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 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 an even number 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 supeior to 1) without any possibility to return 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

We first recall the following proposition which will be necessary to prove the next lemma .

Proposition 2.1. *For all* $(m, n) \in \mathbb{N}^2$ *such that* $(m, n) \neq (1, 0)$*,* $(m, n) \neq (2, 1)$ *and* $2^m - 3^n > 0$*, we have*

$$2^m - 3^n \neq 1.$$

Proof. According to the Catalan's conjecture proven in 2002 by Preda Mihăilescu. □

Let l be a positive integer:

a- If l is an odd integer then the next odd integer will be reached after those 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 those two operations:
   - divide l by 2 until we have the first odd number.
   - triple the odd number resulting from the first operation and add one.

We will call this passage from l supposed to be odd (resp.even)to the next odd (resp. even) integer a step.

Lemma 2.1. *For every positive integer l strictly superior to 1 and diffrent 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, $j \in \{1, ..., i\}$ be the number of divisions by 2 after the $j-$ieth step.

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

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

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

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

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 a 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 has a sens 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 eaquation (I) becomes $3^2 l = 2^{m_1+m_2} l - 2^{m_1} - 3$ 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 according to proposition 2.1. 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 according to proposition 2.1.

We use the same idea for $i \geq 3$, the equality (I) becomes

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

In the other hand we have
$(2^{\sum_{j=1}^{i-1} m_j}+3\times 2^{\sum_{j=1}^{i-2} m_j}+...+3^{i-2}\times 2^{m_1}+3^{i-1})(2^{m_i}-3) = 2^{\sum_{j=1}^{i} m_j}-3^i+2^{m_i}\times(3\times 2^{\sum_{j=1}^{i-2} m_j}+...+3^{i-2}\times 2^{m_1}+3^{i-1})-3\times(2^{\sum_{j=1}^{i-1} m_j}+3\times 2^{\sum_{j=1}^{i-2} m_j}+...+3^{i-2}\times 2^{m_1})$ (B).

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

Recall that $3\times l+1 = 2^{m_1}l_1$, hence

$(2^{\sum_{j=1}^{i} m_j}-3^i)\times(2^{m_i}\times l-2^{m_1}\times l_1) = 2^{m_i}\times(3\times 2^{\sum_{j=1}^{i-2} m_j}+...+3^{i-2}\times 2^{m_1}+3^{i-1})-3\times(2^{\sum_{j=1}^{i-1} m_j}+3\times 2^{\sum_{j=1}^{i-2} m_j}+...+3^{i-2}\times 2^{m_1})$

If $2^{m_i} > 2^{m_1}$ and $l_1-1 = 0 \mod 3$ 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 $l\times 2^{m_i-m_1} = 3\times f_1+1$.This equivalent to say that the Syracuse process starting from f_1 reachs l thus f_1 is among the loop performed by the Syracuse process starting fom l. This means that $f_1 = l_{i-1}$ which is absurd.

If $2^{m_i} > 2^{m_1}$ and $l_1+1 = 0 \mod 3$ 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

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

and thus

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

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

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

Thus

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

It follows that

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

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

If h is odd then $m_1 = m_i$ and $l = 3 \times h + 2$ which contradicts our hypothesis.

If h is even then $m_i = m_1 + 1$ and $l = \frac{3\times h+1}{2}$ therefore $2 \times l = 3 \times h + 1$ which is also absurd because $2 \times l$ is even and $3 \times h + 1$ 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 performes a loop starting from l, then it performes a loop starting from each l_i , then we can repeat the same argumentation to conclude that $2^{m_1} = 2^{m_2} = .. = 2^{m_{i-1}} = 2^{m_i}$ and thus $m_1 = m_2 = .. = m_{i-1} = m_i$.

It follows that equality I is of the forme

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

But this is true if and only if $l = 1$ because $2^{im} - 3^i > 1$ according to proposition 2.1, but l is supposed to be diffrent of 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 exept for $r = 1$ and in this case $l = 4$.

□

Remark 2.1. *a- The lemma 2.1 conffrms that the only loops performed by the Srycuse 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 :

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

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 a non-empty (in fact it 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 $\lim_{i \to +\infty} q^i = 0$. 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)