Paper analyzes the number of zeros in the binary representation of a natural number. The analysis is carried out using the concept of the fractional part of a number, which naturally arises when finding a binary representation. This idea relies on the fundamental property of the Riemann zeta function, which is constructed using the fractional part of a number. Understanding that the ratio of the fractional and integer parts, by analogy with the Riemann zeta function, expresses the deep laws of numbers, will explain the essence of this work. For the Syracuse sequence of numbers that appears in the Collatz conjecture, we use a binary representation that allows us to obtain a uniform estimate for all terms of the series, and this estimate depends only on the initial term of the Syracuse sequence. This estimate immediately leads to the solution of the Collatz conjecture.
Collatz conjecture.

Asset Durmagambetov, Aniyar Durmagambetova

* Correspondence: aset.durmagambet@gmail.com; Tel.: +77787286399

Keywords: Collatz conjecture; binary representation; Syracuse sequence

1. Introduction

Paper analyzes the number of zeros in the binary representation of a natural number. The analysis is carried out using the concept of the fractional part of a number, which naturally arises when finding a binary representation. This idea relies on the fundamental property of the Riemann zeta function, which is constructed using the fractional part of a number. Understanding that the ratio of the fractional and integer parts, by analogy with the Riemann zeta function, expresses the deep laws of numbers, will explain the essence of this work. For the Syracuse sequence of numbers that appears in the Collatz conjecture, we use a binary representation that allows us to obtain a uniform estimate for all terms of the series, and this estimate depends only on the initial term of the Syracuse sequence. This estimate immediately leads to the solution of the Collatz conjecture.

2. Materials and Methods

This work is based on the following methods of analysis of the Syracuse sequence:

1. Analysis of simple cases of natural numbers starting from which the Syracuse sequence quickly converges to one
2. A process of expansion of a natural number in powers of two is created.
3. The proximity to the completion of decomposition is analyzed at each stage
4. The number of zeros in the binary expansion of a natural number is calculated
5. It is shown that the number of powers of two prevails in the doitic expansions in the Syracuse sequence
6. Based on these results, it is shown that the Syracuse sequence converges to one

3. Results

In this work we present the final solution to the Collatz conjecture formulated in [1].

The Collatz conjecture concerns integer sequences generated as follows:

Start with any positive integer \( a_0 \). Every next term is defined as

\[
a_{n+1} = \alpha_n a_n + \beta_n.
\]

Where \( n \geq 0 \), and if \( a_n \) is even then \( \alpha_n = 0.5, \beta_n = 0 \) if \( a_n \) is odd, then \( \alpha_n = 3, \beta_n = 1 \).

The conjecture is that regardless of \( a_0 \), the sequence will always reach 1. The conjecture is named after Lothar Collatz, who introduced the idea in 1937. It is also known as the 3n + 1 problem, the 3n + 1 conjecture, the Ulam conjecture (after Stanisław Ulam), Kakutani’s problem (after Shizuo Kakutani), the Thwaites conjecture (after Sir Bryan Thwaites), Hasse’s algorithm (after Helmut Hasse), or the Syracuse problem.

In this work, we obtained a uniform estimate for the Syracuse sequences and proved that every 4n steps the sequences come down to a number smaller than the starting term, from which follows the solution of the Collatz problem.
4. Results

Our idea of the proof is to obtain a uniform estimate for the Syracuse sequence described in Introduction. Here and below, we will always mean by $a_n$ n-term of the sequence. For definiteness, we assume that

$$a_0 = 2^{n+1} a_n, a_1 = 2^n a_n, a_2 = 2^{n-1} a_n, ..., a_{n-1} = 2a_n, a_n, ...$$

Let us formulate some well-known results that we will use.

**Theorem 1.** Let $a_n$ be any positive integer and $\gamma_i \in \{0, 1\}$, then we can express $a$ as the sum of the powers of 2 represented by each $\gamma_i = \begin{cases} 1 & \text{(2-base number)} \\ 0 & \end{cases}$

$$a_n = \sum_{i=0}^{n} 2^i \gamma_i$$

(2)

According to the sequence generation rule, it is enough to consider the odd numbers, since even numbers will always become odd. Hence, we can assume that for any $a_0$, after the last appearance of a zero coefficient $\gamma_i \in \{0, 1\}$, the rest are not zero, as they would disappear from dividing by 2. Thus, without losing generality of our reasoning, we can assert that it suffices to consider numbers $a_n$ of the following form:

$$a_n = \sum_{i=k+2}^{n} 2^i \gamma_i + \sum_{i=0}^{k} 2^i, \ n > k > 2$$

**Theorem 2.** Let

$$a_n = 2^n + 1, \ n > 0, \ n \geq m > 0$$

and $a_{n+k}$ is generated by sequence generation rule (1) Then

$$a_{n+2m} = 3^m \cdot 4^n / 2 - m + 1$$

(3)

**Proof.** Using the rule (1) step by step we get proof.

**Theorem 3.** Let

$$a_n = \sum_{i=0}^{n} 2^i, \ n > 0, \ n \geq m > 0$$

Then

$$a_{n+2m} = 3^m 2^{n-m} - 1$$

(4)

**Proof.** Using the formula of the sum of the geometric progression, we get

$$a_n = \sum_{i=0}^{n} 2^i = 2^{n+1} - 1$$

Using the formula (1) we get

$$a_{n+1} = 3 \cdot 2^n - 3 + 1 = 3 \cdot 2^n - 2$$

$$a_{n+2} = 3 \cdot 2^n - 1, \ a_{n+2m} = 3^m 2^{n-m} - 1$$

Consider $3^u$ as $u = 12, 13, 14, 15, 25$ in binary representation,
We can expect the number of zeros in such numbers to be quite large. And if the Syracuse sequence stumbles upon these numbers, we expect it to turn back to 1. In the following section we will prove this point.

**Theorem 4.** Let

$$a_n = \sum_{i=0}^{n} 4^i$$

then

$$a_{2n+1} = 1$$

**Proof.**

$$a_{n+1} = 3 \sum_{i=1}^{n} 4^i + 4$$

$$a_{n+2} = 3 \sum_{i=0}^{n-1} 4^i + 1$$

$$a_{n+2} = 3 \sum_{i=1}^{n-1} 4^i + 4$$

$$a_{n+3} = 3 \sum_{i=0}^{n-2} 4^i + 1$$

after (n+1)-steps we have proof $\square$

Consider $a_n = \sum_{i=0}^{n} 4^i$ as $n = 10, 12, 20$ in binary representation,

$$a_{10} = (1001000100)$$

$$a_{12} = (100100100100)$$

$$a_{20} = (100100100100100100)$$

Binary representation helps to understand the idea of this work. We will demonstrate several times how the Syracuse sequence turns into a combination of powers of triples, which in turn have a lot of zeros in binary representation. The zeros, in turn, tend to return the Syracuse sequence to its initial position.

**Theorem 5.** Let

$$a_n = \sum_{i=0}^{n} \gamma_i 2^i, \; \gamma_i \in \{0, 1\}$$

$m$-is number of non-zero elements $\gamma_i$ then

$$a_{n+2m} = 3^m a_n \cdot 2^{-n-m} \tag{5}$$

**Proof.** Proof implies from theorem 3 $\square$
Theorem 6. Let
\[ a_n = 4^{n+2} - 3 \sum_{i=0}^{n} 4^{i}(-1)^{i} \]
then
\[ a_n = 4^{n+2} - 4^{n+1}(-1)^{n+1} + 1 \quad (6) \]
\[ a_{2n+1} = 3^{n+4} - 3^{n+1}(-1)^{n+1} \quad (7) \]

Proof. Applying rule (1) we get second assertion of Theorem

Theorem 7. For \( a_n \) defined by formulas
\[ a_n = \sum_{i=0}^{n} 2^{i}, \ n \geq m > 0 \]
then
\[ a_{3n} = 3^n - 1 \]

Theorem 8. Let
\[ x \in N \]
then \( \exists \)
\[ (a_1, a_2, ..., a_j) \in R^j, j \in N \]
\[ x = \sum_{i=1}^{j-1} 2^{[a_i]} + 2^{a_j} \]

Proof. Let \( j=1 \) then we can take
\[ a_1 = \log_2(x) \]
\[ x = 2^{a_1} \]
Now let \( j=2. \) Then we can take
\[ a_1 = \log_2(x), \ a_2 = \log_2(x - 2^{[a_1]}) \]
\[ 2^{a_1} = 2^{[a_1]} + 2^{a_2} \]
from which we have
\[ 2^{a_1} = \sum_{i=1}^{j-1} 2^{[a_i]} + 2^{a_j} \]
Other statement is simple.

Theorem 9. Let \( x \in N, \ [a_j] - [a_{j+1}] = \delta_j > 0 \)
\[ x = \sum_{i=1}^{j-1} 2^{[a_i]} + 2^{\delta_j} \]
\[ x = \sum_{i=1}^{j} 2^{[a_i]} + 2^{\delta_{j+1}} \]
Then

\[ \delta_j \ln 2 = -\ln \epsilon_j - \ln n2 + \epsilon_{j+1} \ln 2 + o(2^{-\delta_j + \epsilon_{j+1}}) \]

\[ 2^{\delta_j} = (2^{\epsilon_j} - 1)2^{\delta_j} \]

**Proof.**

\[ 2^{\alpha_j} = 2^{\alpha_{j+1}} + 2^{[\alpha_j]} \]

\[ 2^{\sigma_j} = 2^{-\delta_j + \epsilon_{j+1}} + 1 \]

\[ \epsilon_j \ln 2 = 2^{-\delta_j + \epsilon_{j+1}} + o(2^{-\delta_j + \epsilon_{j+1}}) \]

\[ 2^{\delta_j} = (2^{\epsilon_j} - 1)2^{\delta_j} \]

Theorem 10. Let \( x \in N, [\alpha_j] - [\alpha_{j+1}] = \delta_j > 0 \)

\[ \epsilon_1 < 1/2 \]

\[ x = \sum_{i=1}^{j-1} 2^{[\alpha_i]} + 2^{\alpha_j} \]

\[ x = \sum_{i=1}^{j} 2^{[\alpha_i]} + 2^{\alpha_{j+1}} \]

\[ \sigma_j = 1 - \epsilon_j \]

Then

as \( \delta_j = 1 \)

\[ \sigma_j^{\delta_j+1} \ln 2 = \frac{2\sigma_j \ln 2}{1 - \sigma_j^{\delta_j+1} \ln 2} + o(\sigma_j^{2} / 4) \]

and as \( \delta_j > 1 \)

\[ \sigma_j^{\delta_j+1} \ln 2 = \frac{2\sigma_j \ln 2 - \sigma_j^{\delta_j} \ln 2 + o(2\sigma_j^{2} / 4)}{[1 - \sigma_j^{\delta_j+1} \ln 2]} \]

**Proof.** From

\[ 2^{\alpha_j} = 2^{-\delta_j + \epsilon_{j+1}} + 1 \]

we can rewrite

\[ 2^{1-\sigma_j} = 2^{-\delta_j + 1 - \sigma_{j+1}} + 1 \]

after logarithmization we get

\[ \ln(2^{1-\sigma_j}) = \ln 2 - \delta_j \ln 2 \]

Computing as \( \delta_j = 1 \)

\[ \ln(2^{-\sigma_j}) = \ln((1 - \sigma_j^{\delta_j+1} + \sigma_j^{2} \ln 2 / 2) + 1) \]

\[ \ln(2 - \sigma_j^{\delta_j+1} + \sigma_j^{2} \ln 2 / 2) = \ln 2 + \ln(1 - \sigma_j^{\delta_j+1} / 2 + \sigma_j^{2} \ln 2 / 4) \]
Theorem 11. Let

\[ x \in \mathbb{N}, \quad x = \sum_{i=1}^{j-1} 2^{[a_i]} + 2^j \]
\[ \alpha_j = [\alpha_j] \]

Then

the number of zeros in the binary representation \( C_z \) is calculated by the following formula

\[ C_z = \sum_{i=1}^{j-1} [\delta_i - 1] + \alpha_j - 1 \]

Proof.

\[ C_z = \sum_{i=1}^{j-1} [\alpha_i - \alpha_{i+1} - 1] + \alpha_j - 1 \]

By definition \( \delta_i \)

\[ C_z = \sum_{i=1}^{j-1} [\delta_i - 1] + \alpha_j - 1 \]

Let introduce \( \mu_k, \nu_k \) for

\[ x = \sum_{i=0}^{n} \gamma_i 2^i \]

by following rule

\[ \gamma_k + \gamma_{k+1} = 1, \text{ and } \gamma_k + \mu_k + \gamma_{k+1} = 1, \text{ or, } \gamma_0 = 1 \text{ or } \gamma_n = 1 \text{ and } \prod_{i=1}^{\mu_k} \gamma_i = 1; \]

\[ \gamma_j + \gamma_{j+1} = 1, \text{ and } \gamma_j + \mu_j + \gamma_{j+1} = 1, \nu_j = \sum_{i=j}^{i} (1 - \gamma_i) \]

Theorem 12. Let

\[ x = 3^n = 2^{|\alpha| + \{|\alpha|\}} \]

\[ x = \sum_{i=1}^{\mu_k} \gamma_i 2^i \]

\[ \{|\alpha|\} > 1/2 \]

\[ n_\nu - \text{count of} \nu, \ n_\mu - \text{count of} \mu \]

\[ U_\mu = \{\mu_1, \mu_2, ..., \mu_n\} \quad U_\nu = \{\nu_1, \nu_2, ..., \nu_n\} \]

then \( \forall \delta_i, \exists \mu_k \text{ Such that} \)

\[ \delta_i = \mu_k - 1 \]

\( \forall \delta_\nu, \exists \mu_\mu \text{ Such that} \)

\[ i_\nu < j_\nu \quad j_{\nu+1} < i_{\nu+1} \]

\[ n_\nu = n_\mu + 1 \]

\[ n_\nu = \sum_{i=1}^{n_\mu} \mu_i + \sum_{i=1}^{n_\nu} \nu_i = \sum_{i=1}^{n_\mu} (\nu_i + \mu_i) + n_\nu \]

Proof. Proof issue from definition \( \mu, \nu \) and conditions the Theorem

Theorem 13. Let

\[ x = 3^n = 2^{|\alpha| + \{|\alpha|\}} \]
\[ x = \sum_{i=1}^{n^*} \gamma_i 2^i \]
\[ n^* = n \times \left\lfloor \frac{\ln(3)}{\ln(2)} \right\rfloor \]
\[ |\{x\} - 1/2| \leq 0.1 \tag{8} \]

then
\[ x = \sum_{i=1}^{n} \gamma_i * 2^k, \gamma_i \in \{0, 1\} \]
\[ x = \sum_{i=1}^{n} \gamma_i * 2^k, \gamma_i \in \{0, 1\} \]
\[ \sum_{\gamma_i=0} 1 \geq n^*/2 - 5 \]

**Proof.** Solving equation
\[ 3^n = 2^\alpha \]
we get
\[ \alpha = n / \ln(3) / \ln(2) \tag{9} \]
We can rewrite
\[ 3^n = 2^{[\alpha]+\{\alpha\}} \]
Using Theorem 8, we create a sequence
\[ \epsilon_i, m_i, \epsilon_1 = \{\alpha\} \]
\[ 2^{\epsilon_1} = \sum_{k=0}^{i-1} 2^{|\alpha_l|} - \alpha_1 + 2^{n_1 - \alpha_1} \]
Suppose
\[ \sum_{\gamma_i=0} 1 = 0 \]
then by Theorem 10 we have
\[ \sigma_{j+1} n_2 = \frac{2\sigma_j n_2}{1 - \sigma_{j+1} n_2} + o(\ln 2\sigma_{j+1}^2/4) \]
\[ 2^{-1} \sigma_{j+1} n_2 = \frac{\sigma_j n_2}{1 - \sigma_{j+1} n_2} + 2^{-1} \times o(\ln 2\sigma_{j+1}^2/4) \]
After repeating j times we get
\[ 2^{-j} \sigma_{j+1} n_2 = \frac{\sigma_j n_2}{\prod_{k} (1 - \sigma_{k+1} n_2/2)} + \sum_{k=0}^{j} 2^{-k} \times o(\ln 2\sigma_{k+1}^2/4) \]
By Theorems (9-10) and condition of the current Theorem proceed
\[ \ln 2/2 < \sigma_1 n_2 < o(\ln 2\sigma_{k+1}^2/4) \]
repeating, we get in case of our conditions
\[ \sum_{\gamma_i=0} 1 \geq 2 \]
Let 

\[ P_k = \prod_{1}^{k}(1 - \sigma_{k+1}ln2/2) \]

then by Theorem 10 we get 

\[ P_n 2^{-n^*} \sigma_{n^*+1} = \ln 2 \sigma_1 - \ln 2 \sum_{k=1}^{n^*} P_k 2^{-n^* + \sum_{i=k}^{n^*} (\mu_k + \nu_k)} - \ln 2 * P_{n^*} 2^{-n^* + \sum_{i=1}^{n^*} (\mu_k + \nu_k)}2^{n^*} + \ln 2 \sigma_1 + o(\max) \]

Finally we get 

\[ \ln 2 * \sigma_1 = \ln 2 * \sum_{i=1}^{n^*} (\mu_k + \nu_k)2^{n^*} - \ln 2 * P_{n^*} 2^{-n^* + \sum_{i=1}^{n^*} (\mu_k + \nu_k)}2^{n^*} + o(\max) \]

Computing \( s \), where 

\[ s = \ln 2 * \sum_{k=1}^{n^*} P_k 2^{-n^* + \sum_{i=k}^{n^*} (\mu_k + \nu_k)} + \ln 2 * P_{n^*} 2^{-n^* + \sum_{i=1}^{n^*} (\mu_k + \nu_k)}2^{n^*} + o(\max) \]

we get 

\[ n^* < \sum_{i=1}^{n^*} \mu_k < n^*/2 - 1 \]

\[ s < \ln 2 * \sum_{k=1}^{n^*/2-1} P_k 2^{-2k} \leq \ln 2 * \frac{1}{4} \frac{1}{1/4} = \ln 2/3 \]

the other side 

\[ \ln 2/2 < \sigma_1 \ln 2 < \ln 2/3 \]

immediately we get 

\[ \sum_{i=1}^{n^*/2} \mu_k > n^*/2 \]

Now by Theorem 13 me can using counting \( e_i = 1 - \sigma_i \) for counting \( \delta_i \) and then we get. 

\[ \sum_{\gamma_i=0}^{i=0} 1 \geq n^*/2 - 1 \]

\[ \Box \]

**Theorem 14. Let** 

\[ a_n = \sum_{i=0}^{n} \gamma_i 2^i, \quad n > 1000, \quad \gamma_i \in \{0,1\} \]

**then** 

\[ a_{8n} < a_n \]

**Proof.** Consider our formula for 3 in binary representation, and we see the participation of 4' creates an intermittent sequence of zeros and ones, and zero has about the same number of ones, more precisely, zeros are not less than the number of ones minus 2 and after applying Theorem 3-17 we get proof. In more detail, the estimation process consists of replacing 3' in \( a_{n+i} \) by formula 7 which does not contain powers of the triple which
allows one to evaluate the resulting terms of the Syracuse sequence. as a result, we get the following estimate. Let’s introduce operators defined formulas

\[ P f = f / 2 \]
\[ T f = 3f + 1 \]
\[ Z f = 3f \]

Let’s consider all possible scenarios of the behavior of the Syracuse sequence, the same possible scenarios can be written in the following form

\[ a_{n+1} = T_1 T_2 \ldots T_n a_n \]
\[ T_i \in \{ P, T \} \]
\[ R_i \in \{ Z, P \} \]
\[ a_{n+1} = R_1 R_2 \ldots R_n a_n + A \]

Let’s introduce

\[ m = \sum_{R_i = Z} 1 \]

and compute

\[ \sum_{R_i = P} 1 = n - m + m = n \]

By rules of Collatz we have after 2n steps

\[ a_{n+1} = 3^m / 2^n a_n + B_n \]

where

\[ B_n = \sum_{j=1}^{[a_1]} 3^{\sum_{R_i = Z, i \neq j} 1} / 2^{\sum_{R_i = P, i \neq j} 1 + \sum_{R_i = P, i = j} 1} \]

\[ B_n \leq \sum_{j=1}^{[a_1]} 3^j / 2^j < 23^n / 2^n \leq 2(3/4)^n a_n \]

\[ A = a_{2n} = 3^n (a_n * 2^{-n} + B_n) = (a_n * 2^{-n} + B_n) 3^n \]

\[ A = \sum_{i=0}^{[a_1]} \gamma_i 2^i, \quad \gamma_i \in \{0, 1\}, \quad \alpha_1 = m * \ln 3 / \ln 2 + \ln (2^{-n} a_n) \]

Let

\[ m^* \text{ is count of non zeros of } \gamma_i \]
\[ l^* \text{ is count of zeros of } \gamma_i \]

by theorem 12 we will have

\[ m^* \leq [\alpha_1] / 2 + 5 = [m \ln 3 / \ln 2] / 2 + 5 \]
\[ l^* \geq [\alpha_1 / 2 - 5 = [m \ln 3 / \ln 2] / 2 - 5 \]

After \([\alpha_1]\) steps applying rules of Collatz we have

\[ a_{2n+\alpha_1} \leq 3^m 3^{\alpha_1 / 2} 2^{-\alpha_1} 2^5 \leq 3^m q_1 * a_n \]

where

\[ q_1 = 3^m 3^{\alpha_1 / 2} 2^{-\alpha_1} 2^5 \]
Repeating the process 3 times and using \( n > 1000 \), we get
\[
q_3 < 1
\]

we get
\[
a_{8n} < a_n
\]
\[
\square
\]

**Theorem 15.** Let
\[
a_n = \sum_{i=0}^{n} \gamma_i 2^i, \quad n > 1000, \quad \gamma_i \in \{0, 1\}
\]
then for \( a_n \) Collatz conjecture is true

**Proof.** Proof follows from theorem 3-14
\[
\square
\]

5. Conclusions

Our assertion proves that after \( 2n \) of steps the sequence comes to a number less than the start one, from which follows the solution of the Collatz conjecture.

**References**


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.