IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 17, Issue 4 Ser. I (Jul. – Aug. 2021), PP 43-53
www.iosrjournals.org

Investigation of the Collatz problem

Vadim Nikolayevich Romanov
Doctor of Technical Sciences, Professor
St.-Petersburg, Russia; vromanvpi@mail.ru

Abstract
The article analyzes the properties of the sequences 3k – 1 and 3k + 1 at their separate and joint use in iterative processes. The Collatz problem is considered as a special case of the problem of determining the optimal iterative process using both sequences 3k – 1 and 3k +1, which achieves 1 in the minimum number of steps (iterations). It is proved that the process P_2*, using the sequence 3k + 1, cannot diverge or go in loops, so it always reaches 1, but in general this requires a large number of iterations. Process* P_1*, using the sequence 3k – 1, cannot diverge, but can go in loops. An estimate is obtained for the number of iterations required to establish the absence of divergence of the processes* P_1 *and* P_2*.*

Keywords and phrases: *Collatz problem, problem of representation of integers, iterative process, optimal algorithm.*
Classification code: 11A67

Date of Submission: 28-06-2021 Date of Acceptance: 12-07-2021

I. Introduction

The Collatz problem, or the $3k + 1$ problem, is one of the unsolved problems in number theory. It is as follows. Take an arbitrary natural number. If it is even, then we divide it by 2, and if it is odd, then we multiply it by 3 and add 1. This process repeats with the resulting number. It is required to find out whether, in this case, 1 is always achieved in a certain (finite) number of iterations. Numerous computer calculations have been performed to verify the correctness of this statement, but the question remains open. In this article, we prove that the iterative process described above always achieves 1, and give an explanation of this phenomenon. Consider the Collatz problem as a special case of the more general problem of determining an iterative process that makes it possible to reach 1 in the least number of steps (iterations), if we start with an arbitrary natural number. We compare three iterative processes: $P_1 = 3k – 1$, $P_2 = 3k + 1$ and the combined process $P_3 = [(3k – 1) - (3k + 1)]$. The P_1 process is based on the sequence $F_1 = (3k – 1)$, the P_2 process is based on the sequence $F_2 = (3k + 1)$, the combined process P_3 alternately uses the sequences $(3k – 1)$ and $(3k + 1)$. Consider the properties of these sequences for different values of k. It is easy to verify (see [1, 2]) that the triplet of numbers ($x = 6k$, $y = 9k^2 – 1$, $z = 9k^2 + 1$), where k is an even number, is a solution of Fermat's quadratic equation $x^2 + y^2 = z^2$. Similarly, the triplet of numbers ($x = 3k$, $y = (9k^2 – 1)/2$, $z = (9k^2 + 1)/2$), where k is an odd number, is also a solution of Fermat's quadratic equation. The quantity $y = 9k^2 – 1$ can be represented as the product $y = (3k – 1)(3k + 1)$. For even k, the sequences $(3k – 1)$ and $(3k + 1)$ allow us to obtain all odd numbers, including primes, except for 3 and multiples of 3. For even k, the sequence $(3k – 1)$ is the arithmetic progression of the form $(3k – 1) = 5 + 6l$, and the sequence $(3k + 1)$ is the arithmetic progression of the form $(3k + 1) = 7 + 6l$, where $l = 0, 1, 2, 3$, etc. The quantity $y = (9k^2 – 1)/2$ can be represented as the product $y = 1/2(3k – 1)(3k + 1)$. In the expression $y = 1/2(3k – 1)(3k + 1)$, if we divide by 2 a factor that is divisible only by 2, then we obtain an odd number. The second factor will be an even number divisible by some power of 2. For odd k, the sequences $(3k – 1)$ and $(3k + 1)$ allow us to obtain all even numbers, except for multiples of 3. For odd k, the sequence $(3k – 1)$ is the arithmetic progression of the form $(3k – 1) = 2 + 6l$, and the sequence $(3k + 1)$ is the arithmetic progression of the form $(3k + 1) = 4 + 6l$, where $l = 0, 1, 2, 3$, etc. In the sequence F_1, even numbers with even l are divisible only by 2, and numbers with odd l are divisible by 4 or a higher power of 2. For $l = 1, 5, 9, 13$, etc., even numbers are divisible by 8. For $l = 5, 13, 21, 29$, etc., even numbers are divisible by 16, and so on. In the sequence F_2, numbers with odd l are divisible only by 2, and numbers with even l are divisible by 4 or a higher power of 2. For $l = 2, 6, 10, 14$, etc., even numbers are divisible by 8; for $l = 2, 10, 18$, etc., even numbers are divisible by 16, and so on. Therefore, in sequences F_1 and F_2, even numbers divisible by some power of 2 occur as often as numbers divisible only by 2. The even numbers in these sequences are shifted relative to each other by 2. For every odd k, one of the numbers $(3k – 1)$ or $(3k + 1)$ is divisible only by 2, and the second of them is divisible by some power of 2, i.e. by 4, 8, etc. depending on the value of k. When the number k changes, the numbers in the sequences F_1 and F_2 are alternately divisible only by 2 or a power of the number 2. We use the properties of the sequences $(3k - 1)$ and $(3k + 1)$ for a comparative analysis of the iterative processes P_1, P_2 and P_3. This analysis allows us to understand and explain the Collatz problem.

For convenience, we introduce some definitions. The smallest odd number, which is the beginning and the end of the loop, we call the center of attraction. The set of numbers obtained in an iterative process that ends with a center of attraction we call the area of attraction of this center. If in the iterative process that uses only one sequence $(3k- 1)$ or $(3k + 1)$ and starts with an arbitrary odd number k_0, there is at least one number from the area of attraction of the given center, then we will say that the number k_0 is included in the area of attraction of this center. If an iterative process starts with an arbitrary odd number k_0 and ends with 1, then we will say that k_0 is included in the area of attraction of the number 1. The set of numbers in an iterative process, starting with k_0 and ending with the center of attraction, we call the trajectory corresponding to the number k_0. The reason that does not allow reaching 1 is the presence of a center of attraction other than 1. This center can be a finite odd number (one or more) or infinity. An odd number, not equal to 1 or infinity, that is the center of attraction, forms a loop. The number 1 and infinity do not form loops (loops are degenerate). If the center of attraction is 1, then we say that the iterative process converges. If the center of attraction is infinity, then we say that the iterative process diverges. If the center of attraction is a finite number forming a loop, then we say that the iterative process has looping. We are going to prove that the process P_2 cannot diverge or go in loops.

II. Study of iterative process P_3

We formulate the problem in the following form: "It is required to determine a combined iterative process P_3 that uses alternately both sequences F_1 and F_2 and allows achieving 1 in the least number of steps (iterations) if we start with an arbitrary odd number k". Considering only odd numbers does not reduce the generality, since by consecutive division of even number by 2, we can always get an odd number. If at some step we get an even number equal to a power of 2, i.e. 2^n, then the problem will be solved. The sequence $(3k - 1)$ contains odd powers of 2: $2^1 = 2$ ($k = 1$), $2^3 = 8$ ($k = 3$), $2^5 = 32$ ($k = 11$), $2^7 = 128$ ($k = 43$), $2^9 = 512$ ($k = 171$), $2^{11} = 2048$ ($k = 683$), $2^{13} = 8192$ ($k = 2731$), etc. The sequence $(3k + 1)$ contains even powers of 2: $2^2 = 4$ ($k = 1$), $2^4 = 16$ ($k = 5$), $2^6 = 64$ ($k = 21$), $2^8 = 256$ ($k = 85$), $2^{10} = 1024$ ($k = 341$), $2^{12} = 4096$ ($k = 1365$), $2^{14} = 16384$ ($k = 5461$), etc. Let k is an odd number. Three cases are possible: 1). $k = (2^n + 1)/3$; 2). $k = (2^n - 1)/3$; 3). k is an arbitrary odd number not corresponding to the first or second case. In the first case, the values of k correspond to odd powers of 2, therefore, the process P_1, based on the sequence $(3k-1)$, immediately leads to the desired result, since $(3k-1) = 2^n$ and consecutive division by 2 gives 1; the result is achieved in the least number of steps. In the second case, the values of k correspond to even powers of 2, therefore, the process P_2, based on the sequence $(3k+1)$, immediately leads to the desired result, since $(3k+ 1) = 2^n$ and consecutive division by 2 gives 1; the result is achieved in the least number of steps. Consider the third (general) case.

Lemma 1 is valid: "Combined iterative process P_3, using alternately both sequences $(3k- 1)$ and $(3k+ 1)$, always achieves 1 if at each step we choose such a sequence $(3k- 1)$ or $(3k+ 1)$, which gives a decrease of initial number k". We prove that such a strategy exists. Indeed. Let k is an arbitrary odd number. Put $k = 2t - 1$, where $t = 1, 2, 3$, etc., then $(3k- 1) = 2(3t - 2)$. Therefore, the number $(3k- 1)$ is divisible only by 2, if t is an odd number, and it is divisible by some power of the number 2, if t is an even number. For $(3k + 1)$ we obtain $(3k+ 1) = 2(3t - 1)$. Therefore, the number $(3k + 1)$ is divisible only by 2, if t is an even number, and is divisible by some power of the number 2, if t is an odd number. If $(3k- 1)$ or $(3k + 1)$ is divisible only by 2, then after division by 2, we obtain an odd number k_1, that is more than the initial number k. Indeed, for $(3k- 1)$ when t is odd, we have after division by 2: $k_1 = (3t -2) \geq k = (2t - 1)$. Equality takes place only when $k_1 = k = 1$ for $t = 1$, but then $(3k-1) = 2$, $k = (2^n + 1)/3$ for $n = 1$, and we have the first case. Similarly, for $(3k + 1)$ when t is even, we have after division by 2: $k_1 = (3t - 1) > k = (2t - 1)$. If $(3k- 1)$ or $(3k + 1)$ is divisible by some power of the number 2, that is, at least by 4, then after division by this power, we obtain an odd number k_1 that is less than the initial number k. Indeed, for $(3k- 1)$, when t is even, we put $t = 2l$, where $l = 1, 2, 3$, etc. After division by the power of 2, i.e. by 4, we have $k_1 = (3l -1) < k = (4l -1)$. Similarly, for $(3k + 1)$, when t is odd, we put $t = 2l - 1$, where $l = 1, 2, 3$, etc. After division by the power of 2, i.e. by 4, we obtain $k_1 = (3l - 2) \leq k = (4l -3)$. Equality takes place only when $k_1 = k = 1$ for $t = l = 1$, but then $(3k+ 1) = 4$, $k = (2^n - 1)/3$ for $n = 2$, and we have the second case. Since the smallest odd number is 1, then after consecutive decreasing an arbitrary odd number k using the process described above, we always achieve 1. So we have proved Lemma 1.

It follows from the course of the proof of Lemma 1 that using only one sequence $(3k-1)$ or $(3k+ 1)$ in the iterative process increases the number of steps required to achieve 1, i.e. this strategy is not optimal. We are not saying here that in this case it is always possible to achieve 1. Thus, in the general case, the optimal strategy is to use at each step such a sequence $(3k- 1)$ or $(3k + 1)$ that gives a decrease of the initial odd number k.

We consider three examples. Put $k = 2731$, then we have the first case, since $(3k-1) = 2^{13}$. By sequentially dividing this number by 2 we achieve 1. It takes 14 operations to achieve 1, where operation means calculating a number $(3k- 1)$ or dividing by 2. If the sequence $(3k + 1)$ is used in this case, approximately 114 operations is required to achieve 1. Put $k = 5461$, then we have the second case, since $(3k+ 1) = 2^{14}$. It takes 15 operations to achieve 1, where operation means calculating a number $(3k + 1)$ or dividing by 2. Process P_1, using the sequence $(3k - 1)$, also achieves 1, but the number of operations is approximately 78. Put $k = 107$, which corresponds to the third case. We write the iterative process P_3 in detail. We calculate $(3k - 1) = 320$ and $320/64 = 5$; $(3k + 1) = 322$ and $322/2 = 161$. Put $k = 5$, then $(3k - 1) = 14$ and $14/2 = 7$; $(3k + 1) = 16$ and $16/16 = 1$ (end of procedure). If we use the sequence $(3k - 1)$, then for $k = 7$ we have $(3k - 1) = 20$ and $20/4 = 5$, i.e. a loop is obtained and reaching 1 is impossible. When we use only the sequence $(3k + 1)$, much more operations (about 100) are required to achieve 1. In this case, the combined iterative process P_3 is optimal in the number of operations. Using the sequence $(3k - 1)$ leads to looping at some k, which does not allow reaching 1.

Now we determine whether the process P_3 can diverge or go in loops, and under what conditions. It follows from the course of the proof of Lemma 1 and the properties of the sequences F_1 and F_2 that the process P_3 can diverge and does not reach 1 if at each step, we choose such a sequence $F_1 = (3k - 1)$ or $F_2 = (3k + 1)$, which gives an increase of the initial number k. However, at arbitrary step of this divergent process, that is, for arbitrary k, we can make the process P_3 reach 1 if we return to the conditions of Lemma 1. Since the iterative process P_3 always achieves 1 under the conditions of Lemma 1, the numbers from the sequences F_1 and F_2 in this iterative process cannot simultaneously at the same number k belong to the area of attraction of infinity. Therefore, processes P_1 and P_2 cannot simultaneously diverge.

It follows from the course of the proof of Lemma 1 and the properties of the sequences F_1 and F_2 that the processes P_1 and P_2 cannot have the same number k as their center of attraction. Therefore, the process P_3 cannot have looping.

Let us find out whether one of the processes P_1 or P_2 can diverge, if the process P_3 achieves 1. Since the process P_1 can have looping for some values of k (see below), the process P_2 cannot diverge simultaneously for the same k. Otherwise, the process P_3 cannot reach 1 if the conditions of Lemma 1 are satisfied. From the properties of the sequences F_1 and F_2 considered above, it follows that the numbers in each of them are alternately divisible either by 2 or by a power of 2. Therefore, when these sequences are used in an iterative process, none of them has numbers that are divisible only by 2 in an infinite segment of the sequence. We can only talk about the divergence, on average, in the set of finite segments of each sequence. The length of the segment depends on the ending of the number k and its value. This kind of divergence means that on an infinite number of segments of the sequence $F_2(k)$, numbers divisible only by 2 will, on average, prevail over numbers divisible by powers of 2. At the same time, the opposite picture is observed for the sequence $F_1(k)$, which follows from the properties of the sequences F_1 and F_2 discussed above. If in given segment, $F_2(k)$ is divisible only by 2 and initial k increases, then in the same segment, $F_1(k)$ with the same value of k is divisible by a power of 2, so that initial k decreases. We will show that the divergence of the processes P_1 and P_2 is impossible, since this contradicts the properties of the sequences F_1 and F_2.

III. Proof of the Lemma on the divergence of the processes P_1 and P_2

Lemma 2 is valid: "The iterative process P_2 (respectively, P_1) cannot diverge, i.e. cannot have infinity as the center of attraction". We summarize the properties considered above of the sequences F_1 and F_2, which we use in the proof. For odd k, even numbers in these sequences that divisible only by 2 or a power of 2 are evenly distributed and occur equally often. Even numbers are shifted relative to each other by 2, so they are alternately divisible only by 2 or by a power of 2. When k is even, odd numbers are shifted relative to each other by 2. In the iterative process P_2 (respectively, P_1) after each division by 2 of an even number l belonging to one of the sequences F_1 or F_2, the quotient $l/2$ changes its belonging to the sequence from F_1 to F_2 or from F_2 to F_1, respectively. If an even number l belongs to the sequence F_1 or, respectively, to the sequence F_2 and is divisible only by 2, then the odd number $l/2$ obtained after dividing by 2 will belong to the sequence F_2, or, respectively, to the sequence F_1, but already for even k. If an even number l belongs to the sequence F_1 or, respectively, to the sequence F_2 and is divisible by the power of the number 2, then an even number $l/2$ will belong to the sequence F_2, respectively, to the sequence F_1; the number $l/4$ will belong to the sequence F_1, respectively, to the sequence F_2, etc.

Now, we shall prove Lemma 2 by induction. In the interval from 2^1 to 2^2, there is one odd number 3. It is easy to verify that $P_2(3)$ and $P_1(3)$ do not diverge, so Lemma 2 is valid for this interval. Take the interval $2^n \ldots 2^{n+1}$. Suppose that for an arbitrary initial odd number k_0 from this interval the process $P_2(k_0)$, as well as $P_1(k_0)$, does not diverge, i.e. cannot have infinity as the center of attraction. In addition, for process P_1, a looping can also occur at a finite value of k. Suppose that the Lemma 2 is valid for all intervals $2^n \ldots 2^{n+1}$, if $n \le m$. We shall prove the validity of Lemma 2 for an arbitrary initial odd number from the interval $2^{n+1} \ldots 2^{n+2}$, if $n = m$. With this approach, we do not need to prove the convergence of the process P_2 (respectively, P_1); it is enough to show that after a finite number of iterations we reach the odd number k, for which Lemma 2 is valid. In the considered interval, there are 2^n odd numbers from $2^{n+1} + 1$ to $2^{n+2} - 1$. An arbitrary odd number k_0 from the interval $2^{n+1} \ldots 2^{n+2}$ can be represented in two forms $k_0 = 2^{n+1} + a = 2^{n+2} - b$, where a and b are odd numbers Between a and b there is the relation $a + b = 2^{n+1}$. We write a and b in the form $a = 2t_1 - 1$, $b = 2t_2 - 1$, where t_1, t_2 are natural numbers. From this we obtain the relation $t_1 + t_2 = 2^n + 1$. Hence, t_1 and t_2 have different parity. We use both of these representations.

We perform calculations for the smallest and the largest values of k from the interval $2^{n+1} \dots 2^{n+2}$. Put $k_0 = 2^{n+1} + 1$, then $P_2(k_0) = 3\cdot2^{n-1} + 1 = k_1 < k = 2^{n+1} + 1$. The number k_1 belongs to the interval $2^n \dots 2^{n+1}$; therefore, by assumption, the process $P_2(k_1)$ does not diverge. If we continue the process P_2, then we obtain a series of decreasing values of k on the given segment of the process P_2. Thus, the Lemma 2 is valid for process $P_2(k_0)$, if k_0 is the smallest odd number from the considered interval. For this value of k, verification of the validity of the Lemma 2 requires only one iteration. For the largest value $k = 2^{n+2} - 1$ from the considered interval, the calculations are more laborious. Put $k_0 = 2^{n+2} - 1$, then $P_2(k_0) = 3\cdot2^{n+1} - 1 = k_1 > 2^{n+1} - 1$. If we continue the process P_2, then we obtain a series of increasing values of k on the given segment of the process P_2. On the segment of the process P_2, consisting of $n + 1$ iterations, the function F_2 is divisible only by 2. It is the largest segment of the process P_2 in considered interval with this property. After $n + 2$ steps (iterations), we reach the number $k^{n+2} = 3^{n+2} - 1$. We can write the last expression in the form $k^{n+2} = F_1(k) = 3k - 1$, where $k = 3^{n+1}$. $F_1(k)$ is divisible only by 2 if n is an odd number, and $F_1(k)$ is divisible at least by 8 or by a higher power of 2 (depending on the value of k) if n is an even number.

We determine the position (order of magnitude) of the number k^{n+2}. Numerical estimates give $k^{n+2} < 2^{(n+2)\ln3/\ln2} < 2^{1,6(n+2)}$ (upper estimate), then if n is an odd number, $k^{n+2}/2 < 2^{1,6(n+2)}/2$. We continue the process P_2, taking $k^{n+2}/2$ as the initial value. For brevity, we call an iteration consisting of calculating $3k + 1$ followed by dividing only by 2, a single iteration, and an iteration consisting of calculating $3k + 1$, then dividing by a power of 2, a multiple iteration. When we perform a single iteration, the odd number increases approximately 3/2 times; when we perform multiple iteration with division by 4, the odd number decreases 3/4 times; when we perform multiple iteration with division by 8, the odd number decreases 3/8 times, and so on. It follows from the properties of the sequences F_1 and F_2 that the value of k will alternately increase and then decrease, and on a sufficiently large but finite segment of the process P_2 we reach a value k_N, for which the process P_2 cannot diverge. For k_N, four cases (conditions) are possible. 1). k_N is less than $2^{n+1} + 1$, and for it the Lemma 2, by assumption, is valid. 2). k_N belongs to the process $P_2(k_0)$, which does not diverge (k_N itself does not belong to the previous interval). 3). Process $P_1(k_N)$ has looping, i.e. k_N belongs to the trajectory of numbers 5 or 17 (see below). 4). $F_2(k_N)$ is equal to a power of 2; for the process $P_1(k^{n+1}/2)$, the function $F_2(k_N)$ is replaced by $F_1(k_N)$. These conditions are also valid for the process $P_1(k^{n+1}/2)$. In all these cases, the process P_2 cannot diverge. The first case occurs frequently, and the last case is rare. We use the first case as the main one in the proof.

We estimate the number of iterations required to reach k_N from the previous interval. Consider the "worst" case when n is an odd number. For even $n + 2$, according to (1), the number of iterations required to reach k_N is not more than for odd $n + 1$. Per one iteration, in which F_2 and F_1 are divisible only by 2, the odd number is increased 3/2 times. So, the odd number $k^{n+2}/2 = (3^{n+2} - 1)/2$ is more than the initial number $2^{n+2} - 1$, approximately $1/2\cdot(3/2)^{n+2}$ times, and it is more than the number $k_N = 2^{n+1} - 1 < 2^{n+1} + 1$, approximately $(3/2)^{n+2}$ times. We determine the number of iterations required to return to the number $2^{n+1} - 1$ in the iterative process P_2, starting at number $k^{n+2}/2$. From the properties of even numbers and the sequences F_1 and F_2 considered above, it follows that each four of consecutive even numbers (tetrad) contain 2 numbers divisible only by 2, one number divisible by 4 and one number divisible by 8 or a higher power of 2. Each eight consecutive even numbers contain four numbers divisible by 2, two numbers divisible by 4, one number divisible by 8 and one number divisible by 16. If n is a sufficiently large number, then these properties (ratios) are, on average, valid for an arbitrary sample of four, respectively, 8 even numbers. A segment of the iterative process P_2, consisting of four iterations, in which F_2 and F_1 take values from tetrad of even numbers with the indicated properties, gives a decrease of the initial odd number at least $3^4/2^7$ times, i.e. about 3/5 times. Consequently, one iteration in this case gives, on average, the decrease $(3/5)^{1/4}$ times. For a segment of the iterative process P_2, consisting of eight iterations, the corresponding estimates are, respectively, $3^8/2^{15}$ or $3^8/2^{14}$ and for one iteration $(1/5)^{1/8}$ or $(2/5)^{1/8}$. We use segments with four iterations to obtain estimates that are more practical. The equation for determining the required number of iterations x has the form $(3/2)^{n+2}\cdot(3/5)^{x/4} = 1$. The solution of this equation is $x = [4(n+2)\ln 3/2]/(\ln 5/3)$. After substitution of numerical values, we obtain $x = 3.2(n+2)$. So, the total number of iterations required to reach the number k_N is equal to $x_0 = (n + 2) + x = 4.2(n + 2) = O(n)$. The number x_0 consists of two summands. Firstly from odd number $2^{n+2} - 1$ we reach odd number $(3^{n+2} - 1)/2$ ($n + 2$ iterations), than from $(3^{n+2} - 1)/2$ we reach number $k_N = 2^{n+1} - 1$ (x iterations). Put $n = 2^m + 1$, then number of iterations to reach k_N can be large, although finite. In the considered case, we cannot prove Lemma 2 directly by calculations, although it follows from the properties of the functions F_1 and F_2 that the value of k_N is always achieved in a finite number of iterations. Below we obtain general relations that allow us to prove Lemma 2 in this case as well.

Now consider the process P_1 with the same values of k as for the process P_2. We put $k_0 = 2^{n+2} - 1$. The first iteration gives $P_1(k_0) = 3\cdot2^n - 1 = k_1 > 2^{n+1} + 1$. The second iteration gives $P_1(k_1) = 3^2\cdot2^{n-2} -$

$1 = k_2 > 2^{n+1} + 1$. The third iteration gives $P_1(k_2) = 3^3 \cdot 2^{n-4} - 1 = k_3 < 2^{n+1} + 1$. The odd number k_3 belongs to the interval 2^n ... 2^{n+1}; therefore, by assumption, the process $P_1(k_3)$ does not diverge. Thus, the Lemma 2 is valid for process $P_1(k_0)$, if k_0 is the largest odd number from the considered interval. The result for $P_1(2^{n+2} - 1)$ is similar to the result for $P_2(2^{n+1} + 1)$, but the proof requires slightly more iterations (three instead of one). Put $k_0 = 2^{n+1} + 1$. The proof is similar to the proof for $P_2(2^{n+2} - 1)$ and requires, as for P_2, much more steps (iterations). The first iteration gives $P_1(2^{n+1} + 1) = 3 \cdot 2^n + 1 = k_1 > 2^{n+1} + 1$. If we continue the process P_1, then we obtain a series of increasing values of k on the given segment of the process P_1. On the segment of the process P_1, consisting of n iterations, the function F_1 is divisible only by 2. It is the largest segment of the process P_1 in considered interval with this property. After $n + 1$ steps (iterations) we reach the number $k^{n+1} = 3^{n+1} + 1$. We can write the last expression in the form $k^{n+1} = F_2(k) = 3k + 1$, where $k = 3^n$. $F_2(k)$ is divisible only by 2 if n is an odd number, and $F_2(k)$ is divisible by 4 if n is an even number. If we continue the process P_1, taking $k^{n+1}/2$ as the initial value, then for odd n, we obtain the result similar to the result for process $P_2(k^{n+2}/2)$. The behavior of the process P_1 is determined by the properties of the functions F_1 and F_2. Therefore, after some finite number of iterations, we reach the value k_N for which Lemma 2 is valid, but we cannot prove this directly by calculations. For odd $n + 2$, according to (2), the number of iterations is not more than for even $n + 1$. According to (3), for the process $P_1(2^{n+1} + 1)$ in the considered case, when $n + 1$ is an even number, the estimate of total number of iterations required to reach the number k_N does not exceed the estimate obtained for the process $P_2(2^{n+2} - 1)$, when $n + 2$ is an odd number. So, the result for $P_1(2^{n+2} - 1)$ is similar to the result for $P_2(2^{n+1} + 1)$, and the result for $P_1(2^{n+1} + 1)$ is similar to the result for $P_2(2^{n+2} - 1)$. This is explained by the fact that the initial value 2^{n+1} of the considered interval differs from the final (last) value 2^{n+2} of the interval by a factor 2. When we divide or multiply by 2, as follows from the properties of the functions F_1 and F_2, the form of the sequence changes, i.e. F_1 goes to F_2 and vice versa, depending on the parity of n.

We write relations that establish the connection between the processes $P_1(k^{n+1}/2)$ and $P_2(k^{n+2}/2)$. If n is an even number, then the following relation is valid

$$(3^{n+2} - 1)/2 = 3[(3^{n+1} - 1)/2] + 1 . \tag{1}$$

If n is an odd number, then the following relation is valid

$$(3^{n+2} + 1)/2 = 3[(3^{n+1} + 1)/2] - 1 . \tag{2}$$

If n is an odd number, then we have the relation

$$\{3[(3^{n+2} - 1)/2] + 1\}/2 = 3\{[3((3^{n+1} + 1)/2) - 1]/2\} - 1 . \tag{3}$$

Equality (1) reduces the process $P_2(k^{n+2}/2)$ to the process $P_2(k^{n+1}/2)$ for even n. Equality (2) reduces the process $P_1(k^{n+2}/2)$ to the process $P_1(k^{n+1}/2)$ for odd n. Equality (3) reduces the process $P_2(k^{n+2}/2)$ to the process $P_1(k^{n+1}/2)$ for odd n.

We also give results on the convergence of the process $P_1(k^{n+1})$ and the presence of loops for different n. For the interval 2^2 ... 2^3, the number $3^2 + 1 = 10$ belongs to the area of attraction of the number 5. For the interval 2^3 ... 2^4, according to (2), the number $3^3 + 1 = 28$ belongs to the area of attraction of the number 5. For the interval 2^4 ... 2^5, the number $3^4 + 1 = 82$ belongs to the area of attraction of the number 17. For the interval 2^5 ... 2^6, according to (2), the number $3^5 + 1 = 244$ belongs to the area of attraction of the number 17. For the interval 2^6 ... 2^7, the number $3^6 + 1$ belongs to the area of attraction of the number 1. For the interval 2^7 ... 2^8, according to (2), the number $3^7 + 1$ belongs to the area of attraction of the number 1. For the interval 2^8 ... 2^9, the number $3^8 + 1$ belongs to the area of attraction of the number 17. For the interval 2^9 ... 2^{10}, according to (2), the number $3^9 + 1$ belongs to the area of attraction of the number 17. For the interval 2^{10} ... 2^{11}, the number $3^{10} + 1$ belongs to the area of attraction of the number 17 and so on. For the interval 2^{14} ... 2^{15}, the number $3^{14} + 1$ belongs to the area of attraction of the number 1. For the interval 2^{15} ... 2^{16}, according to (2), the number $3^{15} + 1$ belongs to the area of attraction of the number 1. We can assume that this tendency remains. So the process $P_1(k^{n+1})$ converges to 1 for pairs of numbers $(3^6 + 1)$ and $(3^7 + 1)$, $(3^{14} + 1)$ and $(3^{15} + 1)$, etc., separated from each other by large intervals, and for the rest of the numbers of this type, the process $P_1(k^{n+1})$ has looping with the center of attraction equal to 17 (see Section IV below). In all above examples, the number of iterations required to prove Lemma 2 is less or much less than the estimate x_0 obtained above.

Using the above analysis, we prove Lemma 2 for an arbitrary odd k_0 from the interval 2^{n+1} ... 2^{n+2}. Consider the process $P_2(k_0)$. Even numbers can be disregarded, since after reduction by 2 or a power of 2, we reach an odd number k, which belongs to the previous interval, and for this k, Lemma 2 is valid by assumption. Indeed, the largest even number in the considered interval is $2^{n+2} - 2$. Put $a = 2^{n+1} - 2$, then $k_0 = 2^{n+2} - 2$, and after dividing by 2, we reach $k = 2^{n+1} - 1 < 2^{n+1}$, i.e. the number k belongs to the previous interval for which, by assumption, Lemma 2 is valid. We put $k_0 = 2^{n+1} + a$, where a is an arbitrary odd number from the interval 2^{n+1} ... 2^{n+2}; $1 \le a \le 2^{n+1} - 1$. We write the number a in the form $a = 2t - 1$, where t can take values 1, 2, 3, … , 2^n. The number of iterations required to prove Lemma 2 mainly depends on the parity of the number n and t and on the position of the initial odd number in the considered interval. From the subsequent analysis, we will see that the number of iterations does not exceed the estimate x_0 obtained above. Let t is an even number. We put $t = 2l$, where $l = 1, 2$, etc., 2^{n-1}. Then $a = 4l - 1$. The values of a form the arithmetic progression $3 + 4l_1$, where $l_1 = 0, 1, 2$, etc., $2^{n-1} - 1$. For small values of $t << 2^n$, when $a << 2^{n+1} - 1$, the behavior of the process $P_2(2^{n+1} + a)$ is determined by the properties of the process $P_2(a)$. It can be verified by direct calculations that the value k_N from the previous interval is reached in a finite number of iterations. In particular, for $t = 2$, already the second iteration leads to the required result. For large values of a, put $a = 2^{n+1} - b$, where b corresponds already to odd values of t_1. We put $t_1 = 2l - 1$, where $l = 1, 2, 3$, etc. Then $b = 4l - 3$ and takes the values 1, 5, 9, 13, etc., forming an arithmetic progression $1 + 4l_2$, where $l_2 = 0, 1, 2$, etc. In this case, for small values of l_2, when $b << 2^{n+1} - 1$, the behavior of the process $P_2(2^{n+1} + a)$ is determined by the properties of the process $P_2(-b) = -P_1(b)$. For the values of b, which are the centers of attraction for the process $P_1(b)$, i.e. for $b = 1, 5, 17$, as well as for odd numbers from the area of attraction of centers 5 or 17, obtained during subsequent iterations, the number of iterations required to reach k_N from the previous interval can be large, but it certainly does not exceed the above estimate x_0. For the largest even number $t = 2^n$, which corresponds to $b = 1$, we have the case already considered above. The length of the chain of iterations, on which $F_2(a)$ is divisible only by 2, varies from 1 at $t = 2$ to $n + 1$ at $t = 2^n$. In the latter case, the above analysis for $P_2(2^{n+2} - 1)$ remains valid, and we cannot prove Lemma 2 directly by computations.

We obtain the general relation allowing us to prove Lemma 2 for even values of t. To do this, we represent the first iteration $P_2(2^{n+1} + a)$ as a sum of terms that are divisible only by 2. We have $P_2(2^{n+1} + a) = (3 \cdot 2^{n+1} + 3a + 1)/2 = [(3 \cdot 2^{n+1} - 3 + 1) + (3a + 1) + 2]/2$. For even values of t, the terms $(3 \cdot 2^{n+1} - 3 + 1)$ and $(3a + 1)$ are divisible only by 2. After cancellation by 2, we obtain the following relation

$$P_2(2^{n+1} + a) = P_2(2^{n+1} - 1) + P_2(a) + 1 \quad (4)$$

For $P_2(2^{n+1} - 1)$, Lemma 2 is valid, by assumption, since $2^{n+1} - 1$ belongs to the previous interval. $P_2(a)$ does not diverge, since a belongs to one of previous intervals; $3 \le a \le 2^{n+1} - 1$. $P_2(a) = 5 + 6l_1$; $5 \le P_2(a) \le 3 \cdot 2^n - 1$. Therefore, the process $P_2(2^{n+1} + a)$ cannot diverge. In particular, from (4) we obtain $P_2(2^{n+2} - 1) = P_2(2^{n+1} + (2^{n+1} - 1)) = P_2(2^{n+1} - 1) + P_2(2^{n+1} - 1) + 1$. So, we have proved Lemma 2 for process P_2 in the interval $2^{n+1} \dots 2^{n+2}$ if t is an even number. Now consider the process $P_1(2^{n+1} + a)$ for even values of t. For small values of $t << 2^n$, when $a << 2^{n+1} - 1$, the behavior of the process $P_1(2^{n+1} + a)$ is determined by the properties of the process $P_1(a)$. In this case, $3a - 1 = 12l - 4$ is divisible at least by 8 if l is an odd number, and is divisible by 4 if l is an even number. Therefore, for odd l, the process quickly reaches the number k_N from the previous interval. For even l, more iterations are required to reach k_N than for odd l. For some values of a from the area of attraction of centers 5 or 17, the number of iterations required to reach k_N can be large, but it certainly does not exceed the estimate x_0 obtained above. For large values of a, put $a = 2^{n+1} - b$, where b corresponds already to odd values of t_1. In this case, for small values of l_2, when $b << 2^{n+1} - 1$, the behavior of the process $P_1(2^{n+1} + a)$ is determined by the properties of the process $P_1(-b) = -P_2(b)$. Therefore, the process quickly reaches the number k_N. To obtain general relations for the process $P_1(k_0)$, we represent k_0 as $k_0 = 2^{n+2} - b$, where b is an odd number from the interval $2^{n+1} \dots 2^{n+2}$ for even values of t; $3 \le b \le 2^{n+1} - 1$. We write the first iteration as the sum of terms that are divisible only by 2. We have $P_1(2^{n+2} - b) = (3 \cdot 2^{n+2} - 3b - 1)/2 = [(3 \cdot 2^{n+2} - 3 + 1) - (3b + 1) + 2]/2$. For even values of t, the expressions $(3 \cdot 2^{n+2} - 3 + 1)$ and $(3b + 1)$ are divisible only by 2. After cancellation by 2, we obtain the following relation

$$P_1(2^{n+2} - b) = P_2(2^{n+2} - 1) - P_2(b) + 1 \quad (5)$$

In (5) $P_2(2^{n+2} - 1)$ is determined from (4). It follows from (4) that the process $P_2(2^{n+2} - 1)$ does not diverge; $P_2(b)$ does not diverge, since b belongs to one of the previous intervals. Therefore, the process $P_1(2^{n+2} - b)$ cannot diverge. Thus, we have proved Lemma 2 for the process P_1 for even values of t.

Now let t is an odd number. We put $t = 2l - 1$, where $l = 1, 2, 3$, etc., 2^{n-1}. Then $a = 4l - 3$ and a takes the values 1, 5, 9, 13, etc., $2^{n+1} - 3$, forming an arithmetic progression $1 + 4l_2$. Consider the process $P_1(2^{n+1} + a)$. For small values of $t << 2^n - 1$, when $a << 2^{n+1} - 3$, the behavior of the process $P_1(2^{n+1} + a)$ is determined by the properties of the process $P_1(a)$. In this case, $3a - 1$ is divisible only by 2. The length of the chain of iterations on which $F_1(a)$ is divisible only by 2 varies from n at $t = 1$ to 1 at $t = 2^n - 1$. For the values of a, which are the centers of attraction for the process $P_1(a)$, i.e. for $a = 1, 5, 17$, as well as for the odd numbers from the area of attraction of the centers 5 or 17 obtained during subsequent iterations, the number of iterations required to reach k_N from the previous interval can be large, but it does not exceed the above estimate x_0. For large values of a, put $a = 2^{n+1} - b$, where b corresponds already to even values of t_1. We put $t_1 = 2l$, where $l = 1, 2, 3$, etc. Then $b = 4l - 1$ and b takes the values 3, 7, 11, etc., forming the arithmetic progression $3 + 4l_1$, where $l_1 = 0, 1, 2$, etc. In this case, for small values of l_1, when $b << 2^{n+1} - 1$, the behavior of the process $P_1(2^{n+1} + a)$ is determined by the properties of the process $P_1(-b) = -P_2(b)$. Therefore, the process quickly reaches the number k_N. To obtain general relations for the process $P_1(k_0)$, we represent k_0 as $k_0 = 2^{n+2} - b$, where b is an odd number from the interval $2^{n+1} \dots 2^{n+2}$ for odd values of t; $1 \le b \le 2^{n+1} - 3$. Take the smallest odd number $t = 1$, then $b = 1$. After three iterations, we have $P_1P_1P_1(2^{n+2} - 1) = 3^3 \cdot 2^{n-4} - 1 < 2^{n+1} - 1$; therefore, the number $3^3 \cdot 2^{n-4} - 1$ belongs to the previous interval for which Lemma 2 is valid by assumption. Take the largest odd number $t = 2^n - 1$, then

$b = 2^{n+1} - 3$. After the first iteration, we have $P_1(2^{n+2} - (2^{n+1} - 3)) = P_1(2^{n+1} + 3) = 3{\cdot}2^{n-2} + 1 < 2^{n+1} + 1$; therefore, the number $3{\cdot}2^{n-2} + 1$ belongs to the previous interval for which Lemma 2 is valid by assumption. So, we have proved Lemma 2 for the largest and smallest values of t if t is odd. Consider the general case for arbitrary b. For odd t, the values $b = 2t - 1$ form an arithmetic progression $1 + 4l_2$, where $l_2 = 0, 1, 2$, etc., $2^{n-1} - 1$. Then $(3b - 1) = 2 + 12l_2$ is divisible only by 2 We write the first iteration as a sum of terms that are divisible only by 2. We have $P_1(2^{n+2} - b) = (3{\cdot}2^{n+2} - 3b - 1)/2 = ((3{\cdot}2^{n+2} - 3 + 1) - (3b - 1))/2$. After cancellation by 2, we obtain the following relation

$$P_1(2^{n+2} - b) = P_2(2^{n+2} - 1) - P_1(b) \tag{6}$$

Each term on the right side of (6) is an odd number. Therefore, both sides of (6) can be simultaneously cancelled by 2 or a power of 2, depending on the value of b. It follows from (4), that the process $P_2(2^{n+2} - 1)$ does not diverge; $P_1(b)$ does not diverge, since b belongs to one of previous intervals; $1 \le b \le 2^{n+1} - 3$. $P_1(b) = 1 + 6l_2$; $1 \le P_1(b) \le 3{\cdot}2^n - 5$. In addition, it follows from (6) that $P_1(2^{n+2} - b) < P_2(2^{n+2} - 1)$, since $P_2(2^{n+2} - 1) > P_1(b)$. Therefore, the process $P_1(2^{n+2} - b)$ cannot diverge. So, we have proved Lemma 2 for process P_1 for odd values of t. Consider process $P_2(2^{n+1} + a)$ for odd t. We put, as above, $t = 2l - 1$, where $l = 1, 2, 3$, etc., 2^{n-1}. Then $a = 4l - 3$ and a takes the values 1, 5, 9, 13, etc., $2^{n+1} - 3$, forming the arithmetic progression $1 + 4l_2$. The behavior of the process $P_2(2^{n+1} + a)$ is determined by the properties of the process $P_2(a)$. In this case, $3a + 1 = 12l - 8$ is divisible at least by 8 for even l and divisible by 4 for odd l. We can verify directly by calculations that in the first case, for any a, the process $P_2(2^{n+1} + a)$ reaches the number k_N in several iterations. In the second case, i.e. for odd l, for small values of a, a larger number of iterations is required, but no more than n. For large values of a, put $a = 2^{n+1} - b$, where b corresponds already to even t_1. We put $t_1 = 2l$, where $l = 1, 2, 3$, etc. Then $b = 4l - 1$ and b takes the values 3, 7, 11, etc., forming the arithmetic progression $3 + 4l_1$, where $l_1 = 0, 1, 2$, etc. In this case, for small values of l_1, when $b << 2^{n+1} - 1$, the behavior of the process $P_2(2^{n+1} + a)$ is determined by the properties of the process $P_2(-b) = -P_1(b)$. For the values of b, which are the centers of attraction for the process $P_1(b)$, namely, for $b = 17$, as well as for odd numbers from the area of attraction of centers 5 or 17 obtained during subsequent iterations, the number of iterations required to reach k_N from the previous interval can be large, but it does not exceed the above estimate x_0. In particular, for $t = 1$ ($l = 1$) and $t = 2^n - 1$ ($l = 2^{n-1}$), only one iteration is required to reach the number k_N from the previous interval. Thus, Lemma 2 is valid for the smallest and largest odd number t in the interval under consideration. To obtain a general relationship, we write the first iteration as the sum of terms that are divisible only by 2. We have $P_2(2^{n+1} + a) = (3{\cdot}2^{n+1} + 3a + 1)/2 = ((3{\cdot}2^{n+1} + 3 - 1) + (3a - 1))/2$. After cancellation by 2, we obtain the relation

$$P_2(2^{n+1} + a) = P_1(2^{n+1} + 1) + P_1(a) \tag{7}$$

Each term on the right side of (7) is an odd number. Therefore, both sides of (7) can be simultaneously cancelled by 2 or a power of 2, depending on the value of a. It follows from (5), that the process $P_1(2^{n+1} + 1)$ does not diverge, since $P_1(2^{n+1} + 1) = P_1(2^{n+2} - b)$ for $b = 2^{n+1} - 1$. The process $P_1(a)$ does not diverge, since a belongs to one of previous intervals; $1 \le a \le 2^{n+1} - 3$. $P_1(a) = 1 + 6l_2$; $1 \le P_1(a) \le 3{\cdot}2^n - 5$. Therefore, the process $P_2(2^{n+1} + a)$ cannot diverge. Thus, we have proved Lemma 2 for P_2 in the considered interval for odd numbers t. We give in addition some relations that establish a connection between the processes P_1 and P_2 and allow mutual verification of the results obtained. We have obtained these relations in the same way as the previous ones. For odd values of t, the following equalities are valid

$$P_1(2^{n+1} + a) = P_1(2^{n+1} + 1) + P_1(a) - 1 \tag{8}$$

$$P_2(2^{n+2} - b) = P_2(2^{n+2} - 1) - [P_1(b) - 1]. \tag{9}$$

In equality (9), $P_2(2^{n+2} - b) < P_2(2^{n+2} - 1)$. For even values of t, the following equalities are valid

$$P_2(2^{n+1} + a) = P_1(2^{n+1} + 1) + P_2(a) - 1 \tag{10}$$

$$P_1(2^{n+1} + a) = P_1(2^{n+1} + 1) + P_2(a) - 2 \tag{11}$$

$$P_2(2^{n+2} - b) = P_2(2^{n+2} - 1) - [P_2(b) - 2]. \tag{12}$$

In (11) the terms $P_1(2^{n+1} + 1)$ and $P_2(a)$ are odd numbers. Therefore, both sides of (11) can be simultaneously cancelled by 2 or a power of 2, depending on the value of a. In equality (12) $P_2(2^{n+2} - b) < P_2(2^{n+2} - 1)$. In (12) both sides of equality can be simultaneously cancelled by 2 or a power of 2, depending on the value of b. In equalities (4) – (12), all values are obtained using a single iteration, i.e. by division an even number only by 2. In equalities (5), (6), (9) and (12), b can be replaced by a, using the relation $a + b = 2^{n+1}$. Then, in some cases, when $2^{n+2} - b > 3{\cdot}2^n$, i.e. $2^{n+2} - b$ more than the middle of the interval under consideration, the equalities are simplified and the validity of Lemma 2 becomes obvious. For example, put in (5) $b = 2^{n+1} - a$, where $a = 1$. Then (5) is transformed to the form $P_1(2^{n+2} - b) = P_1(2^{n+1} + 1) = P_2(2^{n+2} - 1) - P_2(2^{n+1} - 1) + 1$. From (4) we have $P_2(2^{n+2} - 1) = P_2(2^{n+1} - 1) + P_2(2^{n+1} - 1) + 1$. Finally we obtain the equality $P_1(2^{n+1} + 1) = P_2(2^{n+1} - 1) + 2$. Since $2^{n+1} - 1$ belongs to the previous interval, then for $P_2(2^{n+1} - 1)$ Lemma 2 is valid by assumption; therefore, it is also valid for $P_1(2^{n+1} + 1)$. Relations (4) –(12) are valid for arbitrary n. Due to relations (9), (12) for process P_2 and relations (5), (6) for process P_1, the number of iterations required to reach k_N from the previous interval does not exceed the estimate x_0 obtained above. Thus, Lemma 2 is proved for the interval $2^{n+1} \ldots 2^{n+2}$. This implies its validity for the processes P_2 and P_1 that start at an arbitrary odd number. Process $P_2(2^{n+2} – 1)$, as well as process $P_1(2^{n+1} + 1)$, has exponential complexity, since the number k_N from the previous interval is reached in O(n) iterations; the process $P_3(2^{n+2} – 1)$, as well as the process $P_3(2^{n+1} + 1)$, has polynomial complexity, since k_N is reached in 1 iteration. We use the results of this section to consider the looping problem.

IV. Lemma on looping. Solving the Collatz problem

So, in the iterative process P_2, the reason preventing the achievement of 1 can be looping. We will prove that looping is not possible in this process. Consider in more detail odd numbers and their areas of attraction in the iterative processes P_2 and P_1.

In the iterative process P_1 using the sequence $(3k – 1)$, the number 3 has the trajectory (3, 8, 4, 2, 1), i.e. 3 is included in the area of attraction of the number 1. The number 5 has the trajectory (5, 14, 7, 20, 10, 5), i.e. 5 forms the loop and this number is the center of attraction. For the number 7, the trajectory has the form (7, 20, 10, 5), i.e. 7 is included in the area of attraction of the number 5 and participates in the formation of the loop. Similarly, we determine that the area of attraction of the number 5, in addition to the number 7, includes the numbers 9, 13, 19, 27, 35, 47, 51, 63, 75, 81, 89, 93, 107, etc. All these numbers lead to looping in the iterative process P_1. Take the number 17. For it, the trajectory has the form (17, 50, 25, 74, 37, 110, 55, 164, 41, 122, 61, 182, 91, 272, 136, 68, 34, 17), so the number 17 forms a new loop independently from the number 5, and 17 is the center of attraction. The area of attraction of number 17 includes numbers 21, 23, 25, 31, 33, 37, 41, 45, 49, 55, 61, 67, 73, 79, 83, 91, 99, etc. The area of attraction of number 1 includes numbers 11, 15, 29, 39, 43, 53, 57, 59, 65, 69, 71, 77, 85, 87, 95, 97, 101, 103, 105, etc. The trajectories of large numbers always contain a part of the trajectory of smaller numbers, which makes it easier to determine their belonging to the area of attraction of a given center of attraction. Therefore, if the trajectory of a large number includes a smaller number for which the area of attraction is already known, then the large number also belongs to the area of attraction of this center. The presence in the iterative process P_1 of two centers of attraction other than 1, namely 5 and 17, is a sufficient reason that this process does not always reach 1. The presence of other centers does not change this conclusion. The number 5 ($k = 2$) is the smallest number in the sequence $(3k – 1)$, the numbers 3, 7, 9, 13 and 15 do not belong to this sequence. The number 11 ($k = 4$) corresponds to the first case, i.e. it forms the power of number 2; so it does not form loop. The number 17 ($k = 6$) is the next smallest number in this sequence, that is not included in the area of attraction of number 5 ore number 1; therefore, 17 is the center of attraction. Numbers included in the area of attraction of these centers lead to a looping of the iterative process P_1. We can assume that there are no other centers (see Lemma 3). In this case, since the iterative process P_1 cannot diverge, we always reach numbers (even or odd) that belong to the area of attraction of one of the numbers 1, 5 or 17, which are the smallest odd numbers attainable in the iterative process P_1. The following relations allow us to determine the values of k at which $P_1(k)$ does not go in loops. For values $k = (2^n + 1)/3$ corresponding to the first case, there is no looping in the iterative process P_1. Put $k_1 = (2^{m1}k + 1)/3$; where m_1 runs through the values 1, 2, 3, 4, etc. for each value of k. If, for a given k, we select $m1$ so that k_1 is an integer, then there is no looping for this value of k_1. Put $k_2 = (2^{m2}k_1 + 1)/3$; where $m2$ runs through the values 1, 2, 3, 4, etc. for each value of k_1 obtained above. If we select $m2$ for given k_1 so that k_2 is an integer, then there is no looping for this value of k_2. We can continue this process. Using these values, we can determine other values of k that do not have looping. If there is no looping at a known value k_1, then assuming $k_2 = (((2^{m1}k_1 + 1)/3)2^{m2} + 1)/3\ldots$ etc. and choosing at each step the values $m1$, $m2$, etc. so that the expressions in parentheses are divisible by 3, we obtain the values of k_2 for which there is no looping. The above relations allow, at least theoretically, to obtain all values of k for which the process $P_1(k)$ does not go in loops. We cannot establish a simple general rule, which would allow us to predict the appearance of loop. For example, at $k = 5461$ there is no looping, since this number is in the area of attraction of number 1, and when $k = 341$ there is looping, since this number is in the area of attraction of number 5. It is easy to check that in these examples the number k_N is achieved in several iterations.

We can explain the cause of looping as follows. For even k, the smallest odd number in the sequence F_1 is 5 (for $k = 2$). Then, if in the iterative process P_1 using this sequence, we reach the number 5 after division of even number, for example, 20/4 = 5, then further decrease the number in this process is impossible. Put $k = 5$, we obtain $(3k – 1) = 14$ and 14/2 = 7; put $k = 7$, we obtain $(3k – 1) = 20$, 20/4 = 5. The loop is formed, since $5 \neq (2^n + 1)/3$. In the sequence F_2, the smallest odd number for even k is 7 (for $k = 2$). If in the iterative process P_2 using this sequence, we reach the number 7 after division of even number, for example, 28/4 = 7, then further decrease the number in this process is possible, since number 7 is in the area of attraction of number 1. Put k = 7, we obtain $(3k + 1) = 22$ and 22/2 = 11. If we apply the optimal iterative process P_3 using both sequences, then put $k = 11$, and we obtain $(3k – 1) = 32$ and 32/32 = 1 (end of procedure). If we use only the sequence $(3k + 1)$, then it takes more operations to reach 1. For $k = 11$ we get $(3k + 1) = 34$ and 34/2 = 17. Put $k = 17$, we get (3k +

1) = 52 and 52/4 = 13. Put $k = 13$, we obtain $(3k + 1) = 40$ and 40/8 = 5. But $5 = (2^n - 1)/3$, so for $k = 5$, we obtain $(3k + 1) = 16$ and 16/16 = 1. There is no looping, although the iterative process P_2 is not optimal. Similar reasoning is valid for the number 17 ($k = 6$). If in the iterative process P_1 we reach the number 17 after division of even number, for example, 68/4 = 17, then further decrease the number in the iterative process P_1 is impossible, since the number 17 is not included in the area of attraction of the number 5 or the number 1. In the iterative process P_2, further decrease the number is possible, since 17 is included in the area of attraction of the number 1.

Now consider the process P_2 in more detail. In this process, the number 3 has the trajectory (3, 10, 5, 16, 8, 4, 2, 1), i.e. 3 is included in the area of attraction of number 1. The number 5 has the trajectory (5, 16, 8, 4, 2, 1), which is part of the trajectory of the number 3, i.e. 5 is included in the area of attraction of number 1. For the number 7, the trajectory is (7, 22, 11, 34, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 1), it contains the trajectory of the number 5, i.e. 7 is included in the area of attraction of number 1. Similarly, if the trajectories of numbers more than 7 have a common part with the trajectory of the number 7, then these numbers are included in the area of attraction of the number 1. In particular, the trajectories of odd numbers from 9 to 107 contain a part of the trajectory of the number 7, and each subsequent number contains a part of the trajectory at least one of the previous numbers. **Lemma 3** is valid: "The iterative process P_2 based on the sequence $(3k + 1)$ cannot go in loops". We shall prove Lemma 3 by induction. We use the notation and results of the previous section. In the interval from 2^1 to 2^2, there is one odd number 3. The process $P_2(3)$ does not go in loops. Therefore, Lemma 3 is valid for this interval. Take the interval 2^2 ... 2^3. In this interval, there are two odd numbers 5 and 7. The process $P_2(5)$, as well as the process $P_2(7)$ does not go in loops. Therefore, Lemma 3 is valid for this interval. Take the interval 2^n ... 2^{n+1}. Suppose that for an arbitrary initial odd number k_0 from this interval the process $P_2(k_0)$ does not go in loops. Suppose that the Lemma 3 is valid for all intervals 2^n ... 2^{n+1}, if $n \le m$. We shall prove the validity of Lemma 3 for an arbitrary initial odd number k_0 from the interval 2^{n+1} ... 2^{n+2}, if $n = m$. Since the process $P_2(k_0)$ cannot diverge, the numbers obtained in this process cannot infinitely move away from k_0. Therefore, it suffices to prove that k_0 cannot be a new center of attraction for the process $P_2(k_0)$. This statement is equivalent to the fact that either $P_2(k_0)$ cannot return to k_0, or k_0 belongs to the trajectory of the number $k_1 < k_0$.

It follows from the previous section that if $k_0 = 2^{n+1} + a$, then for small values of a the process $P_2(k_0)$ reaches, in several iterations, the number k_N, $k_N < 2^{n+1} + 1 < k_0$, from the previous interval for which Lemma 3 is valid by assumption. Therefore, the number k_0 cannot be a new center of attraction. In particular, for $a = 1$, the number k_N is reached in one iteration. Similarly, if $k_0 = 2^{n+2} - b$, then for some small values of b the number k_N is reached in several iterations. For all such numbers k_0, Lemma 3 is valid. If $b = 1$, then a large number of iterations of order $O(n)$ are required to reach k_N, and it is impossible to prove the validity of Lemma 3 directly by calculations. It follows from relations (9), (12) that $P_2(2^{n+2} - b) < P_2(2^{n+2} - 1)$; therefore, the number of iterations in the process $P_2(2^{n+2} - b)$ required to reach the number k_N is is not more than the estimate x_0 obtained in the previous section for the process $P_2(2^{n+2} - 1)$.

Now we prove that an arbitrary odd number k_0 from the considered interval cannot be a new center of attraction for the process $P_2(k_0)$. For the process $P_2(k_0)$ three cases are possible: 1). $2k_0$ belongs to the sequence $F_2 = (3k + 1)$; 2). $4k_0$ belongs to the sequence F_2; 3). k_0 is divisible by 3. In the first and second cases, it is sufficient to consider the smallest odd and even powers of 2, respectively. In the first case, k_0 belongs to the sequence F_1, but already for even k. We have $2k_0 = 3k_1 + 1$, then $k_1 = (2k_0 - 1)/3 < k_0$. Consequently, k_0 cannot be a new center of attraction since it belongs to the trajectory of the number k_1 in the process P_2: $P_2(k_1) = k_0$. The number k_1 does not necessarily belong to the previous interval; it can belong to the considered interval. In the third case, k_0 belongs to the sequence $3k$ for odd k. We have after the first iteration $P_2(k_0) = (3k_0 + 1)/2^m = k_1$. If m is even, then k_1 belongs to the sequence F_2, and if m is odd, then k_1 belongs to the sequence $F_1 = 3k - 1$ (here k is an even number). Continuing the process P_2, we obtain a series of odd numbers that alternately belong to F_2 or F_1 and, therefore, are not divisible by 3. Therefore, in the process $P_2(k_0)$ we cannot return to the number k_0, and k_0 cannot be a new center of attraction. The second case is reduced to the first or third case. In the second case, k_0 belongs to the sequence F_2. We have $4k_0 = 3k_1 + 1$, then $k_1 = (4k_0 - 1)/3 > k_0$. If k_1 belongs to the sequence F_1, then $2k_1$ belongs to the sequence F_2, and we obtain, as in the first case, $k_2 = (2k_1 - 1)/3 < k_0$. If k_1 belongs to F_2, then $4k_1$ belongs to F_2, hence $k_2 = (4k_1 - 1)/3 > k_1$. If k_2 again belongs to F_2, then $k_3 = (4k_2 - 1)/3 > k_2$, and if k_2 belongs to F_1, then $k_3 = (2k_2 - 1)/3 < k_2$, etc. The numbers k_1, k_2, k_3, etc. form a trajectory along which the number k_0 can be reached in the process P_2. The trajectory is determined in a unique manner by the value of the number k_0. The trajectory of this "reverse" process always ends with the number k_0 and begins with a multiple of 3 belonging to the sequence $F_3 = 3 + 6l$, where $l = 0, 1, 2$, etc. Indeed. The odd numbers $4k_0 - 1$, $2k_1 - 1$, $4k_1 - 1$, $4k_2 - 1$, $2k_2 - 1$ and others, obtained in the above "reverse" process, belong to the sequence F_3, that is, are divisible at least by 3. Therefore, among these numbers there is always a number divisible at least by 9. As a result, on a finite segment of this "reverse" process, we obtain a series of odd numbers belonging alternately to F_1 and F_2, and one number belonging to F_3. The number k_0 can be reached in the process P_2 from a number belonging to F_3, but a number belonging to F_3 cannot be reached in the process $P_2(k_0)$, since this process does not contain numbers divisible by 3. Therefore, k_0 cannot be a new center of attraction. After a finite number of iterations, we reach the number k_l, which satisfies one of the conditions: 1). $k_l < k_0$, 2). k_l is divisible by 3. The first condition can be satisfied because the odd numbers from the sequences F_1 and F_2 alternate, and the chain of numbers from the sequence F_2 always has a finite (limited) length. The second condition is satisfied because odd numbers divisible by 3 are evenly distributed along the number axis and have the shortest repetition period compared to other numbers. If numbers from sequence F_2 prevail over numbers from F_1, then the second condition is satisfied earlier than the first one. If numbers from sequence F_1 prevail over numbers from F_2, then the first condition is satisfied earlier than the second one, or both conditions can be satisfied simultaneously. In any case, the above "reverse" process always ends with an odd number divisible by 3, if the number k_0 is taken as the beginning. In other words, arbitrary odd number k_0 from the sequences F_1 and F_2 can be obtained from an odd number divisible by 3 using the process P_2, but it is impossible to obtain a number divisible by 3 from the number k_0 in the process $P_2(k_0)$. Let us show that numbers divisible by 3 regularly appear in the "reverse" process described above. If k_0 belongs to F_2, then k_0 can be represented as $k_0 = 7 + 6l$ (see Section I). Hence we obtain $k_1 = (4k_0 - 1)/3 = 9 + 8l$. The number k_1 belongs to F_3, i.e. is divisible by 3, for $l = 3l_1$, where $l_1 = 0, 1, 2$, etc., and then the proof ends, since the second condition is satisfied. The belonging of the number k_1 to the sequence F_2 or F_1 depends on its position relative to the number divisible by 3. So k_1 belongs to F_2 for $l = 2 + 3l_1$, where $l_1 = 0, 1, 2$, etc.; k_1 belongs to F_1 for $l = 1 + 3l_1$. Let k_1 belongs to F_2. Continuing the "reverse" process, we obtain $k_2 = (4k_1 - 1)/3 = (35 + 32l)/3 = 33 + 32l_1$. The number k_2 is divisible by 3 for $l_1 = 3l_2$, and then the proof ends, since the second condition is satisfied. The number k_2 belongs to F_2 for $l_1 = 2 + 3l_2$; k_2 belongs to F_1 for $l_1 = 1 + 3l_2$. Let k_2 again belongs to F_2. We obtain $k_3 = (4k_2 - 1)/3 = 129 + 128l_2$. The number k_3 is divisible by 3 for $l_2 = 3l_3$, and then the proof process ends, since the second condition is satisfied. The number k_3 belongs to F_2 for $l_2 = 2 + 3l_3$; k_3 belongs to F_1 for $l_2 = 1 + 3l_3$. If k_3 again belongs to F_2, then the relationships that determine whether k_4 belongs to F_1, F_2, and F_3 remain the same as for k_3. The general expression has the form $k_n = (2^{2n+1} + 1) + 2^{2n+1}l_n$, where $l_n = 0, 1, 2$, etc., and for any n, the relationships that determine whether k_n belongs to F_1, F_2, and F_3 remain the same as above. Now let k_1 belongs to F_1, that is, $l = 1 + 3l_1$. Continuing the "reverse" process, we reach $k_2 = (2k_1 - 1)/3 = (17 + 16l)/3 = 11 + 16l_1$. The number k_2 is divisible by 3 for $l_1 = 1 + 3l_2$, and then the proof process ends, since the second condition is satisfied. The number k_2 belongs to F_2 for $l_1 = 2 + 3l_2$; k_2 belongs to F_1 for $l_1 = 3l_2$. Here the first condition is satisfied before the second one, and the second condition can be satisfied simultaneously with the first one. Now let k_n is an odd number closest to k_0, which belongs to F_1, such that $k_0 < k_n$, and all numbers preceding k_n belong to F_2. Here the result depends on n, i.e. on how many preceding numbers belong to F_2. We write at once the general expression for k_n. We have $k_n = (2^{2n+3} + 1)/3 + 2^{2n+2}l_n$, where $l_n = 0, 1, 2$, etc. for any n. The terms $(2^{2n+3} + 1)/3$ corresponding to different values of n form the sequence 11, $11 + 2^5$, $11 + 2^5 + 2^7$, $11 + 2^5 + 2^7 + 2^9$, etc. If $n = 1$, then one number belonging to F_2, namely, the number k_0 itself, precedes k_n; if $n = 2$, then two numbers belonging to F_2 precede k_n, etc. In this case, for an arbitrary n, n iterations F_1 are required to satisfy the first condition. We consider only one iteration F_1 for each n. Hence, for $n = 1$, we obtain $k_1 = 11 + 2^4l_1$. The number k_1 is divisible by 3 for $l_1 = 1 + 3l_{12}$, and then the proof process ends, since the second condition is satisfied; k_1 belongs to F_2 for $l_1 = 2 + 3l_{12}$, k_1 belongs to F_1 for $l_1 = 3l_{12}$, which coincides with the above result. For $n = 2$ we obtain $k_2 = 43 + 2^6l_2$. The number k_2 is divisible by 3 for $l_2 = 2 + 3l_{21}$, and then the proof process ends; k_2 belongs to F_2 for $l_2 = 3l_{21}$, k_2 belongs to F_1 for $l_2 = 1 + 3l_{21}$. For $n = 3$, we have $k_3 = 171 + 2^8l_3$. The number k_3 is divisible by 3 for $l_3 = 3l_{31}$, and then the proof process ends; k_3 belongs to F_2 for $l_3 = 1 + 3l_{31}$, k_3 belongs to F_1 for $l_3 = 2 + 3l_{31}$. For $n = 4$ we obtain $k_4 = 683 + 2^{10}l_4$. The number k_4 is divisible by 3 for $l_4 = 1 + 3l_{41}$, and then the proof process ends; k_4 belongs to F_2 for $l_4 = 2 + 3l_{41}$, k_4 belongs to F_1 for $l_4 = 3l_{41}$. The belonging of numbers k_n to sequences F_1, F_2 and F_3 changes with n and repeats with period 3. Therefore, after a certain number of steps, we always reach an odd number divisible by 3, and then the proof process ends. Similarly, the possibility is considered when F_2 and F_1 alternate, but in such a way that the first condition is not satisfied. Of course, not all of the possibilities discussed above are realized simultaneously, since the trajectory is completely (in a unique manner) determined by the value of k_0. Thus, in all three cases, k_0 cannot be a new center of attraction. Thus, Lemma 3 is proved for the interval $2^{n+1} \ldots 2^{n+2}$. This implies its validity for the process P_2, which starts at an arbitrary odd number. Below we will explain this behavior of the process P_2. Using the results of the previous section and carrying out arguments similar to those given above, we can also conclude that k_0 cannot be a new center of attraction for the process $P_1(k_0)$, if $n + 1 \geq 5$. However, for the process P_1, Lemma 3 is not valid, since k_0 (respectively, k_N or k_1) can belong to the area of attraction of numbers 5 or 17.

Consider the process $P_2(k_0 = 2^{n+2} - 1)$ as an example (see Section III). If the number $n + 2$ is even, then $k_0 = 2^{n+2} - 1$ is divisible by 3 (case 3). If the number $n + 2$ is odd, then k_0 belongs to the sequence F_2 (case 2). Therefore, the number k_0 cannot be a new center of attraction. Sometimes it is possible to verify the validity of Lemma 3 for $P_2(k_0 = 2^{n+2} - 1)$ by calculations. If $n + 2$ is an odd number, then the process $P_1(k_0)$ and, therefore, $P_1(k^{n+2}/2)$, can go in loops (see Section III). To establish this, several iterations are sufficient, since the process

$P_1(k_0)$ converges quickly. If $P_1(k_0)$ goes in loops, then $P_2(k_0)$ cannot go in loops, since this would contradict the conditions of Lemma 1.

The reason for the difference between the processes P_1 and P_2 is as follows. An iterative process goes in loops if some number (even or odd) repeat in it. The pair of numbers $\{(3k - 1), (3k + 1)\}$, when k is even, forms a representation of an even number $6k$ as a sum of two odd terms, and if k is odd, this pair forms a representation of an even number as a sum of two even terms. Let k is an arbitrary initial odd number. This is the center of representation for the even number $2k$. Then the even numbers $(3k - 1)$ and $(3k + 1)$ form the representation of the even number $6k$ with the center of the representation equal to $3k$. The number $(3k - 1)$ from the sequence F_1 is always located to the left of the center of the representation of the corresponding even number, and the number $(3k + 1)$ from the sequence F_2 is always located to the right of the center. When k changes, the numbers to the left of the center belonging to the sequence F_1 can repeat in the iterative process P_1, which leads to looping. At the same time, this problem does not exist for the process P_2, since the numbers to the right of the center belonging to the sequence F_2 are always replaced with new ones in the iterative process P_2 and cannot be repeated, so there is no looping.

Since according to Lemma 2 and Lemma 3, the iterative process P_2 cannot diverge or go in loops, then **Corollary 1** from Lemmas 2 and 3 is valid: "The iterative process P_2 based on the sequence $(3k + 1)$ always reaches 1".

V. Conclusion

Thus, the optimal iterative process P_3 for achieving 1 uses both sequences $(3k - 1)$ and $(3k + 1)$. The process P_2 that uses only a sequence $(3k + 1)$ also achieves 1, but in general case it requires much more iterations. The process P_1 that uses only the sequence $(3k - 1)$ in the general case does not allow reaching 1, since this process can go in loops. The above reasoning and conclusions, *mutatis mutandis*, remain valid when k takes negative values. In this case, the relations $P_1(-k) = -P_2(k)$ and $P_2(-k) = -P_1(k)$ are valid. The sequence $(3k - 1)$ is replaced by the sequence $[-(3|k| + 1)]$, and the sequence $(3k + 1)$ is replaced by the sequence $[-(3|k| - 1)]$, where $|k|$ is the modulus of the number k. Center of attraction 5 is replaced by (–5), center 17 by (–17), center 1 by (–1). The optimal iterative process P_3 uses both sequences $[-(3|k| + 1)]$ and $[-(3|k| - 1)]$; it achieves (–1) in the minimum number of iterations. The iterative process based on the sequence $[-(3|k| + 1)]$ also achieves (–1), but in more iterations. The process using the sequence $[-(3|k| - 1)]$ does not always reach (–1), since at $k = -5$ and $k = -17$ there is looping.

References

[1]. Romanov V.N. Study of fundamental problems of number theory. Saint Petersburg, Publishing house "Asterion", 2015 (in Russian).

[2]. Romanov V.N. Fermat quadratic equation and prime numbers. IOSR Journal of Mathematics, V. 15, Issue I (Ser. II), 2019, pp. 19 – 23.

Vadim Nikolayevich Romanov. "Investigation of the Collatz problem." *IOSR Journal of Mathematics (IOSR-JM)*, 17(4), (2021): pp. 43-53.