A probabilistic proof for the Syracuse conjecture

El ghazi Imad (✉ elimadimad@gmail.com)

Research Article

Keywords: The Syracuse(Collatz) conjecture

Posted Date: March 16th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1065511/v6>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A probabilistic proof for the Syracuse conjecture

Imad El ghazi*

2021

Abstract

We prove the veracity of the Syracuse conjecture by establishing that from an arbitrary positive integer different from 1 and 4, the Syracuse process will never return (after $i \geq 1$ steps) to any positive integer already reached and we conclude using a probabilistic approach.

Classification : MSC: 11A25

1 Introduction

The Syracuse conjecture is an idea introduced by Lothar Collatz in 1937. It is also known as the $3n+1$ problem and has been studied by many mathematicians as J.J. O'Connor, J.J. Robertson, E.F. in [1] and T. Tao in [2], since its first appearance. We consider the following operation on an arbitrary positive integer l :

- If l is even, divide it by two.
- If the l is odd, triple it and add one.

The Collatz (or Syracuse) conjecture is: This process will eventually reach the number 1, regardless of which positive integer is chosen initially.

We can also understand this process by the following:

If l is a positive even integer (when l is a positive odd integer we get to the even case by tripling l and adding one to the result of the last multiplication) we

* author e-mail address: elimadimad@gmail.com

divide it by 2 until we get an odd number, this last one we triple it and we add one, or we continue dividing l by two, until we get to 1. This last case is possible just when l is of the form $l = 2^n$ with $n \in \mathbb{N}^*$. In fact when l is odd by tripling it and adding one, what we do is trying to get to a positive even integer of the form $3l+1 = 2^n$ ($n \in \mathbb{N}$ an even integer). Of course, half the numbers of the form 2^n can be written $3k+1$, k been a positive odd integer, the other half is of the form $3k-1$.

The Syracuse process can be modeled as a random variable taking its values in the set of positive integers (strictly superior to 1 and different of 4) without any possibility to return (after $i \geq 1$ steps) to a positive integer reached before.

Using this random walk modelization of the Syracuse process and by a Bernoulli trial argument we prove that the Syracuse conjecture is true.

2 Main results

Let l be a strictly positive integer:

a) If l is an odd integer then the next odd integer will be reached after these two operations:

- triple l and add one.
- divide $3l+1$ by 2 until we have the second odd integer.

b) If l is an even integer then the next even integer will be reached after these two operations:

- divide l by 2 until we have the first odd number.
- triple the odd integer resulting from the first operation and add one.

We will call this passage from an odd positive integer (resp. even) to the next odd (resp. even) integer a step.

Lemma 2.1. *For every positive integer l strictly superior to 1 and different of 4, the Syracuse process starting from l will never return to l after $i \geq 1$ steps.*

Proof. We first suppose that l is a positive odd integer.

Let m_j , $j \in \{1, \dots, i\}$ be the number of divisions by 2 after the j -ieth step, thus $m_j \geq 2$.

After i steps, we have l_i the i -ieth odd number reached :

$$l_i = \frac{1}{2^{m_i}} \left(\frac{3}{2^{m_{i-1}}} \left(\frac{3}{2^{m_{i-2}}} \left(\dots \left(\frac{3}{2^{m_2}} \left(\frac{3}{2^{m_1}} (3l+1) + 1 \right) + 1 \right) \dots \right) + 1 \right) + 1 \right)$$

If the process returns (after i steps) to l then we have:

$$l \times 3^i = l \times 2^{\sum_{j=1}^i m_j} - 2^{\sum_{j=1}^{i-1} m_j} - 3 \times 2^{\sum_{j=1}^{i-2} m_j} - \dots - 3^{i-2} \times 2^{m_1} - 3^{i-1}. \quad (I)$$

We first remark that l can not be a multiple of 3 otherwise $2^{\sum_{j=1}^{i-1} m_j}$ will be divisible by 3. Since we can generate an analogic equality to (I) for every l_i then none of the positive odd integers l_i is a multiple of 3.

If $i = 1$ then since l is the first odd positive integer reached (after one step) we have :

$$3l = 2^{m_1}l - 1$$

this leads to the equality:

$$l(2^{m_1} - 3) = 1$$

The last equality is meaningful in $\mathbb{N}$ if and only if $l = 1$ and $m_1 = 2$ which is absurd because l is supposed to be strictly superior to 1.

If $i = 2$, the equation (I) becomes $3^2l = 2^{m_1+m_2}l - 2^{m_1} - 3$ and hence

$$(2^{m_1+m_2} - 3^2)l = 2^{m_1} + 3.$$

In the other hand we have

$$(2^{m_1} + 3)(2^{m_2} - 3) = 2^{m_1+m_2} - 3^2 - 3 \times 2^{m_1} + 3 \times 2^{m_2} \quad (A)$$

by multiplying both sides by l we have

$$l \times (2^{m_1} + 3)(2^{m_2} - 3) = 3 \times l \times (2^{m_2} - 2^{m_1}) + l \times (2^{m_1+m_2} - 3^2)$$

hence

$$l \times (2^{m_1} + 3)(2^{m_2} - 3) = 3 \times l \times (2^{m_2} - 2^{m_1}) + 2^{m_1} + 3$$

which implies that $2^{m_1} + 3$ divide $3 \times l \times (2^{m_2} - 2^{m_1})$, since $(2^{m_1+m_2} - 3^2)l = 2^{m_1} + 3$ then $(2^{m_1+m_2} - 3^2)$ divides $3 \times (2^{m_2} - 2^{m_1})$ but $2^{m_1+m_2} - 3^2 - 3 \times (2^{m_2} - 2^{m_1}) = (2^{m_2} + 3) \times (2^{m_1} - 3) > 0$ for $m_1 > 2$ and since $2^{m_1+m_2} - 3^2 > 1$ according to proposition 2.1, we deduce that $2^{m_1} + 3$ can not divide $3 \times l \times (2^{m_2} - 2^{m_1})$ and hence the equality (I) is absurd. When $m_1 = 1$ the equality (I) becomes $(2^{m_2+1} - 3^2)l = 5$ which is also absurd since $l \neq 1$ and 5 is prime. Finally when $m_1 = 2$ the equality (I) becomes $(2^{m_2+2} - 3^2)l = 7$ which is absurd too since $l \neq 1$ and 7 is prime.

For $i \geq 3$, the equality (I) becomes

$$l \times (2^{\sum_{j=1}^i m_j} - 3^i) = 2^{\sum_{j=1}^{i-1} m_j} + 3 \times 2^{\sum_{j=1}^{i-2} m_j} + \dots + 3^{i-2} \times 2^{m_1} + 3^{i-1}.$$

Recall that $3 \times l + 1 = 2^{m_1} l_1$ and $3 \times l_{i-1} + 1 = 2^{m_i} l$, thus $2^{m_i} l - 2^{m_1} l_1$ is a multiple of 3.

If $2^{m_i} > 2^{m_1}$ and $l_1 - 1 = 0 \pmod{3}$, then $l \times 2^{m_i-m_1} - 1$ is a whole number multiple of 3. Thus there exists f_1 a positive odd integer such that $l \times 2^{m_i-m_1} = 3 \times f_1 + 1$. One consider the reverse process consisting in finding a couple $m_0 \in \mathbb{N}^*$ and p such that $2^{m_0} \times l = 3p + 1$ where p is a positive odd integer, thus in our situation, we know that the couples (m_i, l_{i-1}) and $(m_i - m_1, f_1)$ satisfy both the equality $2^{m_0} \times l = 3p + 1$ therefore the reverse process is not a loop since there is at least two possible choises. The loop made by the Syracuse process from 1 illustrates the last claim. It is clear that the Syracuse process from 1 consists in finding a couple $a \in \mathbb{N}^*$ and q such that $2^a \times q = 3 \times 1 + 1$ where q is an odd positive integer of course in this case $a = 2$ and $q = 1$, the reverse process from 1 is to find a couple $b \in \mathbb{N}^*$ and q' such that $2^b \times 1 = 3 \times q' + 1$ where q' is an odd positive integer, in this case $b = 2$ and $q' = 1$ and therefore the couple (b, q') is unique.

If $2^{m_i} > 2^{m_1}$ and $l_1 + 1 = 0 \pmod{3}$, then $l \times 2^{m_i-m_1} + 1$ is a whole number multiple of 3. Thus there exists f_2 a positive odd integer such that $l \times 2^{m_i-m_1} = 3 \times f_2 - 1$ hence $2^{m_i} \times l = 3 \times (2^{m_1} \times f_2 - \frac{2^{m_1}+1}{3}) + 1$ therefore we have

$$2^{m_i} \times f_2 - \frac{2^{m_1} + 1}{3} = l_{i-1}$$

and thus

$$\begin{aligned} \frac{3 \times l_{i-1} + 1}{2^{m_1}} &= 3 \times f_2 - 1 \\ f_2 &= \frac{l_{i-1}}{2^{m_1}} + \frac{1}{3 \times 2^{m_1}} + \frac{1}{3} \end{aligned}$$

so f_2 is not a whole number because otherwise there exist $1 \leq k_2 < 2^{m_1}$ and h_2 two positive integers such that $l_{i-1} = h_2 \times 2^{m_1} + k_2$ thus $\frac{1}{3 \times 2^{m_1}} + \frac{1}{3} + \frac{k_2}{2^{m_1}} = f_2 - h_2$, therefore $3 \times k_2 + 2^{m_1} + 1 = 3 \times 2^{m_1} \times (f_2 - h_2)$ and $3 \times k_2 = 3 \times 2^{m_1} \times (f_2 - h_2) - 2^{m_1} - 1$. It follows that $3 \times k_2 + 2^{m_1} + 1 > 6 \times 2^{m_1}$ when $f_2 - h_2 \geq 2$ but we know that since $1 \leq k_2 < 2^{m_1}$ then $3 \times k_2 + 2^{m_1} + 1 < 4 \times 2^{m_1} + 1$ and therefore $f_2 - h_2 = 1$.

Thus

$$2^{m_1} + 3 \times k_2 + 1 = 3 \times 2^{m_1}$$

It follows that

$$3 \times l_{i-1} = 3 \times h_2 \times 2^{m_1} + 3 \times 2^{m_1} - 2^{m_1} - 1$$

$$3 \times l_{i-1} + 1 = 2^{m_1} \times (3 \times h_2 + 2)$$

If h_2 is odd then $f_2 - h_2 > 1$ and $m_1 = m_i$ which contradicts our hypothesis. If h_2 is even then $m_i = m_1 + 1$ and $l = \frac{3 \times h_2 + 1}{2}$ therefore $2 \times l = 3 \times h_2 + 1$ which is also absurd because $2 \times l$ is even and $3 \times h_2 + 1$ is odd.

We use the same arguments when $2^{m_1} > 2^{m_i}$ to conclude that 2^{m_1} can not be strictly superior to 2^{m_i} . We can deduce that $2^{m_1} = 2^{m_i}$.

Because the Syracuse process performs a loop starting from l , then it performs a loop starting from each l_i , then we can repeat the same argumentation to conclude that $2^{m_1} = 2^{m_2} = \dots = 2^{m_{i-1}} = 2^{m_i}$ and thus $m_1 = m_2 = \dots = m_{i-1} = m_i = m$.

It follows that equality (I) is of the forme

$$l \times (2^{im} - 3^i) = (2^m)^{i-1} + 3 \times (2^m)^{i-2} + \dots + 3^{i-2} \times 2^m + 3^{i-1}.$$

Therefore

$$l \times (2^m - 3) = 1.$$

But this is true if and only if $l = 1$ and $m = 2$ which contradicts the hypothesis that $l \neq 1$. We can deduce that equality (I) is absurd for $i \geq 1$.

If l is even, let $r = \frac{l}{2^{m_1}}$, $m_1 \in \mathbb{N}^*$ be the first odd number reached. If we suppose that the process returns to l after i steps then it will reach r again, which is absurd according to what precedes except for $r = 1$ and in this case $l = 4$. □

Remark 2.1. *a- The lemma 2.1 confirms that the only loops performed by the Syracuse process are:*

$$1 \longrightarrow 4 \longrightarrow 1,$$

and

$$4 \longrightarrow 1 \longrightarrow 4.$$

b- Let $l \neq 1$ be a positive odd integer, the lemma 2.1 states that starting from l the Syracuse process will never comeback to l . Let $(l_k)_{k \geq 1}$, $l_k \neq 1$ be the sequence of odd integers reached by the Syracuse poces starting from l . Each positive odd integer l_k can be considred as a starting point for the Syracuse process, then according to the lemma 2.1, the Syracuse process starting from l_k can never comeback to l_k . It follows that the Syracuse process

starting from l can never comeback to any l_k , $k \geq 1$. It is then legitimate to consider the Syracuse process starting from an odd positive integer l as a drawing without replacement in the set of positive odd integers.

Theorem 2.1. *Starting from an arbitrary positive integer the Syracuse process will always reach the value 1.*

Proof. According to the Lemma 2.1, starting from an integer l , the Syracuse process will never come back to l after $i \geq 1$ steps. therefore starting from an arbitrary odd positive integer l , the Syracuse process can be assimilated to a random walk in the set of odd integers (without any possibility to comeback to any of the positive odd integers reached before), we will denote this random variable Y_l .

Remark 2.2. *When l is even then the first odd integer reached ($r = \frac{l}{2^{m_1}}$, $m_1 \in \mathbb{N}^*$) will be the starting point of the random walk of the Syracuse process.*

Let Y_l be a random variable taking values in the set $\{s = 2k + 1, k \in \mathbb{N}^*\}$, without coming back to any value reached before.

Let A be the set of positive odd integers of the forme $\frac{2^n - 1}{3}$ for $n > 2$ such that n is even. Concretely :

$$A := \left\{ \frac{2^n - 1}{3} / n \text{ is even and } n > 2. \right\}$$

Remark 2.3. *The arbitrary odd integer l is assumed not to belong to A .*

Consider the Bernoulli trial with two possible outcomes:

- "Failure" if $\{Y_l \in A\}$,
- "Success" if $\{Y_l \notin A\}$.

Let $0 \leq q \leq 1$ be the probability of the event "Success", then $1 - q$ is the probability of the event "Failure". Since the set A is an infinite subset of the set of odd numbers, the probability $1 - q$ is strictly superior to 0 and therefore $0 < q < 1$.

The probability that Y_l does not reach the set A after $i \geq 1$ steps is equal to q^i , so the probability of the event Y_l never reaching the set A is equal to $\lim_{i \rightarrow +\infty} q^i = 0$. It follows that the occurrence of the first "failure" after a finite number of the previously mentioned Bernoulli trials, is a certain event.

This means that Y_l will necessarily reach a positive odd integer belonging to A , after a finite number of steps in the set of the positive odd integers.

Once such a positive odd integer $s = \frac{2^{n_0}-1}{3}$ (for some positive even integer $n_0 > 2$) reached, the next operation in the Syracuse process is to multiply s by 3 and to add 1, then we get to the even integer 2^{n_0} , after n_0 divisions by 2, we get to the value 1.

According to what have been proved before, we deduce that starting from an arbitrary integer the Syracuse process will always reach the value 1. □

References

- [1] O'Connor, J.J.; Robertson, E.F.: "Lothar Collatz". St Andrews University School of Mathematics and Statistics, Scotland (2006).
- [2] Tao,T.: Almost all Collatz orbits attain almost bounded values. Arxiv(2019)