Article

Not peer-reviewed version

Using Binary and Bezout Identity to Prove Collatz Conjecture

Jishe Feng *

Posted Date: 6 November 2025

doi: 10.20944/preprints202502.1743.v2

Keywords: binary string; bezout indentity; Collatz conjecture; natrual tree

Article

Using Binary and Bezout Identity to Prove Collatz Conjecture

Ji she Feng

School of Mathematics and Information Engineering, Longdong University, Qingyang, Gansu, 745000, China; E-mail: gsfjs6567@126.com.

Abstract

We propose a novel framework utilizing a full binary tree structure to systematically represent the set of natural numbers, which we classify into three subsets: pure odd numbers, pure even numbers, and mixed numbers. Within this framework, we employ a binary string representation for natural numbers and develop a comprehensive composite methodology that integrate both odd- and even-number functions. Our investigation centers on the iterative dynamics of the Collatz function and its reduced variant, which effectively serves as a pruning mechanism for the full binary tree, enabling rigorous examination of the Collatz conjecture's validity. To establish a robust foundation for this conjecture, we ingeniously incorporate binary strings into an algebraic formulation that fundamentally captures the intrinsic properties of the Collatz sequence. Through this analytical framework, we demonstrate that the sequence generated by infinite iterations of the Collatz function constitutes an eventually periodic sequence, thereby providing a discuss of this long-standing mathematical conjecture that has remained unresolved for 87 years.

Keywords: binary string; bezout indentity; Collatz conjecture; natrual tree

---

1. Introduction

In the study of number theory, odd and even numbers are a fundamental pair of ideas. The set of natural numbers can be divided into odd and even sets. There are many conjectures that attempt to generalize the fact of different kinds of natural numbers discovered in a restricted range to the entire infinite set of natural numbers. This article will discuss the famous Collatz conjecture, which states that for each natural number n, if it is even, divided by 2, if it is odd, multiplied by 3, and added 1, and so on, the eventual value must be 1. It is also referred to as the $3n+1$ conjecture and was put forth in 1937 by Lothar Collatz, also known as the $3n+1$ problem. Because it is an extremely simple to state, extremely hard to solve problem, the mathematician Paul Erdos once said of this conjecture: "Mathematics may not be ready for such problems" [1,2].

Inspired by Euler's dot-line graph in graph theory for solving the Konigsberg Seven Bridge problem, we have confidence that a similar solution can be found. Leveraging our knowledge of piecewise and iteration functions, binary strings, and full binary trees, we utilize binary strings to illustrate the step-by-step progression of odd-function and/or even-function iterations, concealed within the Decimal Number System. While the prune strategy of the full binary tree which is equivalent to the Collatz function, we are convinced that the existing framework is adequate to substantiate the conjecture.

In the natural number set, we say a pair notions even and odd, if and only if a number n has a remainder of 0 or 1 upon division by 2. The number 1 is the smallest big natural number. The result of the finite times of the iteration of even-number and /or odd-number is n which concealed the procedure of the iteration of even-number and /or odd-number from 1. For the binary string of n represents the procedure of iteration of even-number and /or odd-number from 1.

The Collatz conjecture talks about the inverse procedure of a given natural number n to 1. In binary string the function $3n+1$ is equivalent to $n+(2n+1)$ which is the add function of two binary strings, one is shifted to left one bit and appends 1 in the last bit, another one is itself. In binary string the function $\frac{n}{2}$ is to delete all zeros in right last substring. Those facts are the key opinions to discuss the Collatz conjecture.

For the Collatz conjecture, we can describe it as a piecewise function:

$$T(n)=\begin{cases} 3n+1, & \text{if } n \text{ is odd, the result must be even,} \\ \frac{n}{2} & \text{if } n \text{ is even, the result is either odd or even.} \end{cases} \tag{1}$$

The following sequence is obtained via the iteration of the Collatz function:
$\Lambda=\{n,T(n),T(T(n)),T(T(T(n))),\cdots\}=\{n,T(n),T^2(n),T^3(n),\cdots\}$.
Consequently, the Collatz conjecture can be stated as follows:

Collatz conjecture 1: For any natural number n, there is a finite natural number m, and the sequence Λ always leads to the integer 1, that is, $T^m(n)=1$.

The series Λ is an infinite sequence of ultimately period [4,5]. So we give another statement of the Collatz conjecture as the following.

Collatz conjecture 2: The series Λ is an infinite sequence of ultimately period, the preperiod $\eta(n)$ varies with the initial value n, but the ultimately period is always $\{1,4,2\}$.

2. A Graph and Algebra Representation of the Natural Numbers

Natural numbers are gradually formed in the long-term practice of human beings, which is mainly used to represent the number of things and the order of things. Arabic numbers 1,2,3,... and decimal values are the generally accepted representations all over our world. Strictly speaking, the infinite set of all natural numbers N = (0,1,2,3,...) Together with the addition (+) and multiplication $(\times)$ operations defined therein, they form the natural number system. The decimal system is very convenient in daily life, and the process of natural numbers is encapsulated by ten Arabic numbers. Gottfried Wilhelm Leibniz [11] in 1703 gave a paper talking about the binary string role, the invention of electronic computers in 1946 prompts us to re-understand and think about the meaning and function of natural numbers in binary string.

2.1. The Iteration of Odd-Number and /or Even-Number Function

A natural number is considered even if it can be divided by 2; if not, it is considered odd. According to the Peano axiom, the smallest natural number is 1. The set $N=\{1,2,3,\cdots\}$ of natural numbers can be divided into odd and even sets; in this paper, we will use the usual definition of natural numbers.

$\{$**natural number**$\}=\{$**odd number**$\}\bigcup\{$**even number**$\}$.

In the set of natural numbers where 1 is the smallest odd number and 2 is the smallest even number, we can use the function $n=2k-1$ to indicate that n it is an odd, and the function $n=2k$ to indicate that it is an even, where k is any natural number.

We introduce two functions: **odd function** $O(x)=2x+1$ and **even function** $E(x)=2x$, where x is any natural number in N.

We define a strictly increase monotonically piecewise function $f(n)$, from a natural number n it generates two cases: odd or even numbers:

$$f(n)=\begin{cases} 2n+1=O(n), & \text{the result is odd number,} \\ 2n \quad\;\, =E(n), & \text{the result is even number.} \end{cases} \tag{2}$$

Definition 1 A natural number n is obtained by finite iterations of *the odd-number function* $O(x) = 2x + 1$ or (and) *the even-number function* $E(x) = 2x$ several times, namely

$$n = f(f(\cdots f(1))) = f^k(1),$$

the function per function f is either an odd number function $O(x)$ or an even number function $E(x)$. Although there are only two simple cases, their iterative results are numerous and diverse.

For example, $f(1) = O(1) = 3, f(1) = E(1) = 2$,
$7 = f^2(1) = 2 \cdot 3 + 1 = 2 \cdot (2 \cdot 1 + 1) + 1) = O(O(1))$,
$189 = 2 \cdot 94 + 1 = 2 \cdot (2 \cdot (47) + 1 = 2 \cdot (2 \cdot (2 \cdot 23 + 1)) + 1 = 2 \cdot (2 \cdot (2 \cdot (2 \cdot 11 + 1) + 1)) + 1$
$= 2 \cdot (2 \cdot (2 \cdot (2 \cdot (2 \cdot 5 + 1) + 1) + 1)) + 1 = 2 \cdot (2 \cdot (2 \cdot (2 \cdot (2 \cdot (2 \cdot 2 + 1) + 1) + 1) + 1)) + 1$
$= 2 \cdot (2 \cdot (2 \cdot (2 \cdot (2 \cdot (2 \cdot (2 \cdot 1) + 1) + 1) + 1) + 1)) + 1 = f^7(1) = O(E(O(O(O(O(E(1))))).$

Any natural integer n is the value of the finite-time iteration function of the odd-number and even-number functions starting at 1.

Definition 2 The inverse functions $f^{-1}(x)$,

$$f^{-1}(n) = \begin{cases} \frac{n+1}{2}, & \text{If } n \text{ is odd number and } n > 1, \\ \frac{n}{2}, & \text{If } n \text{ is even number.} \end{cases} \tag{3}$$

Therefore, If $n = f(f(\cdots f(1))) = f^r(1)$, then $f^{-r}(n) = 1$.

2.2. Use the Binary String to Represent the Process of Iteration from 1

By employing binary string representations for natural numbers, we establish a more transparent framework for characterizing the iterative processes of both odd-number and even-number functions originating from 1. Within this framework, the binary string serves as an explicit encoding of the functional iteration sequence, where each bit position systematically indicates the application order of the fundamental operations $O(x)$ (odd-number function) and $E(x)$ (even-number function).

The binary string representation of a natural number encodes its corresponding sequence of odd-even functional iterations. Specifically, in this encoding scheme, each bit position from left to right corresponds to a distinct iteration step, where the appearance of a 1 in the i^{th} bit position signifies the application of the odd number function $O(x)$ at the i^{th} iteration, while a 0 in the i^{th} bit position indicates the application of the even number function $E(x)$ at that step.

To convert a given natural number n to its binary string, we can use the following iterate steps:
1. Iterate over the number n, repeatedly dividing it by 2 and keeping track of the remaining.
2. Append the remainder to the binary string from right to left.
3. Continue dividing n by 2 until it reaches 0.

For instance, the procedure of the iteration function from 1 to 60 is shown in Figure 1.

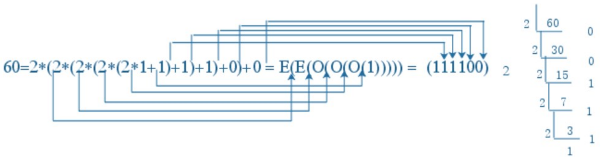


Figure 1. Natural number $60 = (111100)_2$ is obtained starting 1 through the composition of five even-number and odd-number functions.

2.3. Use an Algebra Expression to Represent the Process of the Iteration from 1

For a given natural number n, we obtain its binary string $n = (1 \times \cdots \times)_2$ and algebra expression as $n = 2^r + 2^m + \cdots$, Thus for the multiplication

$$2n = 2^{r+1} + 2^{m+1} + \cdots$$

and

$$3n+1=(2n+1)+n=2^{r+1}+2^{m+1}+\cdots+2^0+2^r+2^m+\cdots.$$

For example, for $7=(111)_2=2^2+2^1+2^0$, there are

$$2\cdot 7=(1110)_2=2^3+2^2+2^1,$$

$$3\cdot 7+1=2\cdot 7+7+1=22=(1110)_2+(111)_2=(10110)_2=2^4+2^2+2^1.$$

2.4. *Use a Graph to Represent the Process of the Iteration from 1 and a Natural Number Tree*

In order to give an intuitive impression, we provide **a full binary tree** to represent the procedure of iteration function of the odd-functions $O(x)$ and /or even-functions $E(x)$ of a given natural number n, the root is the smallest number 1. For per vertex, its left-child is an even number which double itself, its binary string is appended by 0, right-child is an odd number which double itself and add 1, its binary string is appended by 1. The full binary tree, as in Figure 2, is a very good representation of some natural numbers, we can call it as a **natural number tree** [7].

Proposition 1 A binary string's length indicates its level in the full binary tree, and a binary string's length minus one represents the iterate times of the odd- and /or even-number function.

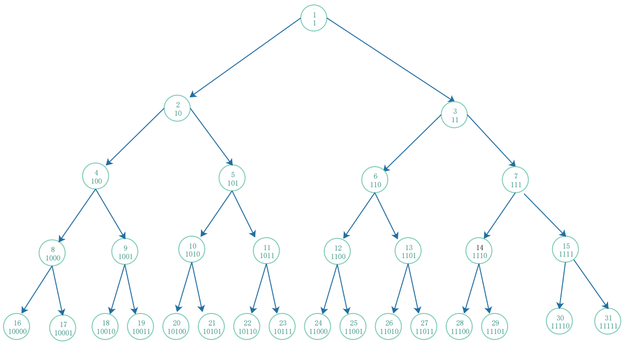


Figure 2. The representation of natural number set is a full binary tree.

For a natural number n, its binary string, composed of 0s and 1s from left to right, represents the path starting from the root node 1, and tracing down to the current node n in the full binary tree. In this tree, each node only has one path leading from the root to itself. As an illustration, the procedures for the composite odd number 21 and the even number 28 are depicted in Figure 3.

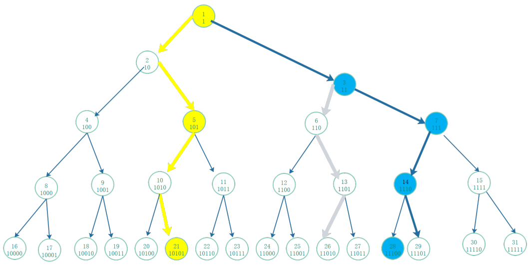


Figure 3. $21=(10101)_2$ and $28=(11100)_2$ comes from the path from root 1 walk to 10101 and 11100 accordingly appending 1 or 0 to the nodes in succession.

2.5. *Another Partition of the Natural Number Set*

To analyze the iterative functions of the odd-function and /or even-function, we introduce a new partition of the natural number set within the full binary tree. We designate distinct names to the numbers located on the left path, right path, and between the two paths in the tree.

We give the definitions of three kinds of natural number:

Definition 3(i) A natural number, $O^m(1) = 2^m - 1 = 2^{m-1} + 2^{m-2} + \cdots + 2 + 1 = (11\cdots1)_2$, is obtained by applying the odd-number function $O(x)$ m iterations. We call it as **pure odd number**. For instance, those are pure odd numbers: $3 = (11)_2, 7 = (111)_2, 15 = (1111)_2, 31 = (11111)_2, 63 = (111111)_2, \cdots$. These are located in the full binary tree of Figure 2, which is in the right path.

(ii) A natural number, $E^m(1) = 2^m = (10\cdots0)_2$, is obtained by applying the even-number function $E(x)$ m iterations. We call it as **pure even number**. For instance, those are pure even numbers: $2 = (10)_2$, $4 = (100)_2$, $8 = (1000)_2$, $16 = (10000)_2$, $32 = (100000)_2$, $64 = (1000000)_2$. Those are located in the left path of the full binary tree of the Figure 2.

(iii) The natural number obtained by the iteration of odd function $O(x)$ and /or even function $E(x)$, we call it as **mixed number**. Based on whether the digit at the end is 1 or 0, we can classify mixed numbers into **mixed even numbers** and **mixed odd numbers**. Such as, $18 = (10010)_2, 28 = (11100)_2$ are mixed even numbers, $67 = (1000011)_2, 309 = (100110101)_2$ are mixed odd numbers. Those are in the inside of the left and right paths of the full binary tree, the Figure 2.

In particular, the natural numbers obtained by the finite alternately iterations of the odd function $O(x)$ and the even function $E(x)$, namely, $[E(O(1))]^m = (101\cdots101)_2$. Such as $5 = (101)_2, 21 = (10101)_2, 85 = (1010101)_2, 341 = (101010101)_2$,
$1365 = (10101010101)_2, 5461 = (1010101010101)_2, \cdots$.
We call this mixed odd number as **hard numbers**.

Proposition 2 The set of natural numbers can be divided into three sets:
{natural number}=
{**pure even number**}∪{**pure odd number**}∪{**mixed number**},
where {**mixed number}={mixed even number**}∪{**mixed odd number**}.

Example (1) 60, 97 are mixed numbers.
(2)64, 1180591620717411303424 are pure even numbers.
(3)63, 1180591620717411303423 are pure odd numbers.

When we convert those natural numbers from decimal to binary, the facts are obvious.
(1) $60 = (111100)_2$ is a mixed even number, $97 = (1100001)_2$ is a mixed odd number.
(2) $64 = 2^6 = (1000000)$, $1180591620717411303424 = 2^{70} = (10000\ldots0)$ are pure even numbers:
(3) $63 = (111111)_2$, $1180591620717411303423 = 2^{70} - 1 = (11\ldots1)$ are pure odd numbers.

Thus for a given natural number $n = f^r(1)$, we first convert it to binary string which length is $r+1$, secondly in the level r of the natural number tree and from the top 1 along the path to the binary string. The traversal path in the full binary tree from the root down along the arcs, for each natural integer n, is its binary string $1 \times \times$, where the left-child appended 0 for each node is an even number and the right-child appended 1 for each node is an odd number. For instance, in Figure 3, $21 = (10101)_2$ originates at the root 1 and proceeds down 2,5,10, ultimately reaching 21. To the nodes, $1 \to 10 \to 101 \to 1010 \to 10101$, 0,1,0,1 are appended. In addition, for $28 = (11100)_2$, the appendix 1,1,0,0 is added to the nodes, $1 \to 11 \to 111 \to 1110 \to 11100$, accordingly in dicimal it traces from the root 1 down 3,7,14, and ultimately reaches 28.

3. Find the Traversal Path to Its Root in the Natural Number Tree

3.1. *The Iteration of Piecewise Function*

A piecewise function is a mathematical function that is defined by different rules or formulas over different intervals or regions of its domain. The Collatz function, denoted by $T(n)$, can be expressed as a piecewise function, with separate cases for odd and even numbers.

An iterative function is a function that is repeatedly applied to its own output. In other words, the output of the function is used as the input for the next iteration of the function. Iterative methods involve using iterate functions to repeatedly update an initial estimate or solution until a desired level of accuracy is achieved.

In order to proof the Collatz conjecture 1 in section 1, finding the beingness and finiteness of the number m in the expression $T^m(n) = 1$ for a natural number n is the main challenge. Iteration is the key to Collatz conjecture, and although there are only two cases where piecewise functions are combined with iterative functions, the result is difficult to control.

For given natural k, the iterative formula $f^k(1) = n$ results n, we know that k is the length of the binary string of n minus 1, it is also the level of the full binary tree. In decimal notation, we represent n, which obscures the iteration process of odd- and /or even-number functions. When n is represented as a binary string, it can be used to understand how odd- and /or even-number functions iterative work.

For given natural n, the iterative inverse formula $f^{-k}(n) = 1$ apply from two cases according party in the procedure. We can use an iterative method to search for the backtracking path of its parent node's natural tree, and eventually we will definitely return to its root node 1.

3.2. The Collatz Function and the Iterative Collatz Function

As the Collatz function $T(n)$ (1), we can get the result about the $f(x)$ and $f^{-1}(x)$

$$T(n) = \begin{cases} 3n+1 = (2n+1)+n = f(n)+n, & \text{if } n \text{ is odd number,} \\ \frac{n}{2} = f^{-1}(n) & \text{if } n \text{ is even number.} \end{cases} \tag{4}$$

The iteration of the Collatz function is the key topic in discuss the proof procedure. The iterative process of the Collatz function$T^{2k}(x)$, where k is the number of bits of the last substring of the binary string of n. When n is an odd number, although the value is increased by 1.5 times when expressed in binary form, we find that every time the Collatz function is applied twice, that is $\frac{3n+1}{2}$, although it increases by approximately 1.5 times, the length of the trailing substring will decrease by one bit with each iteration. When the length of the trailing substring is 1, the subsequent operation must be $\frac{3n+1}{2^m}, m > 1$, and the result must have $\frac{3n+1}{2^m} < n$. Then, we continue to repeat the case of odd numbers. The trailing substring of the binary string of must either be a single digit or a certain finite number of digits.

Thus, we have identified the root causes of the increase and decrease in the Collatz iteration sequence: We have discovered that the process of increase is finite, while the number of decreases, although relatively small, has a significant effect.

$$T^2(n) = \begin{cases} \frac{3n+1}{2}, & \text{n is odd number,} \\ \frac{3}{2}n+1, & \text{n is even, and n/2 is odd number,} \\ \frac{n}{2^2}, & \text{n is even, and n/2 is even number.} \end{cases} \tag{5}$$

There are many different points for piecewise functions when comparing the Collatz function $T(x)$ with the inverse function $f^{-1}(x)$, and the some iterative Collatz function $T^m(x)$ with the inverse function iteration $f^{-k}(x) = O(E(\ldots(E(x)))$.

1) For any natural number x function $f^{-1}(x)$ and $f^{-r}(x) = (f^{-1}(n))^r$ are strictly monotonically decreasing.

2) The function $T(x)$ is increasing in the case x is an odd, in the other case, is decreasing.

3) The function $T^m(x)$ that describes the procedure of the iterative function of $T(x)$, is wavy when x is a pure or mixed odd number and decreases when x is pure even or mixed even.

$T^m(n: even)$	$\frac{n}{2}$	$\frac{n}{2^2}$	$\frac{n}{2^3}$	$\frac{n}{2^4}$	$\cdots$
$T(n: even)$	$T^1(n)$	$T^2(n)$	$T^3(n)$	$T^4(n)$	$\cdots$
$monotonicity$	$\downarrow$	$\downarrow\downarrow$	$\downarrow\downarrow\downarrow$	$\downarrow\downarrow\downarrow\downarrow$	$\cdots$
$compare$	$RT(n) < n$	$RT(n) < n$	$RT(n) < n$	$RT(n) < n$	$\cdots$

$T^m(n: odd)$	$\frac{3n+1}{2}$	$\frac{3n+1}{2^2}$	$\frac{3n+1}{2^3}$	$\frac{3n+1}{2^4}$	$\cdots$
$T(n: odd)$	$T^2(n)$	$T^3(n)$	$T^4(n)$	$T^5(n)$	$\cdots$
$monotonicity$	$\uparrow\downarrow$	$\uparrow\downarrow\downarrow$	$\uparrow\downarrow\downarrow\downarrow$	$\uparrow\downarrow\downarrow\downarrow\downarrow$	$\cdots$
$compare$	$RT(n) > n$	$RT(n) < n$	$RT(n) < n$	$RT(n) < n$	$\cdots$

The function $T^m(x)$ describes the iterative procedure of $T(x)$. The wavy function is increasing at first, then goes through one or more decreasing processes, either as "increase – decrease – increase" or "increase – decrease $\cdots$ decrease – increase." For example, the iterated sequence of Collatz functions is plotted in Figures 4 and 5, where the starting values are pure odd $255 = 2^8 - 1 = (11111111)_2$ and mixed odd numbers $97 = (1100001)_2$, respectively.

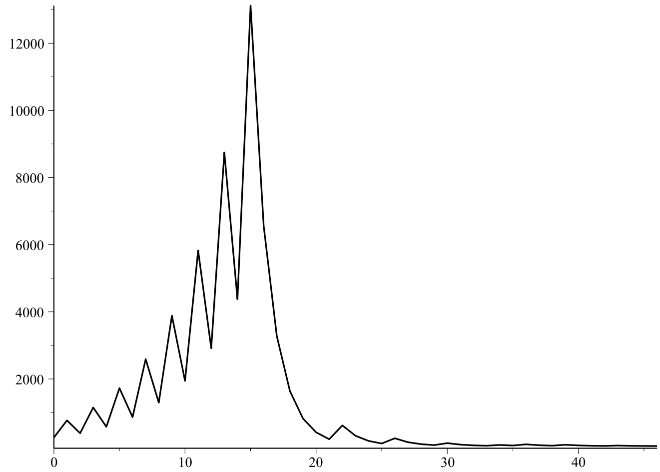


Figure 4. Point plot of a sequence of 47 iterations of the Collatz function for pure odd 255.

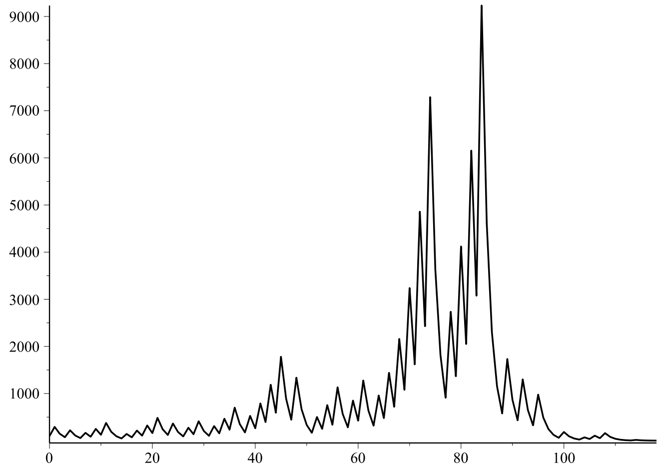


Figure 5. Point plot of a sequence of 118 iterations of the Collatz function for mixed odd 97.

For a given natural number n, the Collatz iterative function converges to 1 in a finite number of steps and cycles indefinitely between numbers 1, 4, and 2 for an infinite number of iterations.

The iterative process of the Collatz function, the expression form of $T^m(n)$ is: the numerator is a power of 3, and the denominator is the sum of powers of 2. The power of 3 in the numerator increases successively from 1 to a certain value, while the power of 2 in the numerator sometimes increases continuously and sometimes jumps. For example, $T^{6+13}(9) = 1$, where 6 times $3x + 1$ and 13 times $\frac{x}{2}$, the algebra expression in Figure 7: $T^{6+13}(9) = T^{19}(9) = \frac{1}{2^4} + \frac{3}{2^7} + \frac{3^2}{2^9} + \frac{3^3}{2^{10}} + \frac{3^4}{2^{11}} + \frac{3^5}{2^{13}} + \frac{3^6}{2^{13}} \cdot 9$.

3.3. Using Binary String to Explore the Collatz Conjecture

In the book [8] authors look at the radix 2 representations of n and $J(n)$ to solve the Josephus problem. Suppose the binary expansion is

$$n = (b_m b_{m-1} \cdots b_1 b_0)_2;$$

that is,

$$n = b_m 2^m + b_{m-1} 2^{m-1} + \cdots + b_1 2 + b_0,$$

where each b_i is either 0 or 1 and where the leading bit b_m is 1. When $n = 2^m + l$, there are

$$\begin{aligned} n &= (b_m b_{m-1} \cdots b_1 b_0)_2; \\ l &= (0 b_{m-1} \cdots b_1 b_0)_2; \\ 2l &= (b_{m-1} \cdots b_1 b_0 0)_2; \\ 2l + 1 &= (b_{m-1} \cdots b_1 b_0 1)_2; \\ J(n) &= (b_{m-1} \cdots b_1 b_0 b_m)_2; \end{aligned}$$

If we start with n and iterate the J function $m + 1$ times, we are doing $m + 1$ one-bit cyclic shifts and delete 0 in the leading bit.

We use the binary string of the natural number n to proceed the iteration of Collatz function $T^{r+1}(n) = \frac{3n+1}{2^r}$ for odd number n as the followig procedure

$$n = (b_m b_{m-1} \cdots b_1 1)_2;$$

$$\begin{aligned} 2n &= (b_m b_{m-1} \cdots b_1 10)_2; \\ 2n + 1 &= (b_m b_{m-1} \cdots b_1 11)_2; \\ 3n + 1 = n + (2n + 1) &= (b_m b_{m-1} \cdots b_1 1)_2 + (b_m b_{m-1} \cdots b_1 11)_2 \\ &= (a_k a_{k-1} \cdots b_1 0)_2; \\ \frac{3n+1}{2} &= (a_k a_{k-1} \cdots b_1)_2; \end{aligned}$$

Key point 1 We find that for odd number $n = (b_m b_{m-1} \cdots b_1 1)_2$, the Collatz iteration $\frac{3n+1}{2} = (a_k a_{k-1} \cdots b_1)_2$ reduces the length of the substring is reduced by one, the number of additional bits in the left side of the binary string does not exceed one. The value $\frac{3n+1}{2} > n$, this is the key in the series of the Collatz iteration.

Key point 2 After some iteration $\frac{3m+1}{2}$ when the length of the substring is one, the binary string $3k + 1$ must end with at least two zeros, this results that there is $\frac{3k+1}{2^t}$ where $t \geq 2$ and the value $\frac{3n+1}{2^t} < n$, this is the key in the series of the Collatz iteration. The length of the substring in the odd number's binary string, $\frac{3k+1}{2^t}$ is either 1 or r.

Key point 3 If n is an even number, i.e. a mixed even number, which end-substring has r zero, the iteration of Collatz function $T^r(n) = \frac{n}{2^r}$ is deleting all m zeros at the end of the binary string.

$$n = (b_m b_{m-1} \cdots b_1 0 \cdots 00)_2;$$

$$\frac{n}{2} = (b_m b_{m-1} \cdots b_1 0 \cdots 0)_2;$$

$$\frac{n}{2^2} = (b_m b_{m-1} \cdots b_1 0 \cdots)_2;$$

$$\cdots$$

$$\frac{n}{2^r} = (b_m b_{m-1} \cdots b_1)_2;$$

As we know that for n is an even number, the parity of $\frac{n}{2}$ is uncertain, it is either even or odd. For n is an odd number, the parity of $3n+1$ is certain, it is an even number. If we start with n and do some iterations of the T function some times, choice the next function from $\frac{T(n)}{2}$ and $3T(n)+1$ according the parity of $T(n)$. After shifting the binary string one bit to the left sequentially, fill the empty positions with 1s. For the example in Figure 9, the scratch paper of the sequence of $7 = (111)_2$

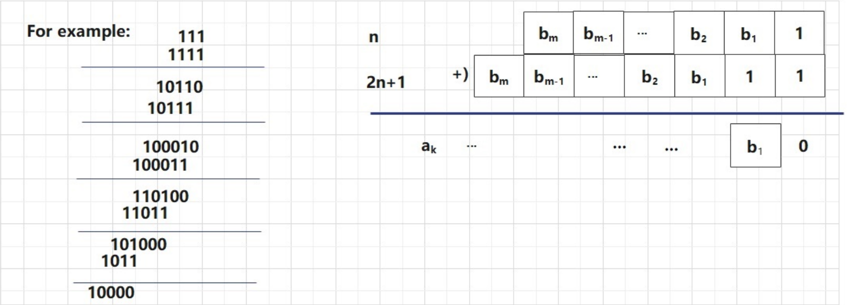


Figure 6. The scratch paper of the sequence of $7 = (111)_2$

In the iterate squence $T^k(n)$, for $t = \frac{3m+1}{2}$, if its length of the end-substring in the binary string is h which is bigger 1, then for the next odd number $\frac{3t+1}{2}$, its length of the end-substring is $h-1$. then there is a relation

$$\frac{3t+1}{2} = \frac{3\frac{3m+1}{2}+1}{2} = \frac{3^2}{2^2}m + \frac{3}{2^2} + \frac{1}{2}$$

For the general, we can get an algebraic expression. For example, there are

$$\begin{aligned} T^{91}(10027) \quad &= \quad \frac{1}{2^4} + \frac{3}{2^9} + \frac{3^2}{2^{10}} + \frac{3^3}{2^{11}} + \frac{3^4}{2^{14}} + \frac{3^5}{2^{18}} + \frac{3^6}{2^{20}} + \frac{3^7}{2^{22}} + \frac{3^8}{2^{26}} + \frac{3^9}{2^{27}} \\ &\quad + \frac{3^{10}}{2^{28}} + \frac{3^{11}}{2^{29}} + \frac{3^{12}}{2^{32}} + \frac{3^{13}}{2^{33}} + \frac{3^{14}}{2^{34}} + \frac{3^{15}}{2^{35}} + \frac{3^{16}}{2^{36}} + \frac{3^{17}}{2^{37}} + \frac{3^{18}}{2^{39}} \\ &\quad + \frac{3^{19}}{2^{42}} + \frac{3^{20}}{2^{43}} + \frac{3^{21}}{2^{44}} + \frac{3^{22}}{2^{45}} + \frac{3^{23}}{2^{48}} + \frac{3^{24}}{2^{49}} + \frac{3^{25}}{2^{53}} + \frac{3^{26}}{2^{56}} + \frac{3^{27}}{2^{58}} \\ &\quad + \frac{3^{28}}{2^{60}} + \frac{3^{29}}{2^{61}} + \frac{3^{30}}{2^{61}} \cdot 10027 \\ &= \quad 1 \end{aligned}$$

$$\begin{aligned}
T^{111}(27) &= \frac{1}{2^4}+\frac{3}{2^9}+\frac{3^2}{2^{10}}+\frac{3^3}{2^{11}}+\frac{3^4}{2^{14}}+\frac{3^5}{2^{18}}+\frac{3^6}{2^{20}}+\frac{3^7}{2^{22}}+\frac{3^8}{2^{26}}+\frac{3^9}{2^{27}}\\
&\quad+\frac{3^{10}}{2^{28}}+\frac{3^{11}}{2^{29}}+\frac{3^{12}}{2^{32}}+\frac{3^{13}}{2^{33}}+\frac{3^{14}}{2^{34}}+\frac{3^{15}}{2^{35}}+\frac{3^{16}}{2^{36}}+\frac{3^{17}}{2^{37}}+\frac{3^{18}}{2^{39}}\\
&\quad+\frac{3^{19}}{2^{40}}+\frac{3^{20}}{2^{42}}+\frac{3^{21}}{2^{43}}+\frac{3^{22}}{2^{44}}+\frac{3^{23}}{2^{47}}+\frac{3^{24}}{2^{49}}+\frac{3^{25}}{2^{50}}+\frac{3^{26}}{2^{51}}+\frac{3^{27}}{2^{52}}\\
&\quad+\frac{3^{28}}{2^{54}}+\frac{3^{29}}{2^{55}}+\frac{3^{30}}{2^{56}}+\frac{3^{31}}{2^{58}}+\frac{3^{32}}{2^{59}}\\
&\quad+\frac{3^{33}}{2^{61}}+\frac{3^{34}}{2^{63}}+\frac{3^{35}}{2^{64}}+\frac{3^{36}}{2^{65}}+\frac{3^{37}}{2^{66}}\\
&\quad+\frac{3^{38}}{2^{67}}+\frac{3^{39}}{2^{69}}+\frac{3^{40}}{2^{70}}+\frac{3^{41}}{2^{70}}\cdot 27\\
&= 1
\end{aligned}$$

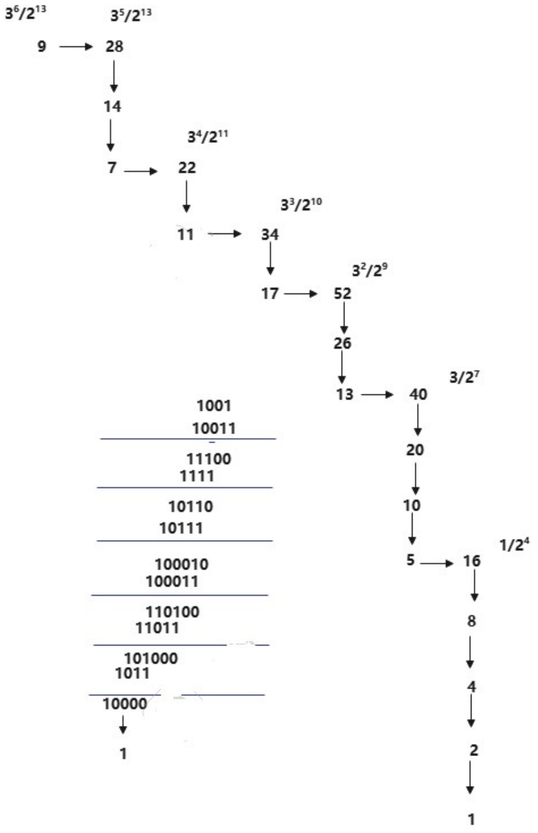
$$\begin{aligned}
T^{41}(911) &= \frac{1}{2^4}+\frac{3}{2^9}+\frac{3^2}{2^{10}}+\frac{3^3}{2^{11}}+\frac{3^4}{2^{14}}+\frac{3^5}{2^{18}}+\frac{3^6}{2^{20}}+\frac{3^7}{2^{22}}+\frac{3^8}{2^{26}}+\frac{3^9}{2^{27}}\\
&\quad+\frac{3^{10}}{2^{28}}+\frac{3^{11}}{2^{29}}+\frac{3^{12}}{2^{29}}\cdot 911\\
&= 1
\end{aligned}$$

We adopt the binary representation method for specific natural numbers and use mathematical experimental methods to obtain their Collatz sequences. The following are three kinds forms to describe the Collatz sequences respectively: (i)algebra expression, (ii)tabular, and (iii)scratch paper. We use the binary representation method for specific natural numbers and employ mathematical experimental techniques to generate their Collatz sequences. The Collatz sequences for the numbers $n = 10027, 9$ are presented in three different formats:

(A) For the formula $T^{91}(10027) = 1$, we apply the mathematical software Maple get the sequence $T^i(10027), i = 0..210$ are algebra expression, and in decimal and binary as the follows.

		$\frac{3^{30}}{2^{61}}\cdot 10027$	
$10027=(10011100101011)_2 \rightarrow$	$(111010110000010)_2 \rightarrow$	$(11101011000001)_2$	$\frac{3^{29}}{2^{61}}$
$15041=(11101011000001)_2 \rightarrow$	$(1011000001000100)_2 \rightarrow$	$(10110000010001)_2$	$\frac{3^{28}}{2^{60}}$
$11281=(10110000010001)_2 \rightarrow$	$(1000010000110100)_2 \rightarrow$	$(10000100001101)_2$	$\frac{3^{27}}{2^{58}}$
$8461=(10000100001101)_2 \rightarrow$	$(110001100101000)_2 \rightarrow$	$(110001100101)_2$	$\frac{3^{26}}{2^{56}}$
$3173=(110001100101)_2 \rightarrow$	$(10010100110000)_2 \rightarrow$	$(1001010011)_2$	$\frac{3^{25}}{2^{53}}$
$595=(1001010011)_2 \rightarrow$	$(11011111010)_2 \rightarrow$	$(1101111101)_2$	$\frac{3^{24}}{2^{49}}$
$893=(1101111101)_2 \rightarrow$	$(101001111000)_2 \rightarrow$	$(101001111)_2$	$\frac{3^{23}}{2^{48}}$
$335=(101001111)_2 \rightarrow$	$(1111101110)_2 \rightarrow$	$(111110111)_2$	$\frac{3^{22}}{2^{45}}$
$503=(111110111)_2 \rightarrow$	$(10111100110)_2 \rightarrow$	$(1011110011)_2$	$\frac{3^{21}}{2^{44}}$
$755=(1011110011)_2 \rightarrow$	$(100011011010)_2 \rightarrow$	$(10001101101)_2$	$\frac{3^{20}}{2^{43}}$
$1133=(10001101101)_2 \rightarrow$	$(110101001000)_2 \rightarrow$	$(110101001)_2$	$\frac{3^{19}}{2^{42}}$
$425=(110101001)_2 \rightarrow$	$(10011111100)_2 \rightarrow$	$(100111111)_2$	$\frac{3^{18}}{2^{39}}$
$319=(100111111)_2 \rightarrow$	$(1110111110)_2 \rightarrow$	$(111011111)_2$	$\frac{3^{17}}{2^{37}}$
$479=(111011111)_2 \rightarrow$	$(10110011110)_2 \rightarrow$	$(1011001111)_2$	$\frac{3^{16}}{2^{36}}$
$719=(1011001111)_2 \rightarrow$	$(100001101110)_2 \rightarrow$	$(10000110111)_2$	$\frac{3^{15}}{2^{35}}$
$1079=(10000110111)_2 \rightarrow$	$(110010100110)_2 \rightarrow$	$(11001010011)_2$	$\frac{3^{14}}{2^{34}}$
$1619=(11001010011)_2 \rightarrow$	$(1001011111010)_2 \rightarrow$	$(100101111101)_2$	$\frac{3^{13}}{2^{33}}$
$2429=(100101111101)_2 \rightarrow$	$(1110001111000)_2 \rightarrow$	$(1110001111)_2$	$\frac{3^{12}}{2^{32}}$
$911=(1110001111)_2 \rightarrow$	$(101010101110)_2 \rightarrow$	$(10101010111)_2$	$\frac{3^{11}}{2^{29}}$
$1367=(10101010111)_2 \rightarrow$	$(1000000000110)_2 \rightarrow$	$(100000000011)_2$	$\frac{3^{10}}{2^{28}}$
$2051=(100000000011)_2 \rightarrow$	$(1100000001010)_2 \rightarrow$	$(110000000101)_2$	$\frac{3^{9}}{2^{27}}$
$3077=(110000000101)_2 \rightarrow$	$(10010000010000)_2 \rightarrow$	$(1001000001)_2$	$\frac{3^{8}}{2^{26}}$
$577=(1001000001)_2 \rightarrow$	$(11011000100)_2 \rightarrow$	$(110110001)_2$	$\frac{3^{7}}{2^{22}}$
$433=(110110001)_2 \rightarrow$	$(10100010100)_2 \rightarrow$	$(101000101)_2$	$\frac{3^{6}}{2^{20}}$
$325=(101000101)_2 \rightarrow$	$(1111010000)_2 \rightarrow$	$(111101)_2$	$\frac{3^{5}}{2^{18}}$
$61=(111101)_2 \rightarrow$	$(10111000)_2 \rightarrow$	$(10111)_2$	$\frac{3^{4}}{2^{14}}$
$23=(10111)_2 \rightarrow$	$(1000110)_2 \rightarrow$	$(100011)_2$	$\frac{3^{3}}{2^{11}}$
$35=(100011)_2 \rightarrow$	$(1101010)_2 \rightarrow$	$(110101)_2$	$\frac{3^{2}}{2^{10}}$
$53=(110101)_2 \rightarrow$	$(10100000)_2 \rightarrow$	$(101)_2$	$\frac{3}{2^{9}}$
$5=(101)_2 \rightarrow$	$(10000)_2 \rightarrow$	$(1)_2$	$\frac{1}{2^{4}}$

(B) For the formula $T^{19}(9) = 1$, we get the sequence $T^i(9), i = 0..18$ are in decimal and binary as the following scratch paper as in Figure 7. The horizontal arrow line means the 3n+1 operation, and Vertical arrow line meams the n/2 operation. There 6 $3n + 1$ denoted by 3^i where $i = 0..6$, and 13 $n/2$ by 2^j where $j = 4..13$ in algebra expression.



$$\frac{1}{2^4}+\frac{3}{2^7}+\frac{3^2}{2^9}+\frac{3^3}{2^{10}}+\frac{3^4}{2^{11}}$$
$$+\frac{3^5}{2^{13}}+\frac{3^6}{2^{13}}\cdot 9=1$$

Figure 7. The scratch paper of a sequence of 9 iterations of the Collatz function for odd 9 in decimal and binary forms, and its algebra expression.

We observe the procedure of iterative Collatz function, namely the reduced Collatz function (7), i.e., the Collatz sequences, which we pay close attention to the zeros in the right-hand side of binary strings of an even number and the *end-substring* between the first 0 encountered from right to left which is made of 1. For instance, for 1011001 the end-substring is 1, for 1011001111 the end-substring is 1111, for 11111 the end-substring is itself 11111. There are many properties of the end substrings.

In the Collatz sequence represented by a binary string looking backward from 9:

(1) If there are several zeros at the end, remove one at a time until all zeros are deleted, and the number becomes an odd, namely $\frac{x}{2^r} < x$, and in the algebra expression $\frac{1}{2^4}$.

(2) When the number of bits at the end-substring is $r(r > 1)$, the adjacent binary string must have only a single zero at the end. Remove this zero to make it the next odd number, and the number of bits at the end-substring is $r - 1$, continues these steps until only a 1 is left at the end. Namely, $\frac{3x+1}{2} > x$, $\cdots$, and in the algebra expression $\frac{3^2}{2^9} + \frac{3^3}{2^{10}} + \frac{3^4}{2^{11}}$.

(3) If the number of digits at the end-substring is only one, the adjacent binary string must end with several zeros. Deleting these zeros in sequence will result in the following two scenarios at the end of the binary string:

(i) When the length of the end-substring is one bit 1, then the adjacent Collatz sequent is $\frac{3x+1}{2^k} < x$, where $k > 1$, accordingly the Collatz sequent is decrease.

(ii) when the length of the end-substring is more than one bit 1, then the adjacent Collatz sequent is $\frac{3x+1}{2} > x$, accordingly the Collatz sequent is increase.

From this, we can obtain the following result: For any positive odd number, the number of bits in the trailing substring of its binary representation is necessarily finite, and thus the increasing process in its Collatz sequence is also finite. After several increments, it will eventually decrease.

3.4. Bezout Identity and Its Extension

In order to explain the algebraic expression, we associate with Bezouts identity, a fundamental theorem in number theory established by Etienne Bezout in 1779. Bezout's Identity, also known as Bezout's Lemma, is a fundamental theorem in number theory that describes a linear relationship between the greatest common divisor (GCD) of two integers and the integers themselves.

Bezout's Identity (Bezout's lemma) states that for any two integers a and b, there exists x and y such that:

$$ax + by = gcd(a, b)$$

where $gcd(a, b)$ is the greatest common divisor of a and b.

The extension of the Bezout identity is stated as the following, For some integers $a_1, a_2, \cdots, a_n$, if $\gcd(a_1, a_2, \cdots, a_n) = 1$, then there are infinite many integers $x_1, x_2, \cdots, x_n$ such that

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 1.$$

For $gcd(3, 2) = 1$, there are positive integers m, n, such that $gcd(3^m, 2^n) = 1$, $gcd(3^m, 3^{m-1}, \cdots, 3, 1) = 1$, $gcd(2k+1, 1, 2^m, 2^r, \cdots, 2^m) = 1$ and $gcd(2k+1, 2^m, 2^r, \cdots, 2^m) = 1$.

We get another extension of Bezout identity as the following:

Another extension of Bezout identity: For some natural n, there are two natural numbers m and q, such that there is a linear combination expression of $gcd(3^m, 3^{m-1}, \cdots, 3, 1) = 1$ and vector $(2k+1, 1, 2^m, 2^r, \cdots, 2^m)$ and $2k+1, 2^m, 2^r, \cdots, 2^m)$.

$$3^m \cdot n + 3^{m-1} \cdot 2^0 + \cdots + 3 \cdot 2^m + 2^r = 2^q.$$

or

$$3^m \cdot n + 3^{m-1} \cdot 2^g + \cdots + 3 \cdot 2^m + 2^r = 2^q.$$

Namely, there is the algebraic expression

$$T^{m+q}(n) = \frac{3^m}{2^q} \cdot n + \frac{3^{m-1}}{2^{q-h}} + \cdots + \frac{1}{2^{q-r}} = 1$$

which is dividing the left side of the equation by the right side. For example, the above algebraic expressions, $T^{91}(10027)$, $T^{111}(27)$, $T^{41}(911)$ and $T^{19}(9)$.

4. Two Methods to Talk About the Proof of the Collatz Conjecture

We give a constructive proof. For a gave natural number n, we firstly convert it to binary string, next based on whether the end is 1 or 0, perform the iteration of the following two operations: One is when the last bit is 1, firstly we append 1 to the last bit, and add with itself, the end of the obtained binary string has at least one zero. Next, we erase all the zeros at the end. One is when the last substring is zero, we delete all zeros at last and get the last bit is 1.

For any positive integer n, when it is a pure even, its binary string only has two parts, lead bit is 1, right are all 0, we delete all zeroes, get 1. when it is a pure odd, its binary string has not zero, all 1. Its result of the Collatz faction must be a mixed even number. Thus we always assume natural number be a mixed odd number.

(i) For a special class of mixed odd numbers, the hard number $\frac{4^k-1}{3} = (101\cdots101)_2$, then its Collatz sequent result is

$$a_k = \frac{4^k-1}{4-1} = 4^{k-1} + 4^{k-2} + \cdots + 4 + 1 = (101\cdots101\cdots101)_2,$$
$$T(a_k) = 3a_k + 1 = 4^k = 2^{2k} = (10\cdots0)_2, T^{2k+1}(a_k) = 1.$$

This means that the Collatz conjecture is valid for this case.

(ii) For a general mixed odd number n, its Collatz sequent has the following feature:

(1) The length of the trailing substring is r, the adjacent binary string must have only a single zero at the end. Remove this zero to make it the next odd number which, and the number of bits at the end-substring is $r-1$, continues these steps until only a 1 is left at the end.

(2) When the length of the trailing substring is 1, the next Collatz sequent number is smaller, and length of the trailing substring (a) only one 1, (b) some many 1.

4.1. Bezout Identity

Another extension of Bezout identity: For some natural n, there are two natural numbers m and q, such that there is a linear combination expression of $gcd(3^m, 3^{m-1}, \cdots, 3, 1) = 1$ and vector $(2k+1, 1, 2^m, 2^r, \cdots, 2^m)$ and $2k+1, 2^m, 2^r, \cdots, 2^m)$.

$$3^m \cdot n + 3^{m-1} \cdot 2^0 + \cdots + 3 \cdot 2^m + 2^r = 2^q.$$

or

$$3^m \cdot n + 3^{m-1} \cdot 2^g + \cdots + 3 \cdot 2^m + 2^r = 2^q.$$

Namely, there is the algebraic expression

$$T^{m+q}(n) = \frac{3^m}{2^q} \cdot n + \frac{3^{m-1}}{2^{q-h}} + \cdots + \frac{1}{2^{q-r}} = 1$$

4.2. The Traversal Path in the Natural Tree

From our real-life experience, we know that when we need to cut down a huge tree, because it is too large, we can only cut off the manageable branches step by step, and finally reach the root to remove the entire tree.

We use the iteration of the Collatz function on a complete binary tree as the pruning algorithm. If a node can be visited, then the subsequent nodes reached through the iteration of the Collatz function will definitely not be on the subtree starting from this node.

Pruning algorithm Starting from one's own node on the natural tree, through the iteration of the Collatz function, each time moving to the next node, the standing branches (subtrees) of oneself are cut off.

Funding: Foundation item: Educational technology innovation project of Gansu Province (No. 2022A-133); and "The Practice and research of Mathematics Teaching Reform and Open Teaching in primary and secondary schools" (Longdong University Horizontal Project: 901324030218).

References

1. Lagarias, J. C. (Ed.). (2010). *The Ultimate Challenge: The* $3x+1$ *Problem*. American Mathematical Society: Providence, RI, USA, 2010; pp. xiv+344, ISBN 978-0-8218-4940-8.
2. Kaufman, R. A reduced forward Collatz algorithm: How binary strings change their length under 3x+1. *arXiv* 2023, *arXiv:2301.07466*.
3. Sternberg, L. Predictable trajectories of the reduced Collatz iteration and a possible pathway to the proof of the Collatz conjecture. *arXiv* 2022, *arXiv:2209.14230*.

4. Ganesan, G. (2021). Linear recurrences over a finite field with exactly two periods. *Advances in Applied Mathematics*, 127, 102-108.
5. Quijada, D. (2015). *Periods of linearly recurring sequences*. Bachelor thesis, Washington and Lee University.
6. Sukusu, B. (2023). Proof of the Collatz Conjecture. *Theoretical and Mathematical Applications*, 13(3), 1-17.
7. Feng, J. (2023). Proof the Collatz Conjecture by a new view of natural numbers. DOI: 10.20944/preprints202309.1200v1.
8. Graham, R. L., Knuth, D. E., & Patashnik, O. (1994). *Concrete Mathematics: A Foundation for Computer Science* (2nd ed.). Addison-Wesley Professional.
9. Li, T.-Y., & Yorke, J. A. (1975). Period Three Implies Chaos. *The American Mathematical Monthly*, 82(10), 985-992.
10. Sharkovsky, A. N. (1964). Coexistence of Cycles of Continuous Mapping of the Line into Itself. *Ukrainian Mathematical Journal*, 16, 61-71.
11. Année. Histoire de l'Academie Royale des Sciences. Paris, 1705; pp. 85–89.