidham for interesting conversations and feedback on my Collatz Conjecture Proof.

References

[1] Lagarias, J. C. (2010). The Collatz conjecture. Chaos, 20(4), 041102. 1

[2] Lagarias, J. C. (1985). The 3x+1 problem and its generalizations. American Mathematical Monthly, 92, 3–23. 1169 5.

[3] Lagarias, J. C. (2004). The 3x+1 problem: Annotated bibliography (1963–1999). In de Gruyter Series in 1170 Nonlinear Analysis and Applications 6 (pp. 189–299). Berlin: Walter de Gruyter. 1171 6.

[4] Marshall, S. (2022), Two Proofs of the Collatz Conjecture.

[5] Lei Li, Makoto Takahashi, A Simple Verification and a New Estimation for L. Collatz Conjecture, Information, 6:5(2003), pp.519-524.

[6] Masahiro Ameya, Lei Li, On the Verification and the Estimation for. L. Collatz Conjecture, Information, 21:10-11(2018), pp.2289-2310.

[7] Lei Li, On Two Expansions of the Collatz Problem, Information, 25:2(2022),pp. 109-146.

[8] Lei Li, Analysis of the Average Number of Steps for the Collatz Problem, Information, 25:3(2022)

[9] L. Rade, R. D. Nelson, Solution of Mathematical Puzzle by Computer, Kaibundo Publishing, Tokyo, 1992.

[10] Interesting Kakutani’s Conjecture, Computer World, Beijing, 1985.

[11] Manuel V.P. Garcia and Fabio A. Tal, A Note on the Generalized 3n+1 Problem, Acta Arithmetica XC 3(1999), pp.245-250.

[12] Terence Tao, Almost, all orbits of the Collatz map attain almost bounded values, https://doi.org/10.48550/arXiv.1909.03562

[13] David Barina, Convergence Verification of the Collatz Problem, The Journal of Supercomputing, 77(2021),pp.2681-2688. https://doi.org/10.1007/s11227-020-03368-x
[14] Mercedes Orus-Lacort and Christophe Jouis, Analyzing the Collatz Conjecture Using the Mathematical Complete Induction Method, Mathematics,2022, 10, 1972. https://doi.org/10.3390/math10121972
[15] Ramon Carbo-Dorca, Boolean Hypercubes, Mersenne Numbers, and the Collatz Conjecture, Journal of Mathematical Sciences and Modelling, 3 (3) (2020), pp.120-129
[16] Jean-Paul A, T. Tao and the Syracuse conjecture, La Gazette des Mathematiciens, Number 168, April 2021.
[17] Michael R. Schwob, Peter Shiue, and Rama Venkat, Novel Theorems and Algorithms Relating to the Collatz Conjecture, International Journal of Mathematics and Mathematical Sciences, Volume 2021, Article ID 5754439, https://doi.org/10.1155/2021/5754439
[18] Robert Deloin, A Brand-new Approach to Collatz Conjecture, Theoretical Mathematics and Applications, Vol. 12, No. 1, 2022, pp.9-14.
[19] YouTube "The Simplest Math Problem No One Can Solve" from Dr. Derek Muller at Veritasium Link: https://www.youtube.com/watch?v=094y1Z2wpJg

TABLE 1: Number below Power of 10 vs. Convergence Steps

Power of 10	Actual Number	No. of steps
$10 = 10^1$	9	19
$100 = 10^2$	97	118
$1000 = 10^3$	871	178
$10000 = 10^4$	6171	261
$100000 = 10^5$	77031	350
$1000000 = 10^6$	837799	524
$10000000 = 10^7$	8400511	685
$100000000 = 10^8$	63728127	949
$1000000000 = 10^9$	63728127	986
$10000000000= 10^{10}$	9780657630	1132
$100000000000= 10^{11}$	75128138247	1228
$1000000000000= 10^{12}$	989345275647	1348

Source: Collatz Conjecture, wikipedia
Last Number A284668 in the OEIS