Some considerations on the total stopping time for the Collatz problem

Nicola Fabiano[a], *Nikola* Mirkov[b], *Stojan* Radenović[c]

[a] University of Belgrade, "Vinča" Institute of Nuclear Sciences - National Institute of the Republic of Serbia, Belgrade, Republic of Serbia,
e-mail: nicola.fabiano@gmail.com, **corresponding author**,
ORCID iD: https://orcid.org/0000-0003-1645-2071

[b] University of Belgrade, "Vinča" Institute of Nuclear Sciences - National Institute of the Republic of Serbia, Belgrade, Republic of Serbia,
e-mail: nmirkov@vin.bg.ac.rs,
ORCID iD: https://orcid.org/0000-0002-3057-9784

[c] University of Belgrade, Faculty of Mechanical Engineering, Belgrade, Republic of Serbia,
e-mail: radens@beotel.rs,
ORCID iD: https://orcid.org/0000-0001-8254-6688

https://doi.org/10.5937/vojtehg72-50306

FIELD: mathematics
ARTICLE TYPE: original scientific paper

Abstract:

Introduction/purpose: The Collatz conjecture has been considered and the stopping time needed for the recursive transformation to end has been investigated.

Methods: A statistical analysis on the stopping time has been used.

Results: The statistical approach shows that the probability of finding an infinite stopping time, that is a violation of the Collatz conjecture, is extremely low.

Conclusion: Picking precisely one particular atom in the Universe is still more favorable, by more than 61 orders of magnitude, than encountering an infinite total stopping time.

Key words: Collatz conjecture, recurrences, statistical analysis, curve fitting.

Introduction and definitions

The Collatz conjecture starts from a very simple function. For $N \in \mathbb{N}$ define $C(N)$ as

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

This function applied recursively creates a sequence. Translating $C(N)$ to a sequence $\{a_i\}_{i\in\mathbb{N}}$, applying recursively the operation starting from a positive integer N, one could write a_i as follows:

$$a_i = \begin{cases} N & \text{for } i = 0 \\ C(a_{i-1}) & \text{for } i > 0 \ , \end{cases} \tag{2}$$

so that a_i is obtained by iterating the application of C for i times, $C(C(C(\ldots C(a_0))))$, written as $a_i = C(N)^i$.

Starting the sequence from $N = 12$, for instance, one has:

$$12, 6, 3, 10, 5, 16, 8, 4, 2, 1,$$

while the same procedure for $N = 3333$ leads to

$$3333, 10000, 5000, 2500, 1250, 625, 1876, 938, 469, 1408, 704, 352, 176, 88, 44, \\ 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1 \ .$$

In both cases, the sequence ends in the same manner with the number 1. The number of steps necessary to end the sequence started with N, that is, to reach number 1, is called stopping time. The conjecture of Collatz, stated almost a hundred years ago in 1937 (MacTutor, 2024), is that for every $N \in \mathbb{N}$ the stopping time is finite. Up to now, there is no proof of the conjecture, or a counterexample disproving it. There is a huge amount of literature on the subject; for a short list, see, for example (Applegate & Lagarias, 1995a,b; Fabiano et al., 2021, 2023; Guy, 2004; Kurtz & Simon, 2007; Lagarias, 1985; Weisstein, 2024) and the references therein. Empirical evidence, that is, numerical simulation, confirms the conjecture that the stopping time is finite for every N, and that there does not exist another ending sequence except from the one described above with the final number 1.

In this work, some results obtained for the stopping time and its consequences will be discussed.

Total stopping time

Collatz transformations (1) and (2) show that even numbers generate smaller numbers. In particular, powers of 2, 2^k, have the shortest total stopping times equal to k. Odd numbers clearly add more steps to the sequence, thus increasing total stopping times. Therefore, even numbers generated from N tend to shorten the sequence, while odd numbers generate larger numbers that increase the stopping time. Let us define $E(N)$ as the number of even numbers generated by the sequence starting from the value N, and $O(N)$ as the number of odd numbers generated in the same sequence. Define the total stopping time $T(N)$ as

$$T(N) = \max_{N' \leq N} T(N') \tag{3}$$

for $N, N' \in \mathbb{N}$, being the largest stopping time for all numbers smaller or equal to N, which clearly has the property that $T(N) = E(N) + O(N)$. To provide some examples, the Collatz sequences for the first few numbers of $\mathbb{N}$ are, respectively,

$N = 3$:

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

that is, $E(3) = 5,\ O(3) = 3$, so $T(3) = 8$.

$N = 4$:

$$4, 2, 1,$$

$E(4) = 2,\ O(4) = 1$, and $T(4) = 8$ because the loop of $N = 3$ is longer than the one of $N = 4$.

$N = 5$:

$$5, 16, 8, 4, 2, 1,$$

$E(5) = 4,\ O(5) = 2$, and $T(5) = 8$ once again for the same reason of the case $N = 4$.

From Collatz transformations (1) and (2), one could write the relation

$$2^{E(N)} = 3^{O(N)} \cdot N \cdot \mathrm{Res}(N), \tag{4}$$

where $\mathrm{Res}(N)$ is called the residue, and is defined as

$$\mathrm{Res}(N) = \prod_{0 \leq i < T(N)}{}' \left(1 + \frac{1}{3C(N)^i}\right) , \tag{5}$$

where $\prod'$ means that the product is taken on all odd values of $C(N)^i$.

Fabiano, N. et al., Some considerations on the total stopping time for the Collatz problem, pp.1019–1028

The coefficient R is defined as

$$R = \frac{E(N)}{O(N)} , \tag{6}$$

sometimes the inverse of R is called “completeness”. The following theorem could be stated:

THEOREM 1. *For every* $N \in \mathbb{N},\ N > 1$*:*

$$R > \frac{\ln(3)}{\ln(2)} \approx 1.58 . \tag{7}$$

Proof. Looking at eqs. (4) and (6) and taking the logarithm, it is possible to write

$$\frac{E(N)}{O(N)} = \frac{\ln(3)}{\ln(2)} + \frac{(\ln(N) + \ln \mathrm{Res}(N))}{O(N)\ln(2)} > \frac{\ln(3)}{\ln(2)} . \tag{8}$$

□

The investigation shall now proceed further on the behaviour for large N of the residue (4). From eq. (1), $C(N) \geq N/2$, therefore

$$\mathrm{Res}(N) \leq \prod_{0\leq i<T(N)}{}' \left(1 + \frac{2^i}{3N^i}\right) .$$

Rewriting the upper bound on the residue in the following manner

$$\mathrm{Res}(N) \leq \exp \sum_{0\leq i<T(N)}{}' \ln\left(1 + \frac{2^i}{3N^i}\right) \tag{9}$$

and taking the limit, one ends up with the result

$$\lim_{N\to+\infty} \ln(\mathrm{Res}(N)) = 0 . \tag{10}$$

The value of R (6) therefore could assume a range of values between $\ln(3)/\ln(2)$ and k, for $N = 2^k$, $N > 2$. The minimal value is reached in the case of an infinite total stopping time, while the value of k is reached when N equals powers of 2, 2^k. Conversely, the completeness has values in the range $[1/k, \ln(2)/\ln(3))$.

Our analysis starts from the results shown in Figure $4a$ of (Fabiano et al., 2021, 2023) and further extends the analysis adding more points from

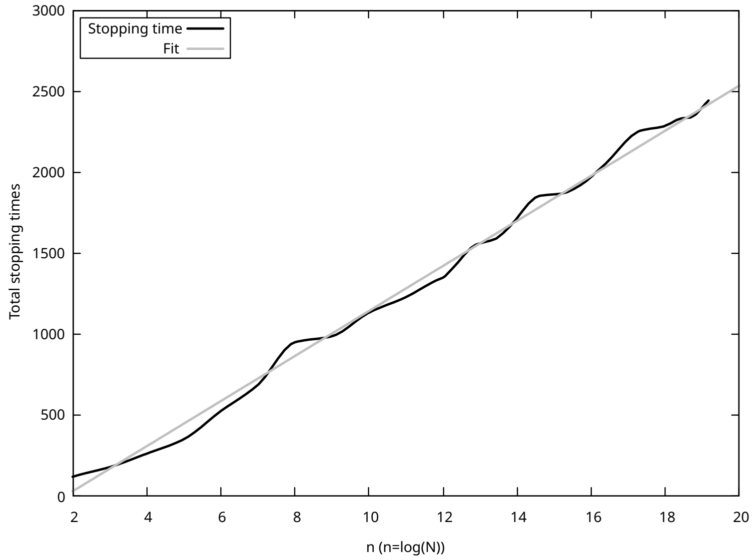

Figure 1 – Total stopping time as a function of $n = \log(N)$, compared to a linear fit

numerical analysis and also using the available data for the largest total stopping times discovered up to now (Roosendaal, 2024) for N of the order of 10^{20}.

Figure (1) shows the plot of the total stopping time as a function of n, $n = \log(N)$, together with a linear fit of the curve. There is clearly a behaviour with increasing n that reproduces remarkably well a linear function, from the fit $139.2n - 249.8$ with the correlation coefficient $r = 0.9972$ quite close to 1. Thus, the total stopping time does not departure significantly from its minimal value encountered for the discussed case of powers of 2, a purely linear function of $\log(N)$.

As a result, it is possible to infer that the behaviour of the total stopping time for large N is proportional to its log, that is,

$$T(N) \sim \log(N) \; . \tag{11}$$

Figure (2) presents the ratio of even to odd numbers for large N, formula (6), of the total stopping time when applying the Collatz procedure. It is apparent that this ratio decreases with growing N while seemingly approaching an asymptotic value larger than zero.

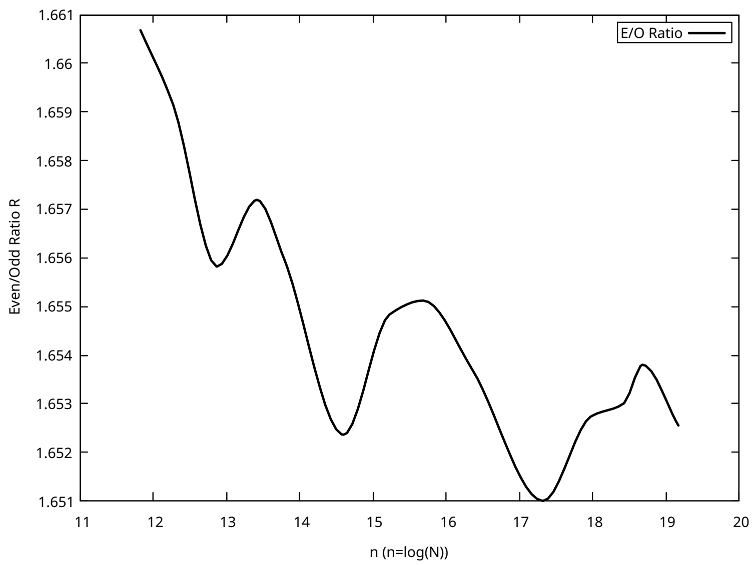

Figure 2 – Even to odd ratio for the total stopping time as a function of $n = \log(N)$

The statistics from the data of Figure (2) gives for the average of R and its standard deviation:

$$\overline{R} \pm \sigma_R = 1.65485 \pm 0.00272805 \quad (12)$$

We assumed that the results of eq. (12) represent a normal distribution. The hypothesis on the values of R being normally distributed has been supported by many tests effectuated on the data set, like Kolmogorov–Smirnov, Pearson χ^2, Shapiro–Wilk, Anderson–Darling and Kuiper, which did not contradict it.

The distance of $\overline{R}$ from the value of $\ln(3)/\ln(2)$ is given by 25.6181 times the standard deviation σ_R, that is, $\overline{R} - 25.6181\sigma_R = \ln(3)/\ln(2)$. The probability for this event to happen is smaller than 10^{-143}.

As a comparison term, the number of atoms in the Universe is of the order of 10^{82}, meaning that picking precisely one particular atom in the Universe is still more favorable, by more than 61 orders of magnitude, than encountering an infinite total stopping time. The current agreement for a discovery in particle physics should have at least a five sigma, 5σ, discrepancy with the already known physics of the Standard Model. Such an event has a probability to occur of less than 10^{-6}.

This result coming from statistical analysis infers that the probability of finding an infinite total stopping time is extremely tiny, being “zero” when compared to other already very small quantities encountered in Science.

It must be stressed however that this result for the finiteness of the total stopping time cannot rule out the possibility of existence of some values of N that under the Collatz procedure enter a different kind of loop not ending with 1.

References

Applegate, D. & Lagarias, J.C. 1995. Density bounds for the $3x+1$ problem. I. Tree-search method. *Mathematics of Computation*, 64(209), pp.411-426. Available at: https://doi.org/10.1090/S0025-5718-1995-1270612-0.

Applegate, D. & Lagarias, J.C. 1995. Density Bounds for the $3x+1$ Problem. II. Krasikov Inequalities. *Mathematics of Computation* 64(209), pp.427-438. Available at: https://doi.org/10.1090/S0025-5718-1995-1270613-2.

Fabiano, N., Mirkov, N. & Radenović, S. 2021. Collatz hypothesis and Planck’s black body radiation. *Journal of Siberian Federal University. Mathematics & Physics*, 17(3), pp.408-414 [online]. Available at: https://www.mathnet.ru/eng/jsfu/v17/i3/p408 [Accessed: 4 April 2024].

Fabiano, N., Mirkov N., Mitrović, Z.D. & Radenović S. 2023. Chapter 3: Collatz Hypothesis and Kurepa’s Conjecture. In: *Advances in Number Theory and Applied Analysis*, pp.31-50. Available at: https://doi.org/10.1142/9789811272608_0003.

Guy, R.K. 2004. *Unsolved Problems in Number Theory, Third Edition*. Springer Science & Business Media. ISBN: 978-0387-20860-2.

Kurtz, S.A. & Simon, J. 2007. The Undecidability of the Generalized Collatz Problem. In: Cai, JY., Cooper, S.B. & Zhu, H. (Eds.) *Theory and Applications of Models of Computation. TAMC 2007. Lecture Notes in Computer Science*, 4484. Berlin, Heidelberg: Springer. Available at: https://doi.org/10.1007/978-3-540-72504-6_49.

Lagarias, J.C. 1985. The $3x+1$ Problem and its Generalizations. *The American Mathematical Monthly*, 92(1), pp.3-23. Available at: https://doi.org/10.1080/00029890.1985.11971528.

MacTutor. 2024. Collatz conjecture. *MacTutor* [online]. Available at: https://mathshistory.st-andrews.ac.uk/Biographies/Collatz [Accessed: 4 April 2024].

Roosendaal, E. 2024. *On the $3x+1$ problem* [online]. Available at: http://www.ericr.nl/wondrous/delrecs.html [Accessed: 4 April 2024].

Weisstein, E.W. 2024. Collatz Problem. *MathWorld–A Wolfram Web Resource* [online]. Available at: https://mathworld.wolfram.com/CollatzProblem.html [Accessed: 4 April 2024].

Algunas consideraciones sobre la detención del tiempo total para el problema de Collatz

Nicola Fabiano[a], **autor de correspondencia**, *Nikola* Mirkov[a], *Stojan* Radenović[b]

[a] Universidad de Belgrado, Instituto de Ciencias Nucleares ”Vinča” - Instituto Nacional de la República de Serbia, Belgrado, República de Serbia

[b] Universidad de Belgrado, Facultad de Ingeniería Mecánica, Belgrado, República de Serbia

CAMPO: matemáticas
TIPO DE ARTÍCULO: artículo científico original

Resumen:

Introducción/objetivo: Se ha considerado la conjetura de Collatz y el tiempo de parada necesario y se ha investigado la detención del tiempo necesaria para que finalice la transformación recursiva.

Métodos: Se ha utilizado un análisis estadístico de la detención del tiempo.

Resultados: El enfoque estadístico muestra que la probabilidad de encontrar una detención del tiempo infinito, es decir, una violación de la conjetura de Collatz, es extremadamente baja.

Conclusión: Escoger precisamente un átomo particular en el El universo es aún más favorable, en más de 61 órdenes de magnitud, que que encontrar una detención del tiempo total infinita.

Palabras claves: Conjetura de Collatz, recurrencias, análisis estadístico, ajuste de curvas.

Некоторые соображения относительно общего времени остановки для задачи Коллатца

Никола Фабиано[a], **корреспондент**, *Никола* Мирков[a], *Стоян* Раденович[б]

[a] Белградский университет, Институт ядерных исследований «Винча» − Институт государственного значения для Республики Сербия, г. Белград, Республика Сербия

[б] Белградский университет, факультет машиностроения, г. Белград, Республика Сербия

РУБРИКА ГРНТИ: 27.15.17 Элементарная теория чисел,
27.43.17 Математическая статистика
ВИД СТАТЬИ: оригинальная научная статья

Резюме:

Введение/цель: В данной статье рассмотрена гипотеза Коллатца и исследовано время остановки, необходимое для завершения рекурсивного преобразования.

Методы: Был использован статистический анализ времени остановки.

Результаты: Статистический подход показывает, что вероятность нахождения бесконечного времени остановки, то есть нарушения гипотезы Коллатца, чрезвычайно мала.

Вывод: Выбор именно одного конкретного атома во Вселенной большего на 61 порядок вероятнее, чем вероятность столкнуться с бесконечным общим временем остановки.

Ключевые слова: гипотеза Коллатца, рекуррентность, статистический анализ, подгонка кривой.

Нека разматрања о укупном времену заустављања за Колацов проблем

Никола Фабиано[a], **аутор за преписку**, *Никола* Мирков[a], *Стојан* Раденовић[б]

[a] Универзитет у Београду, Институт за нуклеарне науке "Винча" – Национални институт Републике Србије, Београд, Република Србија

[б] Универзитет у Београду, Машински факултет, Београд, Република Србија

ОБЛАСТ: математика
КАТЕГОРИЈА (ТИП) ЧЛАНКА: оригинални научни рад

Сажетак:

Увод/циљ: Размотрена је Колацова претпоставка и потребно време за заустављање рекурзивне трансформације.

Методе: Коришћена је статистичка анализа времена заустављања.

Резултати: *Статистички приступ показује да је вероватноћа проналажења бесконачног времена заустављања, што нарушава Колацову хипотезу, изузетно ниска.*

Закључак: *Вероватноћа одабира тачно једног атома у целокупном универзуму је за више од 61 реда величине вероватнија од наилажења бројног низа са бесконачним временом заустављања у Колацовом проблему.*

Кључне речи: *Колацова хипотеза, понављања, статистичка анализа, апроксимација криве.*

Paper received on: 08.04.2024.
Manuscript corrections submitted on: 23.09.2024.
Paper accepted for publishing on: 24.09.2024.