INTERNATIONAL JOURNAL OF PURE SCIENCE RESEARCH STUDIES (IJPSRS)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3139-8995
EDUSCHOOL ACADEMIC RESEARCH & PUBLISHERS



Total Stopping Times of the Collatz Map up to 10^7: A Computational Test of the Random-Walk Model

Lejo J Manavalan

Assistant Professor and Research Guide, Department of Mathematics, Little Flower College Guruvayur, Kerala, India

Article Information

Received: 22nd July 2026
Received in revised form: 25th August 2026
Accepted: 30th September 2026
Available online: 5th November 2026

Volume: 2
Issue: 4
DOI: https://doi.org/10.63090/IJPSRS/3139.8995.0026

Abstract

The Collatz or 3x + 1 conjecture asserts that repeated application of the map $T(n) = \frac{n}{2}$ for even n and $T(n) = \frac{(3n+1)}{2}$ for odd n eventually reaches 1 from every positive integer. Although unproven, it is supported by a heuristic random-walk model which predicts that the total stopping time σ(n), the number of iterations needed to reach 1, has mean approximately 6.952 ln n and a standard deviation proportional to √(ln n). This study computes σ(n) exactly for every integer up to 10^7 and tests these predictions. Trajectories were computed in C using memoisation and validated against direct iteration and known records. The longest trajectory below 10^7 begins at 8,400,511 and takes 429 steps of T (685 steps of the standard map), and the highest point reached, 3.02×10^{13}, occurs in the trajectory of 6,631,675. A least-squares fit of binned mean stopping times against ln n over 10^3 to 10^7 gives a slope of 6.817 ($R^2 = 0.9996$), within 2 per cent of the predicted 6.952. The variance of σ(n) grows in proportion to ln n, with a coefficient of 91 to 97 compared with the predicted 101.4. The standardised stopping times are close to Gaussian in the centre, with 68.7 per cent within one standard deviation, but are positively skewed, with skewness about 0.6 and a long right tail. The results support the random-walk model as a first approximation while identifying systematic departures in the variance and the tail.

Keywords: Collatz Conjecture, 3x + 1 Problem, Total Stopping Time, Random Walk, Stochastic Model, Experimental Mathematics, Dynamical Systems

1. INTRODUCTION

The Collatz problem, also known as the 3x + 1 problem, is one of the most famous unsolved problems in mathematics. Starting from a positive integer, one halves it if it is even and triples it and adds one if it is odd; the conjecture asserts that this process always eventually reaches 1.[1] Despite its elementary formulation, the problem has resisted proof, and it has become a test case for methods in number theory, dynamical systems and probability.[2]

Computational evidence for the conjecture is extensive: convergence has been verified for all starting values below 2^{68}.[3] On the theoretical side, Terras and Everett independently showed that almost all integers have a finite stopping time, in the sense that their trajectories eventually fall below their starting value,[4,5] and Tao proved that almost all orbits, in the sense of logarithmic density, attain almost bounded values.[6] A complete proof, however, remains out of reach.

An influential heuristic treats the parity sequence of a trajectory as if it were random, leading to a random-walk model for the logarithm of the iterates.[7,8] This model makes precise predictions about the statistics of the

total stopping time, but these predictions are not theorems. This study tests them against exact computation. Its objectives are:

- To compute the total stopping time of every integer up to 10^7;
- To identify extreme trajectories in this range;
- To compare the growth of the mean and variance of stopping times with the model's predictions; and
- To examine the shape of the distribution of standardised stopping times.

2. LITERATURE REVIEW

2.1. Definitions

Following common practice, we use the accelerated map[1,3]

$$T(n) = \frac{n}{2} \text{ if n is even;} \quad T(n) = \frac{3n+1}{2} \text{ if n is odd,} \tag{1}$$

which combines each odd step with the halving that necessarily follows it. The total stopping time of n is

$$\sigma(n) = \min \{ k \geq 0 : T^k(n) = 1 \}, \tag{2}$$

and the stopping time is the least k such that $T^k(n) < n$. The number of steps of the standard map, in which 3n + 1 and halving are counted separately, equals σ(n) plus the number of odd steps.

2.2. The Random-Walk Model

If the parity of successive iterates is treated as a sequence of independent fair coin tosses, each application of T multiplies the current value by approximately $\frac{3}{2}$ or $\frac{1}{2}$ with equal probability. The logarithm of the iterate then performs a random walk with increments $\ln\left(\frac{3}{2}\right)$ *and* $\ln\left(\frac{1}{2}\right)$, *whose mean is*

$$m = \frac{1}{2}\ln\left(\frac{3}{4}\right) \approx -0.14384 \tag{3}$$

and whose variance is $v = \frac{(\ln 3)^2}{4} \approx 0.30174$. The expected time for the walk to descend from ln n to 0 is $\frac{\ln n}{|m|}$, giving the prediction[8,7]

$$E[\sigma(n)] \approx \mu \ln n, \qquad \mu \frac{2}{\ln(4/3)} \approx 6.9521 \tag{4}$$

Standard results on first-passage times of random walks further imply that the variance of the passage time grows as $\frac{v \ln n}{|m|^3}$, so that

$$\text{var}[\sigma(n)] \approx \kappa \ln n, \quad \kappa = \frac{v}{|m|^3} \approx 101.39 \tag{5}$$

and that the standardised total stopping time is approximately normally distributed.[8] Crandall gave an early heuristic analysis of this kind,[9] and Lagarias and Weiss developed rigorous versions of stochastic models for the problem.[7] Kontorovich and Lagarias emphasised that these models generate conjectures rather than proofs, and that nothing has been rigorously proved about the average size of the total stopping time.[8] Systematic comparisons of the model's predictions for the mean, variance and distribution with exact data over a large range are therefore of interest.

3. METHODS

3.1. Computation

The total stopping time σ(n), the number of odd steps and the largest value attained were computed for every n from 1 to $N = 10^7$ in the C programming language using memoisation. For each n, the trajectory under T was followed only until it first fell below n, at which point the previously computed values for the smaller number were added. Because the stopping time is finite for every n in the range, this procedure yields σ(n) for all n in increasing order. Values were stored as 16-bit integers for stopping times and 64-bit integers for maxima, and the computation required less than one second. Statistical analysis was carried out in Python with NumPy.[10]

3.2. Validation

The memoised results were checked against direct iteration of T for n = 2, 3, 7, 27, 97, 871, 837,799, 8,400,511 and 9,999,999, with exact agreement. The trajectory of 27 requires 70 steps of T, or 111 steps of the

standard map, and reaches a maximum of 4,616 under T, corresponding to 9,232 under the standard map, in agreement with well-known values. The standard-map step counts for the record holders 837,799 (524 steps) and 8,400,511 (685 steps) also agree with established results.

3.3. Statistical Comparison

To estimate the growth rate of the mean, the range 10^3 to 10^7 was divided into 40 intervals of equal logarithmic width, and the mean of σ(n) in each interval was regressed on the mean of ln n by least squares. The variance was examined in five windows, $[10^3, 2 \times 10^3)$, $[10^4, 2 \times 10^4)$, $[10^5, 2 \times 10^5)$, $[10^6, 2 \times 10^6)$ and $[5 \times 10^6, 10^7)$, by computing Var σ(n)/E[ln n]. Finally, standardised stopping times

$$z = \frac{\sigma(n) - \mu \ln n}{\sqrt{\kappa \ln n}} \tag{6}$$

were computed for $5 \times 10^6 \leq n \leq 10^7$ and their distribution was compared with the standard normal.

4. RESULTS

4.1. Extreme Trajectories

Every integer up to 10^7 reaches 1, consistent with earlier verification.[3] Table 1 lists notable trajectories. The largest total stopping time below 10^7 is attained by n = 8,400,511, with 429 steps of T (685 of the standard map), and the largest below 10^6 by 837,799, with 329 steps of T (524 of the standard map). The highest value reached by any trajectory starting below 10^7 is 30,171,305,459,816 under T, attained from n = 6,631,675; under the standard map this corresponds to 60,342,610,919,632. The largest stopping time, that is, the longest time before first falling below the starting value, is 246 steps, attained by n = 8,088,063. There are 54 delay records, integers whose total stopping time exceeds that of every smaller integer, up to 10^7.

Table 1: Notable trajectories of the Collatz map for $n \leq 10^7$

n	σ(n), steps of T	Odd steps	Standard-map steps	Stopping time	Maximum under T
27	70	41	111	59	4,616
837,799	329	195	524	105	1,487,492,288
6,631,675	362	214	576	222	30,171,305,459,816
8,088,063	356	210	566	246	8,077,577,336
8,400,511	429	256	685	214	79,712,307,440

Source: Author's computation. 837,799 and 8,400,511 have the largest total stopping times below 10^6 and 10^7; 6,631,675 reaches the highest maximum and 8,088,063 has the largest stopping time below 10^7.

4.2. Growth of the Mean

Figure 1(a) shows σ(n) for a sample of n up to 10^7 together with the curve 6.952 ln n. The least-squares regression of binned means on ln n gives a slope of 6.817 and an intercept of 0.51, with $R^2 = 0.9996$. The fitted slope is 1.9 per cent below the predicted value μ = 6.952. In the window $5 \times 10^6 \leq n \leq 10^7$, σ(n) falls short of μ ln n by 1.40 steps on average. The mean of σ(n) over all $n \leq 10^7$ is 103.7 steps. Table 2 shows that the ratio of σ(n) to ln n remains between 6.84 and 6.86 across windows from 10^3 to 10^7.

Table 2: Statistics of total stopping times in selected windows compared with the random-walk model

Window	Mean σ(n)/ln n	Var σ(n)/ln n	Mean Z	SD of Z	Skewness of Z
$10^3 – 2 \times 10^3$	6.842	96.6	−0.030	0.976	0.53
$10^4 – 2 \times 10^4$	6.859	91.0	−0.029	0.946	0.51
$10^5 – 2 \times 10^5$	6.844	90.9	−0.037	0.946	0.57
$10^6 – 2 \times 10^6$	6.857	91.8	−0.036	0.951	0.62
$5 \times 10^6 – 10^7$	6.863	92.6	−0.035	0.955	0.64
Model prediction	6.952	101.4	0	1	—

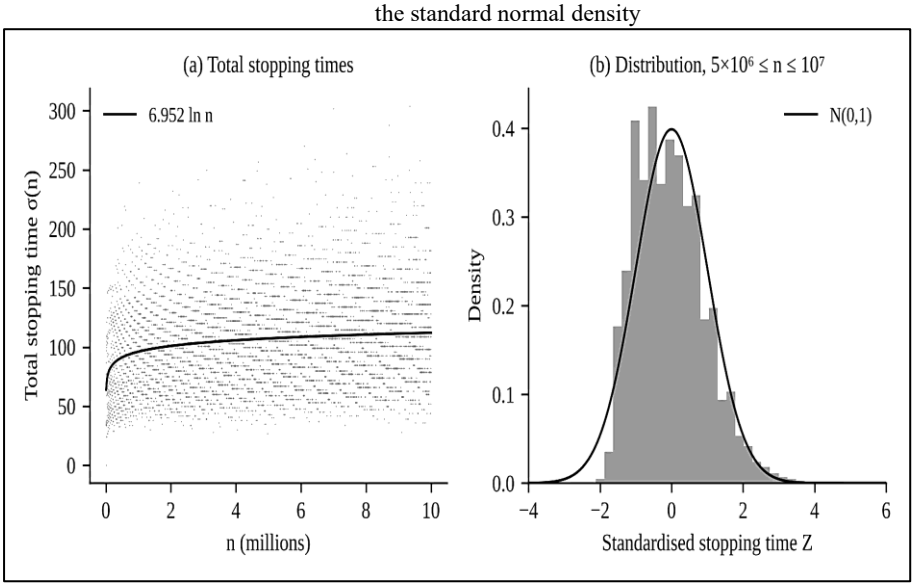
Source: Author's computation. $z = \frac{\sigma(n) - \mu \ln n}{\sqrt{\kappa \ln n}}$ *with μ = 6.952 and κ = 101.39 from Equations (4) and (5).*

4.3. Variance and Distribution

The variance of σ(n) grows approximately in proportion to ln n, as the model predicts, but with a smaller coefficient. Across the five windows, Var σ(n)/ln n lies between 90.9 and 96.6, about 5 to 10 per cent below the predicted κ = 101.39 (Table 2). Consequently, the standard deviation of the standardised variable Z is about 0.95 rather than 1.

Figure 1(b) compares the distribution of Z for $5 \times 10^6 \leq n \leq 10^7$ with the standard normal density. The centre of the distribution is close to Gaussian: 68.7 per cent of values lie within one unit of zero, compared with 68.3 per cent for a normal distribution. However, the distribution is markedly asymmetric. Its skewness is about 0.6, and its range extends from −2.18 to 7.91, so that the right tail is much longer than the left. The fraction of values within two units of zero is 97.0 per cent, higher than the normal 95.4 per cent, reflecting the truncated left tail.

Figure 1: (a) Total stopping times σ(n) for a sample of $n \leq 10^7$ with the predicted mean 6.952 ln n; (b) distribution of standardised stopping times for $5 \times 10^6 \leq n \leq 10^7$ compared with the standard normal density



Source: Author's computation. Panel (a) shows every 997th integer; panel (b) uses every fifth integer in the window.

4.4. Stopping Times

The distribution of stopping times follows the simple parity structure of the map. Exactly half of all integers greater than 1, the even ones, have stopping time 1; 75 per cent have stopping time at most 2; and 87.5 per cent at most 5. The proportions with stopping time at most 10 and 20 are 93.75 and 97.39 per cent respectively. These values reflect the fact that the stopping time is determined by the residue of n modulo powers of 2, as shown by Terras.[4]

5. DISCUSSION

The results provide substantial support for the random-walk model as a description of Collatz trajectories, while revealing systematic departures that deserve attention.

First, the predicted logarithmic growth of the mean is clearly confirmed, and the fitted slope agrees with the predicted constant to within 2 per cent. The small shortfall, together with the negative offset of about 1.4 steps, is consistent with boundary effects near the end of a trajectory: the model treats the descent to 1 as a continuous random walk, whereas real trajectories terminate deterministically through small values such as 16, 8, 4, 2 and 1, where the parity sequence is fixed rather than random.

Second, the variance grows in proportion to ln n, as predicted, but is about 9 per cent smaller than the model suggests. One possible explanation is that the parity sequences of Collatz trajectories are not fully independent. For example, after an odd step $\frac{3n+1}{2}$, the parity of the next iterate is determined by the residue of n modulo 4, so long runs of particular parities are constrained by congruence conditions. Such dependence can reduce variability relative to independent coin tossing.

Third, the distribution of standardised stopping times is positively skewed with a long right tail, as expected of first-passage times, which cannot be much smaller than the minimum number of halvings required but can be much larger when a trajectory makes a long excursion. The model's Gaussian limit applies as ln n →

∞, but because ln n grows so slowly, the range 10^7 corresponds to ln n ≈ 16, far from the asymptotic regime. The skewness observed here did not decrease between 10^3 and 10^7, suggesting that convergence to normality, if it occurs, is very slow.

These observations do not bear directly on the truth of the conjecture, which concerns every trajectory rather than typical ones. They show, however, that the heuristic picture that underlies belief in the conjecture accurately captures the typical behaviour of trajectories, while its quantitative predictions for variance and tails need refinement.

6. CONCLUSION

This study has computed the total stopping time of the Collatz map for every integer up to 10^7 and compared the results with the random-walk model. All trajectories reach 1. The mean total stopping time grows as 6.817 ln n, within 2 per cent of the predicted 6.952 ln n, and the variance grows in proportion to ln n with a coefficient about 9 per cent below the prediction. The standardised stopping times are near-Gaussian in the centre but positively skewed, with a long right tail.

Future work could:

- Extend the computation to 10^{10} or beyond to track the convergence of the variance coefficient and skewness;
- Develop refined stochastic models that incorporate dependence between successive parities and boundary effects near 1;
- Compare the results with the 5x + 1 map, for which the model predicts divergent trajectories; and
- Analyse the distribution of the maximum excursion of trajectories.

ACKNOWLEDGMENTS

[Add acknowledgments.] Conflict of interest: none declared. Funding: none received. Data availability: the code used for the computations can be provided by the authors on request. Use of artificial intelligence: the author(s) used Claude (Anthropic) to assist with computation, literature exploration and drafting; all computations, results and references were checked by the author(s), who take full responsibility for the manuscript.

REFERENCES

1. Lagarias JC. 1985. The 3x+1 problem and its generalizations. Am Math Monthly. 92(1):3–23.
2. Lagarias JC, editor. 2010. The ultimate challenge: the 3x+1 problem. Providence (RI): American Mathematical Society.
3. Bařina D. 2021. Convergence verification of the Collatz problem. J Supercomput. 77(3):2681–2688. doi:10.1007/s11227-020-03368-x.
4. Terras R. 1976. A stopping time problem on the positive integers. Acta Arith. 30(3):241–252.
5. Everett CJ. 1977. Iteration of the number-theoretic function f(2n) = n, f(2n + 1) = 3n + 2. Adv Math. 25(1):42–45.
6. Tao T. 2022. Almost all orbits of the Collatz map attain almost bounded values. Forum Math Pi. 10:e12. doi:10.1017/fmp.2022.8.
7. Lagarias JC, Weiss A. 1992. The 3x+1 problem: two stochastic models. Ann Appl Probab. 2(1):229–261. doi:10.1214/aoap/1177005779.
8. Kontorovich AV, Lagarias JC. 2009. Stochastic models for the 3x + 1 and 5x + 1 problems. arXiv:0910.1944 [math.NT].
9. Crandall RE. 1978. On the '3x + 1' problem. Math Comp. 32(144):1281–1292.
10. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, et al. 2020. Array programming with NumPy. Nature. 585(7825):357–362. doi:10.1038/s41586-020-2649-2.