Proof of the Collatz Conjecture

Kurmet Sultan

Abstract: It is proved that if the calculation of the Collatz function $C(n)$ starts with odd numbers of the form $6m \pm 1, \ m \in \mathbb{N}$, then at each iteration the integer will have the form $6n \pm 1$ or will be equal to 1. Further, it is proved that during the reverse calculation by the formula $((6n \pm 1) \cdot 2^q - 1)/3$, increasing the exponent of two by 1 at each iteration, then each number of the form $6n \pm 1$ will correspond to an infinite number of alternating integers of the form $3n, 6m - 1$ and $6m + 1$ which are the preceding numbers. Then it is shown that if you build a graph by connecting the numbers $6n \pm 1$ with their previous numbers, you get a tree graph. A graph tree, each vertex of which corresponds to numbers of the form $6m \pm 1$, is a proof of the Collatz conjecture, since any of its vertices will be assigned 1.

Keywords: Collatz conjecture, $3n + 1$ problem, Syracuse, graph.

AMS Classification: 11D09

1 Introduction

The Collatz conjecture, also known as the $3\,n + 1$problem, the Syracuse problem, is one of the unsolved problems in mathematics. The following papers devoted to the $3\,n + 1$ problem [1, 2, 3, 4] can be noted. The book [5] provides a proof of the correctness of the Collatz conjecture for natural numbers, as well as the absence of cycles, except for the known three cycles, of the Collatz function for negative integers.

The Collatz function $C(n)$ is defined on natural numbers as follows:

$$(1.1) \quad C(n) = \begin{cases} n/2, & if \quad n - is\ even, \\ 3n + 1, & if \quad n - is\ odd. \end{cases}$$

To explain the Collatz conjecture, we take any natural number n, if it is even, then divide it by 2, and if it is odd, then multiply by 3 and add 1 (we get $3\,n + 1$). We perform the same operations with the resulting number, and so on. The Collatz conjecture is that whatever initial number n is taken, sooner or later we will get it.

2 Start number

By the condition of the conjecture, the calculation of the Collatz function can be started from any natural number. However, it is obviously more efficient to start with an odd number, since any even number when divided by 2 (one or more times) will turn into an odd number.

Since odd numbers can be divided into odd numbers that are divisible by 3 and odd numbers that are not divisible by 3, the following question arises:

Question 1. What odd numbers are more efficient to start the calculation with?

It is known that all-natural numbers, with the exception of 1, can be represented by the formulas: 1) $3n$; 2) $3n-1$; 3) $3n+1$, where $n = 1, 2, 3 \ldots$. It is also known that all odd numbers indivisible by 3 can be represented as $6m-1$; $6m+1$, where $m = 1, 2, 3 \ldots$.

If odd numbers of the form $3n$ are multiplied by 3 and 1 is added, then it is obvious that the result will be numbers of the form $3n+1$. Moreover, if a number having the form $3n+1$is odd, then, of course, it will be a number of the form $6m+1$. And if a number of the form $3n+1$ is an even number, then, when dividing by 2 (one or more times) until an odd number is obtained, a number of the form $6m-1$or $6m+1$is formed, since a number of the form $3n+1$ is not divisible by 3 .

Thus, it can be argued that all-natural numbers that are multiples of 3 through one operation $(3n)3+1$, and division by certain powers of two, in the case of an even number, turns into odd numbers of the form $6m-1$ and $6m+1$. The exceptions are the numbers, which after the operation $(3n)3+1$will be equal to the power of two.

From the above, it follows that the calculation of the Collatz function is more efficient to start with odd numbers that look like $6m-1$or $6m+1$.

Now let's answer the following question:

Question 2. What kind of numbers are formed from numbers of the form $6m-1$ and $6m+1$ as a result of calculating the Collatz function?

It is clear that all numbers of the form $6m-1$ and $6m+1$ are odd numbers of the form $3n-1$ or $3n+1$, since they are not multiples of 3. If you multiply such numbers by 3 and add 1, then it is natural to get even numbers of the form $3m+1$. And when dividing even numbers of the form 3m + 1 by two (one or more times) until an odd number is obtained, then the resulting odd numbers will look like $6n-1$ or $6n+1$, since numbers of the form $3m+1$are not divisible by 3.

It follows from the above that as a result of calculating the Collatz function based on numbers of the form $6m-1$and $6m+1$, numbers of the form $6n-1$ or $6n+1$ are formed, i.e. number does not change.

Note. Different letters m and n are used in the notation of numbers of the same form $6n \mp 1$ *and* $6m \mp 1$ *to distinguish between the argument (variable) and the particular value of the Collatz function.*

Thus, on the basis of the foregoing, we can state that in order to prove the Collatz conjecture, it is necessary to establish connections between numbers of the form $6m \mp 1$ and 1.

3 Forward and backward calculations

3.1 Direct calculation

To establish connections between numbers of the form $6m \mp 1$ and 1, we first perform a direct calculation based on the number 1 and all elements of the set

(3.1) $K = \{k | k = 6m \mp 1, m \in N\}$.

The actual number 1 is also an element of the set K obtained at $m = 0$, but we decided not to associate with the number -1, which is also obtained at m=0, the number 1 is selected from the set K.

For direct calculation, we use the formula

(3.2) $C = \frac{3x+1}{2^q}$; $C, x, q \in \mathbb{N}$.

It is not difficult to establish that if $x = 1, q = 2$ then $C = 1$. This means that when calculating the Collatz function using the number 1, a cycle is formed. By the way, this is the only cycle formed when calculating the Collatz function, the proof of which will be given later.

It follows from what was said in Chapter 2 that if $x = 3n$, $n \in N$ then by formula (3.2) we get $C = 6m \mp 1$, except for the cases when $C = 1$. Obviously, there is only one natural solution for every x (or n). It also follows from Chapter 2 that if $x = 6n \mp 1$, then $C = 6m \mp 1$, except when $C = 1$, where $n, m \in \mathbb{N}$. In this case, too, there is only one natural solution for every x (or two solutions for every n).

This means that in a direct calculation, a number of the form $6n \mp 1$ will certainly be associated with only one number of the same type, or the number 1.

3.2 Back calculation

The reverse calculation of the Collatz function for odd natural numbers is carried out using the formula

(3.3) $\bar{C} = \frac{x \cdot 2^q - 1}{3}$; $\bar{C}, x, q \in \mathbb{N}$.

If $x = 1$, then there are infinitely many natural solutions such that $\bar{C} = 3n$ and $\bar{C} = 6n \mp 1$, except for one case when $\bar{C} = 1$, where $n \in \mathbb{N}$.

For formula (3.3), if $x = 3n,\ n \in \mathbb{N}$, then the inverse Collatz function $\bar{C}$ has no natural solution. This also proves that the direct calculation of the Collatz function using odd multiples of 3 results in a number like $6m \mp 1$.

If for the reverse calculation by formula (3.3) we use numbers of the form $x = 6m \mp 1$, then there are infinitely many natural solutions such that $\bar{C} = 3n$ and $\bar{C} = 6n \mp 1$, where $n, m \in \mathbb{N}$.

This means that when calculating backwards, a number of the form $6m \mp 1$ will certainly be associated with infinite numbers of the same form, if you do not take into account numbers that are multiples of 3.

From chapters 3.1 and 3.2 it follows that each element of the set (3.1) certainly has an infinite connection with other elements of the same set, and some of them have a connection with 1. This means that all elements of the set K will be connected with each other using the Collatz function and number 1.

Thus, the essence of the Collatz conjecture is to compare the elements of the set $K = \{k | k = 6m \mp 1, m \in N\}$ to each other and to the number 1, in a special way.

3.3 Set of predecessor numbers

It follows from the inverse calculation formula (3.3) that each element of the set (3.1) and the number 1 has an infinite number of predecessor numbers, since the power of two q can be increased indefinitely, i.e. each element of the set (3.1) and the number 1 has its own set of predecessor numbers.

Definition 1. The number-precursor of a natural number k is a natural number g, if for x=g according to the formula (3.2) the number k is obtained.

For example, 13 is the predecessor of the number 5, since if we take x =13 and q =3, then by the formula (3.2) we get $\frac{3\cdot 13+1}{2^3} = 5$. In turn of 5 is the predecessor number for 1.

The set of predecessor numbers of the elements of the set (3.1) and the number 1 can be represented, respectively, by the following formulas

(3.4) $K_g = \{k|\ k = 6m \mp 1;\ \ m \in N\}$ – is the set of numbers of predecessors of one element of the set (3.1);

(3.5) $K_1 = \{k_1|\ k_1 = 6m \mp 1;\ \ m \in N\}$ - is the set of numbers of predecessors of the number 1.

The sets K_g and K_1 differ from the set K in that the elements of the set K_g and K_1 are numbers of the form $k = 6m \mp 1$ associated only with one element of the set K and the number 1, and the elements of the set K are all numbers of the form $k = 6m \mp 1$.

In this case, the set of predecessor numbers corresponding to each element of the set (3.1) are subsets of the set (3.1). In other words, the union of all sets of predecessor numbers of the elements of the set (3.1) and the number 1 forms the set (3.1), i.e. $K = K_G \cup K_1$, where

$K_G = K_5 \cup K_7 \cup K_{11} \cup K_{13} \dots$.

Thus, we can assert that any element of the set (3.1) is the predecessor number of only one element of the set (3.1) or the number 1, since on the basis of any element of the set (3.1) by formula (3.2) it is possible to obtain only one other element of the set (3.1) or number 1.

Note that each set K_g and set K_1 are disjoint, i.e. the elements of the sets K_g and K_1 will not match.

This is proved as follows: let two elements belonging to two sets K_{gi} and K_{gj} be equal, then we get the following equality

$\frac{(6m_1 \mp 1)\cdot 2^{q_1} - 1}{3} = \frac{(6m_2 \mp 1)\cdot 2^{q_2} - 1}{3}$ or

(3.6) $\frac{6m_1 \mp 1}{6m_2 \mp 1} = \frac{2^{q_2}}{2^{q_1}}$.

The last equality does not have a solution in integers, since the right side of the equality is equal to an even number or 1 (if $q_1 \leq q_2$), and the left side of the equality is equal to an odd number or non-integer, since $6m_1 \mp 1 \neq 6m_2 \mp 1$. It is obvious that (3.6) does not have a solution in integers even for $q_1 > q_2$.

It is easy to prove that the elements of the set K_1 are not contained in any set K_g, since,

$\frac{1\cdot 2^{q_1} - 1}{3} \neq \frac{(6m \mp 1)\cdot 2^{q_2} - 1}{3}$ or

(3.7) $\frac{1}{6m \mp 1} \neq \frac{2^{q_2}}{2^{q_1}}$.

Inequality (3.7) is true for $q_1 \leq q_2$ and $q_1 > q_2$.

Appendix 1 contains Tables 1 and 2, in which the elements of the set (3.1) and their predecessor numbers, as well as the predecessors of the number 1, are represented by arithmetic progression formulas. The possibility of representing the elements of the set (3.1) and their predecessor numbers by an arithmetic progression confirms the interconnectedness of all Collatz numbers and the number 1.

4 Collatz function graph

4.1 Micrograph

It follows from Chapters 3.1 and 3.2 that each element of the set (3.1) can be represented as a micrograph, and each micrograph is connected from above to infinite elements of the set (3.1), and from below to only one element of the set (3.1) or the number 1.

Definition 2. A micrograph is a graph consisting of one vertex and edges connecting the vertex of the micrograph with other vertices.

To explain what has been said in more detail, we take any element of the set (3.1) or the number 1, they correspond to the variable x in formulas (3.2) and (3.3), then using formula (3.2) we first perform a direct calculation, then we get one number of the form $6n \mp 1$ or number 1. After that, using the same element, we will carry out the reverse calculation according to the formula (3.3), then, if we increase the power of two q, we will get an infinite number of numbers of the form $6m \mp 1$. The number 1 is also a micrograph, however, unlike other vertices, it has a lower edge directed towards itself, i.e. it has a contour.

This can be represented graphically as follows:

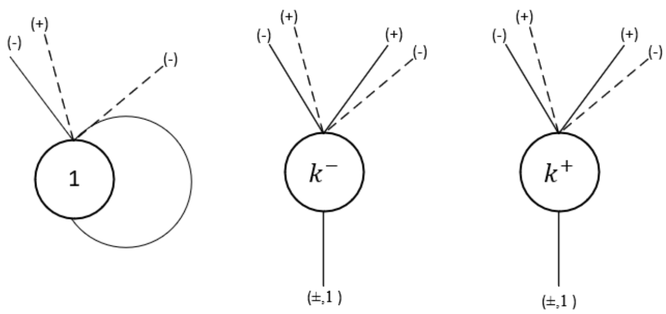

Figure 1. Micrographs, with two pairs of upper edges: $k^- = 6m - 1$ и $k^+ = 6m + 1$.

Note: the indices (-) and (+) in the letters of the vertices correspond to the sign of the formulas.

In Figure 1, for each micrograph, two pairs of edges are shown, their actual number is infinite, however, for the presentability of the graph, a limit should be set in advance on the number of edges for each vertex, as well as on the number of iterations, i.e. the height of the tree and its density should be limited. In particular, you should limit the number of iterations in the reverse calculation, because otherwise the tree will grow indefinitely.

The reason for showing the upper edges in pairs is described in [5], but it does not affect the proof of the Collatz conjecture, so we accept them without explanation.

4.2 Graph tree

Further, if we combine the micrographs described in chapter 4.1, we get a tree graph similar to Figure 2,

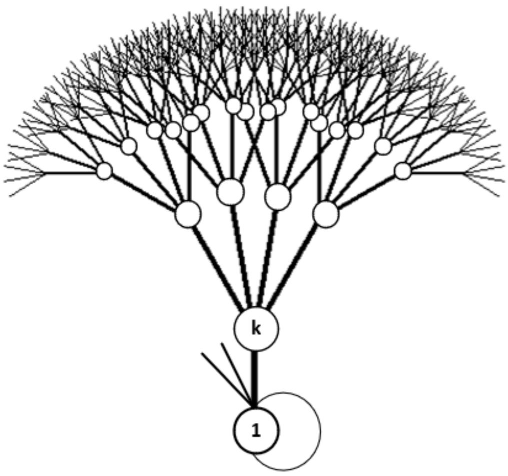

Figure 2. Graph-tree, in which the number of upper edges is limited to 4.

Note: 1) Vertex 1 also has four top edges and one bottom edge, but for vertex 1, the bottom edge and one top edge form a contour; 2) For presentability, the top two edges of vertex 1 are shown without continuation, i.e. they are cut off.

Note that we use the terms top edges and bottom edge instead of using the terms incoming arc and outgoing arc, because in our case the direction of the arc depends on the calculation method. For example, if direct calculation is used according to the formula (3.2), then all arcs will be directed downwards, and if the reverse calculation method is used according to formula (3.3), then all arcs will be directed upwards.

Figure 3 shows a directed tree graph containing even numbers (2n), multiples of 3 (3n) and non-multiples of 3 (k^- and k^+) odd numbers. In this graph, all arcs are directed downwards, because formula (3.2) is used to calculate the nodes of the graph.

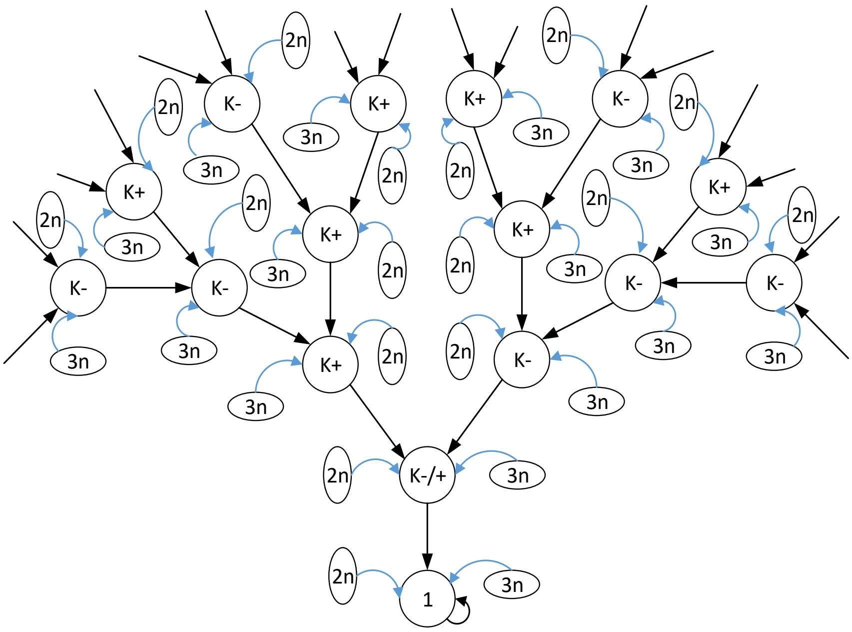

Figure 3. A directed tree graph containing even numbers (2n), multiples of 3 (3n), and non-multiples of 3 (k^- and k^+) odd numbers.

It follows from Figure 2 that the size of the graph tree corresponding to the Collatz function for positive integers is determined by the number of edges at one vertex V and the iteration level L, so we will use the sign $G_{V,L}$ to show the size of the graph. It should be clear that for the graph-tree corresponding to the Collatz function, $V = \infty$, $L = \infty$, since the graph-tree is built on the basis of all elements of the set (3.1).

The calculation of the predecessor numbers of the number 1 will be called the first level of the iteration $L = 1$, the calculation of the predecessor numbers of numbers that are the predecessor numbers of the number 1 will be called the second level of the iteration $L = 2$, and so on.

4.3 Algorithm for constructing graph-tree

The Collatz graph can be built in different ways: 1) by combining pre-formed micrographs; 2) according to the "top-down" method based on a given limit number; 3) according to the “bottom to top” method, with a different number of upper edges of vertices or an equal number of upper edges; 4) in a mixed way.

At the same time, in our opinion, the most convenient and understandable way is to build a Collatz graph according to the “bottom-up” method with an equal number of upper vertex edges, since in this case the tree graph will be fractal.

In view of the foregoing, we will describe below the construction of a graph according to the "bottom-up" method with an equal number of upper vertex edges.

Exercise. It is required to construct the Collatz graph $G_{4,4}$ for positive integers. This means that each vertex of the graph must have four edges, and the iteration level must be 4.

To build a graph-tree $G_{4,4}$ we use the inverse calculation formula (3.3), then based on the three predecessor numbers of the number 1 (the number 1 forms a loop, so only 3 branches will grow from the number 1, instead of 4) we get 12 vertices, then on based on 12 vertices, we get 48 vertices, etc. In this case, we will get a tree graph similar to the graph shown in Figure 2, but unlike it, the two cut edges of vertex 1 continue and will look the same as the first edge of vertex 1.

At the same time, it is obvious that any pair of vertices of the constructed tree graph $G_{4,4}$ will have a route, and all of them will be connected to 1. If there were no constraint $L = 4$, the tree graph would grow infinitely.

It should be noted that each subsequent pair of predecessor numbers will be 64 times more than the previous pair of predecessor numbers plus 21, i.e. the predecessor numbers grow very quickly. For example, the first pair of predecessor numbers of 5 are 13 and 53, and the next pair of predecessor numbers are 853 and 3413, which can be represented as 13·64+21=853 and 53·64+21=3413. Details of the patterns of predecessor numbers are given in [5]. It follows from the above that it is more practical to build a tree graph by limiting the number of vertex edges to a small number. When assigning the size of a graph, it is desirable that the number V, indicating the number of upper edges of each vertex, be even, an explanation for this is given in [5].

Note that if the tree graph is built in three dimensions, then the graph will be much more presentable.

4.4 Graph Properties

Based on what was stated in Chapter 4.2, we list the properties of a graph expressing the Collatz functions for positive integers:

1) The graph of the Collatz function for positive integers is a tree, with root 1 and a single contour 1-1;

1.1) There is a route between any pair of graph vertices.

Obviously, the form of the Collatz graph described in Chapter 4.2 does not depend on the boundary numbers V and L, i.e. for any V >1 and L >1 the Collatz graph will be a tree, so there is a route between any pair of graph vertices.

If we apply the terms of Graph Theory, then the resulting Collatz Function Graph for positive integers is a strongly connected directed graph, since there is a (directed) path from any vertex to any other in it. In other words, based on any number of the form $6m \mp 1$ and 1, you can calculate any number corresponding to any vertex of the graph, if you know the route between these two vertices.

1.2) There is no cycle in the graph, except for cycle 1-1.

By definition, a Tree is a connected acyclic graph. Connectivity means the presence of a route between any pair of vertices, acyclicity means the absence of cycles. Since the Collatz graph is a tree, there will be no contours (loops) in the graph, except for the contour 1-1.

The proof of the absence of a contour (loop) based on the graph is given in [5], where it is shown that the formation of a contour in the Collatz graph is possible only if at least one vertex has two lower edges. And this is possible only if for a given natural x by formula (3.2) it would be possible to obtain two odd integers. Of course, this cannot be, so the contour (or loop) in the graph cannot be formed. The work [5] also provides a proof of the absence of cycles, except for the known three cycles, for the Collatz function for negative integers.

2) All elements of the set (3.1) are vertices of the Collatz graph.

If micrographs are created on the basis of all elements of the set (3.1), then, obviously, all these micrographs are connected to each other, since they are all elements of the same set. However, suppose that one vertex (one element of the set K) is not connected with the graph, i.e. Let's say there is a single graph. However, a single graph can exist only if there is an element x of the set (3.1) that is a solution to the following two equalities: 1) $\frac{3x+1}{2^q} = x$; 2) $\frac{x \cdot 2^q - 1}{3} = x$.

Since these equalities have a natural solution only for $x = 1$ and $q = 2$, we can assert that there is no single graph.

If one element of the set (3.1) cannot be unconnected, then the number of separate vertices connected to each other must be infinitely many, of course, if any can exist at all. If so, then they will certainly be connected with the vertices of the Collatz graph, since the graph $G_{V,L}$ for $V = \infty, L = \infty$ contains all elements of the set (3.1). The proof of the interconnectedness of all elements of the set (3.1) is also the representation of predecessor numbers by arithmetic progression formulas and the Collatz number relationship scheme given in Appendices 1 and 2. From the relationship scheme of Collatz numbers and arithmetic progression formulas, it follows that any element of the set (3.1) can be calculated based on another element of the set, which proves that the vertices of the Collatz graph are strongly connected.

The work [5] also provides a proof of the interconnectedness of all elements of the set (3.1) by studying numbers of the form $k = 6m \mp 1$ and their subsequent numbers of the same form, obtained by direct calculation using formula (3.2).

Conclusion

Thus, we have proved that the graph of the Collatz function for positive integers is a tree with root 1 and contains all elements of the set $K = \{k | k = 6m \mp 1, m \in N\}$. This means that the Collatz conjecture is correct and has been proven.

Acknowledgements

I thank all the buyers of my book The Hardest Problem of mathematics: 3 n + 1. Ingenious Solution.

References

[1] J. C. Lagarias, The Ultimate Challenge: The 3x+1 Problem. American Mathematical Society, 2010.

[2] R. K. Guy, Unsolved Problems in Number Theory (3rd ed.). Springer-Verlag, 2004.

[3] L. Collatz, On the motivation and origin of the (3n + 1) – Problem, J. Qufu Normal University, Natural Science Edition, 12(3) (1986) 9–11.

[4] Tao, T. (2022). Almost all orbits of the Collatz map attain almost bounded values. *Forum of Mathematics, Pi*, 10, E12. doi:10.1017/fmp.2022.8

[5] K. Sultan, The Hardest Problem of mathematics: 3 n + 1. Ingenious Solution, ISBN-13:979-8840633250, 2022.

Appendix 1. Representation of predecessor numbers by arithmetic progression formulas

When considering the relationship between the elements of the set (3.1) and their predecessor numbers, it is convenient to consider them as different sets G and K, although in fact we show the relationship between the elements of the set K and the number 1, since $G = K$. As a result of the study of the elements of the set (3.1) and their predecessor numbers calculated by formula (3.3), the regularity described below is established.

The elements of the set G, taking into account their frequency, can be divided into 6 types, which are determined by the six formulas of the arithmetic progression given in the second column of Table 1.1, and the first pairs of predecessor numbers $k = 6m \mp 1$ calculated by formula (3.3) can be divided into 12 types, which are also determined by arithmetic progression formulas. Arithmetic progression formulas corresponding to 6 types of elements $g = 6n \mp 1$ and 12 types of elements $k = 6m \mp 1$, respectively, of sets G and K are given in Table 1.

Table 1.1. Formulas for calculating the first two pairs of predecessor numbers

		Formulas of predecessor numbers $k^{\mp} = 6n \mp 1$ with exponents of 2 given below					
	g	q =1	q =2	q =3	q =4	q =5	q =6
0	1		1		5		
1	$5 + 18t$			$13 + 48t$		$53 + 192t$	
2	$7 + 18t$				$37 + 96t$		$149 + 384t$
3	$11 + 18t$	$7 + 12t$		$29 + 48t$			
4	$13 + 18t$		$17 + 24t$				$277 + 384t$
5	$17 + 18t$	$11 + 12t$				$181 + 192t$	
6	$19 + 18t$		$25 + 24t$		$101 + 96t$		

Note. The number 1 on the first interval q: 1, 2, 3, 5, 6 corresponds to two predecessor numbers 1 and 5, which are given on the line with the serial number 0.

Table 1.2 shows the elements of the set (3.1) and their predecessor numbers calculated by the formulas of arithmetic progression, which are also elements of the set (3.1). The number 1 and its predecessor numbers are shown on the top line. Table 1.2, unlike Table 1.1, shows formulas for four predecessor numbers for each element of set (3.1). Table 1.1 and 1.2 shows the interconnectedness of all elements of the set (3.1) and the number 1.

Table 1.2. Calculating predecessor numbers with arithmetic progression formulas

		3+12t	1+24t	13+48t	5+96t	53+192t	21+384t	213+768t	85+1536t	853+3072t	341+6144t	3413+12288t	1365+24576t
		7+12t	9+24t	29+48t	37+96t	117+192t	149+384t	469+768t	597+1536t	1877+3072t	2389+6144t	7509+12288t	9557+24576t
		11+12t	17+24t	45+48t	69+96t	181+192t	277+384t	725+768t	1109+1536t	2901+3072t	4437+6144t	11605+12288t	17749+24576t
g		1	2	3	4	5	6	1	2	3	4	5	6
1			1		5		21		85		341		1365
5	1	3		13		53		213		853		3413	
7	2		9		37		149		597		2389		9557
11	3	7		29		117		469		1877		7509	
13	4		17		69		277		1109		4437		17749
17	5	11		45		181		725		2901		11605	
19	6		25		101		405		1621		6485		25941
23	1	15		61		245		981		3925		15701	
25	2		33		133		533		2133		8533		34133
29	3	19		77		309		1237		4949		19797	
31	4		41		165		661		2645		10581		42325
35	5	23		93		373		1493		5973		23893	
37	6		49		197		789		3157		12629		50517
41	1	27		109		437		1749		6997		27989	
43	2		57		229		917		3669		14677		58709
47	3	31		125		501		2005		8021		32085	
49	4		65		261		1045		4181		16725		66901
53	5	35		141		565		2261		9045		36181	
55	6		73		293		1173		4693		18773		75093

Please note that this is only the initial part of the table of infinite size, both rows and columns of the table are infinite.

Appendix 2. Relationships between Collatz numbers

2.1. Relationships between Collatz numbers

As shown in Appendix 1, numbers, depending on their predecessor numbers, can be divided into 6 types, with each type of numbers corresponding to two types of predecessor numbers. The six types of numbers are expressed by the following arithmetic progression

(2.1) $g_i = p_i + 18t$, where $p_i = 5, 7, 11, 13, 17, 19$.

Note. In Appendix 1, the sixth type of numbers was expressed by the arithmetic progression formula $g_6 = 1 + 18t$, which made it possible to show the numbers associated with the number 1. If we use formula (2.1), then the sixth type of number will be obtained with p_i =19 and t =0, i.e. in this case, in order for all numbers to be of the form $6n \mp 1$, we excluded the number 1.

Next, we will show that all Collatz numbers are associated with the initial six types of numbers of the form $g_i = 6n \mp 1$ and twelve predecessor numbers corresponding to the first power-of-two interval. In other words, it will be shown below that all Collatz numbers are related to the numbers given in Table 1.1.

It should be noted that in this chapter we show the relationship of the elements of the set (3.1) and two pairs of predecessor numbers for each element, although the number of predecessor numbers is infinite. At the same time, if necessary, it is possible to show the connection of other predecessor numbers with other elements of the set (3.1) using the method described below.

Figure 2.1 shows a diagram of the relationship between numbers of the form $g_i = 6n \mp 1$ and the first members of the arithmetic progression p_i, as well as two pairs of initial predecessor numbers k_{i1} and k_{i2} and other predecessor numbers k_{i1t} and k_{i2t}.

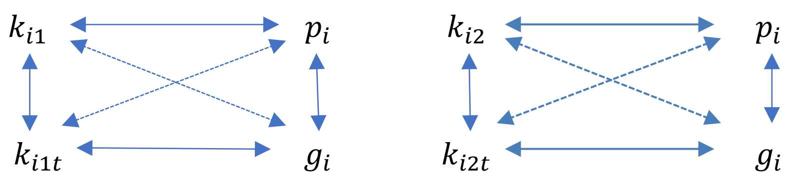

Figure 2.1. Diagram of the relationship between numbers and predecessor numbers

As follows from the diagram, based on the first numbers of the arithmetic progression p_i $(5, 7, 11, 13, 17, 19)$, you can calculate other numbers k_{i1}, k_{i2}, k_{i1t}, k_{i2t} and g_i, if the multiplier t $(t = 0, 1, 2, \ldots)$.

It is known that the calculation of the Collatz function is carried out in several iterations, and at each subsequent iteration, the number obtained at the previous iteration is used as an argument. This means that when calculating the Collatz function, the status of the numbers k_{ij} and g_i will change.

With such an operation, in some cases there are difficulties in establishing the types of numbers, therefore, in Table 2.1. the subtypes of the above 6 types of numbers and their preceding numbers are given.

Table 2.1 shows the transition of the formulas for k_{ij} (given in column 2 of the table) to the formulas g_i (given in column 9 of the table), that is, to the formulas of the arithmetic progression (2.1).

The formulas for the k_{ijl} subtypes in Table 2.1 are derived as follows. From the sequence of numbers obtained by the above 12 formulas of the arithmetic progression k_{ij}, $i = 1, 2, 3, 4, 5, 6$; $j = 1, 2$, we find the remainder modulo 18. Then each k_{ij} formula will fall into three subtypes, i.e. for 3 formulas. Therefore, in this case, instead of 12 k_{ij} formulas, we get 36 formulas of the form $k_{ijl} = p_i + 18t_l$; $i = 1, 2, 3, 4, 5, 6$; $j = 1, 2$; $l = 1, 2, 3$.

In this case, two formulas k_{ijl} will correspond to one formula g_{il}, while the form of six formulas does not change, only for each pair of formulas k_{ijl} there will correspond a formula g_{il} with a multiplier t_{il}, i.e. in this case, the factor t_{il} of the formula g_{il} depends on k_{ijl}, i.e. $t_{il} = f(k_{ijl})$.

Thus, the representation of numbers k_{ij} in the form

(2.2) $$k_{ijl} = p_i + 18t_{ijl},$$

as well as numbers g_{il} in the form

(2.3) $$g_{il} = p_i + 18t_{il},$$

allows you to establish more accurate links between the numbers k_i and g_i, through direct links of their subtypes k_{ijl} and g_{il}.

The exact connection between the numbers k_{ijl} and g_{il} is provided by the factors of these numbers t_{ijl} and t_{il}, which are determined by the formulas

(2.4) $$t_{ijl} = a_{ijl} + b_{ij} \cdot s;$$

(2.5) $$t_{il} = c_{il} + 3s,$$

where $s = 0, 1, 2, \ldots$ - multiplier coefficient (calculation number g_i minus 1); $c_{il} = 0,\ 1,\ 2.$

Since l only takes three values (1, 2, 3), i.e. there are three subtypes k_i, and any factor starts from 0 and takes natural values, then the formula for the factor of the number g_{il} has only three types

$t_{i1} = 0 + 3s$, $t_{i2} = 1 + 3s$, $t_{i3} = 2 + 3s$.

As can be seen from formulas (2.2) and (2.3), as well as (2.4) and (2.5), only if the coefficient *s* is equal, the two numbers k_{ijl} and g_{il} will correspond to each other.

The s coefficient formulas were established as follows.

First, after establishing the subtypes k_{ijl} by finding the residuals by divisor 18 from the sequence of numbers k_{ij}, the multiplier t_{ijl} is calculated. To do this, the calculated residuals by the divisor 18 are subtracted from the sequence of numbers k_{ij}, then divided by the number 18, then as a result of such operations, the multiplier t_{ijl} will be obtained.

After that, knowing that each k_{ijl} corresponds to the number g_{il}, comparing the numbers k_{ijl} and g_{il}, we find the coefficient *s* corresponding to them. Further, we represent the factor t_{ijl} in the form of formula (2.4). A system of 36 formulas showing the relationships between the numbers k_i and g_i, as well as their subtypes k_{ijl} and g_{il}, is shown in Table 4.1.

Table 2.1. System of 36 formulas for Collatz functions

Note: s = 0, 1, 2, ...

$\#k_{ij}$	k_{ij}	k_{ijl}	t_{ijl}	$\#k_{ijl}$	g_{il}	t_{il}	$\#g_{il}$	g_i
1	2	3	4	5	6	7	8	9
		$k_{111} = 13 + 18t_{111}$	0+8s	1.1.1	$g_{11} = 5 + 18t_{11}$	0+3s	1.1	
1.1	$k_{11} = 13 + 48t$	$k_{112} = 7 + 18t_{112}$	3+8s	1.1.2	$g_{12} = 5 + 18t_{12}$	1+3s	1.2	
		$k_{113} = 1 + 18t_{113}$	6+8s	1.1.3	$g_{13} = 5 + 18t_{13}$	2+3s	1.3	$g_1 = 5 + 18t$;
		$k_{121} = 17 + 18t_{121}$	2+32s	1.2.1	$g_{11} = 5 + 18t_{11}$	0+3s	1.1	(1)
1.2	$k_{12} = 53 + 192t$	$k_{122} = 11 + 18t_{122}$	13+32s	1.2.2	$g_{12} = 5 + 18t_{12}$	1+3s	1.2	
		$k_{123} = 5 + 18t_{123}$	24+32s	1.2.3	$g_{13} = 5 + 18t_{13}$	2+3s	1.3	
		$k_{211} = 1 + 18t_{211}$	2+16s	2.1.1	$g_{21} = 7 + 18t_1$	0+3s	2.1	
2.1	$k_{21} = 37 + 96t$	$k_{212} = 7 + 18t_{212}$	7+16s	2.1.2	$g_{22} = 7 + 18t_2$	1+3s	2.2	
		$k_{213} = 13 + 18t_{213}$	12+16s	2.1.3	$g_{23} = 7 + 18t_3$	2+3s	2.3	$g_2 = 7 + 18t$;
		$k_{221} = 5 + 18t_{221}$	8+64s	2.2.1	$g_{21} = 7 + 18t_1$	0+3s	2.1	(2)
2.2	$k_{22} = 149 + 384t$	$k_{222} = 11 + 18t_{222}$	29+64s	2.2.2	$g_{22} = 7 + 18t_2$	1+3s	2.2	
		$k_{223} = 17 + 18t_{223}$	50+64s	2.2.3	$g_{23} = 7 + 18t_3$	2+3s	2.3	

Continuation of Table 2.1

		$k_{311} = 7 + 18t_{311}$	0+2s	3.1.1	$g_{31} = 11 + 18t_{31}$	0+3s	3.1	$g_3 = 11 + 18t;$ (3)
3.1	$k_{31} = 7 + 12t$	$k_{312} = 1 + 18t_{312}$	1+2s	3.1.2	$g_{32} = 11 + 18t_{32}$	1+3s	3.2	
		$k_{313} = 13 + 18t_{313}$	1+2s	3.1.3	$g_{33} = 11 + 18t_{33}$	2+3s	3.3	
		$k_{321} = 11 + 18t_{321}$	1+8s	3.2.1	$g_{31} = 11 + 18t_{31}$	0+3s	3.1	
3.2	$k_{32} = 29 + 48t$	$k_{322} = 5 + 18t_{322}$	4+8s	3.2.2	$g_{32} = 11 + 18t_{32}$	1+3s	3.2	
		$k_{323} = 17 + 18t_{323}$	6+8s	3.2.3	$g_{33} = 11 + 18t_{33}$	2+3s	3.3	
		$k_{411} = 17 + 18t_{411}$	0+4s	4.1.1	$g_{41} = 13 + 18t_{41}$	0+3s	4.1	$g_4 = 13 + 18t;$ (4)
4.1	$k_{41} = 17 + 24t$	$k_{412} = 5 + 18t_{412}$	2+4s	4.1.2	$g_{42} = 13 + 18t_{42}$	1+3s	4.2	
		$k_{413} = 11 + 18t_{413}$	3+4s	4.1.3	$g_{43} = 13 + 18t_{43}$	2+3s	4.3	
		$k_{421} = 7 + 18t_{421}$	15+64s	4.2.1	$g_{41} = 13 + 18t_{41}$	0+3s	4.1	
4.2	$k_{42} = 277 + 384t$	$k_{422} = 13 + 18t_{422}$	36+64s	4.2.2	$g_{42} = 13 + 18t_{42}$	1+3s	4.2	
		$k_{423} = 1 + 18t_{423}$	58+64s	4.2.3	$g_{43} = 13 + 18t_{43}$	2+3s	4.3	

Continuation of Table 2.1

		$k_{511}=11+18t_{511}$	0+2s	5.1.1	$g_{51}=17+18t_1$	0+3s	5.1	$g_5=17+18t;$ (5)
5.2	$k_{51}=11+12t$	$k_{512}=5+18t_{512}$	1+2s	5.1.2	$g_{51}=17+18t_2$	1+3s	5.2	
		$k_{513}=17+18t_{513}$	1+2s	5.1.3	$g_{51}=17+18t_3$	2+3s	5.3	
		$k_{521}=1+18t_{521}$	10+32s	5.2.1	$g_{51}=17+18t_1$	0+3s	5.1	
5.2	$k_{52}=181+192t$	$k_{522}=13+18t_{522}$	20+32s	5.2.2	$g_{51}=17+18t_2$	1+3s	5.2	
		$k_{523}=7+18t_{523}$	31+32s	5.2.3	$g_{51}=17+18t_3$	2+3s	5.3	
		$k_{611}=7+18t_{611}$	1+4s	6.1.1	$g_{61}=19+18t_{11}$	0+3s	6.1	$g_6=19+18t;$ (6)
6.1	$k_{61}=25+24t$	$k_{612}=13+18t_{612}$	2+4s	6.1.2	$g_{62}=19+18t_{12}$	1+3s	6.2	
		$k_{613}=1+18t_{613}$	4+4s	6.1.3	$g_{63}=19+18t_{13}$	2+3s	6.3	
		$k_{621}=11+18t_{621}$	5+16s	6.2.1	$g_{61}=19+18t_{11}$	0+3s	6.1	
6.2	$k_{62}=101+96t$	$k_{622}=17+18t_{622}$	10+16s	6.2.2	$g_{62}=19+18t_{12}$	1+3s	6.2	
		$k_{623}=5+18t_{623}$	16+16s	6.2.3	$g_{63}=19+18t_{13}$	2+3s	1.3	

2.2. Examples

Let's show the connections between the numbers k_{ijl} and g_{il}, which belong to the same subtype.

Example 2.1. Let two numbers $k = 17$ and $g = 13$ be given, which are directly related to each other through the function $g = (3k + 1)/2^q$, i.e. if the number 17 is used to calculate the Collatz function, then as a result of the calculation we get the number 13.

According to Table 2.1, the numbers $k = 17$ and $g = 13$ are related to each other through their following subtypes $k_{411} = 17$ $(t_{411} = 0, s = 0)$ и $g_{41} = 13$ $(t_{41} = 0, s = 0)$.

Next, we will show how these two numbers 17 and 13 are related to other numbers related to these types of numbers, for this we will take the multiplier equal to $s = 1$, then using the formulas given in Table 4.1 we will obtain the following numbers

$k_{411} = 17 + 18t_{411}$; $s = 1 \to t_{411} = 0 + 4 \cdot 1 = 4$; $k_{411} = 17 + 18 \cdot 4 = 89$.

The number $k_{411} = 89$, with $s = 1$ corresponds to the number $g_{41} = 13 + 18t_{41}$, with the multiplier $t_{41} = 0 + 3 \cdot 1 = 3$; $g_{41} = 13 + 18 \cdot 3 = 67$.

Next, starting from the number 89, we calculate the Collatz numbers, then combine the resulting numbers into a common chain and get:

89→**67**→101→19→29→11→**17**→**13**→ 5→ 1.

As you can see, the numbers 89 and 67 form a common chain with the numbers 17 and 13.

Example 2.2. Of course, the larger the coefficient of the multiplier *s*, the greater will be the value of the numbers corresponding to the coefficient, so the length of the chain of numbers and the number of calculations is also greater. Below we will show an example based on the above numbers with $s = 4$.

$k_{411} = 17 + 18t_{411}$; $s = 4 \to t_{411} = 0 + 4 \cdot 4 = 16$; $k_{411} = 17 + 18 \cdot 16 = 305$.

The number $k_{411} = 305$, at $s = 4$ corresponds to the number $g_{41} = 13 + 18t_{41}$, with the multiplier $t_{41} = 0 + 3 \cdot 4 = 12$ therefore $g_{41} = 13 + 18 \cdot 12 = 229$.

Further, if we calculate the Collatz numbers starting from the number 305, then link the resulting numbers, we get the following chain of numbers:

305→ **229** → 43 → 65 → 49 → 37→7→11→**17**→**13**→ 5→ 1.

As you can see, the numbers 305 and 229 also form a common chain with the numbers 17 and 13.

It should be noted that sometimes pairs of numbers belonging to the same subtype do not always form a common chain of numbers. This is due to the fact that for different coefficients s the subtypes of numbers differ, in addition, changing the statuses of the numbers k_(ijl) and g_il will lead to an even greater change in the chain of numbers. The above is demonstrated by example 2.3.

Example 2.3. Below we show an example based on the above numbers for s = 5.

$k_{411} = 17 + 18t_{411}$; $s = 5 \rightarrow t_{411} = 0 + 4 \cdot 5 = 20$; $k_{411} = 17 + 18 \cdot 20 = 377$.

The number k_{411} =377 corresponds to the number $g_{41} = 13 + 18t_{41}$, with the factor $t_{41} = 0 + 3 \cdot 5 = 15$ for s = 5, so $g_{41} = 13 + 18 \cdot 15 = 283$.

Further, if you calculate the Collatz numbers, starting with the number 377, then combine the resulting numbers, you get the following chain of numbers:

377 → **283** → 425 → 319 → 479 → 719 → 1079 → 1619 → 2429 → 911 → 1367 → 2051 → 3077 → 577 → 433 → 325 → 61 → 23 → 35 → 395 → 53 → 5 → 1.

In this case, the numbers 377 and 283 do not form a common chain with the numbers 17 and 13, although they belong to the same subtype, since there is a branch formed by the number 53, which is the predecessor of the number 5, like the number 13.

Thus, the main meaning of the formulas given in Table 2.1 is that they show that all the numbers k_{ijl} and g_{il} are related to the number 5, which, in turn, is related to the number 1.