A proof of the Collatz (3x+1) conjecture

Xingyuan Zhang

(Independent scholar, Suzhou, P. R. China, 215131, e-mail:502553424@qq.com)

Contents

Abstract In this paper we had given an elementary proof of the Collatz conjecture, it holds. By detailed analysis of the properties of both forward and inverse operations of the proposition, we had some important conclusions: 1, there are no cycles except 1 to 1, and for a given odd it either goes to infinity or returns to 1 in forward operations; 2, there hasn't any triple in the forward path numbers; 3, there have an infinity number of inverse path numbers which had been defined as similar numbers between them in one time of inverse operation; 4, to do inverse operations (defined as reverse tracing) repeatedly from the odd 1, it will obtain all of the odds; 5, for any odd obtained by tracing, to do forward operations, it must return to 1 along the reverse tracing path.

Key words and phrases conjecture, tracing, path number, similar number, source number

1. Introductions and statements

The Collatz conjecture is the 3x+1 conjecture, also known as Kakutani's problem. It has not been proved since it was proposed [1]. Its operation rules are: for any given positive integer n, if even, then to $n/2$; if odd, then to $3n+1$. To do it repeatedly, n will eventually return to 1.

In this paper, we called Collatz conjecture as Collatz proposition, or proposition for short.

According to the operation rules, an even number will be transformed into an odd firstly, so we take odds directly to analyze for the operations and use positive integers when counting.

2. Operational rules and analysis of the operation properties

In operations, there are many new odds and they form an operation path.

Definition 1

(a) The operation process from an odd to a new odd is called one time of operation; times of operations are called continuous operations; in one operation, divided by 2 is called one time of local operation;

(b) The new odd obtained after one operation is called a path number;

(c) One operation done in the order of the proposition is called one time of forward operation.

Next, we give the formula and analyze the properties of the forward operations.

2.1 The operational formula

For any given odd number n, let p be its path number, then according to the operation rules, we have

$$p = \frac{3n+1}{2^k} \tag{2.1}$$

Where $k \in N$ and 2^k is a divisor.

Here, we called formula (2.1) as the forward operation formula of the proposition. To analysis formula(2.1) , it is not difficult to obtain: for any given odd number n, there has only one path number p corresponding to n; the value of k is determined by the odd number n, k can be expressed as the times of divided by 2, for an example, when it equals to 1, it means in one operation, that there is one local operation, and when it equals to 3, there are three local operations; for two different odds, the local operations is the same or different because the value of k in the divisor can take all of the positive integers, therefore, there may be infinite odd numbers that they will all get the same path number after one operation, and these new odd numbers may be some correlation properties with each other.

2.2 The properties of the path numbers

2.2.1 Numerical comparison of n and p

Suppose $p = n$, then from formula (2.1) we have

$$n = \frac{1}{2^k - 3}. \tag{2.2}$$

It can be seen that equation (2.2) has only one positive integer solution 1 when k is equal to 2. From this, we can draw a conclusion (conclusion (1)): for the odd number 1, its path number is equal to itself; for any given positive integer n except 1, its path number p isn't equal to n itself, that is $p \neq n$.

Thus it can be seen that there has only one cycle 1-4-2-1 of which n to be taken 1 in one time of operation. If get 1, we stop to do operations.

In a numerical size relationship from formula (2.1), when $n \neq 1$, we can get what as fallows

If, $k = 1 \quad p > n$

If, $k = 2 \quad p < n$

If, $k \geq 3 \quad p < n$

Obviously, the larger k is, the greater the change rate of the path number is.

Here, for an odd, its path number becomes smaller quickly or even goes to 1 directly when $k \geq 3$. It's to do with what whether all odds can go back to 1 and that is what we are going to research in this paper.

2.2.2 The triples

An odd number n can be expressed as

$$n = 2x + 1.$$

Where $x = 0$ or $x \in N$.

To do one operation for n, suppose we can get a path number $3p$, where p is an odd, then from formula (2.1) we have the following equation

$$3p = \frac{3(2x+1)+1}{2^k} = \frac{6x+4}{2^k}.$$

To simply, then we have

$$x + \frac{2}{3} = 2^{k-1} p. \qquad (2.3)$$

Obviously, the equation (2.3) doesn't hold for integers, so we can get the following conclusion (conclusion (2)): the path number is not a triple, but a non-triple; these triples were skipped in operations.

Thus in operations, only the starting point can be a triple.

2.2.3 Changes of the values of two adjacent odds

Let n be an odd and expressed as $2x+1$, where $x = 0$ or $x \in N$, thus, one of its adjacent odd numbers can be expressed as

$$2x+1+2.$$

To do operation for $2x+1$, and the divisor is used 2 firstly, then we get a middle number as fallow

$$\frac{3(2x+1)+1}{2} = 3x+2.$$

To do operation for $2x+1+2$, then we also get a middle number as fallow

$$\frac{3(2x+1+2)+1}{2} = 3x+5.$$

Obviously, in these two numbers above, one is odd and the other is even. They both increase firstly, since the even can be divided by 2 again, so it will decrease finally. From this, we can get a conclusion (conclusion (3)): for two path numbers of two adjacent odds, if one becomes larger, the other must become smaller.

2.3 Narrow paths

Definition 2 Let n be an odd, to do operations continuously for it, in every operation process, if the times of local operations doesn't exceed 2, i.e., $k \le 2$, then we called this section of the operations path as a narrow path which is composed of n and its path numbers.

On the narrow path, the numerical change rate of the path numbers is the smallest, that is, the range of changes is the narrowest.

A complete narrow path has a starting point and an ending point, which we will discuss further below (at section 6.1).

3. The similar numbers and their properties

From the analysis at 2.1, it's known that the same path number can be obtained when doing one operation for two different odd numbers respectively. For example, if doing one operation for 7 and 29, they both get 11. For 7, the value of k in formula (2.1) is 1, and for 29, the value of k is 3.

Definition 3 Suppose, there are two odd numbers n_1 and n_2 which their path numbers are both p, then, we called that n_1 is a similar number of n_2, or n_2 is a similar number of n_1, that is, they are similar each other, and denoted $n_1 \backsim n_2$, or, $n_2 \backsim n_1$.

For example, 29 is a similar number of 7, or 7 is a similar number of 29, i.e., 29$\backsim$7.

Obviously, the similar numbers are caused by different values of k.

Next, we analyze the properties of the similar numbers.

3.1 The relationship between similar numbers

Suppose, there are two similar numbers n_1 and n_2, where $n_2 > n_1$, to do one operation on each of them, we can get the path numbers p_1 and p_2. According to formula (2.1), we have

$$p_1 = \frac{3n_1 + 1}{2^{k_1}}$$

And

$$p_2 = \frac{3n_2 + 1}{2^{k_2}}.$$

Where $k_1 \in N$, and $k_2 \in N$.

Now, let $p_1 = p_2$, then we have

$$\frac{3n_1 + 1}{2^{k_1}} = \frac{3n_2 + 1}{2^{k_2}}. \quad (3.1)$$

Since $n_2 > n_1$, we can get $k_2 > k_1$, that is, $k_2 - k_1$ are positive integers.

From equation (3.1), n_2 can be obtained, that is

$$n_2 = \frac{2^{k_2}}{2^{k_1}} n_1 + \frac{1}{3}\left(\frac{2^{k_2}}{2^{k_1}} - 1\right) = 2^{k_2 - k_1} n_1 + \frac{1}{3}\left(2^{k_2 - k_1} - 1\right). \quad (3.2)$$

Obviously, for equation (3.2), if n_2 to be an integer, $2^{k_2 - k_1} - 1$ must be a triple, and there is a minimum value 3 in triples and at this time $2^{k_2 - k_1}$ takes 4.

Now, let $k_2 - k_1 = 2t$, i.e., $k_2 - k_1$ takes evens, where $t \in N$, thus we have

$$2^{k_2 - k_1} - 1 = 2^{2t} - 1 = (2^t + 1)(2^t - 1). \quad (3.3)$$

As it can be seen from equation (3.3) that there is a triple in three continuous positive integers of $2^t - 1$, 2^t and $2^t + 1$. For 2^t, if t is even, $2^t - 1$ is a triple and we can derive that $2^t + 1$ is not a triple (the smallest gap of two triples is 3); if t is odd, $2^t + 1$ is a triple, and we can also derive that $2^t - 1$ is not a triple, that is, when $k_2 - k_1$ takes odds, $2^{k_2 - k_1} - 1$ hasn't any triple factor. Thus $k_2 - k_1$ must take even numbers ($k_2 \geq 3$), that is, $2^{k_2 - k_1}$ must take a power of 4, and then n_2 has integer solutions in equation (3.2).

Now, we analyze some cases with 4 and its 4 multiples to find some similar numbers respectively,

a) when $2^{k_2-k_1}$ takes 4, we have the second (n_1 is the first)

$$n_2 = 4n_1 + 1$$

b) when $2^{k_2-k_1}$ takes 16, we have the third

$$n_3 = 16n_1 + 5 = 4(4n_1 + 1) + 1 = 4^2 n_1 + 4^1 + 4^0$$

c) when $2^{k_2-k_1}$ takes 64, we have the fourth

$$n_4 = 64n_1 + 21 = 4\left[4(4n_1 + 1) + 1\right] + 1 = 4^3 n_1 + 4^2 + 4^1 + 4^0 .$$

Here, we had got three similar numbers of n_1 in turn. As it can be seen that the formula above is an iterative formula, thus more generally, we can deduce an iterative formula as follow

$$n_{1+i} = 4^i n_1 + \sum_{j=1}^{i} 4^{j-1} . \tag{3.4}$$

Where n_1 takes odds, and i and j takes positive integers. Using formula (3.4), we can get an infinite number of similar numbers of n_1 .

As examples, we can find some similar numbers of the original few odd numbers.

i. Let n_1 =1, then from (3.4) we have

$$n_{1+i} = 4^i + \sum_{j=1}^{i} 4^{j-1} . \tag{3.5}$$

From (3.5) we can obtain a sequence generated by 1 as follow

1, 5, 21, 85, 341⋯

ii. Let n_1 =3, then we can again obtain a sequence generated by 3 as follow

3, 13, 53, 213, 853⋯

iii. Let n_1 =7, then we can also obtain a sequence generated by 7 as follow

7, 29, 117, 469, 1877⋯

When n_1 =5, the sequence generated by 5 is already in the first sequence.

It can be seen that when the gap between two similar numbers is the smallest, the relationship between two similar numbers is as fallow

$$n_2 = 4n_1 + 1 \tag{3.6}$$

Here, we called formula (3.6) as the formula of the similar numbers, and also, two similar numbers when they have the smallest gap between them as the adjacent similar numbers. By using formula (3.6), we can also find out the numberless similar numbers of n_1 one by one and it's easy to do. We use only this formula in this paper. The similar numbers of any odd number can be found one-timely if we use the iterative formula (3.4).

This relationship can be verified by doing operations for n_1 and n_2 separately.

a) For n_1, we have the path number

$$\frac{3n_1+1}{2^k}.$$

b) For n_2, we have a middle even number

$$\frac{3(4n_1+1)+1}{2^k}=4\left(\frac{3n_1+1}{2^k}\right).$$

Obviously, the number on the right above can be divided by 4 again, thus we also have a path number

$$\frac{3n_1+1}{2^k}.$$

As it can be seen that when we do operations for n_1 and n_2 respectively, we get the same path number, so they are similar numbers to each other.

To do inverse operation for formula (3.6), then we have

$$n_1=\frac{n_2-1}{4}\quad, \tag{3.7}$$

where $n_2 \geq 5$, obviously, if n_1 to be a positive integer and an odd, then it is an adjacent similar number of n_2, and $n_1 < n_2$. Here Formula (3.7) is called the inverse operation formula of similar numbers.

3.2 The sets of similar numbers

It is obvious from formula (3.6) that any odd number can generate an infinite number of similar numbers in turn.

Definition 4

(a) Suppose, the similar numbers generated by the odd number n_1 in turn are n_2, n_3, $\cdots$, n_i, where $i \in N$ and $i \geq 2$, then we called the infinite set composed of n_1 and n_i as an infinite set of similar numbers, or a set for short;

(b) We called n_i as the previous similar number of n_{i+1}, and n_{i+1} as the next similar number of n_i;

(c) We called n_1 as the generating number of a set (the first); called the set as number n_1 set.

For examples, when the generating number is 1, it can generate 5, 21, 85, 341, 1365⋯an infinite number of similar numbers (see section 3.1), 1 and all of its similar numbers constitute a set, this set is called number 1 set, in which any similar number returns directly to 1 after one operation; 1 is the path number of itself and also the rests, so it's the operational value of number 1 set. In the same way, 3 can generate similar numbers such as 13, 53, 213, 853⋯ (see section 3.1), they constitute the number 3 set. In a set, all odds constitute an increased sequence. Specifically, if every odd in number 1 set be added with 3 to the right in turn, then number 1 set becomes

number 3 set (missed the first).
For understanding, there is a table of sets of similar numbers below (see Tab. 1 Table of sets of similar numbers).

Tab. 1 Table of sets of similar numbers

n\c	Path numb.	Comp.	Set numb.	Similar numbers (arranged small to large)						
				1	2	3	4	5	6	7
1	1	=	1	1	5	21	85	341	1365	5461
2	5	＞	3	3	13	53	213	853	3413	13653
3	1	＜	5	5	1	1	1	1	1	1
4	11	＞	7	7	29	117	469	1877	7509	30037
5	7	＜	9	9	37	149	597	2389	9557	38229
6	17	＞	11	11	45	181	725	2901	11605	46421
7	5	＜	13	13	3	3	3	3	3	3
8	23	＞	15	15	61	245	981	3925	15701	62805
9	13	＜	17	17	69	277	1109	4437	17749	70997
10	29	＞	19	19	77	309	1237	4949	19797	79189
11	1	＜	21	21	5	5	5	5	5	5
12	35	＞	23	23	93	373	1493	5973	23893	95573
13	19	＜	25	25	101	405	1621	6485	25941	103765
14	41	＞	27	27	109	437	1749	6997	27989	111957
15	11	＜	29	29	7	7	7	7	7	7
16	47	＞	31	31	125	501	2005	8021	32085	128341
17	25	＜	33	33	133	533	2133	8533	34133	136533
18	53	＞	35	35	141	565	2261	9045	36181	368641
19	7	＜	37	37	9	9	9	9	9	9
20	59	＞	39	39	157	629	2517	10069	40277	161109
21	31	＜	41	41	165	661	2645	10581	42325	169301
22	65	＞	43	43	173	693	2773	11093	44373	177493
			……							
				+2	+8	+32	+128	+512	+2048	+8192

notes

1. Each row is a set of similar numbers (only 7 odds), the first similar number of each set is the successive odd numbers starting from 1, which is the generating number of these set, and each set takes this odd number as the number of the set;
2. Comparison represents the relationship between the generating number and the first path number, and the size is continuously distributed in pairs (conclusion 3); the path number is also the operational value of the set;
3. The rows of 5, 13, 21, 29⋯are subsets (underlined); from the second in a row, the odds are all the previous similar number of the first, thus each column contains any odd (some repeated, such as 5 in the second column);
4. The triples are shaded; in a set, there are two non-triples between two triples (see section 3.3).

3.3 The smallest gap between two triples in a set

In a continuous series of odd numbers, two adjacent triples are separated by two non-triples. 3 is the smallest triple, and 3 added by 6 every time, it becomes another triple (odd). Similarly, in a set, two adjacent triples are separated by two non-triples.

To verify as follows:

A triple can be expressed as $3n$, where n is an odd, so according to formula (3.6) the following three similar numbers are as follows in turn:

The first $4(3n)+1=12n+1=3(4n)+1$

The second $4(12n+1)+1=48n+5=3(16n+2)-1$

The third $4(48n+5)+1=192n+21=3(64n+7)$

As it can be seen, that the first and the second are both not triples, and the third is exactly a triple.

3.4 The effects of the similar numbers in operations

In the continuous operations, when a path number has a similar number smaller than itself, the next path number will quickly become smaller, or even directly back to 1. For example, the number 853 (in Tab. 1, row 2) is similar to 3, which returns to 5 quickly after one operation; the number 1365 (in Tab. 1, row 1) is similar to 1, which returns to 1 directly. For operations, if A∽B and B is bigger than A, to do one time of operation for B is that for A, such as, for 1365 is that for 1. Those numbers from the second in a set are the ending-numbers of a narrow path (see definition 2, the times of local operations exceeds 2).

It is not difficult to see that the similarity existing in odds is of great significance in judging this proposition.

4. The analysis of cycles in forward operations

In the continuous series of odd numbers, we divided every four odds into a group starting with 1, and a group is marked g_t, where $t \in N$, thus the expressions of four odds in g_t are as follows,

$$n_1 = 1+8(t-1) = 8t-7 \tag{4.1}$$

$$n_2 = 3+8(t-1) = 8t-5 \tag{4.2}$$

$$n_3 = 5+8(t-1) = 8t-3 \tag{4.3}$$

$$n_4 = 7+8(t-1) = 8t-1 \tag{4.4}$$

Here, we can get the odds in a group, such as 1, 3, 5 and 7 in g_1, and 9, 11, 13 and 15 in g_2, and so on. According to the formula (2.1), the path number p_1 of n_1 is as follow

$$p_1 = \frac{3n_1+1}{2^k} = \frac{3(8t-7)+1}{2^k} = \frac{24t-20}{2^k} = \frac{6t-5}{2^{k-2}} \quad . \tag{4.5}$$

As the molecules is an odd in equation (4.5), thus if p_1 to be an integer and an odd, then k must take 2, so we get

$$p_1 = 6t-5 \, . \tag{4.6}$$

Obviously, n_1 has decreased after one time of operation.

As the same,

$$p_2 = \frac{3n_2+1}{2^k} = \frac{3(8t-5)+1}{2^k} = \frac{24t-14}{2^k} = \frac{12t-7}{2^{k-1}}. \tag{4.7}$$

That is

$$p_2 = 12t-7. \tag{4.8}$$

Obviously, n_2 has increased after one time of operation.

And for p_4, we also have

$$p_4 = \frac{3n_4+1}{2^k} = \frac{3(8t-1)+1}{2^k} = \frac{24t-2}{2^k} = \frac{12t-1}{2^{k-1}}. \tag{4.9}$$

That is

$$p_4 = 12t-1 \tag{4.8}$$

Obviously, n_4 has also increased.

Now, we analyze n_3, and it can be seen that it has similar numbers less than it. Here is the proof. Let its previous similar number be n_{3s}, according to formula (3.7), and then we have

$$n_{3s} = \frac{n_3-1}{4} = \frac{8t-3-1}{4} = 2t-1. \tag{4.9}$$

It can be seen that equation (4.9) had got any odd, and for a part of odds, they also have smaller similar numbers, for example, the third odd 21 in g_3 has two smaller similar numbers 5 and 1 (see Tab. 1).

Here, the path number p_3 of n_{3s} is as follow from formula (2.1)

$$p_3 = \frac{3n_{3s}+1}{2^k} = \frac{3(2t-1)+1}{2^k} = \frac{6t-2}{2^k} = \frac{3t-1}{2^{k-1}}. \tag{4.10}$$

For equation (4.10), when t takes evens and as analyzed for equation (4.5), we can get

$$p_3 = 3t-1. \tag{4.11}$$

And when t takes odds, the right side of equation (4.10) has another local operation and even more, therefore p_3 will again become smaller, so $3t-1$ can be regarded as the maximum value of p_3 when t takes odds. The aim of what we do here is for analyzing below the gap between n_3 and p_3 after taking the similar number n_{3s} of n_3. Obviously, n_3 has decreased after one time of operation, that is, there is a distance between n_3 and p_3. We will analyze under what conditions that p_3 is equal to n_3 after taking similar numbers.

Next, we analyze the cycles in continuous forward operations. Firstly we take n_1 in g_t to analyze under what conditions that the path numbers of n_1 is equal to it. After one time of operation, its path number p_1 is in another group or in the same, after second operation, the path number of p_1 is also in a group, and so on. Obviously, when a path number returns to a certain group, it can still be represented by an expression of one of the four odds in a group, and it also has its path number. Now we compare n_1 with each path number of four odds, that is, using p_1, p_2, p_3 and p_4 respectively, to determine if it is equal to n_1. If any of them is not equal to n_1, then it can be determined that there is no cycle in continuous operations, and conversely, there is a cycle. For example, for the second odd 27 in g_4, its path number 41 is the first odd in g_6, and the odd 41 has its path number 31 in g_4, and so on. Obviously, both 41 and 31 are different from 27. In fact, what we need to determine is whether all of the path numbers of 27 is equal to 27. If not, then we can determine that there hasn't any cycle in the continuous operations starting from 27. We make the same comparison with the path numbers of the rest three odds in a group.

4.1 Comparisons n_1 with four path numbers

a) Let $p_1 = n_1$, then we have

$$6t - 5 = 8t - 7 . \tag{4.12}$$

From equation (4.12), we get $t = 1$, that is, only in g_1, the path number of the odd 1 is equal to itself, it is a small cycle, and that is also the conclusion (1) obtained in section 2.2.1, and we can get $p_1 \neq n_1$ when $t \geq 2$.

b) Let $p_2 = n_1$, then we have

$$12t - 7 = 8t - 7. \tag{4.13}$$

From equation (4.13), we get $t = 0$, that is, $p_2 \neq n_1$.

c) Let $p_3 = n_1$, then we have

$$\frac{3t-1}{2^{k-1}} = 8t - 7 . \tag{4.14}$$

It can be seen that the equation (4.14) holds when $t = 1$, and $k = 2$. This means that in g_1, $n_1 = 8t - 7 = 1$ (equation (4.1)) and $n_3 = 8t - 3 = 5$ (equation (4.3)), as 5 has a smaller similar number 1, so we take 1 to do one time of operation and get its path number 1, so here, 1 is just p_3, i.e., it is the path number of n_3, thus $p_3 = n_1$. Because the previous similar number decreases, so we take the maximum value $3t - 1$ of p_3 to compare with $8t - 7$, and it can be seen that $8t - 7$ is greater than $3t - 1$ when $t \geq 2$. Now we get $p_3 \neq n_1$ when $t \geq 2$.

d) Let $p_4 = n_1$, then we have

$$12t - 1 = 8t - 7 . \tag{4.15}$$

From equation (4.15), we get

$$t = -\frac{3}{2} .$$

Since it is not a positive integer solution, so we can get $p_4 \neq n_1$.

From the analysis of a) to d), a conclusion (conclusion (a)) can be concluded: any path number is not equal to itself when keep doing operations starting with n_1, and $t \geq 2$.

4.2 Comparisons n_2 with four path numbers

a) Let $p_1 = n_2$, then we have

$$6t - 5 = 8t - 5. \tag{4.16}$$

From equation (4.16), we get $t = 0$, so $p_1 \neq n_2$.

b) Let $p_2 = n_2$, then we have

$$12t - 7 = 8t - 5. \tag{4.17}$$

From equation (4.17), we can get

$$t = \frac{1}{2}.$$

Since it is not a positive integer solution, so we have $p_2 \neq n_2$.

c) Comparisons n_2 with p_3. Here, we take the maximum value $3t-1$ of p_3, and then we have

$$n_2 = 8t - 5. \tag{4.2}$$

Analyzing $3t-1$ and $8t-5$, it is not difficult to see that, $8t-5$ is greater than $3t-1$ for any $t \geq 1$, so $p_3 \neq n_2$.

d) Let $p_4 = n_2$, then we have

$$12t - 1 = 8t - 5. \tag{4.18}$$

From equation (4.18), we get $t = -1$, since it is not a positive integer solution, so $p_4 \neq n_2$.

From the analysis of a) to d), a conclusion (conclusion (b)) can be concluded: any path number when keep doing operations starting with n_2, is not equal to itself when $t \geq 2$.

4.3 Comparisons n_3 with four path numbers

a) Let $p_1 = n_3$, then we have

$$6t - 5 = 8t - 3. \tag{4.19}$$

From equation (4.19), we can get $t = -1$, since it is not a positive integer solution, so $p_1 \neq n_3$.

b) Let $p_2 = n_3$, then we have

$$12t - 7 = 8t - 3. \tag{4.20}$$

From equation (4.20), we get $t = 1$, that is, only in g_1, the path number 5 of the second odd 3, is equal to the third odd 5. Obviously, the left side of equation (4.20) is greater than the right side when $t \geq 2$, so we get $p_2 \neq n_3$ when $t \geq 2$.

c) Comparisons n_3 with p_3. Here, we take the maximum value $3t-1$ of p_3, and we have

$$n_3 = 8t - 3. \qquad (4.3)$$

Analyzing $3t-1$ and $8t-3$, it is not difficult to see that, $8t-3$ is greater than $3t-1$ for any $t \geq 1$, so $p_3 \neq n_3$.

d) Let $p_4 = n_3$, then we have

$$12t - 1 = 8t - 3 \qquad (4.21)$$

From equation (4.21), we get

$$t = -\frac{1}{2}.$$

Since it is not a positive integer solution, so we get $p_4 \neq n_3$.

From the analysis of a) to d), a conclusion (conclusion (c)) can be concluded: any path number when keep doing operations starting with n_3, is not equal to itself when $t \geq 2$.

4.4 Comparisons n_4 with four path numbers

a) Let $p_1 = n_4$, then we have

$$6t - 5 = 8t - 1. \qquad (4.22)$$

From equation (4.22), we can get $t = -2$, since it is not a positive integer solution, so $p_1 \neq n_4$.

b) Let $p_2 = n_4$, then we have

$$12t - 7 = 8t - 1. \qquad (4.23)$$

From equation (4.23), we can get

$$t = \frac{3}{2}.$$

Since it is not a positive integer solution, so we get $p_2 \neq n_4$.

c) Comparison n_4 with p_3. Here, we also take the maximum value $3t-1$ of p_3, and we have

$$n_4 = 8t - 1. \qquad (4.4)$$

Analyzing $3t-1$ and $8t-1$, it is not difficult to see that, $8t-1$ is greater than $3t-1$ for any $t \geq 1$, so $p_3 \neq n_4$.

d) Let $p_4 = n_4$, then we have

$$12t - 1 = 8t - 1. \qquad (4.24)$$

From equation (4.24), we get $t = 0$, so $p_4 \neq n_4$.

From the analysis of a) to d), a conclusion (conclusion (d)) can be concluded: any path number when keep doing operations starting with n_4, is not equal to itself.

From the analysis of 4.1 to 4.4 and conclusions (a) to (d) above, now a conclusion (conclusions (4)) can be drawn: to do continuous operation starting from any given odd, every path number is unique, that is, there is no cycles in a forward operation path which may include similar numbers except the cycle 1-4-2-1, and therefore the path number either tends towards infinity or regresses to 1, and it means that a biggest odd will continue to appear in the path if it tends towards infinity, and on the contrary, a minimum odd will continue to appear in the final path.

5. Analysis of the principle of the inverse operations

The forward operation of the proposition is reversible for non-triples. Now, to do one reverse operation for formula (2.1), then we have

$$n = \frac{2^k p - 1}{3}, \tag{5.1}$$

or

$$3n = 2^k p - 1. \tag{5.2}$$

Where $k \in N$, and p takes non-triples (conclusion (2)).

Here, formula (5.1) or (5.2) is called the inverse formula of the proposition, 2^k is called a multiplier, and n is called the inverse path number of the given forward path number p. The formula (5.1) and (5.2) are used in reverse to find the inverse path number n.

Here, we firstly analyze some cases of particular path numbers p.

a) Let $p = 1$, from the formula (5.1), then we have

$$n = \frac{2^k - 1}{3}. \tag{5.3}$$

As it is analyzed in section 3.1, if k is even, $2^k - 1$ is a triple, then n has its positive integer solutions in equation (5.3); if k is odd, it hasn't any positive integer solution. From this, we can find the inverse path numbers such as 1, 5, 21, 85⋯, that is the number 1 set of similar numbers (see Tab.1). Here, it contains 1 itself in the inverse path numbers and there is a cycle 1-1-1.

b) Let $n = p$ and $p > 1$, from the formula (5.1), then we have

$$p = \frac{2^k p - 1}{3}.$$

That is

$$p = \frac{1}{2^k - 3}. \tag{5.4}$$

As it can be seen that equation (5.4) has only one positive integer solution $p = 1$ when

$k=2$. Since $p>1$, so we can draw a conclusion (conclusion (5)): the inverse path number isn't equal to the given forward path number, or $n \neq p$ when $p>1$.

c) Let $p=3t$, that is, p takes triples, where $t \in N$. From formula (5.1), then we have

$$n=\frac{2^k(3t)-1}{3}=2^k t-\frac{1}{3}. \tag{5.5}$$

Obviously, there is no positive integer solution to equation (5.5), so we can draw a conclusion (conclusion (6)): for any triple, it hasn't any inverse path number.

By analyzed in section 2.1 and here, we had known that, there be an infinite number of inverse path numbers, each of them constitutes the source of the known path number p, i.e., p is sourced from the infinite of inverse path numbers.

Next, we analyze the properties of the inverse operations.

5.1 The relationship between the inverse path numbers

From formula (5.1) or (5.2), it can be seen that the inverse path numbers are directly related to the value of k in the multiplier 2^k, and therefore, we use the parity property of the values of k to analyze the relationship between the inverse path numbers. Apparently, for any non-triple p, we can't at the same time get an integer in formula (5.1) when k takes the minimum odd 1 and the minimum even 2; for k and $k+2$, they are the same as an odd or even.

Let p be a non-triple, n_1 be the inverse path number, according to formula (5.1), then we have

$$n_1=\frac{2^k p-1}{3}. \tag{5.6}$$

Where $k \in N$.

To multiply with 4 for two sides in formula (5.6), then we obtain

$$4n_1=4\left(\frac{2^k p-1}{3}\right)=\frac{2^{k+2}p-4}{3}=\frac{2^{k+2}p-1}{3}-1.$$

That is

$$4n_1+1=\frac{2^{k+2}p-1}{3}. \tag{5.7}$$

Comparing formula (5.7) with the similar number formula (3.6), it is not difficult to see, that the right of formula (5.7) is the next similar number of n_1, marked n_2, that is

$$n_2=4n_1+1=\frac{2^{k+2}p-1}{3}$$

That is

$$n_2 = \frac{2^{k+2}p-1}{3}. \tag{5.8}$$

It's generated by added 2 to k in the multiplier, that is, when k takes the next odd or even, we can get the next similar number n_2 of n_1.

Again, to multiply with 4 for two sides in formula (5.8), then we have

$$4n_2 = 4\left(\frac{2^{k+2}p-1}{3}\right) = \frac{2^{k+2+2}p-4}{3} = \frac{2^{k+2+2}p-1}{3} - 1.$$

That is

$$4n_2 + 1 = \frac{2^{k+2+2}p-1}{3}. \tag{5.9}$$

As it can be seen in formula (5.9), that the right is the next similar number of n_2, and marked n_3.

That is

$$n_3 = 4n_2 + 1 = \frac{2^{k+2+2}p-1}{3}. \tag{5.10}$$

In the same way, we can find other similar numbers when k takes again next odd or even. As the value of k increases, there are an infinite number of similar numbers of n_1.

For an odd number n, to $3n+1$, it must be an even and can be divided by 2 one time or more times.

Now, a conclusion (conclusion (7)) can be drawn by the analysis above: in formula (5.1), for a non-triple p, if k takes the smallest odd number 1 (the smallest multiplier is 2), we can get an inverse path number, then k takes the rest odds greater than 1, we can also get an infinite of number of inverse path numbers, and all the inverse path numbers are similar to each other; if, when k takes the smallest odd number 1, we can't get an inverse path number, then k takes the smallest even number 2 (the smallest multiplier is 4) and any even number greater than 2, we can get an infinite of number of inverse path numbers definitely, and they are also similar numbers to each other.

Obviously, all the inverse path numbers constitutes a set of similar numbers.

When k takes the smallest odd 1 or even 2, i.e., multiplier 2^k takes the smallest 2 or 4, the inverse path number is called here the first inverse path number.

5.2 The properties of the first inverse path numbers

Suppose, there are two adjacent similar numbers n_1 and n_2 in a set, where n_2 is known, and $n_2 > n_1$, i.e., n_1 is the previous similar number of n_2, according to formula (5.1), then we have

$$n_2 = \frac{2^k p-1}{3}.$$

According to formula (3.7), we have

$$n_1 = \frac{n_2 - 1}{4}. \qquad (5.11)$$

To substitute n_2 in formula (5.11), then we have

$$n_1 = \frac{\frac{2^k p - 1}{3} - 1}{4} = \frac{2^k p - 1 - 3}{12} = \frac{2^k p}{12} - \frac{1}{3}. \qquad (5.12)$$

Next, we take the smallest multiplier 2 and 4 in formula (5.12) for analysis respectively,

a）when taking 2

$$n_1 = \frac{2p}{12} - \frac{1}{3} = \frac{p}{6} - \frac{1}{3} = \frac{p-2}{6},$$

b）when taking 4

$$n_1 = \frac{4p}{12} - \frac{1}{3} = \frac{p}{3} - \frac{1}{3} = \frac{p-1}{3}.$$

Obviously, $p-2$ is an odd, $p-1$ is an even above, so, neither of these two equations can get an integer, thus, there is no similar number less than n_2, that is, n_1 doesn't exist. Therefore, it can be concluded (conclusion (8)): the first inverse path number must be the generating number of a similar number set (i.e., n_2 is the first here).

Definition 5

(a) When doing the reverse operations according to formula (5.1), we changed the names, called the non-triple p given originally as a primitive number; called the first inverse path number as a source number of p;

(b) Doing one time of reverse operation is called one time of reverse tracing, tracing for short; times of tracing is called continuous tracing.

According to definition 5, one time of tracing is to find out the first source number of a primitive number. Since the rest inverse path numbers of a primitive number are similar to the source number (conclusion (7)), they can be found in turn by using the similar number formula (3.6), and therefore, here we defined only the first of them.

5.3 Analysis of the multipliers of two continuous primitive numbers

As stated in 3.3, for the continuous series of odd numbers, there are only two consecutive non-triples between two adjacent triples, only non-triples are the primitive numbers.

Let $3p$ be a triple, where p is an odd, so, in terms of the increasing value, the first primitive number adjacent to $3p$ is $3p+2$, and the second is $3p+4$. Next, we take the multiplier 2 and 4 for analysis respectively.

5.3.1 Take 2

According to the formula (5.2), we have an equation as fallow

$$3n = 2p - 1. \qquad (5.13)$$

a) Put the first primitive number $3p+2$ into equation (5.13), and then we have

$$3n = 2(3p+2) - 1 = 6p+3.$$

That is

$$n = 2p+1. \tag{5.14}$$

Obviously, there is an integer solution in equation (5.14), and it is the source number of first primitive number.

b) Put the second primitive number $3p+4$ into the equation (5.13), and then we have

$$3n = 2(3p+4) - 1 = 6p+6+1.$$

That is

$$n = 2p+2+\frac{1}{3}. \tag{5.15}$$

Obviously, the equation (5.15) hasn't any integer solution.

5.3.2 Take 4

According to the formula (5.2), we have an equation as fallow

$$3n = 4p-1. \tag{5.16}$$

a) put the first primitive number $3p+2$ into equation (5.16) , then we have

$$3n = 4(3p+2) - 1 = 12p+6+1.$$

That is

$$n = 4p+2+\frac{1}{3}. \tag{5.17}$$

Obviously, the equation (5.17) hasn't any integer solution.

b) put the second primitive number $3p+4$ into equation (5.16) , then we have

$$3n = 4(3p+4) - 1 = 12p+15.$$

That is

$$n = 4p+5. \tag{5.18}$$

Obviously, there is an integer solution in equation (5.18), and it is the source number of the second primitive number.

Thus, based on the analysis of 5.3.1 and 5.3.2, it can be concluded (conclusion (9): in the continuous series of odd numbers, for two continuous primitive numbers, the multiplier of the first is to take 2 and the second is to take 4 definitely.

And as a result, for their two source numbers, the first gets smaller, the second gets larger (be similar to conclusion (3)).

Here, we called the formula (5.13) and (5.16) as the formulas of source numbers.

5.4 Analysis of the multiplier of three consecutive odd numbers

5.4.1 The setting of the multiplier of a triple

Since formula (5.1) doesn't hold for the triples (conclusion (6)), therefore, the multiplier of a triple is set as 0, which means that there is no source number for the triples. Thus, for two consecutive primitive numbers and a triple, i.e. three consecutive odd numbers, as derived from conclusion (9), their multipliers in order are 2, 4 and 0.

5.4.2 A special determinant of odds

According to the conclusion (9), the continuous odd numbers can be arranged in a special determinant in tabular form, and then the special regular of multipliers can be shown. See Tab. 2 Table of multipliers.

Tab. 2 Table of multipliers (2^k)

row h	con-tent	column l											
		1	2	3	4	5	6	7	8	9	10	11	12
	2^k	**2**	**4**	**0**	**2**	**4**	**0**	**2**	**4**	**0**	**2**	**4**	**0**
1	odd											1	3
2		5	7	9	11	13	15	17	19	21	23	25	27
3		29	31	33	35	37	39	41	43	45	47	49	51
4		53	55	57	59	61	63	65	67	69	71	73	75
5		77	79	81	83	85	87	89	91	93	95	97	99
6		101	103	105	107	109	111	113	115	117	119	121	123
7		125	127	129	131	133	135	137	139	141	143	145	147
8		149	151	153	155	157	159	161	163	165	167	169	171
…													
h		n											
notes	1. This table shows consecutive odds in rows and columns; there are 1 and 3 in first row and the multipliers of them are corresponding to 4 and 0; from second row, there are twelve odds in each, and the multipliers of them are corresponding to 2, 4 and 0 four times in order; 2. The triples are with shadows, and they form four columns, the remaining 8 columns are primitive numbers.												

In the table, the multipliers for each row are repeated four times in the order of 2, 4 and 0. Next, we use this table to analyze the method of taking the multipliers.

5.4.3 Location analysis of an odd number in Tab. 2

a） Row number h

Let n be an odd given arbitrarily, then, according to the odd number arrangement in the table, its row number h is given by the following formula

$$h = \frac{n-4}{24} + 1. \tag{5.19}$$

Where, h takes an integer approach large.

Here, we called formula (5.19) the row number formula.

b） Column number l

Let the row number h be known, and then the column number l of the odd number n

in the table is given by the following formula

$$l=\frac{n-4-24(h-2)}{2}. \tag{5.20}$$

Where, l takes an integer approach large also.

Here, we called formula (5.20) the column number formula.

According to the column number l in the table, we can cross-reference to get the multipliers, that is,

When l equals to 1, 4, 7, 10, the odd is a primitive number, and its multiplier is 2;
When l equals to 2, 5, 8, 11, the odd is a primitive number, and its multiplier is 4;
When l equals to 3, 6, 9, 12, the odd is a triple, its multiplier is 0.

Now, by analysis in 5.4.2 and 5.4.3, it can be concluded (conclusion (10)): for any given odd number n (it may be a primitive number or a triple), its multiplier 2^k is obtainable.

As it can be seen from conclusion (10) that the multiplier can be determined by the primitive number its self. Thus, the source number of any primitive number can be found out by using formula (5.13) or (5.16), and an infinite number of similar numbers of the source number can also be found out by using formula (3.6).

In a limited range of odds, there is a simple way to find the source numbers. Firstly, we use 2 directly in formula (5.1) to try to find out the source number n. If n is an integer, so 2 is its multiplier, and p is a primitive number, the integer n is its source number. If n isn't an integer, then use 4 to try secondly, if neither of n is an integer, then p must be a triple, and it has no source numbers.

For two multipliers, more simply, we can determine firstly whether p is a triple (odd), that is, if $p/3$ is not an integer, then to $p+2$, if $(p+2)/3$ is also not an integer, then $p+4$ must be a triple, thus we can get the multipliers of two adjacent primitive numbers by the order of 2, 4 and 0 (conclusion (9)).

6. Analysis of the continuous inverse operations paths

6.1 The continuous tracing path

According to conclusion (10), source number n can be obtained by tracing for a primitive number p. Obviously, if n isn't a triple, then n can be regarded as another primitive number for tracing again. Over and over again, primitive number p with one or more of its source numbers forms a continuous tracing path. For a given primitive number p, if its tracing path hasn't an end of a triple (conclusion (6)), then it will tend towards infinity because the number of odds smaller than p is limited when the path tends to small odd, thus the path will either end at a triple or tend towards infinity. For example, tracing for 1, 1 can be traced itself; tracing for the similar number 5 of 1, we can get the minimum triple 3 (the forward path number of 3 is 5), so there has two odds in this path, and as the same, for the odd 445, 17 times of tracings are required to obtain the triple 27, there are 18 odds in this path. The path from 5 to 3 or from 445 to 27 is a complete successive tracing path, and in the opposite direction, it is a complete narrow path (see section 2.3). Here, we didn't prove that all odds end at a triple when tracing because it does not affect the final conclusion.

6.2 The extended tracing paths

When getting a triple, we take the next similar number (it's not a triple, about the gap, see section 3.3) of the triple to trace again, and as the same, we take the next similar numbers of the rest source numbers to trace again (if it's a triple, skip it, and again take the next), thus we can get some extended tracing paths by using these similar numbers as transitions.

Obviously, we can do inverse operations starting with the odd 1. There are two kinds of operations, one is to find the next similar numbers, and the other is to find the source numbers of the primitive numbers in the similar number sets. Its order is, firstly to find the similar numbers of 1 one by one, that is, we get the number 1 set, then to trace source numbers of the primitive numbers in the set, such as we can get the odd 3 when tracing for the odd 5 of the next similar number of 1, and also we can get the rest similar numbers of 3 then we get the number 3 set. The inverse operations are repeated in this way, and there will be more and more paths spreading out like branches of a tree, and the following paths will all get longer and longer if there is no cycle.

6.3 The analysis of the cycles in continuous inverse operations

Forward operations can start from any odd, and each of its forward path numbers is unique and there hasn't any cycle in the path (conclusion (4). Although some similar numbers were skipped in the forward path, but they can still be regarded as in the path because each of them can be as an operation starting odd. For example, the odd 3 has its similar numbers 13, 53 and so on, and theirs forward path numbers are all 5, obviously, we can do one time of forward operation for 3, 13 and 53 respectively, that is, using each of them as the starting point, thus they had appeared in this path. When we do continuous inverse operations, all of the skipped odds can be found, such as tracing for 5, we can get its source number 3 and from 3 we can get one by one its similar numbers 13 and 53 and so on, that is, we can get number 3 set. From this way, it's not difficult for us to come out that the odds both in the forward and inverse path, such as A to B and B to A, had a one-to-one correspondence, since the similar numbers we take increase in turn, thus we can introduce a conclusion (conclusion (11): there hasn't any cycle in the tracing and extended paths when $p > 1$ and all the extended tracing paths (or the odds in the branches) will tend towards infinity.

Obviously, we can obtain an infinite number of odds when doing continuous inverse operations and if to do forward operations continuously for any odd obtained, it will be back to 1 definitely.

6.4 Analysis of density of the odds obtained and the final conclusion

Now, we analyze the density of the odds obtained when doing continuous inverse operations. Suppose, there is an odd n in the series of odd numbers which hasn't been traced on the paths starting from 1, that is, n had been missed. It is obvious that we can do continuous inverse and forward operations for it. When doing inverse operations continuously, the inverse path numbers in the branches must all tend to infinity (conclusion (11)), and when doing forward operations continuously, the forward path number must also tend to infinity, because if its forward path numbers get smaller, it must eventually reach 1 (conclusion (4)), it shows that there must be a reverse tracing path between 1 and n. From this, for n, both inverse and forward operations path tend to infinity. However, the inverse operations are just the opposite direction based on the same kind of operational rules of the forward directions and the odds had a one-to-one correspondence, so there has only one direction, therefore, the assumption above doesn't hold, thus the odd number n must not only be in the range of the odds obtained by tracing, but it must also be regressed to 1 if doing forward operations for it.

From this analysis, it can be concluded (conclusion (12)): finally, for any positive integer (an even number is transformed into an odd firstly), to do forward operations, it must along the inverse paths in the opposite direction and return to 1. So, the Collatz conjecture holds.

In this paper, the basic operational principle of the conjecture is expounded.

Acknowledgments

I would like to express my thanks to Liansheng Zhang for encouragement and help, who is a professor in Suzhou university of Science and Technology. I also thank the reviewers for comments.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

Reference

[1] Guy, R.K. (2007) Unsolved Problems in Number Theory: The 3x + 1 Problem. Springer Verlag, New York, 330-336.