A probabilistic solution for the Syracuse conjecture

El ghazi Imad (elimadimad@gmail.com)

Research Article

Keywords: The Syracuse(Collatz) conjecture

Posted Date: January 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1065511/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
A probabilistic solution for the Syracuse conjecture

Imad El ghazi*

2021

Abstract

We prove the veracity of the Syracuse conjecture by establishing that starting from an arbitrary positive integer different of 1 and 4, the Syracuse process will never comeback to any positive integer reached before and then we conclude by 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 superior 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 geometric distribution argument we prove that the Syracuse conjecture is true.

2 Main results

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 (even) to the next odd (even) number a step.

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

**Proof.** We first suppose that \( l \) is odd.

Let \( m_j, \ j \in \{1,...,i\} \) be the number of divisions by 2 after the \( j \)-ith step.

After \( i \) steps we have \( l_i \) the \( i \)-ith odd number reached:

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

If the process returns (after \( i \) steps) to \( l \) then we have:
\[ l \times 3^i = l \times 2^{\sum_{j=1}^{i} m_j} - 2^{\sum_{j=1}^{i-1} m_j} - 3 \times 2^{\sum_{j=1}^{i-2} m_j} - \ldots - 3^{i-2} \times 2^{m_1} - 3^{i-1} \quad (I) \]

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

\[ 3l = 2^{m_1}l - 1 \]

implies that:

\[ 3l = 2^{m_1}l - 1 \]

this leads to the equation:

\[ l(2^{m_1} - 3) = 1 \]

The last equation has a sense 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

\[ (2^{m_1 + m_2} - 3^2)l = 2^{m_1} + 3 \]

implies that \( 2^{m_1 + m_2} - 3^2 \) divides \( 2^{m_1} + 3 \).

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

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

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 \( 2^{m_1} + 3 \) divides \( 3 \times l \times (2^{m_2} - 2^{m_1}) \) since \( 3 \wedge (2^{m_1} + 3) = 1 \) and \( 2^{m_1} + 3 > l \). Hence \( 2^{m_1} + 3 \) divides \( 2^{m_2} - 2^{m_1} \) according to equation \((A)\) then

\( 2^{m_1} + 3 \) divides \( 2^{m_1 + m_2} - 3^2 \) which is absurd because \( (2^{m_1 + m_2} - 3^2)l = 2^{m_1} + 3 \) and \( l > 1 \).

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

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

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

By multiplying both sides by \(l\) we have
\[l \times (2\sum_{i=1}^{m_j-1} m_j + 3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1} + 3^{i-1})(2^{m_i} - 3) = l \times (2\sum_{i=1}^{m_j-1} m_j - 3^i) + l \times 2^{m_i} \times (3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1} + 3^{i-1}) - 3 \times (2\sum_{i=1}^{m_j-1} m_j + l \times 3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1})\]

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

which implies that \(2\sum_{i=1}^{m_j-1} m_j + 3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1} + 3^{i-1}\) divides \(2^{m_i} \times (3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1} + 3^{i-1}) - 3 \times (2\sum_{i=1}^{m_j-1} m_j + 3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1})\)

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

Then according to equation (B), \(2\sum_{i=1}^{m_j-1} m_j + 3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1} + 3^{i-1}\) divides \(2\sum_{i=1}^{m_j-1} m_j - 3^i\) which is absurd because \(l \times (2\sum_{i=1}^{m_j-1} m_j - 3^i) = 2\sum_{i=1}^{m_j-1} m_j + 3 \times 2^{i-2} \times 2^{m_1} + ... + 3^{i-2} \times 2^{m_1} + 3^{i-1}\) and \(l > 1\) and \(2\sum_{i=1}^{m_j-1} m_j - 3^i \neq 1\) according to the following remark.

**Remark 2.1.** For all \((m, n) \in \mathbb{N}^2\) \((m, n) \neq (1, 0)\) and \((m, n) \neq (2, 1)\) such that \(2^m - 3^n > 0\), if we suppose that \(2^m - 3^n = 1\) then \(3^n = 2^m - 1 = (2 - 1)(2^{m-1} + ... + 2 + 1)\) hence \(3^n + 1 = 2^m - 1 + 1 = 2^m = 2^m + 3^{i-1} + ... + 2^2 + 2^2\) since \(2^2 = 3^n + 1\) then \(3^n + 1 = 2^m + 3^{i-1} + ... + 2^2 + 2^2\) which is absurd.

If \(l\) is even, let \(r = \frac{l}{2}\) 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 \(i = 1\) and in this case and \(l = 4\).

**Remark 2.2.** • The lemma 2.1 confirms that the only loop performed by the Syracuse process is:

\[1 \rightarrow 4 \rightarrow 1\]

• 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 \((l_k)_{k \geq 1}\) be the sequence of odd integers reached by the Syracuse process starting from \(l\). Each positive odd integer \(l_k\) can be
considered 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 legitim to consider the Syracuse process starting from a 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 2.3.** When \( l \) is even then the first odd integer reached \((r = \frac{l}{2^m}, m \in \mathbb{N}^*)\) 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:

\[
A := \{ \frac{2^n-1}{3} \in \mathbb{N} / n \text{ is even and } > 2. \}
\]

**Remark 2.4.** 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 \( 0 < q < 1 \).

Consider now the random variable \( X_l \), taking value in \( \mathbb{N}^* \cup \{+\infty \} \) and representing the number of success of the previous Bernoulli trials, followed by the first failure. \( X_l \) has a geometric distribution \( \mathbb{P} \) of parameter \( q \), then:
\[
\lim_{m \to +\infty} \mathbb{P}(X_l = m) = \lim_{m \to +\infty} q^{m-1}(1 - q) = 0
\]

So,
\[
\lim_{m \to +\infty} \mathbb{P}(X_l = m) = \mathbb{P}(X_l = \lim_{m \to +\infty} m) = 0
\]

and hence \(\mathbb{P}(X_l = +\infty) = 0\). In other words \(\mathbb{P}(X_l < +\infty) = 1\), i.e., the appearance of the first "failure" after a finite number of the previous 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 odd numbers.

Once such a positive odd integer \(s = \frac{2^n - 1}{3}\) (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.

\[\square\]

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)