



HAL
open science

Collatz conjecture Demonstration of the relationship between the numbers of even and odd steps before reaching 1, and the initial odd value of a Collatz sequence that converges

Farid Baleh

► **To cite this version:**

Farid Baleh. Collatz conjecture Demonstration of the relationship between the numbers of even and odd steps before reaching 1, and the initial odd value of a Collatz sequence that converges. In press. hal-02935616

HAL Id: hal-02935616

<https://hal.science/hal-02935616v1>

Preprint submitted on 24 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Collatz conjecture

Demonstration of the relationship between the numbers of even and odd steps before reaching 1, and the initial odd value of a Collatz sequence that converges

Farid Baleh - farid.baleh@gmail.com
2020/09/11

Summary

The relationship between the numbers of even steps (P), the number of odd steps (I) and the odd initial value u_0 of a compressed Collatz sequence that converges is as follows:

$$I + P = E(\log_2(3^I * u_0)) + 1$$

This relation is equivalent to the expression below:

$$\left\{ \begin{array}{l} I + P = E(I * \log_2(3)) + E(\log_2(u_0)) + 1 \\ \text{or} \\ I + P = E(I * \log_2(3)) + E(\log_2(u_0)) + 2 \end{array} \right.$$

Where:

- $E(x)$ is the integer part of the real number x ;
- $\log_2(x)$ is the base two logarithm of the real number x ;
- $E(\log_2(u_0))$ is the exponent of the biggest power of 2 that is strictly less than the odd natural number u_0 ;
- I+P is the total stopping time of the compressed sequence.

1- Introduction

The Collatz conjecture (or Syracuse conjecture, Ulam conjecture or $3x+1$ problem) claims that the following sequence of natural numbers reaches the value 1 after a certain rank (or that the sequence of its odd numbers tends to 1).

For a natural number $s_0 \geq 1$ and for all natural number n :

$$s_n = \begin{cases} \frac{s_n}{2}, & \text{if } s_n \text{ is even} \\ 3s_n + 1, & \text{if } s_n \text{ is odd} \end{cases}$$

$3p+1$ being even if the natural number p is odd, the compressed sequence (c_n) of the sequence (s_n) is defined as follows:

For a natural number $c_0 \geq 1$ and for all natural number n :

$$c_n = \begin{cases} \frac{c_n}{2}, & \text{if } c_n \text{ is even} \\ \frac{3c_n+1}{2}, & \text{if } c_n \text{ is odd} \end{cases}$$

An uncompressed (or compressed) sequence that converges continues, after a certain rank, with the trivial cycle 1-4-2 (or 1-2) that is infinitely repeated.

2- General expression of an odd element of a compressed Collatz sequence

Considering the extracted sequence (u_n) composed of the odd elements of the sequence (s_n) , two successive elements have the following relationship:

$$u_n = \frac{3u_{n-1} + 1}{2^{k'_n}}$$

Where k'_n is the number of divisions by 2 that occur between u_{n-1} and u_n .

By developing the previous expression:

$$u_n = \frac{3^n u_0}{2^{\sum_{j=1}^n k'_j}} + \frac{3^{n-1}}{2^{\sum_{j=1}^n k'_j}} + \frac{3^{n-2}}{2^{\sum_{j=2}^n k'_j}} + \dots + \frac{3^0}{2^{k'_n}}$$

Then:

$$u_n = \frac{3^n u_0}{2^{\sum_{j=1}^n k'_j}} + \sum_{i=1}^n \frac{3^{n-i}}{2^{\sum_{j=i}^n k'_j}}$$

After factorization of the u_0 multiplier and simplification of the second term:

$$u_n = \frac{3^n}{2^{\sum_{j=1}^n k'_j}} \left[u_0 + \sum_{i=1}^n \left(3^{-i} * 2^{\sum_{j=1}^{i-1} k'_j} \right) \right]$$

Moreover, for all j : $k'_j = 1 + k_j$, because the successor of an odd element of (s_n) is always even; k_j is then the number of divisions by 2 between two successive odd elements of sequence (c_n) .

Therefore: $\sum_{j=1}^n k'_j = n + \sum_{j=1}^n k_j$ and $\sum_{j=1}^{i-1} k'_j = (i-1) + \sum_{j=1}^{i-1} k_j$

Notice also that:

$$u_n = \frac{3u_{n-1} + 1}{2^{1+k_n}}$$

$\sum_{j=1}^n k_j$ is the number of the even elements, at the rank j , of the compressed sequence (c_n) that follow u_0 .

By introducing k_j and after a new factorization, the expression of u_n becomes:

$$u_n = \left(\frac{3}{2}\right)^n * \frac{1}{2^{\sum_{j=1}^n k_j}} \left[u_0 + \frac{1}{3} \sum_{i=1}^n \left(\left(\frac{2}{3}\right)^{i-1} * 2^{\sum_{j=1}^{i-1} k_j} \right) \right]$$

After a shift on the index i , the expression of u_n , depending on an initial value u_0 , of n and of the first n elements of the sequence (k_n) , we have that:

$$u_n = \left(\frac{3}{2}\right)^n * \frac{1}{2^{\sum_{j=1}^n k_j}} \left[u_0 + \frac{1}{3} \sum_{i=0}^{n-1} \left(\left(\frac{2}{3}\right)^i * 2^{\sum_{j=1}^i k_j} \right) \right]$$

And hence:

$$\forall n \geq 0, \quad u_n = \frac{1}{a_n} \left(u_0 + \frac{x_n}{3} \right) \quad (1)$$

With:

$$x_n = \sum_{i=0}^{n-1} \left(\left(\frac{2}{3}\right)^i * 2^{\sum_{j=1}^i k_j} \right) \quad (2)$$

$$a_n = \left(\frac{2}{3}\right)^n * 2^{\sum_{j=1}^n k_j} \quad (3)$$

$\forall n \geq 1$, x_n and a_n are strictly positive. Therefore (k_n) is strictly increasing.

Moreover:

$$a_n = x_{n+1} - x_n \quad \text{and} \quad x_n = \sum_{i=0}^{n-1} a_i$$

a_n is the general term of serie (x_n) .

We also can write that, for all natural number n:

$$a_n = \left(\frac{2}{3}\right)^n * 2^{P(n)} \quad (4)$$

And:

$$x_n = \sum_{i=0}^{n-1} \left(\frac{2}{3}\right)^i * 2^{P(i)}$$

With: $\forall i \in [1, n], P(i) = \sum_{j=1}^i k_j$

P(i) is the number of even elements, at rank i, of the compressed sequence (c_n) that follow u_0 , the corresponding number of odd elements being of course i.

Note also that: $x_0 = 0$ and $a_0 = x_1 = 1$.

3- Relation between the numbers of even steps, odd steps and u_0

a. Expression of I+P as a function of I and a_I

Considering Collatz sequence (u_n) composed of the odd elements of the sequence (s_n) , that converges to 1, i.e. $\lim_{n \rightarrow +\infty} (u_n) = 1$.

This sequence being digital, it exists a rank I for which $u_I = 1$.

Therefore, the formula (1) gives: $u_I = \frac{1}{a_I} \left(u_0 + \frac{x_I}{3}\right) = 1$; then:

$$a_I = u_0 + \frac{x_I}{3} \quad (5)$$

According to (4): $a_I = \left(\frac{2}{3}\right)^I * 2^{P(I)}$

We can note $P=P(I)$; I and P are respectively the odd and even steps of the sequence (c_n) before reaching 1, I+P being then its total stopping time.

By applying the base two logarithms of the two terms of the previous equation, we have that:

$$I + P = \log_2(3^I * a_I) = \log_2 \left[3^I * \left(u_0 + \frac{x_I}{3}\right) \right] = I * \log_2(3) + \log_2 \left(u_0 + \frac{x_I}{3}\right)$$

I+P is an integer that is the sum of the two last terms of this equality, each of them being non integer real numbers. Consequently, by using $E(x)$ as the integer part of the real number x:

$$E(I + P) = I + P = E(I * \log_2(3)) + E\left(\log_2 \left(u_0 + \frac{x_I}{3}\right)\right) + 1 \quad (6)$$

b. Expression of $E(\log_2(a_I))$ as a function of $n_0 = E(\log_2(u_0))$

We have: $u_0 = 2^{n_0} + r_0$, r_0 being a natural number ; n_0 is the exponent of the biggest power of 2 that is strictly less than the odd natural number u_0 .

We can apply the two following inequalities:

- $2^{n_0} \leq u_0 < 2^{n_0+1} \Rightarrow n_0 \leq \log_2(u_0) < n_0 + 1 \Rightarrow n_0 = E(\log_2(u_0))$ by definition of the integer part function of a real number ;
- $0 \leq r_0 < 2^{n_0}$.

Notice that, u_0 being odd, the two above inequalities are strict.

According to (5), $a_I > u_0$ because $x_I > 0$; therefore: $\log_2(u_0) < \log_2(a_I)$; then $n_0 \leq E(\log_2(a_I))$, the logarithm and integer part functions being increasing.

Moreover, it is obvious that, according to (4):

$$\forall n \geq 0, a_{I+n} = \left(\frac{2}{3}\right)^{I+n} * 2^{P(I+n)} = \left(\frac{2}{3}\right)^{I+n} * 2^{P+n} = \left(\frac{4}{3}\right)^n * a_I$$

Because $P(I+n) = P(I) + P(n) = P + n$, $P(n)$ being incremented by 1 from a_{I+n} to a_{I+n+1} .

$$\text{Then: } \forall n \geq 0, \log_2(a_{I+n}) = n * \log_2\left(\frac{4}{3}\right) + \log_2(a_I) \quad (7)$$

(a_{I+n}) is an increasing sequence, with the integer part of the base two logarithm of its first term a_I being greater or equal to n_0 .

We have that: $\log_2\left(\frac{4}{3}\right) \approx 0,415$.

Consequently, a maximum of 3 consecutive terms of the increasing sequence $(E(\log_2(a_{I+n})))$ may be equal.

If we consider the first natural number n_1 for which:

$$E(\log_2(a_{I+n_1})) = n_0 + 1$$

According to (7): $\log_2(a_{I+n_1}) = n_1 * \log_2\left(\frac{4}{3}\right) + \log_2(a_I) = N_1 + \log_2(a_I)$

By noting: $N_1 = n_1 * \log_2\left(\frac{4}{3}\right)$.

By deduction, we get:

$$\begin{cases} E(\log_2(a_{I+n_1})) = E(N_1) + E(\log_2(a_I)) \\ \text{or} \\ E(\log_2(a_{I+n_1})) = E(N_1) + E(\log_2(a_I)) + 1 \end{cases}$$

And then:

$$\begin{cases} n_0 + 1 = E(N_1) + E(\log_2(a_I)) \\ \text{or} \\ n_0 + 1 = E(N_1) + E(\log_2(a_I)) + 1 \end{cases}$$

By taking account the inequality $n_0 \leq E(\log_2(a_I))$, we obtain the two possible following inequalities:

$$\begin{cases} n_0 + 1 \geq E(N_1) + n_0 \\ \text{or} \\ n_0 + 1 \geq E(N_1) + n_0 + 1 \end{cases}$$

Then:

$$\begin{cases} E(N_1) \leq 1 \\ \text{or} \\ E(N_1) \leq 0 \Rightarrow E(N_1) = 0 \end{cases}$$

The integer part of a positive real number being positive or equal to 0, but not negative, we have: $E(N_1) = 0$ or $E(N_1) = 1$. Therefore, the possible values of n_1 are : 0, 1, 2, 3 or 4 (see table below).

n1	N1=n1*log2(4/3)	E(N1)
0	0,00	0
1	0,42	0
2	0,83	0
3	1,25	1
4	1,66	1
5	2,08	2

Investigate the different cases:

Case 1: if $n_1 = 0$, then $E(\log_2(a_I)) = n_0 + 1$ by hypothesis.

Case 2: if $n_1 = 1$, then $E(\log_2(a_I)) = n_0$, because $n_0 \leq E(\log_2(a_I))$ and a_{I+1} is the first term of sequence (a_{I+n}) for which $E(\log_2(a_{I+1})) = n_0 + 1$.

Case 3: if $n_1 = 2$, then $E(\log_2(a_{I+2})) = n_0 + 1$; then $E(\log_2(a_I)) = E(\log_2(a_{I+1})) = n_0$, because $n_0 \leq E(\log_2(a_I))$ and a_{I+2} is the first term of sequence (a_{I+n}) for which $E(\log_2(a_{I+2})) = n_0 + 1$.

Case 4: if $n_1 = 3$, then $E(\log_2(a_{I+3})) = n_0 + 1$; therefore, like for case 3, $E(\log_2(a_I)) = E(\log_2(a_{I+1})) = E(\log_2(a_{I+2})) = n_0$.

Case 5: if $n_1 = 4$, then $E(\log_2(a_{I+4})) = n_0 + 1$; **this case is impossible** because that would imply that $E(\log_2(a_I)) = E(\log_2(a_{I+1})) = E(\log_2(a_{I+2})) = E(\log_2(a_{I+3})) = n_0$, and then that 4 consecutive terms of the sequence $(E(\log_2(a_{I+n})))$ would be equal, the maximum being 3 like precised above (a maximum of 3 consecutive terms of the increasing sequence $(E(\log_2(a_{I+n})))$ can be equal).

Therefore:

$$\begin{cases} E(\log_2(a_I)) = n_0 \\ \text{or} \\ E(\log_2(a_I)) = n_0 + 1 \end{cases}$$

4- Conclusion

The relations (5) and (6) imply that:

$$\begin{cases} I + P = E(I * \log_2(3)) + E(\log_2(u_0)) + 1 \\ \text{or} \\ I + P = E(I * \log_2(3)) + E(\log_2(u_0)) + 2 \end{cases}$$

Which is equivalent to:

$$I + P = E(\log_2(3^I * u_0)) + 1$$

$I+P$ is then the « ceiling » function of the real number $\log_2(3^I * u_0)$.

This last formula is the same that in Ruggiero[2] related to sequence (s_n) , if we add I at each of the two terms of the equation above.

Bibliography

[1] J. C. Lagarias, "The $3x+1$ problem and its generalizations", The American Mathematical Monthly, vol. 92, no. 1, pp. 3–23, 1985.

[2] Rafael Ruggiero, "The relationship between stopping time and number of odd terms in Collatz sequences", arXiv: 1911.01229v2, [math.GM] 6 Nov. 2019.